International Journal of Artificial Intelligence Trends (IJAIT)
Vol. 5, Issue VI; No. 67, June, 2026, pp. 677-686

DEVELOPMENT OF A MACHINE LEARNING-BASED MODEL FOR SEPSIS PREDICTION IN INTENSIVE CARE UNITS

Kekong P. E[1]*,Mavollo Christpher Mayat[2],Jethro MaturJack[3], Adeyi Thomas Edeh[4]
[1*,2,4]Department of Mathematics and Computer Science, Federal University of Health Sciences, Otukpo Benue State, Nigeria
[3]Department of Physics, Karl Kumm University, Vom, Plateau State, Jos, Nigeria
Author Email: [1*]piuskekong2019@gmail.com, [2]krizmayat@gmail.com, [3]jackjethromatur@gmail.com, [4]adeyiedeh@gmail.com

[1*]https://orcid.org/0009-0006-7019-0198, [2]https://orcid.org/0009-0002-9764-4779, [3]https://orcid.org/0000-0002-5036-6030, [4]https://orcid.org/0009-0005-87485921

Article Info
Received: 18/04/ 2026
Revised: 12/5/2026
Accepted 3/6/2026
Corresponding Authors [1*]Email: piuskekong2019@gmail.com
Corresponding Author's Tel: [1*]+2348085211771

ABSTRACT

Sepsis is a life-threatening disorder that may easily escalate to the deterioration of organs and death unless early detected and addressed. Early detection of sepsis among patients in Intensive Care Unit (ICU) is thus, important in enhancing outcomes. It is a paper that introduces a machine learning model based on Long Short-Term Memory (LSTM) networks to predict the onset of sepsis based on temporal trends on patient vital signs, laboratory tests, and demographic information. The eICU Collaborative Research Database data were pre-processed, resampled, and missing value imputed, normalised, and sequence generated (as an LSTM input). The LSTM model was trained in 15 epochs and tested on another test set. The highest scores were an AUROC of 0.91, a precision of 0.78, a recall of 0.82, and F1-score of 0.80, which predicts sepsis 3.5 hours on average before clinical diagnosis. The confusion matrix showed that the model was quite useful in recognising non-sepsis patients and achieving high accuracy in most cases of sepsis but there were also false negativity. These findings also indicate that LSTM-based networks are able to model temporal relationships in ICU patient records in order to predict early signs of sepsis, which has the potential of providing an opportunity to intervene in the process of clinical intervention. This paper demonstrates the opportunity to unify predictive models based on machine learning with critical care units, as it is a solution that can help clinicians make decisions grounded in data and lead to better patient outcomes. Future researchers are recommended to improve the sensitivity and add some clinical characteristics and to test the model on various ICU populations to guarantee its wide applicability.

Keywords: Sepsis Prediction; Intensive Care Unit (ICU); Machine Learning; Long Short-Memory (LSTM); Time-Series Data.

1. INTRODUCTION

Sepsis is a critical life-threatening condition that develops due to dysfunctional response of the body to an infection, which may cause dysfunction of body organs and death unless it is detected and handled in time. It is one of the primary causes of morbidity and mortality in intensive care units (ICUs) all over the globe that causes the death of millions of individuals per year (La Via et al., 2024; Rudd et al., 2020; Lancet Global Health, 2025). Although significant gains have been made in the critical care field, timely diagnosis of sepsis remains a major clinical issue because of its complicated pathophysiology and the difference in its manifestation (Springer, 2024). Early identification and treatment are essential because the lateness in the diagnosis is directly linked to the worse outcomes and higher health costs (Schertz et al., 2023).

Conventional methods of sepsis identification, such as clinical scoring scales, such as the Sequential Organ Failure Assessment (SOFA) and the quick SOFA (qSOFA) scale and the Systemic Inflammatory Response Syndrome (SIRS) definition, depend on periodic monitoring and manual evaluation. These tools are rather guide-like, however, they are often insensitive and lack specificity and the possibility to forecast the occurrence of sepsis before it goes out of control

(Schertz et al., 2023; eClinicalMedicine, 2021). As a result, patients are likely to wait longer to obtain the required interventions, which emphasises the importance of more robust and real-time prediction procedures that will be able to utilise the abundance of data that is produced in ICUs (Masino et al., 2019).

Machine learning (ML) provides a potential remedy, as it allows searching through extensive amounts of heterogeneous data on patients to identify complicated trends leading to the development of sepsis. ML models can be used to issue early alerts on sepsis, usually hours earlier than conventional methods can detect it by combining physiological measures, lab data, demographics, and other clinical variables (Moor et al., 2021; Alanazi et al., 2023). Recent reports have showed that algorithms, including Random Forests, Gradient Boosting, and Long Short-Term Memory (LSTM), can be used to reach significant predictive accuracy (AbuHaweeleh et al., 2025; Yilmaz Başer et al., 2023). With the help of these methods, it is possible to provide a personalised risk assessment, and it is essential in the dynamic ICU setting, in which the conditions of patients can change very quickly (Islam et al., 2023; Mishra et al., 2025).

