

Home Credit Scorecard Model

by AkhasRahmadeyan





Table of Contents



01

Research
Problem

02

Research
Method

03

Data
Preprocessing

04




Visualization and
Business Insight

05

Machine
Learning Model

06

Business
Recommendation



Research Problem



Data Source:

The data used is **application_train** as training data (with TARGET) and **application_test** as testing data (without TARGET).

Objective:

1. Identify characteristics of potential clients who will have difficulty repaying loans and who will not.
 2. Predict client's repayment abilities.
-

Project Background

Many people struggle to get loans due to insufficient or non-existent credit histories. Home Credit strives to broaden financial inclusion for the unbanked population by providing a positive and safe borrowing experience. In order to make sure this underserved population has a positive loan experience, Home Credit makes use of a variety of alternative data to predict their clients' repayment abilities. Doing so will ensure that clients capable of repayment are not rejected and that loans are given with a principal, maturity, and repayment calendar that will empower their clients to be successful.

Data Preprocessing

1. Exploratory Data Analysis (EDA)
2. Remove Duplicates
3. Handling Missing Value (If $<50\%$)
4. Detecting Outliers
5. Label Encoder
6. Feature Selection
7. Handling Imbalance Data

application_train
data

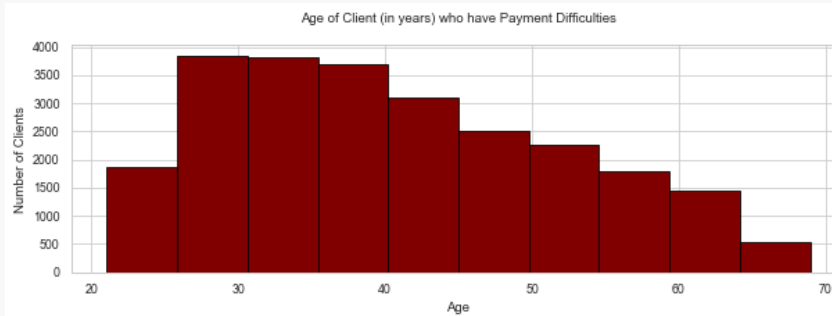
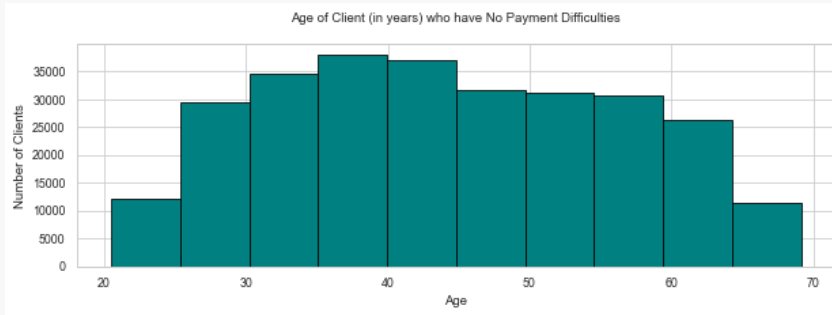
122 **307.511**
Features Record

application_test
data

122 **48.744**
Features Record



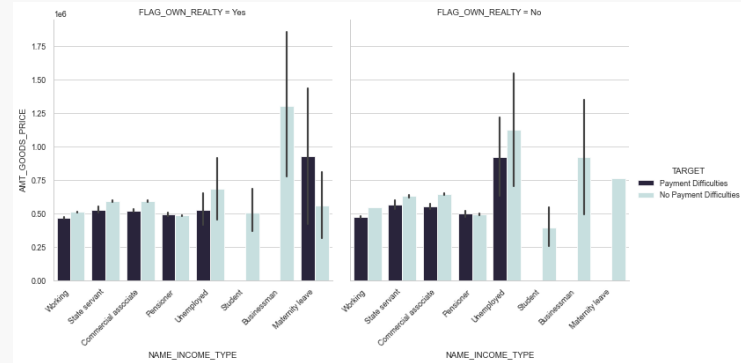
Business Insight (1)



- Most number of clients who apply for loans are in the range of 35-40 years.
- Meanwhile, the number of applicants for clients aged <25 or age >65 is very low
- Clients who have no payment difficulties are clients in the range of 35-45 years. You can target these clients as your priority.
- While clients who have payment difficulties are client the range of 25-35 years

Business Insight(2)

- All student clients have no difficulty repaying the loans whether with cash loan or revolving loan for a low to medium credit amount of the loan.
- All student clients have no difficulty repaying the loans whether with cash loan or revolving loan for a low to medium credit amount of the loan.
- For unemployed clients with cash loans, more than 50% of clients have problems repaying loans with medium credit amounts of the loan. While all unemployed clients with revolving loans have no difficulty repaying the loan.



Machine Learning Model (1)



Phase 1: Data Preparation

- **Feature Selection :**
Chi-Square
- **Feature Encoding :**
Label encoding

Phase 2: Handling Imbalance Data

- Oversampling
- Undersampling:

Phase 3: Training Model

- 80% Training
- 20% Testing
- Logistic Regression
- Decision Tree
- Random Forest
- Support Vector Machine

Phase 4: Evaluation Model




- Accuracy Metrics
- Recall Metrics

Machine Learning Model

Model	ROC
Logistic Regression	0.6728
K-Nearest Neighbor	0.8806
Decision Tree (CART)	0.8826
Random Forest	0.9965
Naïve Bayes	0.6040
K-Nearest Neighbor	0.8806
Multi Layer Perceptron	0.6948




Business Recommendation




- 
- 
- 
1. A client with an income type of student can be said to be a client who is capable of repaying the loans whether with a cash loan or revolving loan (100% of applications approved). But there only 0.005 % of applications come from the student .
 2. A client who works as an accountant can be said to be a client who is capable of repaying the loans (95% of applications approved). But, there is only 3.19% of applications come from an accountant . So do, the client who work as high skill tech staff and manager, they are capable of repaying the loans, but there are only a few applications that come from them



Create a campaign so that more student, accountant, high skill tech staff, manager interested in applying for a loan




Business Recommendation

- 
- 
- 
1. Clients with maternity leaves and cash loans can be said to be a client who is incapable of repaying the loan (100% of applications rejected) . On the contrary, all clients with maternity leave but taking revolving loans to have their applications approved .
 2. For unemployed clients, more than 50% of them have a problem repaying their loans if they take cash loan contracts . Meanwhile, all unemployed client who takes revolving loans is capable of repaying the loan.



Need further analysis, you can survey to find out if there is a problem if a client with maternity leaves or unemployed takes a cash loans contract . So, in the future, if there are clients with that type of income, you can recommend the right contract type so that their applications will be approved





**You can see the entire project
documentation here!**

<https://github.com/akhasjr23/Rakamin> -HomeCredit

THANK YOU!

