

B F S I-Credit Risk Assignment

By
Vishwajeet Singh

B F S I –CREDIT CARD ASSIGNMENT

OBJECTIVE: The objective is to build a statistical model to estimate borrowers' Loss Given Default(LGD)

$$\text{LGD} = \frac{\text{Loan Amount} - (\text{Collateral value} + \text{Sum of Repayments})}{\text{Loan_Amount}}$$

BACKGROUND PROCESS

Credit risk analytics in the context of the banking sector and model a common metric used for estimating the expected credit loss (ECL)

ECL method is used for provisioning the capital buffer to protect banks against possible default of the customers.

The loss given default (LGD) is a measure of the amount of loss that a bank is expected to incur in the event of a

Expected credit loss
= Exposure at default x
Probability of Default x

DATA SOURCES

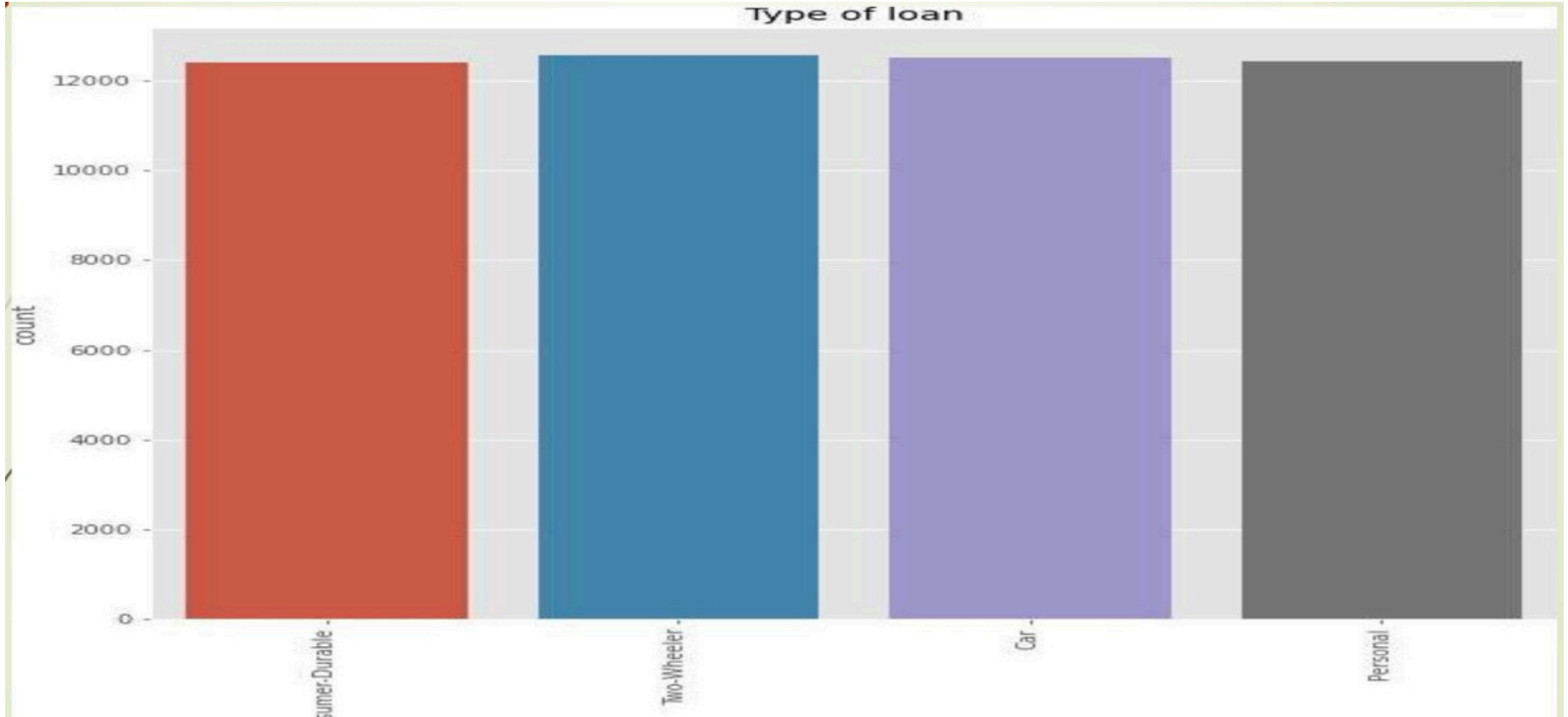
Used 3 Data sets for model Building:

