



# **BANKING MARKETING DEPOSIT TARGET PREDICTION**

by

**Harta, Tahta, Data (HTD)  
Consultant**



**Introduction**

**Background**

**EDA**

**Data Preparation**

**Modeling and Evaluation**

**Business Insights  
and Recommendations**

# **Table Of Contents**

# HTD Consultant : Data Science Div. Team

01



Alfikri Ramadhan

02



Zahra Hanifah

03



Rantika Tresna

04



M. Dhiyaaul F

05



Dhea Putriani

06



Syahid Arbi

07



M. Fadil Biran



Introduction

# We and Our Client



## HTD Consultant

Kami adalah tim dari divisi Data di HTD Consultant. Dengan data analysis dan data science, kami memberikan solusi dari permasalahan bisnis



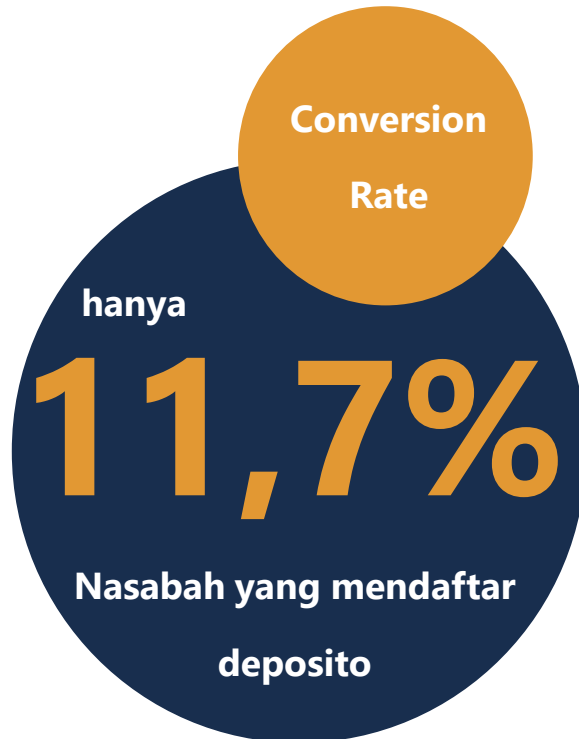
## Portubank

Portubank adalah perusahaan perbankan Internasional yang berbasis di Portugal. Portubank baru saja melakukan campaign telemarketing.



# Problem Statements

**Portubank** telah melakukan strategi **telemarketing** untuk mempromosikan produk **deposito berjangka** ke nasabahnya. Berdasarkan data 45.211 nasabah yang dihubungi :



- **Deposito berjangka** merupakan salah satu produk investasi perbankan dimana nasabah menyetorkan sejumlah uang untuk disimpan dalam kurun waktu tertentu di bank.
- Untuk sebuah bank, deposito berfungsi sebagai pemenuh kebutuhan modal bank dan membantu dalam menjaga posisi likuiditas bank.



**Telemarketing** adalah sebuah strategi marketing dengan menawarkan produk kepada pelanggan melalui telepon

## CONS

- Berkaitan langsung dengan Cost per Acquisiton
- **Cost per Acquisiton** adalah biaya yang dikeluarkan per customer yang sukses convert pada suatu campaign
- Biaya telemarketing adalah **EUR 20 per hour\***

### Pros and Cons Telemarketing

## PROS

- Cara paling efektif dibandingkan metode lain seperti : email marketing, advertisements, atau digital marketing.

Tim marketing perlu memilih **target nasabah yang tepat** untuk dihubungi sehingga dapat **menurunkan cost marketing**.



