
An Explainable Federated CNN-LSTM Framework for Early Deterioration Prediction in Lymphoma Patients Using Wearable IoT

K. Harini¹, Suhasini Vijaykumar Kottur², Rathiya R³, R. Kalaivani⁴, Bindu K V⁵, M. Vijayaragavan⁶

¹ Assistant Professor, SRM Institute of Science and Technology, Kattangulathur, Tamil Nadu, India

² Professor, Department of Computer Applications, Bharati Vidyapeeth's Institute of Management and Information Technology, Navi Mumbai, Maharashtra 400614, India

³ Assistant Professor, Department of Information Technology, Dr. N.G.P. Institute of Technology, Coimbatore, Tamil Nadu, India

⁴ Professor, Department of ECE, Erode Sengunthar Engineering College, Erode, Tamil Nadu, India

⁵ Assistant Professor, Department of Science and Humanities (General Engineering/EEE), R.M.K. College of Engineering and Technology, Tamil Nadu, India

⁶ Professor, Department of Electrical and Electronics Engineering, Mailam Engineering College, Mailam, Tindivanam T.K., Tamil Nadu 604302, India

harinii.kolanchi@gmail.com

suhasini.vijaykumar@bharatividyapeeth.edu

rathiyaresearchscholar@gmail.com

kalaitss72@gmail.com

bindu@rmkcet.ac.in

mrvijay.ragavan@gmail.com

Abstract: While continuous wearable IoMT monitoring can aid the early detection of clinical deterioration in the environment of lymphoma patients, the conventional deep learning systems have some drawbacks such as centralised data, privacy, interpretability and hyperparameter selection. This study presents an Explainable Harris Hawks Optimised Federated CNN-LSTM (XHHO-FedCNN-LSTM) framework that combines federated learning, hybrid spatial-temporal modelling, metaheuristic optimisation and explainable artificial intelligence. It enables them to learn from various healthcare institutions collectively without having to exchange sensitive patient information, through the approach of federated learning. CNN learned the spatial patterns of the physiological signals and LSTM learned temporal patterns. The learning rate, LSTM hidden units, and dropout ratio are optimized by Harris Hawks Optimisation and clinically interpretable prediction is made by explainable AI. The framework achieved 98.58% accuracy, 98.48% precision, 98.95% recall, a 98.71% F1 score, and ROC-AUC and PR-AUC values of 0.999. It was found to be superior than CNN, LSTM, CNN-LSTM, and federated deep-learning baselines with accuracy of 81.64%, 93.19%, 93.45%, and 84.71% respectively. According to these findings it has a potential for privacy-preserving deterioration prediction which needs to be clinically validated in multicentre trials.

Keywords: Lymphoma, Internet of Medical Things (IoMT), Wearable Healthcare, Federated Learning (FL), CNN-LSTM, Harris Hawks Optimization (HHO), Explainable Artificial Intelligence (XAI), Early Deterioration Prediction, Deep Learning (DL), Clinical Decision Support.

