

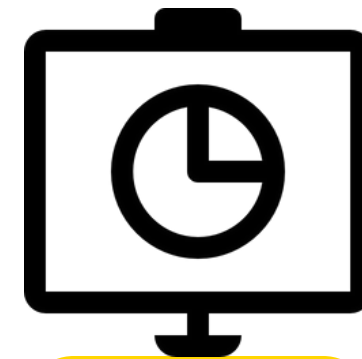
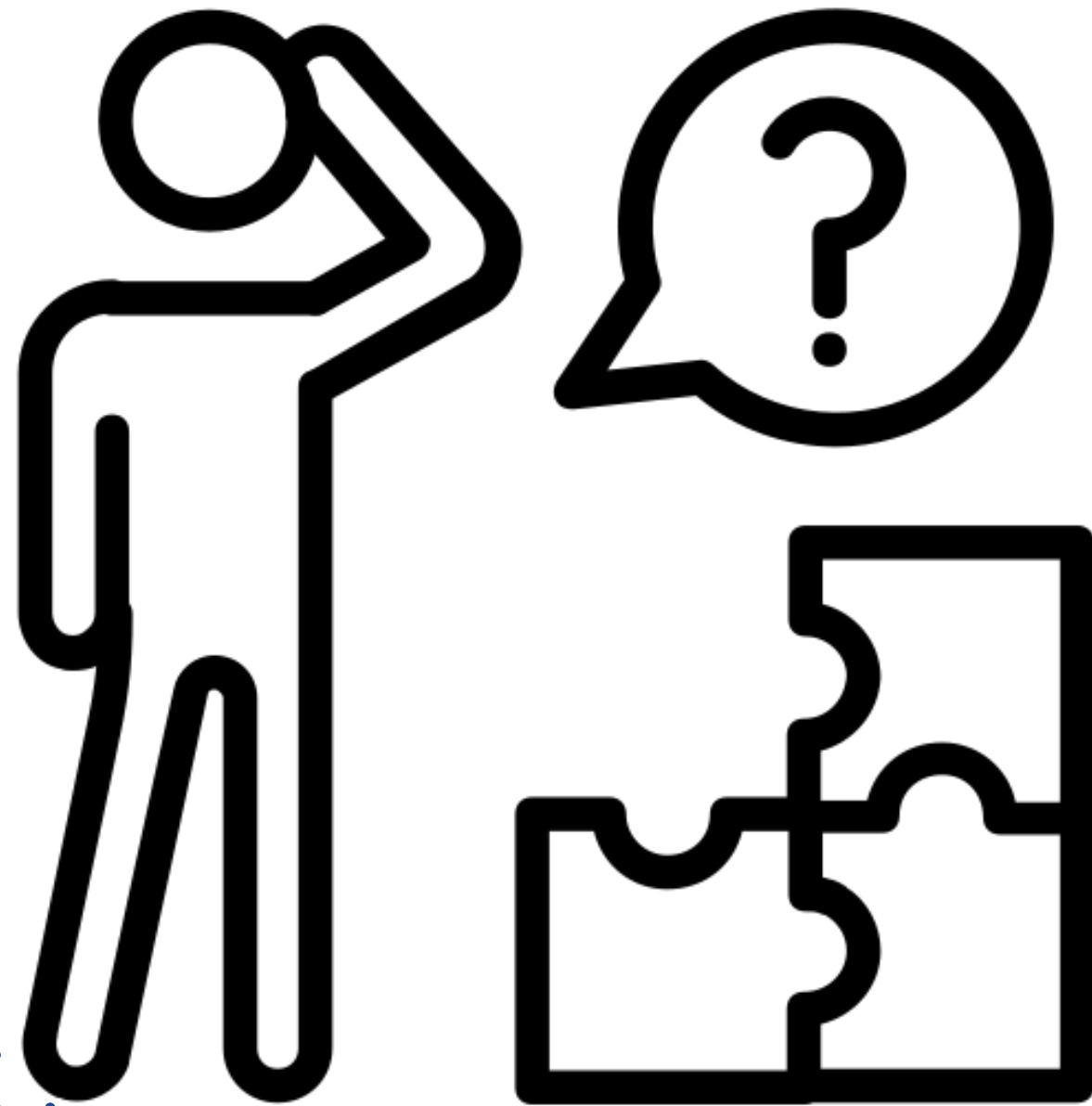
Credit Risk Prediction

by : Zulhaji



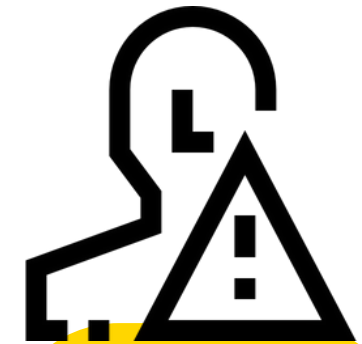
Problems

Credit risk refers to the risk of loss that a lender faces due to a borrower's failure to repay any type of loan or debt



Time

It will takes a long time if we do the assessment manually



Human Errors

It can be happen if we make wrong decisions & will cause considerable losses



Financial Loss

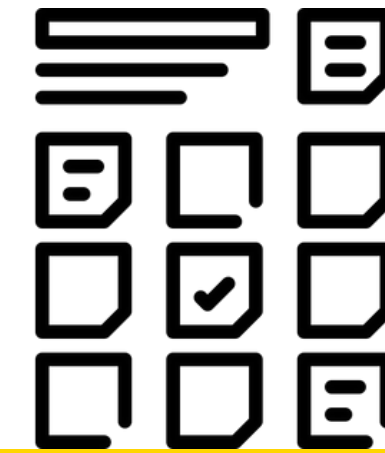
Credit loss is the amount of money lost by the lender when the applicant refuses to pay or runs away with the money owed.

