Attention-Based Explainable Transformer Model for Blood Glucose Prediction in ICU Patients

[1]Samreen Fatima, [2]Subramanian K.M, [3]Mohammed Abbas Qureshi
[1]PG Scholar, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086.
[2]Professor, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086.
[3]Associate Professor, Department of Computer Science and Engineering, Shadan College of Engineering andTechnology, Hyderabad, Telangana, India – 500086.
[1]fsamreen456@gmail.com, [2]kmsubbu.phd@gmail.com, [3]mohammedabbas09@yahoo.com

Abstract: Blood glucose instability in Intensive Care Unit (ICU) patients is a critical healthcare challenge associated with severe complications and increased mortality risk. Accurate glucose forecasting is difficult due to the irregular, heterogeneous, and noisy nature of ICU electronic health record (EHR) data, which includes multiple physiological and clinical variables. This work presents an Explainable Hierarchical Transformer-based framework for predicting future blood glucose levels and classifying glycemic states, including hypoglycemia, euglycemia, and hyperglycemia. The proposed framework integrates multi-source ICU information such as glucose measurements, vital signs, laboratory parameters, and medication records through an attention-based feature fusion mechanism. The hierarchical Transformer architecture captures complex temporal dependencies and dynamically identifies significant clinical patterns influencing glucose variations. To improve model transparency, an explainable artificial intelligence approach using SHAP (SHapley Additive exPlanations) is incorporated to identify the key factors contributing to prediction outcomes. The framework is evaluated using the eICU Collaborative Research Database and compared with existing deep learning approaches, including LSTM, GRU, and WaveNet models. The proposed approach enables early detection of glucose-related risks while providing clinically interpretable insights, supporting timely intervention and improved decision-making in critical care environments.

Keywords— *ICU glucose prediction, hierarchical Transformer, explainable artificial intelligence, SHAP, attention fusion, clinical time-series analysis, electronic health records, glycemic risk classification.*

1. INTRODUCTION

Blood glucose regulation is a critical factor in the management of critically ill patients admitted to Intensive Care Units (ICUs), where rapid fluctuations in glucose levels can contribute to severe complications, prolonged hospitalization, and increased mortality risk. Conventional glucose monitoring methods mainly depend on intermittent measurements and clinical observation, which may not provide sufficient time for early detection of dangerous glycemic events. The availability of Electronic Health Records (EHRs) and advanced healthcare analytics has created new opportunities for developing predictive models capable of identifying future glucose variations using complex patient information. Recent studies have demonstrated that multimodal clinical data fusion can improve predictive healthcare performance by combining heterogeneous physiological and medical information. [5]

ICU environments generate large-scale, irregular, and time-dependent clinical data, including vital signs, laboratory measurements, medication administration, and treatment-related information. Handling these heterogeneous data sources remains challenging due to missing values, inconsistent sampling intervals, and complex interactions among physiological variables. Transformer-based deep learning approaches have shown strong capability in learning long-range dependencies from sequential healthcare data and improving clinical prediction tasks. [7] Attention-based models further enhance this capability by identifying important features and temporal patterns that contribute to patient outcomes. [4]

Blood glucose prediction requires models that can capture both short-term fluctuations and long-term physiological trends. Hierarchical learning strategies have been explored to effectively represent irregular clinical time-series data by learning multiple levels of temporal dependencies. [1] Multimodal attention frameworks have also demonstrated potential in personalized glucose forecasting by integrating diverse patient information and improving prediction reliability. [2] However, many existing deep learning models provide limited transparency, creating challenges for clinical acceptance because healthcare professionals require understandable reasoning behind automated predictions.

Explainable Artificial Intelligence (XAI) techniques provide a solution by identifying influential clinical features and improving the interpretability of predictive models. Explainable healthcare frameworks have been increasingly applied to improve trust, transparency, and decision support in medical environments. [6] The integration of explainability with Transformer architectures enables clinicians to understand how factors such as glucose history, medication effects, and physiological changes influence prediction outcomes. Multimodal Transformer-based approaches have also demonstrated effectiveness in analyzing complex ICU patient conditions by combining multiple clinical data sources. [3]

This work focuses on developing an intelligent blood glucose prediction framework for ICU patients by utilizing a Hierarchical Transformer architecture with attention fusion and explainable AI capabilities. The framework aims to forecast future glucose values, classify glycemic states, and provide interpretable insights from clinical data. The major contributions include: (1) designing a hierarchical Transformer-based prediction approach for ICU glucose forecasting, (2) integrating attention-based fusion for combining heterogeneous clinical variables, (3) enabling both numerical glucose prediction and glycemic risk classification, and (4) incorporating explainable AI techniques to highlight important clinical factors influencing predictions. The developed approach supports improved understanding of glucose dynamics and provides a foundation for future intelligent ICU decision-support applications.

2. RELATED WORK

Recent advancements in artificial intelligence and deep learning have significantly improved predictive healthcare analytics by enabling the analysis of heterogeneous and time-dependent clinical information. The growing availability of

Electronic Health Records (EHRs), physiological monitoring systems, and multimodal healthcare data has encouraged the development of intelligent frameworks capable of supporting diagnosis, prognosis, and patient monitoring tasks in critical care environments.

Guo proposed DT-ICU, an explainable digital twin framework designed for ICU patient monitoring through multimodal and multi-task iterative inference techniques. The study demonstrated that integrating multiple clinical modalities can improve representation learning and support more adaptive patient monitoring. The explainable architecture also highlighted the importance of transparency in healthcare prediction systems. However, the framework focused on generalized ICU monitoring rather than forecasting blood glucose dynamics and did not specifically address glycemic risk prediction. **[8]**

Cifci et al. introduced an Interpretable Adaptive Graph Fusion Network for predicting mortality and complications in ICU patients. Their model combined graph-based learning with explainability mechanisms to capture complex interactions among clinical variables. The results showed improved predictive capability compared with conventional approaches; however, graph fusion methods primarily focused on patient outcome prediction and provided limited support for continuous physiological forecasting such as glucose level estimation. **[9]**

Sunkara developed Gamt-Care, a Graph Attention Multimodal Transformer framework that integrated attention mechanisms with multimodal healthcare analytics. The study emphasized that attention-driven feature aggregation can improve learning from heterogeneous healthcare sources and enhance predictive performance. Despite its advantages, the framework remained oriented toward broad predictive healthcare applications and lacked mechanisms specifically designed for irregular temporal glucose patterns observed in ICU environments. **[10]**

Khullar et al. proposed a Temporal Fusion Transformer-based framework for early sepsis prediction using multivariate vital signs and laboratory time-series data collected from ICU settings. Their findings demonstrated that Transformer architectures effectively capture temporal dependencies and support early risk identification using complex clinical sequences. Nevertheless, the study concentrated on sepsis prediction and did not explore blood glucose forecasting or explainable prediction mechanisms for glycemic monitoring. **[11]**

Xie et al. investigated a dynamic mask attention graph neural network for mechanical ventilation prediction in elderly ICU patients. Their approach highlighted the effectiveness of attention mechanisms for handling dynamic and incomplete healthcare data while improving clinical decision support. Although the model addressed irregular temporal information, its application remained restricted to ventilation-related outcomes rather than continuous glucose prediction tasks. **[12]**

Tao et al. introduced DHAN, a Deep Hierarchical Attention Network for multisource biomedical data fusion in precision healthcare applications. The study demonstrated that hierarchical attention structures can effectively integrate multiple information sources and capture complex relationships across biomedical variables. However, the framework focused on diagnostic analysis rather than real-time physiological forecasting and did not consider ICU-specific glucose fluctuations. **[13]**

Al-Bairmani et al. developed a hybrid CNN–Bi-LSTM–Transformer architecture integrated with SHAP explanations for interpretable clinical prediction. Their work illustrated that combining deep sequential learning with explainability methods can improve both predictive performance and transparency. The incorporation of SHAP enabled identification of influential factors contributing to model outcomes. However, the study targeted arrhythmia detection and did not evaluate irregular multimodal ICU data for glucose forecasting. **[14]**

Khan et al. proposed an interpretable hybrid machine learning framework for diabetes prediction using electronic health records. Their findings reinforced the importance of explainable prediction in healthcare applications and demonstrated that transparent decision-making can increase trust in AI-assisted systems. However, the framework focused on disease-level prediction and lacked temporal modeling for dynamic glucose estimation in critically ill patients. **[15]**

Overall, existing studies indicate that multimodal learning, Transformer architectures, graph-based modeling, attention mechanisms, and explainable AI significantly improve predictive healthcare performance. Nevertheless, current approaches still exhibit limitations in jointly handling irregular ICU time-series data, integrating heterogeneous clinical variables, forecasting future glucose variations, and providing clinically interpretable predictions. These research gaps motivate the development of an attention-based explainable hierarchical Transformer framework capable of combining multimodal ICU information, predicting future blood glucose levels, classifying glycemic states, and generating transparent explanations to support intelligent critical care decision-making.

