Assignment 3

Project: Developing an Ethical Predictive Modelling Pipeline

Objective: Build an end-to-end predictive analytics pipeline for a socially relevant use case like healthcare, education, or public policy. The focus will be upholding principles of privacy, accountability, transparency, and equity.

Data Collection and Privacy (25 points):
- Identify suitable open or synthetic dataset aligned to social good goals
- Implement differential privacy techniques to produce aggregate privacy-preserving data statistics.

Root Cause and Fairness Analysis (25 points):
- Profile datasets along demographic groups identifying any representation skew.
- Apply quantitative bias diagnostic tests and suggest mitigations.

Model Development and Explain ability (25 points):
- Train ML models ensuring balanced performance across user segments.
- Generate model explanations to increase transparency using state-of-the-art methods like LIME and counterfactuals.

Evaluation and Recommendations (25 points):
- Evaluate model behavior along accuracy, privacy loss, and subgroup fairness.
- Demonstrate working prototype to stakeholders along with concrete accountability guidelines for real-world deployment.