Goal, Business Matrix



Goal

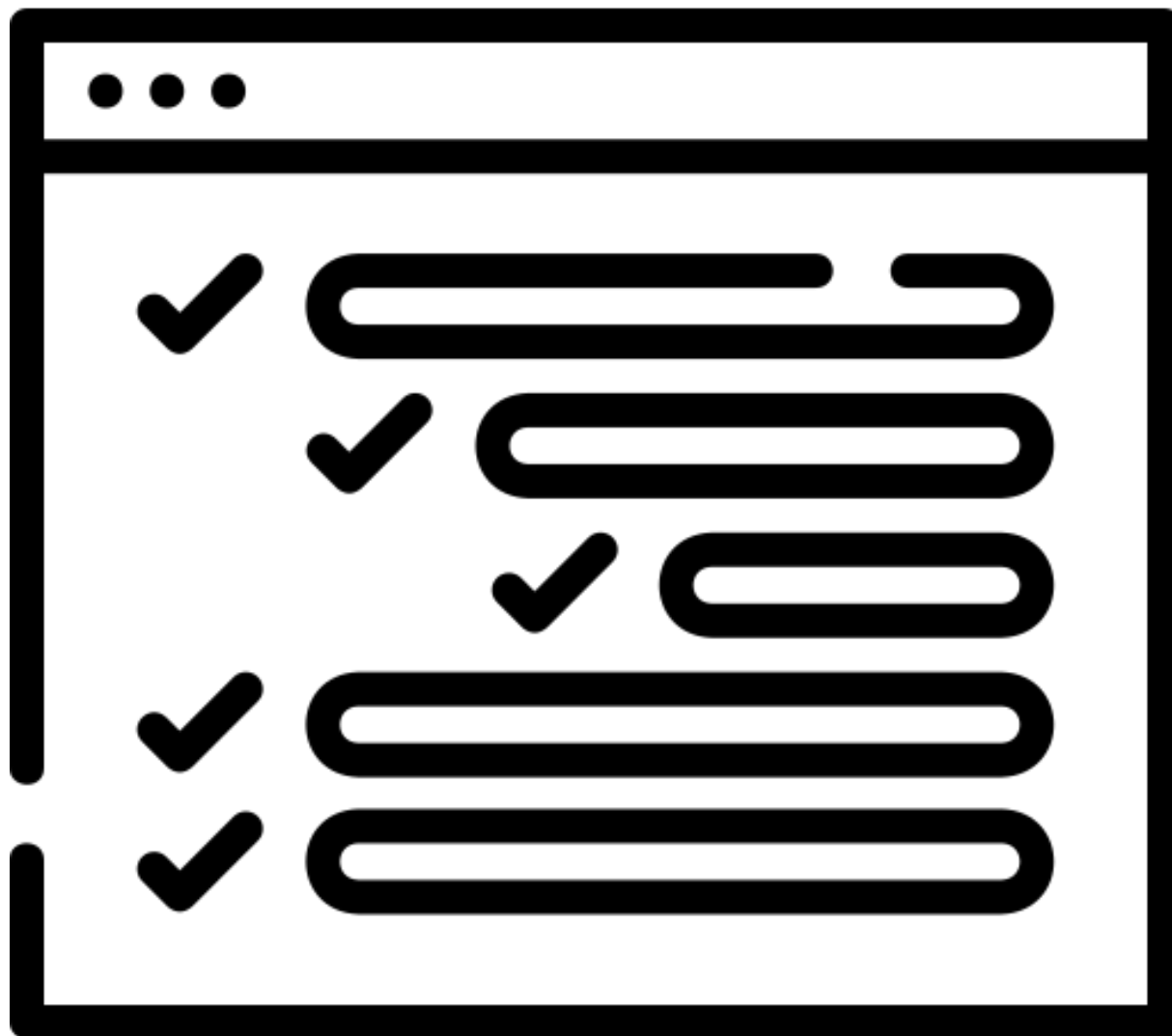
We aim to predict the secure loan before application is approved by implement machine learning algorithms. In addition, this model can be considered by lending companies to provide the amount of interest, the loan term and the loan period



Business Matrix

Loan Status

Table of Contents



- 1 Exploratory Data Analysis
- 2 Data Preprocessing
- 3 Modeling
- 4 Business Recommendation

Customer Classification Status

Good Borrower
Fully paid

Bad Borrower
Charge Off, Default,
Late (16-120 days)