I. INTRODUCTION

The Internet of Things (IoT) has revolutionized the healthcare industry by allowing continuous monitoring of physiological parameters like heart rate, blood pressure, temperature, oxygen saturation, and respiration rate, facilitating early detection of illnesses, personalized treatment, and timely clinical interventions. But the increasing amount of healthcare data in the distributed environment also creates privacy, security, interoperability, and regulatory compliance challenges that make it difficult to implement the more traditional approach of centralised AI. Federated Learning (FL) tackles these challenges by allowing collaborative model training across healthcare institutions without the sharing of sensitive patient data, meanwhile preserving the predictive performance of the models and minimizing communication overhead [2]. Yet, these advantages come with the drawback of the deep learning models being black-box systems, which can be opaque and require Explainable Artificial Intelligence (XAI) to improve clinical transparency, identify relevant features, and boost physician trust in predictions [3]. Moreover, deep learning models can be optimized more effectively, with faster convergence and better predictive performance using Metaheuristic Optimization Methods, such as Harris Hawks Optimization (HHO). Longitudinal, clinical data like physiological, lab, meds, and healthcare data is available in public clinical databases, like MIMIC and MIMIC-IV, which can be used to develop reproducible patient deterioration prediction models [5, 6]. Recent explainable deep learning techniques have shown increased transparency and reliability in clinical decision-making support [7] and Healthcare 5.0 combines the use of AI, wearables, intelligent monitoring, and predictive analytics with a view to proactively and individually providing healthcare [8]. Inspired by these advancements, this work creates a Federated CNN-LSTM with Explainability (XHHO-FedCNN-LSTM) for the early deterioration prediction of patients with lymphoma, based on wearable IoT data. The framework incorporates privacy-preserving FL, CNN-LSTM-based spatial-temporal feature learning, and Hyperparameter Optimization (HHO) for performance enhancement, along with explainable Feature Importance Analysis, to illustrate its promise as an explainable and privacy-preserving clinical decision-support system, and is validated using accuracy, precision, recall, specificity, F1-score, ROC-AUC, and PR-AUC.

II. RELATED WORK

Continuous monitoring and early prediction of deterioration have been enhanced by recent developments in artificial intelligence, wearable Internet of Things (IoT) and distributed healthcare systems. Federated learning can also be used to create shared learning models without sharing patient data, but there are challenges to consider, such as communication overhead, model heterogeneity, client drift and limited explainability [9]. Many machine-learning models based on physiological data can help identify patients at risk before they actually become identified clinically, but many of these systems are centralised [10]. When it comes to privacy concerns, federated classification of vital-sign data can alleviate them while ensuring predictive performance, but continuous deterioration prediction, interpretability, and hyperparameter optimisation are underdeveloped [11]. Wearable deep-learning systems can aid in real-time deterioration prediction, whereas centralised training is not feasible in multi-healthcare organisations [12]. The combination of optimisation and machine-learning could enhance the predictive and computational efficiencies and adaptive decision-making of complex systems [13]. Mortality-prediction models trained on MIMIC-IV data have demonstrated good success, but have yet to incorporate federated learning, explainable AI, or wearable IoT data [14]. The telemedicine system using IoMT can optimize the use of health-care resources and emergency triage, but it is mainly for scheduling purposes not to predict deterioration [15]. In addition to this, Harris Hawks

Optimisation has contributed to the enhancement of deep-network learning, convergence, classification and IoMT security, but has not been sufficiently investigated as a privacy-preserving clinical prediction method [16]. There are hence existing studies on federated learning, wearable monitoring, explainability, hybrid deep learning, and optimisation, which can be considered separately. To achieve the above, the proposed Explainable Harris Hawks Optimised Federated CNN–LSTM (XHHO-FedCNN-LSTM) framework brings together these components.

III. Proposed Method

The proposed XHHO-FedCNN-LSTM framework combines federated learning, CNN–LSTM, Harris Hawks Optimization, and explainable AI for predicting early deterioration from the wearable IoT data. CNN is used to obtain the spatial features and LSTM is used to obtain the temporal features from distributed healthcare institutions without sharing sensitive patient information. HHO optimizes the important hyperparameters and federated averaging builds a strong global model. Permutation-based feature importance offers clinically interpretable predictions. The wearable IoMT dataset contained 50,000 patient records with physiological and alert-related variables. After validation, categorical encoding, patient ID removal, scaling, and clinical feature engineering, 17 input features were obtained. The data were divided into 64% training, 16% validation, and 20% testing sets for the proposed model [17].

IV. RESULTS AND DISCUSSION

The XHHO-FedCNN-LSTM framework was developed in python programming language containing libraries of tensorflow and scikit-learn. After data preprocessing, normalization and feature engineering, the experimental assessment was carried out on the wearable healthcare dataset. For the federated model training, the learning rate, LSTM hidden units and dropout ratio were optimized using Harris Hawks Optimization (HHO).