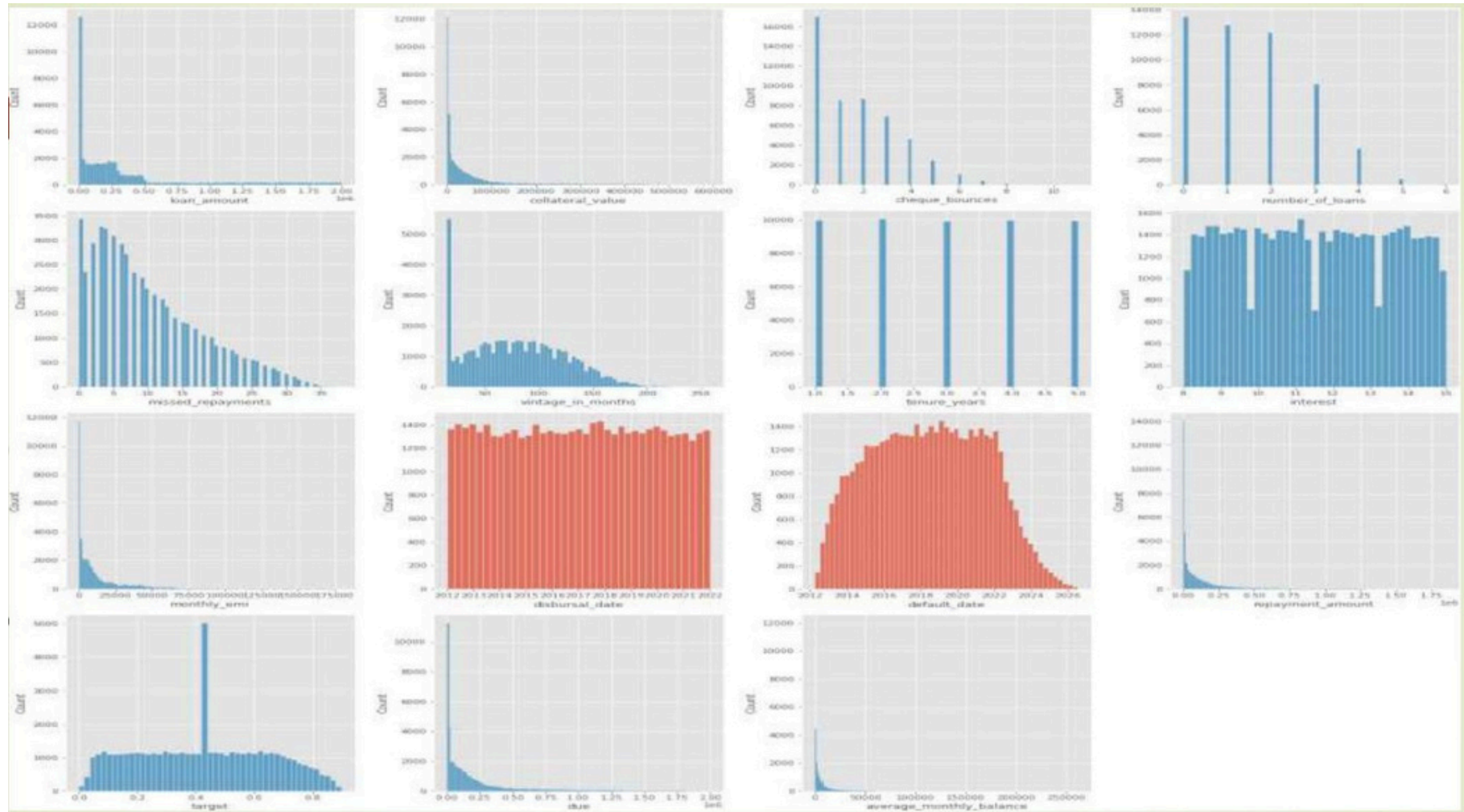
- The main__loan__base data set contains information about loan accounts and other relevant information for the corresponding borrowers.
- The repayment base data set contains information about the repayments received by the banks in the form of EMIs or through other collection efforts
- The monthly__balance__base contains the information pertaining to the monthly balance statements in the borrower's accounts

