

Research on supporting tuberculosis diagnosis based on machine learning

Nguyen Huu Khanh

Abstract: This study presents a machine learning-based approach to support the diagnosis of tuberculosis (TB), addressing the critical need for rapid and accurate detection in resource-limited settings. Leveraging a dataset of clinical features and laboratory results, we developed and evaluated several classification models, including support vector machines, random forests, and gradient boosting algorithms. Feature selection techniques were employed to identify the most predictive indicators, such as patient demographics, symptom profiles, and sputum smear results. The models were trained and validated using cross-validation to ensure robustness and generalizability. Performance was assessed using metrics including accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve. The results demonstrate that the proposed machine learning framework achieves high diagnostic accuracy, significantly reducing the time required for preliminary screening compared to conventional methods. By integrating clinical data with computational analysis, this research offers a reliable, cost-effective decision-support tool for healthcare professionals, potentially improving early detection rates and treatment outcomes for tuberculosis patients.

Keywords: machine learning, tuberculosis diagnosis, support system, medical diagnosis