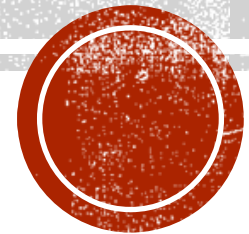


CREDIT SCORE PREDICTION

FINAL PROJECT

Home Credit Indonesia Rakamin

By Muhammad Reiza Indrawan



PROBLEM & SOLUTION

- **Business Problem**

- Resiko pinjaman kredit mengacu pada potensi kerugian finansial yang dihadapi pemberi pinjaman ketika peminjam gagal membayar kembali pinjaman mereka tepat waktu. Hal ini memberikan konsekuensi yang signifikan bagi pemberi pinjaman dan peminjam.

- **Objective**

- Mengetahui apa saja yang mempengaruhi client dalam mengalami kesulitan untuk membayar.
- Mengurangi kesempatan peminjam yang berpotensi mengalami kesulitan untuk membayar



DATA UNDERSTANDING

Jumlah Data: 307511

Jumlah Kolom : 122

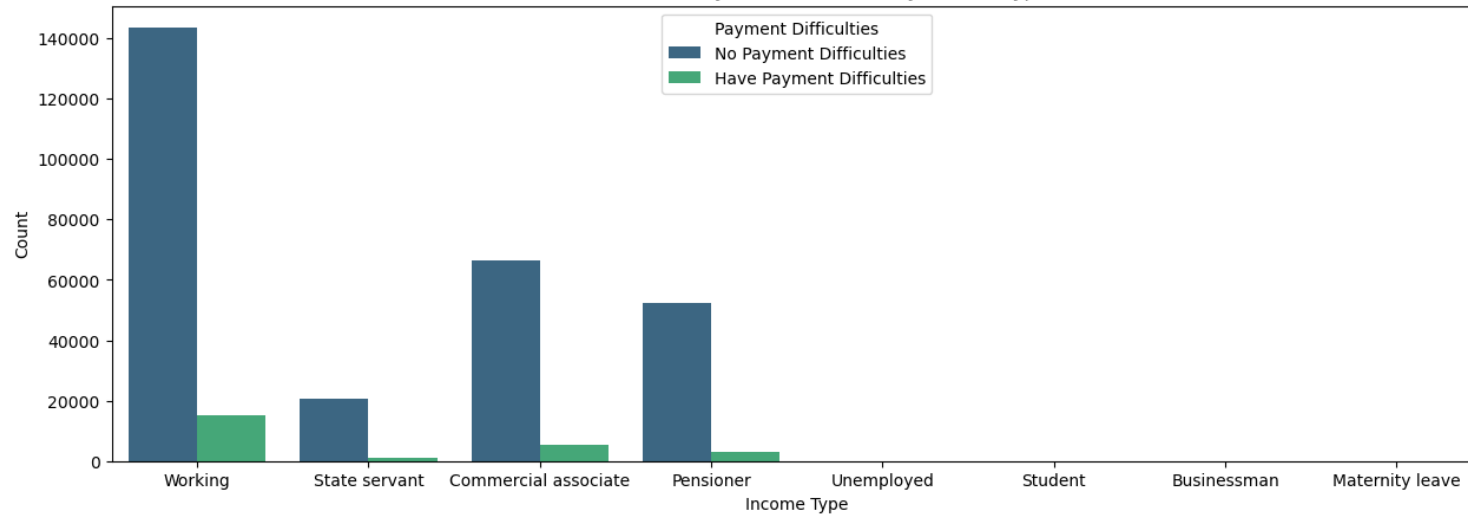
dtypes : Float64(65), int64(41), Object(16)

Dataset yang dipakai
merupakan data application
train yang diperoleh dari
Home Credit Indonesia



BUSINESS INSIGHT

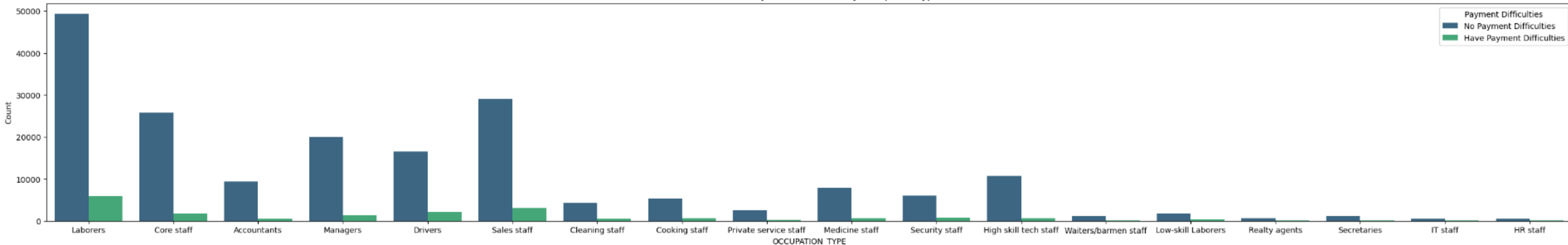
Distribution of Payment Difficulties by Income Type

