

A vertical bar on the left side of the slide with a gradient from orange at the top to blue at the bottom.

LENDING APPROVAL PREDICTION

FINAL PROJECT - RADEK KILAT BUANA

TABLE OF CONTENTS

1. Business Background and Objectives
2. Data Understanding
3. Data Preparation & Manipulation
4. Exploratory Data Analysis (EDA)
5. Feature Engineering
6. Modelling
7. Conclusion & Recommendation

Business Background & Objectives

Business Background:

Dengan kemajuan teknologi yang pesat dan permintaan akan layanan keuangan khususnya pembiayaan dimana perusahaan harus meningkatkan proses pemberian pinjaman melalui analisa prediktif yang cepat & tepat. Dengan memanfaatkan data dan algoritma Machine Learning diharapkan dapat menyederhanakan proses persetujuan kredit serta meminimalkan resiko kredit yang akan ditimbulkan kedepannya.

Objectives:

Membuat sebuah Machine Learning Model yang dapat memprediksi dengan cepat & tepat sebuah aplikasi kredit apakah akan disetujui atau ditolak.

Data Understanding

Dataset di peroleh dari Kaggle dengan link berikut:

<https://www.kaggle.com/datasets/architsharma01/loan-approval-prediction-dataset>

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4269 entries, 0 to 4268
Data columns (total 13 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   loan_id                              4269 non-null   int64
1   no_of_dependents                     4269 non-null   int64
2   education                            4269 non-null   object
3   self_employed                       4269 non-null   object
4   income_annum                        4269 non-null   int64
5   loan_amount                         4269 non-null   int64
6   loan_term                           4269 non-null   int64
7   cibil_score                         4269 non-null   int64
8   residential_assets_value            4269 non-null   int64
9   commercial_assets_value             4269 non-null   int64
10  luxury_assets_value                 4269 non-null   int64
11  bank_asset_value                    4269 non-null   int64
12  loan_status                         4269 non-null   object
dtypes: int64(10), object(3)
memory usage: 433.7+ KB
```

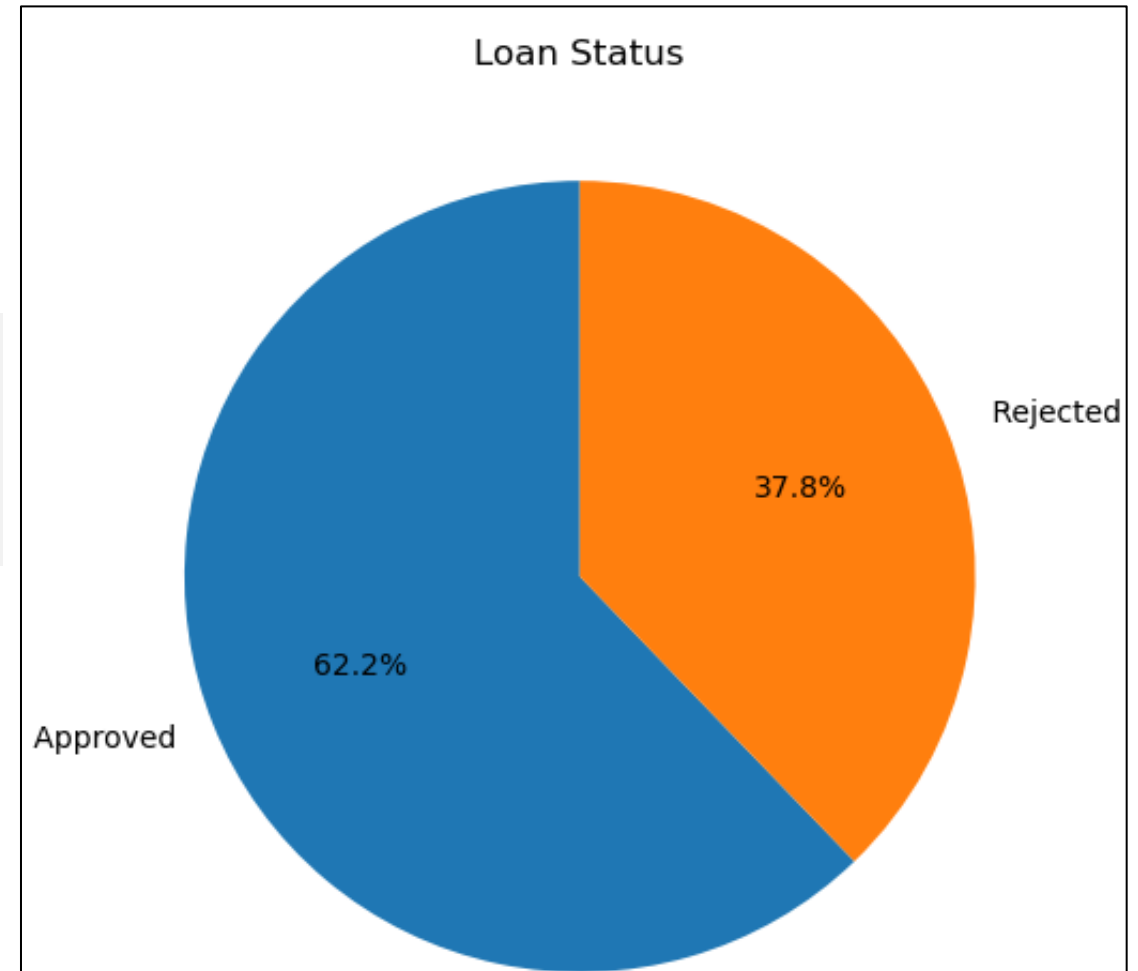
Data Preparation & Manipulation

- Melakukan pengecekan Missing Value
- Melakukan pengecekan Duplicate data
- Melakukan pengecekan terhadap Outlier

Exploratory Data Analysis (EDA)

Berapakah jumlah aplikasi yang disetujui atau ditolak dari dataset?

