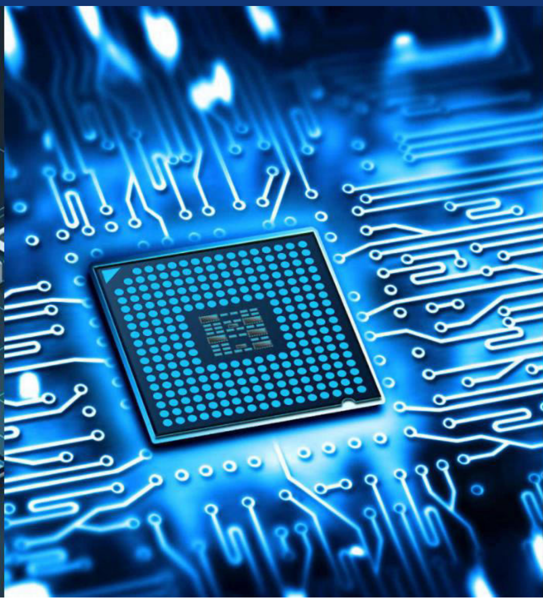
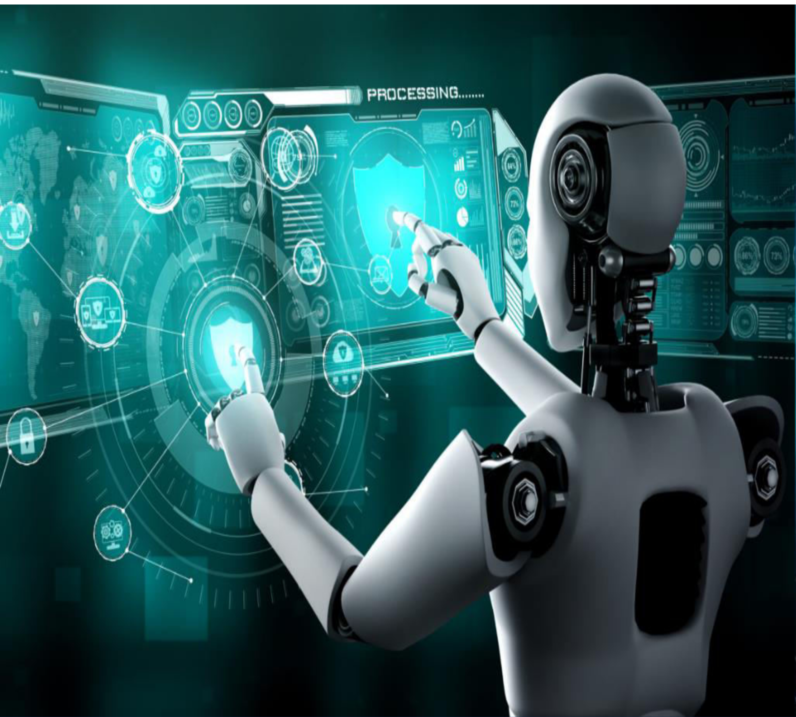
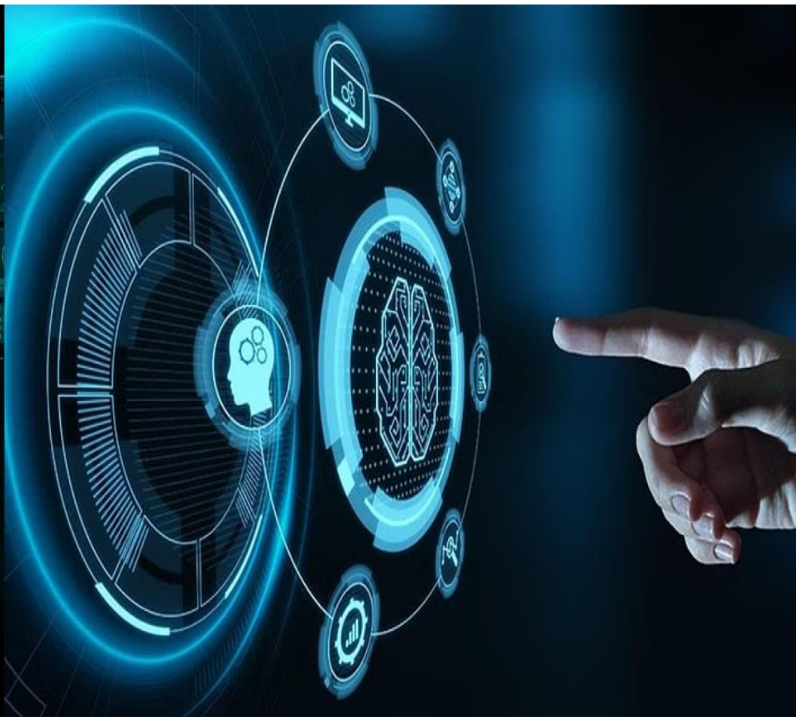
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 14, Issue 3, March 2026