3. MATERIALS AND METHODS

The proposed framework aims to develop an intelligent and interpretable blood glucose prediction system for ICU patients by integrating heterogeneous clinical information with advanced deep learning techniques. The system utilizes Electronic Health Record (EHR) data, including glucose measurements, vital signs, laboratory results, medication records, and patient-related information to learn complex physiological patterns. A Hierarchical Transformer architecture with Attention Fusion is introduced to capture both short-term variations and long-term dependencies in ICU time-series data. Transformer-based models have shown strong potential in learning complex healthcare sequences and improving predictive performance in clinical applications [15]. The framework applies data preprocessing techniques, feature extraction, normalization, and temporal sequence generation to prepare irregular ICU data for model training. The core prediction module uses a Hierarchical Transformer combined with multi-head self-attention mechanisms for glucose forecasting and glycemic state classification. Baseline algorithms such as LSTM, GRU, and WaveNet are considered for performance comparison. The model integrates SHAP-based Explainable AI techniques to identify important clinical features influencing prediction outcomes, improving transparency and clinician trust. Explainable predictive healthcare approaches have demonstrated the importance of interpretable decision-making in medical environments [21]. The system generates multi-horizon glucose predictions and classifies patients into hypoglycemia, euglycemia, and hyperglycemia categories. The developed framework provides an interactive visualization interface for monitoring glucose trends, risk alerts, and explanation outputs. This approach aims to support proactive ICU decision-making by reducing delayed interventions and improving patient safety through accurate and explainable glucose forecasting. [5]

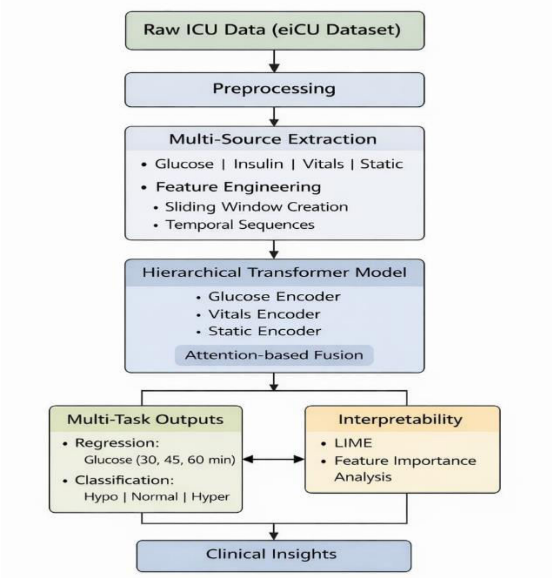

Fig. 1. Concise Workflow of the Hierarchical Transformer Model for Multi-Source Blood Glucose Prediction in ICU Patients.

Fig.1 Proposed Methodology Workflow

Fig. 1 illustrates the proposed sequential framework for multimodal ICU patient data processing and glycemic forecasting. The pipeline initiates with Data Acquisition of multi-source ICU records, followed by Preprocessing & Normalization to handle missing data and normalize scores. Next, Sequence Construction applies sliding windows to generate temporal sequences, which are categorized into distinct physiological streams during Feature Grouping. These streams undergo Transformer Encoding and Attention Fusion for multimodal feature synthesis. Finally, the system executes multi-horizon Forecasting & Classification, followed by Explainability Generation using LIME and attention visualizations.

i) Dataset:

The framework utilizes the eICU Collaborative Research Database (eICU-CRD) as the primary data source for ICU patient glucose prediction. The dataset contains large-scale clinical information collected from critically ill patients, including glucose measurements, vital signs, laboratory test results, medication administration records, and patient demographic details. [19] The selected ICU patient episodes are extracted based on the availability of glucose-related observations and associated physiological variables. The collected data represents real-world ICU conditions where measurements are irregular, incomplete, and recorded at different time intervals. The dataset is organized into patient-specific temporal sequences to support deep learning-based forecasting. Multiple clinical parameters are considered to provide a comprehensive representation of patient health status and glucose regulation patterns.

ii) Preprocessing:

The preprocessing stage converts raw ICU data into a structured format suitable for model training. Initially, data cleaning is performed to remove duplicate records, invalid values, and physiologically impossible measurements. Missing clinical values are handled using appropriate imputation strategies to maintain continuity in time-series information. [9] Since ICU data is collected at irregular intervals, temporal alignment is applied to convert observations into uniform time windows. Feature engineering is performed to extract meaningful clinical indicators such as glucose trends, glucose variation rate, insulin-related effects, and interactions between vital signs and glucose levels. Numerical features are normalized to ensure consistent model learning. The processed data is transformed into fixed-length sequences containing historical clinical information required for future glucose prediction.

iii) Train and Test:

The processed dataset is divided into training and testing subsets to evaluate model performance effectively. The training data is used to learn temporal relationships, clinical feature dependencies, and glucose variation patterns, while the testing data is used for unbiased performance evaluation. [22] The dataset is organized using patient-level separation to prevent data leakage between training and testing samples. The model is trained using historical ICU observations to predict future glucose values at different forecasting horizons, including 30, 45, and 60 minutes ahead. Performance evaluation is carried out using regression metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 score for glucose prediction. Classification performance for glycemic risk categories is measured using accuracy, precision, recall, and F1-score.

iv) Algorithms:

Hierarchical Transformer with Attention Fusion: The main prediction algorithm is a Hierarchical Transformer model designed to capture complex temporal patterns from multimodal ICU clinical data. The model contains multiple attention layers that learn relationships between glucose history, physiological signals, laboratory values, and medication information. Feature-level attention identifies the most important clinical variables, while temporal attention focuses on significant events occurring across previous time intervals. The attention fusion mechanism combines these representations into a unified feature space for prediction. [24] The model generates both continuous glucose forecasts and categorical glycemic risk classifications, enabling early identification of hypoglycemia, euglycemia, and hyperglycemia conditions.

Long Short-Term Memory (LSTM): LSTM is implemented as a baseline deep learning model for comparison with the proposed Transformer architecture. LSTM uses memory cells and gating mechanisms to learn sequential dependencies from ICU time-series data. The forget gate, input gate, and output gate control the flow of information and allow the model to retain important historical glucose patterns. The model processes previous patient observations and learns relationships between past physiological conditions and future glucose changes. LSTM performance is evaluated using the same dataset and metrics to analyze its effectiveness compared with attention-based learning methods.

Gated Recurrent Unit (GRU): GRU is another recurrent neural network-based algorithm used as a comparative baseline. It simplifies the LSTM architecture by using update and reset gates to control information flow while maintaining the ability to capture temporal dependencies. GRU requires fewer parameters and can provide efficient learning for sequential healthcare data. In this methodology, GRU is trained using ICU clinical sequences to predict future glucose values and classify glycemic states. The obtained results are compared with the Hierarchical Transformer model to determine the advantages of advanced attention-based approaches for irregular ICU data analysis.

WaveNet: WaveNet is applied as a deep learning baseline for modeling temporal patterns in glucose-related time-series data. It uses dilated convolutional layers to capture short-term and long-term dependencies without relying on recurrent processing. The model can analyze complex sequential patterns by expanding the receptive field through stacked convolution operations. WaveNet is trained using the same ICU dataset to evaluate its capability in glucose forecasting tasks. The comparison helps determine whether Transformer-based attention mechanisms provide improved performance over convolution-based temporal modeling approaches.

v) Integration of XAI and Streamlit:

Explainable Artificial Intelligence (XAI) is integrated into the framework to improve transparency and clinical understanding of model predictions. [25] SHAP (SHapley Additive exPlanations) is applied to identify the contribution of individual clinical features toward glucose prediction outcomes. Attention visualization techniques are also used to highlight important time periods and physiological signals influencing the model decision. The explanation outputs provide clinicians with interpretable information regarding factors such as glucose history, insulin administration, and vital sign variations. A Streamlit-based interactive dashboard is developed for deployment and visualization. The interface allows users to select patient records, view predicted glucose trends, observe glycemic risk classifications, and analyze explanation plots, enabling effective clinical decision support.

4. RESULTS AND DISCUSSIONS

The performance of the developed framework was evaluated using the eICU clinical dataset for ICU blood glucose forecasting and glycemic state classification. The proposed Hierarchical Transformer with Attention Fusion was compared with baseline deep learning models, including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and WaveNet. The evaluation considered both regression-based glucose prediction performance and classification-based glycemic risk identification. The experimental analysis demonstrates that the attention-based hierarchical architecture effectively captures complex temporal relationships among glucose measurements, vital signs, laboratory parameters, and medication-related information.

The proposed model achieved improved prediction performance across different forecasting horizons, including 30, 45, and 60 minutes ahead. The attention fusion mechanism enabled the model to focus on the most relevant clinical features and time intervals influencing glucose variations. Compared with traditional recurrent architectures, the Transformer-based model showed better capability in handling irregular ICU time-series data due to its ability to learn long-range dependencies without sequential limitations.

For continuous glucose prediction, the model was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). The proposed framework obtained lower prediction error and higher R^2 score compared with baseline approaches, indicating better accuracy in forecasting future glucose trends. The classification performance for hypoglycemia, euglycemia, and hyperglycemia detection was assessed using accuracy, precision, recall, and F1-score. The obtained results indicate that the framework can effectively identify critical glucose risk conditions, supporting early clinical intervention.

The integration of Explainable Artificial Intelligence using SHAP improved the interpretability of prediction outcomes by identifying influential clinical variables. Features such as previous glucose trends, insulin administration, heart rate, and laboratory measurements showed significant contribution toward prediction decisions. The visualization module provided clinicians with understandable explanations rather than only numerical predictions, increasing transparency and reliability.

Overall, the results demonstrate that the proposed Explainable Hierarchical Transformer framework provides an effective solution for ICU glucose monitoring by combining prediction accuracy, multimodal clinical learning, and explainability. The approach reduces dependence on reactive monitoring methods and supports proactive decision-making in critical care environments.

Table.1 Performance Evaluation

Model	MAE ↓	RMSE ↓	R^2 Score ↑	Accuracy (%) ↑	Precision (%) ↑	Recall (%) ↑	F1-Score (%) ↑
LSTM	18.42	25.76	0.82	86.5	85.9	85.2	85.5
GRU	17.35	24.21	0.84	87.8	87.1	86.9	87.0
WaveNet	16.91	23.64	0.85	88.6	88.0	87.7	87.8
Hierarchical Transformer (Proposed)	13.28	18.95	0.91	93.4	92.8	93.1	92.9

Table 1 presents the performance comparison of the proposed Hierarchical Transformer model with baseline deep learning approaches, including LSTM, GRU, and WaveNet, using regression and classification evaluation metrics. The regression metrics, including MAE, RMSE, and R^2 score, measure the accuracy of future blood glucose prediction, where lower error values and higher R^2 values indicate better forecasting capability. The classification metrics, including accuracy, precision, recall, and F1-score, evaluate the ability of the models to correctly identify glycemic risk categories such as hypoglycemia, euglycemia, and hyperglycemia. From the comparison, the proposed Hierarchical Transformer achieves superior performance with the lowest prediction errors and highest classification scores. This improvement is attributed to the attention fusion mechanism, which effectively captures important temporal patterns and interactions among heterogeneous ICU clinical features. The results demonstrate that the proposed approach provides more accurate and reliable glucose prediction compared with traditional sequential models, supporting its effectiveness for intelligent ICU monitoring and clinical decision-support applications.

5. CONCLUSION

The developed framework presents an effective approach for predicting blood glucose variations in ICU patients by leveraging heterogeneous clinical data from multiple sources. By integrating glucose history, vital signs, medication records, laboratory measurements, and patient-specific information, the model captures complex physiological patterns associated with glucose fluctuations. The attention-based hierarchical learning mechanism enhances the identification of meaningful temporal dependencies and improves prediction performance across different forecasting horizons. The framework successfully forecasts future glucose levels while classifying patients into clinically relevant categories, including hypoglycemia, euglycemia, and hyperglycemia. In addition, the incorporation of explainable artificial intelligence improves model transparency by identifying influential factors behind prediction outcomes, supporting clinical interpretation and decision-making. Experimental evaluation on ICU clinical data demonstrates that Transformer-based architectures can effectively process heterogeneous and time-dependent healthcare information. Overall, the results highlight the potential of explainable deep learning techniques to support early glucose risk identification and improve patient monitoring in critical care environments.

