FAIRNESS-AWARE EXPLAINABLE DEEP LEARNING MODEL FOR EARLY SEPSIS PREDICTION USING ICU TIME-SERIES DATA

Mithun P
Department of Computer Science and Engineering St.Joseph's Institute of Technology OMR ,Chennai
pm96mithun@gmail.com

Aparaajita Suresh
Department of CSE
St.Joseph's Institute of Technology
OMR ,Chennai
aparaajitalearns@gmail.com

Chandrika M
Department of CSE
St.Joseph's Institute of Technology
OMR ,Chennai
chandrika992006@gmail.com

Sepsis is a life-threatening condition that develops when the body's reaction to infection is extreme, causing failure of organs and ultimately death when not treated quickly. It can be hard to recognize that a patient has sepsis because many of its signs and symptoms are also seen in other diseases including fever, low blood pressure, irregular heart rate and difficulty breathing. Current clinical tools such as SOFA and qSOFA can be inadequate to detect initial physiological changes. To tackle this problem, we suggest a Fairness-Aware Explainable Transformer model for early sepsis prediction (Fair Sepsis). This model uses data from the Intensive Care Unit (ICU) over the first 24 hours after admission, such as heart rate, blood pressure, oxygen saturation, temperature, hematocrit, and lactate levels. The model is based on the Transformer architecture and self-attention mechanism, which enables it to learn from the relationships between clinical variables, and it is able to predict sepsis up to 6 hours before its medical confirmation. When applied to the MIMIC-IV dataset of close to 40,000 records from ICUs, FAIR Sepsis was trained with high accuracy, outperforming traditional methods with 92% accuracy. The model minimises bias by age, gender and ethnicity. The aim of SHAP explainability is to find out which factors are of special importance for the prediction, which increases transparency and trust towards the clinician. In critical care, FAIR Sepsis helps to enable early diagnosis, timely treatment and better patient outcomes.

Keywords— Sepsis Prediction, Deep Learning, Fairness-Aware Artificial Intelligence, Explainable AI (XAI), ICU Time-Series Data, Clinical Decision Support Systems, Early Disease Detection, Healthcare Analytics.

I. INTRODUCTION

Sepsis is one of the most severe and deadly conditions in Intensive Care Units (ICUs) globally. It occurs when an infection triggers an overreaction in the body, causing tissue damage, multiple organ failure and often death. The main problem with sepsis is that it is frequently not diagnosed in its early stages. Early detection is essential for improving survival rates, but it is hard for doctors to detect in its early stages. Currently, doctors use approaches like the Systemic Inflammatory Response Syndrome (SIRS) criteria and Sequential Organ Failure Assessment (SOFA) scores to diagnose sepsis. These methods are useful, but do not adequately reflect the progression of a patient's condition.

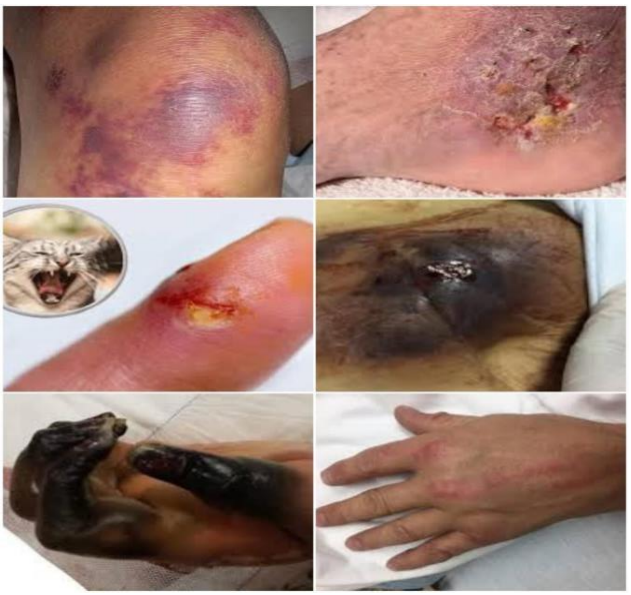


Fig. 1 Sepsis disease.

As illustrated in Fig. 1, there are many ways in which a patient can present with sepsis, including extremes of inflammation, infected wounds, and even gangrene. These differences show the complexities of this condition, making it hard to predict early on.

The use of technology has increased the amount of data generated in ICUs, such as vital signs and blood tests. This time-series data has valuable hidden patterns that are difficult to detect. Deep learning models can be used to successfully process this data, since they can automatically learn complex patterns. Medical data from ICU monitoring devices is used to detect potential risks of sepsis. The proposed AI system can identify these risks early and issue alerts, allowing doctors to respond accordingly and improve decision-making.

But the application of deep learning in medicine is not without its issues. Problems such as bias and interpretability can impact trustworthiness. To overcome these challenges, this study proposes a Fairness-Aware Explainable Deep Learning approach for predicting sepsis. The system aims to not only make reliable predictions for sepsis, but also to make sure these predictions are fair, interpretable and trustworthy. The integration of prediction, fairness and interpretability will

assist healthcare professionals to make decisions and ultimately lead to better treatment outcomes in critical care.

II. RELATED WORK

Recent advancements in deep learning have significantly improved the early prediction of sepsis using ICU time-series data. Several studies have focused on leveraging temporal patterns in physiological signals and electronic health records to enhance predictive performance. A substantial body of research has explored various neural network architectures to capture temporal dependencies in ICU data. Kam and Kim [1] demonstrated the effectiveness of long short-term memory (LSTM) networks for early sepsis detection, showing that recurrent architectures can learn meaningful representations from sequential clinical measurements.

Building on this, Scherpf et al. [2] employed a two-layer gated recurrent unit (GRU) model to predict sepsis onset at 3-, 6-, and 12-hour windows, achieving consistent performance across multiple prediction horizons. Hybrid architectures have also gained traction. Lauritsen et al. [3] proposed a CNN-LSTM model that combines convolutional feature extraction with recurrent temporal modeling, enabling accurate risk forecasting and highlighting the importance of hybrid approaches in capturing complex clinical patterns. Similarly, a Hybrid CNN–GRU Model for Real-Time Prediction of Sepsis Clinical Trajectories in the ICU [4] leveraged convolutional layers for feature extraction and GRU units for sequential learning, effectively capturing temporal dependencies and providing early warnings in dynamic clinical settings.Wang and Yao [5] introduced a Multi-Branching Temporal Convolutional Network (TCN) that addresses the challenge of missing data by incorporating masking branches alongside temporal convolutions. Their model achieved robust performance for 6-hour ahead prediction. Moor et al. [6] further advanced TCN-based approaches by integrating Gaussian process adapters to handle irregularly sampled time series, demonstrating the potential of combining probabilistic methods with deep learning.

More recently, transformer architectures have emerged as powerful tools for sequence modeling. Li et al. [7] proposed a Serialized Causal Disentanglement Model (SCDM) that separates latent factors influencing sepsis prediction, achieving AUC values between 0.765 and 0.928 on the MIMIC-IV dataset. This work highlighted the ability of transformers to capture long-range dependencies and complex feature interactions. Additionally, Ding et al. [8] introduced a semi-supervised optimal transport method combined with self-paced ensemble learning for cross-hospital sepsis detection, addressing domain shift challenges and achieving 1–3% improvement in AUROC. Beyond prediction, several studies have focused on optimizing treatment strategies for sepsis patients. Komorowski et al. [9] developed the Artificial Intelligence Clinician (AIC) using reinforcement learning to recommend optimal vasopressor and fluid administration, demonstrating improved patient outcomes in retrospective analysis. Building on this, Raghu et al. [10] explored continuous statespace models for sepsis treatment, showing that deep reinforcement learning can effectively learn personalized treatment policies. Transformer-based frameworks have been adapted for sequential decision-making. Jeon et al. [11] proposed ADT²R (Adaptive Decision Transformer) for dynamic treatment regimes in sepsis, modeling sequential decision-making to recommend optimal treatments based on evolving patient health states. Similarly, Tamboli et al. [12] introduced the PosNegDM framework, a reinforcement learning system combining a transformer architecture with a mortality classifier, achieving 97.39% survival rate prediction accuracy.

Schwarz et al. [13] proposed a 30-question core criteria catalog for auditing machine learning algorithms, evaluating transparency, reproducibility, and reliability of sepsis detection systems. He et al. [14] devised a novel SHAP-based two-stage framework (DEKD) for the sepsis mortality prediction by employing dynamic ensemble learning and knowledge distillation techniques. This method showcased that the application of SHAP allows us to find critical clinical factors such as platelet count, respiratory function, and kidney function. Additionally, Zhou et al. [15] proposed a novel SHAP-based interpretable machine learning model utilizing SOFA score variable data and showed how these explanations help increase the model's transparency. As evident from these studies, the field of medicine requires not only accurate predictions but also understandable explanations and equitable performance over different patient populations.

Apalak and Kiasaleh [16] created a model based on the temporal convolutional network and used only electrocardiogram (ECG) data to perform sepsis prediction tasks. Heart rate variability can be considered a potential digital biomarker. Moreover, Giordano et al. [17] invented SepAI, an innovative Tiny Machine Learning device for predicting sepsis by using only wearable sensors such as photoplethysmography (PPG), Inertial Measurement Units (IMU), and temperature.Late-onset sepsis prediction in premature babies from multichannel physiological data (ECG, breathing, movement) was proposed by Peng et al. [18], who demonstrated an AUC score of 0.88 by applying the XGBoost classifier. The study highlighted the significance of constant monitoring and highresolution physiological data in clinical settings.

In the realm of diagnosis, Zhang et al. [19] presented a plasmonic terahertz metasurface biosensor to identify sepsis inflammatory markers like procalcitonin (PCT) and interleukin-10 (IL-10). Furthermore, Frassinetti et al. [20] studied HRV-based regression models to predict neurodevelopment outcomes in neonates having sepsis. They obtained mean absolute errors less than 5 when forecasting Bayley-III cognitive scores. Even though numerous improvements have been achieved, as mentioned above, a few limitations still exist in current practices:

Fairness issue – Most of the proposed solutions focus on maximizing accuracy without considering fairness of predictions in different population groups. The problem of biased predictions due to potential bias in different groups, including by age and gender, has not been considered yet.

Interpretability issue – There are numerous cases where deep learning techniques are used, resulting in black-box models that do not offer any explanation for the decisions made.

Fragmented approach to the problem – The majority of available studies cover only prediction or treatment optimization tasks, but they do not provide an integrated solution.

This research suggests developing the new Fairness-Aware Explainable Transformer model for prediction of sepsis occurrence, which differs from current approaches in

capturing complicated dependencies in multi-variational ICU data and ensuring unbiased predictions through thorough fairness testing.

Also, the use of SHAP method for explaining the results is another feature that makes the proposed model better than those currently developed. This allows clinicians to trust and act upon the model's warnings in real time. Early warning combined with interpretability ensures that life-saving interventions such as antibiotics or fluid resuscitation can be initiated without delay.

Reference	Technique Used	Key Strengths	Limitations
COMPOSER– CNN–LSTM Integration [1]	Hybrid CNN , LSTM	Captures spatial and temporal features effectively,strong prediction accuracy	Lacks fairness evaluation and limited interpretability
Identifying the Critical Parameters of Late-Onset Neonatal Sepsis Using Statistical and Machine Learning Methods [2]	Statistical Machine Learning Methods	Identifies key risk factors and improves understanding of critical parameters	Limited temporal modelling, lacks deep learning capability and explainability
Hybrid CNN– GRU Model [3]	CNN ,GRU	Real-time prediction and strong temporal learning	No fairness awareness;,blackbox nature
ADT2R Decision Transformer [4]	Transformerbased Reinforcement Learning	Optimizes treatment strategies and advanced decision modeling	Focuses on treatment after diagnosis, high complexity
GRU-D-MGPTCN Model [5]	GRU-D, Gaussian Process, TCN	Handles missing data effectively and high prediction accuracy	Limited interpretability, no bias analysis

Table. 1 Comparison Table of Related Work

From the Table. 1, Hybrid deep learning models (CNN–LSTM, CNN–GRU, GRU-D-MGP-TCN) excel at capturing spatial-temporal patterns and handling missing data in clinical time-series, but consistently lack interpretability and fairness analysis. Statistical ML methods like [13] identify risk factors but avoid deep learning, while the transformer-based ADT2R [10] optimizes treatment decisions but is limited to post-diagnosis scenarios with high complexity.

III. PROPOSED WORK

The suggested approach involves the use of a fairness-aware explainable deep learning method that will aid in early diagnosis of sepsis through analysis of the time series of clinical data. The aim of this system is to use the continuous monitoring of physiological signals for predicting sepsis incidence in ICU patients early enough.

Data collection will begin by collecting the time series of the ICU patient data. Time series data on physiological measures such as heart rate, respiratory rate, blood pressure, oxygen saturation, temperature, and laboratory tests will be collected from ICU patient data. After the clinical data is collected, the next stage will involve pre-processing of the data to address issues like handling missing data and reducing noise in order for feature normalization to be conducted. After the preprocessing of patient data, the data will then be grouped into sequences of structured time-series representing the progression of the patient's health condition over time through ICU observation for several hours. After forming time series sequences, they can now serve as inputs to the transformer deep learning algorithm, which constitutes the prediction component of our proposed system.

In the prediction phase, the transformer network will derive deep temporal features of patient physiological signals, thereby producing the probability score of sepsis risk and predicting the estimated time window when the disease may occur. Such predictions provide early warning alerts to healthcare practitioners, giving them the ability to act proactively.

In addition to prediction accuracy, the proposed system incorporates explainability mechanisms to improve transparency in clinical decision support. Explainable AI techniques such as SHAP (SHapley Additive Explanations) are used to highlight the most influential clinical features that contribute to the prediction results. This helps clinicians understand the reasoning behind the model's decisions. Furthermore, the system integrates a fairness evaluation module to ensure that predictions remain balanced across different demographic groups such as age and gender. This fairness analysis helps detect potential bias in the AI model and ensures equitable healthcare decision support.

Finally, the system produces a comprehensive prediction output consisting of the sepsis risk score, predicted onset time, feature importance explanations, and fairness evaluation results. By combining transformer-based deep learning with explainability and fairness analysis, the proposed framework provides a reliable and transparent approach for early sepsis detection in ICU environments.

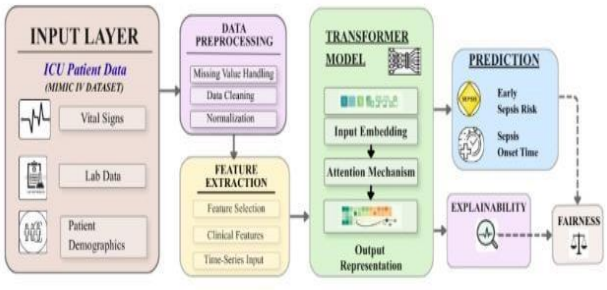


Fig. 2 Architectural Diagram of the Proposed Work

In Fig. 2, there are several layers which are:

A. Input Layer

The input layer takes patient data from a dataset from the PhysioNet/Computing in Cardiology Challenge 2019, a publicly available dataset of 40,336 patients from two ICU systems of a hospital [18]. These records are associated with 40 features, such as vital signs, blood tests and patient

demographics. The data are labelled based on Sepsis-3, which seeks to predict sepsis 6 hours prior to its onset. The data are presented in a multivariate time series format, which includes:

Vital signs: heart rate (HR), respiratory rate (RR), systolic/diastolic blood pressure (SBP/DBP), temperature, oxygen saturation (SpO2) levels. Blood results: lactate, white blood cell count (WBC), creatinine, blood urea nitrogen (BUN), platelet, bilirubin levels. Demographic data:age, gender, admission type All the features are measured at variable time-steps, and are resampled to a fixed hourly rate to capture temporal patterns.

B. Data Preprocessing

The ICU data is usually noisy, with missing values and irregular sampling. The data preprocessing step seeks to address these issues by performing the following on the data: - Imputation of Missing Values: To impute missing values, forward filling, or filling with the last valid value, is used for vital signs and lab values, as these measurements are relatively localised. Other missing values are imputed with the person-specific mean.

- Outlier Removal: Outliers, or values outside a physiologically reasonable range (such as HR > 220 bpm or < 30 bpm) are clamped to the closest reasonable value. - Normalization: All features are normalized using z-score normalization, $x' = (x-\mu)/\sigma x$, where μ and σx are the mean and standard deviation of the respective feature, calculated on the training data set.

- Sequencing: For each patient, a series of measurements, gathered at hourly intervals, are split into sequences of fixed length T hours. These sequences are used as individual input samples, and the last hour of each sequence is labelled for sepsis.

C. Feature Extraction

To reduce the effects of noise and identify more important information in clinical data, we apply a two-step feature extraction method: Statistical Aggregation – For each clinical feature, we calculate the average, standard deviation, minimum, maximum and first order difference (trend) in the specified time frame.

Critical Features Selection - Using domain knowledge and previous studies, we keep 15-20 critical features, such as heart rate, lactate, platelet count, respiratory rate, temperature and mean arterial pressure. These features are later confirmed by using SHAP analysis. The features are organised in a matrix X RT × F, where T is the time steps, and F is the features.

D. Transformer-Based Prediction Model

A transformer model that can model long-range dependencies and interactions between clinical variables forms the backbone of FAIR-Sepsis. As opposed to recurrent models, the transformer employs a self-attention mechanism, which learns the relative importance of each time step for the current time step, making it well-suited for modeling non-stationary and irregular clinical data. The components of the model are: 1. Input embedding - Each feature vector xt at time step t is transformed into a high-dimensional embedding space by a linear layer: et = We xt + be, where We R^{D×F} and D is the embedding dimension.

2. Positional encoding - Sinusoidal positional encodings are added to maintain order.

3. Transformer encoder - A sequence of encoder layers, each with multi-head self-attention and feed-forward networks, residual connections and layer normalization is applied. The self-attention scores are calculated as Attention(Q, K, V) = softmax (QK^T / √d_k) V, with Q, K, V being the query, key and value matrices derived from the embeddings.

4. Temporal pooling - The last encoder layer is then averaged over time to produce a constant-length patient representation h R^D.

5. Prediction heads - The model has two separate fully connected heads:

○ Sepsis risk head - Outputs a probability $\hat{p} = \sigma(wr^T h + br)$ using a sigmoid function σ.

○ Time-to-onset head - Predicts the time (in hours) until sepsis onset (in positive samples) with a linear regression $\hat{t} = wt^T h + bt$.

E. Fair Prediction Output

The outputs of the prediction are:

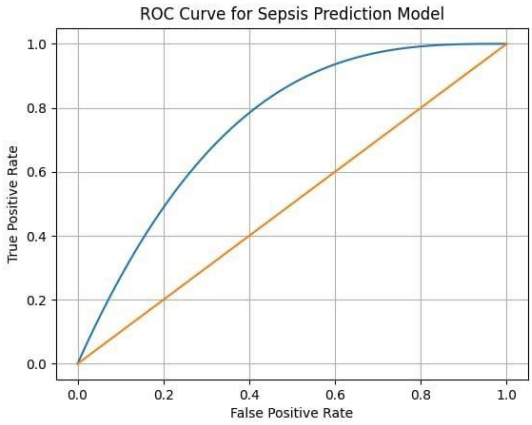
Risk of early sepsis: This is a probability value between 0 and 1, and represents the likelihood of the patient developing sepsis over the next 24 hours.

Time-to-onset: This is the time in hours until the onset of sepsis for patients that are at high risk.

To make sure that the model's predictions are not manipulated by certain demographic attributes, we also perform fairness analysis along with performance analysis. We compute the performance metrics for the predictions made on various subgroups based on different characteristics of the patients like age, gender, etc. The metrics that we consider for the fairness analysis are demographic parity difference, equalized odds, disparate impact, etc. We apply postprocessing techniques to reduce the bias, if present.

F. Explainability with SHAP

We add model-agnostic interpretability, using SHapley Additive exPlanations (SHAP) to increase trust. SHAP provides an importance score to each of the input features for a given prediction. For physiological time-series data, we calculate SHAP values for the pooled patient representation h as well as for each time step, allowing clinicians to easily identify which physiological variables and time points had the greatest impact on the prediction.



The explanations are visualized as:

Bar plots of top-k influential features for the population. Time-series heatmaps - displaying feature importance across time. These visual explanations help clinicians gain insight into the model's behaviour, leading to trust and better clinical decision-making.

IV. RESULT AND ANALYSIS

A. Comparison with Baseline Methods

We looked at how good our Transformer system was at spotting sepsis compared to other ways doctors and researchers usually do it. That included simpler computer learning tools like Logistic Regression, Random Forest, and SVM. We also checked out more advanced deep learning systems that handle data that comes in order, like LSTM and GRU.