Background





## Goal

- **Meningkatkan conversion rate** nasabah yang diberikan campaign untuk mendaftar ke deposito



## Objective

- **Membuat model machine learning** yang dapat memprediksi nasabah yang akan mendaftar ke deposito
- Memberikan insight dan rekomendasi aksi yang dapat **mengoptimalkan** tim marketing bank memilih target nasabah



## Business Metrics

- **Conversion Rate**
- **Cost Per Acquisition**



Age  
Job  
Marital  
Education  
Default  
Balance  
Housing  
Loan  
Contact  
Day  
Month  
Duration  
Campaign  
Pdays  
Previous  
Poutcome  
**y (target)**

## Features

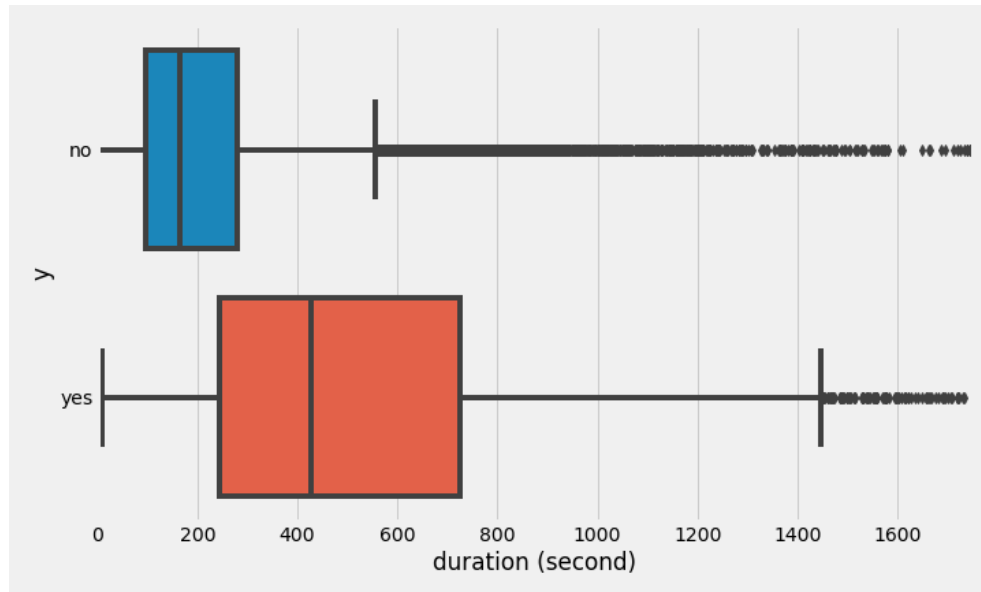
### Dataset Info

- Terdapat **45211 baris** dan **17 kolom**
- Terdapat 10 feature numerical dan 7 feature categorical
- 'y' merupakan **feature target** dengan value yes/no (ket. customer subscribe deposito berjangka)
- **Tidak ada missing values** maupun **data duplikat**



