

0 0 0

PRESENTATION

CREDIT APPLICATION PREDICTION

LINK TO GITHUB:

[HTTPS://GITHUB.COM/BRYANT05/VIX_HCI](https://github.com/Bryant05/VIX_HCI)

BRYAN TAMIN

PROBLEM

Problem Statement

The huge amount of applications for Credit at Home Credit Indonesia (HCI) has made the evaluation of the applications very difficult. Even sometimes the decision that has been made is incorrect which made Home Credit Indonesia suffer losses because the customers defaulted.

From the data Home Credit Indonesia has suffered Rp. 13.846.851.949 losses from customers that failed to pay their debt which covers 8.07% of total customers. The reason for this is the wrong decision when reviewing the credit applications.

Role

As a Data Scientist Intern at Home Credit Indonesia, I will analyse the data and make machine learning model which can predict credit application correctly and give insight to prevent losses and increase the profit of Home Credit Indonesia.

Objective

Make a machine learning model which can predict whether credit applications should be accepted or not so that it can be a guide when evaluating the applications.

Business Metrics

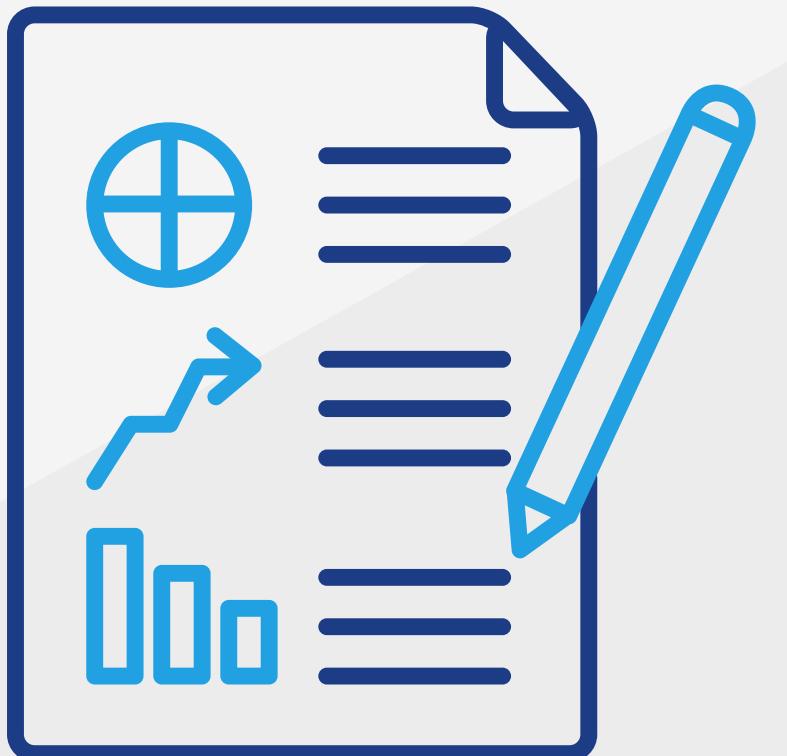
Default Rate

o o o o

DATASET

The dataset contains historical data and transaction of credit application at Home Credit Indonesia

o o o o



The Dataset Consists of 7 Tables:

- **application**
Loan and loan applicant info
- **bureau**
Application data from other institutions
- **bureau_balance**
monthly balance info in Credit Bureau
- **previous_application**
Previous application in Home Credit
- **POS_CASH_balance**
monthly balance info in Home Credit
- **Instalments_payments**
Past payment data in Home Credit
- **credit_card_balance**
Monthly Balance of previous application in Home Credit

Amount Credit

The bigger the amount of credit customers apply for,
The Less Chance They Will Default



