Loan Default Prediction

**Data Analysis:**

First, we check the Data type of the features of the dataset to know whether it is ***categorical data*** or ***numerical data***

Then we view find the null data inside the Dataset.

**Data Filtering:**

Converting the null data into its mean value so that further no problem is occurred for the model when it transforms the train data.

The Categorical Data which named as proof\_submit is converted into integer or numeric data with the help of one hot encoding.

At last, we remove the data or dataset Feature which show less effect or no effect to the ***Target Data*** or ***Prediction Data.***

**Feature Engineering:**

Here we try the Various combination of column to derive a feature in which the target data and prediction data is highly dependent.

**Model Selection:**

Transformed data were trained on 3 models:

* Logistic Regression
* Decision Tree
* Naive Bayes

All of three models Multinomial Behaves the best and give the highest accuracy among other respective model may it is because of its probability prediction feature.

Naïve Bayes behave as the best model among the all three datasets.

**Final Result:**

The predictions predicted by the model is converted to a pandas DataFrame which further convert it into csv file.