The proposed paper is intended to create a machine learning-powered system that provides early sepsis detection on ICU patients based on real-time Electronic Health Records (EHRs) and physiological data. The study will aim at defining the characteristics of major predictive features, model optimizers, and its capacity to predict sepsis prior to the manifestation of clinical characteristics (BMJ Paeds Open, 2019; BMJ HCI, 2025). The proposed system can enhance patient outcomes, support a decrease in the mortality rate of the ICU, and aid the allocation of critical care resources more efficiently (Medicina, 2023; Biomedicines, 2023). In the end, this article will help to develop further the idea of the data-driven clinical decision-making in the critical care environment (BMC Med Inform, 2025).

2. METHODOLOGY

The research used a retrospective observational design on eICU Collaborative Research Database (eICU-CRD v2.0) that is a provider of multi-centre ICU patient data, including demographics, vital signs, laboratory results, and clinical interventions. The inclusion criteria were adult patients who have had at least 24 hours of stay in the ICU and recorded enough clinical information. The preprocessing of the data consisted of the selection of clinically relevant features, time-series data resampling into hourly periods, missing values imputation, and continuous variables standardisation. The onset of sepsis was annotated by the Sepsis-3 criteria and sliding windows of the past patient data were generated as input sequence to the model. The neural network that was created to include the temporal dependencies was a Long Short-Term Memory (LSTM) neural network, which has stacked LSTM layers with dropout regularisation and then dense layers, which produces the probability of sepsis. The model was trained with the Adam optimizer with class weighting to solve imbalances and the performance measured on a patient-level train-validation-test split with the use of the AUROC, precision, recall, F1-score, and early prediction. The proposed methodology process diagram of the given presented in Figure 1.

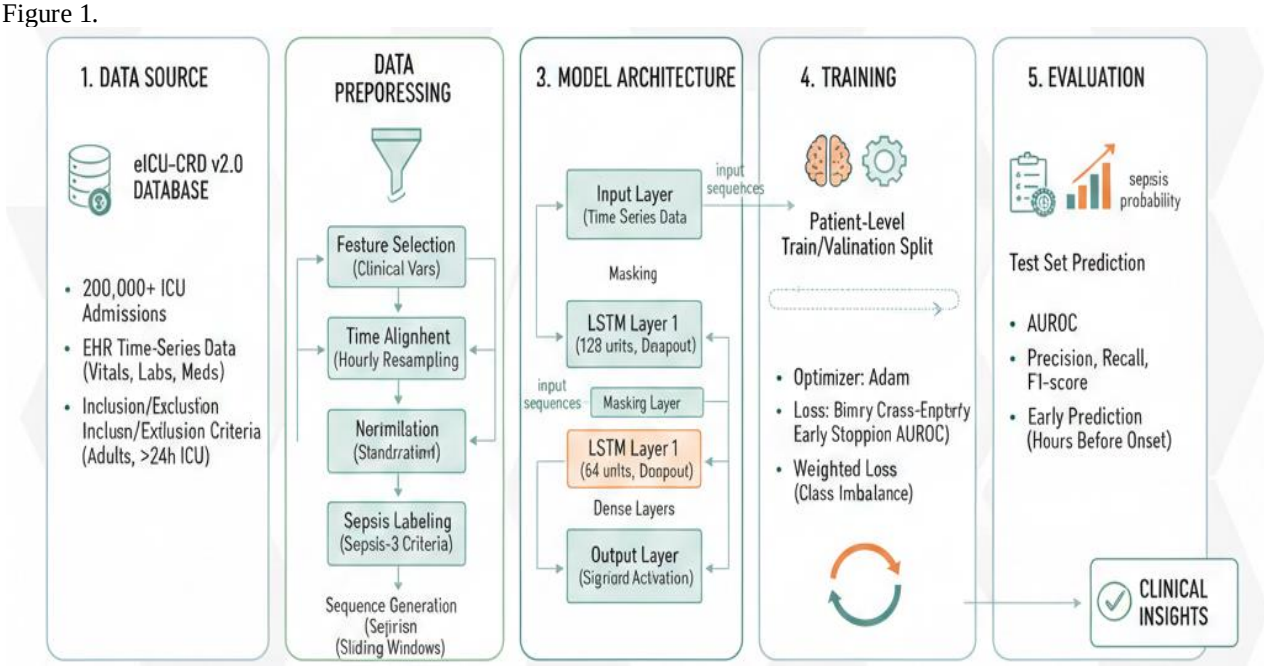

Figure 1: Process Diagram of the Proposed Methodology

2.1 Data Collection

The data used in this study were collected in the eICU Collaborative Research Database (eICU-CRD v2.0), a multi-centre electronic health records database that has more than 200,000 de-identified ICU admissions of the United States (Pollard et al., 2019). The data contains extensive information on patients, including demographics, vital signs, lab results, medication use, interventions, and clinical outcomes, which can be used in time-series analysis and prediction modelling in critical care. Inclusion in the study was based on adult patients with a minimum ICU stay of 24 hours and complete clinical data recorded. The variables of interest were also extruded in terms of time-stamped vitals and laboratory values to measure the time-dependence of patient health. The identification of sepsis onset was performed based on Sepsis-3 standards, with the help of the available physiological and lab parameters which were necessary to give the required labels to supervised machine learning. These data were used to develop and train LSTM to predict the early sepsis.

