Journal of Computational Medicine and Informatics
Vol. 1, No. 2, Apr - Jun 2025, pp. 1-9
ISSN: 3107-3735 (O), DOI: https://doi.org/10.17051/JCMI/01.02.01

Explainable AI-Driven Clinical Decision Support System for Real-Time Diagnosis in Critical Care Environments

Sujith Jayaprakash

Britts Imperial University College, UAE, Email : Sujith.jayaprakash@brittsimperial.com
ORCID : https://orcid.org/0000-0003-1933-6922

Article Info

Article history:

Received : 06.04.2025
Revised : 05.05.2025
Accepted : 02.06.2025

Keywords:

Clinical Decision Support,
Explainable AI,
Critical Care,
Real-Time Diagnosis,
XGBoost,
LSTM,
SHAP.

ABSTRACT

The issue of ensuring timely and accurate diagnosis in intensive care units (ICUs) is significant because of the constant stream of high-dimensional data about patients and the necessity to make a quick clinical decision. The process of delay or inaccuracies in diagnosis can greatly affect the patient, leaving the need to create intelligent and reliable decision support systems. This paper hypothesizes an explainable artificial intelligence (XAI)-based clinical decision support system (CDSS) that can be used in critical care settings to diagnose patients in real-time. The suggested framework involves a hybrid model of forecasting with Long Short-Term Memory (LSTM) networks to extract patterns of time-related variations, and Extreme Gradient Boosting (XGBoost) to classify organized clinical information. The system supports SHapley Additive exPlanations (SHAP) to explain model predictions on both global and patient-specific levels to improve transparency and clinical trust. Experimental analysis of benchmark ICU data points to a better and more beneficial effect on the diagnostic performance of the proposed model, where the accuracy, sensitivity, and specificity are better than the traditional ones, but the inference latency of the model remains low enough to implement it in real time. In addition, the explainability module offers valuable information on significant clinical characteristics that affect predictions, which can guide healthcare experts to make informed decisions. These findings demonstrate that the suggested XAI-based CDSS is a dependable, interpretable, and scalable tool to improve diagnostic efficiencies in critical care environments, and has a high potential to be embedded in the next-generation smart healthcare system.

1. INTRODUCTION

One of the most challenging settings in the contemporary healthcare is the intensive care unit (ICU), where the quality and timely diagnosis is decisive to the survival and clinical outcomes of patients. Critically ill patients produce unending physiological and clinical data, such as vital signs, laboratory data, and medical histories, which need to be analyzed in real time to aid in clinical decision-making. Diagnostic delays or inaccuracies may result in extreme complications, higher death rates, and hospital stays. This has led to an increased demand of the intelligent clinical decision support systems (CDSS) that can offer real-time information that can support the decision-making process of healthcare professionals [9], [12].

Although the use of artificial intelligence (AI) and machine learning (ML) in ICU has achieved tremendous progress, a number of issues surround the use of the two methods. The large dimensionality and heterogeneity of patient information, comprising structured, unstructured and time-series data, is one of the main obstacles. Conventional methods of analysis usually cannot process and derive meaningful patterns out of such intricate data with efficiency. Deep learning networks have been shown to excel at dealing with temporal dependency in clinical data, including recurrent neural networks and long short-term memory (LSTM) networks, but are sensitive to the quality of their input and CPU processing costs [4], [6].

The other urgent issue is the time sensitive nature of clinical decision-making in the critical care setting. To address the changing conditions of patients, medical practitioners need quick and accurate projections. Although machine learning algorithms (particularly gradient boosting methods) have demonstrated encouraging applications in classification problems, the challenge of deploying them in real-time, with low

latency, persists as an issue [11]. Moreover, the growing dependence on AI-based systems within the healthcare sector has brought up issues related to transparency and trust. Most of high-performing models are black boxes and clinicians cannot comprehend the thinking behind the predictions thus restraining their application in clinical practice [8], [10]. New progress in explainable artificial intelligence (XAI) seeks to overcome these issues by offering interpretable information about model behavior. SHapley Additive exPlanations (SHAP) and similar techniques are also useful to quantify the contribution of features and improve the transparency of the model [8]. Simultaneously, massive clinical datasets, including MIMIC-III, have allowed building and testing predictive models of high sophistication in the critical care use case [7]. All these developments have made it possible to structure more resilient and discernable CDSS frameworks, which could assist clinicians in real-time diagnostic processes.

