Research Article

Parkinson's Disease Diagnosis in Cepstral Domain Using MFCC and Dimensionality Reduction with SVM Classifier

Atiqur Rahman,[1] Sanam Shahla Rizvi ,[2] Aurangzeb Khan,[1] Aaqif Afzaal Abbasi,[3] Shafqat Ullah Khan,[4] and Tae-Sun Chung [5]

[1]*Department of Computer Science, University of Science and Technology Bannu, Bannu, Pakistan*
[2]*Raptor Interactive (Pty) Ltd., Eco Boulevard, Witch Hazel Ave, Centurion 0157, South Africa*
[3]*Department of Software Engineering, Foundation University Islamabad, Islamabad 44000, Pakistan*
[4]*Department of Electronics, University of Buner, Buner, Pakistan*
[5]*Department of Software, Ajou University, Suwon, Republic of Korea*

Correspondence should be addressed to Tae-Sun Chung; tschung@ajou.ac.kr

Parkinson's disease (PD) is one of the most common and serious neurological diseases. Impairments in voice have been reported to be the early biomarkers of the disease. Hence, development of PD diagnostic tool will help early diagnosis of the disease. Additionally, intelligent system developed for binary classification of PD and healthy controls can also be exploited in future as an instrument for prodromal diagnosis. Notably, patients with rapid eye movement (REM) sleep behaviour disorder (RBD) represent a good model as they develop PD with a high probability. It has been shown that slight speech and voice impairment may be a sensitive marker of preclinical PD. In this study, we propose PD detection by extracting cepstral features from the voice signals collected from people with PD and healthy subjects. To classify the extracted features, we propose to use dimensionality reduction through linear discriminant analysis and classification through support vector machine. In order to validate the effectiveness of the proposed method, we also developed ten different machine learning models. It was observed that the proposed method yield area under the curve (AUC) of 88%, sensitivity of 73.33%, and specificity of 84%. Moreover, the proposed intelligent system was simulated using publicly available multiple types of voice database. Additionally, the data were collected from patients under on-state. The obtained results on the public database are promising compared to the previously published work.

1. Introduction

After Alzheimer's disease (AD), Parkinson's disease (PD) is the world's second most prevalent neurodegenerative disorder [1–3]. It has been reported that PD prevails at a rate of 0.3% in of the entire population in industrialized countries, while in elder population (60 or above age), the PD prevalence rate is 1% [1]. Impairments in voice have been reported to be the early biomarkers of the disease. Additionally, the proposed intelligent system has the capability to be used as an instrument for prodromal diagnosis. Notably, patients with REM sleep behaviour disorder (RBD) represent a good model as they develop PD with a high

probability. It has been shown that slight speech and voice impairment may be a sensitive marker of preclinical PD [4–7].

People with PD face numerous symptoms including movement impairments (gait and tremors), poor balance, bradykinesia which is slowness of movement, and rigidity [8–12]. As discussed above, the lack of reliable tests for diagnosis of PD has made the diagnosis of PD a challenging task [13–15]. However, recent research reported that PD patients manifest impairments in voice and speech. However, these voice defects cannot be detected in clinics by medical practitioners. Hence, automated signal processing tools are required to capture these impairments in voice and

to detect PD in its early stages. Recent research shows that machine learning and signal processing algorithms are successful in automated disease detection through automated risk factors extraction and classification [16–19]. Motivated by these studies, in this paper, we also attempt to develop a method based on machine learning and signal processing algorithms for PD detection.

The automated disease detection methods discussed above motivated us to develop automated model for PD detection using signal processing algorithms for feature extraction from voice signals and machine learning algorithms for classification. Hence, we collected a voice dataset, namely, Pak-Voice-PD that contains multiple types of vowel phonations for two types of subjects, i.e., healthy and PD patients. Numerical features are extracted using mel-frequency cepstral coefficients (MFCCs). In order to obtain better PD detection performance, we project the MFCC features to lower dimensional space using linear discriminant analysis (LDA) approach. Finally, numerous machine learning models are developed with the goal of obtaining an optimal learning model. Through performance analysis, we pointed out that support vector machine with linear and radial basis function (RBF) kernels provide optimal performance. Hence, in this study, we propose automated PD detection based on MFCC-LDA-SVM hybrid approach. The working of the proposed MFCC-LDA-SVM model is depicted in Figure 1.

The main contributions of this study are as follows:

(1) Collection of a relatively larger dataset: the collected database has relatively larger number of multiple types of voice phonations or samples.

(2) Construction of unbiased machine learning models for the automated detection of PD.

(3) In this paper, we developed MFCC-LDA-SVM model for PD detection problem. To the best of our knowledge, no previous studies have explored development of MFCC-LDA-SVM model for PD detection based on voice data.

(4) The proposed method, namely, MFCC-LDA-SVM has better performance than ten other machine learning models and many recently published studies.

The remaining of the manuscript presents related work in Section 2 and material and methods in Section 3. The evaluation and validation methods are briefly discussed in Section 4. Section 5 presents results of the proposed model and its discussion. Section 6 is about conclusion.

2. Related Work

During the last decade, various machine learning systems are proposed for the automated diagnosis of Parkinson's disease (PD) [20]. Resul [13] conducted a comparative study of different classification methods for effective diagnosis of the PD. Decision Tree, Regression, DMneural, and Neural Networks were evaluated for PD detection on the basis of performance scores. Neural network obtained the highest

classification score of 92.9% as compared to rest of classifiers. Tsanas et al. [21] presented speech signal processing algorithms for the prediction of PD symptom severity using random forests and support vector machines. The proposed algorithms were reported to have achieved classification accuracy of 99% using 10 dysphonia features. Kaya et al. [22] developed an entropy-based discretization method where support vector machines, C4.5, k-nearest neighbors, and naive boys were used as classifiers for the detection of PD. The proposed method was developed without using any preprocessing method. The discretization method improved the classification for diagnosis of PD by 4.1% to 12.8%.