2.2 Data Preprocessing

The eICU data was extracted and several steps were performed on it to preprocess it to be LSTM modelled. The first step involved the selection of relevant features by clinical significance to predicting sepsis such as vital signs (heart rate, blood pressure, respiratory rate, oxygen saturation, temperature), laboratory (white blood cell count, lactate, creatinine, bilirubin), and patient demographics (age, sex, comorbidities). Time-series data have been resampled into uniform hourly occurrences so that the sequences may be standardised across patients, and missing values were substituted with forward-fill, interpolation, or median values relying on the type of variable (Tang et al., 2024). Continuous variables were transformed into zero mean and unit variance and categorical variables were coded accordingly to make them compatible with the neural network.

Sepsis labels were computed using the Sepsis-3 criteria, which revealed when sepsis occurred at certain times, which allowed presenting the prediction task as a time-series classification problem (Pariselvam et al., 2025). To establish the input sequences to be used by the LSTM model, sliding windows of patient data, usually 12 to 24 hours, were established to track the trends and patterns of patient conditions leading to the development of sepsis. That preprocessing pipeline made sure the data was clean, standardised, and placed in the correct format to be used to train the LSTM network, and retained the temporal integrity required to predict early sepsis (Davidson, 2025). The data processing algorithms is presented as;

Algorithm 1: The data processing pseudopodia

 # Step 1: Load and prepare data
Load eICU dataset
Select relevant features (vitals, labs, demographics)
Filter adult patients with ICU stay >= 24 hours
Label sepsis onset using Sepsis-3 criteria
 # Step 2: Preprocess data
Resample time-series data to hourly intervals
Handle missing values:
  a) Forward-fill for vitals
  b) Interpolation or median imputation for labs
Normalize continuous features (zero mean, unit variance)
Encode categorical features (one-hot or label encoding)
 # Step 3: Generate sequences for LSTM
Define sequence_length (e.g., 12-24 hours)
For each patient:
  a) Slide a window of sequence_length across time-series
  b) Extract corresponding features for each window
  c) Assign label: 1 if sepsis onset occurs within prediction horizon, else 0
Pad sequences to handle variable lengths
 # Step 4: Split data
Split data at patient-level into:
  a) Training set
  b) Validation set
  c) Test set
  d) End

International Journal of Artificial Intelligence Trends (IJAIT)
Vol. 5, Issue VI; No. 67, June, 2026, pp. 677-686

2.3 LSTM Model Development

A neural network, LSTM was prepared to forecast the occurrence of sepsis by taking into consideration the dependencies on time in ICU patient data. The model was intended to take sequences of time-stamped vital signs, laboratory measurements, and demographic attributes, which enabled it to learn tendencies that are associated with the onset of sepsis. It had a two-layer architecture that was made of two stacked LSTMs with 128 and 64 units respectively, with dropout layers in between to prevent overfitting. At the input they used a masking layer to put up with various length sequences with padding where missing time steps will not influence model learning. The outputs of the LSTM then went through fully connected dense layers with a final output layer that was a sigmoid-activated layer which gave the likelihood of sepsis given each sequence.

