



Credit Risk Prediction

By Hafshoh Nur Adzikra

Table Of Content

01

**Problem
Research**

03

**Exploratory
Data Analysis**

02

**Data
Processing**

04

Modeling

05

**Business
Recommendation**

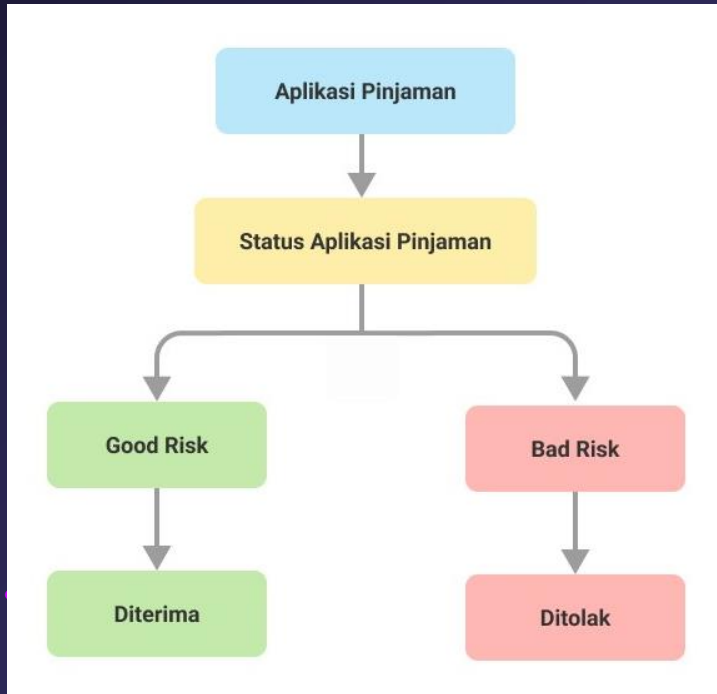


01

Problem Research



Business Understanding



- ❑ Perusahaan pinjaman memiliki masalah pada produk pinjaman yang ditandai dengan meningkatnya credit risk dibandingkan tahun sebelumnya.
- ❑ Perusahaan harus memutuskan untuk menyetujui atau menolak aplikasi berdasarkan profil pemohon.
- ❑ Perusahaan harus menghitung secara kuantitatif risiko kerugian akibat gagal bayar atau pelunasan awal.

Problem Statement

Secara kuantitatif terdapat dua risiko yang disebabkan oleh pemohon, yaitu :

- ❖ Good risk jika pemohon kemungkinan besar akan melunasi pinjaman.
- ❖ Bad risk jika pemohon kemungkinan tidak akan melunasi pinjaman.



Business Objective



Objectives

- Mengidentifikasi pola risiko
- Melakukan tindakan terhadap hasil yang didapat
- Menetapkan algoritma machine learning
- Memprediksi aplikasi pinjaman
- Mengambil keputusan secara otomatis