Manda and Sairam [23] proposed a method for the early diagnosis of the PD based on the detection of dysphonia. A novel inference system measures the severity of disease through feature selection method based on support vector machines and ranker search method. Hariharan et al. [24] presented a hybrid intelligent system that consists of preprocessing through model-based clustering, feature selection using sequential forward selection, and linear discriminant analysis. For the classification purpose, least-square support vector machine (LS-SVM), probabilistic neural network (PNN), and general regression neural network (GRNN) are deployed. The maximum classification accuracy of 100% was achieved by the proposed method for Parkinson's dataset. Bhalchandra et al. [25] designed a system for early detection of Parkinson's disease (PD) using image processing to compute cheap-based features. Parkinson's progression markers initiative (PPMI) dataset was used along with a striatal binding ratio (SBR) to differentiate between the two types of subjects using discriminant analysis (DA) and support vector machine (SVM). The newly developed system observed the classification accuracy of 99.42%.

Saloni and Gupta [26] developed an algorithm for the detection of PD using clinical voice data. Voice features were used for the classification through support vector machines. The proposed algorithm achieved the accuracy of 100% for subset of features derived from the algorithm. Huang et al. [27] presented a framework for the prediction of Alzheimer's disease (AD) using nonlinear supervised sparse regression-based random forest (RF). The probabilistic paths are assigned using proposed soft-split technique to test sample in RF for more accurate prediction. The proposed soft-split sparse regression-based RF helped to estimate the missing scores. The proposed method demonstrated superior performance as compared to the traditional RF and regression models. Al-Fatlawi et al. [28] adopted deep belief network (DBN) for automated diagnosis of Parkinson's disease (PD). Voice data of Parkinson's disease patients are used for the experiments. The DBN classifier was composed of two stacked restricted Boltzmann machines (RBMs). The first stage is an unsupervised learning that used RBMs to eliminate the problems of the random value of initial weight. The second stage is a supervised learning based on the backpropagation algorithm for fine tuning. The accuracy reported by the proposed method was 94%.

Benba et al. [29] studied the discrimination between the two groups of people (patients with PD and healthy subjects) based on multiple types of voice samples. Human factor
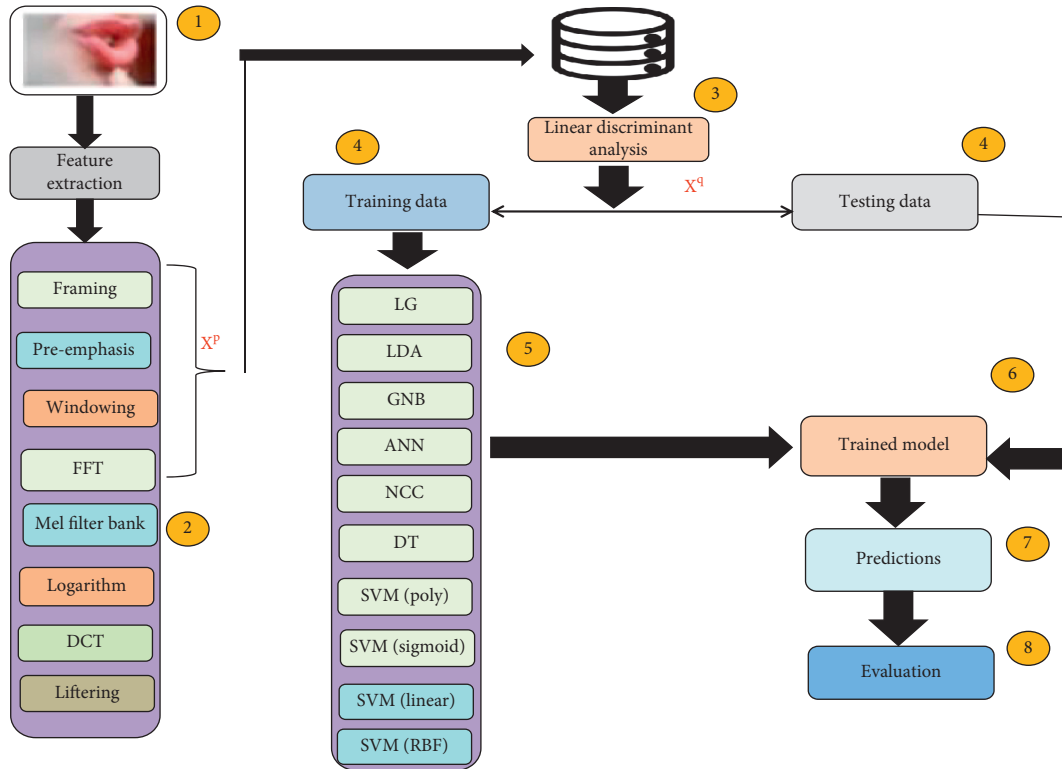
FIGURE 1: Block diagram of the proposed method with the numbering representing the flow of the whole process.

cepstral coefficients (HFCC) were used in the study. Voice print of the each voice recording was calculated for average value through the extracted HFCC. SVM with various kernels (RBF, Polynomial, Linear, and MLP) is deployed for the classification. The best accuracy of 87.5% was achieved through the linear kernel of SVM. Vaiciukynas et al. [30] adopted phonation corresponding to multiple types of vowel and speech tasks to pronounce short sentences in Lithuanian language. Random forest (RF) algorithm is utilized for the individual feature sets and decision-level fusion. It was pointed out that decision-level fusion provides better performance. Naranjo et al. [31] proposed a method for tracking Parkinson's disease (PD) through Bayesian linear regression approach. The proposed method was suitable for the handling of replicated measurements. Li et al. [32] designed a hybrid feature learning algorithm for classification of PD. Hybrid features were developed through combining features and segments. Different methods were deployed for the selection of efficient hybrid features. The classification is made on the basis of selected hybrid features.

Zhang et al. [33] proposed a telediagnosis method through smart phone and machine learning-based Parkinson's disease detection. Time frequency features, stack autoencoders (SAE), and k-nearest neighbor were used for the automated classification of the PD. The classification accuracy reported through proposed method was in the range from 94.00%–98.00%. In another study, Upadhya et al. [34] adopted Single Taper Smooth (STS) window and Thomson Multitaper (TMT) windowing techniques for MFCC and PLP voice feature extraction. For classification,

neural network classifier was deployed for the classification of the subjects at the early stage of PD. Wu et al. [35] designed a feature learning technique for automatically learning about the extracted voice features. Spherical k-means model was deployed to train the two class sample space (PD patients and healthy subjects). The proposed method obtains the mean pooling accuracy of 95.35%. Ali et al. [20] studied the hand tremor abnormality detection associated with the risk of development of Parkinson's disease using a Chi2-based feature selection and Adaboost-based classification. Khan et al. [36] proposed a method for the prediction of cancer and Parkinson's disease. The proposed method utilized the wavelet-based neural networks for the prediction of cancer. The proposed evolutionary wavelet neural network was deployed on various biomedical benchmark datasets for breast cancer and Parkinson's disease, while 10-fold cross-validation scheme was used for performance evaluation metric. The accuracy achieved by the proposed method was 90%.