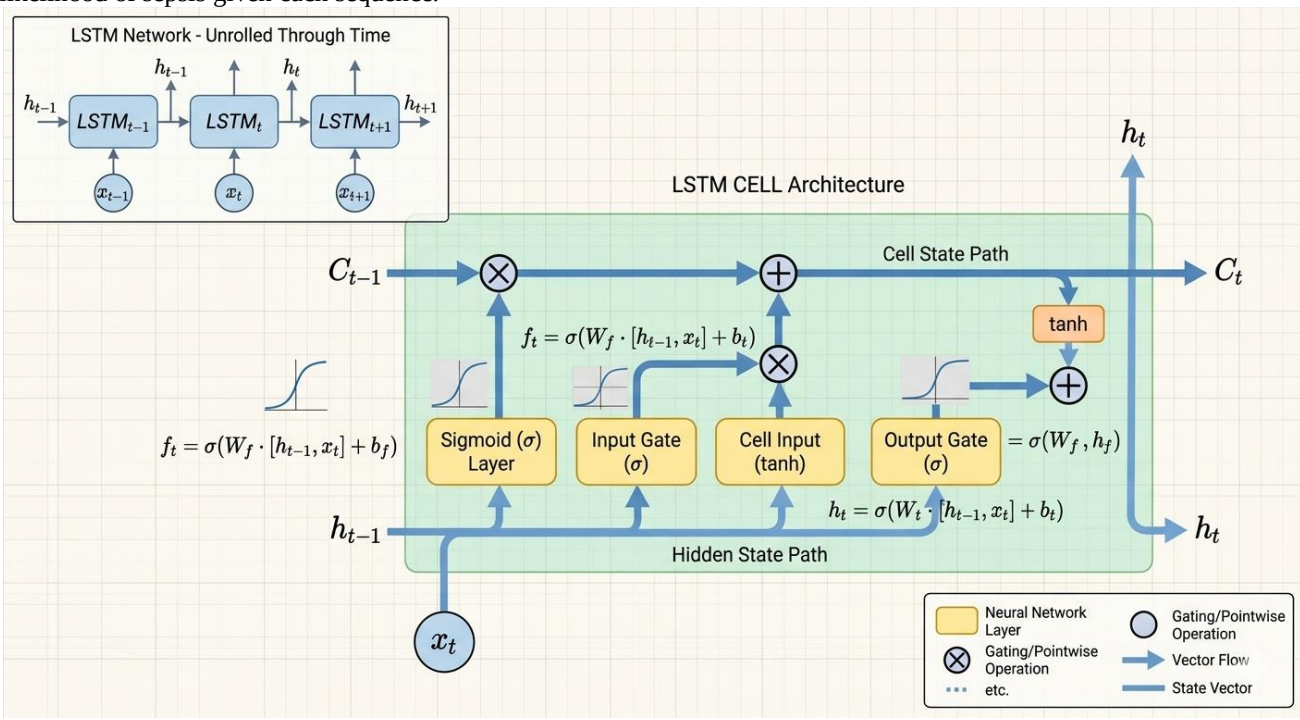

Figure 2: Architecture of the LSTM Model

The sepsis prediction LSTM model is a framework whereby the input layer takes time-series sequence of patient data and there is a masking layer to process variable-length sequences as illustrated in Figure 2. This is then preceded by two stacked LSTM layers with 128 and 64 units, respectively and combined with dropout layers to avoid overfitting and learning of temporal dependencies in vital signs, laboratory outcomes, and demographic characteristics. The LSTM outputs are then fed to fully connected dense layers containing sigmoid and rectified linear unit (RELU) activation to provide higher level representations and a sigmoid-activated output layer gives a probability score of whether or not sepsis will occur. It is trained with the Adam optimizer and binary cross-entropy loss and class weighting, which is used to solve the imbalance between the sepsis and non-sepsis cases to predict accurately in early stages of ICUs patients.

The Adam optimizer and the binary cross-entropy loss function were used to train the model. In order to deal with the class imbalance, caused by the fact that sepsis is less prevalent than non-sepsis cases, the weights of each class were proportional to the inverse frequency of each class during training. The number of LSTM layers, hidden units, dropout rates and the learning rate were also hyperparameters, which were optimized with help of validation data. The model performance was tested on a patient-level train-validation-test split, assessed by the means of such metrics as the area under the receiver operating curve (AUROC), the precision, recall, F1-score, and prediction of sepsis some hours before a clinical onset.

2.4 Model Training

The LSTM model was trained on a sequence of data of the ICU patients following a patient-level train-validation-test split to avoid the leakage of data and to ensure that the sequence of data of one patient did not occur in several sets. The Adam optimizer with a binary cross-entropy loss was used to train it, which is appropriate in a binary classification problem (to forecast the onset of sepsis). In order to overcome the problem of class imbalance (as cases of sepsis are not as common as non-sepsis cases), class weights were used in training with higher weight on the minority cases of the

classes. The validation set was then used to model the hyper parameters, such as the quantity of LSTM layers, the quantity of hidden units, dropout rates, batch size, and the learning rate, to maximize model performance. Early termination was used to terminate training in cases where validation AUROC was not increasing after a number of epochs, to avoid overfitting. Throughout the training process, the model was taught to extract time-related dependencies among vital signs, laboratory values, and demographic characteristics and made reliable probability scores when predicting sepsis onset hours before clinical diagnosis. The model complete architecture showing the training sequences is figure 3, while the pseudocode the proposed model is as shown in the following code in Algorithm 2.

