

Virtual Internship Experience

Loan Credit Risk.



Created by:

Umbara Sakti Mihadja

- umbaramihardja@gmail.com
- [linkedin.com/in/umbarasm](https://www.linkedin.com/in/umbarasm)



OUTLINE

☐ Problem Statement

☐ Exploratory Data Analysis

☐ Machine Learning Model

☐ Model Simulation

☐ Recommendation





Problem Statement.

Current Problem.

Bank Partners akan perlu melakukan penentuan credit risk pada nasabah yang akan melakukan pinjaman, tetapi dalam pelaksanaannya terdapat beberapa hambatan besar, diantaranya:

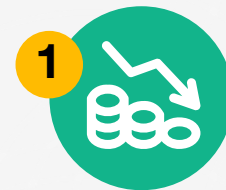
1. **Cost yang tinggi** dalam menentukan credit risk karena memerlukan SDM dan waktu yang banyak
2. **Human error** dapat terjadi, sehingga penentuan credit risk yang salah akan menimbulkan kerugian yang cukup besar.

Goals.



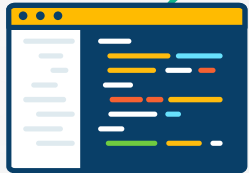
Menurunkan cost yang diperlukan untuk penentuan credit risk seorang nasabah.

Business Metrics.



Operational Cost

Objective.



Membangun model *Machine Learning* yang dapat memprediksi kategori risiko nasabah yang mengambil loan



Menemukan faktor penting dan karakteristik utama dari nasabah yang mengambil loan



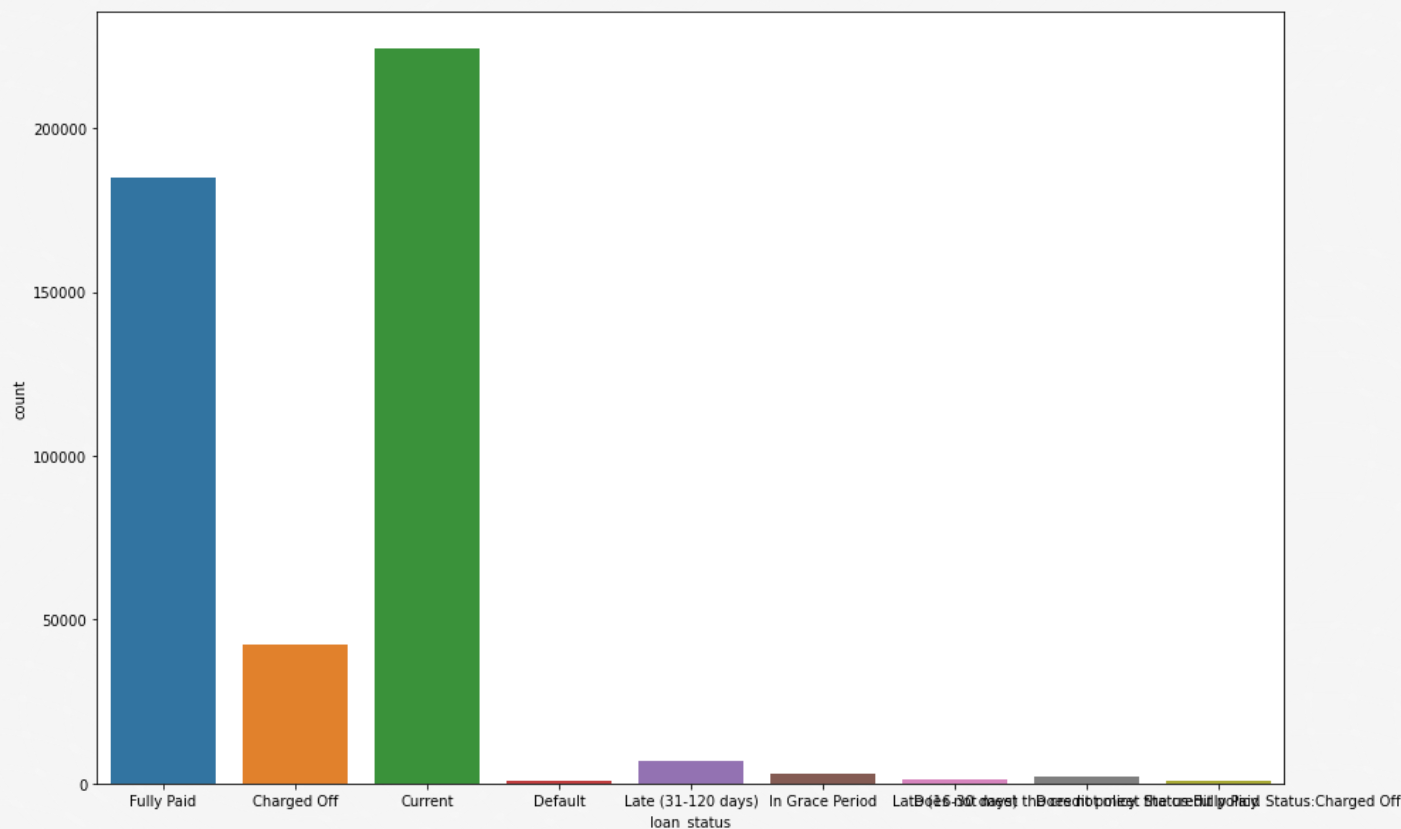
Memberikan rekomendasi yang akan meningkatkan efektivitas perusahaan dalam menentukan pengambilan keputusan terhadap nasabah yang mengambil loan



Exploratory Data Analysis & Insight.

Data Overview.

