

# HOME CREDIT DEFAULT RISK

---

Helmy Ramadhani Putra Yongki

# PROBLEM STATEMENT

Banyak orang berjuang untuk mendapatkan pinjaman karena riwayat kredit yang tidak mencukupi atau tidak ada. Dan sayangnya, populasi ini sering dimanfaatkan oleh pemberi pinjaman yang tidak dapat dipercaya.

Home credit merupakan perusahaan penyedia credit ingin memperluas inklusi keuangan bagi masyarakat yang belum tersentuh layanan perbankan dengan memberikan pengalaman meminjam yang positif dan aman.

Untuk mencapai hal tersebut, Home credit ingin dapat memastikan calon nasabah yang mampu melakukan pelunasan tidak ditolak ketika melakukan pengajuan pinjaman, dan pinjaman dapat diberikan dengan principal, maturity, dan repayment calendar yang akan memotivasi pelanggan dengan menggunakan berbagai data alternatif, seperti telekomunikasi dan informasi transaksional untuk memprediksi kemampuan pembayaran klien mereka.

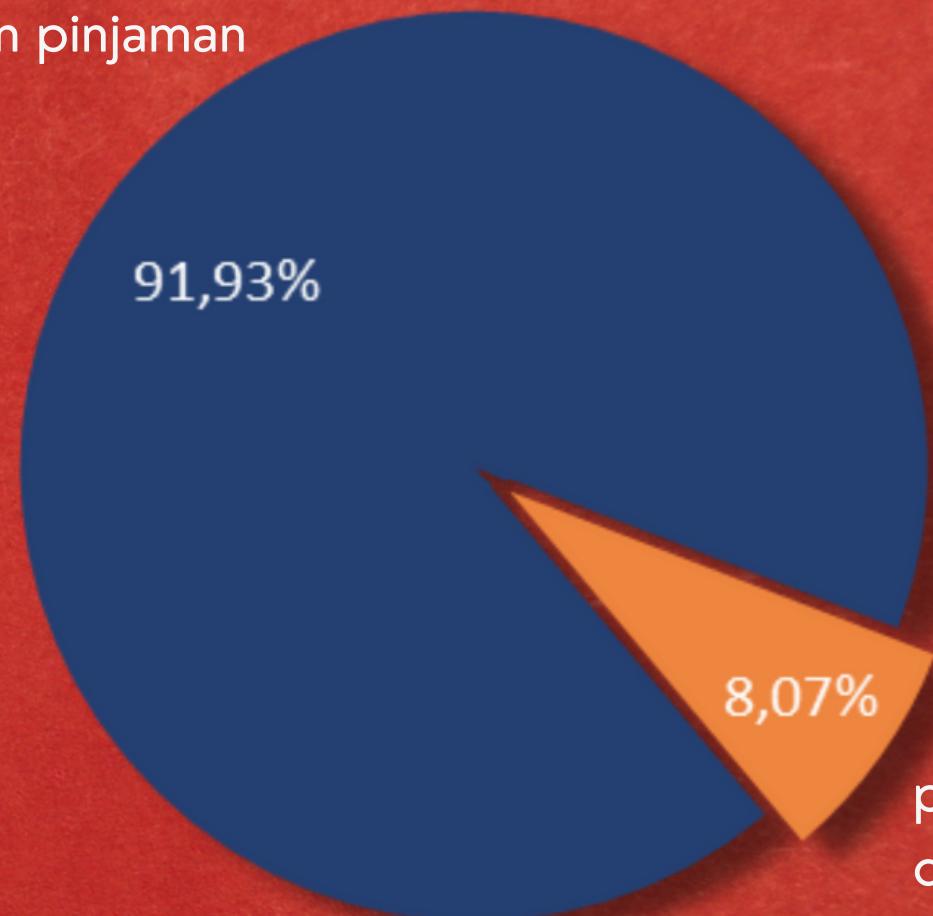
pengajuan pinjaman  
diterima

91,93%

8,07%

pengajuan pinjaman  
ditolak

Pada data pengajuan pinjaman terdapat 307511 nasabah. terdiri dari 91,93% berindikasi akan membayar pinjaman tepat waktu dan 8,07% akan mengalami kesulitan dalam membayar kembali pinjaman