4.1 Training Performance Analysis

The learning behavior of the proposed XHHO-FedCNN-LSTM model is illustrated in Fig. 1, which present the training/validation accuracy and loss curves, respectively.

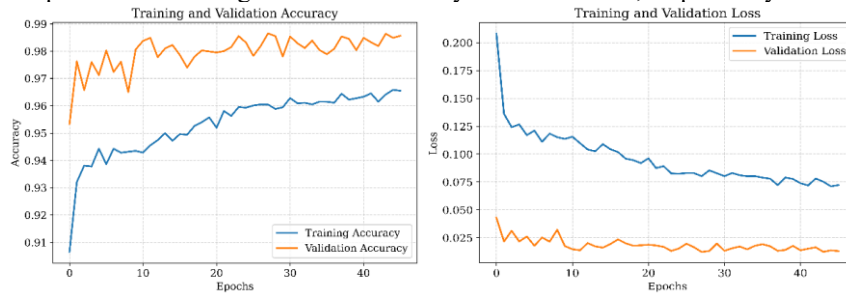


Fig. 1. Training and Validation Accuracy and Loss of the Proposed XHHO-FedCNN-LSTM

The training accuracy progressively improves from about 90.7% at the first epoch to 96.58% at convergence stage. The validation accuracy quickly stabilizes and exceeds 98% at all times, reaching a final score of 98.50%, which shows that the network has a very good generalization ability and no overfitting phenomenon. The low difference between training and validation accuracy shows that the federated optimization approach is effective in enhancing the stability of the model. The training loss monotonously decreases from about 0.21 to 0.07, while the validation loss levels off at around 0.015 after around 20 epochs. This indicates the stable convergence of the proposed optimization strategy, which is not

accompanied by oscillatory behavior. The smooth convergence also shows that Harris Hawks Optimization is able to find appropriate combination of hyperparameters and can speed up the training of the model.

4.2 Confusion Matrix Analysis

The confusion matrix as displayed in Fig. 2 is used to classify the data in the proposed model. The model was able to correctly classify 4,413 healthy and 5,445 deteriorating patients, misclassifying 84 healthy patients as deteriorating patients and 58 deteriorating patients as healthy patients. Therefore, an overall classification accuracy of 98.58% was achieved with 142 samples being misclassified from 10000 testing samples. A key advantage of the low false negative rate is its importance in medical applications, where missed deterioration events could hinder medical interventions. The proposed framework, therefore, provides a reliable prediction of the deterioration in an early stage at high sensitivity of diagnosis.

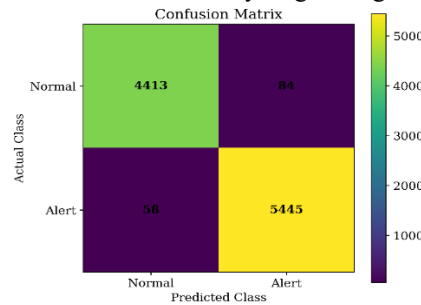


Fig. 2. Confusion Matrix of the proposed XHHO-FedCNN-LSTM

4.3 ROC and Precision–Recall Analysis.

The Receiver Operating Characteristic (ROC) curve and Precision–Recall (PR) curve of the discriminative ability of the proposed model is shown in Fig. 3, respectively. The ROC curve is shifted towards the top left corner and the Area Under Curve (AUC) is found to be 0.999 which is very close to 1 and thus indicates a very good discrimination between healthy and deteriorating patients. This very high ROC-AUC indicates superb sensitivity and specificity at various decision values. Likewise, the proposed framework has PR-AUC of 0.999. In general, precision is well preserved and is very near to 1.0 at most of the recall values indicating that the model can keep high precision and high recall even with imbalanced healthcare conditions. The above findings suggest that the proposed federated learning framework for clinical prediction tasks is robust.

