Credit Risk Management

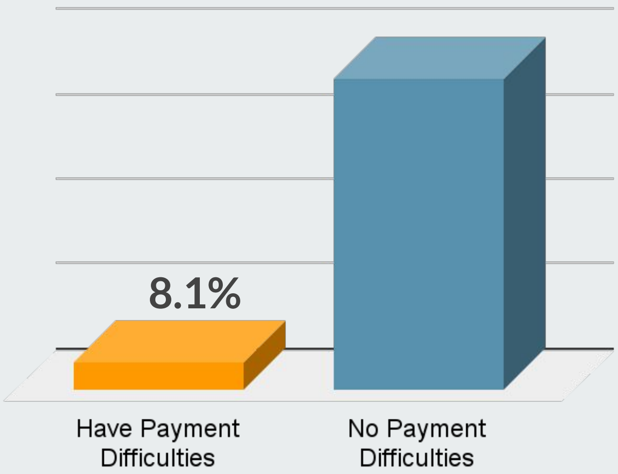
Maximizing Profitability Through the Scorecard Model

Dina Ramadhani

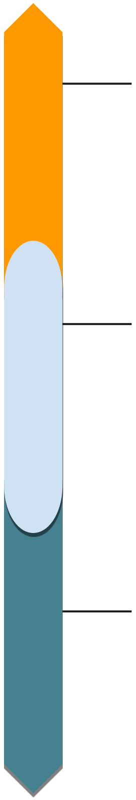
BUSINESS PROBLEM

HOME CREDIT
Kamu Bisa!

Percentage of Applicants



The company lost $ 13,847M due to 8.1% default credits



Market Expansion

High Risk, High Reward
Credit score - Low

Profit Maximization

Credit score - Optimum

"**Customer-centric** companies are **60% more profitable**."
– Deloitte

Risk Minimization

Low Risk, Low Reward
Credit score - High

FEATURE ENGINEERING

HOME CREDIT
Kamu Bisa!

DATASET

Application Train

Previous Application

Bureau

Pos Cash Balance

1 Application Train

- ANNUITY_PER_INCOME
- AGE
- YEARS_EMPLOYED
- YEARS_REGISTRATION

2 Previous Application

- P_AVG_REFUSED_AMT_CREDIT
- P_PREV_APP_COUNT
- P_REFUSED_COUNT
- P_REFUSED_INTEREST_RATE

3 Bureau

- B_AVG_AMT_CREDIT_SUM
- B_BUREAU_COUNT
- B_BAD_DEBT_COUNT

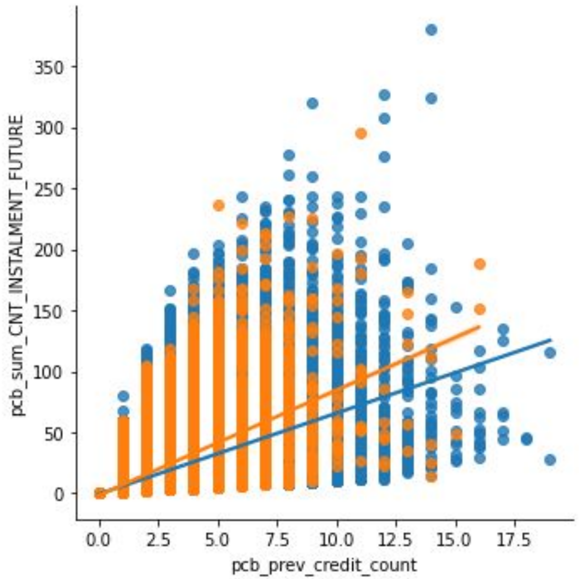
4 Pos Cash Balance

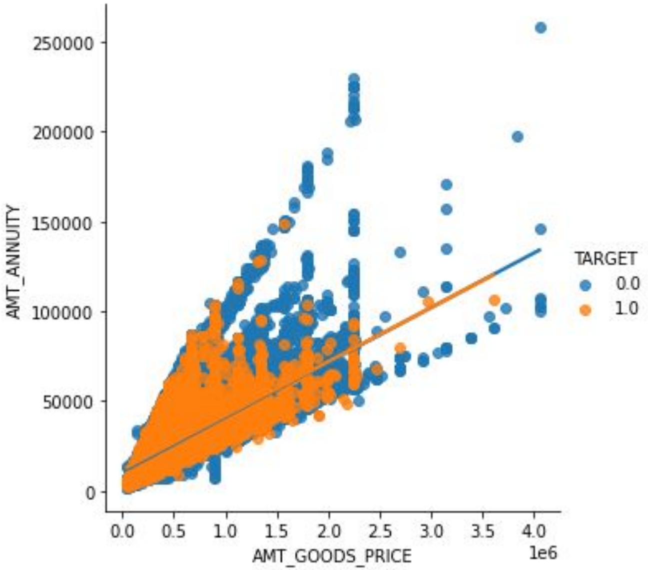
- PCB_SUM_CNT_INSTALLMENT_FUTURE
- PCB_PREV_CREDIT_COUNT

❖ **Description:** Each row represent client demographics, current and previous application, credit bureau record, and installment history.

❖ **Dataset shape:** 307,511 unique clients' ID and 135 features (after feature engineering)

INSIGHTS

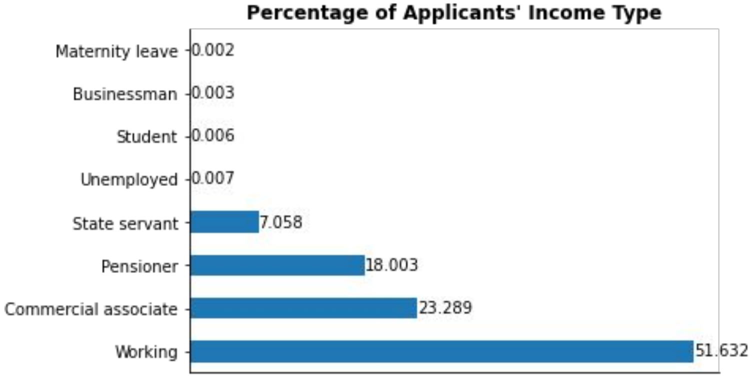


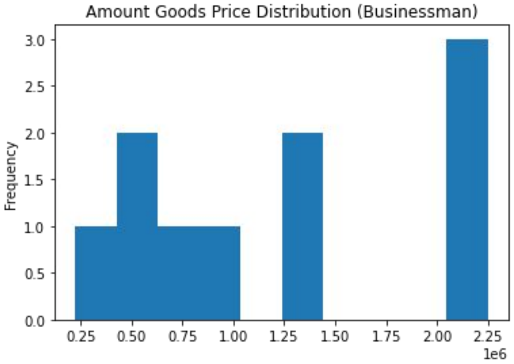


As the amount of goods price increase, the amount of annuity becomes higher. With the same number of credit, the **clients with loan repayment difficulties** have the **higher number of installment** in the future. They tend to choose a longer loan payment duration so the **amount of annuity becomes easier to pay off**.

Action: We need to provide a longer loan payment duration option for the potential default credit application.

INSIGHTS

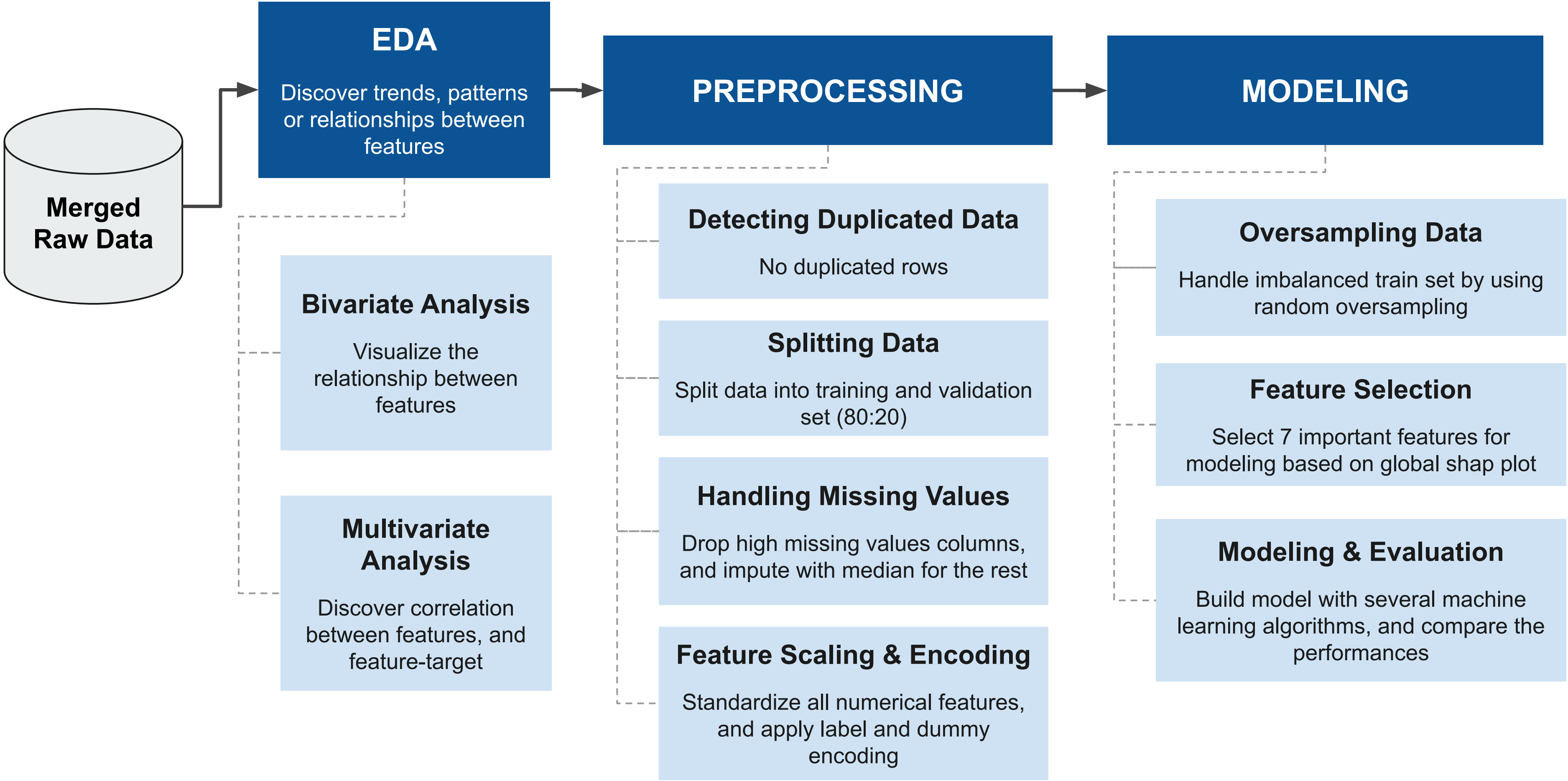




The client with income type **student and businessman** have no loan repayment difficulties. These segment have **accepted application rate 100%**, means they're actually a **promising clients**. However there are only 0.006% and 0.003% of application come from these segments respectively.

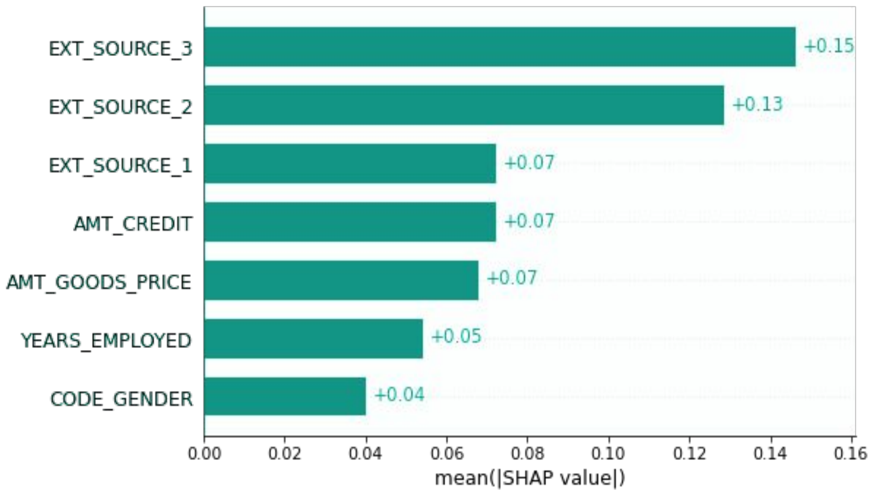
Action: We need to create **new loan product** that **targeting student and businessman** segment specifically. According to the amount of goods price distribution, we know that these segments **have a different kind of products they are interested in**.

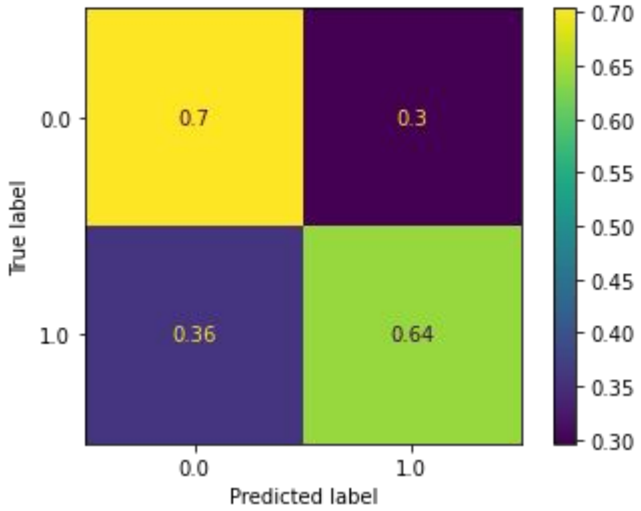
- **Student** : This segment interested in **low-valued products**. **50% of clients** used credit loan for goods at least **$ 452,250** in price. The average goods price is $ 458,250.
- **Businessman** : This segment interested in **high-valued products**. **50% of clients** used credit loan for goods at least **$ 1,125,000** in price or **2.5x higher than student segment**. The average goods price is $ 1,228,500.


Merged Raw Data
EDA
Discover trends, patterns or relationships between features
Bivariate Analysis
Visualize the relationship between features
Multivariate Analysis
Discover correlation between features, and feature-target
PREPROCESSING
Detecting Duplicated Data
No duplicated rows
Splitting Data
Split data into training and validation set (80:20)
Handling Missing Values
Drop high missing values columns, and impute with median for the rest
Feature Scaling & Encoding
Standardize all numerical features, and apply label and dummy encoding
MODELING
Oversampling Data
Handle imbalanced train set by using random oversampling
Feature Selection
Select 7 important features for modeling based on global shap plot
Modeling & Evaluation
Build model with several machine learning algorithms, and compare the performances

BEST MODEL: XGBOOST

HOME CREDIT
Kamu Bisa!





Performance Model

With classification threshold 0.5, the model is able to predict up to **64% of the existing potential default credits**, with the risk of losing promising credit applications due to predicting errors of 30%. It means for every 20 applications which are predicted as a potential default credit, there are 17 which are actually not potential default credits.

Accuracy (Train)	: 73%
Accuracy (Validation)	: 70%
Validation Recall (Class 0)	: 70%
Validation Recall (Class 1)	: 64%

HOME CREDIT
Kamu Bisa!

CREDIT RISK MANAGEMENT STRATEGY

	Profit Maximization	Market Expansion
% of Good Credit predicted correctly	92%	100%
% of Bad Credit predicted correctly	41%	0%
Probability Threshold for Approvals	0.2989	0.0105

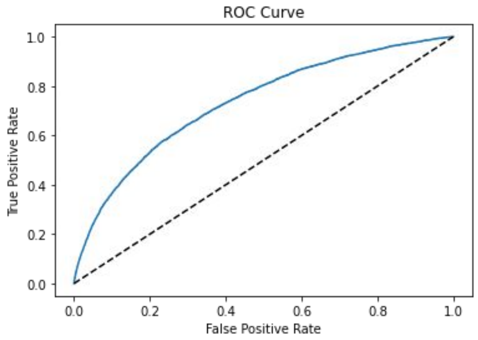
SCORECARD

		Sensitivity	Specificity	
Decile	Probability_Threshold	Cumm. Good %	Cumm. Bad Avoided %	Profit to Business
1	87.60%	11%	95%	805219015
2	80.18%	21%	89%	1583746297
3	74.04%	32%	86%	2402515261
4	68.07%	42%	82%	3208061520
5	62.14%	53%	76%	3965966786
6	55.34%	63%	70%	4725916293
7	48.24%	73%	65%	5474501170
8	40.03%	83%	56%	6148340180
9	29.89%	92%	41%	6720781697
10	1.05%	100%	0%	6814752557

