Loan Approval Prediction Report,

created by Ilnaza Saifutdinova, 60982580
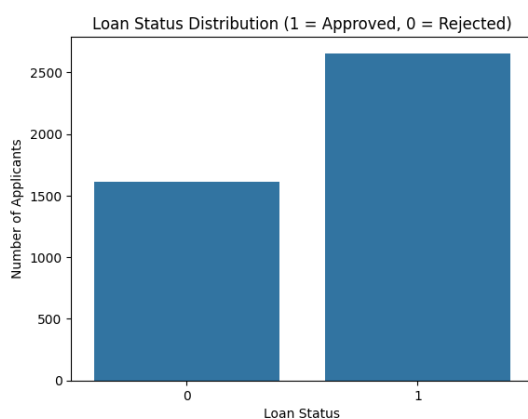
I.   Model Performance Implementation

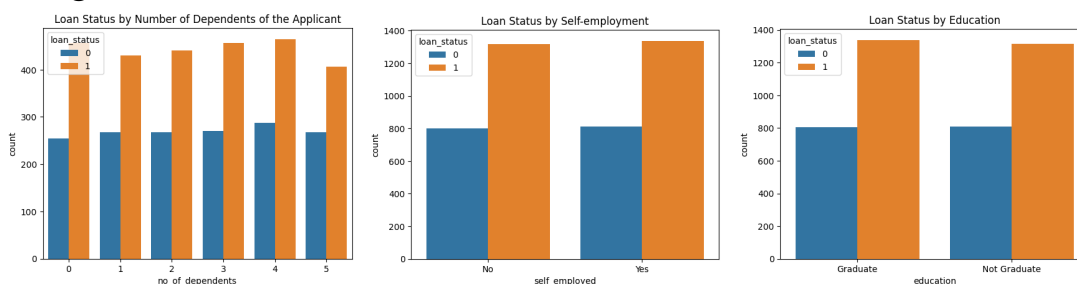| Model	CV Accuracy, %	Test Accuracy, %	Precision, %	Recall, %	F1-score, %	ROC AUC Score, %
Logistic Regression classifier	91.63	93.00	96	92	94	93.16
Decision Tree classifier	97.31	98.22	98	98	98	97.70

