



ANALISIS DAN PREDIKSI RISIKO NASABAH GAGAL BAYAR

PROJECT-BASED INTERNSHIP BY RAKAMIN ACADEMY



FINAL TASK

PRESENTED BY SHIBA SALSABILLA

PROBLEM RESEARCH

Home Credit adalah perusahaan pembiayaan berbasis teknologi yang menyediakan layanan kredit tanpa kartu, seperti cicilan untuk pembelian barang elektronik, furnitur, smartphone, dan produk konsumen lainnya. Saat ini Home Credit ingin mengoptimalkan sistem prediksi client kredit.

Client sebenarnya layak, namun ditolak karena skor kredit rendah.

Client tidak layak (berisiko gagal bayar), namun tetap disetujui.

Akibatnya **kerugian finansial** bagi perusahaan dan **menghambat inklusi keuangan**



FINAL TASK



GOALS

Memprediksi nasabah yang berpotensi mengalami kesulitan membayar kredit agar risiko kerugian dapat diminimalkan.



OBJECTIVE

- Membangun model klasifikasi untuk mendeteksi potensi kesulitan bayar.
- Mengatasi ketidakseimbangan data dengan teknik resampling.
- Melakukan seleksi fitur penting untuk meningkatkan performa model.
- Melakukan evaluasi model menggunakan metrik yang sesuai.



METRICS

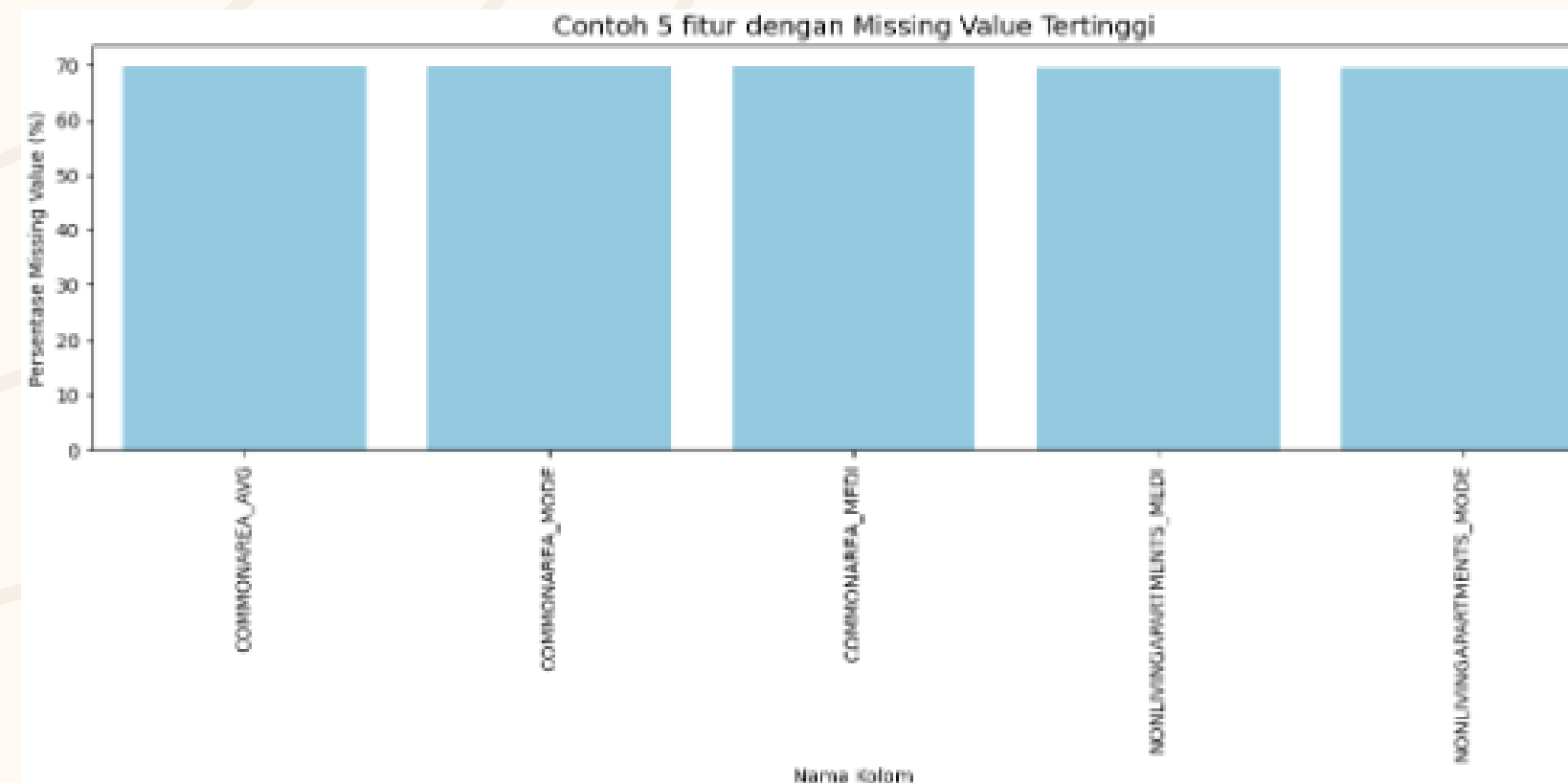
F1-Score, Precision & Recall, ROC-AUC Score, Confusion Matrix