Goals

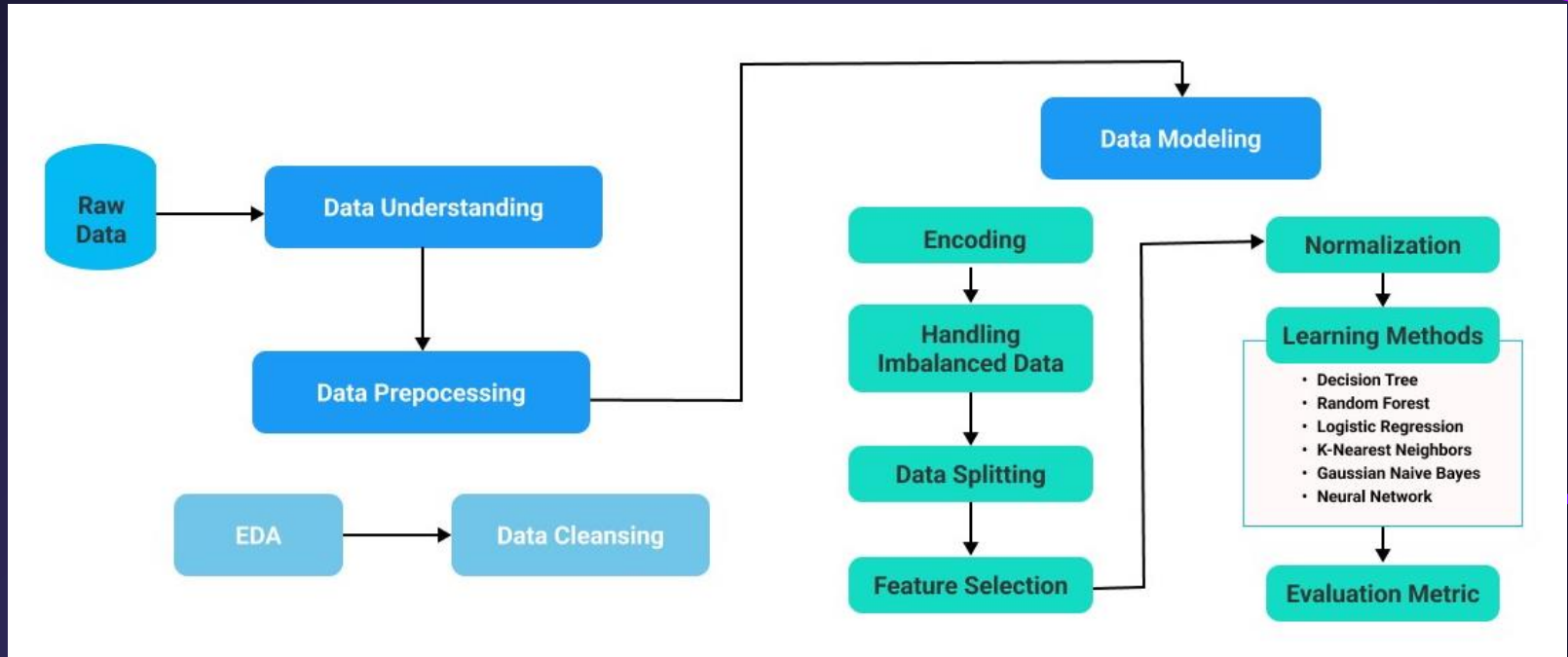
- Membuat machine learning dapat memprediksi risiko dari setiap pemohon
- Diharapkan model dapat menurunkan credit risk buruk pada perusahaan



