UNIVERSITÀ DEGLI STUDI DI PALERMO

Department of Engineering
Doctoral Course in Information and Communication Technologies
Doctoral Thesis

An Information-Theoretic Framework for the Machine Learning Detection of Physiological States from Cardiovascular Signals

PhD Candidate:
Marta Iovino

Supervisor:
Dr. Riccardo Pernice
Co-supervisor:
Prof. Luca Faes

XXXVIII Cycle
Academic Year 2024–2025

Contents

Abstract

Cardiovascular variability analysis represents a powerful non-invasive tool for assessing autonomic regulation, while its interpretation remains challenged by the nonlinear, multi-scale, and highly interdependent nature of physiological signals. Traditional approaches often rely on univariate or pairwise descriptors and treat the stages of the machine-learning (ML) pipeline as loosely connected steps, limiting both interpretability and physiological insight.

This thesis proposes an integrated analytical framework grounded in information theory (IT), where signal processing, feature extraction (FE), feature selection (FS), feature importance (FI), and classification are treated as interconnected stages of a unified pipeline for detecting, dissecting, and interpreting the information contained in cardiovascular variability time series. Within this framework, several methodological developments are introduced to organise and analyse this information in a systematic and interpretable way. Information-theoretic measures are employed at multiple levels: entropy, Mutual Information (MI), and Conditional Mutual Information (CMI) are used to characterise cardiovascular time series and physiological interactions; information-based selection criteria guide the identification of compact and non-redundant feature subsets; finally, high-order information decomposition frameworks are adopted to quantify unique, redundant, and synergistic contributions of features to predictive tasks.

From a methodological standpoint, the thesis introduces and validates information-theoretic approaches for FS and FI analysis and classification, explicitly accounting for feature interactions beyond first-order effects and enabling the identification of compact and non-redundant feature subsets. In addition, high-order FI frameworks based on CMI enable the decomposition of feature relevance into cooperative, redundant, and synergistic contributions. These methods complement classical ML models by improving interpretability while maintaining competitive predictive performance.

The proposed analytical framework is evaluated across multiple experimental and applied scenarios, including assessment of stress (postural, mental, driving, and work-related stress) and of sex-related differences in cardiovascular regulation. Results consistently show that postural stress elicits clearer autonomic signatures than cognitive stress, that ultra-short-term features can retain discriminative capability when carefully selected, and that sex-related differences emerge primarily through coordinated patterns of cardiac and

vascular features rather than isolated markers. These findings highlight how different stressors engage distinct autonomic regulatory mechanisms and show that coordinated cardiovascular interactions provide more reliable indicators of physiological state than isolated variability indices. Applications in real-world and large-cohort settings further demonstrate the generalizability of the framework and highlight the critical role of preprocessing and inter-subject normalization.

Overall, the results demonstrate that information theory provides a unifying and physiologically meaningful foundation for cardiovascular variability analysis and ML-based modelling. By integrating methodological rigor with physiological interpretability, the proposed framework may advance the understanding of autonomic regulation and support the development of transparent, interaction-aware models for biomedical and human-centered applications.

List of Abbreviations

ACC	Accuracy
AI	Artificial Intelligence
AIC	Akaike Information Criterion
ANS	Autonomic Nervous System
AP	Arterial Pressure
AR	Autoregressive
AUC	Area Under the ROC Curve
BREATH	Respiratory Signal
CE	Conditional Entropy
CM	Confusion Matrix
CMI	Conditional Mutual Information
CMIHiFi	Conditional Mutual Information High-Order Interaction FI
CR	Cardio Respiratory
CV	k-Fold Cross-Validation
cVar	Coefficient of Variation
DBP	Diastolic Blood Pressure
ECG	Electrocardiographic Signal
F1	F1-score
FE	Feature Extraction
fHF	Peak Frequency within the High-Frequency Band
FI	Feature Importance
FN	False Negative
FP	False Positive
FS	Feature Selection
H	Entropy
HF	High Frequency
HFn	High Frequency Power Normalised Units
HiFi	High-Order Interaction FI
HRV	Heart Rate Variability
HUT	Head-up Tilt
Ir	Irrelevant

IT	Information Theory
kCMI-FS	k-Nearest Neighbour Conditional Mutual Information FS
kNN	k-Nearest Neighbours
KSG	Kraskov–Stögbauer–Grassberger Estimator
LDA	Linear Discriminant Analysis
LF	Low Frequency
LFn	Low Frequency Power Normalised Units
LIC	Local Information Classifier
LOCO	Leave-One-Covariate-Out
LOSO	Leave-One-Subject-Out Cross-Validation
MA	Mental Aritmetics
MEAN	Mean Inter-beat Interval
MeanHR	Mean Heart Rate
MI	Mutual Information
ML	Machine Learning
mRMR	Minimum Redundancy Maximum Relevance
NB	Naive Bayes
NN	Neural Network
NNi	Normal-to-Normal Interval
PID	Partial Information Decomposition
pNN50	Percentage of Successive NN Intervals Differing by More Than 50 ms
PP	Peak-to-Peak Intervals of the PPG
PPG	Photoplethysmographic Signal
PRE	Precision
PRV	Pulse Rate Variability
PSD	Power Spectral Density
R	Redundant Information
Rd	Redundant
REC	Recall
RESP	Respiratory Time Series
REST	Resting state
RF	Random Forest
Rl	Relevant
RMSSD	Root Mean Square of Successive Differences
ROC	Receiver Operating Characteristic
RR	R-R Intervals of the ECG
RSA	Respiratory Sinus Arrhythmia

S	Synergistic Information
SBP	Systolic Blood Pressure
SDNN	Standard Deviation of NN Intervals
SE	Self Entropy
SHAP	SHapley Additive exPlanations
SPE	Specificity
ST	Short Term
SVB	Sympathovagal Balance (LF/HF Ratio)
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
TPR	True Positive Rate (Sensitivity)
U	Unique Information
UST	Ultra Short Term
VAR	Variance of NN Intervals
VLF	Very Low Frequency
XAI	Explainable AI

Part I

Preface

1

Introduction

Over the past decades, the continuous advancement of data acquisition, computational modelling, and *Artificial Intelligence* (AI) has profoundly transformed the landscape of biomedical research. Among the various computational paradigms, *Machine Learning* (ML) has emerged as one of the most powerful frameworks for extracting knowledge from complex and high-dimensional data. ML methods are capable of identifying non-linear relationships, recognising patterns, and predicting outcomes across heterogeneous datasets, thereby providing a flexible and data-driven alternative to classical statistical models [26, 31, 114, 127]. Their growing application in medicine, neuroscience, and physiology has opened new avenues for understanding biological systems and supporting diagnostic and prognostic decision-making. This multidisciplinary convergence has led to the emergence of *computational physiology* [54] as a distinct field, integrating data-driven inference, biomedical engineering, and theoretical modelling for the quantitative study of physiological regulation. Notably, the effectiveness of ML approaches critically depends on the availability of large-scale and heterogeneous datasets, a requirement that has become increasingly attainable in recent years thanks to the widespread diffusion of public biomedical repositories and the continuous growth of high-resolution data acquired through wearable and unobtrusive sensing technologies.

Despite their predictive power, many ML approaches remain inherently limited by their lack of transparency and interpretability. Complex models such as neural networks or ensemble architectures often behave as ”black boxes”, providing accurate predictions but little insight into the mechanisms underlying the observed phenomena [156, 174, 181, 204]. In biomedical contexts, where the ultimate goal is not merely prediction but understanding, this lack of clarity poses a major barrier. Physiological data are not abstract numerical

entities; they are the manifestation of regulatory processes that sustain life and generate physiopathological states [29, 58, 82]. Therefore, models used to analyse them should not only fit the data but also reflect the causal and informational structure of the systems from which the data arise [65, 66, 105, 185, 189].

Recent advances in explainable artificial intelligence (XAI) have sought to address this issue, yet most methods remain post-hoc and lack physiological interpretability [10, 14, 156]. Within this context, *Information Theory* (IT) provides a rigorous mathematical foundation for quantifying uncertainty, dependence, and information transfer in complex systems [8, 52, 67, 107, 215]. Originally developed to describe the limits of communication and coding, IT has evolved into a general language for understanding how systems store and exchange information. Its core quantities (entropy, mutual information (MI), and conditional MI) describe how variability in one variable constrains or informs the variability of another. In physiological systems, whose states evolve with time, these measures capture the efficiency and directionality of regulatory processes, translating dynamical fluctuations into interpretable informational relationships [12, 65, 66, 189, 225, 232]. Information-theoretic measures are often introduced in the context of dynamical systems to describe how information is stored, transferred, and modified over time. However, the same framework can also be applied more generally to quantify statistical dependencies between variables. In many biomedical and physiological applications, information-theoretic concepts are therefore used to quantify associations between variables that describe properties or conditions of a system rather than its temporal evolution. Typical examples include the relationship between physiological markers and clinical conditions, such as the association between cardiovascular variability and autonomic regulation or the dependence between metabolic status and cardiovascular risk. In such settings, the variables of interest are most often continuous, while the target or outcome variable may be discrete [246].

In this context, information theory has been widely used as a theoretical framework for both feature selection and feature importance, overcoming limitations of traditional statistical and pairwise approaches by quantifying complex interactions between variables through entropy and mutual information [36, 52, 172]. Recent developments in interpretable machine learning have further sought to embed information-theoretic reasoning directly within the learning process itself, moving beyond post-hoc explanations. In this direction, novel frameworks integrate entropy-based feature relevance and uncertainty quantification to achieve intrinsic interpretability in biomedical modelling.

ML can be viewed as a process that transforms raw data into knowledge through successive analytical stages, each performing a specific task that captures, organizes, and interprets the information contained in the data [31, 38, 114, 127], as shown in Figure 1.1.

Figure 1.1: General representation of a multi-stage machine-learning pipeline. Raw data are first acquired and transformed into quantitative descriptors, which are progressively refined through feature extraction and feature selection. The resulting feature set is then used to train predictive models, completing the core sequence of the learning process and enabling the subsequent interpretation of model outcomes.

Similar multi-stage pipelines have been extensively adopted in biomedical signal analysis, particularly for cardiovascular and neural data, where feature extraction, selection, and classification are explicitly structured as sequential learning stages [38, 50, 236].

The first stage, *Feature Extraction* (FE), converts complex physiological signals into quantitative descriptors that summarise relevant properties of the system across temporal, spectral, and informational domains [12, 22, 66, 105, 114, 127, 189].

The second stage, *Feature Selection* (FS), identifies the most informative and non-redundant subset of these descriptors, reducing dimensionality while preserving predictive content [31, 36, 38, 64, 114, 120, 127, 172, 191, 228, 236]. In recent years, Information-based feature selection methods have been proposed to improve robustness and data efficiency in mixed-type biomedical datasets, offering principled control of redundancy and higher-order dependencies [12, 236].

The third stage, *learning or model training*, establishes a mapping between the selected features and the target variables, typically by optimising a loss function that minimises prediction error or, in information-theoretic terms, by reducing the residual uncertainty about the system's state [31, 38, 88, 114, 127]. Beyond these steps, the *Feature Importance* (FI) analysis quantifies the relative contribution of each variable or feature group to model performance and physiological interpretation. FI provides an essential link between predictive modelling and understanding, as it translates the statistical relevance of features into physiological meaning by identifying which aspects of variability drive classification or prediction outcomes [12, 38, 225, 246, 247].

Finally, the *validation and interpretation* phase assesses the model's generalisability and consolidates the physiological insights gained from the analysis, ensuring that the extracted and learned information reflects genuine regulatory mechanisms rather than artefactual correlations [31, 38, 114, 127, 156, 204].

In this framework, both IT and ML address the same fundamental problem: how to reduce uncertainty and extract structure from data. The theoretical framework of IT naturally complements the structure of a typical ML pipeline. Information theory provides a principled foundation for measuring and interpreting the informational structure of data,

while machine learning offers computational tools to model and generalise it. Their integration enables the development of analytical frameworks that are not only predictive but also interpretable, bridging the gap between data-driven modelling and theoretical understanding [12, 38, 66, 189]. This combination is particularly valuable in physiology, where signals reflect the coordinated activity of multiple feedback systems, and where interpretability is as crucial as predictive accuracy [12, 38, 66, 189].

In biomedical signal processing, this integration of ML and IT has become particularly relevant. Physiological recordings are nonlinear, noisy, and multivariate signals that exhibit dynamics across multiple time scales. The challenge lies in extracting information that reflects meaningful physiological mechanisms while managing redundancy and preserving interpretability [12, 38, 65, 111, 112, 161, 189]. Feature-based ML frameworks grounded in information theory address this challenge by identifying variables that maximise the information about the physiological or clinical condition of interest [40, 65, 175, 236].

Among physiological applications, the cardiovascular system offers an exemplary testbed for methodological innovation. It represents a core component of autonomic regulation, whose continuous adjustments manifest as fluctuations in heart rate, arterial pressure, and peripheral circulation. These fluctuations, collectively known as *cardiovascular variability*, reflect the dynamic balance between the sympathetic and parasympathetic branches of the autonomic nervous system (ANS) [12, 66, 108, 189, 212]. From a systems perspective, such variability emerges from a network of interacting regulatory subsystems, as emphasised by the framework of network physiology [22, 100, 100], where cardiovascular oscillations are interpreted as dynamic signatures of multiorgan coordination. Physiological signals such as electrocardiogram, photopletismogram, and respiration encode these oscillations and can be analysed to infer both local and systemic properties of autonomic control [38, 66, 148, 175, 232].

1.1 Problem Statement

Despite the rapid progress and increasing adoption of machine learning (ML) and information-theoretic (IT) approaches in biomedical sciences, several key challenges remain when these methods are applied to physiological data. The promise of ML lies in its ability to model complex, nonlinear, and high-dimensional systems such as the human body. However, its application to physiology often reveals an intrinsic tension between predictive performance, interpretability, and physiological relevance. Most learning algorithms are designed to minimise prediction error or maximise classification accuracy, but they rarely provide insight into the mechanisms underlying the observed patterns. In

biological systems, where variability reflects organised regulatory processes rather than random noise, this limited interpretability reduces the scientific and clinical value of purely data-driven models [38, 100, 156, 204]. Modern ML algorithms can capture non-linear and multivariate dependencies, but frequently operate as black-box models, offering limited transparency about how inputs are transformed into outputs. As a consequence, their predictions often lack physiological interpretability, making them difficult to validate and to translate into mechanistic insight [38, 156, 204, 236]. Moreover, information-theoretic frameworks for analysing relations among variables, as well as for decomposing predictive information into unique, redundant, and synergistic components [225, 246], provide a principled basis for interpreting complex multivariate dependencies. However, these approaches have only rarely been integrated within machine-learning pipelines for physiological data analysis, leaving open the challenge of systematically exploiting such information-theoretic tools to improve both interpretability and predictive modelling.

Several studies have suggested that the gap between predictive accuracy and physiological interpretability could be addressed by embedding information-theoretic criteria directly within the learning process, allowing model decisions to be related to quantifiable measures of informational relevance [12, 65]. Such approaches move beyond purely post-hoc explanations and point toward intrinsically interpretable learning strategies, which are particularly appealing in biomedical contexts where understanding regulatory and causal mechanisms is as important as prediction itself. However, these ideas have often been explored in a fragmented or application-specific manner, without a unified framework that systematically integrates information-theoretic principles across the different stages of physiological data analysis.

A further limitation concerns the fragmentation of analytical workflows. In most studies, preprocessing, feature extraction, selection, and classification are performed as separate tasks using different tools, each relying on distinct assumptions and criteria. This compartmentalised approach hampers reproducibility and prevents a unified understanding of how information flows across the stages of the data analysis pipeline. Moreover, physiological datasets typically include a large number of correlated features derived from time, frequency, and information domains. Without explicit mechanisms to quantify and control redundancy, models risk overfitting or misrepresenting the structure of physiological variability [12, 38, 64, 120, 138, 191, 228, 236].

Another critical issue concerns generalisability. Many existing models are highly dataset-specific, optimised for particular populations or experimental protocols, and thus fail to extend to new contexts or conditions. Physiological signals are inherently variable across individuals, time, and measurement modalities, making it essential that computational methods remain robust to inter-individual and contextual variability

[38, 50, 66, 176, 236].

Finally, while ML and IT share a common foundation in the quantification of information and uncertainty, their integration in physiological data analysis has often remained heuristic and fragmented. In many biomedical applications, information-theoretic measures are employed in a descriptive or post-hoc manner, without a coherent framework linking them systematically to the learning process itself [16, 38, 65, 66, 225]. This separation limits the potential of information theory not only as an analytical tool, but also as a guiding principle for model design, interpretation, and generalisation [204].

A convergence between machine learning and information theory would therefore offer a promising path toward analytical frameworks capable of jointly addressing predictive performance, interpretability, and physiological meaning, while preserving coherence across the different stages of the physiological data analysis pipeline.

1.2 Aims of the Thesis

The central aim of this thesis is to develop and validate a comprehensive *information-theoretic machine learning framework* for the analysis and classification of physiological states. Building upon the conceptual link between machine learning and information theory, the work seeks to formalise their convergence within a unified analytical architecture. In particular, the thesis aims to formalise this convergence into a structured analytical workflow in which the different stages of physiological data analysis—signal preprocessing, feature extraction, feature selection, feature importance analysis, and classification—are coherently integrated within a common information-theoretic framework. To this end, this thesis builds upon and extends recent methodological developments that integrate information-theoretic principles directly into ML pipelines for biomedical data analysis [17, 94, 95, 96, 97, 98, 99, 130, 131, 206].

Information-theoretic measures for the analysis of physiological systems have been extensively investigated within the BitLab research group[1], which has contributed to the development of methodological tools and software implementations for estimating entropy, mutual information, and conditional information measures in physiological signals. These developments provide part of the methodological basis for the analyses conducted in this thesis.

The first objective is to establish a *unified theoretical foundation* connecting the principles of machine learning (ML) and information theory (IT) within a common formalism. While ML provides adaptive models capable of capturing nonlinear and high-dimensional

[1] https://sites.google.com/community.unipa.it/bitlab

relationships, IT offers a rigorous mathematical framework to quantify uncertainty, dependence, redundancy, and synergy [52, 172, 215, 225, 247]. Within this perspective, learning can be interpreted as a process of information extraction and uncertainty reduction, where each analytical stage progressively refines the representation of the system's state [65, 66, 225, 246]. This thesis investigates how this perspective can be operationalised through information-based feature selection and information decomposition methods tailored to physiological data [94, 96, 97, 98, 99, 131]. In this context, information decomposition frameworks based on conditional mutual information and partial information decomposition allow the explicit quantification of unique, redundant, and synergistic contributions among features [17, 99, 131, 164]. These approaches provide a principled link between statistical modelling and physiological interpretation, enabling models that are both predictive and mechanistically meaningful.

The second objective is to address the estimation of information-theoretic quantities in mixed continuous–discrete settings, which naturally arise in classification problems where physiological features are continuous while the target variable is discrete. In particular, this work investigates information-theoretic feature selection and feature importance methods capable of reliably quantifying conditional information in such mixed-variable contexts, enabling the identification and interpretation of informative physiological descriptors [17, 98, 99, 131, 164].

A third objective is to design a *standardised and reproducible pipeline* for physiological signal analysis, ensuring that each stage—from preprocessing to classification—is grounded in an information-theoretic rationale. In this sense, the thesis does not treat these analytical stages as independent operations, but rather as interconnected components of a coherent methodological framework aimed at systematically extracting, organising, and interpreting information from physiological signals. The proposed framework integrates multidomain feature extraction, information-based feature selection, feature-importance analysis, and classification within a coherent structure, overcoming the fragmentation that characterises many existing workflows [17, 94, 95, 96, 97, 98, 99, 130, 131, 206]. This unified approach consolidates recent efforts toward reproducible, interpretable, and data-efficient pipelines in computational physiology.

Another key aim of this thesis is to enhance the effectiveness of existing information-theoretic tools across the different stages of the machine learning pipeline. This includes enabling context-efficient estimation of information-theoretic quantities in mixed continuous–discrete settings, which are typical of classification problems involving physiological data. In addition, the thesis develops and applies information-theoretic methods capable of exploiting the high-order structure of interactions within multivariate feature spaces. In particular, information-based feature selection and feature-importance analysis are em-

ployed to quantify feature relevance and to decompose predictive information into unique, redundant, and synergistic components [97, 98, 99, 130, 131, 206]. By embedding these principles directly within the learning process, the proposed framework enables the identification of informative physiological descriptors and reveals how groups of variables jointly encode system states, improving both interpretability and robustness in physiological data analysis.

Finally, the framework is validated through multiple experimental applications addressing distinct physiological questions, including the analysis of autonomic regulation under stress [17, 95, 97, 98, 130] and the investigation of sex-related differences in cardiovascular control [96, 164]. These applications are not treated as independent case studies, but as complementary scenarios through which the different components of the proposed analytical workflow are progressively tested and evaluated. Across these scenarios, the proposed approach is used to assess the ability of information-theoretic machine learning to combine predictive accuracy, interpretability, and physiological plausibility within a coherent analytical architecture. Although cardiovascular physiology represents the primary validation domain considered in this thesis, the proposed analytical framework is general and can be extended to the study of other complex dynamical systems characterised by multivariate time-series data.

Overall, the thesis aims to demonstrate that information theory can serve not only as a set of analytical tools but also as a unifying conceptual framework for the design of interpretable machine learning pipelines in physiological signal analysis.

1.3 Scientific Contributions of the Thesis

The present thesis proposes the systematic integration of information-theoretic tools within the machine learning pipeline applied to physiological data. The main methodological and applicative contributions developed in this work are outlined in the following.

The first research theme concerns signal processing and feature extraction from physiological recordings. Although these stages rely largely on established signal processing and variability analysis techniques, the thesis contributes to their systematic integration within a unified analytical workflow for feature extraction [95, 97]. In particular, preprocessing procedures and multidomain feature extraction strategies are designed and validated to ensure the reliable estimation of cardiovascular variability descriptors suitable for subsequent information-theoretic analysis.

The second research theme concerns the development of information-theoretic approaches for feature selection. In this context, the thesis investigates a new methodology

based on Conditional Mutual Information (CMI) estimation, enabling the identification of informative and non-redundant feature subsets while accounting for nonlinear dependencies among physiological variables. In particular, the thesis introduces and evaluates the k-Nearest Neighbours Conditional Mutual Information Feature Selection (kCMI-FS) framework [98], which relies on a novel estimator of conditional mutual information specifically designed for mixed continuous–discrete feature spaces [17]. This enables reliable estimation of conditional information in physiological datasets and supports the identification of compact and informative feature subsets.

A third research theme focuses on the analysis of feature relevance through information-theoretic feature importance analysis. Within this line of research, the thesis investigates high-order information decomposition frameworks that quantify the contributions of individual features and feature groups to the predictive information about the target variable. In particular, the thesis develops the Conditional Mutual Information High-Order Feature importance (CMIHiFi) framework [99, 131, 164], which enables the decomposition of predictive information into unique, redundant, and synergistic components. These approaches allow the investigation of high-order interaction structures within multivariate physiological datasets and provide deeper insight into how groups of variables jointly encode information about the system's state.

The last research theme addresses interpretable classification strategies within the proposed analytical workflow. In this context, both traditional machine-learning models and information-theoretic approaches are considered and systematically compared. In particular, the thesis investigates the implementation of the Local Information Classifier (LIC) [130, 206] within the proposed analytical pipeline. This comparative strategy allows the methodological contributions introduced in this thesis to be validated against established techniques, ensuring that improvements in interpretability and predictive performance can be properly assessed.

These methodological developments are validated through several experimental applications involving physiological signal analysis, including stress assessment and the investigation of sex-related differences in cardiovascular regulation. These applications allow the proposed analytical workflow to be tested across different experimental conditions while highlighting the physiological interpretability of the obtained results.

The research presented in this thesis is supported by a series of peer-reviewed scientific publications reporting the conceptual development, methodological implementation, data analysis, and experimental validation of the proposed approaches. For clarity, Table 1.1 summarises the main research themes addressed, the corresponding methodological developments, and the scientific publications in which these contributions are presented.

Further details on the publications and on the specific contributions associated with them are reported in Appendix B.1.

Table 1.1: Overview of the main research themes addressed in this thesis, the corresponding methodological contributions, the thesis chapters where they are discussed, and the associated scientific publications. Further details are reported in Appendix B.1.

Research Theme	**Contribution**	**Chapter**	**Associated Publications**
Signal Processing and Feature Extraction	Design, and validation of preprocessing procedures and multidomain feature extraction strategies for cardiovascular variability analysis, enabling robust physiological descriptors suitable for information-theoretic modelling.	Ch. 2, Ch. 3	main:[97], other:[94, 95, 178]
Information-theoretic Feature Selection	Development and validation of a new CMI-based feature selection algorithm capable of identifying informative and non-redundant subsets in mixed-type physiological datasets.	Ch. 4	main:[98], other:[96]
High-order Feature Importance	Development and application of information decomposition framework (HiFi and CMIHiFi) to quantify unique, redundant, and synergistic contributions of physiological features.	Ch. 5	main:[131], other:[99, 164]
Interpretable Classification	Development, validation, and application of a classifier grounded in the information-theoretic principle of local entropy, and comparison between traditional machine-learning models	Ch. 6	main:[130], other:[94, 95, 97, 206]
Applications to Physiological Analysis	Validation of the proposed framework across multiple physiological applications, including stress assessment and the analysis of sex-related differences in cardiovascular regulation.	Ch. 7	[17, 74, 94, 95, 96, 97, 98, 99, 130, 131, 164, 178, 206]

1.4 Outline of the Thesis

The structure of this thesis mirrors the methodological progression of the proposed framework, following a logical sequence that moves from theoretical foundations (Part II) to practical applications (Part III). After the overview provided in this first chapter, the subsequent parts expand upon the concepts introduced here, guiding the reader through the stages of the *information-theoretic machine learning frameworks* pipeline and its application to physiological data.

The conceptual structure of this analytical workflow is summarised in Figure 1.1, which outlines the successive stages of the proposed information-theoretic machine learning pipeline. Each stage progressively transforms physiological data into interpretable representations by quantifying, reducing, and ultimately explaining the flow of information within the system.

Chapter 2 introduces the physiological and methodological background underlying the analysis of cardiovascular variability. It describes the nature of physiological signals and explains how the electrical, mechanical, and respiratory components of cardiovascular regulation can be monitored through time series suitable for quantitative analysis. Particular attention is devoted to signal acquisition and preprocessing, ensuring methodological rigour and reproducibility across datasets.

Chapter 3 focuses on feature extraction from physiological time series. It details how variability can be characterised in the time, frequency, and information domains, describing the physiological meaning of each class of features and explaining how these indices provide complementary insights into autonomic control. This chapter lays the foundation for the subsequent methodological developments, as the extracted features form the informational substrate of the analytical framework.

Chapter 4 addresses feature selection, discussing how redundancy and irrelevance can distort both interpretability and predictive performance. It reviews classical and modern selection methods and introduces the information-theoretic approach for identifying informative and non-redundant feature subsets through conditional mutual information estimation to capture nonlinear dependencies among physiological variables while preserving interpretability.

Chapter 5 expands the focus from selection to interpretation, presenting high-order information decomposition framework for feature-importance analysis. Here, the total predictive information shared between features and physiological states is decomposed into unique, redundant, and synergistic components, offering a nuanced understanding of how groups of features cooperate to represent autonomic regulation. This chapter exploits

information theory to aid ML model interpretation.

Chapter 6 introduces models and interpretable strategies for information-based decision-making. By quantifying local information content rather than relying on geometric separation, these classifiers link predictive performance directly to the informational structure of the data. Their performance is benchmarked against traditional models, highlighting their advantages in transparency and adaptability.

Chapter 7 presents the experimental applications of the framework to multiple physiological contexts, including the classification of autonomic responses to postural and mental stress, and the analysis of sex-related differences in cardiovascular regulation using information-theoretic decomposition. Across these applications, the framework demonstrates its generality, robustness, and physiological interpretability.

Finally, *Chapters 8* and *9* conclude the thesis by synthesising the main findings, discussing theoretical and methodological implications, and outlining directions for future research. The closing discussion highlights how information theory can serve not only as a mathematical tool but also as a conceptual bridge between data-driven modelling and physiological understanding, fostering interpretable and generalisable models in computational physiology.

2

Physiological Signals and Time Series Analysis

Physiological signals represent the measurable expression of the dynamical processes that sustain life. They originate from the continuous exchange of energy and information within and between the organs of the human body, reflecting the integrated activity of multiple regulatory subsystems that operate across diverse temporal and spatial scales. Recent advances in information-driven modelling have further reinforced this view, showing that structured variability can be quantitatively interpreted as the signature of multi-scale information flow within complex physiological networks [8, 22, 66, 67, 107]. Far from being random fluctuations, these signals exhibit structured variability that encodes the adaptive state of the underlying physiological mechanisms [22, 66, 67, 82, 118].

The quantitative study of physiological dynamics constitutes the foundation of *computational physiology*, a discipline that seeks to characterise biological function using mathematical, physical, and statistical tools [54]. Through the analysis of biosignals, it becomes possible to describe how living systems maintain homeostasis, respond to perturbations, and coordinate their functions through feedback and feedforward regulation.

Physiological signals can be recorded from a wide range of systems, including the brain (e.g. electroencephalography and functional near-infrared spectroscopy), muscles (electromyography), cardiovascular and respiratory systems (e.g. electrocardiography, photoplethysmography, blood pressure, and airflow), or even metabolic and endocrine processes. Despite their different origins, these signals share a common analytical framework: each one represents a temporal sequence of measurements whose fluctuations can reveal the structure, coordination, and adaptability of the underlying control processes.

Within this perspective, physiological variability is not regarded as noise, but as an intrinsic feature of biological organisation that carries information about the state of the organism.

In the context of computational physiology [54], physiological recordings are treated as *time series*, ordered sequences of observations sampled at regular or irregular time intervals. At a first level, these time series correspond to continuously sampled signals, such as the electrocardiogram, which can be viewed as a sequence of electrical potentials, the photoplethysmogram as a sequence of optical intensity values related to blood volume changes, or the respiratory signal as the temporal evolution of thoracic expansion or airflow. Analysis at this level captures waveform morphology, spectral content, and fast dynamical components of the underlying physiological processes. Crucially, physiological time-series analysis also relies on a second and more physiologically meaningful level of representation, obtained by extracting event-based or beat-to-beat variables from continuous signals. These derived time series are defined on the intrinsic temporal scale of the heartbeat and encode the moment-to-moment regulation of cardiovascular and autonomic function. Examples include interval-based, pressure-based, or averaged cycle-wise quantities, whose physiological interpretation and extraction procedures are detailed in the following sections. *Time-series analysis* thus provides the mathematical and statistical framework to study both continuously sampled signals and event-based physiological sequences, enabling the quantification, modelling, and interpretation of variability, rhythmicity, and regulatory responses across multiple temporal scales.

The transformation of physiological recordings into time-series representations also enables multiscale analysis. This multiscale temporal structure is a hallmark of physiological systems and underlies their ability to adapt and remain coherent across diverse conditions [22, 100]. In physiological systems, relevant information is not confined to a single characteristic timescale but is distributed over a hierarchy of temporal resolutions, reflecting the coexistence of fast regulatory mechanisms and slower integrative processes. Short-term fluctuations often reflect rapid autonomic adjustments, while slower oscillations correspond to hormonal, thermal, or circadian regulation. Modern analyses based on *Network Physiology* and information theory have demonstrated that such multiscale fluctuations reflect coordinated communication among physiological subsystems rather than isolated organ activity [131, 164].

The analysis of physiological signals, therefore, spans multiple levels of description, from local single-organ dynamics to global network coordination across organ systems. This integrative view has led to the development of Network Physiology [22, 100], a conceptual framework that describes the human organism as a complex network of interacting physiological subsystems that exchange information and collectively maintain functional

stability. Within this framework, variability is interpreted as an emergent property of coordinated interactions, rather than as an independent characteristic of individual organs. Recent studies have emphasised how these cardiovascular dynamics can be explored not only through linear time-domain and spectral methods but also by adopting complexity-based and information-theoretic perspectives [12, 147, 232, 249]. Recent frameworks extend this concept by combining network-level modelling with information-theoretic machine learning, allowing the estimation of directed and high-order dependencies among cardiovascular and respiratory variables [96, 99, 131, 164].

Moreover, in recent years, the analysis of physiological variability has been enhanced by the integration of modern sensing and analytical techniques, including photoplethysmography, wearable biosensors, and data-driven modelling frameworks that enable real-time physiological monitoring and interpretation [60, 157, 168, 197, 201, 209]. These technological advances have expanded the accessibility of cardiovascular and autonomic measurements, allowing the study of physiological dynamics in everyday environments rather than controlled laboratory conditions.

While this chapter introduces concepts that apply broadly to any physiological system, the cardiovascular and respiratory domains provide a particularly clear and accessible model for quantitative analysis. These systems are characterised by well-defined oscillatory dynamics, measurable coupling mechanisms, and clinical relevance, making them ideal testbeds for the development and validation of analytical methods. Although a wide range of physiological signals can be investigated within computational physiology, this thesis focuses primarily on cardiovascular and respiratory measurements, which represent the most widely used and clinically established markers of autonomic regulation. Other physiological signals, such as galvanic skin response or electrodermal activity, may also provide complementary information about autonomic function but are beyond the scope of the present work.

The chapter is organised as follows. The first Section 2.1 outlines the physiological foundations of cardiovascular and respiratory signals, describing how the autonomic nervous system (ANS) regulates cardiac and respiratory dynamics and how these processes manifest in measurable biosignals such as the electrocardiogram, photoplethysmogram, and respiration. The second Section 2.2 details the *signal acquisition and preprocessing pipeline*, illustrating how raw physiological recordings are transformed into artefact-free, stationary time series suitable for quantitative analysis. Together, these sections focus on the integrated cardiorespiratory system to illustrate how physiological variability can be measured, processed, and interpreted through modern computational and information-theoretic approaches.

2.1 Physiological Foundations of Cardiovascular and Respiratory Signals

Among all physiological systems, the cardiovascular and respiratory networks represent one of the most accessible and informative models for investigating autonomic regulation. The heart, blood vessels, and lungs form a tightly coupled dynamical system whose oscillations and interactions reveal the capacity of the organism to adapt to internal and external demands. Through continuous exchanges of mechanical, electrical, and chemical signals, these subsystems regulate oxygen delivery, arterial pressure, and energy balance, ensuring stability across varying conditions. Variations in heart rhythm, vascular tone, blood pressure, and respiration, collectively described as cardiovascular and respiratory variability, provide quantitative markers of autonomic balance and adaptive function [61, 105, 175, 212].

Traditionally, cardiovascular and respiratory oscillations have been analysed using time-domain and frequency-domain methods, which quantify the amplitude and spectral distribution of physiological variability and provide well-established markers of autonomic modulation [143, 175, 212]. Time-domain indices describe overall variability and short-term fluctuations, while spectral measures capture rhythmic components associated with sympathetic, parasympathetic, and baroreflex activity. These classical approaches have yielded valuable insights into autonomic regulation and remain a cornerstone of physiological signal analysis. Building upon these foundations, recent studies have shown that cardiovascular oscillations can be further characterised using information-theoretic and network-based approaches capable of describing directed and multivariate interactions among physiological variables, including heart rate, respiration, and blood pressure [20, 65, 105, 176].

The *autonomic nervous system* (ANS) acts as the primary regulatory controller of these processes, coordinating sympathetic and parasympathetic activity to maintain cardiovascular homeostasis. By modulating heart rate, vascular resistance, and respiratory rhythm, the ANS ensures appropriate control of arterial pressure, which is continuously stabilised through the baroreflex loop. The resulting temporal fluctuations in cardiac, respiratory, and pressure signals thus reflect the integration of multiple feedback pathways, central and peripheral, that sustain physiological equilibrium [66, 100].

A central tool for quantifying autonomic fluctuations is the *electrocardiogram* (ECG), which records the electrical activity generated by cardiac depolarisation and repolarisation. The R peaks of the QRS complex provide precise temporal markers for deriving *R–R intervals*, the basis of *Heart Rate Variability* (HRV), one of the most widely used indices

of autonomic modulation [142, 143, 212]. Importantly, HRV is a state-dependent measure that reflects the current physiological and pathophysiological condition of the individual, being modulated by factors such as autonomic balance, metabolic demands, ageing, and disease [212]. HRV reflects the beat-to-beat adjustments of cardiac activity mediated by sympathetic and parasympathetic inputs: high variability indicates vagal predominance and adaptive capacity, whereas reduced variability reflects sympathetic activation or reduced flexibility.

At rest, parasympathetic activity predominates, promoting slower heart rates and enhanced variability. Under stress, exercise, or orthostatic challenge, sympathetic activation accelerates the heart and reduces HRV. Classical tilt tests have demonstrated that baroreflex unloading suppresses high-frequency vagal components of HRV and increases low-frequency oscillations linked to sympathetic and baroreflex dynamics [175, 240, 241]. HRV therefore provides a non-invasive marker of autonomic balance and an integrative indicator of physiological adaptability [171, 213].

Alongside the ECG, the *Photoplethysmogram* (PPG) represents the optical recording of pulsatile blood volume changes, enabling the extraction of *Pulse Rate Variability (PRV)*. PRV shares many features with HRV but is also modulated by peripheral vascular properties such as arterial stiffness and pulse transit time [15, 97, 151, 175]. Comparative studies show strong HRV–PRV agreement at rest, with divergence under peripheral vasoconstriction, motion artefacts, or altered arterial compliance [15, 97, 151, 175]. Information-theoretic analyses have recently enabled quantitative mapping of these divergences across autonomic conditions [74, 97, 98].

Technological advances such as remote PPG and multimodal wearable sensors now enable the estimation of PRV and stress responses from ultra-short recordings or even video signals [14, 28, 70, 195], supporting continuous, unobtrusive monitoring of autonomic activity in everyday settings [55, 102, 135].

Beyond cardiovascular electrical and optical measures, *arterial pressure* (AP) signals provide a complementary and physiologically rich window into autonomic regulation. The arterial pressure waveform reflects the interaction between cardiac output, arterial stiffness, peripheral resistance, and wave reflections. From this waveform, beat-to-beat estimates of *systolic blood pressure* (SBP) and *diastolic blood pressure* (DBP) can be derived, forming two additional time series that capture pressor dynamics. Variability in SBP and DBP reflects baroreflex sensitivity, sympathetic vascular control, and mechanical properties of the arterial tree. These pressure-derived series, therefore, play a crucial role in assessing autonomic and vascular regulation, complementing HRV and PRV by providing direct markers of vascular tone and blood pressure control.

Respiration adds another layer of modulation. The cyclic influence of breathing on heart rate, accelerating during inspiration and decelerating during expiration, produces the so-called *Respiratory Sinus Arrhythmia* (RSA), a robust indicator of vagal control. RSA arises from both mechanical interactions (changes in venous return and stroke volume) and central coupling mechanisms between respiratory and cardiovascular control centres [61, 66, 185, 211]. Together with the baroreflex, these mechanisms link respiration, heart rate, and arterial pressure into an integrated feedback system that preserves homeostatic stability. In the context of this thesis, respiratory signals are analysed as modulators of cardiovascular variability, enabling the investigation of cardiorespiratory coupling and autonomic regulation across different experimental conditions.

Gender and clinical condition further modulate autonomic dynamics. Women typically exhibit higher resting heart rates but enhanced vagally mediated HRV components compared with men, reflecting sex-related differences in hormonal and endothelial regulation [4, 119, 199]. Alterations in HRV are consistently reported in cardiovascular disease, heart failure, and metabolic syndrome [9, 40, 145]. These findings highlight the complementary value of HRV, PRV, SBP variability, and DBP variability as biomarkers of autonomic function, cardiovascular risk, and stress responsiveness in both laboratory and real-world environments.

2.2 Acquisition and Preprocessing

Physiological recordings must be carefully processed to obtain artefact-free, synchronised, and stationary time series suitable for quantitative analysis. Standard protocols for heart rate variability (HRV) analysis were first established by the Task Force (1996) and remain the reference framework for contemporary studies [143, 212]. More recent works have refined these procedures to accommodate short-term, multimodal, and wearable applications [62, 97, 151, 175, 209, 212, 222]. Advances in signal processing have improved the reliability of peak detection and artefact correction in noisy conditions or when using wearable devices, enabling robust estimation of cardiac, respiratory, and vascular variability from short-term or non-contact recordings [62, 84, 116, 171].

The preprocessing of physiological recordings follows a sequential pipeline aimed at transforming raw biosignals into reliable beat-to-beat time series suitable for quantitative analysis, as illustrated in Fig. 2.1. This pipeline typically includes signal acquisition (ECG, PPG, AP, BREATH), filtering and artefact suppression, fiducial point detection, artefact correction, resampling and normalisation, and the selection of stationary analysis segments. While the first stages are required to obtain a reliable beat-to-beat series, the final step of segment selection is not always necessary and is therefore represented with dashed

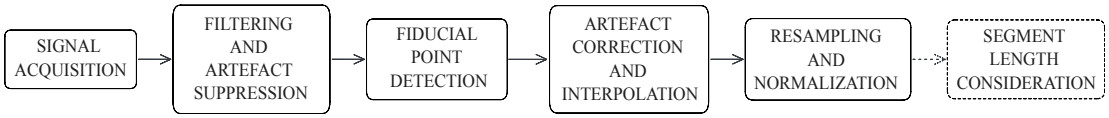


Figure 2.1: Overview of the preprocessing pipeline used to transform raw physiological recordings into beat-to-beat time series suitable for quantitative analysis. The workflow includes signal acquisition, filtering and artefact suppression, fiducial point detection, artefact correction, resampling and normalisation, and the selection of stationary segments used for feature extraction and information-theoretic analysis.

lines in Fig. 2.1, as it depends on the specific experimental design and analysis objectives. Reliable extraction of beat-to-beat time series from raw physiological recordings is not trivial, as these signals are often affected by noise, motion artefacts, sensor displacement, or peripheral physiological changes. Such disturbances may distort waveform morphology or introduce spurious peaks, potentially leading to incorrect interval estimation and artificial variability. For this reason, preprocessing procedures must carefully balance noise suppression with the preservation of physiologically meaningful signal components. Accurate preprocessing is essential because errors introduced at this stage may propagate to the subsequent estimation of variability indices and information-theoretic measures derived from physiological time series.

1. Signal acquisition. Physiological signals are typically acquired using non-invasive sensors that capture the electrical, optical, mechanical, or pressure-related manifestations of cardiovascular and respiratory activity. The *Electrocardiogram* (ECG) records cardiac electrical potentials generated by myocardial depolarisation and repolarisation. Surface electrodes detect voltage differences originating from the propagation of the cardiac action potential, producing a characteristic waveform of P wave, QRS complex, and T wave. Sampling rates between 250 and 1000 Hz ensure precise R-peak detection and reduce aliasing of high-frequency components. The *Photoplethysmogram* (PPG) is an optical signal reflecting changes in peripheral blood volume. Light emitted from an LED is absorbed or reflected by vascular tissue, and the resulting waveform includes key features such as the systolic peak and the dicrotic notch, both linked to arterial elasticity and vascular compliance [15, 149, 151]. Sampling frequencies of 100–500 Hz are typically sufficient when combined with appropriate filtering to mitigate motion and ambient-light artefacts [62, 231]. The *Respiratory* (BREATH) activity can be measured using airflow sensors, respiratory inductive plethysmography, or strain gauges, or estimated indirectly using ECG-derived or PPG-derived respiration algorithms [13, 61]. Typical sampling rates range from 25 to 100 Hz. The *Arterial pressure signals* provide a mechanical representation of blood pressure dynamics, capturing the interaction between cardiac ejection, arterial stiffness, and vascular resistance. Continuous non-invasive arterial pressure can be

acquired using tonometric devices, volume-clamp methods (e.g., Finapres), applanation tonometry, or, more intermittently, via oscillometric cuffs. From the pressure waveform, beat-to-beat values of *systolic blood pressure (SBP)* and *diastolic blood pressure (DBP)* are extracted, forming two additional time series that complement HRV and PRV by providing direct information on vascular control, baroreflex function, and autonomic modulation of arterial tone. Sampling rates of 100–500 Hz are commonly used for continuous pressure recordings to preserve the morphology of the pressure upstroke and dicrotic notch. Synchronised acquisition of ECG, PPG, respiration, and arterial pressure is essential for constructing multimodal datasets that capture central cardiac control, peripheral vascular modulation, and pressory dynamics.

2. *Signal filtering and artefact suppression.* The recorded signals are affected by various sources of noise, including baseline wander, muscular interference, and motion artefacts. In particular, baseline wander, electrode motion, muscle activity, and ambient-light interference represent common sources of contamination in ECG and PPG recordings. If not properly addressed, these disturbances may lead to misidentification of physiological events or to biased estimates of cardiovascular variability indices. Band-pass filtering is therefore applied according to the spectral characteristics of each signal: 0.5–40 Hz for ECG, 0.3–15 Hz for PPG, and below 1 Hz for BREATH. In line with established recommendations for cardiovascular signal processing, fourth-order Butterworth or Chebyshev type II filters are commonly adopted, as they provide a suitable trade-off between sharp frequency selectivity and minimal waveform distortion [134]. Notch filters at 50 or 60 Hz are used to suppress power-line interference [84]. In wearable or ambulatory applications, adaptive and wavelet-based filtering techniques are increasingly used, sometimes in conjunction with accelerometer signals, to identify and correct motion-induced artefacts [62, 231].

3. *Fiducial point detection.* Once filtered, fiducial points are detected in each signal: R peaks in the ECG and P peaks in the PPG. Accurate identification of fiducial points is critical, as even small detection errors can propagate to the derived interval series and introduce spurious fluctuations in variability measures. Robust detection algorithms, therefore, combine morphological analysis with adaptive thresholds and temporal constraints to ensure physiologically plausible beat sequences. The classical Pan–Tompkins algorithm and its more recent derived implementations remain among the most widely used methods for R-peak detection, combining differentiation, integration, and adaptive thresholding [84, 222]. For PPG signals, peak detection often relies on derivative analysis, adaptive thresholds, or wavelet-based approaches to locate the systolic maxima robustly, even under variable waveform morphology [62, 150]. Respiratory cycles are identified using zero-crossing or envelope detection methods applied to airflow, thoracic expansion,

or derived respiratory signals [13]. In the latter case, a respiratory time series is commonly derived by sampling the respiratory signal at the occurrence of ECG R peaks, yielding a beat-synchronous respiratory sequence that is temporally aligned with RR intervals and facilitates joint cardiorespiratory analysis. From these landmarks, discrete sequences of RR, PP, and inter-breath (RESP) intervals are obtained, representing the instantaneous frequency of each physiological rhythm. SBP and DBP are extracted from local maxima and minima of each pressure (AP) beat, while additional fiducial points (e.g., the dicrotic notch) support the evaluation of arterial compliance and wave reflection.

4. Artefact correction and interpolation. Physiological recordings may include ectopic beats, missed detections, or transient artefacts. These outliers are identified using standard physiological and statistical criteria, such as absolute interval limits and deviations exceeding a fixed multiple of the local standard deviation (e.g., three standard deviations from the median or mean), and are subsequently corrected by interpolation within physiologically plausible bounds, ensuring continuity of the interval time series [62, 97, 175]. Quality-control indices based on signal morphology and local signal-to-noise ratio are commonly used to identify and exclude corrupted segments [62, 231].

5. Resampling and normalisation. Because inter-beat, inter-pulse, and beat-to-beat blood pressure series are irregularly spaced, they are resampled to obtain uniformly spaced time series for spectral or information-theoretic analysis. The choice of the resampling frequency represents an important methodological parameter, as it determines the temporal resolution of the derived series and may influence the estimation of spectral or information-theoretic indices. Cubic-spline interpolation followed by resampling at 1 or 4 Hz is commonly used [143, 240]. Normalisation techniques (e.g., z-score transformation) reduce inter-individual amplitude differences and emphasise relative fluctuations. To ensure stable statistical properties, detrending and visual inspection are used to verify weak stationarity across analysed epochs [12, 249].

6. Segment length considerations. The temporal length of the analysed segments represents a critical methodological choice, as different time scales capture distinct aspects of autonomic regulation. Long-term recordings (typically 24 h) are considered the reference standard for HRV analysis, as they encompass circadian rhythms, sleep–wake cycles, and slow regulatory processes, providing an assessment of autonomic function [143, 212]. Short-term recordings (approximately 5 min) are widely adopted as a practical compromise between physiological representativeness and statistical robustness. Although they do not capture very slow dynamics, 5-minute segments have been shown to reliably reflect autonomic tone and to discriminate between physiological and pathophysiological states across a broad range of conditions, which explains their extensive use in experimental and clinical studies [41, 212]. More recently, ultra-short-term recordings (from tens of seconds

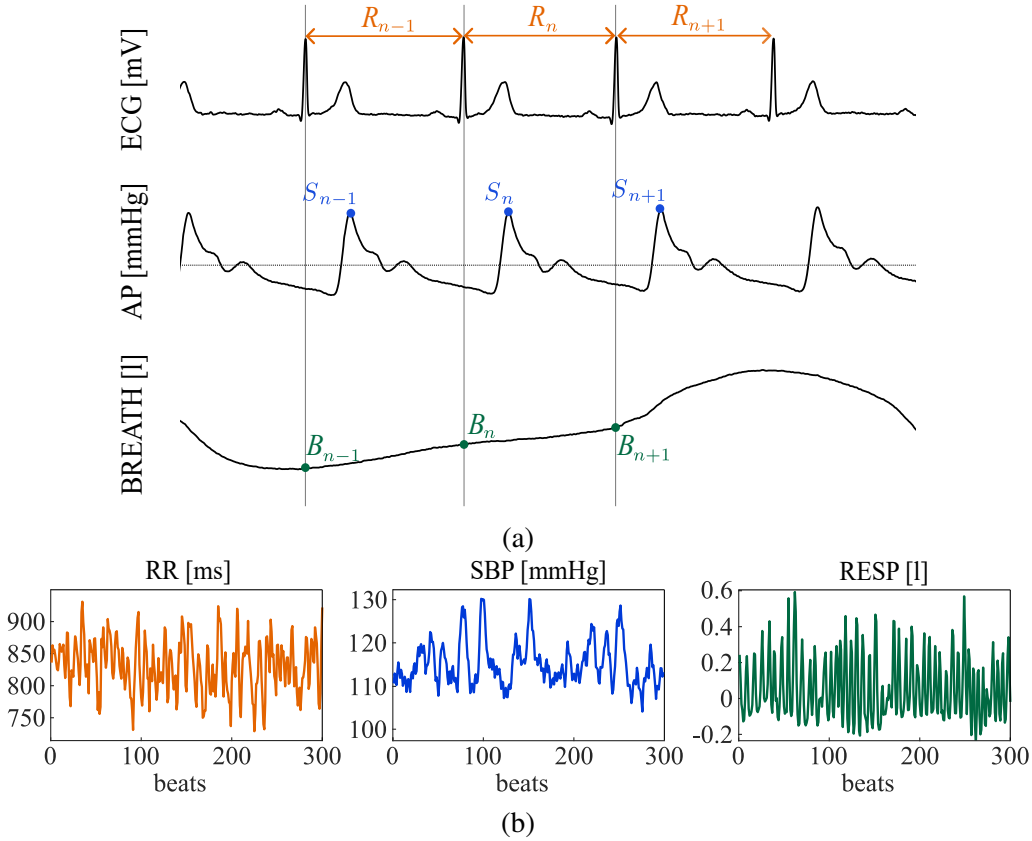


Figure 2.2: Example of signals and time series extraction. Top panels (a) show representative *ECG* traces recorded, with detected R peaks marked with dots, the *AP* waveform, and the *BREATH* signal with the corresponding R peaks marked with dots. From these peak positions, beat-to-beat interval series were derived: R-R intervals (*RR*) from the ECG, the systolic pressure (*SBP*) from the Ap, and the RESP time series from the respiratory signal. The bottom panel (b) shows the resulting time series.

to about 1 min) have gained increasing attention in wearable, ambulatory, and real-time monitoring scenarios, where prolonged acquisitions are often impractical [28, 116, 240]. When carefully preprocessed and interpreted with awareness of their limitations, even ultra-short segments can support reliable estimation of selected HRV, PRV, and beat-to-beat SBP/DBP indices, particularly for tracking relative changes rather than absolute autonomic levels [151, 175, 213].

Although these recordings contain fewer data points, their appropriate preprocessing and synchronisation make them suitable for assessing rapid autonomic adaptations and transient physiological responses [151, 175]. The outcome of this preprocessing pipeline is a set of synchronised, artefact-free, and stationary time series that accurately represent cardiovascular and respiratory dynamics, providing a robust basis for subsequent feature extraction and information-theoretic analysis. An example of physiological signals and the corresponding beat-to-beat time series extracted from them is shown in Fig. 2.2.

In summary, this chapter presented the physiological and methodological background underlying the analysis of cardiovascular and respiratory dynamics. It described how electrical, optical, and mechanical signals, such as the ECG, PPG, AP and respiratory recordings, can be acquired, filtered, and transformed into high-quality time series that faithfully represent autonomic activity. Through careful preprocessing, artefact correction, and synchronisation, these signals become suitable for quantitative investigation of the temporal organisation and coordination that characterise autonomic regulation. By establishing a unified view of physiological signal generation and processing, the chapter laid the foundation for the analytical developments that follow.

While the signal processing procedures described in this chapter largely rely on established protocols for physiological signal analysis, they provide the necessary methodological foundation for the information-theoretic developments introduced in the following chapters. In particular, the extracted cardiovascular and respiratory time series constitute the input variables for the feature extraction, feature selection, and information decomposition frameworks developed and evaluated throughout this thesis.

Part II

Information-Theoretic Framework

3

Feature Extraction

The process of extracting meaningful features from physiological time series lies at the heart of computational physiology and biomedical signal analysis [12, 20, 65, 66, 105, 143, 147, 175, 212, 216, 232, 249]. *Features* are numerical descriptors designed to summarise the dynamical behaviour of physiological signals, such as heart rate, pulse, or arterial pressure, in a compact and physiologically interpretable form [40, 97, 142, 143, 171, 175]. The extraction transforms complex raw data into structured quantitative representations suitable for statistical modelling and machine learning [66, 97, 105, 175, 216].

In cardiovascular research, the analysis of beat-to-beat fluctuations in heart rate and pulse period has long been recognised as a powerful non-invasive tool for assessing ANS regulation [66, 105, 142, 186, 212]. These fluctuations, collectively known as HRV and PRV, provide insights into the balance and adaptability of sympathetic and parasympathetic modulation. However, the physiological signals recorded from humans are inherently non-linear, noisy, and influenced by multiple feedback mechanisms acting over different time scales. As a result, the extraction of robust and physiologically interpretable features becomes a crucial step to reveal underlying regulatory mechanisms and to distinguish between physiological and pathological conditions [12, 16, 65, 175, 216, 232, 249]. Recent methodological contributions have extended classical feature extraction by embedding information-theoretic and network-based criteria, thus capturing multi-organ dependencies and synergistic effects among cardiovascular variables [97, 98, 131, 164].

Feature extraction (FE) serves as a bridge between raw signal dynamics and higher-level physiological interpretation. By condensing complex temporal information into a set of statistical, spectral, and information-theoretic features, it reduces data dimensionality while preserving the relevant informational content [66, 67, 105, 249]. Each feature cap-

tures a specific aspect of cardiovascular dynamics, such as the magnitude of variability, rhythmic oscillations at distinct frequencies, or the degree of regularity and complexity of the underlying process. This representation enables a structured mapping from raw temporal sequences to a feature space that encapsulates the functional behaviour of cardiovascular control systems, providing a foundation for subsequent feature selection, classification, and physiological interpretation [76, 216]. The extraction of multidomain features represents not only a preprocessing step but also a conceptual bridge between raw physiological dynamics and interpretable learning, as recent information-driven models have shown [74, 98, 99, 164].

From a theoretical standpoint, feature extraction links quantitative time-series analysis with the physiological interpretation of the biological function. It provides a mathematical framework for describing how variability arises from interacting feedback loops within the autonomic nervous system, baroreflex control, and cardiorespiratory coupling [20, 66, 67, 105, 176]. Consequently, the set of features derived from RR or PP intervals can be regarded as a multidimensional fingerprint of autonomic regulation, summarising how sympathetic and parasympathetic branches cooperate to maintain homeostasis [12, 16, 66, 76, 105].

This chapter presents a structured overview of the main features typically derived from cardiovascular time series, with particular attention to their mathematical formulation, physiological rationale, and interpretive value. We distinguish three broad categories of descriptors:

- **Time-domain features** (Sec. 3.1), which quantify the overall magnitude of heart rate fluctuations and their short-term variability directly from the time series [41, 66, 142, 143, 171, 175, 212];
- **Frequency-domain features** (Sec. 3.2), which decompose the variability into oscillatory components reflecting distinct physiological rhythms such as respiration and baroreflex modulation [66, 67, 108, 142, 175, 212];
- **Information-theoretic features** (Sec. 3.3), which characterise the complexity, predictability, and dynamical organisation of cardiovascular signals from a probabilistic perspective [12, 16, 66, 122, 232, 249].

Each category offers a distinct but complementary view of autonomic regulation, capturing both linear and non-linear properties of cardiac control. Together, these feature families provide a comprehensive quantitative framework for assessing cardiovascular variability and for constructing models aimed at the classification and interpretation of physiological states in both experimental and clinical contexts [12, 67, 105, 232].

To define the adopted measures, let us first denote X_n as the random variable obtained by sampling the stochastic process X at the present time n. Under the assumptions of stationarity and ergodicity, each finite-length time series can be regarded as a realisation of the underlying process, so that statistical properties can be consistently estimated from a single realisation of X. Moreover, in order to approximate the infinite past history of the process, it is customary to assume that X behaves as a Markov process of order m, whereby the relevant dependence of the present state X_n on its history can be captured by the finite-dimensional vector $X_n^m = [X_{n-1}, X_{n-2}, \ldots, X_{n-m}]$ describing its past m states.

3.1 Time-Domain Features

Time-domain indices provide the most direct description of heart rate variability and are computed from the raw sequence of inter-beat intervals without signal transformation [40, 66, 142, 143, 171, 212]. They quantify overall variability, beat-to-beat fluctuations, and the statistical distribution of cardiac intervals [67, 175, 232]. Measures such as SDNN, RMSSD, and pNN50 are widely recognised as sensitive markers of autonomic balance and parasympathetic modulation [40, 66, 97, 142, 212]. Time-domain features include the MEAN, MeanHR, SDNN, VAR, RMSSD, pNN50, and cVar, which are defined respectively as follows [97, 142, 143, 175, 212]:

- **Mean Inter-beat Interval (MEAN)**:

$$MEAN = \frac{1}{N}\sum_{n=1}^{N} X_n, \tag{3.1}$$

  where X_n represents the n-th RR (or PP) interval and N is the total number of analysed heartbeats. The mean inter-beat interval provides a fundamental estimate of the cardiac period and reflects the overall balance between sympathetic and parasympathetic regulation of heart rate. Longer mean intervals indicate predominant parasympathetic activity (lower heart rate), whereas shorter mean intervals reflect sympathetic dominance [212].
- **Mean Heart Rate (MeanHR)**:

$$MeanHR = \frac{60}{MEAN}, \tag{3.2}$$

  corresponding to the average number of heartbeats per minute. As heart rate and inter-beat intervals are inversely related, MeanHR serves as a complementary measure of autonomic balance, with higher rates indicating increased sympathetic activation or reduced vagal tone [212].

- **Standard Deviation of NN Intervals (SDNN)**:

$$SDNN = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(X_n - MEAN)^2}, \tag{3.3}$$

which quantifies the dispersion of all normal-to-normal (NN) intervals, i.e., excluding abnormal and ectopic beats. SDNN represents the total variability of the heart period over the recording segment and incorporates both sympathetic and parasympathetic influences. When computed from 24-hour recordings, it reflects slower fluctuations associated with circadian rhythms, thermoregulation, and hormonal modulation, whereas in shorter recordings (e.g., 5 min) it is primarily influenced by parasympathetically mediated respiratory sinus arrhythmia [66, 142, 212]. Low SDNN values are indicative of reduced autonomic flexibility and are associated with increased cardiovascular morbidity and mortality [40, 171].

- **Variance (VAR)**:

$$VAR = SDNN^2, \tag{3.4}$$

representing an alternative statistical formulation of overall variability. While SDNN is the most commonly reported measure, variance conveys equivalent information regarding the total spread of the NN intervals [66, 232].

- **Root Mean Square of Successive Differences (RMSSD)**:

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N-1}(X_{n+1} - X_n)^2}, \tag{3.5}$$

which expresses the root mean square of the differences between successive NN intervals. RMSSD is particularly sensitive to rapid, beat-to-beat variations and provides a robust index of parasympathetic (vagal) modulation of heart rate [66, 212]. It reflects the amplitude of respiratory sinus arrhythmia and, because of its stability and independence from slow trends, is widely regarded as a reliable short-term marker of vagal tone [105].

- **Percentage of Successive NN Intervals Differing by More than 50 ms (pNN50)**:

$$pNN50 = \frac{100}{N-1}\sum_{n=1}^{N-1} I(|X_{n+1} - X_n| > 50), \tag{3.6}$$

where $I(\cdot)$ is the indicator function. This index represents the proportion of adjacent NN intervals that differ by more than 50 ms and, like RMSSD, predominantly reflects short-term parasympathetic modulation. The pNN50 is strongly correlated

with both RMSSD and the high-frequency (HF) component of spectral HRV. A decrease in pNN50 denotes reduced vagal activity and diminished cardiac adaptability to physiological demands [212].

- **Coefficient of Variation (cVar)**:

$$cVar = \frac{SDNN}{MEAN}, \tag{3.7}$$

which normalises the standard deviation by the mean inter-beat interval. The coefficient of variation expresses the magnitude of variability relative to the average heart period, allowing comparison of HRV levels across individuals or conditions with different mean heart rates. From a physiological perspective, lower cVar values indicate a more rigid, less adaptable cardiac control, whereas higher values are associated with flexible and efficient autonomic regulation.

Time-domain features thus provide information on cardiac autonomic regulation. Measures such as SDNN and VAR reflect global variability integrating both sympathetic and parasympathetic contributions, whereas RMSSD and pNN50 capture high-frequency oscillations primarily driven by vagal activity. The MEAN, MeanHR, and cVar offer a summary of central tendency and relative dispersion, supporting the interpretation of HRV in terms of overall autonomic balance and cardiovascular adaptability.

3.2 Frequency-Domain Features

Frequency-domain features aim to decompose the variability into oscillatory components that reflect underlying physiological rhythms, such as breathing or vasomotor activity [66, 67, 105, 175]. This decomposition is typically achieved through spectral analysis, either using Fourier transforms or autoregressive models [66, 67, 105, 175, 232].

Prior to the computation of frequency- and information-domain features, time series undergo a preprocessing stage aimed at enhancing stationarity. Slow baseline fluctuations are attenuated using a zero-phase high-pass IIR filter with a cutoff frequency of 0.015 Hz [105, 175], and the series is subsequently demeaned by subtracting its average value. Stationarity is assessed by verifying weak stationarity conditions and by checking the stability of the mean and variance across shorter sub-windows. Spectral analysis of cardiovascular variability can be performed using both nonparametric and parametric approaches. Nonparametric methods, such as the periodogram or Welch's averaged periodogram, estimate the power spectral density directly from the Fourier transform of the signal and are widely used because of their simplicity and minimal modelling assumptions. However, their frequency resolution and robustness can be limited when applied to short or noisy

physiological recordings. For this reason, and in line with previous studies [66, 175, 232], spectral analysis in this thesis is performed using a parametric autoregressive (AR) modelling approach, which provides smoother spectral estimates and improved resolution in short data segments. Specifically, the preprocessed time series is modelled as:

$$X_n = \sum_{m=1}^{p} a_m X_{n-m} + W_n, \tag{3.8}$$

where the coefficients a_m describe the linear dependence of the present observation of the time series, X_n, on its previous samples X_{n-m}, with delays $m = 1, \ldots, p$. The parameter p denotes the model order, while W_n represents a Gaussian white noise process with zero mean and variance σ_W^2. The model parameters are estimated using the ordinary least squares approach, fixing the maximum model order to $p = 10$, according to previous studies [66, 175].

Spectral analysis allows the variability of the beat-to-beat series to be decomposed into oscillatory components occurring at different frequencies, each associated with specific physiological mechanisms. The power spectral density (PSD) of the autoregressive process is then derived in the frequency domain from the estimated AR coefficients as:

$$P(\Omega) = \frac{\sigma_W^2}{|A(\Omega)|^2}, \tag{3.9}$$

where $A(\Omega) = 1 - \sum_{m=1}^{p} a_m e^{-j\Omega m}$ is the frequency-domain representation of the autoregressive polynomial, with $\Omega = 2\pi f T_s$ denoting the normalised angular frequency, T_s the sampling period, and $j = \sqrt{-1}$.

The power spectral density (PSD) of the inter-beat interval time series is computed and integrated over the conventional frequency bands [66, 142, 212]: very low frequency (VLF: 0.003–0.04 Hz), low frequency (LF: 0.04–0.15 Hz), and high frequency (HF: 0.15–0.4 Hz). It is worth noting that the quantitative estimation of spectral indices may depend on several methodological choices adopted during preprocessing and spectral modelling. Parameters such as the resampling frequency of the beat-to-beat series, the length of the analysed segment, and the order of the autoregressive model may influence the estimated spectral distribution. In particular, parametric approaches such as AR modelling generally provide smoother spectra and improved frequency resolution for short physiological recordings, whereas nonparametric Fourier-based methods rely on fewer modelling assumptions but may exhibit lower resolution when applied to short data segments. Consequently, while the LF and HF bands provide a physiologically meaningful representation of autonomic oscillations, their absolute power values and

relative contributions may vary depending on the adopted estimation framework and parameter configuration.

The VLF component is generally associated with long-term regulatory mechanisms, including thermoregulation, hormonal influences, and endothelial activity [212]. However, reliable estimation of VLF power requires long-duration recordings, typically several hours or 24-hour segments, and its physiological interpretation in short-term recordings remains limited and often ambiguous. The LF component reflects a combination of sympathetic and parasympathetic influences and is strongly modulated by the baroreflex control of heart rate [105, 175]. While LF power can be meaningfully assessed in short-term (~5-minute) recordings, its interpretation as a marker of sympathetic activity alone is controversial, as it is sensitive to both autonomic branches and to experimental conditions such as posture and breathing rate. The HF component, also referred to as the respiratory band, corresponds to respiratory sinus arrhythmia (RSA) and is mediated predominantly by parasympathetic (vagal) activity [212]. HF power can be robustly estimated even in relatively short segments, provided that respiration is within the canonical frequency range and reasonably stable, although ultra-short recordings may still suffer from reduced spectral resolution.

In addition to absolute spectral powers, normalized units (LFn, HFn) are derived to express the relative contribution of each band with respect to total power in the LF and HF ranges, thus reducing inter-individual differences in total variability. The LF/HF ratio is also computed as an index of sympathovagal balance, although its interpretation remains debated because of the complex and non-linear interactions between autonomic branches [12, 66, 67, 108].

Overall, spectral features offer a physiologically grounded representation of autonomic dynamics. They enable differentiation of long-term regulatory influences (VLF), baroreflex and mixed autonomic oscillations (LF), and respiratory–vagal modulation (HF). Frequency-domain analysis, therefore, complements time-domain measures by revealing the rhythmic structure of autonomic control and by providing sensitive markers for experimental manipulations such as rest, stress, and paced breathing [66, 67, 105, 175, 232].

3.3 Information-Domain Features

Traditional features capture the amplitude and frequency structure of variability [12, 16, 66, 67, 122, 232, 249]. Information-theoretic features go one step further, quantifying the predictability, complexity, and dependency structure within the time series. These features are motivated by the idea that healthy systems tend to be dynamically complex,

neither too regular nor too random. Loss of this complexity has been linked to ageing, stress, and disease [97, 105, 175]. Recent studies have emphasised that these information-domain measures serve as quantitative markers of adaptive complexity, revealing how information is generated, stored, and transferred within cardiovascular regulation [17, 98, 164]. In particular, the integration of conditional and self-entropy measures within machine learning workflows have improved the physiological interpretability of stress and autonomic classification models [74, 99].

To characterise the dynamical complexity of cardiovascular and respiratory signals, three information-theoretic features are computed: entropy (H), conditional entropy (CE), and self-entropy (SE). Entropy-based measures, such as CE, have been shown to be sensitive to varying levels of sympathetic activity during postural stress, effectively capturing changes in cardiac system dynamics associated with different physiological and clinical conditions [12, 16, 16, 66, 175, 232, 249].

The Entropy quantifies the information carried by the stationary system at the current state and is defined as follows:

$$H = H(X_n) = -\mathbb{E}[\log p(x_n)], \tag{3.10}$$

where $\mathbb{E}[\cdot]$ represents the expectation operator, $p(\cdot)$ is the probability density function, and x_n is a realization of X_n. To have insight on the dynamical evolution of the system, it is possible to quantify the new information carried by the present state X_n which cannot predicted from its past state X_n^m, through the Conditional Entropy (CE) and the information carried by the present of the target process that can be predicted by its own past, through the Self Entropy (SE) [16, 66, 232, 249]:

$$CE = H(X_n \mid X_n^m) = H(X_n, X_n^m) - H(X_n^m) = -\mathbb{E}\left[\log \frac{p(x_n, x_n^m)}{p(x_n^m)}\right], \tag{3.11}$$

$$SE = H - CE = H(X_n) - H(X_n \mid X_n^m) = \mathbb{E}\left[\log \frac{p(x_n, x_n^m)}{p(x_n)p(x_n^m)}\right]. \tag{3.12}$$

The linear estimator represents a model-based approach relying on the assumption that the analysed variables follow a Gaussian distribution. Given a generic d-dimensional zero-mean Gaussian variable X_n^d, its probability density function is:

$$p(x_n^d) = \frac{1}{\sqrt{(2\pi)^d |\Sigma_{X^d}|}} e^{\left(-\frac{1}{2} x_n^{d\top} \Sigma_{X^d}^{-1} x_n^d\right)}, \tag{3.13}$$

where $\Sigma_{X^d} = \mathbb{E}[X_n^d X_n^{d\top}]$ is the covariance matrix of X_n^d. For a stationary zero-mean Gaussian process $\{X_n\}$, the linear parametric representation is given by an AR model of order m as in Eq.3.8. The entropy of the past vector X_n^m and of $X_{n+1}^{m+1} = [X_n, X_n^m]$ are computed as:

- **Entropy - linear estimation (H_{lin})**:

$$H_{\text{lin}} = \tfrac{1}{2}\ln(2\pi e|\hat{\Sigma}_X|), \tag{3.14}$$

  where $\hat{\Sigma}_X$ is the estimated covariance matrix. Then, the conditional entropy is expressed in terms of the variance of the AR prediction error:

- **Conditional Entropy - linear estimation (CE_{lin})**:

$$CE_{\text{lin}} = \tfrac{1}{2}\ln(2\pi e|\hat{\Sigma}_X|) - \tfrac{1}{2}\ln(2\pi e|\hat{\Sigma}_{X_n^m}|) = \tfrac{1}{2}\ln(2\pi e\,\hat{\sigma}_W^2), \tag{3.15}$$

  where $\hat{\sigma}_W^2$ is the prediction error variance obtained after identifying the regression coefficients of the AR model via the least squares method, that coincides with the partial covariance of $\hat{\Sigma}_X$ given $\hat{\Sigma}_{X_n^m}$, being $\hat{\Sigma}_X$ the estimated covariance matrix.