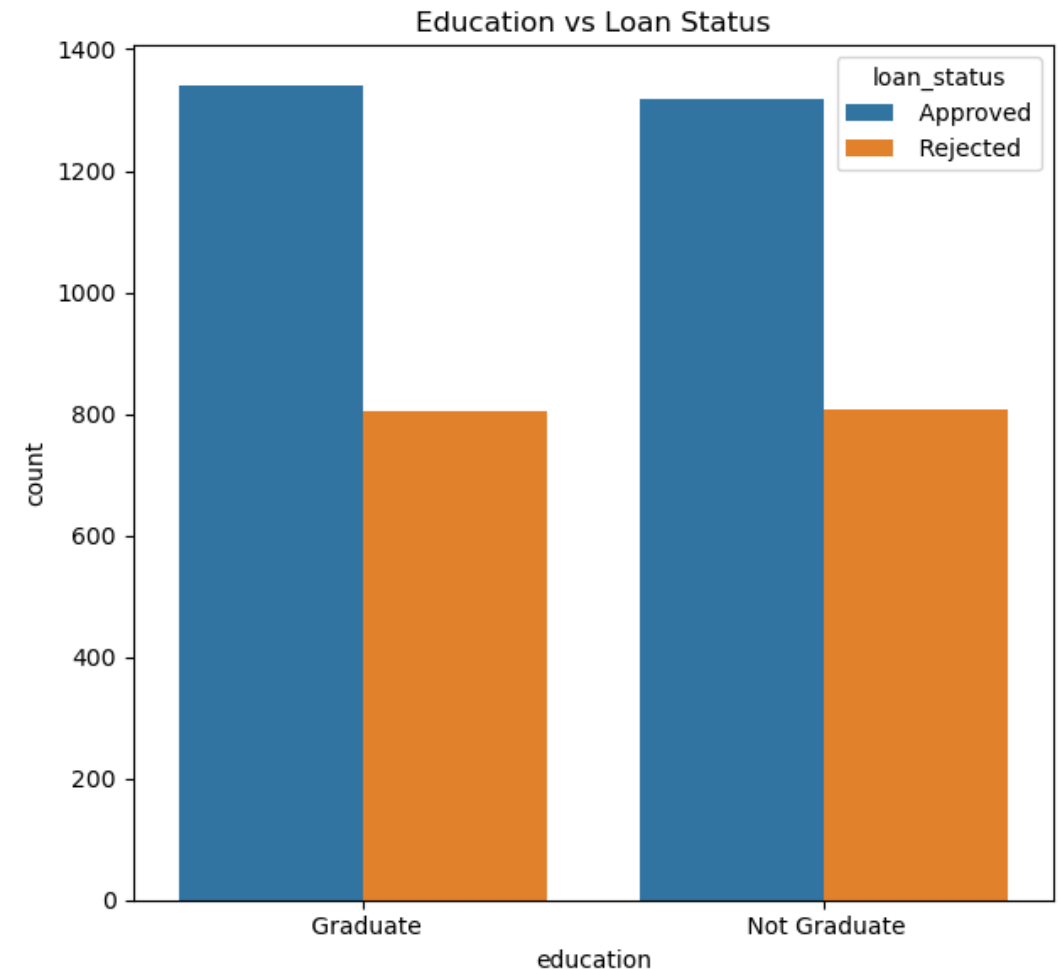
Dari jumlah total aplikasi kredit sebanyak 4.269, ada sebanyak 2.656 aplikasi yang di setujui (62,2%) ada sebanyak 1.613 aplikasi yang di tolak (37.8%)



Exploratory Data Analysis (EDA)

Apakah tingkat pendidikan mempengaruhi persetujuan kredit?

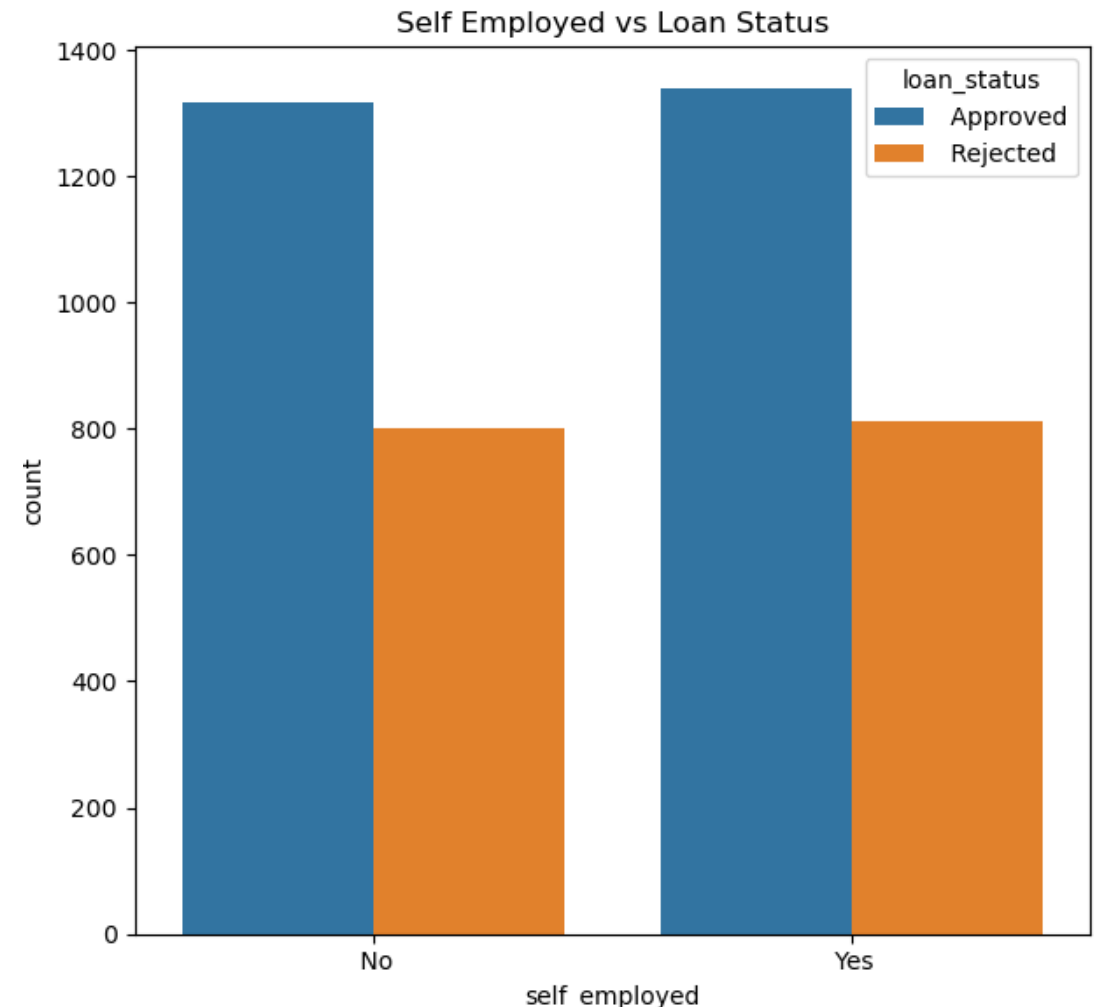
