MIDI 2025 Conference

A Clustering-Based Machine Learning Approach for Developing a Unified Metric to Assess Post-COVID-19 Rehabilitation Outcomes

Abstract

Assessing rehabilitation outcomes in post-COVID-19 patients is challenging due to multiple health factors and the lack of standardized evaluation procedures.

This paper proposes a universal metric for machine learning applications, representing patient outcomes in an interpretable three-dimensional space: Subjective Measures and Physical Fitness, Oxygenation, and Exercise Tolerance. The method flexibly integrates both subjective and objective measures and does not rely on a fixed set of tests. Feature standardization is applied in two stages, followed by K-Means clustering to group patients based on rehabilitation outcomes. The approach provides transparent and interpretable clustering results, bridging complex machine learning analysis with clinical understanding. By combining clinical relevance with interpretable machine learning, the method supports personalized rehabilitation strategies, identifies patient patterns, and enhances treatment outcome prediction. The study is based on 898 patients treated at the Specialist Hospital of St. John Paul II, Głuchołazy, Poland.