Project Report

# Project Title:

📌 Loan Approval Prediction Using Machine Learning

# 1. Objective:

The primary goal of this project is to develop a machine learning model that can accurately predict whether a loan application will be approved or not based on customer profile data. This helps financial institutions make data-driven decisions, improve processing speed, and minimize human bias.

# 2. Problem Statement:

Manual loan approval processes are time-consuming and inconsistent. There is a need for a reliable and automated system that evaluates an applicant’s eligibility using historical data and predicts the likelihood of loan approval.

# 3. Dataset Details:

• Source: [Kaggle / Internal Dataset]

• Total Records: ~600 entries

• Features Include:

- Gender, Marital Status, Dependents  
- Education, Employment (Self\_Employed)  
- ApplicantIncome, CoapplicantIncome, LoanAmount  
- Loan\_Amount\_Term, Credit\_History  
- Property\_Area  
- Loan\_Status (Target Variable: Y/N)

# 4. Tools & Technologies Used:

• Language: Python

• IDE: Jupyter Notebook

• Libraries:

- Pandas, NumPy for data handling

- Matplotlib, Seaborn for visualization

- Scikit-learn for modeling and evaluation

# 5. Methodology:

A. Data Preprocessing:

- Handled missing values using mean/mode imputation

- Converted categorical variables using Label Encoding

- Normalized numerical features where necessary

B. Exploratory Data Analysis (EDA):

- Analyzed trends like credit history vs. approval

- Visualized correlations between income, loan amount, and approval

C. Model Training:

- Applied multiple algorithms:

• Logistic Regression  
 • Decision Tree Classifier  
 • Random Forest Classifier

- Split data into train/test sets (80/20)

D. Evaluation Metrics:

- Accuracy Score  
- Confusion Matrix  
- Precision and Recall

# 6. Results & Observations:

• Best Model: Random Forest Classifier (or Logistic Regression depending on your results)  
• Accuracy Achieved: ~80%+  
• Credit history and applicant income were found to be the most influential factors in loan approval.

# 7. Conclusion:

The project successfully demonstrates a predictive ML model that automates loan approval decisions. With high accuracy and explainability, the model can serve as a core tool for banks and NBFCs to reduce turnaround time and increase objectivity in loan processing.

# 8. Future Enhancements:

• Deploy model via a web app using Streamlit or Flask  
• Integrate advanced models like XGBoost  
• Use SHAP/LIME for model explainability  
• Automate model retraining with new data

# 9. Author:

Vikas Kumar