



HOME CREDIT INDONESIA DATA SCIENTIST
PROJECT BASED INTERNSHIP PROGRAM

HOME CREDIT SCORE CARD MODEL

Presented by: Andi Sadri Agung





OVERVIEW

01

PROBLEM
RESEARCH

02

DATA
PREPROCESING

03

VISUALIZATION &
BUSINESS INSIGHT

04

MACHINE LEARNING
MODEL

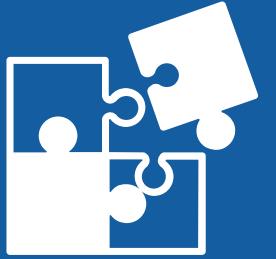
05

BUSINESS
RECOMMENDATION





PROBLEM RESEARCH



PROJECT BACKGROUND

Home Credit aims to enhance financial inclusion by offering a positive and secure borrowing experience to those with limited or no credit history. To achieve this goal and ensure a favorable loan experience, Home Credit employs alternative data sources to assess clients' repayment capabilities. This approach guarantees that eligible clients are not denied loans and that loans are structured with terms and schedules that empower clients for success.

DATA SOURCE

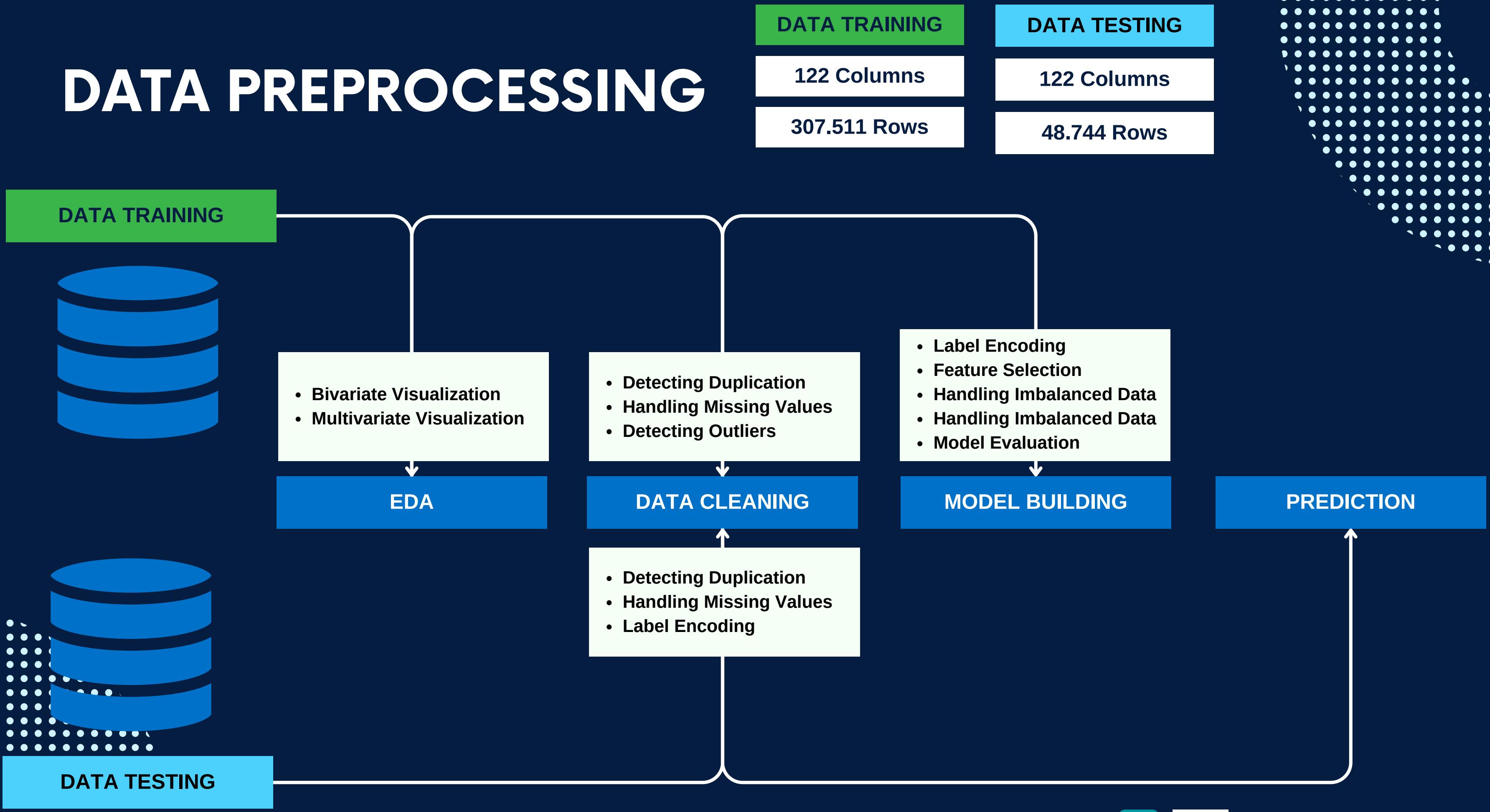
The data comprises two primary tables: one for training, including the "TARGET" variable, and the other for testing, which lacks the "TARGET" variable, with each table stored in a separate file.

