IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AI-Powered Early Sepsis Detection System For ICU With Clinical Explainability

[1]Patan Muskan, [2] Rabiya Khan, [3]Munazira Banu,[4] Prajwal Shiragumpi, [5]Krishna Meena

[1]Student, [2]Student, [3]Student, [4]Student, [5]Assistant Professor

[1]Department of Computer Science and Engineering,
[1]HKBK College of Engineering, Bangalore, India

Abstract: Sepsis is a life-threatening condition resulting from an infection that leads to systemic inflammation and organ dysfunction, often occurring in critically ill ICU patients. Timely detection is vital to enhance patient outcomes, as delayed intervention can result in serious complications and higher mortality rates. This paper proposes an AI-powered early sepsis detection system that leverages advanced machine learning models, specifically Long Short-Term Memory (LSTM) networks and Temporal Convolutional Networks (TCN), to predict sepsis in ICU patients based on clinical data from the MIMIC-III/IV datasets. The system integrates Explainable AI techniques, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), to provide transparent insights into prediction factors, fostering clinician trust and informed decision-making. A real-time dashboard offers actionable alerts and visualizations, aiding clinicians in timely interventions. By improving prediction accuracy, optimizing resource allocation, and reducing ICU mortality rates, the system demonstrates significant potential for enhancing clinical outcomes and generating cost savings in critical care settings.

Index Terms - Sepsis detection, Intensive Care Unit (ICU), Machine Learning, Deep Learning, Long Short-Term Memory (LSTM), Temporal Convolutional Network (TCN), Explainable AI (XAI), SHAP, LIME, MIMIC-III, MIMIC-IV, Clinical Decision Support.

I. INTRODUCTION

Sepsis is a fatal condition occurring when the body gets overwhelmed and responds uncontrolled to an infection. It remains a significant challenge in modern healthcare, especially within Intensive Care Units (ICUs). Despite advancements in medical science, sepsis is still linked with high rates of illness, death, and rising healthcare costs, affecting not only patients but also their families and the entire healthcare system [1], [2]. One of the major hurdles in managing sepsis is its complex and unpredictable nature, which often leads to delayed diagnosis and treatment—factors that can have devastating consequences for patients [3], [4].

Early detection is critical. Timely intervention can prevent severe complications, lower mortality rates, and significantly improve patient outcomes. However, the early signs of sepsis are often subtle and easy to miss, making it difficult for even experienced clinicians to respond quickly and effectively [2], [5]. This is where technology, such as artificial intelligence, plays a paradigm-shifting role [6].

With recent advancements in AI and machine learning, there is growing potential to support clinicians by rapidly analyzing vast amounts of clinical data, identifying patterns, and predicting patient deterioration before it becomes critical [1], [7]. By using sophisticated models such as Long Short-Term Memory (LSTM) networks and Temporal Convolutional Networks (TCNs), along with explainable AI techniques like SHAP and LIME, we can build systems that not only detect sepsis early but also offer insights into why and how decisions are made [2], [3].

This paper introduces a novel AI-driven approach to early sepsis detection—one that combines cutting-edge algorithms with explainability to support real-time clinical decision-making in the ICU. The goal is to improve nursing care, maximize the efficiency of healthcare resource utilization, and ultimately save lives. By integrating AI into the frontline of critical care, we move closer to a future where sepsis no longer claims so many lives, and where technology and human expertise work hand in hand to deliver smarter, faster, and more compassionate care [4], [7].

Table 1: Model Performance on MIMIC-III and MIMIC-IV Datasets

Model	Dataset	Accuracy	Precision	Recall	F1-Score	AUROC
LSTM	MIMIC-III	0.85	0.83	0.78	0.80	0.89
TCN	MIMIC-III	0.87	0.85	0.81	0.83	0.93
LSTM	MIMIC-IV	0.84	0.82	0.77	0.79	0.90
TCN	MIMIC-IV	0.88	0.86	0.83	0.84	0.94

Methodologies

Source and Preprocessing

We utilize the MIMIC-III and MIMIC-IV databases, publicly available datasets containing de-identified health records of over 60,000 ICU patients. These datasets include a comprehensive range of clinical variables such as vital signs, lab results, medication
records, and demographic information. The following preprocessing steps were applied:

- **Patient Selection:** Adult patients (18 years) were selected. Only ICU stays with sufficient temporal granularity and relevant sepsis-related variables were included.
- **Labeling:** Sepsis onset times were annotated using the Sepsis-3 criteria, which define sepsis as suspected infection plus an acute increase of 2 SOFA (Sequential Organ Failure Assessment) points. Each time-series was labeled with a binary indicator representing the presence or absence of sepsis within a future prediction window (e.g., 6 hours).
- **Missing Data Handling:** Time-series gaps were addressed using forward and backward imputation. For non-temporal data, median imputation or mode imputation (for categorical variables) was used.
- **Time Windowing:** Data were segmented into hourly intervals up to 48 hours before sepsis onset or ICU discharge, allowing the models to capture progression trends over time.
- **Normalization:** Continuous features were normalized using z-score standardization to stabilize training and minimize scale-related bias.

Feature Engineering

To build a reliable and clinically meaningful prediction model for early sepsis detection, we constructeda rich and well-curated set of features. These features were drawn from both static (non-temporal) and dynamic (temporal) sources, enabling the model to learn from both baseline patient characteristics and evolving physiological trends:

Static Features: These include patient-specific attributes that remain constant throughout the ICU stay, such as age, gender, type of admission (e.g., elective or emergency), and pre-existing comorbidities like diabetes or hypertension. These variables were encoded as categorical or binary features to ensure compatibility with the model input.

• **Temporal Features:** These are clinical measurements recorded over time, typically at hourly intervals. We included vital signs such as heart rate, systolic/diastolic/mean blood pressure, respiratory rate, body temperature, oxygen saturation (SpO2), white blood cell count, lactate levels, and other lab values. These variables capture real-time physiological fluctuations and are key indicators of a patients clinical trajectory.

• **Derived Variables:** To enhance the model's ability to detect subtle signs of deterioration, we engineered additional features such as rolling averages, first- and second-order differences (deltas), and slopes representing the rate of change in key vitals and labs. These derived features help highlight emerging patterns that might precede clinical signs of sepsis.

Explainable AI (XAI) Integration

To address the "black box" nature of deep learning and foster clinician trust, we integrated modelagnostic interpretability tools that help make our system's predictions transparent and understandable to healthcare professionals:

• **SHAP (SHapley Additive exPlanations):**SHAP determines the contribution of each feature by assigning it an importance score based on its influence on the model's prediction. At a global level, SHAP identifies which variables most influence the model across the entire dataset. On a local level, patient-specific SHAP plots clearly illustrate how factors like heart rate, lab values, or blood pressure contributed to a particular prediction, helping clinicians visualize what drove the model's decision.

• **LIME (Local Interpretable Model-Agnostic Explanations):** LIME builds simple, interpretable surrogate models around individual predictions. These models highlight the most influential features for each specific case—such as sudden changes in vitals or elevated lactate level offering an easily understandable rationale for the model's output.

Model Evaluation and Validation

To assess performance, we used both standard classification metrics and clinical utility measures:

• **Classification Metrics:** Accuracy, precision, recall, F1-score, AUROC (Area Under the Receiver Operating Characteristic Curve), and AUPRC (Area Under the Precision-Recall Curve).

• **Timeliness Metrics:** Early Detection Score (EDS), defined as how many hours before clinical onset the model correctly identified sepsis.

• **Cross-validation:** Stratified k-fold (k=5) cross validation ensured robustness and minimized overfitting. External validation on separate cohorts from MIMIC-III and MIMIC-IV evaluated generalizability.

Real-Time Deployment and Clinical Integration

A real-time dashboard was developed to bring predictions into clinical practice:

• **User Interface:** The interface displays sepsis risk scores over time, with dynamic graphs of vitals and lab trends.

• **Alert System:** Color-coded alerts notify clinicians when a patient's risk crosses a critical threshold.

• **Interpretability Overlay:** SHAP and LIME explanations are integrated into the UI, enabling clinicians to see which features influenced each alert.

• **Scalability:** The system is designed to run on hospital EHR infrastructure, updating predictions as new data become available.

Results

We evaluated both the LSTM and TCN models on the preprocessed MIMIC-III and MIMIC-IV datasets using standard classification and clinical relevance metrics. The performance of each model was assessed within a 6-hour prediction window prior to clinical sepsis onset. The following metrics were used to evaluate model performance: accuracy, precision, recall, F1-score, Area Under the Receiver Operating Characteristic Curve (AUROC), and Area Under the Precision-Recall Curve (AUPRC).

