



# ANALISI DATA KAMPANYE PEMASARAN BANK

Data Science-Manajemen Rekayasa



# Latar belakang ?



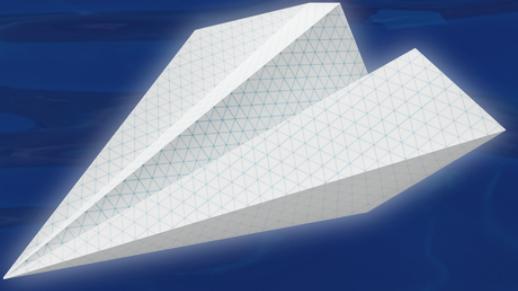
Kebutuhan untuk meningkatkan  
efektifitas kampanye pemasaran



Respon terhadap persaingan dan  
dinamika pasar



Dukungan untuk analisis berbasis  
data dan pengambilan keputusan



# Tujuan analisis & Manfaat



1. Memahami faktor-faktor yang memengaruhi keputusan klien untuk berlangganan deposito
2. Mengetahui Klien seperti apa (usia, pekerjaan, dll.) yang cenderung berlangganan?
3. Bagaimana efektivitas metode kontak (telepon vs seluler)
4. Apakah kondisi ekonomi memengaruhi keputusan klien?

**Manfaat:** Membantu bank merancang kampanye yang lebih tepat sasaran dan efisien.

# Data yang digunakan



## Data klien

Usia, pekerjaan, status pernikahan, pendidikan, status kredit, pinjaman.



## Data kampanye

Tipe kontak (telepon/seluler), bulan, hari, durasi panggilan, jumlah kontak, hasil kampanye sebelumnya



## Data ekonomi

Tingkat variasi pekerjaan, indeks harga konsumen, indeks kepercayaan konsumen, suku bunga Euribor, jumlah karyawan



## Variabel target

(yes/no, apakah klien berlangganan deposito berjangka).

# Exploratory Data Analysis (EDA)



## Definisi

EDA adalah proses awal untuk mengeksplorasi data guna menemukan pola, anomali, dan wawasan tanpa asumsi awal.



## Tujuan

1. Memahami distribusi data (misalnya, usia klien, durasi panggilan).
2. Mengidentifikasi hubungan antar variabel (misalnya, durasi panggilan vs keputusan berlangganan).
3. Mendeteksi anomali atau data yang tidak konsisten.

## Kelebihan

1. Memberikan pemahaman awal tentang data tanpa memerlukan model kompleks.
2. Membantu mengidentifikasi pola, tren, dan anomali.
3. Mudah dilakukan dengan visualisasi sederhana.

## Langkah-langkah EDA

Identifikasi korelasi (misalnya, heatmap untuk variabel numerik).

Visualisasi data (menggunakan grafik seperti bar plot, box plot, scatter plot).

4

Inspeksi data (memeriksa struktur, tipe data, dan missing values)

Analisis univariat (distribusi variabel tunggal, misalnya, histogram usia atau pie chart untuk y)

2

Analisis bivariat (hubungan antar variabel, misalnya, durasi panggilan vs y).

3

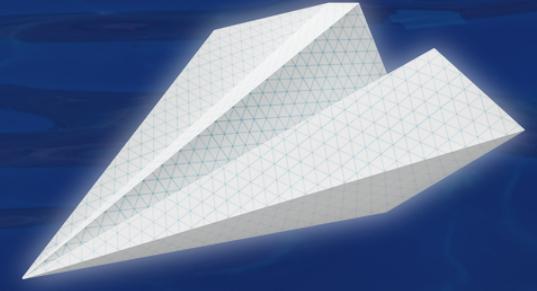
## Kekurangan

1. Tidak memberikan hasil prediktif (hanya deskriptif).
2. Bergantung pada interpretasi subjektif.
3. Tidak dapat menangani hubungan kompleks antar variabel tanpa metode lanjutan.

# Logistic Regression



Metode machine learning untuk memprediksi probabilitas klien berlangganan ( $y = \text{yes/no}$ ).



## Langkah

1

Encoding variabel kategorikal  
(misalnya, job, marital).

2

Membagi data menjadi set pelatihan  
dan pengujian.

3

Melatih model dan mengevaluasi  
akurasi.

## Kelebihan

1. Memberikan hasil prediktif yang jelas.
2. Mudah diinterpretasikan untuk data biner.

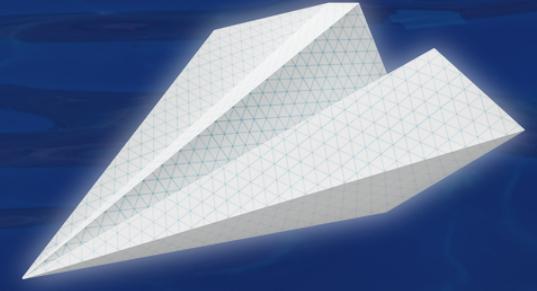
## Kekurangan

1. Membutuhkan preprocessing data yang ekstensif.
2. Kurang efektif untuk hubungan non-linear.

# Random Forest



Metode ini digunakan untuk membangun model yang kuat dan stabil dan akurat



## Langkah

1

Inisiasi model

2

Training

3

Prediksi



## Kelebihan

1. Akurasi tinggi.
2. Tahan Overfitting

## Kekurangan

1. Terlalu sederhana
2. Kurang akurat

# XGBoost

Metode ini digunakan untuk membangun model kuat, stabik dan akurat yang tahan terhadap data yang tidak sempurna

## Langkah

1

Inisiasi model

2

Training

3

Prediksi

## Kelebihan

1. Peforma terbaik
2. Fleksibel

## Kekurangan

1. kompleks
2. Rangan Overfitting

# Kesimpulan



Model Prediksi Berhasil Dibangun: Model XGBoost terbukti paling efektif untuk memprediksi nasabah potensial, dengan performa F1-Score 0.52 (untuk kelas 'Yes') dan ROC AUC Score 0.79. Ini menunjukkan model mampu mengidentifikasi target dengan cukup andal.



# Terima Kasih