Ambigus
Current, In grace period

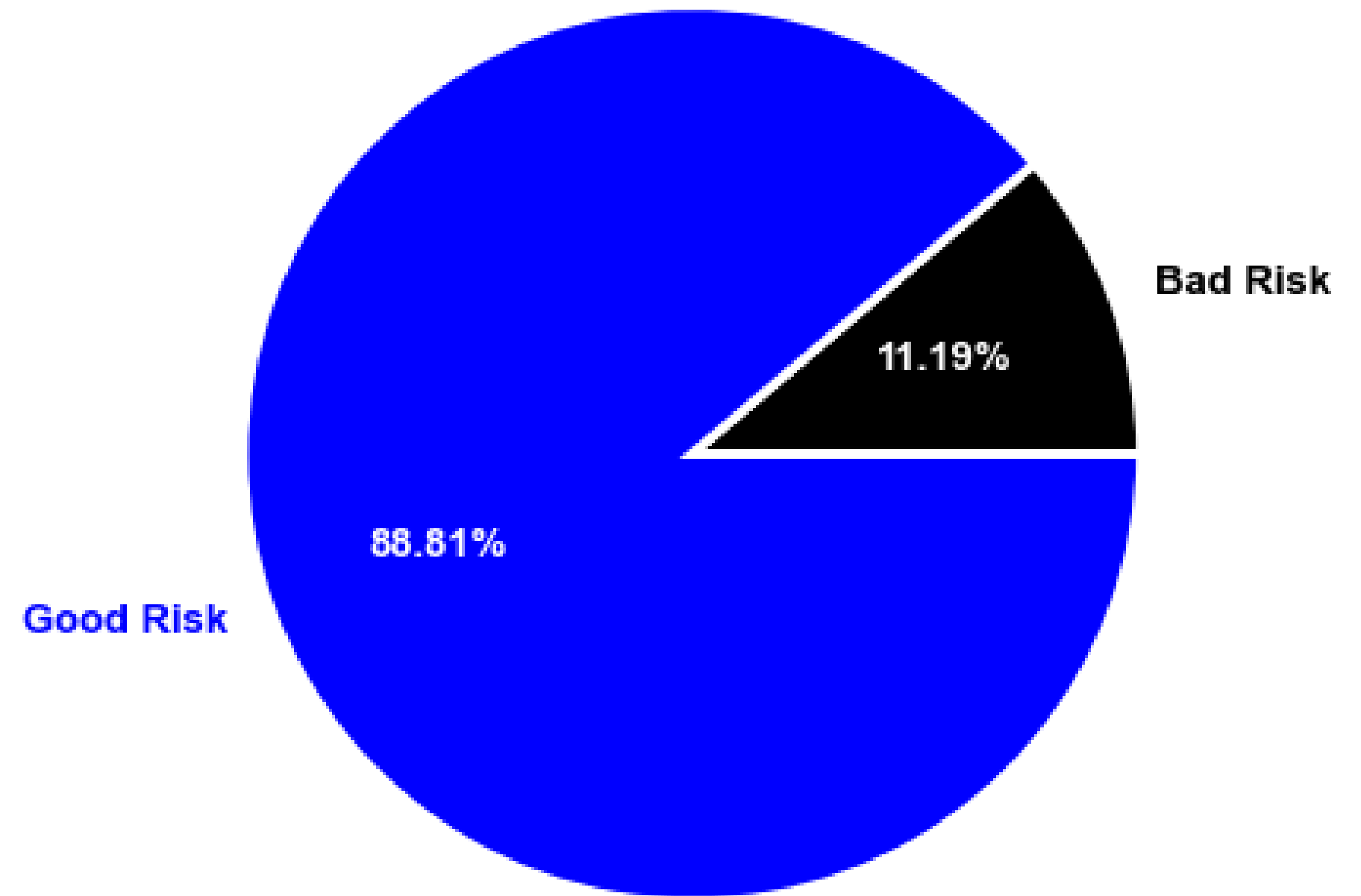


The loan's problem

We get the imbalance data which have 88,81% good borrower and 11,19% bad borrower from 2007 until 2014



Loan Status



The Number of Applicants by Loan Status

There are about 48% applicants with loan status of Current, followed by loan status of Fully Paid with 39.6%.