- **Self Entropy - linear estimation (SE_{lin})**:

$$SE_{\text{lin}} = \tfrac{1}{2}\ln(2\pi e|\hat{\Sigma}_X|) - \tfrac{1}{2}\ln(2\pi e\,\hat{\sigma}_W^2) = \tfrac{1}{2}\ln\left(\frac{|\hat{\Sigma}_X|}{\hat{\sigma}_W^2}\right), \tag{3.16}$$

A widely used model-free approach for entropy estimation is based on nearest-neighbour (kNN). The underlying idea is that the probability density around a given sample can be approximated from the distance to its nearest neighbours.

Consider $N - d + 1$ observations of a d-dimensional continuous random variable X_n^d. For a fixed number k of neighbour samples, the local probability density around a point x_n^d can be expressed as [16, 122]:

$$p(x_n^d) = \frac{k}{(N-d)\,c_{d,L}\,\epsilon_{n,k}^d}, \tag{3.17}$$

where $\epsilon_{n,k}$ is the distance (according to the maximum norm) between x_n^d and its k-th nearest neighbour, and $c_{d,L}$ is the volume of the d-dimensional unit ball under the chosen norm.

To avoid biases introduced by combining entropies of variables with different dimensions (e.g. X_n^m), we adopt the formulation of Kraskov, Stögbauer, and Grassberger [122]. In this setting, the entropy and the conditional entropy are estimated as

- **Entropy - kNN estimation (H_{knn})**:

$$H_{\text{knn}} = \psi(N) - \psi(k) + \frac{1}{N}\sum_{n=1}^{N}(\ln \epsilon_n), \tag{3.18}$$

- **Conditional Entropy - kNN estimation (CE_{knn})**:

$$CE_{\text{knn}} = -\psi(k) + \frac{1}{N-m}\sum_{n=1}^{N-m}\Big(\ln \epsilon_{n,k} + \psi(N_{X_n^m}+1)\Big), \tag{3.19}$$

  where $\psi(\cdot)$ denotes the digamma function, ϵ_n is twice the distance from (x_n) to its k-th nearest neighbor in the one-dimensional space $\epsilon_{n,k}$ is the maximum-norm distance of the point (x_n, x_n^m) to its k-th nearest neighbour in the $(m+1)$-dimensional space, and $N_{X_n^m}$ is the number of points whose distance from x_n^m is smaller than $\epsilon_{n,k}/2$ in the m-dimensional space.

- **Self Entropy - kNN estimation (SE_{knn})**:

$$SE_{\text{knn}} = \psi(N) + \psi(k) - \frac{1}{N-m}\sum_{n=1}^{N-m}\Big(\psi(N_{X_n^m}+1)\Big) - \frac{1}{N}\sum_{n=1}^{N}\Big(\psi(N_X)\Big), \tag{3.20}$$

The two key parameters of the estimator are the embedding dimension m, which controls the reconstruction of the process history, and the number of neighbours k, which determines the trade-off between variance (too small k) and bias (too large k). In this thesis, the number of neighbours and the embedding dimension were set to $k = 10$ and $m = 2$, respectively [16, 16, 175, 249]. The choice of $m = 2$ balances the ability to capture dynamic information with the limitations of short physiological time series, while the selection of $k = 10$ has been shown to provide stable and reliable entropy estimates in HRV analysis.

These measures reveal temporal complexity and offer insights into autonomic adaptability. Their estimation is computationally more intensive but especially useful in short, non-stationary signals. Together, these measures characterise complementary aspects of cardiovascular dynamics: entropy (H) reflects the overall variability and richness of the system's states, conditional entropy (CE) quantifies the degree of unpredictability and external influence acting on the system beyond its own past, while self-entropy (SE) captures the intrinsic temporal organisation and memory of the physiological process. In this sense, higher SE is associated with stronger internal regulatory structure, whereas increased CE indicates a greater contribution of external or unmodelled influences, providing a nuanced description of autonomic adaptability and control mechanisms.

Tables 3.1-3.3 summarise the set of HRV features described in the previous sections,

providing a concise overview of the features adopted in the time, frequency, and information domains. These measures capture complementary aspects of cardiac dynamics, from the overall amplitude of variability to the distribution of oscillatory power and the complexity of temporal dependencies.

Table 3.1: Time-domain features.

Parameter	Unit	Description
MEAN	ms	Mean inter-beat interval; represents the average cardiac cycle length and overall autonomic balance.
MeanHR	bpm	Mean heart rate; inverse of the mean inter-beat interval, reflecting the combined sympathetic and parasympathetic drive.
SDNN	ms	Standard deviation of NN intervals; quantifies the total variability of heart rate over the recording period.
VAR	ms^2	Variance of NN intervals; equivalent to the squared SDNN, representing overall variability amplitude.
RMSSD	ms	Root mean square of successive RR interval differences; a robust marker of short-term, vagally mediated variability.
pNN50	%	Percentage of successive NN intervals differing by more than 50 ms; index of parasympathetic modulation.
cVar	–	Coefficient of variation (SDNN/MEAN); expresses variability relative to mean heart period, allowing inter-subject comparison.

Table 3.2: Frequency-domain features.

Parameter	Unit	Description
VLF power	ms^2	Absolute power of the very-low-frequency band (0.003–0.04 Hz); reflects slow regulatory processes such as thermoregulation and hormonal modulation.
LF power	ms^2, %	Absolute or relative power of the low-frequency band (0.04–0.15 Hz); associated with baroreflex and mixed autonomic modulation.
HF power	ms^2, %	Absolute or relative power of the high-frequency band (0.15–0.4 Hz); corresponds to respiratory sinus arrhythmia and parasympathetic activity.
LF/HF ratio	%	Ratio of LF-to-HF power; commonly interpreted as an indicator of sympathovagal balance (with caution).
Total Power	ms^2	Total spectral variance of NN intervals within the 0–0.4 Hz range, representing overall autonomic modulation.
$\mathbf{f}_{HF}$	Hz	Frequency of the HF peak; identifies the dominant respiratory frequency.

Feature Extraction transforms physiological time series into quantitative descriptors that encapsulate the dynamics of autonomic regulation. By jointly analysing time-domain, frequency-domain, and non-linear features, it becomes possible to construct a multidimensional representation of cardiovascular control. Each feature captures a complementary

Table 3.3: Information-domain features.

Parameter	Unit	Description
H	nat or bit	Entropy; measures the average information content or uncertainty of the system's instantaneous state.
CE	nat or bit	Conditional entropy; quantifies the new information carried by the present state that cannot be predicted from the past, reflecting system unpredictability.
SE	nat or bit	Self entropy; represents the intrinsic information generation of the process, indicating its degree of dynamical dependence and regularity.

aspect of cardiac variability, from overall fluctuation amplitude and rhythmic spectral components to complex, scale-invariant patterns of self-regulation. Together, these metrics provide a comprehensive and physiologically grounded framework for assessing autonomic function, facilitating the interpretation, comparison, and classification of cardiovascular dynamics in both research and clinical applications.

Beyond its descriptive value, feature extraction serves as a critical methodological bridge between signal-level processing and model-based interpretation. The extracted features constitute the informational substrate on which feature selection, importance analysis, and classification are subsequently built. The following chapter, therefore, focuses on identifying the most informative and non-redundant subsets of these features, ensuring that subsequent machine learning analyses remain both efficient and physiologically interpretable.

4

Feature Selection

Feature Selection (FS) represents a cornerstone of modern machine learning and statistical modelling, particularly in biomedical contexts where datasets often exhibit high dimensionality and limited sample sizes. In such conditions, selecting a subset of informative variables is essential to mitigate overfitting, enhance generalisation, and improve model interpretability. The main objective of FS is to identify features that contribute maximally to predicting the target variable while discarding redundant, irrelevant, or noisy information [33, 45, 85, 93, 123, 169, 235, 236]. In biomedical signal analysis, this step is particularly critical due to the inherently high correlation among physiological indices and the limited number of available observations per subject [40, 66, 83, 175, 189]. Multimodal biomedical datasets, which combine electrocardiogram, photoplethysmogram, and respiration measurements, pose additional challenges for feature selection due to cross-modality redundancy and diverse signal characteristics [74, 236]. Advanced FS algorithms help reduce dimensionality and ensure that the most informative and complementary features are retained across modalities.

The field of FS exemplifies this convergence between information theory and machine learning. From early heuristic strategies such as the sequential forward and backward selection methods [2, 3, 191] to wrapper-based approaches that evaluate subsets through model performance [120, 256], the objective has consistently been to maximise predictive information while minimising redundancy. Subsequent work introduced criteria explicitly grounded in IT, such as mutual information and conditional mutual information, which quantify how much information a feature contributes about the target variable beyond what is already provided by others [47, 64, 83, 124, 162, 228, 236, 247]. These measures have become essential in modern FS algorithms, providing a unifying principle that

combines statistical relevance with interpretability. Recent reviews and comparative analyses confirm that information-theoretic selection methods outperform purely statistical or distance-based criteria when dealing with nonlinear and multivariate physiological data [33, 45, 235, 236, 237, 247].

Beyond *improving classification performance*, feature selection plays a pivotal role in *ensuring interpretability* of machine learning models, particularly in biomedical applications where the ultimate goal is to *support diagnostic and decision-making processes*. By explicitly identifying the subset of features that truly carry discriminative information, FS not only simplifies the learning model but also enhances its physiological and clinical interpretability [14, 230]. In such contexts, FS algorithms must effectively handle redundancy, noise, and mixed data types to ensure the development of reliable and trustworthy health information systems. Collinearity among physiological features, measurement artefacts, and population biases are recurrent issues in biomedical FS [40, 175]. Information-theoretic and partial information decomposition (PID)-based selection methods are particularly effective in overcoming these pitfalls and in enhancing robustness across heterogeneous cohorts.

In the general case, let $\mathbf{X} = [X_1, X_2, \ldots, X_M]$ denote an M-dimensional random feature vector and let Y denote the target random variable. The available dataset is composed of N observations $\mathcal{D} = \{(\mathbf{x}_n, y_n)\}_{n=1}^{N}$,, where $\mathbf{x}_n = [x_{n1}, x_{n2}, \ldots, x_{nM}]$ and y_n are realisations of $\mathbf{X}$ and Y, respectively. The goal of feature selection is to identify an optimal subset of variables $\mathbf{X}_S \subseteq \mathbf{X}$, with cardinality $S <= M$, that maximises a given objective function $J(\mathbf{X}_S)$, typically associated with predictive performance, information content, or statistical relevance:

$$\mathbf{X}_S^* = \arg \max_{\mathbf{X}_S \subseteq \mathbf{X},\, |\mathbf{X}_S| = S} J(\mathbf{X}_S). \tag{4.1}$$

Since the number of possible feature subsets grows exponentially with the total number of features 2^M [172, 236], exhaustive optimisation of Eq. (4.1) is computationally infeasible for high-dimensional data. Consequently, heuristic or approximate procedures are adopted to explore the feature space and identify a near-optimal subset that balances computational cost and representational power [2, 93, 238].

The feature selection process can be described as a systematic sequence of four interrelated stages [45, 123, 235]. The first is *subset generation*, where candidate combinations of features are created according to a specific search strategy. The second is *subset evaluation*, which measures the discriminative ability or information content of each candidate subset using a defined criterion. The third stage, the *stopping criterion*, determines when the search process should be terminated typically when no further improvement is achieved or a predefined number of features is reached. Finally, the fourth stage, *validation*, assesses

the stability and generalisability of the selected feature subset, ensuring that the result is not an artefact of sampling variability or model bias. This structured framework provides a unified perspective for comparing different FS algorithms and highlights the trade-off between optimality, computational efficiency, and interpretability [33, 83, 169].

Within this framework, two conceptual components play a pivotal role: the *search strategy* and the *evaluation criterion*. The search strategy defines how the space of all possible feature subsets is explored. Because the search space grows exponentially with dimensionality, different strategies have been developed to manage this complexity. Exhaustive search guarantees the global optimum but is feasible only for small feature sets. Deterministic or sequential strategies, such as Sequential Forward Selection and Sequential Backward Elimination, construct or prune feature subsets in a greedy fashion, iteratively adding or removing one feature at a time based on the incremental gain in $J(\mathbf{X}_S)$ [191, 235, 237]. Stochastic methods, such as genetic algorithms, simulated annealing, and particle swarm optimisation, introduce randomness to escape local minima, allowing broader exploration of the search space at the expense of computational efficiency. More recently, hybrid and metaheuristic strategies have emerged, combining deterministic and probabilistic components to achieve a more balanced exploration–exploitation trade-off, particularly suited for nonlinear and multimodal data distributions [2, 3, 93].

The evaluation criterion, in turn, determines how the quality of each candidate subset is assessed. Evaluation can be based on the intrinsic statistical relationship between features and the target variable, on the predictive accuracy of a specific learning model, or on model-derived metrics obtained during training. Accordingly, FS methods are traditionally grouped into three main families: *filter*, *wrapper*, and *embedded* approaches [36, 45, 123, 138, 235]. Filter methods rely on statistical measures such as correlation, consistency, or mutual information (MI) to rank and select features independently of the classifier [83, 124, 236, 247]. They are model-agnostic and computationally efficient, making them ideal for high-dimensional biomedical datasets. Wrapper methods, by contrast, evaluate subsets based on the performance of a specific predictive model, such as accuracy, F1-score, or cross-validation error, thus capturing feature–model interactions but at the cost of increased computational complexity and risk of overfitting [120, 256]. Embedded methods integrate feature selection directly into the learning process itself, as in LASSO or Elastic Net regression, where regularisation forces uninformative coefficients toward zero [33, 93]. These approaches yield sparse, interpretable models but depend on model-specific assumptions, such as linearity or additivity. In recent years, novel approaches based on deep learning and attention mechanisms have enabled feature selection to be performed automatically within neural network architectures, facilitating the analysis of high-dimensional physiological data streams in wearable and telemonitoring applications

[14, 156]. While these methods offer flexibility and scalability, explicit feature selection remains crucial for model interpretability and clinical validation.

Beyond these classical categories, hybrid and ensemble approaches have gained traction by combining the generalisation ability of filter methods with the precision of wrapper-based optimisation. Hybrid strategies typically perform an initial ranking of features using filter-based criteria, followed by an iterative refinement driven by a model-dependent wrapper objective. This two-stage design aims to balance computational efficiency with sensitivity to complex feature interactions. Ensemble methods further extend this idea by aggregating the outcomes of multiple selection algorithms, thereby improving robustness and reducing variance in the selected feature subsets [33, 247]. In this thesis, the focus is placed on *filter-based feature selection* methods. This choice is motivated by their model-agnostic nature, computational efficiency, and suitability for high-dimensional physiological feature spaces, where interpretability and robustness across datasets are critical. In particular, filter criteria grounded in information theory allow the relevance and redundancy of features to be quantified independently of the classifier, providing a principled and physiologically meaningful basis for feature selection.

In the context of biomedical signal analysis, and particularly in cardiovascular research, feature selection serves both methodological and physiological purposes [39, 40, 93]. A single short-term recording of heart period or pulse signal can yield dozens of indices derived from time-domain variability, frequency spectra, or nonlinear information measures. While this diversity enables a detailed characterisation of autonomic function, the inclusion of redundant or collinear features can obscure the interpretation of physiological mechanisms. FS mitigates this issue by retaining the most informative and non-overlapping indices, simplifying subsequent modelling while preserving the essential information required for physiological interpretation [66, 83, 189]. In this sense, feature selection not only acts as dimensionality reduction but also as an informational lens, isolating the variables that convey the largest and most physiologically meaningful share of uncertainty reduction [98, 99, 131].

Moreover, FS can act as a form of implicit regularisation. By limiting the feature space to variables that carry meaningful information, it reduces variance in model estimates and enhances robustness, an effect particularly valuable in biomedical studies, where sample sizes are small and measurement noise is inherent [33, 252]. Practically, FS brings several advantages: it alleviates the curse of dimensionality, improves computational efficiency in real-time or embedded systems, and increases model transparency by focusing analysis on physiologically interpretable parameters, consistent with the principles of XAI in healthcare [14, 156, 230]. Feature selection algorithms must also be evaluated in terms of computational cost, especially for real-time embedded systems or wearable devices,

where low-latency decisions and energy efficiency are critical. Hybrid and filter-based FS strategies maintain scalability and speed in these contexts. The reliability of FS is inherently limited by sample size, noise, and missing data, which frequently affect physiological recordings [33, 40]. Regularization through feature selection mitigates these effects and promotes model stability.

In summary, FS can be viewed as a trade-off between informativeness, redundancy, and computational tractability. The selection of an appropriate method depends on the problem domain, the characteristics of the data, and the interpretability requirements of the application. In physiological signal analysis, interpretability is often as critical as predictive accuracy. For this reason, model-agnostic and information-theoretic methods are particularly advantageous, as they capture the statistical relationships between features and physiological states without depending on specific model assumptions. Such approaches not only improve predictive performance but also facilitate physiological understanding by highlighting which indices, such as measures of autonomic balance, baroreflex sensitivity, or signal complexity, carry the most discriminative information. In this thesis, feature selection is performed using a k-nearest-neighbour Conditional Mutual Information–based feature selection framework (kCMI-FS). This approach ranks candidate features according to their conditional mutual information with the target variable, explicitly accounting for redundancy with respect to already selected features through a nonparametric kNN estimator. Overall, the combination of theoretical rigour, statistical efficiency, and interpretability offered by the proposed information-theoretic framework establishes kCMI-FS as a reliable foundation for subsequent classification and physiological interpretation tasks discussed in this thesis [83, 93, 235, 236, 247].

The remainder of this chapter is organised as follows. Section 4.1 introduces traditional information theoretic-based feature selection approaches such as *mRMR*, and *AIC-based selection*, outlining their theoretical principles and application to physiological datasets (Subsec. 4.1.1, 4.1.2). Section 4.2 presents the advanced information-theoretic framework developed in this thesis, focusing on the *kCMI-FS algorithm*, which extends mutual information analysis to capture conditional, nonlinear, and synergistic dependencies among features. Finally, the chapter concludes with a comparative discussion summarising the relative advantages, limitations, and interpretability of the implemented methods.

4.1 Traditional Information-Theoretic Approaches

Traditional feature selection methods have long provided the foundation for building interpretable and efficient machine learning models, particularly in biomedical applications where model transparency and robustness are essential. Within an information-theoretic

perspective, feature relevance is defined in terms of statistical dependence between features and target variables, leading to a powerful and model-independent framework that does not rely on specific functional forms and is therefore able to capture both linear and nonlinear relationships directly from data [67, 83, 141, 247]. The theoretical basis of these approaches is rooted in the concept of mutual dependence between random variables, which provides a unified and axiomatic foundation for measuring relevance and redundancy [52, 141, 247]. By quantifying how much uncertainty about the target variable is reduced by observing a feature, information-theoretic feature selection establishes a direct and model-free link between statistical dependence and predictive value, a property that is particularly appealing for biomedical data where assumptions of linearity, stationarity, and Gaussianity rarely hold.

Historically, Mutual Information (MI) has represented the core quantity underlying a wide class of traditional information-theoretic feature selection algorithms, due to its ability to quantify statistical dependence between individual features and the target variable in a model-independent manner [236]. Early criteria such as Mutual Information Maximization (MIM) [133] operationalise this idea by ranking candidate features solely according to their individual MI with the target. In practice, this results in a univariate relevance ranking, where features are selected independently of one another, implicitly assuming that high individual relevance implies overall usefulness. While computationally efficient and easy to interpret, this strategy neglects inter-feature dependencies and frequently leads to the selection of highly redundant predictors, especially in physiological datasets where multiple indices may reflect the same underlying regulatory mechanism.

To partially address this limitation, several extensions introduce heuristic redundancy control mechanisms based on pairwise interactions between features. A prominent example is the minimum Redundancy Maximum Relevance (mRMR) algorithm [172], which implements a greedy forward-selection strategy that explicitly balances relevance to the target with redundancy among already selected features. At each iteration, candidate features are evaluated by combining their mutual information with the target and their average pairwise mutual information with the current feature subset, thereby favouring variables that are both informative and minimally redundant. Related methods such as Mutual Information Feature Selection (MIFS) [23] adopt a similar principle by penalising relevance with a redundancy term weighted by a user-defined parameter, whereas Joint Mutual Information (JMI) [251] extends this idea by accounting for the joint contribution of candidate features together with those already selected. In JMI, relevance is evaluated in terms of how much information a candidate feature provides about the target when considered jointly with the selected subset, partially capturing cooperative effects beyond simple pairwise redundancy.

Despite their widespread adoption and favourable computational properties, MI-based methods with heuristic redundancy control remain fundamentally limited by their reliance on low-order dependency structures and greedy selection strategies. Redundancy is typically approximated through pairwise interactions, and higher-order relationships among multiple features are not explicitly modelled. As a result, these approaches may still over-penalise correlated but informative variables or fail to identify features whose relevance emerges only through interactions with multiple predictors.

A more conservative line of traditional methods attempts to mitigate redundancy by explicitly relying on Conditional Mutual Information (CMI). Algorithms such as Conditional Mutual Information Nearest Neighbour estimate (CMINN) [228] and Conditional Mutual Information Maximization (CMIM) [71] evaluate candidate features based on the minimum amount of information they provide about the target when conditioned on each of the already selected variables. Operationally, this strategy ensures that a feature is selected only if it contributes information that cannot be explained by any single previously selected predictor. From an information-theoretic perspective, conditional mutual information has been shown to provide an optimal criterion for feature selection, as it simultaneously quantifies the unique contribution of a candidate feature and its synergistic interactions with already selected variables, while explicitly discounting redundant information [247]. This conservative criterion tends to favour features whose contribution provides additional information beyond that already captured by the selected set. Rather than penalising synergistic effects, conditioning enables the explicit preservation of features that contribute information jointly with others, while discarding variables whose apparent relevance can be fully explained by redundancy with the existing predictors.

Related criteria, including Double Input Symmetrical Relevance (DISR) [155], Conditional Infomax Feature Extraction (CIFE) [136], and Interaction Capping (ICAP) [103] further refine the trade-off between relevance and redundancy through alternative combinations of MI and CMI terms [36, 236]. These methods differ mainly in how relevance and redundancy are weighted or capped during the selection process, with the common goal of preventing excessive penalisation of informative features while controlling overlap among predictors. Although these approaches improve robustness to redundancy compared with purely MI-based rankings, they still rely on low-order conditional structures and heuristic aggregation rules. Consequently, their ability to capture complex multivariate and high-order interactions remains limited, particularly in physiological systems characterised by tightly coupled and nonlinear regulatory dynamics.

The intrinsic limitation of MI-based feature selection lies in the fact that Mutual Information, by construction, does not distinguish between unique, redundant, or synergistic contributions. As a consequence, MI-based criteria may overestimate the relevance of

correlated features or fail to identify variables whose importance emerges only through interactions with others. Conditional Mutual Information overcomes this limitation by quantifying the additional information provided by a candidate feature once the information conveyed by already selected features has been accounted for [83, 141, 247]. This conditional formulation enables redundancy to be discounted and interaction effects to be revealed, providing a more principled basis for feature relevance assessment in biomedical datasets, where physiological variables are often interdependent and share overlapping informational content due to common regulatory mechanisms.

In parallel to MI- and CMI-based filters, information-theoretic model selection approaches such as Akaike Information Criterion (AIC)-based feature selection [5] adopt a parametric perspective, evaluating feature subsets by explicitly balancing goodness of fit and model complexity. Although conceptually distinct from MI-based criteria, AIC-based methods share the same overarching objective of controlling redundancy and preventing overfitting through principled penalisation of model complexity, and are particularly useful when a probabilistic model of the data is available.

In practice, feature selection in physiological data is further complicated by missing data, artefacts, short recording durations, and limited sample sizes, all of which can significantly affect the reliability of low-order dependency estimates, as demonstrated in wearable and ambulatory monitoring studies [40, 97]. Moreover, because most traditional approaches rely on pairwise or low-order conditional structures, they may fail to capture the high-order or synergistic relationships that characterise physiological systems governed by interacting regulatory mechanisms [17, 131]. Nevertheless, information-theoretic feature selection remains both principled and data-efficient, as it relies on a rigorous statistical definition of relevance while requiring no explicit model fitting or parameter estimation. This balance between interpretability, computational tractability, and robustness is particularly valuable for the analysis of physiological signals characterised by complex dependency structures and limited data availability [83, 98, 247]. These considerations motivate the development of more advanced information-theoretic feature selection frameworks that explicitly quantify conditional and multivariate dependencies, moving beyond heuristic redundancy control toward principled decompositions of information into unique, redundant, and synergistic components, as introduced in the following section.

The next subsections describe in detail the minimum Redundancy Maximum Relevance (mRMR) algorithm 4.1.1and the AIC-based feature selection approach 4.1.2, which represent two representative and complementary traditional strategies within the information-theoretic feature selection framework.

4.1.1 Minimum Redundancy Maximum Relevance (mRMR)

The Minimum Redundancy Maximum Relevance (mRMR) criterion is one of the most established filter-based approaches for feature selection. Proposed by Peng *et al.* [172] and summarized in Algorithm 1, it aims to identify features that are both highly informative about the target variable and minimally redundant with each other. The rationale is that an optimal feature subset should maximise the dependency between the selected features and the class labels while minimising inter-feature correlations [83, 169, 236].

In the general information-theoretic framework, the relevance of a subset of features $\mathbf{S} = \{x_1, x_2, \ldots, x_m\}$ to a class variable C can be quantified by their mutual information:

$$D(\mathbf{S}, C) = I(\mathbf{S}; C) = \int_{\mathbf{S}} \sum_{c \in C} p(\mathbf{s}, c) \log \frac{p(\mathbf{s}, c)}{p(\mathbf{s})p(c)} \, d\mathbf{s}. \tag{4.2}$$

Maximising $I(\mathbf{S}; C)$ corresponds to the *max-dependency* criterion, which seeks the subset of features that jointly share the largest amount of information with the target variable. However, computing $I(\mathbf{S}; C)$ directly is computationally intractable, as it requires estimating high-dimensional joint probability densities. To address this issue, Peng *et al.* proposed two approximations, *max-relevance* and *min-redundancy*, which can be combined into a tractable optimisation problem [93, 172].

The *max-relevance* criterion seeks features that exhibit the largest average mutual information with the class variable:

$$D = \frac{1}{|\mathbf{S}|} \sum_{x_i \in \mathbf{S}} I(x_i; C), \tag{4.3}$$

being x_i the candidate feature to be added to or deleted from the subset of selected features while the *min-redundancy* criterion minimises the average mutual information among selected features:

$$R = \frac{1}{|\mathbf{S}|^2} \sum_{x_i, x_j \in \mathbf{S}} I(x_i; x_j), \tag{4.4}$$

being x_j denotes another feature belonging to the subset $\mathbf{S}$ distinct from x_i. The mRMR criterion then combines these two quantities into a single objective, defined either as a difference or as a ratio:

$$\text{Difference form:} \quad \Phi = D - R, \tag{4.5}$$

$$\text{Quotient form:} \quad \Phi = \frac{D}{R}. \tag{4.6}$$

At each step, the feature that maximises Φ is selected and added to the subset $\mathbf{S}$, providing

Algorithm 1 mRMR Feature Selection

Input: Feature set $X = \{x_1, x_2, \ldots, x_M\}$, class variable C, desired subset size S
Output: Selected subset S^*

Initialize $S^* \leftarrow \varnothing$
for each feature $x_i \in X$ **do**
    Compute $I(x_i; C)$ ▷ Mutual Information between feature and class
Select $x_{\max} = \arg\max_{x_i} I(x_i; C)$
$S^* \leftarrow S^* \cup \{x_{\max}\}$
while $|S^*| < S$ **do**
    for each $x_j \in X \setminus S^*$ **do**
        Compute the mRMR criterion:

$$\Phi(x_j) = I(x_j; C) - \frac{1}{|S^*|} \sum_{x_i \in S^*} I(x_j; x_i)$$

    Select $x_k = \arg\max_{x_j} \Phi(x_j)$
    Update $S^* \leftarrow S^* \cup \{x_k\}$
return S^*

a greedy but efficient approximation to the intractable max-dependency problem.

In the original implementation by Peng *et al.* [172], mutual information is computed using discrete probability distributions. Because the features analysed in this thesis are continuous, they are discretised prior to MI estimation, following the conventional histogram-based approach. This quantisation step enables MI to be computed from relative frequencies within discrete bins, providing a practical and computationally efficient implementation. It should be noted, however, that discretisation is not a theoretical requirement of the mRMR framework itself but rather a computational necessity associated with discrete estimators [169, 236].

The mRMR algorithm was implemented using the publicly available toolbox by Peng *et al.*, adapted to process continuous physiological features through data quantisation prior to mutual information estimation. All computations were performed using the histogram-based estimator originally proposed in [172], ensuring methodological consistency with the classical definition of the mRMR criterion.

4.1.2 The AIC-based Feature Selection

Feature selection based on the Akaike Information Criterion (AIC) represents a model-based approach that balances model fit and complexity [5]. Originally introduced within the framework of information theory and maximum likelihood estimation, AIC provides an objective measure to compare competing statistical models by penalising overparameteri-

sation [45, 238]. The rationale is to identify the simplest model that adequately describes the data, thereby preventing overfitting while maintaining explanatory power [31, 238].

Formally, the AIC is defined as:

$$\mathrm{AIC} = 2p - 2\ln(\hat{L}), \tag{4.7}$$

where p denotes the number of estimated model parameters and $\hat{L}$ is the maximised likelihood function. The optimal configuration is obtained by minimising the AIC value, which balances the opposing objectives of goodness of fit and model simplicity.

The AIC-based feature selection implemented in this thesis and summarised in Algorithm 2, follows the approach proposed by [238], relying on a multinomial logistic regression model to relate continuous features to a categorical target variable. Logistic regression is particularly appropriate here because the classification tasks addressed, such as sex discrimination or stress-state classification, involve categorical outcomes rather than continuous responses [2, 3]. Unlike linear regression, which assumes a continuous dependent variable, logistic regression models the conditional probability of class membership and provides interpretable coefficients that quantify the contribution of each feature to class discrimination.

For each possible combination of features, a multinomial logistic regression model [31, 238] was fitted using continuous predictors and categorical labels as inputs. The model estimates the posterior probabilities of class membership as:

$$P(C = k \mid \mathbf{x}) = \frac{\exp(\beta_{0k} + \boldsymbol{\beta}_k^\top \mathbf{x})}{\sum_{j=1}^{K} \exp(\beta_{0j} + \boldsymbol{\beta}_j^\top \mathbf{x})}, \qquad k = 1, \ldots, K, \tag{4.8}$$

where $\boldsymbol{\beta}_k = [\beta_{1k}, \beta_{2k}, \ldots, \beta_{Mk}]^\top$ are the regression coefficients for class k, and β_{0k} is the intercept. Model parameters are estimated by maximising the log-likelihood:

$$\ln \hat{L} = \sum_{n=1}^{N} \sum_{k=1}^{K} y_{nk} \ln P(C = k \mid \mathbf{x}_n), \tag{4.9}$$

where y_{nk} equals 1 if sample n belongs to class k, and 0 otherwise.

For each subset of features, the corresponding AIC value was computed according to Eq. (4.7). The subset minimising AIC identifies the most informative and parsimonious configuration for describing class membership.

The AIC-based selection relies on logistic regression, which directly handles continuous input features and categorical output variables. Unlike mutual-information-based approaches that require discretisation of continuous features, logistic regression models

the conditional probability of each class given the continuous predictors. This makes it particularly well-suited for the physiological datasets analysed in this work, where the features are continuous (e.g., HRV and PRV indices) and the targets correspond to categorical conditions such as stress or sex [39, 93].

The AIC-based procedure was implemented using a multinomial logistic regression via iterative reweighted least squares. For each combination of the ten candidate features, a separate logistic regression model was fitted, resulting in $2^{10} - 1 = 1023$ model configurations. The AIC value was computed for each model as:

$$\mathrm{AIC} = 2p - 2\log\mathrm{L}, \tag{4.10}$$

where logL is the log-likelihood, and p is the total number of fitted coefficients, including intercepts.

Although computationally demanding, this exhaustive search guarantees an exact identification of the global AIC minimum, avoiding suboptimal solutions that may arise from heuristic methods. The subset corresponding to the lowest AIC value was retained as the optimal configuration, providing both statistical parsimony and physiological interpretability.

Algorithm 2 AIC-Based Feature Selection

Input: Feature set $X = \{x_1, x_2, \ldots, x_M\}$, class variable C, classification model type (e.g., multinomial logistic regression)
Output: Optimal subset S^*_{AIC}
  Initialize $S^*_{\mathrm{AIC}} \leftarrow \varnothing$, bestAIC $\leftarrow +\infty$
  for each non-empty subset $S \subseteq X$ **do**
    Train the model using features S to predict C
    Compute the log-likelihood $\hat{L}$ and the number of parameters p
    Evaluate the Akaike Information Criterion:

$$\mathrm{AIC}(S) = 2p - 2\ln(\hat{L})$$

    if AIC(S) < bestAIC **then**
      bestAIC $\leftarrow$ AIC(S)
      $S^*_{\mathrm{AIC}} \leftarrow S$
  return S^*_{AIC}

TTaken together, mRMR and AIC-based methods provide complementary perspectives on feature relevance in physiological signal analysis. The mRMR algorithm quantifies the trade-off between relevance and redundancy using mutual information, offering an information-theoretic and computationally efficient filter-based approach. In contrast,

AIC-based selection promotes model parsimony by penalising unnecessary complexity, ensuring that retained features contribute to statistically optimal models capable of describing the underlying physiological dynamics.

When applied to biomedical signals, these approaches collectively highlight the importance of balancing mathematical rigour, data-driven relevance, and physiological interpretability: mRMR identifies globally informative and minimally redundant features, while AIC-based selection enforces statistical optimality through explicit model evaluation. Together, these traditional techniques establish the methodological foundation upon which more advanced information-theoretic frameworks, such as Conditional Mutual Information-based feature selection, have been developed, as discussed in the following section. The future of feature selection in physiology will increasingly rely on interpretable, high-order, and model-agnostic strategies—capable of revealing physiologically meaningful associations in ever more complex and multimodal data environments [14, 131, 156].

4.2 Developed Information-Theoretic Feature Selection Method

Building upon the limitations of traditional information-theoretic feature selection methods, this section introduces a new framework grounded in Conditional Mutual Information. Unlike heuristic MI-based criteria, this approach defines feature relevance in rigorous probabilistic terms, explicitly accounting for conditional and multivariate dependencies among features. The resulting methodology provides a principled and model-independent basis for selecting informative and non-redundant feature subsets in complex physiological datasets.

4.2.1 The k-Nearest Neighbour Conditional Mutual Information Feature Selection (kCMI-FS)

The proposed method described in this subsection and part of the work published in [98], termed *kCMI-FS* (for *k-nearest neighbour Conditional Mutual Information Feature Selection*), represents an information-theoretic approach designed to identify the most informative and non-redundant subset of features in datasets comprising continuous variables and discrete class labels [17]. The name of the algorithm reflects its two key components: the use of CMI as the measure of feature relevance, and the adoption of a non-parametric k-nearest neighbour (*kNN*) strategy for estimating information quantities directly from the data. In this framework, the “k” parameter controls the local neighbour-

hood size used to estimate the underlying probability distributions, thus influencing the sensitivity of the algorithm to local data structure and sample density.

By combining CMI with kNN estimation, kCMI-FS provides a principled and model-free approach to feature selection, capable of capturing both linear and nonlinear dependencies as well as high-order and synergistic interactions among variables. Unlike traditional MI-based techniques that only measure pairwise relevance, the conditional formulation enables the method to assess the unique contribution of each feature once the information carried by previously selected features has been accounted for. This property makes kCMI-FS particularly well-suited for physiological datasets, where features often exhibit interdependence and overlapping informational content due to shared physiological mechanisms.

The kCMI-FS algorithm introduced in this thesis extends traditional MI-based frameworks by incorporating a significance-driven forward-selection scheme that leverages k-nearest neighbour estimation of CMI to handle mixed-type data composed of continuous features and discrete class variables. Unlike conventional estimators, it maintains estimation accuracy without discretisation, ensuring robustness and data efficiency in high-dimensional biomedical contexts [99, 122, 247].

Let us consider a set of M continuous features collected in a matrix $\mathbf{X}$ of dimension $N \times M$, with N being the number of observations, and C a discrete variable encompassing the classes associated to each instance of $\mathbf{X}$. The goal is to identify a subset of features, defined as the selected features, $\mathbf{X_S} \subseteq \mathbf{X}$ of dimension $N \times M_s$, with $M_s \leq M$, preserving as much as possible the information necessary for accurate classification. Specifically, we employ information-theoretic methods for feature selection, which aim to measure the contribution of each feature $X_i \in \mathbf{X}$, with $i \in [1, ..., M]$, to the prediction of the class variable C, while accounting for redundancy with other features. To this end, we exploit CMI to quantify the information shared between the discrete class variable and the continuous feature variables.

The feature selection using CMI is carried out through an iterative step-by-step forward-selection algorithm, calculating the MI between each continuous feature X_i and the discrete target variable C, denoted as $I(C; X_i)$. In the first iteration, the feature exhibiting the highest MI value is selected and tested for statistical significance. To assess significance, a surrogate test is performed under the null hypothesis of independence between the candidate feature and C. The null distribution is obtained by repeatedly recomputing MI after randomly shuffling the values of the candidate feature, thereby destroying any potential dependency with C. A distribution of surrogate MI values is produced by repeating this procedure 100 times, after which the original measure is compared to it

using a preset significance level. If the hypothesis is rejected, the candidate feature is added to the set of selected features, $\mathbf{X_s}$.

After this, the algorithm calculates the CMI between the class variable C and each remaining feature $X_j \in \mathbf{X} \setminus \mathbf{X_s}$, conditioning on the already selected features in $\mathbf{X_s}$:

$$I(C; X_j|\mathbf{X}_s) = I(C; X_j, \mathbf{X_s}) - I(C; \mathbf{X_s}). \tag{4.11}$$

In this case, the feature with the highest computed CMI among all candidates is identified, for which a surrogate test is performed under the null hypothesis of independence between the candidate and C given $\mathbf{X_s}$. If rejected, this feature is incorporated into the selected set of features $\mathbf{X_s}$. The evaluation process is repeated by applying 4.11 iteratively and enlarging the set $\mathbf{X_s}$ until no significant CMI value emerges. A 5% significance level is used to assess the statistical significance of both MI and CMI terms. Algorithm 3 formally describes the steps of the proposed method.

Estimation strategy. The MI between the class variable C with alphabet A_C and a generic feature variable X_i with continuous domain D_{X_i} can be obtained as the weighted sum of MI terms related to the specific outcome of the discrete variable:

$$I(C; X_i) = \sum_{c \in A_C} p(c) I(c; X_i), \tag{4.12}$$

with $p(c)$ being the probability that the variable C assumes the class c. The specific information term $I(c; X_i)$ can be expressed as:

$$I(c; X_i) = \int_{X_i \in D_{X_i}} p(x_i \mid c) \ln \frac{p(x_i \mid c)}{p(x_i)} \, dx_i, \tag{4.13}$$

where $p(x_i)$ and $p(x_i \mid c)$ denote the probability density of X_i considering all the possible outcomes of the variable C or a specific outcome c, respectively. Similarly to [17, 51, 203, 255], this specific information term can be computed by using the kNN estimation method, considering the association of a feature variable with a specific class as a criterion for defining the search space. Specifically, the kNN approach proposed in [122] is here adopted to minimize the estimation bias arising from entropy terms derived from probabilities and conditional probabilities, i.e., $p(x_i)$ and $p(x_i|c)$, but also deriving from combining MI terms computed on different dimensional spaces for estimating CMI measures.

Fig. 4.1 depicts the graphical illustration of the computation of the MI (panel a) and CMI (panel b) for a discrete class variable and continuous features. For the first step of the

Algorithm 3 kCMI-FS: Conditional Mutual Information Feature Selection

Input: Feature set $\mathbf{X} = \{x_1, x_2, \ldots, x_M\}$, class variable C, significance level α
Output: Selected subset $\mathbf{S}^*$

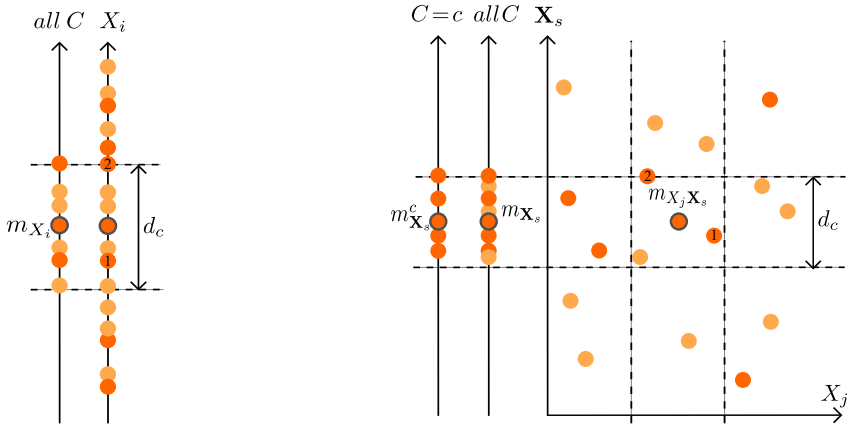
```
Initialise S* ← ∅
for each feature x_i ∈ X do
    Estimate I(C; x_i) via the kNN estimator
Select x_k ← arg max_{x_i ∈ X} I(C; x_i)
Test significance of I(C; x_k | S*) = I(C; x_k) using surrogate data (level α)
while I(C; x_k | S*) passes the significance test do
    S* ← S* ∪ {x_k}
    if X \ S* = ∅ then
        break
    for each x_j ∈ X \ S* do
        Estimate I(C; x_j | S*) via the kNN estimator
    Select x_k ← arg max_{x_j ∈ X\S*} I(C; x_j | S*)
    Test significance of I(C; x_k | S*) using
        surrogate data (level α)
return S*
```

FS algorithm described in Algorithm 3 that requires estimating the MI terms $I(C; X_i)$, k nearest neighbours are searched for each sample $x_i \in X_i$ among the N_c samples associated to the class c. That is, for each class-specific sample x_i^n, where $\mathcal{I}_c$ denotes the set of indices corresponding to samples from class c and $n \in \mathcal{I}_c$, the distance $d_c(x_i^n, k)$ from the specific observation of x_i^n to its k^{th} nearest neighbour within the class c is obtained. The number of samples from X_i that fall within this distance is then counted and denoted by $m_{X_i}(x_i^n)$. The specific MI between the outcome c of the class and the feature X_i is then obtained as:

$$I(c; X_i) = \psi(N) + \psi(k) - \psi(N_c) - \frac{1}{N_c} \sum_{n \in \mathcal{I}_c} \psi(m_{X_i}(x_i^n) + 1), \tag{4.14}$$

where $\psi(\cdot)$ is the digamma function.

Similarly, for estimating CMI quantities involved in the FS iterative procedure (Fig. 4.1b), the neighbourhood search is initially performed in the conditioned space $[X_j, \mathbf{X}_s]$ for each observation of $[x_j, \mathbf{x}_s]$ belonging to the class c to obtain the distance $d_c(x_j^n \boldsymbol{x}_s^n, k)$. Here, the expression $x_j^n \boldsymbol{x}_s^n$ denotes the concatenation into a single joint vector of the corresponding samples from the combined features $[X_j, \mathbf{X}_s]$. Next, the number of samples within the determined range is counted in the unconditioned joint space, accounting for all possible classes, i.e., $m_{X_j \mathbf{X}_s}(x_j^n \boldsymbol{x}_s^n)$. Regarding the set of selected features $\mathbf{X}_s$, the search is performed on both the unconditioned and conditioned space, producing the number of neighbours $m_{\mathbf{X}_s}(\boldsymbol{x}_s^n)$ and $m^c_{\mathbf{X}_s}(\boldsymbol{x}_s^n)$, respectively, obtained as depicted in Fig. 4.1b. The specific counterparts of the two MI terms in Eq.(4.11) are computed as:



(a) (b)

Figure 4.1: Graphical illustration of the nearest-neighbour estimation method (k=2) used to compute the MI $I(c; X_i)$ (Eq. 4.14, Fig. 4.1a) and CMI $I(c; X_j|\mathbf{X_s})$ (Eq. 4.17, Fig. 4.1b) for a discrete class variable and continuous features. This figure has been modified from [98]

$$I(c; X_j, \mathbf{X}_s) = \psi(N) + \psi(k) - \psi(N_c) - \frac{1}{N_c} \sum_{n \in \mathcal{I}_c} \psi(m_{X_j\mathbf{X}_s}(x_j^n \boldsymbol{x}_s^n) + 1), \tag{4.15}$$

$$I(c; \mathbf{X}_s) = \psi(N) - \psi(N_c) + \frac{1}{N_c} \sum_{n \in \mathcal{I}_c} (\psi(m^c_{\mathbf{X}_s}(\boldsymbol{x}_s^n) + 1) - \psi(m_{\mathbf{X}_s}(\boldsymbol{x}_s^n) + 1)), \tag{4.16}$$

thus obtaining:

$$I(c; X_j|\mathbf{X}_s) = \psi(k) + \frac{1}{N_c} \sum_{n \in \mathcal{I}_c} \Big(\psi(m_{\mathbf{X}_s}(\boldsymbol{x}_s^n) + 1) - \psi(m^c_{\mathbf{X}_s}(\boldsymbol{x}_s^n) + 1) - \psi(m_{X_j\mathbf{X}_s}(x_j^n \boldsymbol{x}_s^n) + 1)\Big). \tag{4.17}$$

Computational Complexity. The computational cost of kCMI-FS arises from two sources: (i) the estimation of CMI using a mixed-type kNN strategy, and (ii) the greedy forward-selection procedure that repeatedly evaluates candidate features. Although the individual components rely on well-established operations such as kNN searches and range queries [27, 37], their combination defines the overall scalability of the algorithm. For each candidate feature, the computation of $I(Y; X_j \mid \boldsymbol{X}_s)$ involves several nearest-neighbour evaluations in the joint space $[X_j, \boldsymbol{X}_s]$ and its marginals. The greedy selection process considers a search of $M - M_s$ features at each iteration, and each evaluation requires both the computation of CMI and the generation of N_{surr} surrogates for statistical testing. Assuming that the selection terminates at $M^* < M$ features, summing over all

iterations and retaining only the dominant terms in N and dimensionality, the resulting overall cost scales as $O\big(M\, M^*\, N \log N\big)$.

Evaluation of the kCMI-FS estimation. This paragraph describes the evaluation of the estimator for MI and CMI presented above in simulation settings in which the theoretical values can be computed, so as to allow analysis of bias and variance of the estimator.

We simulate $S_1 \sim \mathcal{U}(a_1, b_1)$ and $S_2 \sim \mathcal{U}(a_2, b_2)$ as two continuous source variables with uniform distribution, where $a_1, a_2 < 0 < b_1, b_2$, and $S_3 \sim \mathcal{N}(0, \sigma^2)$ as a normally distributed variable with zero mean and variance σ^2. The three source variables are assumed to be independent. Then, a target variables is defined through a logic-based formulation: $c = (s_1 \geq 0) \wedge (s_2 \geq 0)$, where $\wedge$ denotes the logical AND, while s_1, s_2, and c represent realisations of the corresponding random variables. The resulting binary target variable C thus captures the ensemble of the outcomes of the logical operation across realisations. To simulate realistic classification scenarios, the three source variables S_1, S_2, S_3, were combined to generate the input features as follows: $X_1 = S_1$, $X_2 = S_2 + S_3$, and $X_3 = S_3$. In this configuration, X_1 represents a relevant feature, X_2 is a noisy combination of relevant and irrelevant sources, and X_3 is fully irrelevant. Two binary target variables were then derived to represent the outputs of separate simulated classification tasks, allowing us to assess the feature selection method's ability to discriminate informative inputs under varying levels of redundancy and noise.

The theoretical values expected from the FS process for each iteration were computed and compared with the estimates derived from realisations generated using a specified set of parameters of the experimental setup, i.e., $a_1 = -1.5$, $b_1 = 0.5$, $a_2 = -0.5$, $b_2 = 0.5$, with varying $\sigma \in \{0, 0.1, 0.2, \ldots, 3\}$. Moreover, different values of $k \in \{2, 5, 10, 20\}$ were taken into account to evaluate the sensitivity of MI and CMI on the estimation parameter. A total of $N = 1000$ realisations were generated to compute the outcomes, from which the information-theoretic measures are calculated. The procedure was repeated across 100 simulations, enabling the estimation of the confidence intervals of each measure.

Fig. 4.2 presents the results in terms of estimation accuracy of both MI and CMI measures for the logical AND operation, as a function of noise level (σ) and at different values of the parameter k. Following the FS approach described Algorithm in 3, the initial iteration evaluates the MI between the outcomes C and each feature X_i, for $i \in \{1, 2, 3\}$, i.e., $I(C; X_i)$, as shown in the panels (4.2a), (4.2b) and (4.2c) in Fig. 4.2. Since only X_1 and X_2 contain information present in the target variable C, the resulting MI measures exhibit higher positive values. For the given parameters, feature X_1 exhibits consistently the highest MI value and would be initially added to the feature subset. Due to the outside

influence, the noisy feature X_2 exhibits decreasing MI values when increasing the noise level, while X_3 produces values always around 0, being independent of the target C. In the subsequent iteration - shown in the panels (4.2d), (4.2e) of Fig. 4.2, respectively - the algorithm computes the CMI between C and each remaining feature, conditioned on X_1, $I(C; X_j \mid X_1)$. Considering that X_3 is still independent of C with the conditioning, the following selected feature would be X_2. The process finishes with the final step involving the computation of CMI for $I(C; X_3 \mid X_1, X_2)$, illustrated in Fig. 4.2f.

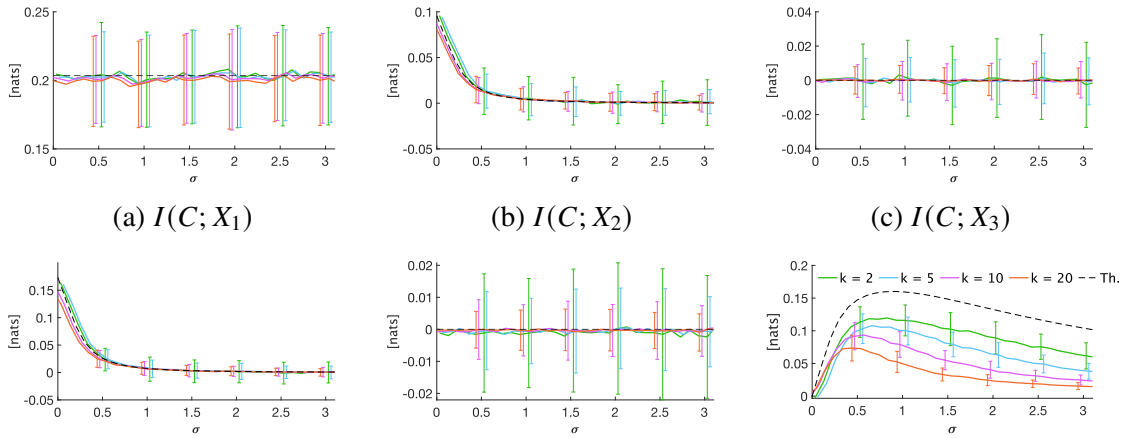
The results shown in Fig. 4.2 demonstrate that increasing σ in S_3 progressively diminishes the relevance of feature $X_2 \equiv S_2 + S_3$, making it harder to determine the true outcome c, which depends on both S_1 and S_2. In the initial selection stages, the values of $I(C; X_3)$ and $I(C; X_3|X_1)$ remain low, correctly indicating that the confounding effect does not introduce spurious relevance for X_3. However, after the selection of both X_1 and X_2, the CMI $I(C; X_3|X_1, X_2)$ evidences an initial increase along with the increase of σ, as X_3 now helps recover the original knowledge of S_2. At higher σ levels, the confounding effect of source S_3 in X_2 begins dominating, making it progressively more independent of S_2, which leads to a decrease in the CMI.

Observing the effects of the nearest neighbour parameter k, although in most cases it did not influence the estimation trends (panels 4.2a-4.2e), we notice that in Fig. 4.2f the choice of k can significantly impact the estimation, leading to a general underestimation compared to the theoretical value. The estimation bias is potentially related to both the number of samples used for the estimation, as well as to the increase in dimension when considering the neighbourhood search [12, 16].

Overall, the obtained findings confirm that both MI and CMI estimations are sensitive to the presence of noise and the choice of the neighbourhood parameter k. An appropriate selection of k is thus essential to balance the bias-variance trade-off, particularly in noisy conditions, and to ensure robust detection of both informative and non-informative features.

Comparison with traditional CMI-based estimators. This section focuses on evaluating and comparing three FS algorithms: CMINN developed by Tsimpiris *et al.* [228], the proposed kCMI-FS method based on either the kNN estimation or through the binning approach (BIN) after discretising all variables, both of them implemented within the FS strategy described above. The CMINN and BIN approaches were taken into account to demonstrate the efficacy of our kNN-based method, which is capable of handling a mix of discrete and continuous variables, in contrast to traditional approaches that exclusively rely on the same types of variables for both target and features.

Herein, we consider five synthetic datasets (*A*-*E*), three of which were adapted from



(a) $I(C; X_1)$ (b) $I(C; X_2)$ (c) $I(C; X_3)$

(d) $I(C; X_2|X_1)$ (e) $I(C; X_3|X_1)$ (f) $I(C; X_3|X_1, X_2)$

Figure 4.2: Theoretical (dashed lines) and estimated (solid lines) MI in panels 4.2a, 4.2b, 4.2c and CMI in panels 4.2d, 4.2e, 4.2f for the AND operation, over 1000 realisations and 100 simulations as a function of the noise level σ. Each solid line represents the average MI or CMI for different values of the number of neighbours $k = 2, 5, 10, 20$, and the error bars denote the $\pm 2std$ over the simulations. This figure is taken from [98].

[228] and two newly generated, to evaluate the performance of the considered estimators in feature selection tasks under heterogeneous data-generating scenarios. Each dataset includes a combination of relevant (Rl), redundant (Rd), and irrelevant (Ir) features. A complete description of the feature construction, target definitions, and relevance structure of all datasets is reported in Table 4.1.

For each dataset, the generated features $\mathbf{X}$ were standardised using z-score normalisation, while the response variable C, whenever not originally defined as discrete (i.e., *Datasets (A-C)*), was binarised into two classes using an equidistant partitioning method [228]. To assess the robustness of the feature selection procedures, simulations were performed for different sample sizes $N = [50, 100, 150, 250, 500, 1000, 2000]$ and number of neighbours $k = [2, 5, 10, 20]$ in the kNN-based entropy estimation. For each dataset and parameter configuration, 100 independent realisations were generated. For the BIN estimator, the number of bins was set to $b = 6$, as in [16, 249].

Each dataset was carefully designed to target the performance of a specific aspect of feature selection: *Dataset A* provides a baseline test in a low-dimensional setting with clearly defined relevant and irrelevant features; *Dataset B* introduces redundancy among features to evaluate whether algorithms can avoid selecting derived or functionally overlapping variables; *Dataset C* increases the complexity by introducing multiple optimal predictor sets (which contain features (X_4, X_5, X_6) and any two of (X_1, X_2, X_3) and additional redundant groups with varying correlation, to test whether the methods can consistently identify equally informative but correlated features; *Dataset D* allows the detection of interactions

Table 4.1: Detailed description of the simulated datasets, including feature generation mechanisms and target definitions. This table is taken from [98].

Dataset	Feature Construction	Target / Class Definition	Type of Features
A	Features: X_1–$X_{22} \sim \mathcal{N}(0,1)$ with constant pairwise correlation $r = 0.5$.	$C = 0.5(-3X_1 + 2X_2 + E_1) + 0.5(3X_3 + 2X_4 - 4X_5 + E_2)$, $E_1, E_2 \sim \mathcal{N}(0,1)$.	Rl: $X_1, \ldots, X_5$ Rd: / Ir: $X_6, \ldots, X_{22}$
B	Predictors $F_1, F_2, F_3 \sim \mathcal{N}(0,1)$. Features: $X_1 = F_1$, $X_2 = F_2$, $X_3 = 0.2F_1 + 0.3F_2 + 2F_3$, $X_4 = 0.1F_1^2 + 0.1F_2^2$. $X_5, X_6 \sim \mathcal{N}(0,1)$.	$C = X_1 + X_2 + 0.2X_3 + 0.3X_4 + E$, $E \sim \mathcal{N}(0,1)$.	Rl: X_1, X_2, X_3 Rd: X_4 Ir: X_5, X_6
C	Predictors $F_1, \ldots, F_5 \sim \mathcal{N}(0,1)$. Features: $X_1 = F_1$, $X_2 = F_2$, $X_3 = F_1F_2$, $X_4 = F_3$, $X_5 = F_4^2$, $X_6 = F_1F_5$. X_7–X_{12} correlated 0.8 with X_1–X_6, X_{13}–X_{18} correlated 0.4, $X_{19}, \ldots, X_{30} \sim \mathcal{N}(0,1)$.	$C = \sum_{i=1}^{6} b_i X_i + E$, $b_i = 1/r_i$, $E \sim \mathcal{N}(0,1)$.	Rl: $X_1, \ldots, X_6$ Rd: $X_7, \ldots, X_{18}$ Ir: $X_{19}, \ldots, X_{30}$
D	Features: $X_1 \sim \mathcal{U}[-5.5, 1.5]$, $X_2 \sim \mathcal{U}[-3.5, 2.5]$, $X_3 = X_1 + \epsilon$, $\epsilon \sim \mathcal{N}(0,1)$. X_4–$X_{13} \sim \mathcal{U}[-5.5, 1.5]$.	$C = \text{sign}(X_1) \cdot \text{sign}(X_2)$.	Rl: X_1, X_2 Rd: X_3 Ir: $X_4, \ldots, X_{13}$
E	Features: $X_1, \ldots, X_4 \sim \mathcal{U}[-5.5, 1.5]$, $X_5, \ldots, X_8 \sim \mathcal{U}[-3.5, 2.5]$. $X_9, \ldots, X_{12}$ obtained by adding $\mathcal{N}(0,1)$ noise to X_1–X_8. X_{13}–$X_{15} \sim \mathcal{N}(a, b-a)$, X_{16}–$X_{24} \sim \mathcal{N}(c, d-c)$.	$C = \text{sign}((X_1 + X_2)(0.5X_3 - X_4) + (X_8 - X_6) + (X_7 + 0.4X_5))$.	Rl: $X_1, \ldots, X_8$ Rd: $X_9, \ldots, X_{12}$ Ir: $X_{13}, \ldots, X_{24}$

between discrete variables through sign-based rules, highlighting the performance of the tested methods to capture non linear, discrete patterns; finally, *Dataset E* scales up dimensionality and redundancy, serving as a high-dimensional challenge to evaluate whether algorithms can still detect all relevant variables among a large number of distractors.

The selection behaviour obtained for kCMI-FS with different configurations of N and k is reported in Fig. 4.3. These results provide practical guidance for parameter selection. As shown, feature recovery improves at increasing sample sizes ($N = 500$) and number of neighbours ($k = 10$), with stable performance observed for N between 500 and 2000. Since simulations with $N = 2000$ entail a substantially higher computational burden, the subsequent comparative analyses were conducted using $N = 1000$ and $k = 10$, as suggested in the literature [122, 228], given the best trade-off between estimation stability and computational costs. For very small sample sizes ($N < 500$), performance becomes less stable due to the increased variance of kNN-based entropy estimation. In such cases, smaller values of k may help reducing variance effects, although larger sample sizes remain preferable for reliable feature selection. This behaviour is consistent with

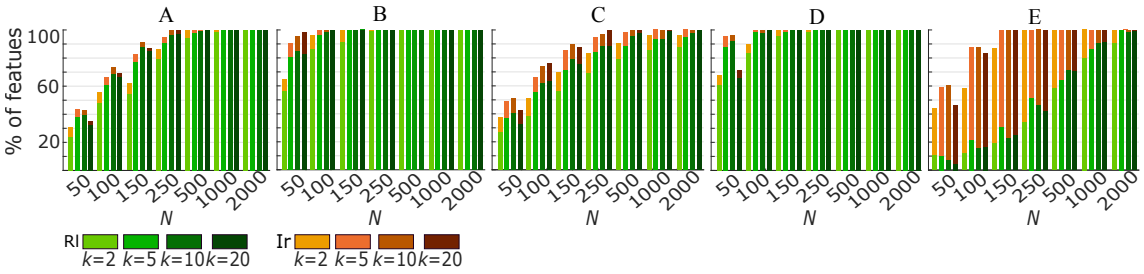
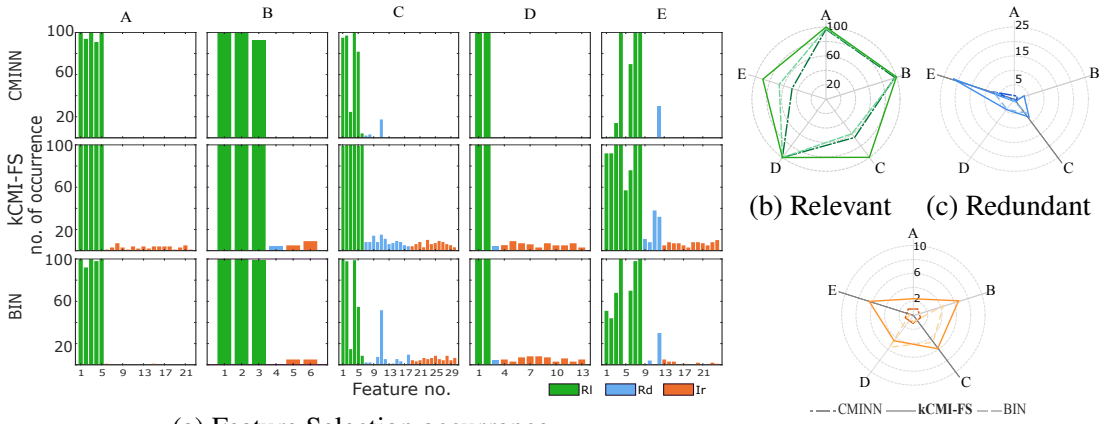


Figure 4.3: Feature selection performance for kCMI-FS across synthetic datasets. Stacked bar plots report the selection behaviour across the five synthetic datasets (panels *A*-*E*) at increasing sample size $N = [50, 100, 150, 250, 500, 1000, 2000]$. Each bar represents the mean percentage of correctly selected relevant features (green shades) and erroneously selected irrelevant features (orange shades), averaged over 100 independent simulations. For each dataset and value of N, the four grouped bars regard four different values of k, $k = [2, 5, 10, 20]$ shown in ascending order and progressively darker tones. This figure is taken from [98].

the statistical properties of kNN-based entropy estimators: as the sample size decreases, the neighbourhood must remain sufficiently local to preserve its interpretability, thus favouring smaller values of k. Similar remarks have been made in [72] using partial mutual information from mixed embedding (PMIME) and related kNN-based information-theoretic selection procedures.

Fig. 4.4a depicts the results in terms of the number of times each feature was selected by the three algorithms over 100 runs of the five synthetic datasets. The comparative analysis across all datasets reveals consistent patterns in the behaviour of the evaluated FS estimation methods (kCMI-FS, CMINN, and BIN), highlighting their respective strengths and limitations depending on the complexity and structure of the dataset. In addition, Fig. 4.4 provides a summary of the feature selection behaviour of the three evaluated algorithms across the five synthetic datasets, reporting the proportion of relevant, redundant, and irrelevant features selected. This visual representation highlights the comparative performance of each estimation method. The values shown in the radial plot correspond to the average percentages of features correctly identified as relevant, mistakenly selected as irrelevant, and additionally included as redundant. For each method and dataset, we recorded the total number of times each feature was selected and categorised it based on its ground-truth role (relevant, redundant, or irrelevant). The final percentages were obtained by dividing the number of selections within each category by the total number of features in that category.

In *Dataset A*, all methods correctly identify the small set of relevant features, confirming their effectiveness in low-dimensional, low-redundancy settings. In detail, all three methods recovered between 97% and 100% of the relevant features, with virtually



(a) Feature Selection occurrence

(b) Relevant

(c) Redundant

(d) Irrelevant

Figure 4.4: Feature selection performance across synthetic datasets. Selection occurrence (4.4a) for the three algorithms (rows) across the five synthetic datasets (columns) generated 100 times. Relevant, redundant, and irrelevant features are identified by green, blue, and orange bar plots, respectively. Radar plots show the average selection behaviour of three methods (CMINN, kCMI-FS, BIN) across the five synthetic datasets (A–E). Each panel summarizes the percentage of (4.4b) relevant features correctly selected, (4.4c) redundant features selected, and (4.4d) irrelevant features erroneously selected. This figure is taken from [98].

no inclusion of irrelevant or redundant ones. In *Dataset B*, kCMI-FS tends to sporadically include such features due to its sensitivity to functional dependencies, contrary to CMINN. Nonetheless, the relevant features were always correctly selected by kCMI-FS, while CMINN missed a few times feature X_3. In terms of irrelevance, only kCMI-FS and BIN showed small inclusion percentages (7% and 5%, respectively), while CMINN avoided all spurious selections. In *Dataset C*, when challenging the selection process with correlated feature groups and multiple optimal predictor sets, while CMINN potentially captures all the sets except for X_6 (and just a couple of redundant ones), kCMI-FS completely captures all the relevant features, indicating an inclusion of redundant predictors, as well as a selection of few redundant-irrelevant features, due to its sensitivity to ambiguous but informative signals. In detail, kCMI-FS selected nearly all relevant features (99.7%), compared to 65.7% for CMINN and 59.3% for BIN. However, this came with a slight increase in redundant (8.6%) and irrelevant features (6.3%). For *Dataset D*, the discrete, rule-based structure of relevant features is consistently captured by all methods, but kCMI-FS and BIN occasionally include redundant features, likely due to estimation sensitivity under low signal-to-noise conditions. Still, all methods achieved 100% recall of relevant features, and inclusion of irrelevant ones remained modest (kCMI-FS: 5%, BIN: 6%). Finally, *Dataset E* allows to demonstrate how kCMI-FS excels in recovering all the relevant features despite high redundancy and dimensionality, though at the cost of occasionally selecting additional noisy or redundant variables. This result aligns with its

design objective, i.e. robustness in complex, high-dimensional feature spaces. kCMI-FS retrieved 89.6% of the relevant features, greatly outperforming BIN (66.4%) and CMINN (48.1%), while including 6.8% irrelevant and 22.3% redundant features. Contrarily, BIN and especially CMINN algorithms often fail to recognize relevant features.

Overall, the kCMI-FS-based approach consistently identified all relevant features in all datasets, even at the cost of often including redundant ones. This behaviour is especially evident in datasets *B* and *C*, reflecting the property of the method to estimate more intricate dependencies and higher-order interactions among variables. It is worth noting that kCMI-FS was the only method able to detect all relevant features in more structured datasets, such as *Dataset E*. The inclusion of redundant or irrelevant features is less likely due to the presence of distractors, since truly irrelevant features should not interfere with greedy selection, but may instead reflect limitations in sample size or instability in estimating high-dimensional dependencies. This suggests that future extensions could benefit from multistage or backward pruning strategies.

Conversely, CMINN exhibited a more conservative behaviour, consistently selecting a reduced set of highly relevant features across all datasets. Its performance is particularly strong in settings where redundancies are limited (e.g., *Dataset D*), but it may underrepresent the information structure in more complex scenarios (e.g., *Datasets C* and *E*), where correctly detecting synergistic effects could improve interpretability and performance. This behaviour can be ascribed to the continuous-only kNN estimator adopted in CMINN which, despite its robustness, may be less suited to the mixed nature of discrete outcomes and continuous inputs occurring in many practical problems. On the other hand, its conservative bias ensures perfect avoidance of irrelevant features in all datasets.

The BIN method reported greater variability across datasets, both in terms of selected features and consistency across iterations. This is especially evident in *Dataset C*, where the binning-based estimator retrieved the most relevant features but with higher inclusion of redundant features compared to CMINN, and a moderate amount of irrelevant ones, even if less than kCMI-FS. Its selection rate of redundant features was of 6.9% in *Dataset C* and 8.5% in *Dataset E*, while irrelevant feature inclusion remained generally low. This behaviour suggests that, while effective at detecting broadly informative patterns, the method may be susceptible to discretisation artefacts that reduce its selectivity, especially in situations of strong redundancy and noise.

In summary, while CMINN provides a reduced and more interpretable subset of features suitable for simpler models, the kCMI-FS method demonstrates superior adaptability in more intricate settings, offering a wider information content, even if somewhat slightly irrelevant or redundant. The BIN method, despite its flexibility, requires cautious appli-

cation due to its higher variability. These remarks underline the importance of aligning the feature selection strategy with the data structure and the aims of the analysis, whether prioritising parsimony and stability or maximising the information content captured.

Table 4.2: Comparison of the feature selection methods employed (traditional information-theoretic approaches).

Method	Principle	Category	Advantages	Limitations
MIM	Ranks features according to their individual mutual information (MI) with the target variable.	Filter-based (information-theoretic)	Simple and computationally efficient; model-independent; effective for preliminary relevance screening.	Ignores redundancy; unable to capture interactions or conditional dependencies.
MIFS	Maximises relevance via MI while penalising redundancy through pairwise MI with selected features.	Filter-based (information-theoretic)	Introduces redundancy control; computationally efficient.	Heuristic redundancy penalisation; limited to pairwise dependencies.
JMI	Maximises joint mutual information between candidate features, selected variables, and the target.	Filter-based (information-theoretic)	Partially captures cooperative effects; improved robustness to redundancy.	Limited to low-order interactions; greedy selection.
CMIM	Selects features by maximising the minimum conditional mutual information.	Filter-based (information-theoretic)	Conservative redundancy control; reduces redundant predictors.	May discard synergistic features; limited high-order sensitivity.
DISR	Balances relevance and redundancy using symmetrical relevance measures.	Filter-based (information-theoretic)	Improved stability compared with MI-based methods.	Heuristic formulation; low-order dependency structure.
CIFE	Combines MI and CMI to promote informative and non-redundant features.	Filter-based (information-theoretic)	Better relevance–redundancy trade-off.	Heuristic weighting; limited synergy detection.

Table 4.2: Comparison of the feature selection methods employed (continued).

