



FLOUPIN

PAYMENT DEFAULT
PREDICTION

Final Project



Welcome!!

BANK JANATA

Thank you for trusting Floupin
Team to solve the problem

Meet Our Team



Rafif Noor Audryan

Data Scientist



**Alfiyan Faissofi
Hidayat**

Data Scientist



Arief Mis Hargi

Data Scientist Lead



Nanda Dwi Putra

Data Scientist



Firly Putri Fardilla

Data Scientist



Agenda Style

01

Problem Business

Business Understanding | Business Purpose | Business Metrics

02

Data Analysis

Exploratory Data | Data Pre-processing | Data Insight

03

Modeling

Model Evaluation | Modeling | Feature Importances

04

Business Recommendation & Simulation

Business Flow | Business Recommendation | Simulation

05

Result

Result



PROBLEM BUSINESS

Business Understanding | Business Purpose | Business Metrics

Background

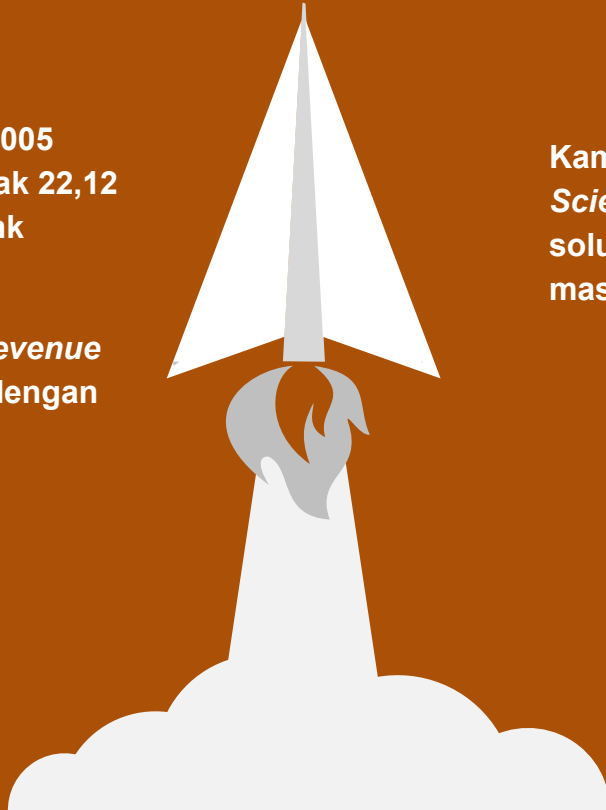


Bank Janata Taiwan pada tahun 2005 memiliki problem dimana sebanyak 22,12 % dari keseluruhan *customer* Bank Janata gagal bayar (*default*).

Hal ini tentu dapat menurunkan *revenue* Bank Janata jika tidak ditangani dengan baik.



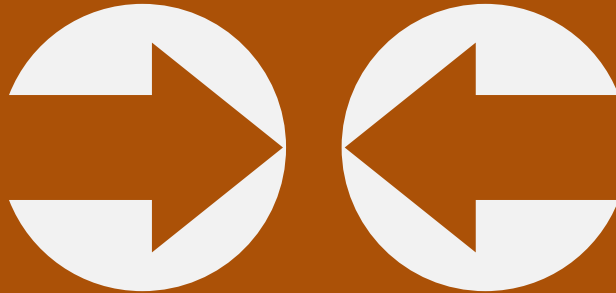
Kami Floupin sebagai *Team Data Scientist* akan membantu memberikan solusi kepada *Team Business* terhadap masalah yang dihadapi Bank Janata



Goals & Objective

Goals

- Menurunkan *Default Rate*
- Meningkatkan *Revenue* Bank
- Menurunkan kerugian Bank.



Objective

Memprediksi *customer* yang gagal bayar di bulan berikutnya

Business Metrics



Payment Default Rate

Menghitung rasio dari nasabah yang gagal bayar di bulan berikutnya



Lalu bagaimana
tim Floupin
membantu
Bank Janata
dalam
mengatasi
problem
tersebut?