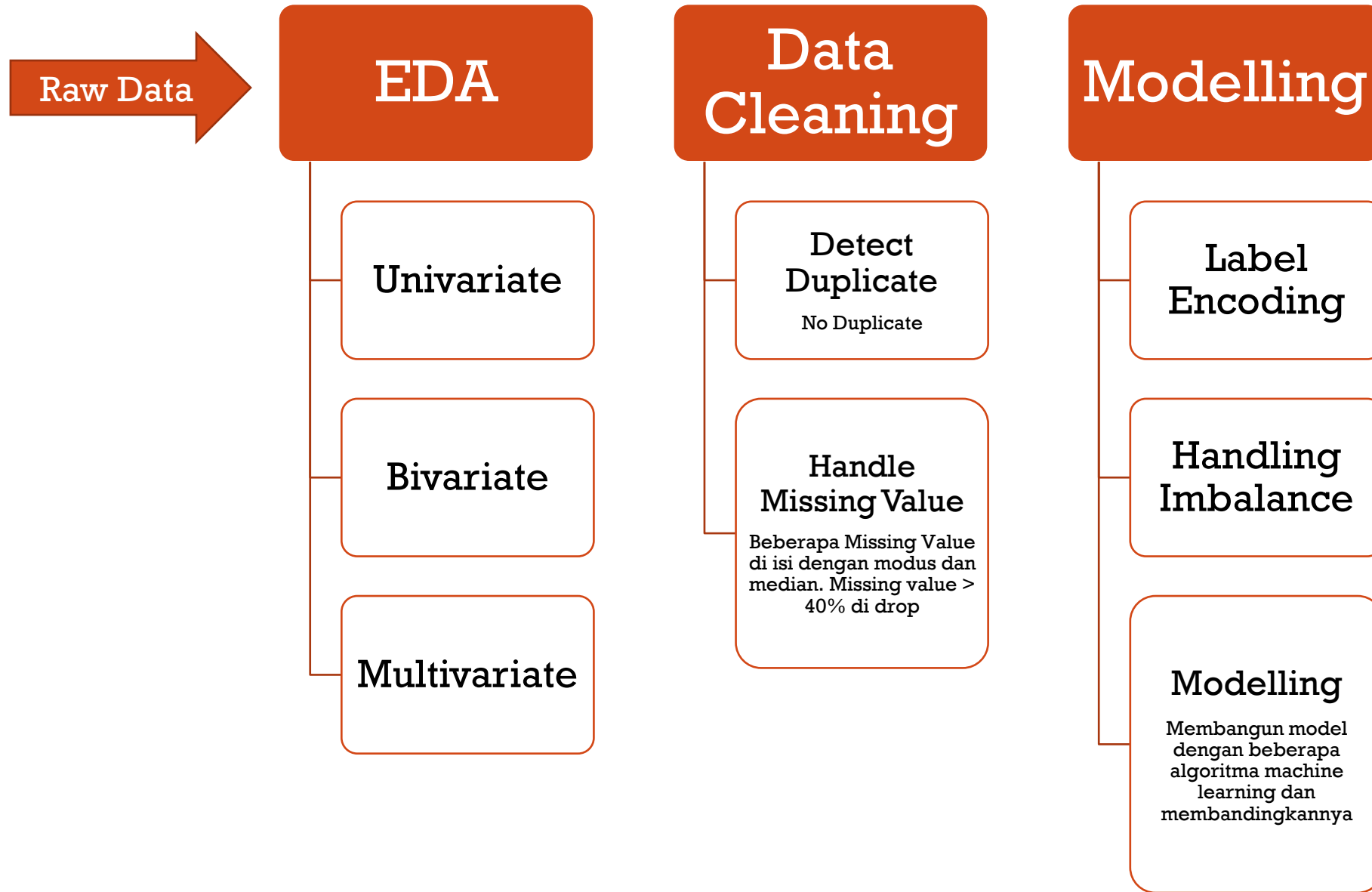


Working income type client menjadi yang paling tinggi dibanding client lainnya 90%nya tidak memiliki payment difficulties dan sisanya memiliki payment difficulties

Client dengan occupation type Laborers, sales staff dan drivers menjadi yang paling tinggi dibandingkan client dengan occupation type lainnya yang disertai memiliki payment difficulties.

