



ADD
COMPANY NAME

HOME CREDIT SCORE CARD MODEL

By Daffa Andika Zain

github:

https://github.com/DaffaZain/Home_Credit_Score_Card_Model_recommendation





HOME CREDIT
INDONESIA

BACKGROUND

Banyak orang berjuang untuk mendapatkan pinjaman karena kredit yang tidak mencukupi atau tidak ada Sejarah. Kredit Rumah berupaya memperluas inklusi keuangan bagi yang tidak memiliki rekening bank populasi dengan memberikan pengalaman meminjam yang positif dan aman. Untuk membuat yakin populasi yang kurang terlayani ini memiliki pengalaman pinjaman yang positif. Kredit Rumah membuat berbagai data alternatif untuk memprediksi kemampuan pembayaran klien mereka. Melakukan jadi akan memastikan bahwa klien yang mampu membayar kembali tidak ditolak dan bahwa pinjaman adalah diberikan dengan pokok, jatuh tempo, dan kalender pelunasan yang akan memberdayakan mereka klien untuk menjadi sukses.

DATA SOURCE, OBJECTIVE, ACTIONS

Data Source

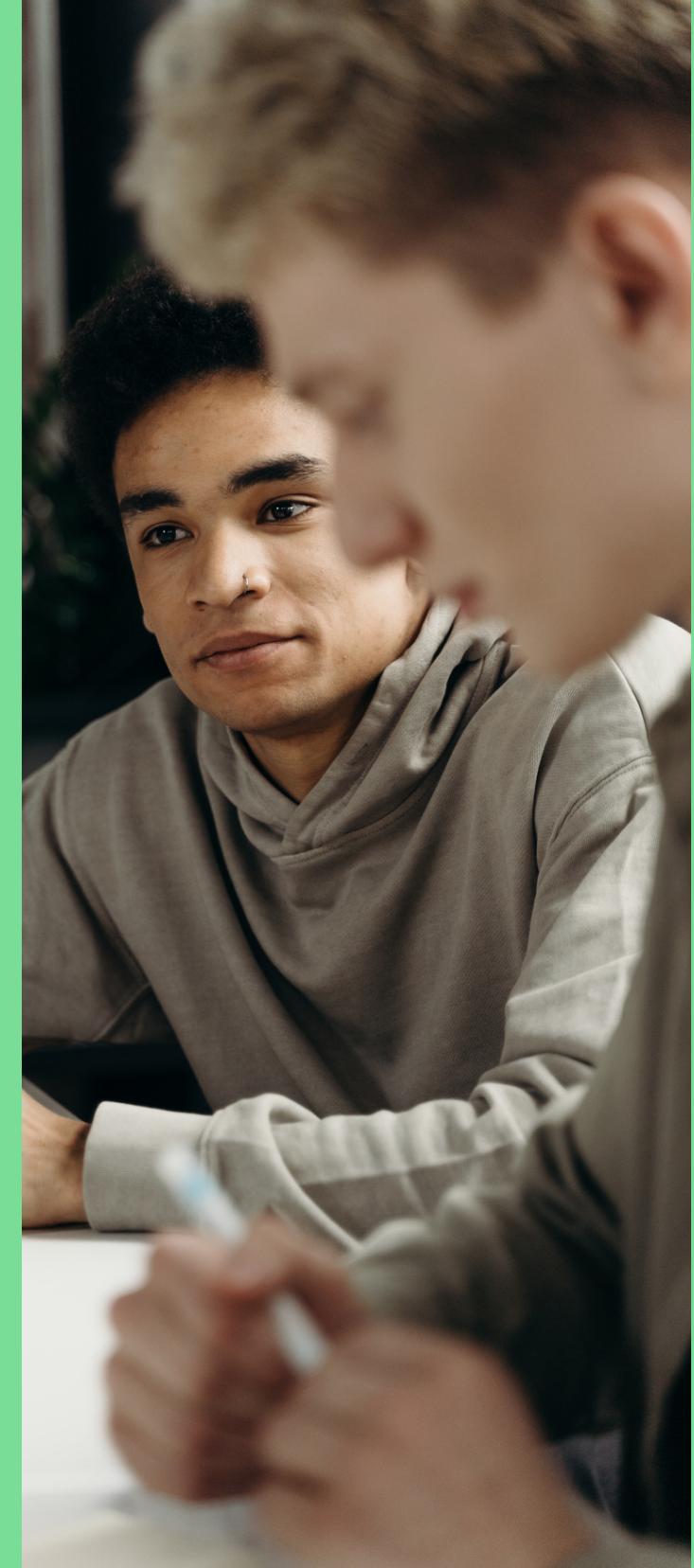
Data yang digunakan adalah Data train dan Data test. dimana dipecah menjadi dua file untuk data train (dengan TARGET) dan uji (tanpa TARGET).