Braga et al. [37] presented a methodology for early detection of Parkinson's disease by using free-speech recording in uncontrolled background conditions. Machine learning (ML) algorithms along with signal and speech processing techniques were used for the early detection of the disease. For classification, support vector machine (SVM) and random forest (RF) were deployed. The accuracy reported by SVM (RBF) was 92.38% and 99.94% for RF. Recently, Ali [3] developed a hybrid intelligent system that carries out acoustic analysis of voice signals for automatically detecting Parkinson's disease (PD). Linear discriminant

analysis (LDA) was adopted for the dimension reduction and genetic algorithm (GA) for fine tuning the parameter of neural network. Leave one subject out (LOSO) validation scheme was used to avoid the subject overlap. The proposed intelligent system achieved the classification accuracy of 80%. Mostafa et al. [38] presented a Multiple Feature Evaluation Approach (MFEA) and classification machine learning methods (Neural networks, Decision tree, SVM, and Random forest) based on the voice disorders analysis. The performance of the proposed method was evaluated through 10-fold cross-validation metric. The proposed system reported accuracy for SVM was 95.43%. Eskidere et al. [39] proposed a novel random subspace classifier ensemble and obtained 74.17% accuracy under 10-fold CV. Vadovský and Parali [40] utilized decision tree based methods, namely, C4.5, C5.0, Random Forest, CART, and obtained PD detection accuracy of 66.5% under 4-fold cross-validation. Kraipeerapun and Amornsamankul [41] proposed stacking of complementary neural networks (CMTNN) and obtained classification accuracy of 75% under 10-fold cross-validation.

The main problems in these studies were the inappropriate validation scheme that causes artificial subject overlap and baisedness in the developed models [2, 42]. Hence, the obtained results are biased due to the subject overlap between training and testing datasets. In order to develop unbiased machine learning models, Sarkar et al. proposed to use a more practical validation scheme, namely, Leave One Subject Out (LOSO) cross-validation [42]. Under their proposed LOSO approach, they trained and tested KNN and SVM classifiers on multiple types of speech data collected from two classes, i.e., healthy and PD patients and achieved 55% of PD detection accuracy, which are unbiased and more practical results. The same LOSO approach was adopted by Canturk and Karabiber in [43]. In order to improve the PD detection while developing unbiased machine learning methods, they explored integration of four different feature selection methods with six different machine learning models. They obtained best performance of 57.5 using LOSO approach. Recently, Ali et al. [44] proposed a multimodal approach under the LOSO approach and obtained unbiased performance of 70% classification accuracy using time frequency features.

3. Materials and Methods

3.1. Data Acquisition. In this study, we collected voice and handwritten-based database from two types of populations, i.e., PD patients and healthy subjects. The database was collected after the approval of ethical review board of Lady Reading Hospital (Medical Teaching Institution), Pakistan (Ref. No: 174/LRH, 2019). The database was collected from 160 subjects (60 PD patients and 100 age matched healthy subjects). The ages of the PD group of patients range from 43 to 88 with mean 68.3 and standard deviation of 10.4, while the ages of the healthy group of subjects range from 45 to 86 with mean 61.3 and standard deviation of 8.7. Moreover, the PD group contains data of 19 females and 41 males, while the data of healthy group contain 21 females and 79 males. The

data collection process was carried out using smart phones. The phone was kept at a distance of 10 cm from each subject during recording of the voice phonations. Each subject was asked to pronounce sustained phonations "a," "o," and "u." Consequently, the database contains $160 \times 3 = 480$ voice samples. Out of these 480 samples, 300 samples belong to healthy subjects and the remaining 180 samples belong to the patient group. The statistical information about the collected data have been reported in Table 1. Moreover, apart from using our own collected data, we also performed experiments on a bench mark dataset, namely, "multiple Types of Speech Dataset" [2].

3.2. Proposed Method. In this paper, we propose a three stage automated approach for PD detection. The first stage uses MFCC approach for feature extraction. The second stage is about dimensionality reduction through LDA, while the third stage is classification. In order to obtain better results, we explore the feasibility of various machine learning models at the third stage of the system. Hence, we developed ten different machine learning models. Based on the performance analysis, we pointed out that our proposed method, namely, MFCC-LDA-SVM approach, provides optimal PD detection. The proposed approach is depicted in Figure 1. The working of each stage of the proposed learning system is briefly discussed as follows.

3.2.1. Feature Extraction through MFCC. For extracting numerical features from the voice samples, we utilized the MFCC method. The MFFC algorithm establishes the relationship between perceived frequency and pitch of a pure tone as a function of its acoustic frequency. A subjective pitch is measured in the mel scale in units called mel. The mel for a given frequency f in Hz can be calculated using the following approximate formula [45]:

$$f_{\text{mel}} = 2595 \times \log_{10}\left(1 + \frac{f_{\text{Hz}}}{700}\right). \qquad (1)$$

Framing: according to [46], it takes a long period of time to examine the voice signals. This is because the voice signals are not stationary. Hence, it is necessary to move on with a short time analysis (generally, from 10 ms to 30 ms). The rate of movement of the voice articulators is limited by physiological limitations and can be considered stable within an interval from 10 to 30 ms. Therefore, the analysis of voice signal is carried out within uniform frames of this interval. In frame blocking, the voice signal is divided into frames of N samples. Neighboring frames should be separated by $M(M < N)$.

Pre-emphasis: in this step, we emphasize the higher frequencies by applying the first-order difference equation to the voice samples. This is to increase the energy in the voice signal.

The difference equation to voice signal $(S_n, n = 1, \ldots, N)$ is given in equation (2) [47] as follows:

$$s'_n = s_n - k \times s_{n-1}, \qquad (2)$$

TABLE 1: Summarized statistical information about the voice database.