OBJECTIVE & ACTION

- 1.Distinguish traits of clients likely to face loan repayment challenges from those who won't.
- 2.Forecast clients' repayment capacities through data cleaning, visualization for insights, and machine learning model development.
- 3.Offer strategic recommendations to enhance clients' loan approval rates and success for the company.

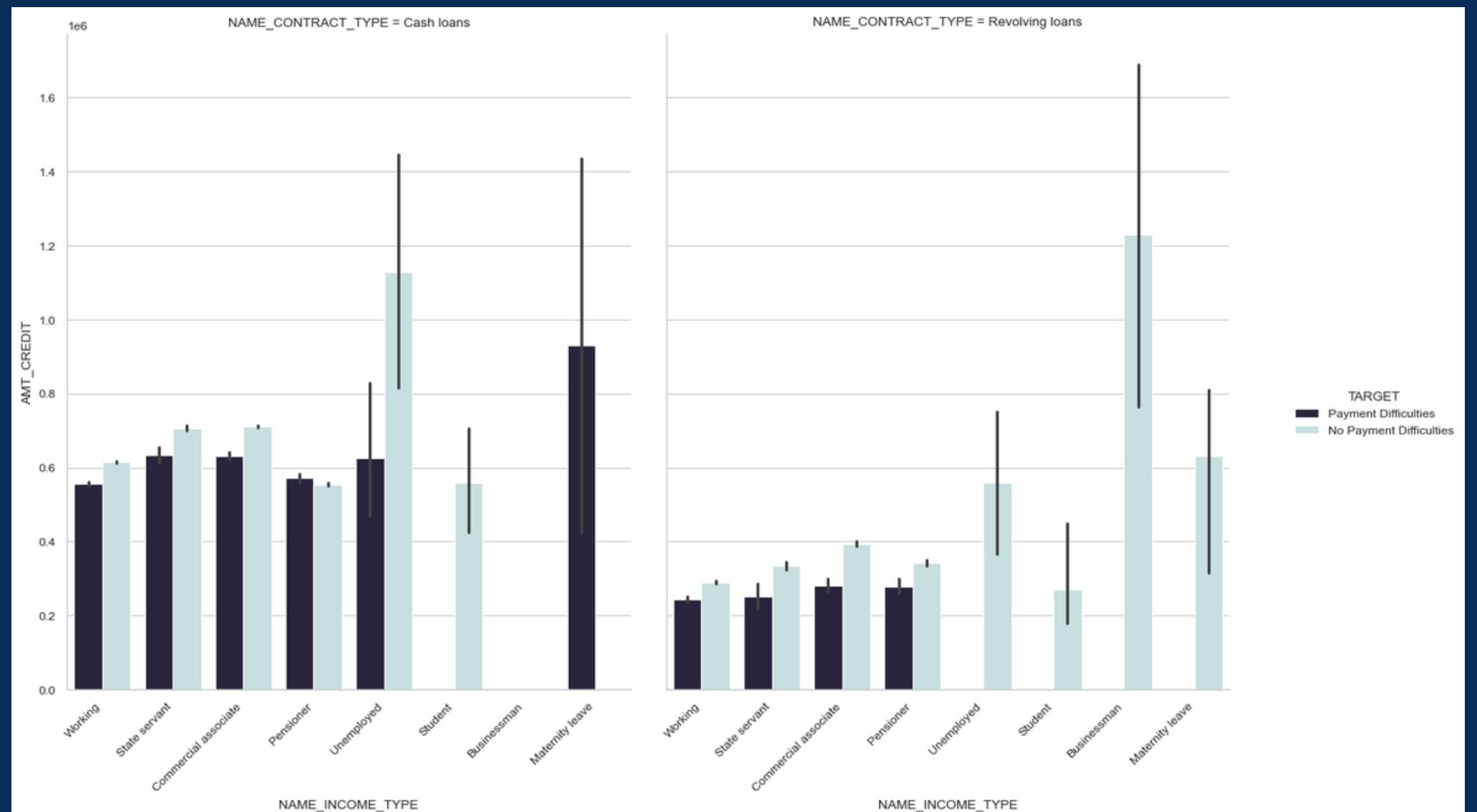


DATA PREPROCESSING



VISUALIZATION & BUSINESS INSIGHT

Income Type of Client

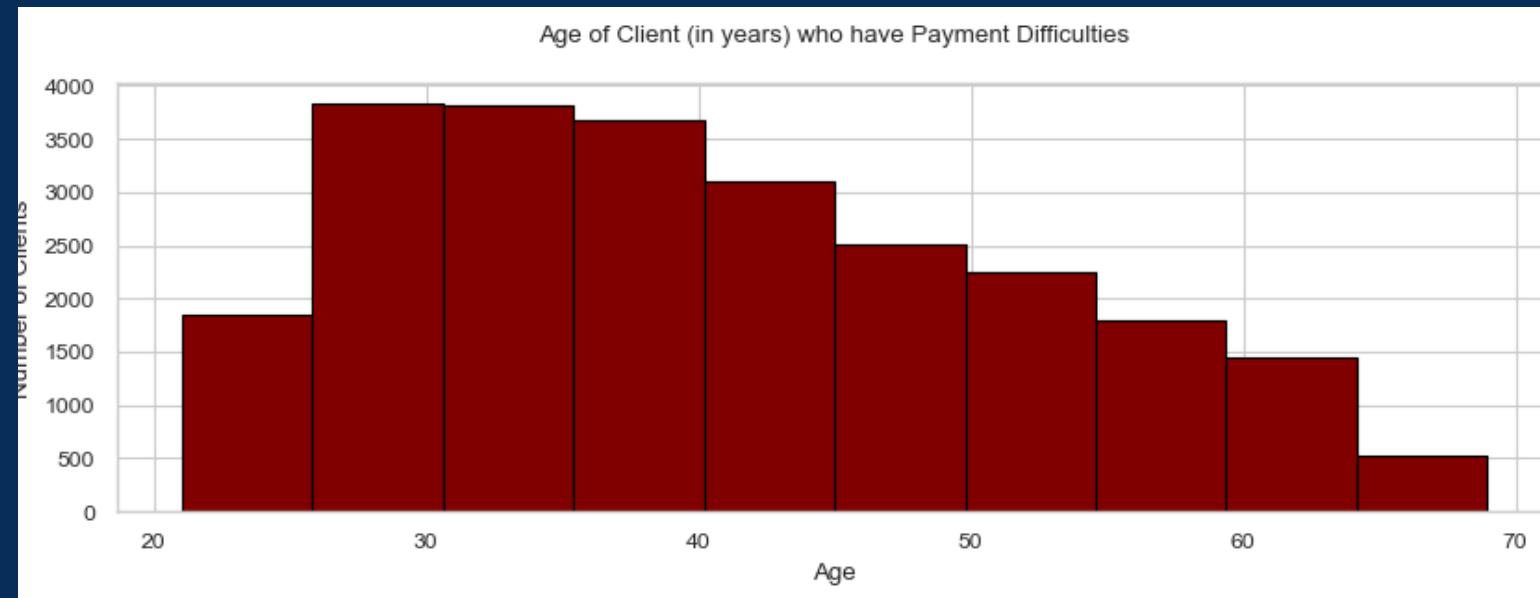