Status of Loan

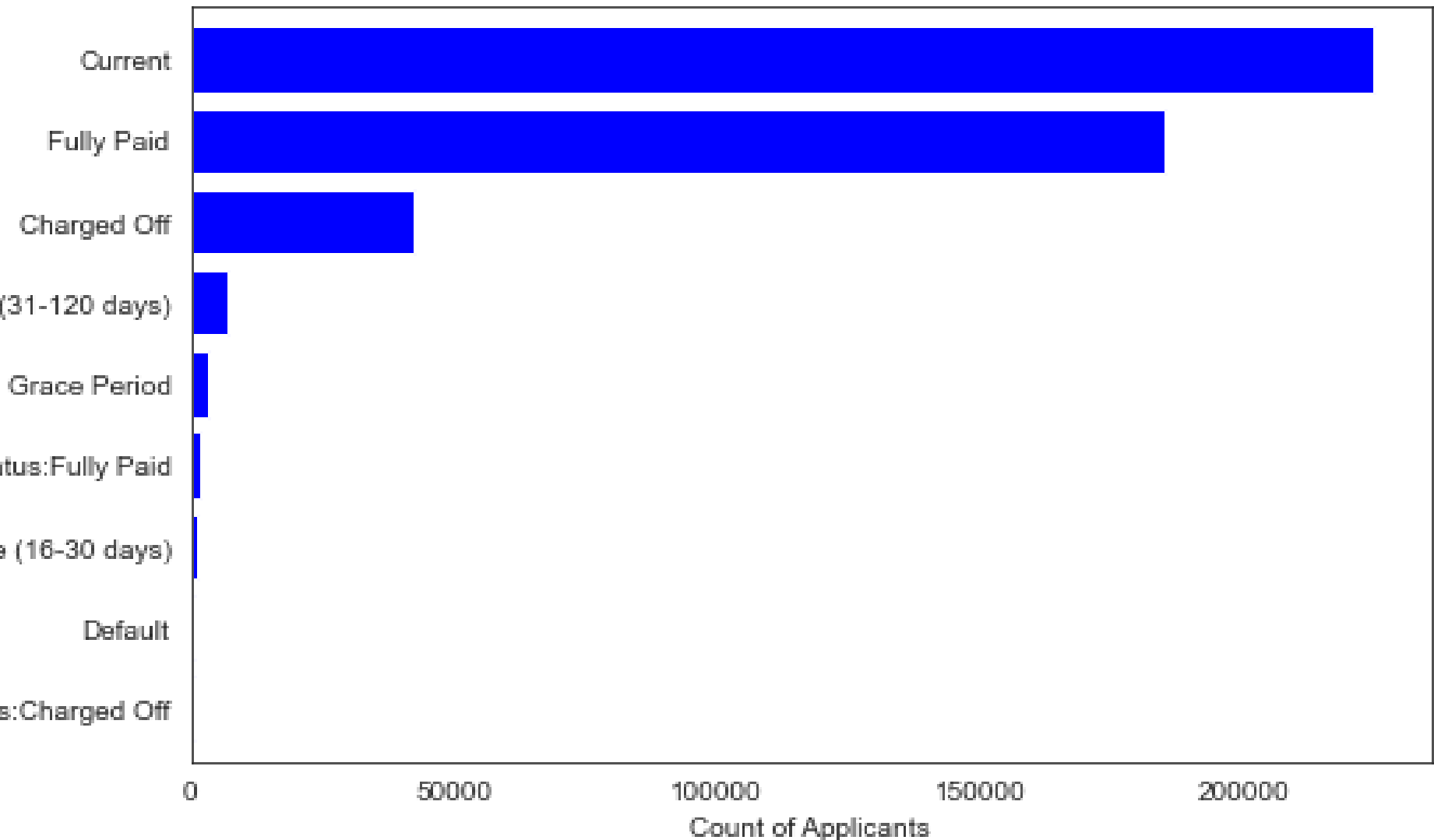
Does not meet the credit policy. Status:Fully Paid

Late (16-30 days)

Default

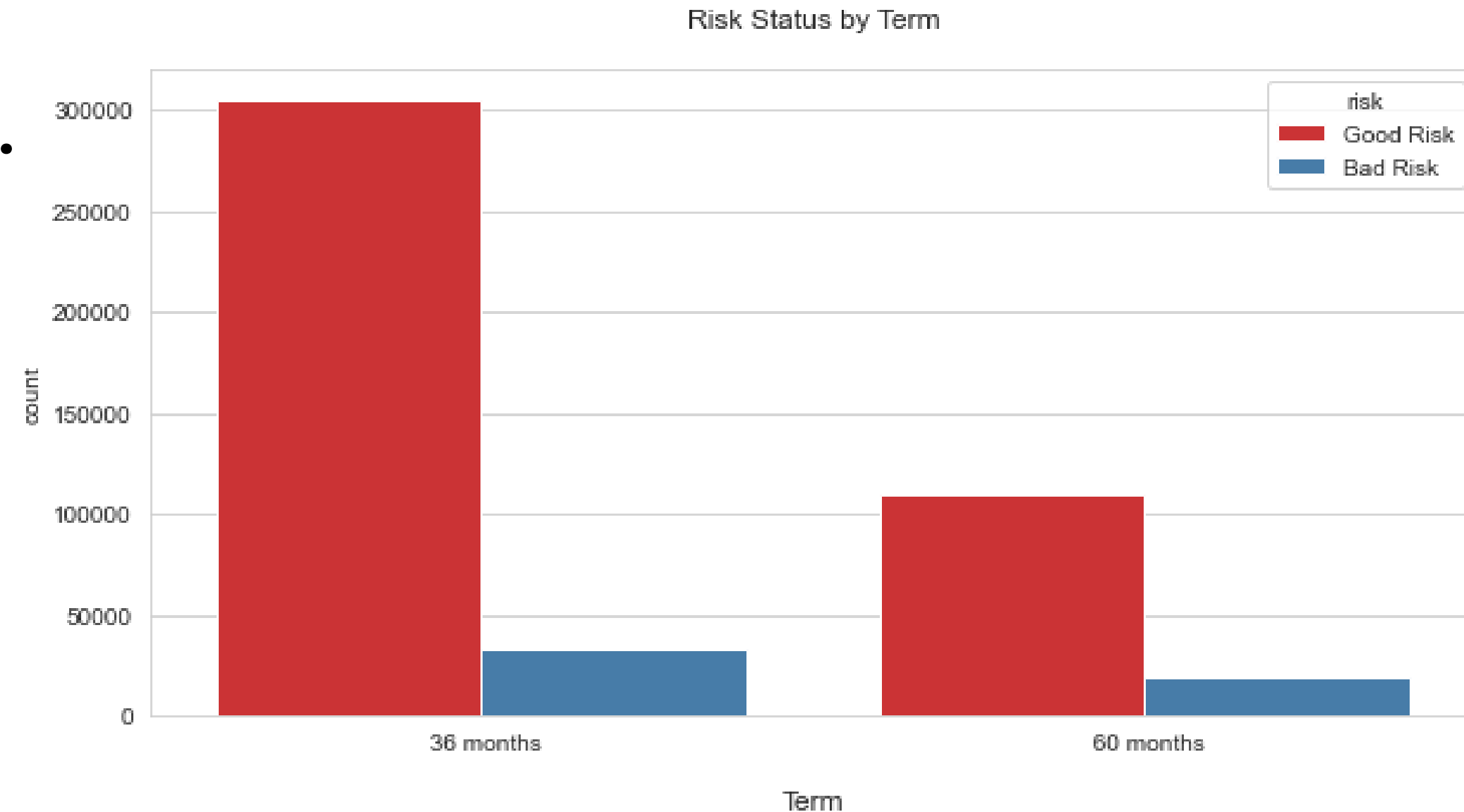
Does not meet the credit policy. Status:Charged Off