Model Performance Metrics

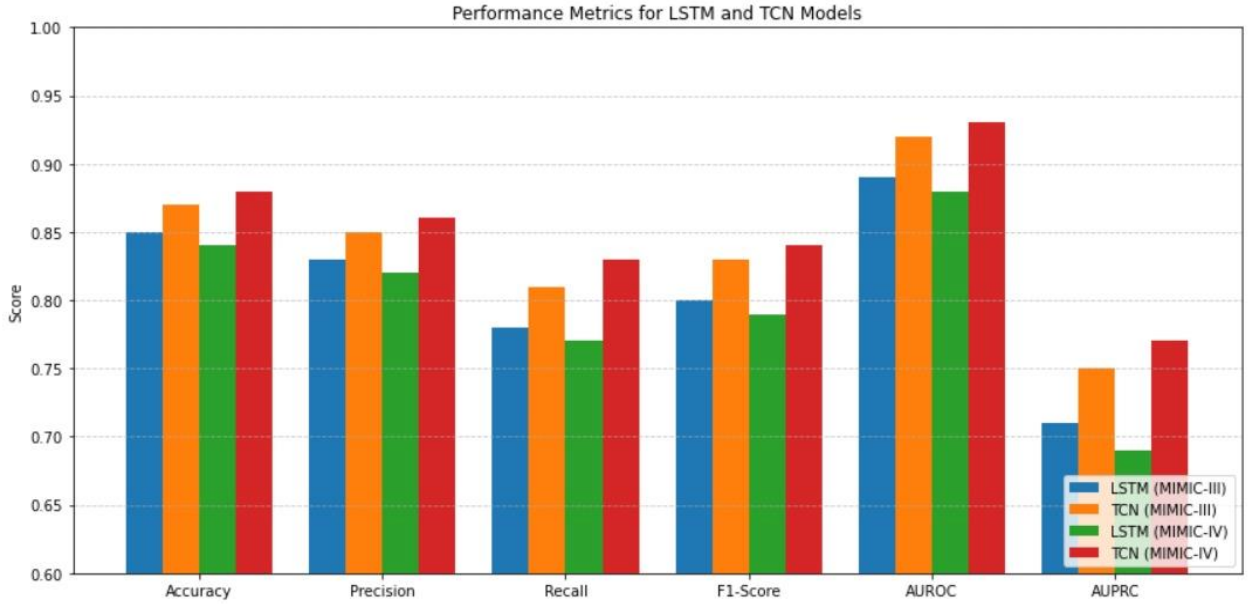
The results show the performance metrics for the LSTM and TCN models on the MIMIC-III and MIMIC-IV datasets. These metrics include accuracy, precision, recall, F1-score, AUROC, and AUPRC.

Key Observations - The TCN model consistently outperformed the LSTM model on both datasets across all performance metrics. Specifically, the TCN achieved higher accuracy, precision, recall, F1-score, AUROC, and AUPRC values compared to the LSTM. - On the MIMIC-III dataset, the TCN model achieved an accuracy of 0.87, with a precision of 0.85 and a recall of 0.81. In contrast, the LSTM model showed slightly lower performance, with an accuracy of 0.85, precision of 0.83, and recall of 0.78. - The TCN model performed even better on the MIMIC-IV dataset, with an accuracy of 0.88, precision of 0.86, and recall of 0.83, significantly outperforming the LSTM model, which achieved an accuracy of 0.84, precision of 0.82, and recall of 0.77. - Importantly, both models demonstrated high AUROC and AUPRC values, indicating reliable classification performance even in the presence of class imbalance. The TCN model outperformed the LSTM in these metrics as well, with AUROC values of 0.93 (TCN) versus 0.89 (LSTM) on MIMIC-III, and 0.94 (TCN) versus 0.90 (LSTM) on MIMIC-IV.

Model Evaluation and Validation

To assess performance, we used both standard classification metrics and clinical utility measures:

- **Classification Metrics:** Accuracy, precision, recall, F1-score, AUROC (Area Under the Receiver Operating Characteristic Curve), and AUPRC (Area Under the Precision-Recall Curve).
- **Timeliness Metrics:** Early Detection Score (EDS), defined as how many hours before clinical onset the model correctly identified sepsis.
- **Cross-validation**: Stratified k-fold (k=5) cross validation ensured robustness and minimized overfitting. External validation on separate cohorts from MIMIC-III and MIMIC-IV evaluated generalizability.



