# **AI Development Workflow Assignment Part 1: Short Answer Questions**

## **1. Problem Definition**

Hypothetical AI Problem: Predicting student dropout rates in universities.  
  
Objectives:  
1. Identify at-risk students early based on behavioral and academic patterns.  
2. Enable targeted interventions (e.g., counseling, academic support).  
3. Improve student retention rates and institutional reputation.  
  
Stakeholders:  
- University administration  
- Students  
  
Key Performance Indicator (KPI): Accuracy of dropout prediction within a semester (e.g., >85%)

## **2. Data Collection & Preprocessing**

**Data Sources:**  
1. Learning Management System (LMS) logs (e.g., Moodle, Blackboard)  
2. Student Information System (SIS) (grades, attendance, demographics)  
  
Potential Bias: Socioeconomic or cultural bias may skew results if the model overemphasizes students from certain backgrounds as "at-risk."  
  
**Preprocessing Steps:**  
1. Handle missing data (e.g., imputation or removal).  
2. Normalize continuous features (e.g., GPA, attendance).  
3. Encode categorical features (e.g., department, residence status) using one-hot encoding.

## **3. Model Development**

Chosen Model: Random Forest  
  
Justification: It handles both categorical and numerical data well, is interpretable, and robust to overfitting.  
  
Data Split: 70% training, 15% validation, 15% test  
  
Hyperparameters to Tune:  
1. n\_estimators – number of trees in the forest (impacts performance and overfitting).  
2. max\_depth – limits tree depth to avoid overfitting on training data.

## **4. Evaluation & Deployment**

Evaluation Metrics:  
1. F1 Score – balances precision and recall, important for imbalanced dropout cases.  
2. ROC-AUC – good for measuring class separation.  
  
Concept Drift: Refers to a change in the data distribution over time (e.g., new learning platforms or COVID-19 impacting student behavior).  
Monitoring: Use a dashboard to track model accuracy and re-train periodically.  
  
Technical Deployment Challenge: Scalability – Ensuring the model processes data from thousands of students in near real-time without lags.

# **Part 2: Case Study Application**

## **Problem Scope**

Problem: Build an AI model to predict which patients are at high risk of readmission within 30 days of discharge.  
  
Objectives:  
1. Reduce preventable readmissions.  
2. Improve post-discharge care and resource allocation.  
  
Stakeholders:  
- Hospital administrators  
- Physicians and care coordinators

## **Data Strategy**

Data Sources:  
1. Electronic Health Records (EHRs): diagnosis, treatment, vitals, discharge notes.  
2. Demographic Data: age, gender, income, location.  
 **Ethical Concerns:**  
1. Patient Privacy: Ensuring data anonymization and secure handling (e.g., encryption).  
2. Algorithmic Bias: Bias in training data could lead to incorrect risk classification, especially for minorities.  
  
**Preprocessing Pipeline:**  
1. Missing value imputation (e.g., mean/mode for vitals).  
2. Feature Engineering:  
 - Length of stay  
 - Number of chronic conditions  
 - Discharge type (home vs rehab)  
3. Normalization (for lab values, age).  
4. One-hot encode categorical variables (diagnosis code, gender).

## **Model Development**

Model: Logistic Regression  
  
Justification: It’s interpretable, efficient, and widely used in healthcare for binary classification.  
  
Hypothetical Confusion Matrix:

|  |  |  |
| --- | --- | --- |
|  | Predicted Readmit | Predicted No Readmit | |
| Actual Readmit | 80 | 20 |
| Actual No Readmit | 30 | 70 |

Precision = 80 / (80 + 30) = 0.727  
Recall = 80 / (80 + 20) = 0.80

## Deployment

Integration Steps:  
1. Develop REST API for model using Flask or FastAPI.  
2. Integrate API with hospital EHR system.  
3. Trigger model after discharge form is filled.  
4. Show prediction on physician dashboard with explanation.  
  
Compliance:  
- Ensure adherence to HIPAA by:  
 - Encrypting data in transit and at rest.  
 - Logging all access events.  
 - Hosting on HIPAA-compliant cloud infrastructure (e.g., AWS HIPAA-eligible services).

## Optimization

Overfitting Mitigation Strategy:  
- Use Regularization (L2) in logistic regression and cross-validation to tune hyperparameters.

# Part 3: Critical Thinking

## Ethics & Bias

Impact of Bias:  
If training data underrepresents certain ethnic groups or chronic conditions, the model may underestimate their readmission risk, leading to inadequate care and worsening health disparities.  
  
Bias Mitigation Strategy:  
Use fairness-aware learning techniques like reweighing samples or adversarial debiasing. Also, perform subgroup analysis to evaluate model fairness across demographic groups.

## Trade-offs

Interpretability vs Accuracy:  
Complex models like XGBoost offer higher accuracy but are harder to explain. In healthcare, interpretability is often prioritized due to ethical and legal accountability.  
  
Impact of Limited Resources:  
Hospitals with limited computational power may prefer lightweight models (e.g., Logistic Regression, Decision Trees) that are easier to deploy and maintain compared to deep learning models.

# Part 4: Reflection & Workflow Diagram

## Reflection

Most Challenging Part:  
Data preprocessing was the most challenging due to dealing with missing, inconsistent, and sensitive medical data.  
  
Improvement Plan:  
With more time and resources, I would use automated ML pipelines, consult domain experts for better feature engineering, and deploy a real-time data ingestion and feedback loop.

## AI Development Workflow Diagram

