Title: AI-Powered ICU Deterioration Prediction with Clinical Explanations
AI CoLab Fall 2025 Internship

I. Team & Roles:

Project Management Lead: Kimberly Rivera Velez
Technical Lead: Kimberly Rivera Velez

List all team members:
- Zane Al-Dalli: Research in data, Presentation editor, Poster presentation (introduction)
- Atika Fatima: Research in Data,Presentation editor, Literature review, Poster presentation(Objectives)
- Hiba Sheikh: Literature Review, Presentation editor
- Sharfuddin Al Pasha Chowdhury: Literature Review, Presentation editor, Dataset Understanding.
- Kimberly Rivera Velez: Project Management, Data Cleaning and Analysis, AI/ML, Poster presentation (methods)

Ⅰ. Dataset description:

The MIMIC-III Clinical Database (version 1.4) is a large, public hospital dataset created by the MIT Laboratory for Computational Physiology. It includes detailed, de-identified information from over 40,000 adult patients who stayed in the intensive care units at Beth Israel Deaconess Medical Center between 2001 and 2012.

The database has 26 linked tables that store different types of clinical information such as patient details, hospital and ICU admissions, lab test results, vital signs, medications, procedures, and doctors' notes. Each patient, hospital stay, and ICU stay has its own ID number so the data can be connected correctly.

MIMIC-III is often used in research to study diseases, predict patient outcomes, and build artificial intelligence or machine learning models that help improve healthcare. The data are de-identified to protect patient privacy, and all users must complete required training and sign a data use agreement before accessing it.

This dataset is valuable because it provides real clinical information that can be used to find patterns, test models, and explore how different factors affect patient recovery and survival in the ICU.

II. Group Project Overview

Summary:

This project studies patient deterioration in the Intensive Care Unit (ICU) using the MIMIC-III Clinical Database (version 1.4). The main goal is to identify early warning signs and patterns that show when a patient's condition may begin to worsen during their ICU stay.

The project involves exploring patient records, vital signs, lab results, and treatments to find key risk factors linked to deterioration. The data will be cleaned, organized, and analyzed to create a structured dataset that can be used for predictive modeling and clinical research.

This project can help patients by allowing doctors and nurses to detect health decline sooner, respond faster, and prevent serious complications or deaths. It can also support the design of AI-based early warning tools that alert healthcare teams before a patient's condition becomes critical. In the long term, this work can improve patient safety, treatment planning, and overall outcomes in critical care settings.

Purpose:

The central purpose of this project is to significantly improve patient safety and outcomes within the Intensive Care Unit (ICU) by implementing a proactive, intelligent early warning system. Rather than waiting for clear clinical signs of distress, the system utilizes advanced machine learning to continuously analyze high-fidelity vital sign waveform data—such as ECG and Arterial Blood Pressure (ABP)—to predict patient deterioration, like cardiac arrest or septic shock, hours in advance. Crucially, the project is designed to be clinically trustworthy: it integrates explainable AI (XAI) to identify the specific physiological features driving the prediction, and then leverages Large Language Models (LLMs) to translate that technical data into clear, actionable clinical alerts and recommendations for the ICU staff. Ultimately, this dual approach ensures that clinicians are not just alerted to a risk, but are empowered with the reason for the risk, providing a critical time buffer for targeted interventions and shifting the standard of care from reactive response to proactive prevention.

Objectives :

- **Data Acquisition & Filtering**: Identify all usable ≥ 4-hour segments with ECG/ABP.
- **Feature Engineering**: Extract robust vital sign trend features like HR change and Shock Index.
- **Dataset Preparation**: Create a balanced, labeled dataset for ML training.
- **Develop a Predictive ML Model**: Random Forest classifier analyzes patterns.
- **Ensure Clinical Explainability**: Identify the top 3 contributing vital sign features.
- **Generate Clinical Explanations**: LLM to create actionable alerts/recommendations.

Success Criteria:

- **Model Performance:** Achieve ≥ 85% accuracy and AUC ≥ 0.85 in predicting ICU patient deterioration using ≥ 4-hour ECG + ABP waveform segments, with no patient overlap between training and test sets.
- **Explainability:** Identify the top 3 most predictive features (e.g., BP change, BP percent change, HR late mean) and generate clear clinical rationales for every prediction using LLM-based explanations.
- **System Efficiency:** Produce complete ML predictions and LLM-generated clinical summaries for each patient segment within 30 seconds of data input through the Streamlit dashboard.
- **Validation Scope:** Successfully analyze and validate on 90 + patient segments from ≥ 25 unique ICU patients, maintaining a clinically realistic deterioration rate (~ 23 %) and demonstrating full data integrity and compliance with the MIMIC-III data-use agreement.