Objektive

1. Identifikasi karakteristik klien potensial yang akan mengalami kesulitan membayar kembali pinjaman dan siapa yang tidak mau. 2. Memprediksi kemampuan pembayaran klien.

Steps

1. Melakukan pembersihan data, dan visualisasi untuk wawasan bisnis.
2. Membangun model dengan algoritma pembelajaran mesin.
3. Memprediksi kemampuan pembayaran klien pada uji data aplikasi dengan model pembelajaran mesin terbaik.
4. Memberikan rekomendasi bagi perusahaan untuk meningkatkan kemampuan nasabahnya dalam mengajukan pinjaman





Data PreProcessing on Data Train

EDA
Temukan pola, dan
struktur
Dataset

DATA CLEANING

MODEL BUILDING

Bivariat
Visualisasi

Visualisasi dari hubungan
antara 2 Fitur

Mendeteksi Duplikasi
Tidak ada baris
duplicat

Pengkodean Label

Multivarian
Visualisasi

Visualisasi dari
hubungan lebih banyak
dari 2 fitur

Menangani Nilai yang Hilang
Ada beberapa kolom yang
dijatuhkan dan sisanya
diperhitungkan

Pemilihan Fitur

Menangani Data yang
Tidak Seimbang

Mendeteksi Outlier
Ada beberapa kolom yang
memiliki outlier, tetapi diputuskan
outlier tidak akan dihapus