Berdasarkan chart disamping, dapat disimpulkan bahwa tidak ada perbedaan yang signifikan antara tingkat pendidikan dalam persetujuan kredit.



Exploratory Data Analysis (EDA)

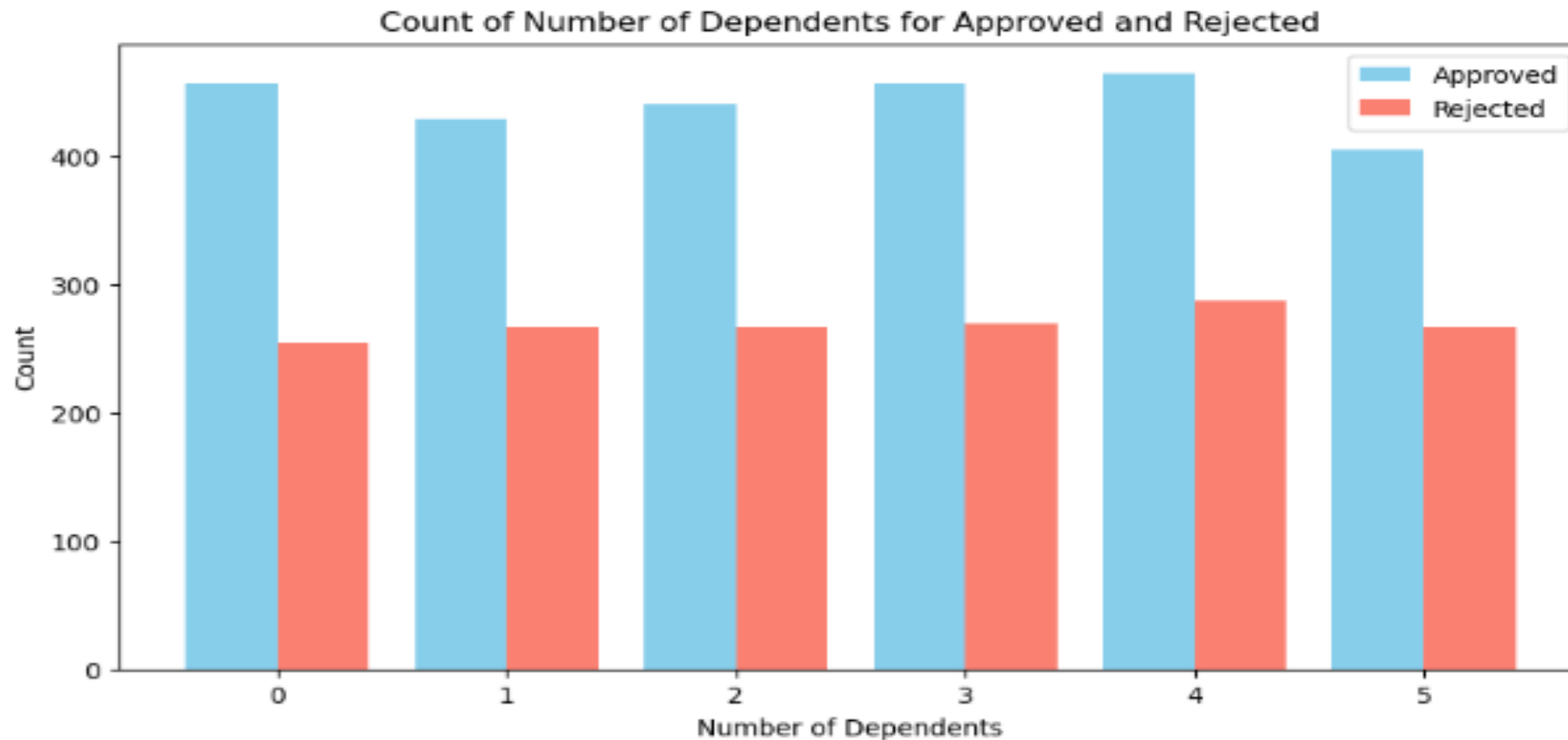
Apakah bekerja sendiri atau dengan pasangan mempengaruhi persetujuan kredit?