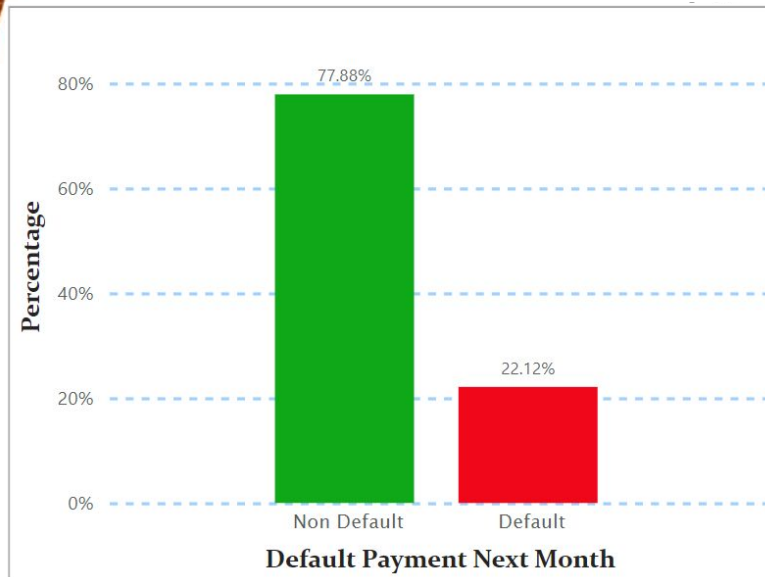


DATA ANALYSIS

Exploratory Data | Data Pre-Processing | Data Insight

Customer Overview

- Data customer terdiri dari 21.000 customer dan 25 atribut
- Sebagian besar data bersifat numerik
10 atribut kategorik
15 atribut numerik



Data Timpang dengan persentase

77,88% : 22,12%

Meskipun data didominasi oleh nasabah yang tidak *default*, kegagalan bayar nasabah menjadi penyebab meningkatnya *default rate* dan membuat *revenue* dari bank menurun

Customer Profile



Gender

Male : 39%
Female : 61%

Education

Graduate School : 35%
University : 47%
High School : 17%
Others & Unknown : 2 %

Age Group

Young (< 25 yo) : 9%
Middle (25 - 50 yo) : 82%
Senior (50 - 60 yo) : 1%
Elderly (> 60 yo) : 8%

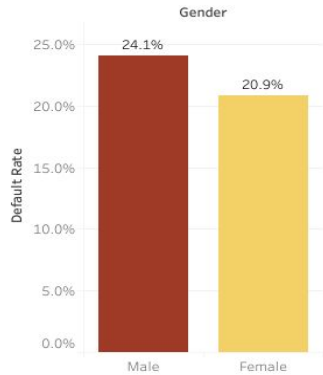
Segment

Low_limit (< 7jt) : 48%
Medium_limit (7jt - 12jt) : 26%
High_limit (> 12jt) : 26%

Siapa saja yang lebih berpotensi untuk default ?

