

HOME CREDIT SCORECARD MODEL

Fiorent Arie Hernanti

Project Based Internship - Rakamin Academy



PROBLEM STATEMENT

Home Credit dihadapi tantangan dalam memastikan bahwa pelanggan yang mampu melakukan pelunasan tidak ditolak ketika mengajukan pinjaman. Dengan begitu, perusahaan dapat menjaga bahkan meningkatkan profitabilitasnya.

Business Metrics

Approval Rate

Meningkatkan approval rate untuk client yang memiliki kemampuan pelunasan yang tinggi.

Goal

Meminimalkan penolakan terhadap customer yang mampu melakukan pelunasan dan meminimalisir jumlah customer yang memiliki kesulitan pembayaran pinjaman

Objective

1. Menganalisis faktor-faktor yang mempengaruhi keberhasilan atau tidaknya pelunasan pinjaman.
2. Membangun model machine learning untuk memprediksi kemampuan customer melakukan pelunasan pinjaman

DATASET INFORMATION

Data disediakan oleh Home Credit, yang merupakan perusahaan pembiayaan berbasis teknologi yang menyediakan layanan pinjaman kepada clientnya.

1

Application_train/test : Tabel utama yang menyimpan informasi mengenai setiap pengajuan pinjaman di Home Credit. Pada data application_train terdapat TARGET yang menunjukkan 0 (pinjaman berhasil dilunasi) dan 1 (pinjaman tidak dilunasi)

2

Bureau : Informasi kredit customer sebelumnya dari lembaga keuangan lain. Setiap kredit yang dilakukan sebelumnya memiliki baris tersendiri di bureau, namun satu pinjaman di application data dapat memiliki beberapa kredit sebelumnya.

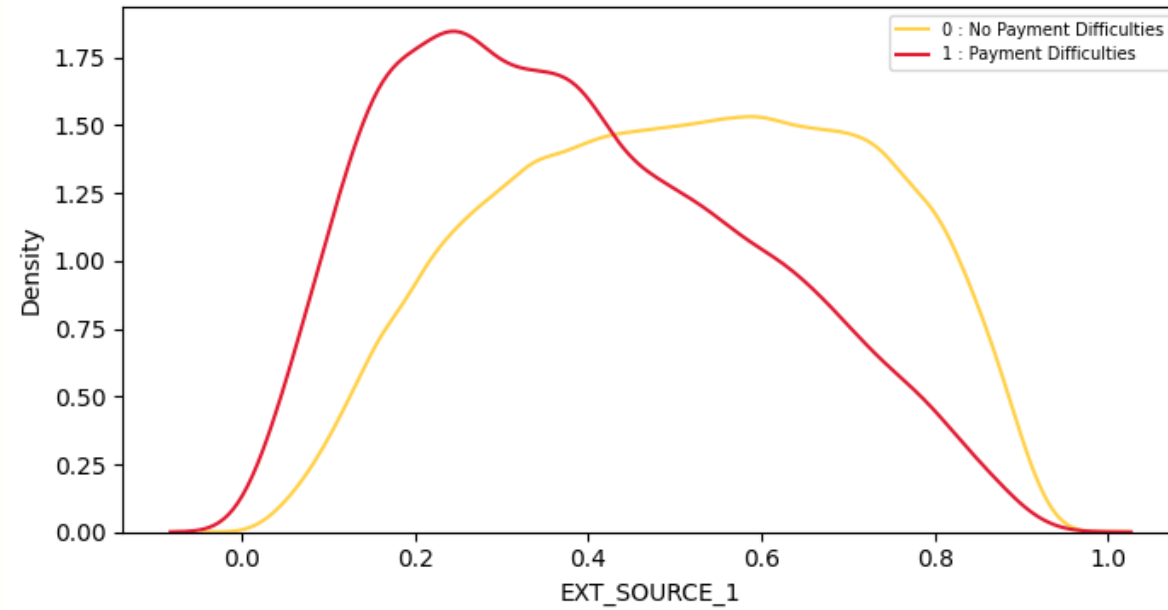
3

Previous_application : Menyimpan informasi mengenai pengajuan pinjaman yang sebelumnya diajukan oleh client di Home Credit.

