Week 4 Assignment: Understanding the AI Development Workflow

Course: AI for Software Engineering **Duration:** 7 days **Total Points:** 100

Part 1: Short Answer Questions

1. Problem Definition - Hypothetical AI Problem: Predicting student dropout rates - Objectives: 1. Identify students at high risk of dropping out early 2. Provide personalized interventions 3. Optimize allocation of academic support resources - Stakeholders: University administration, Academic advisors - KPI: Dropout rate reduction percentage within one academic year

2. Data Collection & Preprocessing - Data Sources: Student academic records, Demographics and enrollment data - Potential Bias: Underrepresentation of certain student groups - Preprocessing Steps: Handle missing values, Normalize features, Encode categorical variables

3. Model Development - Model Choice: Random Forest (handles categorical/numerical data, robust to overfitting) - Data Split: Training 70%, Validation 15%, Test 15% - Hyperparameters: n_estimators (number of trees), max_depth (tree depth control)

4. Evaluation & Deployment - Metrics: Accuracy, F1-Score - Concept Drift: Changes in student behavior over time; monitor metrics and retrain as needed - Technical Challenge: Scalability to handle large student datasets

---

Part 2: Case Study Application

Problem Scope - Problem: Predict patient readmission within 30 days - Objectives: Reduce readmissions, prioritize high-risk patients, optimize resources - Stakeholders: Hospital management, Healthcare providers

Data Strategy - Data Sources: EHRs, Patient demographics - Ethical Concerns: Patient privacy (HIPAA), Bias against minorities - Preprocessing Pipeline: Handle missing values, Normalize features, Feature engineering

Model Development - Model Choice: Gradient Boosting Classifier - Confusion Matrix (Hypothetical): | | Pred Readmit | Pred No Readmit | |----------------|--------------|----------------| | Actual Readmit | 80 | 20 | | Actual No Readmit | 15 | 85 | - Precision: 0.842, Recall: 0.8

Deployment - Steps: Deploy REST API, connect to EHR, schedule retraining - Compliance: Encrypt data, access control, logging

Optimization - Method: Regularization and cross-validation to prevent overfitting

---

Part 3: Critical Thinking

Ethics & Bias - Biased training data may misclassify minority patients - Mitigation: Bias detection, oversampling underrepresented groups, fairness-aware algorithms

Trade-offs - Interpretability vs Accuracy: Complex models are accurate but less interpretable; simpler models are interpretable but may underperform - Limited Resources: Use smaller models or reduce features for hospitals with limited computational power

---

Part 4: Reflection & Workflow Diagram

Reflection - Challenge: Handling biased/incomplete data - Improvement: Use diverse datasets, advanced feature engineering, automated monitoring pipelines

Workflow Diagram

```
[Problem Definition] --> [Data Collection] --> [Data Preprocessing] --> [Model Development] --> [Evaluation] --> [Deployment] --> [Monitoring & Maintenance]
```

- Stages include KPI definition, bias handling, model tuning, performance metrics, concept drift monitoring

---

References 1. CRISP-DM Framework: https://www.spss.com/crisp-dm 2. IBM AI Fairness 360: https://aif360.mybluemix.net/ 3. Healthcare AI Ethics Guidelines, HIPAA