02

Data Processing

Data Processing





03



Exploratory Data Analysis



Tentang Dataset

Loan Data

75

466,285

Numbers of features

Number of rows

0

Mising Value

0

Duplicates

Data Type

int64, float64, datetime64[ns], object

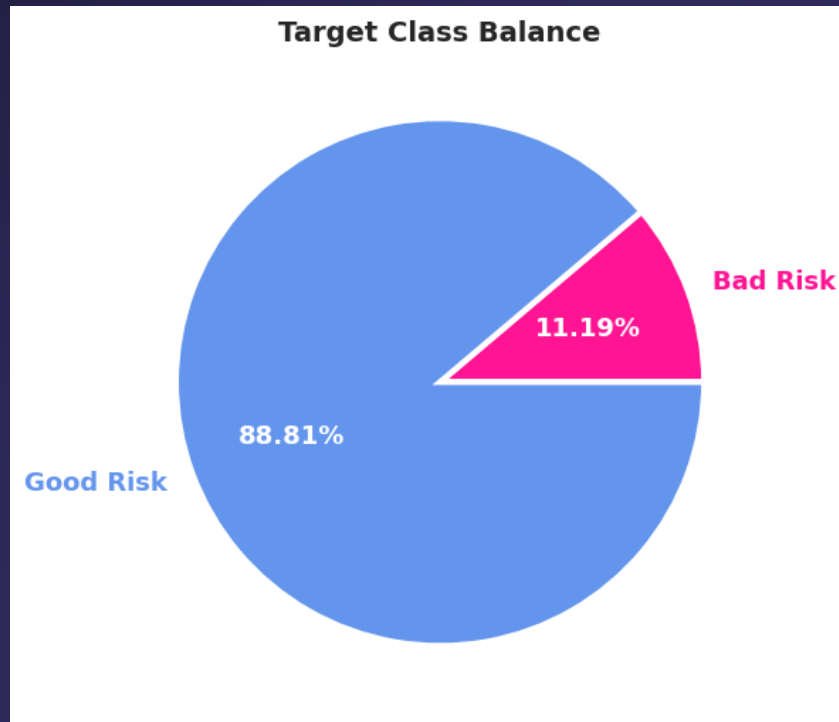
Data Numerikal

Terdapat 53 fitur numerikal

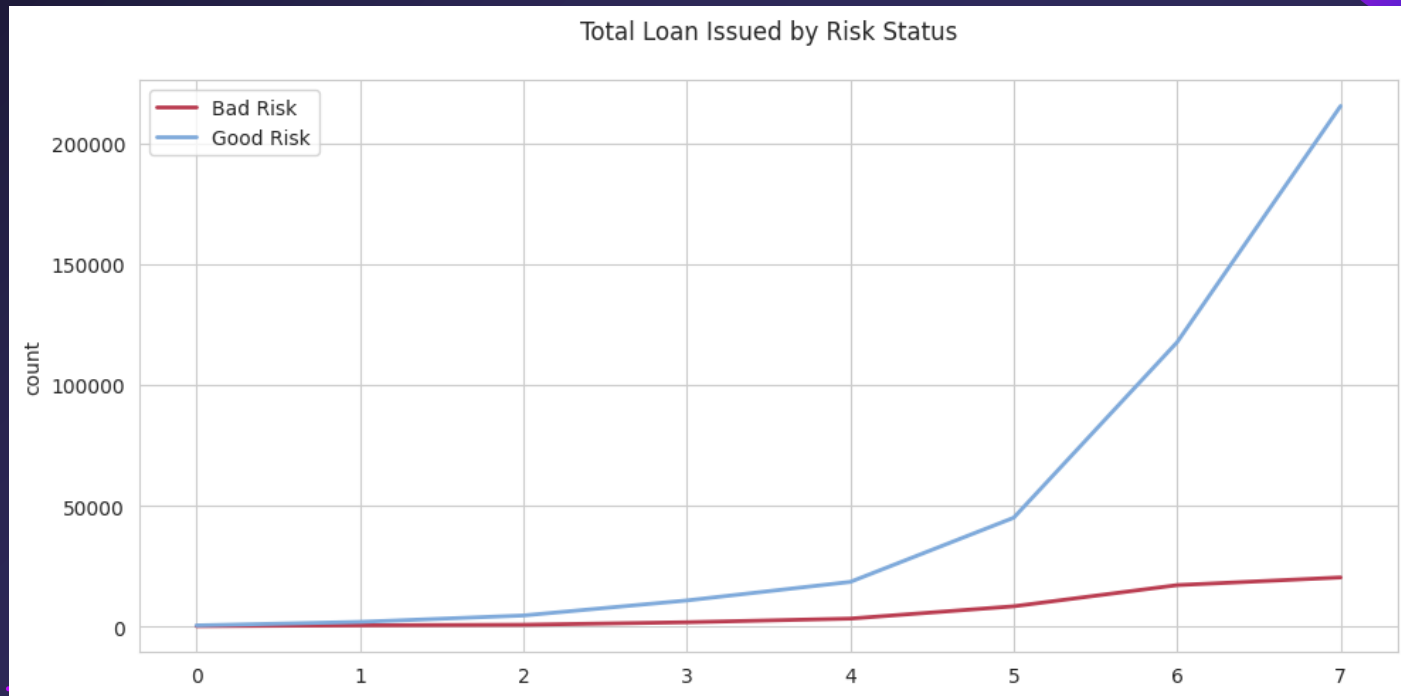
Data Kategorikal

Terdapat 17 fitur categorical

Insight



Insight





04

Modeling

Classification Model

Learning Method	Training Accuracy	Testing Accuracy	Error Margin
Decision Tree	99.98%	98.32%	1.66%
Random Forest	99.98%	99.11%	0.87%
Logistic Regression	88.38%	88.41%	0.03%
K-Nearest Neighbor	97.61%	96.7%	0.91%
Gaussian Naïve Bayes	75.08%	74.98%	0.1%
Neural Network	98.99%	98.72%	0.27%
XGBoost Classifier	99.39%	99.08%	0.31%
Gradient Boosting Classifier	97.56%	97.52%	0.04%