# **GOAL**

Meningkatkan probabilitas nasabah untuk mendapatkan pinjaman dengan memperhatikan data atau variabel alternatif

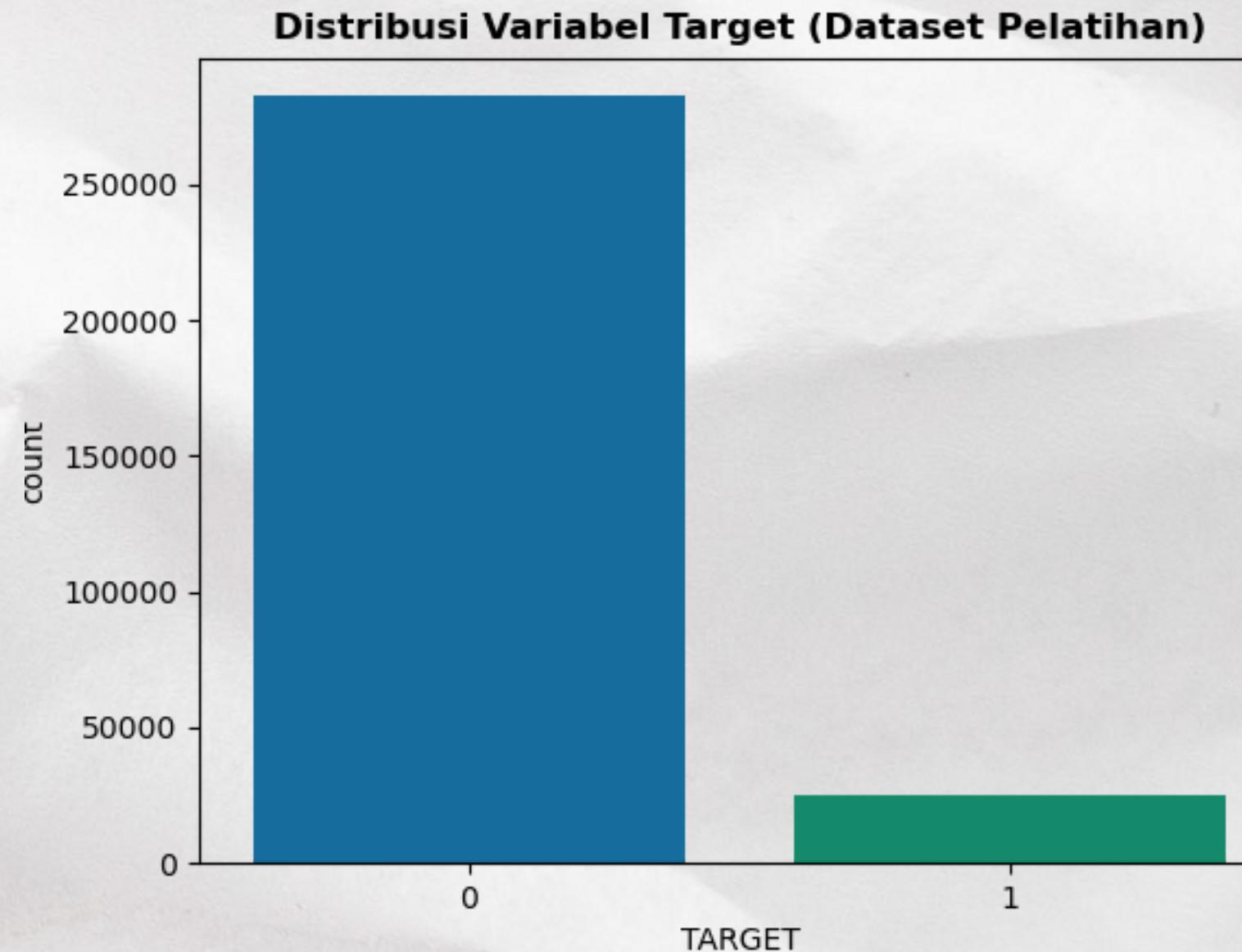
# **OBJECTIVE**

Mendeteksi atau memprediksi nasabah yang diperkirakan mampu melakukan pelunasan tidak ditolak ketika melakukan pengajuan pinjaman