JANATA Bank Customer Overviews on September 2005

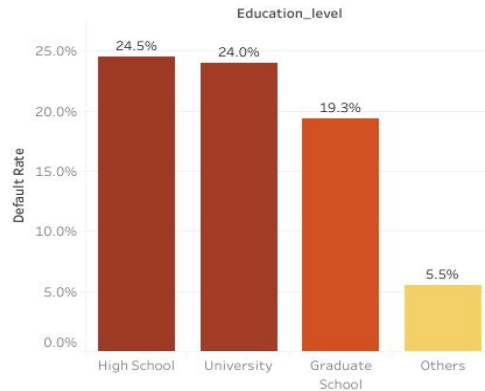
Default Probability by Gender



Male

Nasabah yang *default* tertinggi didominasi oleh customer bergender **Pria**

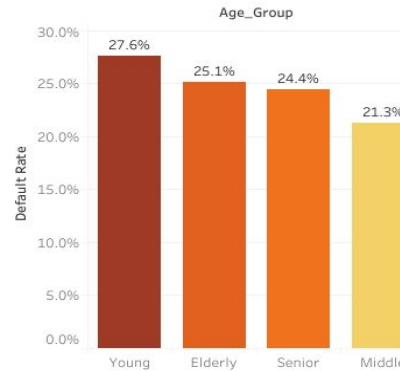
Default Probability by Education Level



High School

Nasabah dengan edukasi **sekolah menengah** lebih banyak *default*

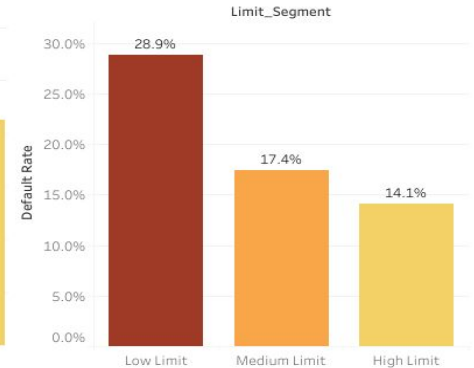
Default Probability by Age



Young

Nasabah yang *default* tertinggi yaitu kelompok usia **muda**

Default Probability by Limit Segment



Low Limit

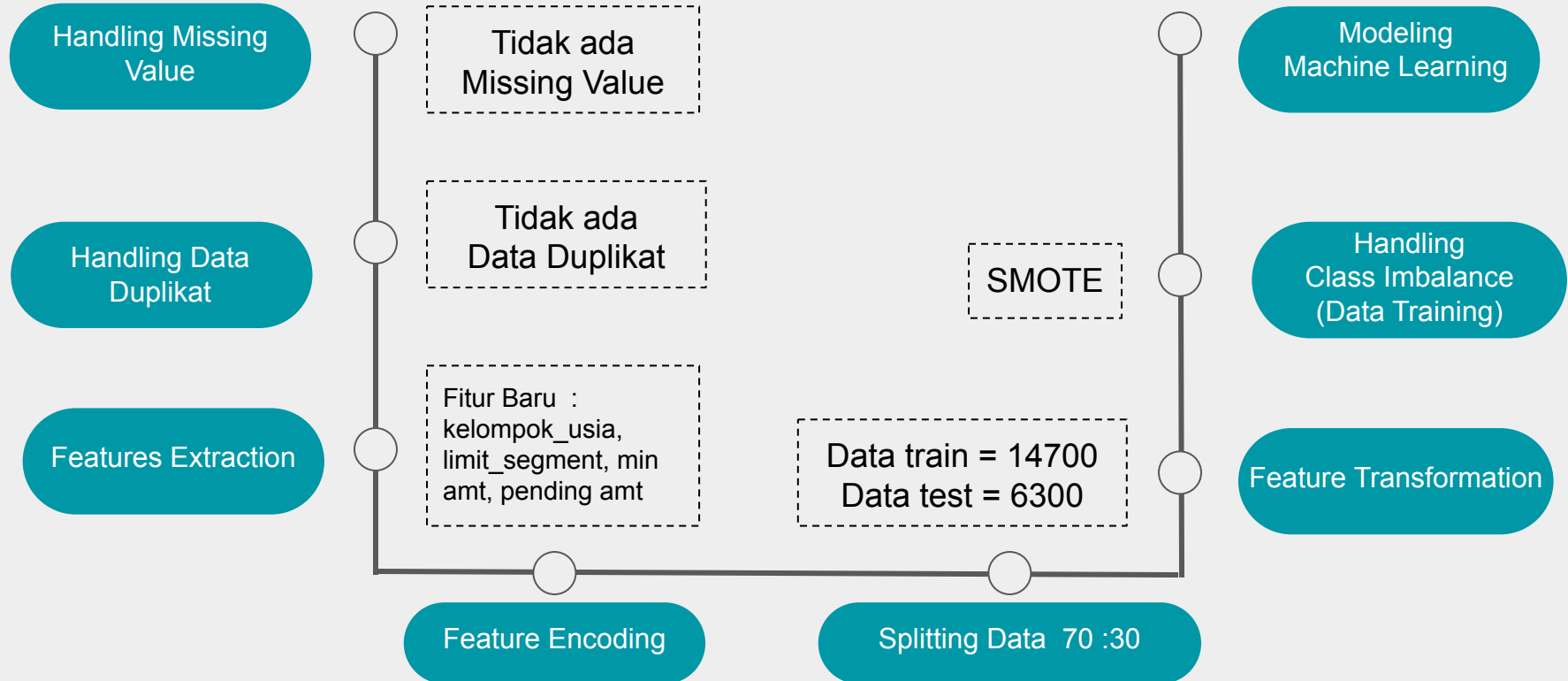
Nasabah dari segmen **low_limit** banyak yang *default*