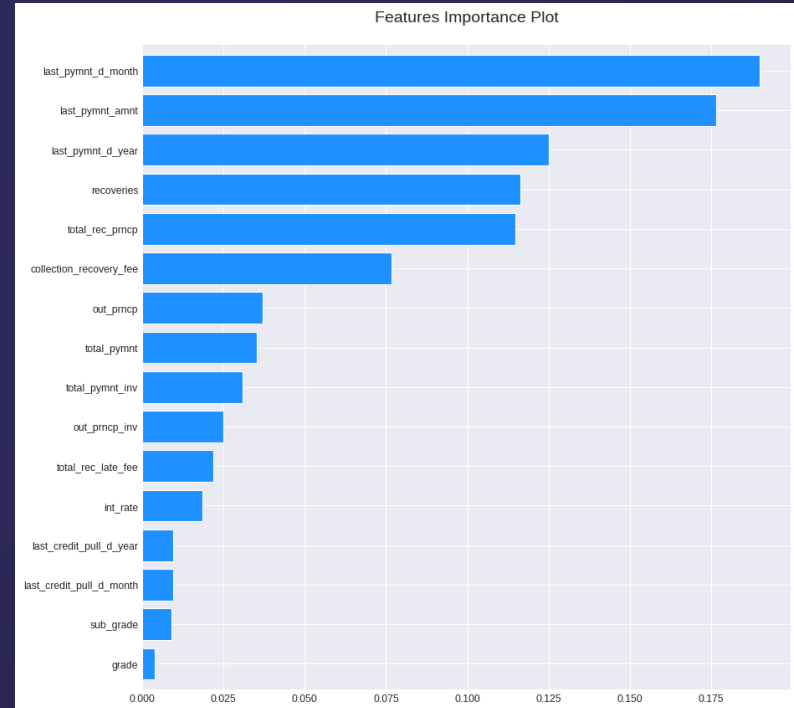
Final Model

Random Forest

Training Accuracy	Testing Accuracy	Error Margin
99.98%	99.08%	0.9%

Alasan pemilihan model

Perbedaan akurasi pada saat dilakukan pelatihan dan pengujian sangat kecil. Random Forest memiliki generalisasi yang lebih baik dibandingkan model lainnya





05



Business Insights & Recommendation



Business Insights

Total Loss Suffered By The Company

Loan Status	Total Loss	% of Total Loss	Total Applicants	Average Loss
Charged Off	\$574,356,330	83.49%	43.236	\$13,284
Late	\$102,293,296	14.87%	8.118	\$12,600
Default	\$11,299,446	1.64%	832	\$13,581

Status Charged off mengalami kerugian yang paling besar dari semuanya dengan total kerugian 574 juta dari 43.236 pemohon

Untuk status Late, mengalami kerugian sebesar 14% dengan total kerugian 102 juta dari 8.118 pemohon

Untuk status Default hanya mengalami kerugian 1,64% dengan total kerugian 11 juta dari 832 pemohon. Namun status pinjaman ini memiliki rata-rata kerugian tertinggi yaitu \$13,581 perpemohon

Business Recommendation



Untuk machine learning

Perusahaan dapat melihat ciri – ciri terpenting seperti bulan pembayaran terakhir, jumlah pembayaran terakhir, dan lain sebagainya untuk mengidentifikasi apakah pemohon memiliki kemungkinan menguntungkan atau merugikan.

From insight

Perusahaan dapat mengambil keputusan mutlak seperti menolak pinjaman, mengurangi jumlah pinjaman dan lain sebagainya untuk menghindari dan mengurangi total kerugian yang diderita oleh persahaan



Documentation

Kamu dapat melihat dokumentasi dan penjelasan lebih lengkap pada link di bawah ini :

https://github.com/sseiraa/Final-Project_Id.x-Partners

The background is a solid dark purple. In the top-left corner, there is a cluster of small, light purple dots of varying sizes. On the right side, there are several concentric, wavy lines in a lighter shade of purple, creating a sense of depth and movement. A thin, horizontal purple line is positioned directly beneath the word 'Terimakasih!'.

Terimakasih!