Method	Principle	Category	Advantages	Limitations
ICAP	Caps redundancy contribution to prevent excessive penalisation of informative features.	Filter-based (information-theoretic)	Avoids over-penalisation of relevant features; computationally efficient.	Relies on heuristic thresholds; limited theoretical guarantees.
CMI-NN	Estimates conditional mutual information using a non-parametric k-nearest neighbour approach to assess feature relevance conditioned on selected variables.	Filter-based (information-theoretic, non-parametric)	Model-free; handles continuous variables without discretisation; captures nonlinear dependencies more reliably than histogram-based MI/CMI estimators.	Sensitive to neighbourhood size and sample size; computationally demanding; conditional estimation limited to low-order conditioning sets.
mRMR	Maximises mutual information with the target while minimising redundancy among selected variables.	Filter-based (information-theoretic)	Balances relevance and redundancy; widely validated in physiological data analysis.	Requires discretisation of continuous features; limited nonlinear interaction detection.
AIC-based FS	Selects the subset of features that minimises the Akaike Information Criterion.	Embedded/model-based	Penalises overfitting; statistically rigorous; interpretable coefficients.	Model-dependent; assumes specific functional forms.
kCMI-FS	Iteratively selects features based on conditional mutual information estimated via a non-parametric kNN approach.	Filter-based (information-theoretic, non-parametric)	Captures nonlinear and high-order dependencies; explicit redundancy and synergy handling; model-free.	Higher computational cost; sensitive to neighbourhood size; less stable for very small sample sizes.

The methods described in this chapter collectively illustrate how information-theoretic principles, originally developed within the context of communication and statistical inference, can be effectively adapted to the analysis of physiological signals. By framing feature selection in terms of statistical dependence, these approaches provide a unified and model-independent foundation for feature selection, supporting both predictive modelling and mechanistic interpretation in biomedical applications [96, 97, 98].

This chapter has reviewed the theoretical foundations and practical implementations of information-theoretic feature selection methods employed in this thesis, highlighting their role in identifying informative and non-redundant representations of cardiovascular dynamics. Across the different approaches discussed, the central objective remains the same: to select subsets of features that preserve the relevant informational content of physiological signals while mitigating redundancy, noise, and overfitting. A comparative overview of the key properties, strengths, and limitations of the considered methods is provided in Table 4.2.

Traditional information-theoretic approaches based on mutual information and low-order conditional dependencies offer a computationally efficient and interpretable starting point for feature relevance assessment. Methods such as mRMR and related MI- and CMI-based criteria approximate the balance between relevance and redundancy through pairwise or heuristic formulations, providing practical solutions in small-sample and noisy settings commonly encountered in physiological data analysis. In parallel, information-theoretic model selection strategies based on the Akaike Information Criterion (AIC) adopt a complementary perspective, explicitly enforcing model parsimony and statistical optimality when an underlying probabilistic model can be specified [238]. Together, these approaches represent well-established baselines for information-driven feature selection.

Building upon these foundations, the developed information-theoretic framework introduced in this chapter extends classical MI-based selection by adopting Conditional Mutual Information as a core quantity. By explicitly quantifying the additional information contributed by a candidate feature once the effect of previously selected variables has been accounted for, the kCMI-FS algorithm enables a principled treatment of redundancy, synergistic interaction effects, and nonlinear dependencies [98, 141, 247]. This formulation is particularly well suited to physiological datasets, where features often arise from tightly coupled regulatory mechanisms and where relevance may emerge only conditionally or through multivariate interactions.

Overall, feature selection emerges as a conceptual bridge between statistical learning and physiological interpretation. The combination of traditional information-theoretic cri-

teria, model-based selection, and fully conditional information-theoretic methods adopted in this work provides a comprehensive and coherent framework for analysing complex cardiovascular data. Rather than merely reducing dimensionality, these methods facilitate the extraction of compact, interpretable, and physiologically meaningful representations, forming the methodological basis for the classification and interpretation tasks addressed in the subsequent chapters of this thesis.

5

Feature Importance

Feature Importance (FI) represents an intermediate analytical stage that bridges the selection of relevant features and their integration into supervised classification models. Its primary role is to provide a quantitative framework for interpreting how individual and collective features influence predictive outcomes.

Following the feature extraction (FE) phase (Chapter 3), which transforms raw physiological time series into quantitative descriptors of cardiovascular dynamics, and the feature selection (FS) stage (Chapter 4), which identifies a compact and non-redundant subset of variables, FI focuses on assessing the relative contribution of each selected feature to model predictions. In this sense, FI acts as a conceptual and methodological bridge between statistical modelling and physiological interpretation, enabling an explicit evaluation of which features drive the model's decisions and how their interactions shape predictive behaviour. Beyond its interpretative role, FI can also be exploited as a feedback mechanism to refine the feature selection process itself. By quantifying how relevance is distributed across selected variables, FI enables the identification of features that contribute marginally, redundantly, or only in specific multivariate configurations. This information can be leveraged to iteratively adjust the selected feature set, promoting the removal of weak or redundant predictors and the prioritisation of features that exhibit stable and physiologically meaningful contributions. Recent studies have shown that coupling feature selection with information-based importance analysis can improve robustness and interpretability in physiological classification tasks, effectively integrating selection and interpretation within a unified analytical workflow.

The motivation for FI is closely linked to the broader requirement for transparency in machine-learning systems. Within the framework of XAI, interpretability refers to the

extent to which a human can understand the internal mechanisms of a model, whereas explainability concerns the ability to justify why a specific prediction has been produced [117, 158, 217]. Although these concepts partially overlap, both underscore the need for models whose decisions can be examined, communicated, and validated. This requirement is particularly critical in biomedical applications, where predictive outputs must be physiologically plausible, support clinical reasoning, and comply with regulatory standards for transparency and accountability [32, 146, 152]. FI directly contributes to these objectives by quantifying how predictive information is distributed across input variables.

In biomedical signal analysis, and especially in studies of cardiovascular variability, understanding feature relevance is not only a methodological necessity but also a physiological one. Cardiovascular indices are often interdependent, reflecting coupled autonomic mechanisms, nonlinear dynamics, and feedback regulation acting across multiple time scales. Consequently, accurate estimation of the importance of each given feature is essential for interpreting the physiological processes underlying model predictions. Beyond performance optimisation, FI facilitates the translation of machine-learning outputs into physiologically meaningful insights, thereby enhancing both explainability and clinical interpretability [156, 204].

Traditional approaches to feature importance quantify relevance through model-specific parameters or perturbation-based analyses. Coefficient-based methods interpret standardized regression coefficients as indicators of relevance under assumptions of linearity and weak collinearity. Tree-based approaches, such as Random Forests and Gradient Boosted Trees, assess importance through reductions in node impurity associated with feature-based splits [34]. Model-agnostic techniques, including permutation importance, evaluate the impact of disrupting the association between a feature and the target variable on predictive performance. Finally, explainability frameworks such as SHAP (SHapley Additive exPlanations) adopt a game-theoretic perspective to attribute a fair contribution to each feature [140].

Despite their widespread adoption, these methods primarily capture first-order effects and often fail to represent the complex dependency structure characteristic of physiological data. Features derived from cardiovascular time series are rarely independent; many are statistically correlated or physiologically coupled, sharing redundant or complementary information. As a result, classical importance measures may overestimate the relevance of correlated predictors or fail to detect collective effects that emerge only through feature interactions. To address these limitations, recent work has introduced information-theoretic formulations of feature importance, capable of quantifying both individual and higher-order contributions within a unified probabilistic framework [225, 246].

Within this context, the High-Order Interaction Feature Importance (HiFi) framework represents a first step beyond traditional approaches by explicitly accounting for cooperative effects among predictors [163, 164]. Its information-theoretic extension, Conditional Mutual Information HiFi (CMIHiFi) [131], further generalizes this perspective by defining feature relevance directly in probabilistic terms, independently of any predictive model. By decomposing predictive information into *unique*, *redundant*, and *synergistic* components, these frameworks shift feature importance from univariate relevance scoring toward a structured description of multivariate interactions. Such a decomposition is particularly suited to physiological datasets, where features derived from time, frequency, and information domains encode distinct yet interdependent aspects of autonomic regulation.

More broadly, feature importance integrates the goals of statistical learning and physiological modelling. It provides a quantitative means to interpret how models allocate relevance across features while offering insight into how physiological mechanisms combine to generate observable patterns in cardiovascular variability. In this thesis, feature importance is not treated solely as a post-hoc tool for model explainability, but is developed as a central analytical component for physiological interpretation. The original contribution lies in the development and systematic adoption of information-theoretic feature importance algorithms capable of capturing high-order effects, namely the High-order Interaction Feature Importance (HiFi) [163, 164] framework and its fully information-theoretic extension based on Conditional Mutual Information (CMIHiFi) [131]. These methods explicitly decompose feature relevance into unique, redundant, and synergistic components, providing a principled way to analyse how cardiovascular indices cooperate, overlap, and interact in shaping autonomic function. This perspective extends beyond the predominantly first-order importance measures commonly employed in the literature and enables a multivariate interpretation of physiological regulation grounded in information theory.

The remainder of this chapter is structured as follows. Section 5.1 introduces traditional FI methods, including coefficient-based, tree-based, permutation, and SHAP approaches, highlighting their theoretical foundations and limitations in high-dimensional physiological datasets. Moreover, it presents the high-order feature importance (HiFi) framework (Sec. 5.1.2). Section 5.2.1 presents an the information-theoretic verson of the Hifi framework, describing the CMIHiFi methodology and its mathematical formulations.

5.1 Feature Importance Methods

Traditional approaches to feature importance aim to quantify the relevance of each variable by assessing its influence on model behaviour or predictive performance. Although

they differ in implementation and underlying assumptions, these methods share a common conceptual foundation: feature relevance is primarily evaluated through marginal or first-order effects, whereby each variable is treated as an individual contributor to the prediction task. As a consequence, interactions among features are either absorbed implicitly by the learning algorithm or ignored altogether, and feature importance is typically interpreted as an additive contribution.

5.1.1 Traditional Approaches

Coefficient-based feature importance, commonly adopted in linear and logistic regression models, interprets standardised model coefficients as direct measures of relevance [35]. In this setting, the coefficient magnitude reflects the sensitivity of the model output to variations in the corresponding feature, providing a transparent and computationally efficient importance estimate. However, this interpretation is strictly valid only under assumptions of linearity, independence among predictors, and correct model specification. When predictors are correlated or the underlying relationships are nonlinear, coefficient estimates may become unstable and fail to reflect the true contribution of individual features.

Impurity-based feature importance, widely used in tree-based models such as Random Forests and Gradient Boosted Trees, quantifies relevance through the cumulative reduction in node impurity (e.g., Gini index or entropy) induced by feature-based splits [34]. Because such models recursively partition the feature space, impurity-based importance can capture certain nonlinear effects and interactions. Nonetheless, importance scores are intrinsically model-specific and are known to be biased toward features with many potential split points or higher variability, complicating their interpretation and comparison across different models.

Permutation-based feature importance provides a model-agnostic alternative by directly linking relevance to predictive performance [34]. By measuring the degradation in model accuracy following random permutation of a feature, this approach estimates how much predictive information is lost when the association between that feature and the target is disrupted. While permutation importance is applicable to a wide range of models and can capture nonlinear dependencies, it is sensitive to correlations among features, as redundant predictors may compensate for the permuted variable, leading to underestimation of relevance. Furthermore, repeated permutations can be computationally demanding and unstable in small-sample regimes.

Explainability frameworks such as *SHAP* (SHapley Additive exPlanations) extend traditional importance measures by grounding feature attribution in cooperative game theory

[140]. SHAP values quantify the marginal contribution of a feature averaged over all possible subsets of variables, ensuring desirable properties such as consistency and local accuracy. This formulation enables both instance-level explanations and global summaries of feature relevance. However, SHAP remains fundamentally additive in nature, focusing on how features contribute individually to predictions rather than explicitly characterising interaction or synergy structures, and its computational cost grows rapidly with model complexity.

A key limitation shared by all these traditional approaches is the absence of an explicit mechanism to disentangle independent, overlapping, and cooperative contributions among features. As a result, feature importance scores may conflate unique information with redundancy or fail to reveal relevance that emerges only through feature interactions. In complex systems such as physiological datasets, where inter-feature correlations and synergistic relationships are intrinsic, this limitation can lead to misleading interpretations of variable relevance. These considerations motivate the development of feature importance frameworks that move beyond marginal or additive attributions, explicitly accounting for redundancy and interaction effects, as addressed by the methods introduced in the following sections.

5.1.2 High-Order Interaction Feature Importance (HiFi)

The *High-Order Interaction Feature Importance* (HiFi) framework was introduced to extend traditional feature importance beyond marginal effects by explicitly accounting for higher-order interactions among predictors [163]. The method is grounded in a predictability-based perspective and formalizes feature relevance in terms of the contribution of a driving variable to the prediction of a target variable, conditional on different subsets of other predictors.

Let Y be a scalar target variable, X a candidate driving feature, and $\mathbf{Z} = \{Z_1, \ldots, Z_N\}$ the set of remaining predictors. The starting point of the HiFi framework is the Leave-One-Out Covariates (LOCO) measure, which quantifies the gain in predictability obtained by including X in the regression of Y given $\mathbf{Z}$ [132]. The LOCO measure is defined as:

$$L_{\mathbf{Z}}(X \to Y) = \varepsilon(Y \mid \mathbf{Z}) - \varepsilon(Y \mid X, \mathbf{Z}), \tag{5.1}$$

where $\varepsilon(Y \mid \mathbf{Z})$ denotes the mean squared prediction error of Y given the set of predictors $\mathbf{Z}$.

A key observation underlying HiFi is that the predictive contribution of X depends on the particular conditioning set $\mathbf{Z}$. Accordingly, the LOCO measure is evaluated across all possible subsets $\mathbf{z} \subseteq \mathbf{Z}$, revealing how the relevance of X changes when different variables

are accounted for.

To characterize this dependence, HiFi introduces two extremal quantities:

$$L_{\min}(X \to Y) = \min_{\mathbf{z} \subseteq \mathbf{Z}} L_{\mathbf{z}}(X \to Y), \tag{5.2}$$

$$L_{\max}(X \to Y) = \max_{\mathbf{z} \subseteq \mathbf{Z}} L_{\mathbf{z}}(X \to Y), \tag{5.3}$$

which represent the minimum and maximum predictability gains achievable by conditioning on subsets of $\mathbf{Z}$.

The unconditional predictability gain of X is defined as:

$$L(X \to Y) = \varepsilon(Y) - \varepsilon(Y \mid X), \tag{5.4}$$

where $\varepsilon(Y)$ is the prediction error obtained without conditioning on any predictor other than X.

Using these quantities, the total predictive contribution of X can be decomposed into three additive components:

$$U(X \to Y) = L_{\min}(X \to Y), \tag{5.5}$$

$$R(X \to Y) = L(X \to Y) - L_{\min}(X \to Y), \tag{5.6}$$

$$S(X \to Y) = L_{\max}(X \to Y) - L(X \to Y). \tag{5.7}$$

This decomposition satisfies the additive identity:

$$L_{\max}(X \to Y) = U(X \to Y) + R(X \to Y) + S(X \to Y), \tag{5.8}$$

which formally separates the predictive contribution of X into unique, redundant, and synergistic components.

The term U quantifies the portion of predictive information carried exclusively by X, R represents information redundantly shared with other features, and S captures information that emerges only through the interaction of X with other variables. This decomposition generalizes the classical concept of feature relevance to include higher-order dependencies, thus providing a richer interpretation of multivariate relationships.

Computing the minima and maxima in Eqs. (5.2)–(5.3) for all subsets $\mathbf{Z}$ is computationally intractable for large feature sets. To address this limitation, HiFi employs a greedy approximation strategy to identify the conditioning subsets $\mathbf{Z}_{\min}$ and $\mathbf{Z}_{\max}$. Starting from an empty conditioning set, variables are iteratively added to the set so as to minimize or

maximize the predictability gain:

$$Z_j = \arg \min_{Z \in \mathbf{Z} \setminus \mathbf{z}_{j-1}} L_{\mathbf{z}_{j-1} \cup \{Z\}}(X \to Y), \tag{5.9}$$

$$Z_j = \arg \max_{Z \in \mathbf{Z} \setminus \mathbf{z}_{j-1}} L_{\mathbf{z}_{j-1} \cup \{Z\}}(X \to Y). \tag{5.10}$$

At each iteration, the statistical significance of the observed change in predictability is assessed via surrogate-based permutation testing under the null hypothesis of conditional independence. The greedy search terminates when no further statistically significant changes are detected, yielding approximations of the subsets attaining $L_{\min}$ and $L_{\max}$. This procedure enables the identification of the variables that contribute to redundancy or synergy alongside the driving feature.

The HiFi framework therefore provides a general methodology to identify and quantify the structural interactions among predictors. It establishes a foundation for feature importance that is not limited to individual contributions, but also captures the cooperative informational architecture of the feature space.

Conceptually, HiFi represents a transition from first-order feature importance toward a multivariate interpretation of relevance. While still formulated in terms of predictability and regression error, it introduces a structured decomposition inspired by recent developments in higher-order interaction analysis and partial information decomposition. This perspective reveals how feature relevance may arise not only from isolated effects, but also from cooperative or overlapping relationships among predictors.

The predictability-based nature of HiFi, however, ties the decomposition to the choice of regression model and error metric. This limitation motivates a fully information-theoretic reformulation, in which feature relevance is defined directly in terms of conditional mutual information, independently of any predictive model. Such a formulation is provided by the Conditional Mutual Information HiFi (CMIHiFi) framework, introduced in the following section.

5.2 Novel Information-Theoretic Feature Importance Method

Traditional feature importance approaches quantify the contribution of individual features to model prediction by evaluating changes in predictive performance or model parameters. However, such approaches are inherently limited to first-order dependencies and cannot disentangle how multiple features jointly contribute to the target variable.

In multivariate systems, such as those encountered in physiological or biological data, variables often interact in a nonlinear and cooperative manner. To address these limitations, a family of information-theoretic frameworks has been introduced to explicitly quantify higher-order interactions in feature relevance. Among these, the *Conditional Mutual Information HiFi* (CMIHiFi) [131] provides a principled and model-independent formulation for decomposing the informational contribution of each feature into unique, redundant, and synergistic components.

5.2.1 Conditional Mutual Information HiFi (CMIHiFi)

Building upon the conceptual structure of HiFi (Subsec 5.1.2), the *Conditional Mutual Information HiFi* (CMIHiFi) framework [131] reformulates the entire method in purely information-theoretic terms, removing the dependence on any predictive model. Instead of using an explicit performance metric such as the prediction error in Eq. (5.1), CMIHiFi directly quantifies the informational dependencies among variables through empirical estimation of the conditional mutual information.

Let us consider a system with $N + 2$ variables, out of which one represents the class variable Y, while the rest represent continuous features. For the analysis, we isolate the variable X, denoted as the source feature, for which we assess the importance in relation to the other features collected in $\mathbf{Z} = \{Z_1, \ldots, Z_N\}$. We want to determine the importance of X to explain Y in the context of the variables $\mathbf{Z}$. The LOCO equivalent in information-theoretic terms with the CMI would be

$$I(Y; X|\mathbf{Z}) = I(Y; X, \mathbf{Z}) - I(Y; \mathbf{Z}). \tag{5.11}$$

Intuitively, the level of CMI increase shows the level of contribution of source X, already investigated in [247] through PID-based reasoning, showing that this measure provides the unique and synergistic effect obtained from the feature, providing a well-principled criterion for feature selection. Following the same reationale of the HiFi method devised in [163], we look for sets of variables in $\mathbf{Z}$ that minimize or maximize the CMI between the class Y and the source X, i.e., $I(Y; X|\mathbf{Z}_{min})$ and $I(Y; X|\mathbf{Z}_{max})$, so as to disentangle the unique contribution of X in determining Y from that deriving from its redundant and synergistic interaction with the variables in $\mathbf{Z}$. Specifically, the maximum information shared between the class Y and the source X can be decomposed as follows [220]:

$$\begin{aligned} I(Y; X|\mathbf{Z}_{max}) = S(Y; X|\mathbf{Z}) + R(Y; X|\mathbf{Z}) \\ + U(Y; X|\mathbf{Z}), \end{aligned} \tag{5.12}$$

with

$$\begin{aligned} S(Y;X|\boldsymbol{Z}) &= I(Y;X|\boldsymbol{Z}_{max}) - I(Y;X), \\ R(Y;X|\boldsymbol{Z}) &= I(Y;X) - I(Y;X|\boldsymbol{Z}_{min}), \\ U(Y;X|\boldsymbol{Z}) &= I(Y;X|\boldsymbol{Z}_{min}), \end{aligned} \tag{5.13}$$

representing the synergistic, redundant, and unique contribution, respectively.

Search algorithm. Due to the exponential computational cost of an exhaustive search, which requires evaluating all possible conditioning subsets (scaling as $O(2^n)$), a greedy approach is adopted to obtain $\boldsymbol{Z}_{min}$ and $\boldsymbol{Z}_{max}$ in practice. Specifically, at each iteration of the sequential procedure, variables are added to the two subsets one at a time, minimizing or maximizing the CMI quantities with respect to the set of previously selected variables, reducing the computational complexity to approximately $O(n^2)$ evaluations. In practice, for high-dimensional systems, the greedy search stops after $m \ll n$ iterations, further reducing the complexity to $O(mn)$.

The overall flow of the greedy variable selection is illustrated in Fig. 5.1, which summarizes the minimization procedures used to determine $\boldsymbol{Z}_{min}$, while the selection of $\boldsymbol{Z}_{max}$ follows a similar approach. The two searches follow the same iterative structure, differing only in the selection criterion applied to the candidate feature V_j (minimizing or maximizing the CMI). The procedure is initialized representing the subsets of selected variables as the empty set, $\boldsymbol{Z}_0 = \emptyset$.

Then, to determine the set $\boldsymbol{Z}_{min}$, at any given step j ($1 \leq j \leq n$) the variable $V_j \in \boldsymbol{Z} \setminus \boldsymbol{Z}_{j-1}$ is determined such that

$$V_j = \underset{V \in \boldsymbol{Z} \setminus \boldsymbol{Z}_{j-1}}{\arg\min}\, I(Y;X|\boldsymbol{Z}_{j-1}, V). \tag{5.14}$$

Estimation approach. To perform the searching algorithm and find $\boldsymbol{Z}_{min}$ and $\boldsymbol{Z}_{max}$, at the step j we need to compute two CMI terms which need to be compared to quantify the predictive improvement brought by the candidate variable V, i.e.,

$$\begin{aligned} I(Y;X|\boldsymbol{Z}_{j-1}, V) &= I(Y;X,\boldsymbol{Z}_{j-1},V) \\ &\quad - I(Y;\boldsymbol{Z}_{j-1},V), \\ I(Y;X|\boldsymbol{Z}_{j-1}) &= I(Y;X,\boldsymbol{Z}_{j-1}) \\ &\quad - I(Y;\boldsymbol{Z}_{j-1}). \end{aligned} \tag{5.15}$$

We propose to estimate these two CMI values using a mixed variable approach based on the kNN estimator [122]. Considering the class variable Y discrete, with alphabet A_Y,

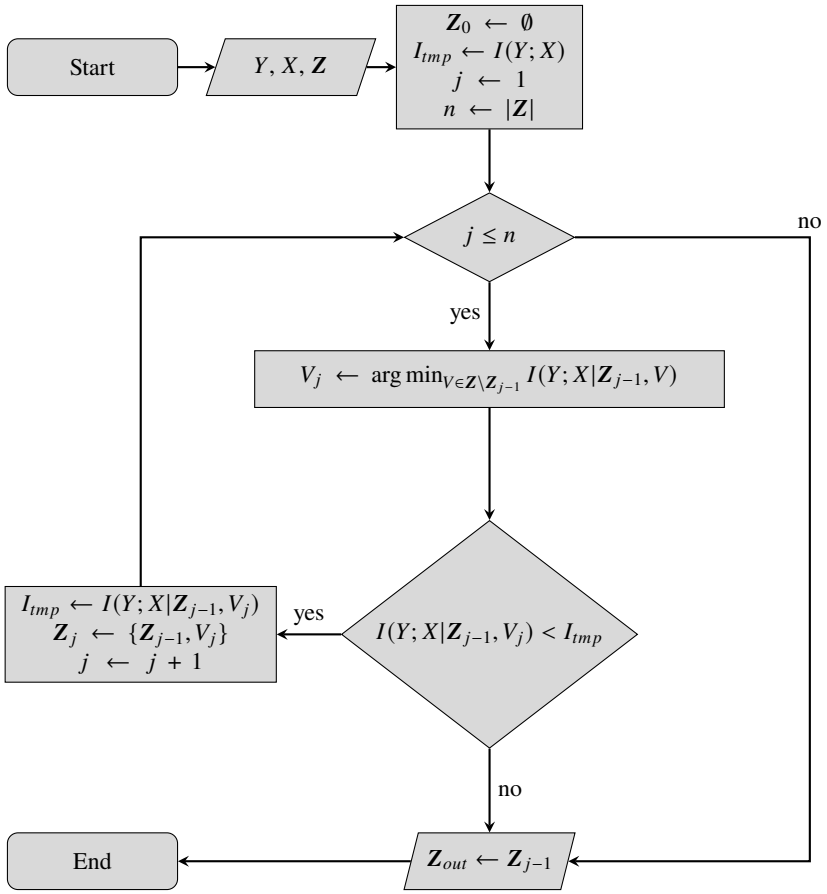


Figure 5.1: Flowchart of the theoretical greedy selection procedure for determining $\mathbf{Z}_{min}$. At each iteration, a candidate variable V_j is selected, provided it has a minimal CMI. n represents the number of features in $\mathbf{Z}$. Selection of the $\mathbf{Z}_{max}$ set is similar, where V_j is selected as the feature that maximizes the CMI out of the set of candidate features, $V_j \leftarrow \arg\max_{V \in \mathbf{Z} \setminus \mathbf{Z}_{j-1}} I(Y;X|\mathbf{Z}_{j-1},V)$, along with the comparison $I(Y;X|\mathbf{Z}_{j-1},V_j) > I_{tmp}$ before the update step. This figure has been adapted from [131].

and the source variables X and $\mathbf{Z}$ continuous, with domains D_X and $D_{\mathbf{Z}}$, respectively, the mixed kNN estimation approach detailed in [17] is used to compute the four MI terms in (5.15). In order to minimize the potential estimation bias which can arise when dealing with different neighborhood spaces [249], the neighborhood search is performed in the highest-dimensional space $\{X, \mathbf{Z}_{j-1}, V\}$, after which the determined distances are projected to the lower-dimensional spaces $\{\mathbf{Z}_{j-1}, V\}$, $\{X, \mathbf{Z}_{j-1}\}$, and $\{\mathbf{Z}_{j-1}\}$.

In detail, each MI term is calculated as the weighted sum of the specific MI relevant to the outcome y of the class Y, with weights given by the probabilities $P(Y = y) \equiv P(y)$ and specific MI terms computed as:

$$I(Y = y; X, \mathbf{Z}_{j-1}, V) = \psi(N) - \psi(N_y) + \psi(k) - \langle \psi(m_{x,z_{j-1},v} + 1) \rangle, \tag{5.16}$$

$$I(Y = y; \mathbf{Z}_{j-1}, V) = \psi(N) - \psi(N_y) + \langle \psi(m^y_{z_{j-1},v} + 1) \rangle - \langle \psi(m_{z_{j-1},v} + 1) \rangle, \tag{5.17}$$

$$I(Y = y; X, \mathbf{Z}_{j-1}) = \psi(N) - \psi(N_y) + \langle \psi(m^y_{x,z_{j-1}} + 1) \rangle - \langle \psi(m_{x,z_{j-1}} + 1) \rangle, \tag{5.18}$$

$$I(Y = y; \mathbf{Z}_{j-1}) = \psi(N) - \psi(N_y) + \langle \psi(m^y_{z_{j-1}} + 1) \rangle - \langle \psi(m_{z_{j-1}} + 1) \rangle, \tag{5.19}$$

with $\langle \cdot \rangle$ being the mean value, $\psi(\cdot)$ is the digamma function, N the total number of samples, N_y the number of samples associated to the outcome $Y = y$, k the number of neighbors searched for each sample in the space $\{X, \mathbf{Z}_{j-1}, V|Y = y\}$, referring to samples in $\{X, \mathbf{Z}_{j-1}, V\}$ associated to the outcome $Y = y$, $m^y_{z_{j-1},v}$, $m^y_{x,z_{j-1}}$, and $m^y_{z_{j-1}}$ the number of samples in $\{\mathbf{Z}_{j-1}, V|Y = y\}$, $\{X, \mathbf{Z}_{j-1}|Y = y\}$ and $\{\mathbf{Z}_{j-1}|Y = y\}$ for the determined distance, respectively, and $m_{x,z_{j-1},v}$, $m_{z_{j-1},v}$, $m_{x,z_{j-1}}$, and $m_{z_{j-1}}$ those counted considering all samples in $\{X, \mathbf{Z}_{j-1}, V\}$, $\{\mathbf{Z}_{j-1}, V\}$, $\{X, \mathbf{Z}_{j-1}\}$ and $\{\mathbf{Z}_{j-1}\}$.

Given that the set $\mathbf{Z}_{min}$ minimizing the shared information between X and Y determines the unique contribution of the source variable to the class, and that the redundant contribution of the variables in $\mathbf{Z}_{min}$ is obtained as the difference between the $I(Y; X)$ and $I(Y; X|\mathbf{Z}_{min})$, it is necessary to assess first the statistical significance of $I(Y; X)$ to determine if there is any potential unique or redundant contribution. Therefore, in the initial step, the MI term $I(Y; X)$ is computed as

$$I(Y; X) = \sum_{y \in A_Y} P(y) I(Y = y; X), \tag{5.20}$$

with the specific MI term $I(Y = y; X)$ being

$$I(Y = y; X) = \psi(N) - \psi(N_y) + \psi(k) - \langle \psi(m_X + 1) \rangle. \tag{5.21}$$

Here, k is the number of neighbors searched in the space $\{X|Y = y\}$, and m_X is the number of samples counted in the same range for the space $\{X\}$. A surrogate test is performed to test the null hypothesis of independence between Y and X. To simulate the null distribution, the samples of X are randomly shuffled, and the MI in (5.20) is

recomputed. This procedure is repeated N_{surr} times, producing a distribution of surrogate values. The significance of the original measure is then evaluated by comparing it to the 95^{th} percentile of the surrogate distribution. If the observed value exceeds this threshold, the null hypothesis is rejected and the MI is assumed to be greater than 0. Otherwise, this would indicate that the empty set is the final set $\boldsymbol{Z}_{min} = \emptyset$, thus shortening the search procedure.

Similarly, for determined values of $I(Y; X|\boldsymbol{Z}_{j-1}, V_i)$ and $I(Y; X|\boldsymbol{Z}_{j-1})$, a statistical analysis of their difference is performed to test the null hypothesis that the candidate V_i does not alter the conditional information shared by X and Y given $\boldsymbol{Z}_{j-1}$; if the null hypothesis is verified, V_i will not be included in the sets of selected variables $\boldsymbol{Z}_{min}$ or $\boldsymbol{Z}_{max}$ and the search is terminated. Specifically, significance of $I(Y; X|\boldsymbol{Z}_{j-1}, V_i) - I(Y; X|\boldsymbol{Z}_{j-1})$ is tested for the $\boldsymbol{Z}_{max}$ search and of $I(Y; X|\boldsymbol{Z}_{j-1}) - I(Y; X|\boldsymbol{Z}_{j-1}, V_i)$ for the $\boldsymbol{Z}_{min}$ search. With the same null hypothesis, a test is done by randomly permuting the investigated variable V_i N_{surr} times, producing a distribution of surrogate values of the difference measure. If the original measurement exceeds the 95^{th} percentile threshold, the difference is deemed significant, and the contributing variable V_i is added to the appropriate set.

Once $\boldsymbol{Z}_{min}$ and $\boldsymbol{Z}_{max}$ are found, all the terms needed to obtain the contributions in (5.13), i.e., $I(Y; X)$, $I(Y; X|\boldsymbol{Z}_{min})$ and $I(Y; X|\boldsymbol{Z}_{max})$, are computed again considering as neighbor search space $\{X, \boldsymbol{Z}_{min} \cup \boldsymbol{Z}_{max}\}$. Specifically, equations (5.18) and (5.19) are employed, replacing $\boldsymbol{Z}_{j-1}$ with $\boldsymbol{Z}_{min}$ and $\boldsymbol{Z}_{max}$.

Validation on Gaussian Mixture Models. To establish a quantitative reference for evaluation, we introduce a Gaussian-based framework in which the conditional mutual information (CMI) can be computed theoretically. Assuming that the conditional joint distribution $\{X, \boldsymbol{Z} \mid Y = y\}$ is multivariate Gaussian, it allows for the derivation of ground-truth CMI values via Monte Carlo (MC) estimation. These theoretical measures provide a benchmark against which the proposed estimator can be validated. The following subsections detail the derivation of these theoretical values and the validation experiments designed to assess the estimator's performance under controlled dependency structures.

Theoretical computation. Under the Gaussian-based framework, we derive the theoretical values of the CMI for a known system. With the assumption that for a given discrete realization $Y = y$, the underlying joint conditioned space $\{X, \boldsymbol{Z}|Y = y\}$ follows a multivariable Gaussian distribution, the joint distribution $p(X, \boldsymbol{Z})$ becomes a Gaussian Mixture Model (GMM), formed by the class weights and conditional probability densities, i.e.,

$$p(X, \boldsymbol{Z}) = \sum_{y} P(y) p(X, \boldsymbol{Z}|Y = y). \tag{5.22}$$

To evaluate the CMI values, starting from the definition

$$I(Y;X|\boldsymbol{Z}) = \iiint p(y,x,z)\log\left(\frac{p(z)p(y,x,z)}{p(x,z)p(y,z)}\right)dydxdz, \tag{5.23}$$

we have

$$I(Y;X|\boldsymbol{Z}) = \sum_{y} P(y)\iint p(x,z|y)\log\left(\frac{p(z)p(x,z|y)}{p(x,z)p(z|y)}\right)dxdz, \tag{5.24}$$

with each of the terms, apart from $P(y)$, representing the density for either a multivariate Gaussian distribution or a GMM. The integral represents the specific CMI for a given realization $Y = y$, $I(Y = y; X|\boldsymbol{Z})$.

Given the initial distribution assumption constraints, this formulation allows us to simulate complex, multimodal dependencies in the joint space while retaining full knowledge of the underlying generative processes. Since closed-form expressions for the individual specific cross-entropy terms, specifically $H_{X,\boldsymbol{Z}|Y=y}(X,\boldsymbol{Z})$ and $H_{X,\boldsymbol{Z}|Y=y}(\boldsymbol{Z})$, are not available in this setting, we resort to high-precision numerical estimation using the Monte Carlo (MC) method [154], with the integral approximated as

$$I(Y = y;X|\boldsymbol{Z}) = \hat{\mathbb{E}}_{p(x,z|y)}\left[\log\left(\frac{p(z)p(x,z|y)}{p(x,z)p(z|y)}\right)\right]. \tag{5.25}$$

To obtain this estimate, we draw many samples from the conditional distribution space $\{X,\boldsymbol{Z}|Y = y\}$, and for each sample we evaluate the entire log-ratio expression in (5.25). The MC estimate of $I(Y = y; X|\boldsymbol{Z})$ is then computed as the average of these expressions across all sampled instances.

Computation on simulated data. This section presents a series of experiments on synthetically generated data designed in a way such that the underlying generative models are fully specified, allowing to capture the distinct components of information in structured settings. The goals of these experiments are to investigate the true values of the information-theoretic contributions of individual feature variables as a function of the simulation parameters, and to quantify the bias and variance of the proposed estimator

relative to the theoretical values of the information measures. To this end, first a series of simple setups that focus on showcasing each possible type of individual contributions of the feature variables to the target (i.e., unique, redundant, and synergistic) were construct, and then an additional, more complex simulation covering all possible types of interaction was designed.

In detail, let consider a class variable Y and a set of m feature variables $\boldsymbol{X} = \{X_1, \dots, X_m\}$; the features are assumed to have a multivariate Gaussian distribution when associated with each specific outcome of the class, i.e. $\boldsymbol{X}|Y = y \sim \mathcal{N}(\mu_{\boldsymbol{X}}^{(y)}, \Sigma_{\boldsymbol{X}}^{(y)})$, where $\mu_{\boldsymbol{X}}^{(y)}$ and $\Sigma_{\boldsymbol{X}}^{(y)}$ are the mean vector and the covariance matrix of $\boldsymbol{X}$ evaluated when $Y = y$. Inspired by the examples in [163, 247], a set of simulations is outlined to induce behaviors that elicit specific types of information contributions. This is achieved by imposing changes in the mean of a feature variable across class states (i.e., assigning different values of $\mu_{X_i}^{(y)}$ at varying y) in order to reproduce relevant contributions, introducing different levels of high-order interaction determined by the correlation among features (i.e., by the off-diagonal elements of $\Sigma_{\boldsymbol{X}}^{(y)}$), and imposing no changes in the mean and variance of a feature across states to reproduce an absence of contribution. In the first three simulations only $m = 2$ features are considered, showing how unique, synergistic, or redundant behaviors can be isolated through appropriate choices of the simulation parameters. In the final simulation, with $m = 6$, it is demonstrated how combined changes in the distribution parameters of several features can give rise to coexisting contributions. The kNN estimation method is applied to simulations consisting of $N_{sample} = 1000$ samples per realization $Y = y$, considering the typical setting $k = 10$ for the number of neighbors [21, 248]. The statistical significance of the selected variables is tested by generating $N_{surr} = 100$ surrogate sequences. Each simulation is repeated $N_{sim} = 100$ times to assess the confidence intervals of each measure, while theoretical values are obtained using $N_{MC} = 10^6$ Monte Carlo samples for each realization $Y = y$.

Unique components. First, let consider a 2-class ($\mathcal{A}_Y = \{y_1, y_2\}$), 2-variable ($\boldsymbol{X} = \{X_1, X_2\}$) setup with the following model parameters:

$$\begin{aligned} P(y_1) &= P(y_2) = 0.5, \\ \mu_{\boldsymbol{X}}^{(y_1)} &= [1, 1], \\ \mu_{\boldsymbol{X}}^{(y_2)} &= [-1, 1], \\ \Sigma_{\boldsymbol{X}}^{(y_1)} &= \Sigma_{\boldsymbol{X}}^{(y_2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned} \tag{5.26}$$

The class-specific distributions are $X_1|Y = y_1 \sim \mathcal{N}(1, 1)$, $X_1|Y = y_2 \sim \mathcal{N}(-1, 1)$ and $X_2|Y = y_1 \equiv X_2|Y = y_2 \sim \mathcal{N}(1, 1)$, showing that the mean of X_1 shifts between the two

classes while the distribution of X_2 remains the same. The diagonal covariance implies that the two features are uncorrelated.

The results of the analysis are shown in Fig. 5.2(a), documenting that the predictive information is fully determined by the purely unique contribution of X_1, which can be ascribed to the shift in mean between the two class labels; The feature X_2, being uncorrelated with X_1 and equally distributed within the two classes, displays no relevancy to the class variable Y; the uncorrelation between the features explains also the absence of redundant or synergistic effects. The distribution of the information measures estimated over N_{sim} realizations shows the accuracy of the estimates, which exhibit a small negative bias and limited variability.

Synergistic components. This simulation reproduces again a 2-class, 2-variable setup, now characterized by the following parameters:

$$\begin{aligned} P(y_1) &= P(y_2) = 0.5, \\ \mu_X^{(y_1)} &= \mu_X^{(y_2)} = [0, 0], \\ \Sigma_X^{(y_1)} &= \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \\ \Sigma_X^{(y_2)} &= \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}, \end{aligned} \tag{5.27}$$

corresponding to identical standard normal distributions of the two feature variables within the two classes ($X_i|Y = y_j \sim \mathcal{N}(0, 1)$, $i, j = 1, 2$). Moreover, as seen in the covariance matrix, the two features are positively correlated when the class takes the value y_1, and anticorrelated when the class takes the value y_2.

The results for this setup, reported in Fig. 5.2(b), document the absence of unique contributions, reflecting the fact that the class-specific mean and variance remained unchanged across the realizations of Y. However, their joint interaction, here captured by the class-specific covariance, carries discriminative information about the class, which is encoded in the significant synergistic contribution observed in the absence of redundancy. In this case, synergy arises by the fact that the relationship between the features differs across the realizations of Y, so that both X_1 and X_2 need to be known to predict the class.

Redundant components. Another 2-class, 2-variable setup is then considered, featuring

the following parameters:

$$\begin{aligned} P(y_1) &= P(y_2) = 0.5, \\ \mu_X^{(y_1)} &= [1, 1], \\ \mu_X^{(y_2)} &= [-1, -1], \\ \Sigma_X^{(y_1)} &= \Sigma_X^{(y_2)} = \begin{bmatrix} 1 & 0.99 \\ 0.99 & 1 \end{bmatrix}. \end{aligned} \tag{5.28}$$

In this case, the class-specific distributions are the same for the two features, i.e. $X_i|Y = y_1 \sim \mathcal{N}(1, 1)$, $X_i|Y = y_2 \sim \mathcal{N}(-1, 1)$, $i = 1, 2$, showing a shift in the mean between the two classes. Moreover, the features are strongly correlated as documented by the covariance value of 0.99.

The analysis of this configuration reveals a fully overlapping contribution of the two features to the prediction of the class labels, as seen in Fig. 5.2(c) where the predictive information is entirely redundant, with zero values for the unique and synergistic components. These results are explained by the fact that the two features share the same class-dependent change in the mean value.

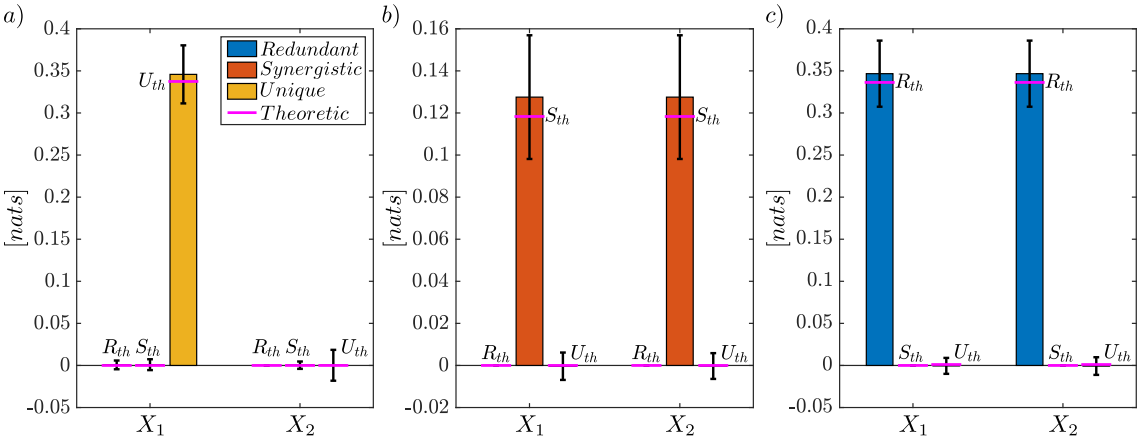


Figure 5.2: Decomposition into unique, redundant and synergistic components of the maximum information shared between the binary class variable Y and the feature X_i ($i = 1, 2$) for the simulated systems with parameters given by Eq. 5.26 (a), Eq. 5.27 (b), Eq. 5.28 (c). Barplots represent the mean values (with ±2 SD) of the various information contributions, for features X_1 and X_2 showcasing unique (a), synergistic (b), and redundant (c) predictive information; magenta lines represent the theoretical values. This figure has been taken from [131].

Multiple interactions. The last simulation, which displays a setup with more complex interactions, including combined contributions of multiple components, is designed to portray the robustness of the proposed estimation strategy, considering a higher-dimensional interaction system. Specifically, a 2-class ($\mathcal{A}_Y = \{y_1, y_2\}$), 6-variable ($\boldsymbol{X} = \{X_1, \ldots, X_6\}$)

system is formed with the following model parameters:

$$
\begin{aligned}
P(y_1) &= P(y_2) = 0.5, \\
\mu_X^{(y_1)} &= [1, 0.5, 0.5, 0, 1, 0], \\
\mu_X^{(y_2)} &= [-1, -0.5, -0.5, 0, -1, 0], \\
\Sigma_X^{(y_1)} &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.25 & 0 & 0 & 0 \\ 0 & 0.25 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \\
\Sigma_X^{(y_2)} &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.25 & 0 & 0 & 0 \\ 0 & 0.25 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -0.5 & 0 \\ 0 & 0 & 0 & -0.5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.
\end{aligned}
\tag{5.29}
$$

According to 5.29, the joint distribution of the features display the following characteristics: the feature X_1 exhibits different class-specific mean while remaining uncorrelated with the other features; X_2 and X_3 have the same different class-specific means and additionally display a positive correlation for both realizations of Y; while X_4 does not exhibit a change in the class-specific mean, X_5 does, and they jointly have a class-dependent covariance value; the feature X_6 has no class-specific change in the probability distribution and is uncorrelated with the other features.

Fig. 5.3 shows the results of the kNN estimator applied to the defined system, for which several expected contributions are observed as follows. The change in the mean value across class realizations enables the variables X_1, X_2, X_3, and X_5 to be predictive for the class, thereby providing unique contribution, with a higher magnitude for X_1 and X_5 which display the largest mean shift. Moreover, since these features exhibit similar class-dependent changes in the mean, their contribution is also partially redundant; it is noted that redundancy arises not only for the correlated variables X_2 and X_3, but also for X_1 and X_5 which are not correlated with each other or with X_2 and X_3. Additionally, a synergistic contribution is detected for X_4 and X_5, reflecting the change in sign of their correlation across the output values of the class variable Y. Finally, X_6 does not provide any contribution to Y, as expected by the fact that its probability density is not class-dependent

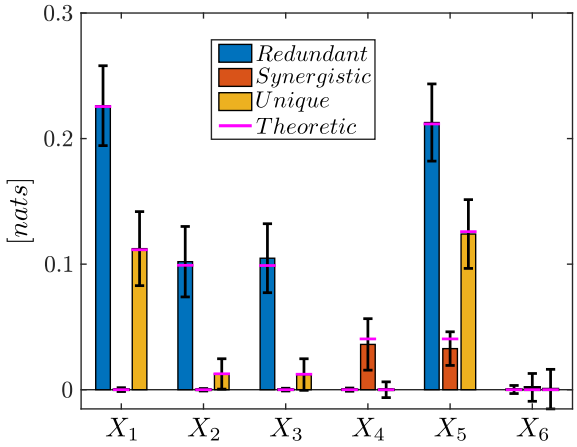


Figure 5.3: Decomposition into unique, redundant and synergistic components of the maximum information shared between the binary class variable Y and the feature X_i ($i = 1, \ldots, 6$) for the simulated systems with parameters given by Eq. 5.29. Barplots represent the mean values (with ± 2 SD) of the various information contributions estimated across several realizations, while magenta lines represent the theoretical values. This figure has been taken from [131].

and it is uncorrelated with the other features. Remarkably, the accuracy of the kNN estimation remains high also in this higher-dimensional system, as documented by the very small bias and limited variance of the estimates of unique, redundant and synergistic components computed across N_{sim} realization (barplots and errorbars in Fig. 5.3).

Fig. 5.4 depicts the selection of the variables included for this simulation into the conditioning sets $\boldsymbol{Z}_{min}$ and $\boldsymbol{Z}_{max}$ during the search for redundant and synergistic contributions to the class label, both theoretical based on the exact CMI values and practical based on the estimated values and the surrogate test for significance. Moreover, Fig. 5.5 shows the order according to which the variables are selected by the sequential procedure for the formation of the sets $\boldsymbol{Z}_{min}$ and $\boldsymbol{Z}_{max}$.

The group of features X_1, X_2, X_3 or X_5 are selected into the conditioning set $\boldsymbol{Z}_{min}$ when one of them is taken as source variable (Fig. 5.4a, columns 1,2,3,5), in an order generally prioritizing X_1 and X_5 followed by X_2 and X_3 (Fig. 5.5a). This selection is expected given the previously discussed marginal predictive capability of these features, which leads to redundant contributions. Moreover, X_4 is also included theoretically into the set $\boldsymbol{Z}_{min}$ when X_1, X_2, or X_3 are taken as sources; in the practical estimation this selection occurred only in a fraction of realizations (89%,16% and 17% for X_1, X_2, and X_3, Fig. 5.4a) and at the third or fourth iteration of the searching procedure (Fig. 5.5a). The inclusion of X_4 is more nuanced, as it exhibits a synergistic interaction with X_5, and its selection along with X_5 further decreases the CMI.

With regard to the procedure of maximization of the CMI, the only significant selec-

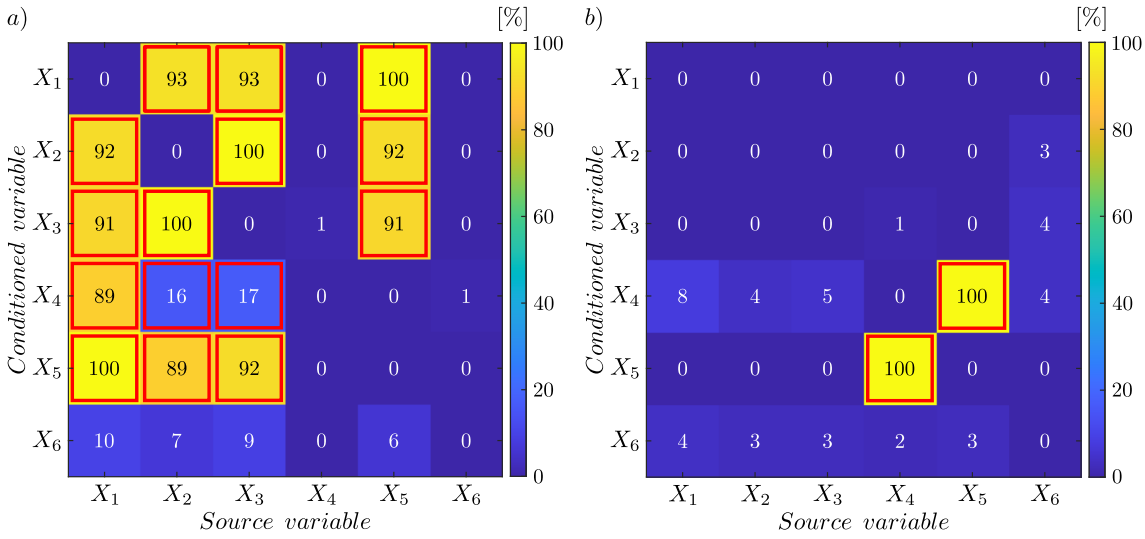


Figure 5.4: Selection of the features to be included in the conditioning set during the search for redundancy and synergy in the simulated system with parameters given by Eq. 5.29. Heatmaps show, for any source variable (columns), the selection of the conditioning variables (rows; cells with red border: theoretical selection; cell color and number: percentage of selections over N_{sim} realizations) included within the set $\boldsymbol{Z}_{min}$ (a) or within the set $\boldsymbol{Z}_{max}$ (b). In panel a, the source variables X_1, X_2, X_3, and X_5 consistently contribute to minimizing the CMI value, with X_4 further decreasing it, except when the starting variable is X_5. Panel b shows that the only synergy-induced CMI increase happens between X_4 and X_5. This figure has been taken from [131].

tions involve X_4 and X_5, which are always reciprocally included into $\boldsymbol{Z}_{max}$ at the first iteration (Fig. 5.4b, Fig. 5.5b). This selection is meaningful as it relates to the synergistic contribution to the class labels expected for the features X_4 and X_5. Only minor spurious inclusions, not expected theoretically and related to false positive selections, are observed in the other cases.

Application to non-Gaussian data. To investigate the behavior of the kNN CMI estimator when dealing with distributions for which the theoretical values of the information components cannot be computed, two validation scenarios were designed: one using synthetic data and the other using real-world data. In the synthetic case, the results were interpreted according to the expected interaction effects based on the experimental setup, while in the real-world case, the findings were validated against known analyses reported in the literature.

Synthetic data. In this setup, four independent uniform distributions are simulated, $X_i \sim \mathcal{U}(-1, 1)$, $i \in 1, \ldots, 4$. The discrete class variable Y is defined as $Y = \Theta(X_1 \cdot X_2) + \Theta(X_3) + \Theta(X_4)$, where $\Theta(\cdot)$ denotes the Heaviside function, resulting in a class alphabet $A_Y = 0, 1, 2, 3$. Across $N_{sim} = 100$ simulations, the proposed estimator receives as inputs the class variable Y and the set of feature variables $X_1, \ldots, X_4$, together with two additional

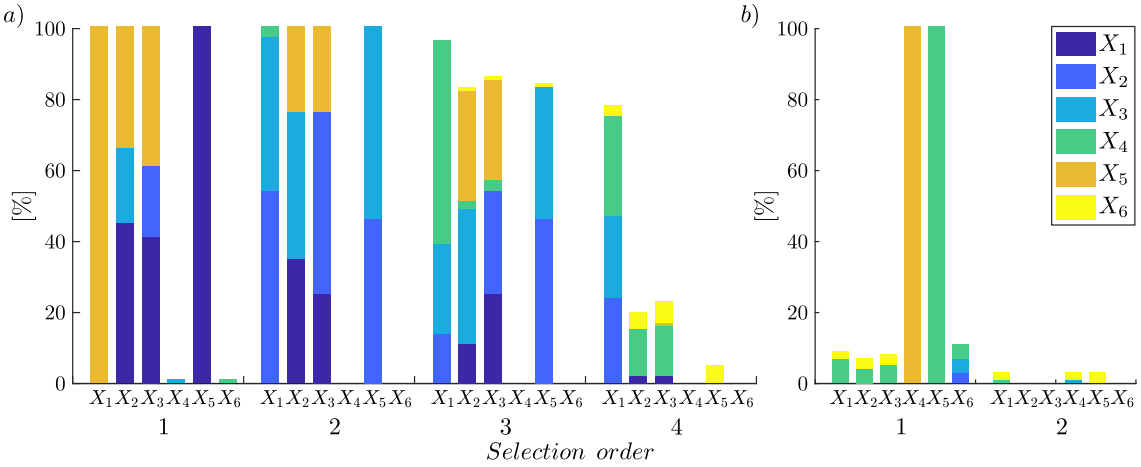


Figure 5.5: Order and percentage of occurrence of the features selected as conditioning variables for the MI between each source feature and the target class during the search for redundancy (a, inclusion into $\boldsymbol{Z}_{min}$) and synergy (b, inclusion into $\boldsymbol{Z}_{max}$). The height of each bar quantifies the percentage of times over N_{sim} realizations for which a variable was selected for the source feature X_i ($i = 1, \ldots, 6$) at each iteration $j \in \{1, 2, \cdots\}$; the color information represents which variable was selected. This figure has been taken from [131].

features: X_5, defined as $X_5 = X_4 + M$ (a noisy version of X_4 with $M \sim \mathcal{N}(0, 1)$), and an independent uniform variable $X_6 \sim \mathcal{U}(-1, 1)$. Under this configuration, a synergistic effect is expected between X_1 and X_2, as they interact multiplicatively (similarly to the XOR logic gate), while X_3 and X_4 are expected to provide unique contributions. Moreover, due to the additive formulation of the class variable Y, a certain degree of synergistic interaction among these four variables is also anticipated. The variable X_5, being a noisy version of X_4, is expected to share with it a certain amount of redundant predictive information. Finally, X_6, being independent of all other variables, is expected to show no contribution to the target class. To evaluate the sensitivity of the measures to the sample size, the analysis was repeated using $N_{sample} = 1000$ and $N_{sample} = 300$ as sample sizes.

The results reported in Fig. 5.6 confirm the expected synergistic interaction between X_1 and X_2, as well as the unique contributions of X_3 and X_4. In addition, redundancy is correctly identified for X_4, since X_5 represents a noisy version of it. Finally, X_6, which is independent of the other variables, shows no contribution with respect to the class. The expected synergistic effect arising from the additive interaction of X_1, X_2 with X_3 and X_4 is also confirmed. Notably, the observed synergistic effect assumes lower values when a smaller sample size is used.

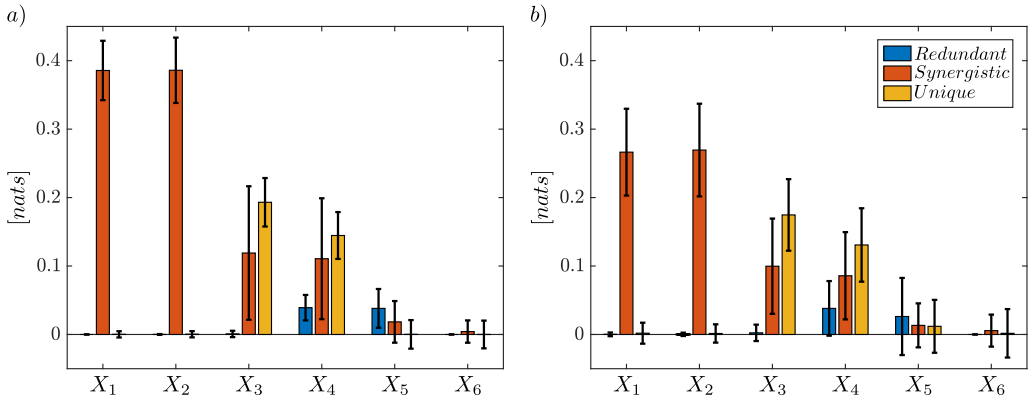


Figure 5.6: Decomposition into unique, redundant and synergistic components of the maximum information shared between the 4-class variable Y and each of the six source feature variables simulated as in Sect. 4.1. Barplots represent the mean value (±2 SD) of the various information contributions estimated across several realizations lasting $N_{sample} = 1000$ (a) and $N_{sample} = 300$ (b). This figure has been taken from [131].

Table 5.1: Comparison of feature-importance methods employed.

Method	Core idea	Strengths	Limitations
Coefficient-based FI	Uses model coefficients (e.g., linear or logistic regression weights) as a measure of feature relevance.	Simple and interpretable; computationally efficient; directly linked to model parameters.	Model-dependent; assumes linearity; unreliable in the presence of correlated predictors or nonlinear effects.
Impurity-based FI	Measures feature importance from the decrease in node impurity (e.g., Gini or entropy) in tree-based models.	Fast; naturally integrated in tree ensembles; captures some nonlinear effects.	Biased toward variables with many splits; model-specific; limited interpretability across models.
Permutation FI	Quantifies importance by measuring performance degradation after random permutation of a feature.	Model-agnostic; directly linked to predictive performance; captures nonlinear dependencies.	Sensitive to feature correlations; computationally expensive; unstable with small datasets.
SHAP	Attributes feature contributions based on Shapley values from cooperative game theory.	Consistent and locally accurate explanations; applicable to black-box models; strong theoretical guarantees.	High computational cost; approximation required for complex models; local explanations may not generalise globally.
HiFi	Decomposes feature contribution into unique, redundant, and synergistic components via predictability loss.	Explicitly quantifies higher-order dependencies; structured decomposition of multivariate relevance.	Combinatorial search cost; relies on greedy approximations; predictability-based and model-dependent.
CMIHiFi	Reformulates HiFi purely in terms of Conditional Mutual Information using non-parametric kNN estimation.	Fully information-theoretic; model-independent; captures redundancy and synergy in mixed discrete–continuous data.	High computational cost; sensitive to sample size and neighbourhood parameter; requires careful density estimation.

The HiFi and CMIHiFi frameworks represent a paradigm shift in the quantification of feature relevance. Unlike conventional approaches, they do not evaluate importance through model-specific parameters or perturbations of predictive accuracy, but through a direct decomposition of the informational structure underlying the data. By distinguishing between unique, redundant, and synergistic components, these methods allow one to understand whether a feature contributes independent predictive content, shares information with others, or acts cooperatively within a multivariate context.

From a theoretical standpoint, both HiFi and CMIHiFi satisfy desirable properties of additivity, symmetry, and model independence. They provide an additive decomposition of total information, remain invariant under invertible transformations of the variables, and require no assumptions on the underlying data distribution. The equivalence between the predictability-based and purely information-theoretic formulations further unifies regression-based and entropy-based perspectives within a single framework.

Ultimately, HiFi and CMIHiFi establish a rigorous foundation for understanding high-order feature interactions and for extending the concept of feature importance beyond first-order dependencies. Their ability to reveal the multivariate informational architecture of the feature space makes them particularly suitable for analysing complex datasets where redundancy and synergy coexist, a scenario that traditional importance metrics are unable to capture.

The framework presented in this chapter establishes feature importance as a key analytical layer bridging feature selection and supervised classification. Rather than representing the final stage of the analytical process, FI serves as a transitional framework that connects the identification of informative features with their interpretative and predictive roles within learning models. Its primary purpose is to quantify, in a principled way, how the selected features contribute, individually or jointly, to the model output, thereby enriching the understanding of the underlying information structure that drives predictive performance.

The comparative synthesis in Table 5.1 highlights a clear methodological progression. Traditional FI methods offer straightforward and computationally efficient means of ranking variables. However, their interpretive scope is inherently constrained: they measure the magnitude of association or performance change attributable to each feature in isolation, without distinguishing whether that contribution is unique, shared, or cooperative. Consequently, such approaches often fail to capture the multivariate dependencies and informational overlap that characterize complex systems.

The information-theoretic framework introduced through the HiFi and CMIHiFi methodologies provides a substantial conceptual advance in this respect. By redefining feature

relevance in terms of mutual and conditional mutual information, these methods capture not only the amount of information each feature contributes about the target variable but also the structure of how that information is distributed among multiple predictors. This decomposition into unique, redundant, and synergistic components allows a granular understanding of the collective organization of information within the feature space. Unlike traditional rankings, the resulting representation conveys how predictive knowledge emerges from independent, overlapping, or cooperative effects.

Beyond its theoretical formulation, the information-theoretic view of FI promotes a more general shift in perspective. Feature relevance ceases to be a property tied to a specific model or algorithm; instead, it becomes an intrinsic characteristic of the data distribution itself. By leveraging conditional mutual information and non-parametric estimators such as the kNN approach, the CMIHiFi framework quantifies dependencies without imposing linearity or parametric assumptions, ensuring invariance under reparameterisation and applicability across models. The trade-off, however, lies in increased computational cost and sensitivity to sampling density, which necessitate careful methodological control to ensure robust estimation.

From a broader standpoint, the introduction of high-order feature importance reshapes the interpretative function of feature analysis. It allows one to move beyond the question of "which features matter most" toward the more fundamental question of "how features matter, and in what combinations". This perspective integrates explainability and information theory into a unified analytical lens, positioning FI as a methodological and conceptual bridge between feature selection and classification. It offers a principled way to connect quantitative modelling with mechanistic understanding, ensuring that subsequent classification analyses are not only accurate but also interpretable in terms of the informational architecture of the underlying system.

In summary, feature importance provides a rigorous, model-agnostic means to examine the informational roles of features within predictive frameworks. The transition from traditional to information-theoretic methodologies marks a conceptual maturation: from assessing relevance through model performance to dissecting the structure of information itself. In the context of this thesis, FI thus serves as a critical interpretative step linking the extraction and selection of features with their integration into supervised classification models, forming the theoretical foundation for the analyses that follow.

6

Classification Methods

The classification stage represents the final step of the analytical framework developed in this thesis. After extracting, selecting, and characterising features derived from cardiovascular time series, the objective is to construct models capable of discriminating among physiopathological states or between demographic markers based on these descriptors. Classification thus represents the transition from the descriptive domain, where signal dynamics are quantified through statistical and nonlinear indices, to the inferential domain, where these features are combined to produce meaningful and reproducible decisions [26, 31, 88, 111, 114, 127].

In the context of physiological signal analysis, classification aims to map multivariate feature representations into categorical outcomes, i.e., stress versus rest, male versus female, or healthy versus pathological conditions [39, 40, 56, 79, 119, 135, 168]. From a methodological perspective, classification plays a dual role. On one hand, it provides a quantitative validation of the relevance of the extracted features: the ability to correctly distinguish physiological conditions serves as indirect evidence that the chosen features encode meaningful physiological information. On the other hand, it offers a practical framework for predictive modelling in biomedical applications, enabling the automated identification of states or pathophysiological patterns that may be difficult to discern through conventional qualitative assessment or univariate analysis alone [14, 230].

However, classification in the biomedical domain presents several inherent challenges. Physiological datasets are often characterised by small sample sizes, high inter-subject variability, and non-Gaussian distributions. These properties make conventional statistical assumptions inadequate and increase the risk of model overfitting [66, 82, 85, 189]. Furthermore, physiological signals are influenced by multiple regulatory mechanisms act-

ing at different temporal scales, resulting in nonlinear and multidimensional dependencies among features. Consequently, selecting a suitable classifier requires balancing flexibility, interpretability, and robustness to noise and data imbalance [93].

In this thesis, we focus on *supervised learning methods*, which represent the most widely adopted approach for physiological classification tasks [39, 40, 111, 127]. In supervised settings, each sample in the dataset is associated with a known label, allowing the model to learn how specific feature patterns correspond to predefined classes. This labelled structure provides the ground truth required to infer a mapping capable of assigning the correct category to unseen observations.

Supervised learning stands in contrast to unsupervised methods, where no labels are provided and the goal is to uncover latent structure, and to reinforcement learning, in which models optimise sequential decisions based on reward signals rather than explicit class assignments [114, 127]. A typical supervised workflow comprises three stages: training, validation, and testing. During training, the model adjusts its parameters to approximate the relationship between features and labels; validation helps regulating model complexity and reducing over-fitting by evaluating performance on data not used for parameter optimisation; finally, testing assesses the model's capacity to generalise to previously unseen samples [114, 127].

Supervised classification forms the core of most physiological state recognition pipelines [3, 31, 114, 127, 135, 196]. The problem can be formally expressed as the estimation of a function:

$$f : \mathbb{R}^d \rightarrow \{1, 2, \ldots, C\}, \tag{6.1}$$

where each feature vector, $\mathbf{X_N} \in \mathbb{R}^d$, is assigned to one of C predefined classes. This mapping is learned from a labelled dataset $\mathcal{D} = \{(\mathbf{X}, Y)\}$, in which each observation $\mathbf{x}_N$ is associated with a known physiological label y_N. The classifier then generalises this learned relationship to new, unseen data, providing a predictive model that captures the discriminative structure of the underlying physiological processes [24, 48, 63, 101].

The learning process typically involves minimising a loss function $\mathcal{L}(f(\mathbf{x}_i), y_i)$ over the training dataset, where $\mathcal{L}$ measures the discrepancy between predicted and actual labels. Once the parameters of the model are optimised, the classifier can be evaluated on unseen data to assess its generalisation ability. Reliable performance estimation requires robust validation strategies, such as k-fold cross-validation or leave-one-subject-out approaches, particularly important in small and heterogeneous biomedical datasets [31, 114, 120, 127].

Several families of algorithms are considered, ranging from classical linear approaches to modern nonlinear and information-theoretic models. Specifically, we describe *Linear Discriminant Analysis* (LDA), *Support Vector Machines* (SVM), *k-Nearest Neighbours*

(kNN), *Naive Bayes* (NB), *Neural Networks* (NN), *Random Forests* (RF), and we introduce a new information-theoretic approach denoted as *Local Information Classifier* (LIC) [31, 114, 127]. Each of these methods embodies distinct theoretical assumptions and inductive biases, offering complementary perspectives on how physiological information can be represented and discriminated. The choice of the classification algorithm to be employed depends on the nature of the data, the number of available samples, and the required trade-off between accuracy, interpretability, and computational complexity. Linear models offer transparency and analytical tractability, whereas nonlinear or ensemble methods provide greater flexibility to capture complex decision boundaries. Through their comparison, this chapter aims to highlight the conceptual and methodological differences among these classifiers, as well as their implications for physiological interpretability and predictive reliability. The following sections provide a detailed overview of supervised classification principles, the mathematical foundations of the implemented algorithms, and the validation procedures used to assess their performance. The use of supervised models enables the establishment of quantitative and reproducible relationships between feature patterns and physiological labels.

The remainder of this chapter is organised as follows. Section 6.1 introduces the set of traditional supervised classifiers implemented in this work, detailing their theoretical formulation and mathematical properties, including linear, nonlinear, probabilistic, and ensemble approaches. Section 6.2 presents the information-theoretic *Local Information Classifier* (LIC), a model-free method based on local entropy and mutual information estimation, particularly suited for biomedical datasets characterised by non-Gaussian and heterogeneous distributions. Section 6.3 describes the adopted procedures for training, testing, and validation, as well as the performance metrics used to assess the robustness and generalisation of the classifiers. Finally, the chapter concludes with a comparative discussion summarising the relative advantages, limitations, and interpretability of the implemented methods, and placing classification within the broader methodological framework established through feature extraction, selection, and importance.

6.1 Traditional Classifiers

6.1.1 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is one of the most established and interpretable classification techniques [3, 31, 114, 127, 135, 196]. It assumes that the data from each class $c \in \{1, \ldots, C\}$ follow a multivariate Gaussian distribution with a common covariance matrix Σ but class-specific mean vectors $\boldsymbol{\mu}_c$. The class-conditional probability density is

expressed as

$$p(\mathbf{x} \mid y = c) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}_c)\right), \tag{6.2}$$

where $\mathbf{x} \in \mathbb{R}^d$ denotes a d-dimensional feature vector, $\boldsymbol{\mu}_c$ is the mean vector of class c, and Σ is the covariance matrix shared across all classes. The determinant $|\Sigma|$ acts as a normalization factor, while the quadratic form in the exponent represents the squared Mahalanobis distance between the observation $\mathbf{x}$ and the class mean $\boldsymbol{\mu}_c$.

Applying Bayes' theorem [113] yields the discriminant function

$$\delta_c(\mathbf{x}) = \mathbf{x}^\top \Sigma^{-1} \boldsymbol{\mu}_c - \frac{1}{2} \boldsymbol{\mu}_c^\top \Sigma^{-1} \boldsymbol{\mu}_c + \ln \pi_c, \tag{6.3}$$

where $\pi_c = P(y = c)$ represents the prior probability of class c. The first term corresponds to a linear projection of the feature vector weighted by the inverse covariance matrix, the second term is a class-dependent offset related to the magnitude of the class mean, and the logarithmic prior term incorporates prior information about class prevalence.

The predicted class corresponds to the largest discriminant value,

$$\hat{y} = \arg\max_c \ \delta_c(\mathbf{x}). \tag{6.4}$$

Because the covariance matrix is common to all classes, the resulting decision boundaries are linear hyperplanes in the feature space.

LDA is computationally efficient and performs well when classes are approximately linearly separable and follow Gaussian distributions. Its linearity and interpretability make it particularly suitable for biomedical applications where the contribution of each feature must be clearly understood [39, 111].

6.1.2 Support Vector Machines (SVM)

Support Vector Machines (SVM) are powerful classifiers designed to maximize the margin between different classes [114, 127]. For a binary classification problem with training samples $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, where $\mathbf{x}_i \in \mathbb{R}^d$ denotes a d-dimensional feature vector and $y_i \in \{-1, +1\}$ is the corresponding class label, SVM seeks the separating hyperplane

$$\mathbf{w}^\top \mathbf{x} + b = 0, \tag{6.5}$$

where $\mathbf{w}$ is the normal vector defining the orientation of the hyperplane and b is a bias term controlling its offset from the origin. The margin is defined as the distance between

the hyperplane and the closest training samples, and maximizing this margin corresponds to minimizing the norm of $\mathbf{w}$.