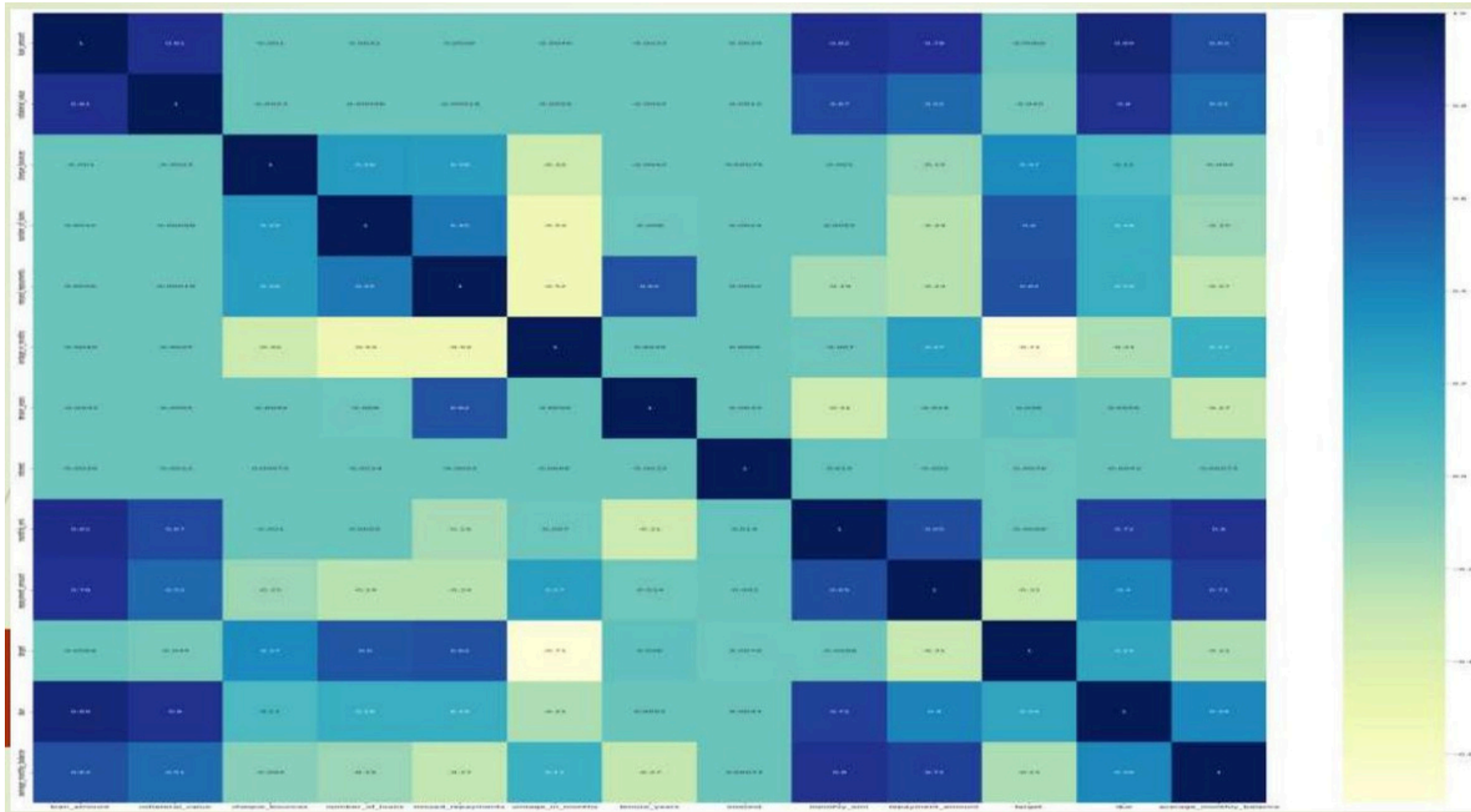
PROCESSING OF DATA

- For each data set converted Data types if necessary
- Merging the data sets and created target variable(LGD)
- Exploratory Data Analysis has been performed
- Variable Transformation
- Dummy Encoding Scaling
- using Standard Scale