FINAL TASK

DATA PREPROCESSING

Handling Missing Value

Menghapus kolom dengan >40% missing value,
Imputasi dengan median untuk variabel numerik
dan modus untuk variabel kategorik



Feature Engineering

Konversi hari ke tahun:

DAYS_BIRTH → AGE

DAYS_EMPLOYED → EMP_LENGTH etc

Important ratio

$\text{INCOME_CREDIT_RATIO} = \text{INCOME} / \text{CREDIT}$

$\text{ANNUITY_INCOME_RATIO} = \text{ANNUITY} / \text{INCOME}$

Interaksi

$\text{AGE_EMP_INTERACTION} = \text{AGE} \times \text{EMP_LENGTH}$

FINAL TASK

DATA PREPROCESSING

Transformasi dan Encoding

- **Log Transformation** untuk mengatasi skewness: LOG_AMT_INCOME_TOTAL, LOG_AMT_CREDIT
- **Encoding biner:** NAME_CONTRACT_TYPE, CODE_GENDER, FLAG_OWN_CAR, FLAG_OWN_REALTY
- **Penyederhanaan kategori:** Kategori langka pada NAME_TYPE_SUITE, NAME_INCOME_TYPE, OCCUPATION_TYPE dikelompokkan ke dalam 'Other', 'Office Job', dll.
- **Drop data anomali** pada isi gender 'XNA'