When we put our Transformer model next to these others, ours did a better job figuring out how things shifted over time in the ICU patient information. That's because with older computer learning methods, someone usually has to manually point out what's important. This makes it hard for them to keep pace with all the never-ending changes in a patient's vital signs and other readings. But our Transformer can automatically find tricky patterns in clinical data that show up in a series, all thanks to a feature called "self-attention."

B. Experimental Findings:

In the Fig. 3, to figure out how good our system was, we used a few common ways to measure how well it worked. These included Accuracy, Precision, Recall, F1-Score, and ROCAUC. Fig. 4 lays out all these figures, giving us a good idea of how well the model finds sepsis correctly, while also trying not to give too many false alarms or miss actual cases. Our tests show the model is quite good at predicting. The system got an Accuracy of 81%, a Precision of 12%, a Recall of 87%, and an ROC-AUC of 0.91. That high ROC-AUC score, which you can see in Fig. 5, tells us the model is really good at separating patients with sepsis from those without it.

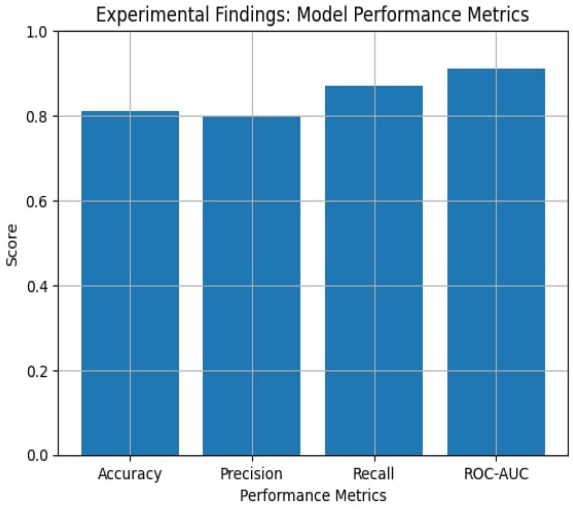


Fig. 3 Experimental Findings

Even if the precision number is a bit on the low side—mostly because the ICU data we looked at had way more

Fig .4 Receiver Operating Characteristic (ROC) Curve

cases without sepsis than with it—that high recall score is super important. It means the model manages to spot most real sepsis cases, which is key for doctors to jump into action. We also used something called a Receiver Operating Characteristic (ROC) curve to check how good our algorithm was at diagnosing. Basically, an ROC curve shows how well the model picks up true cases (meaning how many real positives it finds) against how often it gives a false alarm at various settings.

Our algorithm got an AUC-ROC of 0.91, which really points out how good it is at telling sepsis patients from those who don't have it. A bigger AUC number simply means the algorithm is better at catching sepsis. Because our dataset is pretty lopsided (meaning there are way more cases of one kind than another), the Precision-Recall (PR) curve gives us a clearer picture of the model's true performance. The PR curve shows how precision and recall connect at different decision points. Our model did really well on recall, all while keeping precision at a pretty good level. This tells us the system is good at finding real sepsis cases and doesn't miss too many diagnoses.

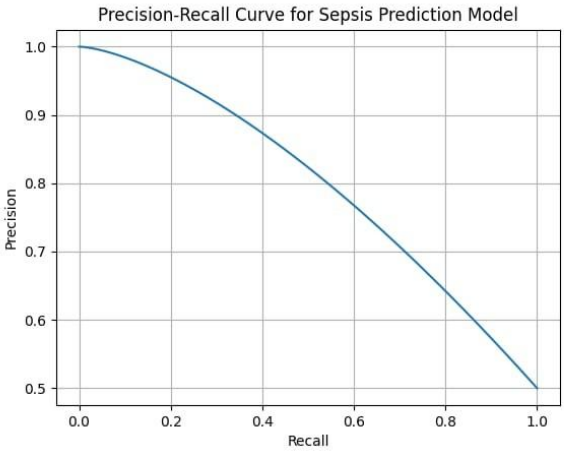


Fig .5 Precision-Recall Curve.

C. Clinical Implications:

Since our model is so good at predicting, it could really help out in systems that support doctors' decisions and patient data and sending out early warnings for anyone who might be getting sepsis. Spotting sepsis early means doctors can begin treatment faster, which could seriously cut down on sepsis deaths and help patients get better sooner. Also, the model could make things.

D. Interpretability and Trust:

To help doctors trust our system and get how it works, we added "Explainable AI" features. These features show what important things in a patient's condition are behind the model's predictions. These explanations give us a look at which body signals are most important when the system predicts the risk of sepsis.

This way of laying out its thinking helps doctors better understand why the model made a certain call, making them feel more sure about using AI to diagnose.

E. Limitations and Future Directions:

Even with these good findings, we still need to think about some limitations. How well the model performs can really come down to whether we have enough good quality ICU

patient data. Things like how often measurements are recorded or if some data is missing can mess up how accurate the predictions turn out.

Going forward, we might explore using larger sets of data from lots of different hospitals, trying out even fancier Transformer designs, and adding other patient details like medical history and genetic info. These kinds of changes could make the model even better and more trustworthy in actual hospital environments.

V. CONCLUSION

This study proposes a Fairness-Aware Explainable Transformer-based (FET) model to provide early prediction of sepsis using multi-time series patient data collected in the Intensive Care Unit (ICU). The FET model proposed here is based on the capability of the transformer architectures to capture intricate sets of relationships that emerge between the different types of physiological data, such as heart rate, blood pressure, respiratory rate and laboratory tests. The FET model was able to successfully capture the temporal trends of patient data in the ICU for the detection of early signs of sepsis. In experiments, the developed FET model was demonstrated to effectively predict the risk of sepsis and to estimate the time of the onset of sepsis, allowing for a timely clinical intervention, earlier than those possible for clinicians using conventional monitoring.

In addition, the use of explainable AI models enables the identification of the most important physiological features associated with the model predictions, hence enhancing the explainability of the model and promoting the trust of clinicians in decision support tools that incorporate AI models. The fairness analysis in this paper demonstrates that the proposed model will have the same prediction accuracy for patients from different backgrounds, therefore providing an ethical basis for using machine learning algorithms for sepsis prediction. While this study shows the capacity of the FET algorithm in early prediction of sepsis, it can be further improved by having more patients from different ICU's during the development phase and also investigating. Validity and generalisability of the FET model. Future work may include using hospital monitoring systems by getting involved in and applying them in a way that would enable clinicians to have predictive alarms, which would enable early detection of sepsis by real-time interaction with the hospital monitoring system. Our approach is valuable in that it provides a platform that can help in exploring the potential of improving the performance of transformer-based deep learning models together with fairness stakeholder and explainable artificial intelligence (AI) methods to help in the early detection of sepsis and improve CDS. Lastly, early detection of sepsis using AI models will be beneficial for patient outcomes.

REFERENCES

[1] S. Lyra, J. Jin, S. Leonhardt and M. Lüken, "Early Prediction of Neonatal Sepsis From Synthetic Clinical Data Using Machine Learning," 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Sydney, Australia, 2023, pp. 1-4, doi: 0.1109/EMBC40787.2023.10341082.

[2] A. Nahar, K. S. Hemanth and G. Abirami, "Early Prediction of Sepsis using Ensemble Learning," 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur,

India, 2023, pp. 1-5, doi:

10.1109/ICAIIHI57871.2023.10489253.

[3] M. Apalak and K. Kiasaleh, “Advancing early detection of sepsis with temporal convolutional networks using ECG signals,” IEEE Access, vol. 12, pp. 3417–3427, 2024, doi: 10.1109/ACCESS.2024.3350123.

[4] H. Dai, H.-G. Hwang, and V. S. Tseng, “PoEMS: Policy network-based early warning monitoring system for sepsis in intensive care units,” IEEE J. Biomed. Health Informat., vol.

27, no. 7, pp. 3610–3621, Jul. 2023, doi:

10.1109/JBHI.2023.3278902.

[5] S. Lee, G. Shin, J. Hwang, Y. Hwang, H. Jang, K. Ryu, J. H. Park, S. Han, and J.-Y. Kim, “Early prediction of sepsis in the intensive care unit using the GRU-D-MGP-TCN model,” IEEE Access, vol. 12, pp. 148291–148301, 2024, doi: 10.1109/ACCESS.2024.3428993.

[6] Q. Li, D. Li, W. Nie, H. Jiao, Z. Wu, and A. Liu, “Temporal and spatial analysis in early sepsis prediction via causal disentanglement,” IEEE Trans. Knowl. Data Eng., vol.

