AI Case Study and Implementation Report

Part 1: Short Answer Questions

1. Problem Definition:

Hypothetical AI Problem: Predicting employee turnover in a tech company.

Objectives:

- Identify employees likely to resign within 6 months.
- Understand key factors contributing to turnover.
- Help HR make data-driven retention decisions.

Stakeholders:

- Human Resources Department
- Company Executives

Key Performance Indicator (KPI):

- Recall on the positive class (correctly identifying employees who leave)

2. Data Collection & Preprocessing:

Data Sources:

- Internal HR records (attendance, performance reviews, tenure)
- Employee surveys (job satisfaction, work-life balance)

Potential Bias:

Survey data may underrepresent unhappy employees who choose not to respond.

AI Case Study and Implementation Report

Preprocessing Steps:

- Handle missing data using imputation.
- Encode categorical variables.
- Normalize numerical features.

3. Model Development:

Model: Random Forest (handles mixed data types and is interpretable).

Data Split:

- 70% Training
- 15% Validation
- 15% Test

Hyperparameters:

- n_estimators: Number of trees
- max_depth: Limit tree depth to avoid overfitting

4. Evaluation & Deployment:

Evaluation Metrics:

- Precision: Avoid false positives.
- Recall: Ensure high true positive rate.

Concept Drift:

Change in data distribution over time.

Monitoring: Use drift detectors, periodic retraining.

Deployment Challenge:

Scalability in real-time prediction environments.

Part 2: Case Study Application

Problem Scope:

Objective: Predict hospital readmissions within 30 days.

Stakeholders: Administrators, doctors, nurses.

Data Strategy:

Sources: EHRs, demographics, medication logs.

Ethical Concerns: Patient privacy, bias against age/economic status.

Preprocessing: Clean missing data, encode features, engineer (e.g., length of stay).

Model Development:

Model: Logistic Regression (transparent and interpretable).

Confusion Matrix:

	Pred Yes	Pred No
Actual Yes	50	20
Actual No	30	100

Precision: 0.625

Recall: 0.714

Deployment:

Steps: API integration, flagging high-risk patients.

Compliance: Encrypt data, audit access, de-identify training data.

Optimization:

Regularization or Cross-validation to reduce overfitting.

Part 3: Critical Thinking

Ethics & Bias:

Bias in training data can neglect high-risk groups, worsening outcomes.

Mitigation: Fair sampling and reweighting underrepresented groups.

Trade-offs:

Interpretability vs. Accuracy: Deep models may outperform but lack transparency.

Resource Limits: Lightweight models preferred in constrained environments.

Part 4: Reflection & Workflow Diagram

Reflection:

Most challenging: Addressing bias fairly.

Improvement: Stakeholder input and model retraining cycles.

Workflow Diagram:

Problem Definition -> Data Collection -> Preprocessing -> Modeling -> Evaluation -> Deployment ->

Monitoring

References

1. IBM AI Fairness 360 Toolkit: https://aif360.mybluemix.net/

2. Kaggle Breast Cancer Dataset: https://www.kaggle.com/uciml/breast-cancer-wisconsin-data

3. Scikit-learn Documentation: https://scikit-learn.org/

4. Testim.io: https://www.testim.io/