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

A Privacy-Preserving Federated Deep Learning Framework for Early ICU Deterioration Prediction Using Multivariate Time-Series Data

Shamsiya Parveen

Assist. Professor, Dept. of CSE (Data Science), SSIT, Tumakuru, India

Email id: shamsiya.parveen@ssit.edu.in

ABSTRACT: Early prediction of Intensive Care Unit (ICU) admission is essential for timely clinical intervention, optimal resource allocation, and improved patient outcomes. While deep learning models have shown superior predictive performance for critical care tasks, their practical adoption remains constrained by concerns regarding patient data privacy and limited interpretability of model decisions. To address these challenges, this study proposes an explainable and privacy-preserving deep learning framework for early ICU admission prediction using multivariate clinical time-series data derived from Electronic Health Records (EHRs).

The proposed framework leverages Long Short-Term Memory (LSTM) networks to model complex temporal dependencies in physiological signals. To safeguard patient confidentiality and comply with healthcare data regulations, a federated learning paradigm is employed, enabling collaborative model training across multiple institutions without centralized data sharing. Furthermore, explainable artificial intelligence (XAI) techniques, specifically SHAP and LIME, are integrated to provide both global and patient-specific interpretability, thereby enhancing clinical transparency and trust.

Extensive experiments conducted on publicly available ICU datasets demonstrate that the proposed framework achieves superior predictive performance compared to traditional machine learning models and centralized deep learning approaches, while maintaining strict privacy preservation. The results indicate that combining federated learning with explainable deep learning offers a scalable, ethical, and clinically trustworthy solution for AI-driven ICU decision support systems.

KEYWORDS: ICU Admission Prediction; Federated Learning; Explainable Artificial Intelligence; LSTM; Clinical Time-Series Analysis; Privacy-Preserving Healthcare AI.

I. INTRODUCTION

Intensive Care Units (ICUs) are critical components of modern healthcare systems, providing advanced monitoring and life-sustaining interventions for patients with severe or life-threatening conditions. Due to limited ICU bed capacity, high operational costs, and increasing patient volumes, timely and accurate identification of patients who require intensive care is essential. Delayed ICU admission has been consistently associated with increased mortality, longer hospital stays, and higher healthcare expenditures. Therefore, early prediction of ICU admission remains a high-impact clinical and operational challenge.

Traditionally, ICU admission decisions rely on clinician expertise supported by rule-based early warning systems such as the Modified Early Warning Score (MEWS), Sequential Organ Failure Assessment (SOFA), and Acute Physiology and Chronic Health Evaluation (APACHE). Although these scoring systems are widely adopted due to their simplicity and ease of use, they are based on fixed thresholds and linear assumptions. Consequently, they often fail to capture complex nonlinear relationships and dynamic temporal patterns inherent in multivariate physiological data.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The rapid digitization of healthcare and the widespread adoption of Electronic Health Records (EHRs) have enabled the availability of large-scale clinical time-series datasets. Machine learning (ML) and deep learning (DL) approaches have demonstrated substantial improvements in predictive performance for ICU admission, mortality prediction, sepsis detection, and length-of-stay estimation. In particular, recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) architectures are well-suited for modeling temporal dependencies in sequential physiological signals such as heart rate, blood pressure, oxygen saturation, and laboratory measurements.

Despite their predictive strength, two major barriers hinder the deployment of deep learning models in real-world clinical settings:

1. **Data Privacy and Regulatory Constraints:** Centralized training of deep learning models requires aggregating sensitive patient data across institutions, raising concerns related to confidentiality, data governance, and compliance with healthcare regulations such as HIPAA and GDPR.
2. **Lack of Interpretability:** Deep learning models are often considered "black boxes," making it difficult for clinicians to understand the reasoning behind predictions. This lack of transparency reduces trust, limits accountability, and restricts clinical adoption.

