**Project 4 Proposal Team 1**

**Title: Creating Home Loan App to Predict Loan Approval**

**Alberto Medina; Pam Alhousseiny; Mason Plikerd; Parmeet Arora**

**Objective:** The goal of this project is to develop a machine learning-based application that predicts home loan approval status. The app will be designed to assist lenders in making data-driven decisions by assessing various applicant features, thereby streamlining the loan approval process and reducing manual evaluation time.

**Introduction**

In the current financial landscape, the ability to efficiently assess loan applications is crucial for lending institutions. This proposal outlines the development of a Home Loan App that utilizes machine learning algorithms to predict loan approval based on various applicant attributes. By leveraging data-driven insights, the app aims to streamline the loan approval process and enhance decision-making for lenders.

**Objectives**

The primary objective of this project is to create a predictive model that can accurately determine the likelihood of loan approval based on specific applicant variables. The app will serve both potential borrowers and lending institutions by providing quick and reliable assessments of loan applications.

**Dataset**

The dataset for this project will be sourced from Kaggle, specifically the "Home Loan Approval" dataset, which can be found at the following link: [Home Loan Approval Dataset](https://www.kaggle.com/datasets/rishikeshkonapure/home-loan-approval?select=loan_sanction_train.csv). This dataset contains comprehensive information on loan applicants, including their demographics and loan details.

**Variables**

The following variables will be utilized in the predictive model:

1. **Gender**: The gender of the applicant, which may influence loan approval rates.
2. **Marital Status**: Whether the applicant is single, married, or divorced, potentially affecting financial stability.
3. **Dependents**: The number of dependents the applicant has, which can impact their financial obligations.
4. **Education**: The educational background of the applicant, which may correlate with income levels and job stability.
5. **Income**: The applicant's income, a critical factor in assessing repayment capability.
6. **Loan Amount**: The total amount of the loan requested, which will be compared against the applicant's financial profile.
7. **Term of Loan**: The duration of the loan, which can affect monthly repayment amounts and overall risk.

**Methodology**

1. **Data Preprocessing**:
   * Load the dataset into a Pandas DataFrame.
   * Handle missing values and encode categorical variables (e.g., gender, marital status, education) into numerical formats.
   * Scale numerical features (e.g., income, loan amount) to ensure uniformity.
2. **Model Selection**:
   * Implement various machine learning algorithms, including logistic regression, decision trees, and random forests, to identify the most effective model for predicting loan approval.
3. **Model Evaluation**:
   * Split the dataset into training and testing subsets.
   * Evaluate model performance using metrics such as accuracy, precision, recall, and F1-score.
4. **Deployment**:
   * Develop a user-friendly interface for the Home Loan App that allows users to input their details and receive instant feedback on loan approval likelihood.

**Expected Outcomes**

The Home Loan App is expected to:

* Provide accurate predictions of loan approval based on applicant data.
* Reduce the time taken for loan processing and enhance the efficiency of lending institutions.
* Offer insights into factors influencing loan approval, aiding in better decision-making.

**Conclusion**

This proposal outlines the creation of a Home Loan App that leverages machine learning to predict loan approval. By utilizing the specified variables and the dataset from Kaggle, the app aims to provide valuable insights and streamline the loan approval process for both borrowers and lenders.