

# Final Task Machine Learning Modelling for Home Credit Indonesia

Laras Puji Pramesty (Bachelor Of Information Technology)

Virtual Internship Experience Data Scientist at Home Credit Indonesia  
by Rakamin Academy

# 01 – Problem Research

- Banyaknya Klien yang berjuang untuk mendapatkan pinjaman Kredit
- Mengidentifikasi dan memprediksi karakteristik klien yang memiliki Kesulitan pembayaran dan yang tidak memiliki kesulitan pembayaran credit
- Membuat Model Scoring untuk menganalisis, memahami dan model credit risk dapat diatasi secara optimal
- Dapat menentukan klien di Home Credit Indonesia, pada saat mengajukan aplikasi diterima atau ditolak

# 02 - Data Pre-Processing

## EDA (Exploratory Data Analysis)

Visualization between 2  
features and more than  
features

## Data Cleaning

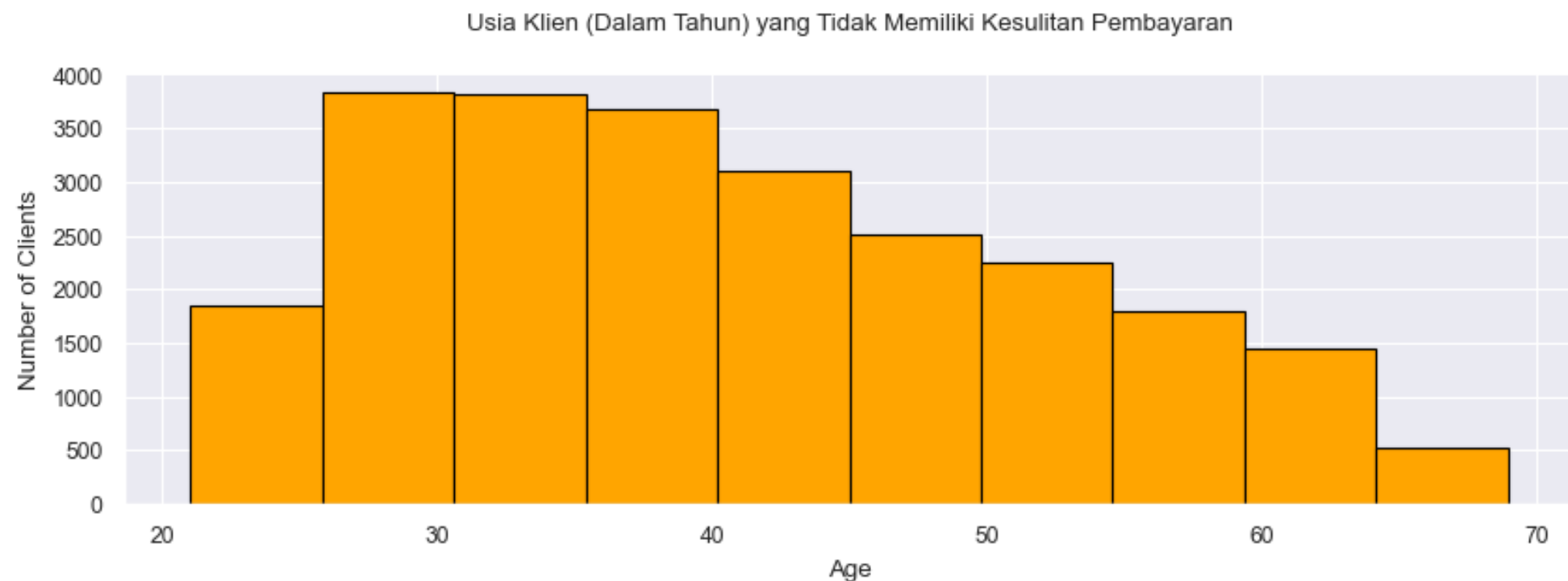
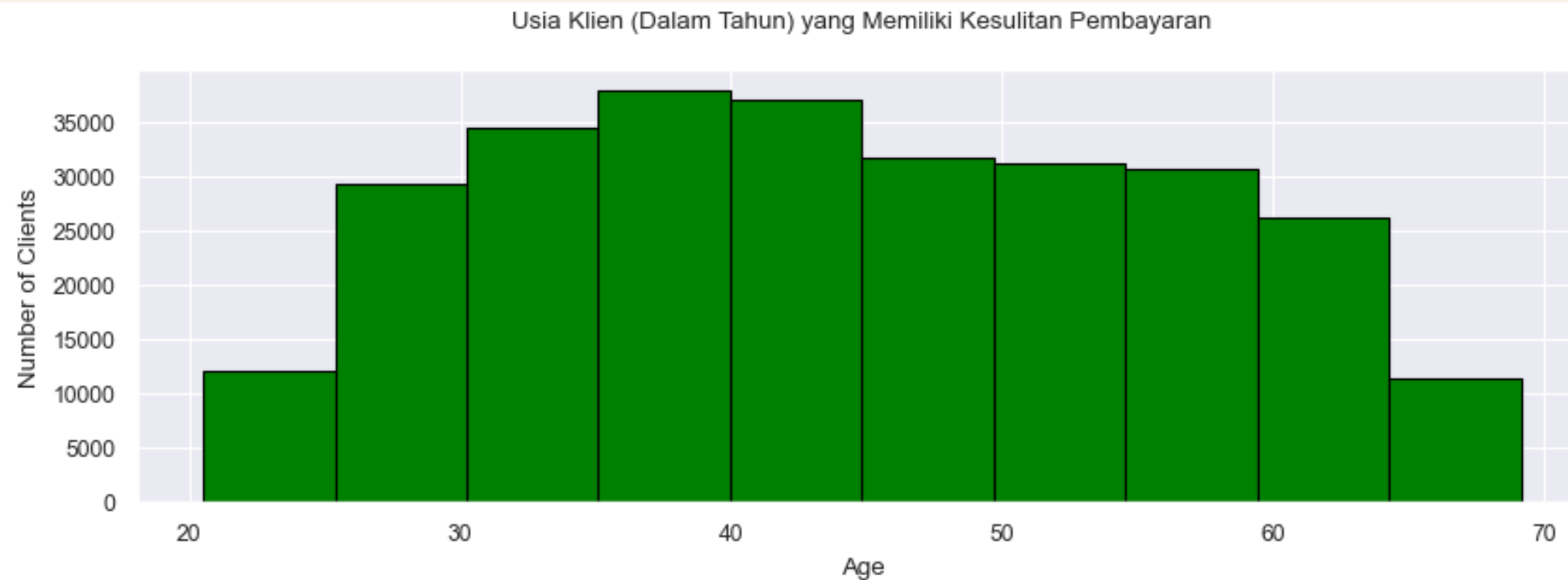
- detecting duplication
- handling missing value
- Detecting outliers