Membangun Model

Evaluasi Model



Data PreProcessing on Data Test

DATA CLEANING

Mendeteksi Duplikasi
Tidak ada baris
duplicat

Menangani Nilai yang Hilang

Ada beberapa kolom yang
dijatuhkan dan sisanya
diperhitungkan

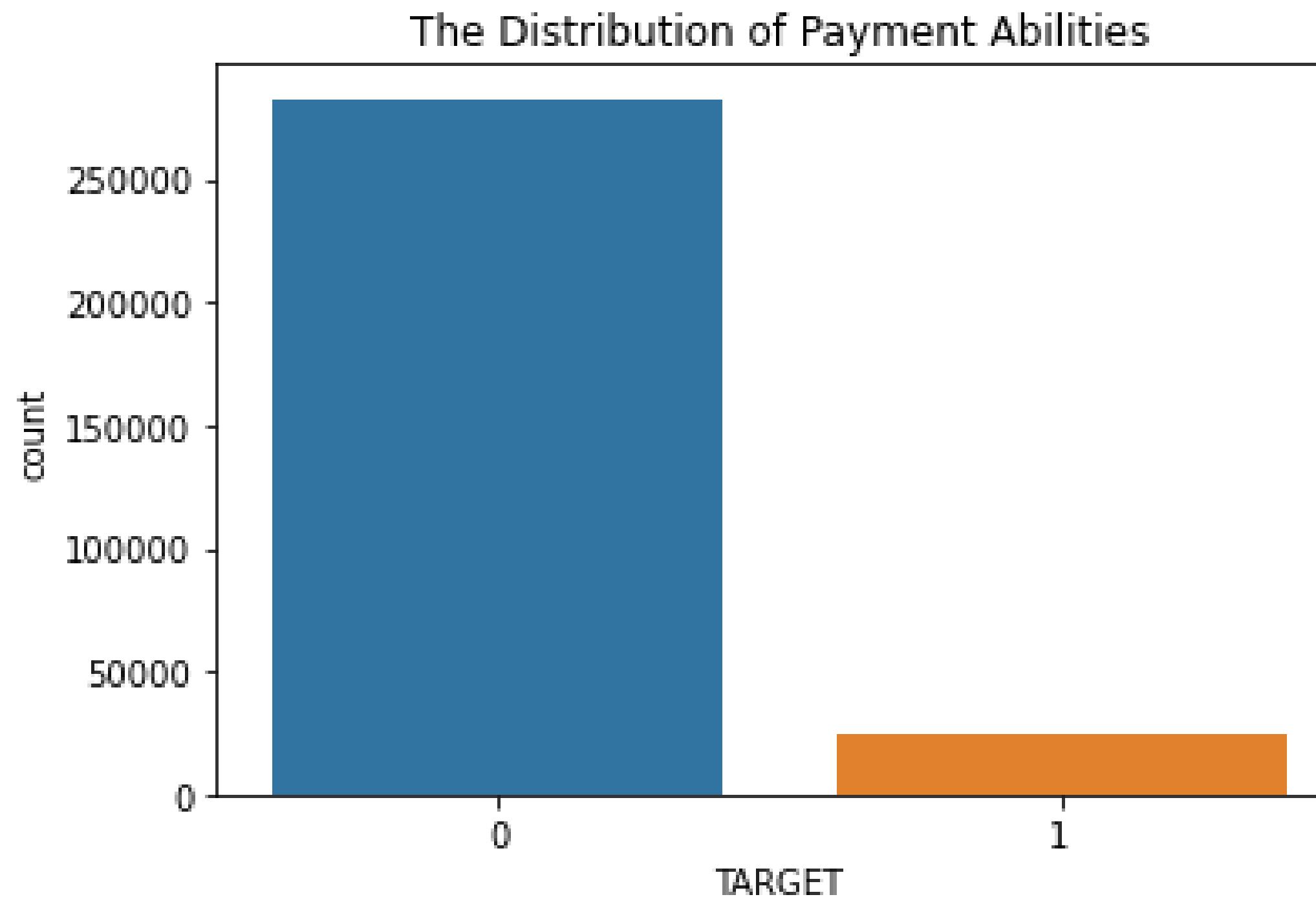
Pengkodean Label

Ubah non-numerik menjadi
numerik Label

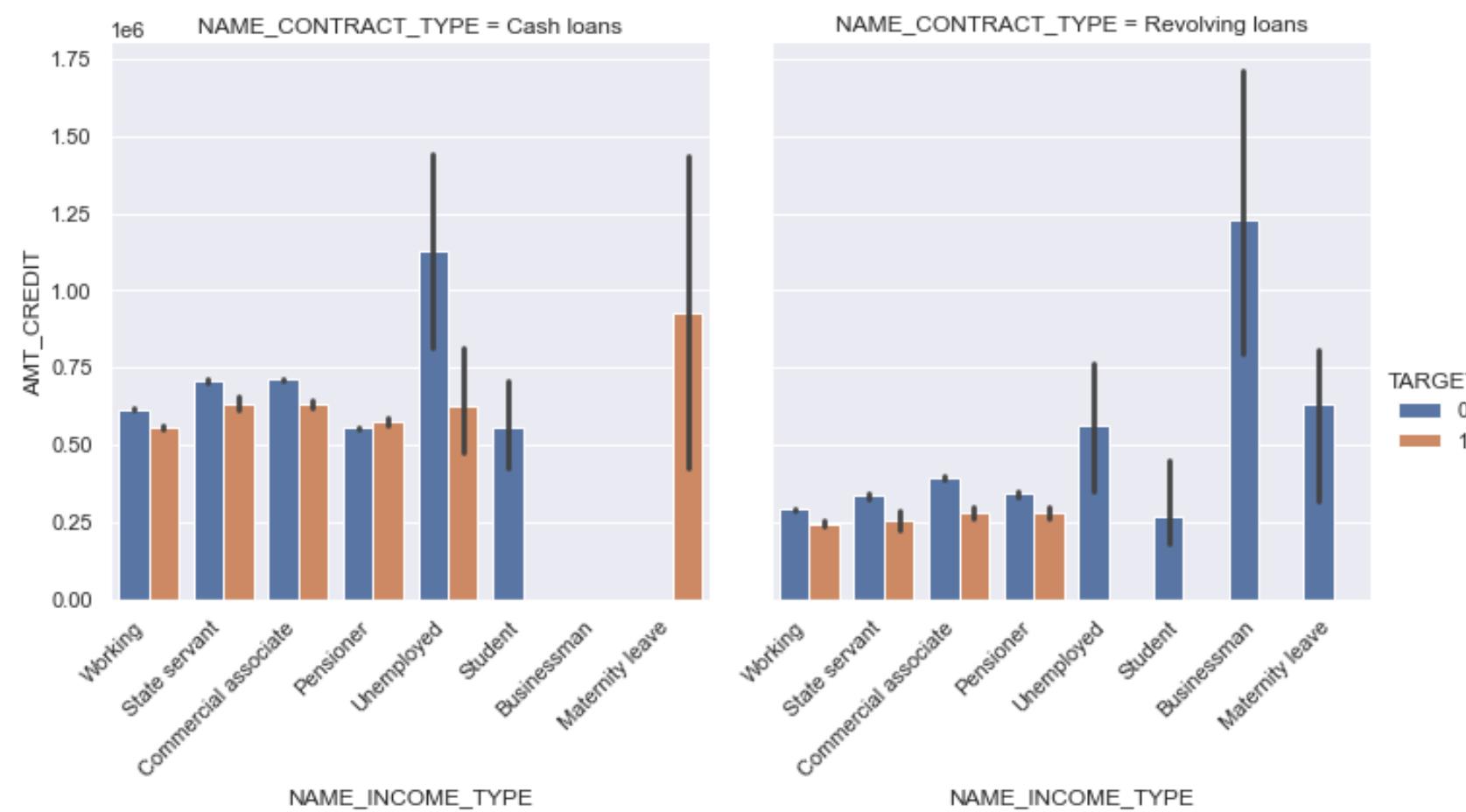
PREDICTION

Prediksi kemampuan
pembayaran klien
dengan yang terbaik
model pembelajaran
mesin yang diperoleh
sebelumnya