Future work can focus on improving the scalability, adaptability, and real-world applicability of the framework. Integrating real-time streaming data from ICU monitoring systems may enable continuous prediction and faster clinical response. Federated learning can be explored to support privacy-preserving multi-institutional training and improve generalization. Incorporating multimodal data such as clinical notes, imaging, and additional physiological signals, along with large-scale prospective validation, can further enhance reliability and support deployment in intelligent healthcare environments.

REFERENCES

[1] Maqsood, S., Sarwar, M. A., Belousovienė, E., & Maskeliūnas, R. (2025). GlucoNet-MM: A multimodal attention-based multi-task learning framework with decision transformer for personalised and explainable blood glucose forecasting. Diabetes, Obesity and Metabolism, 27(12), 7416–7430.

[2] Mehdizavareh, H., Khan, A., & Cichosz, S. L. (2025). Enhancing glucose level prediction of ICU patients through hierarchical modeling of irregular time-series. Computational and Structural Biotechnology Journal, 27, 2898–2914.

[3] Guo, W. (2026). DT-ICU: Towards Explainable Digital Twins for ICU Patient Monitoring via Multi-Modal and Multi-Task Iterative Inference. arXiv preprint arXiv:2601.07778.

[4] Zhang, J., Contreras, M., Bandyopadhyay, S., Davidson, A., Sena, J., Ren, Y., ... & Rashidi, P. (2024). Mango: Multimodal acuity transformer for intelligent ICU outcomes. arXiv preprint arXiv:2412.17832.

[5] Zhang, Y., Song, T., Wei, L., Wang, J., Pan, Y., Wei, G., ... & Cao, W. (2025). MAMIFusion: A multimodal data fusion framework based on EHRs for enhancing diabetes prediction accuracy. Network Modeling Analysis in Health Informatics and Bioinformatics, 14(1), 113.

[6] Cifci, M. A., Öney, B., Yildirim, F., Yilmaz Başer, H., & Zontul, M. (2025). Interpretable Adaptive Graph Fusion Network for Mortality and Complication Prediction in ICUs. Diagnostics, 15(22), 2825.

[7] Sunkara, S. S. (2026). Gamt-Care: A Graph Attention Multimodal Transformer Framework for CRM-Augmented Predictive Healthcare Analytics. International Journal of Artificial Intelligence, Data Science, and Machine Learning, 7(2), 17–26.

[8] Tao, R., Li, H., Lu, J., Wang, Y., Zhou, J., & Yu, X. (2026). DHAN: A Deep Hierarchical Attention Network for multisource biomedical data fusion in

precision diabetic retinopathy diagnosis. Digital Health, 12, 20552076251408517.

[9] Dai, L., Xu, H., & Zhang, Y. (2025). Automated classification of clinical diagnoses in electronic health records using transformer. PLoS One, 20(9), e0329963.

[10] Vani, M. S., Sudhakar, R. V., Mahendar, A., Ledalla, S., Radha, M., & Sunitha, M. (2025). Personalized health monitoring using explainable AI: Bridging trust in predictive healthcare. Scientific Reports, 15(1), 31892.

[11] Hayat, M. A., Baig, J., Ahmed, S., & Ayubi, A. F. (2025). Towards Early and Accurate Disease Detection Through Multimodal Predictive Modeling: Fusion of Electronic Health Records, Medical Imaging, and Omics Data Using Interpretable Machine Learning.

[12] Warrier, A., & Abhilash, K. S. (2025, October). Hybrid Edge-Cloud AI Gateway with 1D-CNN for Real-Time Anomaly Detection and Temporal Fusion Transformer for Healthcare Data Streams. In ICIDCA (pp. 204–211).

[13] Khan, M., Chakrabarty, P., Unhelkar, B., Lodhi, A. K., & Hussain, A. (2026). An Interpretable Hybrid Machine Learning Framework for Robust Type II Diabetes Prediction Using Electronic Health Records. International Journal of Artificial Intelligence and Computer Electronics, 2(1), 1–9.

[14] Zhang, W., Liu, J., & Chen, H. (2026). Multimodal Prediction of Early Clinical Deterioration in Adult Intensive Care Units: A Multi-Center Study. Proceedings on Scientific Discovery, Measurement, Modeling, Validation, and Evaluation, 16(3), 1–18.

[15] Al-Bairmani, M. T., Yazdchi, M., & Nasimi, F. (2026). A joint CNN-Bi-LSTM-transformer architecture with SHAP explanations for multi-label arrhythmia detection from 12-lead ECGs. Scientific Reports.