BUSINESS INSIGHT

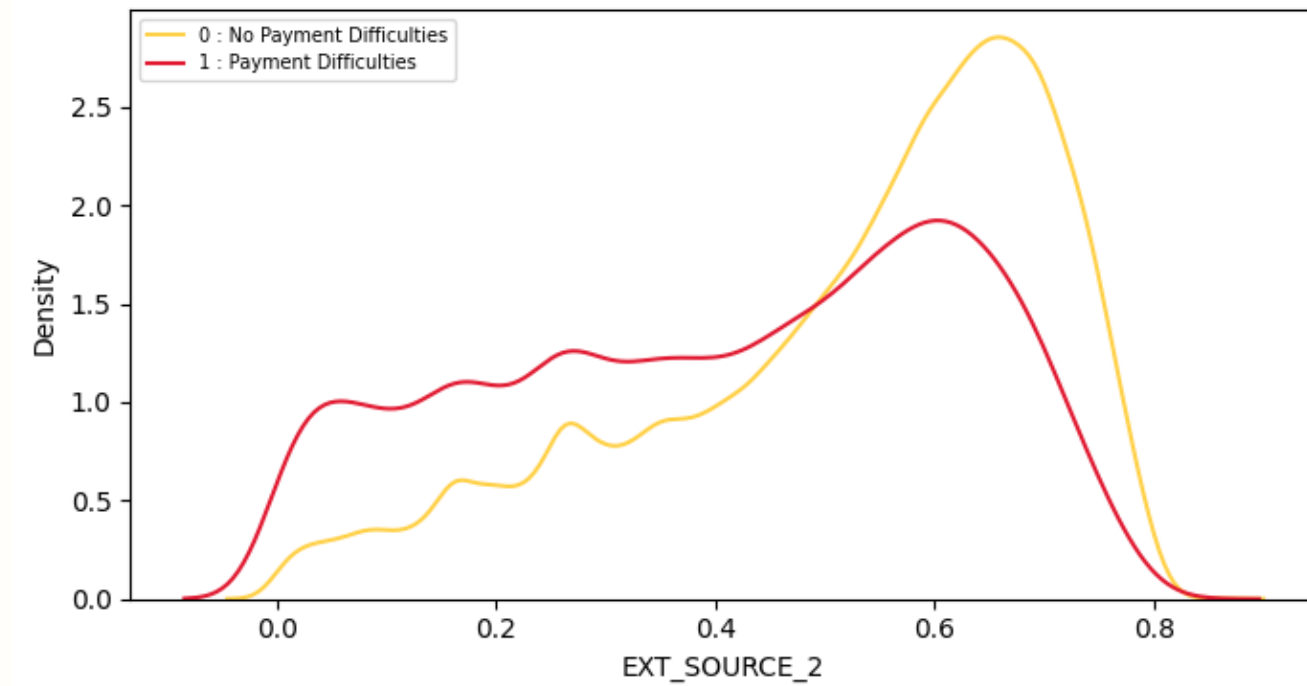
3

Distribution of EXT_SOURCE_1 by Target



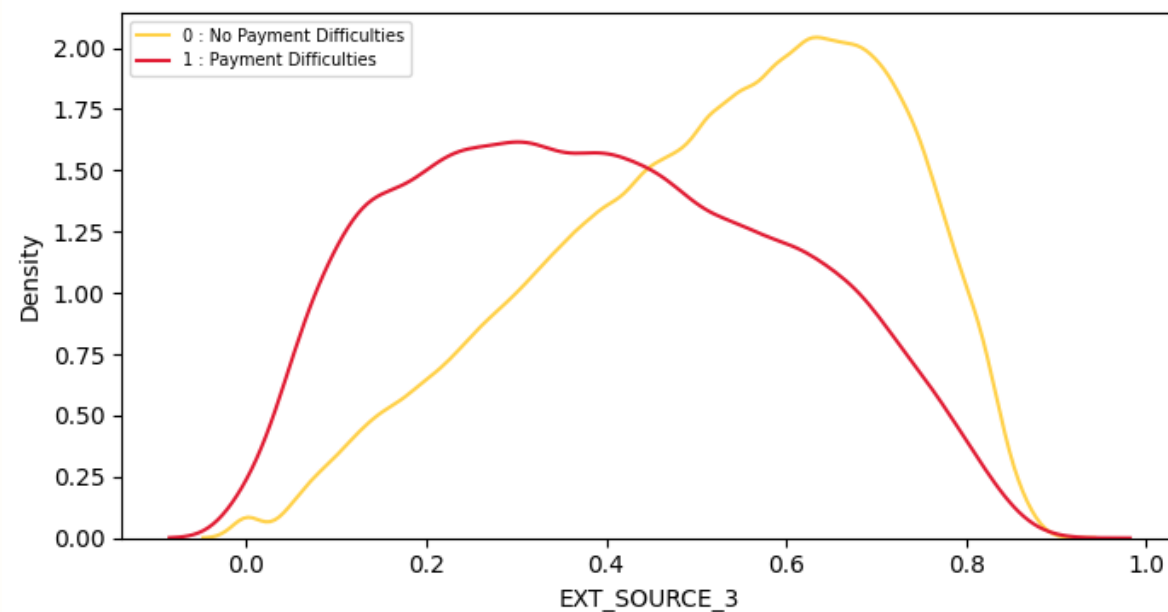
semakin kecil nilai EXT_SOURCE_1,
semakin besar kemungkinan
gagal melakukan pelunasan

Distribution of EXT_SOURCE_2 by Target