- All student clients, regardless of the type of loan (cash loan or revolving loan) and for low to medium credit amounts, have no difficulty repaying their loans.
- For clients with the income type of maternity leave and cash loans, all of them face difficulties repaying the loans, specifically for medium credit amounts. However, clients with maternity leave income and revolving loans do not encounter any repayment issues.
- Among unemployed clients with cash loans, more than 50% experience problems repaying loans, particularly for medium credit amounts. In contrast, all unemployed clients with revolving loans do not face any repayment difficulties.

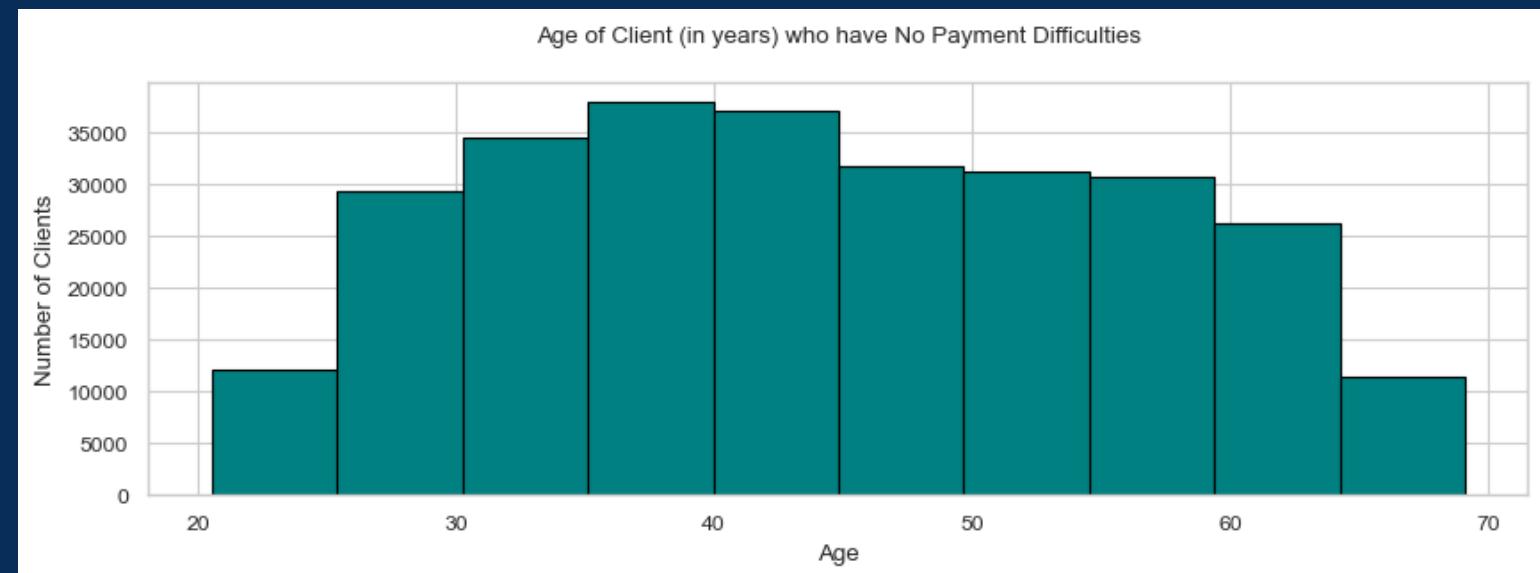


VISUALIZATION & BUSINESS INSIGHT

Age of Client



- Clients aged between 35 and 45 years, with no payment difficulties, should be prioritized as they represent a low-risk group.
- On the other hand, clients between 25 and 35 years old tend to face payment difficulties and should be subject to more in-depth assessment and support.



