



partners



# CREDIT RISK LOAN PREDICTION



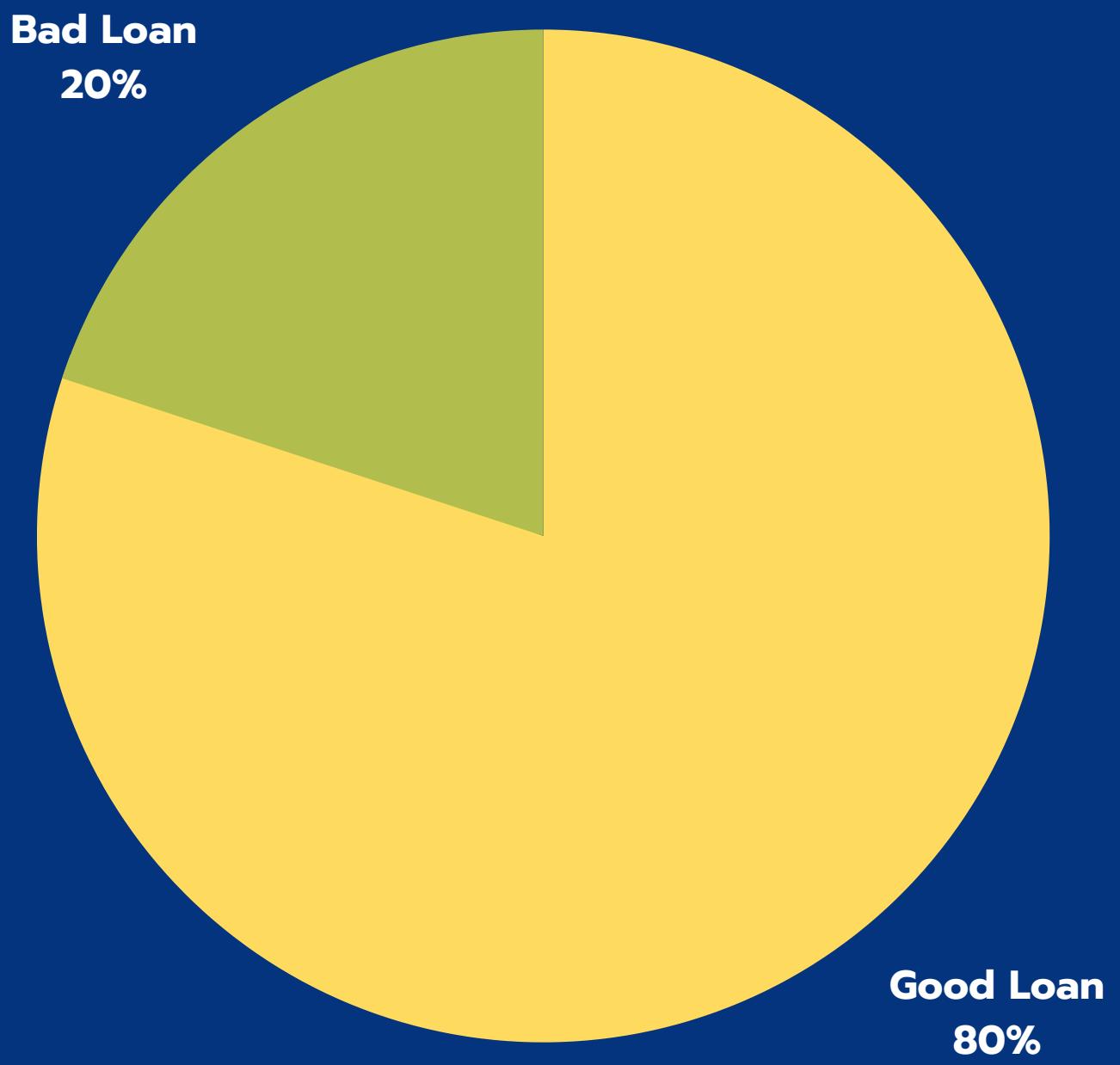
Gustav Bagus Samanta

# Problem and Goals

Company has recorded a bad loan rate at 20%. 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. High bad loan rate causes a company loss.

The objective is to make a machine learning model which able to predict that loan will become a bad loan or good loan, we aim to predict the secure loan before application is approved by machine learning algorithms.

.....



