



# HOME CREDIT SCORE CARD MODEL

## TASK 5 VIX Rakamin Academy

By Pitri Handayani

# TABLE OF CONTENTS



01

Problem Research

04

Machine Learning  
Implementation

02

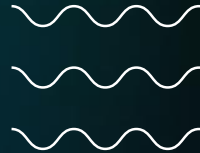
Data Pre-Processing

05

Business  
Recommendation

03

Data Visualization





01

Problem Research

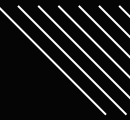
---



# Introduction


Perusahaan Pemberi kredit atau pinjam meminjam membutuhkan analisa awal apakah nasabah yang akan dibiayai mempunyai kemungkinan pembayaran kredit/pinjaman bermasalah atau tidak. Hal tersebut dilakukan perusahaan untuk menghindari kerugian dan mendatangkan keuntungan besar.

Pada Dataset ini kita akan melakukan modeling data dari Perusahaan Home Credit.





# Problem



Bagaimana membuat model yang dapat memprediksi kemampuan pembayaran Pelanggan sehingga dapat membantu tim menentukan pinjaman yang diajukan diterima atau ditolak, kemudian dapat memaksimalkan target pasar selanjutnya.





# Data Source

Dataset menggunakan file csv  
application train



# Action

1. Data Visualization
2. Data Cleaning
3. Modeling
4. Rekomendasi



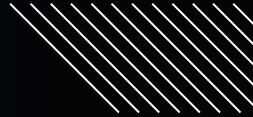


02

# Data Preprocessing

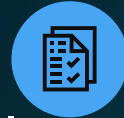
---

# Raw Data



## Data

Jumlah Kolom : 122  
Jumlah Baris : 307.511



## Cleaning

1. Tidak ada nilai duplicate
2. Beberapa kolom di drop dan Imputed



## Model

1. Label Encoding
2. Feature Selection
3. Handel Imbalance data
4. Model Building
5. Evaluasi







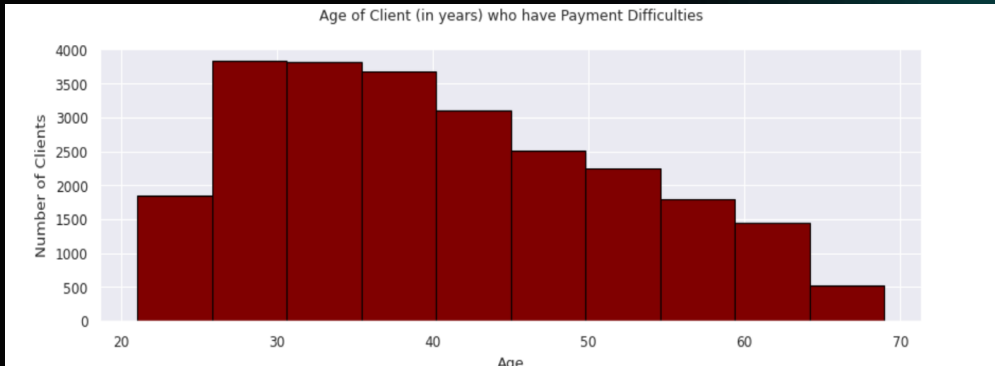
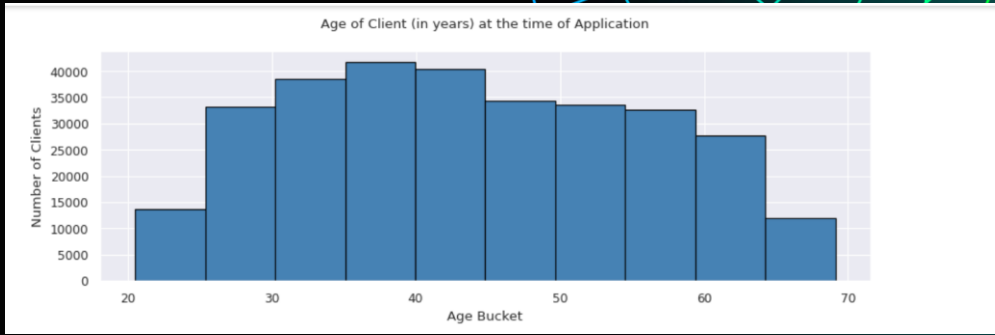
03