Skor EXT_SOURCE_3 > 0.5 cenderung melunasi pinjamannya

Distribution of EXT_SOURCE_3 by Target



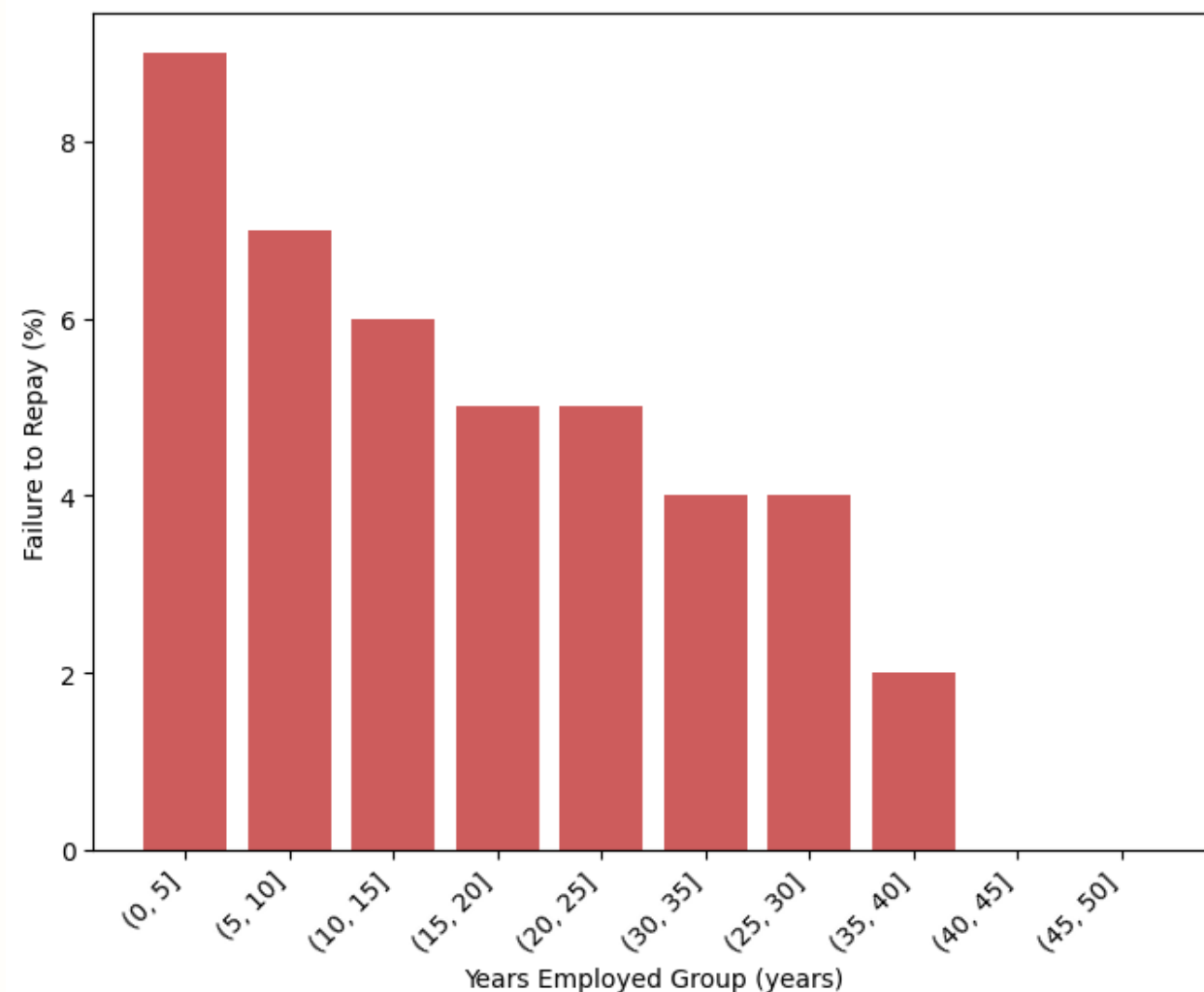
EXT_SOURCE_2 dalam
rentang kurang dari 0 - 0.5
berkemungkinan tinggi
tidak melunasi pinjamannya

