

Mortgage Loan Project Description Document

This machine learning project aims to predict whether a borrower will become ever delinquent on their mortgage loan , Ever delinquent means that the borrower has missed payments for a certain period of time, typically 90 days or more.

The project will involve cleaning and analyzing a dataset containing information about past loans and creating a predictive model using machine learning techniques. The final model will be used to identify high-risk borrowers and help financial institutions prevent loan default.

The dataset for this project contains information about past mortgage loans, including details about the borrower's credit history, employment status, income, loan amount, loan term, and other factors that may affect the likelihood of delinquency. The dataset also includes information about whether each loan became ever delinquent or not.

The project will involve the following steps:

- **Data cleaning and preprocessing:** The dataset will be cleaned and preprocessed to remove missing values, outliers, and other anomalies that could affect the accuracy of the model.

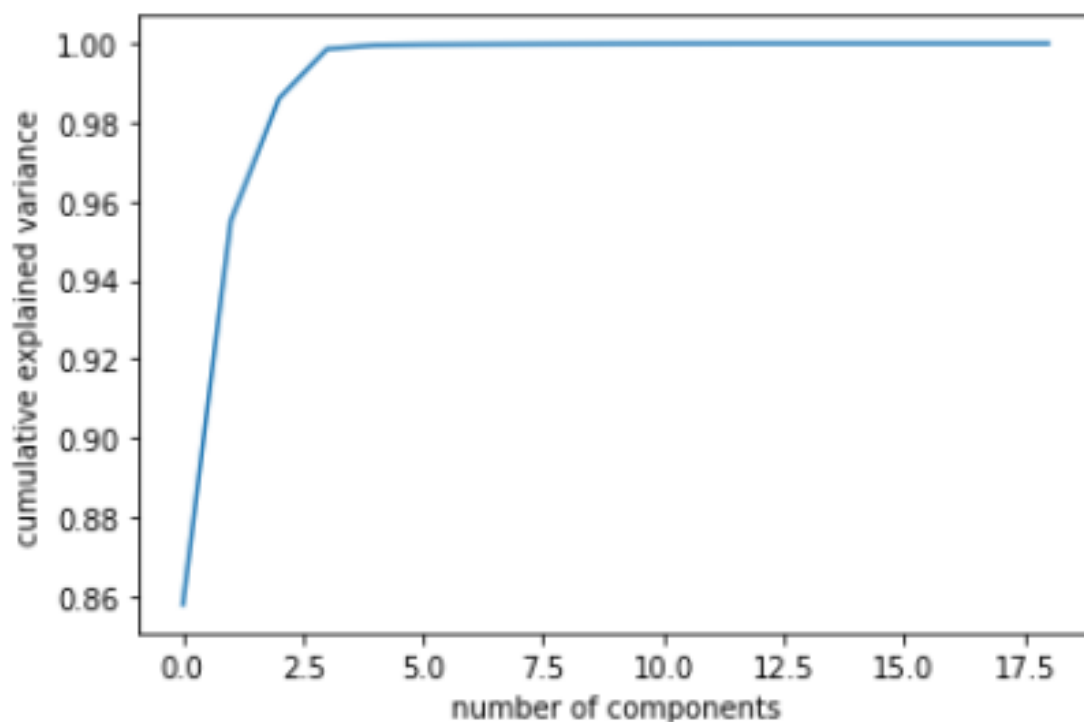
In this stage there is some EDA and preprocessing including data cleaning . We remove outliers and check the null values then handle it .

Use EDA techniques as showing profile report , plots , relations and some statistics of the data .

- Feature engineering: New features will be created based on domain knowledge or statistical analysis of the data to improve the predictive power of the model .

In this stage there is feature engineering which we made new features as credit range , LTV range and repay range which help in predicting our target according to domain knowledge .

Then , we made PCA graph and made some testing which show that better number of PCA is 6 .



- Model selection: Different machine learning models will be tested and evaluated based on their performance metrics such as accuracy, precision, recall, and F1 score.

In this stage there is the model building , we made 3 machine learning classifications models

SVC , Random Forest and Decision Tree which show different accuracies and performance

Random Forest					SVC				
Classification report:					Classification report:				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.93	0.95	0.94	139060	0	0.88	0.67	0.76	139060
1	0.79	0.70	0.74	34527	1	0.32	0.62	0.42	34527
accuracy			0.90	173587	accuracy			0.66	173587
macro avg	0.86	0.83	0.84	173587	macro avg	0.60	0.64	0.59	173587
weighted avg	0.90	0.90	0.90	173587	weighted avg	0.76	0.66	0.69	173587

Decision Tree

Classification report:				
	precision	recall	f1-score	support
0	0.80	1.00	0.89	139060
1	0.00	0.00	0.00	34527
accuracy			0.80	173587
macro avg	0.40	0.50	0.44	173587
weighted avg	0.64	0.80	0.71	173587

We chose the model which has the highest AUC curve and more sharp after testing .

- **Deployment:** The model will be deployed in a production environment such as a web application , where it can be used to predict the likelihood of ever delinquency for new mortgage loan applications.

In this stage we deployed the chosen model which is Random Forest in addition to deploy the PCA model and standardization model to easily convert the input of the user about the asked features into the shape which processed to the model to make the better accuracy .

We make a Django app which contain a list to enter the features then making preprocessing to it using the saved joblib files generated from the notebook then it print the result in another web page that show the statues of the borrower is Ever Delinquent or not .

Main input page list

TEAM B

Home

Welcome To Our Mortgage Loan Project

Enter The Credit Score...

Enter The MIP...

Enter The Units...

Enter The OCLTV...

Enter The DTI...

Enter The OrigURS...

Enter The LTV...

Enter The OrigInterestRate...

Enter The OrigLoanTerm...

Is Your LoanPurpose is C... YES=1 NO=0...

Is Your LoanPurpose is N... YES=1 NO=0...

Is Your LoanPurpose is D... YES=1 NO=0...

FirstTimeHomebuyer... YES=1 NO=0...

Enter The LTVRange...

Is Your CreditRange is Excellent... YES=1 NO=0...

Is Your CreditRange is Fair... YES=1 NO=0...

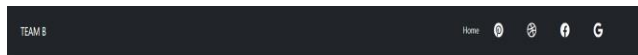
Is Your CreditRange is Good... YES=1 NO=0...

Is Your CreditRange is Poor... YES=1 NO=0...

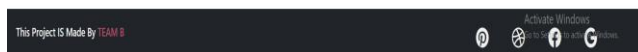
Is Your RepayRange is High... YES=1 NO=0...

Activate Windows
Go to Settings to activate Windows.

Not Ever Delinquent Result



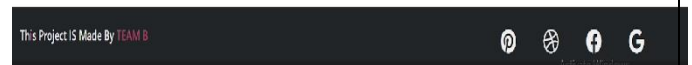
The User you Want To Classify is NOT Ever Delinquent



Ever Delinquent Result



The User you Want To Classify is Ever Delinquent



The outcome of this project will help financial institutions to identify high-risk borrowers and take appropriate measures to mitigate the risk of loan default.