Loan Approval Prediction using Python

This project involves building a machine learning model to predict whether a loan application will be approved or not, based on various applicant attributes. By analyzing historical loan data, the model aims to assist financial institutions in making informed lending decisions.

Project Overview

Loan approval prediction is a classic classification problem in machine learning. It requires analyzing key factors such as:

- Applicant income.
- Credit history
- Employment status
- Loan amount
- Marital status
- Education level

The goal is to train a model that can accurately predict the loan status (Approved or Not Approved) for new applicants.

Data set

- Source: https://www.kaggle.com/datasets/ninzaami/loan-predication
- Type: Structured tabular data
- Target variable: Loan_Status (Y/N)

Tools and Libraries

- Python
- Google Colab
- Pandas
- Scikit-learn
- Matplotlib

Machine Learning Model

- Algorithm used: Support Vector Machine (SVM)
- Model Evaluation:
  -Accuracy: 83%
  - Classification Report: Includes precision, recall, and F1-score
  - Confusion Matrix
  -ROC curve

Visualizations

- Data distribution and missing values
- Count plots for categorical features
- ROC Curve and Confusion Matrix

Results

The SVM model achieved strong predictive performance, especially in identifying approved loans. The final evaluation shows:

Classification Report:

	precision	recall	f1-score	support
N	0.94	0.49	0.64	35
Y	0.80	0.99	0.89	75
Accuracy			0.83	110
Macro avg	0.87	0.74	0.76	110
Weighted avg	0.85	0.83	0.81	110

Key Learnings

- Preprocessing and handling missing values.
- Exploratory data analysis (EDA)
- Training and evaluating a classification model.
- Visualizing model performance using ROC and confusion matrix.
- Understanding the practical application of SVM in finance.