INSIGHT

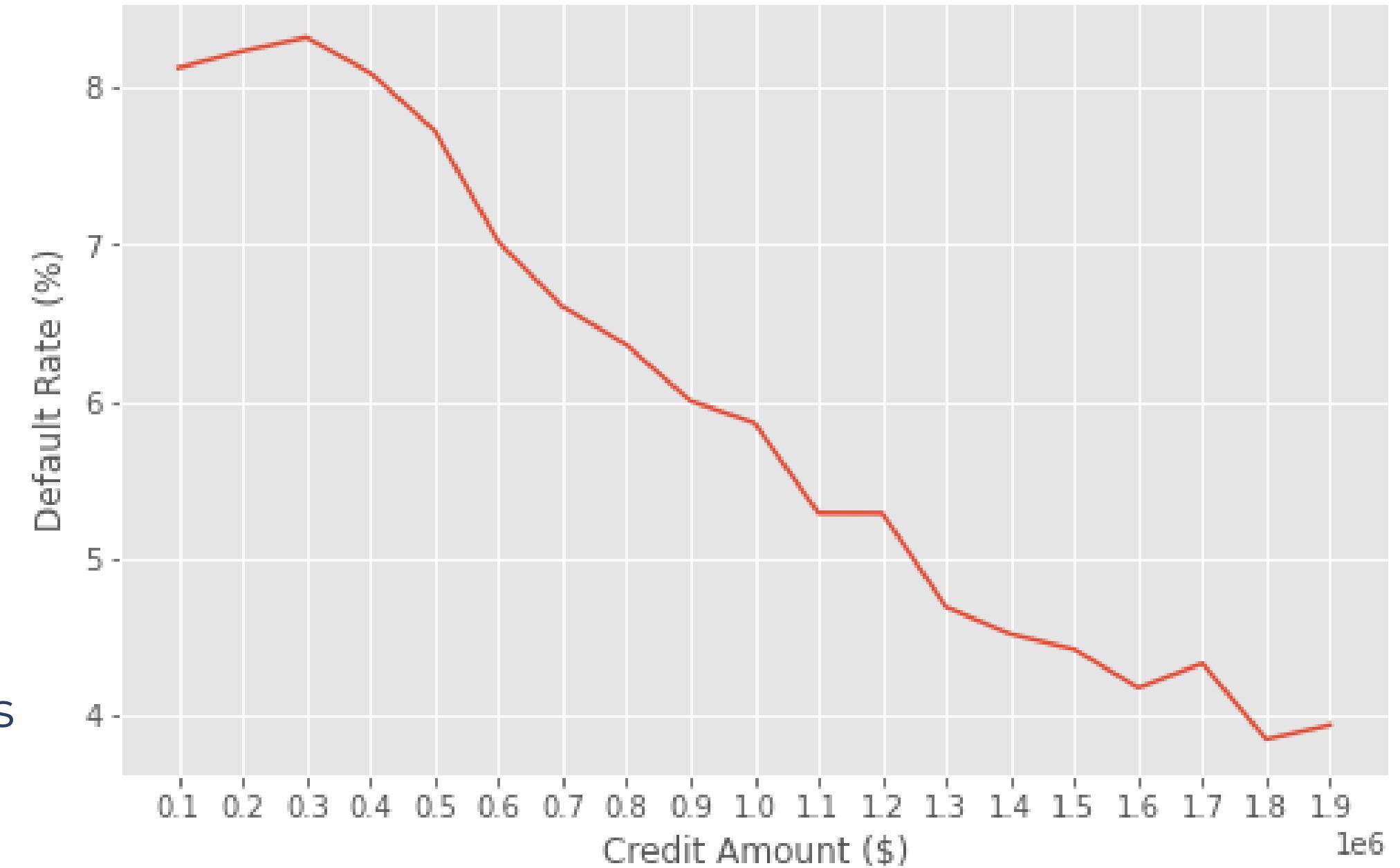


Age

The Older The Customers, The Less Chance They Will Default

CREDIT AMOUNT

