**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 id, income, loan amount, loan term etc.

Data Overview: The dataset has total of 4270 row records and total of 13 columns features, including both numerical and categorical data.

Missing Values: there are some missing values in columns like loan amount, term, approval etc.

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 :

The models considered for this task are LightGBM, Decision Tree, XGBoost, Random Forest, AdaBoost, Extra Trees, Logistic Regression, KNN, Naive Bayes, MLP

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 :

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

MODELS ACCURACY

LightGBM 0.977594

Decision Tree 0.970519

XGBoost 0.970519

Random Forest 0.966981

AdaBoost 0.963443

Extra Trees 0.936321

Logistic Regression 0.831368

KNN 0.588443

SVC 0.604953

Naive Bayes 0.604953

MLP 0.604953

8. Conclusion :

LightGBM is the best model which given accuracy more than 95%. LightGBM, Decision Tree, XGBoost are the best models.