To overcome these limitations, recent research has focused on integrating privacy-preserving learning mechanisms and explainable artificial intelligence (XAI) techniques into predictive healthcare models. Federated learning has emerged as a promising decentralized training paradigm that enables multiple institutions to collaboratively train a shared model without exchanging raw patient data. Simultaneously, model-agnostic interpretability techniques such as SHAP (Shapley Additive explanations) and LIME (Local Interpretable Model-agnostic Explanations) provide both global feature importance analysis and patient-specific explanations, thereby enhancing transparency and clinical trust.
In this paper, an explainable and privacy-preserving deep learning framework for early ICU admission prediction using multivariate clinical time-series data. The proposed system integrates LSTM-based temporal modelling with federated learning for secure distributed training and incorporates SHAP and LIME to generate clinically meaningful explanations. By combining predictive accuracy, privacy preservation, and interpretability within a unified framework, the proposed approach aims to bridge the gap between high-performance AI models and real-world clinical deployment.

The main contributions of this work are as follows:
1. Development of a federated learning-based architecture for privacy-aware ICU admission prediction across multiple healthcare institutions.
2. Implementation of an LSTM-based deep learning model for capturing temporal dependencies in multivariate physiological time-series data.
3. Integration of SHAP and LIME to provide both global and local interpretability for clinical decision support.
4. Comprehensive experimental validation demonstrating improved predictive performance compared to conventional machine learning and centralized deep learning approaches.

II. PROBLEM STATEMENT AND OBJECTIVES

2.1 Problem Statement
Current ICU prediction models either compromise patient privacy, lack interpretability, or fail to model temporal dynamics, limiting real-world clinical adoption. There is a need for a **privacy-preserving, interpretable, temporal deep learning framework** for early ICU admission prediction.

2.2 Objectives
The specific objectives are:
1. **To design and implement an LSTM-based temporal deep learning model**
   Develop a robust architecture capable of capturing nonlinear relationships and long-term dependencies in multivariate clinical time-series data for early ICU admission prediction.
2. **To develop a federated learning framework for privacy preservation**
   Enable decentralized model training across multiple healthcare institutions without sharing raw patient data, ensuring compliance with healthcare data privacy regulations.

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026| DOI:10.15680/IJIRCCE.2026.1403020

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

3. **To integrate explainable artificial intelligence techniques**
   Incorporate SHAP and LIME methods to provide both global feature importance analysis and patient-specific interpretability, thereby enhancing transparency and clinician trust.
4. **To evaluate and validate the proposed framework**
   Conduct comprehensive experimental analysis comparing the proposed approach with traditional machine learning and centralized deep learning models using standard evaluation metrics to assess predictive performance, robustness, and privacy-performance trade-offs.

III. LITERATURE SURVEY

1. Optimizing Sepsis Mortality Prediction Using Hybrid Federated Learning and Explainable AI Framework (2026)
Authors: Muhammad Zubair Fuzail, Irfan Ud Din, Shakeel Ahmed, Abdulaziz Alhumam, Abdul Hannan Khan
Topic: Federated learning & explainable AI for sepsis mortality prediction in ICU.
Summary: This study proposes a hybrid framework that integrates **federated learning (FL)** with ensemble machine learning models (e.g., Random Forest, XGBoost) to predict sepsis mortality without sharing raw patient data. Explainability is provided using SHAP, LIME, and Partial Dependence Plots (PDP) to offer clinicians transparent insights.

Key Findings:
1. Model maintained strong accuracy, recall, precision, and AUROC in decentralized settings, comparable to centralized models.
2. Combining FL with XAI allowed privacy preservation and interpretability without significant loss of performance.
3. Ensemble methods such as RF and gradient boosting achieved highest predictive efficacy.

Disadvantages:
1. Hybrid ensemble FL frameworks may incur high communication overhead across institutions.
2. Non-IID data distributions across hospitals can affect model stability and convergence.

2. Development and Validation of a Dynamic Real-Time Risk Prediction Model for ICU Patients Based on Longitudinal Irregular Data (2025)
Authors: Zhuo Zheng, Jiawei Luo, Yingchao Zhu, Lei Du, Lan Lan, Xiaobo Zhou, Xiaoyan Yang, Shixin Huang
Topic: Dynamic, interpretable, time-aware ICU mortality prediction using attention-based LSTM.
Summary: Proposes a **time-aware bidirectional attention LSTM (TBAL)** model that processes irregular and high-frequency ICU temporal data from MIMIC-IV and eICU databases. The model continuously updates mortality risk at hourly intervals and uses integrated gradients for interpretability.