To account for non-separable data, SVM introduces slack variables $\xi_i \geq 0$, which quantify the degree of violation of the margin constraints for each training sample. Specifically, $\xi_i = 0$ indicates that the sample is correctly classified and lies outside the margin, $0 < \xi_i < 1$ corresponds to a correctly classified sample within the margin, and $\xi_i > 1$ denotes a misclassified observation. The resulting optimization problem is formulated as

$$\min_{\mathbf{w},b,\boldsymbol{\xi}} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N} \xi_i, \tag{6.6}$$

subject to

$$y_i(\mathbf{w}^\top \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \tag{6.7}$$

where the regularization parameter $C > 0$ controls the trade-off between maximizing the margin and penalizing classification errors. Larger values of C enforce stricter adherence to the training data, while smaller values allow a wider margin at the cost of increased misclassification.

For nonlinear classification problems, SVM employs a kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$ that implicitly maps the input data into a higher-dimensional feature space, where a linear separating hyperplane can be constructed. This kernel-based formulation enables nonlinear decision boundaries without explicitly computing the feature-space transformation. Commonly used kernels include the radial basis function (RBF),

$$K_{\text{RBF}}(\mathbf{x}_i, \mathbf{x}_j) = \exp\big(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2\big), \tag{6.8}$$

where the parameter $\gamma > 0$ controls the spatial scale of the kernel and determines the influence range of individual training samples, and the polynomial kernel,

$$K_{\text{poly}}(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^\top \mathbf{x}_j + 1)^p, \tag{6.9}$$

where the degree p regulates the complexity of the resulting decision function.

SVMs are particularly effective in high-dimensional feature spaces and often perform well with limited sample sizes, making them a popular choice for physiological data classification [40, 93, 97].

6.1.3 k-Nearest Neighbours (k-NN)

The *k-Nearest Neighbours* (k-NN) classifier is a non-parametric, instance-based method that assigns a class label to a sample based on the labels of its k nearest neighbours in the feature space [31, 52, 114, 127]. Given a test instance $\mathbf{x} \in \mathbb{R}^d$, the distance to each training sample $\mathbf{x}_i$ is computed, commonly using the Euclidean metric,

$$d(\mathbf{x}, \mathbf{x}_i) = \sqrt{\sum_{j=1}^{d} (x_j - x_{ij})^2}, \tag{6.10}$$

where x_j and x_{ij} denote the j-th feature of the test sample and of the i-th training sample, respectively. This distance measure quantifies the similarity between observations in the feature space.

The predicted class label is determined by majority voting among the k nearest neighbours,

$$\hat{y} = \arg\max_{c} \sum_{i \in \mathcal{N}_k(\mathbf{x})} \mathbb{I}(y_i = c), \tag{6.11}$$

where $\mathcal{N}_k(\mathbf{x})$ denotes the set of indices corresponding to the k closest training samples to $\mathbf{x}$, and $\mathbb{I}(\cdot)$ is the indicator function, which equals one if its argument is true and zero otherwise. The parameter k controls the locality of the decision rule, with smaller values leading to more flexible but potentially noisy classifiers, and larger values yielding smoother decision boundaries.

The k-NN algorithm is simple and effective when the feature space is well-structured. However, it is sensitive to the choice of k and to feature scaling, as distance-based comparisons can be strongly influenced by differences in feature magnitudes. These aspects must be carefully controlled in physiological datasets.

6.1.4 Naive Bayes (NB)

Naive Bayes is a probabilistic classifier grounded in Bayes' theorem and characterized by the assumption of conditional independence among features given the class label [31, 114, 127]. For a feature vector $\mathbf{x} = (x_1, \ldots, x_d)$, the posterior probability of class c is computed as

$$P(y = c \mid \mathbf{x}) = \frac{P(y = c) \prod_{j=1}^{d} P(x_j \mid y = c)}{P(\mathbf{x})}, \tag{6.12}$$

where $P(y = c)$ denotes the prior probability of class c, $P(x_j \mid y = c)$ is the class-conditional probability of the j-th feature, and $P(\mathbf{x})$ is a normalization constant independent of the class label.

The classification rule follows from maximizing the posterior probability,

$$\hat{y} = \arg\max_{c} \; P(y = c) \prod_{j=1}^{d} P(x_j \mid y = c), \tag{6.13}$$

which highlights how the contribution of each feature is combined multiplicatively under the independence assumption. Despite this strong simplification, Naive Bayes often performs well in high-dimensional and small-sample scenarios, as the number of parameters to estimate grows only linearly with the number of features.

In biomedical applications, the Gaussian Naive Bayes variant is commonly adopted, assuming that continuous features follow class-specific normal distributions. This formulation provides a simple and computationally efficient baseline for exploratory analyses of physiological data [14, 230].

6.1.5 Neural Networks (NN)

Artificial Neural Networks (NN) are flexible nonlinear models inspired by biological neural systems [2, 26, 31, 114, 127]. They consist of layers of interconnected neurons, where each neuron computes a weighted combination of its inputs followed by a nonlinear activation function. For the l-th layer, the transformation is given by

$$z^{(l)} = f\left(\mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}\right), \tag{6.14}$$

where $\mathbf{W}^{(l)}$ is the weight matrix, $\mathbf{b}^{(l)}$ is the bias vector, and $\mathbf{a}^{(l-1)}$ denotes the activations of the previous layer. The nonlinear activation function $f(\cdot)$, such as the Rectified Linear Unit (ReLU), sigmoid, or hyperbolic tangent, enables the network to model complex nonlinear relationships.

The set of network parameters is denoted by $\Theta = \{\mathbf{W}^{(l)}, \mathbf{b}^{(l)}\}$ and is optimized by minimizing a loss function over the training data. For classification tasks, this is typically the cross-entropy loss,

$$\min_{\Theta} \; \mathcal{L} = -\sum_{i=1}^{N}\sum_{c=1}^{C} \mathbb{I}(y_i = c) \; \log \hat{P}(y_i = c \mid \mathbf{x}_i; \Theta), \tag{6.15}$$

where $\hat{P}(y_i = c \mid \mathbf{x}_i; \Theta)$ denotes the predicted posterior probability for class c. Neural networks can approximate highly nonlinear decision boundaries, but they require sufficient training data and appropriate regularization strategies to prevent overfitting. In physiological classification tasks, shallow architectures with one or two hidden layers are often preferred due to limited dataset sizes [39, 93].

6.1.6 Random Forest (RF)

Random Forest (RF) is an ensemble learning method that combines the predictions of multiple decision trees to improve robustness and generalization performance [114, 127]. Each tree is trained on a bootstrap sample of the original dataset, and at each internal node a random subset of features is considered when determining the optimal split. This randomness reduces correlations among individual trees and mitigates overfitting.

For classification tasks, the final prediction is obtained by majority voting across the ensemble,

$$\hat{y} = \arg\max_{c} \sum_{m=1}^{M} \mathbb{I}\big(h_m(\mathbf{x}) = c\big), \tag{6.16}$$

where $h_m(\mathbf{x})$ denotes the class predicted by the m-th tree and M is the total number of trees in the forest. This aggregation strategy stabilizes the prediction by averaging over multiple diverse models.

Random Forests typically exhibit low variance and strong generalization capabilities compared to single decision trees. Moreover, they naturally provide estimates of feature importance based on split statistics, which are valuable for model interpretation. These properties make RFs particularly suitable for biomedical applications, where both predictive accuracy and interpretability are essential [39, 97].

While these classical classifiers provide complementary perspectives on decision boundaries, their performance and interpretability can vary depending on data heterogeneity and class imbalance. For this reason, recent information-theoretic approaches have been proposed to complement traditional methods with estimators grounded in entropy and mutual information, capable of capturing high-order dependencies in physiological data [17, 98, 130, 131].

6.2 Novel Information-Theoretic Classification Method

Beyond conventional discriminative algorithms, information-theoretic approaches provide a powerful and conceptually grounded alternative framework for classification [215, 236]. Rather than relying on explicit decision boundaries, these methods quantify uncertainty and statistical dependence between variables through, enabling the identification of informative patterns even in the presence of nonlinear and high-dimensional relationships. This perspective is particularly appealing in biomedical applications, where interpretability, robustness to model assumptions, and sensitivity to complex interactions are essential.

Within this framework, the *Local Information Classifier* (LIC) represents a recent and significant contribution, combining the interpretability of information measures with

the flexibility of non-parametric, data-driven estimation [130, 206]. The central idea underlying the LIC is to assign each observation to the class that minimizes its local uncertainty or, equivalently, maximizes its local information content. This idea builds on the concept of local information content introduced in the context of non-parametric entropy estimators such as those of Kozachenko and Leonenko [121], Victor [239], and related formulations for local entropy computation [8, 122]. While these estimators have been used almost exclusively for computing global (ensemble-averaged) entropy, their pointwise formulation can intuitively be beneficial for developing Bayesian-type classifiers. Previous research has explored integrating probability density estimation techniques into Bayesian classifiers using non-parametric kernel density estimation [89, 109] or probability estimation trees [192], yet the explicit use of local information remains much less explored.

In contrast to traditional classifiers, which rely on global statistical properties such as mean, variance, or covariance structures, the LIC evaluates information quantities in the local neighborhood of each sample. This localized focus, together with the non-parametric implementation via k-nearest neighbors estimators, enables the model to adapt naturally to heterogeneous datasets and irregular class boundaries. Such characteristics make the LIC particularly suitable for biomedical applications, where data distributions are often non-Gaussian, multimodal, and not easily described by parametric models [66, 189].

6.2.1 The Local Information Classifier (LIC)

Let us consider a classification experiment where d features are used to assign the experiment outcomes to M classes. The experiment is modeled by a mixed random variable $Y = [\mathbf{X}, C]$ composed by a d-dimensional variable $\mathbf{X}$ taking values in the continuous domain $\mathcal{D}_X = \mathbb{R}^d$ and a scalar variable C taking values in the discrete alphabet $\mathcal{A}_C = c_1, ..., c_M$; the realizations of $\mathbf{X}$ and C contain respectively the values taken by the d features and by the class label for a given experiment run. To complete the notations, the probability distribution of the observed mixed variable is denoted as $p(y_i) = p(x, c_i) \equiv P[X = x, C = c_i]$, the probability of observing the i^{th} class label is denoted as $p(c_i) \equiv P[C = c_i]$, and the probability density inside the i^{th} class is denoted as $p_i(x) := p(x|c_i) \equiv P[X = x|C = c_i]$.

Here, the LIC propose to exploit the probability densities defined above to set out a classifier based on the concept of local information content, following the rationale that information is inversely related to probability. In statistical learning theory, a classifier is a function $f(\cdot)$ which is applied to the feature values $x \in \mathcal{D}_X$ observed in a new experiment run and returns the output $f(x) = c_i \in \mathcal{A}_C$. The classifier that we consider in this work selects a given class by minimizing the information that would be contained in a new run of the experiment if it was labeled with that class. Specifically, the classifier output,

written in shorthand notation as f_j, is obtained by defining the classification function as

$$\begin{aligned} f_j &\equiv f(x) = c_j, \\ j &= \underset{i=1,\ldots,M}{\operatorname{argmin}}\, h(x, c_i), \end{aligned} \tag{6.17}$$

where

$$h(x, c_i) = -\log p(x, c_i), \tag{6.18}$$

is the local joint entropy, or local information content associated with the joint probability $p(x, c_i)$, quantifying the pointwise information contained in the outcome $y_i = [x, c_i]$ of the observed variables. Note that, to favor computation of the local joint entropy (6.18), such entropy can be expanded as

$$h(x, c_i) = h(x|c_i) + h(c_i), \tag{6.19}$$

which evidences the local entropy of x given the class c_i, $h(x|c_i) = -\log p(x|c_i)$, and the local entropy of the class, $h(c_i) = -\log p(c_i)$.

The proposed classifier decides on the class associated with the minimum local information content. It is important to note that, since minimizing the local information content corresponds to maximizing the joint probability of the features and the tested class, this classifier implements an information-theoretic version of the Bayes rule [113]. Indeed, the Bayes decision rule selects the class c_j for which $p(c_j|x) = p(x, c_j)/p(x)$ is maximum, which corresponds to searching for the minimum local joint entropy $h(x, c_j)$.

The estimator adopted makes use of k-nearest neighbor metrics exploiting the intuitive notion that probability density is inversely related to the distance between data observations. This notion has been formalized in terms of information content in several works defining non-parametric entropy estimators [121, 122, 239]. Using our notation, the local entropy of a generic point x of the space spanned by the observations of the analyzed d-dimensional random variable X can be estimated, considering N realizations of X collected in the $N \times d$ data matrix $\boldsymbol{x} = [x_1 \cdots x_N]^{\mathrm{T}}$, as follows [122]:

$$\hat{h}(x) = \psi(N + 1) - \psi(k) + \log v_d + d \log \epsilon(x, k) \tag{6.20}$$

where $\epsilon(x, k)$ is twice the distance between the reference point x and its k^{th} neighbor inside $\boldsymbol{x}$, $\psi(x) = \frac{\log \Gamma(x)}{\mathrm{d}x}$ is the digamma function with $\Gamma(x)$ the gamma function, and v_d is the volume of the d-dimensional unit ball under a given norm ($v_d = 1$ for the maximum norm, and $v_d = \pi^{d/2}/\Gamma(1 + d/2)/2^d$ for the Euclidean norm).

Considering the supervised learning context of the classification experiment, the data in $\boldsymbol{x}$ are associated with an N-dimensional vector of class labels $\boldsymbol{c} = [c_1 \cdots c_N]^{\mathrm{T}}$ in a way such that N observations of the mixed variable Y describing the experiment are collected in the $(N \times (d+1))$ data matrix $\boldsymbol{y} = [\boldsymbol{x}, \boldsymbol{c}]$. Now, assuming that N_i observations are labelled as belonging to the class c_i $(i = 1, \ldots, M; \sum_{i=1}^{M} N_i = N)$, we consider the sub-matrix that contains the rows of $\boldsymbol{x}$ associated with the label c_i in $\boldsymbol{y}$. This $N_i \times d$ matrix, denoted as $\boldsymbol{x}_i = [x_{i,1} \cdots x_{i,N_i}]^{\mathrm{T}}$, contains the realizations of X given the i^{th} class, and is used considering (6.20) to estimate the local conditional entropy of the observation $x_{i,n}$ given the class c_i as

$$\hat{h}(x_{i,n}|c_i) = \psi(N_i) - \psi(k) + \log v_d + d \log \epsilon_i(x_{i,n}, k) \tag{6.21}$$

where $\epsilon_i(x_{i,n}, k)$ is twice the distance from $x_{i,n}$ to its k^{th} neighbor inside $\boldsymbol{x}_i$. Then, the entropy of X given c_i is estimated as the average information content over all observations in the class:

$$\hat{H}(X|c_i) = \frac{1}{N_i} \sum_{n=1}^{N_i} \hat{h}(x_{i,n}|c_i). \tag{6.22}$$

To estimate the local conditional entropy at the point $x_{i,n}$ given the class c_m, with $m \neq i$, one needs to consider the $N_m \times d$ data sub-matrix $\boldsymbol{x}_m = [x_{m,1} \cdots x_{m,N_m}]^{\mathrm{T}}$ containing the realizations of X associated with the m^{th} class. In this case the application of (6.20) results in the estimate

$$\hat{h}(x_{i,n}|c_m) = \psi(N_m + 1) - \psi(k) + \log v_d + d \log \epsilon_m(x_{i,n}, k), \tag{6.23}$$

where $\epsilon_m(x_{i,n}, k)$ is twice the distance from $x_{i,n}$ to its k^{th} neighbor inside $\boldsymbol{x}_m$. The use of (6.21) and (6.23), the latter iterated for all classes other than c_i, together with the estimation of the local entropy of the class, $\hat{h}_{c_i} = -\log(N_i/N)$, allows to compute (6.19) for each class, and then to find the minimum estimated local entropy needed in (6.17) to assign the class, i.e. $g_j \equiv g(x) = c_j$.

As a comparison to LIC, due to its similar nature, we use the Bayes classifier [202]. The classifier selects the class c_j which produces the maximum discriminant function $f_i(x, c_i)$ on a given feature input x, noted as

$$\begin{aligned} g_j &\equiv g(x) = c_j, \\ j &= \underset{i=1,\ldots,M}{\operatorname{argmax}}\ f(x, c_i), \end{aligned} \tag{6.24}$$

where the discriminant function is the class posterior probability, which can be written as

$$f(x, c_i) = p(c_i|x) = \frac{p(x|c_i)p(c_i)}{p(x)}, \tag{6.25}$$

after applying the Bayes rule [113]. For the goal of maximization, the $p(x)$ term can be ignored, giving the form of the maximum a posteriori probability (MAP) estimation

$$j = \underset{i=1,\ldots,M}{\operatorname{argmax}}\ p(x|c_i)p(c_i). \tag{6.26}$$

This general form of the classifier can be simplified by introducing certain assumptions, such as feature independence for a given class, leading to the well-known Naive Bayes. Additional assumptions, such as that the underlying distribution for each of the features is assumed Gaussian, lead to the version of the Gaussian Naive Bayes. In this study, however, since we are constructing models based on single feature inputs, we denote the used Naive Bayes classifier as the Gaussian Bayes classifier (GB). With regard to this, the distribution parameters for the unimodal Gaussian distributions are obtained by estimating the mean and variance of the input training data features for the given classes. The complete computational procedure of the Local Information Classifier is summarised in Algorithm 4.

Algorithm 4 Local Information Classifier (LIC)

Input: Training samples $\boldsymbol{x}$ with labels $\boldsymbol{c}$, number of classes M, number of neighbors k
Output: Predicted class label for a test sample x

Preprocessing:
for each class c_i, $i = 1, \ldots, M$ **do**
  Extract class-specific dataset $\boldsymbol{x}_i = \{x_{i,1}, \ldots, x_{i,N_i}\}$
  Estimate class prior: $\hat{p}(c_i) = N_i/N$

For a test sample x**:**
for each class c_i **do**
  Find k-nearest neighbors of x within $\boldsymbol{x}_i$
  Compute distance $\epsilon_i(x, k)$ to the k-th neighbor
  Estimate local conditional entropy: $\hat{h}(x \mid c_i) = \psi(N_i) - \psi(k) + \log v_d + d \log \epsilon_i(x, k)$
  Estimate class-information term: $\hat{h}(c_i) = -\log \hat{p}(c_i)$
  Compute local joint entropy: $\hat{h}(x, c_i) = \hat{h}(x \mid c_i) + \hat{h}(c_i)$
Decision rule: $g(x) = \arg\min_{i=1,\ldots,M} \hat{h}(x, c_i)$
return $g(x)$

Because the LIC relies exclusively on local information quantities, it is fully non-parametric and model-independent. This characteristic allows the classifier to adapt naturally to heterogeneous datasets and irregular class boundaries, handling non-stationary and non-Gaussian feature distributions that are typical of physiological data [66, 97, 189].

Moreover, since LIC's decision mechanism is based on interpretable quantities, local entropy and mutual information, it enables direct inspection of which regions in the feature space carry the highest discriminative information. This interpretability is particularly valuable in biomedical contexts, where understanding the physiological meaning of model decisions is as important as their accuracy [14, 230].

The advantages of LIC have been demonstrated both in simulation studies and in real-world applications to cardiovascular variability data. When tested on real physiological recordings involving rest, orthostatic, and mental stress conditions, LIC achieved stable and interpretable classification outcomes based on single and multiple feature configurations [39, 93]. These results confirmed that the model's ability to exploit localized information structures makes it well-suited for distinguishing between physiological states characterized by subtle, nonlinear differences in autonomic regulation.

In summary, the Local Information Classifier integrates the theoretical rigor of information theory with the adaptability of non-parametric estimation. By minimizing local entropy and maximizing information content, it provides a principled and interpretable approach to classification that is especially suited to the complexity and variability of biomedical datasets [66, 189, 236]. Its formulation bridges probabilistic inference and information theory, yielding a flexible framework capable of describing the underlying informational architecture of physiological dynamics with both conceptual clarity and practical utility. Its performance across heterogeneous experimental contexts, including stress classification, further demonstrates the versatility of this information-based approach [130, 206].

Validation on synthetic dataset. The evaluation of the proposed LIC method was conducted through two controlled numerical simulations. In particular, two synthetic datasets were generated, each consisting of two classes with known probability distributions. Simulation 1 used two-dimensional Gaussian distributions with distinct means and covariances for the two classes, representing a scenario with moderate class overlap. Simulation 2 employed two-dimensional Weibull distributions to create a more challenging scenario with non-Gaussian, asymmetric distributions. In both simulations, datasets of equal class size were generated, and multiple realisations were used to assess statistical robustness.

For each simulation dataset, the LIC method classified samples based on the local mutual information estimate computed between the feature vector and the class label, using a k-nearest-neighbour entropy estimator. Classification performance was evaluated by assigning each sample to the class providing the largest local information value. The results obtained with LIC were compared against those of a Gradient Boosting classifier

trained on the same datasets.

Across both simulations, classification performance was assessed in terms of accuracy, sensitivity, and specificity, averaged across multiple repetitions to ensure statistical reliability. All analyses followed the methodological definitions and preprocessing steps detailed in the methodological chapters of this thesis.

To investigate the predictive capabilities of the proposed LIC model, we utilized $k = 10$ nearest neighbors along with the maximum norm distance metric for the LIC, and employed a form of cross-validation. The performances of both the LIC and the GB classification models were assessed in terms of sensitivity (SE) and specificity (SP) for each class, along with computation of the overall accuracy.

The first simulation reproduces a classification experiment with a scalar feature distributed within three equally probable classes according to a one-dimensional Gaussian distribution. For classes 1 and 3, the mean and variance remained constant, whereas for class 2, the mean was fixed while the variance σ_2 was varied, taking the values as specified in Fig. 6.1 a). For all combinations of the parameters, the theoretical values of the accuracy, SE, and SP were computed considering the prior knowledge of the distribution of the feature within each class. Moreover, the performance parameters were computed implementing both the LIC and the GB classifiers over 100 realizations of the simulation, each producing $N_i = 150$ samples for each class ($i = 1, 2, 3$). The lower sample number serves as a real-case scenario for classification problems with fewer instances. Given this, we employed a leave-one-out cross-validation approach for each realization to assess the classifiers.

The performance metrics of both LIC and GB models exhibit a decline with increasing the variance of the distribution within the second class, as shown in Fig.6.1 b). Nevertheless, the estimated values of the metrics remain in close alignment with the theoretical values, demonstrating the ability of both LIC and GB to achieve the expected performance. Marginal enhancements in accuracy are detected when the variance of the second class is altered, suggesting that both estimators have managed to capture some new information, possibly pertaining to sampling noise.

In the second simulation, we adopted the same three-class scheme but employed a one-dimensional Weibull distribution for each class in place of the Gaussian. Each class is characterized by specific scale and shape parameters, which remained constant except for the shape parameter k_2 of the second class (see Fig. 6.2 a). As k_2 increases, narrowing the shape of the probability of the feature in the second class, as expected both the accuracy and the SE of the first and second classes increase, less affecting their SP metrics and the third class (Fig. 6.2 b). Importantly, when compared to theoretical values, the performance

metrics of the LIC model outperform those of the GB model in effectively capturing the underlying distribution (Fig. 6.2 b).

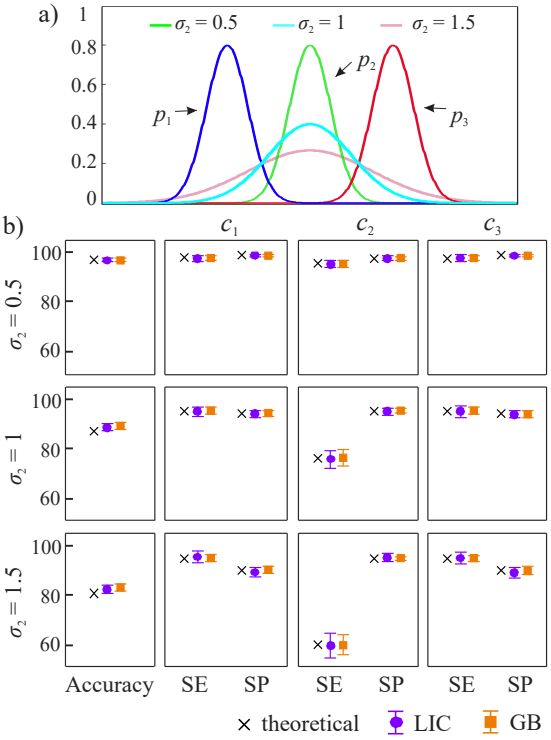


Figure 6.1: Simulation 1 generates realizations of a random variable containing $N = 150$ values of a single feature drawn from a Gaussian distribution with mean μ and variance σ^2. a) Theoretical profiles of the simulated within-class probability densities (p_1: $\mu_1 = -2$, $\sigma_1 = 0.5$; p_2: $\mu_2 = 0$, $\sigma_2 \in \{0.5, 1, 1.5\}$; p_3: $\mu_3 = 2$, $\sigma_3 = 0.5$). b) Overall accuracy, sensitivity (SE), and specificity (SP) computed for the LIC and GB classifiers through leave-one-out cross-validation, reported as mean ± standard deviation over 100 realizations. This figure has been modified from [130].

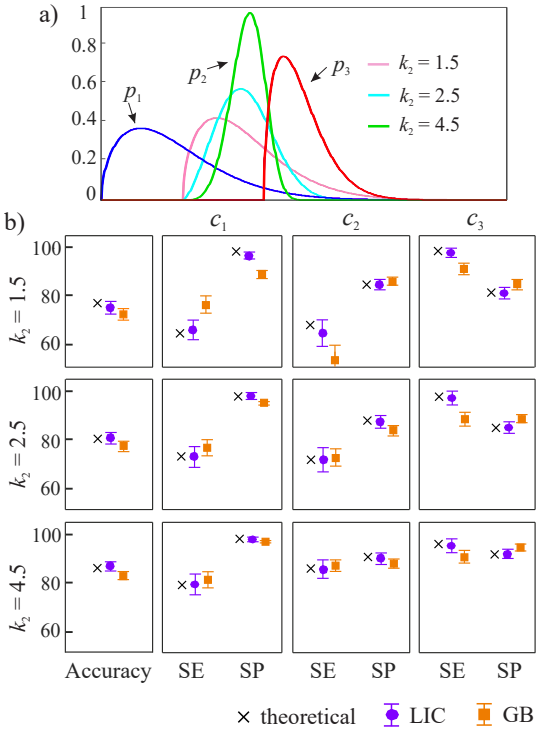


Figure 6.2: Simulation 2 generates realizations of a random variable containing $N = 150$ values of a single feature drawn from a Weibull distribution with scale parameter λ and shape parameter k. a) Theoretical profiles of the simulated within-class probability densities (p_1: $\lambda_1 = 2, k_1 = 1.5$; p_2: $\lambda_2 = 1.75$, $k_2 \in \{1.5, 2.5, 4.5\}$; p_3: $\lambda_3 = 1, k_3 = 1.5$). b) Performance metrics computed as in Simulation 1. This figure has been modified from [130].

Simulation experiments involving classes characterized by Gaussian distributions yielded the expected results, demonstrating good performances of the LIC alongside the Gaussian Bayes classifier with a priori knowledge of the underlying class distributions. On the other hand, as expected, the LIC outperforms the Gaussian classifier in scenarios featuring non-Gaussian distributions.

6.3 Training, Validation, and Performance Evaluation

The evaluation of classification models represents a crucial stage in the analytical workflow, ensuring that the results obtained reflect generalizable and physiologically meaningful patterns rather than artifacts of model fitting [2, 2, 3, 3, 31, 214, 214]. In biomedical and physiological research, where datasets are often limited in size and characterized by high inter-subject variability, rigorous training, validation, and testing procedures are essential to achieve robust conclusions and to prevent overfitting [2, 3, 114, 120, 127, 214]. This section outlines the methodological principles adopted in this thesis to train and evaluate the classifiers described previously, emphasizing the rationale for model selection, cross-validation strategies, and performance assessment metrics.

Training and Model Optimization. During the training phase, the classifier learns the functional mapping between the input features $\mathbf{X} = [X_1, X_2, \ldots, X_d]$ and the class labels $Y \in \{1, 2, \ldots, C\}$ by minimizing a loss function that quantifies the discrepancy between predicted and actual labels. In general form, the objective function is expressed as

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}(f(\mathbf{x}_i), y_i), \tag{6.27}$$

where $f(\mathbf{x}_i)$ denotes the model prediction for the i-th sample, and $\mathcal{L}(\cdot)$ represents the loss function appropriate for the task. For classification problems, the cross-entropy loss is commonly adopted [31, 114, 127]:

$$\mathcal{L}_{CE} = -\sum_{i=1}^{N} \sum_{c=1}^{C} \mathbb{I}(y_i = c) \log \hat{P}(y_i = c|\mathbf{x}_i), \tag{6.28}$$

where $\hat{P}(y_i = c|\mathbf{x}_i)$ is the predicted probability that sample $\mathbf{x}_i$ belongs to class c, and $\mathbb{I}(\cdot)$ is the indicator function. This formulation penalizes misclassifications logarithmically, ensuring that the model not only assigns the correct label, but also provides confident predictions.

Hyperparameter optimization is an integral component of model training, as parameters such as the number of neighbors in kNN, the kernel function and regularization constant in SVM, or the number of trees in Random Forests significantly influence model performance. Hyperparameter tuning is usually carried out through grid search within nested cross-validation schemes to ensure unbiased estimation of performance. The outer loop estimates model generalization, while the inner loop identifies the optimal hyperparameter configuration based on validation accuracy. This hierarchical design prevents the information leakage that can occur when tuning and testing are performed on the same

data, which is especially critical when working with small biomedical datasets [135].

Validation and Testing Protocols. To obtain unbiased estimates of classifier performance and generalization ability, several validation schemes were adopted depending on the characteristics of each dataset and classification task. In the simplest case, the dataset is randomly divided into separate training and test subsets, typically following an 80/20 or 70/30 split. The model is trained on the training set and evaluated on the independent test set, ensuring that the test data remain completely unseen during parameter estimation and model selection [120].

However, because of the limited size and heterogeneity typical of physiological datasets, simple train–test splits may not provide stable or representative performance estimates. To address this issue, k-fold cross-validation (CV) is employed as a more robust validation strategy. The dataset is partitioned into k approximately equal subsets (folds), and the model is iteratively trained on $k-1$ folds and tested on the remaining one. The process is repeated k times so that each fold serves once as a test set, and the overall performance is computed as the average of the fold-wise metrics:

$$\bar{M} = \frac{1}{k} \sum_{j=1}^{k} M_j, \tag{6.29}$$

where M_j denotes the chosen performance metric (e.g., accuracy, F1-score, or AUC) for the j-th fold.

For inter-subject classification tasks, such as those involving HRV or PRV features across multiple participants, the *leave-one-subject-out* (LOSO) cross-validation is also employed [39, 40]. In LOSO-CV, all samples from one subject are excluded during training and used exclusively for testing, and the procedure is repeated until each subject has served as a test case once. This approach provides a stringent and physiologically meaningful measure of generalization, as it assesses the model's ability to classify data from entirely unseen individuals, a critical requirement for applications in personalized or clinical monitoring [93].

Regardless of the chosen scheme, strict data separation between training, validation, and test phases was maintained throughout all analyses. Preprocessing operations such as normalization, feature selection, and dimensionality reduction were performed exclusively within each training fold to prevent information leakage from the test data, ensuring a methodologically sound evaluation.

Performance Metrics. Classifier performance was assessed using a set of complementary quantitative metrics that capture different aspects of prediction quality [2, 3, 31,

114, 127, 214]. The most straightforward measure is the overall accuracy, defined as

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{6.30}$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively. While accuracy provides a general measure of correctness, it can be misleading when class distributions are unbalanced, as often occurs in biomedical applications.

To account for this, precision, recall, and F1-score were also computed. Precision (positive predictive value) measures the proportion of correctly classified positive instances among all samples predicted as positive:

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{6.31}$$

whereas recall (sensitivity) quantifies the proportion of true positives correctly identified among all actual positives:

$$\text{Recall} = \frac{TP}{TP + FN}. \tag{6.32}$$

Their harmonic mean, the F1-score, provides a balanced summary between these two quantities:

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{6.33}$$

In binary problems, these metrics were complemented by the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate (TPR) against the false positive rate (FPR) across all possible classification thresholds. The area under this curve (AUC) was used as a scalar indicator of discriminative power [31, 114, 127]:

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR})\, d(\text{FPR}). \tag{6.34}$$

For multiclass problems, metrics were computed using the one-vs-all strategy and averaged according to macro- or weighted schemes depending on class balance [31, 114, 127]. In addition, confusion matrices were analysed to identify systematic misclassifications and to evaluate how specific physiological states or subject groups were differentiated [39, 40, 135]. Whenever possible, statistical significance between classifiers was tested using non-parametric paired tests, such as the Wilcoxon signed-rank test, applied to fold-wise accuracy distributions.

In addition to global metrics, the analysis of the confusion matrix provides valuable insight into the detailed behaviour of each classifier. The confusion matrix summarizes

the correspondence between predicted and actual class labels, displaying true and false classifications for each category. Its diagonal elements represent correctly classified instances, whereas off-diagonal terms highlight systematic misclassifications and inter-class overlaps. In biomedical applications, such as cardiovascular state recognition, these patterns can reveal physiologically meaningful relationships, for instance, partial confusion between rest and mild stress conditions may reflect genuine transitional autonomic states rather than model errors [39, 40]. Quantitative measures derived from the confusion matrix, including per-class precision, recall, and specificity, were examined to identify which physiological conditions were most or least distinguishable. This interpretative analysis complements scalar metrics like accuracy and F1-score, allowing a physiologically informed assessment of classifier behaviour and reliability.

Robustness and Reliability Considerations. Beyond conventional performance metrics, special attention was devoted to the robustness and reproducibility of the models [2, 3, 114, 127, 214]. Robustness refers to the ability of the classifier to maintain consistent predictions in the presence of data perturbations, such as measurement noise, missing samples, or small variations in feature scaling. This property was examined through controlled resampling and bootstrapping analyses, allowing the estimation of confidence intervals for each performance metric. In parallel, reproducibility was ensured by fixing random seeds, documenting hyperparameter configurations, and performing repeated cross-validation runs with stratified folds to mitigate the effects of random sampling. The combination of these procedures ensures that the reported classification results reflect genuine discriminative structure in the data rather than chance-level fluctuations or dataset-specific biases. Through rigorous validation and comprehensive performance assessment, the proposed framework guarantees both statistical reliability and physiological interpretability.

Table 6.1 provides a comparative summary of the classification algorithms implemented in this work, highlighting their theoretical assumptions, computational demands, and interpretability. Each model embodies a distinct inductive bias and operational principle, offering specific trade-offs between flexibility, robustness, and transparency. Linear methods such as LDA provide clear probabilistic foundations and interpretable decision boundaries, well-suited for small and structured datasets. Nonlinear algorithms like SVMs and Neural Networks extend modelling capability to complex and high-dimensional data, albeit with reduced interpretability and higher computational cost. Instance-based and ensemble methods such as kNN and Random Forests balance simplicity and adaptability, while the information-theoretic LIC introduces a novel, model-free perspective grounded in local entropy and mutual information estimation.

Overall, the examined classifiers offer complementary analytical perspectives rather than competing alternatives. Linear models provide interpretability and analytical clarity, nonlinear and ensemble approaches capture intricate physiological interactions, and the information-theoretic formulation offers a principled, data-driven view of variability. Together, they outline a continuum of modelling paradigms through which cardiovascular complexity can be quantitatively assessed and physiologically interpreted.

From a broader standpoint, classification represents the culmination of the methodological pipeline developed in this thesis. It integrates the preceding analytical stages—feature extraction, feature selection, and feature importance—into predictive frameworks capable of distinguishing among physiological or physiopathological states. Feature extraction transforms raw cardiovascular signals into structured descriptors of variability and complexity; feature selection identifies the most informative and non-redundant indices; feature importance quantifies their unique and cooperative contributions. Classification consolidates these outcomes, translating the multivariate structure of the feature space into discriminative decision boundaries that reflect the underlying physiological organisation.

The comparative evaluation of classifiers demonstrates that predictive accuracy, interpretability, and physiological meaning can coexist within an information-based framework that regards classification as a process of uncertainty reduction and knowledge extraction. In this context, the LIC represents a synthesis of traditional and modern paradigms, combining the local adaptability of non-parametric estimators with the explanatory power of information-theoretic measures. It achieves competitive performance while preserving conceptual transparency, bridging the gap between data-driven prediction and physiological understanding.

Ultimately, the analytical framework established in this thesis, from feature computation to model validation, illustrates how the flow of information can be systematically traced from physiological signals to predictive models. The transition from descriptive analysis to classification mirrors the informational dynamics of the cardiovascular system itself, moving from variability and interaction toward coherent functional states. This conceptual continuity underscores the central contribution of the present work: the establishment of an integrated, information-driven methodology that unifies quantitative modelling with physiological interpretation, laying the groundwork for future advances in computational physiology and clinical decision support.

Table 6.1: Comparison and summary of the main characteristics of the implemented classifiers.

Classifier	Model Type / Principle	Strengths	Limitations	Interpretability
LDA	Linear parametric model assuming Gaussian class distributions with shared covariance.	Fast, analytically interpretable; performs well on linearly separable and well-balanced data.	Requires normality and homoscedasticity; limited for nonlinear patterns.	High, coefficients directly show feature contribution.
SVM	Kernel-based discriminant model maximizing class margin in feature space.	Handles nonlinear and high-dimensional data; strong generalization even with small samples.	Requires kernel and hyperparameter tuning; less transparent decision process.	Medium, depends on kernel and support vectors.
kNN	Instance-based, non-parametric; assigns class by majority vote among k nearest neighbors.	Simple, model-free, adaptive to local data structure.	Sensitive to scaling, noise, and irrelevant features; computationally heavy at prediction.	Low, no explicit model representation.
NB	Probabilistic model based on Bayes' theorem with feature independence assumption.	Very fast; effective for small datasets and high-dimensional features.	Independence assumption often unrealistic in physiological data.	High, probabilities directly interpretable.
NN	Nonlinear parametric model of layered artificial neurons optimized via gradient descent.	Captures complex nonlinear relationships and hierarchical representations.	Requires large datasets and careful regularization; prone to overfitting.	Low–Medium, hidden representations difficult to interpret.
RF	Ensemble of decision trees trained with bagging and random feature subsets.	Robust to noise and feature correlations; provides feature importance; low variance.	Reduced interpretability due to ensemble structure; may require many trees.	Medium, offers feature importance but global structure opaque.
LIC	Non-parametric, information-theoretic classifier minimizing local entropy and maximizing mutual information.	Model-free, adaptive to non-Gaussian and multimodal data; interpretable through information measures.	Computationally intensive for large datasets; depends on neighborhood parameter k.	High, directly quantifies local information and uncertainty.

Part III

Implementation of Physio-Pathological State Classification from Cardiovascular Signals

7

Experimental Applications and Case Studies

This chapter presents a comprehensive overview of the experimental applications developed throughout this doctoral research, illustrating the practical implementation of the methodological framework introduced in the previous chapters (3-6). The main objective is to demonstrate how advanced techniques for feature extraction, feature selection, feature importance assessment, and classification can be effectively applied to physiological data for the quantitative characterisation of cardiovascular dynamics under different experimental conditions.

Although the individual studies address different physiological questions and experimental datasets, they are all grounded in a common methodological perspective based on the analysis of physiological time series through signal processing, feature extraction, and information-driven modelling.

Depending on the specific objectives of each study, different components of the analytical workflow are employed. These may include signal preprocessing and beat-to-beat time series construction, extraction of time-, frequency-, or information-domain features, information-theoretic feature selection or feature importance analysis, and classification models for discriminating physiological states. Not all stages are required in every application; rather, each case study emphasises the components that are most relevant to the specific physiological problem under investigation.

Within this perspective, the different applications presented in this chapter should be interpreted as complementary case studies that explore and validate individual components of the proposed analytical framework under different experimental conditions and data

modalities.

The chapter is organised into four main sections, each addressing a specific aspect within the context of applied studies. Specifically, it is structured according to the main *physiological applications* addressed in this work, each representing a distinct experimental scenario in which the developed algorithms and analytical techniques were tested and validated. First, the datasets and experimental protocols are described in Sec. 7.1, detailing the acquisition of cardiovascular signals and the preprocessing steps applied to obtain standardised time series. The following sections focus on *Stress characterisation and classification* (7.2), presenting the application of signal processing, feature extraction, and machine learning methods for the discrimination of physiological stress induced by postural and cognitive challenges. Then, Sec.7.3 investigates the topic of gender differences in cardiovascular regulation, focusing on the sex-dependent patterns in the dynamical complexity and information content of heart rate and blood pressure variability using information-theoretic and classification approaches. Finally, additional applications are presented, including experimental validation studies and further applications of the algorithms introduced in the previous chapter that are not directly related to the two aforementioned topics (Sec. 7.4).

Overall, this chapter serves to bridge the methodological advances introduced earlier with their experimental validation, illustrating how the proposed analytical workflow can be applied and adapted across different physiological scenarios while preserving a common methodological structure.

7.1 Datasets

- **Dataset 1 (D1):** The first dataset includes cardiovascular recordings collected at the Department of Physiology, Jessenius Faculty of Medicine, Comenius University, Martin, Slovakia. The study was conducted to investigate cardiovascular control at rest and under different stress types, and comprises recordings from 127 healthy young individuals (75 females; mean age: 18.6 ± 3.3 years; mean body mass index: 21.4 ± 2.2 kg/m^2) [97, 232]. All participants were non-smokers, normotensive, and free from cardiovascular, neurological, or metabolic disorders, and none were under chronic pharmacological treatment. The dataset was not acquired within the present thesis work but represents a historical dataset made available through established research collaborations.

  The study protocol was approved by the Ethics Committee of the Jessenius Faculty of Medicine, Comenius University, Martin, Slovakia, and written informed consent

was obtained from all participants (and from legal guardians for minors). The experiments adhered to the Declaration of Helsinki and followed standard ethical guidelines for research in human physiology.

Electrocardiogram signals were recorded using a horizontal bipolar thoracic lead (Cardiofax ECG-9620, Nihon Kohden, Japan), while continuous beat-to-beat arterial blood pressure (AP) was measured using a volume-clamp photoplethysmographic device (Finometer Pro, FMS, Netherlands). Moreover, respiratory volume through respiratory inductive plethysmography (RespiTrace 200; NIMS) was simultaneously and non-invasively recorded. All signals were synchronously acquired at a sampling frequency of 1 kHz to ensure precise temporal alignment between cardiac and vascular signals. More details on the protocol and experimental acquisition are reported in [106].

The experimental protocol comprised five consecutive phases alternating resting and stress conditions:

- **REST:** 15 minutes of baseline rest in the supine position on a motorized bed.
- **HUT (Head-Up Tilt):** 8 minutes of orthostatic stress induced by tilting the motorised bed to a 45° upright position.
- **REST-2:** 10 minutes of supine rest for recovery.
- **MA (Mental Arithmetic):** 6 minutes of cognitive stress induced by a mental arithmetic task involving rapid summation and parity decision of three-digit numbers projected on a screen (WIN 5 PMT software, PsychoSoft, Brno, Czech Republic).
- **REST-3:** 10 minutes of final recovery in the supine position.

Throughout the recording, subjects were instructed to avoid movement and speech and to maintain spontaneous breathing to minimise artefacts. All acquisitions were performed in a quiet, temperature-stable environment, with the subjects comfortably lying on the motorised table.

For each subject and experimental condition, time series of 300 heartbeats were extracted from the acquired ECG and arterial pressure signals following the standard short-term analysis protocol. The n-th RR interval (RR_n) was defined as the time interval between the n-th and $(n + 1)$-th QRS complexes detected in the ECG signal. Similarly, the n-th pulse-to-pulse interval (PP_n) was defined as the time interval between two consecutive systolic arterial pressure peaks. This procedure ensured that the RR, PP, SBP and DBP time series had the same length for each subject and physiological condition. The n-th sample of the systolic blood pressure

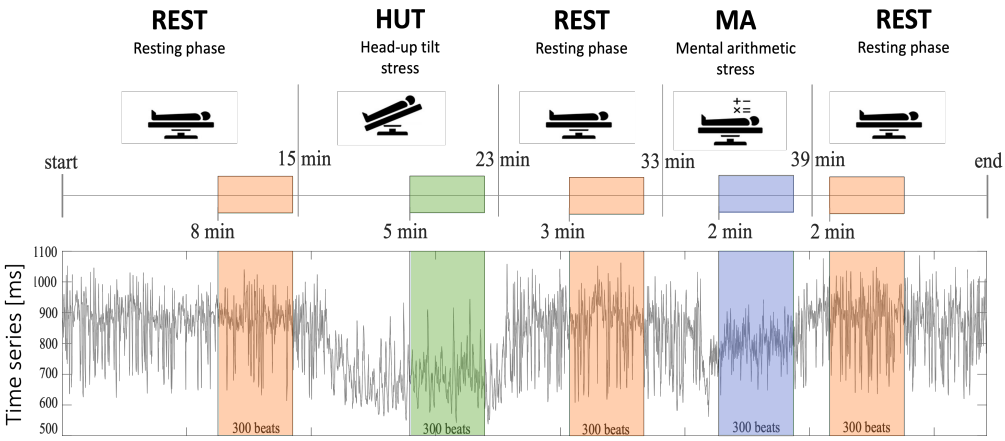


Figure 7.1: Schematic illustration of the experimental protocol, including baseline resting (REST), orthostatic stress (HUT), second resting (REST-2), mental arithmetic (MA) and the final third recovery phase (REST-3). Representative RRI and PPI time series, extracted respectively from ECG and blood pressure recordings. Colour boxes indicate the windows taken into account for short-term (300 points) analysis. This figure has been modified from [97].

time series (SBP_n) was defined as the maximum arterial pressure value occurring within the cardiac cycle identified by consecutive R peaks, while the diastolic blood pressure sample (DBP_n) corresponded to the minimum pressure value within the same interval. The respiratory signal was synchronised with the cardiac rhythm by sampling the respiratory volume signal at the onset of each detected heartbeat, thus obtaining the $RESP_n$ time series. The 300-beat windows were extracted from the recordings after allowing sufficient time for physiological stabilisation. Specifically, segments were selected approximately 8 minutes after the beginning of the REST phase, 3 minutes after the start of the HUT phase, 3 minutes after the beginning of the REST-2 phase, 2 minutes after the start of the MA phase, and 2 minutes after the beginning of the REST-3 phase. This strategy was adopted to avoid transient responses occurring immediately after posture or task transitions and to ensure that the analysed segments reflected stable physiological conditions. All analysed segments were verified to be free of artefacts, including those associated with the calibration procedure of the Finometer device. The Finometer calibration, which temporarily interrupts the continuous measurement of arterial pressure, occurred only during the final minute of the REST and REST-2 phases and was therefore excluded from the analysed windows. Each time series was subsequently corrected for outliers exceeding three standard deviations from the mean. Isolated outliers were replaced using linear interpolation, whereas sequences of consecutive outliers were corrected using cubic spline interpolation. After interpolation, the corrected

time series were visually inspected to ensure signal consistency. Overall, only 0.6% of the samples in each time series required correction, ensuring a reliable basis for the subsequent analyses.

- **Dataset 2 (D2):** The second dataset, denoted as *HYPOL* (Healthy Young POLes), publicly accessible at `https://hypol.ump.edu.pl`, consists of recordings from 278 healthy Polish adults (149 females and 129 males; mean age: 23.8 ± 2.6 years; mean body mass index: 21.98 ± 2.66 kg/m^2) [86]. Participants were recruited from seven independent research projects conducted at the Department of Cardiology–Intensive Therapy and Internal Medicine, Poznan University of Medical Sciences, Poland. Each participant underwent a comprehensive medical screening, including physical examination, medical history evaluation, and resting ECG to exclude cardiovascular or systemic diseases, metabolic disorders, and the use of medication affecting autonomic function.

  All procedures complied with the Declaration of Helsinki and were approved by the local Bioethics Committee of the Poznan University of Medical Sciences Written informed consent was obtained from all participants prior to inclusion in the study.

  Recordings were acquired in the supine position using a bipolar chest-lead ECG and a continuous finger arterial pressure signal obtained with Portapres 2 or Finapres Nova devices (FMS, Netherlands). ECGs were digitised using Porti 5 or Porti 17 (TMSI, Netherlands) at sampling rates of 1600 Hz or 2048 Hz, while the blood pressure waveforms were sampled at 100–200 Hz depending on the instrumentation. Each acquisition lasted up to 45 minutes, and approximately 30 minutes of high-quality, artefact-free data were retained for each subject.

  The dataset includes synchronised beat-to-beat cardiovascular time series: RR intervals (from ECG), interbeat intervals (from the pressure waveform), systolic (SBP), diastolic (DBP), and mean blood pressure (MBP) values. The data represent spontaneous cardiovascular dynamics at rest in a homogeneous population of healthy young adults, making the HYPOL database an ideal reference for physiological and methodological studies on heart rate and blood pressure variability.

  The *HYPOL* database is widely employed for the study of heart rate variability, blood pressure variability, baroreflex function, and gender-related differences in cardiovascular regulation [86, 96].

  It should be noted that the preprocessing of the physiological signals and the extraction of the beat-to-beat time series were already performed by the authors [86] of the HYPOL database and provided as part of the publicly available dataset. In the present thesis, these precomputed RR, SBP, and DBP time series were used directly,

and segments of 300 consecutive beats were extracted after discarding the initial 60 beats in order to obtain standardised short-term (~5 min) recordings suitable for the subsequent analyses.

- **Dataset 3 (D3):** The third dataset was collected within two complementary studies conducted at the A.R.N.A.S. Ospedali Civico Di Cristina Benfratelli Hospital in Palermo, Italy, and coordinated by the Department of Engineering, University of Palermo. The dataset was designed to investigate physiological stress and autonomic regulation in healthcare workers exposed to high occupational load through cardiovascular variability analyses. Within this study, the candidate actively contributed to the design of the experimental protocol and participated in the acquisition of the physiological recordings. The candidate was also responsible for the preprocessing, signal processing, and analysis of the collected data within the analytical framework developed in this thesis.

  A total of 60 shift-working nurses (42 females; mean age: 50.0 ± 3.0 years) participated in the study. All subjects were actively employed in hospital wards with rotating day–night shifts and were free from acute or chronic medical conditions. Inclusion criteria comprised active employment under a shift-work schedule for at least five years, absence of pharmacological treatments affecting autonomic or metabolic regulation, and no history of cardiovascular or endocrine disease. The average body-mass index (BMI) at baseline was 25.6 ± 4.6 kg/m^2, consistent with an overweight but otherwise healthy population.

  Participants were divided into two groups of equal size ($n = 30$ each):

  - **Intervention group (INT):** participated in a structured wellness program comprising educational seminars, nutritional counselling, and guided physical activity sessions aimed at improving lifestyle, stress management, and metabolic health.
  - **Control group (CONT):** received general informational material about healthy habits without participating in any supervised intervention.

  All participants underwent two experimental sessions: a *baseline* (PRE) assessment and a *follow-up* (POST) after approximately six months, allowing longitudinal evaluation of both physiological and compositional adaptations.

  The study protocol was approved by the Ethics Committee of the A.R.N.A.S. Civico Hospital (protocol no. 59/2023) and by the Bioethics Committee of the University of Palermo (ref. no. 129/2023) and conducted in accordance with the Declaration

of Helsinki. Written informed consent was obtained from all participants prior to enrolment.

Participants completed two validated questionnaires to assess perceived health and sleep quality at baseline: the SF-12 Health Survey and the Pittsburgh Sleep Quality Index (PSQI). The SF-12, a short version of the SF-36 Health Survey [244], provides two summary indices, the Physical Component Summary (PCS12) and the Mental Component Summary (MCS12), calculated according to Ottoboni et al. [166] and interpreted using normative data from Andersen et al. [7]. The PSQI [194] evaluates sleep quality through seven components (COM 1–7), whose sum yields a global score from 0 to 21, with higher scores indicating poorer sleep quality.

Descriptive characteristics and results of questionnaires (i.e., PCS12, MCS12 and PSQI scores) for both groups at PRE are reported in Table 7.7. With regard to PCS12 and MCS12, an independent samples t-test was performed to assess baseline differences between CONT and INT groups, after verifying normality of data distribution using the Shapiro–Wilk test. The analysis did not reveal any statistically significant differences between groups at the start of the intervention. Both groups exhibited a physical health comparable to the general population according to [7], but a mental health score lower than the general population, compatible with a condition of work-related stress.

The experimental sessions involved non-invasive acquisition of cardiovascular recordings using the *Multi Parametric Sensor Acquisition System (MPSAS)*, a portable multimodal device previously developed in [233] and available at the University of Palermo, enabling synchronous high-resolution recording of electrocardiographic (ECG), photoplethysmographic (PPG), and respiratory (RESP) signals. The ECG was acquired through standard three-lead chest electrodes (RA, LA, LL), while the PPG was measured in transmission mode on the index or middle finger using a 940 nm infrared LED and OPT101 photodiode.

Each acquisition session comprised two short-term (5-minute) recordings per subject under different respiratory conditions:

- **Spontaneous breathing (REST):** participants remained seated in a quiet room, breathing spontaneously and instructed to avoid movement or speech.
- **Controlled breathing (CB):** a paced breathing exercise with 5 s inspiration, 2 s pause, and 5 s expiration (5 cycles/min), designed to elicit parasympathetic activation and relaxation.

The dataset thus integrates multimodal physiological data acquired from a homoge-

	CONT	INT
Age (years)	51.13 ± 5.96	54.42 ± 5.98
Weight (Kg)	68.58 ± 17.73	68.92 ± 11.32
Height (cm)	164 ± 0.09	165 ± 0.06
BMI (Kg/m^2)	25.18 ± 5.04	25.33 ± 3.84
PCS12	49.39 ± 7.77	51.64 ± 8.21
MCS12	42.74 ± 10.17	43.52 ± 10.16
Global PSQI	10.21 ± 3.51	8.73 ± 1.73
COM 1	1.16 ± 0.76	1.00 ± 0.89
COM 2	1.12 ± 1.11	1.03 ± 0.95
COM 3	1.37 ± 0.82	1.23 ± 0.81
COM 4	3.00 ± 0.00	3.00 ± 0.00
COM 5	1.54 ± 0.77	1.23 ± 0.51
COM 6	1.29 ± 0.69	0.84 ± 0.46
COM 7	0.70 ± 0.90	0.38 ± 0.63

Table 7.1: Descriptive characteristics of groups. Data are presented as mean ± standard deviation; CONT: Control group; INT: Intervention group; BMI: body mass index; PCS12: Physical Component Summary; MCS12: Mental Component Summery; PSQI: Pittsburgh sleep quality index questionnaire; COM 1: Subjective sleep quality; COM 2: Sleep latency; COM 3: Sleep duration; COM 4: Sleep efficiency; COM 5: Sleep disturbance; COM 6: Use of sleep medication; COM 7: Daytime dysfunction.

neous population of healthcare workers before and after a lifestyle intervention. It includes:

- ECG, PPG, and RESP signals sampled at 1 kHz in REST and CB conditions;
- Derived beat-to-beat R–R and pulse–pulse interval time series for HRV and PRV analyses

ECG and PPG signals were synchronously recorded using the MPSAS system at a sampling rate of 1 kHz. The ECG signal was filtered using a zero-phase 4^{th}-order Butterworth bandpass filter with cut-off frequencies of 0.1 Hz and 20 Hz, allowing for the removal of baseline drift and high-frequency noise while preserving QRS morphology. R peaks were detected using a simplified version of the Pan-Tompkins algorithm [195], and R-R intervals (RRI) were calculated as the temporal difference between successive R peaks, for both ECG leads.

The PPG signal was filtered using a zero-phase 4^{th}-order passband Butterworth filter with cut-off frequencies of 0.1 Hz and 15 Hz, to suppress motion artefacts and enhance the pulsatile component. Pulse-to-pulse intervals (PPI) were extracted by detecting local maxima and minima of the PPG waveform. Since the distributions of R-R interval time series extracted from the two ECG leads and between the P-P

interval time series computed from minima and maxima of PPG signal did not show any statistically significant differences according to t-test, only the ECG lead I and the maxima PPI time series were retained for the subsequent analyses. For each physiological condition and subject (i.e. REST, CB), 300-point-long time series (~ 5 min of the original signal) were extracted.

The dataset, therefore, represents an original experimental component of this doctoral research and was used to support the development and validation of the signal processing and feature extraction procedures employed throughout the thesis. This dataset provides a unique framework for studying stress-related autonomic adaptations, the impact of controlled breathing on vagal modulation, and the interplay between metabolic composition and cardiovascular regulation in a real-world occupational context.

Table 7.2: Comparative summary of the three datasets (D1–D3) used in this thesis.

Category	Dataset 1 (D1)	Dataset 2 (D2)	Dataset 3 (D3)
Population	127 healthy young subjects (75 females, 52 males).	278 healthy young adults (149 females, 129 males).	60 shift-working hospital nurses (42 females, 18 males).
Age (years)	18.6 ± 3.3.	23.8 ± 2.6.	50.0 ± 3.0.
Body mass index	21.4 ± 2.2 kg/m^2.	21.98 ± 2.66 kg/m^2.	25.6 ± 4.6 kg/m^2.
Health status	Healthy, non-smokers, no medication.	Healthy; screened for cardiovascular, respiratory, and systemic disorders; no chronic medication (except contraception).	Healthy; no chronic disease; minimum of 5 years of shift work, thus potentially affected by work-related stress.
Experimental protocol and recording	Five-phase protocol (REST, HUT, REST-2, MA, REST-3); total duration ≈ 49 min.	Single recording in supine rest (30–45 min; ≈ 30 min of high-quality data).	Two sessions (PRE and POST), each including REST and controlled breathing; four 5-min segments in total.
Signals and acquisition	ECG (bipolar thoracic lead, 1 kHz) and continuous beat-to-beat arterial blood pressure (Finometer Pro).	Bipolar chest-lead ECG (1600 or 2048 Hz) and continuous blood pressure waveforms (Portapres 2 or Finapres Nova, 100–200 Hz).	Three-lead ECG (RA–LA–LL, 1 kHz), infrared photoplethysmography (940 nm, 1 kHz), and respiration signal.
Derived time series	RR, SBP, DBP, PP.	RR, SBP, DBP, PP.	RR, PP.

7.2 Stress Characterization and Classification

This chapter applies the analytical framework developed in the previous sections to the study of acute physiological stress. The focus is not on reintroducing concepts already discussed in previous chapters, but rather on illustrating how these tools are employed across different datasets and methodological contexts to capture stress-induced changes in autonomic regulation.

The stressors considered in these studies encompass both mental and physical challenges. Psychological stress has been described as the feeling experienced when individuals are overwhelmed and struggle to meet external demands [97], whereas orthostatic stress represents a physical perturbation caused by the effects of gravity on the distribution of circulating blood during a change in body position. These well-characterised perturbations evoke quantifiable shifts in sympathovagal balance and form a controlled basis for assessing the discriminative value of physiological features.

The American Physiological Association has defined stress as ”the pattern of specific and non-specific responses an organism makes to stimuli events that disturb its equilibrium” [78]. Excess stress has been recognised as one of the most significant pathogenic factors of modern life [231]. In everyday settings, psychological stress typically arises when individuals perceive that environmental or cognitive demands exceed their adaptive capacity [97]. Orthostatic stress, by contrast, represents a well-controlled physical perturbation: the gravitational shift in blood volume triggered by a change in posture induces predictable autonomic adjustments aimed at maintaining arterial pressure. These two classes of stressors, cognitive and postural, provide a complementary framework for evaluating how different analytical techniques capture sympathovagal reconfiguration under varying regulatory demands.

Across the applications presented in this chapter, the emphasis is placed on the obtained results and their interpretation. Although the methodological structure of each study (datasets, acquisition protocols, and preprocessing) is briefly outlined for completeness, detailed descriptions are omitted as they have already been discussed in the dedicated chapters of this thesis. Instead, attention is devoted to how different analytical strategies, ranging from classical machine learning models to information-theoretic measures, interaction-based indices, and normalization techniques, behave when applied to the discrimination of resting, postural, and mental stress conditions.

The studies examined here collectively demonstrate how cardiovascular and cardiorespiratory dynamics respond to diverse stressors, how these responses differ across experimental paradigms, and how various analytical approaches capture such differences. Taken

together, they provide a comprehensive view of the practical performance, limitations, and interpretability of the methods introduced earlier, showing their applicability both in controlled laboratory settings and in more ecologically valid scenarios.

The results in this section are presented through a series of application studies, each corresponding to a specific experimental context and stress paradigm, and each accompanied by its own results and focused discussion. The first six studies address the classification of acute stress conditions induced by orthostatic and mental challenges, providing a systematic evaluation of how different feature sets and analytical strategies discriminate resting and stress states under controlled laboratory conditions. These are followed by two additional studies targeting more ecologically valid scenarios: one focused on driving-related stress and one on work-related stress in healthcare professionals. While each study is discussed individually in its respective section, a comprehensive and integrative discussion that compares findings across all stress paradigms and analytical approaches is deferred to Chapter 8, where the broader physiological and methodological implications are examined.

7.2.1 Comparison of Machine Learning Approaches for Physiological States Classification Using Heart Rate and Pulse Rate Variability Indices

Introduction. This application [94] evaluates the ability of cardiovascular variability features to discriminate physiological states associated with rest, orthostatic stress, and mental stress. The work compares heart-rate-based and pulse-rate-based indices and examines how different machine learning classifiers and a feature selection step influence classification performance.

Materials and Methods. The analyses were performed on the dataset **D1**, introduced in the chapter (7.1), which includes ECG and PPG recordings from healthy young participants monitored during three controlled physiological conditions (REST, head-up tilt (HUT), and mental arithmetic stress (MA)). From these signals, short-term inter-beat interval series were extracted to obtain HRV (from ECG) and PRV (from PPG) time series.

A set of cardiovascular variability features was computed from each time series. These included time-domain indices (MEAN, SDNN, RMSSD), frequency-domain indices (LF and HF powers and their normalized components, LF/HF ratio, respiratory peak frequency), and information-theoretic indices (entropy, conditional entropy, and self-entropy). Their definitions and physiological interpretations are described in the methodological Chapter 3.

The classification task involved discriminating among resting, orthostatic, and mental stress states. Four supervised machine-learning classifiers were employed: LDA, k-NN,

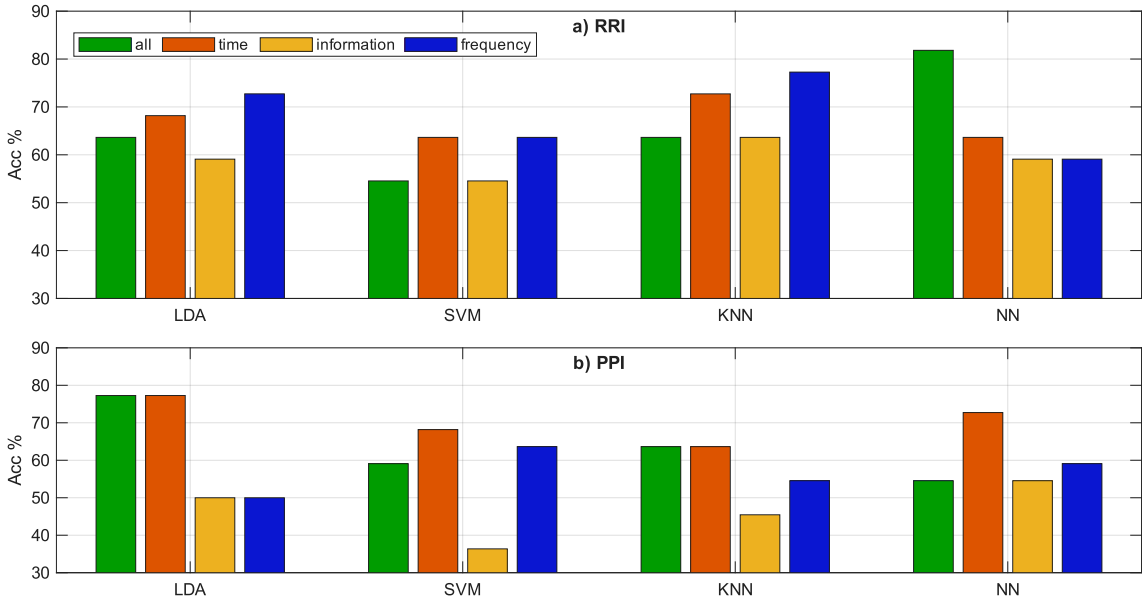


Figure 7.2: Classification accuracy of the four ML algorithms for (a) RRI and (b) PPI indices, considering the features altogether (blue), or grouped according to their domain (orange: time; yellow: information; purple: frequency). This figure has been modified from [94].

SVM, and a NN. In addition to analysing all computed features, a feature-ranking stage based on the mRMR method [172] (Sec. 4.1.1) was performed to evaluate the effect of reducing the feature set. Classification performance was assessed using cross-validation, and results were expressed in terms of accuracy and standard per-class metrics (Chapter 6).

Results and discussion. In this paragraph we first investigate the performance of the four ML algorithms in classifying the different physiological states, using either RRI or PPI features, all together or grouped according to the domain of calculation of the feature (time, frequency or information). Subsequently, a feature selection phase through mRMR was performed, in order to assess how the classification accuracy varies by removing one by one feature according to the order established by the feature selection algorithm, once again comparing the use of either RRI or PPI indices.

Classification results without feature selection Figure 7.2 shows the classification accuracy of the four ML algorithms (LDA, SVM, kNN, NN) in classifying the physiological state (rest, orthostatic or mental stress) when using all the 12 HRV or PRV features, or using only a group of features according to their domain (time, frequency and information).

The accuracy values computed using PRV features extracted from PPG signals are overall slightly lower than those achieved obtained with HRV indices. This behaviour suggests that the difference between RRI and PPI affects the capability of discriminating

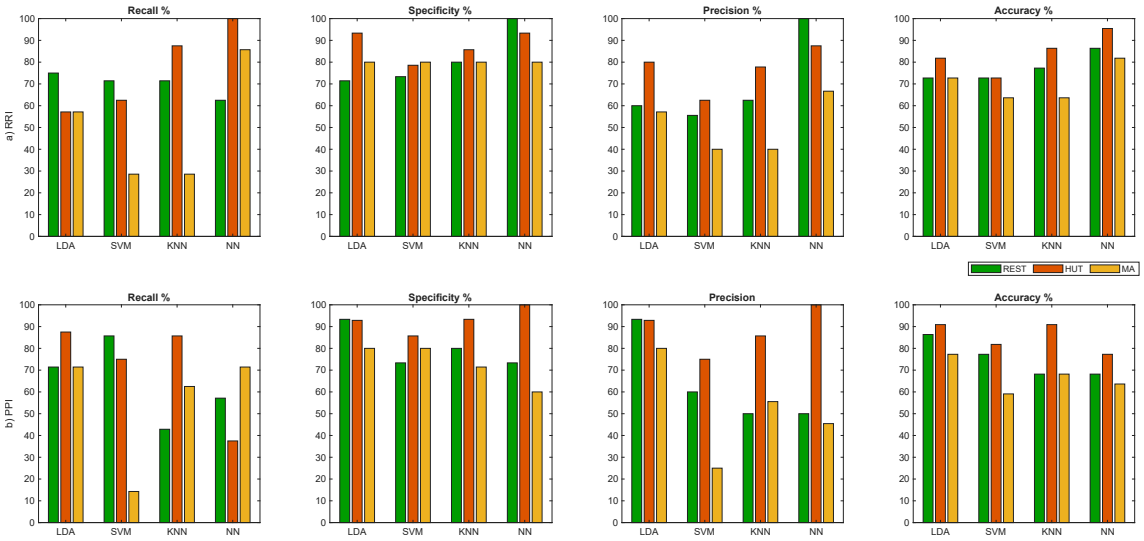


Figure 7.3: Recall, specificity, precision and accuracy metrics by class (blue: REST; orange: HUT; yellow: MA) computed on (a) RRI and (b) PPI features. This figure has been modified from [94].

among classes, even if it does not happen in all cases. Results also evidence that using only information-domain indices produces a lower accuracy for both RRI and PPI (almost always below 60%), and not allowing correct classification of the three physiological stages. On the other hand, time- and frequency domain indices allow to achieve higher values of accuracy, the former performing always better in the case of PPI, the latter overall better in the case of RRI features (apart from NN). Considering all the HRV features together and comparing the ML classifiers, NN achieves the best results, performing however worse when considering only certain groups of features. This behaviour is expected, given that NN needs a larger amount of data and this need is not met when the features are considered separately by domains. The accuracy obtained by LDA classifier on PPI features was similar to that achieved by NN using RRI indices, thus supporting the use of PPG for identifying stress. Figure 7.3 depicts the results in terms of metrics by class, including recall, specificity, precision, and accuracy, to understand which stress state is best classified by each ML algorithm considering all the 12 features together.

Overall, recall values are lower for PPI than RRI, with the maximum value obtained in the case of HUT class (100%). With regard to PPI indices, the highest recall value (87.5%) is obtained with LDA classifier again for HUT class. Analysing the other metrics, we highlight that higher values of specificity, precision and accuracy are achieved with the NN classifier for RRI and with LDA for PPI. Nonetheless, the NN classifier is able to reach the maximum specificity and precision (100%) also for PPI with regard to HUT class. Summarising, the HUT class is classified by all the ML algorithms and for both RRI and PPI indices better than MA that in turn achieves lower values on average. This may

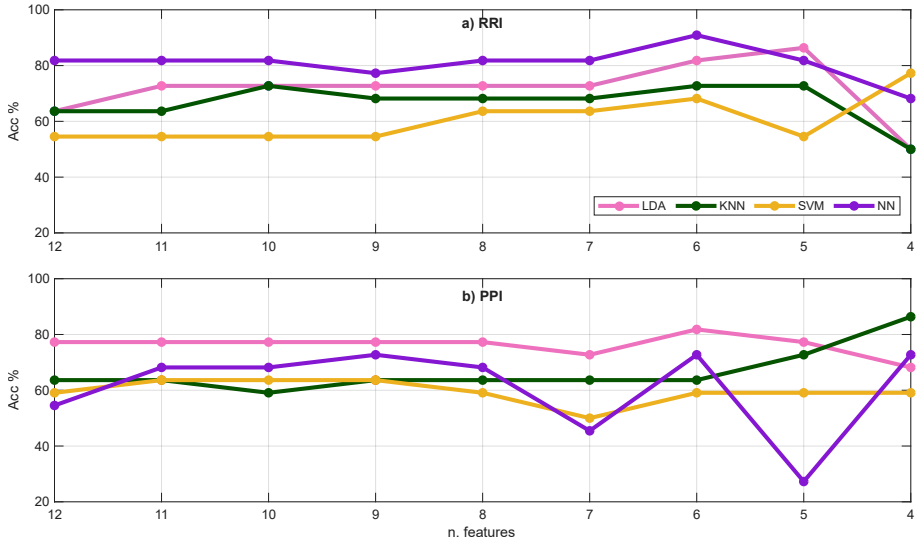


Figure 7.4: Accuracy of the 4 ML classifiers at decreasing number of selected features for (a) RRI and (b) PPI data. This figure has been modified from [94].

be ascribed to the fact that orthostatic stress produces a stronger sympathetic activation and/or parasympathetic inhibition if compared to the mental one, in agreement with previous findings in the literature [175].