EDA



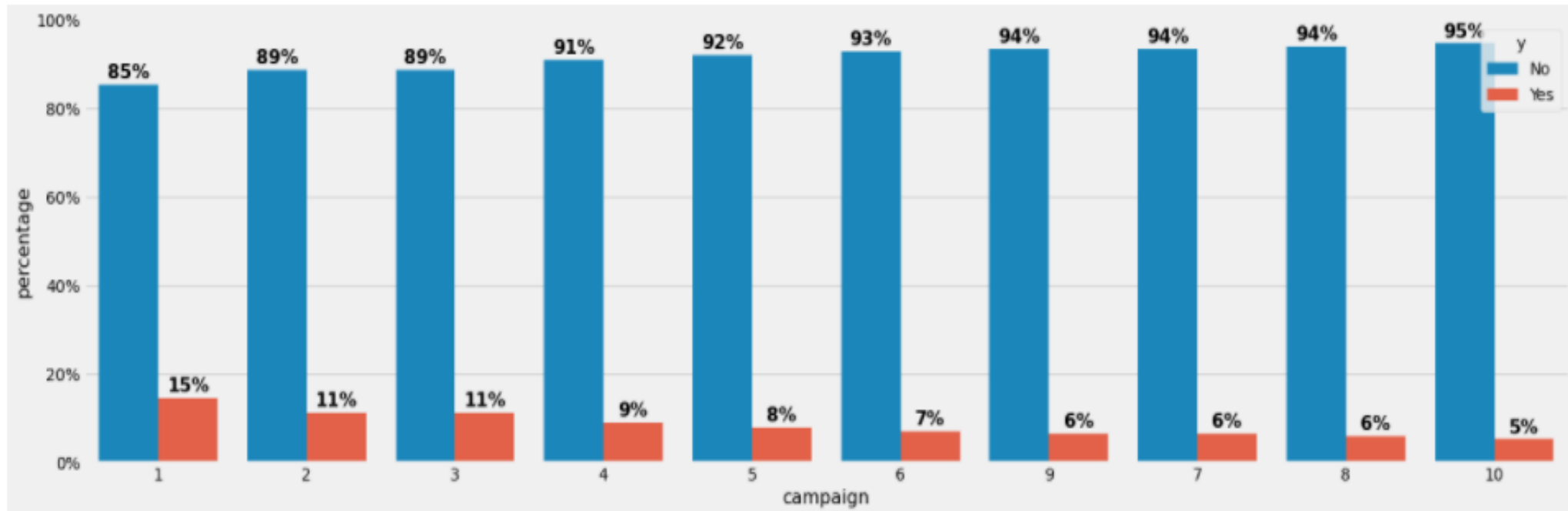


**Dapatkan atensi customer di 5 menit awal telepon!!**

**Duration** didefinisikan sebagai durasi kontak berlangsung (dalam detik). Dari EDA, kita dapat menyimpulkan:

- Customer dengan durasi kontak lebih lama cenderung untuk subscribe ke deposit
- Terdapat 'intersection' antara customer yang subscribe deposito dan tidak subscribe yaitu pada waktu 4 – 5 menit. Menurut kami ini adalah 'decisive moment'. Jika sampai waktu 5 menit customer masih belum tertarik, maka sebaiknya telepon dihentikan agar dapat menghemat waktu dan cost
- Tim marketing bisa membuat semacam "guidelines" yang menstandarisasi pesan yang disampaikan ke customer berdasarkan history data customer yang mendaftar deposito