Feature Scaling

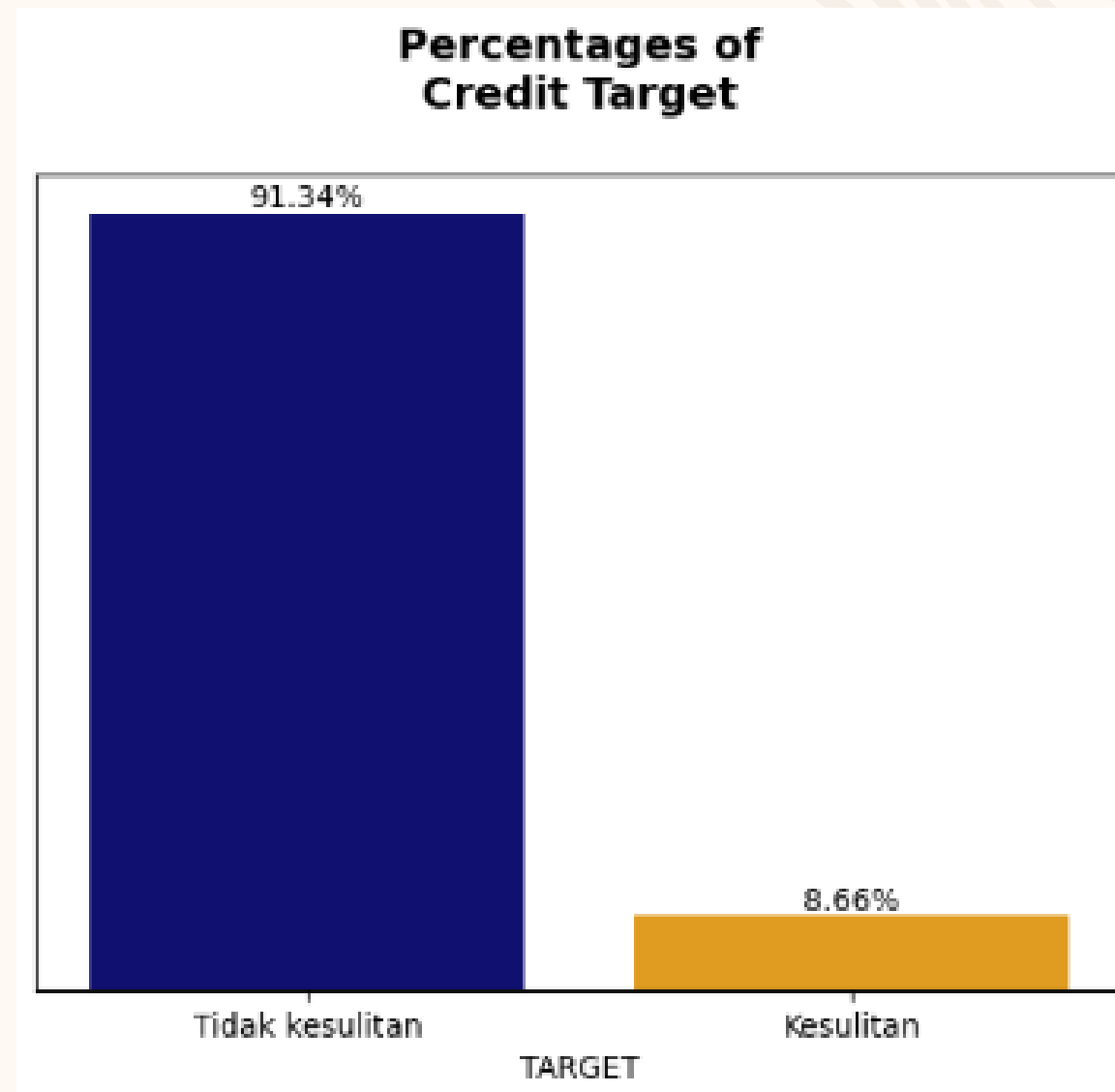
Menambahkan RobustScaler untuk menormalkan data fitur, agar fitur punya skala yang sama dan tahan terhadap outlier

Handling Imbalance Data

Percobaan pada 3 metode dan evaluasi dengan classification_report_imbalanced, di dapatkan **Random Over Sampling** memberikan performa terbaik berdasarkan metrik evaluasi.

FINAL TASK

DATA VISUALIZATION



Dataset yang digunakan merupakan data client kredit Home Credit Indonesia. Terdiri dari 122 variabel dan 307511 data.