Classification results with mRMR feature selection Figure 7.4 depicts the results in terms of accuracy obtained through the 4 ML classifiers performing first a feature selection phase using the mRMR algorithm. Table 7.3 lists the feature importance ranking according to mRMR (in descending order) for RRI and PPI indices, i.e. the order of minimally redundant and maximally relevant features. Fig.7.4(a) and (b) show the accuracy of each classifier, respectively for RRI and PPI indices, as the input features, starting from all 12, are gradually removed one by one, as in the table in inverse order (less important features are removed first). Also in this case, NN is the classifier with the best results for RRI (with a maximum accuracy of 90.9%), while LDA is the best for PPI (81.8%). For most classifiers and for both RRI and PPI indices, there is a slight increase in accuracy when reducing the features, with a peak around 6 or 5 features (apart from NN for PPI which presents an irregular trend). Overall, the maximum accuracy achieved with feature selection is higher than without (cf Fig. 7.2), with a more marked increase for RRI than PPI. For instance, for NN applied to RRI features the maximum accuracy increases from 81.8% to 90.9%, while for LDA applied to PPI the increase is from 77.2% to 81.8%. In [79] authors applied 3 out of 4 classifiers used in this work (kNN, LDA, SVM) and mRMR feature selection algorithm on RRI data in various conditions of stress corresponding to a range of everyday life conditions, obtaining accuracy values of 66.7% (kNN), 65.6% (LDA) and 73.6% (SVM) without pairwise transformation method. Even if the database and the stress conditions are not the same and the order of features selected by mRMR is

different, our accuracy values are in line and sometimes higher.

n	1	2	3	4	5	6	7	8	9	10	11	12
RRI	HF	SE	LF	CE	MEAN	HFn	RMSSD	LFn	SDNN	SVB	H	fHF
PPI	HF	SE	LF	MEAN	RMSSD	HFn	CE	SDNN	LFn	SVB	H	fHF

Table 7.3: Feature importance ranking according to mRMR algorithm (in descending order). This table has been taken from [94].

Conclusion. This work presented a comparison of the performances of various ML algorithms for classifying stress and rest states using either HRV or PRV features. Our results evidence that HRV features allow for achieving higher accuracy than PRV and that it is easier to discriminate postural than mental stress from rest. Moreover, the NN classifier outperforms the other ones, even if LDA works better on PPI data. Finally, a feature selection phase through the mRMR algorithm allows to increase accuracy.

7.2.2 Classification of Physiological States through Machine Learning Algorithms Applied to Ultra-Short-Term Heart Rate and Pulse Rate Variability Indices on a Single-Feature Basis

Introduction. This study [95] evaluates the feasibility of classifying stress states using machine learning algorithms applied to short-term (ST) and ultra-short-term (UST) HRV and PRV features computed individually. The work extends previous analyses by assessing how classification performance changes as the length of cardiovascular variability time series decreases from approximately five minutes down to only ten heartbeats. With the increasing adoption of wearable devices for continuous ECG- and PPG-based monitoring, UST signal analysis has become particularly valuable: shorter windows are less susceptible to motion artefacts, better preserve local stationarity, crucial for HRV and PRV estimation, and enable real-time assessment of physiological stress during daily activities where longer, artifact-free recordings are rarely feasible.

Materials and Methods. The analyses were conducted on the **D1** dataset described in the chapter 7.1, which includes ECG and arterial blood pressure recordings acquired from healthy young adults during three physiological conditions: resting baseline (REST), head-up tilt (HUT), and mental arithmetic (MA). Inter-beat interval series were derived to obtain HRV (RRI) and PRV (PPI), and short-term segments of 300 beats were first analysed. UST evaluation was then performed by progressively shortening the time series in steps down to 10 beats.

From each time series length, twelve cardiovascular variability features were computed, covering time-, frequency-, and information-domain indices already introduced in the

Table 7.4: Classification accuracy values using the three ML classifiers on a single feature basis: comparison of the three-class classification (REST vs. HUT vs. MA) and the two-class classifications (REST vs. HUT or MA) using either RRI and PPI-based features. The values around or below the random guess threshold are reported in bold. This table has been taken from [95].

			MEAN	RMSSD	SDNN	f_{HF}	HF	HF_n	LF	LF_n	SVB	CE	H	SE
REST vs HUT vs MA	NB	*RRI*	61.40	55.26	46.93	35.53	49.12	52.63	36.40	53.51	46.05	51.32	45.18	52.63
		PPI	62.28	55.26	44.74	36.84	46.05	50.00	41.23	52.19	46.93	52.19	42.98	54.82
	SVM	*RRI*	58.77	53.51	42.11	43.86	50.44	52.63	**33.33**	49.56	46.49	51.32	39.47	52.19
		PPI	58.77	52.19	39.91	38.60	46.05	50.44	**32.89**	47.37	44.74	49.56	38.60	51.32
	NN	*RRI*	65.35	44.30	35.09	35.09	45.18	49.56	34.65	46.93	49.12	52.19	41.67	49.56
		PPI	65.35	47.81	37.28	37.72	43.42	43.42	**28.07**	45.61	45.18	45.61	41.67	42.54
REST vs HUT	NB	*RRI*	88.82	77.63	66.45	53.29	71.05	80.26	53.29	80.26	67.11	76.32	67.76	81.58
		PPI	90.13	77.63	65.13	51.97	69.08	77.63	**50.00**	76.32	66.45	74.34	65.79	76.97
	SVM	*RRI*	88.16	76.32	66.45	54.61	69.74	78.29	53.29	80.26	61.18	78.29	67.11	81.58
		PPI	88.82	76.97	65.13	**50.66**	67.76	74.34	51.97	75.00	65.79	74.34	66.45	76.97
	NN	*RRI*	86.84	76.97	61.18	55.26	73.68	72.37	53.95	78.95	78.29	75.66	60.53	79.61
		PPI	85.53	76.97	65.79	**50.66**	67.11	69.74	**50.66**	71.71	70.39	71.05	57.24	73.68
REST vs MA	NB	*RRI*	73.68	62.50	61.18	**48.68**	61.18	60.53	55.92	59.21	60.53	**48.68**	61.84	**50.00**
		PPI	73.03	62.50	61.18	54.61	61.84	61.84	55.92	59.87	61.18	53.95	59.87	55.92
	SVM	*RRI*	73.03	59.87	61.18	51.32	54.61	61.84	53.95	58.55	51.97	51.97	61.18	**49.34**
		PPI	73.03	63.16	61.18	56.58	54.61	61.84	51.32	59.87	56.58	**50.00**	60.53	**50.00**
	NN	*RRI*	76.32	**50.00**	53.29	53.95	54.61	52.63	**43.42**	**42.76**	57.89	51.32	51.97	**50.66**
		PPI	77.63	51.97	55.26	**46.05**	**50.66**	57.89	**42.76**	51.97	51.97	**50.00**	**49.34**	**44.74**

methodological chapters. Three commonly used machine learning algorithms, NB, SVM, and NN classifiers, were trained separately for HRV and PRV and for each time series length, using a single-feature classification approach. A 5-fold cross-validation procedure (80% training, 20% testing) was employed, and accuracy, recall, and specificity were computed for all classifiers and features.

Results. Table 7.4 reports the accuracy values obtained for the three ML approaches in both the three-class and two-class classifications on a single-feature basis, to assess which state is best classified by each feature, using either HRV or PRV indices. In classification tasks, accuracy refers to the percentage of correct predictions the classifier makes. In a two-class classification problem, there are only two possible classes to which data can belong, hence the random guess accuracy is 50% (the same applies to recall and specificity). On the other hand, in a three-class classification problem, there are three possible classes to which the data can belong, thus the random guess threshold is equal to 33.33%. Values around or below the random guess threshold are reported in bold.

The accuracy values of the three-class classification (REST vs. HUT vs. MA) (roughly 50%) were lower than those obtained for the two-class approach (REST vs. HUT and REST vs. MA) which achieved average values of roughly 70% and 60%, respectively. When classifying REST and MA, the only feature to achieve accuracy values higher than 65% was the MEAN, with values of 73.68%, 73.03%, and 76.32% achieved for the

Table 7.5: Recall (REC) and specificity (SPE), for the HUT class computed using RRI and PPI-based indices on a single-feature basis for the REST vs. HUT classification. Values lower than the 65% threshold are in bold, while features with accuracy higher than 65% (cf. Table 7.4) are reported in grey. This table has been taken from [95].

			MEAN	RMSSD	SDNN	f_{HF}	HF	HF_n	LF	LF_n	SVB	CE	H	SE
REC	NB	*RRI*	88.16	92.11	77.63	**50.00**	96.05	81.58	69.74	82.89	**39.47**	72.37	73.68	77.63
		PPI	89.47	92.11	77.63	**39.47**	94.74	81.58	**63.16**	78.95	**40.79**	72.37	72.37	75.00
	SVM	*RRI*	84.21	92.11	85.53	**60.53**	96.05	77.63	96.05	80.26	**27.63**	75.00	78.95	81.58
		PPI	86.84	90.79	84.21	**42.11**	96.05	71.05	93.42	75.00	**35.53**	71.05	80.26	75.00
	NN	*RRI*	88.16	82.89	73.68	**60.53**	78.95	73.68	**57.89**	75.00	73.68	75.00	68.42	80.26
		PPI	85.53	76.32	77.63	**53.95**	72.37	72.37	**56.58**	72.37	67.11	67.11	65.79	75.00
SPE	NB	*RRI*	89.47	**63.16**	**55.26**	**56.58**	**46.05**	78.95	**36.84**	77.63	94.74	80.26	**61.84**	85.53
		PPI	90.79	**63.16**	**52.63**	**64.47**	**43.42**	73.68	**36.84**	73.68	92.11	76.32	**59.21**	78.95
	SVM	*RRI*	92.11	**60.53**	**47.37**	**48.68**	**43.42**	78.95	**10.53**	80.26	94.74	81.58	**55.26**	81.58
		PPI	90.79	**63.16**	**46.05**	**59.21**	**39.47**	77.63	**10.53**	75.00	96.05	77.63	**52.63**	78.95
	NN	*RRI*	85.53	71.05	**48.68**	**50.00**	68.42	71.05	**50.00**	82.89	82.89	76.32	**52.63**	78.95
		PPI	85.53	77.63	**53.95**	**47.37**	61.84	67.11	**44.74**	71.05	73.68	75.00	**48.68**	72.37

NB, SVM, and NN classifiers, respectively. Higher accuracy values were obtained when distinguishing between REST and HUT classes (in grey), with multiple features exceeding an accuracy of 65% for RRI and PPI-based indices.

Table 7.5 illustrates the recall and specificity results for each ML algorithm considering each feature individually for the HUT class in the REST vs. HUT classification. The lowest values were obtained for f_{HF} (both metrics) and for SDNN, LF and HF (with regard to specificity only). Such trends are similar to those found for accuracy (reported in Table 7.5), with values around the random guess (50%) for SDNN, f_{HF} and LF. Overall, the specificity values were lower than recall, often below 60%.

For the further analyses focusing on shorter time series, we decided to consider only the best features in terms of accuracy, by setting a threshold of 65% allowing at least 15% improvement over a random guess [160]. Features that exceeded the 65% accuracy threshold (cf. Table 7.5) for all the classifiers and for both HRV and PRV indices were CE, SE, HF, HF_n, LF_n, MEAN and RMSSD and are reported in grey in Table 2. Such features achieved as well, specificity and recall values for the HUT class in most cases higher than 80%.

Figure 7.5 depicts the accuracy of the three ML classifiers and the most significant features while decreasing the length of RRI and PPI time series. The information-based features were computed down to 60 heartbeats while the frequency-domain features were computed down to 20 heartbeats (black vertical lines). All features showed a decreasing trend with heartbeats, from a few up to a maximum of 10 percentage points. Only the time domain features continued to yield higher accuracy values for all the classifiers for both RRI and PPI-based indices till a few heartbeats, with values around 80% (RMSSD) or even 90% (MEAN).

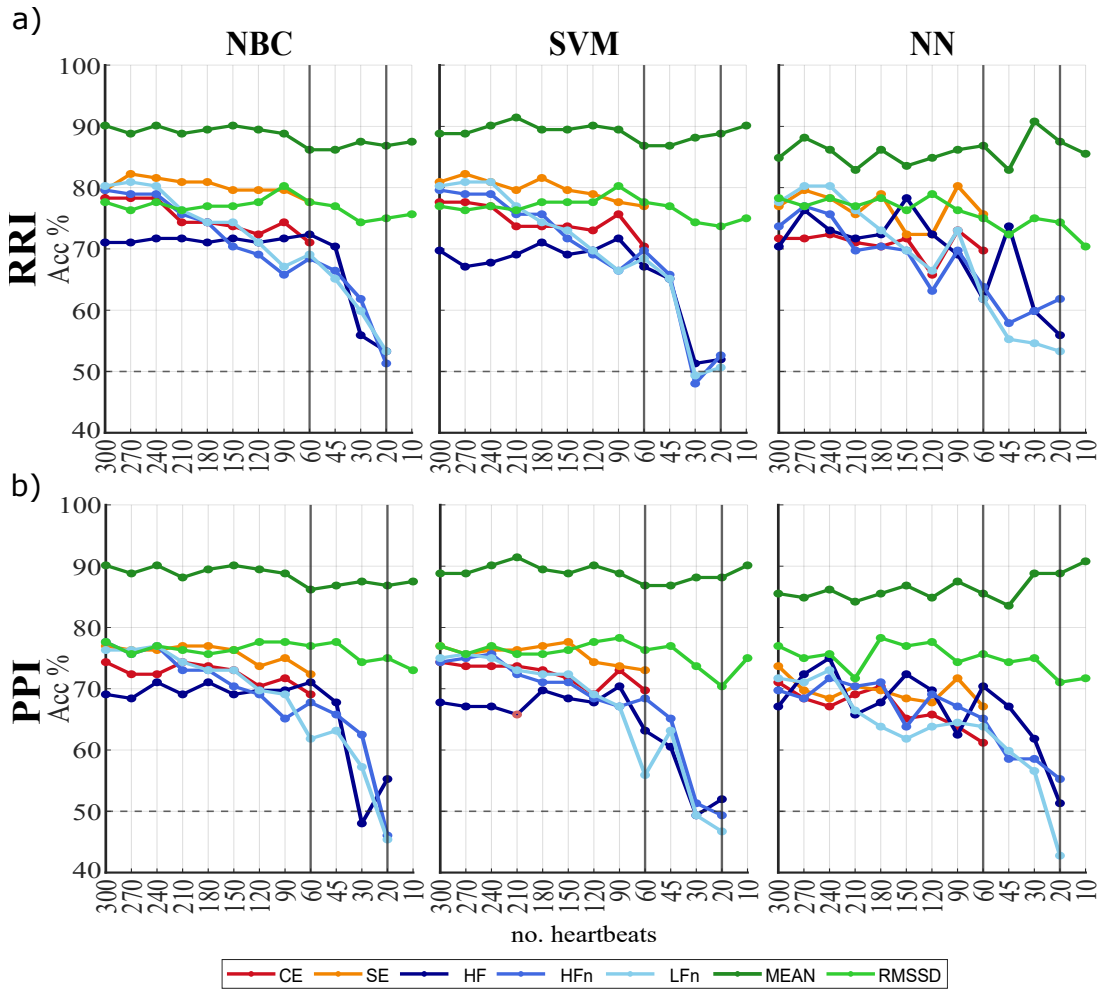


Figure 7.5: Accuracy of the three ML classifiers on a single-feature basis at decreasing time series length (heartbeats) for (a) HRV and (b) PRV indices. This figure has been adapted from [95].

Discussion. First, our attention will be devoted to examining the viability of classifying physiological states by utilising various machine learning (ML) algorithms based on individual features of standard short-term HRV or PRV. Both three-stage (REST vs. HUT vs. MA) and two-stage classification (REST vs. HUT or vs. MA) will be taken into account. Subsequently, the viability of employing shorter time series (i.e. UST analysis) as a substitute for the short-term standard will be discussed, focusing only on REST-HUT classification.

Classification using short-term HRV and PRV features. The accuracy values obtained with the three-class classification (Table 7.4) were indicative of an arbitrary classification primarily attributable to the classifiers' inability to effectively distinguish between the REST and MA classes. This observation is confirmed by the classification results obtained for the two-class REST vs. MA classification approach, where very low accuracy values were obtained. The classification accuracy was higher when distinguishing between the REST and HUT classes, for either RRI or PPI-based indices. This finding is deemed reliable as it has been previously demonstrated that the orthostatic stress condition is more marked than MA stress [240], given the stronger sympathetic activation and the concurrent parasympathetic withdrawal [170, 175].
In light of these results, the focus was shifted to REST vs. HUT classification, taking in this case also into account per-class metrics for a more detailed assessment of the effectiveness of the classification model. The results shown in Table 7.5 highlighted recall values consistently above 60%, reinforcing the effectiveness of the classification model for discriminating between the REST and HUT classes. On the other hand, only the frequency-domain HF and SVB features, strongly related to each other, exhibited low values across all three classifiers and for both RRI and PPI data. Finally, similar accuracy and per-class metrics results were obtained using either HRV or PRV indices (cf. Tables 7.4 and 7.5).

REST vs HUT classification using Ultra-Short-Term features. To assess the feasibility of using UST analysis, only the features that yielded accuracy values exceeding 65% for all three classifiers were selected, i.e. MEAN and RMSSD w.r.t. time-domain, HF, HF_n and LF w.r.t. frequency-domain and finally CE and SE w.r.t. information-domain features. Using the above-mentioned seven features, the three classifiers were trained and tested by varying the length of the series from the initial 300 heartbeats down to 10 heartbeats. Although according to [41, 212] a minimum of 1 minute is needed to reliably estimate HF and at least 2 minutes for the LF component with regard to frequency-domain HRV/PRV analysis, we have adopted the approach followed by several other studies which have instead computed such features on even shorter time series down to 10 s [28, 70, 116]. In our case, it was feasible to compute the information domain features using 60 heartbeats as

a minimum. On the contrary, the frequency features could only be computed up to 20 beats, while they started already to lose their discriminating capability from 60 beats downwards. Our results highlight that the accuracy values are almost constant when considering time- and information-domain UST features (until they can be computed). On the other hand, frequency features accuracy values exhibited a decreasing trend, particularly when going below 60 heartbeats. Moreover, the three classifiers exhibited similar accuracy levels, and this was also observed when considering either HRV or PRV features. As already demonstrated by [41], not all UST indices are good short-term HRV and PRV surrogates, and generally, time domain features (e.g. MEAN and RMSSD) proved more effective in classifying the two physiological states. Although employing different methods and datasets, these results are in agreement with ours [28, 41, 70]. Herein, MEAN and RMSSD time domain features yielded higher accuracy values for RRI and PPI in all the classifiers, underscoring their high potential for classifying postural stress.

Conclusion. This work presented a comparison of the performances of various ML algorithms to classify physiological states using either HRV or PRV indices on a single-feature basis. Our results indicated comparable results for RRI or PPI-based indices and confirmed that postural stress is easier to discriminate than mental stress. Our findings identified some reliable UST HRV and PRV features (e.g. MEAN and RMSSD) that can be employed for detecting postural stress even for very few heartbeats (~10 s). UST frequency-domain HRV and PRV indices were instead worse short-term surrogates, with decreased accuracy for shorter (<60 s) time series.

7.2.3 Comparison of Automatic and Physiologically-based Feature Selection Methods for Classifying Physiological Stress Using Heart Rate and Pulse Rate Variability Indices

Introduction. This study [97] examines and compares two complementary feature-selection strategies: an automatic, information-theoretic approach and a physiologically driven selection, in their ability to identify HRV and PRV features that best discriminate among resting, orthostatic, and mental stress conditions. The study investigates how the choice of feature subset affects classification accuracy across multiple machine-learning algorithms, thereby providing insight into the relative advantages of data-driven versus physiology-informed approaches in stress-state discrimination.

Materials and Methods. The study was conducted on the **D1** dataset described in the Sec. 7.1, consisting of ECG and arterial blood pressure recordings collected from healthy young adults during three distinct physiological conditions: resting baseline (REST), head-up tilt (HUT), and mental arithmetic (MA). From these recordings, HRV (RRI) and PRV (PPI) time series were extracted, and a set of short-term cardiovascular variability features

was computed across time-, frequency-, and information-domain indices, as specified in the methodological Chapter 3.

Two feature selection methods were evaluated, an automatic information-theoretic approach based o Akaike criterium (AIC-FS, 4.1.2), and a physiologically based selection derived from known autonomic signatures of different stressors.

The selected features were then used as inputs to five machine learning classifiers: LDA, SVM, k-NN, NN, and NB Classifiers as described in 6.1. Classification performance was quantified through performance metrics and confusion matrices.

Results. In the following, the results will be presented, focusing on classification according to the AIC-FS algorithm or the physiologically-based feature selection method.

Akaike feature selection method. The results in Fig. 7.6 show the frequency of selection for each feature across the 10 CV folds for (a) RRI and (b) PPI time series using the AIC-FS algorithm. Regarding the RRI time series, features such as CE, SE, HFn, and MEAN demonstrated higher selection frequencies, more than 70% over the 10-fold, while for the PPI time series features like SE, fHF, HFn, MEAN, and RMSSD were more commonly selected.

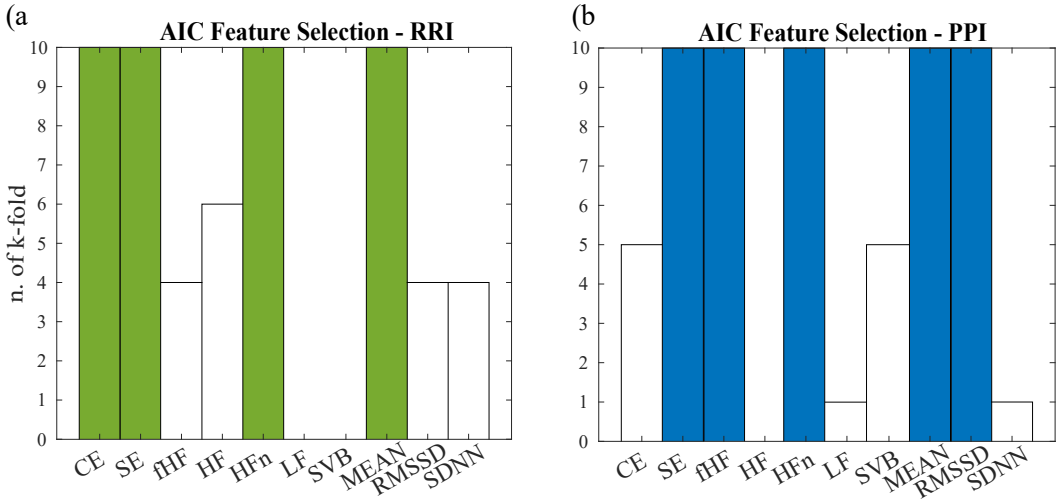
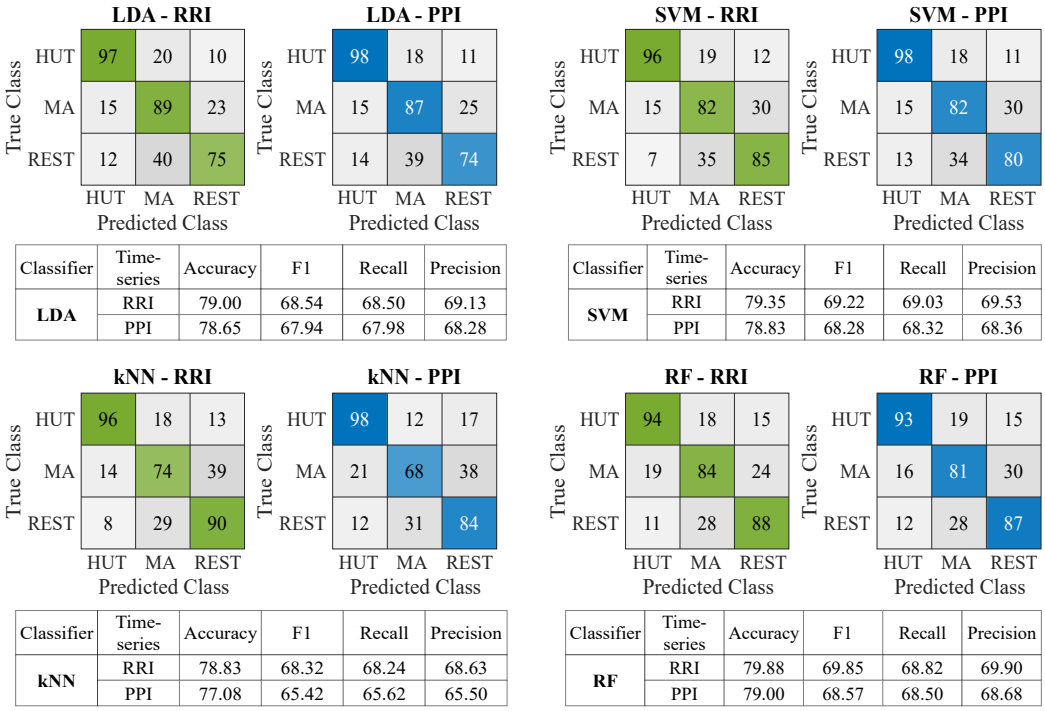


Figure 7.6: Frequency of selected features using Akaike Information Criterion across 10-fold cross-validation. The bar plot illustrates the number of times each feature was selected during the feature selection process. The coloured bars represent the features that were selected more frequently (> 70%) for the RRI and PPI time series. This figure has been adapted from [97].

As shown in the confusion matrices of Fig. 7.7, the HUT class consistently ranks best, achieving higher True Positive values than the other two classes. On the other hand, the high values shown in grey in the confusion matrix suggest that both MA and REST phases have been more often incorrectly classified. Across the four classifiers (LDA, SVM, KNN,

and RF) the RF generally achieves the highest results for both time series (Fig. 7.7). In contrast, kNN showed the weakest performance across all computed metrics compared to the other classifiers. However, the Wilcoxon signed-rank test for the four metrics (accuracy, recall, precision, and F1 score) across the classifiers for each time series (RRI and PPI) indicated that the differences were not statistically significant. Additionally, the Kruskal-Wallis test revealed no significant overall differences for each measure between the RRI and PPI time series. These findings suggest that, despite the observed differences in performance, the classifiers do not significantly differ in their performance when applied to the RRI and PPI time series.



Classifier	Time-series	Accuracy	F1	Recall	Precision
LDA	RRI	79.00	68.54	68.50	69.13
	PPI	78.65	67.94	67.98	68.28

Classifier	Time-series	Accuracy	F1	Recall	Precision
SVM	RRI	79.35	69.22	69.03	69.53
	PPI	78.83	68.28	68.32	68.36

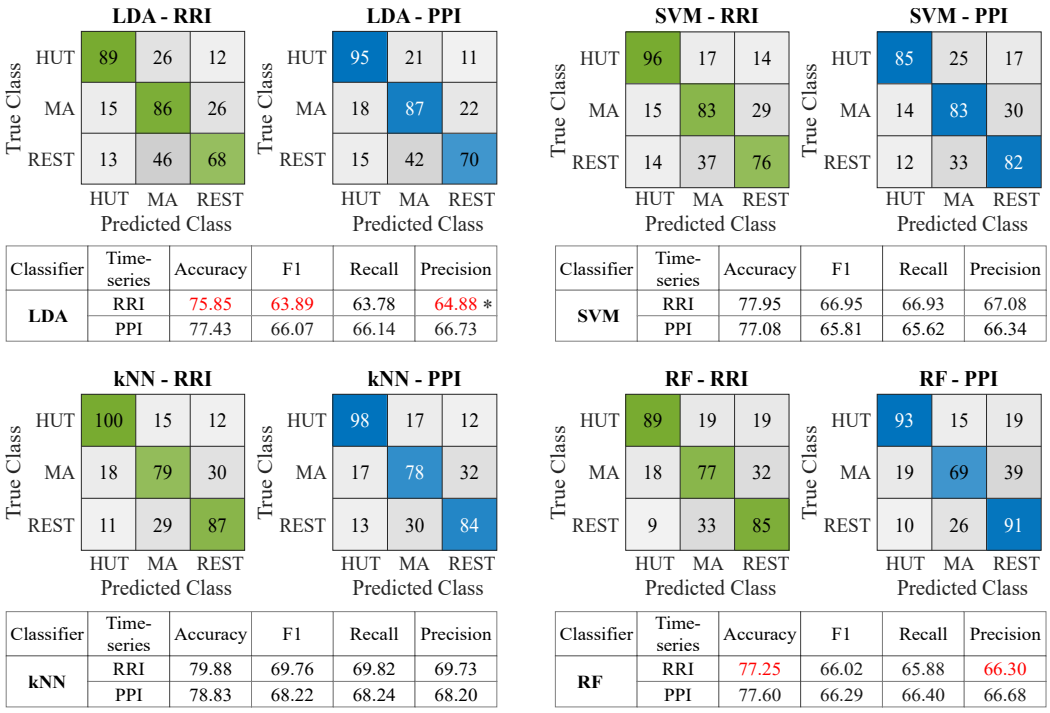
Classifier	Time-series	Accuracy	F1	Recall	Precision
kNN	RRI	78.83	68.32	68.24	68.63
	PPI	77.08	65.42	65.62	65.50

Classifier	Time-series	Accuracy	F1	Recall	Precision
RF	RRI	79.88	69.85	68.82	69.90
	PPI	79.00	68.57	68.50	68.68

Figure 7.7: Confusion matrices, tables of the classification performance results (Accuracy, Recall, Precision and F1-score) for each classifier (LDA, SVM, kNN, RF) and each time series (RRI and PPI). No statistically significant differences were observed in the comparisons between RRI and PPI, or across the classifiers. This figure has been adapted from [97].

Physiologically-based feature selection method. Fig. 7.8 presents the classification values achieved by the four machine learning classifiers, taking into account the proposed most informative features from a physiological perspective (CE, HF, LF, MEAN, SDNN). The results of the confusion matrices show that the HUT class regularly gets the highest rating, as reported in Fig. 7.7. It routinely earns higher True Positive values than the other two states. The latter appears to have been more often wrongly categorised, as shown by the high values given in grey in the confusion matrix of Fig. 7.8. The classifier achieving the best performance considering the most physiologically informative features

is the kNN classifier which outperforms the others. Nonetheless, the results of statistical tests indicated once again no significant overall differences between classifiers, suggesting that, when considering each metric individually across both time series, the classifiers do not significantly differ in their performance. When comparing values obtained using RRI or PPI time series for each classifier and metric, significant differences ($p < 0.05$) were acived for precision when considering the LDA classifier ($p = 0.031$, $*$ in Fig. 7.8).



Classifier	Time-series	Accuracy	F1	Recall	Precision
LDA	RRI	75.85	63.89	63.78	64.88 *
	PPI	77.43	66.07	66.14	66.73

Classifier	Time-series	Accuracy	F1	Recall	Precision
SVM	RRI	77.95	66.95	66.93	67.08
	PPI	77.08	65.81	65.62	66.34

Classifier	Time-series	Accuracy	F1	Recall	Precision
kNN	RRI	79.88	69.76	69.82	69.73
	PPI	78.83	68.22	68.24	68.20

Classifier	Time-series	Accuracy	F1	Recall	Precision
RF	RRI	77.25	66.02	65.88	66.30
	PPI	77.60	66.29	66.40	66.68

Figure 7.8: Confusion matrices, tables of the classification performance results (Accuracy, Recall, Precision and F1-score) using only the features with the most informative physiological meaning (CE, HF, LF, MEAN, SDNN), for each classifier (LDA, SVM, kNN, RF) and each time series (RRI and PPI). Statistical test: *, $p < 0.05$, pairwise Wilcoxon test between the distribution of metrics computed for RRI vs PPI across the 10-fold CV. Red values, $p < 0.05$, unpaired Wilcoxon test between distribution of metrics computed using the AIC-FS vs physiological-based FS across the 10-fold CV. This figure has been adapted from [97].

Comparing the results from the AIC feature selection method and the physiologically-based feature selection method, it is possible to notice overal higher accuracy values and F1 scores when using the AIC-FS method. In general, AIC feature selection appears to enhance the performance of most classifiers on both time series. Also in this case, the Wilcoxon signed-rank test was employed between distributions of values computed using each approach to statistically validate the results and compare the performance of classifiers using the two feature selection methods. When considering the RR time series, significant differences were observed in accuracy for LDA ($p = 0.027$) and RF ($p = 0.039$) (highlighted in red in Fig. 7.8) with the physiological-based feature selection method showing lower

values. Precision demonstrated significant differences for LDA ($p = 0.027$) and RF ($p = 0.004$), again with lower values for the physiologically-based approach, while recall showed no significant differences for any classifier. The F1 score indicated a significant difference for LDA ($p = 0.020$), while no other classifiers showed significant differences in this metric. It is worth noting that, concerning the PPI time series, the performance of classifiers using the two feature selection methods did not exhibit statistically significant differences in accuracy, recall, and F1 score.

Discussion. This study aimed to comprehensively compare the results of an automatic feature selection method using the Akaike Information Criterion and a physiologically-based feature selection approach in classifying different physiological states from cardiovascular variability time series. The objective was to gain insights into their performance and to identify the most suitable HRV and PRV features (in time-, frequency- and information-theoretic domain) for the best discrimination among physiological states, highlighting the different physiological mechanisms involved during postural and mental stress accompanied by a shift in the sympathovagal balance.

The feature selection process using the AIC was repeated for each cross-validation fold to ensure a selection specifically tailored to each training set, thus minimizing potential overfitting and enhancing the model's performance. Our results shown in Fig. 7.6 indicate that certain features are consistently chosen more frequently than others. Specifically, for the RRI time series features such as CE, SE, HFn and MEAN demonstrated higher selection frequencies, suggesting their greater relevance in predicting the categorical target variable. Different features like SE, fHF, HFn, MEAN and RMSSD were more commonly selected concerning PPI time series. It is interesting to underline that each set includes features belonging to all three domains, i.e. time, frequency, and information, highlighting the diverse nature of the relevant features for each time series. The different selection of the features suggests that RRI and PPI time series present specific aspects that can be better explained from a peculiar features set.

In the time domain, the MEAN was always selected across the cross-validation folds when considering both RRI and PPI time series. These results are expected, given that the MEAN feature is widely recognized as the basic cardiovascular measure expressing the overall balance between parasympathetic and sympathetic control, as the SE features of the information domain [187]. Moreover, the selection of RMSSD using the PPI time series can be related to the fact that RMSSD is more responsive to the parasympathetic nervous system (PNS) activity being, in turn, less influenced by respiratory frequency than HF power [173, 193]. The frequency-domain analysis reveals differential behaviour of the various indices across the two types of time series. For the PPI time series, the fHF and the HFn features are always selected across the 10-folds, while only the

HFn is always selected for the RRI time series. This different behaviour is difficult to physiologically interpret, even if we can hypothesize that the selection of fHF for PPI is related to the strong influence of respiration on PEP that affects PPI time series (and not RRI series). It is widely known that information-theoretic metrics reveal the complexity of physiological processes and thus cardiovascular time series, varying across different physio-pathological conditions [82, 184, 186]. Entropy-based measures, i.e CE and SE, have been shown to respond to different degrees of sympathetic activity during postural stress and can effectively describe variations in the cardiac system activity linked to diverse physiological and clinical conditions [175, 249]. The CE is a measure of complexity that quantifies the dynamics of a time series by assessing the information in the current data point that is not explained by previous points, while the SE is the portion of information which can be derived from its past history [105]. A rise in CE reflects an increase in series complexity, while higher SE indicates a more regular and predictable series [19]. Our results demonstrate that both CE and SE indices are consistently selected across all folds for RRI time series, suggesting that complexity measures represent features that can effectively explain differences among physiological states. The same remarks apply to the PPI time series, but with the difference that only SE was consistently chosen.

The difference between HRV and PRV selected indices may be related to the distorting effect of the non-constant pre-ejection period (PEP), depending mainly on left ventricular contractility and on pulse transit time (PTT), which has also been shown to exhibit physiological variability [44]. Respiration may also affect ventricular loading and thus PEP, as well as non-physiological factors like an inaccurate detection of blood pressure maxima [175]. The confounding effect of PEP and PTT may also have a role in the complexity of the PPI time series, which in turn could be responsible for the non-selection of the SE index.

The results of the AIC-FS method are in quite good accordance with the proposed more physiologically-related feature subset (Fig. 7.8) which includes the MEAN, the information-theoretical CE, and both frequency-domain LF and HF, being the latter mutually more independent. This suggested subset also contains the SDNN, which conversely did not appear to be a feature selected by the automatic AIC-based method. Consistent feature combinations are identified in the studies of [41, 92, 102, 148, 190, 229, 242] which focus on stress and rest detection protocols similar to ours. Nevertheless, they encompass additional features such as pNN50 (i.e., the percentage of differences between duration of successive RR intervals > 50 ms), providing avenues for further exploration in subsequent research of our study.

Our results evidence (cfr. Fig. 7.7 and Fig. 7.8) that the comparison across the four distinct classifiers reveals no statistically significant differences in their performance.

Moreover, the consistency between RRI and PPI outcomes in all except for one metric, alongside the uniformity across various classifiers, underscores the robustness and reliability of the analysis and demonstrates the viability of substituting features derived from HRV by the ones from PRV, confirming previous findings obtained in other works with classical statistical analyses [91, 175, 210]. Moreover, our results (both regarding the AIC-based approach, Fig. 7.7 and the physiological-based selection, Fig. 7.8) show that when considering all the classifiers and both time series, the class with the best classification performance is HUT, followed by REST. This confirms previous findings evidencing that postural stress is easier to be discriminated than mental workload and the different nature of the two stressors [175, 182].

Overall, our results suggest that SVM and kNN classifiers produce similar results either using the automatic or the physiologically-based feature selection approaches, while RF and LDA are more feature-dependent, with lower performance metrics when considering physiologically-based features computed from RRI time series. This suggests that automatic feature selection enhances performance for some metrics and classifiers, but not for the PPI-based analysis. Interestingly, while the automatically selected features differ from the physiologically selected ones (except for the MEAN feature), they hold similar meanings. For instance, both CE and SE indicate sympathetic tone, while SDNN and RMSSD reflect vagal control. These findings suggest that incorporating physiological knowledge into machine learning approaches yields results comparable to blind automatic classification, while also enhancing explainability. The importance of employing a feature selection phase was also emphasized in a previous study 7.2.1 considering a smaller version of the dataset and the backward-type Minimal Redundancy Maximal Relevance (mRMR) feature selection algorithm [57]. The results are in agreement and evidence that the inclusion of the feature selection phase (either AIC-FS or mRMR) leads to better classification accuracy.

A limitation of our study consists in the fact that the differences between the suggested set of most informative physiologically-based features and the automatic feature selection set could stem from distinctive traits in the subjects, i.e., inter-subject variability of the features, and the relatively low amount of available data. Subsequent research endeavours should involve the application of the identified feature set classification to more ”realistic” datasets, i.e. recorded from real-life or within clinical settings, rather than within controlled experimental environments as in this work, so as to validate the study’s potential for practical applications. Moreover, the obtained results from the automatic feature selection could be improved by performing a more sophisticated feature interaction analysis. Additionally, our study specifically evaluates the effects of postural stress induced by head-up tilt and mental stress elicited by mental arithmetic tasks, while further studies should take

into account different types of stressors and stress levels to prove the generability of our findings.

Conclusions. This study demonstrates that specific stress classification is shown to be feasible, especially for postural stress compared to the rest state, being however more difficult to distinguish between postural and mental stress since both stressors evoke similar sympathetic activation. Our results highlight that the choice of the classifier did not impact classification performances. The automatic AIC-based feature selection method overall achieved slightly better classification than the physiology-driven approach for LDA and RF algorithms. Notably, PPI-based classification, even if with a slightly different set of features, demonstrated similar performance to RRI-based one for almost all metrics and classifiers. In perspective, these results support the feasibility of implementing PPI-based algorithms on wearable devices for outpatient settings monitoring. Additionally, our findings underscore the importance of incorporating physiological knowledge into ML models, enhancing explainability while maintaining robust classification performance. The outcome of our investigation holds potential implications for advancing stress assessment methodologies, not only within clinical settings but also for broader contexts of health monitoring and well-being evaluation.

7.2.4 A Classification Method Based on Local Information and Nearest Neighbour Entropy Estimation

Introduction. This application [130] applies the *Local Information Classifier (LIC)* to the problem of stress-state discrimination. The LIC, described in detail in Chapter 6.2 is a non-parametric, information-theoretic classifier that assigns each sample to the class minimizing its local information content, defined as the negative logarithm of the joint feature–class probability. By leveraging local probability density estimation, the method provides an interpretable alternative to conventional machine-learning models, while retaining flexibility in handling complex and nonlinear data distributions.

Materials and Methods. The evaluation of the proposed LIC method was conducted through a real-data application involving HRV features. In particular, HRV and PRV features (MEAN, RMSSD, CE, HF) were extracted from short-term heart rate variability time series acquired in a physiological experiment (**D1** dataset). The Gaussian Bayes (GB) classifier was again used as a baseline for comparison. Classification performance was assessed in terms of accuracy, sensitivity, and specificity, averaged across multiple repetitions to ensure statistical reliability. All analyses followed the methodological definitions and preprocessing steps detailed in the methodological chapters of this thesis (Ch. 6.2).

Results. To investigate the predictive capabilities of the proposed LIC model, we

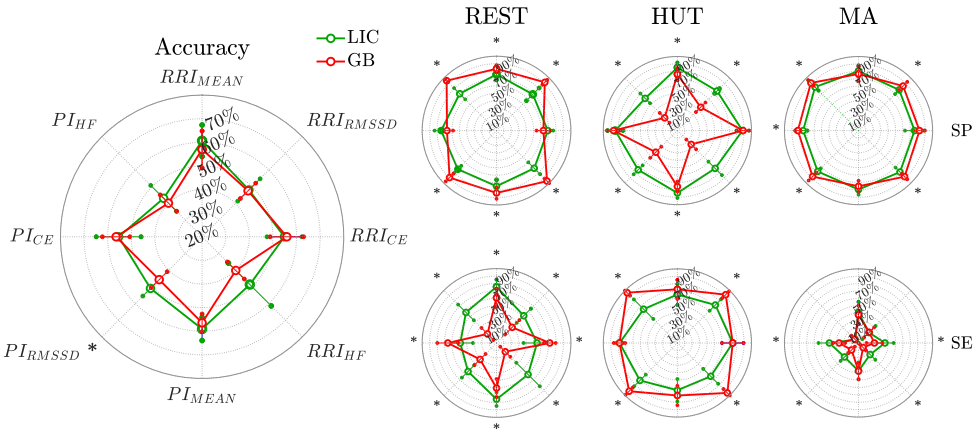


Figure 7.9: Polar plot comparative assessment of the performances of the LIC and GB classifiers in terms of accuracy scores and within-class sensitivity (SE) and specificity (SP) scores performed on a single-feature basis. The plots show values of the mean ± standard deviation for all cases. Statistically significant differences between the cross-validation scores of the models for each feature are denoted with an asterisk (∗) symbol, determined with the pairwise t-test using a significance $\alpha = 0.05$. The figures highlight an overall similar behavior of the scores when considering individual features for the LIC, while GB exhibits larger variations. This figure has been adapted from [130].

evaluated it using real-world datasets and comparing with the GB classifier. In both cases, we utilized $k = 10$ nearest neighbors along with the maximum norm distance metric for the LIC, and employed a form of cross-validation. The performances of both the LIC and the GB classification models were assessed in terms of sensitivity (SE) and specificity (SP) for each class, along with the computation of the overall accuracy.

The classifiers were aimed at addressing a prediction task focused on distinguishing between the two stress and control resting conditions: the orthostatic stress induced by HUT, mental stress induced by the MA test, and the control initial REST state. Features were extracted from the beat-to-beat time series of RR intervals (RRI) and pulse intervals (PI) as described in Chapters 2 and 3. For both RRI and PI, the selected features included: the MEAN, the RMSSD, the CE and the fHF (Chapter 3). A 10-fold cross-validation scheme is applied in this analysis to obtain a distribution over the evaluated measures. Two comparisons are then made: the first to assess the difference between the LIC and GB models, and the second to compare the effectiveness of the selected feature within one model. All the comparisons are made utilizing a pairwise t-test with a significance $\alpha = 0.05$, under the null hypothesis that the difference in means is zero. An additional Bonferroni correction of 28 is applied, concerning the multiple comparisons of model performances for the individual 8 features.

As depicted in Fig. 7.9, the scores closely align and show similar performances for a given feature across both RRI and PI time series. In terms of accuracy, the LIC

demonstrates slightly better performances across most of the observed features, achieving an averaged accuracy of 52.57 ± 5.98% (mean ± standard deviation), while, in comparison, the GB classifier follows closely with slightly lower performances with an accuracy of 49.91 ± 7.41%. The greater variation observed for the GB classifier may arise from its erroneous initial assumption of Gaussianity, a characteristic likely absent in the data. However, only the models built with RMSSD features from the PI series show statistically significant differences.

In evaluating individual classes, both classifiers demonstrate comparable metric values for certain cases. Specifically, the LIC demonstrates an overall better SP for the HUT class (with only both CE features showing non-significant differences), while also showing better SE for the REST (with all comparisons being statistically significant) and MA (with both CE, RRI_{HF}, and PI_{RMSSD} showing significant differences) classes. In contrast, the GB classifier displays large variations in SP across the features concerning the HUT class, and in SE concerning the REST class, suggesting a trade-off between the measures for certain features. These variations reflect the inability of GB classification performed with features yielding low SE scores and consequently very high SP scores (e.g., HF power and $RMSSD$). Both classifiers show underwhelming performances in predicting the MA class, reflecting the difficulty of modeling the mental stress, potentially because it does not elicit marked changes in the main time-domain, spectral and complexity indexes extracted from heart rate variability when compared to the resting state [107].

Concerning within-model feature comparison, as shown in Fig. 7.10, the LIC model shows the highest 60.67 ± 6.67% accuracy when utilizing the RRI_{MEAN} feature, however only with a statistical difference to scores when using RRI_{RMSSD}, PI_{RMSSD} and PI_{HF} features. The lowest performing feature was PI_{HF}, producing an accuracy score of 43.05 ± 7.69%, significantly different than the two $MEAN$ features. On the other hand, the GB classifier achieves its peak accuracy of 56.98 ± 7.83% employing the same RRI_{MEAN} feature, showing statistically significant differences to PI_{RMSSD} and HF features. The lowest accuracy score of 40.16 ± 4.94% was obtained with the same PI_{HF} feature, statistically different from models using $MEAN$ and CE features. These results align with the similarly used setup in [95], for the selected features.

Discussion and Conclusions. This work introduces a novel classification model grounded in information theory, of which we provide both the theoretical background and a practical implementation strategy. The proposed model, minimizes the entropy linked to a given sample for a specific class and, as such, implements Bayesian classification. Its practical implementation highlights the efficient utilization of non-parametric entropy estimation, leveraging the k-nearest neighbors approach.

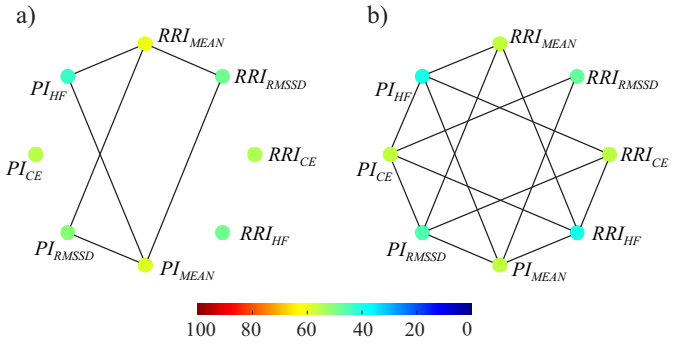


Figure 7.10: Graph representation of mean accuracy across individual features with the a) LIC and b) GB model. Connections between features represent statistically significant differences in cross-validation results, determined with the pairwise t-test with significance $\alpha = 0.05$, with an additional Bonferroni correction for 28 total comparisons within a model. This figure has been adapted from [130].

The efficacy of the proposed model was first tested on a real-world dataset, focusing on a physiological stress classification task with three distinct classes, evaluated on a single-feature classification basis. While the models showed similar behavior in terms of accuracy, an overall difference is present when comparing SE and SP scores, with the LIC being overall better in SP for HUT and SE for REST and MA. Across individual features, we find that the $MEAN$ feature, representing an overall influence of the parasympathetic and sympathetic control on the heart rate [241, 254], and the CE, representing the complexity change with a phase change [68, 175], show better overall performances.

Moreover, both classifiers encountered challenges in predicting the mental-arithmetic class, as evidenced by lower sensitivity scores. This finding underscores the inherent complexity associated with modeling this physiological state, which would be probably better tackled considering additional features focusing on the interaction between different variability series in the spirit of Network Physiology [100].

While the LIC model displays promising performances compared to GB, it is essential to acknowledge that the experimental setup didn't include an exhaustive hyperparameter search aimed at optimizing model performance for both cases nor investigated a more complex system setup, along with the comparison to state-of-the-art models while also investigating the use of multi-feature inputs. Even so, in the real-data application, it can distinguish to a certain degree between the three classes while having comparable results to previous research [95]. Most importantly, the proposed methodology opens possibilities for facilitating a more comprehensive exploration of the underlying data structures by applying an information-theoretic examination of entropy distributions between the classes, potentially revealing novel insights, and by utilizing the k-nearest neighbor search, remaining a comparable lightweight classifier.

7.2.5 Multi-Feature Classification of Physiological Stress in Cardiovascular and Cardiorespiratory Interactions

Introduction. This study [206] investigates the classification of physiological stress through cardiovascular and cardiorespiratory interactions using the information-theoretic framework introduced in Chapter 6.2 and Section 7.2.4, where the *Local Information Classifier* is described. The LIC, together with other local-information indices, provides a principled way to characterise both variability within each physiological signal and directed interactions across them.

Physiological stress induces distinct modulations in autonomic control mechanisms, influencing key variables such as heart period, systolic arterial pressure (SAP), and respiration. These effects propagate through the baroreflex and respiratory sinus arrhythmia pathways, altering both cardiovascular and cardiorespiratory variability. Because these modulations can differ subtly across stress types, accurate discrimination requires classification methods capable of capturing complex univariate and bivariate physiological relations.

Building on this premise, the study implements a multi-feature classification framework based on local information theory. The approach incorporates multiple indices describing both internal variability of each signal and their directed interactions, with the objective of distinguishing postural (orthostatic) and mental stress from rest. By relying on local estimates of joint and conditional entropies, the method explicitly leverages information-theoretic descriptions of physiological mechanisms to enhance discriminability across conditions.

Materials and Methods. All analyses were performed on the **D1** dataset described in the previous section (Sec. 7.1, including simultaneously acquired ECG, arterial pressure (AP), and respiratory (RESP) signals recorded during REST, head-up tilt (HUT), and mental arithmetic (MA). For each subject and condition, stationary segments of RRI, SAP, and RESP time series were extracted using the preprocessing pipeline already detailed earlier in the thesis.

A set of thirteen features was computed from these series, including time-domain descriptors (MEAN, STD), frequency-domain indices in the LF and HF bands, and information-theoretic measures assessing cardiovascular and cardiorespiratory interactions (transfer entropy SAP $\rightarrow$ RRI and RESP $\rightarrow$ RRI, conditional self-entropy, mutual information rate). Their definitions and computation procedures follow the methods introduced in the methodological chapters (Ch. 2 and 3).

Classification was performed using the Local Information Classifier (LIC) [130] 6.2, which assigns each observation to the class maximising the locally estimated informa-

tion content. The classifier was applied separately to cardiovascular-based features and cardiorespiratory-based features. A 10-fold cross-validation scheme was used, and performance was assessed in terms of overall accuracy and class-wise sensitivity and precision.

Results and Discussion. Table 7.6 reports the results in terms of the performance metrics of multi-feature classifications, i.e. accuracy, sensitivity and precision in the three experimental phases, computed separately for cardiovascular and cardiorespiratory variability using the LIC. As reported, the best overall accuracy is achieved for cardiorespiratory analysis, while the class with the best performance in terms of sensitivity and precision is the HUT for both cardiovascular and cardiorespiratory analysis. As shown by the confusion matrices (Fig. 7.11, the HUT class consistently exhibits the best ranking, achieving higher True Positive values if compared to the other classes.

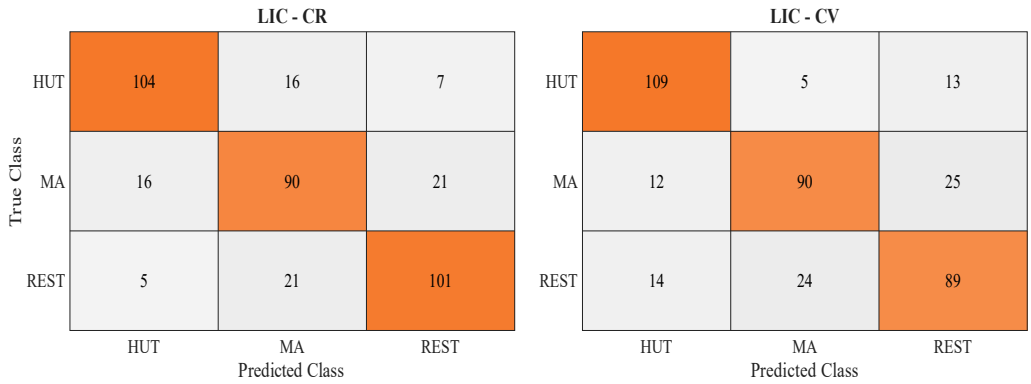


Figure 7.11: Confusion matrices for cardiovascular and cardiorespiratory variability in REST, HUT, and MA using the LIC. This figure has been adapted from [206].

Accuracy (%)	**Sensitivity (%)**			**Precision (%)**		
	REST	*HUT*	*MA*	*REST*	*HUT*	*MA*
CV	75.6	70.1	85.8	70.9	80.7	75.6
CR	77.4	79.5	81.9	70.8	83.2	70.9

Table 7.6: Classification performance (accuracy, sensitivity, and precision; bottom) for cardiovascular and cardiorespiratory variability in REST, HUT, and MA using the LIC. This table has been adapted from [206].

Physiologically, although the analyzed time series are the same across the three experimental conditions, the patterns of physiological responses prove to be more easily identifiable during postural stress than during mental stress, thus facilitating the classification process. These results are in accordance with previous findings reported in the literature [182] that exhibited a better discrimination of postural stress using traditional statistical analyses. Physiologically this may be due to the stronger autonomic nervous system response during postural stress [182]. More specifically, during the HUT phase

changes in AP activate the baroreflex mechanism to modulate heart rate, producing simplified cardiovascular dynamic [182]. This mechanism could generate more pronounced patterns of variability than those associated with a MA task. Moreover, for CR interactions, the reduced involvement of the RSA mechanism during postural stress allows a better classification of this type of stress compared to cognitive load [182].

Conclusions. the present activity highlighted the feasibility of multi-feature classification to obtain discrimination of physiological stress based on cardiovascular and CR variability indices. Our results document a more accurate classification of the HUT phase probably due to the higher sensitivity of the LIC classifier to physiological patterns and mechanisms that are more peculiar of postural stress. The accurate detection of different stress states could be crucial in clinical settings for the early diagnosis of cardiovascular and cardiorespiratory diseases.

7.2.6 Assessing Feature Importance in Cardiovascular Variability for the Classification of Physiological Stress

Introduction. This study [99] examines the contribution of individual cardiovascular variability features to the discrimination of physiological stress conditions. While previous analyses focused primarily on classification performance, this work shifts the attention toward interpretability, aiming to identify which HRV and PRV indices provide the strongest discriminative power across resting, postural, and mental stress. The analysis is carried out within a unified framework that evaluates feature importance through multiple complementary strategies, including model-based assessments and information-theoretic relevance measures. The objective is to quantify the specific role of each variability descriptor and to determine the degree to which feature importance remains stable across different classification paradigms and signal modalities.

Materials and Methods. All analyses were conducted on the **D1** dataset described in Sec. 7.1, which includes simultaneous ECG and arterial blood pressure recordings acquired during REST, head-up tilt (HUT), and mental arithmetic (MA). From these recordings, short-term time series were extracted for four physiological signals: RR from ECG, PP from PPG and SDB and DBP from AP. All preprocessing steps follow the procedures already detailed in the methodological section of this thesis (Chapter 2).

For each signal and condition, a standard set of cardiovascular variability features was computed across time, frequency, and information domains. Since these indices have been fully introduced earlier in the thesis (Chapter 3), only their names are reported here and no computational details are repeated.

Feature importance (FI) and feature selection (FS) were performed using the CMIHiFi

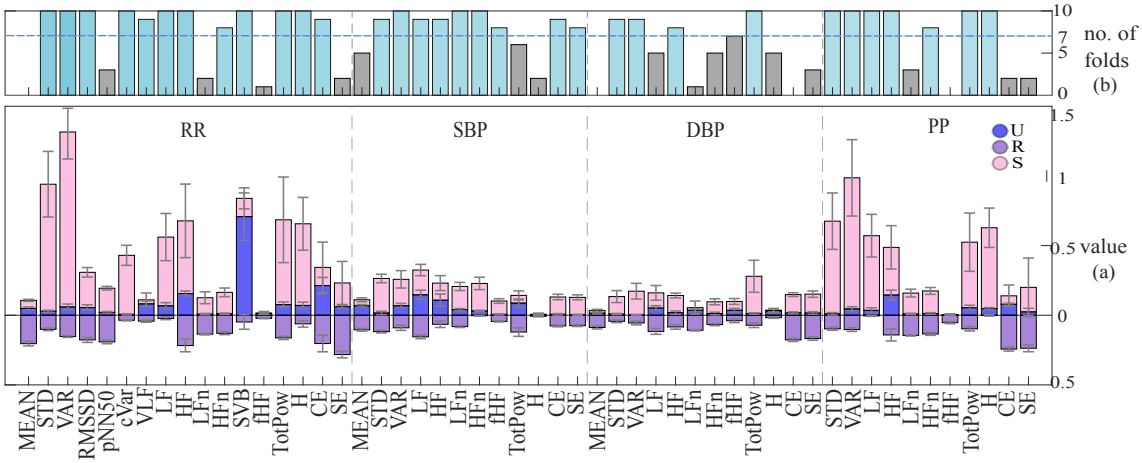


Figure 7.12: Feature Importance (FI) and Feature Selection (FS) results. (a) The bar plot shows the average values (column) and standard deviation (error bar) of the unique information (U), redundancy (R), and synergy (S) for each cardiovascular feature across the 10 folds. R values are reported using a secondary axis to highlight their role in the selection criterion, specifically against the sum of U and S. (b) The number of folds (out of 10) in which each feature was selected is displayed based on the selection criterion ($I = U + S - R > 0$). In light blue, features selected in more than 7 folds out of 10. This figure has been adapted from [99].

algorithm (Sec. 5.2.1), an information-theoretic framework that decomposes the relevance of each feature into unique and shared contributions with respect to the target classes. HiFi was applied separately to the HRV, PRV, SBP, and DBP feature sets to evaluate how the discriminative role of each feature changes across modalities and stress conditions.

The resulting feature-importance profiles and selected feature subsets were subsequently used within classical classification models to assess their ability to discriminate REST, HUT, and MA. All analyses employed cross-validation procedures consistent with the classification pipeline adopted throughout this thesis.

Results and Discussion. Figure 7.21 summarises the results of the FI and FS analyses. Panel (a) presents the average (bar column) and standard deviation (error bar) of the unique information, redundancy, synergy, and total CMI for each cardiovascular feature across the 10-fold cross-validation. Panel (b) illustrates the number of folds (out of 10) in which each feature was selected based on the selection criteria. These results provide insights into the contribution and stability of each feature in stress classification across different folds.

Among the analysed cardiovascular features, the SVB computed from RR time series consistently exhibited the highest unique information across folds, indicating its strong individual contribution to the classification of stress conditions. This suggests that RR alone provides relevant and non-redundant information about the output variable. Interestingly, SBP and DBP presented a balanced profile, with moderate levels of unique and

synergistic information, suggesting that their features contribute both individually and in interaction with other features. PP, while exhibiting relatively low unique information, displayed the highest synergy among all features (e.g., VAR), indicating that its relevance emerges predominantly through joint interactions with the rest of the feature set. This pattern suggests that PP features may act as modulatory or complementary variables in the presence of others rather than providing standalone predictive power. Overall, these results highlight heterogeneous contributions among cardiovascular features, with RR playing a dominant and independent role and PP contributing primarily in a synergistic fashion.

The highly unique information observed in entropy-based features (e.g., SE, CE for both RR and PP) and frequency components (e.g., HF, LF) can be physiologically explained by their strong link to ANS control. In this context, SE and CE provide complementary insights into the complexity and regularity of heart rate dynamics, which are known to vary under different stress conditions [107]. Similarly, frequency domain indices reflect both sympathetic and parasympathetic modulation, which changes significantly between rest and stress [144].

Redundant information is higher among features that capture similar physiological mechanisms. High redundancy is often seen in time- and information-domain features that co-vary due to their shared physiological underpinnings [188].

Features showing high synergy, such as VAR and some entropy indices, likely capture interaction effects between different regulatory components (e.g., sympathetic modulation combined with non-linear heart rate dynamics) [208]. Synergistic information highlights combinations of features that together reveal patterns not evident from individual variables alone. This suggests that complex physiological responses to stress are better captured when considering the joint behaviour of features.

Based on the adopted selection criterion, several features were consistently selected across the 10 folds of cross-validation, indicating a robust and stable contribution to the classification task. The proposed FS method yields approximately a 20% reduction in the number of features, highlighting the physiological relevance of the retained variables. RR-derived features STD, VAR, LF, and CE are selected with high frequency, while features such as pNN50, fHF, SE, and MEAN are rarely or never selected. This suggests that the most consistently selected features capture dynamic aspects of heart rate variability, such as temporal variability (e.g., pNN50), low-frequency oscillations (e.g., VLF, LF) [212], and signal complexity (e.g., CE) [107]. These results indicate that stress-related autonomic regulation is more effectively characterised by fluctuating and irregular patterns in cardiovascular signals rather than by static or aggregate measures (e.g., MEAN).

A similar trend is observed for the SBP derived features. STD, VAR, LF, HF, and SE

	TP	FP	FN	TN	Acc	Pre	Spe	F1
REST	334	68	47	186	81.9	83.1	87.7	85.3
HUT	103	28	24	480	91.9	78.6	81.1	79.9
MA	73	29	54	479	86.9	71.6	57.5	63.7
Tot					80.3	77.8	75.4	76.3

Table 7.7: Confusion Matrix and Classification performance, expressed in terms of accuracy (Acc), specificity (Spe), precision (Pre), and the F1-score for each condition (REST, HUT, and MA) and overall (values in percentage). This table has been taken from [99].

are selected with high frequency, while features such as MEAN, H, and TotPow are rarely or never selected, indicating that variability and frequency-domain information are more informative than total power or high-frequency components in the SBP signal.

Within the DBP features, a broad set of features, MEAN, LF, HFn, fHF, H, CE, and SE, are rarely or never selected across folds. This reinforces the relevance of both spectral and entropy-based descriptors in diastolic pressure, underlining the complex physiological modulation of this signal during stress.

PP-derived features also show high consistency in selection, particularly LF, HF, STD, VAR, TotPow and H, again emphasising the role of both time and frequency content and signal variability. In contrast, the fHF and some high-level descriptors such as CE and SE appear to be less informative.

This pattern supports the idea that stress-related changes in autonomic function are better captured through dynamic, oscillatory, and irregular signal components. For instance, the consistent selection of LF and HF features across DBP and PP points to the role of relative power in distinguishing sympathetic/parasympathetic balance. Similarly, the consistent selection of SE and CE suggests that complexity-related descriptors are essential for capturing subtle modulations in cardiovascular control [232], particularly during mental and postural stress.

Overall, the most robustly selected features reflect variability (e.g., STD, VAR, H) and frequency components rather than simple mean or total power. This reinforces the hypothesis that dynamic aspects of cardiovascular signals, such as oscillatory behaviour and irregularity, carry more discriminative information for stress classification than static features. The results highlight not only the direct influence of individual features but also the complex interactions between physiological markers, offering a more interpretable assessment of ANS dynamics.

The results of the classification performance further validate these findings and are summarised in Table 7.7. The classification performance is reported in terms of accuracy (Acc), precision (PRE), specificity (Spec), and the F1-score for each condition (REST,

HUT, and MA). The results highlight the strong performance of the classification model in distinguishing between stress and non-stress states across different physiological conditions, demonstrating the importance of dynamic cardiovascular features for accurate stress detection. For the REST condition, the model exhibited good accuracy, successfully distinguishing between stress and non-stress states, as found in [97]. In the HUT condition, where participants underwent a head-up tilt test, the model performed even better, especially in terms of specificity, suggesting its ability to identify non-stress states correctly. However, in the MA condition, which involved mental arithmetic, the performance was lower, particularly in specificity, confirming that cognitive load posed a greater challenge for accurate stress detection [97].

Taken together, these results highlight the direct influence of individual features and the complex interactions between physiological markers, offering a more interpretable assessment of ANS dynamics. The selection of features emphasising variability, spectral content, and entropy, combined with the classification results, reinforces the importance of dynamic cardiovascular regulation metrics for stress classification. The classification performance across the different conditions supports the utility of advanced feature decomposition methods to isolate informative components in physiological data and suggests that our approach offers a reliable and context-sensitive way to assess stress based on ANS activity.

Conclusions. This application presented a comprehensive framework that integrates information decomposition and classification to identify stress-related physiological markers. The proposed method combines CMI analysis with a principled selection criterion to isolate features conveying unique and synergistic information while discarding redundancy. The results demonstrate that cardiovascular features reflecting variability, spectral content, and entropy, particularly those derived from RR, DBP, and PP signals, are more consistently informative than static measures like mean or total power values.

Classification outcomes confirm the effectiveness of the selected features, especially in distinguishing REST from orthostatic conditions. Although MA stress detection was more challenging, overall performance supports the validity of the decomposition strategy. Features with high unique information, such as RR-based entropy and spectral markers, contributed independently, whereas redundancy among features like CE and SE reflected shared physiological content. The consistency of selected features across folds highlights the relevance of dynamic, non-linear traits, such as variability, complexity, and oscillatory patterns, as key indicators of autonomic regulation under stress.

These findings emphasise not only the informativeness of individual features but also the importance of capturing complex interactions among physiological signals. By leveraging

information-theoretic tools, the proposed approach enables a richer and more interpretable assessment of ANS dynamics, enhancing both detection and understanding of stress responses. The framework also offers a systematic path for Hi-Fi quantification, improving feature interpretability and revealing hidden dependencies in physiological data.

7.2.7 A Signal Normalization Approach for Robust Driving Stress Assessment Using Multi-Domain Physiological Data

Introduction. This application [74] investigates the use of an inter-subject normalization framework for robust assessment of driving-induced stress using multi-domain physiological signals. The methodological foundation of this work has been described in the previous chapters (see Chapters 3 and 6). Here, we focus on its experimental validation on a publicly available multimodal driving-stress dataset. The assessment of stress in naturalistic settings is strongly affected by inter-individual variability in physiological responses, which limits model generalization across subjects. To address this issue, the proposed approach applies a subject-specific normalization directly at the signal level, aligning ECG and respiratory dynamics to individual resting rates and thus enhancing the discriminability of features extracted for stress classification.

Material and Methods. The data used in this study belongs to the database of J.A. Healey and colleagues [90], composed of multimodal synchronized physiological data and video recorded on 16 subjects in different driving conditions.

Three different stress states were elicited during the experimental session, each one associated with a specific driving condition: low stress (resting, no driving), high stress (city driving) and medium stress (highway driving). The low-stress state was collected on two 15 minutes sessions respectively at the beginning and at the end of the experiment (herein referred as resting1 and resting2). In this condition, the subjects were sitting in a garage with their eyes closed, while the car was parked and idle. The medium-stress state was monitored while driving on the highway in two sessions (highway1 and highway2). The high stress was induced by driving in the congested streets of Boston (three driving phases named city1, city2 and city3). Depending on the traffic, acquisitions lasted from 50 to 90 min including the resting state.

The signals taken into account in this study were the ECG waveform, acquired through lead II configuration (sampling frequency of 496 Hz), and the respiration signal recorded through an elastic Hall effect sensor monitoring the chest cavity expansion (acquired at 31 Hz). Only subjects who had both ECG and respiratory data available for each stress state were considered, thereby excluding six subjects from the analysis.

The data preprocessing procedure prepared the raw physiological data for the successive

feature extraction step. Following previous works [73], the preprocessing in this study consisted in filtering the raw physiological data and then partitioning the data based on the labelled stress state.

The raw ECG signal was filtered using a zero-phase passband Butterworth filter (0.1–115 Hz cutoff frequencies). For the raw respiratory signal, its relative mean value was first removed, and then the signal was filtered with a zero-phase bandpass Butterworth filter (0.01–5 Hz cutoff frequencies). Both the ECG and the respiratory signals were then resampled to 250 Hz, ensuring high data quality and reducing computational processing. The filtered signals were subsequently organized into a structure containing the three primary stress phases based on the reported labels: resting, highway driving, and city driving.

Subsequently, each signal was segmented into 20-second windows with a 5-second overlap, thus performing a data augmentation, as this increased the number of samples available for classification. The choice of a 20-second window length was made to maximize the number of windows while preserving temporal information within each window [81, 218, 223]. The 5-second overlap was chosen to ensure independence between samples, which, in turn, enhances the model's ability to generalize during stress state prediction. According to the study by Farias da Silva and colleagues [69], a 5-second overlap was found to be the best compromise for the overlapping window data augmentation technique.

In the following analyses, two different sets of data originating from the same database were utilized to emphasize the effects of the inter-subject normalization procedure. The first dataset comprised raw signals subjected to the preprocessing procedure without any additional modifications, as described in this paragraph. The second dataset was generated by applying the inter-subject normalization process before the feature extraction step. Specifically, for each driver, the resting phases were considered for subject normalization. From now on, the term *original signal* will be used to identify the first dataset, while the second dataset will be identified by the term *subject-normalized signal*.

This application employs an inter-subject normalization algorithm to account for inherent physiological variability among individuals during stress classification. During resting states, characterized by low stress, each driver exhibits unique physiological characteristics. For instance, a healthy individual's resting heart rate typically ranges from 60 to 100 beats per minute (bpm), while their respiratory rate can vary from 12 to 20 breaths per minute [1, 11]. This physiological diversity persists across all stress levels and can significantly impact classification accuracy.

Consider an example: a heart rate of 80 bpm might be recorded for one subject

during rest, but the same heart rate could indicate moderate stress in another subject. This observation extends to other physiological conditions and stress levels. Without addressing this inter-subject variability, classification algorithms trained and tested with inconsistent data can lead to misclassification errors.

To mitigate this issue, this application presents a novel resampling procedure applied to raw ECG and respiratory signals. This approach extends previous work by Gasparini and colleagues [75], offering a multidomain solution. The primary goal of this procedure is to reduce inter-subject variability in specific physiological features during the resting state. This reduction improves the ability to first identify changes in these features across different stress levels and, indirectly, extends the normalization to all other features extracted from the signals, as the signals themselves are transformed into a new normalized domain.

The normalization procedure involves selecting a specific feature from each physiological signal (ECG and respiration) to serve as the basis for normalization. Subsequently, a new sampling frequency is assigned to the original signal. This adjustment ensures consistency of the selected feature among all subjects within the resting phase. The inter-subject normalization procedure for both ECG and breath signals is further detailed in the following subsections.