Berdasarkan chart disamping, dapat disimpulkan bahwa tidak ada perbedaan yang signifikan antara tingkat bekerja sendiri ataupun pasangan ikut bekerja terhadap persetujuan kredit.



Exploratory Data Analysis (EDA)

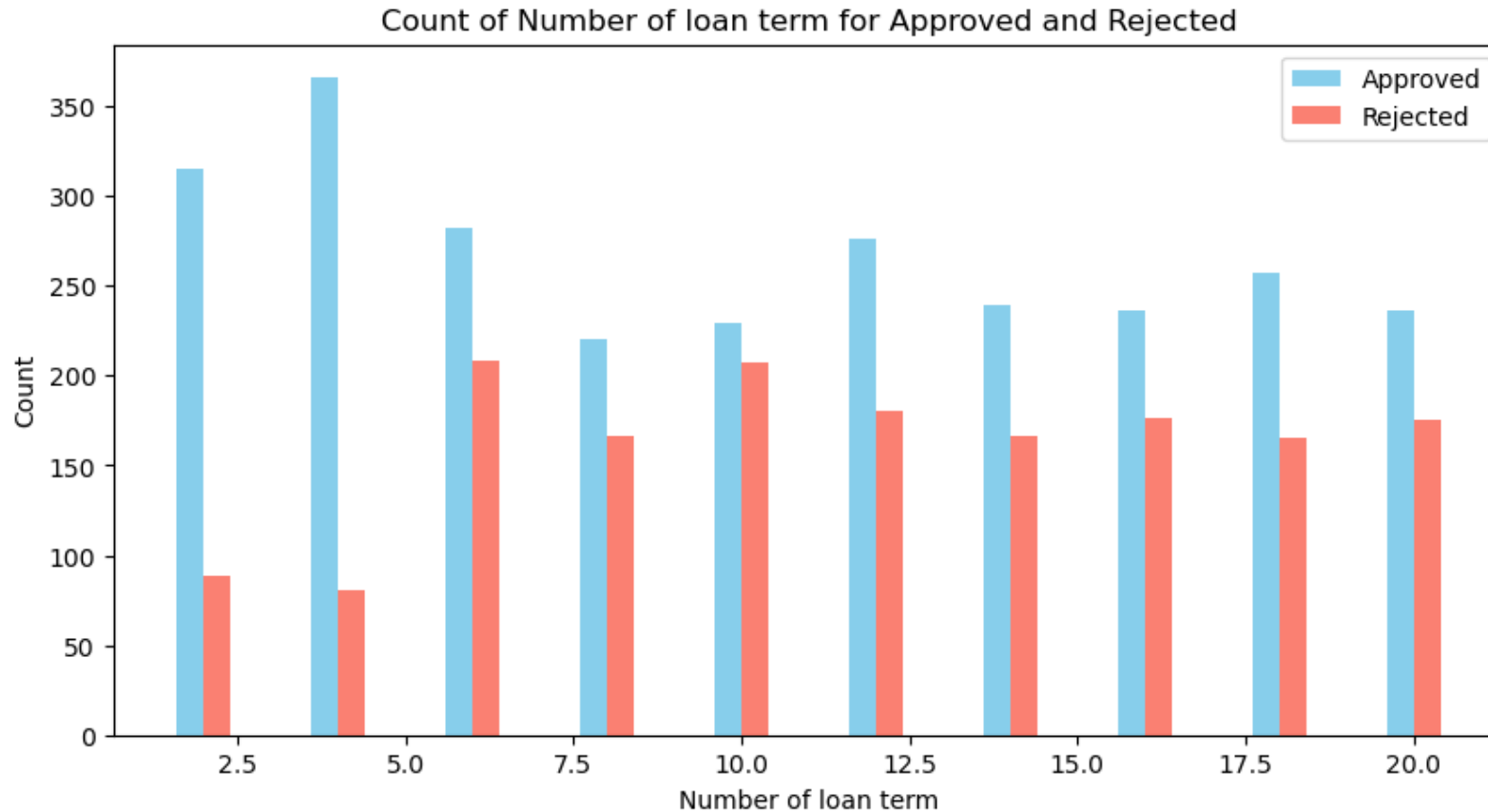
Apakah jumlah tanggungan mempengaruhi persetujuan kredit?



