Part 1: Short Answer Questions

1. Problem Definition

Hypothetical AI Problem: Predicting student dropout rates in online courses.

Objectives:

- Identify students at high risk of dropping out early.
- Enable targeted interventions to reduce dropout rates.
- Improve overall course completion rates.

Stakeholders:

- Educational institutions (administrators, instructors).
- Students.

Key Performance Indicator (KPI):

- Accuracy of dropout prediction (percentage of correctly identified at-risk students).

---

2. Data Collection & Preprocessing

Data Sources:

- Student activity logs from the Learning Management System (LMS).
- Demographic data collected during student registration.

Potential Bias:

- Overrepresentation of certain demographics (e.g., age groups, regions) could bias predictions against underrepresented groups.

Preprocessing Steps:

1. Handle missing data by imputing or removing incomplete records.
2. Normalize numerical features like time spent on platform to standardize scale.
3. Encode categorical variables such as course type using one-hot encoding.

---

3. Model Development

Model Choice: Random Forest — because it handles mixed data types well, is robust to overfitting, and offers feature importance insights.

Data Split:

- 70% training set
- 15% validation set
- 15% test set

Hyperparameters to Tune:

- Number of trees (n_estimators): affects model complexity and accuracy.
- Maximum tree depth (max_depth): controls overfitting and generalization.

---

4. Evaluation & Deployment (8 points)

Evaluation Metrics:

- F1 Score: balances precision and recall, useful when class distribution is imbalanced.
- ROC-AUC: measures overall discrimination ability of the model across thresholds.

Concept Drift:

- Concept drift is when the statistical properties of input data change over time, causing model performance to degrade.
- Monitor drift by tracking model accuracy on new data continuously and retraining periodically.

Technical Deployment Challenge:

- Scalability: Ensuring the AI system can handle increasing data volume and user requests in real-time.

Part 2: Case Study Application (40 points)

Problem Scope

Problem: Predict patient readmission risk within 30 days of hospital discharge to improve care and reduce costs.
Objectives:

- Early identification of high-risk patients.
- Assist care teams in planning follow-up interventions.
- Reduce unnecessary readmissions.
  Stakeholders: Hospital management, healthcare providers, patients.

Data Strategy

Data Sources:

- Electronic Health Records (EHRs) with clinical data (diagnoses, medications, previous admissions).
- Patient demographics (age, gender, socioeconomic status).

Ethical Concerns:

- Patient privacy and data security.
- Potential bias leading to unfair treatment of minority groups.

Preprocessing Pipeline:

- Data cleaning to remove inconsistent or duplicate records.
- Feature engineering: create new features such as number of previous admissions, comorbidity scores, length of stay.
- Handling missing values via imputation.
- Encoding categorical variables (e.g., diagnosis codes).

Model Development

Model Selection: Gradient Boosting Machine (e.g., XGBoost) — chosen for strong predictive power and ability to handle heterogeneous medical data.

Confusion Matrix (Hypothetical):

	Predicted Positive	**Predicted Negative**
Actual Positive	80	20
Actual Negative	10	90

Precision: 80 / (80 + 10) = 0.89 (89%)
Recall: 80 / (80 + 20) = 0.80 (80%)

Deployment

Integration Steps:

- Develop an API to serve predictions in real-time.

- Integrate the API with hospital EHR software workflow.
- Train healthcare staff on interpreting AI outputs.
- Monitor model performance and collect feedback for improvements.

Healthcare Compliance:

- Implement data encryption and access controls per HIPAA.
- Conduct regular audits and staff training on patient data confidentiality.
- Ensure patient consent and data anonymization where needed.

Optimization

Method to Address Overfitting: Use cross-validation during training and implement early stopping in gradient boosting to prevent the model from fitting noise.

Part 3: Critical Thinking

Ethics & Bias

Effect of Biased Training Data:

- Biased data can cause the model to underpredict risk for minority patients, leading to inadequate care and worsened health outcomes.

Bias Mitigation Strategy:

- Use fairness-aware ML techniques such as re-sampling underrepresented groups or adding bias correction constraints.

Trade-offs

Interpretability vs Accuracy:

- Highly accurate models (e.g., deep neural networks) are often less interpretable, which can reduce trust and acceptance in healthcare where decisions must be explainable.