*) Model simulation is applied to validation set

Assumptions

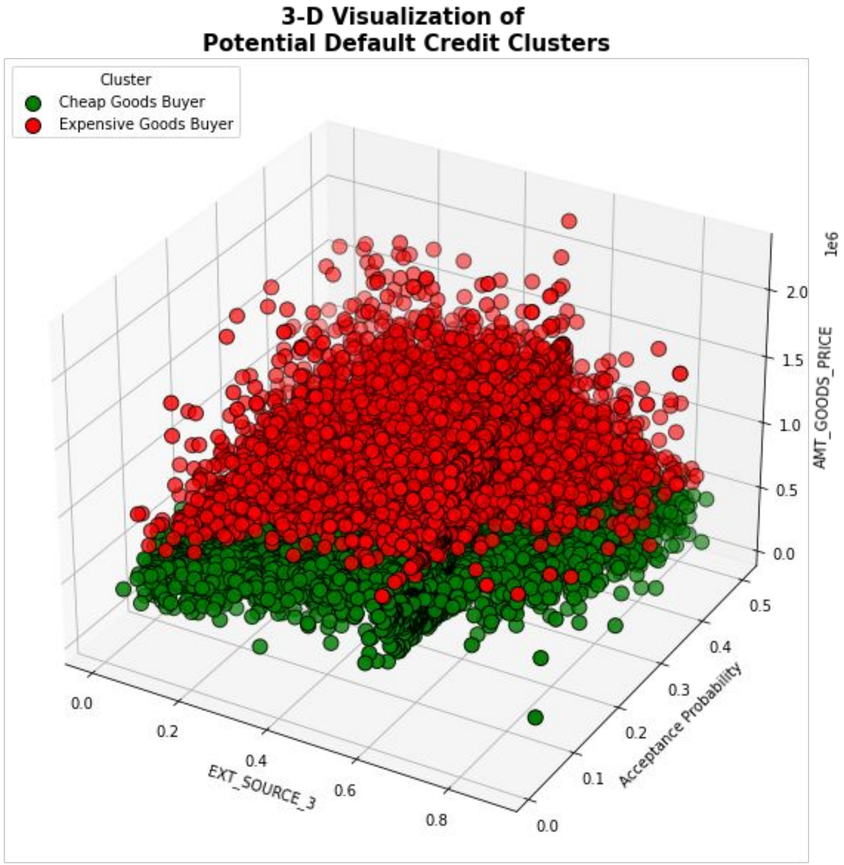
Loss from 1 bad credit : $ 557,778
Profit from 1 good credit : $ 144,636