# **BUSINESS METRICS**

take-up rate

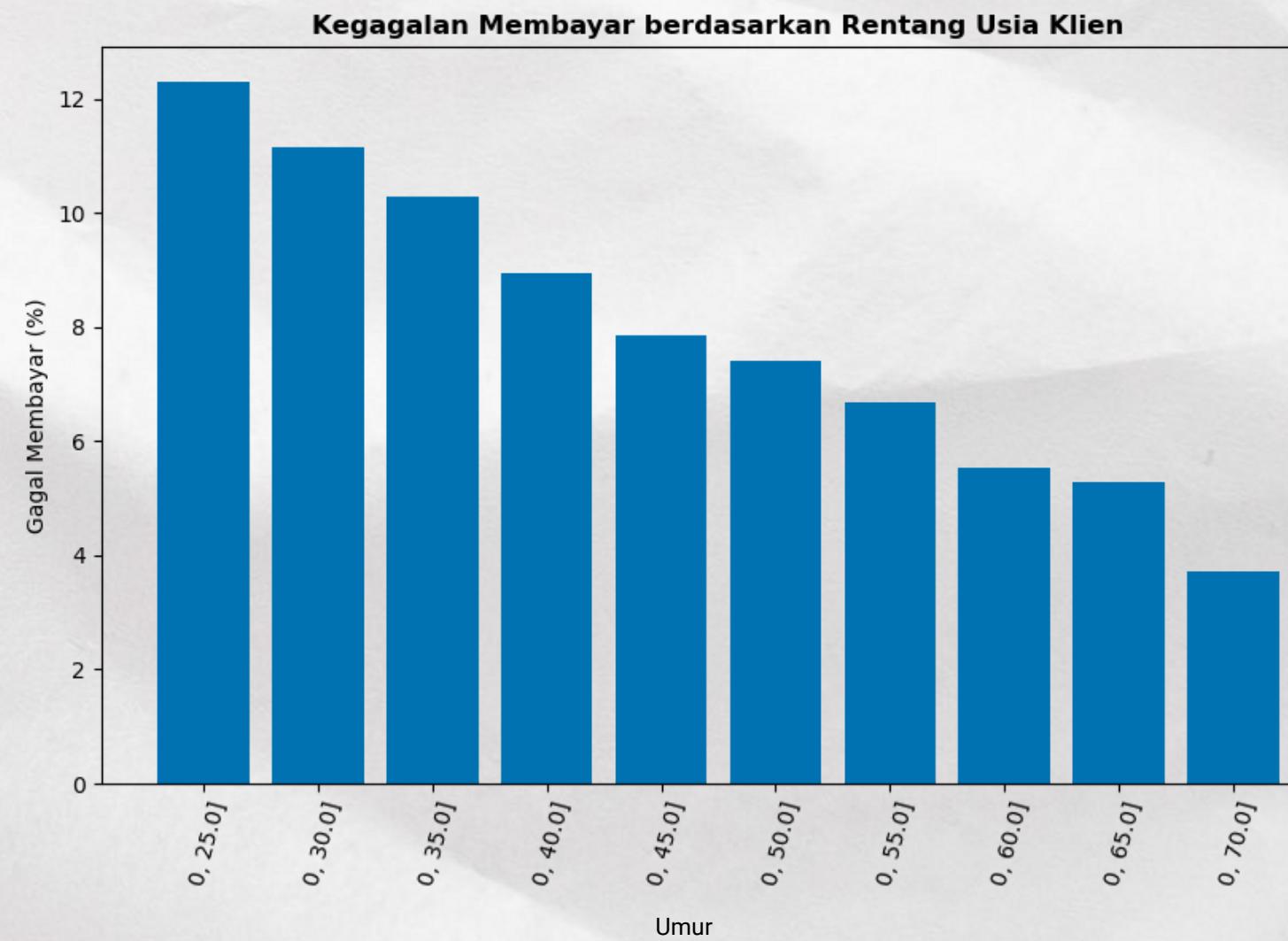
# EXPLORATORY DATA ANALYSIS