II.   Key Graphs
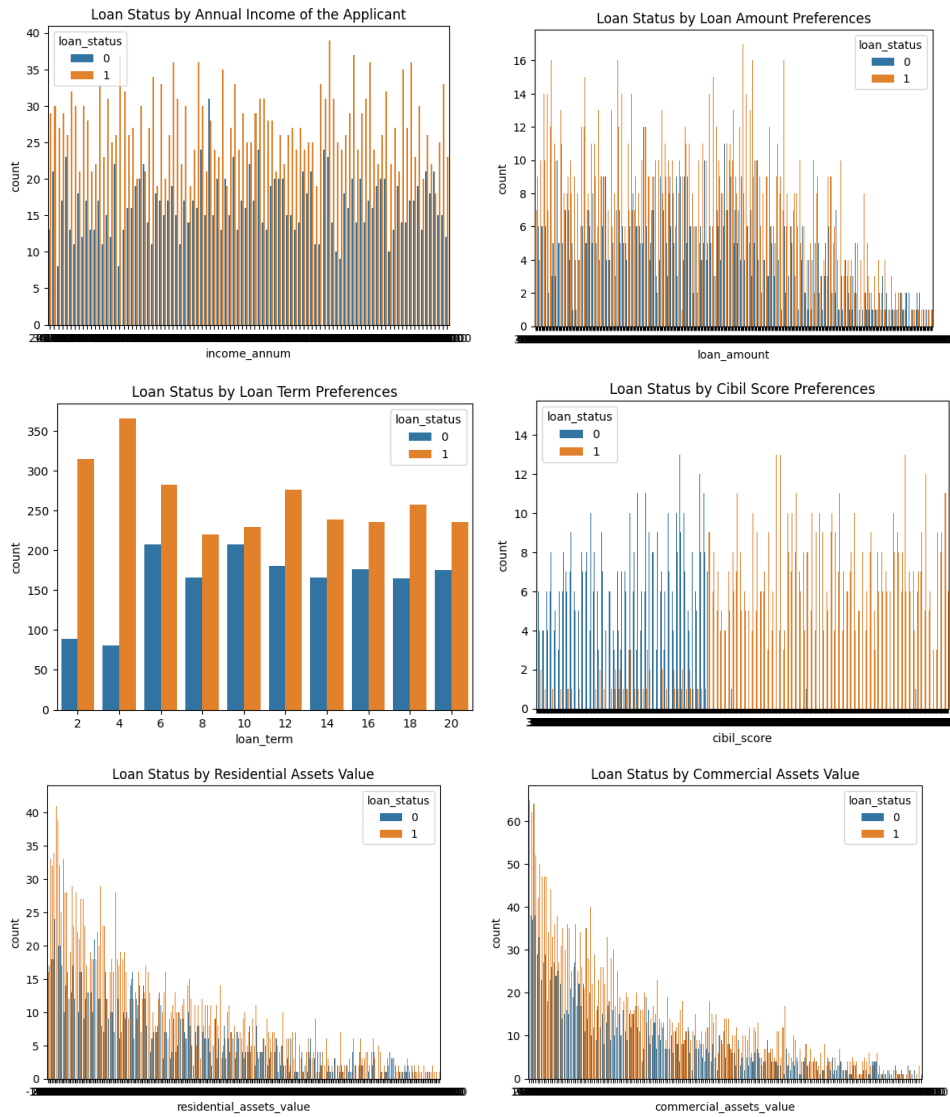
1. Loan Status Distribution

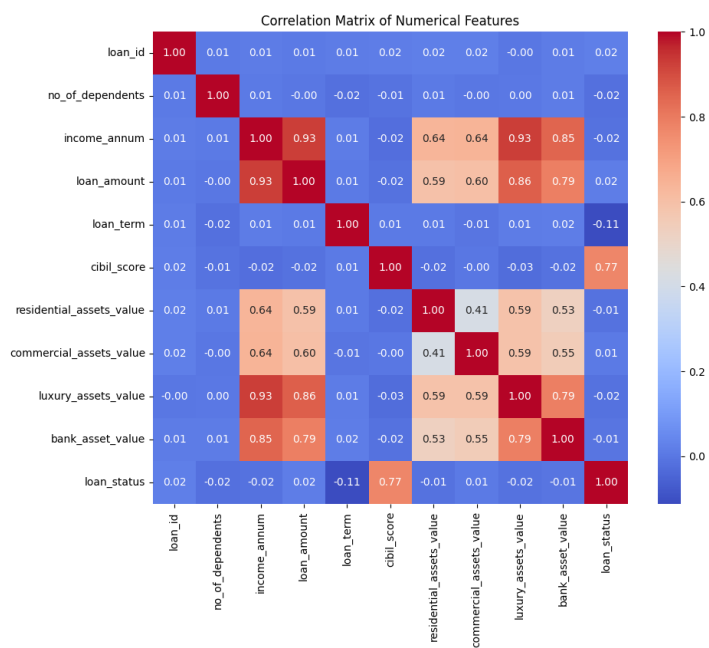A countplot showing the proportion of approved and rejected loans.
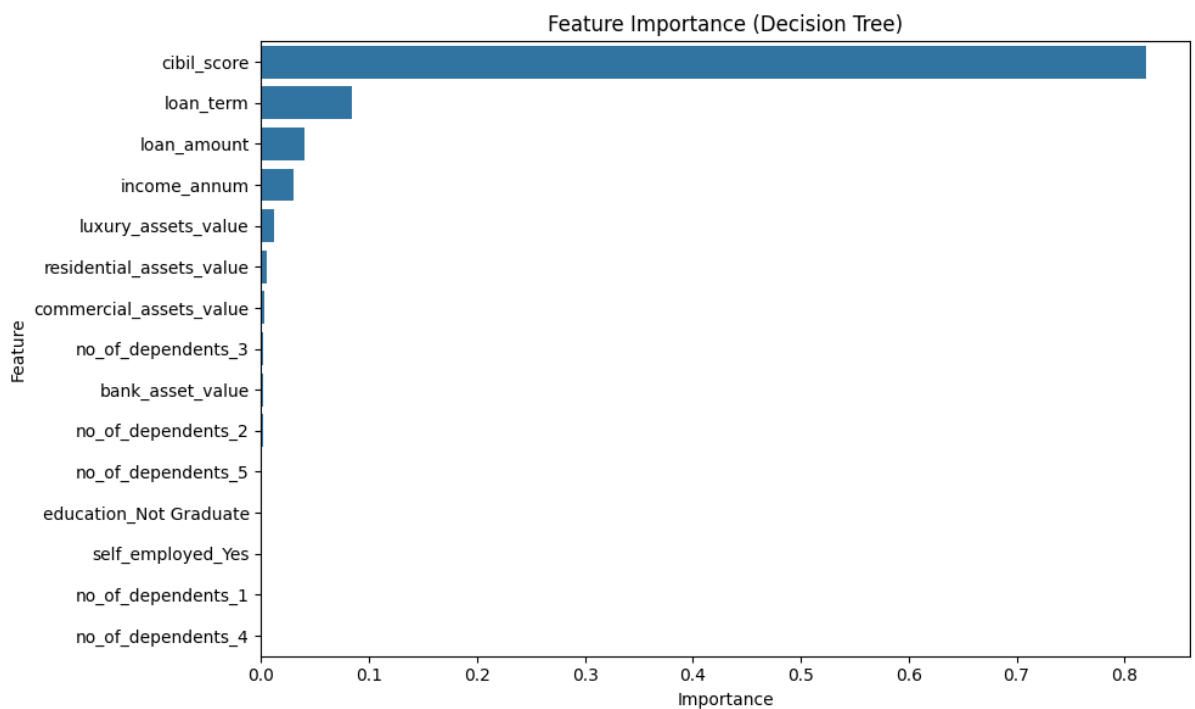

2. Categorical Features vs Loan Status

3. Categorical Features vs Loan Status

4. Correlation Matrix of Numerical Features

5. Feature Importance from Decision Tree classifier

Feature Importance (Decision Tree)

III. Model Choice and Reasoning

This project applies supervised machine learning techniques to predict loan approval outcomes using the Kaggle Loan Approval Prediction Dataset. The process involved thorough data preprocessing, feature encoding, and model training using both Logistic Regression and Decision Tree classifiers. After tuning hyperparameters and evaluating performance through cross-validation and test set metrics, the models were assessed based on accuracy, precision, recall, F1-score, and ROC AUC score. Additional insights were gained from visualizations such as confusion matrices, correlation heatmaps, and feature importance plots. Based on consistently superior results across all metrics, the ***Decision Tree classifier*** was selected as the final model due to its high predictive power and balanced performance on both training and unseen data.