Information	PD patients	Healthy subjects
No. of subjects	60	100
Age ($\mu \pm \sigma$)	68.3 ± 10.4	61.3 ± 8.7
Range of age	43–88	45–86
MDS-UPDRS-III ($\mu \pm \sigma$)	15.4 ± 4.7	—

where k is the pre-emphasis coefficient, and it should be within the range of $0 \leqslant k < 1$. Following the approach of [29], in this work, we used a pre-emphasis coefficient of $k = 0.97$.

Windowing: in order to minimize disrupts at the ends and make them continuous enough to correlate with the beginnings, windowing must be applied. Ideally, there exist several window functions (flat top window, hamming window, and rectangular window); however, the hamming window is used in our study for carrying out windowing. It is used to abate (decrease) signal to zero at the beginning and end of each frame and be represented as follows:

$$s_n' = \left\{ 0.54 - 0.46 . \cos\left(\frac{2\pi(n-1)}{N-1}\right) \right\} . s_n, \qquad (3)$$

where s_n is the voice samples and $n = 1, N$.

Fast Fourier transform: the main purpose of FFT is to have a look at frequency domain information when the given signal information are in time domain. For this purpose, we will have to convert into frequency domain each frame having N samples. Compared to DFT, i.e., discrete Fourier transform, FFT is a faster algorithm on the given set of N samples [46, 47]:

$$s_n = \sum_{k=0}^{N-1} s_k e^{-2\pi jkn/N}, \qquad (4)$$

where $n = 0, 1, 2, \ldots, N - 1$.

Mel scale/filter bank analysis: here, the approximation about the existing energy at each spot is determined. Thus, the spectrums calculated above are mapped on a mel scale using a triangular overlapping window, i.e., triangular filter bank (FB). The FB consists of a number of band pass filters with spacing along with bandwidth which is decided by steady mel frequency time. The mel frequency scale takes a linear spacing for frequency values below 1000 Hz and logarithmic spacing for values above 1000 Hz. To convert a given frequency (f) to a mel frequency (m_f), we used the approximate equation (1) [29].

Logarithm/DCT: with the intension of back conversion to spatial domain from the log mel spectrum, discrete cosine transform is brought into account for evaluating coefficients from the spectrum. Thus, we calculate the MFCC from the amplitudes of the log filter banks [15]:

$$c_i = \sqrt{\frac{2}{N}} \sum_{j=1}^{N} m_j . \cos\left(\frac{\pi i}{N}(j - 0.5)\right). \qquad (5)$$

Liftering: lack of correlation among the cepstral coefficients is the key advantage. However, the fact that the cepstral coefficients of higher order are fairly small is the main problem. Hence, rescaling of the coefficients is necessary in order to have quite similar magnitudes [29, 45]. There is, therefore, the need to apply liftering to the cepstral coefficients using the following equation:

$$c_n' = \left(1 + \frac{L}{2} \times \sin\left(\frac{\pi.n}{L}\right)\right) \times c_n, \qquad (6)$$

where L is the cepstral sine lifter parameter.

3.2.2. Linear Discriminant analysis (LDA).
LDA is a supervised ML technique that is mostly used for classification and dimensionality reduction. The working of LDA is based on linear transformation of data (features) into small dimensional space, for maximum discrimination between classes [48]. LDA, in machine learning, is search for the vectors based on linear combination of features in vector space that separates two or more classes. Furthermore, original data values are plotted on the vectors for evaluation of the classes division. When classes are overlapped on the particular data values, then transformation mechanism is adopted by the LDA for better separation of the classes. To achieve the better separation between the classes, LDA deploys a rule known as the Fisher ratio. The maximum value of the Fisher ratio means maximum distance between the two classes. Equation (7) is the formulation of the Fisher ratio:

$$\frac{(v_1 - v_2)^2}{\rho_1^2 + \rho_2^2}, \qquad (7)$$

where ρ_1 and ρ_2 denote the variance of 1st and 2nd class, while $(v_1 - v_2)$ is the difference between the means of the two classes. $\rho_1^2 + \rho_2^2$ is the sum of classes scatter. For example, δ_m tries to compact two classes by reducing $(v_1 - v_2)$ and δ_s tries to minimize the class scatter. For detailed formulation and discussion about LDA, readers can refer to [3].

LDA has the following two benefits. Firstly, the performance of the predictive model is enhanced by LDA through transforming the original feature dimension into reduced dimensional space, where the class division is maximized. Secondly, time complexity of the predictive model reduced tremendously by LDA. Reduced dimensionality data by the LDA are supplied to the SVM for classification.

3.2.3. Support Vector Machine.
Support vector machines (SVMs) are considered powerful learning methods and have been widely used in different biomedical- and health informatics-related problems [49]. During the training process, the output of an SVM model is an optimal hyperplane that could augment the distance of any class from the nearest training data points. The major reasons that motivate machine learning researchers to use SVM for their problems are as follows. (1) The first reason is that SVMs have powerful generalization capabilities to unseen data. (2) The second reason is the dependence of SVMs on a very small number of hyperparameters [50, 51].

Consider a dataset D_S with S instances, $D = \{(x_i, y_i)|x_i \in R^Q, y_i \in \{-1, 1\}\}_{i=1}^{S}$, where x_i stands for i^{th} instance, Q represents the dimension of the original features space of PD data, and y_i denotes the class labels, i.e., presence or absence of PD disease. The Q value is 20 for the PD dataset considered in this paper. The SVM model determines a hyperplane calculated by $f(x) = \theta^T * x + \delta$, where δ represents the bias and θ denotes the weight vector. Based on training data, the hyperplane $f(x)$ of the SVM model augments the margin whereas curtails (reduces) the classification error. Sum of the distances to one of the closest negative and one of the closest positive instances is regarded as margin. The margin is defined as the sum of the distances between the closest negative and closest positive instances. That is, the hyperplane augments the margin distance $2/\|\theta\|_2^2$.

SVM uses a set of lax variables denoted by $\xi_i, i = 1, \ldots, S$, and a penalty parameter, i.e., $\mathbb{C}$, and attempts to parity the minimization of $\|\theta\|_2^2$ and minimization of the misclassification errors. This fact is formulated as follows:

$$\min_{\theta, \delta, \xi} \underbrace{\frac{1}{2}\|\theta\|_2^2}_{\text{Regularizer}} + \mathbb{C} \underbrace{\sum_{i=1}^{S} \xi_i}_{\text{ErrororLoss}} \quad s.t \begin{cases} y_i(\theta x_i + \delta) \geq 1 - \xi_i \\ \\ \xi_i \geq 0, i = 1, \ldots, S \end{cases}. \tag{8}$$

In equation (8), ξ is lax variable that calibrates the degree of misclassification and Euclidean norm or L_2-norm is the penalty term.