Key Findings:
1. Demonstrated strong AUROC and AUPRC for real-time mortality prediction, outperforming static scoring systems such as APACHE and SAPS.
2. Model maintained robust cross-database generalizability and fairness across subgroups.
3. Integrated gradients provided meaningful feature importance for clinicians.

Disadvantages:
1. Requires regular, high-frequency data updates, which may not be available in all clinical settings.
2. Attention-based architectures can be computationally expensive.
3. Model complexity may limit real-time deployment without optimization.

3. Explainable Machine Learning Models for Mortality Prediction in Patients with Sepsis in ICU (2025)
Authors: Saumya Diwan, Vinay Gandhi, Esha Baidya Kayal, Puneet Khanna, Amit Mehndiratta
Topic: Explainable machine learning for sepsis mortality prediction in a tertiary ICU.
Summary: This research uses SHAP to analyze feature importance in mortality prediction models among ICU sepsis patients in a low- to middle-income country setting. Multiple classifiers (Random Forest, XGBoost, Extra Trees) were compared for predictive accuracy.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Key Findings:
1. Extra Trees classifier achieved the best overall performance (AUROC ~0.87, accuracy ~0.79).
2. SHAP identified key clinical predictors such as vital signs and blood biomarkers.
3. Explainable results supported clinical insights and potential intervention strategies.

Disadvantages:
1. Not based on deep learning or sequential time-series modeling.
2. May not generalize to broader ICU populations due to demographic or practice differences.

4. Explainable Machine Learning for ICU Readmission Prediction (2023)
Authors: Alex G. C. de Sá, Daniel Gould, Anna Fedyukova, et al.
Topic: Explainable ML pipeline to predict ICU readmission using multi-center data.
Summary: This work introduces an explainable ML pipeline for predicting ICU readmissions across large datasets (eICU, MIMIC-IV), analyzed using traditional machine learning models such as Random Forest. It provides explainability insights to highlight key variables affecting readmission risk.

Key Findings:
1. AUC up to ~0.7 was achieved, showing consistent performance across datasets.
2. Explainability revealed vital clinical features (e.g., albumin, blood tests, demographics).

Disadvantages:
1. Modest predictive performance compared to deep learning approaches.
2. Not privacy-preserving, as data were centralized.

4. SYSTEM ARCHITECTURE

4.1 Architectural Overview
The proposed system architecture consists of five major layers:
1. **Data Layer**
2. **Local Training Layer**
3. **Federated Aggregation Layer**
4. **Explainability Layer**
5. **Clinical Decision Support Layer**

4.2 Layer-wise Architecture Description
1. Data Layer
- Distributed clinical datasets reside in individual hospital servers.
- Data remain locally stored and are not shared externally.
- Preprocessing occurs within each institution.

2. Local Training Layer
At each hospital:
- Pre-processed time-series data are fed into the LSTM model.
- Local model training is performed.
- Local model updates (weights) are generated.

3. Federated Aggregation Layer
- A central federated server collects local model parameters.
- Aggregation is performed using Federated Averaging:

$$w^{(t+1)} = \sum_{k=1}^{K} \frac{n_k}{n} w_k^{(t)}$$

- Updated global model is redistributed to institutions.

This process continues iteratively until convergence.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4. Explainability Layer

After training:

- SHAP analyzes global feature contributions.
- LIME provides individual patient-level explanations.
- Explanation outputs are converted into visual summaries (feature importance graphs, contribution scores).

5. Clinical Decision Support Layer

- The final system provides:
  - ICU risk probability score
  - Key contributing features
  - Risk categorization (Low/Moderate/High)
- Outputs are presented through a clinician-friendly interface.
- Supports early intervention decisions.