The Number of Applicants based on Loan Status



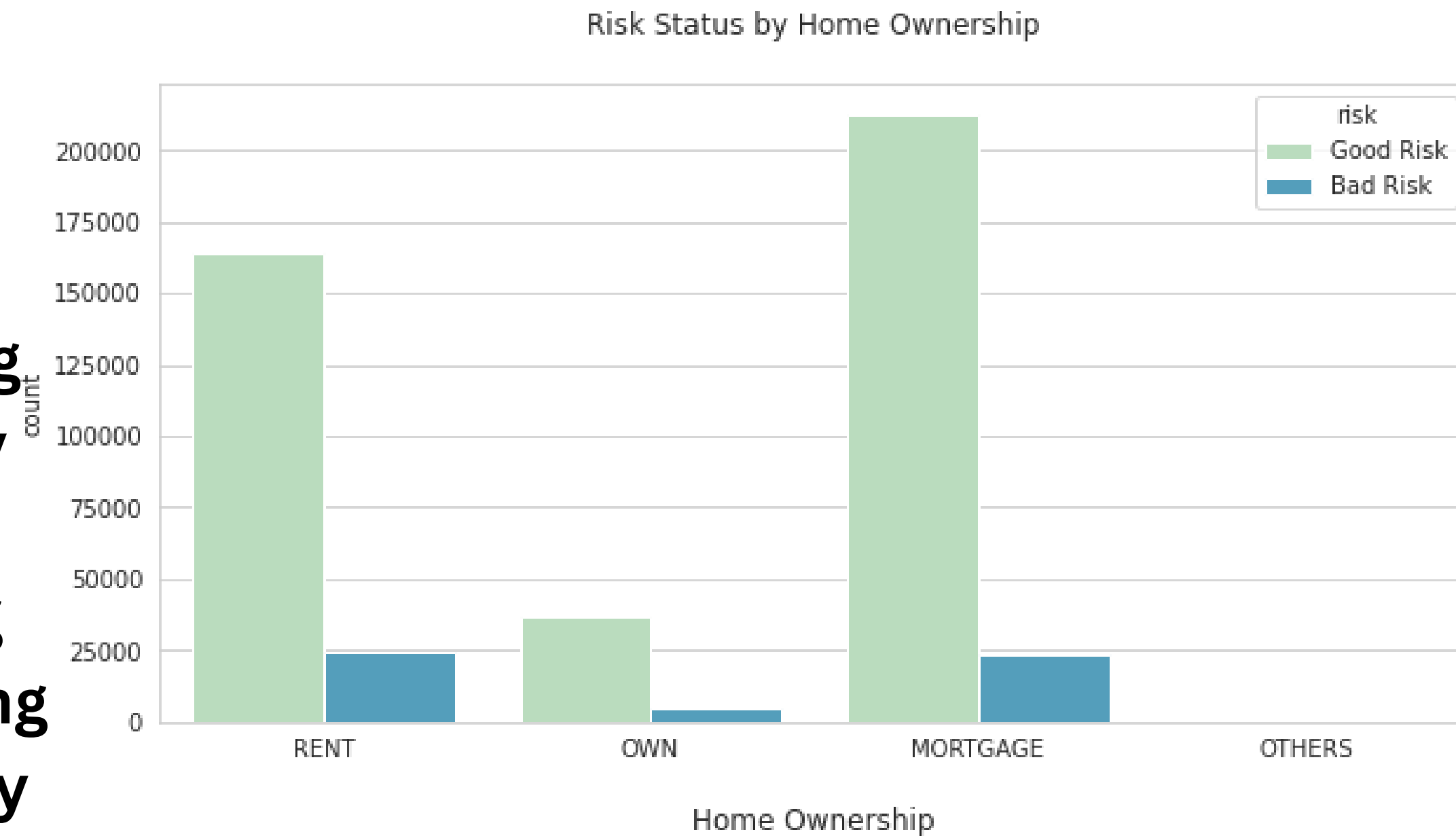
Risk Status by Term

- Loan term tell us about the number of payments on the loan.
- There are only two types of loan terms, either 36 months or 60 months. Most of the loans (73%) are shorter, with a term of 36 months.
- Loans with 36 months period are almost twice as likely to bad risk as loans with 60 months period.