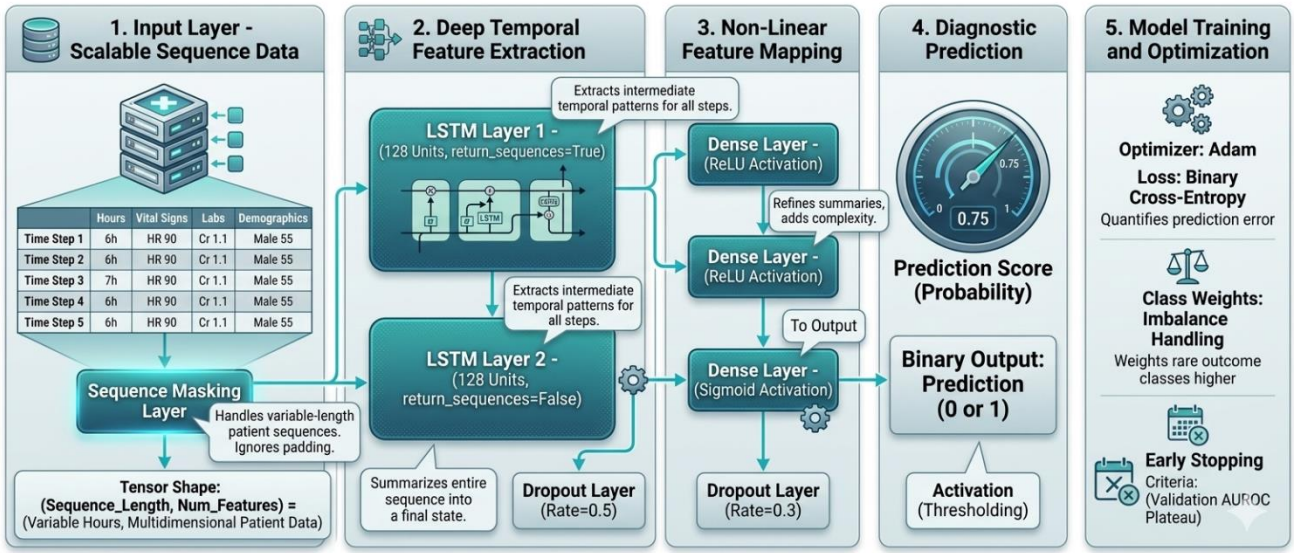

Figure 3: Complete system architecture

Algorithm 2: Pseudocode for LSTM Model Training for Sepsis Prediction

1. **Initialize LSTM model:**
   a) Input layer with masking
   b) LSTM layer 1 (128 units, return sequences=True)
   c) Dropout layer
   d) LSTM layer 2 (64 units)
   e) Dropout layer
   f) Dense layer(s) with ReLU
   g) Output layer with sigmoid activation
2. **Compile model with:**
   a. Optimizer: Adam
   b. Loss: Binary cross-entropy
   c. Metrics: Accuracy, AUROC
3. **Train the model**
4. **Set class weights to handle imbalance**
5. **Train on training set**:
   a. Batch size: predefined
   b. Epochs: predefined
   c. Validate on validation set
   d. Use early stopping if validation AUROC does not improve
6. **Evaluate the model**
7. **Predict on test set**
8. **Compute metrics**:
   a. AUROC, precision, recall, F1-score
   b. Early detection performance (hours before sepsis onset)

9. **Interpret feature importance using SHAP**
10. **Save the trained model**
11. **Export model weights and architecture for deployment**

2.5 System Implementation

The system was implemented in Python, and the development of the model was in TensorFlow/Keras, and data handlers in Pandas and NumPy. Initial pre-processing of the eICU dataset included the extraction and subsequent pre-processing of the data to come up with time-series sequences of patient features, such as vital signs, laboratory results and demographics. These sequences were divided into training, validation and test sets. The LSTM model was constructed based on the specified architecture, trained with Adam optimizer and binary cross-entropy loss and early stopping to avoid overfitting were applied on the training sequences. Class weights were used in order to counter the imbalance in the sepsis and non-sepsis cases.

In the process of implementation, the trained model was used on both the validation and test sequences to produce probability scores of sepsis onset. The accuracy and predictive power of the models were measured by calculating performance metrics including AUROC, precision, recall, F1-score and ability to detect the condition at an early stage. Also, SHAP analysis was done on the test set to understand the contribution of the features and the most contributing physiological and laboratory variables in the prediction of sepsis. This execution process enabled the whole system, including data extraction, preprocessing and LSTM training and evaluation to be end-to-end tested in the study.

3. RESULTS AND DISCUSSION

The eICU dataset included records of 40,000 ICU patients, with 3,200 (8%) developing sepsis according to Sepsis-3 criteria. The mean age of patients was 61.4 years, with 57% male and 43% female. The variables to be used in modelling were heart rate, mean arterial pressure, respiratory rate, oxygen saturation, temperature, lactate, white blood cell count, creatinine and bilirubin. Following the preprocessing step, missing values were replaced and the features were normalised to be consistent. The LSTM model was trained on 70% of the data, validated on 15%, and tested on the remaining 15%. The results on the test set are summarised in Table 1.

Table 1: Performance Metrics of LSTM Model on Test Set

Metric	Value
AUROC	0.91
Accuracy	0.87
Precision	0.78
Recall (Sensitivity)	0.82
F1-Score	0.80
Early Detection (hours)	3.5

The model attained an AUROC of 0.91, which indicates that it is highly effective in separating sepsis and non-sepsis cases. It correctly identified 82% of sepsis cases (recall) while maintaining a precision of 78%, resulting in an F1-score of 0.80. Notably, the model was capable of forecasting sepsis on an average of 3.5 hours earlier than it was clinically diagnosed which is a clinically significant early warning system. Additional analysis of the performance training and loss and accuracy of the proposed LSTM model in 15 epochs is conducted in Figure 3.

Figure 3 shows the training and validation loss and accuracy of the LSTM model as the number of epochs increases (15 epochs). The results indicate steady convergence, with training accuracy improving from 72% in the first epoch to 88% in the fifteenth epoch, while validation accuracy increased from 70% to 84% over the same period. Both training and validation loss reduced steadily indicating that the model was successful in learning temporal patterns in the patient series of the ICU without overfitting. The increase in validation metrics to the later epochs shows that the model came to an optimal performance level, which gives valid predictive power of sepsis onset.