EDA: Monthly EMI also car loan is much higher compared to other loans

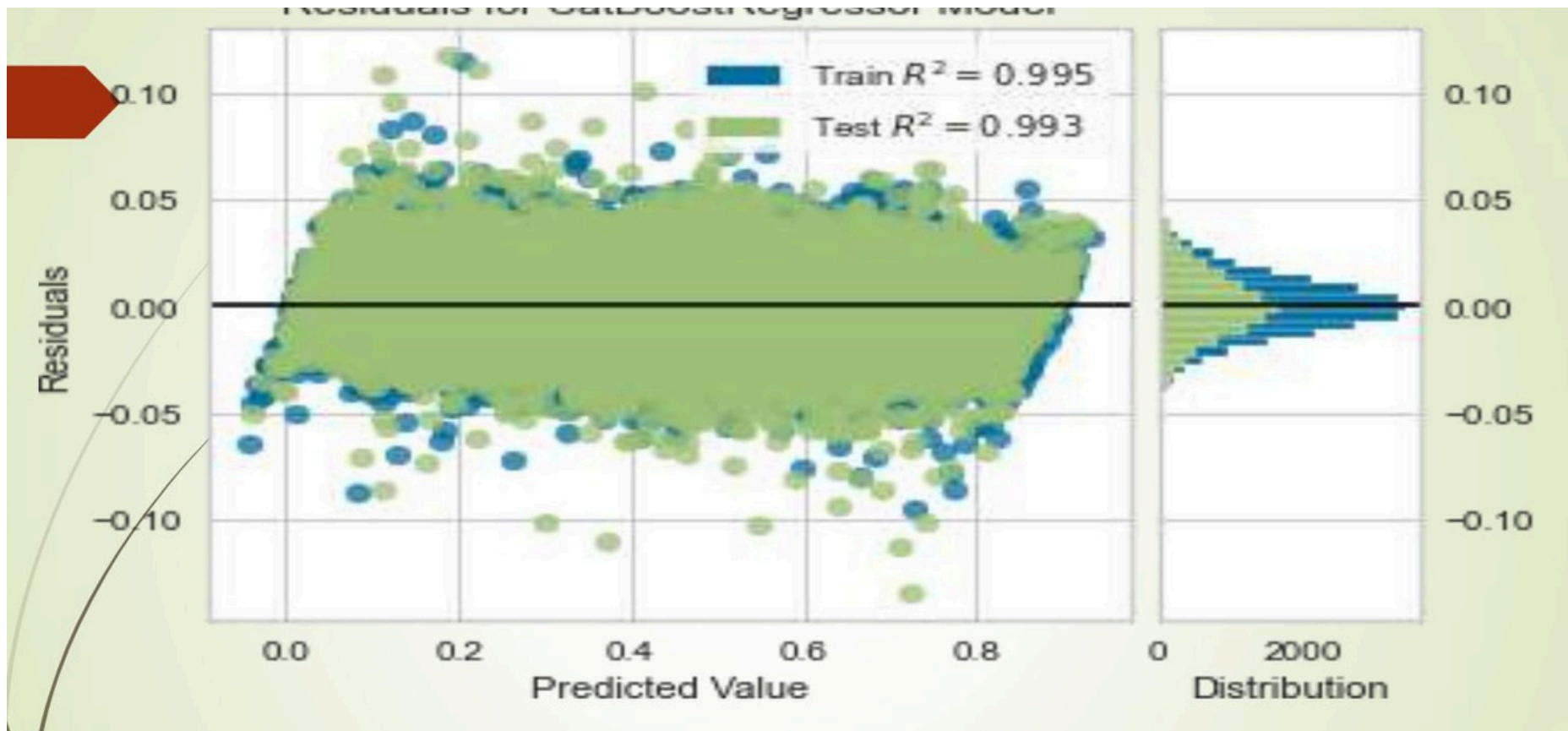






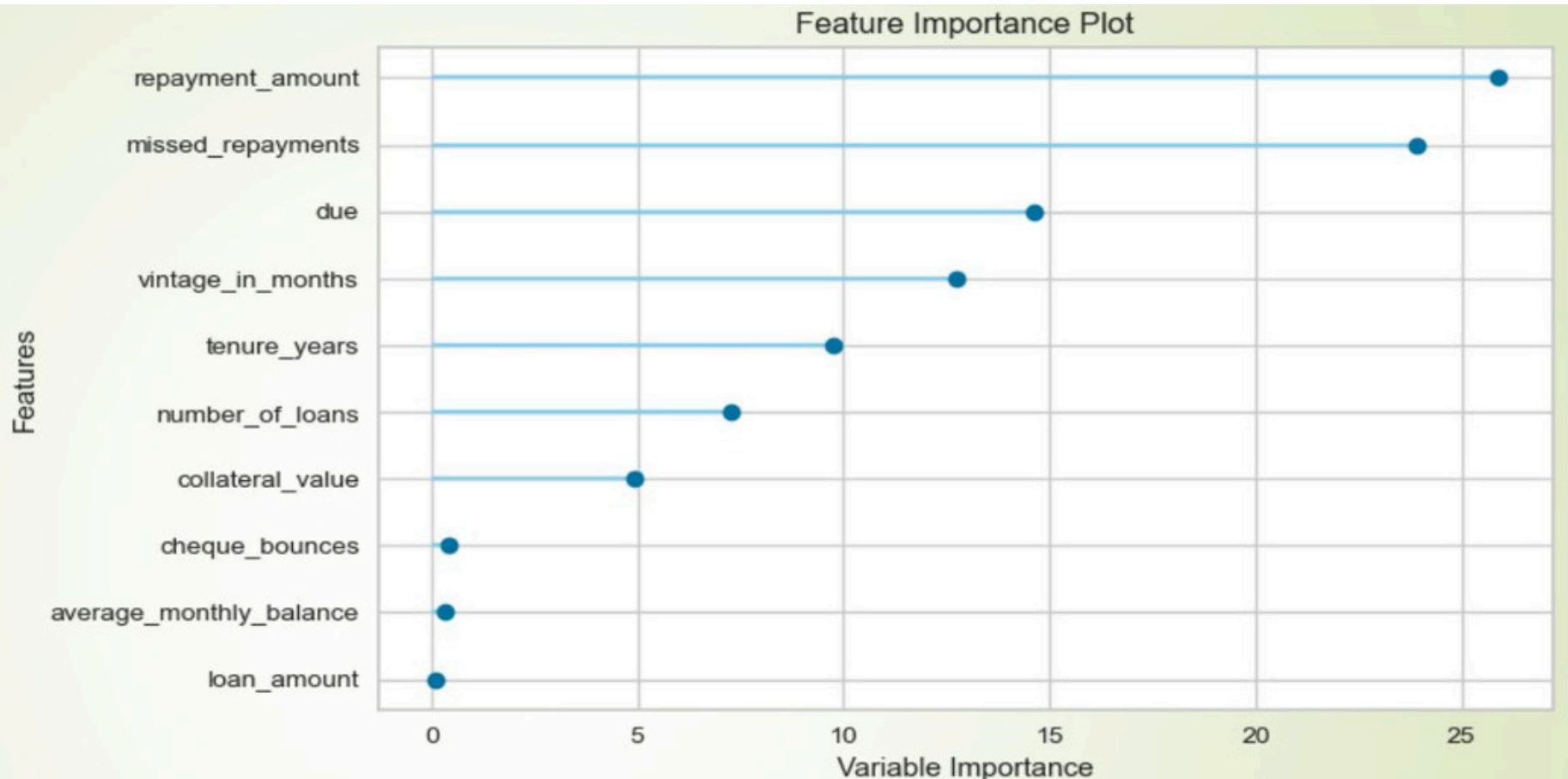
MODEL BUILDING

- Used various models like Multiple Linear Regression, Randomforestregressor, Gradient Boosting Regressor, XGBoost Regressor, Adaboostregressor, Elastic Net: Hybrid Regularized Model, LightGBM for model building.
- Used R Squared as a performance metrics.
- XGBoost has given us 99.5% R squared on test data across the models.



Residuals for catboost regressor

REGRESSION INTERPRETATION



RECOMMENDATIONS

- We should focus more on car and two-wheeler
- Missed Repayment customers with high repayment amount should be highlighted
- Customer's due factors and tenure are another subset of influencers to predict the LGD of the customers