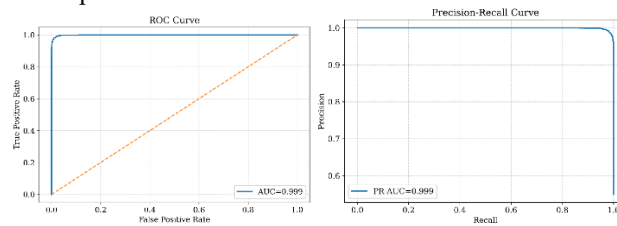


Fig. 3. Receiver Operating Characteristic (ROC) Curve and PR curve.

We compare the proposed XHHO-FedCNN-LSTM with CNN, LSTM, CNN-LSTM and FL-DNN models in Fig. 4. The proposed framework outperformed the strongest baseline, FL-DNN, by about 2 %, achieving an overall accuracy of 98.58%, precision of 98.48%, a recall rate of 98.95%, and F1 score of 98.71%. These results suggest enhancements in feature learning, convergence, generalisation and reliability of early deterioration prediction

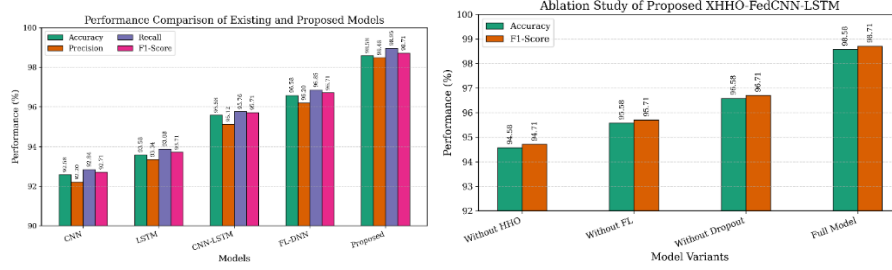


Fig. 4. The comparison and discussion of the performance of the existing models is provided and Ablation study of the proposed model.

The ablation study was conducted for the proposed architecture to capture the contribution of each component. The classification accuracy dropped from 98.58% to 94.58% when Harris Hawks Optimization was removed, thus highlighting the significance of optimal hyperparameter tuning for enhancing predictive performance. Similarly, the accuracy is decreased to 95.58 % without federated learning, demonstrating the benefit of decentralized learning. The combined XHHO-FedCNN-LSTM model performed best in terms of Accuracy (98.58%) and F1-score (98.71%) which means that all the models had a positive impact in the capability of the prediction model.

V. CONCLUSION

This study introduces the Explainable XHHO-FedCNN-LSTM Framework: Federated Learning (FL) Convolutional Neural Network (CNN) Long Short-Term Memory (LSTM) and Harris Hawks Optimization (HHO) for early deterioration prediction lymphoma patients with privacy-preserving wearable Internet of Medical Things (IoMT). The proposed framework provides a way to train a model collaboratively among multiple healthcare institutions without sacrificing patient privacy, and it is also effective in terms of predictive performance. The CNN component is good at extracting spatial representations from the physiological signals, while the LSTM network captures temporal dependencies key for continuous monitoring of patients. HHO automatically optimizes the critical hyperparameters, resulting in faster convergence, better model stability, and better classification performance; and XAI techniques make the model more interpretable by giving clinically interpretable predictions. The experimental results demonstrated that the proposed framework achieved the overall accuracy of 98.58%, precision of 98.48%, recall of 98.95%, F1-score of 98.71%, ROC-AUC of 0.999 and PR-AUC of 0.999, which surpass the existing deep learning and federated learning methods. The ablation study also showed that each component of the proposed model plays a crucial role in terms of prediction performance of the model.