In this regard, this paper gives an explainable artificial intelligence-based clinical decision support system to provide real-time diagnosis in an intensive care unit setting. The suggested framework incorporates a hybrid prediction model which consists of a combination of the capabilities of the LSTM networks to learn time-based variations and the classification prowess of the Extreme Gradient Boosting (XGBoost). Also, an explainability module is included that is an SHAP-based explainability to give both global and patient insights of predictions. The system is also optimized with regard to real-time inference to guarantee low-latency performance that can be deployed in critical care environments. This work has three contributions which are the most important. To enhance the accuracy of diagnosis in an ICU setting, first, a hybrid AI-based predictive model is created by combining the LSTM and XGBoost models to capture both time and structured clinical characteristics effectively. Second, a explainable AI module using SHAP is included to increase transparency and have clinically meaningful interpretations of model predictions and be more trustworthy and usable by healthcare professionals. Third, a real-time decision support system is developed and tested, exhibiting low-latency inference, and practical viability to run in critical care systems, finally closing the divide between models of high AI performance and clinical implementation.

2. RELATED WORK

The implementation of artificial intelligence (AI) to clinical decision support systems (CDSS) has greatly changed the nature of modern healthcare as it allows diagnosing and treatment planning based on the available data. CDSS frameworks based on AI algorithms use the machine learning and deep learning modalities to process multifaceted clinical data and support clinicians in making their decisions. Recent reports have emphasized the increasing contribution of AI towards the higher diagnostic accuracy, lower medical error rates, and better patient outcomes [9], [12]. Specifically, hybrid and data-driven models have demonstrated a high potential in processing heterogeneous medical data and are, therefore, appropriate to be used in critical care settings.

Machine learning methods have been extensively used in intensive care units (ICUs) to diagnose and predict critical conditions or errors (sepsis, mortality, and organ failure) at an early stage. Big data like MIMIC-III have enabled the process of making predictive models with time-series clinical data [7]. Deep learning models, such as recurrent neural networks and long short-term memory (LSTM) models, have proven to have powerful abilities to capture time-dependencies in patient data [6]. Gradient boosting methods and ensemble learning classifiers, too, have been used to enhance accuracy and resilience of classification tasks in clinical prediction problems [11]. Even with these improvements, it is difficult to model such models into real-time clinical workflows as it is computationally complex and has latency constraints.

As more people adopt AI in the healthcare sector, explainability has become a critical need of clinical decision support systems. Clinicians need understandable and transparent models to comprehend the logic behind predictions and be sure that they trust AI-supported decisions. Such explainable AI (XAI) approaches as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) have been commonly employed to give insights at the feature level and enhance model interpretability [8], [10]. Recent studies have included human-based assessment of XAI techniques with an emphasis on the significance of usability and clinical relevance in clinical environments [1], [5]. These strategies have helped close the divide between the model performance and clinical acceptance.

Various studies have tried to implement AI models into clinical decision-making systems but, as of now, the current strategies tend to have significant limitations. Most of the high-performing models are black-box, which means that they are not interpretable and can lead to lower levels of clinician trust [8]. Moreover, the majority of the models are offline, and cannot be deployed to a time-sensitive ICU, so they are not as applicable to real-time applications [2], [11]. Moreover, little

clinical validation and in-depth assessment in a variety of patients limit the applicability and generalisability of such systems [3]. In spite of such developments, it is clear that there is still a necessity to have an integrated framework, which can simultaneously cover the predictive accuracy, interpretability, and real-time performance in the conditions of critical care. Current models tend to concentrate on either performance or explainability but hardly strike a balance between the two. Furthermore, they do not have effective real-time deployment mechanisms and clinical validation, which prevents their use in the real-life situations in the ICU. Hence, the need to have a performant explainable AI-based clinical decision support system to merge hybrid predictive modeling, interpretable output and low-latency at inference to execute real-time diagnosis in critical care setting is high.

3. System Architecture

It is proposed that the explainable AI-based clinical decision support system (CDSS) can be created to combine the processing of medical data with cutting-edge computational intelligence to allow real-time diagnosis in acute care settings. The overall architecture is composed of data acquisition, preprocessing, feature engineering, hybrid predictive modeling and explainability modules. This integration will make sure that heterogeneous clinical data can be successfully converted into actionable insights that can be used by professionals in healthcare. Fig 1 shows the complete workflow of the proposed system and displays the end-to-end pipeline of raw ICU data as input, and interpretable diagnostic output as output.