4. Validation and Evaluation of the Proposed Approach

In order to validate the effectiveness of the proposed approach, we utilized leave-one-subject-out (LOSO) validation scheme in which the data of the one subject (all samples) are left out for testing and the proposed framework is trained on the remaining data. The process is repeated till the point where all the subjects have been tested. At the end, the final accuracy of the model is evaluated by calculating the mean accuracy for all the subjects.

To evaluate the performance of the proposed framework, we utilize some well-known statistical metrics, namely, Mathews Correlation Coefficient (MCC), sensitivity, specificity, and classification accuracy. Classification accuracy gives the precision with which the proposed method can classify all subjects (including patients and healthy). On the contrary, specificity tells us about how precise the model can classify healthy subjects and sensitivity tells us about how precise the developed model can classify patients. If A denotes the number of true positives, B denotes the number of true negatives, C denotes the number of false positives, and D denotes the number of false negatives, then the formulation of these evaluation metrics is given in equations (9)–(12):

$$\text{Acc} = \frac{A + B}{A + B + C + D}, \tag{9}$$

$$\text{Sn} = \frac{A}{A + D}, \tag{10}$$

$$\text{Sp} = \frac{B}{B + C}, \tag{11}$$

$$\text{MCC} = \frac{A \times B - C \times D}{\sqrt{(A + C)(A + D)(B + C)(B + D)}}, \tag{12}$$

where MCC is a value in the range −1 to 1, where −1 denotes the worse case and 1 denotes the best case.

5. Experiment Results

In this section, we discuss the implementation details and the obtained performance of different developed machine learning models for the problem of PD detection based on the voice data. All the experiments were performed using Intel (R) Core (TM) m3-7Y30 CPU @ 1.00 GHz 1.61 Ghz with memory of 8 GB and operating system of 64 bit Windows. All the experiments were performed using Python programming package and scikit-learn library.

The first experiment was performed by extracting the MFCC features from the voice phonations. The extracted MFCC was in the form of a matrix for each voice phonation. The matrix contained 20 columns which act as MFCC features. Following the approach of previous studies, we evaluated mean for each column or MFCC feature along the rows of the matrix. In this way, we obtained a feature vector of size equal to 20 for each voice phonation. Next, we used iterative feature selection before application of LDA for dimensionality reduction. After dimensionality reduction through the LDA model, we applied the resultant feature vectors at the input of machine learning models. The results for each of the developed machine learning models are given in Table 2.

After observing the results given in Table 2, it can be seen that the worst performance was produced by the GNB model and SVM with sigmoid kernel which are 48.12% accuracy and 46.87%, respectively, while the optimal performance is produced by the SVM model with RBF kernel which is 77.5% accuracy, 84% specificity, and 74.33% sensitivity. It means the proposed MFCC-LDA-SVM model can correctly classify 124 subjects out of the total 160 subjects. Similarly, the specificity value of 80% reveals that out of 100 healthy subjects, 80 are correctly classified, while the sensitivity rate of 73.33 reveals the fact that out of 60 PD patients, the proposed model can successfully detect 44 PD patients correctly. These statistical results are more clearly depicted in the confusion matrix given in Figure 2.

The performance of the MFCC-LDA-SVM model is further evaluated in terms of area under the curve (AUC) matrix which was calculated from the receiver operating characteristic curve (ROC curve). The ROC curve for the two models with worse performance and the ROC curve for the two models with optimal performance is given in Figures 3 and 4, respectively. It is important to note that a model with higher AUC is decided as a much better model than those models which are having lower values of AUC. Based on these evaluation criteria, we can see in the Figures 3 and 4 that the proposed MFCC-LDA-SVM is an optimal model when compared with other developed models. Additionally,

TABLE 2: Evaluation of different models in terms of PD detection based on voice data coeff, MFCC coefficients, selected: size of subset of features, Acc (%): percentage of correctly classified subjects, MCC: Mathews correlation coefficient, Sn: sensitivity and Sp: specificity, and HP: hyperparameters

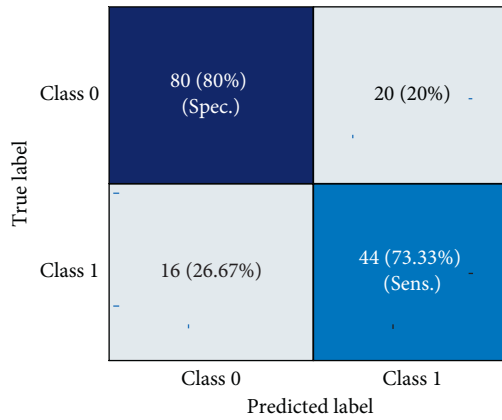| Method	Coefficients	Acc (%)	Sp (%)	Sn (%)	Hyperparameters
LR	7	50.00	60.00	33.33	$C = 0.0001$
DT	12	51.25	60.00	36.66	$d = 10$, $l = 22$
GNB	1	48.12	60.00	28.33	—
LDA	15, 17, 19	62.25	60	66	—
NCC	11	50.00	50	50	—
ANN	11, 17, 18, 20	65.62	85	33	$H = 100$
SVM (Lin)	10, 15, 16	68.75	90	33.33	$C = 10$
SVM (Pol)	8	63.75	90	20	degree = 7
SVM (Sig)	14, 15	46.87	70	8.3	$C = 0.001$
SVM (RBF)	10, 12, 18	77.50	84	73.33	$C = 0.01$, $G = 0.0001$