Berdasarkan chart diatas, tidak terlihat adanya perbedaan yang signifikan sehingga jumlah tanggungan tidak dapat mempengaruhi persetujuan kredit

Exploratory Data Analysis (EDA)

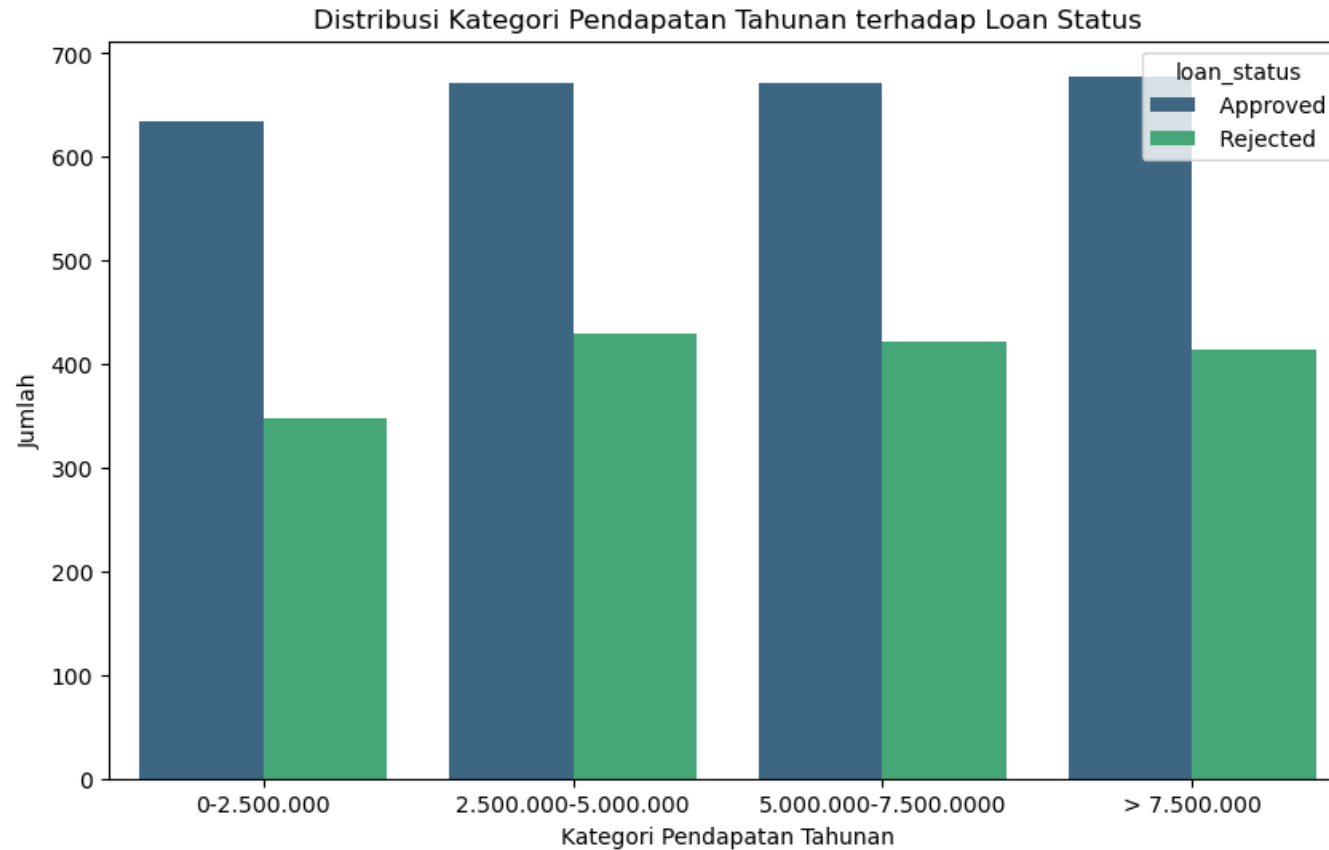
Apakah jangka waktu pinjaman mempengaruhi persetujuan kredit?



Berdasarkan chart diatas, terlihat bahwa untuk pengajuan dengan jangka waktu 2 & 4 bulan memiliki Tingkat persetujuan kredit yang lebih tinggi dibandingkan yang lainnya. Namun secara umum belum dapat disimpulkan bahwa semakin rendah jangka waktu yang diajukan maka persetujuan kredit akan disetujui.

Exploratory Data Analysis (EDA)

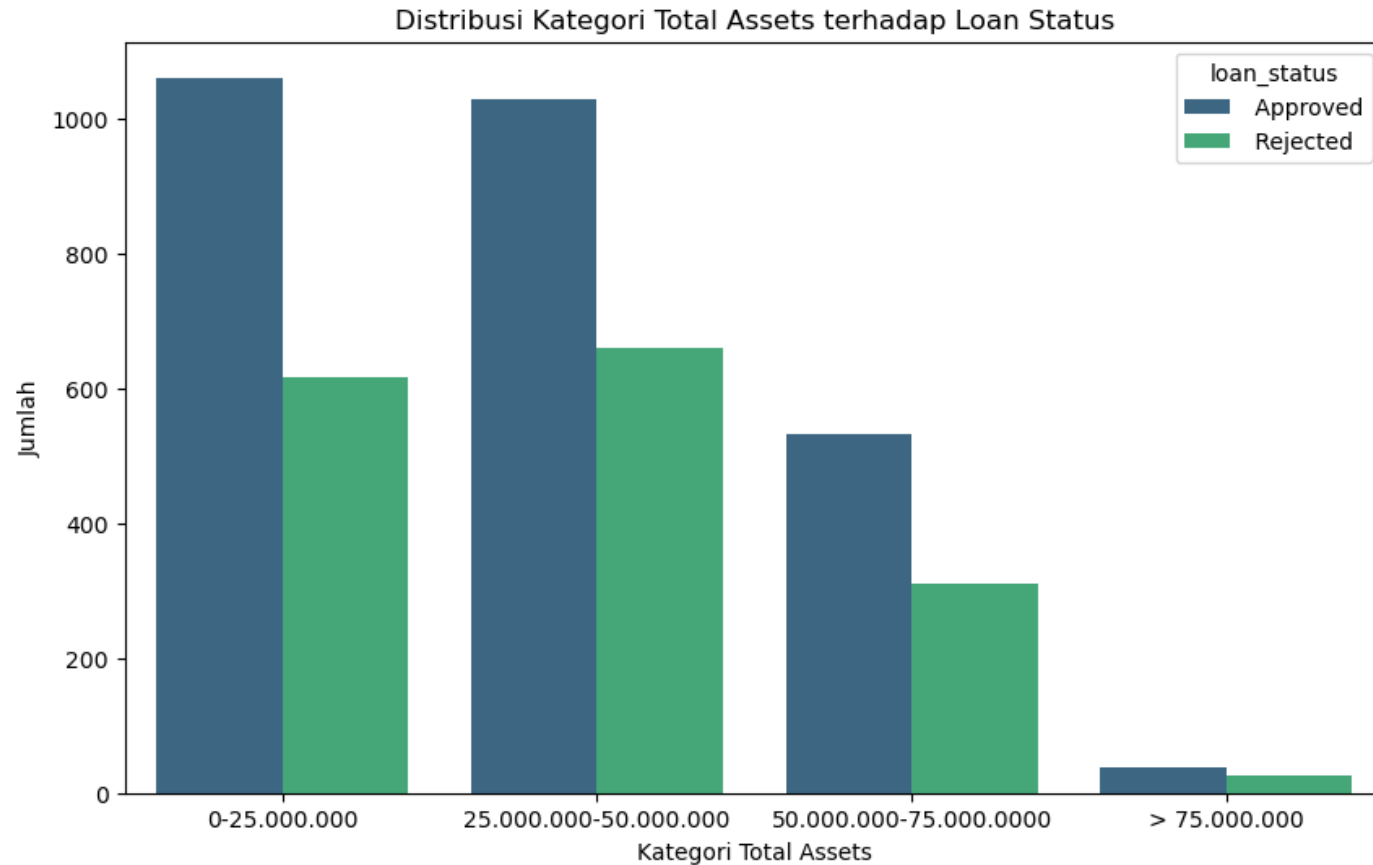
Apakah semakin tinggi pendapatan akan mempengaruhi persetujuan kredit?