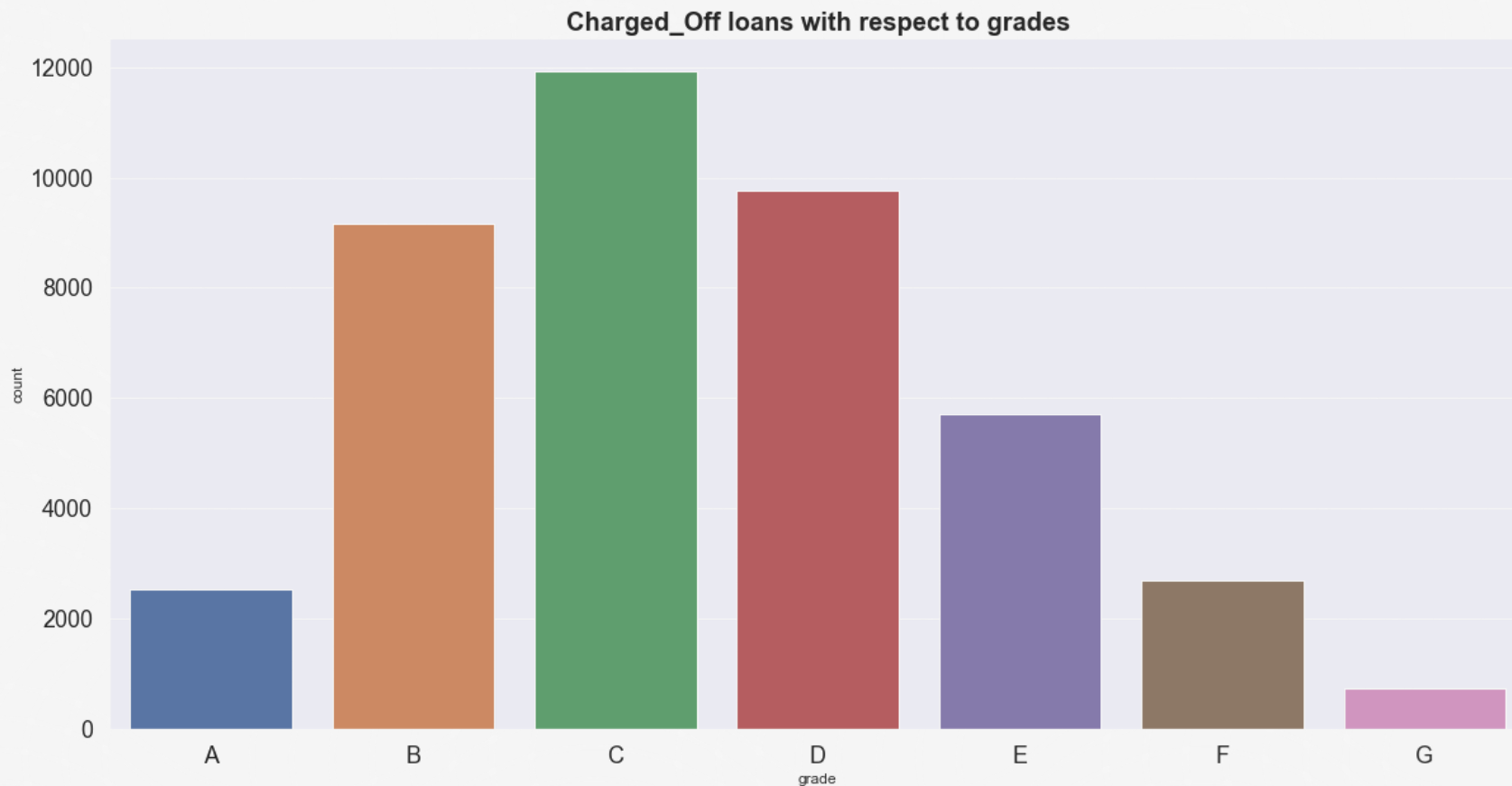
1. Dataset terdiri dari **466.285 baris** dan **76 kolom**
2. Terdapat **54 data numerik** dan **22 data kategorikal**
3. Terdapat **41 kolom** yang memiliki data *null* atau kosong



Target Feature

1. Fully Paid
2. Charged Off
3. Default
4. Current
5. In Grace Period
6. Late (16-30 days)
7. Late (31-120 days)
8. Does not meet the credit policy. Status: Fully Paid
9. Does not meet the credit policy. Status: Charged Off

Data Insight.

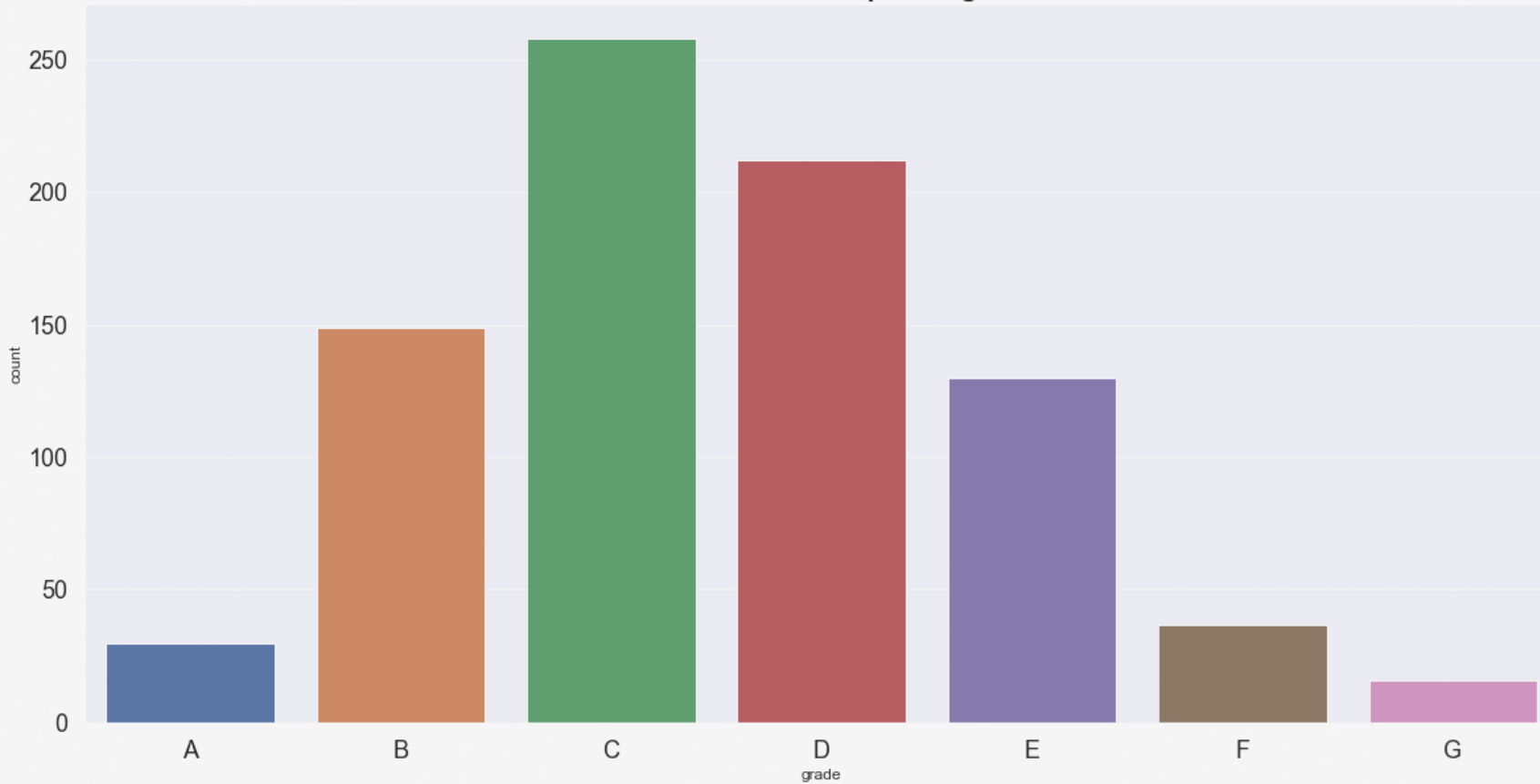


Pada status loan Charge Off, grade Loan Credit (LC) paling tinggi terdapat pada grade C (Upper medium Grade), kemudian diikuti oleh grade D dan B.