# Business Insights

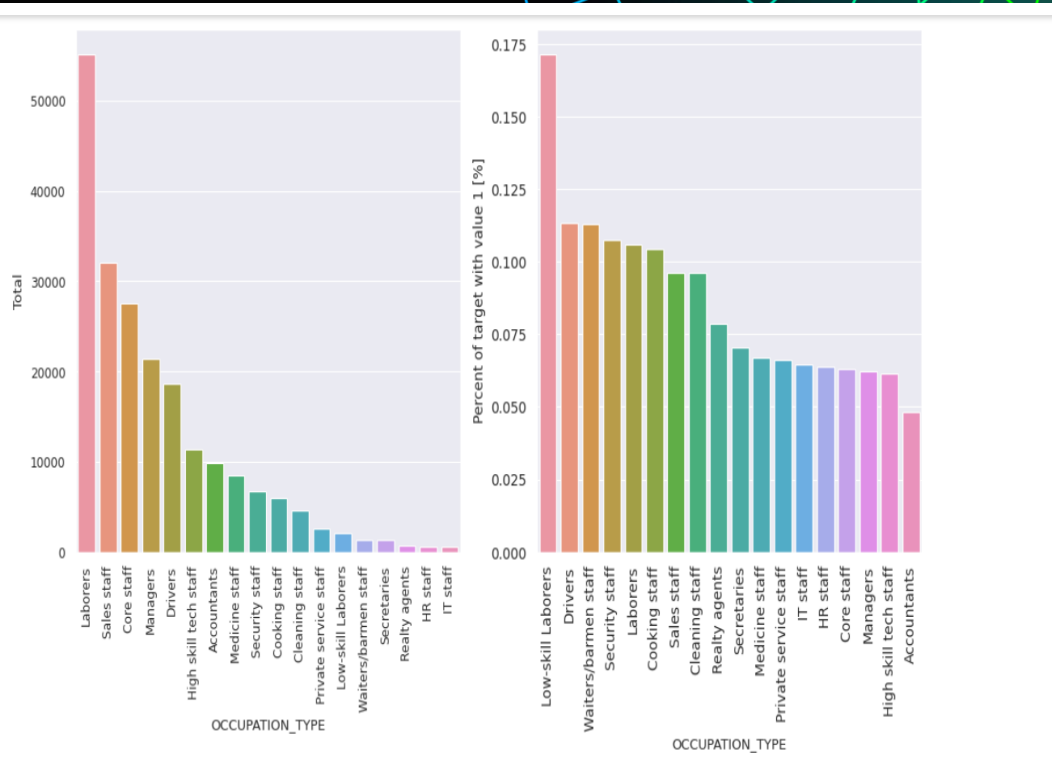
---



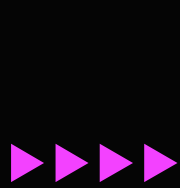


Dari grafik, terlihat bahwa pada umur 35-50 tahun merupakan kategori usia yang memiliki pengajuan pinjaman terbesar dan memiliki pinjaman lunas tanpa mengalami kendala apapun.





Sebagian besar pinjaman diambil oleh Buruh, diikuti oleh sales staff, core staff dan terendah oleh IT staff. Kategori dengan persentase yang tidak mengembalikan pinjaman tertinggi adalah Tenaga Kerja Keterampilan Rendah (di atas 17%), diikuti oleh sopir dan waiters.



# 04



## Machine Learning Model

---





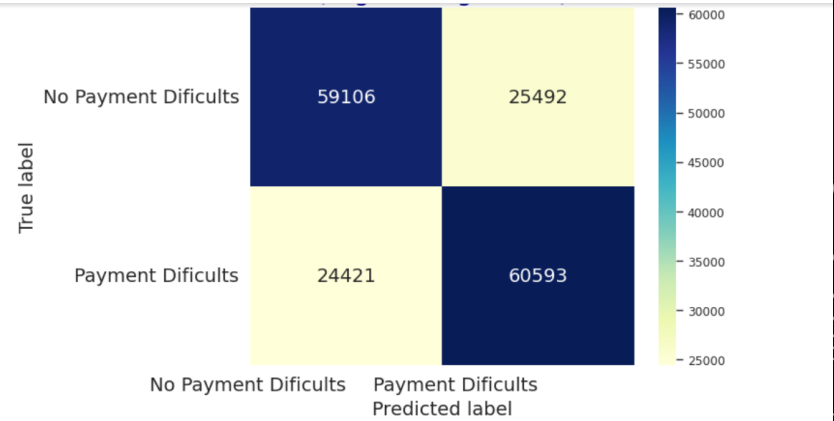
# Logistic Regression

```
[ ] #Performansi Data Testing
```

```
from sklearn.metrics import classification_report
# Predict
y_test_pred = log_model.predict(x_test)
# Print classification report
print('Classification Report Testing Model (Logistic Regression):')
print(classification_report(y_test, y_test_pred))
```