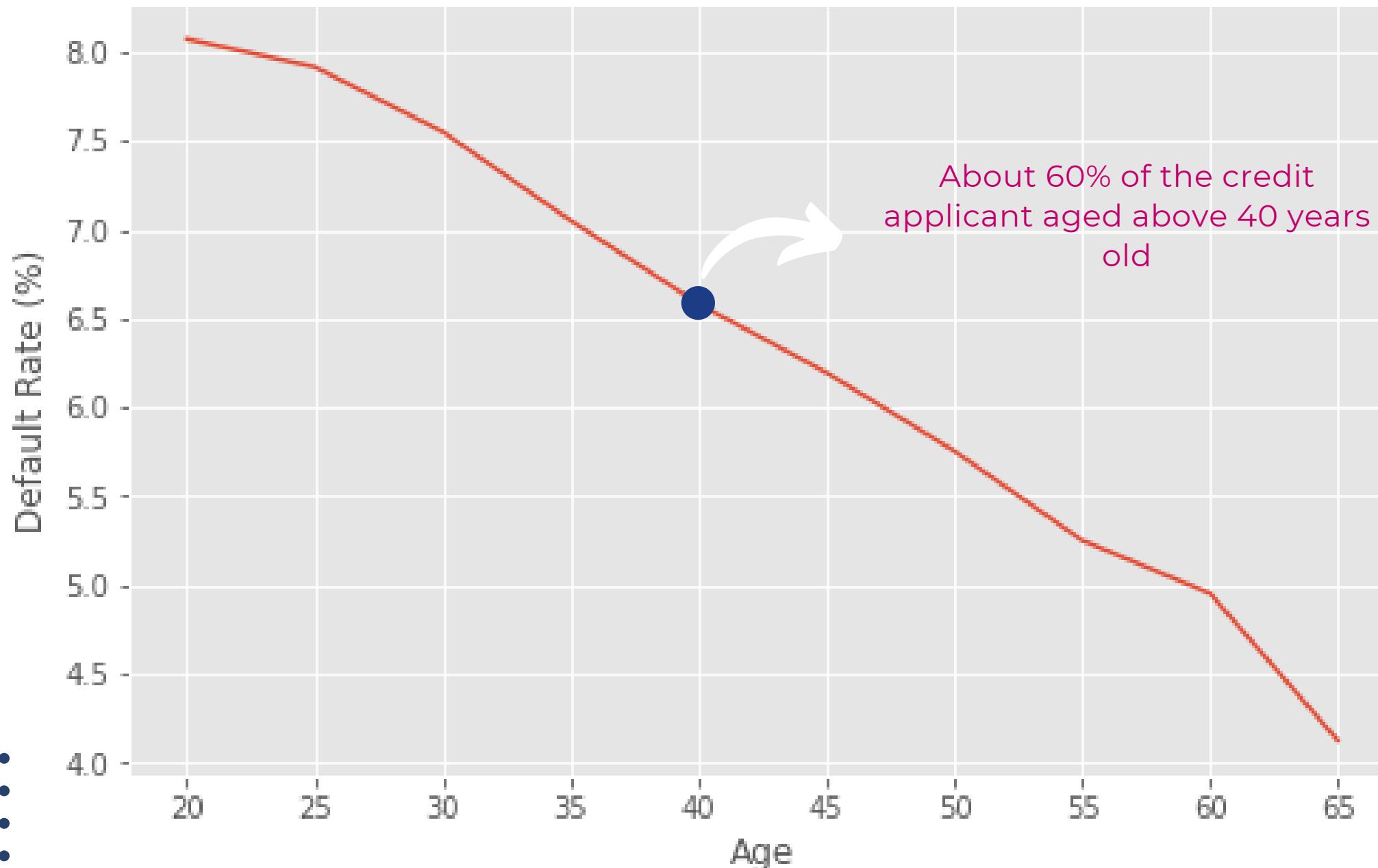
When The Amount of Credit Requested is Above 1 Million, Customer Default Happens Less



