

# CREDIT RISK Prediction Model

Home Credit Indonesia – Data Scientist

Oleh : Zaldi Abdul Hadi



# Problem Identification

## Business Problem

Home Credit harus menjamin bahwa pelanggan yang memiliki kemampuan membayar tidak ditolak, dan bahwa struktur pinjaman (termasuk pokok, tenor, dan jadwal) dirancang secara efektif untuk mendorong dan mendukung keberhasilan pelunasan oleh pelanggan.



## Goals

Pengembangan model prediktif yang akurat untuk menilai risiko setiap peminjam secara objektif guna mengurangi kerugian akibat kredit macet dan meningkatkan efisiensi operasional dalam menyetujui pinjaman.

# Dataset Overview

## Struktur Data Utama

Komponen	Data Training (Train)	Data Testing (Test)	Keterangan
Jumlah Baris	307.511	48.744	Total observasi/kasus peminjam.
Jumlah Kolom	122	121	Jumlah fitur (variabel) yang tersedia.

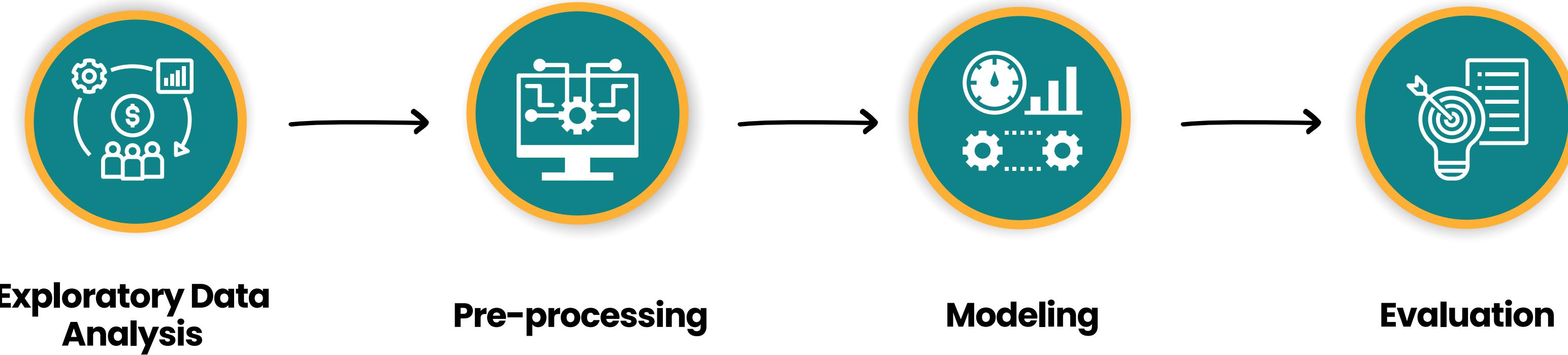
## Variabel Target dan Tujuan Model

Kriteria	Deskripsi
Variabel Target	TARGET
Nilai Target	<b>1</b> = Gagal Bayar
	<b>0</b> = Pembayaran Lancar
Tujuan Model	Memprediksi probabilitas pemohon akan gagal bayar (TARGET = 1) berdasarkan fitur yang ada.

## Fitur Utama

- Data Demografis:** Meliputi usia, jenis kelamin, tingkat pendidikan, dan status keluarga.
- Pekerjaan dan Penghasilan:** Mencakup jenis pekerjaan, stabilitas pekerjaan, dan besaran penghasilan tahunan.
- Kredit Sebelumnya (Past Credit):** Informasi historis mengenai kredit yang pernah dimiliki atau diajukan peminjam (terdapat 83 fitur tambahan lainnya).
- Properti dan Aset:** Data terkait kepemilikan aset, seperti rumah atau mobil.

# Data Pipeline



# Logistic Regression Model Evaluation

	Precision	Recall	F1 - Score
0	0.85	0.84	0.85
1	0.84	0.86	0.85
Accuracy : 0.85			
AUC Score : 0.9213			

# XGBoost Model Evaluation

	Precision	Recall	F1 – Score
0	0.94	1.00	0.97
1	1.00	0.94	0.96
Accuracy : 0.97			
AUC Score : 0.9859			

# LightBGM Model Evaluation

	Precision	Recall	F1 - Score
0	0.94	1.00	0.97
1	1.00	0.93	0.96
Accuracy : 0.97			
AUC Score : 0.9850			

# Business Insight

## Kualitas Kredit Unggul (Tingkat Persetujuan 12.63%)

Tingkat Pendidikan Tinggi terbukti memiliki kualitas kredit terbaik, ditandai dengan tingkat persetujuan tertinggi (12.63%) di seluruh segmen.



## Volume Akuisisi Rendah (Hanya 0.97%)

Segmen berkualitas tinggi ini hanya menyumbang 0.97% dari total volume pengajuan. Ini menunjukkan perusahaan gagal mengakuisisi pelanggan terbaiknya.



## Risiko Historis Rendah

Segmen ini memiliki bukti historis risiko kredit yang sangat rendah, menjadikannya kunci untuk menurunkan Cost of Risk portofolio secara keseluruhan.