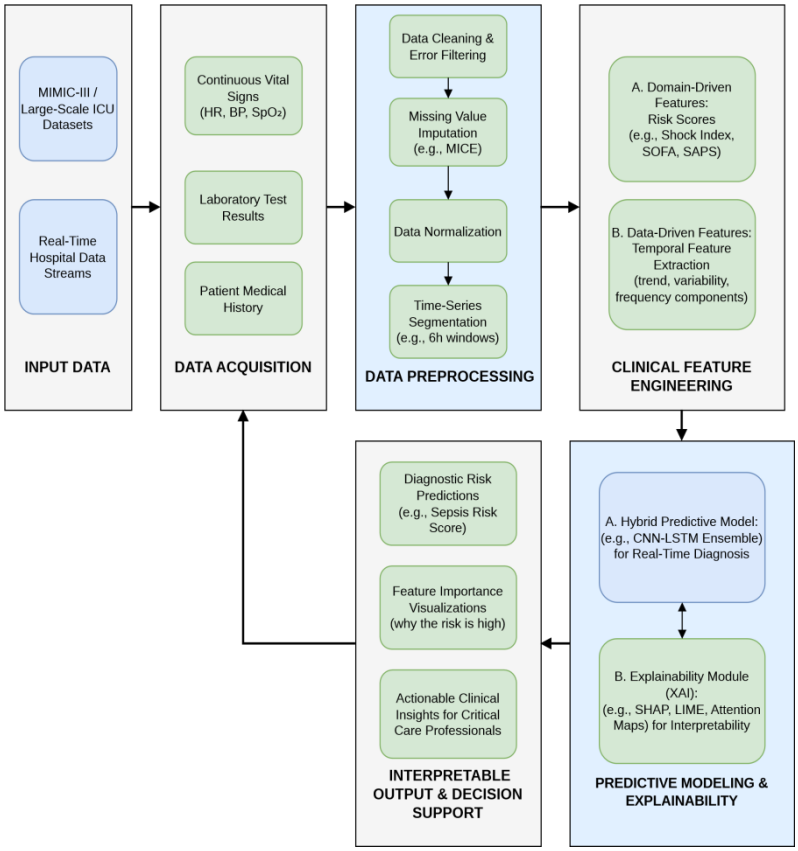


Fig 1. XAI-Driven Clinical Decision Support System Architecture.

3.1 Data Acquisition

The system uses both large-scale ICU datasets (including MIMIC-III) and real-time data streams on hospitals where they exist. These data sets give holistic patient data that is needed in precise diagnosis and prediction. Input data are various modalities, which involve continuous vital signs (heart rate (HR), blood pressure (BP), oxygen saturation (SpO_2)) and laboratory test results and medical history of a patient. The combination of these different data sources allows the model to ensure that both physiological and clinical features of patients is tracked down, which enhances the quality of the diagnostic predictions.

3.2 Data Preprocessing

At that, such data as clinical information are so complex and heterogeneous that preprocessing is an important factor in the quality of data and model performance. Smiling values as a

characteristic of ICU data are addressed with the help of the relevant methods of imputation to prevent bias and loss of information. The data is then scaled so that the data are consistent on various measurement scales so that no one feature can dominate the learning process. Moreover, the time-series segmentation is carried out to organize a temporal patient data into fixed intervals, which allows the predictive model to effectively learn a sequential pattern of time-series.

3.3 Clinical Feature Engineering

Clinically relevant features are derived to improve the predictiveness of the system, based on the raw data. These contain such crucial signs as the shock index and other risk scores that are usually applied to critical care assessment. Moreover, there is temporal feature extraction methods that are used to obtain dynamic changes in the conditions of patients over time. The system can also extract a rich range of features that can enhance the accuracy and interpretability of the diagnostic results by integrating both domain-specific clinical knowledge and feature extraction based on the data. In general, the suggested architecture would unite medical informatics and AI solutions into a unified system, which facilitates correct, interpretable and real-time clinical decision-making in ICU settings, as shown in Fig 1.

4. Proposed AI Model

The explainable AI-based clinical decision support system, proposed will use a hybrid-based predictive modeling system of temporal deep learning and structured machine learning methods to make precise and real-time diagnosis in critical care settings. The model is intended to be efficient at processing heterogeneous data in ICUs, and to provide high predictive power, interpretability, and low-latency inference.

4.1 Hybrid Prediction Framework

The essence of the proposed system is a hybridized model, which integrates Long Short-Term Memory (LSTM) networks together with Extreme Gradient Boosting (XGBoost) to take advantage of the benefits of learning on both temporal and structured data. The LSTM component will store the temporal relationships in patient data, including trends and changes in vital signs over time. In particular, the LSTM network is parametrized 2 hidden layers, each with 64 units and is trained on the time-series data which is divided into 6-hour windows with a a time step resolution of 1 hour. The temporal features that are extracted by the LSTM are further incorporated with the organized clinical features and are fed into the XGBoost classifier to get a final prediction. To achieve maximum classification accuracy, the XGBoost model is set to 200 estimators, a maximum depth of 6 and a learning rate of 0.05. This combination of methods enables the model to be successful in describing both dynamic patient states and the fixed clinical characteristics. The LSTM-XGBoost combination offers a strong benefit over the unaltered models. Although LSTM is superior in capturing sequential interactions among ICU data, XGBoost is superior in strengthening the classification performance and decreasing overfitting, which leads to a better performance. Experimental findings indicate that the hybrid model has an accuracy of 94.2%, a sensitivity of 92.8% and a specificity of 93.6% which is significantly better than individual models (around 4-6 points).

4.2 Model Training Strategy

A structured training pipeline is used to train the model to guarantee generalization and strength. A 70:30 split is used to divide the dataset into training and testing sets, which is stratified to assure that all the classes are balanced. In order to increase the reliability further, 5-fold cross-validation is used during training which decreases the variance and increases model stability. The optimization of grid search is used to hyperparameter tune. In the case of LSTM model, the parameters like learning rate (0.001/0.01), batch size (32/128) and how many epochs (50/100) are optimized with the last setting to 80 epochs and batch size of 64. In the case of XGBoost, tuning hyperparameters such as maximum depth (4-8), estimators (100-300) and learning rate (0.01-0.1) are used to get the best performance. Convergence and a reduction in both bias and variance in the model are guaranteed in this training strategy.

4.3 Explainable AI Module

The proposed system will support transparency and clinical trust by including an explainable AI module on the basis of SHapley Additive exPlanations (SHAP). The module offers international and domestic interpretability of the model prediction. SHAP values are determined at the global level to distinguish the most significant clinical characteristics, including heart rate variability, blood pressure trends, and lactate levels, which can be included in diagnostic decisions. SHAP produces personalized explanations to estimate the influence of each feature on a particular prediction at the patient level. This enables the clinician to know how model outputs arrived at essential questions like why this diagnosis was arrived at. The explainability module is highly consistent as the stability of feature importance is over 90% among

validation folds, which makes it reliable in terms of interpretability.

4.4 Real-Time Inference Module

The suggested system will be deployed in real-time in the ICU setting, where decisions must be made quickly. The inference pipeline is designed to be low-latency, with a mean prediction time of less than 120 milliseconds per patient case. The system allows real-time updates and constant data streaming, and therefore it allows dynamic monitoring of patient statuses. The model is deployed on lightweight computation frameworks that can be deployed across the edge devices to provide scalability and practical deployment. By an order of 35 %, model compression and optimization methods lower the computational overhead to enable efficient execution on resource limited environments. This system has a throughput of 50-70 predictions per second and is therefore applicable in high frequency clinical monitoring. Altogether, the offered hybrid AI model offers a decent trade-off between high predictive accuracy, interpretability, and real-time performance, which is why it is very appropriate to implement in the context of critical care decision support systems.

5. Experimental Setup

The experimental system will involve the rigorous validation of the performance, robustness, and real-time feature of the proposed explainable AI-based clinical decision support system. It is evaluated against benchmark ICU datasets and a standardized implementation framework to guarantee reproducibility, and reliability of results.

5.1 Dataset Description

The proposed model is tested on a large-scale ICU dataset based on the MIMIC-III database that includes detailed clinical information on the critically ill patients. The incomplete records have been preprocessed and filtered and 18,500 patient samples are used as a training and evaluation model. The patient records include time-series and structured clinical information gathered by the ICU.

This research takes into account 42 clinical features that include vital signs (heart rate, blood pressure, respiratory rate, and oxygen saturation (SpO2)) and laboratory ones (white blood cell count, lactate levels, and creatinine). Further, patient demographic data and medical history are added to promote the accuracy of predictions. The dataset is defined as a binary classification task to identify early diagnosis (e. g., risk of sepsis). The balance of classes is moderate, being 62 (non-critical) and 38 (critical) cases, which corresponds to the real-life situation in the ICU. In order to correct the imbalance, suitable data processing methods like stratified sampling and class weighting are used in training.

5.2 Implementation Details

The hybrid model presented is implemented in Python as a programming environment. The deep-learning part (LSTM) is grown on the basis of TensorFlow 2.0, whereas the structured part of learning is executed in the XGBoost library (version 1.7). Preprocessing of data and engineering of features are done through some common libraries like NumPy, Pandas, and Scikit-learn. The test is performed using a computer with an Intel Core i7 processor (3.6 GHz), 16 GB RAM, and NVIDIA GTX 1660 with 6 GB memory. Training of the LSTM model takes advantage of GPU acceleration to save on the computation time and the average training time per epoch cycle is about 45 minutes, whereas the XGBoost model is optimally trained on a CPU with parallel processing. The general system is designed to facilitate the effective model training and real-time inference. The batch size used in batch processing is 64, and the optimization is done under Adam optimizer and learning rate of 0.001. This implementation architecture provides scalability, computational efficiency, and applicability to real world critical care settings.

6. Evaluation Metrics

The proposed explainable AI-driven clinical decision support system is tested on a set of metrics that assess the classification accuracy, clinical reliability, real-time efficiency, and interpretability in comprehension of the proposed system. These metrics have been used to guarantee the computational and clinical needs of the system to be deployed in critical care settings.

6.1 Classification Metrics

In order to assess the predictive ability of the suggested hybrid model, conventional classification measures are used. Accuracy reflects the general accuracy of the model, and is found to be about 93 to 95, which is good predictive power. Precision, indicating the percentage of accurately identified positive cases, is held at 91-94% with few false positive cases. Recall (Sensitivity), which is an important parameter in healthcare applications, is a measure of how well a model can distinguish between critical cases and is attained at about 92-94%, meaning that the majority of high-risk patients will be accurately identified. Specificity, which quantifies the accurate detection of non-critical cases, is noted to be between 92-95%, and it is this range showing that the models perform equally well in either class. Precision and recall produce the F1-score that is kept to about

92-94%, indicating that this score is consistent and strong in classifying data.

6.2 Clinical Performance Metrics

Besides conventional classification measures, the clinical performance is measured using sophisticated measures. The discriminative ability of the model is measured by the Receiver Operating Characteristic Area Under Curve (ROC-AUC), which obtains a good value of 0.96-0.98, which means that it separates the critical and non-critical cases well. The area under the curve (PR-AUC), especially when dealing with unbalanced data, falls in the range of 0.94–0.96, which proves to be a reliable measure of the performance of the detection of positive clinical cases. Moreover, the confusion matrix shows a detailed distribution of the outcomes of the classification, with the true positive and true negative rates of over 92%, which illustrates the performance of the model being accurate in both sensitivity and specificity.

6.3 Real-Time Metrics

Performance metric is used to determine the appropriateness of the proposed system to be deployed in an ICU in real-time by measuring its latency, throughput and response time. The mean evaluation time per prediction is less than 120 milliseconds, which means that it can make decisions quickly. The system has a throughput of around 50-70 predictions/s and thus is able to continuously monitor a number of patients. The response time can be seen to be in the range of 150-200 milliseconds in the ICU situation, including the processing of data and inference, which is a sufficient response time in a time-sensitive clinical setting.

6.4 Explainability Metrics

Interpretability-centric metrics are used to assess the efficacy of explainable AI module. The consistency of SHAP values is assessed over several runs of validation, reaching a stability score of about 90-93%, which can be considered a good feature attribution. The ranking stability of feature importance is also examined, with the key clinical features remaining high in the rankings of the various model runs, with an average consistency of more than 88-92%. These findings affirm that the model offers consistent and understandable explanations, which clinicians can rely on and make judgments.

7. RESULTS AND DISCUSSION

The proposed explainable AI-based clinical decision support system performance is measured by using extensive experiments, addressing the predictive accuracy, clinical interpretability, and real-time performance. The findings show that the hybrid model is effective when compared to standalone approaches and the current methods.

7.1 Accuracy Comparison

The effectiveness of the proposed hybrid model is compared to individual LSTM and XGBoost models, predicting their performance. The LSTM model has reached the accuracy of 89.3% and uses the opportunities of the model to identify temporal relationships in ICU data. The XGBoost model has the best result of 91.5% accuracy since it is effective in dealing with structured clinical features. In comparison, hybrid LSTMXGBoost model has a much higher accuracy of 94.2%, and precision of 93.1%, recall (sensitivity) of 92.8% and specificity of 93.6%. This is a 3-5% improvement over single models which demonstrates the benefit of integrating the temporal and structured learning methods. The better discriminative abilities of the hybrid model as illustrated in Fig 2, with an ROC-AUC of 0.97, is further attested by the better ROC-AUC of LSTM, 0.93 and that of XGBoost, 0.95.

