

Machine Learning nella diagnosi differenziale tra appendicite acuta non complicata e complicata in età pediatrica: sviluppo e validazione di un modello predittivo integrato

Autore: Giorgio CAVATOI ⁽¹⁾, Giulia MAISANO ⁽¹⁾, Fabiola CASSARO ^(1;2), Nunzio TURLACO ⁽¹⁾, Giovanni CICCERI ⁽³⁾, Salvatore VITABILE ⁽³⁾, Pietro ANTONUCCIO ⁽¹⁾, Pietro IMPELLIZZERI ⁽¹⁾, Salvatore ARENA ⁽¹⁾, Carmelo ROMEO ⁽¹⁾.

Affiliazione: ⁽¹⁾ U.O.C. Chirurgia Pediatrica, Dipartimento di Patologia Umana e dell'Età Evolutiva "G. Barresi", Università di Messina, Messina, Italia; ⁽²⁾ Dipartimento di Scienze Biomediche, Odontoiatriche e delle Immagini Morfologiche e Funzionali, Università di Messina, Messina, Italia; ⁽³⁾ Dipartimento di Biomedicina, Neuroscienze e Diagnostica avanzata (Bi.N.D.), Università di Palermo, Palermo, Italia.

Keywords: Appendicite acuta, Intelligenza artificiale, Machine Learning.

Razionale dello studio

L'appendicite acuta è la più frequente urgenza chirurgica pediatrica. L'algoritmo terapeutico tra forma "non complicata" e "complicata" può essere differente. Una classificazione precoce e accurata risulta essere fondamentale per migliorare gli outcomes clinici, ottimizzando le scelte terapeutiche. Errori o ritardi diagnostici potrebbero compromettere la prognosi e aumentare la morbidità. In questo contesto, l'integrazione di modelli di apprendimento automatico (ML) può offrire un valido supporto decisionale nella stadiazione dell'infezione.

Obiettivi

Lo studio si è proposto di sviluppare e validare un approccio diagnostico basato su ML, finalizzato alla classificazione accurata dei casi di appendicite acuta in pazienti pediatrici. La ricerca ha integrato in modo sistematico parametri clinici, dati di laboratorio e reperti ecografici, con l'obiettivo di creare uno strumento predittivo in grado di supportare i chirurghi nell'iter diagnostico.

Materiali e Metodi

È stato condotto uno studio retrospettivo su 387 pazienti pediatrici sottoposti ad appendicectomia con diagnosi istologica di appendicite acuta, classificata in "non complicata" e "complicata". Sono stati raccolti parametri anamnestici, clinici, laboratoristici ed ecografici. Tre modelli di ML (Random Forest, SVM, KNN) sono stati confrontati mediante validazione incrociata ripetuta (metodo StratifiedShuffleSplit), basata su 10 suddivisioni casuali bilanciate tra le classi cliniche, al fine di stimare con robustezza le performance predittive. Le performance sono state valutate mediante le metriche standard di accuratezza, precisione, recall e F1-score. È stata inoltre implementata un'analisi di interpretabilità, che ha permesso di quantificare l'impatto medio di ciascuna variabile sulla decisione del miglior modello e di costruire un ranking delle features più rilevanti.

Risultati

I risultati hanno dimostrato che l'algoritmo Random Forest ha raggiunto un'accuratezza media di 80.9% ($\pm 4.3\%$), una precisione media di 82.1 ($\pm 4.8\%$), una recall media di 86% ($\pm 3.9\%$) e un F1-score media di 83.9% ($\pm 3.3\%$). La particolare efficacia nel ridurre i falsi negativi (recall) rende questo modello particolarmente adatto all'uso clinico, dove la capacità di identificare correttamente i casi positivi è cruciale. La metodica ha evidenziato una robusta rilevanza dei marcatori infiammatori (P.C.R, globuli bianchi) e dei parametri ecografici dell'appendice nell'algoritmo diagnostico di ML.

Conclusioni

L'uso di modelli di ML basati su dati clinici, di laboratorio ed ecografici, può rendere affidabile la diagnosi differenziale tra appendicite acuta non complicata e complicata, fornendo un supporto diagnostico affidabile ed interpretabile, potenzialmente migliorando la tempistica nelle decisioni chirurgiche, riducendo le complicanze. Nella gerarchia decisionale, i marcatori infiammatori risultano essere le variabili più predittive nella diagnosi differenziale.