Berdasarkan chart diatas, tidak ada perbedaan yang signifikan dari empat kategori pendapatan tahunan terhadap persetujuan kredit.

Exploratory Data Analysis (EDA)

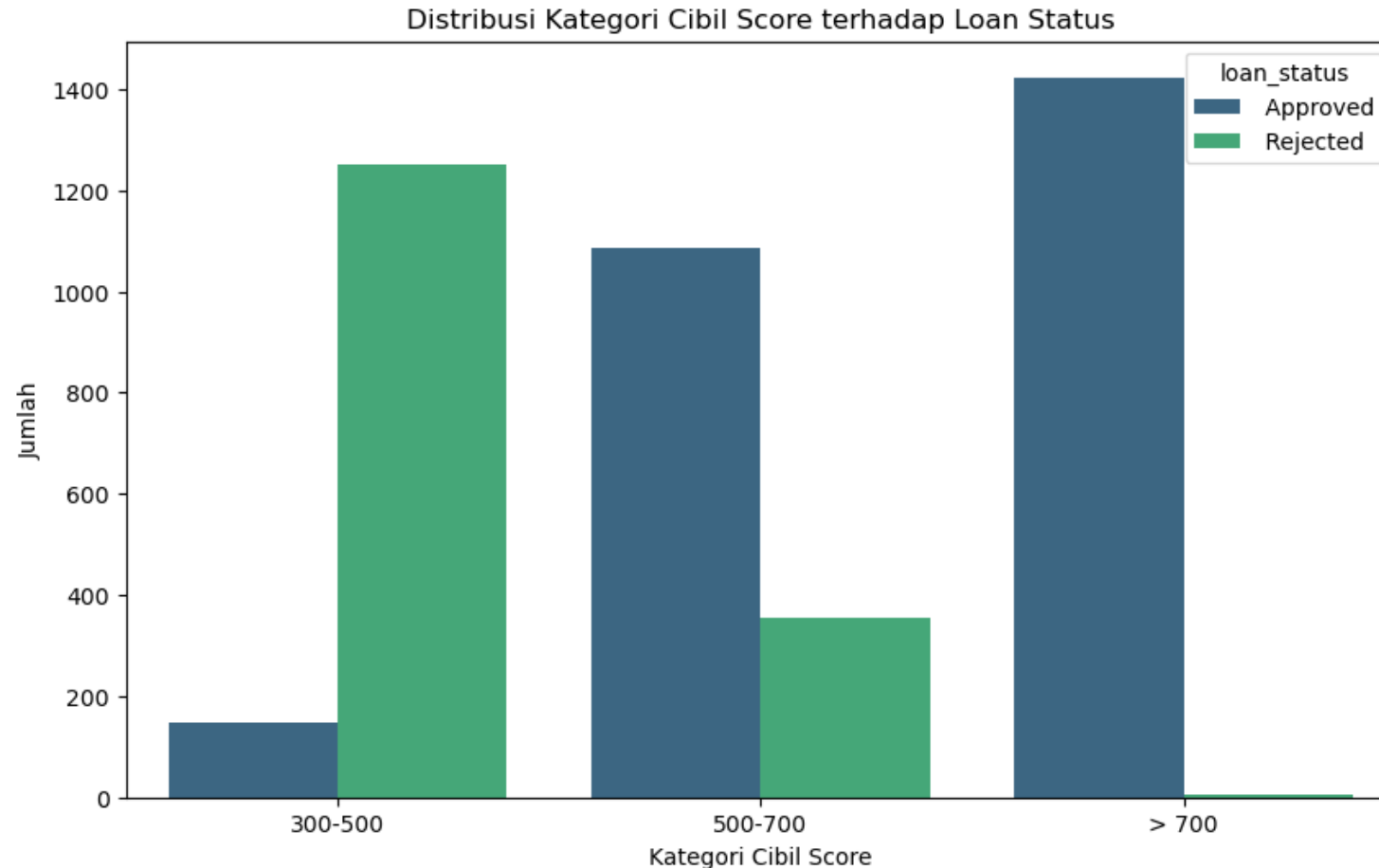
Apakah ada hubungan antara total asset terhadap persetujuan kredit?



Berdasarkan chart diatas, untuk jumlah asset sampai dengan 50.000.000 memiliki jumlah persetujuan kredit yang lebih tinggi dibandingkan dengan jumlah asset diatas 50.000.000

Exploratory Data Analysis (EDA)

Apakah ada hubungan antara score kredit terhadap persetujuan kredit?