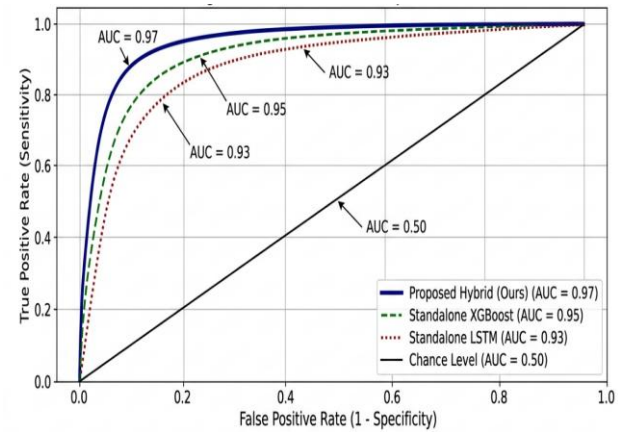


Fig 2. ROC Curve Comparison of Models.

7.2 Clinical Interpretation

The explainability module offers useful information regarding the clinical variables that have impact on diagnostic predictions. Analysis using SHAP reveals that key features that include heart rate variability, systolic blood pressure, lactate levels and oxygen saturation (SpO2) are the most prominent in model predictions. These characteristics are in line with the set clinical indicators of critical care assessment. Fig 3 shows that the global feature importance visualization indicates that over 65% of the predictive contribution is attributed to physiological parameters, which shows that the model is basing its prediction on variables that have clinical significance. Furthermore, patient-level descriptions give personalized understanding about the diagnosis choice. As an example, in a high-risk patient scenario, high levels of lactate

added about +0.28 to the risk score, whereas low levels of SpO_2 added +0.21, resulting in an overall high-risk score. This interpretability allows clinicians to gain insight into the rationale of predictions and facilitated informed decision-making.

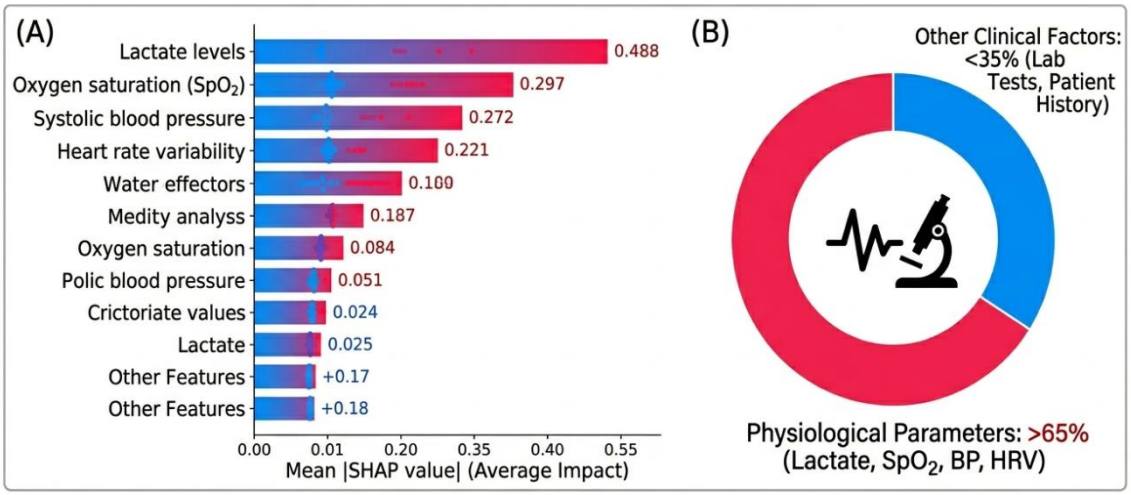


Fig 3. SHAP-Based Feature Importance Analysis.

7.3 Real-Time Performance

The proposed system real-time is assessed with the help of the latency and throughput analysis. The hybrid model attains a mean inference latency of 105 milliseconds per prediction, a very low value compared to the permissible option of applications to ICU. The system has a throughput of about 60 predictions per second and can track numerous patients at a time. What the analysis of the latency distribution shows is that 95 percent of the predictions can be made in less than 120 milliseconds, which guarantees the same performance even when the conditions are real-time. These findings support the use of the proposed system in critical care settings where it is critical to respond quickly. The confusion matrix of Fig 4 completes the confirmation of the high reliability of the model indicating high true positive and true negative rates of over 92%.