TARGET

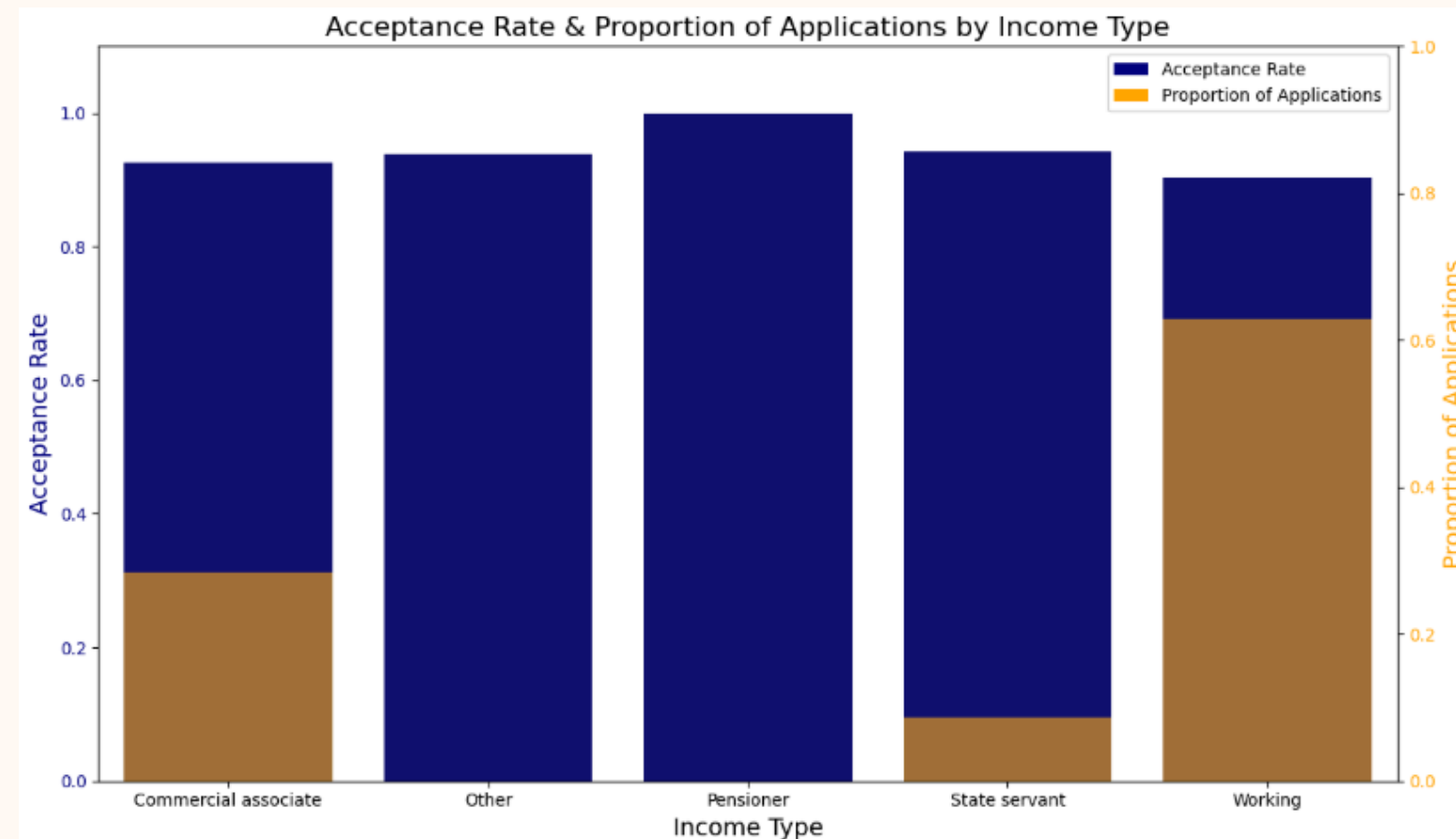
0 : Client tidak memiliki kesulitan bayar kredit

1 : Client memiliki kesulitan bayar kredit

Adanya ketidakseimbangan kelas yang sangat signifikan dalam data.