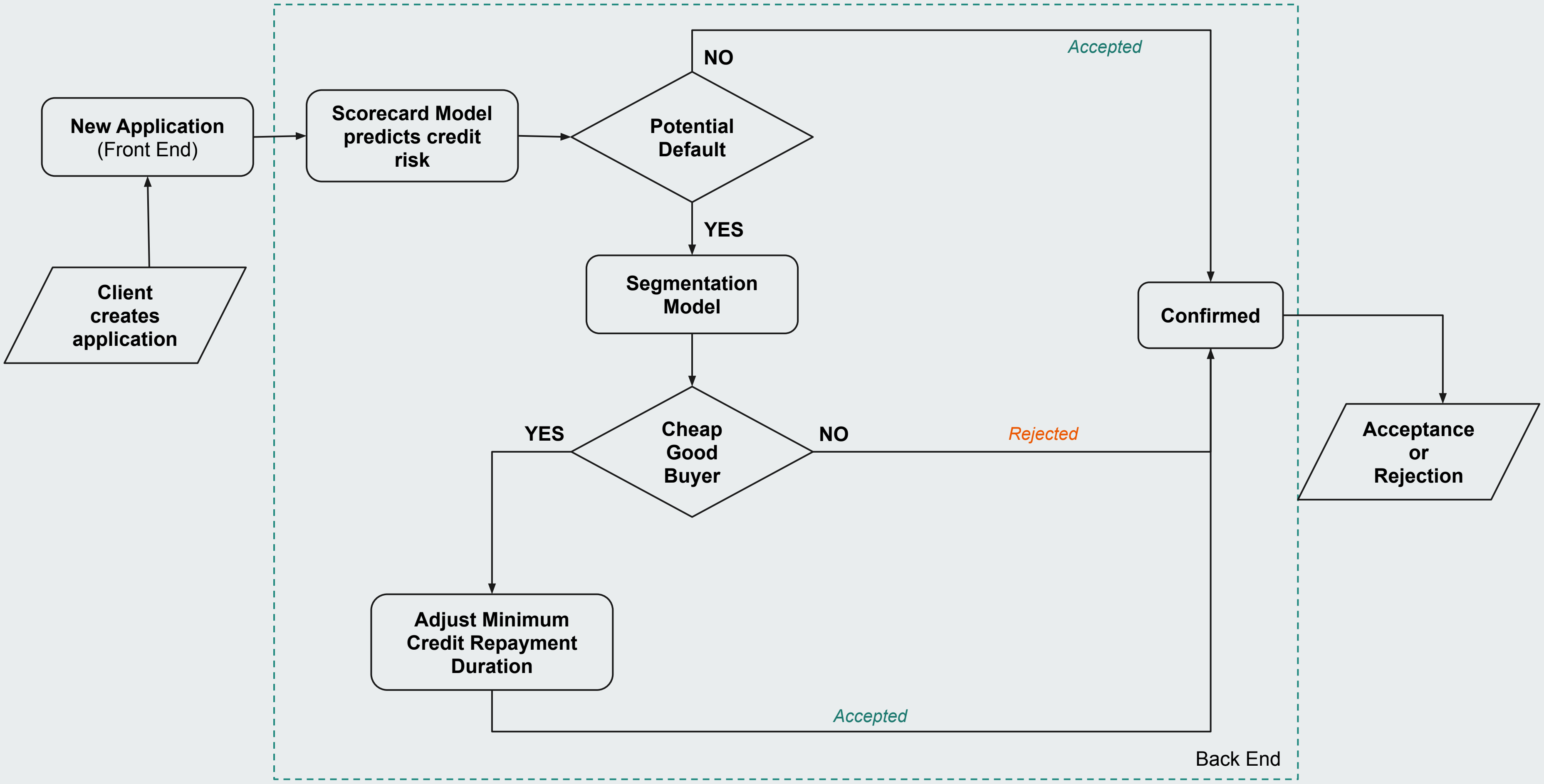
In average, Cheap Goods buyer segment (71.6%) applied for $ 322,396 credit, meanwhile **Expensive Goods buyer** segment (28.4%) applied for $ 818,797 credit or **2.5x higher**. It will cause default risk higher, since the credit annuity the clients must pay off periodically will also increase.

Actions:

1. Consider to **reject** credit application from clients who belong to **Expensive good buyer** segment.
2. **Adjust** minimum credit **repayment duration** for clients who belong to **Cheap goods buyer** segment, so they still can get credit and capable to repay it with light-amount annuity per period.



FLOWCHART


New Application
(Front End)
Client creates application
Scorecard Model predicts credit risk
Potential Default
NO
Accepted
YES
Segmentation Model
Confirmed
Cheap Good Buyer
YES
NO
Rejected
Acceptance or Rejection
Adjust Minimum Credit Repayment Duration
Accepted
Back End

POTENTIAL IMPACT

	Before	After	% Change
Loss From Default Credit	13,846,850,000	3,932,488,132	-71.6%

THANK YOU

HOME CREDIT
Kamu Bisa!

Created by:
Dina Ramadhani
dinachoirotul@gmail.com

Project Documentation

https://github.com/dinachoir/credit-risk-management