Exploratory Data Analysis (EDA)



Business Insight

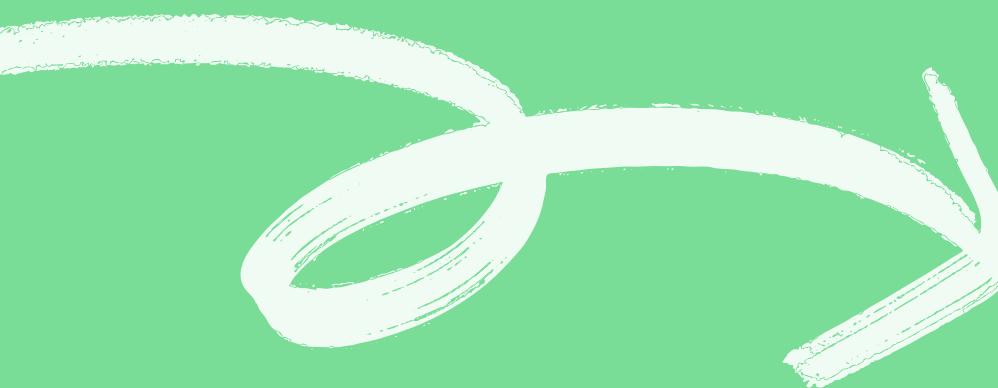


Semua klien siswa tidak mengalami kesulitan membayar kembali pinjaman apakah dengan uang tunai pinjaman atau bergulir pinjaman untuk yang rendah ke kredit menengah jumlah pinjaman.

- Untuk jenis pendapatan cuti hamil dengan pinjaman tunai, semua klien memiliki masalah dalam membayar kembali pinjaman untuk jumlah kredit menengah dari pinjaman. Sementara semua klien dengan cuti hamil dan pinjaman bergulir tidak mengalami kesulitan membayarkan kembali pinjaman.
- Untuk klien yang menganggur dengan pinjaman tunai, lebih dari 50% klien memiliki masalah dalam membayar kembali pinjaman dengan media jumlah kredit pinjaman. Sementara semua klien yang menganggur dengan pinjaman bergulir tidak mengalami kesulitan untuk membayar kembali pinjaman tersebut.

Best Machine Learning Model

Algoritma	Akurasi Skor Training	Akurasi Skor Test	Error Margin	Skor ROC
Logistic Regression	67.16%	67.29%	0.13%	0.6728
Gaussian Naive Bayes	60.24%	60.39%	0.15%	0.604
Decision Tree	100%	83.9%	11.74%	0.8826
Random Forest	100%	99.65%	0.35%	0.9965
K-Nearest Neighbor	91.56%	88.07%	3.79%	0.8806
Neural Network	70.01%	69.48%	0.58%	0.6948



Business recommnedation

Berdasarkan eksplorasi kami dari EDA ke model pembelajaran mesin, berikut adalah beberapa Rekomendasi yang akan berguna untuk meningkatkan bisnis kami :

1. Pinjaman tunai memberi kami lebih banyak klien daripada Pinjaman bergulir tetapi masalahnya membayar kembali pinjaman hampir 2 kali lipat lebih tinggi. Kami dapat membuat kampanye untuk klien kami untuk memilih Pinjaman bergulir sebagai jenis kontrak saat mengajukan pinjaman.
2. Berikan kampanye untuk jenis penghasilan Cuti melahirkan, Pengangguran, Pengusaha, dan pelajar karena jumlahnya sedikit mengajukan pinjaman.
3. Sebagian besar klien kami berada di rentang usia menengah 40-60 diikuti oleh rentang usia muda 20-40 tetapi masalah melunasi pinjaman hampir 2 kali lipat. Kita harus memberikan jenis kontrak yang tepat jika ada yang berusia muda ingin mengajukan pinjaman sehingga mereka tidak akan kesulitan dalam membayar kembali pinjaman.
4. Model terbaik untuk memprediksi Home Credit Score menggunakan **Extra Trees Classifier**. Karena nilai Training dan Test Akurasi tinggi dan tidak berbeda sama sekali. Model ini juga akan bagus jika kita mencoba mengurangi jumlah False Negative (Kondisi dimana klien memiliki masalah dalam membayar kembali pinjaman tetapi prediksi kami memberikan hasil bahwa mereka tidak memiliki masalah) dalam hal ini kami juga akan melihat **Recall score**.