FIGURE 2: Graphical depiction of statistics of the obtained results for the proposed diagnostic system in terms of confusion matrix. Spec: Specificity and Sens: Sensitivity.
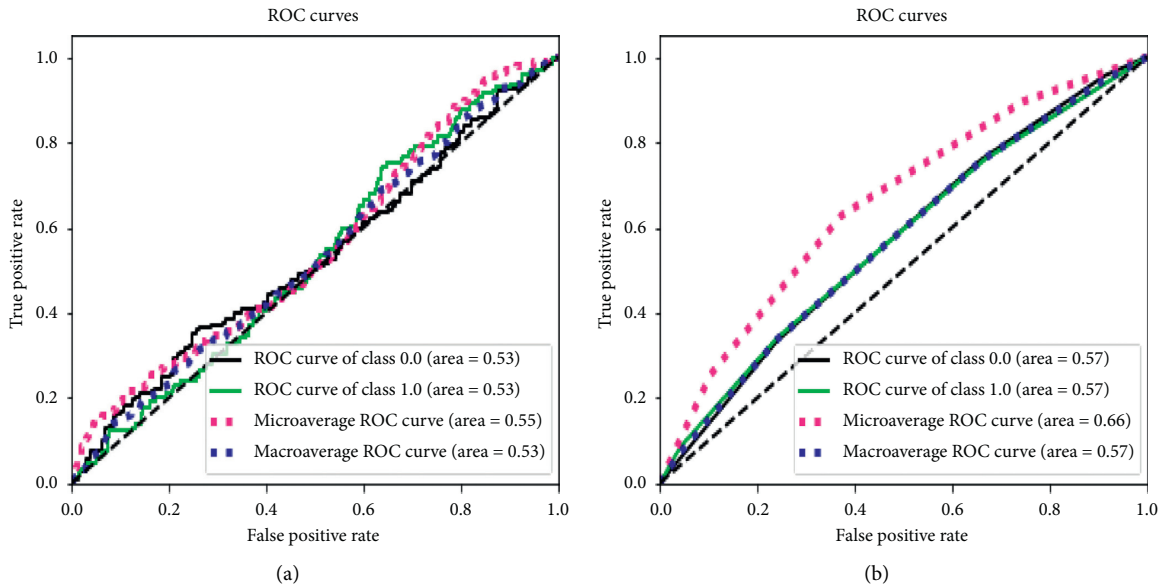

(a)

(b)

FIGURE 3: ROC charts of the two models with least performance. (a) ROC chart of the GNB model. (b) ROC chart of the LR model.

for further validation of the proposed approach, it is compared with recently published studies shown in Table 3.

The data were collected by different individuals who had different smart phones for recording the voice data. It is a well-known fact that spectral characteristics of the microphone can highly influence the results, especially considering that MFCCs have been used in the study. These factors can degrade the performance of the proposed intelligent
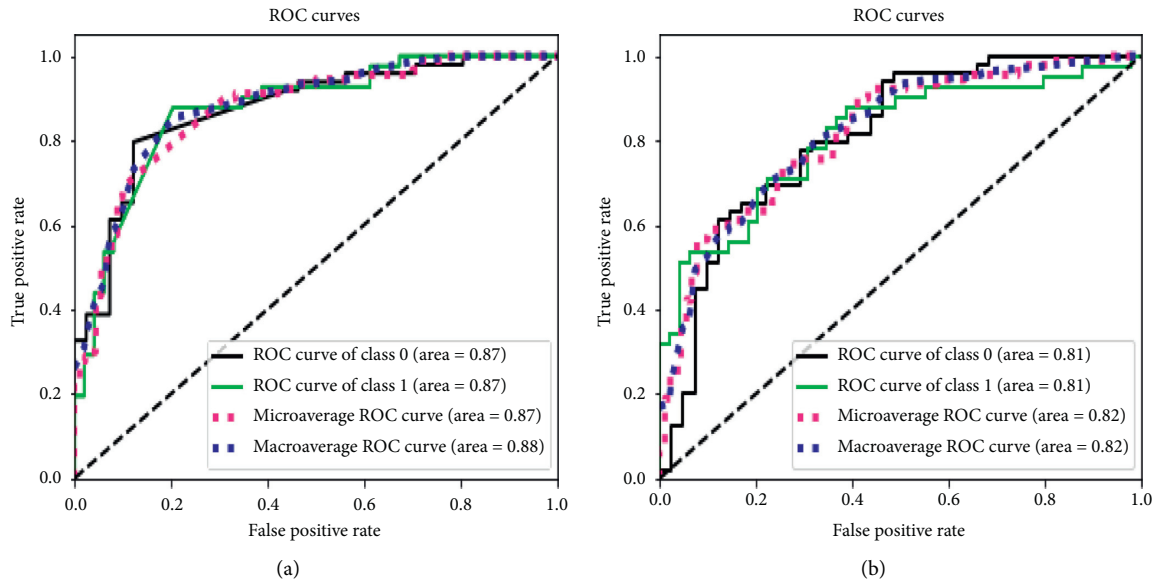
(a)

(b)

Figure 4: ROC charts of the two models with higher performance. (a) ROC chart of the SVM RBF model. (b) ROC chart of the SVM linear model.

Table 3: Performance comparison with recently published work.

Study	Method	Acc (%)	Sen. (%)	Spec. (%)
Sarkar et al. [42]	KNN + SVM	55.00 (LOSO on training database), 68.45 (LOSO on testing database)	60 (Training database)	50 (training database)
Canturk and Karabiber [43]	4 feature selection methods + 6 classifiers	57.5 (LOSO CV), 68.94 (10-fold)	54.28 (LOSO), 70.57 (10-fold)	80 (LOSO), 66.92 (10-fold)
Eskidere et al. [39]	Random subspace classifier ensemble	74.17 (10-fold CV)	Did not report	Did not report
Vadovský and Parali [40]	C4.5 + C5.0 + random forest + CART	66.5 (4-fold CV with pronouncing numbers), 65.86 (5-fold CV with pronouncing numbers)	Did not report	Did not report
Kraipeerapun and Amornsamankul [41]	Stacking + complementary neural Networks (CMTNN)	Average 75% (10-fold CV)	Did not report	Did not report
Ali et al. [44]	Multimodal approach	70	Not reported	Not reported
Benba et al. [29]	HFCC-SVM	87.5	90.00	85.00
Li et al. [32]	SVM + FS	82.50	85.00	80.00
Ali et al. [3]	LDA–NN–GA	95.00	95.00	95.00
Proposed method	MFCC-LDA-SVM	**97.5% (LOSO CV)**	**100.0%**	**97.5%**

system. Furthermore, the shorter length of phonations in PD could be another factor influencing cepstral analysis. To check the strength of our model, we simulated the same model on a publicly available dataset, namely, "Multiple Types of Speech Dataset" [42]. The proposed intelligent system, i.e., MFCC-LDA-SVM obtained outstanding results on the publicly available dataset. Using LOSO CV on the training dataset of the "Multiple Types of Speech Dataset," we obtained 97.5% of accuracy, 100% sensitivity, and 95% specificity. Similarly, the proposed intelligent method produced accuracy of 89.28% on the testing dataset of the "Multiple Types of Speech Dataset."