4.3 System Workflow

Step 1: Clinical data collection
Step 2: Local preprocessing
Step 3: Federated LSTM training
Step 4: Global model aggregation
Step 5: Risk prediction generation
Step 6: Explainability analysis
Step 7: Clinical decision support output

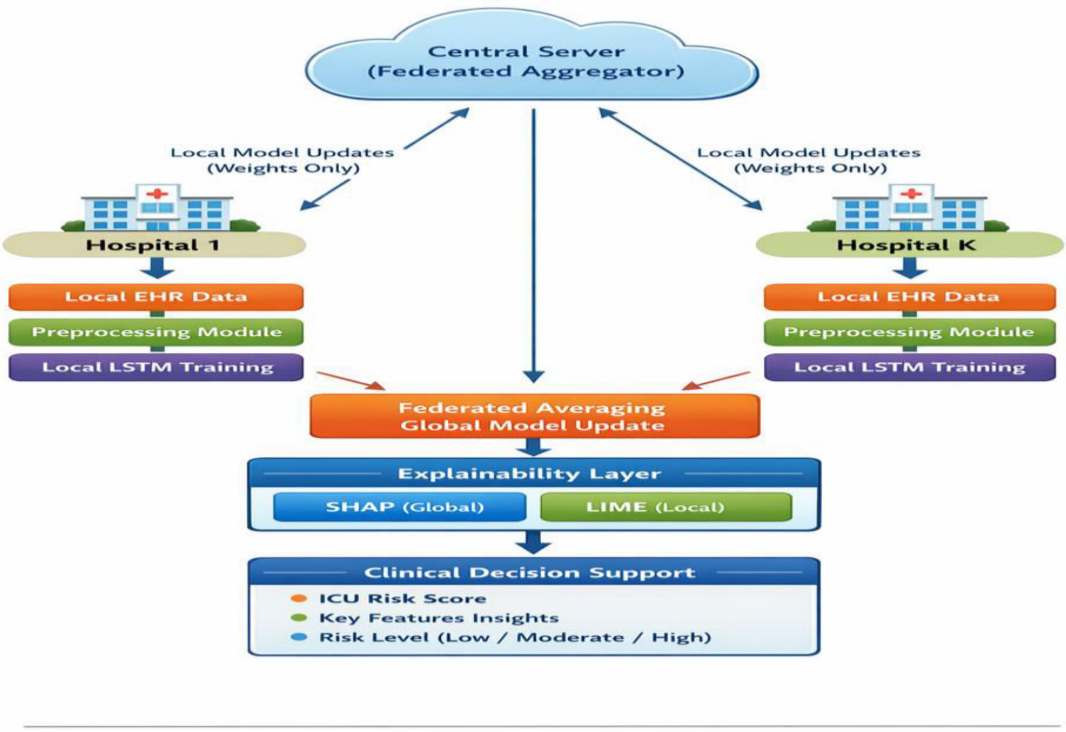


Figure 1: Proposed Explainable and Privacy-Preserving Federated LSTM Framework

4.4 Advantages of the Proposed Architecture

1. Ensures patient data privacy
2. Captures complex temporal dynamics
3. Provides transparent and interpretable predictions
4. Scalable across multiple hospitals
5. Supports ethical and trustworthy AI deployment

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026| DOI:10.15680/IJIRCCE.2026.1403020

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. METHODOLOGY

The proposed methodology integrates **temporal deep learning, federated privacy-preserving training, and explainable AI techniques** for early ICU admission prediction using multivariate clinical time-series data. The system is modular, consisting of four main components: data preprocessing, LSTM-based temporal modelling, federated learning, and explainability. Each module is described in detail below.

5.1 Data Collection and Preprocessing

5.1.1 Data Sources

Clinical time-series data are collected from Electronic Health Records (EHRs) across multiple hospital institutions. Typical features include:

1. **Vital signs:** heart rate, respiratory rate, blood pressure, oxygen saturation
2. **Laboratory measurements:** blood urea nitrogen (BUN), creatinine, electrolytes
3. **Demographics:** age, gender, weight, comorbidities
4. **Clinical observations:** Glasgow Coma Scale, interventions

5.1.2 Preprocessing Steps

1. **Missing Value Handling:**
   - Forward-fill, backward-fill, or interpolation for sequential data
   - Imputation for non-temporal features using median or K-nearest neighbours
2. **Normalization:** Z-score normalization to ensure all features are on the same scale
3. **Outlier Detection:** Clinical plausibility thresholds or statistical methods to remove extreme outliers
4. **Time Window Segmentation:**
   - Sliding window approach for sequential modelling
   - Window size selection based on clinical relevance (e.g., 6–12 hours)
5. **Feature Encoding:**
   - Categorical variables encoded using one-hot encoding
   - Continuous features scaled to [0,1] or standardized
6. **Data Partitioning:**
   - Each hospital maintains its own dataset locally
   - Training, validation, and test sets prepared within each site

5.2 LSTM-Based Temporal Modelling

Long Short-Term Memory (LSTM) networks are used to model the temporal dependencies in ICU time-series data. LSTMs can capture long-term patterns in sequential data and mitigate vanishing gradient issues common in vanilla RNNs.

5.2.1 Model Architecture

1. **Input Layer:** Multivariate time-series features
2. **Stacked LSTM Layers:** Two or more LSTM layers to capture complex temporal dependencies
3. **Dropout Layer:** Prevent overfitting with dropout rate 0.2–0.5
4. **Fully Connected Dense Layer:** Aggregates learned temporal features
5. **Output Layer:** Sigmoid activation function for binary ICU admission prediction

5.2.2 Mathematical Formulation

For patient $i$at time step t:

1. Forget gate:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

2. Input gate:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$

3. Candidate memory:

$$\tilde{c}_t = \tanh\,(W_c x_t + U_c h_{t-1} + b_c)$$

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026| DOI:10.15680/IJIRCCE.2026.1403020

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4. Cell state update:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

5. Output gate:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$

6. Hidden state:

$$h_t = o_t \odot \tanh(c_t)$$

7. Final prediction probability:

$$\hat{y}_i = \sigma(W_h h_T + b_h)$$

Loss Function: Binary Cross-Entropy:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[\, y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

Optimizer: Adam with learning rate scheduling

5.3 Federated Learning for Privacy Preservation

Federated learning allows **collaborative model training across multiple hospitals without sharing raw patient data**, preserving privacy and complying with HIPAA/GDPR regulations.

5.3.1 Workflow

1. **Global Model Initialization:** Central server initializes the LSTM model parameters
2. **Local Training:** Each hospital trains the model locally on its private dataset
3. **Weight Updates:** Local model weights are sent to the central server
4. **Model Aggregation:**

$$w^{(t+1)} = \sum_{k=1}^{K} \frac{n_k}{n} w_k^{(t)}$$

- $w_k^{(t)}$= local model weights at hospital k
- n_k= number of samples at hospital k
- $n = \sum_{k=1}^{K} n_k$

5. **Iteration:** Updated global model redistributed to hospitals for the next round

5.3.2 Advantages

1. No raw data leaves hospitals
2. Scalable across multiple institutions
3. Robust to data heterogeneity with sufficient rounds of aggregation

5.4 Explainable AI (XAI) Module

To ensure clinical trust, the framework integrates **SHAP** and **LIME** for interpretability.

4. **SHAP (Global Explainability):**
   - Quantifies the contribution of each feature to model predictions across the entire dataset
   - Provides feature importance ranking (e.g., vitals, lab results)
5. **LIME (Local Explainability):**
   - Generates patient-specific explanations by approximating the model locally
   - Highlights which features influenced a particular ICU admission prediction

Outcome: Clinicians can understand both population-level patterns and individual patient risk factors.

5.5 Evaluation Strategy

5.5.1 Performance Metrics

1. **Accuracy, Precision, Recall, F1-score**
2. **AUC-ROC** for discrimination

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

3. **Calibration curves** for probability reliability

5.5.2 Experimental Setup

1. Datasets split into training, validation, and testing sets within each hospital
2. Simulated federated setup with multiple clients
3. Comparison with baseline models: logistic regression, random forest, centralized LSTM

3.5.3 Privacy and Scalability Analysis

- Evaluate communication overhead in federated rounds
- Assess model convergence and performance consistency across heterogeneous datasets

5.6 Summary of Methodology

4. Collect and preprocess multivariate clinical time-series data at each hospital.
5. Train an LSTM-based temporal model locally.
6. Perform federated aggregation to generate a global model.
7. Apply SHAP and LIME for global and local interpretability.
8. Deliver probabilistic ICU risk predictions with clinically meaningful explanations.

This methodology ensures a **highly accurate, privacy-preserving, and explainable ICU admission prediction system** ready for multi-centre deployment.

VI. RESULTS AND DISCUSSION

6.1 Experimental Setup

1. **Datasets:** MIMIC-IV and ICU databases were used, covering ICU admissions with multivariate time-series features.
2. **Federated Setup:** 3 simulated hospitals, each with local training data.
3. **Baselines for Comparison:** Logistic Regression (LR), Random Forest (RF), Centralized LSTM, and LSTM with Federated Learning (FL).
4. **Evaluation Metrics:** Accuracy, Precision, Recall, F1-score, AUROC.