The inter-subject normalization procedure on the ECG signal was conducted considering the heart rate as the normalized feature. This ensured that, after normalization, all subjects exhibited the same average heart rate during the resting state.

Starting from the raw ECG signal with a sampling frequency f_c equal to 250 Hz, the Pan–Tompkins algorithm [167] was applied to detect the R peaks during the two resting phases. Subsequently, the R–R interval (RRI) time series were extracted (Figure 1a), from which the heart rate series was derived as the inverse of the RRI time series. The mean heart rate f_h across the two resting phases was then calculated as a reference value using the equation reported in the paper.

The inter-subject normalization procedure aimed to transform all subjects' data into a subject-normalized domain in which each subject's resting heart rate was standardized to a predefined value f_{sh} (70 bpm). This was achieved by resampling the raw ECG signals at a new sampling frequency $f_{s,ch}$ determined as described in the paper. This frequency was then consistently applied to the ECG signals recorded during all other stress phases, preserving temporal relationships across stress levels.

An analogous procedure was applied to the respiratory signal, using the resting breathing frequency as the reference feature. Inspiration peaks were detected (Figure 1c), breath-to-breath intervals were computed, and the mean breath rate f_s was obtained. The new sampling frequency $f_{s,cs}$ was then computed following the equation provided in the

paper and used to resample the respiratory signal across all stress phases.

As with ECG, resampling generated different signal lengths depending on whether the subject's resting breath rate was above or below the chosen reference value (14 breaths/min).

Feature extraction was performed on 20-second windows. Only time-domain features were used, due to the short window length. For ECG, the extracted features included MEAN, STD, RMSSD, SDSD, NN50, and pNN50. For respiration, features included respiratory rate, inspiration and expiration durations, and corresponding areas.

A total of 12 features were extracted for each 20-second window. The original dataset yielded 2979 windows; the normalized dataset yielded 2974 windows. Class imbalance was addressed using the Synthetic Minority Oversampling Technique (SMOTE) [46].

A Gaussian SVM classifier was trained on 80% of the data and tested on the remaining 20%. This procedure was repeated 100 times to reduce variability due to random partitioning. Confusion matrices and performance metrics (accuracy, sensitivity, specificity, precision, and F-measure) were computed for each iteration.

The effects of inter-subject normalization were compared against standardization and min–max scaling. Normality of performance distributions was assessed via chi-square tests. ANOVA and Bonferroni-corrected Student's t-tests were used to evaluate significant differences between normalization procedures.

Table 7.8: Heart rate derived from the original ECG signals and associated resampling frequencies in the subject-normalized domain. This table has been taken from [74].

Subject (n)	**Heart rate (bpm)**	**Resampling frequency (Hz)**
1	67.3	260.0
2	66.2	264.4
3	80.7	216.9
4	71.9	243.4
5	61.3	285.5
6	76.6	228.5
7	61.4	285.0
8	60.7	288.3
9	83.7	209.1
10	62.4	280.4

Results. Following normalization, ECG and respiratory signals were resampled to the new subject-normalized domain using each subject's resting HR or RR. Subsequent classification using the SVM model was performed across four conditions: original data, subject-normalized data, standardized features, and scaled features. Figure 7.13 shows an example confusion matrix from a single iteration.

Table 7.9: Breath rate derived from the original respiratory signal and associated resampling frequencies in the subject-normalized domain. This table has been taken from [74].

Subject (n)	Breath rate (bpm)	Resampling frequency (Hz)
1	14.3	244.8
2	14.9	234.9
3	16.7	209.6
4	13.7	255.5
5	11.8	296.6
6	11.6	301.7
7	16.2	216.0
8	18.1	193.4
9	8.9	393.3
10	13.9	251.8

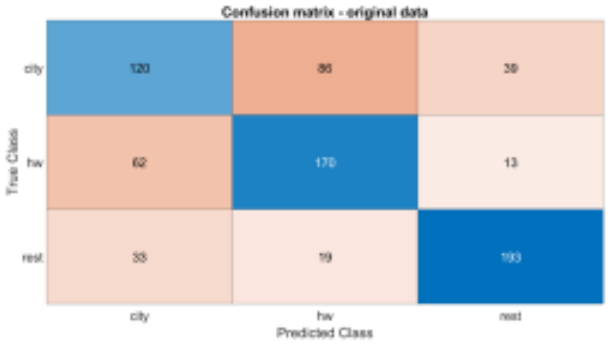


Figure 7.13: Confusion matrix obtained by considering the original testing data in a single iteration. This figure has been taken from [74].

Mean and standard deviation of each performance metric across 100 iterations are reported in Table 7.10. Accuracy results and statistical significance levels are shown in Figure 7.14.

Discussion. This study investigated the impact of inter-subject normalization on electrocardiographic (ECG) and respiratory signals to reduce physiological variability during stress state classification. We selected ECG and respiratory data due to their established relevance in stress research, as evidenced by previous studies [73, 87, 218]. These signals provide valuable insights into an individual's physiological response to stress.

We performed a feature-based multilevel stress classification using a SVM classifier. This classification was conducted after applying inter-subject normalization and two widely used normalization procedures: standardization and scaling. Subsequently, we compared the classification performances derived from these procedures with those obtained from the original, unnormalized data.

When using the original ECG and respiratory signals, our classification model con-

Table 7.10: Performance obtained using features from original signals, subject-normalized signals, standardized features, and scaled features. Superscript "n.s." denotes non-significant differences; "***" indicates $p < 0.001$. This table has been taken from [74].

Metric	Original	Standardization	Scaling	Inter-subject norm.
Precision (%)	68.1 (1.9)	$68.3^{\text{n.s.}}$ (1.9)	$68.2^{\text{n.s.}}$ (2.0)	72.9^{***} (1.9)
Sensitivity (%)	68.5 (1.8)	$68.6^{\text{n.s.}}$ (1.8)	$68.7^{\text{n.s.}}$ (1.9)	73.0^{***} (1.9)
Specificity (%)	84.2 (0.9)	$84.3^{\text{n.s.}}$ (0.9)	$84.3^{\text{n.s.}}$ (0.9)	86.5^{***} (0.9)
Accuracy (%)	68.5 (1.8)	$68.6^{\text{n.s.}}$ (1.8)	$68.7^{\text{n.s.}}$ (1.9)	73.0^{***} (1.9)

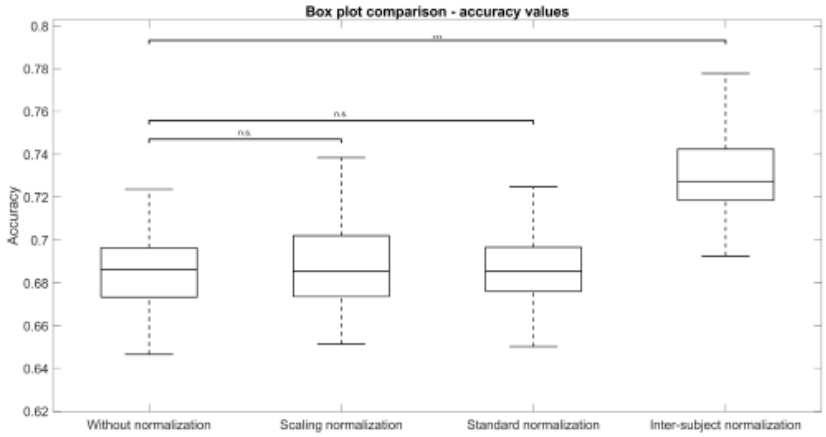


Figure 7.14: Box plot and reported statistical significance obtained through a parametric Student's t-test comparing model accuracy using features derived from original data, scaled features, standardized normalized features, and features obtained from inter-subject normalized signals. This figure has been taken from [74].

sistently achieved a precision, sensitivity, and accuracy of 68%, while specificity was approximately 84%. Applying standardization and scaling techniques yielded comparable performance values (Table 3), suggesting these methods do not significantly alter the model's performance compared to using raw data. This outcome aligns with expectations, as both standardization and scaling primarily adjust data distribution without introducing substantial changes to the underlying information.

The inter-subject normalization procedure successfully reduced inter-subject variability in physiological features during the resting state, effectively aligning subjects to a common physiological domain. This normalization enhances the classification of stress levels by ensuring consistency in physiological data across resting and stressful conditions for different subjects. The results clearly demonstrated the effectiveness of this approach, showing significant improvements over the original data, standardization, and scaling procedures. Precision, sensitivity, and accuracy increased from 68% to 73%, and specificity improved by approximately two percentage points.

The inter-subject normalization procedure operates directly on the original signals, placing them in a transformed space defined by a distinct sampling frequency. This novel procedure ensures that each feature derived from the inter-subject normalized signals is already indirectly normalized. Crucially, this normalization structure depends solely on the specific features used in the normalization process (e.g., heart rate for ECG and breath rate for respiratory signals in this study).

Consequently, each extracted feature becomes strongly interconnected with every other feature. In contrast, traditional feature normalization approaches like standardization and scaling independently modify each feature obtained from the original signal. These methods treat each feature in isolation, which can lead to a loss of consistency between features and varying levels of normalization, ultimately impacting the subsequent classification step.

A critical aspect of implementing this technique is identifying the most appropriate feature for defining the inter-subject normalization rescaling. In this project, we distinctly used the mean heart rate and mean breath rate during resting phases for ECG and respiratory signals, respectively, given their distinct physiological meanings.

Other studies have utilized physiological signals for stress assessment with various experimental protocols. For example, Smets and colleagues [90] achieved approximately 82% classification accuracy using an SVM in a binary stress classification problem by combining ECG and respiratory features. Similarly, Han et al. [46] discriminated between three stress states with 84% accuracy using ECG and respiratory features. These studies often incorporated both time and frequency domain features, necessitating longer signal windows for adequate frequency resolution.

In contrast, our proposed work effectively utilizes 20-second segments and the inter-subject normalization procedure to achieve efficient classification with strong performance, even without considering frequency domain features. This suggests the presented framework could potentially be employed for real-time stress classification based on ultra-short-term recordings.

Conclusions. This study investigated the impact of a subject-normalization procedure on stress classification. It utilized ECG and respiratory physiological signals from 10 drivers across varying stress levels, employing a feature-driven methodology. The findings demonstrate that this novel normalization procedure significantly improves stress classification compared to traditional standardization and scaling methods. The inter-subject normalization approach effectively reduced variability between individuals, leading to consistent results in both resting and stressful states. This enhancement was reflected in improved precision, sensitivity, and accuracy, indicating a greater ability to correctly

identify stress. A key advantage of inter-subject normalization is its direct operation on the original signals, fostering strong interconnectivity among features. This differs from other methods that modify features independently. This novel procedure therefore establishes a deeper link between features, as they originate from a shared signal domain.

7.2.8 Evaluation of Work-Related Stress on Healthcare Operators using Cardiovascular Variability Indices from Portable Multisensor System

Introduction. This study aims to investigate the autonomic nervous system response to stress, exploiting cardiovascular variability and signal complexity measures. Focusing on a population of healthcare workers, particularly vulnerable to chronic occupational stress, we evaluate how stress modulation strategies affect autonomic function across different physiological conditions. By combining traditional HRV indices with entropy-based measures rooted in information theory, we focus on capturing both the magnitude and the structural nuances of autonomic regulation under stress, allowing for a dynamic assessment of physiological adaptation over time. By focusing on a high-risk occupational population and leveraging recent advances in wearable technology and cardiovascular signals analysis, this study holds potential implications for improving the early detection of work-related stress responses through portable and wearable devices directly in workplace settings.

Materials and Methods. The analyses were conducted using the **D3** dataset, which includes recordings from 60 healthcare professionals (42 females; mean age: 50.0±3.0 years) working in rotating shifts at the A.R.N.A.S. Ospedali Civico Di Cristina Benfratelli Hospital in Palermo, Italy. It includes cardiovascular and respiratory recordings acquired before and after a six-month wellness program in a cohort of shift-working nurses. The dataset also contains measurements obtained from the MPSAS prototype, previously developed at the University of Palermo, and enables high-resolution, synchronous acquisition of electrocardiographic (ECG), photoplethysmographic (PPG), and respiratory (RESP) signals [233]. Using this system, the PPG signal was acquired on the left middle finger. HRV and PRV indices were computed from ECG- and PPG-derived beat-to-beat intervals in the time, frequency, and information domains, following the same analytical framework described in Chapter 3.

The analyses were aimed at the evaluation of physiological response to the experimental protocol using data from the MPSAS research device. Of the initial cohort, data from 58 participants (42 females) were retained for analysis, after excluding two participants due to poor signal quality.

To prove that MPSAS system-derived indices can be exploited for distinguishing physiological variations affecting autonomic and respiratory regulation due to breathing and

lifestyle interventions, analyses were performed on the full set of parameters extracted from the device. The analysis encompassed two within-subject conditions, spontaneous breathing (REST) and controlled breathing (CB), as well as two experimental phases, i.e. PRE and POST. Participants were stratified into intervention and control groups to evaluate group-specific effects.

For each index, the distribution of values computed on REST and CB were compared using the Wilcoxon signed-rank test, separately for the PRE and POST phases. Longitudinal differences between PRE and POST were assessed for each breathing condition using the same test. To examine differences between the *INT* and *CONT* groups, the Mann–Whitney U test (ranksum) was used for each condition and phase. All statistical tests were two-tailed, with a significance level set at $\alpha = 0.05$; for within-subject comparisons involving multiple indices, a more stringent threshold of $\alpha = 0.005$ was adopted.

This statistical pipeline enabled a comprehensive evaluation of autonomic and respiratory dynamics modulated by controlled breathing and the intervention program, based on all the metrics extracted from the MPSAS dataset.

Results. Results of the cardiovascular variability indices (extracted from RRI or PPI time series) in the three domains, i.e. time (Fig. 7.15), frequency (Fig. 7.16), and information-theoretic (Fig. 7.17), are analysed across the respiratory conditions (*REST*, *CB*), experimental phases (*PRE*, *POST*), and participant groups (*ALL*, *CONT*, *INT*). Statistically significant differences are indicated in the figures as follows: * denotes significant differences between *REST* and *CB* conditions within the same group, # between the *CONT* and *INT* groups, and § between *PRE* and *POST* phases within the same condition.

Figure 7.15 illustrates the distributions of *MEAN*, *SDNN* and *RMSSD* indices computed from RRI and PPI time series. No statistically significant variations were reported for *MEAN* index across experimental phases or respiratory conditions. On the other hand, across all subjects (*ALL*), both *SDNN* and *RMSSD* were significantly higher during *CB* compared to *REST* within the same phase, reflecting enhanced parasympathetic modulation during guided respiration.

Figure 7.16 presents the distributions of *VLF*, *HF*, and *LF/HF*, computed from both *RRI* and *PPI* series across phases, groups, and breathing conditions. Results for the *LF* power were excluded for brevity, as no significant effects were observed. The *VLF* showed a consistent decrease from *REST* to *CB* across all groups and phases, for both *RRI*- and *PPI*-based indices. This reduction reached statistical significance only during the *PRE* phase, particularly in the *ALL* and *CONT* groups. A similar decreasing trend

was observed for *HF*, likely reflecting the concentration of respiratory modulation into a narrower frequency band under paced breathing, which may reduce spectral power within the traditional high-frequency range (0.15–0.4 Hz). The *LF/HF* ratio increased during *CB* compared to *REST*. A statistically significant difference between the *CONT* and *INT* groups was observed at *PRE* phase for the same index during the *CB* condition, solely for the *RRI*-derived measures. Additionally, a significant increase from *PRE* to *POST* was detected in the *INT* group during *REST*, but only for the *PPI*-based indices. Overall, frequency-domain indices reflected coherent respiratory- and intervention-related modulations, while interpretation of *HF* and *LF/HF* must account for the spectral shifts induced by paced respiration.

Figure 7.17 shows the results of information-theoretic indices extracted from the *RRI* and *PPI* series, i.e., H_{lin}, H_{knn}, and CE_{knn}. Results for CE_{lin} were excluded for brevity, as no significant effects were observed. Across all measures, a decreasing trend from *REST* to *CB* can be observed, with this reduction being significant for CE_{knn} at PRE and H_{lin} at POST, indicating higher regularity and lower entropy during guided respiration. No significant differences were found between *PRE* and *POST*, except for H_{lin} of the *RRI* series, which showed a *PRE*-to-*POST* decrease in the *INT* group under *CB*. When comparing the *INT* and *CONT* groups, the only significant group-level difference emerged for H_{knn}, observed under *REST* conditions in both *RRI*- and *PPI*-derived series.

Discussion. Heart rate variability (HRV) provides a powerful, non-invasive marker of autonomic regulation and its dynamic modulation in response to internal and external stressors. In the present study, we investigated short-term autonomic adaptations to controlled breathing and an experimental intervention targeting stress regulation, using a multimodal acquisition platform in a sample of healthcare professionals exposed to high work-related stress levels, as also demonstrated by the lower values of MCS12 scores if compared to general population. It is widely known that healthcare professionals, and in particular nursing staff, are frequently exposed to high levels of occupational stress due to organisational, emotional, and cognitive demands and a prolonged exposure to such stressors may contribute to autonomic imbalance, thus increasing the risk for cardiovascular and psychological disorders [49, 80].

Our findings confirmed that a biofeedback protocol encompassing a controlled breathing period elicits measurable increases in vagal activity, as reflected by elevated *RMSSD* during guided respiration. This finding aligns with previous literature highlighting the role of slow breathing in activating the parasympathetic system through baroreflex engagement and respiratory sinus arrhythmia [126, 205, 212]. Such autonomic responses are functionally adaptive, promoting restoration and stress recovery. Moreover, the reduction of CE_{knn} confirms previous findings reporting that guided breathing predominantly reduces

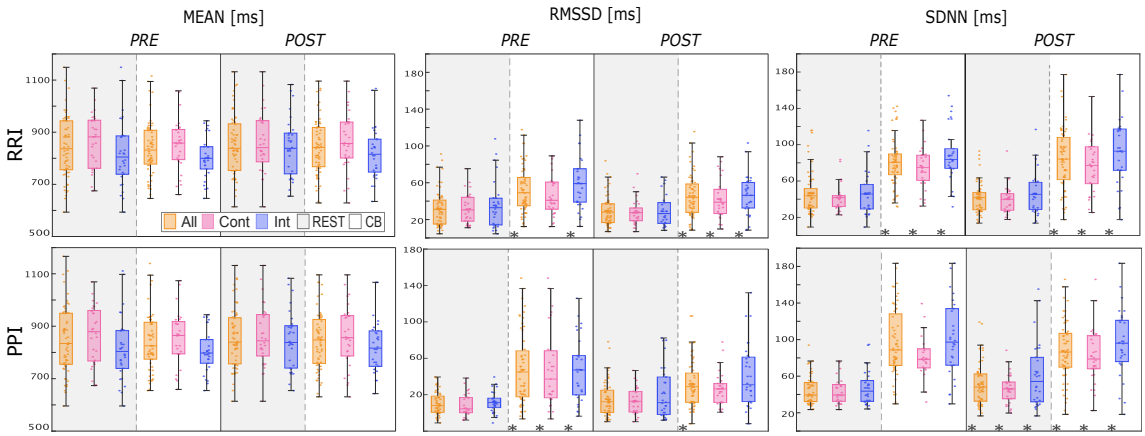


Figure 7.15: Time-domain indices (*MEAN*, *SDNN*, *RMSSD* computed from RRI and PPI series extracted using the *MPSAS* system. Results are reported for three groups: *ALL* (orange), *CONT* (magenta), and *INT* (blue), across both experimental phases (*PRE*, *POST*) and respiratory conditions (*REST*, *CB*). Results of statistical analysis: * denotes significant differences between *REST* and *CB* within the same phase and group according to Wilcoxon signed-rank test, # indicates differences between the *CONT* and *INT* groups under the same phase and condition according to Mann-Whitney U test, and § highlights differences between *PRE* and *POST* within the same group and condition according to Wilcoxon signed-rank test. Bonferroni correction for multiple comparisons was used, with a conservative adjusted threshold of $\alpha = 0.005$ applied to detect significant differences.

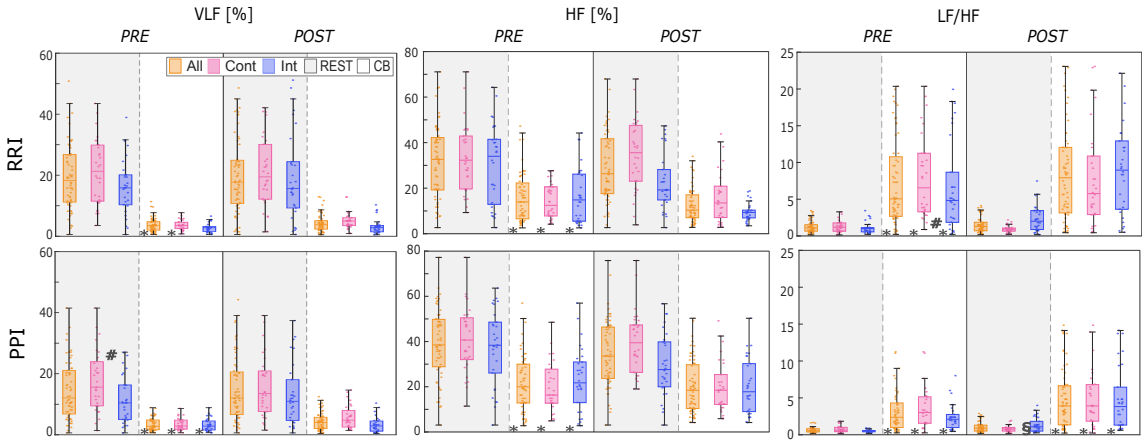


Figure 7.16: Frequency-domain indices (*VLF [%]*, *HF [%]*, *LF/HF*) computed on RRI and PPI series extracted from signals acquired with the *MPSAS* system. Results are reported for three groups: *ALL* (orange), *CONT* (magenta), and *INT* (blue), across both experimental phases (*PRE*, *POST*) and respiratory conditions (*REST*, *CB*). For the information about the statistical analysis please refer to the caption of Figure 7.15.

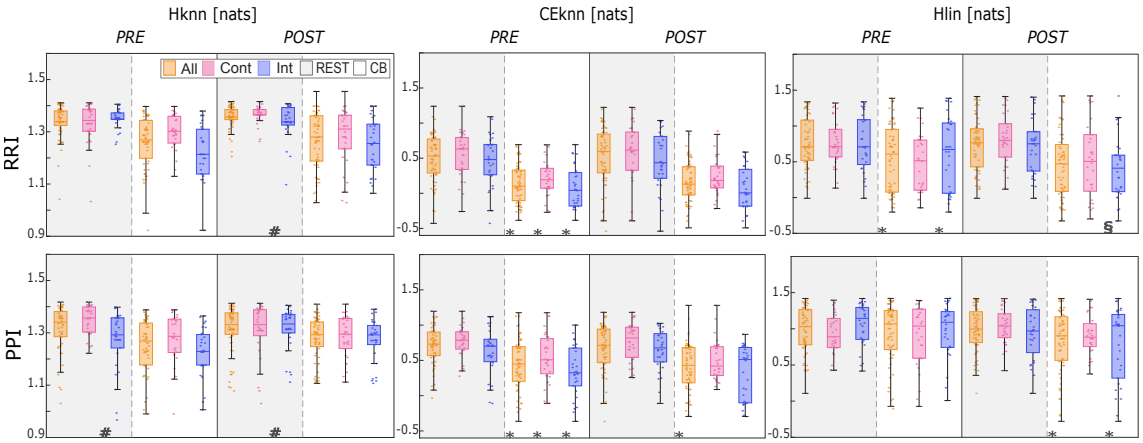


Figure 7.17: Information-theoretic indices (H_{lin}, H_{knn}, CE_{knn}) computed on RRI and PPI series extracted from signals acquired with the *MPSAS* system. Results are reported for three groups: *ALL* (orange), *CONT* (magenta), and *INT* (blue), across both experimental phases (*PRE*, *POST*) and respiratory conditions (*REST*, *CB*). For the information about the statistical analysis please refer to the caption of Figure 7.15.

complexity [16], even if group-related effects are limited to baseline condition. Moreover, in PRE, the higher H_{knn} at *REST* for *INT* group compared to the *CONT* group points to a higher baseline entropy in autonomic regulation.

The reduction of *VLF* during *CB* compared with *REST* reflects the suppression of slow oscillatory components under paced respiration. This reduction likely indicates a withdrawal of non-respiratory autonomic influences, such as baroreflex and thermoregulatory mechanisms, in favour of respiratory-driven oscillations. Particular caution is warranted when interpreting the results of other frequency-domain indices, given the multifactorial determinants of *LF* power and the *LF/HF* ratio, which are influenced by sympathetic modulation, baroreflex dynamics, and breathing patterns [142, 234]. Moreover, in our protocol the paced respiration alters the location of spectral components shifting from *HF* to *LF* band, thus causing a reduction of *HF* normalized power and an increase of *LF/HF* ratio. Therefore, the values of these indices may not reflect pure autonomic variations, but rather the combined influence of respiratory-driven spectral redistribution and intervention-related effects. Moreover, the statistically significant difference in the same index between the *CONT* and *INT* participants during the *PRE* phase, only under the *CB* condition indicates differences in autonomic regulation between groups evident primarily when a paced breathing is elicited, effect that is not anymore observed in the *POST* phase, suggesting an effect of the intervention.

Beyond these respiration-related effects, the intervention group showed enhanced autonomic regulation in terms of sympathovagal activation following the experimental phase, with increases in *LF/HF* and a decrease of H_{lin} from *PRE* to *POST* during spontaneous

breathing (*REST*). The latter finding points to a more irregular signal structure in the intervention group at baseline. The absence of analogous statistically significant differences in the *CONT* group highlights that these effects are specifically induced by the intervention. Moreover, the higher entropy values (H_{knn}) observed in the *INT* group under resting conditions at *POST* (not reported at *PRE* using RRI-based indices) point to a more complex and adaptable autonomic regulation pattern. Increased entropy values in cardiovascular time series have been associated with greater physiological flexibility and resilience against stress-related dysregulation [67, 232]. Taken together, these post-intervention changes suggest an improved ability to regulate autonomic output at baseline, reflecting reduced chronic stress load and greater physiological adaptability. Notably, the control group did not show such changes, supporting the specificity of the intervention effect. It is worth noting that the significant *PRE–POST* increase of *LF/HF ratio* in the *INT* group during *REST* emerged exclusively in the *PPI*-based indices. This finding suggests that the rebalancing of autonomic control induced by the intervention became particularly evident at rest and when assessed through peripheral rather than ECG-derived measures. Such an effect aligns with the notion that peripheral indices can be more sensitive to subtle shifts in autonomic modulation, especially in contexts where the intervention targets global physiological regulation.

Overall, our results demonstrate the usefulness of a multidomain assessment approach for characterising physiological adaptations, and converge to show that paced breathing acts as a strong modulator of vagal activity across all participants, while the intervention appears to selectively enhance autonomic function and signal complexity at rest. Taken together, these results support the usefulness of exploiting multimodal biomedical devices for monitoring of physiological stress, offering the opportunity to track complex physiological inter-system dynamics, particularly in high-stress professions such as healthcare operators.

Conclusions. This study confirmed the proper working of a custom-built multimodal system [233] able to acquire various physiological signals synchronously and its usefulness as a reliable, low-cost device for stress assessment in operational environments. Our results pave the way for the feasibility of using a multisensor portable device for computing short-term cardiovascular variability indices, supporting its use for monitoring autonomic activity on a population of healthcare workers vulnerable to chronic occupational stress.

The results of the analysis revealed distinct patterns of autonomic modulation across respiratory conditions, experimental phases, and participant groups. Time-, frequency-, and information-domain features derived from RRI and PPI series captured the physiological effects of the paced breathing and of the intervention, with group-level comparisons shedding light on both shared and specific adaptations.

7.2.9 Partial Information Decomposition for Mixed Discrete and Continuous Random Variables

Introduction. This application [17] applies the Partial Information Decomposition (PID) framework to investigate the redistribution of information shared between cardiovascular and respiratory dynamics under postural stress. By decomposing mutual information into unique, redundant, and synergistic components, the analysis aims to disentangle the individual and joint contributions of cardiac and vascular signals in encoding respiratory phases during resting and orthostatic conditions.

Materials and Methods. Analyses were performed on a reduced subset of the **D1** dataset (Sec. 7.1), consisting of synchronous ECG, ABP, and RESP recordings collected from healthy young subjects during resting and postural stress conditions. From each acquisition, short stationary segments of heart period (H), systolic arterial pressure (S), and respiration (R) time series were extracted for the analysis of cardio–respiratory coupling through the PID framework.

To investigate how cardiovascular dynamics encode respiratory phases, the respiratory signal was discretized into a binary variable $\bar{Y}$ representing inspiration ($\bar{R}_n = 0$) and expiration ($\bar{R}_n = 1$). Two-lag dynamic patterns were used as continuous source variables: $X_1 = [H_n, H_{n+1}]$ for heart period and $X_2 = [S_n, S_{n+1}]$ for systolic pressure. Mutual information (MI) between $\bar{Y}$ and the joint sources (X_1, X_2) was decomposed into PID components—unique (U), redundant (R), and synergistic (S). Statistical significance was assessed using surrogate data tests and paired Student's t-tests ($\alpha = 0.05$) comparing resting and stress conditions.

Results and Discussion.

Figure 7.18 shows the distribution of MI and PID terms computed between cardiovascular and respiratory phase variables in both REST and STRESS conditions. The statistical significance of each PID term is assessed through a surrogate data approach and the number of subject for which each measure is statistically significant reported below the respective distribution. Moreover, a Student's t-test was applied to assess the statistical significance differences between REST and STRESS phases for each measure ($\alpha = 0.05$). The comparison of the two conditions indicates a weakening of the interaction between the cardiovascular and respiratory systems with postural stress, documented by the statistically significant decrease of I(Y; X1,X2). The application of the PID reveals physiologically meaningful modulations induced by the postural change, such as the dampening of the mechanism of respiratory sinus arrhythmia [59] reflected by the drop of U(Y; X1), and the enhancement of the cyclic intrathoracic pressure changes and the baroreflex mechanism related to ventilation [58] reflected by the rise of U(Y; X2). Even if no variations are

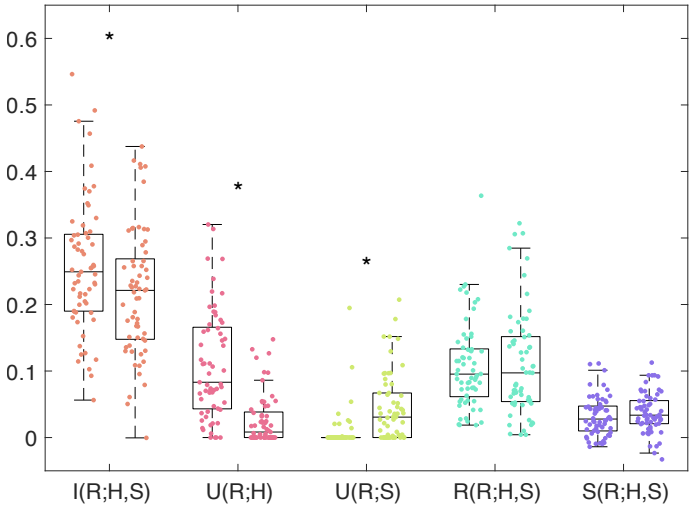


Figure 7.18: Boxplot distributions and individual values of the PID applied to the MI between the binary target variable Y measuring inspiration/expiration and the continuous source variables X1 and X2 measuring respectively heart period and systolic arterial pressure, computed in conditions of resting state and postural stress. *, $p < 0.05$ paired Student's t-test, REST vs. STRESS. The number of subjects for which each measure is statistically significant according to surrogate data analysis is shown below each boxplot. This figure has been adapted from [17].

obtained for the redundant and synergistic terms, the prevailing redundant contribution in both conditions is likely related to the influence of ventilatory activity on both cardiac and vascular dynamics with similar latency [58, 59].

7.3 Sex-related Differences in Cardiovascular Regulation

This section examines how the analytical framework developed in the methodological sections of this thesis can be applied to the study and classification of sex-related differences in cardiovascular autonomic regulation. Rather than revisiting concepts already introduced, such as heart rate and blood pressure variability, information-theoretic measures, or feature selection strategies, the focus here is on illustrating how these tools are leveraged to uncover sex-specific physiological patterns across different datasets, methodological paradigms, and analytical perspectives.

This section examines how the analytical framework developed in the methodological chapters of this thesis can be applied to the study of sex-related differences in cardiovascular autonomic regulation. The primary objective is not to determine the sex of a subject, but rather to use classification and information-based analyses as interpretative tools to investigate how sex influences the structure and relevance of cardiovascular variability descriptors. In this context, feature selection and feature importance play a central role, as

they allow one to identify which physiological indices, or combinations thereof, are most affected by sex-related regulatory mechanisms and how their contributions are organised across cardiac and vascular domains.

Sex is a major determinant of cardiovascular function, influencing autonomic balance, vascular properties, and the mechanisms underlying short-term cardiovascular regulation. The analysis of short-term cardiovascular control exhibits different characteristics based on sex, and it is crucial to understand the differences in how women and men regulate their cardiovascular systems, as these differences significantly impact the risk, manifestation, and treatment outcomes of cardiovascular diseases [77, 119, 164, 179, 198]. Studies of two key indicators of autonomic cardiovascular control, heart rate variability (HRV) and blood pressure variability (BPV), have produced mixed and sometimes contradictory findings. While some analyses show that females tend to have higher high-frequency power and a lower low-frequency/high-frequency ratio than males, indicating greater parasympathetic tone [119, 164], other research reveals inconsistent or even opposite trends, particularly regarding BPV, as denoted in [129, 164, 198, 224]. This ambiguity highlights the complex interplay of biological, hormonal, and physiological factors influencing cardiovascular regulation, making it difficult to draw definitive conclusions about sex differences in autonomic function [164, 179, 198]. Furthermore, the multidimensional and nonlinear nature of cardiovascular signals requires advanced analytical frameworks to uncover subtle regulatory patterns.

Recent work has highlighted how AI and information-theoretic approaches can support the identification of subtle autonomic differences between sexes, by modelling feature relevance, redundancy, and higher-order dependencies among cardiovascular indices [164]. Classic ML models and mutual-information-based feature selection have been employed to explore sex classification from HRV and BPV indices, with preliminary results suggesting that combinations of features, rather than individual markers, may capture better sex-related autonomic patterns [96]. Building on this, more advanced frameworks such as high-order feature importance decomposition have recently been introduced to disentangle the unique, redundant, and synergistic contributions of cardiovascular features to sex prediction [164].

Across the applications presented in this chapter, emphasis is placed on the results and their physiological interpretation. Each study is briefly contextualised in terms of dataset and overall analytical workflow, while methodological details already described in dedicated sections are omitted. The focus is instead directed toward how different analytical strategies, from conditional mutual-information feature selection to synergy-aware importance decomposition and ML classification, capture sex-specific signatures embedded in heart rate, blood pressure, and related variability indices.

The applications presented in this section are organised into a series of studies, each focusing on a specific dataset and analytical perspective. For each study, results and interpretations are discussed locally, with methodological details kept concise and limited to aspects not previously covered. The first set of analyses addresses sex-related differences using HRV- and BPV-based descriptors under resting conditions, while subsequent studies extend the investigation to multivariate and interaction-aware frameworks incorporating feature selection and high-order importance measures. A comprehensive and integrative discussion that compares findings across datasets and methodologies, and places them in a broader physiological and methodological context, is deferred to Chapter 8.

Collectively, these studies illustrate both the complexity and interpretability of sex-related autonomic regulation. They show how variability features encode distinct contributions across cardiac and vascular domains, how higher-order interactions may reveal differences not observable via pairwise analyses, and how ML models generalise across independent cohorts. By integrating results from multiple approaches, the chapter provides a comprehensive assessment of the feasibility, limitations, and physiological meaning of sex classification based on cardiovascular variability.

7.3.1 Characterization of Sex Differences in Cardiovascular Variability Using Univariate and Bivariate Indices

Introduction. Sex-related differences in cardiovascular autonomic regulation have been widely investigated, yet results in the literature remain heterogeneous. Several studies have reported higher vagal activity in females and stronger sympathetic modulation in males, but findings vary considerably depending on the metric and experimental conditions considered. Traditional HRV and BPV indices provide valuable information on autonomic control, but univariate analyses may fail to capture the interplay between cardiac and vascular mechanisms that contributes to sex-specific responses.

This study [178] explores sex differences using a combination of univariate HRV and BPV indices and bivariate interaction measures quantifying cardiorespiratory and cardiovascular coupling. By comparing multiple domains and integrating complementary information sources, the analysis aims to determine which features, or combination of features, best discriminate males from females. In this perspective, the study serves as a preparatory step for the subsequent classification analyses: rather than applying machine-learning models directly, it provides a structured assessment of which autonomic markers are most informative, thereby establishing the physiological and methodological groundwork required for the development of sex-classification frameworks in the following sections.

Materials and Methods. Analyses were conducted on the **D2** dataset, consisting of spontaneous short-term cardiovascular recordings (5 minutes) from healthy young adults. As detailed in Chapter 2, RR intervals (RR), systolic BP (SBP), diastolic BP (DBP), and respiratory signals were preprocessed and used to derive standard univariate HRV and BPV features in the time, frequency, and information domains (Chapter 3).

In addition, this study computed bivariate indices describing interactions between cardiovascular and respiratory variability, including cardiorespiratory coupling and baroreflex-related measures. Statistical comparisons between indices computed on males and females were performed using nonparametric tests with Bonferroni correction for multiple comparisons.

Results. Conventional HRV time-domain analyses evidenced in females statistically significant lower mean RR intervals (852±117 ms vs 894±116 ms) and higher average SBP values (127±27 mmHg vs 120±24 mmHg). Time-domain results of information-theoretic analyses (Fig. 7.19) indicate that females exhibit increased ER with regard to DBP only. Spectral decomposition allowed to highlight statistically significant lower ER values in females in LF band for RR and SBP time series (Fig. 7.19(a)). From MIR results (Fig. 7.19(b)) we infer that females exhibit a decreased dynamical coupling in LF band for all the pairwise interactions between the cardiovascular time series.

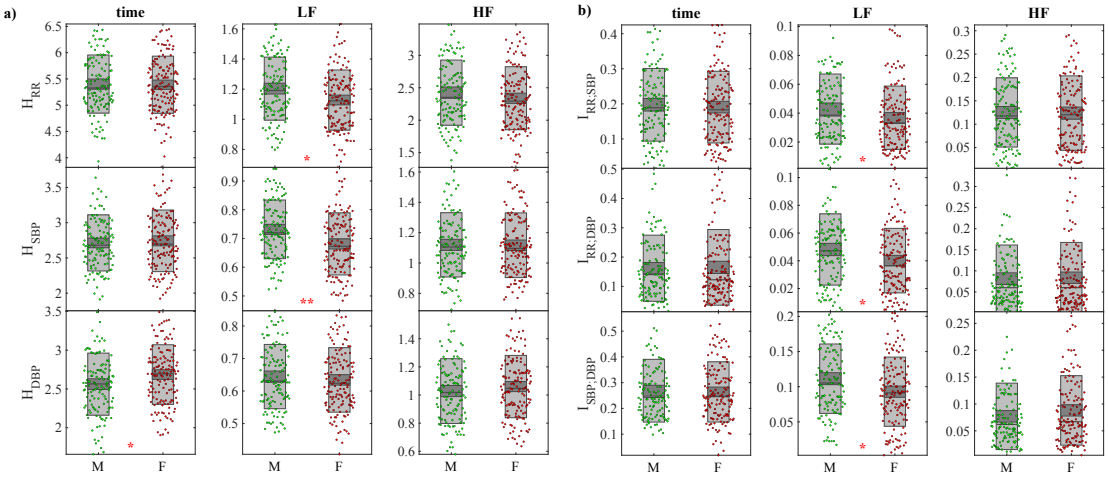


Figure 7.19: Boxplot distributions and individual values of (a) ER and (b) MIR time-domain and spectral measures integrated in LF and HF bands. Statistical test: Student's t-test: *,$p < 0.05$;**,$p < 0.001$, males (M) vs females (F). This figure has been adapted from [178].

Discussion and Conclusions. The results evidence that females exhibit a lower complexity of RR and SBP dynamics within the sympathetic-related band, together with a reduced cardiovascular coupling. The findings explain the lower normalized LF power values reported in females (not shown for brevity), being the spectral ER related to band

power (as detailed in [219]). Overall, these results support the findings in literature that evidence differences in autonomic function associated with blood pressure regulation with lower responsiveness and sympathetic support in young healthy females, eventually resulting in increased predisposition to orthostatic intolerance [200]. Gender-related differences in cardiovascular coupling even during resting supine condition were also evidenced in [200] using a different approach (Joint Symbolic Dynamics). Our results indicate that the lower sympathetic activity in females elicit a decreased cardiovascular coupling occurring mainly within the LF spectral band. Nonetheless, such differences may be related to the young age of the subjects, which represents a limitation of our study. Future activities should complete this work through a spectral analysis and MIR decomposition in its causal components [177].

7.3.2 Conditional Mutual Information-based Feature Selection for Sex Differences Characterization

Introduction. The purpose of this study [96] is to apply a CMI-based feature selection approach, already detailed in Chapter 3, to cardiovascular time series for sex classification, demonstrating that conditioning on previously selected features can improve the identification of informative variables and support physiologically meaningful interpretation.

Materials and Methods. This analysis was conducted using the **D2** dataset, which includes short-term (5-minute) beat-to-beat cardiovascular recordings from 276 young healthy subjects (147 females). The examined spontaneous time series consist of RR intervals (RR), systolic blood pressure (SBP), and diastolic blood pressure (DBP), extracted as described in Chapter 2. A total of 41 features were extracted from each signal across the time, frequency, and information domains (Chapter 3).

Only methodological elements specific to this application are summarised here. A CMI-based forward feature selection procedure was implemented, starting from the computation of univariate mutual information between each feature and the sex variable. Significance thresholds were estimated using surrogate data generated by random permutation of the feature values. Features exceeding the threshold were retained and subsequently evaluated using conditional mutual information conditioned on previously selected features. This iterative procedure continued until no feature exhibited statistically significant CMI.

The selected features were used to train a LDA classifier (Chapter 6.1) validated with 10-fold cross-validation. Classification performance was quantified using accuracy, recall, and F1-score.

Results and Discussion. Results shown in Fig. 7.20 (A) suggest that the feature selection algorithm can identify an initial set of relevant features reaching the stopping

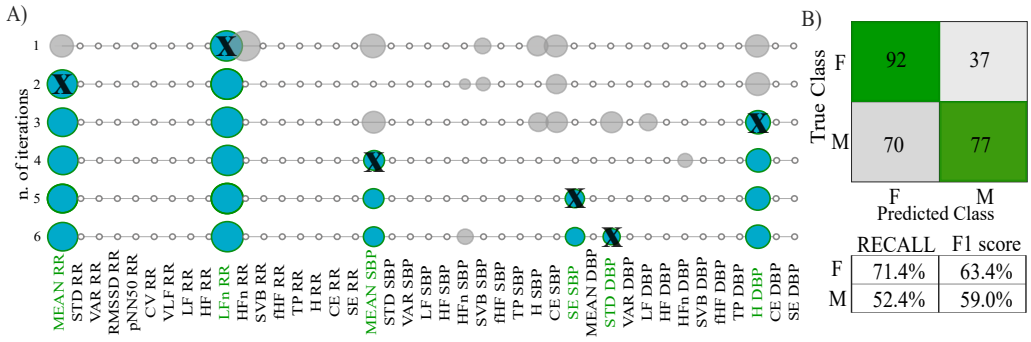


Figure 7.20: (A) Schematic representation of the FS algorithm procedure and results for each iteration. Grey bullets: features with significant MI or CMI values. Bold X: specific feature selected in a given iteration. Orange bullets: all the features selected for each iteration. In orange: features selected after all the steps. The size of the bullets is proportional to the value of MI or CMI. (B) Confusion matrix and tables of the classification performance results (Recall and F1-score) using only the selected features. This figure has been adapted from [96].

criterion after six iterations. In the first iteration, only 8 features (out of the 41 total) achieved significant MI values (grey bullets in step 1 of Fig. 7.20 (A)), of which the LF_n-RR feature exhibited the maximum value (orange bullet and bold X). At the end of all the iterations, a total of 6 significant features were selected, including features from all three domains: LF_n RR, MEAN RR, H DBP, MEAN SBP, SE SBP, and STD DBP, all with a widely known physiological meaning [212]. In particular, the MEAN is recognized as the basic cardiovascular measure expressing the balance between parasympathetic and sympathetic control [212]. The LF power reflects baroreflex activity and not cardiac sympathetic innervation [212]. Information-theoretic metrics (SE and H) reveal the complexity of physiological processes [183]. Some features remained statistically significant over multiple iterations, indicating their robustness in contributing to the feature set. Recall and F1-score results, reported in Fig. 7.20 (B), proved an overall good ability of the classifier to discriminate between sexes, with relatively high values for females (~71% and ~63%, respectively), despite the sub-optimal overall classification accuracy (~62%) that is comparable to value achieved obtained using all the 41 features (~61%).

Conclusions. The study presented an approach employing Conditional Mutual Information for feature selection, suggesting promising results for sex classification. While preliminary, our results highlight the discriminatory power of the selected features and their ability to achieve overall good accuracy comparable to the entire set. The study is primarily concentrated on the performance of the proposed FS approach, rather than the characteristics of the classifier itself, alongside the limited amount of data which could affect the results.

7.3.3 Assessment of Sex-Related Differences in Cardiovascular Control Using High-Order Feature Importance

Introduction. This application [164] applies a recently developed High-Order Feature Importance (HiFi) framework (Chapter 5.1.2, [163]) to quantify unique, redundant, and synergistic information provided by cardiovascular variability features for sex discrimination. By decomposing information contributions beyond pairwise dependencies, this approach aims to clarify which combinations of cardiac and vascular indices most strongly encode sex-related autonomic differences.

Materials and Methods. Two independent datasets were analysed in this study. The first (**D1**) consists of short-term beat-to-beat cardiovascular recordings acquired under controlled resting, postural, and cognitive conditions, while the second (**D2**) includes spontaneous 5-minute recordings from a large cohort of healthy young adults. Both datasets are fully described in Chapter 7.1.

For the analyses, only the first REST phase of the second dataset was considered to ensure uniformity with the first dataset. This choice was justified by the high degree of similarity between the two populations in terms of sex distribution, anthropometric, and clinical characteristics, as both datasets include only healthy young subjects. In the D2 dataset, mean age was 23.9 ± 2.5 years for females and 23.6 ± 2.7 years for males ($p = 0.450$, Wilcoxon rank-sum test), while BMI was 20.8 ± 2.1 and 23.3 ± 2.7 kg/m^2, respectively ($p < 0.001$). In the D1 dataset, mean age was 18.5 ± 2.8 years for females and 18.8 ± 3.9 years for males ($p = 0.744$), and BMI was 21.0 ± 1.8 and 22.0 ± 2.6 kg/m^2, respectively ($p = 0.042$). Between datasets, age distributions differed significantly for both sexes ($p < 0.001$), reflecting the slightly older age of the D1 participants, while BMI differences were not significant for females ($p = 0.361$) but significant for males ($p = 0.005$). All values were within normal physiological limits, confirming that both populations represent healthy young adults with comparable demographic characteristics. A comparison of demographic parameters confirmed the internal homogeneity of each dataset and their overall comparability. Although participants from the D1 dataset were on average older than those from the D1 dataset ($p < 0.001$ for both sexes), all subjects were healthy, normotensive, and exhibited BMI values within the normal physiological range. The minor age difference reflects the different recruitment settings (young adults in D2 vs. late adolescents in D1) rather than heterogeneity within groups. Given the large sample size and the narrow variance of the measured variables, this demographic distinction is unlikely to bias the information-theoretic or classification analyses, which rely primarily on normalized and dimensionless cardiovascular indices.

For each subject, RR intervals (RR), systolic blood pressure (SBP), and diastolic blood

pressure (DBP) time series were extracted, and the complete set of time-, frequency-, and information-domain variability features (41) introduced earlier in the thesis (Chapter 3) was computed.

The HiFi framework was applied to evaluate the contribution of each feature, and groups of features, to sex classification. This method decomposes the mutual information between the feature set and the target variable (sex) into components reflecting unique, redundant, and synergistic contributions, extending classical feature importance analysis to higher-order interactions. Considering sex as the target variable Y and cardiovascular features as the drivers X, the importance of each driver (feature) was estimated by HiFi, and decomposed into its unique contribution, plus the redundant and synergistic contributions that quantify its cooperation with other features. In addition, the corresponding redundant/synergistic multiplet for each driver is also reported. To increase the robustness of the regression model estimation, implemented here as a linear model, we merged the two datasets into a combined sample. This aggregation ensured a broader and more heterogeneous population for model training, enhancing the stability and generalizability of the HiFi decomposition. Also, before HiFi estimation the dataset was normalized using Z-score.

Using the most synergistic features obtained from HiFi and employing machine learning techniques we also aimed to identify the physiological patterns in heart rate and blood pressure variability able to distinguish between female and male subjects. This approach can reveal patterns of physiological features that cannot be assessed through conventional statistics, such as univariate hypothesis tests. Therefore, we performed sex classification based on the synergy-aware ranking of features derived from HiFi. This analysis followed a domain generalization setup: Dataset 1 was used exclusively for model training and feature selection, while Dataset 1 served as the unseen test set. HiFi was re-estimated on D2 data alone, and features were ranked according to their informativeness index $I = U + S - R$. Only features with $I > 0$ were retained for classification. To further improve the reliability of the estimated information contributions, we implemented a bootstrap resampling procedure with $n_{\text{boot}} = 100$ iterations. The final values of U, R, and S for each feature were computed by averaging across bootstrap iterations, thus reducing the impact of sampling variability and potential outliers.

A 10-fold cross-validation procedure was conducted on D2. Within each fold, a repeated hold-out validation (100 repetitions) was performed on the training portion to determine the optimal number of top-ranked features. At each step, a Support Vector Machine (SVM) classifier with default hyperparameters was trained using an increasing number of ranked features, and the subset achieving the highest average accuracy was selected. This optimal feature subset was then used to train a final model on the full

training set and evaluated on the test fold from D2 as well as on the external D1 dataset. Z-score normalization was applied to training data and consistently transferred to the test sets in each iteration.

Classification performance was evaluated using accuracy, recall, and F1-score metrics, computed for both D2 (aggregated over folds) and D1 in each cross-validation iteration. The entire procedure was repeated 100 times to ensure robust performance estimates. Additionally, we reported the selection frequency of each feature across all repetitions and folds to identify those most consistently informative.

This synergy-aware feature evaluation and selection strategy enables the identification of cardiovascular features that are not only individually predictive but also collectively complementary, thereby enhancing both model interpretability and classification performance.

Results. Figure 7.21 illustrates the importance of each cardiovascular feature, estimated via predictability decomposition, as well as using a pure LOCO and pairwise indices (for simple comparison purposes) [163]. These results provide insight into the high-order contribution of each feature to the sex prediction model.

The set of considered variables (41 features extracted from cardiovascular time series) is dominated by a synergistic higher-order dependencies ($meanS = 0.0306 \pm 0.0301$, $meanR = 0.0078 \pm 0.0106$, and $meanU = 0.0006 \pm 0.0026$). Several features from different signals and domains exhibited a high synergistic information: MEAN, STD, RMSSD, pNN50 and CV -from RR-, STD and TP -from SBP-, and STD and TP -from DBP-; confirming the complex interrelation of cardiovascular regulation. The MEAN and HFn RR features present the highest unique information, indicating their strong individual contribution to the sex prediction models. This suggests that these two features in an alone way provide relevant and non-redundant information about the target variable. Among the redundant contributions, LFn and HFn -from RR-, and HFn and CE -from SBP- were the most redundant features. Compared to LOCO and pairwise indices, the decomposition reveals a clearer separation of unique, redundant, and synergistic information, offering a more nuanced interpretation of feature relevance. These results suggest that while traditional feature importance methods like LOCO may identify key features, only decomposition-based techniques uncover the intricate multivariate dependencies driving sex-specific cardiovascular dynamics, due to LOCO can fail in detecting redundancy effects, and pairwise can miss synergistic dependencies.

In addition, the higher-order patterns (redundant and synergistic multiplets of each feature/driver) are shown in Figure 7.22 and 7.23, respectively. The number of features entering a multiplet for synergy or redundancy varies depending on how many iterations

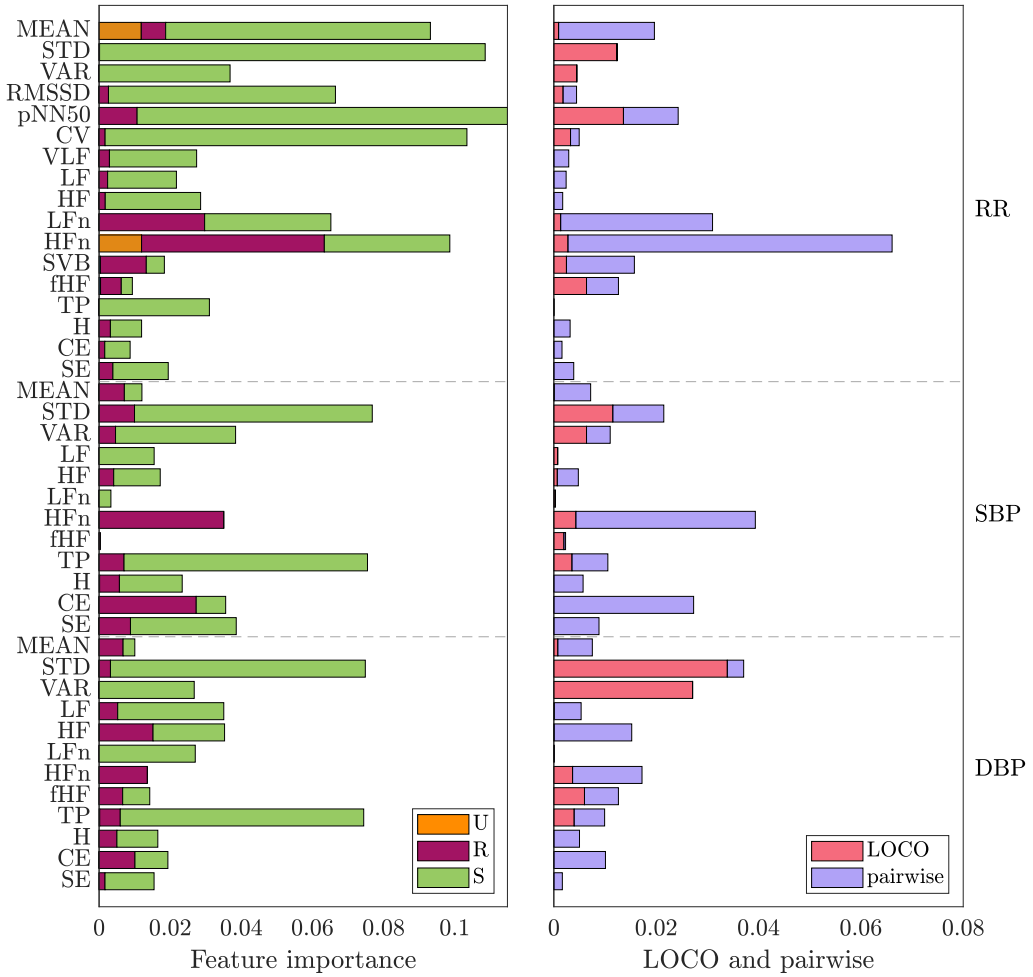


Figure 7.21: Feature importance measure of each cardiovascular variable. The left panel shows the predictability decomposition: unique (U), redundant (R) and synergistic (S) information reported in red, red and green colors, respectively. The panel on the right shows the LOCO and the pairwise indices, reported in light red and yellow colors, respectively. This figure has been adapted from [164].

are done looking for Z_{min} and Z_{max}, and the selected features are presented as colored squares in the raw corresponding to each individual driver. For visualization purposes, we plotted an approximation of the decrement/increment of the LOCO value for each feature of each redundant/synergistic multiplet. Most of the R and S patterns from RR drivers, are also composed of other RR variables. For SBP and DBP drivers, the corresponding multiplets include variables for both blood pressure signals. These findings suggest the presence of two relatively independent regulatory subnetworks: one centered on HRV and the other on BPV, each exhibiting sex-specific interdependencies.

Using HiFi decomposition for sex prediction, several features were consistently selected across the cross-validation and repetitions (see Figure 7.24), indicating a robust and stable contribution to the sex classification. Most of the features correspond to RR series,

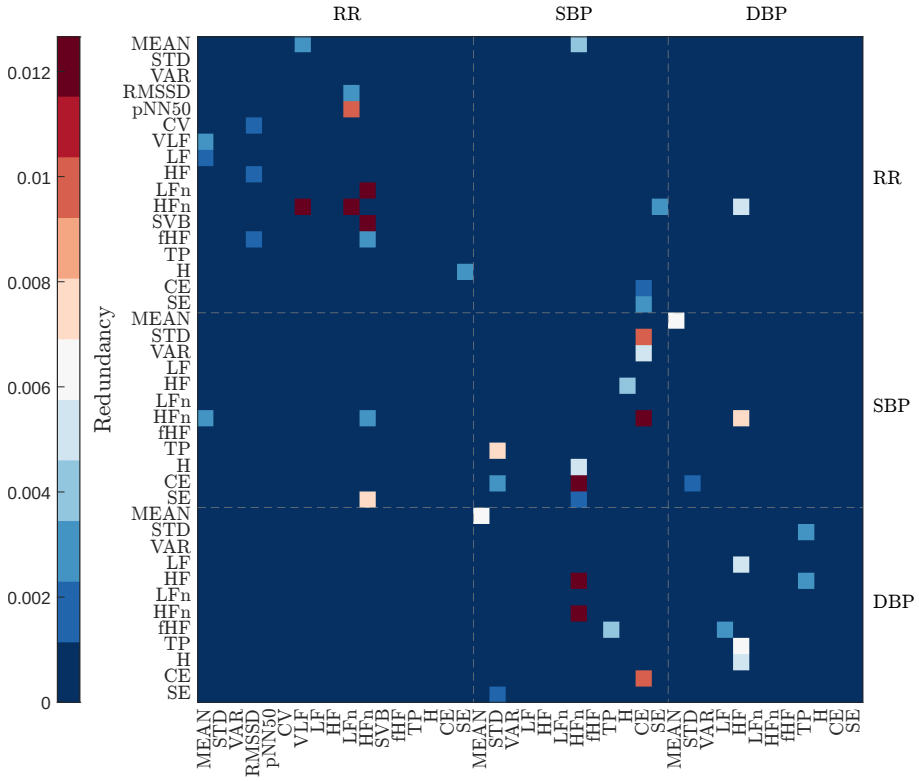


Figure 7.22: The redundant multiplets for each cardiovascular feature (driving variables). The color value representing the decrement of LOCO value due to the inclusion of that variable (column) in the multiplet of a specific driver variable (row). This figure has been adapted from [164]

suggesting that heart rate variability plays a dominant role in sex-related autonomic regulation. Interestingly, STD and TP features from both HR and BP signals were frequently selected, indicating their relevance in capturing amplitude- and power-related physiological fluctuations. On the other hand, entropy-based features, as well as frequency-specific indices (especially VLF and HF), were rarely selected, suggesting limited discriminative power for sex classification in this context.

The results of the classification performance further validate these findings and are summarised in Table 7.11. The classification performance is reported in terms of accuracy, recall, and the F1-score measures for each sex. The overall mean accuracy was around 0.68 with a small variance obtained across all repetitions, for both the validation data and the independent test data. This suggests a stable sex prediction without overfitting. Recall and F1-scores are higher for female than for male class, suggesting that sex-related cardiovascular patterns may be more consistently captured for female subjects in the used datasets.

Discussion and Conclusions. This study presents a novel high-order framework for feature importance and selection, designed to explore sex-related differences in cardio-

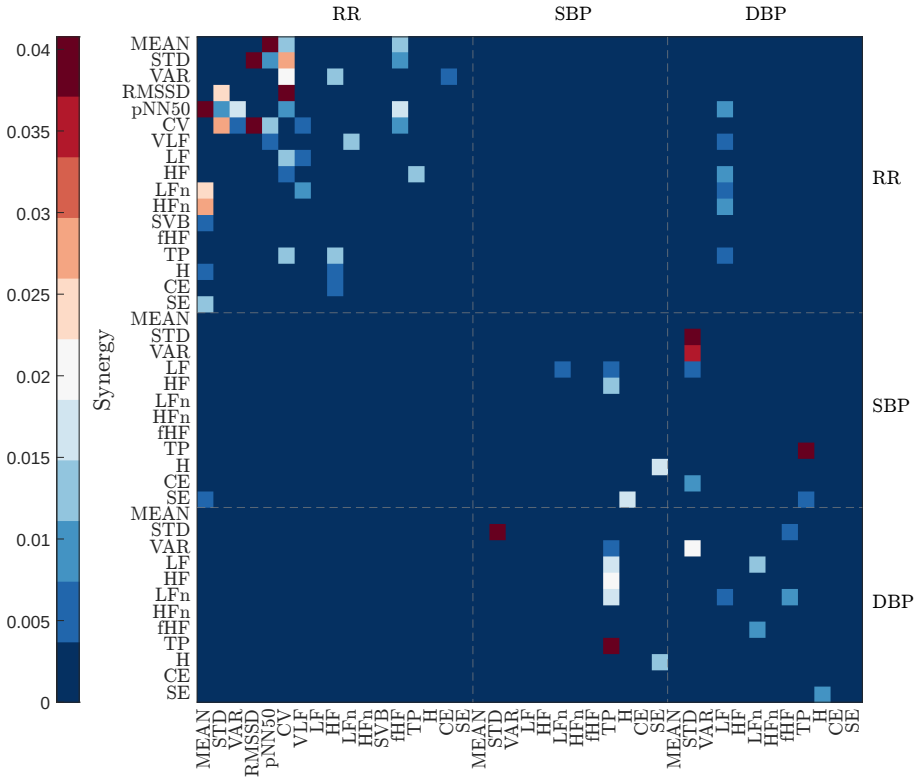


Figure 7.23: The synergistic multiplets for each cardiovascular feature (driving variables). The color value representing the increment of LOCO value due to the inclusion of that variable (column) in the multiplet of a specific driver variable (row). This figure has been adapted from [164]

vascular control through physiological time series. By leveraging an IT decomposition that accounts for unique, redundant, and synergistic contributions of each of the physiological indices extracted from RR, SBP and DBP time series, we go beyond conventional univariate or pairwise analyses and provide deeper insight into the collective role of these indices in sex classification. The application of this methodology to cardiovascular indices derived from RR intervals and blood pressure time series has revealed distinctive patterns in how male and female physiological systems encode information.

Notably, features such as MEAN and HFn extracted from RR intervals exhibited high unique information content, confirming their relevance in capturing sex-specific autonomic modulation. MEAN RR reflects baseline heart rate and is generally higher in females due to greater vagal tone and lower sympathetic drive at rest [119], whereas HFn is a validated marker of parasympathetic activity [212]. Their consistent selection aligns with existing knowledge of sex differences in autonomic balance. Moreover, time-domain features like RMSSD and STD, which reflect beat-to-beat and overall heart rate variability, were frequently involved in synergistic interactions, suggesting that combinations of these features encode higher-order information not apparent when considered individu-

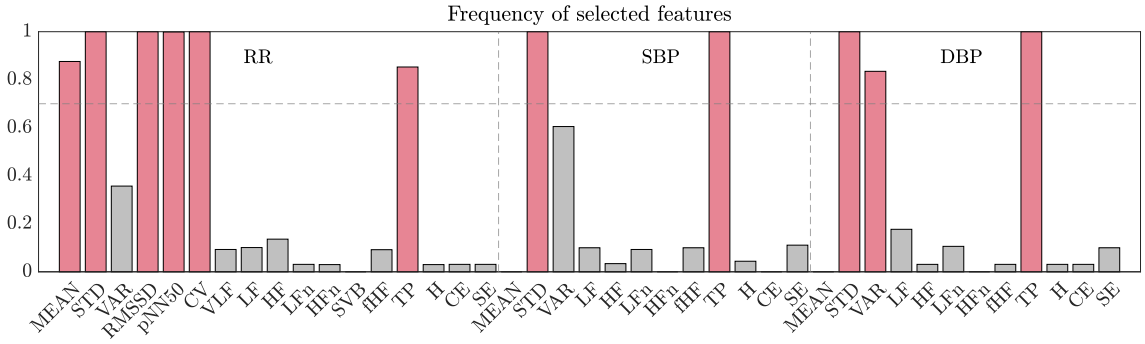


Figure 7.24: Frequency of selected features for sex classification. Black color represents the features that were selected more than 70% across bootstrap with the 10-fold cross-validation. This figure has been adapted from [164]

Table 7.11: Classification metrics. This table has been taken from [164]

	Test data (D_2)			Validation data (subsets from D_1)		
Measures	Whole	M	F	Whole	M	F
Accuracy	0.68 [±0.01]	–	–	0.69 [±0.01]	–	–
Recall	–	0.65 [±0.01]	0.71 [±0.01]	–	0.64 [±0.02]	0.73 [±0.02]
F-score	–	0.63 [±0.01]	0.72 [±0.01]	–	0.65 [±0.02]	0.71 [±0.01]

ally. This suggests that different variability indices provide complementary information, and supports the notion that the autonomic nervous system does not operate in isolation but through complex and nonlinear couplings, particularly in female physiology, where parasympathetic modulation tends to dominate and display higher complexity [186].

Although several HRV and BPV indices are mathematically or physiologically related (e.g., STD and total power; RMSSD, pNN50, and HF power), this interdependence represents coherent aspects of autonomic regulation rather than methodological bias. Indeed, the similarity between features introduces redundancy, but since redundancy is explicitly subtracted in the informativeness index, a high selection frequency indicates that the feature is a truly robust physiological marker rather than an artefact of correlation. The information-theoretic framework employed in this study quantifies and minimizes shared information among variables, thereby reducing the impact of multicollinearity and ensuring that selected features capture meaningful physiological variability.

From the blood pressure time series, features such as DBP-TP and SBP-STD were also prominently involved in synergistic groups. These features reflect vascular tone and baroreflex contributions to cardiovascular control, which are known to differ between sexes due to hormonal and structural factors [179, 198]. Specifically, females often exhibit higher baroreflex sensitivity and lower arterial stiffness, which increases the variability and reactivity of blood pressure signals. The synergistic behavior of these indices suggests that

only in combination do they effectively encode the sex-related physiological information, possibly reflecting sex-specific coordination between cardiac and vascular subsystems.

It should be noted that SBP and DBP originate from the same arterial pressure waveform and are therefore partially correlated. This interdependence was reflected in the redundant components identified within the blood pressure domain. However, these two measures describe distinct physiological mechanisms: SBP primarily reflects cardiac ejection and arterial stiffness, whereas DBP is influenced by vascular properties (peripheral resistance and vascular compliance) and heart rate. Their joint inclusion thus enables a more comprehensive characterization of cardiovascular control. The information-theoretic decomposition explicitly separates redundancy from synergy, revealing that while SBP and DBP features share common information, they also contribute synergistically when combined with HRV indices. Excluding one of them would likely reduce the richness of the observed cross-system interactions, particularly those integrating cardiac and vascular dynamics.

In this context, it is also relevant to discuss the behaviour of the peak frequency within the high-frequency band (fHF), which reflects the dominant respiratory rhythm. One might expect a strong redundancy among fHF values extracted from RR, SBP, and DBP series, since respiration acts as a common driver across cardiovascular signals. However, the respiratory modulation of heart period, systolic, and diastolic pressures originates from distinct physiological mechanisms: in RR intervals, fHF primarily represents respiratory sinus arrhythmia mediated by vagal efferent activity, whereas in SBP and DBP it arises from mechanical and baroreflex influences that introduce phase shifts, amplitude attenuation, and partial decorrelation across signals [66, 186]. These differences in SBP and DBP compared to RR, could contribute to the reduced alignment of HF components and hence to the low redundancy detected in our analysis. Moreover, it is plausible that part of the redundancy associated with fHF is already captured by other HRV features, such as RMSSD and HFn, which were frequently selected as redundant for fHF in RR. This would explain why the inclusion of fHF from BP series did not further reduce conditional information. The very low selection rate of fHF further supports the interpretation that these indices are redundant both among themselves and with other features that more directly represent respiratory-related variability.

The decomposition of feature importance also revealed two functionally distinct but partially interacting modules: one primarily associated with heart rate variability and another with blood pressure variability. This modularity may correspond to the division between parasympathetic-dominated and sympathetic-dominated pathways, respectively. The low degree of cross-module synergy suggests that RR and BP features carry complementary information, highlighting the need to combine them for a more detailed representation of

sex-specific cardiovascular dynamics. Such modular organization is consistent with previous findings showing partial independence between cardiac and vascular control loops [175].

From a classification standpoint, our model achieved satisfactory performance with balanced accuracy and robustness across repetitions and cross-validation folds. Classification performance metrics suggest that the female class achieved a better prediction than the male class, which agrees with prior studies reporting more stable autonomic patterns in women [119]. This discrepancy may be attributed to hormonal influences such as estrogen's role in enhancing vagal tone and baroreflex sensitivity, which increase signal regularity and reduce inter-subject variability in females [179, 198]. However, this point needs to be examined more rigorously in future studies.

The sex-related differences observed in cardiovascular variability are consistent with known neurophysiological and hormonal mechanisms. Females generally exhibit predominant parasympathetic (vagal) modulation and attenuated sympathetic activity, associated with higher heart rate variability and lower blood pressure variability at rest [119, 198]. Estrogen exerts multiple autonomic effects, enhancing vagal tone, baroreflex sensitivity, and endothelial nitric oxide release, while reducing sympathetic vasoconstrictor outflow [179]. Conversely, testosterone increases sympathetic tone and vascular stiffness, contributing to the reduced HRV and elevated BP typically observed in males. These mechanisms likely underlie the stronger and more stable synergistic interactions identified in females, reflecting tighter integration of cardiac and vascular control loops.

Despite the interesting findings, several limitations must be acknowledged. The study focused on resting-state data from a healthy, young population, which may limit generalizability. As discussed by Porta et al. [185, 188, 189], respiration can bias the balance between redundancy and synergy by introducing shared variability among cardiovascular signals. Although in the present study the target variable was sex rather than the temporal evolution of heart period or pressure, respiration may still act as a latent common driver modulating the relationships among HRV and BPV features. In particular, respiration simultaneously affects RR, SBP, and DBP, potentially contributing to sex-related differences in the extracted indices. Females generally display higher spontaneous breathing frequencies and lower tidal volumes than males, reflecting hormonal and anatomical influences on ventilatory control [25, 139]. From this perspective, the recurrent selection of features such as STD and RMSSD may in part reflect sex-specific respiratory modulation, which is physiologically integrated with autonomic regulation.

In conclusion, our high-order information decomposition framework offers a powerful and interpretable approach for investigating sex-related differences in physiological reg-

ulation. By quantifying not only the individual but also the synergistic contributions of features, it bridges the gap between statistical learning and physiological understanding. These results suggest that interpretable and physiologically-informed machine learning models can enhance both the classification of sex-related differences and our comprehension of the underlying mechanisms, opening the way for more personalized and equitable healthcare technologies. Finally, although this study emphasized cardiovascular signals, integration with respiratory, hormonal, and behavioral data could further enhance model accuracy and interpretability. The modular nature of physiological information discovered here supports a broader multivariate framework that could be expanded to multimodal health monitoring, particularly in personalized medicine.