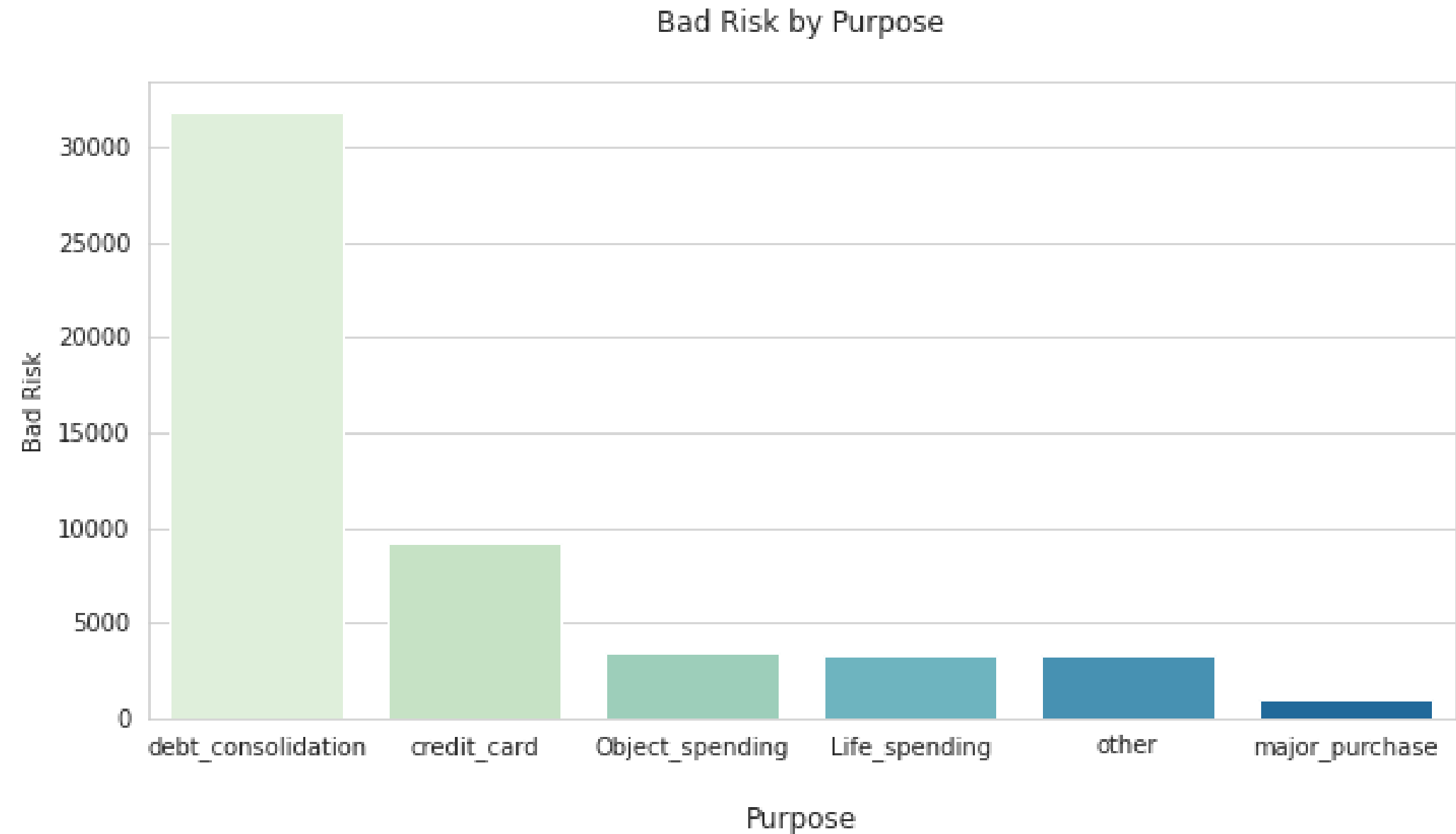
Risk Status by Home Ownership

- The home ownership feature is category provided by the applicant's during registration.
- Most applicants have an existing mortgage (50%) or are currently renting a home (40%).
- Applicants who have an existing mortgage or are currently renting a home have a higher probability of bad risk.



Loan Status by Purpose

- Most of Customer that type Bad Risk is in debt_consolidation 6.8%
- For the second is credit_card
- This shows that for debt reasons, the customer is more likely to be the type of customer that is Bad

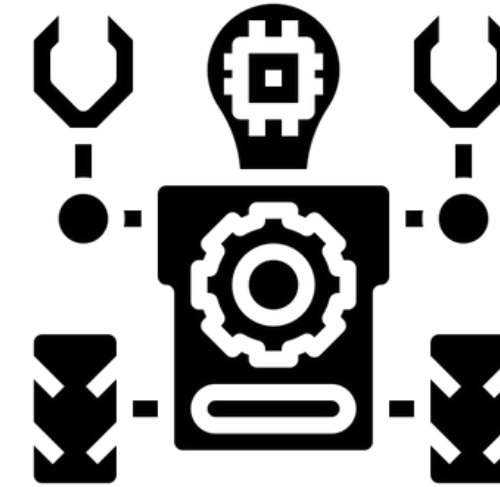


Data Preprocessing



Data Cleansing

- Drop feature with null value $>50\%$
- Impute feature with missing value $<20\%$ with mode & median
- Drop unnecessary feature & feature with 1 unique value
- Drop high correlation feature (>0.7) with target