Limited Computational Resources Impact:

- May need to choose simpler, more interpretable models like logistic regression or decision trees, sacrificing some accuracy for speed and transparency.

Part 4: Reflection & Workflow Diagram

Reflection

Most Challenging Part:

- Designing an ethical, bias-aware data strategy was challenging due to balancing patient privacy and data completeness.

Improvements with More Time/Resources:

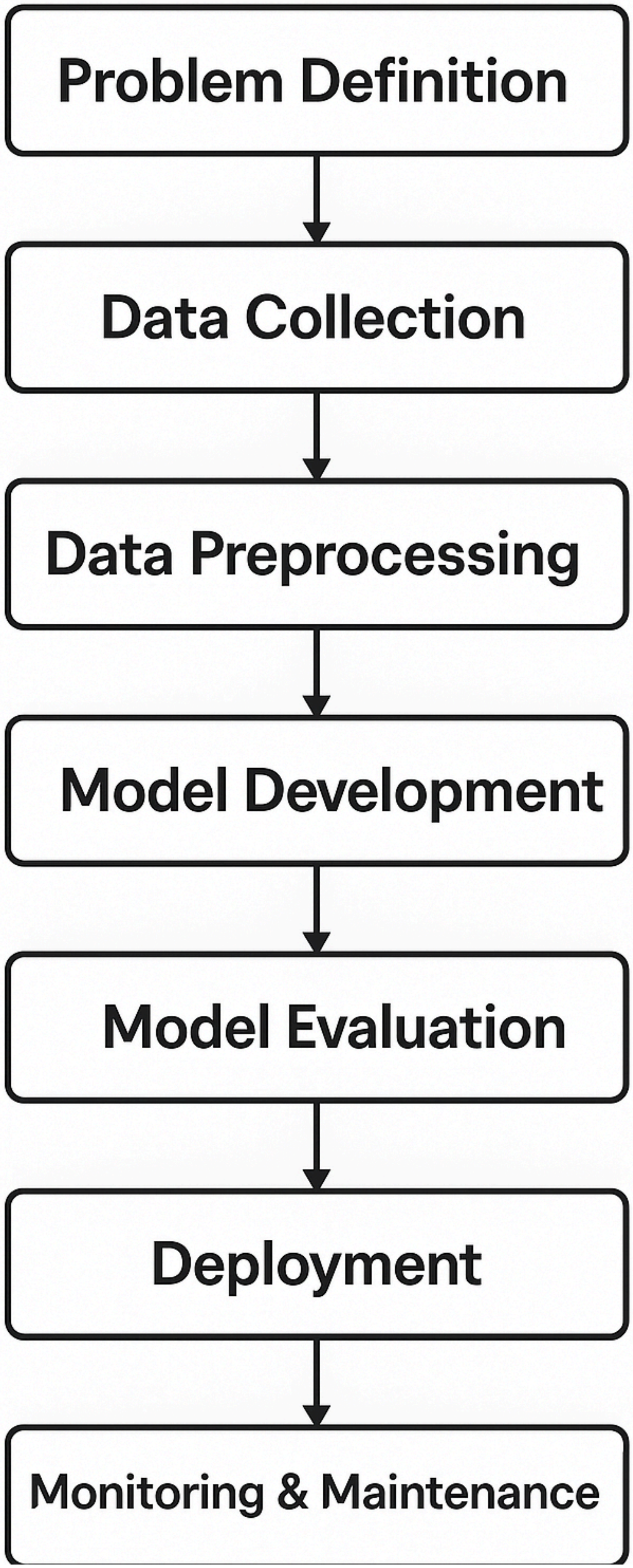
- Implement continuous model monitoring with feedback loops and invest in richer datasets, including social determinants of health.

Diagram (5 points)

A flowchart of AI Development Workflow with stages:

1. **Problem Definition**
2. **Data Collection**
3. **Data Preprocessing**

4. **Model Development**
5. **Model Evaluation**
6. **Deployment**
7. **Monitoring & Maintenance**


Problem Definition
Data Collection
Data Preprocessing
Model Development
Model Evaluation
Deployment
Monitoring & Maintenance