7.4 Further Physiological Applications

This final section presents additional applications that complement the main physiological themes addressed in this thesis and further showcase the flexibility of the analytical framework developed in the previous chapters. While the earlier sections focused on stress characterisation and sex-related differences, two well-defined physiological contexts, here the goal is to illustrate how the same information-theoretic, feature-selection, and classification methodologies can be extended to broader scenarios. These studies do not target a single physiological mechanism but instead highlight the generality of the proposed approaches and their applicability across diverse datasets, experimental protocols, and biomedical objectives.

7.4.1 A principled and data-efficient information-theoretic method for feature selection

Introduction. This application [98] applies the k-nearest neighbour Conditional Mutual Information-based Feature Selection (kCMI-FS) framework described in Chapter 4.2.1 to evaluate its performance across multiple biomedical datasets. The method extends previous implementations by optimizing the estimation of conditional dependencies among continuous variables through a data-efficient kNN approach. Here, it is applied to assess the ability of kCMI-FS to identify informative and non-redundant features across heterogeneous physiological and clinical scenarios, thereby testing its robustness, scalability, and adaptability to diverse data structures.

Material and Methods. The performance of the kCMI-FS method has been evaluated on four biomedical datasets. Table 7.12 summarises the main information on the datasets in terms of number of samples, features and classes, showing how they differ in dimensionality and class structure. This heterogeneity enables a comprehensive evaluation of the

Table 7.12: Information (number of samples, features and classes) on the four biomedical datasets. This table has been taken from [98].

Dataset	**Samples**	**Features**	**Classes**
Breast Cancer (BC)	569	31	2
Parkinson's disease (PA)	195	22	2
TCGA-BRCA (BRCA)	572	50	4
Cardiovascular variability (**D1**)	127	53	3

robustness and adaptability of the kCMI-FS method across diverse and realistic settings.

To benchmark the behaviour of our method against well-established FS strategies, we additionally employed the publicly available Feature Selection Toolbox (FEAST) developed by Brown *et al.* [36], which provides implementations of several widely used mutual-information–based feature-ranking algorithms, also described in Sec. 4.1, including Mutual Information Maximization (MIM), Mutual Information Feature Selection (MIFS), Joint Mutual Information (JMI), Conditional Mutual Information Maximization (CMIM), Double Input Symmetrical Relevance (DISR), Conditional Infomax Feature Extraction (CIFE), Interaction Capping (ICAP), Conditional Redundancy (CondRed), and the Fast Correlation-Based Filter (FCBF). Out of these, CMIM is the most closely related to the proposed algorithm, as it iteratively selects candidate features by maximising the minimum CMI conditioned to a single previously selected feature [71]. This approach, while simpler, potentially could fail to capture high-order interactions that emerge when utilising the whole selected subset. Since the previous algorithms operate on discretised inputs, following the authors' recommendation [36], uniform quantisation with 10 equally spaced bins was carried out on the data. Since the methods generate ranked feature lists, a knee-point strategy to determine the number of features to retain was applied, restricting our comparison to filter-based approaches. This selection was automated using the algorithm proposed by Satopää *et al.* [207]. Moreover, the mRMR FS algorithm [172] and the CMINN and BIN methods already used in Section 4.2.1 were applied, for consistency with the analyses carried out on the synthetic datasets. For the same reason, MI and CMI were estimated in kCMI-FS and CMINN using the kNN approach with parameter $k = 10$.

The classification stage encompassed a Monte Carlo cross-validation (CV) strategy with 10 repetitions. For each repetition, 70% of the data was randomly selected for FS and model training, while the remaining 30% was held out for testing and performance evaluation. In line with [227], FS was embedded within the classification pipeline and evaluated jointly with the classifier. The only exception was the PA dataset, which was partitioned on a per-subject basis, given the limited number of subjects. The selected features were then used to train four classifiers with different inductive biases: Support Vector Machine

(SVM), k-Nearest Neighbors (kNN), Linear Discriminant Analysis (LDA), and Random Forest (RF)[31], implemented with class weighting to mitigate class imbalance. Model hyperparameters were optimised on the training data using the default 5-fold CV procedure of the MATLAB 2023a implementation.

Classification performance was assessed using conventional accuracy and F1-score metrics (Sec. 6.3. In multi-class settings, F1-scores were computed on a per-class basis and macro-averaged to provide an overall performance estimate, assigning equal weight to each class. For each dataset and method, results were summarised by reporting the mean and standard deviation (SD) of each metric across CV repetitions. Statistical comparisons between FS methods were conducted independently for each dataset and performance metric. A Wilcoxon signed-rank test was applied pairwise between the distribution of values obtained for each method, and the best-performing algorithm was defined according to the median performance across CV repetitions. Statistical significance was assessed at three levels: $p < 0.05$, 0.01, and 0.001.

To evaluate whether the proposed kCMI-FS method is dependent on a specific classification model, the four aforementioned classifiers were applied on the selected features. Table 7.14 reports the classification performance across datasets in terms of accuracy and F1-score. A pairwise Wilcoxon signed-rank test across cross-validation folds with Holm correction for multiple comparisons was applied to compare SVM with the other classifiers. Overall, performance variations were generally small and not systematic with respect to model type, without statistically significant differences among the classifiers across datasets in almost all the cases. These results indicate that the features selected by kCMI-FS generalize well across linear, distance-based, and ensemble classifiers, suggesting that the method is not model-dependent. Therefore, for the sake of brevity and to avoid repetitions discussing the results, the subsequent analyses were carried out using the SVM classifier only.

Results. Figure 7.25 reports the heatmaps of FS occurrences across all four datasets. This unified representation allows the comparison of datasets with substantially different dimensionalities and numbers of classes, providing insight into how heterogeneity affects the stability and interpretability of the outcomes. Only features with at least 50% occurrences in at least one method are shown, ensuring that the maps highlight consistently informative features while filtering out low-frequency, noise-driven selections. Table 7.13 summarises, for each dataset and FS method, the number of features selected based on the average 50% inclusion ratio across all folds, to provide a reference for comparing the cardinality of the selected subsets.

Figure 7.26 summarises the classification performance obtained across all datasets and

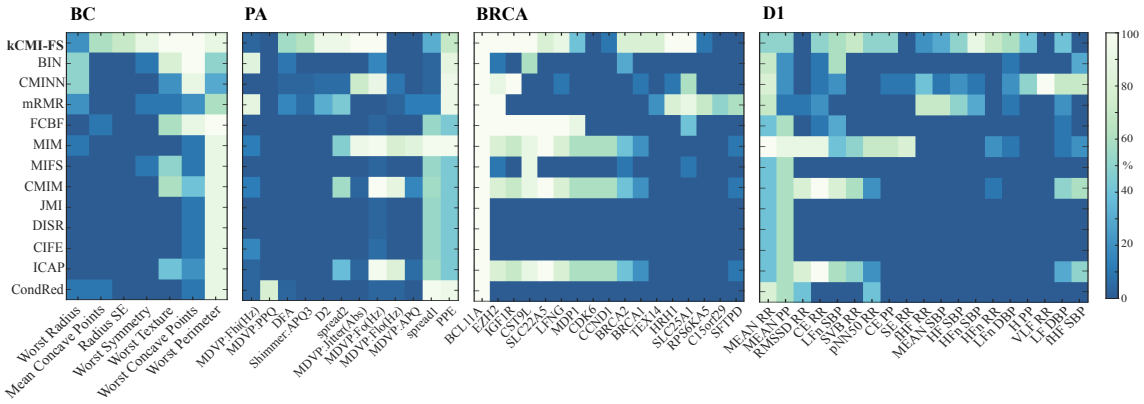


Figure 7.25: Heatmaps of feature-selection occurrences across the four datasets (BC, PA, BRCA, **D1**). For each dataset, panel shows the corresponding heatmap, reporting how many times each feature was selected by each method, using a global colour scale shared by all datasets. Only features with at least 50% occurrences in at least one method are shown. This figure has been adapted from [98].

Table 7.13: Number of features retained by each feature-selection method on the considered datasets. Reported values correspond to the final retained subsets, defined by stability-based selection threshold (≥ 50% of CV repetitions). This table has been taken from [98].

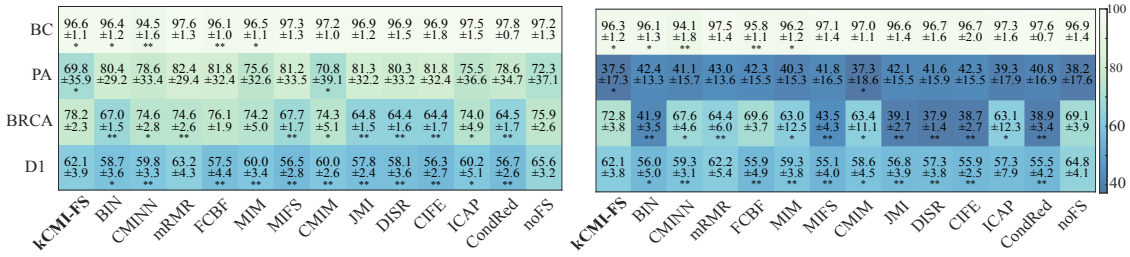
Dataset	**kCMI-FS**	BIN	CMINN	mRMR	FCBF	MIM	MIFS	CMIM	JMI	DISR	CIFE	ICAP	CondRed
BC	6	6	4	2	3	1	3	4	1	1	1	2	3
PA	7	2	3	2	2	6	2	4	1	1	1	3	2
BRCA	9	2	4	7	8	8	3	3	1	1	1	9	1
D1	11	3	4	4	3	7	3	7	1	1	1	7	2

FS methods. Moreover, the same analysis was conducted without applying any FS method (noFS), to assess the impact of FS on classification performance, using the identical CV procedure for consistency.

Breast Cancer dataset. The Diagnostic Wisconsin Breast Cancer (BC) dataset [221] consists of digitised fine needle aspirate images of breast masses and is used to discriminate benign from malignant cases based on a set of real-valued morphological descriptors of cell nuclei [110]. For this dataset, features selected with kCMI-FS (Fig. 7.25) predominantly originate from the group of "worst" morphological descriptors, capturing the most atypical nuclei within each sample. These measures provide a quantitative characterisation of nuclear abnormalities that are often diluted or missed when averaging over regions of interest. In particular, *Worst Texture* reflects chromatin heterogeneity, while *Worst Perimeter* and *Worst Concave Points* quantify irregular nuclear contours, both established cytological markers of malignancy [180, 253]. The inclusion of *Radius SE* further highlights increased within-region variability in malignant nuclei. In contrast to other methods, that predominantly provide stable selections limited to *Worst Perimeter*, kCMI-

Table 7.14: Classification performance across datasets and classifiers. Results are reported as mean ± standard deviation (%). Statistical significance was assessed comparing SVM with the other classifiers across cross-validation folds through pairwise Wilcoxon signed-rank test with Holm correction for multiple comparisons ($* p < 0.05$). This table has been adapted from [98].

	Accuracy [%]				F1-score [%]			
Dataset	SVM	kNN	LDA	RF	SVM	kNN	LDA	RF
BC	96.6 ± 1.1	95.9 ± 1.7	95.8 ± 1.6	95.9 ± 2.1	96.3 ± 1.2	95.6 ± 1.9	95.4 ± 1.8	95.6 ± 2.2
PA	69.8 ± 35.9	74.5 ± 34.4	81.2 ± 33.0	79.2 ± 34.1	37.5 ± 17.3	39.6 ± 16.2	42.1 ± 15.5	41.2 ± 16.2
BRCA	78.2 ± 2.3	75.2 ± 2.0*	78.4 ± 3.1	78.1 ± 2.3	72.8 ± 3.8	68.5 ± 3.5*	74.8 ± 3.6	72.8 ± 4.4
CARDIO	62.1 ± 3.9	60.6 ± 3.9	63.9 ± 3.1	61.8 ± 3.9	61.1 ± 3.8	59.7 ± 4.0	63.3 ± 3.4	61.0 ± 4.3



(a) Accuracy (b) F1-score

Figure 7.26: Heatmaps summarising the classification performance across all datasets and FS methods for the two evaluation metrics: (a) accuracy, (b) F1-score. For each dataset, values indicate mean ± SD over the CV repetitions, while significance markers below the values indicate whether a method performs significantly worse than the best-performing one according to the Wilcoxon signed-rank test. Statistical significance is reported at three levels: $* p < 0.05$, $** p < 0.01$, and $*** p < 0.001$. This figure has been adapted from [98].

FS is consistent with additional well-established pathological findings and confirms their strong discriminative value. Other methods yield a wider number of features, probably due to their CMI calculation algorithm (e.g., for BIN, CMIM) or to a weaker redundancy selection criteria (e.g., for FCBF). The kCMI-FS achieves 96.6% ± 1.1% accuracy and 96.3% ± 1.2% F1 in the BC dataset, close to noFS baseline (97.2% ± 1.1%), as reported in Fig. 7.26. Competing methods such as mRMR and ICAP produce slightly higher accuracies but with inconsistent subsets across repetitions. Methods that select minimal subsets (typically 1 feature) experience a notable decline in performance (down to 91–93%).

Parkinson's Disease dataset. The Oxford Parkinson's Disease Detection (PA) dataset [137] consists of voice recordings collected from healthy individuals and patients diagnosed with Parkinson's disease. A set of acoustic features is extracted to characterise vocal impairments associated with the pathology. Applying kCMI-FS to this dataset yields seven stable features across the repetitions (Fig. 7.25): *MDVP:Fo(Hz)*, *MDVP:Jitter(Abs)*, *Shimmer:APQ3*, *DFA*, *spread2*, *D2*, and *PPE*. Together, these variables capture comple-

mentary aspects of Parkinsonian dysphonia, spanning both classical perturbation metrics and nonlinear dynamical descriptors. The recurrent pitch-, jitter-, and shimmer-related features reflect well-documented phonatory instabilities arising from impaired laryngeal motor control, reduced neuromuscular coordination, and diminished precision in subglottal pressure regulation, reflecting short-term irregularities [18]. On the other hand, the consistent inclusion of *spread2*, *D2*, and *PPE* highlights alterations in the global temporal voice architecture. These descriptors quantify increased aperiodicity, irregular airflow and reduced fine motor modulation patterns, characteristic of Parkinson's disease [18]. The resulting feature subset aligns with previous studies emphasizing both perturbation- and complexity-based measures [6, 115]. The kCMI-FS, CMINN, MIM, CMIM, and ICAP integrate these short-term phonatory irregularities and nonlinear dynamical alterations, yielding a compact representation of voice pathology in Parkinson's disease.

The PA dataset exhibits substantial subject-wise variability (Fig. 7.26), resulting in large standard deviations in performance metrics for all the methods, thus making classification quite unreliable. kCMI-FS achieves an accuracy of 69.8% ± 35.9% and an F1-score of 37.50% ± 17.3%, values comparable to the noFS baseline (72.3% ± 37.1%) and all other competing methods. Algorithms such as JMI, DISR, and CIFE achieve accuracies below 60% due to excessively small feature subsets.

TCGA Breast Invasive Carcinoma dataset. A subset of the TCGA Breast Invasive Carcinoma (BRCA) RNA-sequencing cohort [226], sourced from the Marginal Contribution Feature Importance repository [43], was used for classification of breast cancer molecular subtypes. Applying kCMI-FS yields a compact and stable gene set (Fig. 7.25): *BCL11A*, *IGF1R*, *EZH2*, *BRCA1*, *BRCA2*, *SLC22A5*, *LFNG*, *TEX14*, and *SLC25A1*; all who align closely with established subtype-defining biological processes. These genes collectively reflect key variabilities in breast cancer, like cell proliferation (*BCL11A*, *IGF1R*), chromatin remodelling (EZH2), DNA repair (*BRCA1*, *BRCA2*), Notch signalling modulation (*LFNG*), metabolic rewiring (*SLC25A1*), or subtype-specific prognostic markers (*SLC22A5*, *TEX14*) [53, 159]. Notably, known proliferation-associated genes such as *CDK6* and *CCND1* are not retained, despite their established roles in breast cancer. Within a conditional-information framework, this omission reflects redundancy: once a previously selected gene captures most of the subtype-related information, additional correlated markers provide limited incremental information. Comparatively, only the selections of MIM, CMIM, and ICAP tend to similar selections, albeit with instability across repetitions, potentially failing to account for more nuanced high-order interactions. All algorithms, however, correctly infer the significance of *BCL11A* as a prognostic marker regarding the basal-type molecular subtype.

The kCMI-FS achieves the highest overall performance among all methods (Fig. 7.26),

with 78.2% ± 2.3% accuracy and 72.8% ± 3.8% F1. The noFS baseline performs worse (75.9% ± 2.6%), indicating the benefit of feature selection in transcriptomic data.

Cardiovascular dataset. The (**D1**) dataset [97, 164] includes electrocardiographic (ECG) and blood pressure (BP) recordings acquired under different physiological conditions (i.e. rest, postural and mental stress). Cardiovascular variability features extracted from ECG R-R and BP pulse-pulse intervals capture autonomic variability, yielding a physiologically grounded multi-class classification scenario. Applying kCMI-FS yields a coherent subset of eleven features spanning mean interval measures, nonlinear information-based descriptors, and frequency-domain indices derived from *RR*, systolic, and diastolic blood pressure series (Fig. 7.25). This subset captures complementary aspects of autonomic regulation, reflecting both slow tonic shifts in cardiac timing and faster oscillatory components of sympathetic and parasympathetic modulation [97, 164, 175]. The joint selection of ECG- and BP-derived features indicates that discriminative information is distributed across tightly coupled physiological domains rather than concentrated in a single signal. Overall, all methods select the mean-based ECG features, while kCMI-FS exhibits instabilities across repetitions regarding other selections. MIM, CMIM, and ICAP display similar, more stable selections.

The classification performance between rest and stress states is moderate across all methods due to the high redundancy among cardiovascular variability indices (Fig. 7.26). kCMI-FS achieves 62.1% ± 3.9% accuracy and 62.1% ± 3.8% F1, values slightly lower than noFS baseline (accuracy of 65.6% ± 3.2%) but with a much more compact and stable feature set (11 features). Competing methods either select too few features (JMI, DISR, CIFE: 1 feature) and underperform (56–58% accuracy), or select larger redundant subsets (MIM, ICAP: 7 features) without achieving notable performance gains (~60% accuracy).

Overall remarks. The classification results show that, compared with competing FS methods, kCMI-FS achieves a favourable balance between accuracy, model compactness, and stability across heterogeneous domains, making it particularly suitable for biomedical applications, where interpretability and robustness are essential. Moreover, kCMI-FS is able to achieve higher accuracy than noFS baseline in scenarios where redundant or noisy variables hinder learning, such as the BRCA dataset. The results confirm that kCMI-FS achieves relatively good or superior predictive performance across heterogeneous datasets, while selecting substantially fewer features and maintaining low variability across CV repetitions. Pairwise comparisons with the best-performing method revealed statistically significantly lower accuracy ($p< 0.05$) only for BC and PA datasets.

Discussion and Conclusions. This work introduced kCMI-FS, a feature selection method based on Conditional Mutual Information estimated via a k-nearest neighbour

strategy specifically adapted for mixed-type datasets. While the ability to handle redundancy and capture synergistic dependencies is a known theoretical advantage of CMI-based approaches [247], our contribution lies in developing an estimation framework tailored to scenarios involving discrete class labels and continuous features, conditions common in biomedical classification tasks.

Through real-world applications, we demonstrated that kCMI-FS achieves high accuracy in selecting relevant features while remaining robust to redundancy and noise. Its performance exceeded that of standard kNN-based estimators (such as CMINN [228]) and binning-based methods in complex settings, particularly where the detection of high-order dependencies is required. The method proved especially effective in identifying relevant features in datasets with strong interactions or correlated groups, while tolerating a small proportion of redundant or irrelevant variables, an acceptable trade-off for capturing nuanced informational patterns.

From a practical standpoint, the proposed mixed-variable estimator addresses key limitations of traditional kNN-based methods like KSG [122], which assume continuous distributions across all variables. This assumption is often violated in real-world data, where discrete labels or quantisation effects (e.g., low-precision formats) introduce boundary effects and tie distances. Several approaches have been proposed to address this challenge, including kernel-based and nearest-neighbour methods specifically designed for mixed-type data [51, 153, 228, 255]. Building upon these insights, our modified estimator mitigates these issues by incorporating mixed-type neighbourhood strategies, enabling reliable estimation in hybrid feature spaces while retaining the scalability and flexibility of the kNN framework.

The advantages of kCMI-FS stem primarily from its ability to capture nonlinear and synergistic dependencies, operate natively in mixed discrete–continuous spaces, and produce compact and interpretable subsets across heterogeneous datasets. These strengths make the method particularly suited to domains where interpretability and reliability are essential, such as biomedical decision support. At the same time, the method has some limitations: the greedy forward-selection procedure cannot guarantee globally optimal subsets; the statistical-testing phase increases computational cost compared to simpler filters; and some redundancy may persist when synergistic structures are strong or when features lie on tightly coupled manifolds. This reflects a trade-off between strict sparsity and the preservation of clinically meaningful dependencies. Furthermore, while computational complexity scales well for moderate feature dimensions, the quadratic dependence on the number of selected features suggests that very high-dimensional datasets may benefit from parallelised neighbourhood search or approximate density estimators.

It is worth noting that the sensitivity of our approach to synergistic interactions does not imply, from a theoretical point of view, that redundant or even irrelevant features can be selected. Therefore, we tend to ascribe the tendency towards selection of such features to practical aspects related to estimation. These can include: the greedy forward selection strategy which, being suboptimal, might lead the sequential selection towards local CMI maxima including redundant features; the surrogate-based test for significance, which might induce a more sensitive algorithm like the proposed kCMI-FS to yield CMI estimates above the empirical significance threshold even for irrelevant features. From a clinical interpretability perspective, the inclusion of partially redundant features may complicate the identification of more dominant driver predictors. However, this does not necessarily reduce interpretability, as in many biomedical settings the mechanisms are driven by synergistic interactions, rather than isolated effects. We therefore view this behaviour as a trade-off between strict sparsity and the preservation of clinically meaningful dependency structures.

The comparison with the noFS baseline further clarifies the strengths and limitations of the method. In highly redundant datasets (e.g., BC dataset), the noFS classifier benefits from large model capacity, but at the cost of instability and reduced interpretability. kCMI-FS, conversely, achieves comparable accuracy with lower variance across repetitions and smaller feature sets. In noisier conditions (e.g., PA dataset), performance may slightly decrease in comparison to the full set, but the selected features remain physiologically meaningful and substantially more stable across folds. This behaviour reflects the intrinsic difficulty of small noisy datasets and highlights the importance of methods that preserve interpretability even when predictive performance reaches its natural limits.

Another limitation, commonly occurring in biomedical applications, is the class imbalance, since high skewed class distributions may lead to unreliable local density estimates for minority classes, especially when using kNN estimators with a fixed neighbourhood size[128]. In our experiments, this effect was partially mitigated in imbalanced datasets (e.g. BRCA) by selecting a moderate value of k that provided a stable bias–variance trade-off, and by removing redundant predictors amplifying estimation variance in minority-class regions.

In conclusion, the findings of this study highlight the importance of selecting FS strategies that align with the structure and complexity of the data. Conservative methods may suffice when interpretability and compactness are critical, but in interaction-heavy scenarios, typical of clinical and physiological datasets, methods such as kCMI-FS offer improved adaptability and enhanced information capture. These characteristics are essential for building trust and transparency in machine learning for healthcare, ultimately contributing to more robust decision-making pipelines.

7.4.2 Information-Theoretic Estimation of High-Order Feature Effects in Classification Problems

Introduction. This application [131] builds upon the previously described High-Order Feature Importance (Hi-Fi) framework (Sec. 5.1.2 and its theoretical extension described in Sec. 5.2.1. The Hi-Fi approach was reformulated within an information-theoretic context, where feature contributions are decomposed into unique, redundant, and synergistic informational components, providing a model-free characterization of feature importance beyond marginal effects.

As detailed in the methodological chapters, this formulation links the change in predictive performance induced by omitting a feature to the Conditional Mutual Information (CMI) between the target and the source variable, conditioned on the remaining features. This enables a principled, model-independent quantification of how each feature contributes to prediction both individually and in interaction with others.

Here, only the experimental application of the kNN-based CMI estimator is presented, illustrating its ability to quantify high-order informational effects in realistic classification settings. Specifically, we demonstrate its use in the analysis of biomedical data, showing how the proposed method can identify relevant features and disentangle their unique and redundant informational contributions across complex, multivariate datasets.

Materials and Methods. Analyses were performed on a subset of The Cancer Genome Atlas - Breast Invasive Carcinoma (TCGA-BRCA) RNA sequencing dataset [226] containing gene expression measurements from breast cancer tissues. The data, obtained from the Marginal Contribution Feature Importance (MCI) [43] (GitHub page), comprises 572 samples characterized by the expressions levels of 50 genes, spanning four molecular subtypes: 304 Luminal A (LumA), 114 Luminal B (LumB), 105 Basal, and 49 HER2-enriched. The approach introduced in this work is used to investigate the contribution of gene expressions in determining the molecular subtype, considered as class outputs in the analysis. To estimate the confidence intervals of the informational contribution of each feature, we generate a bootstrapped dataset with 572 samples and $N_{sim} = 100$ sampling repetitions, while ensuring the same class counts. Using the kNN estimator, we fix the number of neighbors at $k = 10$, and employ $N_{surr} = 100$ surrogate samples for the selection criteria.

Results and Discussion. The results shown in Fig. 7.27 display the ranked top contributors according to the kNN CMI between the target class and each observed feature, conditioned on the subset of features that maximizes this measure, $I(Y; X|\boldsymbol{Z}_{max})$. According to [104, 245], the 10 genes: BCL11A, SLC22A5, CDK6, IGF1R, LFNG, CCND1, EZH2, BRCA1, BRCA2, and TEX14, are known breast cancer-associated markers, and

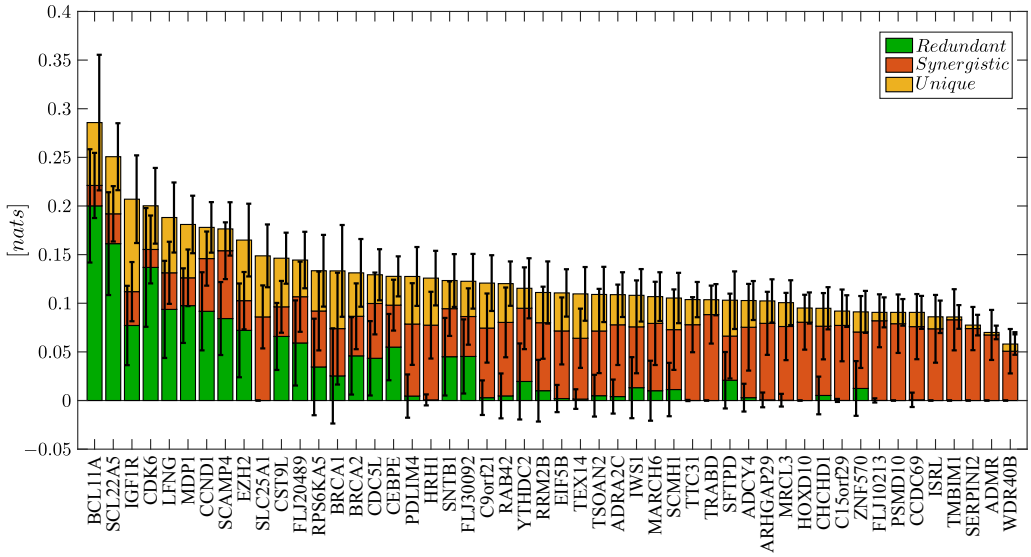


Figure 7.27: TCGA-BRCA feature analysis results. Barplots represent the contributions obtained with the kNN-based estimation approach for each gene, shown according to a decreasing CMI. The $\pm 2std$ confidence interval is shown for the unique, synergistic, and redundant contribution for each gene. This figure has been adapted from [131].

most of these genes exhibit high CMI values based on their mean contributions. The findings, based on the selected set of features, are consistent with previous, similar works in [43, 53, 104], that identify the top contributing set of genes. However, as also noted by Catav *et al.* [43], we observe substantial contributions from MDP1, SCAMP4, SLC25A1, and CST9L genes not previously linked to breast cancer in the literature. SLC25A1, however, has been reported by Catalina-Rodriguez *et al.* [42] to exhibit elevated expression in breast cancer tissues, suggesting a potential relevance to subtype differentiation. Albeit not comparable in this setting, we note the low influence of BRCA1, BRCA2, and TEX14 despite being one of the prognostic markers.

Considering the 10 selected genes, the proposed decomposition allows for the observation of component-wise interactions, noting that the majority of the significant contribution comes from the redundancy components. A similar study by Westphal *et al.* [245] performs a PID-like analysis to determine synergistic and redundant contributions in addition to the MI that each feature has with the target, $I(Y; X_i)$. However, while their results indicate lower contribution of synergy and redundancy, we note that the nature of these components is different. Specifically, they define feature-wise synergy and redundancy (FWS and FWR) through the concept of interaction information. To obtain FWS they maximize $I(Y; X_i; \boldsymbol{Z}^{ms})$, where $\boldsymbol{Z}^{ms}$ is the subset of features for maximum synergy, while the FWR is determined as FWS $- I(Y; X_i; \boldsymbol{Z}_{\setminus X_i})$, where $\boldsymbol{Z}_{\setminus X_i}$ is the subset of features without X_i.

For our approach, the redundancy measure for a given gene indicates shared information

when the inclusion of another gene results in a reduced CMI. Analyzing the selection order (not shown in the work), we observe shared redundancy between BCL11A and SLC22A5, which is expected given their subtype-specific roles as BCL11A is associated with the basal-like phenotype [125], whereas SLC22A5 is linked to luminal phenotypes and is known to be an estrogen receptor (ER) regulated gene [165, 243]. Furthermore, LFNG and BCL11A exhibit redundant information, likely due to their shared association with the basal-like subtype [250].

Conclusion. The application of the CMI-based Hi-Fi framework demonstrated the feasibility of decomposing feature relevance into unique, redundant, and synergistic informational components in real biomedical data. Using a k-nearest neighbour estimator, the method provided a model-free quantification of high-order dependencies, offering an interpretable view of how individual and joint features contribute to classification performance. Analyses performed on the breast cancer gene expression dataset confirmed the capacity of the estimator to capture meaningful interaction patterns consistent with prior biological evidence. Compared with conventional feature relevance measures, the Hi-Fi approach revealed complementary insights by identifying synergistic effects that cannot be explained through marginal associations alone. Overall, this application highlights the potential of the proposed information-theoretic framework as a diagnostic tool for interpreting complex feature interactions in biomedical classification tasks. It demonstrates that a data-efficient CMI estimation can provide physiologically and statistically interpretable markers of relevance, paving the way for future applications in feature selection and model interpretability.

Table 7.15: Overview of the studies included in this thesis, summarizing the physiological applications, datasets, analysed signals, extracted features, and the machine learning or information-theoretic methods adopted across the different chapters.

Application	Topic	Dataset	Signals (time series extracted)	Computed Features	Applied Methods and Models
Stress Characterization and Classification					
7.2.1, [94]	Classification	D1	ECG (RR), PPG (PP)	MEAN, SDNN, RMSSD, LF, HF, LFn, HFn, fHF, SVB, H, CE, SE	LDA, SVM, NN, kNN, mRMR
7.2.2, [95]	Classification	D1	ECG (RR), PPG (PP)	MEAN, SDNN, RMSSD, LF, HF, LFn, HFn, fHF, SVB, H, CE, SE	NB, SVM, NN (UST HRV/PRV)
7.2.3, [97]	Feature Selection Classification	D1	ECG (RR), PPG (PP)	MEAN, SDNN, RMSSD, LF, HF, LFn, HFn, fHF, SVB, H, CE, SE	AIC-FS, LDA, SVM, kNN, RF
7.2.4, [130]	Classification	D1	ECG (RR), PPG (PP)	MeanNN, RMSSD, HF, CE	LIC, kNN, GB
7.2.5, [206]	Classification	D1	ECG (RR), AP (SBP), BREATH (RESP)	MEAN, STD, TE, cSE, MIR	LIC
7.2.6, [99]	Feature Importance Classification	D1	ECG (RR), AP (SBP, DBP), BREATH (RESP)	MEAN, STD, VAR, RMSSD, pNN50, CV, VLF, HF, LF, HFn, LFn, fHF, TP, CE, SE	CMIHiFi
7.2.7, [74]	Feature Extraction Classification	Driving Dataset	ECG (RR), BREATH (RESP)	MEAN, STD, RMSSD, SDNN, pNN50, NNN50	Inter-subject normalization, Relief, SVM
7.2.8	Feature Extraction Statistical Analysis	D3	ECG (RR), PPG (PP)	MEAN, SDNN, RMSSD, LF, HF, LFn, HFn, fHF, SVB, H, CE, SE	–
7.2.9, [17]	Feature Importance	D1	ECG (RR), AP (SBP), BREATH (RESP)	–	PID

Table 7.15: Overview of the studies included in this thesis, summarizing the physiological applications, datasets, analysed signals, extracted features, and the machine learning or information-theoretic methods adopted across the different chapters. (Continued)

Application	Topic	Dataset	Signals (time series extracted)	Computed Features	Applied Methods and Models
Sex-related Differences in Cardiovascular Regulation					
7.3.1, [178]	Feature Extraction Statistical Analysis	D2	ECG (RR), AP (SBP, DBP)	ER, MIR	–
7.3.2, [96]	Feature Selection	D2 + D1	ECG (RR), AP (SBP, DBP)	MEAN, STD, VAR, RMSSD, pNN50, CV, VLF, HF, LF, HFn, LFn, fHF, TP, CE, SE	kCMI-FS, LDA
7.3.3, [164]	Feature Importance Classification	D2 + D1	ECG (RR), AP (SBP, DBP)	MEAN, STD, VAR, RMSSD, pNN50, CV, VLF, HF, LF, HFn, LFn, fHF, TP, CE, SE	HiFi
Further Physiological Applications					
7.4.1, [98]	Feature Selection	Parkinson, Breast Cancer, TCGA-BRCA, D1	ECG (RR), PPG (PP) (for D1)	MEAN, STD, VAR, RMSSD, pNN50, CV, VLF, HF, LF, HFn, LFn, fHF, TP, CE, SE	kCMI-FS, SVM
7.4.2, [131]	Feature Importance	TCGA-BRCA	–	–	CMIHiFi

Table 7.15 provides an overview of the studies included in this thesis, summarizing the considered physiological applications, datasets, analysed signals, extracted features, and the machine learning and information-theoretic methods adopted in each study.

The table highlights both the common methodological perspective adopted throughout the thesis and the application-specific adaptations required by different experimental contexts. While the analysed signals, extracted descriptors, and modelling strategies vary depending on the physiological problem under investigation, the studies share a common analytical approach based on the quantitative analysis of physiological time series through signal processing, feature extraction, and information-driven modelling.

Beyond summarising the experimental settings of each study, the table also illustrates how the different applications relate to the methodological development presented in this thesis. Each case study explores specific components of the analytical workflow—such as signal processing, feature extraction, feature selection, feature importance analysis, or classification—depending on the objectives of the individual study. Taken together, these applications contribute to the progressive development, validation, and refinement of the overall analytical framework proposed for the analysis of physiological signals.

Overall, the applications presented in this chapter demonstrate the versatility, interpretability, and translational potential of the analytical approaches developed throughout this thesis. By combining statistical, information-theoretic, and machine learning methodologies, the proposed framework enables the quantitative characterisation of autonomic and physiological variability across multiple contexts, from stress and sex-related differences to multimodal cardiovascular and network-level analyses.

From a methodological standpoint, the chapter also integrates the various analytical components developed earlier in this thesis. The results on *time series analysis and feature extraction* focus on the derivation and validation of cardiovascular features from heart rate and pulse rate variability, including studies comparing measurement systems and exploring physiological factors such as gender-related differences. The subsequent analyses on *feature selection* apply both traditional and information-theoretic approaches, ranging from physiologically driven methods to conditional mutual information and partial information decomposition to identify the most informative and non-redundant features contributing to classification. The part on *feature importance* complements feature selection by quantifying the relative contribution of individual features or feature groups to model performance, thereby enhancing interpretability and physiological insight into cardiovascular regulation. Finally, the *classification* analyses apply machine learning algorithms to a variety of physiological contexts, including stress discrimination and gender classification. Through this structure, the chapter demonstrates the versatility and robustness of the

proposed computational framework, bridging theoretical development with experimental validation and providing physiologically interpretable insights into autonomic regulation.

These studies collectively highlight how physiological variability, when examined through principled computational tools, provides a powerful window into autonomic regulation and its adaptive mechanisms. The results confirm that the same analytical principles can be effectively applied across heterogeneous datasets and experimental paradigms, supporting both mechanistic understanding and the development of data-driven biomarkers for personalised and non-invasive physiological assessment.

Part IV

Discussion and Conclusions

8

Discussion

This chapter provides an integrated interpretation of the findings presented throughout this thesis, examining how the theoretical and methodological contributions support the experimental and physiological results observed across multiple applications. The discussion is organised around the main axes of the thesis: methodological advances in cardiovascular variability analysis and information-theoretic modelling (Sec. 8.1), their relevance for physiological and clinical interpretation, and the implications that emerge from the assessment of stress-induced or sex-related differences in autonomic control, in Sec. 8.2.1 and 8.2.2 respectively. The chapter concludes by outlining potential directions for future research (Sec. 8.3).

8.1 Theoretical and Methodological Advances

The theoretical core of this thesis focuses on the development and evaluation of analytical tools for empowering the pipeline of machine learning in the context of classification and analysis of physiological states. From a methodological perspective, the work integrates concepts from physiological signal analysis, information theory, and machine learning to build a coherent analytical framework for the study of cardiovascular dynamics. Within this framework, the machine learning pipeline is interpreted not simply as a sequence of computational operations, but as a structured process for extracting, filtering, and organising the information contained in physiological signals. The methodological contributions of the thesis span the main stages of this pipeline, including signal processing, feature extraction (FE), feature selection (FS), feature importance analysis (FI), and classification. Adopting this perspective makes it possible to analyse cardiovascular

variability not only through individual descriptors, but also through the informational relationships that exist among features and across physiological subsystems. The following sections discuss how the methodological choices adopted at each stage of the pipeline contribute to building an information-driven framework for the analysis of physiological data.

Before addressing FE and model construction, it is important to consider the signal-processing stage, which represents the first step of the analytical pipeline adopted throughout this thesis (Chapter 2). Raw physiological recordings must be transformed into reliable beat-to-beat time series before variability descriptors can be computed. This transformation involves preprocessing procedures such as filtering, artefact detection and correction, fiducial point identification, and temporal alignment of physiological signals. These operations ensure that the derived variability series reflects the underlying physiological dynamics rather than measurement noise or signal distortions. Additional methodological choices concern segmentation and normalization, which influence how the physiological dynamics are represented in the analysed time series. In some analyses (Section 7.2.2), segmentation was performed using a fixed number of heartbeats rather than fixed temporal durations. This approach ensures that each analysed segment contains the same number of samples in the beat-to-beat domain, facilitating the computation of variability descriptors, although it introduces small variations in the temporal duration of the segments depending on the subject's heart rate. Another relevant aspect concerns normalization when analysing heterogeneous datasets (Section 7.2.7). Physiological recordings from different individuals often present substantial baseline variability that may affect the magnitude of derived descriptors and influence model generalisation. In the considered applications, signal-level normalization proved effective in reducing inter-subject variability and improving the stability of the subsequent analytical stages. Overall, these preprocessing choices influence how the variability structure of the signals is represented and therefore determine the informational content available for the subsequent feature extraction and modelling stages.

At the FE stage (Chapter 3), the processed time series are transformed into quantitative descriptors capturing different aspects of signal variability. The results of this thesis show that informative representations of the data can be obtained both from relatively short recordings [96, 97, 99, 206] and from ultra-short recordings [95], provided that the selected features are sensitive to the relevant temporal dynamics of the signals. One important methodological observation concerns the robustness of different feature domains with respect to recording length and acquisition conditions. Frequency-domain descriptors, while widely used in variability analysis, tend to be more sensitive to window length and preprocessing choices [95]. In contrast, time-domain and information-theoretic descrip-

tors generally exhibit greater stability when shorter segments are analysed [96, 97, 99]. This behaviour suggests that different feature domains capture complementary aspects of signal dynamics across multiple temporal scales. Within this framework, the set of extracted features was not identical across all studies included in the thesis. Although a common pool of descriptors was available, including time-domain, frequency-domain, and information-theoretic indices, different subsets were considered depending on the structure of the dataset and the specific methodological objective of each study. This adaptive strategy allowed the representation of the data to be tailored to the characteristics of the analysed signals while maintaining methodological consistency across the pipeline.

A central methodological contribution of the thesis concerns FS, where information-theoretic criteria enabled results that were not consistently observable with classical approaches (Chapter 4). Physiological datasets often contain multiple descriptors derived from the same signals (i.e., SBP and DBP), resulting in high levels of correlation and redundancy among features. In such conditions, selecting informative yet complementary descriptors becomes essential to obtain compact representations of the data, improve model generalization, and facilitate interpretability of the resulting models. Traditional FS (Section 4.1) strategies based on Mutual Information typically rely on pairwise relevance measures or heuristic redundancy penalties. Although these approaches provide computationally efficient solutions, they may fail to capture conditional and multivariate dependencies among features, particularly in settings where informative descriptors emerge only once the informational contribution of other variables has been accounted for. To address these limitations, within the context of this thesis, the kCMI-FS (Section 4.2.1, [96, 98]) algorithm was developed and implemented as part of the proposed analytical pipeline, providing a principled information-theoretic framework for FS in classification problems involving continuous physiological descriptors and discrete target variables. It is an information-theoretic FS algorithm based on Conditional Mutual Information and non-parametric kNN estimation [17], designed to quantify the additional information contributed by each candidate feature once the information carried by previously selected features has been accounted for. By evaluating the information contributed by a candidate feature with respect to the target variable while conditioning on already selected features, the method explicitly accounts for redundancy and conditional dependencies within the feature space. This conditional formulation enables the algorithm to identify descriptors that provide additional information beyond what is already explained by the current feature subset. Another key element of the proposed method is the use of a significance-driven forward selection strategy based on surrogate testing. At each iteration, candidate features are included in the selected subset only if their conditional information with respect to the target variable is statistically significant. This procedure reduces the risk of includ-

ing spurious predictors and contributes to the stability of the resulting feature subsets. The combination of conditional information measures, non-parametric estimation, and significance-based selection provides a principled and model-independent strategy for FS in complex datasets. Within the analytical pipeline adopted in this thesis, the kCMI-FS framework therefore plays a crucial role in identifying compact yet informative representations of the data, enabling subsequent modelling stages to operate on feature subsets that preserve the relevant informational structure of the physiological signals. While FS identifies informative descriptors, the subsequent FI analysis further investigates how predictive information is distributed among the selected features and how different variables jointly contribute to the classification task.

Feature importance analysis was addressed through an information-theoretic framework aimed at characterizing how predictive information is distributed across the feature space (Chapter 5). Conventional importance measures commonly used in machine learning, such as model coefficients, permutation scores, or impurity-based rankings, typically quantify the relevance of individual predictors in isolation (Section 5.1). While these approaches are effective for identifying dominant variables, they often fail to capture the dependency structures that arise when predictive information is distributed across multiple interacting features. In many datasets, informative patterns do not originate from single variables alone but emerge from combinations of predictors whose joint contribution cannot be reduced to independent effects. Standard FI measures are generally unable to disentangle these situations, as they produce a single scalar relevance score that conflates different informational mechanisms. To overcome this limitation, this thesis adopts a high-order information-theoretic feature importance framework (Section 5.2). In particular, the CMIHiFi method, based on the HiFi framework (Section 5.1.2, [163, 164]), was implemented and integrated into the proposed analytical pipeline to enable the practical estimation of high-order feature importance in classification problems involving continuous physiological descriptors and discrete target variables (Section 5.2.1, [99, 131]). These approaches quantify the contribution of each feature by decomposing the information shared between the feature set and the target variable into distinct components, including unique, redundant, and synergistic information. This decomposition allows the analysis to distinguish between predictors that independently convey information about the target, those that encode overlapping information already contained in other variables, and those that contribute only through cooperative interactions with other features. From a methodological perspective, this representation provides a more detailed characterization of the informational structure of the dataset compared with traditional importance rankings. By explicitly separating different modes of information contribution, the framework enables a deeper understanding of how predictive information is organised across the feature space

and how complex dependencies among variables influence the modelling process.

The final stage of the analytical pipeline concerns classification, where the selected features are used to construct predictive models for the considered tasks. In this thesis, multiple classifiers were employed across different studies, including linear models (LDA), distance-based approaches (kNN), kernel methods (SVM), ensemble models (RF) (Section 6.1, [97]), and information-driven classifiers such as LIC (Section 6.2, [130, 206]). The use of different classifiers was not intended to identify a single optimal predictive model, but rather to evaluate the robustness of the proposed methodological framework across different modelling paradigms. By testing classifiers based on distinct learning principles, it becomes possible to assess whether the predictive information extracted through the preceding stages of the pipeline can be consistently exploited by different types of models. From a methodological perspective, the results indicate that classification performance is strongly influenced by the quality and informational structure of the extracted features rather than by the complexity of the classifier itself. Once redundancy among features is properly addressed and informative representations of the data are obtained, comparable predictive performance can often be achieved using models of varying complexity [95, 96, 97, 130, 206]. Within the context of this thesis, the LIC classifier was implemented and integrated into the analytical pipeline in order to evaluate the interaction between information-theoretic feature representations and information-driven classification strategies. The LIC framework offers several advantages in the analysis of physiological datasets, including robustness in small-sample regimes, adaptability to non-Gaussian distributions, and transparency in the interpretation of decision boundaries. These characteristics make it particularly suitable for applications where interpretability and statistical robustness are as important as predictive accuracy. Overall, the classification analyses performed in this thesis highlight how the effectiveness of predictive modelling in physiological data analysis depends primarily on the representation and organisation of information within the feature space. The role of the classifier, therefore, becomes coherent with the preceding stages of signal processing, feature extraction, and information-theoretic analysis, which collectively determine the quality of the predictive representation of the data.

Taken together, the methodological approaches described above highlight the advantages of adopting a unified information-theoretic perspective across the different stages of the machine learning pipeline. By integrating signal processing, feature extraction, feature selection, feature importance analysis, and classification within a coherent analytical framework, the overall framework designed in the thesis enables the systematic identification, organisation, and interpretation of the information contained in physiological datasets. Within this framework, the development of the kCMI-FS algorithm, together

with the implementation of the CMIHiFi framework and the LIC classifier, provides the building blocks for analysing complex datasets in the context of ML from an information-driven perspective. These methods were also evaluated using simulated datasets in which key parameters such as sample size, feature dimensionality, redundancy structure, and noise levels were systematically varied. This controlled validation allowed the robustness and behaviour of the algorithms to be assessed across different data regimes, demonstrating their applicability to datasets with different sizes and structural characteristics.

8.2 Physiological and Experimental Findings

The experimental analyses presented in this thesis aim primarily to demonstrate how the proposed methodological framework can be used to identify physiologically meaningful cardiovascular markers across different experimental scenarios. Within this perspective, ML steps and, in particular, the classification task, are not treated as an end in themselves nor as a benchmark for achieving maximal predictive performance. Instead, they serve as an operational tool for evaluating on real physiological datasets the capability of the proposed information-theoretic framework for ML in differentiating physiological conditions or subject groups.

Across the different applications considered in this thesis, the results consistently show that the proposed approach highlights cardiovascular features reflecting known physiological mechanisms. For instance, in the analysis of sex-related differences (Section 7.3), the objective is not to determine the sex of a subject per se, but to identify cardiovascular indices that capture underlying autonomic differences between males and females. Similarly, in the study of stress-related conditions (Section 7.2), the focus is on identifying the physiological markers that contribute to the observed differentiation between experimental states.

Importantly, many of the observed patterns are consistent with findings previously reported in the literature. However, in this thesis, they emerge through a different analytical perspective based on information-theoretic feature analysis. This aspect is particularly relevant, as it shows that the proposed framework is capable of recovering physiologically meaningful patterns while providing additional insight into how predictive information is distributed across cardiovascular descriptors.

8.2.1 Discussion on Application to Stress Characterization and Classification

Across the analyses addressing stress assessment through cardiovascular variability signals, a consistent physiological pattern emerged: orthostatic stress (HUT) was sys-

tematically more distinguishable from resting conditions than mental stress (MA). This observation is fully consistent with well-established physiological knowledge on autonomic responses to orthostatic and cognitive stress. The present analyses therefore do not discover a new physiological phenomenon, but rather confirm these patterns across multiple datasets and analytical pipelines (Sections 7.2.1, 7.2.2, 7.2.4, 7.2.6, and 7.2.3). Importantly, the contribution of the proposed framework lies in identifying which descriptors encode these responses and how their informational relevance emerges across different analytical stages.

Across these analyses, a subset of cardiovascular variability descriptors repeatedly emerged as particularly informative for distinguishing the experimental conditions. These features were consistently identified through mutual-information-based feature selection (mRMR, AIC-FS, kCMI-FS, Sections 4.1.1, 4.1.2, and 4.2.1 respectively) and subsequently confirmed through feature importance analyses within the classification models (Sections 7.2.1, 7.2.2, 7.2.4, and 7.2.3). In particular, indices associated with vagal modulation and autonomic balance, such as RMSSD and the spectral components in the LF and HF bands, were repeatedly selected during the feature selection stage and assigned high importance scores in the trained models. The convergence between feature selection and feature importance analyses indicates that these descriptors capture stable physiological patterns underlying autonomic responses to stress.

From a physiological perspective, the higher separability of HUT from resting conditions can be explained by the regulatory mechanisms underlying orthostatic perturbation. Head-up tilt stress induces a rapid autonomic reconfiguration driven by baroreflex unloading, leading to increased sympathetic activation and reduced parasympathetic modulation [97, 175, 232]. These adjustments produce coordinated changes across multiple cardiovascular variability domains, including HRV, PRV, BPV, and indices reflecting cardiorespiratory interactions (Sections 7.2.1, 7.2.2, 7.2.4, and 7.2.3). At the feature level, postural stress was consistently associated with reductions in vagally mediated time-domain indices such as RMSSD, increases in MEAN heart rate, and marked modulation of low-frequency oscillatory components linked to baroreflex regulation [212]. The coherence of these responses generates well-defined physiological patterns that facilitate the discrimination of orthostatic stress from baseline conditions ([94, 95, 97, 206]).

Consistently, with this interpretation, HRV-related measures associated with short-term vagal regulation showed systematic variations across experimental conditions. In particular, descriptors such as RMSSD and spectral indices in the LF and HF bands, as well as the LF/HF ratio, captured the shift toward sympathetic predominance and vagal withdrawal that accompanies orthostatic stress [97]. These findings are in agreement with extensive literature on autonomic responses to postural perturbations [30, 143, 212, 232].

The value of the present analyses, therefore, lies not in identifying previously unknown physiological markers but in demonstrating how these well-known descriptors repeatedly emerge as informative features across independent datasets and analytical pipelines.

In contrast, mental stress produced more heterogeneous cardiovascular responses across subjects. Although cognitive tasks such as mental arithmetic are known to increase sympathetic activity, the resulting autonomic adjustments tend to be subtler and more variable compared to those induced by postural perturbations. As observed in several analyses (see Sections 7.2.1, 7.2.2, 7.2.4, and 7.2.3), this variability reduces the consistency of the corresponding cardiovascular signatures, making mental stress more difficult to discriminate from resting conditions [97, 206]. Accordingly, no single descriptor emerged as a robust and reproducible marker of MA across datasets. Instead, mutual-information-based feature selection revealed that discriminative information was distributed across multiple descriptors and domains, highlighting the relevance of multivariate patterns rather than isolated physiological markers.

This difference was also reflected in classification performance. Across datasets and validation schemes, HUT versus rest consistently achieved high accuracy and specificity, often exceeding 85% when multivariate feature sets were employed (see Sections 7.2.1, 7.2.2, 7.2.4, and 7.2.3). In contrast, discrimination between MA and rest was generally lower and more variable, typically remaining in the range of 60–70% accuracy across studies. This reduced separability is physiologically coherent with the weaker and more heterogeneous autonomic responses elicited by cognitive stress.

Additional insight into stress-related physiological mechanisms was obtained through information-theoretic analyses of cardiovascular signals. Entropy-based metrics, as discussed in Section 7.2.6, quantify the irregularity, predictability, and dynamical complexity of cardiovascular variability, thereby providing an indirect measure of the adaptability of autonomic regulation [66, 99]. While classical HRV descriptors quantify the amplitude or spectral distribution of oscillations, entropy-based measures estimate the informational content and temporal organization of the underlying dynamics [66]. In this framework, self-entropy reflects the portion of signal variability explained by its internal temporal structure, whereas conditional entropy quantifies the amount of new information generated by the system that cannot be predicted from its past. These descriptors, therefore, provide complementary insight into how autonomic regulation organizes cardiovascular dynamics under stress conditions.

The relevance of information-domain features was also highlighted by the FI procedure, where entropy-based descriptors were selected together with traditional HRV indices (Section 7.2.6). While time- and frequency-domain indices often emerged as the most

discriminative features, reflecting the dominant role of autonomic modulation amplitude and spectral balance in stress responses, entropy-based measures provided additional complementary information. In particular, these descriptors capture aspects of dynamical organization and predictability that are not directly reflected in conventional variability indices. The joint selection of classical HRV features and entropy-based metrics, therefore, suggests that stress-induced physiological changes are distributed across multiple aspects of cardiovascular dynamics rather than being confined to a single descriptor.

A consistent observation across studies concerned the robustness of time-domain indices such as mean RR and RMSSD even when computed from ultra-short-term (UST) segments (Section 7.2.2, [95]). These features preserved discriminative capability across datasets and stress paradigms, supporting their suitability for real-time and wearable monitoring applications. In contrast, frequency-domain indices were more sensitive to window length, spectral estimation choices, and respiratory variability, reflecting the stronger dependence of LF and HF estimates on signal duration and spectral resolution.

The physiological patterns observed in controlled laboratory protocols were further supported by analyses conducted in other contexts related to stress. In the study on work-related stress (Section 7.2.8), cardiovascular variability descriptors revealed modulations consistent with increased sympathetic activation, reduced vagal activity, altered autonomic complexity, and reduced regulatory flexibility during demanding working conditions. Notably, work-related stress was not characterized by large shifts in individual indices alone, but rather by coordinated changes across time-domain and information-domain features, suggesting that baseline autonomic organization may carry relevant information about chronic stress exposure and recovery processes.

The application to a dataset consisting of signals acquired from drivers provided an additional opportunity to evaluate the proposed analytical framework in a realistic operational context. In the studies described in Section 7.2.7 and [74], cardiovascular variability analysis was applied to driving scenarios in order to investigate physiological responses associated with cognitive workload and situational stress during vehicle operation. In this setting, the driving task introduces a combination of cognitive, perceptual, and motor demands that can modulate autonomic regulation. The analyses showed that variations in HRV descriptors reflected changes in attentional load and task complexity, with patterns consistent with increased sympathetic activation and reduced vagal modulation during more demanding driving conditions. Although the magnitude of these changes was generally smaller than that observed during controlled orthostatic stress protocols, the results demonstrated that meaningful autonomic information can still be extracted from cardiovascular signals in real-world scenarios.

Finally, although different classifiers were adopted across the analysed studies, the general physiological patterns remained remarkably consistent (see Sections 7.2.1, 7.2.2, 7.2.4, 7.2.3, and 7.2.7). This observation is relevant from a methodological perspective, as the objective of the proposed framework was not to identify a single optimal predictive model but to verify whether the physiological information encoded by the extracted features remained stable across modelling strategies. The observed consistency, therefore, supports the interpretation that the discriminative structure primarily reflects physiological signal organization rather than classifier-specific effects.

Overall, these findings confirm well-established physiological knowledge on autonomic regulation under stress while demonstrating how these responses are systematically encoded in cardiovascular variability descriptors. The novelty of the present work, therefore, lies not in identifying previously unknown physiological phenomena, but in showing how a structured analytical framework—combining feature extraction, feature selection, feature importance analysis, and machine-learning classification—can reveal the informational structure through which autonomic responses to stress are reflected in cardiovascular signals.

From a broader perspective, the analyses presented in this section support the view that stress-related physiological information is distributed across interacting physiological systems rather than being confined to individual signals or isolated descriptors. By combining cardiovascular variability analysis with multivariate and information-theoretic approaches, the proposed framework enables the identification of coordinated patterns of autonomic regulation across different stress conditions and experimental contexts.

8.2.2 Discussion on Application to Sex-related Differences in Cardiovascular Regulation

Sex-related differences in cardiovascular autonomic regulation have been widely documented in the physiological and biomedical literature. Numerous studies have reported systematic differences between females and males in terms of sympathovagal balance, baroreflex sensitivity, vascular stiffness, and cardiovascular control strategies [119, 129, 198, 224]. In particular, females are generally characterized by higher parasympathetic modulation at rest and lower sympathetic dominance compared to males [119], whereas males tend to exhibit greater vascular stiffness and a relatively stronger sympathetic influence. These differences are thought to arise from a combination of hormonal effects, structural and mechanical properties of the cardiovascular system, and sex-dependent autonomic integration mechanisms [129, 198, 224]. Understanding how these physiological differences are reflected in cardiovascular variability signals is therefore relevant for both basic physiology and the interpretation of variability-based biomarkers in clinical

and experimental studies.

Within this framework, the analyses presented in this thesis explored whether the proposed analytical pipeline could capture differences in the informational structure of cardiovascular variability associated with biological sex. Importantly, the analysed datasets were not originally designed with the specific objective of investigating sex-related autonomic differences. Sex, therefore, represents an additional dimension of variability within the data rather than the primary experimental factor. For this reason, the results should be interpreted cautiously and primarily as exploratory observations. Nevertheless, consistent patterns were observed across several datasets and analytical pipelines (Sections 7.3.2, 7.3.1, and 7.3.3), suggesting that the framework is able to reveal meaningful differences in the organization of cardiovascular variability between females and males.

Widely known findings obtained from traditional HRV analyses were largely confirmed. In particular, females tended to exhibit higher HF-related indices and lower LF/HF ratios compared to males, consistent with a predominance of parasympathetic modulation under resting conditions [96, 119, 164]. These observations are consistent with a large body of literature reporting stronger vagal activity and enhanced baroreflex control in females. The present analyses thus do not identify previously unknown physiological mechanisms, but rather confirm that these well-established patterns can be consistently detected across independent datasets and analytical pipelines.

However, when extending the analysis beyond traditional HRV descriptors to include BPV, the results became more heterogeneous. This behaviour reflects the multi-layered nature of cardiovascular regulation. While HRV mainly reflects cardiac autonomic modulation, BPV is influenced by additional mechanisms such as vascular tone, baroreflex feedback loops, and mechanical properties of the arterial system. Sex-related differences may therefore emerge across several interacting control systems, making them difficult to capture through isolated descriptors alone [129, 198, 224].

In this context, feature selection played an important role in identifying which descriptors carry the most relevant information about sex-related differences. Mutual-information-based feature selection consistently indicated that discrimination between females and males was rarely driven by a single dominant marker. Instead, informative patterns emerged from combinations of features spanning different physiological domains. In particular, subsets combining RR-derived and BP-derived descriptors were frequently selected across datasets (Sections 7.3.2, and 7.3.3), suggesting that sex-related autonomic differences are encoded in the coordinated behaviour of cardiac and vascular regulation rather than in isolated variability measures [164].

Feature importance analyses (Section 7.3.3, [164]) further supported this interpreta-

tion by quantifying the effective contribution of each descriptor within the classification models. The convergence between feature selection and feature importance results allowed classification outcomes to be interpreted in physiological terms. In particular, time-domain HRV descriptors such as MEAN RR, STD, RMSSD, and pNN50 frequently appeared among the most relevant features, often in combination with BP variability descriptors such as STD and total power. These patterns indicate that sex differences are expressed through coordinated modulation of variability across cardiac and vascular loops rather than through isolated changes in a single autonomic branch [96, 164].

Additional insight was obtained through the information-theoretic analyses presented in Section 7.3.3. By decomposing feature relevance into unique, redundant, and synergistic contributions, these analyses revealed that synergistic components dominated the information structure across datasets. This result suggests that a substantial fraction of sex-related information emerges only when features are considered jointly. In other words, the physiological differences between females and males are not fully captured by individual variability indices, but rather by interaction patterns linking multiple cardiovascular signals, as known by the literature [119, 129, 199, 224].

From a physiological perspective, these synergistic structures are consistent with known sex-related differences in autonomic integration. Greater vagal modulation and enhanced baroreflex sensitivity in females may promote tighter coordination between heart rate and blood pressure dynamics, resulting in more stable interaction patterns between cardiac and vascular signals. Conversely, the relatively higher sympathetic tone and vascular stiffness typically observed in males may lead to more fragmented or less synergistic autonomic signatures.

In terms of classification performance, accuracy values remained moderate, typically around 70% across independent datasets and validation schemes (Sections 7.3.2, and 7.3.3). Recall and F1-score were consistently higher for females than for males, suggesting more stable and homogeneous cardiovascular patterns in the female group. Importantly, these performance levels should not be interpreted as an attempt to achieve accurate sex identification, but rather as an analytical tool to investigate how biological sex shapes the informational architecture of cardiovascular variability.

Overall, the results presented in this thesis confirm that sex-related differences in autonomic regulation are reflected in cardiovascular variability signals, in agreement with previously reported physiological observations. The novelty of the present work thus lies not in identifying new physiological differences, but in demonstrating how these differences can be systematically revealed through a structured analytical framework combining feature extraction, feature selection, feature importance analysis, and information-theoretic

interpretation. This approach highlights how sex-related autonomic differences emerge from multivariate interaction patterns spanning cardiac and vascular domains.

8.2.3 Discussion on Other Applications

The additional applications considered in this thesis extend the evaluation of the proposed information-theoretic framework beyond cardiovascular physiology and demonstrate its applicability across heterogeneous data domains. In particular, the kCMI-FS algorithm and the CMIHiFi framework were tested on benchmark datasets and large-scale biomedical data, including Parkinson's disease, breast cancer, and gene expression datasets (Sections 7.4.1and 7.4.2).

These analyses served primarily as methodological validation studies, aimed at assessing whether the proposed information-theoretic approaches remain effective when applied to high-dimensional and non-physiological data. Across these datasets, the methods consistently identified compact subsets of informative features and revealed structured patterns of redundancy and synergy among predictors. This behaviour confirms that the principles underlying conditional information–based FS and high-order FI are not restricted to cardiovascular variability analysis but apply more broadly to classification problems involving complex and correlated feature spaces.

The results obtained in these datasets therefore support the generality of the proposed framework. While the physiological applications discussed in previous sections provide insight into autonomic regulation and cardiovascular interactions, the analyses performed on heterogeneous biomedical datasets highlight the methodological robustness and scalability of the developed algorithms. Together, these findings suggest that the information-theoretic tools introduced in this thesis may be applicable to a wide range of biomedical data analysis problems where interpretability and multivariate dependency structure play a central role.

High-dimensional FI analyses conducted on large cohorts provided additional evidence of the scalability and interpretability of the proposed framework. In these settings, multidomain feature sets comprising time-, frequency-, and information-domain descriptors were systematically ranked and decomposed, thereby enabling identification of population-wide patterns of relevance, redundancy, and synergy. Such analyses demonstrated that information-theoretic FI can be applied not only to small experimental datasets but also to large-scale studies, where interpretability and robustness are critical for extracting physiologically meaningful insights.

Overall, these diverse applications confirm the generalizability of the methodological

framework developed in this thesis. By unifying FE, FS, FI, and classification under a common information-theoretic rationale, the framework proves adaptable to different signal modalities, experimental designs, and levels of complexity. This versatility supports its potential impact across computational physiology, clinical research, and human–machine integration, particularly in contexts where interpretability, robustness, and multivariate interaction analysis are essential.

8.3 Directions for Future Research

The research presented in this thesis primarily focused on the development of an information-theoretic machine learning pipeline for the analysis of physiological data. The experimental applications considered throughout the work were mainly intended to provide an empirical validation of the proposed methodological framework rather than to optimise predictive performance for specific physiological problems.

In this perspective, future research may focus on extending the application of the proposed framework to specific physiological and clinical scenarios, with the goal of maximising predictive performance and practical utility. In particular, further developments are desirable to consolidate and expand the experimental validation of the framework in stress-related applications, including the classification of different stress conditions and the analysis of stress responses across heterogeneous populations and experimental settings.

The research directions outlined below are therefore primarily oriented toward expanding and refining the applicative aspects of the proposed framework, while maintaining the methodological foundations developed in this thesis.

- **Integration of multimodal physiological signals.** Future studies should incorporate additional physiological channels, such as respiration, electrodermal activity, and electromyography, which are known to contain relevant information about stress-related autonomic and behavioural responses. Multimodal integration would enable a more comprehensive characterization of physiological regulation and allow the extraction of additional feature types, including causal descriptors and multivariate or high-order interaction measures, that may further improve the detection and interpretation of stress-related physiological states.

- **Integrated modelling of the cardiorespiratory system.** An important extension of the present work concerns the joint analysis of cardiac, vascular, and respiratory dynamics. Combining heart rate variability, blood pressure variability, and respiratory signals would allow a more complete characterization of autonomic regulation

and cardiorespiratory coupling. Information-theoretic approaches could play a central role in this context, enabling the quantification of multivariate interactions and coordinated regulatory mechanisms across physiological subsystems.

- **Validation on larger and heterogeneous datasets.** Although the methods proposed in this thesis were evaluated across multiple datasets and experimental scenarios, further validation on larger cohorts and multi-centre datasets would allow a more comprehensive assessment of their robustness and generalisability. In particular, future studies could investigate how information-theoretic feature selection and feature importance frameworks behave in the presence of stronger inter-subject variability, heterogeneous acquisition protocols, or partially missing physiological signals.

- **Real-world and longitudinal monitoring.** Most analyses presented in this thesis relied on short-term recordings collected in controlled experimental settings. Extending these approaches to long-term and ambulatory monitoring would allow investigation of circadian rhythms, behavioural influences, and intra-individual variability, which are essential for real-world physiological assessment and wearable health monitoring applications.

- **Standardization of HRV windowing strategies.** Future investigations could systematically compare heartbeat-based and time-based segmentation approaches for ultra-short-term HRV analysis. Such studies would help clarify how window definition influences the stability and physiological interpretation of variability indices, particularly in wearable and real-time monitoring scenarios.

- **Advanced normalization and personalization strategies.** Inter-subject variability represents one of the main limitations for building generalizable models based on physiological signals. In this thesis, a signal-level normalization strategy was successfully applied in the context of driving stress assessment to improve classification robustness across subjects. Future work could investigate the systematic extension of this approach to the other experimental contexts analyzed in this work, in order to evaluate its generalizability across different stress conditions and datasets. More broadly, different normalization strategies could be compared, including signal-level normalization, normalization of HRV indices with respect to mean heart rate, and baseline-relative normalization. In addition, personalized baselines, adaptive calibration procedures, and transfer-learning strategies may further improve the robustness and generalizability of physiological classification models.

Addressing these directions would further strengthen the integration between physiological signal analysis, information-theoretic modelling, and machine learning, expanding the applicability of the framework proposed in this thesis and supporting its validation and extension across a wider range of biomedical and real-world monitoring scenarios.

9

Conclusions

The work presented in this thesis aims to bridge information theory and machine learning for the analysis of physiological states within the framework of explainable artificial intelligence.

From a methodological perspective, this thesis establishes information theory as a unifying and coherent framework underpinning all major stages of the machine-learning pipeline applied to physiological data. Information-theoretic concepts are employed at multiple levels and play distinct roles. First, at the signal level, to characterise cardiovascular time series and physiological interactions through measures such as entropy, Mutual Information, and Conditional Mutual Information; second, at the feature level, to guide feature selection and feature importance analyses through information decomposition, explicitly quantifying unique, redundant, and synergistic contributions among extracted features with respect to the target variable; and finally, at the modelling level, to support classification strategies that remain consistent with the underlying informational structure of the data.

By adopting this multi-level use of information theory, the thesis moves beyond isolated applications of entropy or mutual information and instead provides a unified methodological framework in which feature extraction, feature selection, feature importance, and classification are tightly interconnected. In this framework, features are not treated as independent descriptors but as carriers of structured information whose relevance emerges from their individual, shared, and cooperative relationships. This perspective addresses key limitations of traditional univariate or pairwise analyses and enables a physiologically interpretable and mathematically principled assessment of autonomic regulation.

From an experimental standpoint, the analyses conducted across multiple datasets

confirm that cardiovascular variability carries informative signatures of both physiological perturbations and subject-specific characteristics. Stress-related applications showed that orthostatic stress induces clearer and more robust autonomic signatures than cognitive stress, whereas ultra-short-term features can retain meaningful discriminative power when carefully selected. Studies on sex-related differences revealed that autonomic distinctions between males and females are not fully captured by individual HRV indices, but rather emerge from multivariate interaction patterns involving both cardiac and vascular descriptors. Additional applications, including multimodal network physiology analyses, wearable monitoring scenarios, and high-dimensional biomedical datasets, further demonstrated the versatility and robustness of the proposed information-theoretic framework.

A unifying theme emerging from this thesis is the interplay between methodological rigor and physiological interpretability. A more thorough analysis of cardiovascular variability requires tools that capture subtle interactions while remaining grounded in the biological principles of autonomic regulation. The proposed methods achieve this balance by linking mathematical formalisms to physiological meaning and by demonstrating robustness across independent datasets, experimental conditions, and application domains.

Overall, this thesis contributes to the advancement of cardiovascular variability research by:

- introducing and validating information-theoretic methods for modelling high-order feature interactions in physiological datasets;
- investigating the discriminative value of cardiovascular and multimodal features in different contexts (e.g., stress assessment, evaluation of sex-related differences in cardiovascular regulation);
- evaluating the robustness and limitations of ultra-short-term cardiovascular variability analysis;
- identifying methodological factors, such as inter-subject variability, window length, and feature redundancy, that significantly influence classification performance and physiological interpretation;

In sum, this thesis work emphasizes the importance of combining physiological insight, advanced statistical modelling, and machine learning techniques to improve the analysis and interpretation of autonomic regulation. The results provide both methodological foundations and practical evidence supporting the use of information-theoretic approaches for studying complex physiological systems.

The perspectives outlined in the final discussion chapter indicate several promising directions for expanding this work, particularly in multimodal integration and long-term or real-time physiological monitoring.

By building on the methodological developments and experimental insights presented in this thesis, future research may further advance cardiovascular variability analysis toward precision physiology and practical deployment in biomedical, wearable, and human–machine interaction applications.

References

[1] A. Addeh, F. Vega, P. R. Medi, R. J. Williams, G. B. Pike, and M. E. MacDonald. Direct machine learning reconstruction of respiratory variation waveforms from resting state fmri data in a pediatric population. *NeuroImage*, 269:119904, 2023. doi: 10.1016/j.neuroimage.2023.119904.