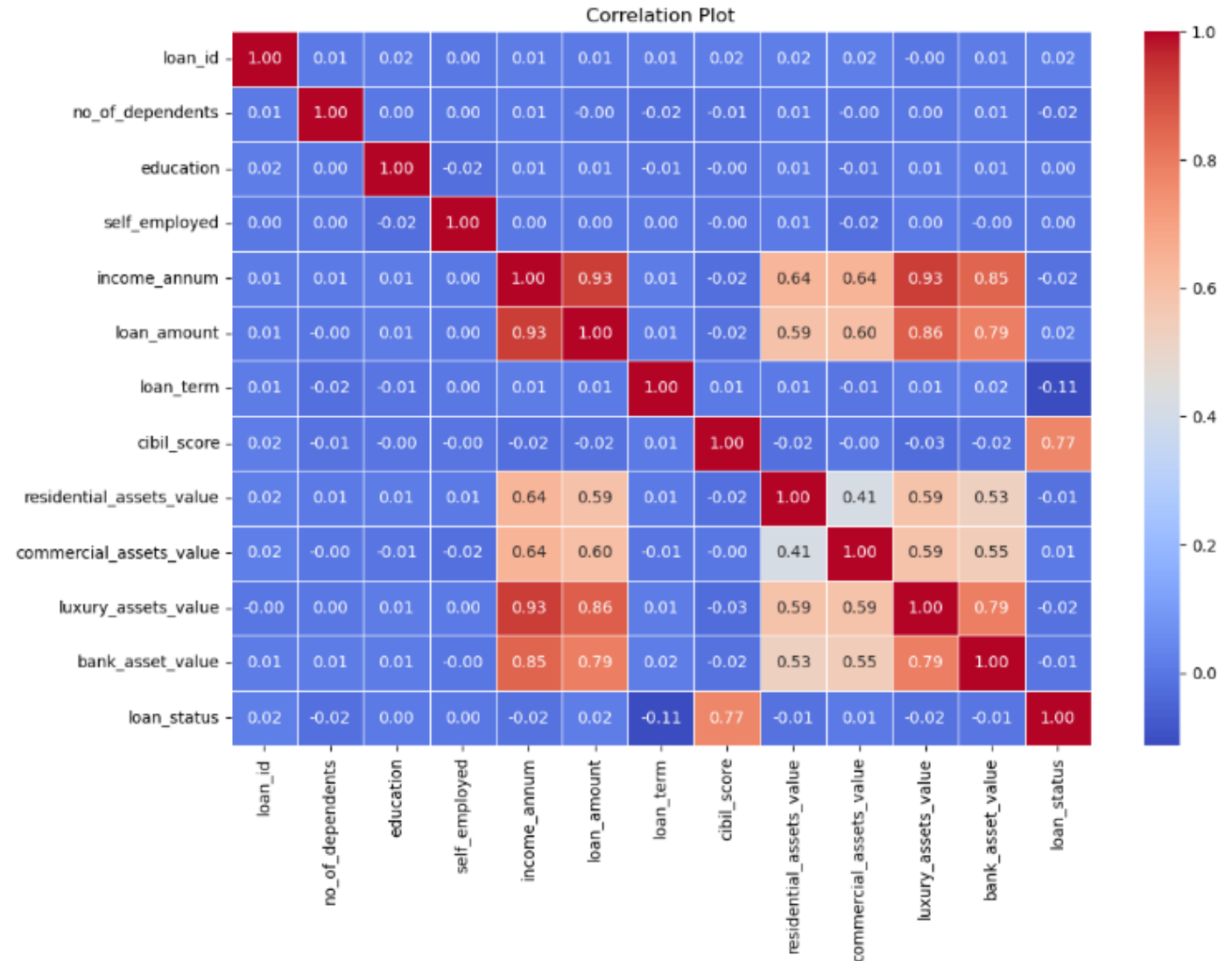


Berdasarkan chart diatas, dapat disimpulkan bahwa semakin tinggi Tingkat score kredit maka semakin besar memiliki peluang untuk disetujui pengajuan kredit.