6. Conclusion

In this study, we considered the challenge of PD detection based on multiple types of voice signals. From each subject,

we recorded three different voice phonations. Signal processing algorithm (MFCC) was utilized to extract numerical features from the voice phonations. The extracted MFFC features were dimensionality reduced through the application of the linear discriminant analysis (LDA) model. At the final stage, numerous machine learning models were developed. It was pointed out that the MFCC-LDA-SVM method produces optimal performance in terms of PD detection. The performance comparison was carried out using different evaluation criteria including classification accuracy, area under the curve (AUC), and receiver operating characteristics curve. The proposed method produced AUC of 87%, PD detection accuracy of 78.5%, sensitivity of 73.33%, and specificity of 80%. Moreover, the proposed intelligent system was also simulated on the publicly available dataset. The obtained results were promising compared to the previous work.

Data Availability

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] L. M. de Lau and M. M. Breteler, "Epidemiology of Parkinson's disease," *The Lancet Neurology*, vol. 5, no. 6, pp. 525–535, 2006.

[2] L. Ali, C. Zhu, M. Zhou, and Y. Liu, "Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection," *Expert Systems with Applications*, vol. 137, pp. 22–28, 2019.

[3] L. Ali, C. Zhu, Z. Zhang, and Y. Liu, "Automated detection of Parkinson's disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 7, pp. 1–10, 2019.

[4] S. Arora, F. Baig, C. Lo et al., "Smartphone motor testing to distinguish idiopathic rem sleep behavior disorder, controls, and pd," *Neurology*, vol. 91, no. 16, pp. e1528–e1538, 2018.

[5] J. Rusz, J. Hlavnička, T. Tykalová et al., "Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder," *Sleep Medicine*, vol. 19, pp. 141–147, 2016.

[6] J. Rusz, M. Novotny, J. Hlavnivcka, T. Tykalova, and E. Ruuvzivcka, "High-accuracy voice-based classification between patients with Parkinson's disease and other neurological diseases may be an easy task with inappropriate experimental design," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 8, pp. 1319–1321, 2016.

[7] J. Rusz, J. Hlavnicka, T. Tykalova et al., "Smartphone allows capture of speech abnormalities associated with high risk of developing Parkinson's disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 8, pp. 1495–1507, 2018.

[8] L. Cunningham, S. Mason, C. Nugent, G. Moore, D. Finlay, and D. Craig, "Home-based monitoring and assessment of Parkinson's disease," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 1, pp. 47–53, 2011.

[9] Z. A. Dastgheib, B. Lithgow, and Z. Moussavi, "Diagnosis of Parkinson's disease using electrovestibulography," *Medical & Biological Engineering & Computing*, vol. 50, no. 5, pp. 483–491, 2012.

[10] G. Rigas, A. T. Tzallas, M. G. Tsipouras et al., "Assessment of tremor activity in the Parkinson's disease using a set of wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 478–487, 2012.

[11] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, L. O. Ramig et al., "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1015–1022, 2009.

[12] S. K. Van Den Eeden, C. M. Tanner, A. L. Bernstein et al., "Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity," *American Journal of Epidemiology*, vol. 157, no. 11, pp. 1015–1022, 2003.

[13] R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1568–1572, 2010.

[14] L. Parisi, N. RaviChandran, and M. L. Manaog, "Feature-driven machine learning to improve early diagnosis of Parkinson's disease," *Expert Systems with Applications*, vol. 110, pp. 182–190, 2018.

[15] L. Naranjo, C. J. Pérez, J. Martín, and Y. Campos-Roca, "A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications," *Computer Methods and Programs in Biomedicine*, vol. 142, pp. 147–156, 2017.

[16] L. Ali, I. Wajahat, N. Amiri Golilarz, F. Keshtkar, and S. A. C. Bukhari, "LDA-GA-SVM: improved hepatocellular carcinoma prediction through dimensionality reduction and genetically optimized support vector machine," *Neural Computing and Applications*, 2020.

[17] T. Meraj, A. Hassan, S. Zahoor et al., "Lungs nodule detection using semantic segmentation and classification with optimal features," *Neural Computing and Applicationsdoi*, 2019.

[18] L. Ali and S. Bukhari, "An approach based on mutually informed neural networks to optimize the generalization capabilities of decision support systems developed for heart failure prediction," *IRBM*, 2020, In press.

[19] L. Ali, S. U. Khan, N. A. Golilarz et al., "A feature-driven decision support system for heart failure prediction based on χ^2 statistical model and Gaussian naive bayes," *Computational and Mathematical Methods in Medicine*, vol. 2019, Article ID 6314328, 8 pages, 2019.

[20] L. Ali, C. Zhu, N. A. Golilarz, A. Javeed, M. Zhou, and Y. Liu, "Reliable Parkinson's disease detection by analyzing handwritten drawings: construction of an unbiased cascaded learning system based on feature selection and adaptive boosting model," *IEEE Access*, vol. 7, pp. 116480–116489, 2019.

[21] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, "Novel speech signal processing algorithms for high-accuracy classification of parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 5, pp. 1264–1271, 2012.

[22] E. Kaya, O. Findik, I. Babaoglu, and A. Arslan, "Effect of discretization method on the diagnosis of Parkinson's disease," *International Journal of Innovative Computing, Information and Control*, vol. 7, pp. 4669–4678, 2011.

[23] I. Mandal and N. Sairam, "Accurate telemonitoring of Parkinson's disease diagnosis using robust inference system," *International Journal of Medical Informatics*, vol. 82, no. 5, pp. 359–377, 2013.

[24] M. Hariharan, K. Polat, and R. Sindhu, "A new hybrid intelligent system for accurate detection of Parkinson's disease," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 3, pp. 904–913, 2014.