Distribution of Payment Difficulties by Occupation Type





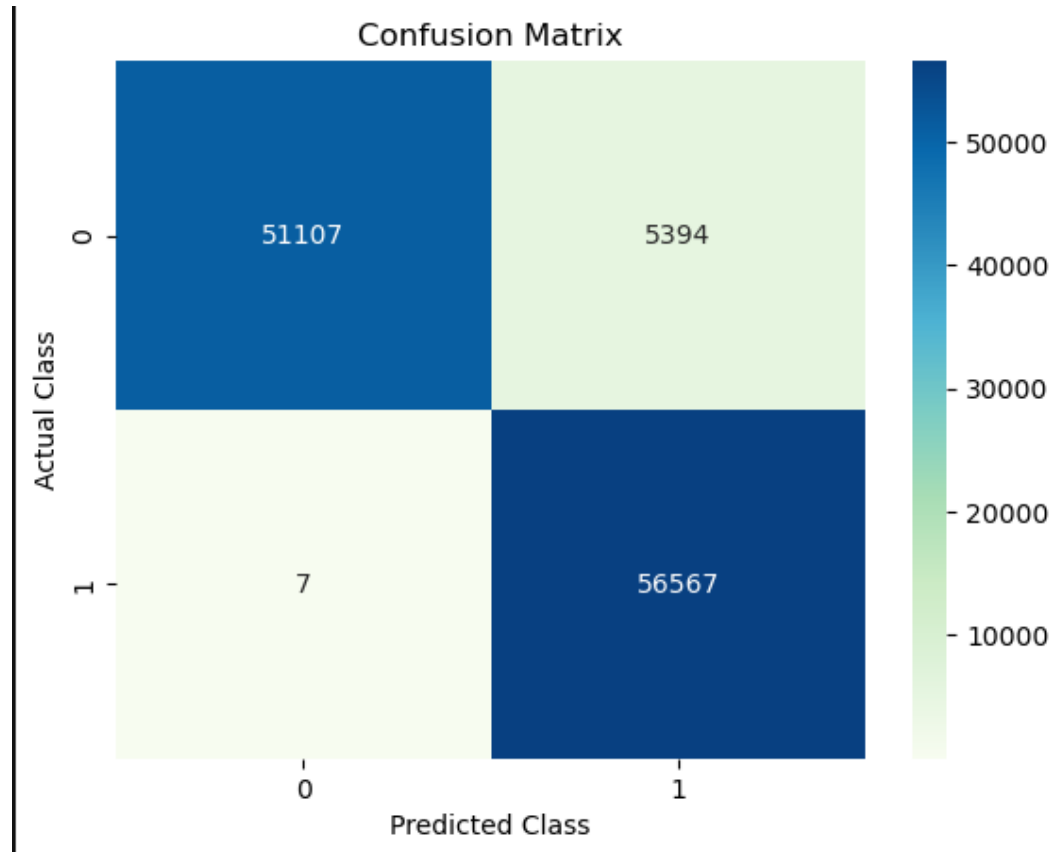
MODELING

No	Model	Akurasi		Recall		Precision		ROC/AUC	
		Train	Test	Train	Test	Train	Test	Train	Test
1	Decision Tree	100	95.22	100	91.29	100	99.99	100	95.22
2	XGB	76.07	74.85	75.37	74.05	77.45	76.56	84.20	82.78
3	Ada Boost	97.55	97.52	98.57	98.49	78.73	78.55	96.87	96.62
4	Logistic Regression	57.94	57.77	58.50	58.33	54.60	54.57	60.72	60.53