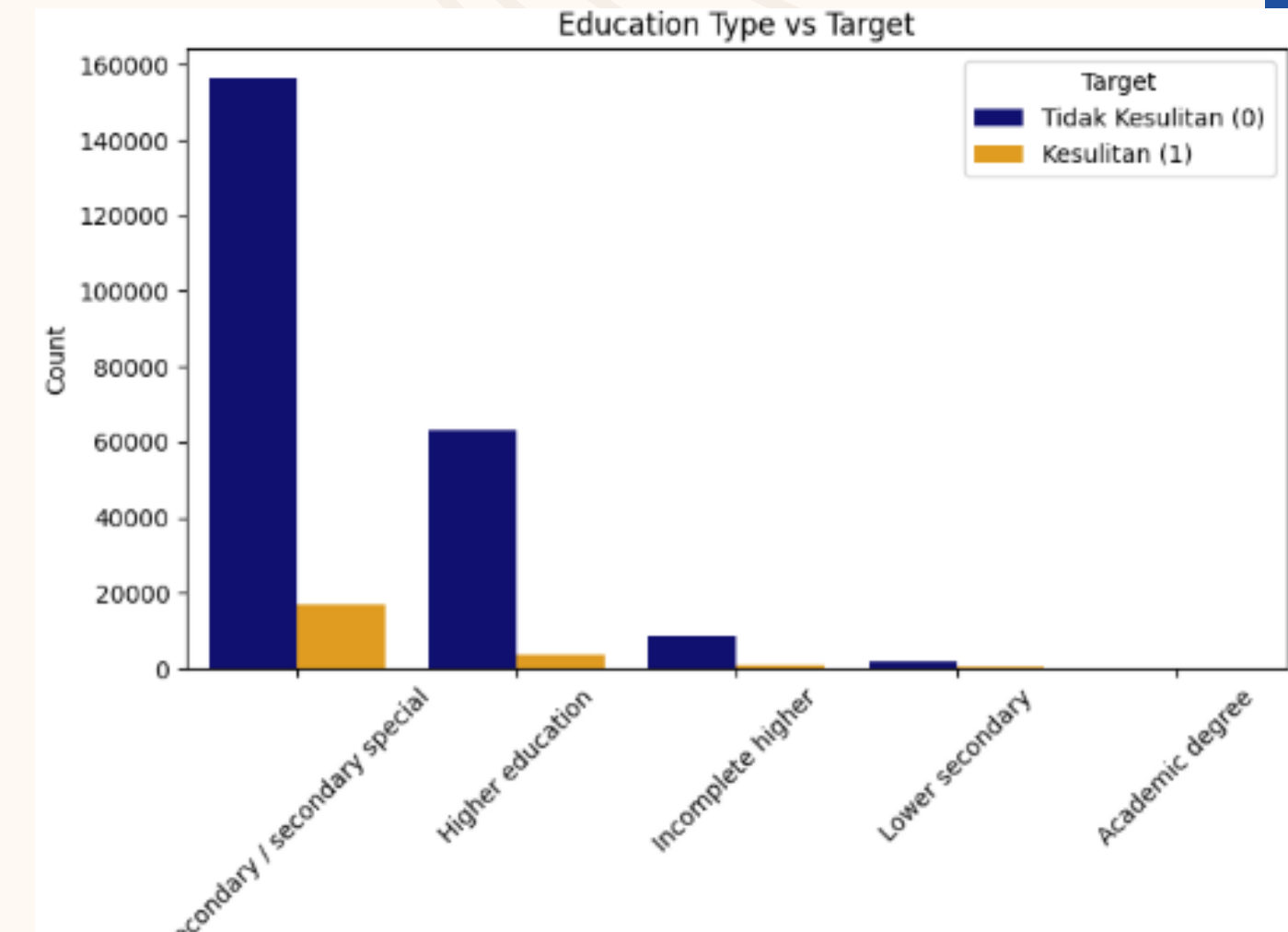
FINAL TASK

SOME INSIGHT



Pensiunan memiliki acceptance rate tertinggi, namun sangat sedikit yang mengajukan kredit, diikuti oleh PNS.

Sementara pekerja aktif mendominasi pengajuan kredit meskipun acceptance rate-nya sedikit lebih rendah



Pendidikan yang tinggi cenderung memiliki pengajuan lebih banyak dan tingkat kesulitan relatif rendah.

Pendidikan menengah ke bawah memiliki persentase kesulitan lebih tinggi, artinya mereka lebih berisiko gagal bayar.

FINAL TASK

DATA MODELLING



Tuning Model	F1-score (class 1)	AUC Test
Logistic Regression	0.24	0.7115
Random Forest	0.26	0.7264
XGBoost	0.26	0.7348

Berdasarkan hasil evaluasi di atas, model **XGBoost** dipilih sebagai model paling optimal dalam memprediksi kelas 1 atau client yang memiliki kesulitan bayar kredit. Evaluasi fokus pada kelas 1 karena sebagai target utama bisnis, jika model salah prediksi kelas 1 sebagai kelas 0, maka akan menimbulkan kerugian finansial, karena nasabah berisiko dianggap aman. Pada kelas 0 cenderung memiliki F1-Score yang tinggi hingga 93%.

FINAL TASK

RECOMENDATION

1. **Targeted marketing untuk Pensiunan dan PNS** dengan sosialisasi khusus (iklan, promosi bunga ringan, referral program)
2. **Wajib pelatihan Literasi Keuangan yang interaktif**, simulasi cicilan dan tips manajemen utang sebelum pencairan kredit untuk client berisiko tinggi (Pendidikan menengah ke bawah)
3. **Kembangkan Produk Kredit dengan Fitur Khusus**, seperti Suku bunga lebih rendah atau cashback jika membayar tepat waktu untuk client berkualitas dan tenor pendek, angsuran ringan, atau grace period awal untuk client berisiko gagal bayar.
4. **Membuat Early Warning System (EWS)** untuk memantau client.

FINAL TASK



[Here for Github](#)

FINAL TASK