Terdapat kecenderungan bahwa semakin kecil skor
eksternal client, semakin besar kemungkinan client
mengalami kesulitan melunaskan pinjamannya

BUSINESS INSIGHT

Masa Bekerja

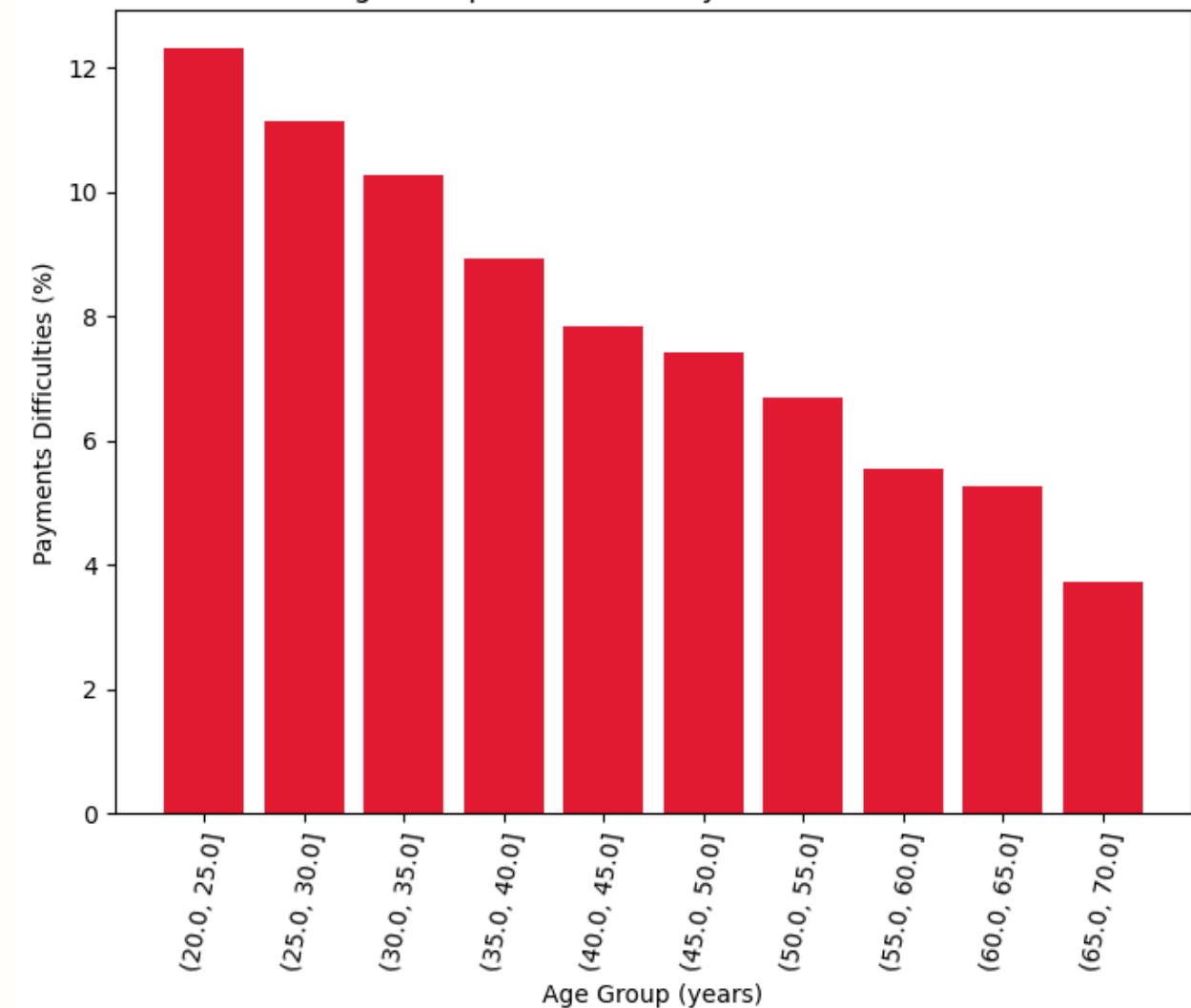
Failure to Repay by Years Employed Group