Feature Engineering

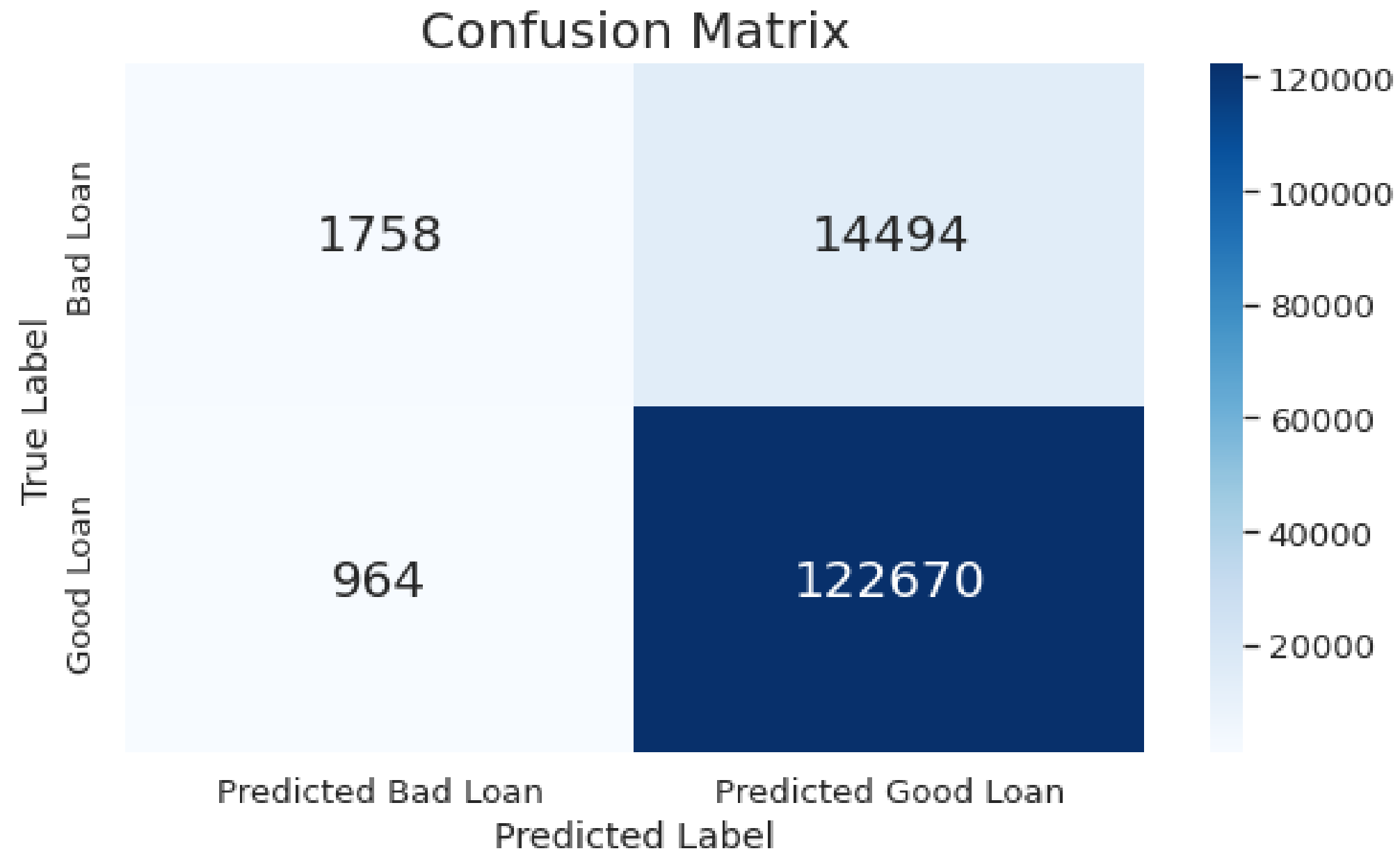
- Change datatype for some feature with incorrect datatype
- Feature selection using weight of evidence and information value
- Encode all features for modeling with label encoding and one hot encoding

