Loan Approval Prediction Project Documentation

1. Project Title :

Loan Approval Prediction

2. Objective :

The objective of this project is to build a machine learning model that predicts loan approval status based on applicants' data. This prediction aims to assist financial institutions in making quick and data-driven decisions on loan applications.

3. Database Description :

Dataset: The dataset contains information on loan applicants, including demographic details, loan amount, credit history, etc.

Data Overview: The dataset has [total number of rows-13] records and [total number of columns-13] features, including both numerical and categorical data.

Missing Values: Details about missing values for each feature.

4. Preprocessing :

Data Cleaning: Handling missing values, removing duplicates, and correcting data types.

Encoding: Encoding categorical features using appropriate techniques.

Normalization/Scaling: Scaling numerical features to improve model performance.

5. Model Selection :

Outline the models considered for this task, such as Logistic Regression, Random Forest, SVM,XGBoost.

6. Training and Validation :

Data Split: Training and testing split ratio.

Cross-validation: Details of cross-validation, if used, to assess model stability.

Hyperparameter Tuning: Parameters tuned to optimize the model performance.

7. Results :

Metrics : Logistic Regression

Accuracy: 0.8617886178861789

Random Forest

Accuracy: 0.8292682926829268

SVM

Accuracy: 0.8373983739837398

XG Boost

Accuracy: 0.7642276422764228

Confusion Matrix: Overview of true positives, false positives, true negatives, and false negatives.

8. Conclusion :

Logistic Regression is the best model which given **Accuracy** more than **85%.**