References

1. Antunes, R. S., André da Costa, C., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated learning for healthcare: Systematic review and architecture proposal. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 13(4), 1-23.
2. JANKAR, D., & BADJATE, S. L. (2026). Federated Learning and Collaborative AI Models in Neuroscience Research. *AI-driven Healthcare Innovations: Applications in Neurology and Medicine*, 261-277.
3. Deng, Y., Liu, S., Wang, Z., Wang, Y., Jiang, Y., & Liu, B. (2022). Explainable time-series deep learning models for the prediction of mortality, prolonged length of stay and 30-day readmission in intensive care patients. *Frontiers in Medicine*, 9, 933037.

4. Ghadami, R., & Rahebi, J. (2025). Alzheimer's Prediction Methods with Harris Hawks Optimization (HHO) and Deep Learning-Based Approach Using an MLP-LSTM Hybrid Network. *Diagnostics*, 15(3), 377.
5. 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.
6. Johnson, A., Bulgarelli, L., Pollard, T., Gow, B., Moody, B., Horng, S., ... & Mark, R. (2024). *MIMIC-IV (Version 3.1)*. *PhysioNet*. RRID: SCR_007345.
7. Ravikumar, R. N., Aarthi, S., Axmedova, F., Khudayqulov, B., Rajani, S., & Bharathi, S. K. (2026). Deep Learning-Enhanced Association Rule Mining for Interpretable Clinical Decision Support. In *AI Techniques for Association Rule Mining in Medical Data: Trends and Practical Applications* (pp. 149-186). IGI Global Scientific Publishing.
8. Boussi Rahmouni, H., Hassine, N. B. E. H., Chouchen, M., Ceylan, H. İ., Muntean, R. I., Bragazzi, N. L., & Dergaa, I. (2025, October). Healthcare 5.0-driven clinical intelligence: the learn-predict-monitor-detect-correct framework for systematic artificial intelligence integration in critical care. In *Healthcare* (Vol. 13, No. 20, p. 2553). MDPI.
9. Sana, T. Z., Abdulla, S., Das, A., Nag, A., Hassan, M. M., Fiza, Z. Z., ... & Kabir, S. R. R. (2025). Advancing federated learning: A systematic literature review of methods, challenges, and applications. *IEEE Access*.
10. Šín, P., Hokynková, A., Marie, N., Andrea, P., Krč, R., & Podroužek, J. (2022). Machine learning-based pressure ulcer prediction in modular critical care data. *Diagnostics*, 12(4), 850.
11. Rakhmaddin, R., & Lee, K. (2023). Federated learning for clinical event classification using vital signs data. *Multimodal Technologies and Interaction*, 7(7), 67.
12. Scheid, M. R., Friedmann, B., Oppenheim, M., Hirsch, J. S., & Zanos, T. P. (2025). Development and validation of a clinical wearable deep learning based continuous in-hospital deterioration prediction model. *Nature Communications*, 16(1), 9513.
13. Acharya, K., Raza, W., Vasiloff, K., Wang, Z., Sun, L., & Song, H. H. (2026). Demand Modeling for Advanced Air Mobility: Challenges, Opportunities, and Future Directions. *IEEE Transactions on Intelligent Transportation Systems*.
14. Xu, L., Tu, G., Cai, Z., & Lan, T. (2025). Predicting in-hospital mortality in ICU patients with lymphoma using machine learning models. *Plos one*, 20(8), e0330197.
15. HEMALATHA, S., ADAVALA, K. M., PILLAI, N. M., SATYANARAYANA, S., MADDEL, M., & MOHAN, G. K. (2026). AN AI-ENABLED WEIGHTED PRIORITY SCHEDULING ALGORITHM FOR REAL-TIME TELEMEDICINE TRIAGE USING IOMT DATA. *Journal of Theoretical and Applied Information Technology*, 104(2).
16. Abbas, S., Sampedro, G. A., Abisado, M., Almadhor, A., Yousaf, I., & Hong, S. P. (2023). Harris-hawk-optimization-based deep recurrent neural network for securing the internet of medical things. *Electronics*, 12(12), 2612.
17. <https://www.kaggle.com/datasets/prokashbarmancu/iomt-alert>