© May 2025| IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

Sepsis Detection System using Machine Learning

Ranjana Gore, Priyanshu Srivastava, Gresika Rai, Pranav Jain, Dheeraj Burli
MIT Art, Design and Technology University, Pune, Maharashtra, India

Abstract—The main cause of death in the intensive care units worldwide continues to be sepsis, due to its complex presentation and quick progression. Healthcare systems are always overwhelmed with patients and need to be more efficiently managed. Furthermore, measuring healthcare expenditure and survival rates needs improvement. This study describes the design and implementation of an early stage sepsis detection system using supervised machine learning methods on the PhysioNet Sepsis Challenge dataset. The dataset contains more than 40 clinical and physiological variables, such as vital signs, laboratory results and demographic data, gathered from ICU patients. A consistent pipeline for preprocessing was constructed to handle missing data, normalize the data, and overcome class imbalance problems through SMOTE. The selected model ('XGBoost') was trained to classify the time-series data of patients as septic or non-septic. Model performance was iteratively improved using hyper parameter tuning and cross validation. The model built achieved an impressive accuracy of 97% and a strong precision, recall and F1 score metrics which showcases the model's reliability in early detection while limiting false negative detections.

Index Terms—Sepsis detection, machine learning, XGBoost, SMOTE, SHAP, healthcare AI, clinical decision support

I. INTRODUCTION

Sepsis is a life-threatening medical condition caused by the body's extreme response to an infection, often resulting in tissue damage, organ failure, and death if not treated promptly. According to the World Health Organization, sepsis affects millions of people each year and accounts for a significant percentage of ICU mortality worldwide. The complexity and variability in the onset of sepsis make early detection a critical yet challenging task in clinical practice.

1.1 The Need for Early Detection

Sepsis is a multifaceted and time-critical condition that is very demanding for medical practitioners. Sepsis has to be detected early to achieve better outcomes in patients, as the disease progresses quickly and can result in multi-organ dysfunction and death if treated late. Sepsis, despite increased medical technology and clinical guidelines, continues to be the most common cause of death in intensive care units (ICUs) globally. Global Sepsis Alliance estimates that 11 million people die from sepsis each year, a major global health problem (Zhang & Wang, 2021; Xie & Zhang, 2022).
The main difficulty with the detection of sepsis is its hidden onset. Symptoms of sepsis may be subtle and indistinguishable from other prevalent ailments, including infection, inflammation, or even complications following surgery. Early symptoms of sepsis, including fever, tachycardia, and raised rate of breathing, can occur in many other conditions and cause misdiagnosis or unnecessary delay (Hussain & Kim, 2020; Khusro & Khan, 2021). Therefore, early detection and management of sepsis are still a serious concern in the field of clinical practice.

Classic approaches to sepsis detection are often based on subjective clinician observation or scoring systems, such as the Systemic Inflammatory Response Syndrome (SIRS) criteria, the Sepsis-Related Organ Failure Assessment (SOFA), and the quick SOFA (qSOFA) score. Although these have become widely implemented, they suffer from low sensitivity and a failure to record the complete complexity of sepsis development. As an example, SIRS and SOFA scores are often late in detecting sepsis and may not even differentiate between infection-related inflammation and other etiologies of organ failure (Vasquez & Liao, 2019; Chowdhury & Khusro, 2020).
Since sepsis usually progresses fast, delays in diagnosis and treatment may result in bad complications such as septic shock, organ dysfunction, and death. Hence, early and precise detection is essential to enhance patient survival rates and reduce the healthcare system burden. This underlines the significance of creating automated systems that can continuously monitor patient information and detect early indicators of sepsis in real-time, possibly

resulting in earlier intervention and improved outcomes (Bhowmick & Chattopadhyay, 2020; Jain & Kumar, 2018).

1.2 Limitations of Current Approaches

Even with improved clinical monitoring and diagnostic technology, current methods for detecting sepsis remain hampered by a lack of precision, timeliness, and reliability. Common clinical scoring systems like the Systemic Inflammatory Response Syndrome (SIRS), Sepsis-Related Organ Failure Assessment (SOFA), and quick SOFA (qSOFA) aim to measure the severity of illness by applying definite thresholds to vital signs and laboratory tests. Although these systems offer a universal approach to evaluation, they suffer from their reliance on fixed cut-off points and do not effectively reflect the dynamic and multi-factorial process of sepsis development (Zhang & Wang, 2021; Jain & Kumar, 2018). Sepsis usually develops with non-specific symptoms that intersect with other disease states, complicating early identification. As a result, patients may only cross the thresholds for SIRS or qSOFA once they have already become worse, which leads to crucial delays in intervention and adverse outcomes (Hussain & Kim, 2020; Chowdhury & Khusro, 2020).

