An Uncertainty-Aware Predictive Modeling Tool Towards Effective Clinical Monitoring of Sepsis Using Electronic Health Records From Hospitalized Patients

Dalal, Shaunak (School: Hershey High School)

Sepsis—a global health priority declared by the World Health Organization—is a life-threatening immune response to infection and the cause of 20% of all deaths. Early recognition is crucial, yet traditional diagnostics lack predictive capability, and machine learning models struggle with irregular data, false alarms, and generalizability outside the Intensive Care Unit (ICU). This study aimed to develop a deep learning tool for predicting sepsis early that adapts to hospitalized patient data over time while minimizing uncertainty. A novel recurrent neural network (RNN)-based model was trained on the MIMIC-IV dataset of 83,813 patients to analyze longitudinal pathophysiological data and estimate sepsis likelihood. A conformal prediction algorithm was then built to assign indeterminate classifications to uncertain cases in testing by assessing whether a new patient fit within the training distribution. The system was externally validated on the eICU-CRD dataset with 82,486 patients to assess performance across hospital settings. Predictions at 24, 12, and 6 hours before sepsis onset achieved accuracies exceeding 96% at each time window. The predicted sepsis probabilities and indeterminate identifications aligned with true clinical status ($p < 0.05$). External validation confirmed similar performance, along with a notable 72% false alarm reduction at the 12-hour window, supporting practical use in low-monitoring environments and diverse patient populations. The proposed tool may enable predictive monitoring of sepsis onset with minimal temporal data while potentially improving resource allocation in hospital settings, which may enhance timely interventions, reduce unnecessary ICU admissions, and contribute to improved patient care.

Awards Won: