

Virtual Internship Experience

Home Credit Scorecard Model
Oleh : M.Irfansyah

**HOME
CREDIT**
Anda Bisa!

Content



1. Problem Research
2. Data Pre-Processing
3. Data Visualization and Business Insight
4. Machine Learning Implementation and Evaluation
5. Business Recommendation

Problem Research



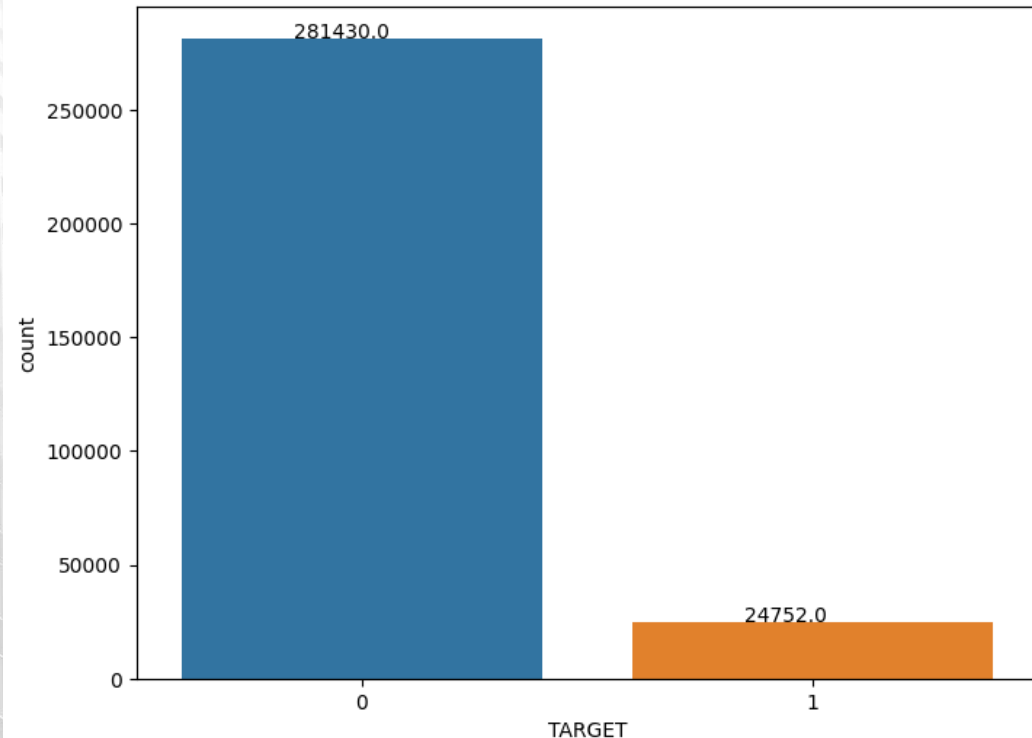
1. Home Credit saat ini sedang menggunakan berbagai macam metode statistik dan Machine Learning untuk membuat prediksi skor kredit. Dengan menggunakan data-data yang telah disediakan diharapkan dapat memastikan pelanggan yang mampu melakukan pelunasan tidak ditolak ketika melakukan pengajuan pinjaman, dan pinjaman dapat diberikan dengan principal, maturity, dan repayment calendar yang akan memotivasi pelanggan untuk sukses.
2. Dataset yang digunakan adalah `application_train.csv` dan `application_test.csv`
3. Model Machine Learning yang digunakan adalah XGBoost Classifier dan Decision Tree Classifier

Data Pre-Processing

1. Menentukan fitur-fitur yang akan digunakan .
2. Melakukan pemeriksaan missing value .
3. Menghapus missing value.
4. Memeriksa adanya data duplikat.
5. Menghapus data duplikat.
6. Mengecek ejaan kata dan menemukan kesalahan pada kolom CODE_GENDER dan NAME_INCOME_TYPE, yang kemudian dihapus.
7. Menambahkan kolom AGE.
8. Mengubah data kategorikal menjadi biner.
9. Membagi dataset menjadi 80% data training dan 20% data testing.