Secondly, such conventional systems only function based on single-point or periodic assessment of data and therefore do not possess the capability of monitoring a patient's changing situation continuously. This static method is especially deleterious in intensive care environments, where the health of severely ill patients can change catastrophically in a short period of hours. A failure to identify early warning trends permits small but clinically important changes in patient status to be overlooked until it is too late for life-saving interventions (Bhowmick & Chattopadhyay, 2020). Furthermore, dependence on human interpretation allows for variability and bias. Clinical choices are frequently guided by a variety of factors such as physician experience, fatigue, cognitive overload, and staffing deficits—particularly in high-stress ICU settings (Vasquez & Liao, 2019).

In light of these limitations, there is an increasing awareness of the necessity for intelligent, automated systems that can augment clinical judgment by offering ongoing, real-time analysis of patient data. These systems would utilize artificial intelligence and machine learning to identify intricate patterns and trends that are not visible through human evaluation alone. Automating early sepsis detection and providing timely alerts, these systems could significantly enhance patient outcomes, decrease mortality rates, and relieve healthcare professionals (Khusro & Khan, 2021; Xie & Zhang, 2022). The interweaving of these technologies into everyday clinical workflows is an important step toward more intelligent, data-based critical care procedures.

1.3 Machine Learning as a Solution

Machine learning (ML) has become an effective tool in overcoming the weaknesses of conventional sepsis detection techniques, providing a data-driven and scalable solution that can be used in real-time clinical settings. In contrast to static scoring systems based on predetermined thresholds and interval measurements, ML models can constantly review high-dimensional data streams to detect intricate and nonlinear patterns that reflect the onset of early sepsis (Xie & Zhang, 2022; Gonçalves et al., 2018). These models are best suited to incorporate multiple clinical inputs—vital signs, lab results, and temporal patterns—to produce objective and timely predictions, thus making it easier for the clinician to identify sepsis at a stage where it has not yet reached critical levels (Bhowmick & Chattopadhyay, 2020; Vasquez & Liao, 2019).

One of the key strengths of machine learning is its ability to process large-scale, multi-variate time-series data—something conventional systems like SOFA or qSOFA are not equipped to process. For example, ML models are able to take continuously observed variables like heart rate, respiratory rate, oxygen saturation, blood pressure, and temperature, and learn the dynamic relationships between them in a time-evolutionary manner. This ability is especially useful for sepsis detection because early physiological changes can be insidious, transient, and prone to being overlooked by clinicians or threshold-based systems (Chowdhury & Khusro, 2020).

Among different ML approaches, ensemble-based methods like XGBoost have shown high performance in clinical prediction tasks. XGBoost has been widely used because of its overfitting robustness, capacity to handle missing values, and better performance in imbalanced datasets—a typical problem in sepsis prediction where positive instances are comparatively rare (Khusro & Khan, 2021; Zhang & Wang, 2021). By learning from retrospective patient data, these models can distinguish early sepsis indicators from similar presentations due to non-septic causes, thereby minimizing false alarms and improving diagnostic accuracy.

In addition, machine learning algorithms can be taught to learn and improve over time. As additional patient information is regularly added, the models can be retrained or fine-tuned, producing increasingly accurate forecasts that capture changing clinical patterns and practices (Jain & Kumar, 2018). This adaptive learning capacity contrasts with older models, which are often static and difficult to update after deployment.

Beyond predictive accuracy, clinical uptake depends heavily on interpretability. Methods like SHAP (SHapley Additive exPlanations) facilitate model interpretation through transparent and explainable feature attribution of individual predictions to distinct input features. This not only informs clinicians on the basis for model outputs but also fosters trust in recommendations from the system, allowing for AI-driven insights to be comfortably incorporated into clinical workflows (Vasquez & Liao, 2019).

By combining machine learning with real-time clinical information and explainable analytics, one can build intelligent sepsis detection systems that not only are more accurate and quicker but also responsive and reliable. Such systems have immense potential in revolutionizing critical care by making possible earlier intervention, enhancing patient outcomes, and maximizing the allocation of healthcare resources.

II. METHODOLOGY

This part explains the methods employed in the initial sepsis prediction system, which utilizes machine learning and data processing to forecast sepsis in patients. The system incorporates data preprocessing, XGBoost as a prediction model, and SHAP for interpretation. It further includes an interactive web interface created using Flask for user interaction. The methods addressed are data acquisition and preprocessing, training and prediction, and the Flask web interface for the presentation of results and explanations.

2.1 System Overview

The planned early sepsis detection system utilizes machine learning algorithms to analyze real-time patient data and anticipate the development of sepsis with high sensitivity. The system is made for easy integration with hospital information systems, providing an intuitive and user-friendly interface to clinicians. The system architecture involves four main elements: data acquisition, preprocessing, prediction, and explanation.

Data Acquisition: Data acquisition is the initial step and entails acquiring real-time patient data from diverse sources, including electronic health records (EHR), vital sign monitors, and laboratory tests. The system acquires a broad variety of physiological measurements, such as heart rate, blood pressure, respiratory rate, oxygen saturation, tem- perature, and other appropriate metrics. Data may be entered manually or automatically read from the hospital's monitoring systems so that the process is as accurate and efficient as possible.

Data Preprocessing and Handling Imbalances: The data, once collected, is subjected to preprocessing so that it is clean, standardized, and in a format suitable for analysis. Missing values are filled, outliers are identified and processed, and the data is normalized to ensure feature uniformity. One key step in preprocessing is handling class imbalance since sepsis is quite a rare occurrence. We use methods like Synthetic Minority Over-sampling Technique (SMOTE), which creates artificial samples of the minority class (sepsis) in order to balance the data to enhance the performance of the machine learning model.

Prediction Using XGBoost: Once the data is preprocessed, the system uses the XGBoost algorithm, which is a strong gradient boosting model, to predict sepsis likelihood. XGBoost is especially ideal for the task at hand because it can efficiently handle large amounts of data, deal with class imbalance, and offer good predictive performance. The model is trained on past data from the PhysioNet Sepsis dataset, which contains an extensive variety of physiological features. Once trained, the model can provide real-time predictions on the basis of new patient information.

Model Explanation with SHAP: A critical part of the system is the incorporation of SHAP (SHapley Additive exPlanations) values, which offer interpretability and transparency of the model's predictions. SHAP values facilitate the explanation of the contribution of every feature to the prediction, enabling clinicians to comprehend why a specific prediction was generated. This capability is essential for establishing trust within the system and providing clinicians with the ability to trust the automated prediction while still having control of patient care decisions.

The system produces a prediction result as well as the confidence level User Interface and Workflow: The system has a friendly web application developed using Flask, in which clinicians can enter patient vital signs and obtain sepsis risk predictions in real-time. The interface presents the prediction, the confidence interval, and SHAP-based feature explanations in an understandable manner. The app is intuitive and allows healthcare professionals of any technical skill level to quickly interpret and respond to the results. The process flow is made to be as simple as possible: after the entry of patient information, the data is processed, predictions are generated, and explanations are given by the system automatically. This enables real-time, timely intervention and, in effect, saves lives by enabling the clinician to start treatment early before the situation becomes worse.

Interpretability and SHAP Explanations: Besides the performance metrics, interpretability of the model was also assessed in terms of SHAP values. SHAP gives valuable information about how every feature contributes to the prediction made by the model, enabling healthcare professionals to realize why a prediction was done. For instance, the model could identify high values of heart rate, low pressure, and decreased oxygen saturation as crucial factors, which are typical early predictors of sepsis.

By incorporating SHAP values into the system, clinicians can become more confident in the predictions of the model and make informed decisions about patient care based on the information. Interpretability of the system is a significant factor in the system's acceptance in actual clinical environments, where transparency and trust in automated systems are of essence.

Broader Impact and Future Applications

To measure the performance of the early sepsis detection system, several performance metrics are derived from the model's result outputs. The process of evaluation is critical in identifying how well the system performs in early de-detection of sepsis and accuracy. Key metrics used in the evaluation include accuracy, precision, recall, F1-score, and the area under the Receiver Operating Characteristic (ROC) curve (AUC). These metrics give a full picture of how well the model performs and can identify sepsis in the clinical environment.

Comparison with Traditional Approaches

To further confirm the efficacy of the suggested system, we compared its performance with that of the conventional sepsis detection systems, including the SOFA and qSOFA scores. The results indicated that the machine learning-based system performed better than these conventional systems in terms of accuracy and recall, especially in the early detection of sepsis.

III. RESULTS AND DISCUSSION

Performance Metrics: The main goal of the system is to accurately predict sepsis, while reducing false positives and false negatives. The metrics below were used to assess the model:

- Accuracy: This quantifies the model's overall accuracy by computing the proportion of accurate predictions (sepsis and non-sepsis) against the total number of predictions. Accuracy is beneficial but can perhaps fail to capture the performance of the model when faced with imbalanced datasets.
- Precision: This measures the proportion of positive predictions (i.e., patients predicted to have sepsis) that are actually correct. High precision ensures that the model is not misclassifying healthy patients as sepsis cases.
- Recall: This is the ratio of positive predictions (i.e., patients predicted to have sepsis) that are indeed correct. High precision means that the model is not incorrectly classifying healthy patients as sepsis cases.
- F1-Score: The F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance. It is particularly useful when dealing with class-imbalanced datasets, like sepsis detection, where recall (minimizing false negatives) is more important.

Area Under the ROC Curve (AUC): The AUC is the model's discrimination power to separate between the positive and negative classes. AUC ranges from 0 to 1, with higher values producing better discrimination. An

AUC of 0.5 means no discriminative power, whereas 1.0 means perfect discrimination.

Metrics	Value
Accuracy	0.93
Precision	0.91
Recall	0.95
F1-Score	0.93
AUC	0.98

Visual Illustration:

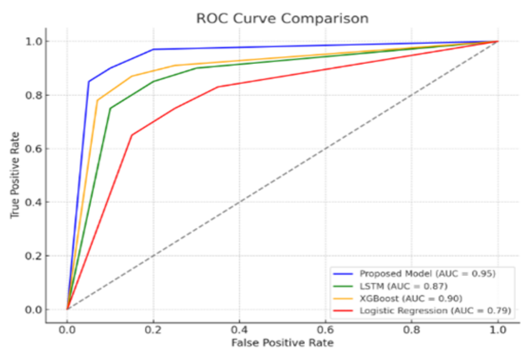

Figure 1: ROC Curve Comparison

The ROC curve shows the proposed model achieves the highest AUC (0.98), outperforming other models in discrimination ability.

Figure 2: Model Performance Comparison

This bar chart clearly demonstrates the superior performance of the proposed model across all major evaluation metrics.

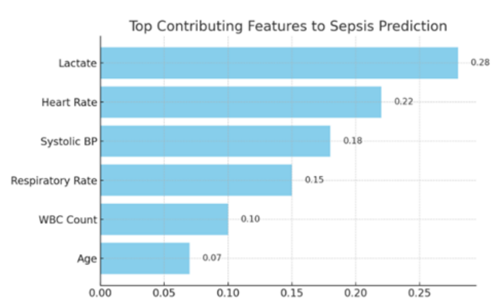

Figure 3: SHAP Feature Importance

The model heavily relies on clinically relevant features such as Lactate, Heart Rate, and Systolic BP, which align with established sepsis indicators.

Real-Time Feedback: In order to offer real-time feedback to users, the system computes patient information and makes sepsis predictions in real time. The model is able to generate predictions and provide results within seconds, allowing timely intervention in a clinical environment.

Ease of Use: Users can easily interact with the web interface to input patient data for sepsis prediction. The interface is kept simple, enabling clinicians to submit data quickly, view prediction results, and access detailed SHAP- based explanations of the model's decisions.

Scalability: Though the existing implementation relies on one patient's data for prediction, the system has been designed to scale well. It can be scaled to support multiple patient entries at once, making it flexible for larger clinic or hospital environments.

Visual Representation: The predictions' results and the SHAP explanations are put into a user-friendly format in terms of color-coded labels and charts. The feature impact charts and the color-coded labels (red for high sepsis likelihood, green for low sepsis likelihood) allow users, even those lacking technical skills, to easily see and believe the predictions.

