Problem Statement

Commuters often receive numerous promotional offers while traveling, but many of these deals are irrelevant or poorly timed, leading to low redemption rates. The challenge is to accurately predict which deals a commuter is likely to **redeem based on their personal characteristics, travel context, and past behavior**. Developing **a model that can understand and anticipate commuter preferences** will enable businesses to deliver **smarter, personalized recommendations** that increase deal acceptance and improve marketing effectiveness.

Approach

To address the challenge of recommending personalized deals to commuters, we first gather and **preprocess a rich dataset** containing user demographics, travel context, environmental conditions, and past deal redemption behavior. We then **engineer features** that capture key patterns such as proximity to offers, timing, and lifestyle habits. Using this data, we train and evaluate several machine learning classification models—like **Logistic Regression, Random Forest, and Gradient Boosting**—to predict the likelihood of a commuter redeeming a deal. The best-performing model will be used to build a dynamic recommendation system that suggests relevant offers based on real-time user context and profile, aiming to **increase deal redemption rates** and **enhance marketing effectiveness**.

Machine Learning Algorithms

The planned machine learning algorithms for this project include classification models such as **Logistic Regression**, **Random Forest**, and **Gradient Boosting Machines (XGBoost or LightGBM)**, chosen for their effectiveness in handling structured data and binary classification tasks. **Feature engineering** will also consider **temporal and spatial proximity** to offers. The development process will include **data cleaning**, **preprocessing**, **exploratory data analysis** to identify user patterns, model training, and performance evaluation using metrics like **precision**, **recall**, **F1-score**, and **ROC-AUC**.

DataSet

The project will utilize a provided, behavior-driven dataset. This dataset features a rich set of attributes such as **trip_purpose, travel_company, current_weather, ambient_temp, time_of_day, offer_type, deal_expiry_window, and various user-centric features like user_gender, user_age_group, salary_range,** and **historical visit frequencies to bars, cafes, and restaurants**. Crucially, it also includes spatial relationship features such as **min_gap_to_offer_X and direction_match/mismatch**, and the target variable **redeemed**, indicating coupon acceptance.