Exploratory Data Analysis (EDA)

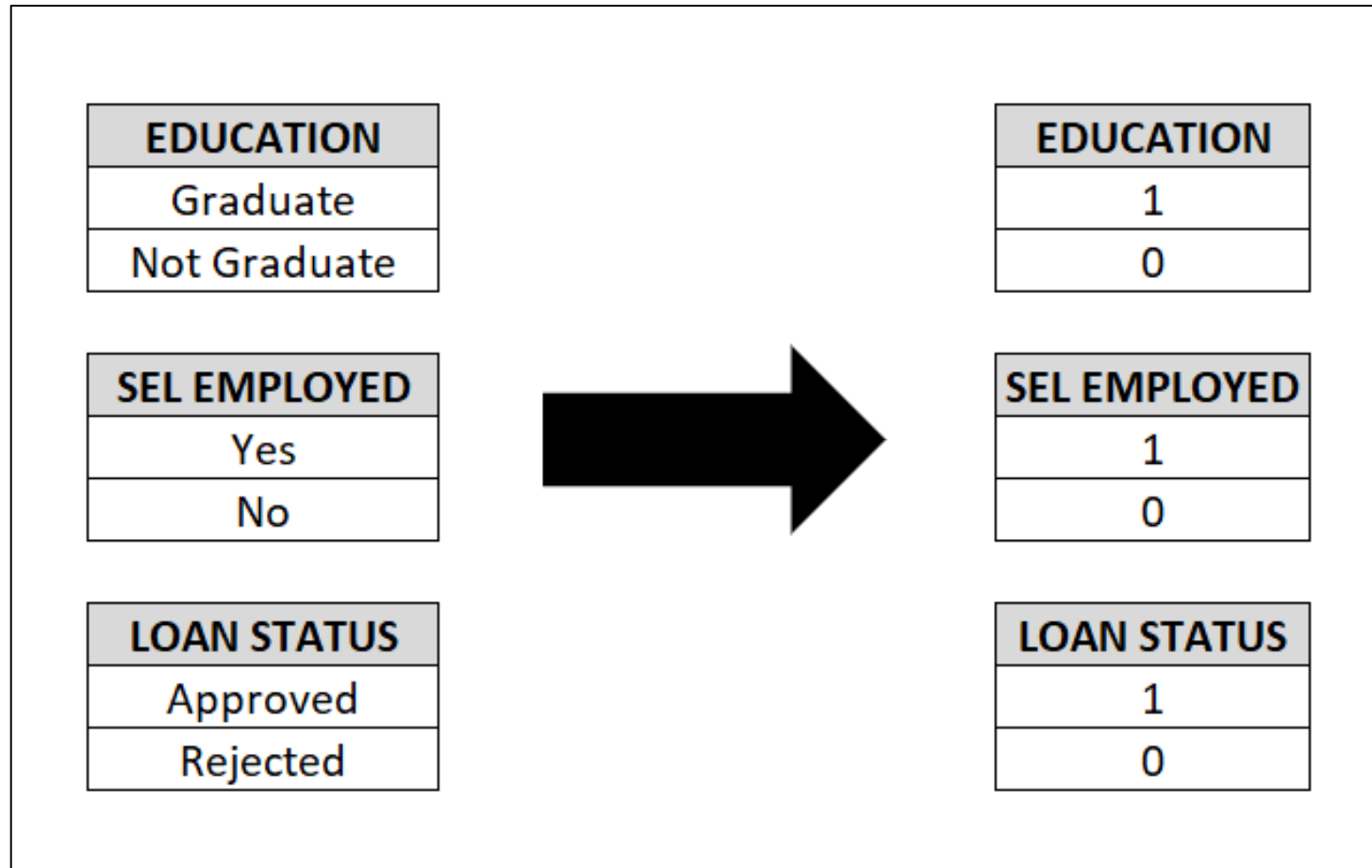
Analisa korelasi tiap kolom

- Income annum memiliki pengaruh yang tinggi terhadap loan amount, artinya semakin tinggi pendapatan maka pengajuan kredit yang diajukan akan semakin besar.
- Income annum juga memiliki hubungan yang cukup besar terhadap seluruh kategori asset, artinya semakin tinggi pendapatan maka asset yang dimiliki semakin besar.
- Cibil score memiliki pengaruh yang cukup besar terhadap loan status.



Feature Engineering

Label Encoding



Modelling

	Accuracy	Recall	Precision	F1 Score	Avg Cross Val
Logistic Regresion	72.59%	94.59%	71.51%	81.45%	72.78%
SVC	62.76%	100%	62.76%	77.12%	62.22%
Decision Tree	97.66%	98.32%	97.96%	98.14%	97.42%
Random Forest	97.98%	99.07%	97.61%	98.33%	98,01%

Random Forest adalah Model terbaik dikarenakan memiliki tingkat Accuracy yang tinggi diantara yang lain.

Joblib

Loan Status Prediction Using...

Loan Approval Prediction

no_of_dependents [1,2,3,4,5]	0
education [1:Gradute,0:Not Graduate]	0
self_employed [1:Yes,0:No]	1
income_annum	4100000
loan_amount	12200000
loan_term	8
cibil_score	417
residential_assets_value	2700000
commercial_assets_value	2200000
luxury_assets_value	8800000
bank_asset_value	330000

Predict

Loan Status Prediction Using...

Loan Approval Prediction

no_of_dependents [1,2,3,4,5]	0
education [1:Gradute,0:Not Graduate]	0
self_employed [1:Yes,0:No]	1
income_annum	4100000
loan_amount	12200000
loan_term	8
cibil_score	417
residential_assets_value	2700000
commercial_assets_value	2200000
luxury_assets_value	8800000
bank_asset_value	330000

Predict

Loan Rejected

Joblib

Loan Status Prediction Using...

Loan Approval Prediction

no_of_dependents [1,2,3,4,5]	4
education [1:Gradute,0:Not Graduate]	1
self_employed [1:Yes,0:No]	1
income_annum	3900000
loan_amount	12300000
loan_term	18
cibil_score	784
residential_assets_value	7600000
commercial_assets_value	6900000
luxury_assets_value	13100000
bank_asset_value	2800000

Predict

Loan Status Prediction Using...

Loan Approval Prediction

no_of_dependents [1,2,3,4,5]	4
education [1:Gradute,0:Not Graduate]	1
self_employed [1:Yes,0:No]	1
income_annum	3900000
loan_amount	12300000
loan_term	18
cibil_score	784
residential_assets_value	7600000
commercial_assets_value	6900000
luxury_assets_value	13100000
bank_asset_value	2800000

Predict

Loan Approved

Conclusion & Recommendation

- Random Forest memiliki keakuratan yang tinggi dalam memprediksi persetujuan pinjaman (97.98%).
- Keseimbangan antara recall (98,8%) & presisi (97,6%) menunjukkan bahwa model tersebut dapat mengidentifikasi baik pengajuan yang benar disetujui maupun ditolak.
- Perlu dilakukan pemantauan & memperbaharui model sesuai dengan perkembangan industri & teknologi kedepannya.



TERIMAKASIH

FINAL PROJECT - RADEK KILAT BUANA