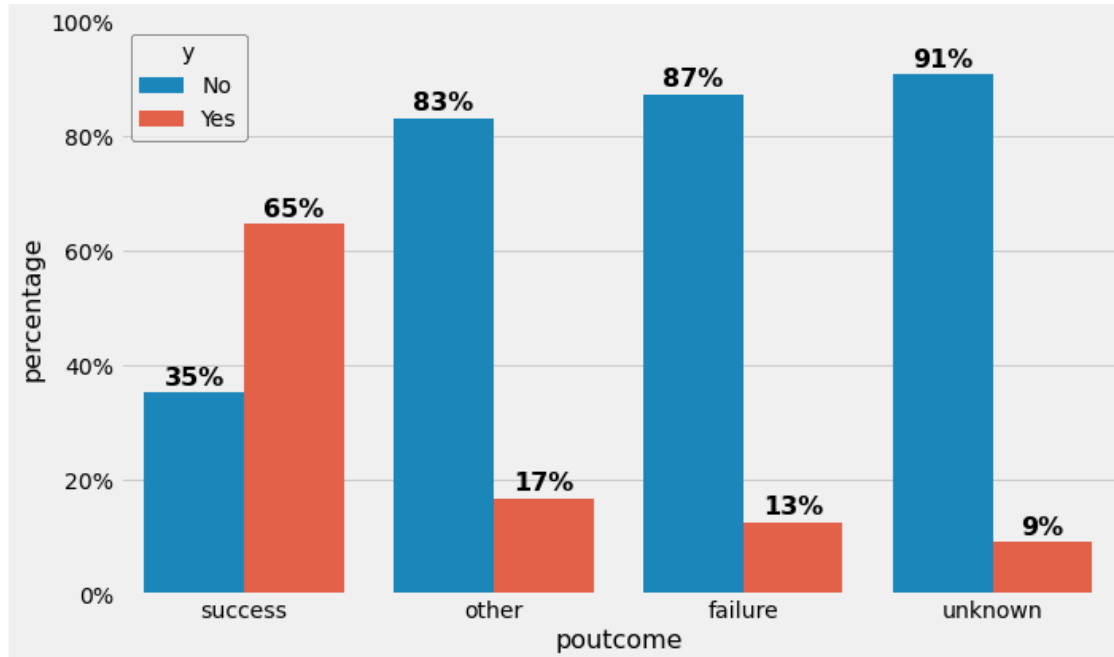




**Campaign** didefinisikan sebagai berapa kali customer dihubungi saat campaign ini. Dari EDA, kita dapat menyimpulkan:

- Jumlah campaign 1 s/d 3 tampaknya merupakan jumlah campaign optimum agar customer mau subscribe
- Jika dalam 3 kali telepon customer masih tidak menunjukkan ketertarikan ke produk deposito, maka disarankan campaign untuk customer ini dihentikan (tidak perlu dihubungi lagi, akan menghemat waktu dan cost)
- Sebaiknya tim marketing mencari metode yang dapat membuat customer tertarik untuk subscribe cukup dalam 3 kali telepon.



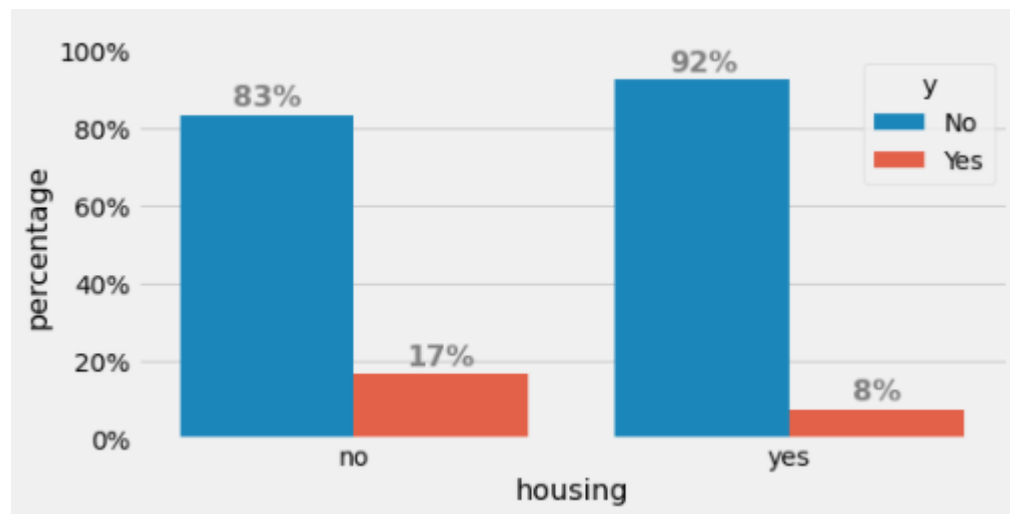
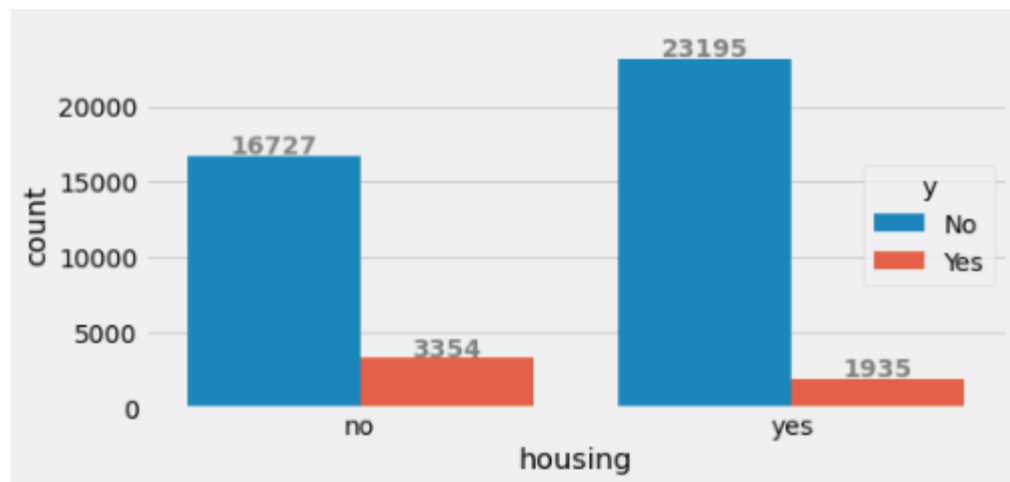


**Customer yang pernah mendaftar deposito sebelumnya cenderung untuk mendaftar lagi**

**Poutcome** didefinisikan sebagai hasil campaign sebelumnya dari customer ini. Dari EDA, kita dapat menyimpulkan:

- Customer dengan hasil campaign sebelumnya “sukses” cenderung untuk subscribe ke deposito
- Kami menyarankan tim marketing untuk menghubungi customer dengan hasil “sukses” pada campaign sebelumnya terlebih dahulu untuk diberikan campaign saat ini.





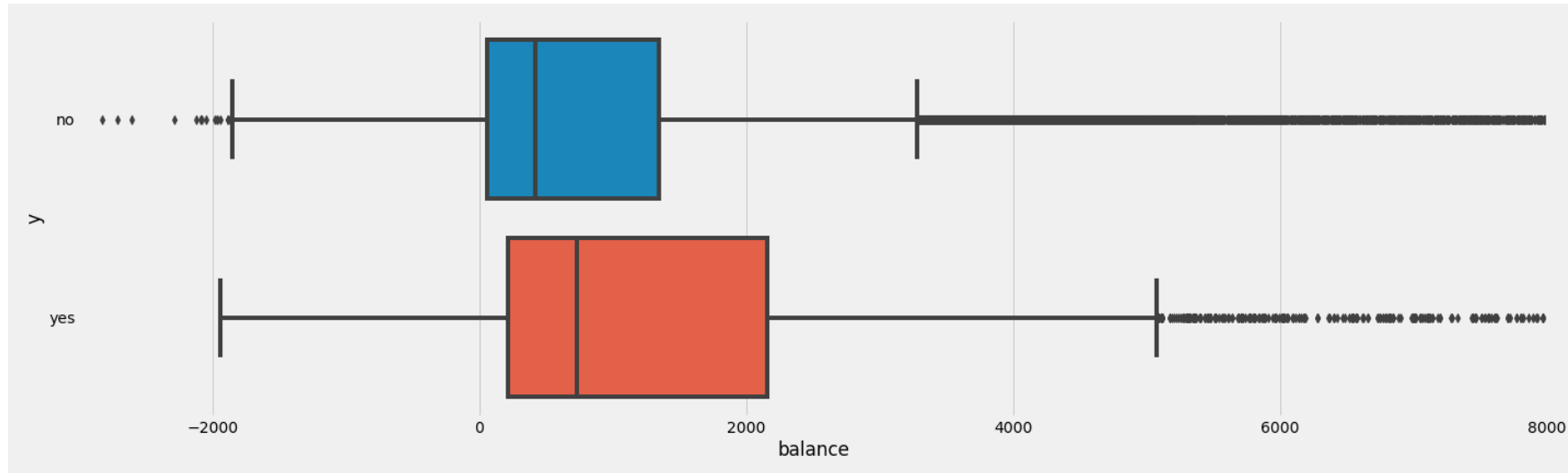
**Housing** didefinisikan sebagai apakah customer memiliki utang KPR atau tidak. Dari EDA, kita dapat menyimpulkan:

- Customer yang tidak memiliki utang KPR cenderung untuk subscribe ke deposito.
- Berdasarkan dataset, tim marketing lebih banyak menargetkan customer yang memiliki utang KPR untuk diberikan campaign deposito. Padahal, customer yang tidak memiliki utang KPR memiliki conversion rate **2 kali lipat lebih tinggi** dibandingkan customer yang memiliki utang KPR.