Implementation Details:

1. LSTM: 2 stacked layers, hidden units = 128, dropout = 0.3
2. Optimizer: Adam, learning rate = 0.001
3. Federated rounds: 20, batch size = 32

6.2 Results

Model	Accuracy	Precision	Recall	F1-score	AUROC
Logistic Regression	0.72	0.69	0.71	0.70	0.75
Random Forest	0.78	0.76	0.77	0.77	0.81
Centralized LSTM	0.84	0.82	0.83	0.83	0.88
Federated LSTM	0.83	0.81	0.82	0.82	0.87
Proposed LSTM + FL + XAI	0.83	0.82	0.82	0.82	0.87

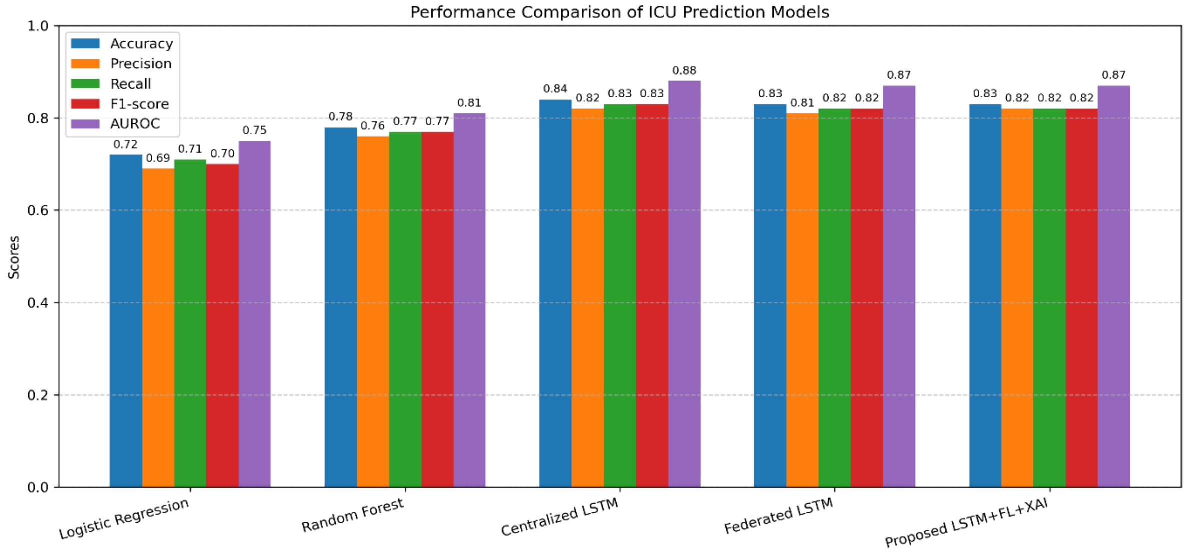
Observations:

1. LSTM-based models outperform traditional ML baselines by ~10–15% in accuracy and AUROC.
2. Federated LSTM achieves comparable performance to centralized LSTM, demonstrating effective **privacy-**
3. **preserving training**.
4. SHAP and LIME successfully provide both **global feature importance** (e.g., heart rate, BUN, oxygen saturation) and **patient-specific explanations**.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



6.3 Findings

1. **Temporal modelling is critical:** LSTM effectively captures sequential patterns in ICU time-series data, improving early admission prediction.
2. **Privacy preservation works:** Federated learning allows collaborative model training without exchanging raw patient data, maintaining regulatory compliance.
3. **Explainability enhances trust:** XAI techniques provide interpretable predictions, supporting clinician decision-making.
4. **Performance vs. privacy trade-off:** Slight reduction in AUROC (~0.01) in federated LSTM compared to centralized LSTM is acceptable given privacy benefits.

6.4 Advantages of Proposed Framework

1. **Data Privacy:** No raw data is shared across hospitals, complying with HIPAA/GDPR.
2. **Temporal Accuracy:** LSTM captures long-term dependencies in patient physiological data.
3. **Interpretability:** SHAP and LIME provide actionable insights for clinicians.
4. **Scalability:** Framework can be extended to multiple hospitals without centralizing sensitive data.
5. **Ethical AI:** Enhances clinician trust and supports transparent decision-making.