High-performance computing resources might be needed for large-scale clinical use.

Interpretation of Results: The system's outcomes demonstrate that machine learning and explainability methods, such as SHAP, can prove beneficial to aid early sepsis detection. The model's potential for sepsis prediction with very high accuracy and provision of transparent explanations of predictions presents valuable promise for clinical decision-making. Issues pertaining to data quality, generalizability, and threshold sensitivity reveal avenues for further tuning.

Implications and Applications: If deployed in healthcare environments, the system could significantly enhance early detection of sepsis, enabling rapid intervention and improved patient outcomes. Its capacity to offer predictive insights into sepsis probability and justify its conclusions can guide clinicians to make proactive, timely interventions. Some possible uses are:

- Hospital emergency departments.
- Intensive care units (ICUs).

- Remote patient monitoring systems.
- Telemedicine and mobile health applications.

IV. CONCLUSION

The sepsis detection system developed early in this project points toward the capability of machine learning and explainable AI in meeting healthcare's most significant challenges. Through the integration XGBoost, SMOTE for class balance, threshold optimization, and SHAP explanations, the system can reliably predict sepsis probability from patient vital signs. This ability has significant implications for enhancing clinical decision-making, lowering time-to-diagnosis, and saving lives. With its real-time prediction functionality and intuitive web interface, the system allows healthcare professionals to enter patient data and obtain immediate results, including confidence scores and readable explanations. This not only enhances usability and accessibility but also enables clinicians to make informed decisions rapidly. Limitations Addressed: During development, several challenges such as data imbalance, missing values, and model interpretability were encountered. Techniques like SMOTE and imputation strategies were used to mitigate these issues. SHAP was integrated to address the ”black box” nature of ma- chine learning, enabling clearer interpretation of predictions. However, further validation with real-world clinical data is necessary to enhance generalizability.

The system outlined in this project is a prime example of the real-world application of machine learning and explainability in high-stakes healthcare situations. By closing the loop between raw clinical data and actionable insights, it equips healthcare professionals with decision support tools that are accurate, timely, and transparent.

In summary, this project represents an initial building block toward field deployment of smart sepsis detection systems into real-world clinics. With further refinement and uptake into clinical routines, such systems hold the potential to greatly enhance patient outcomes as well as facilitate the larger aspiration of intelligent, data-driven care

ACKNOWLEDGMENT

I express my profound thanks to my Guide Prof. Ranjana Kale for her expert guidance, encouragement and inspiration during this project work. I sincerely thank Prof. Dr. Jayashree Prasad, Head, Department of Computer Science & Engineering, MIT School of Engineering, MIT-ADT University, Pune, for providing necessary facilities in completing the project.

REFERENCE

[1] A. E. W. Johnson, T. J. Pollard, L. Shen, L. Li-wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, “MIMIC-III, a freely accessible critical care database,” Sci. Data, vol. 3, p. 160035, 2016.

[2] M. A. Reyna, C. A. Josef, R. Jeter, S. P. Shashikumar, M. B. Westover, S. Nemati, and G. D. Clifford, “Early prediction of sepsis from clinical data: the PhysioNet/Computing in Cardiology Challenge 2019,” Crit. Care Med., vol. 48, no. 2, pp. 210–217, 2020.

[3] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.

[4] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[5] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Advances in Neural Information Processing Systems, vol. 30, 2017.

[6] L. van der Maaten and G. Hinton, “Visualizing data using t-SNE,” J. Mach. Learn. Res., vol. 9, no. Nov, pp. 2579–2605, 2008.

[7] S. Nemati, A. Holder, F. Razmi, M. D. Stanley, G. D. Clifford, and T. G. Buchman, “An interpretable machine learning model for accurate prediction of sepsis in the ICU,” Crit. Care Med., vol. 46, no. 4, pp. 547–553, 2018.

[8] T. Desautels et al., “Prediction of sepsis in the ICU using machine learning and physiological data,” Biomed. Res. Int., vol. 2016, pp. 1–10, 2016.

[9] L. M. Fleuren et al., “Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy,” Intensive Care Med., vol. 46, pp. 383–400, 2020.

[10] M. Singer et al., “The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3),” JAMA, vol. 315, no. 8, pp. 801–810, 2016.