# Table of Contents

**Data Exploration**

**Data Preparation**

**Modelling**

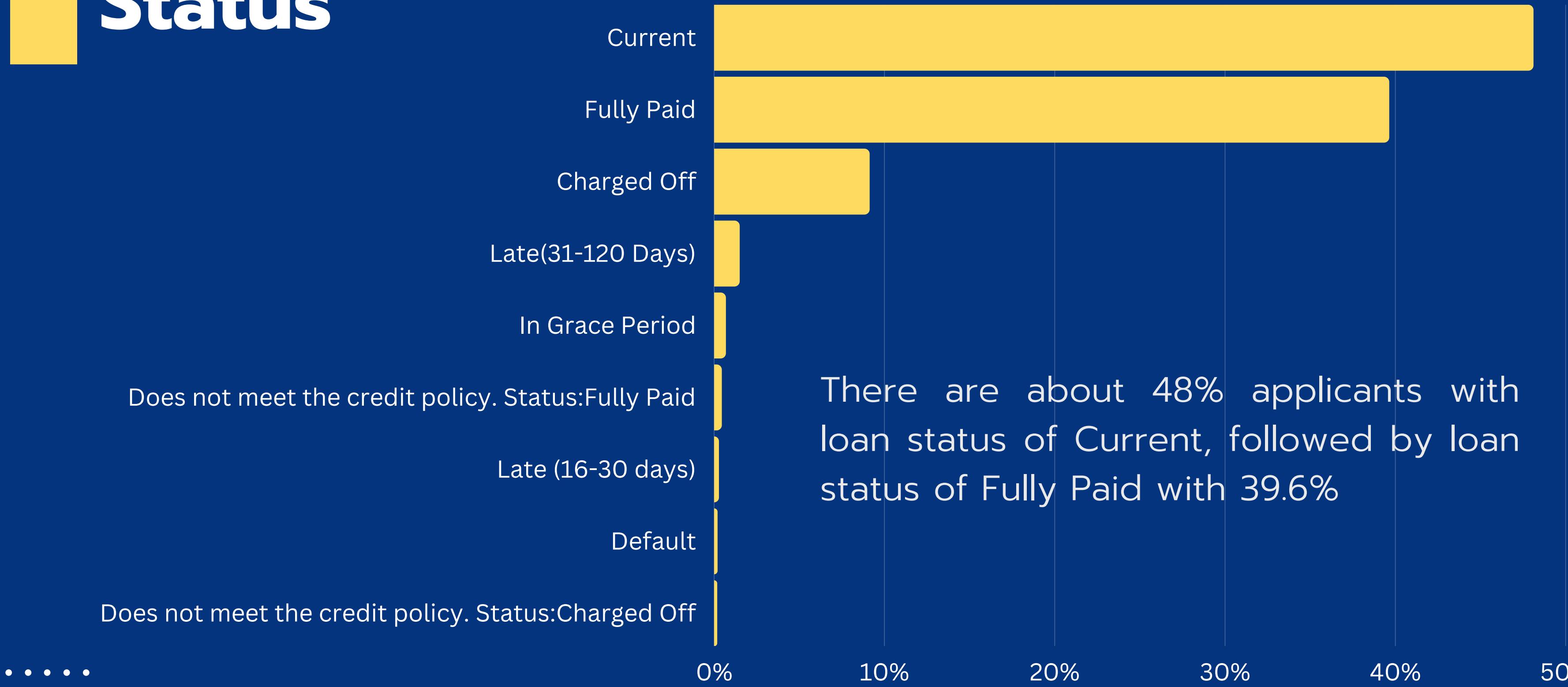
**Business Recommendation**

# Data Exploration

The data that we use is loan record dataset from 2007 until 2014. The datasets are consist of 466,285 rows and 75 columns.



# Loan Status



# Data Preparation

## Cleansing

- Drop feature with null value >25%
- Impute feature with missing value with median
- Drop unnecessary feature



## Feature Engineering

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

# Modelling and Result

Metrics evaluation that important for this model is AUC Score.

We use Logistic Regression, the table beside shows that the trained model has a good performance (AUC > 85%), which is good enough to be implemented.

.....

Performance	
Metrics	Train set
Accuracy	<b>0.982711931</b>
Recall	<b>0.997726535</b>
Precision	<b>0.980999521</b>
AUC	<b>0.991884988</b>

## **Business Recommendation**

**It is important to carefully evaluate the borrower's creditworthiness and the potential risks associated with the loan amount**



# Thank you

Reach me at

[gustavsmnt@gmail.com](mailto:gustavsmnt@gmail.com)

<https://www.linkedin.com/in/gustavsmnt/>



**idx** partners