Data Insight



Kami menemukan beberapa insight dari ciri-ciri *customer* yang berpotensi tinggi default. Untuk mendukung prediksi kami maka diperlukan pembuatan model machine learning. Sebelum melakukan modeling maka diperlukan *pre-processing* data untuk menghasilkan prediksi model yang tidak bias

Data Pre-Processing





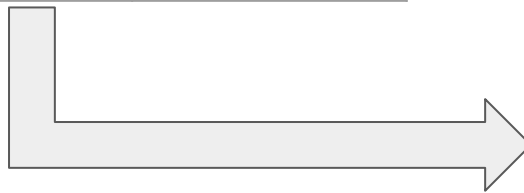
MODELING

Model Evaluation | Modeling | Feature Importances

Model Evaluation

$$\text{Precision} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})}$$

	Positive	Negative
Positive	TP	FN
Negative	FP	TN



<u>Actual</u>	<u>Prediksi</u>	
	Default	Tidak Default
Default	463	945
Tidak Default	233	4659

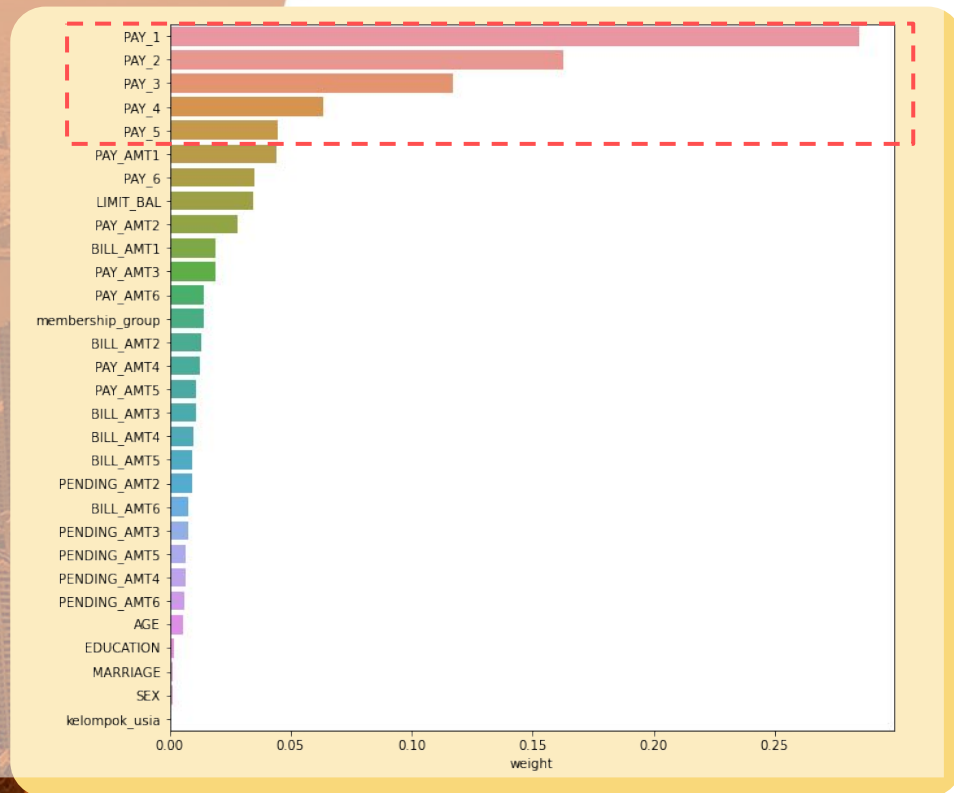
Modeling

Model	Precision	Recall	F1-Score	AUC Test	AUC Train
kNN	63%	24%	34%	74%	79%
Decision Tree	58%	22%	32%	71%	76%
Logistic Regression	36%	29%	32%	58%	74%
Random Forest	66%	36%	46%	78%	85%
XGBoost	64%	36%	46%	78%	91%

Algorithm terbaik : Random Forest Classifier
Hyperparameter : RandomizedSearchCV
CV : 5
Scoring : Precision



Features Importances



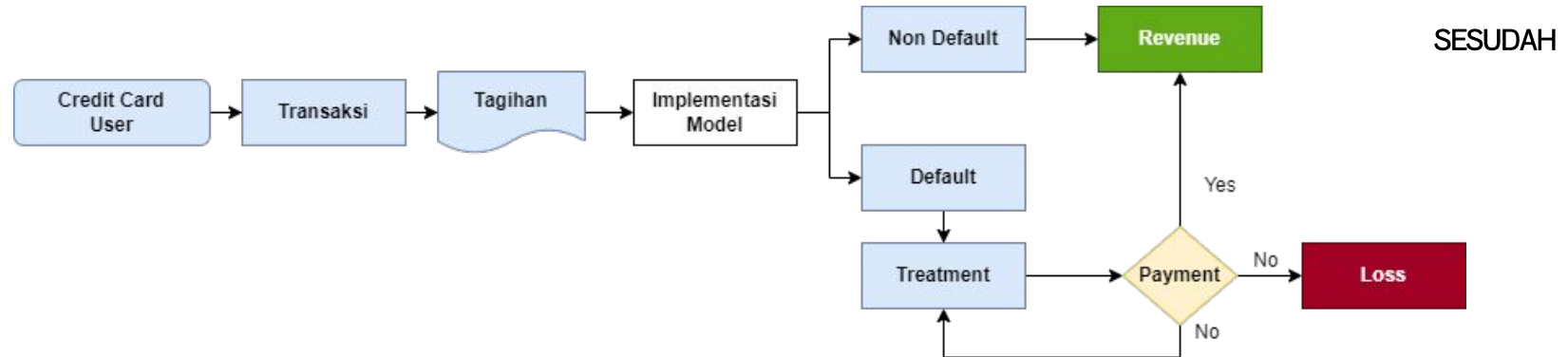
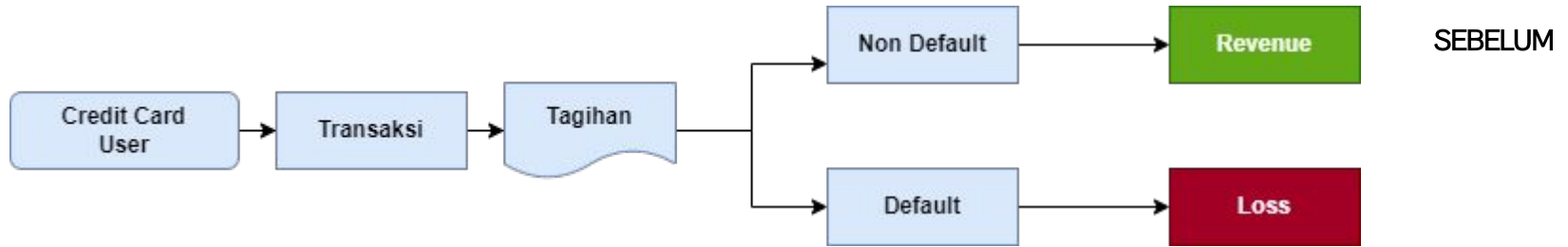
PAY_1 : September
PAY_2 : Agustus
PAY_3 : Juli
PAY_4 : Juni
PAY_5 : Mei