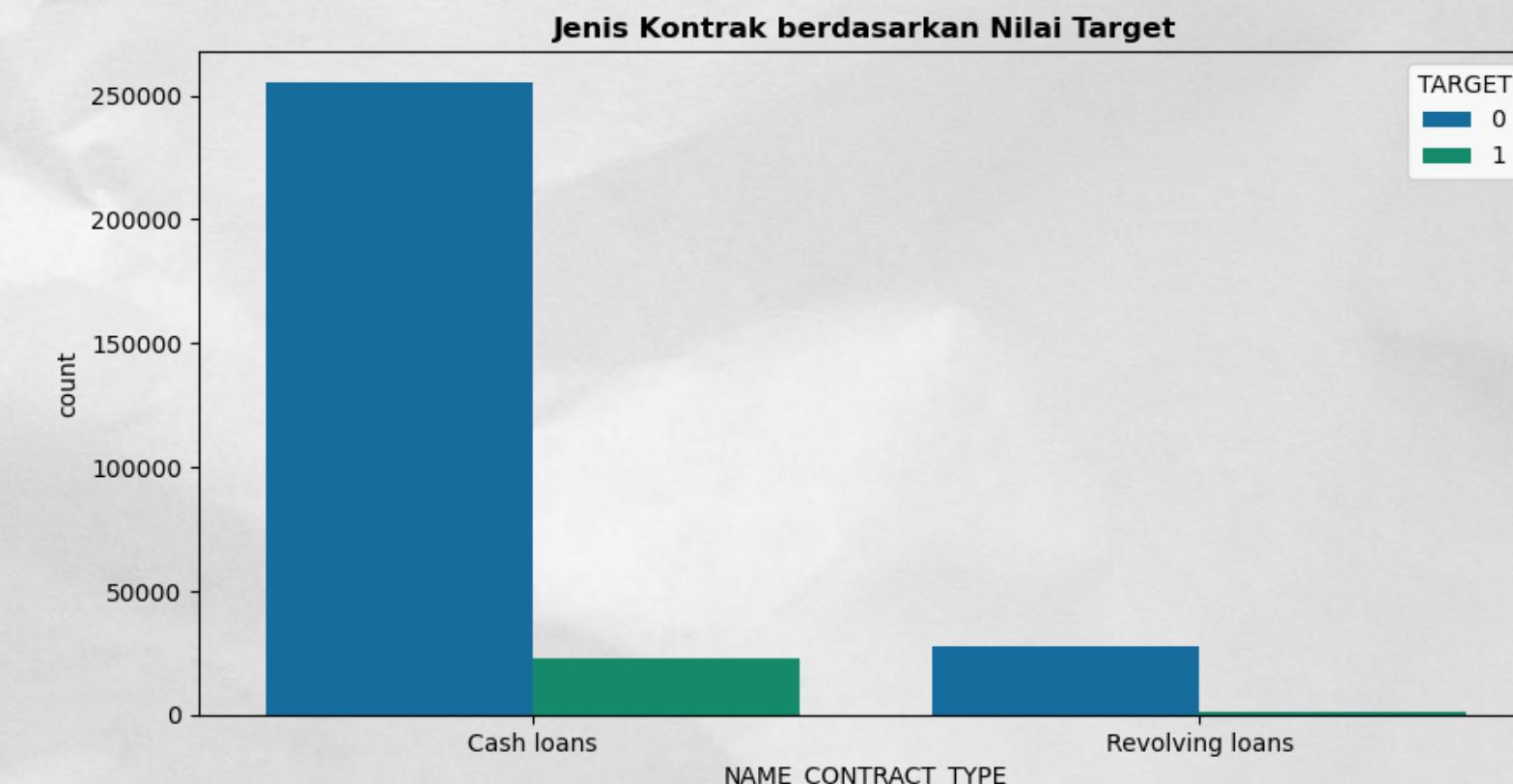
Terdapat masalah kelas yang tidak seimbang pada Features "TARGET"