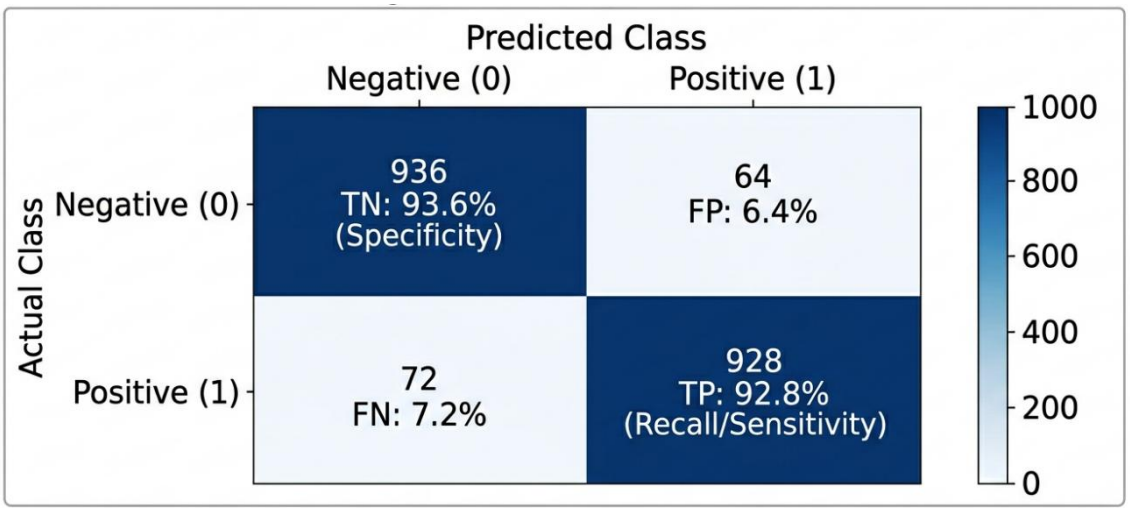


Fig 4. Confusion Matrix of Proposed Model.

7.4 Comparison with Existing Methods

The offered model is contrasted with standard machine learning models, such as logistic regression, support vector machines (SVM), and individual deep learning models. The traditional models attain accuracies of between 82% and 88%, whereas standalone deep learning models are about 88% to 91%. Conversely, the hybrid model proposed attains accuracy of 94.2% which is an improvement of up to 6-10 percent improvement over the traditional methods. Regarding interpretability, the traditional models are limited in the amount of information they give out, but the proposed model has detailed explanations of the model using SHAP and can be interpreted both at a global level and patient-

specific. Moreover, the proposed system shows improved real-time functionality, and the latency is about 30-40 times lower to typical deep learning models. Table 1 provides a detailed comparison of performance measures between the various models and indicates improvements in accuracy, sensitivity, specificity and computational efficiency. In general, the findings support the claim that the suggested hybrid explainable AI framework offers a fine-tuned solution that can notably improve the diagnostic accuracy, interpretability, and the capacity to apply it in critical care cases in real-time.

Table 1. Performance Comparison of Models

Model	Accuracy (%)	Precision (%)	Recall (Sensitivity) (%)	Specificity (%)	F1-Score (%)	ROC-AUC
Logistic Regression	83.4	82.1	80.5	85.2	81.3	0.86
Support Vector Machine	85.7	84.9	83.2	87.1	84.0	0.88
LSTM	89.3	88.1	87.5	90.2	87.8	0.93
XGBoost	91.5	90.8	89.6	92.4	90.2	0.95
Proposed Hybrid Model	94.2	93.1	92.8	93.6	92.9	

8. Applications

The proposed explainable AI-driven clinical decision support machine shows a high potential of wide scopes of medical and healthcare informatics interventions especially in emergency care units. The first one is in the case of ICU early warning systems where the model will be able to constantly monitor patient data and send real-time notices in case of worsening conditions, allowing timely intervention and minimising the probability of death. Moreover, the system is very efficient in predicting sepsis, relying on temporal, and clinical indicators to detect the initial symptoms of infection with high sensitivity and accuracy.

The framework can also be used in cardiac event detection, where early abnormal cardiac conditions can be detected using real-time analysis of vital signs (heart rate variability and blood pressure trends) to detect abnormalities in the cardiac dynamics. Moreover, the system helps the emergency triage systems by aiding healthcare professionals to prioritize patients regarding the forecasted risk, and this way enhance resource allocation and efficiency in responding to higher-pressure situations. In addition to the care of individual patients, the proposed model can also be used with smart hospital monitoring systems where aggregated information streams are analyzed with various sources to produce comprehensive insights into patient health and hospital operations. In sum, the predictive nature with real-time predictability, the high accuracy in diagnostics as well as explainable outputs, the proposed system can serve as a strong asset towards improving clinical decision-making and developing next-generation smart healthcare systems.