Semakin lama masa bekerja, persentase client yang gagal melunaskan pinjaman cenderung menurun. Artinya client dengan masa bekerja yang lebih lama memiliki kecenderungan untuk melunasi pinjamannya.

Usia

Age Groups Who Have Payments Difficulties



Client dengan usia yang lebih muda cenderung mengalami kesulitan pembayaran pinjaman. Sebaliknya, klien dengan usia yang lebih tua cenderung untuk dapat melunasi pinjamannya.

DATA PREPROCESSING

● Feature Engineering

- Converting days to years :
 - DAYS_BIRTH
 - DAYS_EMPLOYED
 - DAYS_REGISTRATION
 - DAYS_ID_PUBLISH

● Handling Invalid Value

- Replace XNA values with NaN
- 1000 YEARS_EMPLOYED to NaN

● Handling Missing Values

- Numeric columns : Median
- Categorical columns : Mode

● Feature Encoding

One Hot Encoding (OHE)

● Feature Selection

Correlation, chi-square

● Handling Class Imbalance

SMOTE

● Feature Transformation

StandardScaler

MODELING

Model	ROC-AUC	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.55	0.51	0.54	0.21	0.30
Decision Tree	0.88	0.88	0.87	0.89	0.88
LightGBM	0.97	0.94	0.98	0.90	0.94

