Section 1: Theory & Foundations (Member 1)

Problem Definition

- Predict 30-day readmission
- Build a model for accurate predictions
- Identify risk factors and support doctors

Data

- EHR and demographic data
- Handle bias (e.g., underrepresentation)
- Preprocessing: imputation, normalization, encoding

Model

- Random Forest
- Data split: 70/15/15
- Tune: n_estimators, max_depth

Evaluation

- Accuracy & F1-score
- Monitor concept drift
- Deployment challenge: legacy systems + privacy

Section 2: Hospital AI Case Study – Problem & Data Strategy (Member 2)

Problem

- Predict 30-day readmission
- Help hospitals reduce costs, improve care

Objectives

1. Prioritize high-risk patients
2. Use resources efficiently
3. Improve care outcomes

Data

- EHR, SDOH, labs, meds, public health

Ethical Concerns

- Bias (e.g. underserved populations)
- Privacy of health data

Preprocessing

- MICE, RFECV, LASSO, encoding, NLP

Section 3: Model & Metrics (Member 3)

Model

- Logistic Regression for interpretability

Confusion Matrix

TP: 50, FN: 10

FP: 15, TN: 75

Precision = 50 / (50 + 15) = 0.769

Recall = 50 / (50 + 10) = 0.833

Overfitting Fix

- Regularization (L2), cross-validation

Section 4: Deployment, Ethics & Trade-offs (Member 4)

Deployment

- Serialize model, build API, connect to EHR
- Create clinician dashboard
- Monitor with tools like Logstash

HIPAA Compliance

- Encrypt data, audit logs, anonymization

Bias

- Use fairness tools, retrain regularly

Trade-offs

- Interpretability (Logistic) vs Accuracy (NNs)
- Lightweight models for low-resource settings

Section 5: Reflection & Lifecycle (Member 5)

Reflection:

Working on this AI project as a team was both rewarding and challenging. The most difficult aspect was coordinating the diverse tasks across all five members each with a different focus and technical depth. Aligning the structure, tone, and formatting of the report took extra effort, especially as we combined theoretical and practical elements.

Another key challenge was understanding the real-world implications of using AI in healthcare. Writing about ethical concerns such as bias and fairness pushed us to think beyond accuracy and into human impact which was new for some of us.

If we had more time or tools, we would have built a working prototype of the model and tested deployment through a simple API. Tools like GitHub for version control or Google Docs for collaborative editing would have streamlined our process significantly. With more visual aids and validation data, we could have made the report even more persuasive.

AI Lifecycle Description:

The AI development lifecycle we followed consists of six core stages:

1. Problem Definition Clearly identifying the healthcare challenge.
2. Data Collection Gathering clinical and social data.
3. Preprocessing Cleaning, normalizing, and encoding data.
4. Modeling Selecting and training a predictive model.
5. Evaluation Using metrics to assess performance.
6. Deployment Integrating into real-world hospital systems.

Below is the visual representation of the lifecycle we followed.