Dapat diindikasikan bahwa nasabah dengan LC dengan grade C berpotensi tinggi untuk berstatus Charge Off.

Data Insight.

Default loans with respect to grades



Pada status loan Default, grade Loan Credit paling tinggi terdapat pada grade C (Upper medium Grade), kemudian diikuti oleh grade D dan B.

Dapat diindikasikan bahwa nasabah dengan LC dengan grade C berpotensi tinggi untuk berstatus Default.

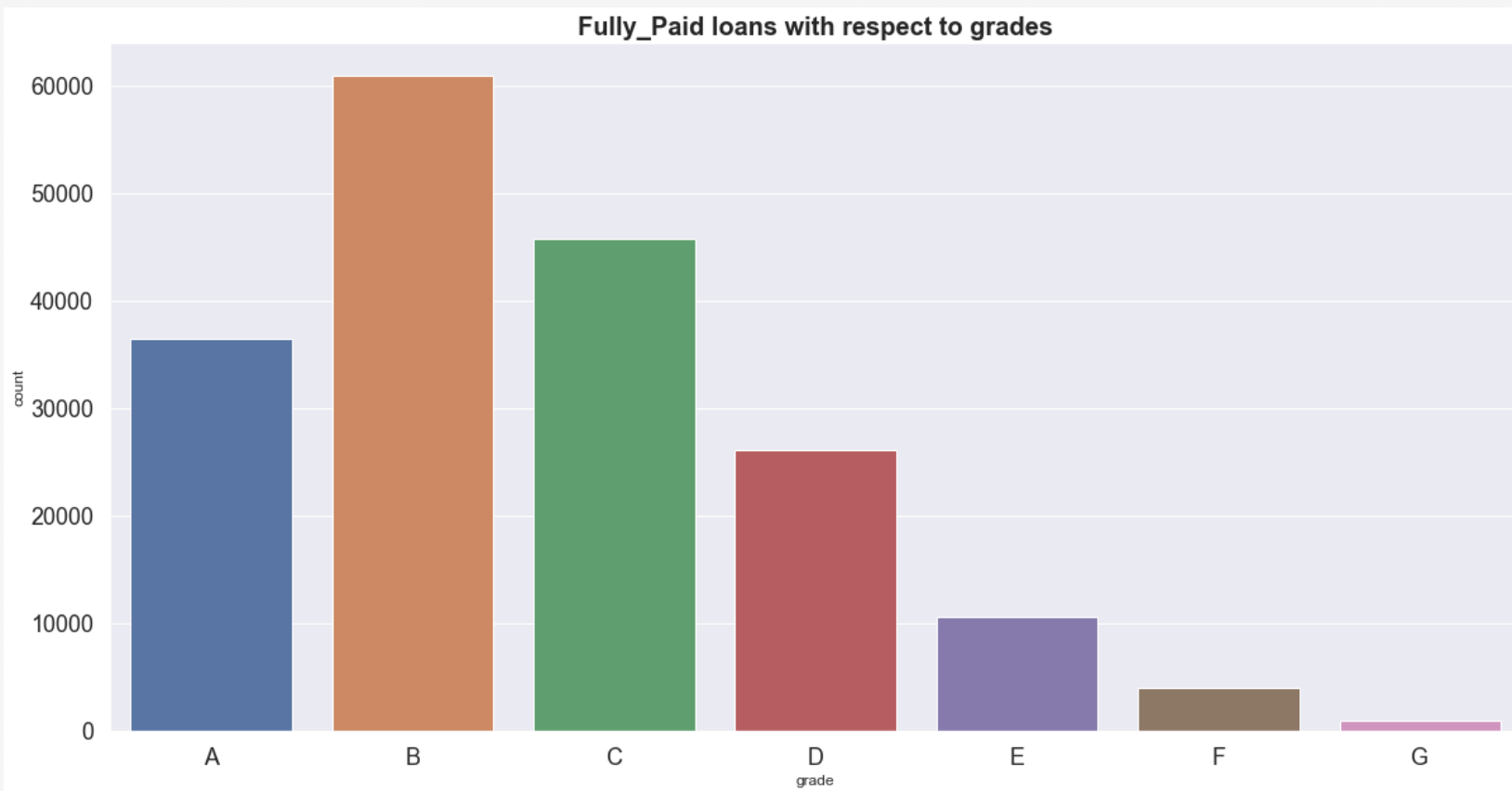
Data Insight.



Pada status loan Does not meet the credit policy. Status:Charged Off, grade Loan Credit paling tinggi terdapat pada grade D (lower medium grade), kemudian diikuti oleh grade E dan C.

Dapat diindikasikan bahwa nasabah dengan LC dengan grade D berpotensi tinggi untuk berstatus Does not meet the credit policy. Status:Charged Off karena sudah memiliki status kurang baik.

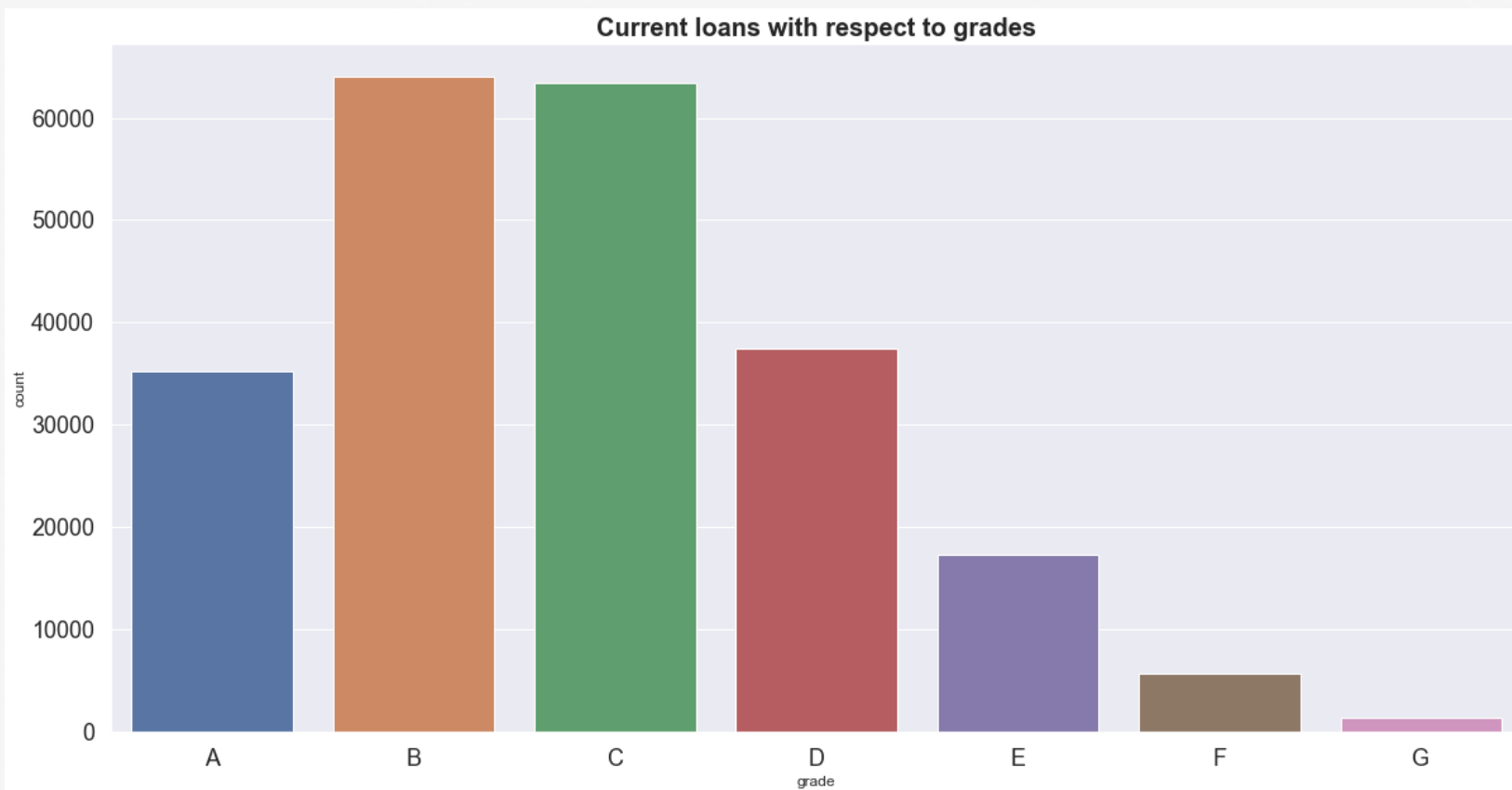
Data Insight.