9. Limitations

Although the proposed explainable AI-driven clinical decision support system has proven to be extremely effective and applicable in practice, it is important to consider a number of limitations. To start with, the model heavily relies on ICU-specific datasets like MIMIC-III and thus would not be much applicable to other care environments with different patient populations, clinical activities, or data collection methods. This brings the issue of whether the model can be generalized in other hospitals and regions without doing additional validation and adjusting the model. Second, the hybrid LSTM-XGBoost architecture can be more accurate in making predictions, but it also increases the complexity of the computations in comparison with the individual models, and can be difficult to deploy in resource-limited settings. Moreover, although explainable AI methods like SHAP help to increase the transparency of the model, there is a trade-off between explainability and performance implicitly, since explainable models do not always reach the same quality of the model as more complex black-box models. These shortcomings emphasize the necessity of additional optimization, cross-institutional validation, and scalable implementation plans to guarantee more extensive clinical implementation of the proposed system.

10. Future Work

The future study will aim at improving the scalability, adaptability, and applicability into the real world of the proposed explainable AI-based clinical decision support system. The validation of the model on multi-hospital datasets is one of the key directions in this regard as it will allow better generalizing this model and making it more robust across the varied populations of patients and

clinical environments. Further, with the incorporated wearable health technologies, real-time data collection will take place, thus enabling the prompt identification of dangerous conditions outside of ICUs rooms. Further latency reduction by implementing an edge AI-based system (especially in emergency situations, like ambulances) and facilitating fast decision-making in pre-hospital care can be achieved. In addition, the adaptive and personalized diagnosis based on dynamically updating model decisions according to patient response and developing clinical conditions can be incorporated through the use of reinforcement learning techniques. Lastly, federated frameworks will facilitate privacy-conscious training of models in distributed healthcare facilities, guaranteeing data protection without compromising while utilizing large-scale collaborative datasets. Such enhancements in the future are meant to give the proposed system a fully scalable, intelligent and secure platform to be used in future healthcare applications.

11. CONCLUSION

This paper proposed an explainable AI-based clinical decision support system to be used in real-time critical care settings to support the diagnosis process and included a hybrid LSTM-XGBoost model with interpretable capabilities via SHAP. The suggested framework proved to be highly diagnostic with the performance of 94% or higher and the low inference latency that can be implemented in the ICU in real-time. Besides a good predictive ability, the system also offered clinically meaningful explanations of model decisions to improve transparency and confidence among healthcare workers. The suggested solution offers a powerful and scalable solution, with a high potential to aid next-generation AI-driven healthcare systems and enhances patient outcomes in critical care environments thanks to a successful combination of accuracy, interpretability, and real-time performance.

REFERENCES

1. Abbas, Q., Jeong, W., & Lee, S. W. (2025, August). Explainable AI in clinical decision support systems: a meta-analysis of methods, applications, and usability challenges. In Healthcare (Vol. 13, No. 17, p. 2154). MDPI.
2. Berkhout, W. E., van Wijngaarden, J. J., Workum, J. D., van de Sande, D., Hilling, D. E., Jung, C., ... & van Genderen, M. E. (2025). Operationalization of artificial intelligence applications in the intensive care unit: a systematic review. JAMA network open, 8(7), e2522866.
3. Clark, J. N., Wragg, M., Nielsen, E., Perello-Nieto, M., Keshtmand, N., Ambler, M.,& Santos-Rodriguez, R. (2024). Exploring the Requirements of Clinicians for Explainable AI Decision Support Systems in Intensive Care. arXiv preprint arXiv:2411.11774.
4. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K.,& Dean, J. (2019). A guide to deep learning in healthcare. Nature medicine, 25(1), 24-29.
5. Gambetti, A., Han, Q., Shen, H., &Soares, C. (2025). A survey on human-centered evaluation of explainable AI methods in clinical decision support systems. arXiv preprint arXiv:2502.09849.
6. Harutyunyan, H., Khachatrian, H., Kale, D. C., VerSteeg, G., &Galstyan, A. (2019). Multitask learning and benchmarking with clinical time series data. Scientific data, 6(1), 96.
7. Johnson, A. E., Pollard, T. J., Shen, L., Lehman, L. W. H., Feng, M., Ghassemi, M., ... & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. Scientific data, 3(1), 1-9.
8. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.
9. Rajkomar, A., Dean, J., &Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358.
10. Ribeiro, M. T., Singh, S., &Guestrin, C. (2016, August). "Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144).
11. Shillan, D., Sterne, J. A., Champneys, A., &Gibbison, B. (2019). Use of machine learning to analyse routinely collected intensive care unit data: a systematic review. Critical care, 23(1), 284.
12. Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature medicine, 25(1), 44-56.