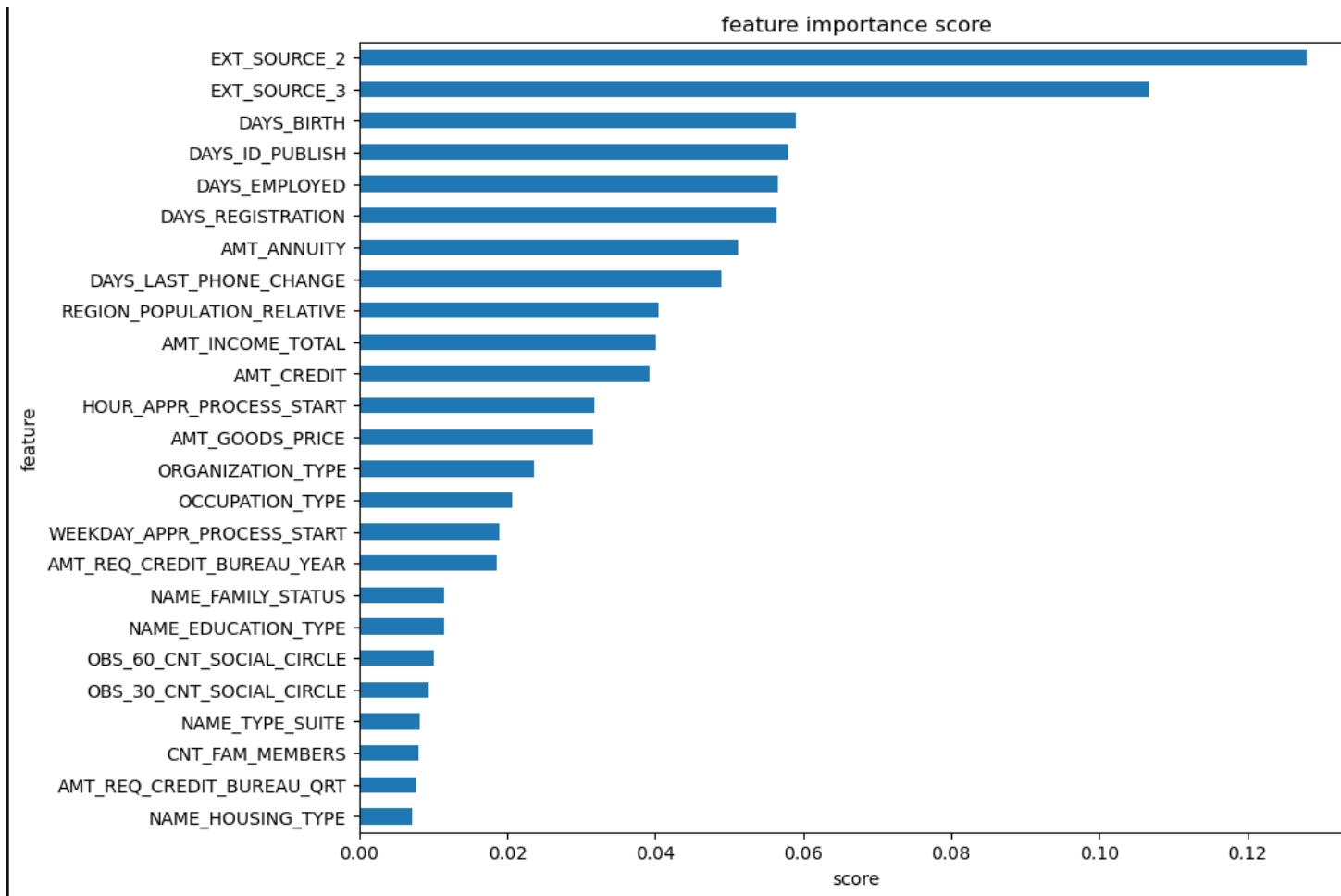


CONFUSION MATRIX

True Positive (TP): customer yang diprediksi kesulitan membayar dan kenyataannya kesulitan membayar.
True Negative (TN): customer yang diperkirakan tidak kesulitan membayar dan kenyataannya tidak kesulitan membayar.
False Positive (FP): customer yang diprediksi kesulitan membayar dan kenyataannya tidak kesulitan membayar.
False Negative (FN): customer yang diprediksi tidak kesulitan membayar dan ternyata kesulitan membayar.



FEATURE IMPORTANCE



EXT_SOURCE_2 dan
EXT_SOURCE_3
merupakan feature yang
paling penting



RECOMMENDATION

- Diperlukan perhatian extra terhadap client yang memiliki income type Working dan juga occupation type Laborers, sales staff dan drivers dengan memberikan syarat yang lebih ketat untuk mengurangi client yang memiliki kesulitan membayar
- Berdasarkan feature importance terlihat bahwa feature yang paling penting yang menentukan apakah client akan kesulitan membayar atau tidak adalah feature EXT_Source_2 dan EXT_Source_3 yang menandakan bahwa terdapat faktor external yang belum terindikasi dan mungkin memiliki feature yang berhubungan dengan credit score.
- Perlu adanya explorasi lebih lanjut yang dapat menambahkan insight dan feature baru score sehingga dapat meningkatkan akurasi dan feature yang lebih tepat yang mempengaruhi credit score.



Click Here To
See more On
Github