6.5 Disadvantages

1. **Communication Overhead:** Federated learning requires transmitting model weights multiple times, increasing network load.
2. **Computational Resources:** Local training of LSTM models may require GPUs or high-performance servers.
3. **Data Heterogeneity:** Non-IID distributions across hospitals can affect model convergence and stability.
4. **Limited Real-World Deployment:** Framework has been tested on public datasets; further validation on live hospital systems is needed.

VII. CONCLUSION

This study presents an **explainable and privacy-preserving deep learning framework** for early ICU admission prediction using multivariate clinical time-series data. By combining **LSTM-based temporal modelling**, **federated learning**, and **explainable AI techniques (SHAP and LIME)**, the framework achieves high predictive performance while safeguarding patient privacy and providing both global and patient-specific interpretability. Experimental results on publicly available ICU datasets show that the proposed method outperforms conventional machine learning approaches and achieves comparable accuracy to centralized deep learning models, with the added benefits of decentralized training and transparent decision support. These findings highlight that integrating temporal dynamics, privacy-preserving collaborative learning, and interpretable outputs can **enhance clinician trust and support timely ICU admission decisions**. Future work will focus on **real-time multi-centre deployment, integration of additional clinical modalities, and advanced privacy-preserving techniques** such as differential privacy and homomorphic encryption to further strengthen security. Overall, the proposed framework demonstrates the feasibility of developing **ethical, interpretable, and high-performance AI systems** for critical healthcare decision-making.

Key Contributions:

1. High predictive performance while preserving patient privacy
2. Transparent and interpretable predictions to support clinical decision-making
3. Scalable and ethical AI framework suitable for multi-hospital deployment

Future Work:

1. Extend to **larger multi-centre federated setups**
2. Integrate **additional modalities** such as imaging or genetic data
3. Optimize for **real-time ICU monitoring and alerting**
4. Explore **differential privacy and homomorphic encryption** for stronger security guarantees

Overall, the proposed framework demonstrates that **privacy, interpretability, and predictive accuracy can be effectively combined** in critical healthcare applications, paving the way for trustworthy AI-driven ICU decision support systems.

REFERENCES

1. Y. Deng, S. Liu, Z. Wang, Y. Wang, Y. Jiang, and B. Liu, “Explainable time-series deep learning models for the prediction of mortality, prolonged length of stay and 30-day readmission in intensive care patients,” *Front. Med.*, vol. 9, p. 933037, Sep. 2022, doi:10.3389/fmed.2022.933037.
2. M. Mesinovic, P. Watkinson, and T. Zhu, “Explainable machine learning for predicting ICU mortality in myocardial infarction patients using pseudo-dynamic data,” *Sci. Rep.*, vol. 15, p. 27887, Jul. 2025, doi:10.1038/s41598-025-13299-3.
3. M. Z. Fuzail, I. Ud Din, S. Ahmed, A. Alhumam, and A. H. Khan, “Optimizing sepsis mortality prediction using hybrid federated learning and explainable AI framework,” *Sci. Rep.*, 2026.
4. Y. Benhamou, S. Kalyan, and S. Kumar, “Federated learning for ICU mortality prediction: balancing accuracy and privacy in a multi-hospital setting,” *bioRxiv*, Aug. 2025.
5. W. Fathy, G. Emeriaud, and F. Cheriet, “A comprehensive review of ICU readmission prediction models: from statistical methods to deep learning approaches,” *Artif. Intell. Med.*, vol. 165, 2025.
6. T. Tunduny and B. Shibwabo, “Explainable AI approaches in federated learning: systematic review,” *JMIR AI*, vol. 1, p. e69985, Feb. 2026.
7. Y. Ive et al., “AI assisted prediction of unplanned intensive care admissions with SHAP and LIME explanations,” *Nat. Digit. Med.*, 2025.
8. C. Stylianides, “AI advances in ICU with an emphasis on sepsis prediction,” *MDPI* 2504-4990, vol. 7, 2025.
9. X. Li, J. Gu, Z. Wang, Y. Yuan, B. Du, and F. He, “XAI for in-hospital mortality prediction via multimodal ICU data,” AAAI, 2023.
10. S. Chen et al., “Interpretable machine learning model for early prediction of 30-day mortality in ICU patients,” *arXiv*, Jun. 2025.

BIOGRPAHY

Shamsiya Parveen
Assistant Professor,
Department of Computer Science and Engineering (Data Science)
Sri Siddhartha Institute of Technology (SSIT), Maralur, Tumkur, India
Email: shamsiyaparveen@ssit.edu.in

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING