Rakamin Academy

HOME CREDIT Anda Bisa!

Home Credit Scorecard Model

Presented by
M.Idris Mangku Jagat

Table of Contents

Problem Research

Background of the Problem

Home Credit wants to predict credit scores using statistical and machine learning methods. So, Home Credit can ensure that customers are able to make payments and are not rejected when applying for a loan, and loans can be given with a principal, maturity, and repayment calendar that will motivate customers to be successful.

Scope and Goals

- The **dataset** used is the **HCI Credit Score** which consists of 2 data, namely **application training (with Target)** and **application test (without TARGET)**.
- The **machine learning models** used are **Logistic Regression**, **Decision Tree**, and **K-Nearest Neighbors (KNN)**.
- **Identify** and **predict customers** who have the **potential to be able to pay off their loans**, and who are experiencing **difficulties in paying off their loans**.

Tools

Python

Google Colaboratory

Methodology

Data Application Train

- 122 Column
- 307,511 Rows

Exploratory Data Analysis (EDA)

Identify data patterns, and characteristics.

Data Preparation

- Checking for Duplicate Data
- Checking for Missing Values
- Checking for Outliers

Modeling

Machine Learning Models

- Logistic Regression
- Decision Tree
- K-Nearest Neighbors (KNN)

Evaluate the best machine learning models.

.

Data Application Test

- 121 Column
- 48,744 Rows

Data Preparation

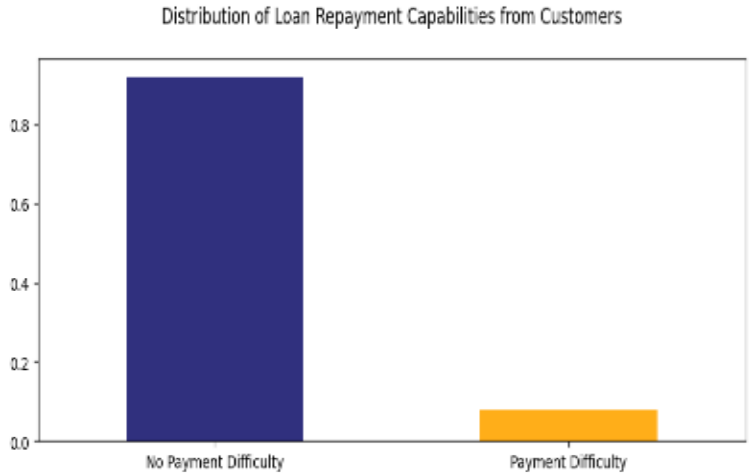
- Checking for Duplicate Data
- Checking for Missing Values
- Checking for Outliers

Prediction

Predict customer payment capabilities with the best machine learning models.

Data Visualization & Business Insight

Total Customers Based on Loan Repayment Ability



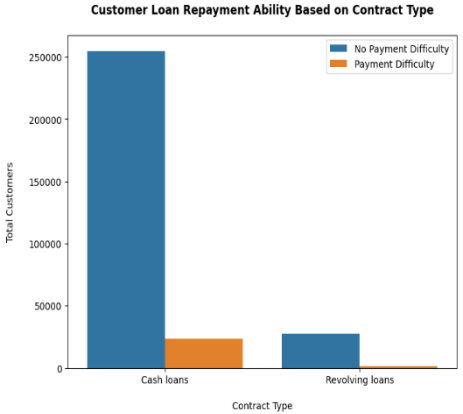
	TARGET	COUNT
0	No Payment Difficulty	282686
1	Payment Difficulty	24825

No Payment Difficulty	0.919271
Payment Difficulty	0.080729

There are **282,686 customers or around 92%**, which shows that customers do **not experience difficulties** in making loan payments within a certain period. Then, there were **24,825 customers or around 8%** who indicated there were **problems in making loan payments**.

Data Visualization & Business Insight

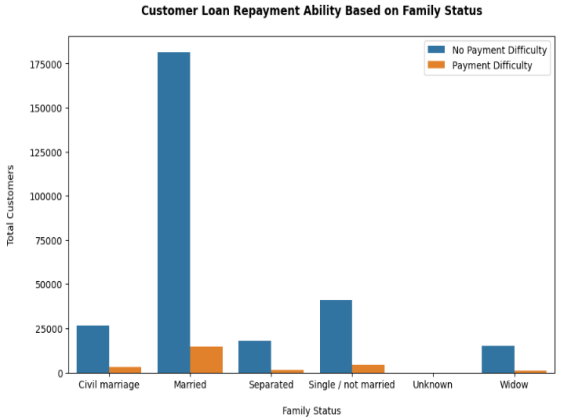
Customer Loan Repayment Ability Based on Contract Type



	NAME_CONTRACT_TYPE	TARGET	TOTAL_CUSTOMERS
0	Cash loans	No Payment Difficulty	255011
1	Cash loans	Payment Difficulty	23221
2	Revolving loans	No Payment Difficulty	27675
3	Revolving loans	Payment Difficulty	1604

There are far **more customers taking cash loans than revolving loans**.

Customer Loan Repayment Ability Based on Family Status

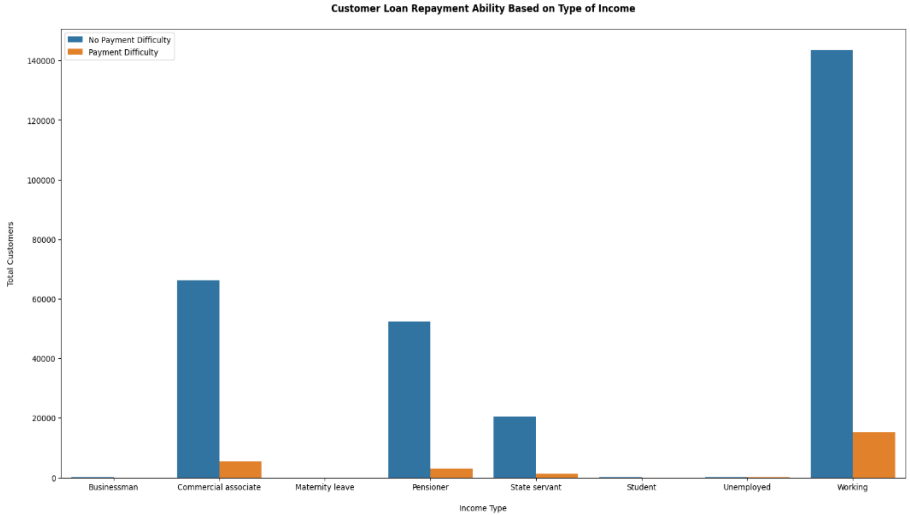


	NAME_FAMILY_STATUS	TARGET	TOTAL_CUSTOMERS
0	Civil marriage	No Payment Difficulty	26814
1	Civil marriage	Payment Difficulty	2961
2	Married	No Payment Difficulty	181582
3	Married	Payment Difficulty	14850
4	Separated	No Payment Difficulty	18150
5	Separated	Payment Difficulty	1620
6	Single / not married	No Payment Difficulty	40987
7	Single / not married	Payment Difficulty	4457
8	Unknown	No Payment Difficulty	2
9	Widow	No Payment Difficulty	15151
10	Widow	Payment Difficulty	937

- It is known that **married customers applied for the most loans**, namely **around 196,432 loan applications**.
- Customers with **civil marriage** and **single** status have the highest level of **difficulty in repaying loans** with a percentage of **around 10%**.

