**Project 4 - Group 5**

Project Proposal

**Project Introduction**

The goal of this project is to develop a machine learning model to predict loan approval decisions for a financial institution. Accurately assessing loan eligibility is crucial for banks and lenders to mitigate risk and make informed decisions. Historically, this process relied on manual evaluation and subjective criteria, often leading to inconsistent decisions and increased default risk.

By leveraging machine learning techniques, we aim to automate and optimize the loan approval process. The model will learn from historical loan application data, identifying patterns and relationships to predict whether a loan should be approved or rejected. This will enable faster, more objective, and data-driven decisions, ultimately reducing defaults and improving profitability for the institution.

**Dataset**

We will be using a loan application dataset sourced from Kaggle. The dataset contains information on past loan applicants, including demographic details, credit history, income, loan amount requested, and the final loan status (approved or rejected).

Key features in the dataset include:

*Number of Dependents, Education, Self Employed, Income Annum, Loan Amount, Loan Term, Cibil Score, Residential - Commercial - Luxury - Bank Assets Value and finally Loan Status*.

The dataset will be split into training and testing sets. The training data will be used to build and optimize the machine learning model, while the test set will evaluate its performance on unseen data.

**Analysis Questions**

* **What are the most significant factors influencing loan approval decisions?**

*By analyzing feature importance and conducting statistical tests, we can identify the key applicant characteristics and financial metrics that have the greatest impact on loan outcomes. This insight can help streamline the application process and guide data collection.*

* **How well can machine learning predict loan approvals compared to traditional methods?**

*We will compare the performance of our machine learning model against a baseline model reflecting current manual approval processes. Metrics such as accuracy, precision, recall, and F1-score will quantify the improvement in predictive power. We hypothesize that machine learning will substantially outperform human judgment.*

* **Can the model provide explainable insights to support its predictions?**

*For high-stakes decisions like loan approvals, transparency and interpretability of the model are critical. We will explore techniques like SHAP (SHapley Additive exPlanations) to understand how each feature contributes to individual predictions. Clear explanations build trust and aid in auditing the model's fairness and reliability.*

*By addressing these questions, our project aims to demonstrate the value of machine learning in optimizing loan approvals. The end product will be a robust, interpretable model that can be integrated into the existing loan application workflow, driving better decisions and financial outcomes.*

Carolina: Create an algorithm that analyzes credit scores and predicts consumer personal-loan eligibility.

Powerpoint Presentation

Abbey: Visualisations **What are the most significant factors influencing loan approval decisions?**

Caroline: Readme  
**How well can machine learning predict loan approvals compared to traditional methods?  
How well can machine learning predict loan approvals compared to traditional methods?**