Modeling & Evaluation

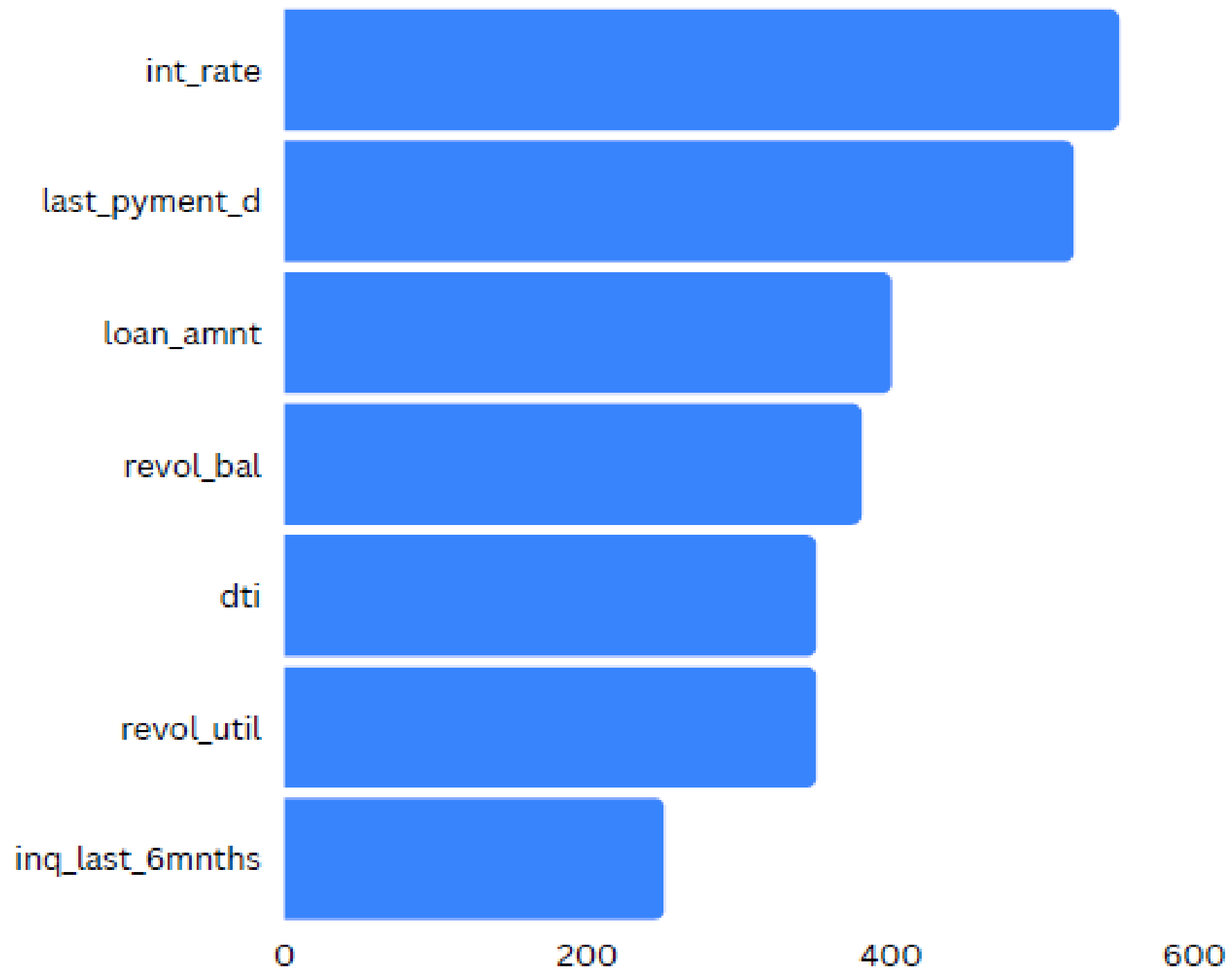
Model	Accuracy	Recall	Precision	AUC
LightGBM	0.889	0.992	0.894	0.859
XGBoost	0.888	0.993	0.982	0.855
Gradient Boosting	0.888	0.992	0.983	0.855
RandomForest	0.887	0.986	0.897	0.847
KNN	0.873	0.983	0.885	0.551

- Metrics evaluation that important for this model is AUC Score
- Best model is LightGBM with the best score & short duration

Modeling & Evaluation



Feature Importance



int_rate

Interest Rate on the loan

revol_bal

Total credit revolving balance

loan_amnt

The listed amount of the loan applied for by the borrower

revol_util

Revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit.

last_pyment_d

Last month payment was received

dti

A ratio calculated using the borrower's total monthly debt payments on the total debt obligations

inq_last_6mnths

The number of inquiries in past 6 months (excluding auto and mortgage inquiries)

Simulation

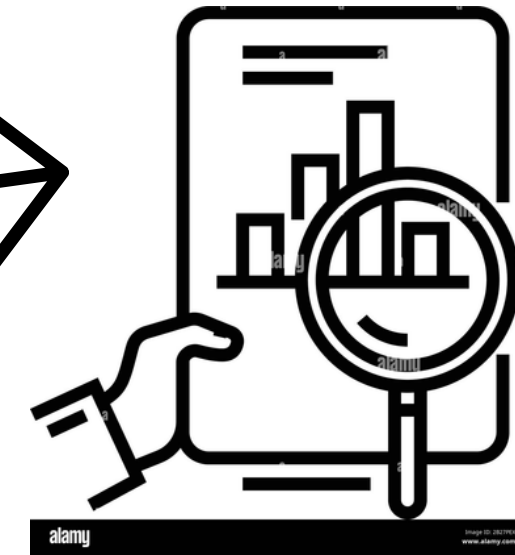
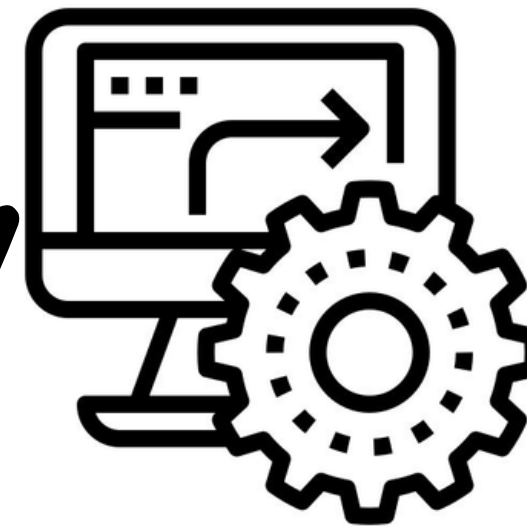


Bank Partners

Credit Loan



Automate Machine Learning
Model



Results

- Bad Loan decrease 21.8% >> 2.4%
- Suppress the loss due to error determining credit risk

Business Recommendation



Loan Amount

It is important to carefully evaluate the borrower's creditworthiness and the potential risks associated with the loan amount to ensure that the loan is viable and sustainable for both the lender and the borrower. This may include conducting a thorough credit risk assessment, implementing a risk-based pricing strategy.

Feature Importance

Focusing on feature importance could be a valuable business opportunity. By identifying the most important features and building a focused and effective model, this could help to reduce credit risk and make more informed lending decisions

Interest Rate

It would be beneficial to offer flexible interest rates that can be adjusted based on the borrower's creditworthiness and payment history. This allows for a more personalized approach to lending and can help the lender manage their credit risk more effectively.



Thank you

Portfolio link

CLICK HERE