**Sebaiknya kita menargetkan customer yang tidak memiliki utang KPR**



## Business Insight Balance



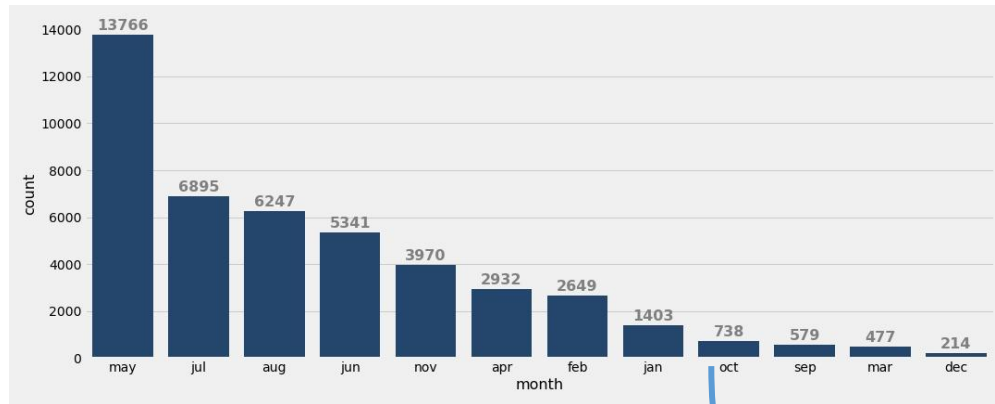
**Balance** didefinisikan sebagai rata-rata saldo bank yang dimiliki seorang customer dalam setahun. Dari EDA, kita dapat menyimpulkan:

- Customer yang memiliki balance lebih tinggi cenderung untuk mendaftar deposito.
- Hal ini masuk akal karena umumnya mereka yang berinvestasi harusnya memiliki sisa uang yang cukup dan tidak memiliki banyak utang pula.

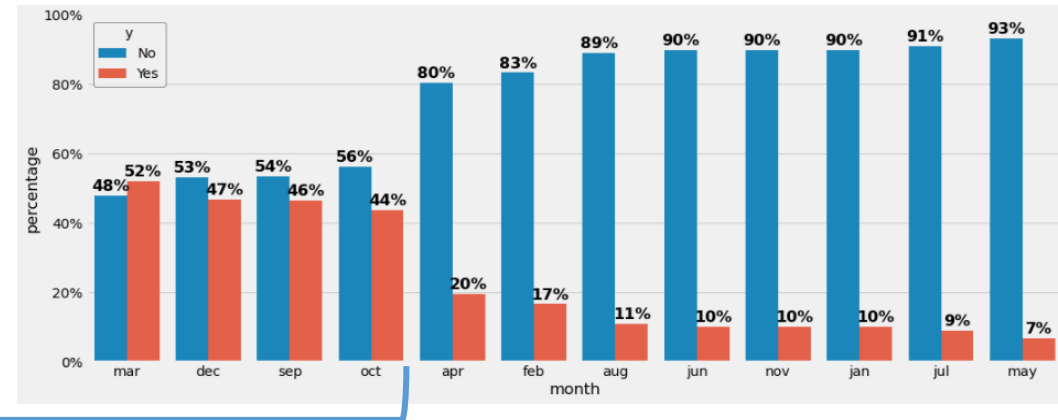


EDA

Jumlah customer dihubungi



Conversion rate customer



- Rata-rata conversion rate customer yang dihubungi pada bulan Maret, September, Oktober dan Desember adalah 47.25%. Sedangkan pada bulan lainnya hanya 11.75%. **Berbeda 3 kali lipat!!**
- Kami menyimpulkan bahwa tim marketing telah mentarget customer yang tepat pada keempat bulan ini. Sayangnya, jumlah customer yang dihubungi justru paling sedikit di antara bulan lain.
- Hal ini menunjukkan pentingnya memilih target yang tepat. Jika banyak customer yang dihubungi tidak potensial untuk diberikan campaign, conversion rate akan sangat rendah. Misalnya seperti bulan Mei, jumlah customer yang dihubungi sangat banyak akan tetapi conversion ratenya sangat rendah.

