

Project Title: Multimodal framework for ICU based on clinical notes and lab result.

Introduction

- **Background**

The Intensive Care Unit (ICU) is a specialized hospital ward that offers comprehensive and continuous care to critically ill patients. As the population of critically ill patients grows, the demand for ICUs has risen significantly, placing strain on already limited and costly intensive care resources [1]. Sometimes some patients have death rates or prolonged length of stay (PLOS). Accurate forecasting can help for improve output and prevent these possibilities. It is considered very important because it allows early intervention to prevent major risks.

- **Methods**

Data source and inclusion criteria

The data for this study were sourced from the Medical Information Mart for Intensive Care IV (MIMIC-IV), an authoritative and comprehensive database [3]. The database contains health records of the Beth Israel Deaconess Medical Center from 2008 to 2019 and includes over 40,000 unique patients from critical care units. The admission information was recorded into various tables, such as demographics, lab results, and ICU information. Compared to its predecessor, MIMIC-III, this dataset contains updated patient information and extends the scope of data captured, thus offering a more current view of patient care. The utilization of MIMIC-IV for our study ensures that our analysis is grounded in the latest available data, facilitating a more accurate and relevant exploration into the factors affecting patient outcomes in intensive care settings. [2]

- **Problem Statement**

In critical care settings like the Intensive Care Unit (ICU), accurate and timely prediction of patient outcomes (e.g., mortality, length of stay) is crucial for optimizing treatment decisions and resource allocation. Traditional models rely predominantly on structured data such as laboratory results and clinical notes,

often ignoring the rich context found in unstructured clinical notes (e.g., physician assessments, nursing observations, and discharge summaries). This results in a loss of valuable semantic information that could enhance decision-making. This project aims to develop a multimodal deep learning framework that integrates clinical notes (unstructured textual data) and lab results (structured numerical data) to improve the prediction of critical outcomes for ICU patients.

- **Objectives**

- ✓ Improve prediction of outcomes like mortality and prolonged stay.
- ✓ Integrate structured and unstructured data for a fuller picture.
- ✓ Enhance reliability and reduce errors in predictions.
- ✓ Support smarter clinical decision-making.

Plan:

1. **Data Collection & Preprocessing**

- **Source:** MIMIC-IV.
- **Lab Data:** Normalize and impute missing values.
- **Clinical Notes:** Clean text, remove PHI, tokenize, and segment by time or event.

2. **Feature Extraction**

- **Structured Data:** Extract features like lab.
- **Unstructured Data:** BioClinicalBERT.

3. **Multimodal Fusion**

- **Early Fusion:** Concatenate features before modeling.
- **Late Fusion:** Train separate models and combine predictions.
- **Hybrid Fusion:** Use attention or belief function theory to weigh modalities adaptively.

4. **Modeling**

- **Architecture:** Use LSTM for temporal data.
- **Training:** Optimize for classification (e.g., mortality, PLOS) using cross-entropy or focal loss.
- **Validation:** Use stratified k-fold or time-based splits.

5. Evaluation

- **Metrics:** AUC, F1-score, precision-recall, calibration curves.
- **Ablation Studies:** Test performance with and without each modality.

6. Deployment & Feedback

- **Integration:** Embed into ICU dashboards or EHR systems.
- **Interpretability:** Use SHAP to explain predictions.
- **Feedback Loop:** Continuously retrain with new data and clinician input.

Table-1 Papers Comparing

Paper	Year	Methods	Dataset	The best model	Result
[1]	2025	Applying ML model.	MIMIC-III	FT-Trans	0.7672
[2]	2024	Applying ML model.	MIMIC-III	AUROC	0.884

References

- [1] Terwiesch, C., KC, D., Kahn, J.M.:Working with capacity limitations: operations management in critical care. *Critical Care* 15, 1–6 (2011)
- [2] Gao, J., Lu, Y., Ashrafi, N., Domingo, I., Alaei, K., Pishgar, M.: Prediction of sepsis mortality in icu patients using machine learning methods. *BMC Medical Informatics and Decision Making* 24(1), 228 (2024)
- [3] MIMIC-IV (Medical Information Mart for Intensive Care, Version 4.0). Laboratory for Computational Physiology, Massachusetts Institute of Technology; 2020. <https://mimic.mit.edu/>. Accessed 22 Feb 2024.