Data Visualization & Business Insight

Customer Loan Repayment Ability Based on Type of Income

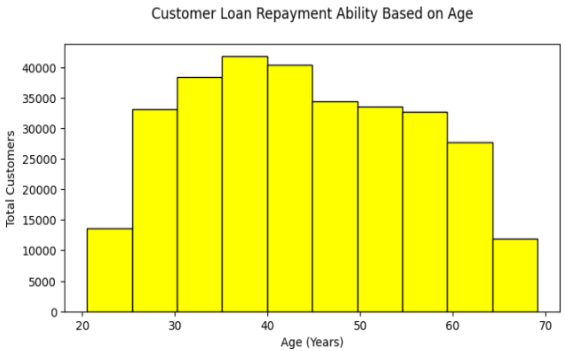


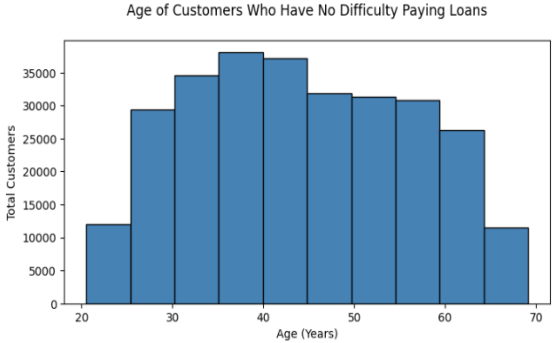
	NAME_INCOME_TYPE	TARGET	TOTAL_CUSTOMERS
0	Businessman	No Payment Difficulty	10
1	Commercial associate	No Payment Difficulty	66257
2	Commercial associate	Payment Difficulty	5360
3	Maternity leave	No Payment Difficulty	3
4	Maternity leave	Payment Difficulty	2
5	Pensioner	No Payment Difficulty	52380
6	Pensioner	Payment Difficulty	2982
7	State servant	No Payment Difficulty	20454
8	State servant	Payment Difficulty	1249
9	Student	No Payment Difficulty	18
10	Unemployed	No Payment Difficulty	14
11	Unemployed	Payment Difficulty	8
12	Working	No Payment Difficulty	143550
13	Working	Payment Difficulty	15224

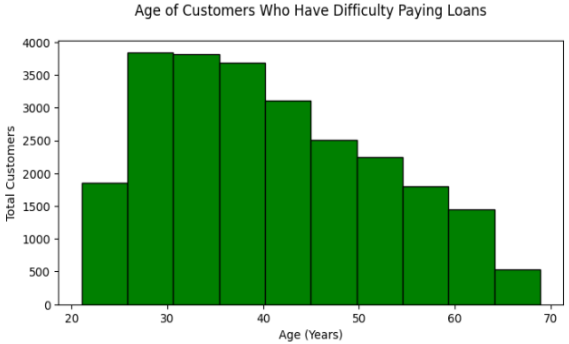
- Customers with **businessman** and **student** income types **do not experience difficulties** in paying off their loans.
- It is known that customers with the **working income type applied for the most loans**, namely **around 158,774 loan applications**.

Data Visualization & Business Insight

Customer Loan Repayment Ability Based on Age







- The customers who **most often apply for loans** are in the **35-40 years** age range, followed by customers in the 40-45 year age range. Meanwhile, customers aged **<25 years and >65 years** are the groups **least likely to apply for a loan**.
- Customers aged **35–45 years** do **not experience difficulty paying loans**. Meanwhile, customers aged **25–35 years experience difficulty in paying loans**.

Machine Learning Implementation & Evaluation

Machine Learning Models	Training Accuracy	Testing Accuracy	ROC Score
Decision Tree	**1**	**0,8831**	**0,8831**
K-Nearest Neighbors (KNN)	0,9156	0,8807	0,8806
Logistic Regression	0,6716	0,6729	0,6729

Based on the accuracy of the prediction results from training, testing and ROC data, the **Decision Tree model is better** than Logistic Regression and K-Nearest Neighbors (KNN).

So, the Decision Tree model will be used as a model to predict customer loan repayment capabilities.

Github link (to see the full project) : https://github.com/midrismj/Home-Credit-Scorecard-Model.git

Business Recommendation

- Customers who are **civil marriage** and **single** have the highest level of **difficulty in paying loans**, so this must be anticipated by companies when providing loans.

- Customers with **businessman** and **student** income types have **no difficulty at all in paying off their loans**, but very few take out loans. Companies can market to these customers so that more people take out loans.

- Customers with an age range of **35-45 years** do **not experience difficulties in paying loans**, so the company can make these customers a priority in providing loans.

Thank You

Let's connect!

Lampung, Indonesia

idrismangkujagat11@gmail.com

https://www.linkedin.com/in/midrismangkujagat/