Early Sepsis Detection Using Multi-Modal Deep Learning on ICU Patient Data

23b1861 Varun Sri Teja, 23b0378 Siddharth, 23b2181 Jishnu
Indian Institute of Technology Bombay

***Abstract*—Sepsis is a life-threatening condition caused by the body's extreme response to infection. Early detection remains one of the most important clinical challenges, as every hour of delayed treatment significantly increases mortality. Traditional rule-based scoring systems such as SOFA or SIRS fail to generalize across patient populations and often lag behind physiological deterioration. Deep learning methods have recently emerged as a powerful alternative, capable of modelling complex physiological dynamics over time.**

This project presents a complete end-to-end deep learning pipeline for early sepsis prediction using ICU patient time-series data. The pipeline includes patient-wise cleaning, missing value handling, sepsis onset computation, early labeling within a six-hour prediction horizon, temporal normalization, sequence generation, and compatibility with multiple neural architectures. Experiments confirm that temporal models can detect early signs of sepsis well before clinical onset. This work lays the foundation for real-time clinical decision support systems.

I. Introduction

Sepsis is one of the leading preventable causes of death in ICU settings worldwide. According to global health estimates, millions of deaths annually are attributed directly or indirectly to sepsis-related complications. Early identification is crucial because treatment efficacy decreases rapidly with time. However, the subtle and nonlinear nature of physiological deterioration makes early detection extremely challenging.

Modern ICUs collect large amounts of continuous patient data including vitals, laboratory results, and intervention logs. Yet clinical systems rarely leverage the full predictive capability of these time-varying signals. Deep learning models, especially recurrent architectures and transformers, have shown enormous potential for representing complex temporal health trajectories.

The purpose of this project is to build a complete preprocessing and modeling framework for early sepsis detection using high-resolution ICU data. Emphasis is placed on generating high-quality temporal sequences that accurately reflect patient physiology.

II. Dataset Description

The dataset consists of ICU patient records sampled hourly. Each row corresponds to a patient-hour record with:

- **Vital signs:** Heart Rate (HR), Oxygen Saturation (O2Sat), Temperature, Systolic and Diastolic Blood Pressure, Respiratory Rate.
- **Laboratory measurements:** WBC count, Platelets, Creatinine, Bilirubin, Lactate, BUN, Chloride, Magnesium, Calcium, AST, etc.
- **Time index:** Hour of ICU stay.
- **Patient_ID:** Unique patient identifier.
- **SepsisLabel:** Binary indicator assigned per hour.

Patients have unequal sequence lengths, ranging from a few hours to several days. This heterogeneity necessitates patient-wise grouping and careful preprocessing.

III. Preprocessing

A. Column Standardization

Column names are cleaned by removing whitespace and standardizing formats. Numerical types are enforced. Erroneous or constant-value columns are removed.

B. Handling Missing Values

ICU datasets naturally contain missing observations due to lab measurement delays or irregular sampling. A two-stage policy is applied:

1) **Forward fill:** For each patient trajectory:

$$x_t^{(i)} = \{ x_t^{(i)}, if observed x_{t-1}^{(i)}, if missing$$

2) **Mean imputation:** Remaining NaNs replaced by feature-wise means from the training set.

This preserves temporal continuity while avoiding information leakage.

C. Early Label Generation

For each patient, the earliest hour where `SepsisLabel = 1` is recorded as the clinical onset. A new label is assigned to all hours for which:

$$(onset\_time - t) \leq 6 hours$$

This converts the task into predicting impending sepsis, not just identifying current sepsis.

D. Patient-Wise Splitting

To avoid contamination across time-series, patients—not individual hours—are split as:

$$70\% train,\ 15\% validation,\ 15\% test$$

E. Normalization

Feature normalization is applied using:

$$x' = \frac{x - \mu_{train}}{\sigma_{train}}$$

IV. Methodology

A. Sequence Construction

Each patient trajectory is represented as a multivariate sequence:

$$X = [x_1, x_2, \ldots, x_T], \quad x_t \in R^F$$

where T varies per patient and F is the number of features.

Sliding windows or full-length sequences are supported depending on the model.

B. Model Compatibility

Although the report focuses on preprocessing, the pipeline supports multiple deep architectures:

- **LSTM/GRU:** capture long-term dependencies.
- **1D CNNs:** excel at local temporal patterns.
- **Transformers:** model global attention and long-range interactions.
- **Hybrid CNN-RNN models:** combine spatial and temporal modeling.

C. Loss and Optimization

Binary cross-entropy is used:

$$\mathcal{L} = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

Optimization typically uses Adam with scheduled learning rates.

V. Experiments

A. Evaluation Metrics

Class imbalance requires robust metrics:

- AUROC
- AUPRC
- Sensitivity/Recall (primary clinical metric)
- F1 Score
- Accuracy

B. Training Protocol

Models are trained on patient sequences with early labels, with validation-based checkpointing. Dropout and weight decay prevent overfitting.

VI. Results

The preprocessing pipeline generated stable sequences suitable for deep temporal modeling. Initial experiments show:

- Models detect rising physiological abnormalities hours before clinical onset.
- Transformers and LSTMs outperform static baselines by large margins.
- Recall improves significantly when using onset-based early labeling.

Qualitative analysis shows that heart rate variability, rising lactate, and abnormal respiratory patterns are strong predictors.

VII. Conclusion

We presented an end-to-end early sepsis detection framework including data cleaning, time-series generation, early labeling, and model preparation. The system is compatible with state-of-the-art temporal deep learning models and replicates modern clinical ML pipelines.

Future work may incorporate:

- Multimodal fusion with clinical notes.
- Transformer-based global attention.
- Uncertainty-aware risk scoring.
- Real-time ICU deployment.

References

[1] Singer et al., "The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)," JAMA, 2016.

[2] Moor et al., "Foundations of Machine Learning for Healthcare," 2021.

[3] Rajkomar et al., "Machine Learning in Medicine," NEJM, 2019.