## Model Building

- Label Encoding
- Feature Selection
- Handling imbalance data
- Model Building
- Model Evaluation

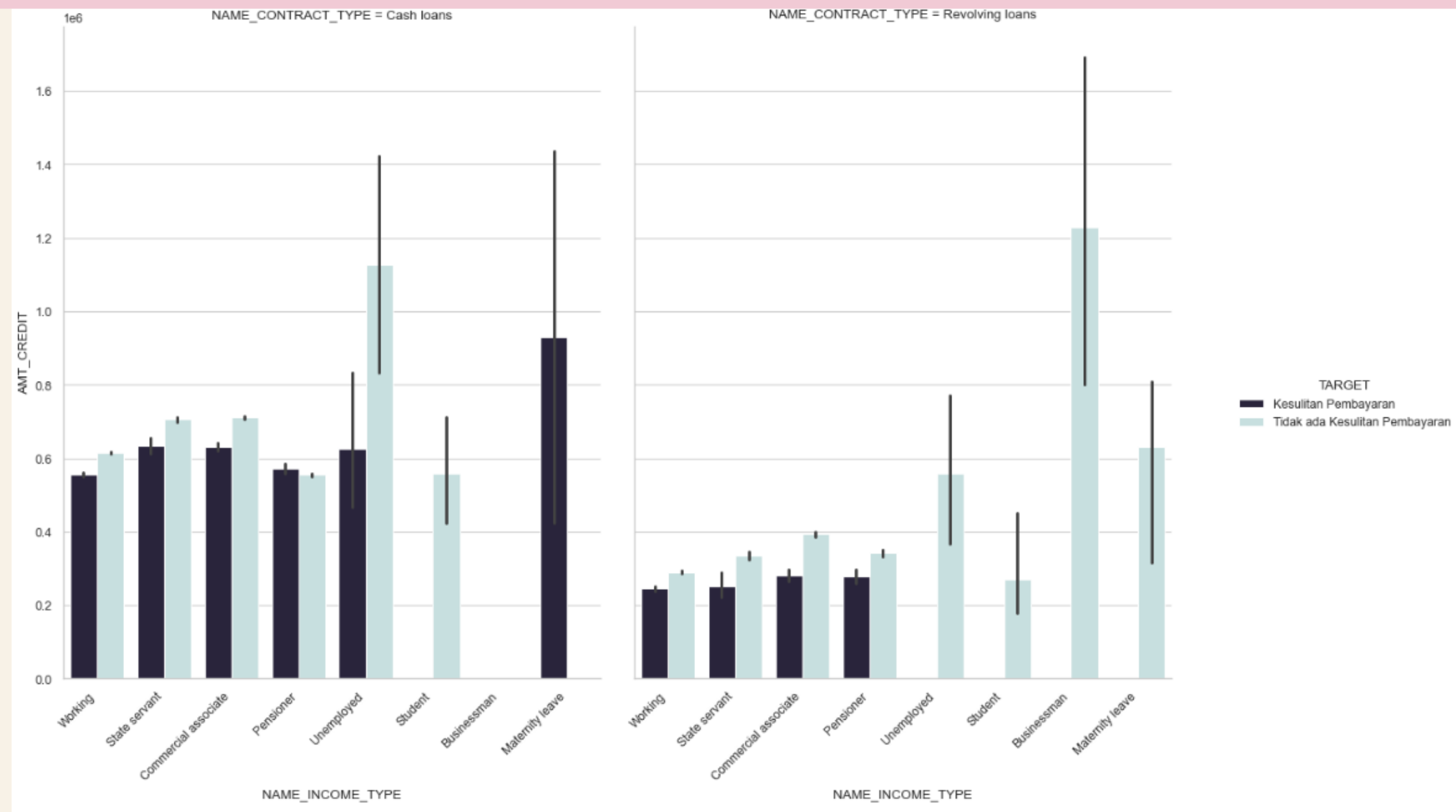


# 03 - Data Visualization and Business Insight



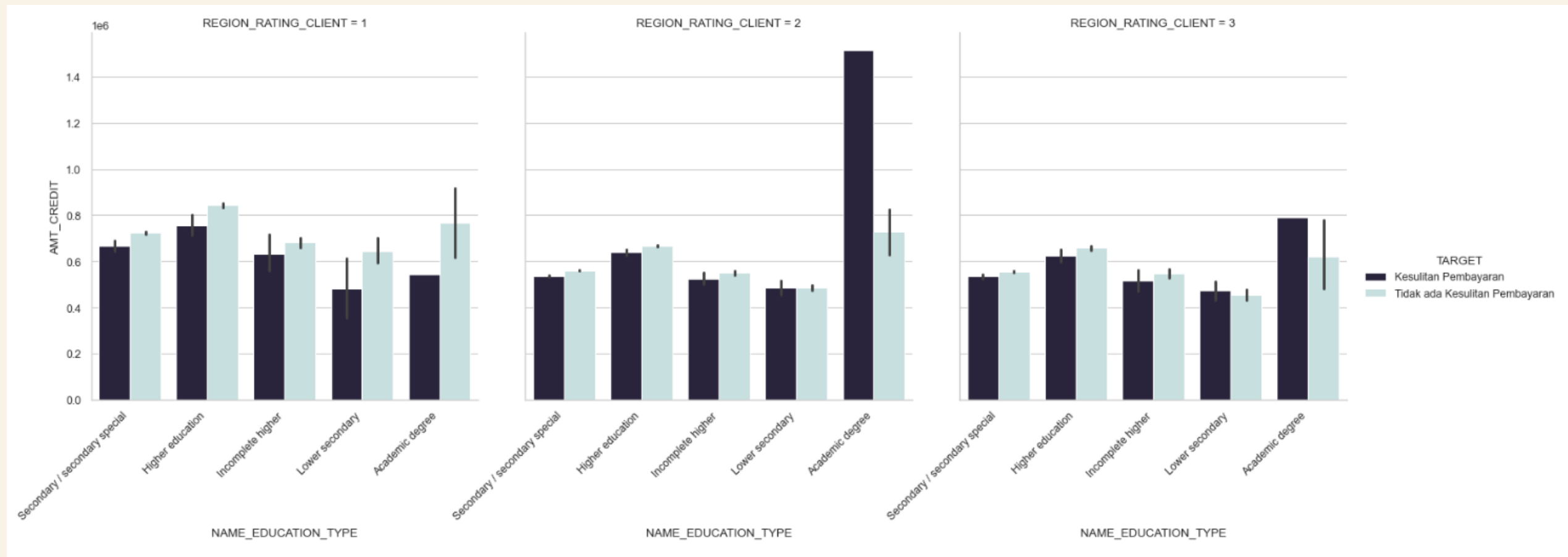
Sebagian besar nasabah yang mengajukan pinjaman berada pada rentang usia 35-40 tahun, diikuti nasabah pada rentang usia 40-45 tahun. Sementara itu, jumlah pelamar untuk klien berusia <25 tahun atau usia >65 tahun sangat rendah. Klien yang tidak mengalami kesulitan pembayaran adalah klien dengan rentang usia 35-45 tahun. Sedangkan nasabah yang mengalami kesulitan pembayaran adalah nasabah dengan rentang usia 25-35 tahun.

## 03 - Data Visualization and Business Insight



- Untuk jenis **maternity leave** dengan pinjaman tunai, semua klien bermasalah dalam membayar kembali pinjaman sebesar kredit menengah jumlah pinjaman. Sementara semua klien dengan maternity leave dan revolving loans tidak mengalami kesulitan pembayaran kembali pinjaman.
- Semua klien mahasiswa tidak mengalami kesulitan membayar kembali pinjaman dengan uang tunai pinjaman atau pinjaman bergulir untuk tingkat kredit rendah hingga menengah jumlah pinjaman.
- Untuk klien menganggur dengan pinjaman tunai, lebih dari 50% klien memiliki masalah dalam membayar kembali pinjaman dengan menengah jumlah kredit pinjaman. Sementara semua klien yang menganggur dengan pinjaman bergulir tidak mengalami kesulitan untuk mengembalikan pinjaman.

# 03 - Data Visualization and Business Insight



Untuk klien yang memiliki gelar akademik dan tinggal di wilayah dengan peringkat 2, memiliki masalah dalam membayar kembali pinjaman untuk jumlah kredit pinjaman yang lebih tinggi. Dan, klien dengan gelar yang sama tetapi tinggal di wilayah dengan peringkat 3 memiliki masalah dalam membayar pinjaman untuk kredit pinjaman dalam jumlah sedang.