[25] N. A. Bhalchandra, R. Prashanth, S. D. Roy, and S. Noronha, "Early detection of Parkinson's disease through shape based

features from [123] I-Ioflupane SPECT imaging," in *Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 963–966, IEEE, Brooklyn, NY, USA, April 2015.

[26] R. Saloni and A. Gupta, "Detection of Parkinson disease using clinical voice data mining," *International Journal of Circuits, Systems and Signal Processing*, vol. 9, 2015.

[27] L. Huang, Y. Jin, Y. Gao et al., "Longitudinal clinical score prediction in Alzheimer's disease with soft-split sparse regression based random forest," *Neurobiology of Aging*, vol. 46, pp. 180–191, 2016.

[28] A. H. Al-Fatlawi, M. H. Jabardi, and S. H. Ling, "Efficient diagnosis system for Parkinson's disease using deep belief network," in *Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1324–1330, IEEE, Vancouver, Canada, July 2016.

[29] A. Benba, A. Jilbab, and A. Hammouch, "Using human factor cepstral coefficient on multiple types of voice recordings for detecting patients with Parkinson's disease," *IRBM*, vol. 38, no. 6, pp. 346–351, 2017.

[30] E. Vaiciukynas, A. Verikas, A. Gelzinis, and M. Bacauskiene, "Detecting Parkinson's disease from sustained phonation and speech signals," *PLoS One*, vol. 12, no. 10, Article ID e0185613, 2017.

[31] L. Naranjo, C. J. Pérez, and J. Martín, "Addressing voice recording replications for tracking Parkinson's disease progression," *Medical & Biological Engineering & Computing*, vol. 55, no. 3, pp. 365–373, 2017.

[32] Y. Li, C. Zhang, Y. Jia, P. Wang, X. Zhang, and T. Xie, "Simultaneous learning of speech feature and segment for classification of Parkinson disease," in *Proceedings of the 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–6, IEEE, Dalian, China, October 2017.

[33] Y. Zhang, "Can a smartphone diagnose parkinson disease? a deep neural network method and telediagnosis system implementation," *Parkinson's Disease*, vol. 2017, Article ID 6209703, 11 pages, 2017.

[34] S. S. Upadhya, A. N. Cheeran, and J. H. Nirmal, "Thomson multitaper MFCC and PLP voice features for early detection of Parkinson disease," *Biomedical Signal Processing and Control*, vol. 46, pp. 293–301, 2018.

[35] K. Wu, D. Zhang, G. Lu, and Z. Guo, "Learning acoustic features to detect Parkinson's disease," *Neurocomputing*, vol. 318, pp. 102–108, 2018.

[36] M. M. Khan, A. Mendes, and S. K. Chalup, "Evolutionary wavelet neural network ensembles for breast cancer and Parkinson's disease prediction," *PLoS One*, vol. 13, no. 2, Article ID e0192192, 2018.

[37] D. Braga, A. M. Madureira, L. Coelho, and R. Ajith, "Automatic detection of Parkinson's disease based on acoustic analysis of speech," *Engineering Applications of Artificial Intelligence*, vol. 77, pp. 148–158, 2019.

[38] S. A. Mostafa, A. Mustapha, M. A. Mohammed et al., "Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease," *Cognitive Systems Research*, vol. 54, pp. 90–99, 2019.

[39] Ö. Eskıdere, A. Karatutlu, and C. Ünal, "Detection of Parkinson's disease from vocal features using random subspace classifier ensemble," in *Proceedings of the 2015 Twelve International Conference on Electronics Computer and Computation (ICECCO)*, pp. 1–4, IEEE, Almaty, Kazakhstan, September 2015.

[40] M. Vadovskỳ and J. Paralič, "Parkinson's disease patients classification based on the speech signals," in *Proceedings of the 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, p. 000321, January 2017, Article ID 000326.

[41] P. Kraipeerapun and S. Amornsamankul, "Using stacked generalization and complementary neural networks to predict Parkinson's disease," in *Proceedings of the 2015 11th International Conference on, Natural Computation (ICNC)*, pp. 1290–1294, IEEE, Zhangjiajie, China, August 2015.

[42] B. E. Sakar, M. E. Isenkul, C. O. Sakar et al., "Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 4, pp. 828–834, 2013.

[43] İ. Cantürk and F. Karabiber, "A machine learning system for the diagnosis of Parkinson's disease from speech signals and its application to multiple speech signal types," *Arabian Journal for Science and Engineering*, vol. 41, no. 12, pp. 5049–5059, 2016.

[44] L. Ali, S. U. Khan, M. Arshad, S. Ali, and M. Anwar, "A multi-model framework for evaluating type of speech samples having complementary information about Parkinson's disease," in *Proceedings of the 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pp. 1–5, IEEE, Swat, Pakistan, July 2019.

[45] A. Benba, A. Jilbab, A. Hammouch, and S. Sandabad, "Voiceprints analysis using MFCC and SVM for detecting patients with Parkinson's disease," in *Proceedings of the 2015 International Conference on Electrical and Information Technologies (ICEIT)*, pp. 300–304, IEEE, Marrakech, Morocco, March 2015.

[46] C. S. Kumar and P. M. Rao, "Design of an automatic speaker recognition system using MFCC, vector quantization and LBG algorithm," *International Journal on Computer Science and Engineering*, vol. 3, no. 8, p. 2942, 2011.

[47] S. Gupta, J. Jaafar, W. F. wan Ahmad, and A. Bansal, "Feature extraction using mfcc," *Signal & Image Processing: An International Journal*, vol. 4, no. 4, pp. 101–108, 2013.

[48] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Prediction, Inference and Data Mining*, Springer-Verlag, New York, NY, USA, 2001.

[49] F. S. Ahmad, L. Ali, Raza-Ul-Mustafa et al., "A hybrid machine learning framework to predict mortality in paralytic ileus patients using electronic health records (EHRs)," *Journal of Ambient Intelligence and Humanized Computing*, 2020.

[50] S. Maldonado, J. Pérez, R. Weber, and M. Labbé, "Feature selection for support vector machines via mixed integer linear programming," *Information Sciences*, vol. 279, pp. 163–175, 2014.

[51] L. Ali, A. Niamat, J. A. Khan et al., "An optimized stacked support vector machines based expert system for the effective prediction of heart failure," *IEEE Access*, vol. 7, pp. 54007–54014, 2019.