Hence, create a campaign or promote credit applications with a credit amount above 1 million to increase revenue and prevent default customer

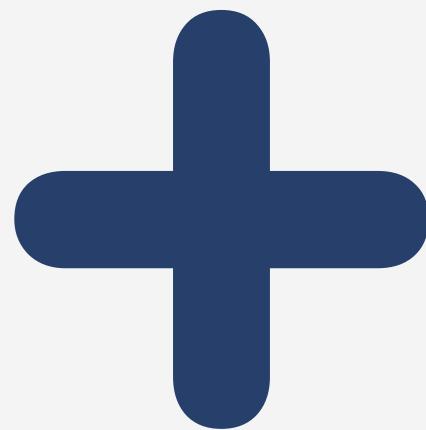
AGE

The Older The
Customers, The Less
Chance They Will
Default



Therefore, create a campaign or promotion for credit application with the applicant aged above 40 years old to increase revenue and prevent default customer

PREPROCESSING PATH



Handle Missing Value
Fill in and drop some of the values in the data

Feature Selection
Select the most important feature from huge number of features

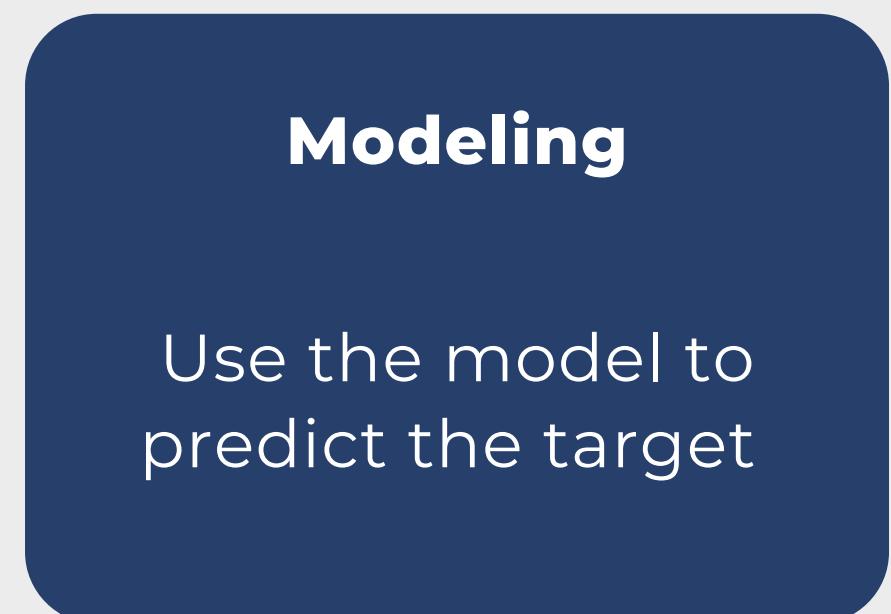
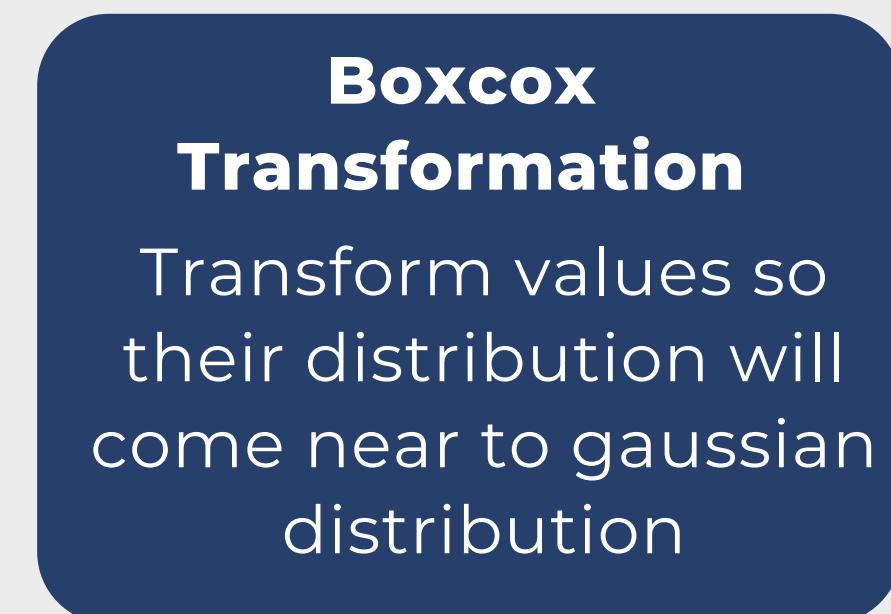
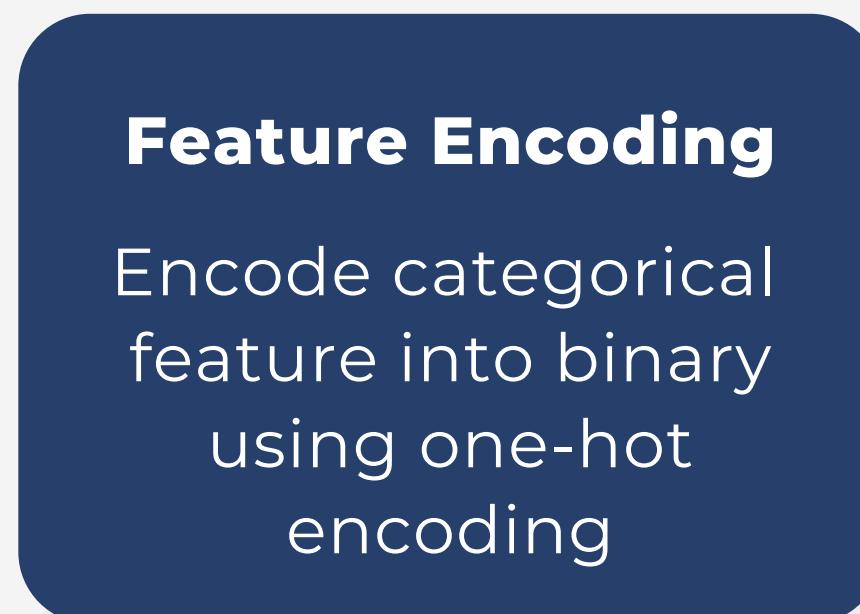
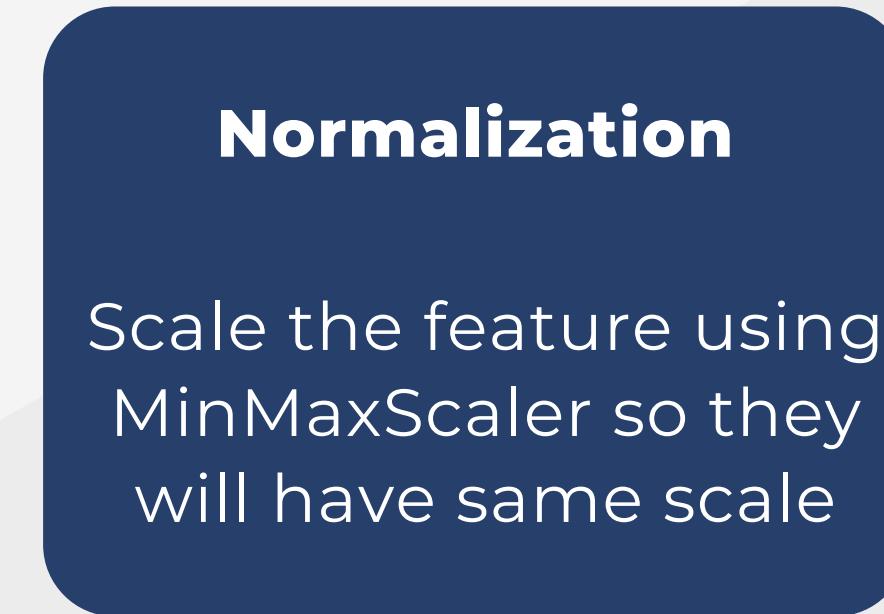
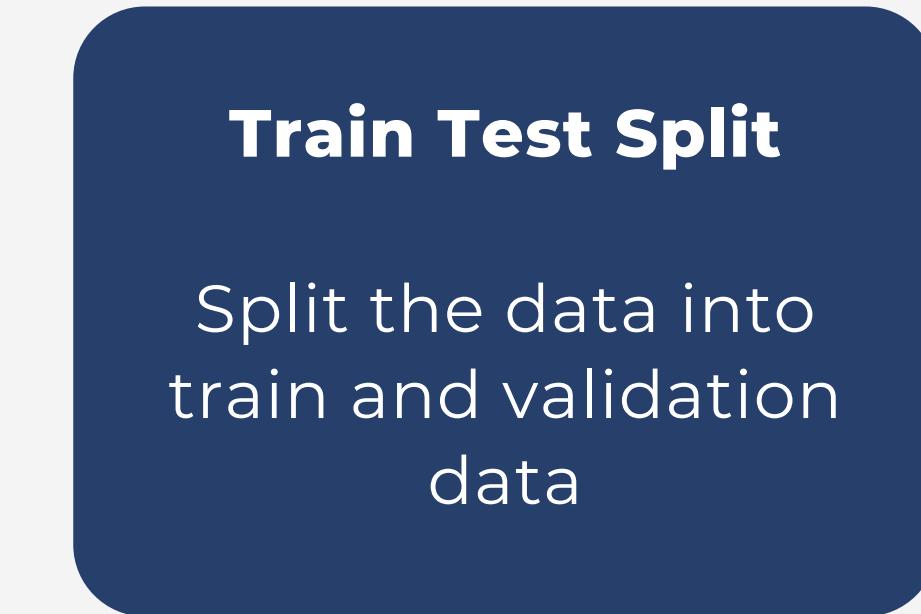
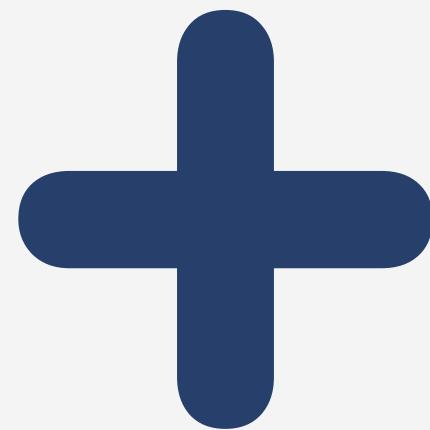
Outlier Handling
Handle outlier using Z-score

Merge Data
Merge the data from one table with another

Feature Engineering
Create new features from the existing ones and fix mismatch value



PREPROCESSING PATH



MACHINE LEARNING

Evaluation Metrics

Precision and
ROC_AUC

Tested Models

- Logistic Regression
- Decision Tree
- Random Forest
- XGBoost
- CatBoost

Best Models

Decision Tree

Recall: 0.68

ROC_AUC: 0.66

BUSINESS RECOMENDATION

Create a campaign or promote credit applications with a credit amount above 1 million



Default rate:
8.07% --> 5.87%

Create a campaign or promotion for credit application with the applicant aged above 40 years old



Default rate:
8.07% --> 6.58%

Use machine learning model prediction for guide



Default rate:
8.07% --> 2,58%

With the recall score 0.68 means the model can predict 68% of the applicant that will default. If used properly It can reduce the default rate by 68%