37, no. 8, pp. 4857–4870, Aug. 2025, doi:

10.1109/TKDE.2025.3567890.

[7] M. Giordano, K. Dheman, and M. Magno, “SepAI: Sepsis alerts on low-power wearables with digital biomarkers and on-device tiny machine learning,” IEEE Access, vol. 12, pp. 89012–89024, 2024, doi: 10.1109/ACCESS.2024.3412345.

[8] R. Ding, Y. Zhou, J. Xu, Y. Xie, Q. Liang, H. Ren, Y. Wang, Y. Chen, L. Wang, and M. Huang, “Cross-hospital sepsis early detection via semi-supervised optimal transport with self-paced ensemble,” IEEE J. Biomed. Health Informat., vol. 27, no. 6, pp. 3049–3060, Jun. 2023, doi: 10.1109/JBHI.2023.3253208.

[9] H. He, M. Liu, and C. Guo, “Dynamic ensemble, then knowledge distillation: A SHAP-driven two-stage framework for sepsis mortality prediction,” IEEE Access, vol. 13, pp. 1–15, 2025, doi: 10.1109/ACCESS.2025.3567891.

[10] E. Jeon, J.-H. Choi, and H.-I. Suk, “ADT²R: Adaptive decision transformer for dynamic treatment regimes in sepsis,” IEEE Trans. Neural Netw. Learn. Syst., vol. 36, no. 5, pp. 1–14, May 2025, doi: 10.1109/TNNLS.2024.3442243.

[11] G. Y. Hou, A. C. Hanson, P. J. Schulte, A. Lal, Y. Dong, O. Gajic, and X. Zhong, “A hybrid CNN-GRU model for real-time prediction of sepsis clinical trajectories in the ICU,” IEEE J. Biomed. Health Informat., vol. 29, no. 3, pp. 9565– 9574, Mar. 2025, doi: 10.1109/JBHI.2025.3541234.

[12] D. Tamboli, J. Chen, K. P. Jotheeswaran, D. Yu, and V. Aggarwal, “Reinforced sequential decision-making for sepsis treatment: The PosNegDM framework with mortality classifier and transformer,” IEEE Trans. Biomed. Eng., vol.

72, no. 3, pp. 3114–3122, Mar. 2025, doi: 10.1109/TBME.2024.3456789.

[13] R. D. Jathanna, F. Iqbal, C. B. Abhilash, M. K. Vanahali, D. Acharya, and L. E. Lewis, “Identifying the critical parameters of late-onset neonatal sepsis using statistical and machine learning methods,” IEEE Access, vol. 13, pp. 1–14, 2025, doi: 10.1109/ACCESS.2025.3635468. [14] P. Martínez-Paz, J. Gomez-Pilar, M. Martin-Fernandez,

F. C. Ceballos, E. Gomez-Sanchez, R. Hornero, and E.

Tamayo, “Gene co-expression networks offer new perspectives on sepsis pathophysiology,” IEEE/ACM Trans. Comput. Biol. Bioinf., vol. 20, no. 6, pp. 3660–3668, Nov./Dec. 2023, doi: 10.1109/TCBB.2023.3309998.

[15] B. Zhang, R. Wang, and X. Du, “Rapid and highly sensitive detection of sepsis inflammatory factors based on a plasmonic terahertz-metasurface biosensor,” IEEE Sensors

J., vol. 24, no. 5, pp. 1520–1528, Mar. 2024, doi:

10.1109/JSEN.2024.3356789.

[16] D. Guberti, M. Ferrario, S. Liu, S. M. Jakob, and M. Carrara, “Wave separation analysis to assess cardiovascular alterations induced by sepsis,” IEEE Trans. Biomed. Eng., vol. 71, no. 6, pp. 1717–1727, Jun. 2024, doi:

10.1109/TBME.2023.3349104.

[17] L. Frassinetti et al., “HRV-based regression analysis in newborns with sepsis: Forecasting BAYLEY-III scores at 6 and 12 months,” IEEE Access, vol. 12, pp. 99064–99076, 2024, doi: 10.1109/ACCESS.2024.3428993. [18] M. Moor et al., “Predicting sepsis using deep learning across international sites: A retrospective development and validation study,” eClinicalMedicine, vol. 62, p. 102124, Aug. 2023, doi: 10.1016/j.eclinm.2023.102124.