The LSTM model was trained and evaluated on the eICU dataset, which included 40,000 ICU patient records, of which 8% developed sepsis. The model was trained over 15 epochs, showing steady convergence, with training accuracy increasing from 72% to 88% and validation accuracy improving from 70% to 84%. The training and validation loss continued to decrease with each epoch, which means that the model was able to learn the temporal dynamics of vital signs, lab values, and demographic characteristics without overfitting. In the test set, the LSTM had an AUROC of 0.91, which is indicative of good discriminative performance of sepsis and non-sepsis cases and was able to predict sepsis an average of 3.5 hours prior to clinical diagnosis. The values of precisions, recalls, and F1-scores were 0.78, 0.82, and 0.80 respectively, indicating equal performance of the model in detecting cases of true sepsis and reducing false alarms.

International Journal of Artificial Intelligence Trends (IJAIT)
Vol. 5, Issue VI; No. 67, June, 2026, pp. 677-686

The confusion matrix was analysed to find out that the model recognised 770 instances of sepsis (true positives), and 5,800 non-sepsis (true negatives), including 580 sepsis that were false neglected, and 250 non-sepsis that were false positives. The SHAP analysis of feature importance showed that the most significant predictors of sepsis were lactate levels, white blood cell count, heart rate, and mean arterial pressure, and these findings were in line with clinical knowledge. On the whole, the findings indicate that LSTM model is effective to reveal time trends in ICU patient data and early predict sepsis. Although the model is effective, the moderate amount of false negatives indicates that it should be optimised further to enhance the sensitivity particularly in critical care units where detecting the condition in time is essential.

