AI Assignment Report

Part 1: Short Answer Questions (30 points)

1. Problem Definition (6 points)

Hypothetical AI Problem: Predicting job application success rates based on resume and job post data.

Objectives:

- Improve job-matching accuracy
- Reduce time spent applying to jobs
- Increase applicant interview success rates

Stakeholders:

- Job seekers
- Recruiting platforms

KPI:

- Job interview conversion rate (percentage of applicants who reach interview stage)

2. Data Collection & Preprocessing (8 points)

Data Sources:

- Job application portals (e.g., Indeed)
- Resume parsing tools

Potential Bias:

- Representation bias due to over-representation of elite institutions or specific regions

Preprocessing Steps:

- Handle missing data (e.g., blank experience entries)
- Normalize job titles (e.g., "Dev" vs "Developer")
- Encode categorical features (e.g., skills, degree)

AI Assignment Report

3. Model Development (8 points)

Chosen Model: Random Forest

Justification: Handles both numerical and categorical data, less prone to overfitting, interpretable.

Data Split Strategy:

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

Hyperparameters to Tune:

- max_depth: Controls model complexity
- n_estimators: Number of trees in the forest

4. Evaluation & Deployment (8 points)

Evaluation Metrics:

- Precision: Measures accuracy of positive predictions (important when false positives are costly)
- ROC-AUC: Evaluates classification ability across all thresholds

Concept Drift:

- Data distribution may change over time (e.g., job market trends)
- Monitoring: Use drift detection tools to track changes in input distribution

Technical Challenge:

- Scalability: Processing large-scale job data in real time