[2] R. Aggrawal and S. Pal. Sequential feature selection and machine learning algorithm-based patient's death events prediction and diagnosis in heart disease. *SN Computer Science*, 1(6):344, 2020.

[3] G. N. Ahmad, S. Ullah, A. Algethami, H. Fatima, and S. M. H. Akhter. Comparative study of optimum medical diagnosis of human heart disease using machine learning technique with and without sequential feature selection. *ieee access*, 10:23808–23828, 2022.

[4] P. Ahmadi, A. Afzalian, A. Jalali, S. Sadeghian, F. Masoudkabir, A. Oraii, A. Ayati, S. Nayebirad, P. S. Pezeshki, M. Lotfi Tokaldani, et al. Age and gender differences of basic electrocardiographic values and abnormalities in the general adult population; tehran cohort study. *BMC Cardiovascular Disorders*, 23(1):303, 2023.

[5] H. Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 2003.

[6] L. Ali, A. Javeed, A. Noor, et al. Parkinson's disease detection based on features refinement through l1 regularized svm and deep neural network. *Scientific Reports*, 14(1333), 2024. doi: 10.1038/s41598-024-51600-y.

[7] H. H. Andersen, A. Mühlbacher, M. Nübling, J. Schupp, and G. G. Wagner. Computation of standard values for physical and mental health scale scores using the soep version of sf-12v2. *Journal of Contextual Economics–Schmollers Jahrbuch*, (1):171–182, 2007.

[8] Y. Antonacci, C. Barà, A. Zaccaro, F. Ferri, R. Pernice, and L. Faes. Time-varying information measures: an adaptive estimation of information storage with application to brain-heart interactions. *Frontiers in Network Physiology*, 3:1242505, 2023.

[9] R. Arora, A. Krummerman, P. Vijayaraman, M. Rosengarten, V. Suryadevara, T. Lejemtel, and K. J. Ferrick. Heart rate variability and diastolic heart failure. *Pacing and clinical electrophysiology*, 27(3):299–303, 2004.

[10] P.-D. Arsenault, S. Wang, and J.-M. Patenaude. A survey of explainable artificial intelligence (xai) in financial time series forecasting. *ACM Computing Surveys*, 57 (10):1–37, 2025.

[11] A. H. Association. All about heart rate (pulse). [Online].

[12] H. Azami, L. Faes, J. Escudero, A. Humeau-Heurtier, L. E. Silva, et al. Entropy analysis of univariate biomedical signals: Review and comparison of methods. *Frontiers in Entropy across the Disciplines: Panorama of Entropy: Theory, Computation, and Applications*, pages 233–286, 2023.

[13] R. Bailón, L. Sörnmo, P. Laguna, et al. Ecg-derived respiratory frequency estimation. *Advanced methods and tools for ECG data analysis*, 1:215–244, 2006.

[14] J. S. Banerjee, M. Mahmud, and D. Brown. Heart rate variability-based mental stress detection: an explainable machine learning approach. *SN Computer Science*, 4(2):176, 2023.

[15] C. Barà, R. Pernice, L. Sparacino, Y. Antonacci, M. Javorka, and L. Faes. Analysis of cardiac pulse arrival time series at rest and during physiological stress. In *2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON)*, pages 926–931. IEEE, 2022.

[16] C. Barà, R. Pernice, C. A. Catania, M. Hilal, A. Porta, A. Humeau-Heurtier, and L. Faes. Comparison of entropy rate measures for the evaluation of time series complexity: Simulations and application to heart rate and respiratory variability. *Biocybernetics and Biomedical Engineering*, 44(2):380–392, 2024.

[17] C. Barà, Y. Antonacci, M. Iovino, I. Lazic, and L. Faes. Partial information decomposition for discrete target and continuous source random variables. *Physical Review E*, 112(1):L012301, 2025.

[18] R. Bárcenas, R. Fuentes-García, and L. Naranjo. Mixed kernel svr addressing parkinson's progression from voice features. *Plos one*, 17(10):e0275721, 2022.

[19] V. Bari, G. Girardengo, A. Marchi, B. De Maria, P. A. Brink, L. Crotti, P. J. Schwartz, and A. Porta. A refined multiscale self-entropy approach for the assessment of cardiac control complexity: application to long qt syndrome type 1 patients. *Entropy*, 17(11):7768–7785, 2015.

[20] V. Bari, B. De Maria, C. E. Mazzucco, G. Rossato, D. Tonon, G. Nollo, L. Faes, and A. Porta. Cerebrovascular and cardiovascular variability interactions investigated through conditional joint transfer entropy in subjects prone to postural syncope. *Physiological Measurement*, 38(5):976, 2017.

[21] C. Barà, R. Pernice, C. A. Catania, M. Hilal, A. Porta, A. Humeau-Heurtier, and L. Faes. Comparison of entropy rate measures for the evaluation of time series complexity: Simulations and application to heart rate and respiratory variability. *Biocybernetics and Biomedical Engineering*, 44(2):380–392, 2024. ISSN 0208-5216. doi: https://doi.org/10.1016/j.bbe.2024.04.004.

[22] A. Bashan, R. P. Bartsch, J. W. Kantelhardt, S. Havlin, and P. C. Ivanov. Network physiology reveals relations between network topology and physiological function. *Nature communications*, 3(1):702, 2012.

[23] R. Battiti. Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks*, 5(4):537–550, 1994.

[24] A. L. Beam and I. S. Kohane. Big data and machine learning in health care. *Jama*, 319(13):1317–1318, 2018.

[25] M. Behan and J. M. Wenninger. Sex steroidal hormones and respiratory control. *Respiratory physiology & neurobiology*, 164(1-2):213–221, 2008.

[26] Y. Bengio, I. Goodfellow, A. Courville, et al. *Deep learning*, volume 1. MIT press Cambridge, MA, USA, 2017.

[27] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.

[28] A. Bernardes, R. Couceiro, J. Medeiros, J. Henriques, C. Teixeira, M. Simões, J. Durães, R. Barbosa, H. Madeira, and P. Carvalho. How reliable are ultra-short-term hrv measurements during cognitively demanding tasks? *Sensors*, 22(17): 6528, 2022.

[29] G. G. Berntson, J. Thomas Bigger Jr, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone, et al. Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, 34(6): 623–648, 1997.

[30] G. E. Billman. The lf/hf ratio does not accurately measure cardiac sympatho-vagal balance, 2013.

[31] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[32] A. A. Biswas. A comprehensive review of explainable ai for disease diagnosis. *Array*, 22:100345, 2024. ISSN 2590-0056. doi: https://doi.org/10.1016/j.array.2024.100345.

[33] A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, and M. Lang. Benchmark for filter methods for feature selection in high-dimensional classification data. *Computational Statistics & Data Analysis*, 143:106839, 2020.

[34] L. Breiman. Random forests. 45 (1): 5–32. *Use of Modified Masood Score with Prediction*, 99, 2001.

[35] J. Bring. How to standardize regression coefficients. *The American Statistician*, 48 (3):209–213, 1994. ISSN 00031305.

[36] G. Brown, A. Pocock, M.-J. Zhao, and M. Luján. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. *The journal of machine learning research*, 13(1):27–66, 2012.

[37] R. A. Brown. Building a balanced k-d tree in $o(kn \log n)$ time. *Journal of Computer Graphics Techniques (JCGT)*, 4(1):50–68, March 2015. ISSN 2331-7418.

[38] S. L. Brunton and J. N. Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.

[39] D. Cardone, D. Perpetuini, C. Filippini, L. Mancini, S. Nocco, M. Tritto, S. Rinella, A. Giacobbe, G. Fallica, F. Ricci, et al. Classification of drivers' mental workload levels: comparison of machine learning methods based on ecg and infrared thermal signals. *Sensors*, 22(19):7300, 2022.

[40] R. Castaldo, P. Melillo, R. Izzo, N. De Luca, and L. Pecchia. Fall prediction in hypertensive patients via short-term hrv analysis. *IEEE journal of biomedical and health informatics*, 21(2):399–406, 2016.

[41] R. Castaldo, L. Montesinos, P. Melillo, C. James, and L. Pecchia. Ultra-short term hrv features as surrogates of short term hrv: A case study on mental stress detection in real life. *BMC medical informatics and decision making*, 19(1):12, 2019.

[42] O. Catalina-Rodriguez, V. K. Kolukula, Y. Tomita, A. Preet, F. Palmieri, A. Wellstein, S. Byers, A. J. Giaccia, E. Glasgow, C. Albanese, and M. L. Avantaggiati. The mitochondrial citrate transporter, cic, is essential for mitochondrial

homeostasis. *Oncotarget*, 3(10):1220–1235, Oct. 2012. ISSN 1949-2553. doi: 10.18632/oncotarget.714.

[43] A. Catav, B. Fu, Y. Zoabi, A. L. W. Meilik, N. Shomron, J. Ernst, S. Sankararaman, and R. Gilad-Bachrach. Marginal contribution feature importance - an axiomatic approach for explaining data. In M. Meila and T. Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 1324–1335. PMLR, 18–24 Jul 2021.

[44] G. S. Chan, P. M. Middleton, B. G. Celler, L. Wang, and N. H. Lovell. Change in pulse transit time and pre-ejection period during head-up tilt-induced progressive central hypovolaemia. *Journal of clinical monitoring and computing*, 21(5):283–293, 2007.

[45] G. Chandrashekar and F. Sahin. A survey on feature selection methods. *Computers & electrical engineering*, 40(1):16–28, 2014.

[46] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.*, pages 321–357, 2002. doi: 10.1002/eap.2043.

[47] g. Cheng, Z. Qin, C. Feng, Y. Wang, and F. Li. Conditional mutual information-based feature selection analyzing for synergy and redundancy. *Etri Journal*, 33(2): 210–218, 2011.

[48] J. Cheng, Q. Zou, and Y. Zhao. Ecg signal classification based on deep cnn and bilstm. *BMC medical informatics and decision making*, 21(1):365, 2021.

[49] F. Chirico, A. A. Afolabi, O. S. Ilesanmi, G. Nucera, G. Ferrari, A. Sacco, L. Szarpak, P. Crescenzo, N. Magnavita, M. Leiter, et al. Prevalence, risk factors and prevention of burnout syndrome among healthcare workers: an umbrella review of systematic reviews and meta-analyses. *Journal of Health and Social Sciences*, 6(4):465–491, 2021.

[50] G. Clifford, J. Behar, Q. Li, and I. Rezek. Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiological measurement*, 33(9):1419, 2012.

[51] F. Coelho, A. P. Braga, and M. Verleysen. A mutual information estimator for continuous and discrete variables applied to feature selection and classification problems. *International Journal of Computational Intelligence Systems*, 9(4):726–733, 2016.

[52] T. M. Cover. *Elements of information theory*. John Wiley & Sons, 1999.

[53] I. Covert, S. M. Lundberg, and S.-I. Lee. Understanding global feature contributions with additive importance measures. *Advances in neural information processing systems*, 33:17212–17223, 2020.

[54] E. J. Crampin, M. Halstead, P. Hunter, P. Nielsen, D. Noble, N. Smith, and M. Tawhai. Computational physiology and the physiome project. *Experimental Physiology*, 89(1):1–26, 2004.

[55] K. M. Dalmeida and G. L. Masala. Hrv features as viable physiological markers for stress detection using wearable devices. *Sensors*, 21(8):2873, 2021.

[56] E. Dhamala, K. W. Jamison, M. R. Sabuncu, and A. Kuceyeski. Sex classification using long-range temporal dependence of resting-state functional mri time series. *Human brain mapping*, 41(13):3567–3579, 2020.

[57] C. Ding and H. Peng. Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology*, 3(02): 185–205, 2005.

[58] A. E. Draghici and J. A. Taylor. The physiological basis and measurement of heart rate variability in humans. *Journal of physiological anthropology*, 35(1):22, 2016.

[59] D. L. Eckberg. Point: counterpoint: respiratory sinus arrhythmia is due to a central mechanism vs. respiratory sinus arrhythmia is due to the baroreflex mechanism. *Journal of applied physiology*, 2009.

[60] C. El-Hajj and P. A. Kyriacou. A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure. *Biomedical Signal Processing and Control*, 58:101870, 2020.

[61] M. Elstad, E. L. O'Callaghan, A. J. Smith, A. Ben-Tal, and R. Ramchandra. Cardiorespiratory interactions in humans and animals: rhythms for life. *American Journal of Physiology-Heart and Circulatory Physiology*, 315(1):H6–H17, 2018.

[62] F. Esgalhado, A. Batista, V. Vassilenko, S. Russo, and M. Ortigueira. Peak detection and hrv feature evaluation on ecg and ppg signals. *Symmetry*, 14(6):1139, 2022.

[63] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean. A guide to deep learning in healthcare. *Nature medicine*, 25(1):24–29, 2019.

[64] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada. Normalized mutual information feature selection. *IEEE Transactions on neural networks*, 20(2):189–201, 2009.

[65] L. Faes and A. Porta. Conditional entropy-based evaluation of information dynamics in physiological systems. In *Directed information measures in neuroscience*, pages 61–86. Springer, 2014.

[66] L. Faes, A. Porta, and G. Nollo. Information decomposition in bivariate systems: theory and application to cardiorespiratory dynamics. *Entropy*, 17(1):277–303, 2015.

[67] L. Faes, A. Porta, G. Nollo, and M. Javorka. Information decomposition in multivariate systems: definitions, implementation and application to cardiovascular networks. *Entropy*, 19(1):5, 2016.

[68] L. Faes, M. Gómez-Extremera, R. Pernice, P. Carpena, G. Nollo, A. Porta, and P. Bernaola-Galván. Comparison of methods for the assessment of nonlinearity in short-term heart rate variability under different physiopathological states. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(12), 2019.

[69] M. A. Farias da Silva, R. L. De Carvalho, and T. da S. Almeida. Evaluation of a sliding window mechanism as dataaugmentation over emotion detection on speech. *Acad. J. Comput. Eng. Appl. Math.*, 2(1):11–18, 2021. doi: 10.20873/uft.2675-3588.2021.v2n1.p11-18.

[70] M. Finžgar and P. Podržaj. Feasibility of assessing ultra-short-term pulse rate variability from video recordings. *PeerJ*, 8:e8342, 2020.

[71] F. Fleuret. Fast binary feature selection with conditional mutual information. *Journal of Machine learning research*, 5(Nov):1531–1555, 2004.

[72] A. Fotiadis, I. Vlachos, and D. Kugiumtzis. The causality measure of partial mutual information from mixed embedding (pmime) revisited. *Chaos (Woodbury, NY)*, 34 (3):033113–033113, 2024.

[73] D. Fruet, C. Bara, R. Pernice, L. Faes, and G. Nollo. Assessment of driving stress through svm and knn classifiers on multi-domain physiological data. In *MELECON 2022 - IEEE Mediterr. Electrotech. Conf. Proc.*, pages 920–925, 2022. doi: 10.1109/MELECON53508.2022.9842891.

[74] D. Fruet, C. Barà, R. Pernice, M. Iovino, L. Faes, and G. Nollo. A signal normalization approach for robust driving stress assessment using multi-domain physiological data. *Eng*, 6(11):288, 2025.

[75] F. Gasparini, A. Grossi, M. Giltri, and S. Bandini. Personalized ppg normalization based on subject heartbeat in resting state condition. *Signals*, 3(2):249–265, 2022. doi: 10.3390/signals3020016.

[76] M. A. Gelen, P. D. Barua, I. Tasci, G. Tasci, E. Aydemir, S. Dogan, T. Tuncer, and U. Acharya. Novel accurate classification system developed using order transition pattern feature engineering technique with physiological signals. *Scientific Reports*, 15(1):15278, 2025.

[77] G. R. Geovanini, E. R. Vasques, R. de Oliveira Alvim, J. G. Mill, R. V. Andreão, B. K. Vasques, A. C. Pereira, and J. E. Krieger. Age and sex differences in heart rate variability and vagal specific patterns–baependi heart study. *Global heart*, 15 (1):71, 2020.

[78] R. J. Gerrig and P. G. Zimbardo. *American Psychological Association: Glossary of Psychological Terms*. Pearson Education, Education, Incorporated (COR), 2002.

[79] G. Giannakakis, K. Marias, and M. Tsiknakis. A stress recognition system using hrv parameters and machine learning techniques. In *2019 8th International conference on affective computing and intelligent interaction workshops and demos (ACIIW)*, pages 269–272. IEEE, 2019.

[80] G. Giorgi, L. I. Lecca, J. M. Leon-Perez, S. Pignata, G. Topa, and N. Mucci. Emerging issues in occupational disease: mental health in the aging working population and cognitive impairment—a narrative review. *BioMed research international*, 2020(1):1742123, 2020.

[81] M. Gjoreski et al. Datasets for cognitive load inference using wearable sensors and psychological traits. *Appl. Sci. Switz.*, 10(11), 2020. doi: 10.3390/app10113843.

[82] A. L. Goldberger, C.-K. Peng, and L. A. Lipsitz. What is physiologic complexity and how does it change with aging and disease? *Neurobiology of aging*, 23(1): 23–26, 2002.

[83] X. Gu, J. Guo, L. Xiao, and C. Li. Conditional mutual information-based feature selection algorithm for maximal relevance minimal redundancy. *Applied Intelligence*, 52(2):1436–1447, 2022.

[84] V. Gupta, N. K. Saxena, A. Kanungo, A. Gupta, P. Kumar, and Salim. A review of different ecg classification/detection techniques for improved medical applications. *International Journal of System Assurance Engineering and Management*, 13(3): 1037–1051, 2022.

[85] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.

[86] P. Guzik, T. Krauze, A. Wykretowicz, and J. Piskorski. Healthy young poles–hypol database with synchronised beat-to-beat heart rate and blood pressure signals: Hypol–cardiovascular time series database. *Journal of Medical Science*, 92(4): e941–e941, 2023.

[87] L. Han, Q. Zhang, X. Chen, Q. Zhan, T. Yang, and Z. Zhao. Detecting work-related stress with a wearable device. *Comput. Ind.*, 90:42–49, 2017. doi: 10.1016/j.compind.2017.05.004.

[88] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, 2nd edition, 2009. ISBN 978-0387848570.

[89] Y.-L. He, R. Wang, S. Kwong, and X.-Z. Wang. Bayesian classifiers based on probability density estimation and their applications to simultaneous fault diagnosis. *Information Sciences*, 259:252–268, 2014. ISSN 0020-0255. doi: https://doi.org/10.1016/j.ins.2013.09.003.

[90] J. A. Healey and R. W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transp. Syst.*, 6(2):156–166, 2005. doi: 10.1109/TITS.2005.848368.

[91] D. Hernando, S. Roca, J. Sancho, Á. Alesanco, and R. Bailón. Validation of the apple watch for heart rate variability measurements during relax and mental stress in healthy subjects. *Sensors*, 18(8):2619, 2018.

[92] S. Huang, J. Li, P. Zhang, and W. Zhang. Detection of mental fatigue state with wearable ecg devices. *International journal of medical informatics*, 119:39–46, 2018.

[93] I. K. Ihianle, P. Machado, K. Owa, D. A. Adama, R. Otuka, and A. Lotfi. Minimising redundancy, maximising relevance: Hrv feature selection for stress classification. *Expert Systems with Applications*, 239:122490, 2024.

[94] M. Iovino, M. Javorka, L. Faes, and R. Pernice. Comparison of machine learning approaches for physiological states classification using heart rate and pulse rate variability indices. In *Proceedings of the Eighth National Congress of Bioengineering*, pages 679–682, 2023.

[95] M. Iovino, I. Lazic, T. Loncar-Turukalo, M. Javorka, R. Pernice, and L. Faes. Classification of physiological states through machine learning algorithms applied to ultra-short-term heart rate and pulse rate variability indices on a single-feature basis. In *Mediterranean Conference on Medical and Biological Engineering and Computing*, pages 114–124. Springer, 2023.

[96] M. Iovino, I. Lazic, C. Barà, L. Faes, and R. Pernice. Conditional mutual information-based feature selection for sex differences characterization. In *2024 13th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, pages 1–2, 2024. doi: 10.1109/ESGCO63003.2024.10766964.

[97] M. Iovino, I. Lazic, T. Loncar-Turukalo, M. Javorka, R. Pernice, and L. Faes. Comparison of automatic and physiologically-based feature selection methods for classifying physiological stress using heart rate and pulse rate variability indices. *Physiological Measurement*, 45:115004, 2024.

[98] M. Iovino, I. Lazic, C. Barà, D. Kugiumtzis, L. Faes, and R. Pernice. A principled and data-efficient information-theoretic method for feature selection. *submitted to IEEE Journal of Biomedical and Health Informatics*, 2025.

[99] M. Iovino, M. Ontivero-Ortega, C. Barà, S. Stramaglia, M. Javorka, R. Pernice, and L. Faes. Assessing feature importance in cardiovascular variability for the classification of physiological stress. In *IEEE International Conference on Omni-layer Intelligent Systems (COINS)*, pages 1–6. IEEE, 2025.

[100] P. C. Ivanov. The new field of network physiology: building the human physiolome. *Frontiers in network physiology*, 1:711778, 2021.

[101] P. Ivaturi, M. Gadaleta, A. C. Pandey, M. Pazzani, S. R. Steinhubl, and G. Quer. A comprehensive explanation framework for biomedical time series classification. *IEEE journal of biomedical and health informatics*, 25(7):2398–2408, 2021.

[102] N. Izzah, A. P. Sutarto, and M. Hariyadi. Machine learning models for the cognitive stress detection using heart rate variability signals. *Jurnal Teknik Industri*, 24(2): 83, 2022.

[103] A. Jakulin. *Machine Learning Based on Attribute Interactions*. PhD thesis, University of Ljubljana, Slovenia, 2005.

[104] J. Janssen, V. Guan, and E. Robeva. Ultra-marginal feature importance: Learning from data with causal guarantees. *arXiv preprint arXiv:2204.09938*, 2024.

[105] M. Javorka, J. Krohova, B. Czippelova, Z. Turianikova, Z. Lazarova, K. Javorka, and L. Faes. Basic cardiovascular variability signals: mutual directed interactions explored in the information domain. *Physiological Measurement*, 38(5):877, 2017.

[106] M. Javorka, F. El-Hamad, B. Czippelova, Z. Turianikova, J. Krohova, Z. Lazarova, and M. Baumert. Role of respiration in the cardiovascular response to orthostatic and mental stress. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 314(6):R761–R769, 2018.

[107] M. Javorka, J. Krohova, B. Czippelova, Z. Turianikova, Z. Lazarova, R. Wiszt, and L. Faes. Towards understanding the complexity of cardiovascular oscillations: Insights from information theory. *Computers in biology and medicine*, 98:48–57, 2018.

[108] M. Javorka, J. Krohova, B. Czippelova, Z. Turianikova, N. Mazgutova, R. Wiszt, M. Ciljakova, D. Cernochova, R. Pernice, A. Busacca, et al. Respiratory sinus arrhythmia mechanisms in young obese subjects. *Frontiers in neuroscience*, 14: 204, 2020.

[109] G. H. John and P. Langley. Estimating continuous distributions in bayesian classifiers, 2013.

[110] A. I. Jony and A. K. B. Arnob. Deep learning paradigms for breast cancer diagnosis: A comparative study on wisconsin diagnostic dataset. *Malaysian Journal of Science and Advanced Technology*, 4(2):109–117, Mar. 2024. doi: 10.56532/mjsat.v4i2.245.

[111] A. Jovic and N. Bogunovic. Classification of biological signals based on nonlinear features. In *Melecon 2010-2010 15th IEEE Mediterranean Electrotechnical Conference*, pages 1340–1345. IEEE, 2010.

[112] A. Jovic and N. Bogunovic. Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artificial intelligence in medicine*, 51(3):175–186, 2011.

[113] J. Joyce. Bayes' Theorem. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2021 edition, 2021.

[114] H. Kerner, J. Campbell, and M. Strickland. Chapter 1 - introduction to machine learning. In J. Helbert, M. D'Amore, M. Aye, and H. Kerner, editors, *Machine Learning for Planetary Science*, pages 1–24. Elsevier, 2022. ISBN 978-0-12-818721-0. doi: https://doi.org/10.1016/B978-0-12-818721-0.00007-0.

[115] M. M. Khan, A. Mendes, and S. K. Chalup. Evolutionary wavelet neural network ensembles for breast cancer and parkinson's disease prediction. *PLOS ONE*, 13(2): e0192192, 2018. doi: 10.1371/journal.pone.0192192.

[116] J. W. Kim, H. S. Seok, and H. Shin. Is ultra-short-term heart rate variability valid in non-static conditions? *Frontiers in physiology*, 12:596060, 2021.

[117] M. Kiruthika, K. Moorthi, M. Anousouya Devi, and S. Abijah Roseline. Chapter 11 - role of xai in building a super smart society 5.0. In F. Al-Turjman, A. Nayyar, M. Naved, A. K. Singh, and M. Bilal, editors, *XAI Based Intelligent Systems for Society 5.0*, pages 295–326. Elsevier, 2024. ISBN 978-0-323-95315-3. doi: https://doi.org/10.1016/B978-0-323-95315-3.00013-9.

[118] R. E. Kleiger, P. K. Stein, and J. T. Bigger Jr. Heart rate variability: measurement and clinical utility. *Annals of noninvasive electrocardiology*, 10(1):88–101, 2005.

[119] J. Koenig and J. F. Thayer. Sex differences in healthy human heart rate variability: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, 64:288–310, 2016.

[120] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial intelligence*, 97(1-2):273–324, 1997.

[121] L. F. Kozachenko and N. N. Leonenko. Sample estimate of the entropy of a random vector. *Problems Inform. Transmission*, pages 95–101, 1987.

[122] A. Kraskov, H. Stögbauer, and P. Grassberger. Estimating mutual information. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 69(6):066138, 2004.

[123] V. Kumar and S. Minz. Feature selection. *SmartCR*, 4(3):211–229, 2014.

[124] N. Kwak and C.-H. Choi. Input feature selection for classification problems. *IEEE transactions on neural networks*, 13(1):143–159, 2002.

[125] E. Katnik, A. Gomu lkiewicz, A. Piotrowska, J. Grzegrzó lka, A. Rusak, A. Kmiecik, K. Ratajczak-Wielgomas, and P. Dziegiel. Bcl11a expression in breast cancer. *Current Issues in Molecular Biology*, 45(4):2681–2698, Mar. 2023. ISSN 1467-3045. doi: 10.3390/cimb45040175.

[126] S. Laborde, E. Mosley, and J. F. Thayer. Heart rate variability and cardiac vagal tone in psychophysiological research–recommendations for experiment planning, data analysis, and data reporting. *Frontiers in psychology*, 8:213, 2017.

[127] K. Lai, N. Twine, A. O'Brien, Y. Guo, and D. Bauer. Artificial intelligence and machine learning in bioinformatics. In S. Ranganathan, M. Gribskov, K. Nakai, and C. Schönbach, editors, *Encyclopedia of Bioinformatics and Computational Biology*, pages 272–286. Academic Press, Oxford, 2019. ISBN 978-0-12-811432-2. doi: https://doi.org/10.1016/B978-0-12-809633-8.20325-7.

[128] P. Laiou, R. G. Andrzejak, and D. Kugiumtzis. Evaluation of causality measures based on non-uniform embedding schemes with application to the cardiovascular system. In *2014 8th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, pages 225–226. IEEE, 2014.

[129] Y. Lan, H. Liu, J. Liu, H. Zhao, and H. Wang. Gender difference of the relationship between arterial stiffness and blood pressure variability in participants in prehypertension. *International Journal of Hypertension*, 2019(1):7457385, 2019.

[130] I. Lazic, G. Mijatovic, M. Iovino, T. Loncar-Turukalo, and L. Faes. A classification method based on local information and nearest neighbor entropy estimation. In *2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON)*, pages 311–316. IEEE, 2024.

[131] I. Lazic, C. Barà, M. Iovino, S. Stramaglia, K. Kasas-Lazetic, N. Jakovljević, and L. Faes. Information-theoretic quantification of high-order feature effects in classification problems. *Chaos, Solitons & Fractals*, 204:117724, 2026.

[132] J. Lei, M. G'Sell, A. Rinaldo, R. J. Tibshirani, and L. Wasserman. Distribution-free predictive inference for regression. *Journal of the American Statistical Association*, 113(523):1094–1111, 2018.

[133] D. D. Lewis. Feature selection and feature extraction for text categorization. In *Speech and Natural Language: Proceedings of a Workshop Held at Harriman, New York, February 23-26, 1992*, 1992.

[134] Y. Liang, M. Elgendi, Z. Chen, and R. Ward. An optimal filter for short photoplethysmogram signals. *Scientific data*, 5(1):1–12, 2018.

[135] H. W. Lim, Y. W. Hau, C. W. Lim, and M. A. Othman. Artificial intelligence classification methods of atrial fibrillation with implementation technology. *Computer Assisted Surgery*, 21(sup1):154–161, 2016.

[136] D. Lin and X. Tang. Conditional infomax learning: An integrated framework for feature extraction and fusion. In *European Conference on Computer Vision*, 2006.

[137] M. A. Little, P. E. McSharry, S. J. Roberts, D. Costello, and I. M. Moroz. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Engineering OnLine*, 6:23 – 23, 2007.

[138] H. Liu and L. Yu. Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on knowledge and data engineering*, 17(4): 491–502, 2005.

[139] A. LoMauro and A. Aliverti. Sex differences in respiratory function. *Breathe*, 14 (2):131–140, 2018.

[140] S. M. Lundberg and S.-I. Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

[141] A. Makkeh, A. J. Gutknecht, and M. Wibral. Introducing a differentiable measure of pointwise shared information. *Physical Review E*, 103(3):032149, 2021.

[142] M. Malik. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use: Task force of the european society of cardiology and the north american society for pacing and electrophysiology. *Annals of Noninvasive Electrocardiology*, 1(2):151–181, 1996.

[143] M. Malik and A. J. Camm. Components of heart rate variability—what they really mean and what we really measure, 1993.

[144] A. Malliani, M. Pagani, F. Lombardi, and S. Cerutti. Cardiovascular neural regulation explored in the frequency domain. *Circulation*, 84(2):482–492, 1991.

[145] J. E. Manson and S. S. Bassuk. Biomarkers of cardiovascular disease risk in women. *Metabolism*, 64(3):S33–S39, 2015.

[146] A. F. Markus, J. A. Kors, and P. R. Rijnbeek. The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the

terminology, design choices, and evaluation strategies. *Journal of Biomedical Informatics*, 113:103655, 2021. ISSN 1532-0464. doi: https://doi.org/10.1016/j.jbi.2020.103655.

[147] A. Martins, R. Pernice, C. Amado, A. P. Rocha, M. E. Silva, M. Javorka, and L. Faes. Multivariate and multiscale complexity of long-range correlated cardiovascular and respiratory variability series. *Entropy*, 22(3):315, 2020.

[148] D. McDuff, S. Gontarek, and R. Picard. Remote measurement of cognitive stress via heart rate variability. In *2014 36th annual international conference of the IEEE engineering in medicine and biology society*, pages 2957–2960. IEEE, 2014.

[149] E. Mejía-Mejía, K. Budidha, T. Y. Abay, J. M. May, and P. A. Kyriacou. Heart rate variability (hrv) and pulse rate variability (prv) for the assessment of autonomic responses. *Frontiers in physiology*, 11:779, 2020.

[150] E. Mejía-Mejía, J. M. May, R. Torres, and P. A. Kyriacou. Pulse rate variability in cardiovascular health: A review on its applications and relationship with heart rate variability. *Physiological Measurement*, 41(7):07TR01, 2020.

[151] E. Mejía-Mejía, K. Budidha, P. A. Kyriacou, and M. Mamouei. Comparison of pulse rate variability and morphological features of photoplethysmograms in estimation of blood pressure. *Biomedical Signal Processing and Control*, 78:103968, 2022.

[152] M. Mersha, K. Lam, J. Wood, A. K. AlShami, and J. Kalita. Explainable artificial intelligence: A survey of needs, techniques, applications, and future direction. *Neurocomputing*, 599:128111, 2024. ISSN 0925-2312. doi: https://doi.org/10.1016/j.neucom.2024.128111.

[153] O. C. Mesner and C. R. Shalizi. Conditional mutual information estimation for mixed, discrete and continuous data. *IEEE Transactions on Information Theory*, 67(1):464–484, 2020.

[154] N. Metropolis and S. Ulam. The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341, 1949. ISSN 01621459, 1537274X.

[155] P. Meyer and G. Bontempi. On the use of variable complementarity for feature selection in cancer classification. In *Evolutionary Computation and Machine Learning in Bioinformatics*, pages 91–102, 2006.

[156] C. Molnar, G. Casalicchio, and B. Bischl. Interpretable machine learning–a brief history, state-of-the-art and challenges. In *Joint European conference on machine learning and knowledge discovery in databases*, pages 417–431. Springer, 2020.

[157] J. L. Moraes, M. X. Rocha, G. G. Vasconcelos, J. E. Vasconcelos Filho, V. H. C. De Albuquerque, and A. R. Alexandria. Advances in photopletysmography signal analysis for biomedical applications. *Sensors*, 18(6):1894, 2018.

[158] S. Nazir, D. M. Dickson, and M. U. Akram. Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks. *Computers in Biology and Medicine*, 156:106668, 2023. ISSN 0010-4825. doi: https://doi.org/10.1016/j.compbiomed.2023.106668.

[159] C. G. A. Network et al. Comprehensive molecular portraits of human breast tumours. *Nature*, 490(7418):61–70, 2012.

[160] A. Noulas, C. Mascolo, and E. Frias-Martinez. Exploiting foursquare and cellular data to infer user activity in urban environments. In *2013 IEEE 14th International Conference on Mobile Data Management*, volume 1, pages 167–176, 2013. doi: 10.1109/MDM.2013.27.

[161] N. Nourbakhsh, Y. Wang, and F. Chen. Gsr and blink features for cognitive load classification. In *IFIP conference on human-computer interaction*, pages 159–166. Springer, 2013.

[162] J. Novovičová, A. Malík, and P. Pudil. Feature selection using improved mutual information for text classification. In *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, pages 1010–1017. Springer, 2004.

[163] M. Ontivero-Ortega, L. Faes, J. M. Cortes, D. Marinazzo, and S. Stramaglia. Assessing high-order effects in feature importance via predictability decomposition. *Physical Review E*, 111(3):L033301, 2025.

[164] M. Ontivero-Ortega, M. Iovino, R. Pernice, C. Barà, M. Javorka, L. Faes, and S. Stramaglia. Assessment of sex-related differences in cardiovascular control using high-order feature importance. *The European Physical Journal Special Topics*, pages 1–12, 2025.

[165] E. Orrantia-Borunda, P. Anchondo-Nuñez, L. E. Acuña-Aguilar, F. O. Gómez-Valles, and C. A. Ramírez-Valdespino. Subtypes of breast cancer. In *Breast Cancer*, pages 31–42. Exon Publications, Aug. 2022.

[166] G. Ottoboni, A. Cherici, M. Marzocchi, and R. Chattat. Algoritimi di calcolo per gli indici pcs e mcs del questinario sf-12. 2017.

[167] J. Pan and W. J. Tompkins. A real-time qrs detection algorithm. *IEEE Trans. Biomed. Eng.*, (3):230–236, 1985.

[168] F. C. Panganiban and F. A. de Leon. Stress detection using smartphone extracted photoplethysmography. In *2021 IEEE region 10 symposium (TENSYMP)*, pages 1–7. IEEE, 2021.

[169] C. Pascoal, M. R. Oliveira, A. Pacheco, and R. Valadas. Theoretical evaluation of feature selection methods based on mutual information. *Neurocomputing*, 226: 168–181, 2017.

[170] J. Paton, P. Boscan, A. Pickering, and E. Nalivaiko. The yin and yang of cardiac autonomic control: vago-sympathetic interactions revisited. *Brain research reviews*, 49(3):555–565, 2005.

[171] L. Pecchia, R. Castaldo, L. Montesinos, and P. Melillo. Are ultra-short heart rate variability features good surrogates of short-term ones? state-of-the-art review and recommendations. *Healthcare technology letters*, 5(3):94–100, 2018.

[172] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8):1226–1238, 2005.

[173] J. Penttilä, A. Helminen, T. Jartti, T. Kuusela, H. V. Huikuri, M. P. Tulppo, R. Coffeng, and H. Scheinin. Time domain, geometrical and frequency domain analysis of cardiac vagal outflow: effects of various respiratory patterns. *Clinical physiology*, 21(3):365–376, 2001.

[174] F. Pereira, T. Mitchell, and M. Botvinick. Machine learning classifiers and fmri: a tutorial overview. *Neuroimage*, 45(1):S199–S209, 2009.

[175] R. Pernice, M. Javorka, J. Krohova, B. Czippelova, Z. Turianikova, A. Busacca, and L. Faes. Comparison of short-term heart rate variability indexes evaluated through electrocardiographic and continuous blood pressure monitoring. *Medical & biological engineering & computing*, 57(6):1247–1263, 2019.

[176] R. Pernice, L. Sparacino, G. Nollo, S. Stivala, A. Busacca, and L. Faes. Comparison of frequency domain measures based on spectral decomposition for spontaneous baroreflex sensitivity assessment after acute myocardial infarction. *Biomedical Signal Processing and Control*, 68:102680, 2021.

[177] R. Pernice, L. Sparacino, V. Bari, F. Gelpi, B. Cairo, G. Mijatovic, Y. Antonacci, D. Tonon, G. Rossato, M. Javorka, et al. Spectral decomposition of cerebrovascular and cardiovascular interactions in patients prone to postural syncope and healthy controls. *Autonomic Neuroscience*, 242:103021, 2022.

[178] R. Pernice, L. Sparacino, M. Iovino, A. Raimondi, Y. Antonacci, and L. Faes. Gender-related differences in time and spectral entropy rate measures of cardiovascular variability. In *2024 13th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, pages 1–2. IEEE, 2024.

[179] S. V. Philbois and et al. Important differences between hypertensive middle-aged men and premenopausal women in cardiovascular autonomic control. *Biology of Sex Differences*, 12(1):24, 2021.

[180] K. J. Pienta and D. S. Coffey. Correlation of nuclear morphometry with progression of breast cancer. *Cancer*, 68(9):2012–2016, 1991.

[181] G. Pinna, R. Maestri, and A. Di Cesare. Application of time series spectral analysis theory: analysis of cardiovascular variability signals. *Medical and Biological Engineering and Computing*, 34(2):142–148, 1996.

[182] H. Pinto, R. Pernice, M. E. Silva, M. Javorka, L. Faes, and A. P. Rocha. Multiscale partial information decomposition of dynamic processes with short and long-range correlations: theory and application to cardiovascular control. *Physiological Measurement*, 43(8):085004, 2022.

[183] H. Pinto, I. Lazic, Y. Antonacci, R. Pernice, D. Gu, C. Bará, L. Faes, and A. P. Rocha. Testing dynamic correlations and nonlinearity in bivariate time series through information measures and surrogate data analysis. *Frontiers in Network Physiology*, 4:1385421, 2024.

[184] A. Porta, S. Guzzetti, R. Furlan, T. Gnecchi-Ruscone, N. Montano, and A. Malliani. Complexity and nonlinearity in short-term heart period variability: comparison of methods based on local nonlinear prediction. *IEEE Transactions on Biomedical Engineering*, 54(1):94–106, 2006.

[185] A. Porta, T. Bassani, V. Bari, G. D. Pinna, R. Maestri, and S. Guzzetti. Accounting for respiration is necessary to reliably infer granger causality from cardiovascular variability series. *IEEE transactions on biomedical engineering*, 59(3):832–841, 2011.

[186] A. Porta, L. Faes, V. Bari, A. Marchi, T. Bassani, G. Nollo, N. M. Perseguini, J. Milan, V. Minatel, A. Borghi-Silva, et al. Effect of age on complexity and causality of the cardiovascular control: comparison between model-based and model-free approaches. *PloS one*, 9(2):e89463, 2014.

[187] A. Porta, B. De Maria, V. Bari, A. Marchi, and L. Faes. Are nonlinear model-free conditional entropy approaches for the assessment of cardiac control complexity superior to the linear model-based one? *IEEE Transactions on Biomedical Engineering*, 64(6):1287–1296, 2016.

[188] A. Porta, V. Bari, B. De Maria, A. C. Takahashi, S. Guzzetti, R. Colombo, A. M. Catai, F. Raimondi, and L. Faes. Quantifying net synergy/redundancy of spontaneous variability regulation via predictability and transfer entropy decomposition frameworks. *IEEE Transactions on Biomedical Engineering*, 64(11):2628–2638, 2017.

[189] A. Porta, F. Gelpi, V. Bari, B. Cairo, B. De Maria, D. Tonon, G. Rossato, M. Ranucci, and L. Faes. Categorizing the role of respiration in cardiovascular and cerebrovascular variability interactions. *IEEE Transactions on Biomedical Engineering*, 69 (6):2065–2076, 2021.

[190] H. F. Posada-Quintero and J. B. Bolkhovsky. Machine learning models for the identification of cognitive tasks using autonomic reactions from heart rate variability and electrodermal activity. *Behavioral Sciences*, 9(4):45, 2019.

[191] P. Pudil, J. Novovičová, and J. Kittler. Floating search methods in feature selection. *Pattern recognition letters*, 15(11):1119–1125, 1994.

[192] Z. Qin. Naive bayes classification given probability estimation trees. In *2006 International Conference on Machine Learning and Applications*, pages 34–42, Los Alamitos, CA, USA, dec 2006. IEEE Computer Society. doi: 10.1109/ICMLA.2006.36.

[193] D. S. Quintana, M. Elstad, T. Kaufmann, C. L. Brandt, B. Haatveit, M. Haram, M. Nerhus, L. T. Westlye, and O. A. Andreassen. Resting-state high-frequency heart rate variability is related to respiratory frequency in individuals with severe mental illness but not healthy controls. *Scientific Reports*, 6(1):37212, 2016.

[194] S. N. Rai and M. S. Sherkhane. Assessment of quality of sleep among urban working women using pittsburgh sleep quality index. *National Journal of Community Medicine*, 8(12):705–709, 2017.

[195] A. Raimondi, A. Busacca, G. C. Giaconia, Y. Antonacci, S. Stivala, L. Faes, and R. Pernice. Effects of electrocardiographic noise on ultra-short term heart rate variability indices. *Computers in Biology and Medicine*, 196:110803, 2025.

[196] S. Rani, S. Shelyag, C. Karmakar, Y. Zhu, R. Fossion, J. Ellis, S. Drummond, and M. Angelova. Differentiating acute from chronic insomnia with machine learning from actigraphy time series data. *Frontiers in Network Physiology*, 2:1036832, 2022.

[197] S. Rao, K. Adithi, S. C. Bangera, et al. An experimental investigation on pulse transit time and pulse arrival time using ecg, pressure and ppg sensors. *Medicine in Novel Technology and Devices*, 17:100214, 2023.

[198] V. Regitz-Zagrosek and et al. Effects of sex and gender on cardiovascular disease. *Nature Reviews Cardiology*, 20:158–172, 2023.

[199] V. Regitz-Zagrosek and C. Gebhard. Gender medicine: effects of sex and gender on cardiovascular disease manifestation and outcomes. *Nature Reviews Cardiology*, 20(4):236–247, 2023.

[200] S. Reulecke, S. Charleston-Villalobos, A. Voss, R. González-Camarena, M. J. Gaitan-Gonzalez, J. Gonzalez-Hermosillo, G. Hernández-Pacheco, and T. Aljama-Corrales. Gender differences in cardiovascular and cardiorespiratory coupling in healthy subjects during head-up tilt test by joint symbolic dynamics. In *2014 36th Ann Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pages 3402–3405. IEEE, 2014.

[201] S. Rinella, S. Massimino, P. G. Fallica, A. Giacobbe, N. Donato, M. Coco, G. Neri, R. Parenti, V. Perciavalle, and S. Conoci. Emotion recognition: Photoplethysmography and electrocardiography in comparison. *Biosensors*, 12(10):811, 2022.

[202] I. Rish et al. An empirical study of the naive bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, pages 41–46. Seattle, USA, 2001.

[203] B. C. Ross. Mutual information between discrete and continuous data sets. *PloS one*, 9(2):e87357, 2014.

[204] C. Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence*, 1(5): 206–215, 2019.

[205] M. A. Russo, D. M. Santarelli, and D. O'Rourke. The physiological effects of slow breathing in the healthy human. *Breathe*, 13(4):298–309, 2017.

[206] R. Saputo, M. Iovino, R. Pernice, M. Javorka, and L. Faes. Multi-feature classification of physiological stress in cardiovascular and cardiorespiratory interactions. In *2024 13th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, pages 1–2. IEEE, 2024.

[207] V. Satopaa, J. Albrecht, D. Irwin, and B. Raghavan. Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st international conference on distributed computing systems workshops*, pages 166–171. IEEE, 2011.

[208] T. Scagliarini, L. Sparacino, L. Faes, D. Marinazzo, and S. Stramaglia. Gradients of o-information highlight synergy and redundancy in physiological applications. *Frontiers in Network Physiology*, 3:1335808, 2024.

[209] F. Scardulla, G. Cosoli, S. Spinsante, A. Poli, G. Iadarola, R. Pernice, A. Busacca, S. Pasta, L. Scalise, and L. D'Acquisto. Photoplethysmograhic sensors, potential and limitations: Is it time for regulation? a comprehensive review. *Measurement*, 218:113150, 2023.

[210] A. Schäfer and J. Vagedes. How accurate is pulse rate variability as an estimate of heart rate variability?: A review on studies comparing photoplethysmographic technology with an electrocardiogram. *International journal of cardiology*, 166(1): 15–29, 2013.

[211] S. Schulz, F.-C. Adochiei, I.-R. Edu, R. Schroeder, H. Costin, K.-J. Bär, and A. Voss. Cardiovascular and cardiorespiratory coupling analyses: a review. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1997):20120191, 2013.

[212] F. Shaffer and J. P. Ginsberg. An overview of heart rate variability metrics and norms. *Frontiers in public health*, 5:258, 2017.

[213] F. Shaffer, Z. M. Meehan, and C. L. Zerr. A critical review of ultra-short-term heart rate variability norms research. *Frontiers in neuroscience*, 14:594880, 2020.

[214] S. A. Shah, A. Ren, D. Fan, Z. Zhang, N. Zhao, X. Yang, M. Luo, W. Wang, F. Hu, M. U. Rehman, et al. Internet of things for sensing: A case study in the healthcare system. *Applied sciences*, 8(4):508, 2018.

[215] C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[216] A. K. Singh and S. Krishnan. Ecg signal feature extraction trends in methods and applications. *BioMedical Engineering OnLine*, 22(1):22, 2023.

[217] L. K. Singh and M. Khanna. Chapter 3 - introduction to artificial intelligence and current trends. In S. Bhatia Khan, S. Namasudra, S. Chandna, A. Mashat, and F. Xhafa, editors, *Innovations in Artificial Intelligence and Human-Computer Interaction in the Digital Era*, Intelligent Data-Centric Systems, pages 31–66. Academic Press, 2023. ISBN 978-0-323-99891-8. doi: https://doi.org/10.1016/B978-0-323-99891-8.00001-2.

[218] E. Smets et al. Comparison of machine learning techniques for psychophysiological stress detection. In *Pervasive Comput. Paradig. Ment. Health*, pages 13–22, 2016. doi: 10.1007/978-3-319-32270-4.

[219] L. Sparacino, Y. Antonacci, G. Mijatovic, and L. Faes. Measuring hierarchically-organized interactions in dynamic networks through spectral entropy rates: theory, estimation, and illustrative application to physiological networks. *arXiv preprint arXiv:2401.11327*, 2024.

[220] S. Stramaglia, L. Faes, J. M. Cortes, and D. Marinazzo. Disentangling high-order effects in the transfer entropy. *Physical Review Research*, 6(3):L032007, 2024.

[221] W. N. Street, W. H. Wolberg, and O. L. Mangasarian. Nuclear feature extraction for breast tumor diagnosis. In *Biomedical image processing and biomedical visualization*, volume 1905, pages 861–870. SPIE, 1993.

[222] B. Subramanian. Ecg signal classification and parameter estimation using multiwavelet transform. *Biomed. Res*, 28(7):3187–3193, 2017.

[223] J. Tervonen, K. Pettersson, and J. Mäntyjärvi. Ultra-short window length and feature importance analysis for cognitive load detection from wearable sensors. *Electron. Switz.*, 10(5):1–19, 2021. doi: 10.3390/electronics10050613.

[224] J. F. Thayer and et al. Gender differences in the relationship between resting heart rate variability and blood pressure variability in healthy adults. *Autonomic Neuroscience*, 202:16–23, 2016.

[225] N. Timme, W. Alford, B. Flecker, and J. M. Beggs. Synergy, redundancy, and multivariate information measures: an experimentalist's perspective. *Journal of computational neuroscience*, 36(2):119–140, 2014.

[226] K. Tomczak, P. Czerwińska, and M. Wiznerowicz. Review the cancer genome atlas (tcga): an immeasurable source of knowledge. *Wspolczesna Onkologia*, 1A:68–77, 2015. ISSN 1428-2526. doi: 10.5114/wo.2014.47136.

[227] I. Tsamardinos and C. F. Aliferis. Towards principled feature selection: Relevancy, filters and wrappers. In C. M. Bishop and B. J. Frey, editors, *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, volume R4 of *Proceedings of Machine Learning Research*, pages 300–307. PMLR, 03–06 Jan 2003. Reissued by PMLR on 01 April 2021.

[228] A. Tsimpiris, I. Vlachos, and D. Kugiumtzis. Nearest neighbor estimate of conditional mutual information in feature selection. *Expert Systems with Applications*, 39(16):12697–12708, 2012.

[229] K. Tsunoda, A. Chiba, K. Yoshida, T. Watanabe, and O. Mizuno. Predicting changes in cognitive performance using heart rate variability. *IEICE TRANSACTIONS on Information and Systems*, 100(10):2411–2419, 2017.

[230] H. Turbé, M. Bjelogrlic, C. Lovis, and G. Mengaldo. Evaluation of post-hoc interpretability methods in time-series classification. *Nature Machine Intelligence*, 5(3):250–260, 2023.

[231] M. Umair, N. Chalabianloo, C. Sas, and C. Ersoy. Hrv and stress: A mixed-methods approach for comparison of wearable heart rate sensors for biofeedback. *IEEe Access*, 9:14005–14024, 2021.

[232] M. Valente, M. Javorka, A. Porta, V. Bari, J. Krohova, B. Czippelova, Z. Turianikova, G. Nollo, and L. Faes. Univariate and multivariate conditional entropy measures for the characterization of short-term cardiovascular complexity under physiological stress. *Physiological measurement*, 39(1):014002, 2018.

[233] S. Valenti. *Multi-Modal Techniques for the Acquisition and Processing Of Physiological Network Signals*. PhD thesis, Università degli Studi di Palermo, 2024. Available at https://tesidottorato.depositolegale.it/handle/20.500.14242/171848.

[234] G. Valenza, J. P. Saul, and R. Barbieri. Heart rate variability in spontaneous and controlled breathing: a hf power vs. parasympathetic activity index study. In *2022 12th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, pages 1–2. IEEE, 2022.

[235] B. Venkatesh and J. Anuradha. A review of feature selection and its methods. *Cybern. Inf. Technol*, 19(1):3–26, 2019.

[236] J. R. Vergara and P. A. Estévez. A review of feature selection methods based on mutual information. *Neural computing and applications*, 24(1):175–186, 2014.

[237] D. Ververidis and C. Kotropoulos. Sequential forward feature selection with low computational cost. In *2005 13th European Signal Processing Conference*, pages 1–4. IEEE, 2005.

[238] D. P. Vetrov, D. A. Kropotov, and N. O. Ptashko. An efficient method for feature selection in linear regression based on an extended akaike's information criterion. *Computational Mathematics and Mathematical Physics*, 49(11):1972–1985, 2009.

[239] J. D. Victor. Binless strategies for estimation of information from neural data. *Phys Rev E Stat Nonlin Soft Matter Phys*, 66(5 Pt 1):051903, Nov. 2002.

[240] G. Volpes, C. Barà, A. Busacca, S. Stivala, M. Javorka, L. Faes, and R. Pernice. Feasibility of ultra-short-term analysis of heart rate and systolic arterial pressure variability at rest and during stress via time-domain and entropy-based measures. *Sensors*, 22(23):9149, 2022.

[241] T. Vybiral, R. J. Bryg, M. E. Maddens, and W. E. Boden. Effect of passive tilt on sympathetic and parasympathetic components of heart rate variability in normal subjects. *The American journal of cardiology*, 63(15):1117–1120, 1989.

[242] C. Wang and J. Guo. A data-driven framework for learners' cognitive load detection using ecg-ppg physiological feature fusion and xgboost classification. *Procedia computer science*, 147:338–348, 2019.

[243] C. Wang, I. P. Uray, A. Mazumdar, J. A. Mayer, and P. H. Brown. Slc22a5/octn2 expression in breast cancer is induced by estrogen via a novel intronic estrogen-response element (ere). *Breast Cancer Research and Treatment*, 134(1):101–115, Jan. 2012. ISSN 1573-7217. doi: 10.1007/s10549-011-1925-0.

[244] J. E. Ware, M. Kosinski, and S. D. Keller. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Medical care*, 34(3):220–233, 1996.

[245] C. Westphal, S. Hailes, and M. Musolesi. Partial information decomposition for data interpretability and feature selection. *arXiv preprint arXiv:2405.19212*, 2024.

[246] P. L. Williams and R. D. Beer. Nonnegative decomposition of multivariate information. *arXiv preprint arXiv:1004.2515*, 2010.

[247] P. Wollstadt, S. Schmitt, and M. Wibral. A rigorous information-theoretic definition of redundancy and relevancy in feature selection based on (partial) information decomposition. *Journal of Machine Learning Research*, 24(131):1–44, 2023.

[248] W. Xiong, L. Faes, and P. C. Ivanov. Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. *Physical Review E*, 95(6), June 2017. ISSN 2470-0053. doi: 10.1103/physreve.95.062114.

[249] W. Xiong, L. Faes, and P. C. Ivanov. Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. *Physical review E*, 95(6):062114, 2017.

[250] K. Xu, J. Usary, P. C. Kousis, A. Prat, D.-Y. Wang, J. R. Adams, W. Wang, A. J. Loch, T. Deng, W. Zhao, R. D. Cardiff, K. Yoon, N. Gaiano, V. Ling, J. Beyene, E. Zacksenhaus, T. Gridley, W. L. Leong, C. J. Guidos, C. M. Perou, and S. E. Egan. Lunatic fringe deficiency cooperates with the met/caveolin gene amplicon to induce basal-like breast cancer. *Cancer Cell*, 21(5):626–641, May 2012. ISSN 1535-6108. doi: 10.1016/j.ccr.2012.03.041.

[251] H. Yang and J. Moody. Data visualization and feature selection: New algorithms for non-gaussian data. In *Advances in Neural Information Processing Systems*, volume 12, 1999.

[252] X. Ying. An overview of overfitting and its solutions. In *Journal of physics: Conference series*, volume 1168, page 022022. IOP Publishing, 2019.

[253] S. Yu, M. Jin, T. Wen, L. Zhao, X. Zou, X. Liang, Y. Xie, W. Pan, and C. Piao. Accurate breast cancer diagnosis using a stable feature ranking algorithm. *BMC Medical Informatics and Decision Making*, 23(1):64, 2023.

[254] L. B. T. Yugar, J. C. Yugar-Toledo, N. Dinamarco, L. G. Sedenho-Prado, B. V. D. Moreno, T. d. A. Rubio, A. Fattori, B. Rodrigues, J. F. Vilela-Martin, and H. Moreno. The role of heart rate variability (hrv) in different hypertensive syndromes. *Diagnostics*, 13(4):785, 2023.

[255] L. Zan, A. Meynaoui, C. K. Assaad, E. Devijver, and E. Gaussier. A conditional mutual information estimator for mixed data and an associated conditional independence test. *Entropy*, 24(9), 2022. ISSN 1099-4300.

[256] H. Zhou, X. Wang, and R. Zhu. Feature selection based on mutual information with correlation coefficient. *Applied intelligence*, 52(5):5457–5474, 2022.

List of Figures

List of Tables

A

Curriculm Vitae

Marta Iovino was born in Italy on December 2nd 1997. In 2016, she enrolled in the Biomedical Engineering program at the University of Palermo (Palermo, Italy), where she received the Bachelor's degree in Biomedical Engineering in 2019. Motivated by a strong interest in the biomedical field she decided to continue her academic path by pursuing a Master's degree in Biomedical Engineering at the Politecnico di Torino (Turin, Italy).

During her Master's thesis, she focused on the analysis of polysomnographic signals and the application of deep learning techniques to sleep analysis. Her Master's thesis, entitled *"Automated LSTM-based Sleep Stage Classification Using Polysomnographic Signal Processing Techniques"*, addressed the development of a fully automatic framework for sleep stage classification. She obtained the Master's degree in Biomedical Engineering in 2022.

In December 2022, she joined the Biosignals and Information Theory Laboratory (BITLAB), Department of Engineering, University of Palermo, as a PhD student in Information and Communication Technologies (cycle XXXVIII), under the supervision of Dr. Riccardo Pernice and Prof. Luca Faes. Her doctoral research is centred on the development and application of machine learning algorithms for the analysis of biological signals, with the aim of characterizing both physiological and pathological states.

Her research activity primarily focuses on the processing of electrocardiographic and photoplethysmographic signals, with particular emphasis on heart rate variability analysis for the classification of different stress conditions. Within this framework, she investigates information-theoretic methods for feature selection and feature importance to enhance the robustness, interpretability, and physiological relevance of artificial intelligence models applied to biomedical data.

During her PhD, she carried out a research period abroad and collaborated with international research groups, contributing to several peer-reviewed publications in the fields of biomedical signal processing, artificial intelligence, and information theory.

B

List of Publication

Articles in internationally reviewed journals

1. **M. Iovino**, I. Lazić, T. Lončar-Turukalo, M. Javorka, R. Pernice, and L. Faes, ”Comparison of Automatic and Physiologically-based Feature Selection Methods for Classifying Physiological Stress Using Heart Rate and Pulse Rate Variability Indices”, *Physiological Measurement*, 2024.

2. C. Barà, Y. Antonacci, **M. Iovino**, I. Lazić, and L. Faes, ”Partial Information Decomposition for Discrete Target and Continuous Source Random Variables”, *Physical Review E*, vol. 112, no. 1, American Physical Society, 2025.

3. D. Fruet, C. Barà, R. Pernice, **M. Iovino**, L. Faes, and G. Nollo, ”A Signal Normalization Approach for Robust Driving Stress Assessment Using Multi-Domain Physiological Data”, *Eng*, MDPI, 2025.

4. I. Lazić, C. Barà, **M. Iovino**, S. Stramaglia, N. Jakovljević, and L. Faes, ”Information-Theoretic Quantification of High-Order Feature Effects in Classification Problems”, *Chaos, Solitons & Fractals*, 2025.

5. M. Ontivero-Ortega*, **M. Iovino***, R. Pernice, C. Barà, M. Javorka, L. Faes, and S. Stramaglia, ”Assessment of Sex-Related Differences in Cardiovascular Control Using High-Order Feature Importance”, *European Physical Journal Special Topics*, 2025. * These authors contributed equally to the work.

Submitted and in preparation articles

6. **M. Iovino***, I. Lazić*, C. Barà, D. Kugiumtzis, L. Faes, and R. Pernice, "A Principled and Data-Efficient Information-Theoretic Method for Feature Selection", *IEEE Journal of Biomedical and Health Informatics*, under revision, 2025.* These authors contributed equally to the work.

7. **M. Iovino**, Y. Antonacci, G. Genova, S. Valenti, G. Volpes, R. Saputo, V. R. Vergara, F. Oddo, T. F. Scognamiglio, A. Scardina, M. Bellafiore, L. Mistretta, A. Parisi, L. Faes, N. Salviato, and R. Pernice, "Evaluation of Work-Related Stress on Healthcare Operators Using Cardiovascular Variability Indices from Commercial and Portable Multisensor Systems", *in preparation*.

8. **M. Iovino**, L.J. Ayoub, M.l Datko, K. Lin, N. N. Slocum, A. Bolender, J. Lee, T. Kaptchuk, B. Kuo, V. Napadow, R. Barbieri, R. Pernice, R. Sclocco, "Exploring the Interaction between Brainstem Autonomic Nuclei and Cardiorespiratory Signals in Functional Dyspepsia", *submitted to Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2026)*.

Articles in proceedings of international conferences

9. **M. Iovino**, M. Ontivero-Ortega, C. Barà, S. Stramaglia, M. Javorka, R. Pernice, and L. Faes, "Assessing Feature Importance in Cardiovascular Variability for the Classification of Physiological Stress", in *Proceedings of the IEEE International Conference on Omni-layer Intelligent Systems (COINS 2025)*, Madison, WI, USA, 2025.

10. **M. Iovino**, G. Genova, F. Oddo, T. F. Scognamiglio, N. Salviato, Y. Antonacci, A. Busacca, and R. Pernice, "A Combined Analysis of Bioimpedentiometry and Cardiovascular Variability Indices for Non-Invasive Classification of Obesity", in *Proceedings of the Ninth National Congress of Bioengineering*, Pàtron Editore, 2025.

11. **M. Iovino**, I. Lazić, C. Barà, L. Faes, and R. Pernice, "Conditional Mutual Information-Based Feature Selection Method for Sex Differences Characterization", in *Proceedings of the 13th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, Zaragoza, Spain, 2024, pp. 1–2.

12. **M. Iovino**, M. Javorka, L. Faes, and R. Pernice, ”Comparison of Machine Learning Approaches for Physiological States Classification Using Heart Rate and Pulse Rate Variability Indices”, in *Proceedings of the Eighth National Congress of Bioengineering*, Pàtron Editore, 2023, pp. 679–682.

13. **M. Iovino**, I. Lazić, T. Lončar-Turukalo, M. Javorka, R. Pernice, and L. Faes, ”Classification of Physiological States through Machine Learning Algorithms Applied to Ultra-Short-Term Heart Rate and Pulse Rate Variability Indices on a Single-Feature Basis”, in *Mediterranean Conference on Medical and Biological Engineering and Computing*, Springer Nature Switzerland, Cham, 2023, pp. 114–124.

14. G. Genova, C. Barà, **M. Iovino**, G. Chiarello, R. Pernice, A. Parisi, G. C. Giaconia, F. Ficuciello, C. Iacono, A. Smaldone, R. La Rocca, A. Busacca, and S. Stivala, ”Design of a Non-Invasive Multiparametric Biomedical Device for Surgeon Monitoring during Robotic-Aided Procedures”, in *Proceedings of the Ninth National Congress of Bioengineering*, Pàtron Editore, 2025.

15. I. Lazić, G. Mijatović, **M. Iovino**, T. Lončar-Turukalo, and L. Faes, ”A Classification Method Based on Local Information and Nearest Neighbor Entropy Estimation”, in *Proceedings of the IEEE 22nd Mediterranean Electrotechnical Conference (MELECON)*, Porto, Portugal, 2024, pp. 311–316.

16. R. Pernice, L. Sparacino, **M. Iovino**, A. Raimondi, Y. Antonacci, and L. Faes, ”Gender-Related Differences in Time and Spectral Entropy Rate Measures of Cardiovascular Variability”, in *Proceedings of the 13th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, Zaragoza, Spain, 2024, pp. 1–2.

17. R. Saputo, **M. Iovino**, R. Pernice, M. Javorka, and L. Faes, ”Multi-Feature Classification of Physiological Stress in Cardiovascular and Cardiorespiratory Interactions”, in *Proceedings of the 13th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, Zaragoza, Spain, 2024, pp. 1–2.

Oral Presentation

- ”Assessing Feature Importance in Cardiovascular Variability for the Classification of Physiological Stress”, *IEEE International Conference on Omni-layer Intelligent Systems (COINS 2025)*, Madison, Wisconsin, USA, 2025.

- ”Conditional Mutual Information-Based Feature Selection Method for Sex Differences Characterization”, *13th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, Zaragoza, Spain, 2024.

- ”Classification of Physiological States through Machine Learning Algorithms Applied to Ultra-Short-Term Heart Rate and Pulse Rate Variability Indices on a Single-Feature Basis”, *Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON)*, Sarajevo, Bosnia and Erzegovina, 2023.

Poster Presentation

- ”A Combined Analysis of Bioimpedentiometry and Cardiovascular Variability Indices for Non-Invasive Classification of Obesity”, *Ninth National Congress of Bioengineering*, Palermo, Italy, 2025.

- ”Comparison of Machine Learning Approaches for Physiological States Classification Using Heart Rate and Pulse Rate Variability Indices”, *Eight National Congress of Bioengineering*, Padova, Italy, 2025.

B.1 Author Contributions

This section summarises the scientific publications produced during the PhD research activity, with particular reference to those publications in which the candidate is not the first author. For these works, the candidate contributed to different methodological, analytical, or interpretative aspects. The contributions are first described according to the main research themes addressed in the thesis (See table 1.1), and subsequently detailed with reference to the individual publications listed above.

Signal Processing and Feature Extraction. Starting from signal processing pipelines and variability analysis tools available within the BitLab research group, the author contributed to the systematic extraction and organisation of cardiovascular variability features from physiological recordings. The work focused on implementing preprocessing pipelines, extracting multidomain descriptors from RR intervals, pulse signals, and blood pressure variability, and assessing the contribution of different feature domains to classification performance. These procedures enabled the creation of consistent feature sets suitable for subsequent information-theoretic analysis and machine learning applications.

Information-theoretic Feature Selection. The author contributed to the development of information-theoretic methods for feature selection in physiological datasets. In particular, the work focused on the development of an estimator of conditional mutual information suitable for mixed-variable contexts, where continuous physiological features are related to discrete class variables. Building on this estimator, the author formulated and designed the conditional mutual information feature selection (kCMI-FS) algorithm, implementing iterative testing procedures and validating the method through simulations and applications to physiological datasets.

High-order Feature Importance. Starting from methodological ideas developed in collaboration with other researchers on the quantification of high-order information contributions, the author implemented these concepts within a machine-learning framework for physiological data analysis. In particular, the author contributed to the development and implementation of algorithms able to assess the unique, redundant, and synergistic contributions of continuous physiological variables to a discrete target variable, based on conditional mutual information (CMIHiFi). The author also carried out simulation studies and applied these methods to cardiovascular variability datasets in order to evaluate the structure of high-order interactions among physiological descriptors.

Interpretable Classification. The author contributed to the application of interpretable classification approaches based on information-theoretic principles. In particular, the author implemented and evaluated the Local Information Classifier within the context of physiological datasets, following methodological developments carried out in collaboration with researchers at the University of Novi Sad during the PhD research period. The work included the application of the classifier to cardiovascular variability datasets, the investigation of parameter selection strategies, and the evaluation of classification performance and interpretability. The development and validation of this classifier are ongoing.

Applications to Physiological Analysis. The author conducted the experimental analyses used to validate the proposed analytical framework in physiological applications. These activities included the preprocessing of datasets, implementation of feature extraction procedures, execution of feature selection and feature importance analyses, and interpretation of the obtained results in the context of physiological regulation. The applications focused in particular on stress assessment and on the investigation of sex-related differences in cardiovascular control.

The contributions described above are further detailed below with reference to the

individual scientific publications produced during the PhD research activity.

In publication (2), **M. Iovino** contributed to the preparation and analysis of the datasets used to test the proposed methodology, as well as to the implementation of the analysis routines and to the revision of the manuscript.

In publication (3), **M. Iovino** contributed to the critical revision of the manuscript, including the discussion of the methodological approach and the interpretation of the results.

In publication (4), **M. Iovino** contributed to the conceptual discussion of the methodological framework and to the application of the proposed method to physiological datasets, including data analysis and interpretation of the results. The author also contributed to the revision of the manuscript.

In publication (16), **M. Iovino** contributed to the preprocessing and organization of the cardiovascular dataset, as well as to the visualization and interpretation of the results and to the revision of the work.

In publication (17), **M. Iovino** contributed to the implementation of the analysis code, the investigation of the physiological applications, and the interpretation and discussion of the results, as well as to the preparation of the manuscript.

In publication (15), **M. Iovino** contributed to the application of the proposed method to physiological datasets, including the analysis of the results and their physiological interpretation, as well as to the writing of the manuscript.

In publication (14), **M. Iovino** contributed to the preprocessing and organization of the cardiovascular dataset, as well as to the visualization and interpretation of the results and to the revision of the work.

C

Acknowledgements

The research presented in this thesis was carried out at the *Biosignals and Information Theory Laboratory (BITLAB), Department of Engineering, University of Palermo, Palermo (Italy)*, headed by *Prof. Luca Faes*. The scientific environment of the laboratory, characterized by continuous methodological discussion and interdisciplinary interaction, provided a fundamental contribution to the development of both the theoretical framework and the applied analyses presented in this work. Several established and newly formed research collaborations played a key role in the advancement of the research activity, either by granting access to consolidated datasets or by offering valuable feedback on the physiological interpretation of the results. The research activity was conducted within the framework of the research project *Sensoristica intelligente, infrastrutture e modelli gestionali per la sicurezza di soggetti fragili (4Frailty)* and the PRIN 2022 research project *High-Order Dynamical Networks in Computational Neuroscience and Physiology: an Information-Theoretic Framework (HONEST)*, under which the author was supported through a research fellowship. These projects provided the institutional and scientific conditions necessary to ensure continuity of the research activity. A substantial part of the research activity was conducted during a research period abroad in strong collaboration with *Ivan Lažić and Prof. Tatjana Lončar-Turukalo* from the *Faculty of Technical Sciences, University of Novi Sad, Serbia*. This collaboration resulted in several joint scientific publications and provided a significant contribution to the development, refinement, and validation of the methodological approaches presented in this thesis, through continuous scientific exchange, shared data analysis, and joint interpretation of the results. Part of the research outlined in this thesis was also conducted in collaboration with the *Department of Physiology, Jessenius Faculty of Medicine, Comenius University, Martin*

(Slovakia), headed by *Prof. Michal Javorka*. *Prof. Javorka* provided the cardiovascular datasets analyzed in this work and dispensed a unique contribution to the physiological interpretation of the results. It is also worth acknowledging *Prof. Przemys law Guzik (Poznan University of Medical Sciences)* for providing the *HYPOL* dataset. The availability of this independent cohort allowed the extension and experimental validation of the proposed analytical framework on a large population of healthy subjects. The dataset on work-related stress in healthcare professionals was provided by *MD. Nicoletta Salviato* within the collaborative project *Obiettivo PSN 2017/4.1.14*, *"Valutazione non invasiva dello stress lavoro-correlato per la prevenzione degli infortuni nelle strutture sanitarie della Regione Siciliana"*. This contribution allowed the analysis of chronic occupational stress in a real-world healthcare context. An additional research period abroad was carried out in the *Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Harvard Medical School, Boston, MA, USA* under the supervision of *Dr. Roberta Sclocco*. This experience provided valuable exposure to an international research environment and contributed to broadening the methodological and scientific perspective of this work, particularly with respect to advanced signal processing and information-theoretic approaches applied to biomedical data. Finally, the signal processing pipelines and the practical implementation tools developed throughout this work were reviewed and discussed with several colleagues of the *BITLAB* group, whose feedback contributed to improving the robustness and clarity of the proposed methodologies.