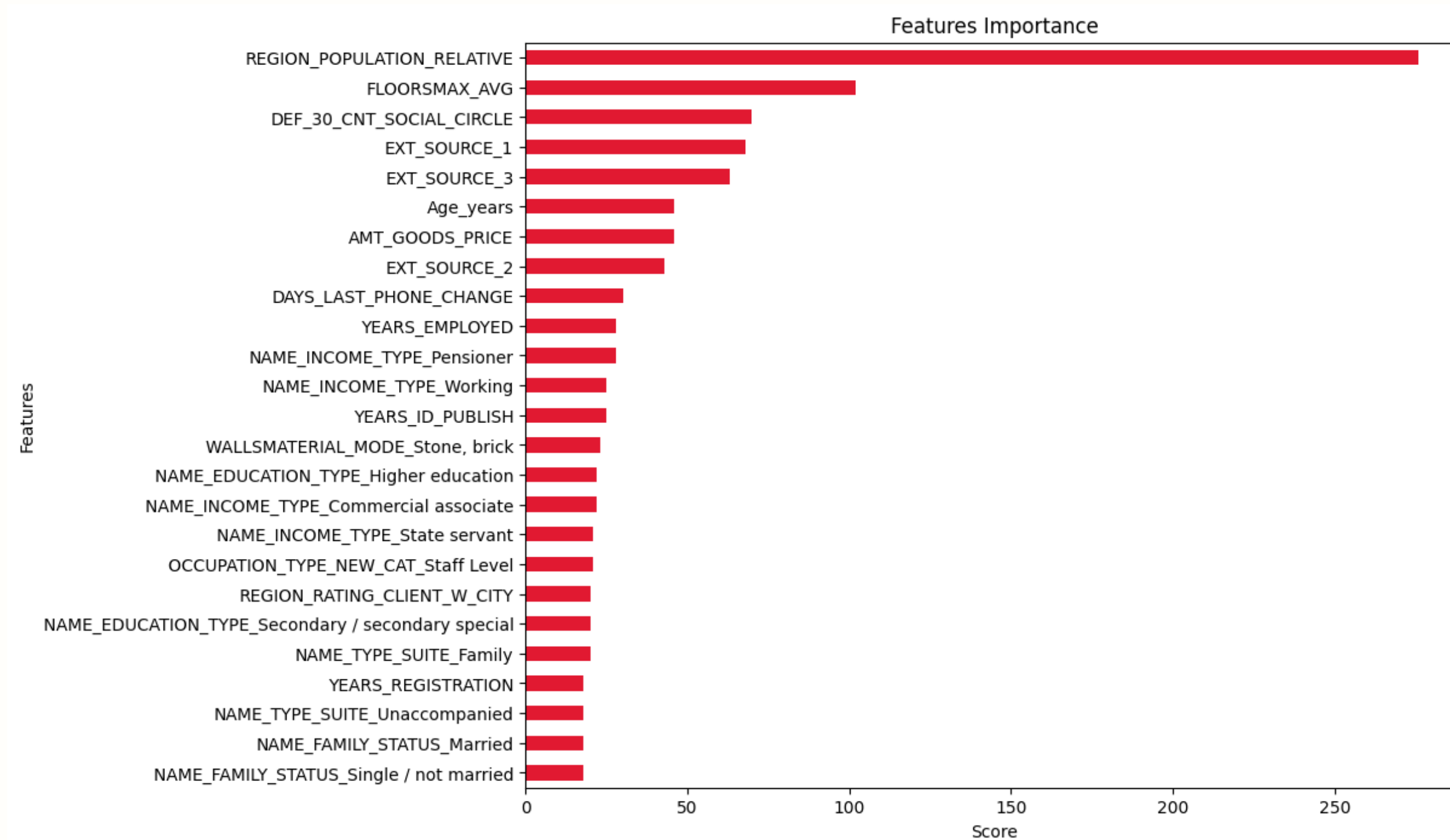
1. Logistic Regression

2. Decision Tree

3. LightGBMClassifier



FEATURES IMPORTANCE



Top 10 Features Importance

1. REGION_POPULATION_RELATIVE
2. FLOORSMAX_AVG
3. DEF_30_CNT_SOCIAL_CIRCLE
4. EXT_SOURCE_1
5. EXT_SOURCE_3
6. Age_years
7. AMT_GOODS_PRICE
8. EXT_SOURCE_2
9. DAYS_LAST_PHONE_CHANGE
10. YEARS_EMPLOYED



BUSINESS RECOMMENDATION

8

- 1. Meningkatkan approval untuk client potensial dengan mempertimbangkan skor eksternalnya karena skor eksternal yang besar cenderung untuk mampu melunaskan pinjamannya. Selain itu, untuk meminimalisir client gagal melunasi pinjamannya, dapat diberikan strategi atau ketentuan khusus untuk mendorong pembayaran tepat waktu.**
- 2. Meningkatkan approval untuk client yang tinggal di wilayah padat populasi sehingga profitabilitas juga dapat meningkat. Namun perlu dibarengi pertimbangan faktor lainnya seperti masa bekerja. Dan pastikan kebijakan tersebut tidak mengabaikan client yang tinggal di wilayah kurang padat.**
- 3. Mempermudah proses pengajuan pinjaman untuk client dengan masa bekerja yang lebih lama, karena mereka cenderung untuk mampu melakukan pelunasan pinjaman. Selain itu, untuk meningkatkan profit perusahaan, client dengan masa bekerja yang lebih lama dapat ditingkatkan limit peminjamannya.**

The slide features a decorative background with three vertical bars on the left: a red bar, an orange bar, and a light gray bar. In the top right and bottom right corners, there are rectangular areas filled with a grid of small, light pink dots.

THANK YOU

Github : https://github.com/fiornth/HomeCredit_Scorecard_Model.git