## Prioritas Strategis: Alihkan Fokus Akuisisi

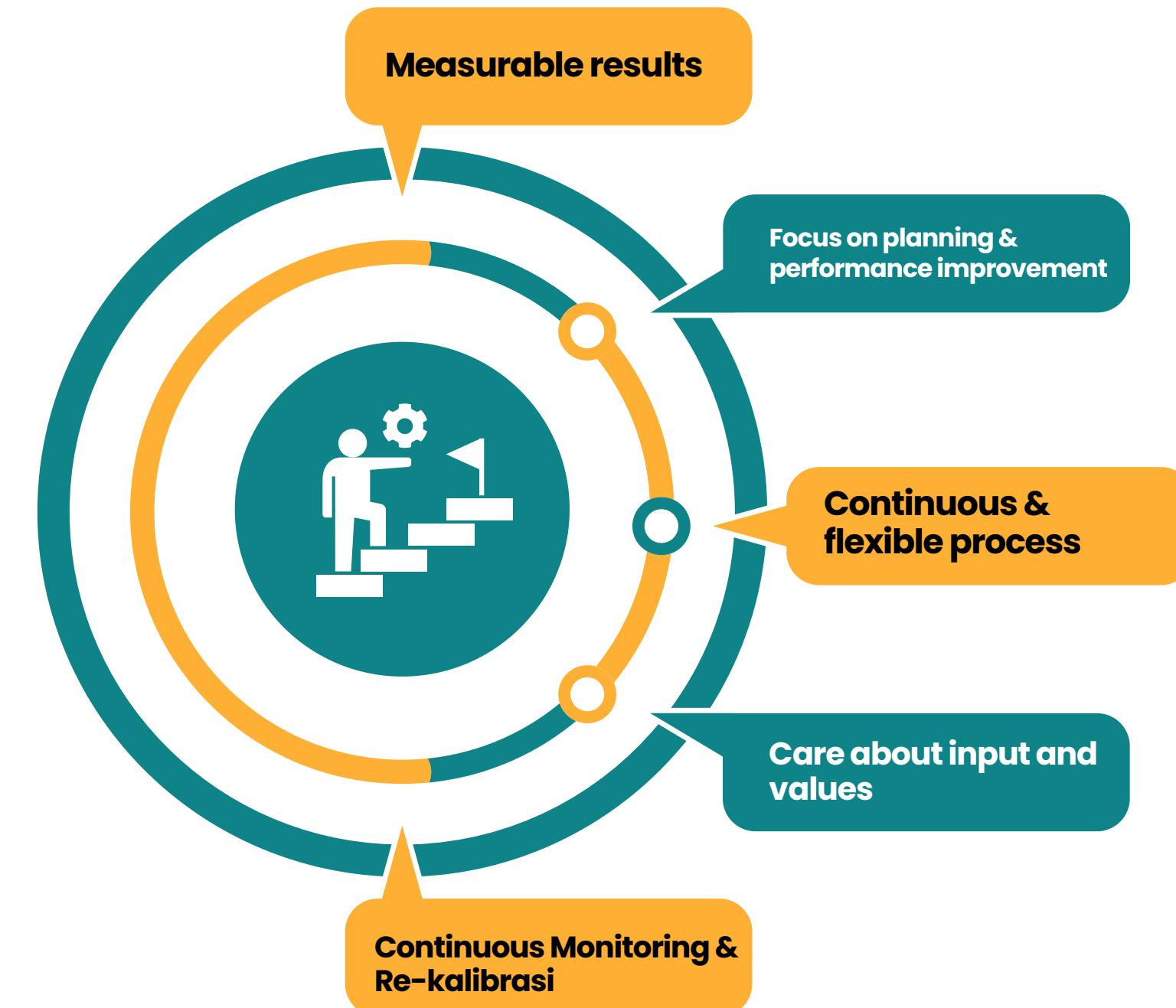
Dibutuhkan perubahan strategi untuk agresif menargetkan segmen berpendidikan tinggi dengan penawaran yang lebih baik, guna menyeimbangkan volume dan kualitas portofolio.



# Business Recommendation

Rekomendasi utama adalah peralihan strategi mendesak menjadi model risk-smart. Perusahaan harus segera meluncurkan Program Akuisisi Khusus yang secara agresif menargetkan segmen berpendidikan tinggi (kualitas superior, risiko rendah) dengan penawaran menarik (suku bunga rendah/batas tinggi) dan proses persetujuan digital yang cepat. Tujuannya adalah mendorong pertumbuhan volume pada segmen sehat ini yang saat ini hanya menyumbang 0.97% volume.

Secara bersamaan, untuk segmen volume tinggi yang berisiko, perusahaan harus memperketat mitigasi risiko dan meningkatkan pricing. Identifikasi variabel risiko tambahan dari EDA dan terapkan persyaratan underwriting yang lebih ketat atau penyesuaian harga strategis untuk menutupi Biaya Risiko (Cost of Risk). Strategi ganda ini akan memastikan pertumbuhan volume diimbangi dengan perbaikan kualitas portofolio secara keseluruhan.



## REFERENCES & GITHUB

### REFERENCES :

[Logistic Regression Documentation](#)

[LGBM Documentation](#)

[XGBoost Documentation](#)

### GITHUB :

<https://github.com/zaldiabdulhadi/rakamin-home-credit/>

Thank You