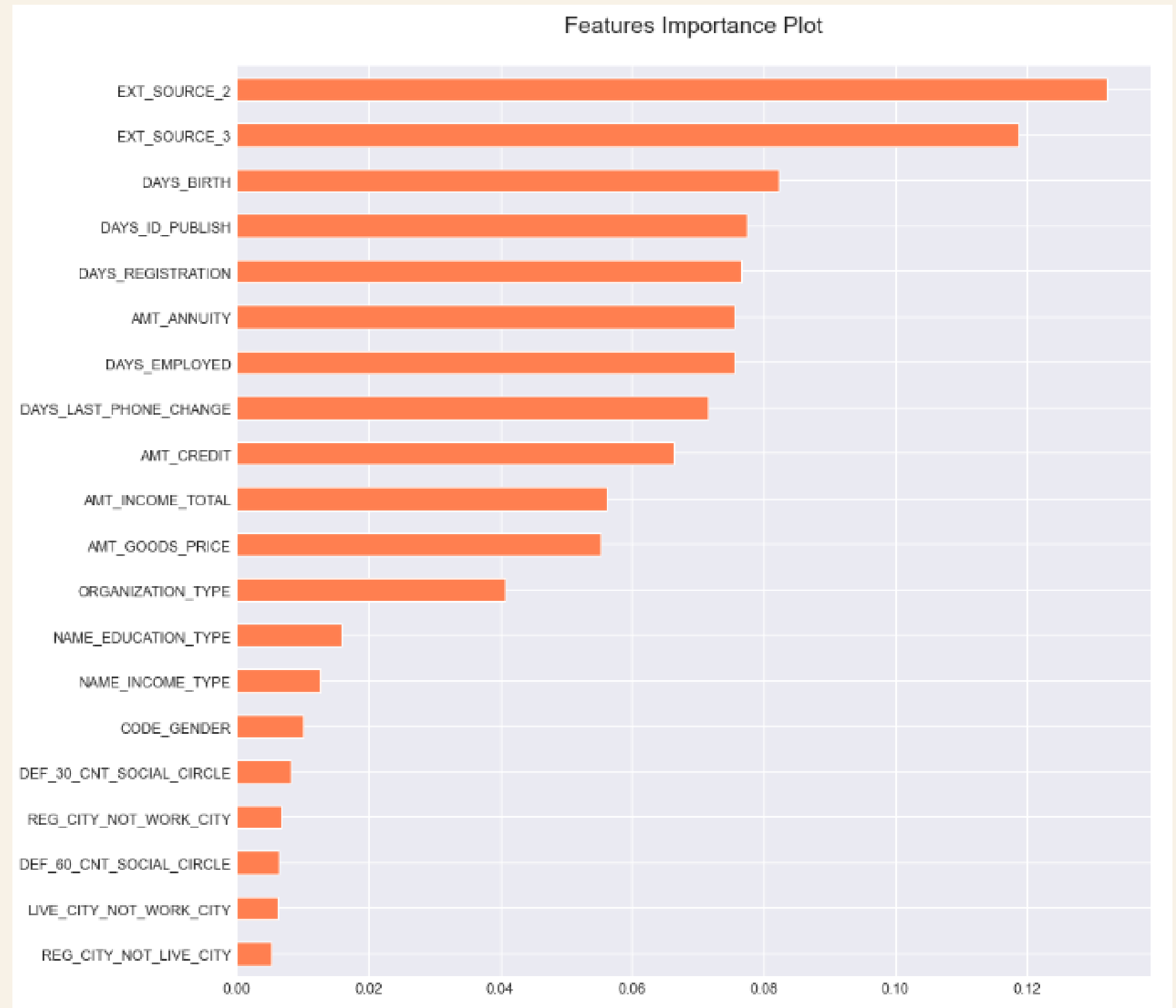


# 04 - Machine Learning Implementation and Evaluation

- Berdasarkan Grafik Disamping terlihat bahwa, terdapat 5 features terpenting yakni:

1. EXT\_SOURCE 2
2. EXT\_SOURCE 3
3. Client's age In Days
4. Days ID Publish
5. Days Registration

- Performance : Random Forest Model memberikan 100% hasil yang tepat dimana hanya terdapat 0.35% error margin



# 04 - Machine Learning Implementation and Evaluation

## Model Comparison

<b>Model</b>	<b>Training Accuracy Score</b>	<b>Testing Accuracy Score</b>	<b>ROC Score</b>
<b>Logistic Regression</b>	<b>67.16%</b>	<b>67.29%</b>	<b>0.6729</b>
<b>Gaussian Naive Bayes</b>	<b>60.24%</b>	<b>60.39%</b>	<b>0.604</b>
<b>Decision Tree</b>	<b>100%</b>	<b>88.3%</b>	<b>0.8823</b>
<b>Random Forest</b>	<b>100%</b>	<b>99.66%</b>	<b>0.9966</b>



## 04 - Machine Learning Implementation and Evaluation

### Model Comparison

Model Algoritma Random Forest memiliki Nilai yang tidak banyak berbeda, dan menjadi model yang sangat bagus dimana tidak terjadi Underfitting maupun overfitting. Jadi Model Random Forest dapat dipilih menjadi model terbaik dalam memprediksi kemampuan Pembayaran Klien

# 05 - Business Recommendation

- Klien dengan tipe penghasilan mahasiswa dapat dikatakan sebagai klien yang mampu membayar kembali pinjaman baik dengan pinjaman tunai atau pinjaman bergulir (100% dari aplikasi disetujui). Namun hanya 0,005% aplikasi yang berasal dari siswa.
- Klien yang berprofesi sebagai akuntan dapat dikatakan sebagai klien yang mampu membayar kembali pinjaman (95% dari aplikasi disetujui). Namun, hanya ada 3,19% aplikasi berasal dari akuntan. Begitu juga, klien yang bekerja sebagai staf teknologi keterampilan tinggi dan manajer, mereka mampu membayar kembali pinjaman, tetapi hanya ada beberapa aplikasi yang berasal dari mereka
- Untuk jenis penghasilan cuti hamil dengan pinjaman tunai, semua klien memiliki masalah dalam membayar kembali pinjaman untuk kredit menengah jumlah pinjaman. Sementara semua klien dengan cuti melahirkan dan pinjaman bergulir tidak memiliki kesulitan mengembalikan pinjaman.
- Untuk klien dengan jenis cuti melahirkan dan pelanggan menganggur dengan pinjaman tunai memiliki masalah membayar pinjaman dengan jumlah pinjaman moderat. Adapun pinjaman bergulir, tidak ada kesulitan dalam membayar kembali meminjamkan
- Persentase tertinggi klien dengan masalah di membayar kembali pinjaman adalah klien dengan pekerjaan low-Skilled worker jobs sebesar 17% dan klien yang bekerja sebagai akuntan, High Skill tech staff, Manager bisa dikatakan klien yang mampu membayar kembali pinjaman tersebut

## 05 - Business Recommendation

- Membuat kampanye untuk siswa, akuntan, high skill tech, manajer, dan klien dengan rentang usia 35 45 tahun ke atas tertarik untuk mengajukan pinjaman
- Perlu analisis lebih lanjut, dengan melakukan survei (seperti Family Status, Tipe Rumah/Rusun dan Status Kepemilikan, Jumlah Anak-anak, Pekerjaan) untuk mengetahui apakah ada masalah saat mengambil kontrak pinjaman tunai, terutama klien cuti melahirkan atau menganggur. Jadi, ke depan, jika Anda memiliki klien dengan itu jenis pendapatan, Anda dapat merekomendasikan jenis kontrak yang tepat agar permohonan mereka disetujui.

# Thanks

Find Out My Project in Github :  
<https://github.com/laraspramesty6770/Home-Credit-Indonesia-Final-project>