- The majority of loan applicants fall within the 35-40 age bracket, indicating a significant demand within this demographic.
- In contrast, there is a notably low number of applicants aged below 25 or above 65, suggesting a relatively smaller client base in these age groups.



HOME CREDIT INDONESIA DATA SCIENTIST
PROJECT BASED INTERNSHIP PROGRAM

MACHINE LEARNING MODEL

Model Comparison

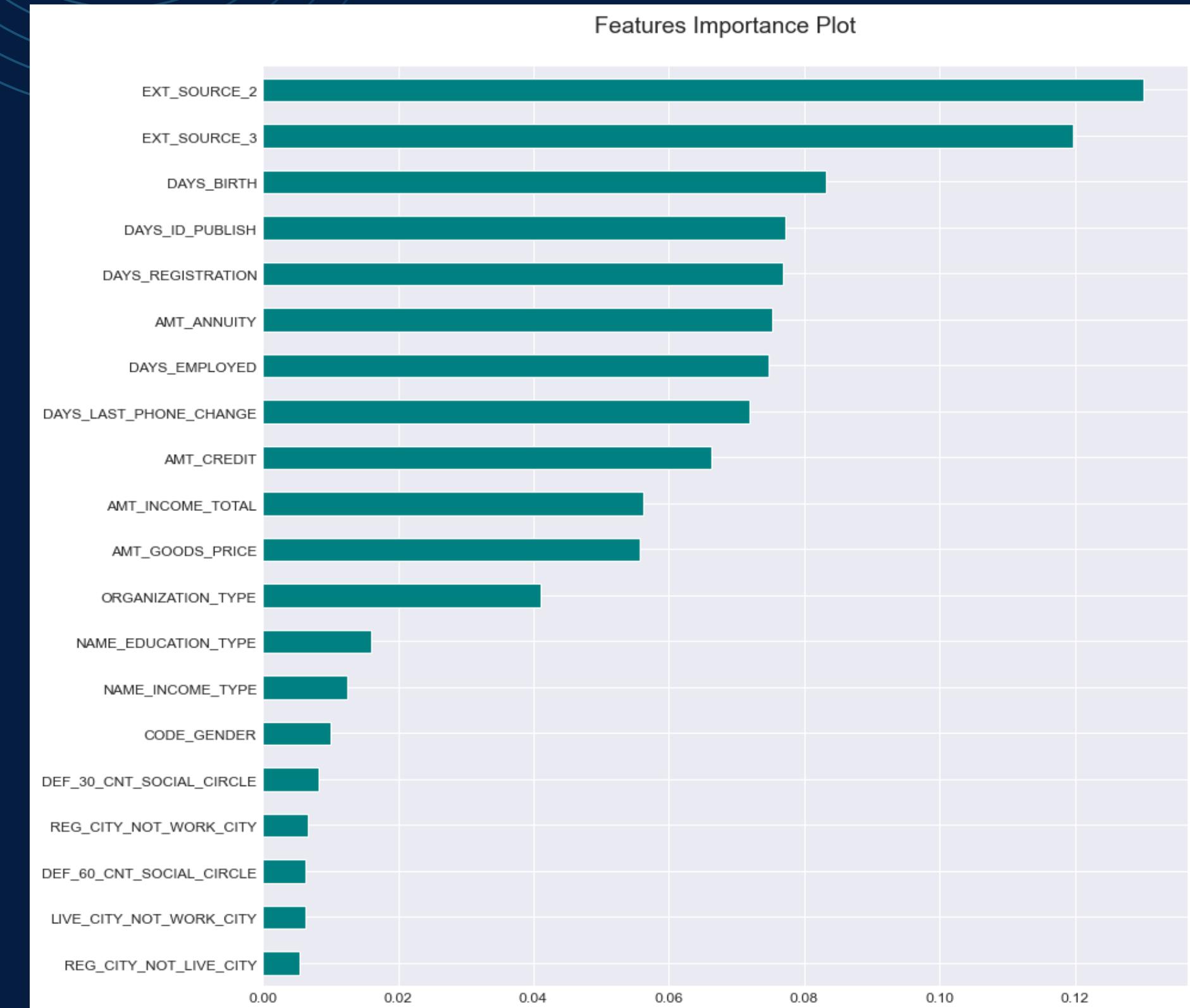
ML Algorithm	Training Accuracy Score	Testing Accuracy Score	Error Margin	ROC Score
1. Logistic Regression	67.16%	67.29%	0.13%	0.6728
2. Gaussian Naive Bayes	60.24%	60.39%	0.15%	0.604
3. Decision Tree	100%	83.9%	11.74%	0.8826
4. Random Forest	100%	99.65%	0.35%	0.9965
5. K-Nearest Neighbor	91.56%	88.07%	3.79%	0.8806
6. Neural Network	70.01%	69.48%	0.58%	0.6948