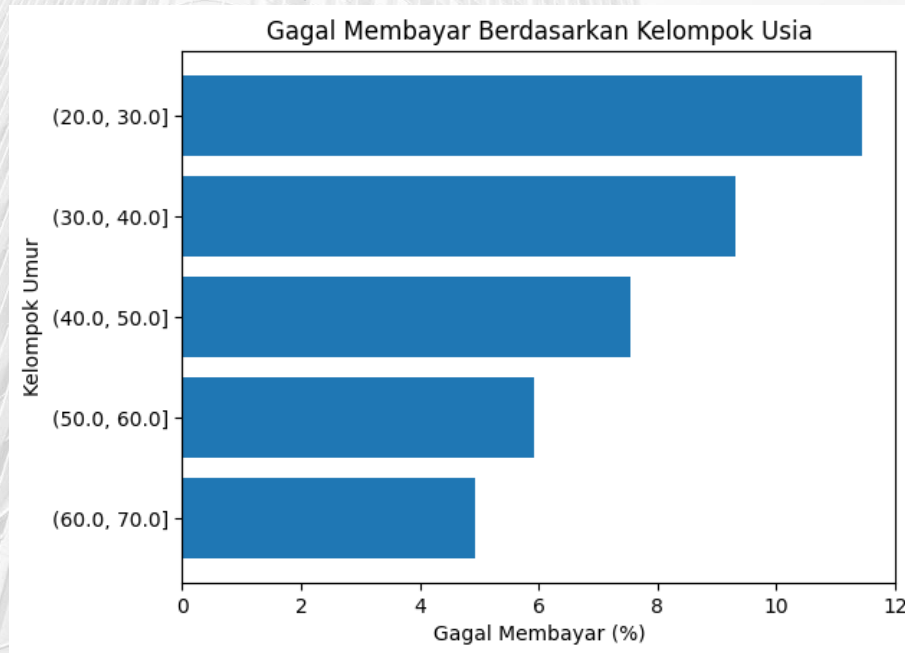
Data Visualization and Business Insight

Mayoritas klien (281.430) berhasil membayar pinjaman tepat waktu, sedangkan 24.752 klien mengalami kesulitan dalam pembayaran.



Data Visualization and Business Insight

Klien yang berusia muda memiliki risiko yang lebih tinggi untuk gagal membayar pinjaman. Tingkat kegagalan pembayaran melebihi 10% untuk kelompok usia terendah, sementara tingkat kegagalan tersebut di bawah 5% untuk kelompok usia tertua.



Machine Learning Implementation and Evaluation



Ada 2 Model Machine Learning menggunakan hyperparameter tuning

1. XGBoost Classifier
2. Decision Tree Classifier

	Model	Score
0	XGBoost Classifier	0.916603
1	DecisionTreeClassifier	0.916358

Setelah menganalisis data Home Credit, ditemukan bahwa model terbaik yang digunakan adalah **XGBoost Classifier**. Model ini memiliki **tingkat akurasi sebesar 0.916603**.

Business Recommendation



1. Home Credit Indonesia perlu memberikan perhatian istimewa kepada pelanggan yang memenuhi kriteria berikut: memilih pinjaman tunai, memiliki pekerjaan, sudah menikah, dan memiliki rumah atau apartemen. Kelompok pelanggan ini memiliki tingkat kemampuan pembayaran yang tinggi. Agar mereka terus menjadi pelanggan kita, kita harus memberikan beberapa benefit kepada mereka seperti mengurangi jumlah angsuran yang harus dibayarkan, meningkatkan batas pinjaman yang dapat diakses oleh pelanggan tersebut dan memberikan fleksibilitas pada jangka waktu pembayaran
2. Berdasarkan visualisasi data, didapat bahwa klien berusia muda memiliki risiko yang tinggi untuk gagal membayar pinjaman. Untuk mengatasi masalah ini, Home Credit Indonesia setidaknya memberikan edukasi keuangan terhadap klien muda, meningkatkan pengawasan dan juga menyesuaikan produk pinjaman.