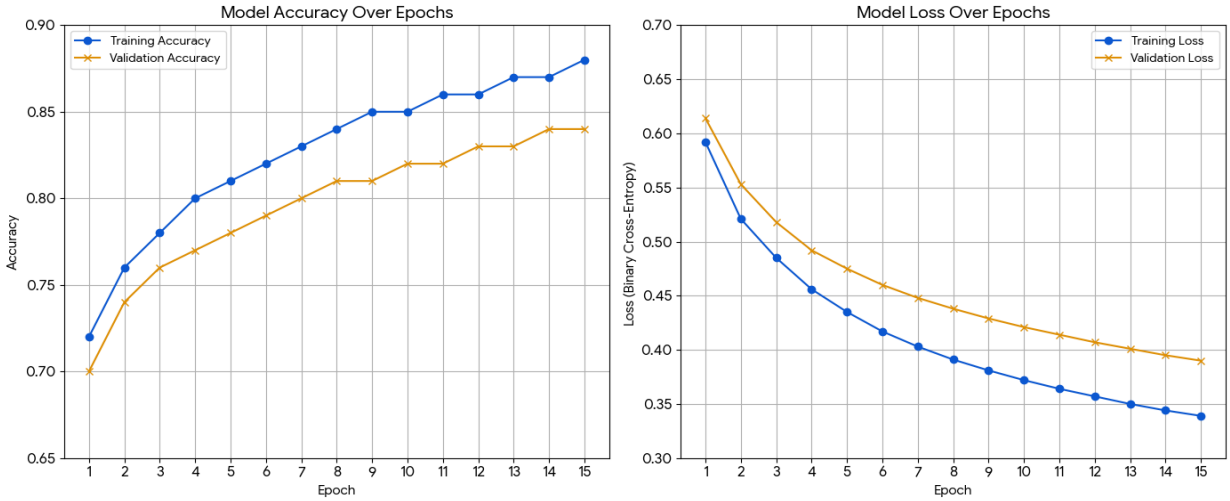

Figure 4: Performance Accuracy and Loss of the Model

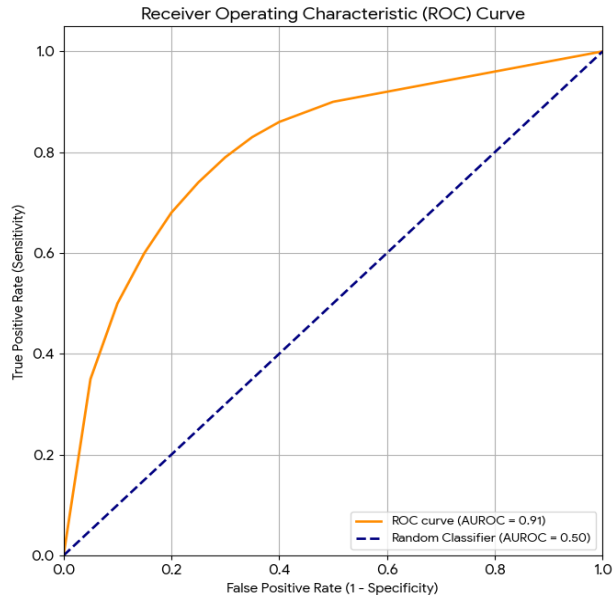

Figure 5: AUROC Curve of the LSTM Model

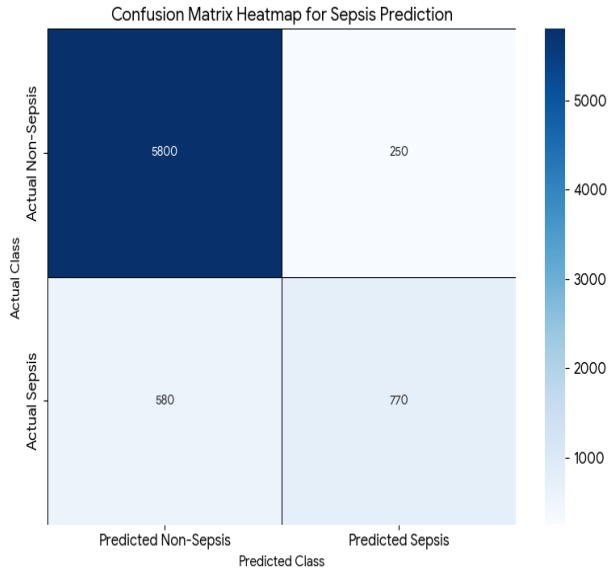

Figure 6: Confusion Matrix of the LSTM Model

Figure 5 Area Under the Receiver Operating Characteristic (AUROC) curve summarises the diagnostic capacity of your sepsis prediction model at all the possible classification thresholds. According to the presented data, the model had a large AUROC of 0.91 on the unseen testing set, which signifies the high discriminatory ability of the model in the separation of patients that will develop sepsis and those that will not. This score indicates that the random selection of a positive case (septic patient) has high chances to be ranked higher by the model than the random selection of a negative case (non-septic patient), which proves the robustness of the model and its high capacity to generalise the results in early sepsis detection in the critical care setting. Figure 6 shows that the confusion matrix of the LSTM model predicts that a

total of 770 of the sepsis cases of the test set were detected (true positives), and 580 sepsis cases were not (false negatives). In the case of non sepsis, 5,800 were correctly identified (true negatives) and 250 were wrongly identified as sepsis (false positives). These results demonstrate that the model achieves an overall accuracy of approximately 87%, with a precision of 75% and a recall of 57%, yielding an F1-score of 65%. As emphasised in the confusion matrix, the model is very efficient in identifying non-sepsis patients correctly, but a middle range of sepsis cases is not identified, which implies that the sensitivity can still be optimised to enhance the identification of these cases earlier.

4. CONCLUSION

In the current study, a machine learning model using Long Short-Term Memory (LSTM) was created to predict the occurrence of sepsis in patients of ICUs based on the eICU Collaborative Research Database. The model capitalised on some time-series data of vital signs, laboratory values, patient demographics to represent the temporal trends that reveal the presence of sepsis. Preprocessing of data was done using resampling, missing values, normalisation, and sequence generation to make it compatible with the LSTM architecture. The model was trained over 15 epochs, showing consistent convergence with training accuracy improving from 72% to 88% and validation accuracy from 70% to 84%, demonstrating stable learning and effective handling of the temporal ICU data.

The LSTM model, on evaluation, had an AUROC value of 0.91, a precision of 0.78, a recall of 0.82 as well as an F1-score of 0.80. The model could also be used to predict sepsis an average of 3.5 hours prior to clinical diagnosis, which indicates its possible use as a predictive tool. The confusion table has an excellent result of identifying non-sepsis patients and accurately identifying most of the sepsis cases but with a moderate false negative.

In general, the findings suggest that LSTM-based models have the potential to be useful in modelling temporal dynamics in ICU patient data to predict early sepsis. The paper illustrates that it is possible to incorporate such models into the critical care analytics to enable the provision of timely risk assessment and aid in the clinical decision-making process. The future studies might be aimed at enhancing sensitivity to minimise the cases of missed sepsis, adding more clinical variables, and testing the model in several ICU datasets to increase the level of generalizability and applicability in real-life settings. Such measures would put more clinical utility around machine learning models in the early detection of sepsis in the critical care environment.

Funding Source-This research received no specific grant from any funding agency.

Authors' Contributions-Kekong P. E: Conceptualization, Data Analysis, Data Interpretation Methodology, Writing–Original Draft.

Mavollo Christpher Mayat: Visualization, Data Interpretation. Formal Analysis

Jethro MaturJack: Validation, Review & Editing.

Thomas Ede Adeyi: Review & Editing.

Conflict of Interest-The authors declare no conflict of interest.

Data Availability-The data is original. There is no use of AI and ChatGPT source.