- Dataset pelatihan terdiri dari 307.511 baris dengan 122 Features. Dataset pengujian terdiri dari 48.744 baris dengan 121 Features (memiliki Features yang sama seperti Dataset pelatihan tetapi dikurangi variabel 'TARGET')
- Dari 121 Features di Dataset latih, 40,99% fitur memiliki lebih dari 50% nilai yang hilang.
- Dari 121 Features terdapat 16 Features bertipe Categorical
- Pada Dataset menampilkan 20 variabel yang terkait dengan apakah klien telah menyediakan formulir tertentu 'FLAG\_DOCUMENT'.

# BUSINESS INSIGHTS

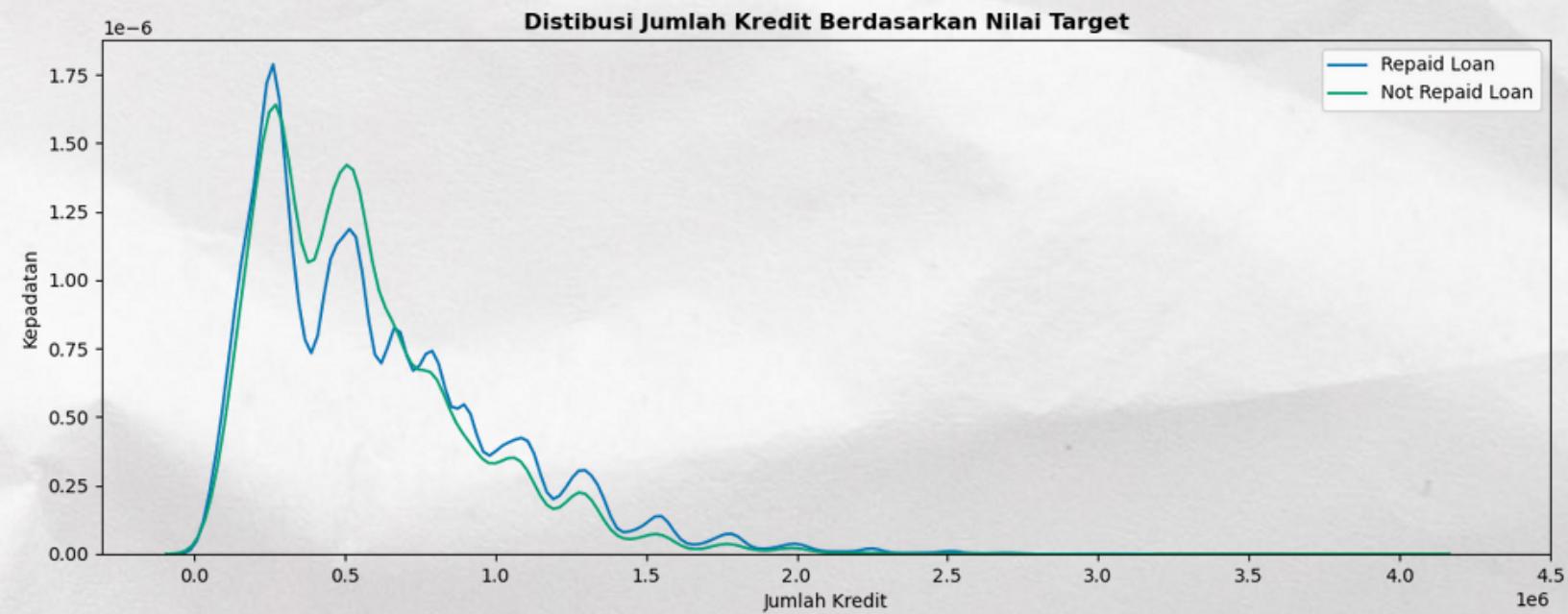


seiring bertambahnya usia klien,  
tingkat nasabah yang mengalami kesulitan  
membayar pinjaman yang terus menurun

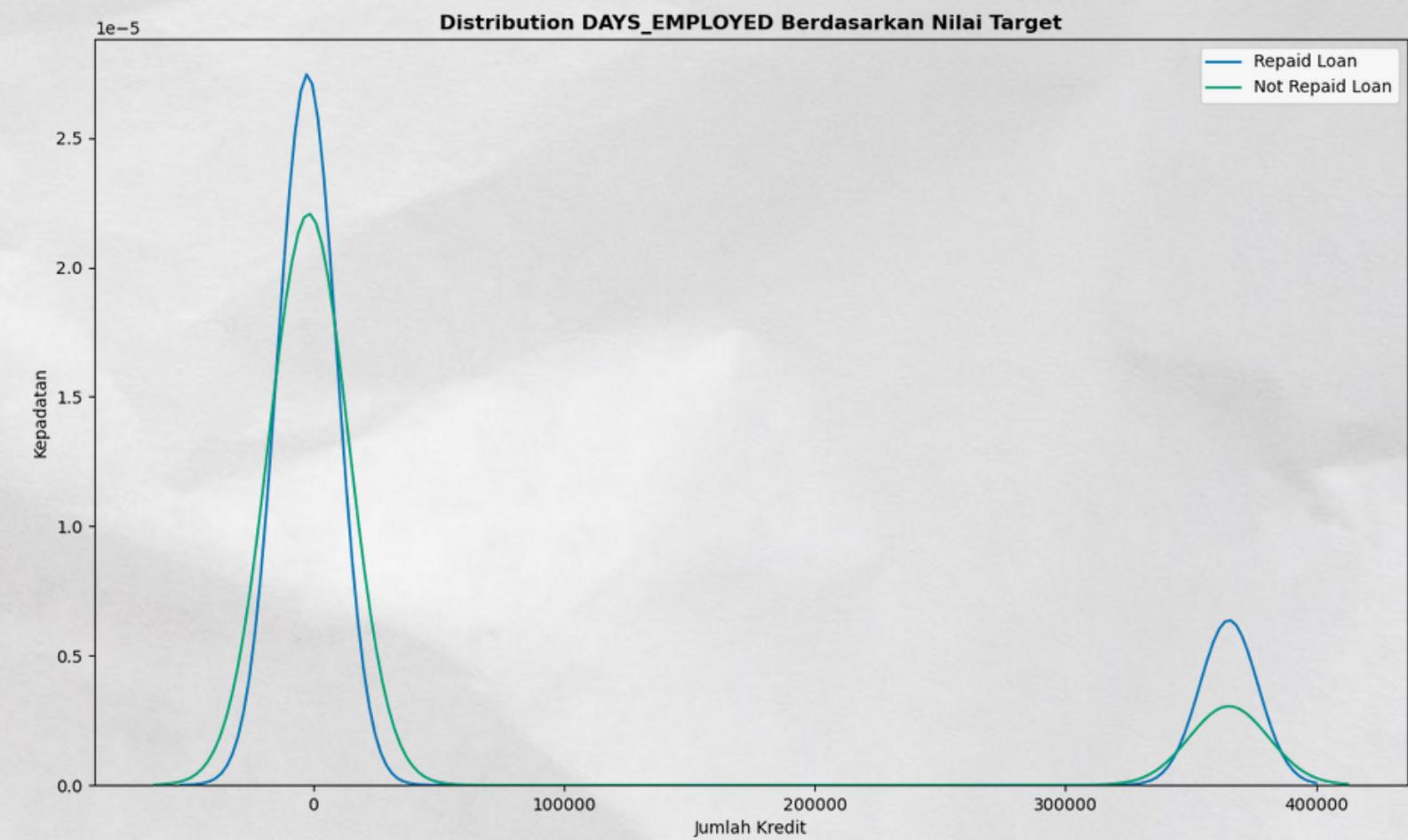


Rasio pinjaman tidak dilunasi (Target =1) untuk  
pinjaman Tunai adalah 8,35%, sedangkan rasio  
pinjaman tidak dilunasi (Target =1) untuk  
pinjaman Revolving adalah 5,48%.

# BUSINESS INSIGHTS



Distribusi jumlah kredit yang dibayar tepat waktu dan mengalami kesulitan membayar adalah serupa. Tetapi terdapat anomali pada jumlah kredit kisaran antara 400.000 hingga 600.000 memiliki rasio pinjaman yang tidak dilunasi lebih tinggi.



Distribusi pinjaman mengalami kesulitan membayar memuncak pada DAYS\_EMPLOYED sekitar nol. Nilai yang sama dengan dan di bawah nol cenderung mengindikasikan bahwa nasabah tersebut menganggur. Sehingga kemampuan seseorang untuk memperbaiki pinjaman mereka tampaknya berkorelasi positif dengan masa kerja di pekerjaan mereka saat ini

# PREPROCESSING

- Handling missing value
- Handling Outlier
- Label Encoder
- Train - test split
- Normalization (MinMaxScaler)
- Handling class imbalance  
(Oversampling)

# MODELING

Logistic Regression

Light GBM

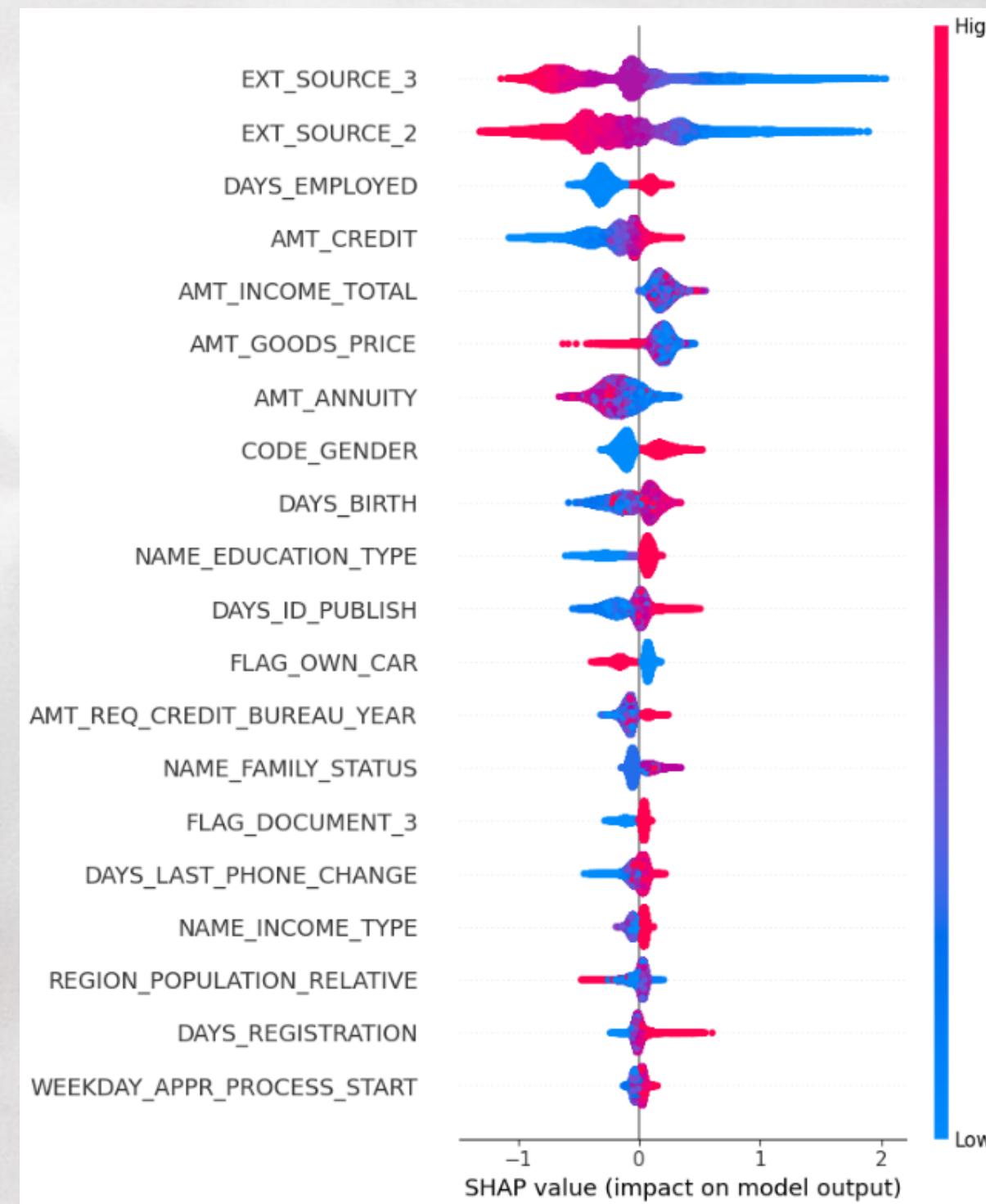
Evaluasi dilakukan dengan menggunakan  
Cross-Validation Score

# EVALUASI MODEL

Dari beberapa model yang dibuat, didapatkan model dengan Tuned Light GBM (dilakukan hyperparameter tuning) memiliki Cross-Validation Score yang lebih besar dibandingkan dengan model yang lainnya dengan nilai 0.742355

Model	Cross-Validation Score
Logistic Regression	0.735775
Tuned Logistic Regression	0.525955
Default Light GBM	0.740487
<i>Tuned Light GBM</i>	0.742355

# FEATURE VALUE



Urutan pengaruh fitur terhadap take-up:

1. EXT\_SOURCE\_3
2. EXT\_SOURCE\_2
3. DAYS\_EMPLOYED
4. AMT\_CREDIT
5. AMT\_INCOME\_TOTAL
6. AMT\_GOODS\_PRICE
7. AMT\_ANNUITY
8. CODE\_GENDER
9. DAYS\_BRITH
10. NAME\_EDUCATION\_TYPE
11. DAYS\_ID\_PUBLISH
12. FLAG\_OWN\_CAR
13. AMT\_REQ\_CREDIT\_BUREAU\_YEAR
14. NAM\_FAMILY\_STATUS
15. FLAG\_DOCUMENT\_3
16. DAYS\_LAST\_PHONE\_CHANGE
17. NAME\_INCOME\_TYPE
18. REGION\_POPULATION\_RELATIVE
19. DAYS\_REGISTRATION
20. WEEDAY\_APPR\_PROCESS\_START

Relasi antara masing-masing fitur dengan take-up sejalan dengan temuan pada business insights

# RESULT

potensi kenaikan  
take-up rate

+4,17%

96,10%

91,93%

Before

After

## insight-based recommendations:

- Mempertimbangkan untuk memberikan kriteria usia sebagai salah satu faktor dalam proses persetujuan pinjaman. Hal ini dapat membantu perusahaan untuk meminimalkan risiko pinjaman yang tidak terbayar.
- Mempertimbangkan untuk memberikan jenis pinjaman yang tepat berdasarkan rasio pinjaman tidak dilunasi. Misalnya, perusahaan dapat memberikan pinjaman Revolving Loans kepada nasabah dengan rasio pinjaman tidak dilunasi lebih rendah.
- Mempertimbangkan untuk memberikan analisis lebih lanjut terhadap anomali pada jumlah kredit kisaran antara 400.000 hingga 600.000 yang memiliki rasio pinjaman tidak dilunasi lebih tinggi. Hal ini dapat membantu perusahaan untuk menentukan strategi yang lebih baik dalam memberikan pinjaman kepada nasabah dengan jumlah kredit tertentu
- Mempertimbangkan untuk memberikan penilaian khusus terhadap nasabah yang memiliki DAYS\_EMPLOYED sekitar nol. Hal ini dapat membantu perusahaan untuk memberikan penanganan khusus terhadap nasabah yang memiliki risiko pinjaman yang lebih tinggi, seperti memberikan program pelatihan atau bantuan pekerjaan kepada nasabah yang menganggur, sehingga dapat memperbaiki kemampuan pembayaran pinjaman mereka.

# REPOSITORY GIT HUB

Link:

<https://github.com/HelmyRamadhani/Home-Credit-Default-Risk>