Pada status loan Default, grade Loan Credit paling tinggi terdapat pada grade B (Upper medium Grade), kemudian diikuti oleh grade C dan A

Dapat diindikasikan bahwa nasabah dengan LC dengan grade B berpotensi tinggi untuk berstatus Fully paid dan perlu menjadi prioritas utama bank.

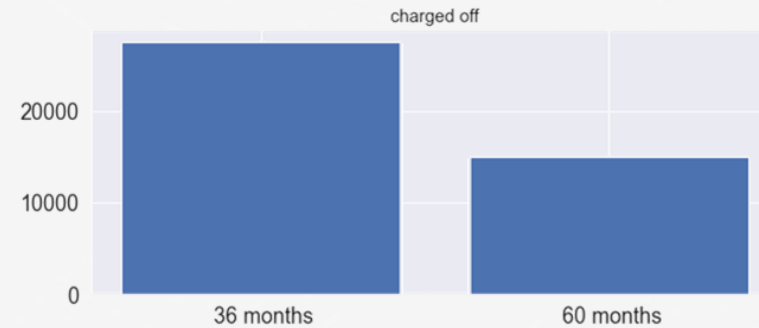
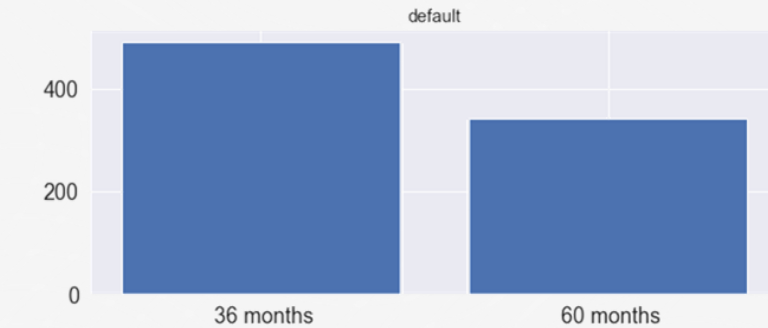
Data Insight.



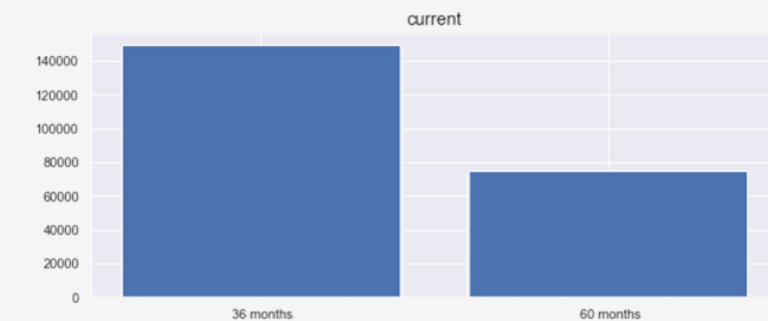
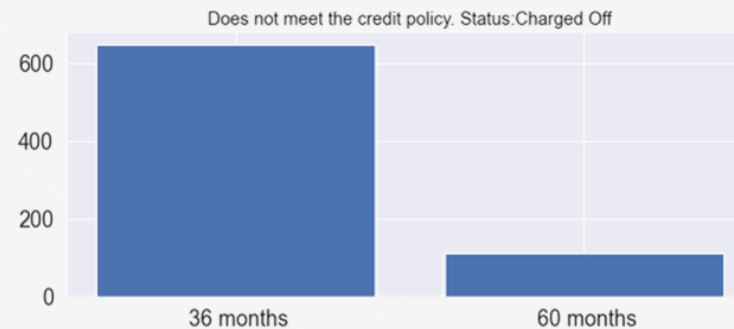
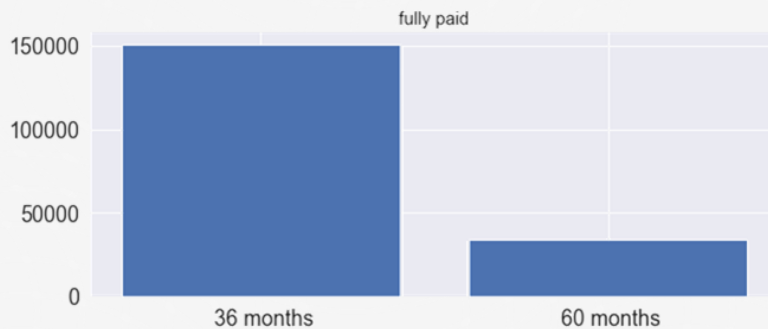
Pada status loan Default, grade Loan Credit paling tinggi terdapat pada grade B-C (high grade - upper medium grade).

Dapat diindikasikan bahwa nasabah yang masih memiliki pinjaman berada pada grade B-C

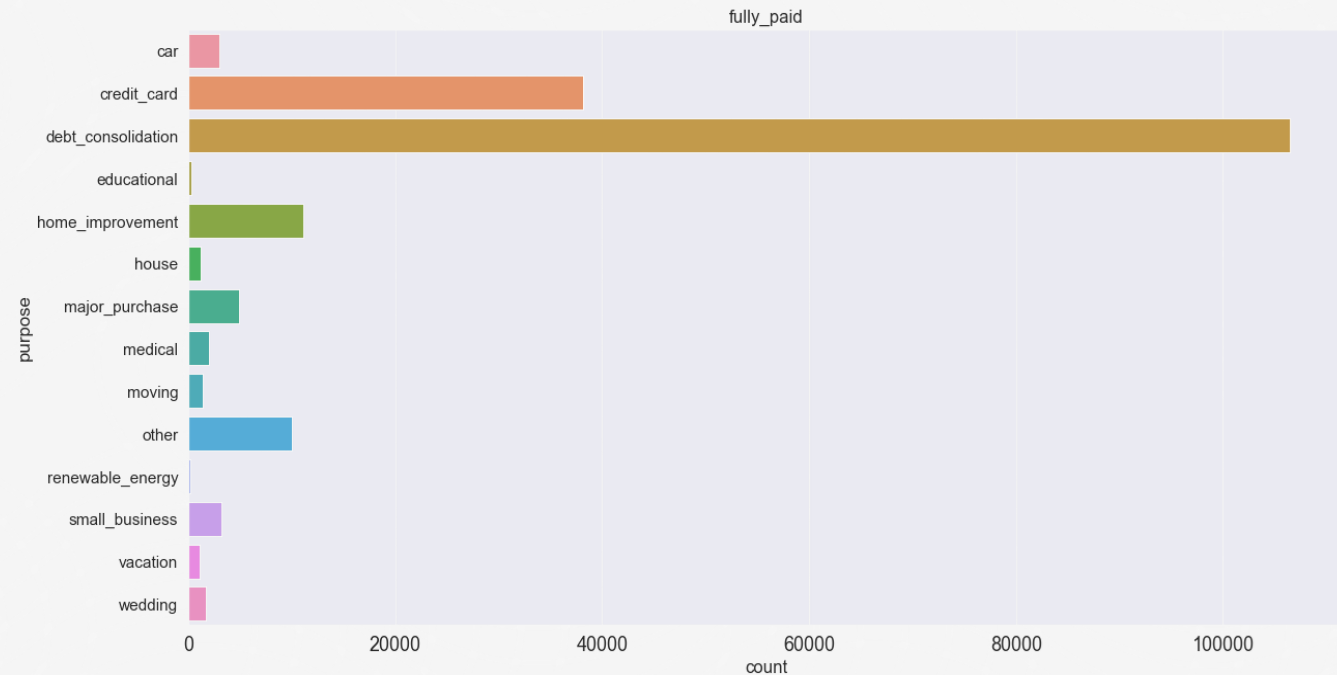
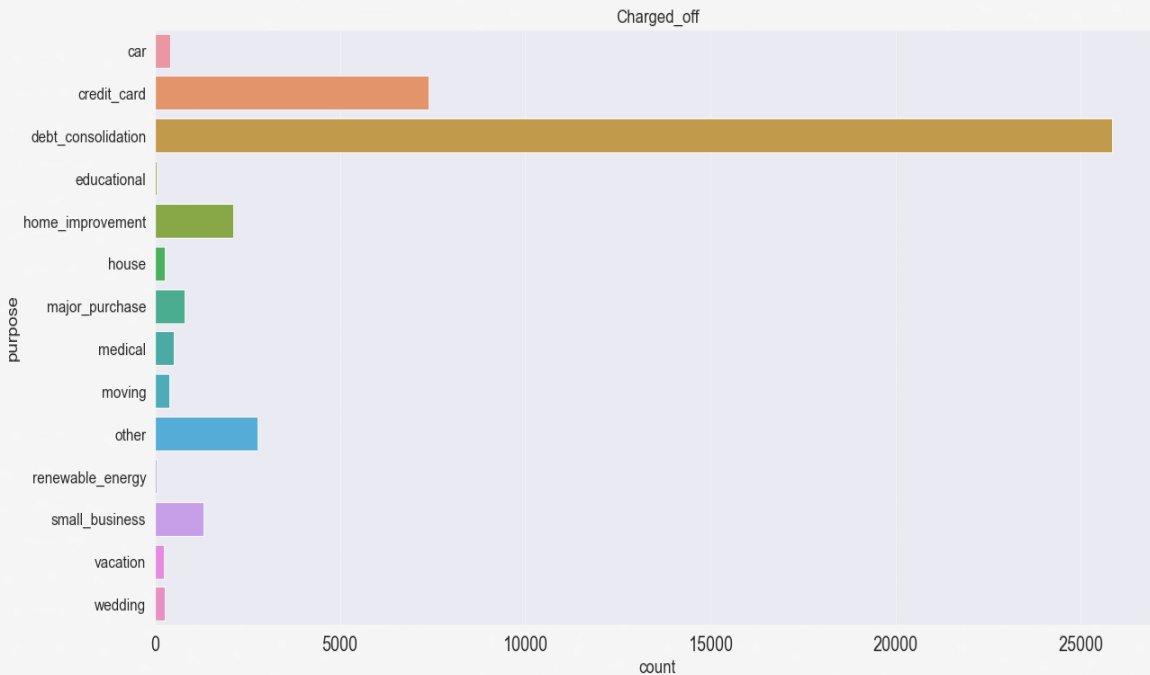
Data Insight.



Pada semua status loan, term loan yang dimiliki oleh nasabah berjangka waktu 36 bulan, sedangkan term loan 60 bulan tertinggi pada status loan default.

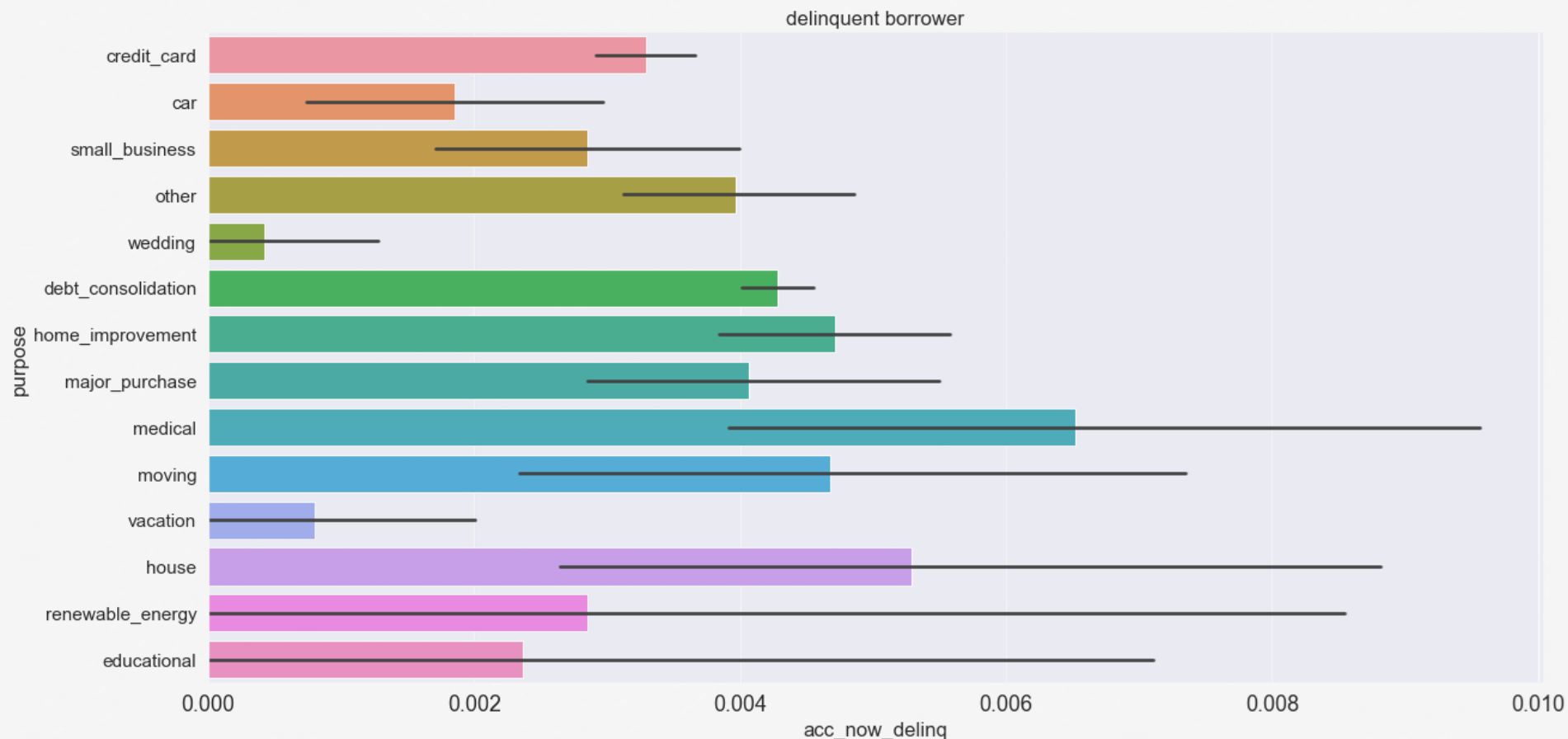


Data Insight.



Pada status loan charge off dan fully paid, tujuan credit loan terbanyak adalah sama, yaitu untuk debt consolidation atau penggabungan beberapa pinjaman tanpa jaminan menjadi satu pinjaman dengan nilai yang lebih besar.

Data Insight.



Nasabah yang menunggak paling banyak terjadi pada pinjaman medis dan rumah. Hal ini tentu menjadi sebuah pertimbangan bank untuk meningkatkan pengawasan kinerja pembayaran loan pada kedua tujuan loan tersebut.

Data Pre-Processing.

Handling Outlier

- Z-Score

Feature Encoding

- Label Encoding
- One Hot Encoding

Split Data Train & Test (80:20)

- Data Train: 36.168
- Data Test: 9.043

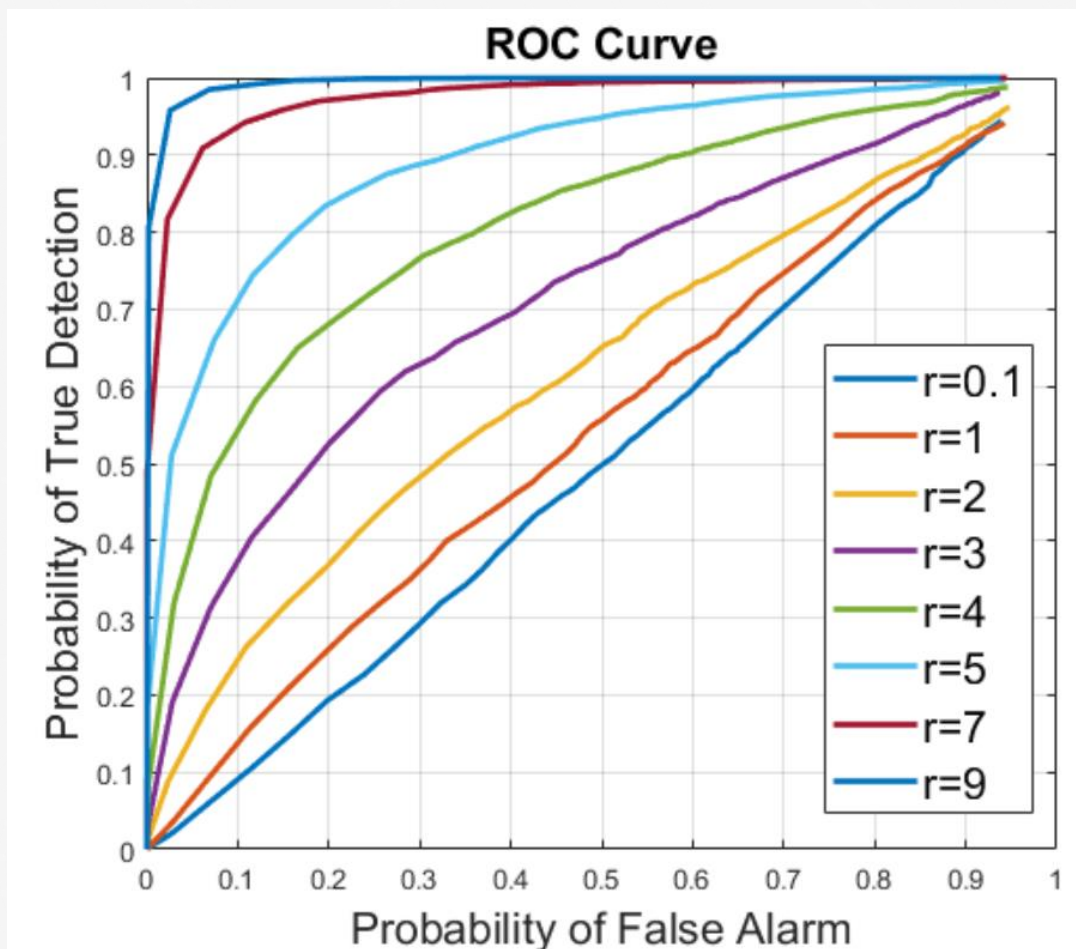
Scalling

- Standarization

Machine Learning Model

Machine Learning Model.

Model Evaluation.



$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\begin{aligned} \text{ROC - AUC} &= \int_0^1 \text{TPR}(\text{FPR}) d\text{FPR} \\ &= \int_0^1 \text{TPR}(\text{FPR}^{-1}(x)) dx \end{aligned}$$

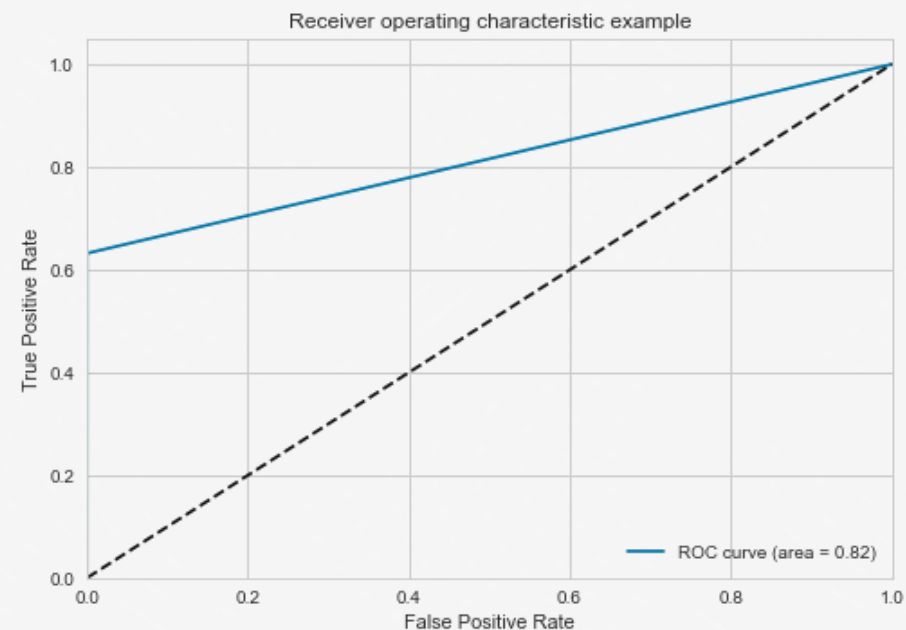
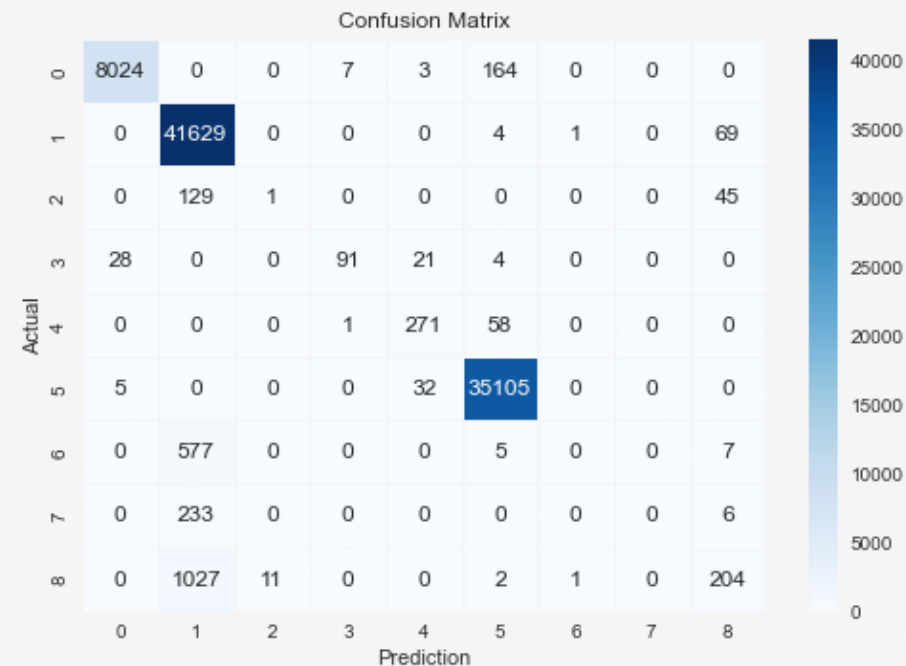
ROC-AUC

AUC (Area Under Cover) merupakan area di bawah kurva (Area under the Curve of) ROC (Receiver Operating Characteristic), suatu kurva yang menggambarkan probabilitas dengan variabel sensitivitas dan kekhususan (*specificity*) dengan nilai batas antara 0 hingga 1. Area di bawah kurva memberikan gambaran tentang keseluruhan pengukuran atas kesesuaian dari model yang digunakan.

Modeling Result.

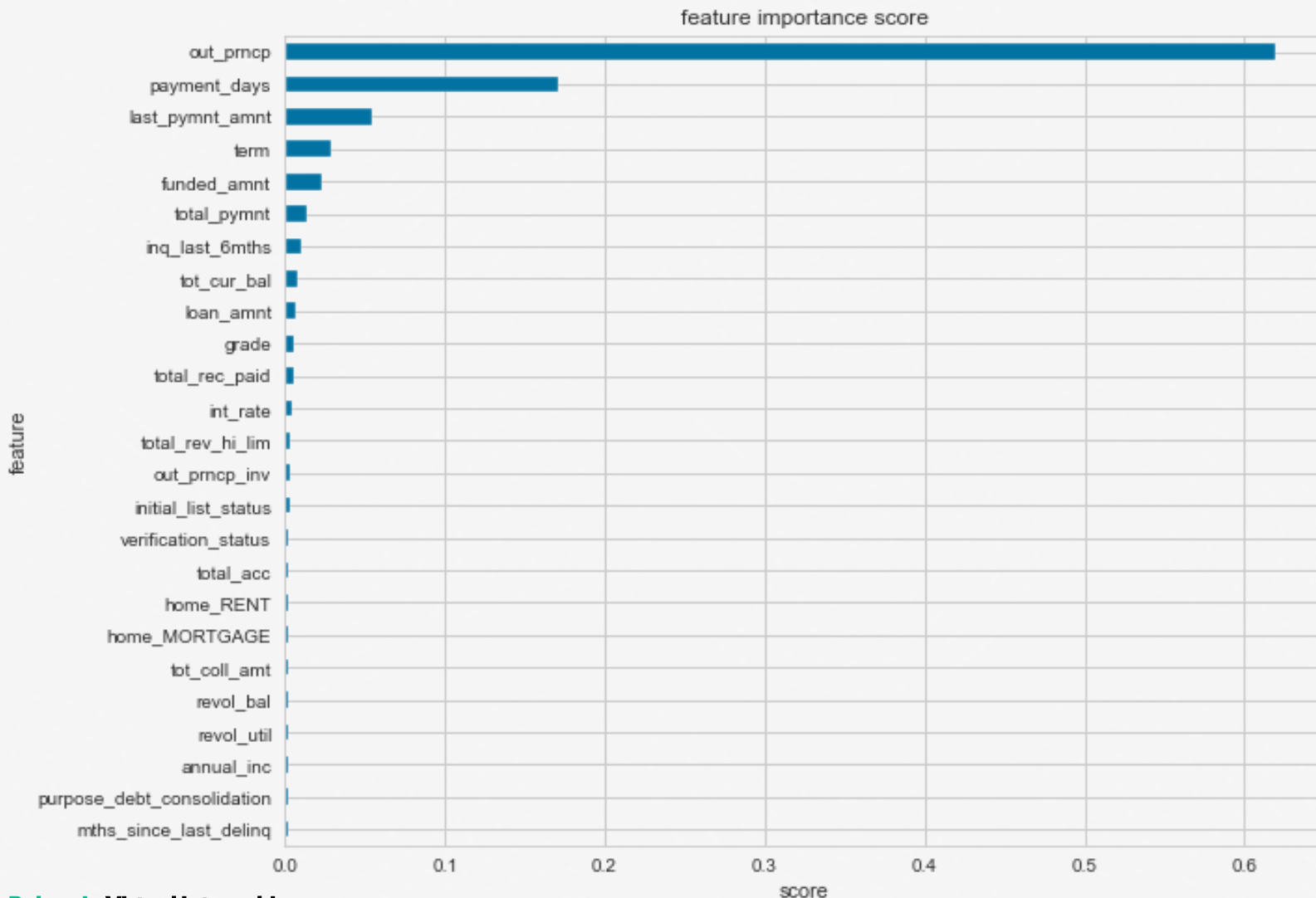
Model	Recall	Precision	F1 Score	AUC
Decision Tree Classifier	0,49	0,49	0,49	0,80
Random Forest Classifier	0,46	0,59	0,49	0,79
XGBoost Classifier	0,51	0,60	0,53	0,82

Algoritma terbaik yang dipilih adalah
XGBoost Classifier



Feature Importance.

Visualisasi *feature importance* menggunakan **dalex**



Dari grafik feature importance di samping, dapat diambil 5 fitur penting yang dapat menjadi dasar **dalam rekomendasi kebijakan**, diantaranya:

1. Out principal: Sisa pinjaman pokok yang didanai
2. Payment days: jumlah hari untuk melakukan pembayaran loan terakhir
3. Last payment amount: Jumlah total pembayaran terakhir diterima
4. Term: Jangka waktu pembayaran loan
5. Funded amount: jumlah total komitmen untuk suatu pinjaman

Business Simulation

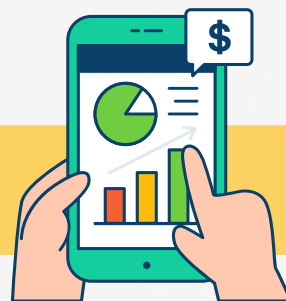
Business Problem.



Bank
Partners



Credit loan

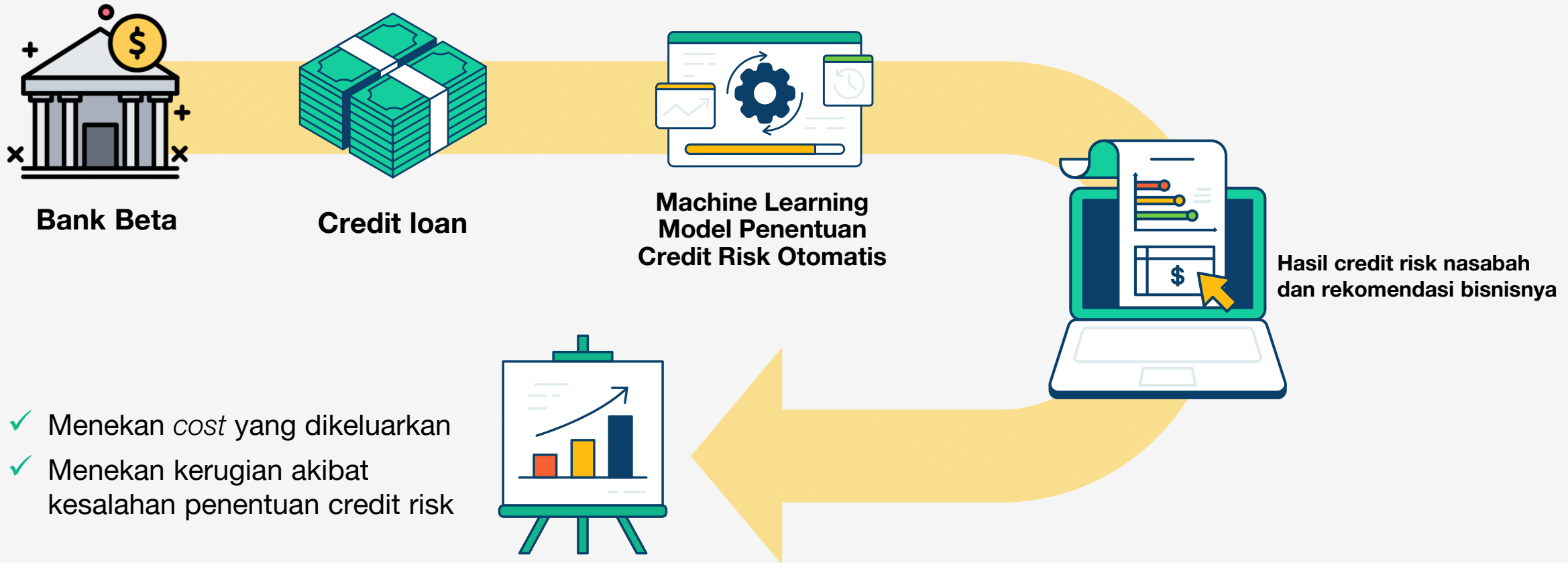


Penelitian Credit
Risk secara manual



- *Cost* tinggi untuk membayar SDM
- Kerugian dalam credit loan yang membesar akibat kesalahan penentuan credit risk

Business Simulation.



Business Recommendation

Business Recommendation.



Out principal

- Meningkatkan pengawasan terhadap sisa pinjaman pokok yang tersedia dengan melakukan pengingat kepada nasabah



Impact Estimation

Kewaspadaan nasabah terhadap sisa pinjaman pokok menjadi tinggi, sehingga tidak terjadi perubahan status loan yang berdampak negatif untuk kedua belah pihak.



Payment Days

- Rata-rata nasabah melakukan pembayaran hingga 2-5 tahun, sehingga bank dapat memberikan keringanan bunga untuk dapat meningkatkan performa nasabah dalam membayar.



Nasabah akan melakukan pembayaran secara teratur ditambah dengan bunga yang lebih ringan meningkatkan performa pembayaran nasabah



Last payment amount

- Pembayaran terakhir yang diterima dapat menjadi indikator bagi bank untuk menentukan performa nasabah



Bank dapat lebih mudah untuk melihat performa nasabah dan memberikan penawaran yang lebih

Business Recommendation.



Term

- Memprioritaskan pembayaran pada jangka waktu 36 bulan



Impact Estimation

Mengurangi nasabah dengan credit risk default atau charge off



Funded amount

- Memberikan penawaran yang meringankan bagi nasabah dengan komitmen yang tinggi



Nasabah akan lebih baik dalam performa dan komitmen yang diberikan



Terima Kasih