BUSINESS RECOMMENDATION & SIMULATION

Business Flow | Business Recommendation | Simulation

Business Flow



Business Recommendation (I)

Handling Customer Default



SMS & Telepon Reminder



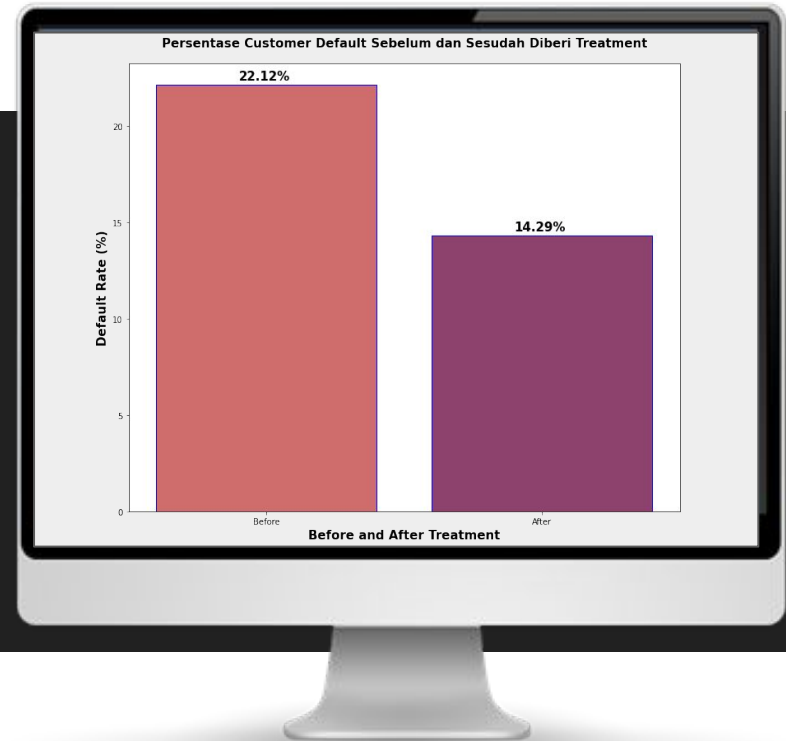
Skema Pembayaran dengan Cicilan



Penghapusan Bunga & Denda



Penagihan bagi Customer yang tidak melakukan pembayaran



Business Recommendation (II)

Handling Customer Non Default

Internal



Point Reward



Insentif bagi Team Reminder



Validasi *Income* guna penentuan Limit

Simulation Business Recommendation

Proyeksi Perhitungan Cost Sebelum Implementasi Model

Data sampel menggunakan data test (6300 data)

Default Rate sebesar 22%

Deskripsi	Nominal
Total Tagihan (a)	104,842,703,382
Tagihan yang berhasil dibayar (Revenue) (b)	82,381,145,025
Tagihan yang gagal dibayar (Loss)	22,461,558,357