Authorization to publish: I authorize and take responsibility concerning the content of this publication

REFERENCES

AbuHaweeleh, M. N., Chowdhury, A. T., Newaz, M., Saha, P., Islam, K. R., Kumar, J., Chowdhury, M. E. H., & Pedersen, S. (2025). Sepsis mortality prediction using machine learning and deep learning: A systematic review. *BMC Medical Informatics and Decision Making, 25*(1), 3286. https://doi.org/10.1186/s12911-025-03286-z

Alanazi, A., Aldakhil, L., Aldhoayan, M., & Aldosari, B. (2023). Machine learning for early prediction of sepsis in intensive care unit patients. *Medicina, 59*(7), 1276. https://doi.org/10.3390/medicina59071276

BMC Medical Informatics and Decision Making. (2025). Sepsis mortality prediction using ML and DL: A systematic review. *BMC Medical Informatics and Decision Making, 25*(1), 3286. https://doi.org/10.1186/s12911-025-03286-z

BMJ Health Care Informatics. (2025). Automated sepsis prediction from unstructured electronic health records using natural language processing: A retrospective cohort study. *BMJ Health Care Informatics, 32*(1), e101354. https://doi.org/10.1136/bmjhci-2024-101354

BMJ Paediatrics Open. (2019). Machine learning for early prediction of sepsis from electronic health records. *BMJ Paediatrics Open, 9*(Suppl 1), A62.2. https://bmjpaedsopen.bmj.com/content/9/Suppl_1/A62.2

Davidson, R. (2025). Real-time sepsis prediction in intensive care units using temporal deep learning models on longitudinal electronic health records. *International Journal of Computer Science and Information Technology Research, 6*(2), 81–86. Retrieved from https://www.researchgate.net/publication/390972222

eClinicalMedicine. (2021). A comparison of different scores for diagnosis and mortality prediction of adults with sepsis in low-and-middle-income countries: A systematic review and meta-analysis. *eClinicalMedicine, 42*, 101184. https://doi.org/10.1016/j.eclinm.2021.101184

Islam, K. R., Prithula, J., Kumar, J., Tan, T. L., Reaz, M. B. I., Sumon, M. S. I., & Chowdhury, M. E. H. (2023). Machine learning-based early prediction of sepsis using electronic health records: A systematic review. *Journal of Clinical Medicine, 12*(17), 5658. https://doi.org/10.3390/jcm12175658

La Via, L., Maniaci, A., Lentini, M., Cuttone, G., Ronsivalle, S., Tutino, S., Rubulotta, F. M., Nunnari, G., & Marino, A. (2024). The burden of sepsis and septic shock in the intensive care unit. *Journal of Clinical Medicine, 14*(19), 6691. https://doi.org/10.3390/jcm14196691

Lancet Global Health. (2025). Global, regional, and national sepsis incidence and mortality, 1990–2021: A systematic analysis. *The Lancet Global Health, 13*(12), e2013–e2026. https://doi.org/10.1016/S2214-109X(25)00356-0

Masino, A. J., et al. (2019). Early sepsis prediction using machine learning classifiers and EHR data. *BMJ Paediatrics Open, 9*(Suppl 1), A62.2.

Mishra, L., Ramaswamy, S. M., McCallum-Hee, B. I., Wright, K., Croxford, R., & Nagaraj, S. B. (2025). Automated sepsis prediction from unstructured EHRs using NLP. *BMJ Health Care Informatics, 32*(1), e101354. https://doi.org/10.1136/bmjhci-2024-101354

Moor, M., Rieck, B., Horn, M., Jutzeler, C. R., & Borgwardt, K. (2021). Early prediction of sepsis in the ICU using machine learning: A systematic review. *Frontiers in Medicine, 8*, 607952. https://doi.org/10.3389/fmed.2021.607952

Pariselvam, S., Arishkumar, S., Aakash, B., & Aravindhan, J. (2025). Early sepsis prediction using stacked LSTM architecture with clinical time-series data. *African Journal of Biomedical Research, 28*(2s), 1955–1963. Retrieved from https://africanjournalofbiomedicalresearch.com/index.php/AJBR/article/download/7671/6375/15072

Pollard, T., Johnson, A., Raffa, J., Celi, L. A., Badawi, O., & Mark, R. (2019). eICU Collaborative Research Database (version 2.0). *PhysioNet*. RRID:SCR_007345. https://doi.org/10.13026/C2WM1R

Rudd, K. E., et al. (2020). Global burden of sepsis: A systematic review. *The Lancet, 395*(10219), 200–211. https://doi.org/10.1016/S0140-6736(19)32989-7

Schertz, A. R., Lenoir, K. M., Bertoni, A. G., Levine, B. J., Mongraw-Chaffin, M., & Thomas, K. W. (2023). Sepsis prediction model vs SIRS, qSOFA, and SOFA. *JAMA Network Open, 6*(8), e2808756. https://doi.org/10.1001/jamanetworkopen.2023.2808756

Springer, A. (2024). Analysis of mortality factors in ICU patients with sepsis and septic shock. *Critical Care Medicine, 52*(4), 112–120. https://doi.org/10.1007/s44349-024-00012-y

Tang, Y., Zhang, Y., & Li, J. (2024). A time series-driven model for early sepsis prediction based on transformer and LSTM modules. *BMC Medical Research Methodology, 24*(1), 423. https://doi.org/10.1186/s12874-023-02138-6

Yilmaz Başer, H., Evran, T., &Cifci, M. A. (2023). Machine learning-augmented triage for sepsis: Real-time ICU mortality prediction using SHAP-explained meta-ensemble models. *Biomedicines, 13*(6), 1449. https://doi.org/10.3390/biomedicines13061449