Milestones:

- **Milestone 1:** Data collection and environment setup completed
- **Milestone 2:** Risk prediction model functional and tested
- **Milestone 3:** Resource matching system integrated and working
- **Milestone 4:** Web interface complete with validation results documented

III. <u>Scope:</u>

In Scope:

- Risk prediction for food insecurity, housing instability, and transportation barriers
- Resource matching for DC/MD/VA metropolitan area
- Web-based prototype interface
- Validation with synthetic patient scenarios
- Basic eligibility filtering for community resources

Out of Scope:

- Integration with actual MedStar EMR systems
- Real patient data (using synthetic/test cases only)
- Advanced medical diagnosis or treatment recommendations
- Real-time resource availability verification
- Multi-language support
- Mobile application development

IV. <u>Communications Plan</u>

Stakeholder Communication Matrix:

Stakeholders	Frequency	Format
All team members	Mondays & Fridays, 3:00 PM (EST)	Zoom, Microsoft teams

Group Meeting Schedule:

- **Team syncs:** Mondays & Fridays, 3:00–3:30 PM (EST) on Zoom.

Communications Ground Rules:

1. Primary day-to-day channel: **Slack and WhatsApp**.
2. Use Zoom for live meetings on Mondays/Fridays, 3 PM (EST).
3. Share brief notes/action items in Slack after each Zoom meeting.
4. Reply to Slack when it is convenient for you to do so.
5. Check Slack messages regularly and reply when you can.

V. Risks & mitigation plan: 4-column table - risk, likelihood (1-5), impact (1-5), mitigation plan:

Risks:	**Top 5–6 Highest Likelihood Risks:**	**Top 5–6 Highest Impact Risks:**	**Mitigation Plan:**	**Responsible:**
Missing or low-quality waveform data (ECG, ABP, RESP)	Likelihood = 4	Impact = 5	Enforce ≥4-hour signal filter; remove incomplete records; interpolate short gaps; maintain data-quality logs.	Data Engineer / ML Lead

Class imbalance — limited deterioration cases (~23%)	Likelihood = 4	Impact = 4	Apply class-weighted training, oversampling, and recall monitoring; validate model sensitivity on rare cases.	ML Engineer
Potential data leakage or overfitting across patients	Likelihood = 3	Impact = 5	Ensure patient-level split; verify no shared IDs; perform independent validation on unseen patients.	ML Engineer / QA Analyst
LLM explanations may misinterpret physiological patterns	Likelihood = 3	Impact = 4	Review random sample of LLM outputs with clinical experts; refine prompts and add factual grounding checks.	Clinical AI Team / NLP Engineer
.Ethical and regulatory compliance (MIMIC-III DUA, HIPAA)	Likelihood = 2	Impact = 5	Verify CITI training completion; restrict data access to approved team	Project PI / Compliance Officer

			members; maintain de-identification compliance.	
Limited generalizability beyond MIMIC-III site	Likelihood = 3	Impact = 4	Plan external validation with MIMIC-IV or eICU datasets; transparently report dataset scope and limitations.	Research Lead

AI CoLab Fall 2025 Internship - Group Meeting Minutes & Progress Report

Meeting Information

Team Name:	Group 4
Date:	07/2025
Meeting #	
Start–End Time:	3:00 pm
Meeting Type (Kickoff / Weekly / Review / Final):	Final
Facilitator:	
Recorder:	
Attendees (Present):	
Absent Members:	

Agenda

Agenda Item	Discussion Summary	Lead

Progress Summary

Project Area	What Was Done Since Last Meeting	Current Status (On Track / At Risk / Delayed)	Supporting Notes
Data Collection / Preparation			
Model Development			
Evaluation & Testing			

Documentation / Presentation			

Action Items

Task	Assigned To	Due Date	Status (Pending / In Progress / Done)	Notes
Getting literature reviews				

Challenges / Risks

Risk or issue	Impact (1–5)	Likelihood (1–5)	Mitigation / Action	Responsible

Next Steps

- [ ] Summarize next key deliverables or milestones

- [ ] Set date and time for next meeting

- [ ] Confirm who will lead next update

Instructor / Mentor Comments (for review)

Reviewed by	Feedback / Recommendations	Date
Omar Aljawfi	Add Data Description	October, 7th, 2025