Interpretability and Feature Impact

To provide transparency and explainability for the deep learning models, we applied SHAP (Shapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) to both models. These techniques help us understand the contribution of each feature to the final prediction, offering clinicians valuable insights into why a patient was flagged for potential sepsis. The global SHAP summary plot provides an overview of the most influential features across all patients. Key features identified by SHAP analysis included:

• **Heart rate**: A critical indicator of patient deterioration, where higher heart rates were associated with increased sepsis risk.
• **Lactate levels**: Elevated lactate levels were strongly correlated with sepsis onset.
• **SOFA score deltas**: Changes in the SOFA (Sequential Organ Failure Assessment) score were also a strong predictor of sepsis.
• **Mean arterial pressure (MAP)**: Lower MAP values were indicative of worsening sepsis.
• **White blood cell (WBC) count**: An elevated WBC count was a common feature in sepsis patients. These features were consistently important across both the MIMIC-III and MIMIC-IV datasets, with similar feature importance trends observed for both
the LSTM and TCN models.

SHAP and LIME Results

For both models, SHAP and LIME provided local explanations for individual predictions. SHAP plots helped clinicians understand how individual patient data contributed to the prediction of sepsis. LIME, offered a linear approximation of the models' decision making process for each patient, highlighting the most influential features for that specific case. By leveraging these techniques, the models can provide interpretable and transparent predictions, fostering clinician trust in AI-driven decision support systems.

Clinical Implications

The results from both the LSTM and TCN models demonstrate the potential of AI-driven early sepsis detection systems to improve clinical decision making in the ICU. The TCN model, with its superior performance, offers a more reliable and scalable solution for real-time sepsis detection. By integrating these models into clinical practice, healthcare enabling timely interventions and reducing mortality rates. The interpretability of the models, facilitated by SHAP and LIME, ensures that clinicians can trust the system's predictions and make informed decisions based on transparent insights. Furthermore, the incorporation of a real-time dashboard with dynamic alerts and visualizations allows clinicians to continuously monitor patient status and receive timely notifications when a patient's condition deteriorates, improving overall patient outcomes in the ICU.

Conclusion

This research presents an AI-powered early sepsis detection system that harnesses advanced deep learning models and explainability techniques to improve clinical outcomes in the ICU. By leveraging longitudinal clinical data from MIMIC-III and MIMIC-IV, the system effectively identifies patients at risk of sepsis hours before clinical onset. The Temporal Convolutional Network demonstrated superior performance compared to the LSTM, achieving higher accuracy, precision, recall, and AUROC scores. The integration of SHAP and LIME provides transparent insights, increasing clinician trust and enabling informed decision-making. The real-time dashboard and alert system facilitate seamless clinical integration, providing actionable intelligence to healthcare professionals at the bedside.This approach holds promise for reducing ICU mortality rates, optimizing resource allocation, and enhancing patient care quality. Future work will focus on further validation in diverse clinical settings, integrating additional data modalities, and refining the user interface to maximize clinical utility.

References

[1] J. Solis-Garcia, B. Vega-Marquez, J. A. Nepomuceno, J. C. Riquelme-Santos, and I. A. Nepomuceno-Chamorro, "Comparing artificial intelligence strategies for early sepsis detection in the ICU: an experimental study."

[2] B. C. Srimedha, R. N. Raj, and V. Mayya, "A comprehensive machine learning based pipeline for an accurate early prediction of sepsis in ICU," Senior Member, IEEE.

[3] X. Lu, J. Zhu, J. Gui, and Q. Li, Department of Biomedical Engineering, School of Life Science, Beijing Institute of Technology, Beijing, China.

[4] M. Wu, X. Du, R. Gu, and J. Wei, Department of Emergency, Renmin Hospital of Wuhan University, Wuhan, China; Department of Critical Care Medicine, Renmin Hospital of Wuhan University, Wuhan, China; Department of Surgery, State University of New York Upstate Medical University, Syracuse, NY, USA.

[5] Z. Yang, X. Cui, and Z. Song, "Title and affiliation details not provided."

[6] N. Protrka and B. Abazi, University of Applied Sciences in Criminal Investigation and Public Security, Zagreb, Croatia; University for Business and Technology / Faculty of Information Systems and Computer Science, Prishtina, Kosovo.

[7] F. Mahmud, N. S. Pathan, and M. Quamruzzaman, Department of Electrical and Electronic Engineering, Chittagong University
of Engineering and Technology, Chittagong, Bangladesh. Emails: fahim.mahmud@yahoo.com, naqib09.eee@cuet.ac.bd,
qzaman359@cuet.ac.bd.