Classification Report Testing Model (Logistic Regression):

	precision	recall	f1-score	support
0.0	0.71	0.70	0.70	84598
1.0	0.70	0.71	0.71	85014
accuracy			0.71	169612
macro avg	0.71	0.71	0.71	169612
weighted avg	0.71	0.71	0.71	169612



Akurasi pada model regresi logistic untuk data testing sebesar 0.71 atau 71%



# Random Forest Classifier

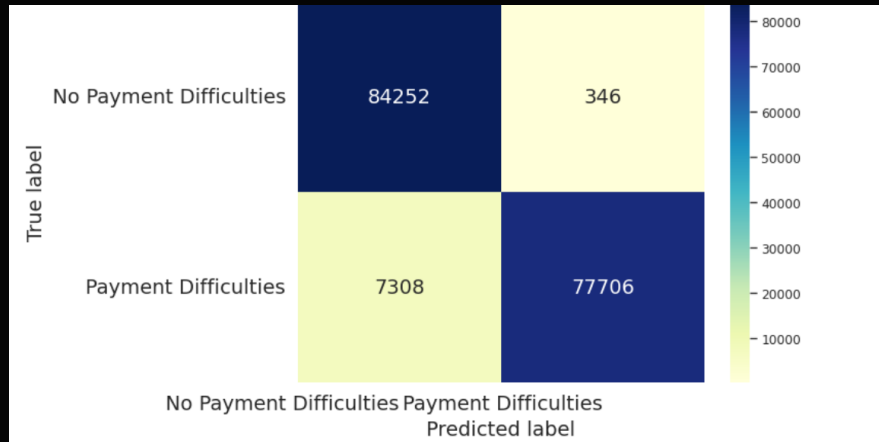


```
#Performa Data Test

from sklearn.metrics import classification_report
# Predict
y_test_pred = rdf_model.predict(x_test)
# Print classification report
print('Classification Report Testing Model (Random Forest Classifier):')
print(classification_report(y_test, y_test_pred))
```

Classification Report Testing Model (Random Forest Classifier):

	precision	recall	f1-score	support
0.0	0.92	1.00	0.96	84598
1.0	1.00	0.91	0.95	85014
accuracy			0.95	169612
macro avg	0.96	0.95	0.95	169612
weighted avg	0.96	0.95	0.95	169612



Akurasi Pada model RF untuk data testing sebesar 0.95 atau 95%



## Kesimpulan

Pada Dataset ini model Random Forest Classifier sangat baik dibandingkan model Regresi Logistik



05


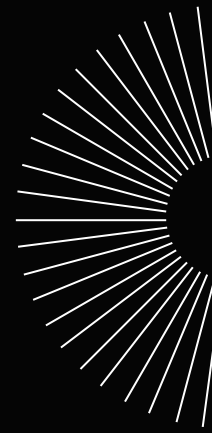



Recommendation

---





- 
- 
1. Target Pasar selanjutnya bisa diarahkan kepada klien yang berumur 35-50 tahun , karena merupakan kategori usia yang memiliki pengajuan pinjaman terbesar dan memiliki pinjaman lunas tanpa mengalami kendala apapun.
  2. Pekerja akuntan dapat dikatakan sebagai pelanggan yang mampu mengembalikan pinjaman dengan baik (hanya 5% yang memiliki masalah pembayaran) tetapi dalam pengajuan pinjaman masih tidak terlalu besar, ini bisa ditingkatkan dengan membuat campaign agar mereka tertarik untuk mengajukan pinjaman.
  3. Dapat membuat iklan yang menarik untuk pekerja dengan posisi IT Staff, HR staff dan realty agents.
- 

# THANKS

Visit my github:

[https://github.com/Phandayani20/Machine Learning/blob/main/TASK\\_5\\_HCI\\_X\\_Rakamin\\_Academy\\_Scored\\_Model.ipynb](https://github.com/Phandayani20/Machine-Learning/blob/main/TASK_5_HCI_X_Rakamin_Academy_Scored_Model.ipynb)



**CREDITS:** This presentation template was created by **Slidesgo**, and includes icons by **Flaticon** and infographics & images by **Freepik**