The Random Forest model exhibits similar prediction accuracy for both the training and test datasets, suggesting that the model performs well without signs of underfitting or overfitting. As a result, the Random Forest model has been selected as the best choice for forecasting clients' repayment abilities.



Best Model

- **Algorithm:** The Random Forest Classifier was utilized for modeling.
- **Performance:** The Random Forest model has achieved a remarkable performance by delivering 100% correct results. It also maintains a very low error margin of only 0.35%.
- **External Data Source 2:** The model's performance score from an external data source is unknown or not provided.
- **The 5 Most Important Features:** The model identifies the following features as the top 5 most important for its predictions: Client's age in days, Days ID published, Days registration. (Two other important features not mentioned in the provided text.)
- **External Data Source 3:** There is no information provided about the model's performance score from external data source 3.



BUSINESS RECOMMENDATIONS .

Based on the information provided, here are some potential solutions or strategies that the lending institution could consider:

1. **Targeted Marketing:** For student clients who have a 100% approval rate but only make up 0.005% of applications, the lending institution could implement targeted marketing campaigns to attract more student clients. Offering loan products with features that appeal to students, such as flexible repayment terms or lower interest rates, could help increase their application numbers.
2. **Expand Accountant Client Base:** Accountants, with a 95% approval rate, are considered strong candidates for loan repayment. To increase the percentage of applications from accountants (currently 3.19%), the institution could explore partnerships or referral programs with accounting organizations or firms to reach more potential clients.
3. **Attract High-Skill Tech Staff and Managers:** Similarly, high-skill tech staff and managers are seen as capable of loan repayment, but their applications are limited. Targeted marketing efforts and networking within the tech industry could help draw more applications from these professional groups.
4. **Evaluate Maternity Leave Clients:** For clients with maternity leaves, consider assessing their financial situations more thoroughly to determine if there are specific challenges they face with cash loans. It might be worthwhile to offer financial counseling or alternative loan products tailored to their needs. For those choosing revolving loans, continue to approve their applications, recognizing their capacity for repayment.
5. **Unemployed Clients:** For unemployed clients, it's important to differentiate between cash loan and revolving loan applicants. If more than 50% of unemployed clients face repayment difficulties with cash loans, consider offering alternative financial products or services that better match their circumstances. For those choosing revolving loans, continue to provide support and encourage responsible borrowing.
6. **Loan Product Diversification:** Consider diversifying the range of loan products offered to cater to different client needs and risk profiles. This could include introducing products with flexible repayment terms, lower interest rates for low-risk clients, and financial literacy programs to help clients manage their finances effectively.
7. **Continuous Monitoring:** Regularly monitor the performance of different client segments and loan products to identify any emerging trends or issues. Adjust lending strategies and product offerings accordingly to optimize client success and minimize risk.





**HOME CREDIT INDONESIA DATA SCIENTIST
PROJECT BASED INTERNSHIP PROGRAM**

Thank's For Watching

Project Repo : <https://github.com/andizadry/HCI-CreditScore-PBI>

Connect with us.



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