Kerugian Bank sebesar
21,4%

Simulation Business Recommendation

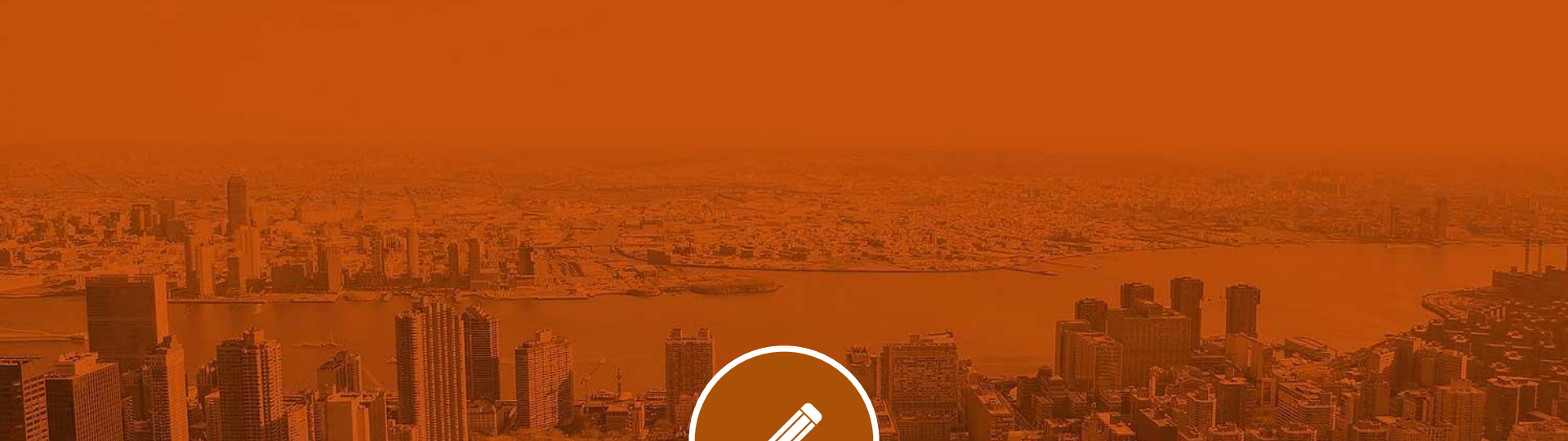
Proyeksi Perhitungan Cost Setelah Implementasi Model

Data sampel menggunakan data test (6300 data)

Default Rate turun menjadi 14%

Deskripsi	Nominal
Total Tagihan (a)	104,842,703,382
Tagihan yang berhasil dibayar (Revenue) (b)	82,381,145,025
Tagihan yang gagal dibayar (Loss)	22,461,558,357
Biaya Operasional (c)	64,403,375
Tagihan yang berhasil dibayar (Revenue) (e)	89,881,816,026
Tagihan yang gagal dibayar (Loss)	14,960,887,356
%Cost Efficiency Ratio (Perbandingan Biaya dengan Revenue) (f = c/e)	0.07%
Δ Revenue (g = e-b)	7,500,671,001
% Δ Revenue (h = g/b)	9.10%

Biaya operasional digunakan untuk melakukan *treatment* bagi customer yang terprediksi *Default* saja.



RESULT

Result

Before

22 %

Default Rate



After

14 %

Default Rate

82 Miliar

Revenue



90 Miliar

Revenue

22 Miliar

Loss



14 Miliar

Loss

Goals

- Potensi **menurunkan** *Default Rate* sebesar **8%** ↓
- Potensi **meningkatkan** *Revenue Bank* sebesar **9%** ↑
- Potensi **menurunkan** *Kerugian bank* sebesar **33%** ↓

An aerial photograph of a city skyline, likely New York City, viewed from a high angle. The image is overlaid with a large, semi-transparent orange circle that has a dashed white border. The text "Thank you" is centered within this circle in a white, sans-serif font. In the bottom center, below the circle, the text "Floupin Team" is visible in a smaller white font. The background shows a dense urban landscape with numerous skyscrapers and a body of water in the distance. The overall color scheme is dominated by the orange overlay.

Thank you

Floupin Team