**Pemilihan target yang tepat akan meningkatkan conversion rate**





Raw Data

Handling  
Outliers

Z-Score

Scaling

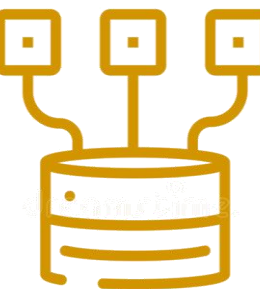
- Normalization
- Standardization

Encoding

- Label Encoding
- One-Hot Encoding

Handling Class  
Imbalance

- Dengan rasio 0.5
- Oversampling
  - SMOTE
  - Undersampling



Data Modeling

Split Data into  
Train and Set

Train : Test =  
80 : 20



Data Preparation



## Classification Algorithms

- Logistic Regression
- kNN
- Decision Tree
- SVM
- Random Forest
- Gradient Boosting
- AdaBoost
- XGBoost
- CatBoost

## Evaluation Metrics

- **Precision**  
Fokus untuk mengurangi **kesalahan dalam memprediksi customer** (*False Positive*) pada saat modeling *machine learning*
- **ROC\_AUC**  
Common use metrics untuk model evaluation data imbalance

<b>Actual No</b>	True Negative	False Positive
<b>Actual Yes</b>	False Negative	True Positive
	<b>Predicted No</b>	<b>Predicted Yes</b>



# Model Evaluation

Handling Imbalance	SMOTE				Over Sampling				Under Sampling			
Model	Train		Test		Train		Test		Train		Test	
	Precision	AUC	Precision	AUC	Precision	AUC	Precision	AUC	Precision	AUC	Precision	AUC
Logistic Regression	0.868	0.955	0.868	0.957	0.784	0.909	0.790	0.911	0.785	0.909	0.802	0.918
KNN	0.856	0.983	0.808	0.949	0.841	0.982	0.771	0.940	0.847	0.920	0.752	0.791
Decision Tree	1.000	1.000	0.829	0.874	1.000	1.000	0.873	0.958	1.000	1.000	0.719	0.784
SVM	0.857	0.964	0.848	0.961	0.811	0.936	0.802	0.928	0.816	0.942	0.795	0.932
Random Forest	1.000	1.000	0.898	0.985	1.000	1.000	0.926	0.998	1.000	1.000	0.798	0.931
Gradient Boosting	0.864	0.967	0.862	0.965	0.793	0.933	0.798	0.932	0.811	0.936	0.801	0.931
AdaBoost	0.854	0.951	0.855	0.9	0.786	0.912	0.787	0.913	0.790	0.913	0.812	0.920
XGBoost	0.959	0.995	0.912	0.985	0.900	0.986	0.857	0.969	0.941	0.993	0.796	0.933
CatBoost	0.962	0.993	0.920	0.986	0.885	0.981	0.853	0.967	0.890	0.979	0.808	0.940

Dari evaluasi di atas, dipilih Random Forest karena memiliki peforma terbaik

# Hyperparameter Tuning (Random Forest)

Oversampling			
Train		Test	
Precision	AUC	Precision	AUC
1.000	1.000	0.926	0.998

Hyperparameter Tuning

Feature Selection

1.000	1.000	0.953	0.998
-------	-------	-------	-------

Improvement precision 2,7%

Confusion Matrix (Before)

Actual	Negative	7198	302
	Positive	45	3706
		Negative	Positive
		Predicted	

Confusion Matrix (After)

Actual	Negative	7321	179
	Positive	56	3695
		Negative	Positive
		Predicted	



# Hyperparameter Tuning & Feature Selection

## Best Parameter:

- `n_estimators : 1800`
- `min_samples_split : 2`
- `min_samples_leaf : 1`
- `max_features : auto`
- `max_depth : 600`
- `bootstrap : False`

## Feature Selection

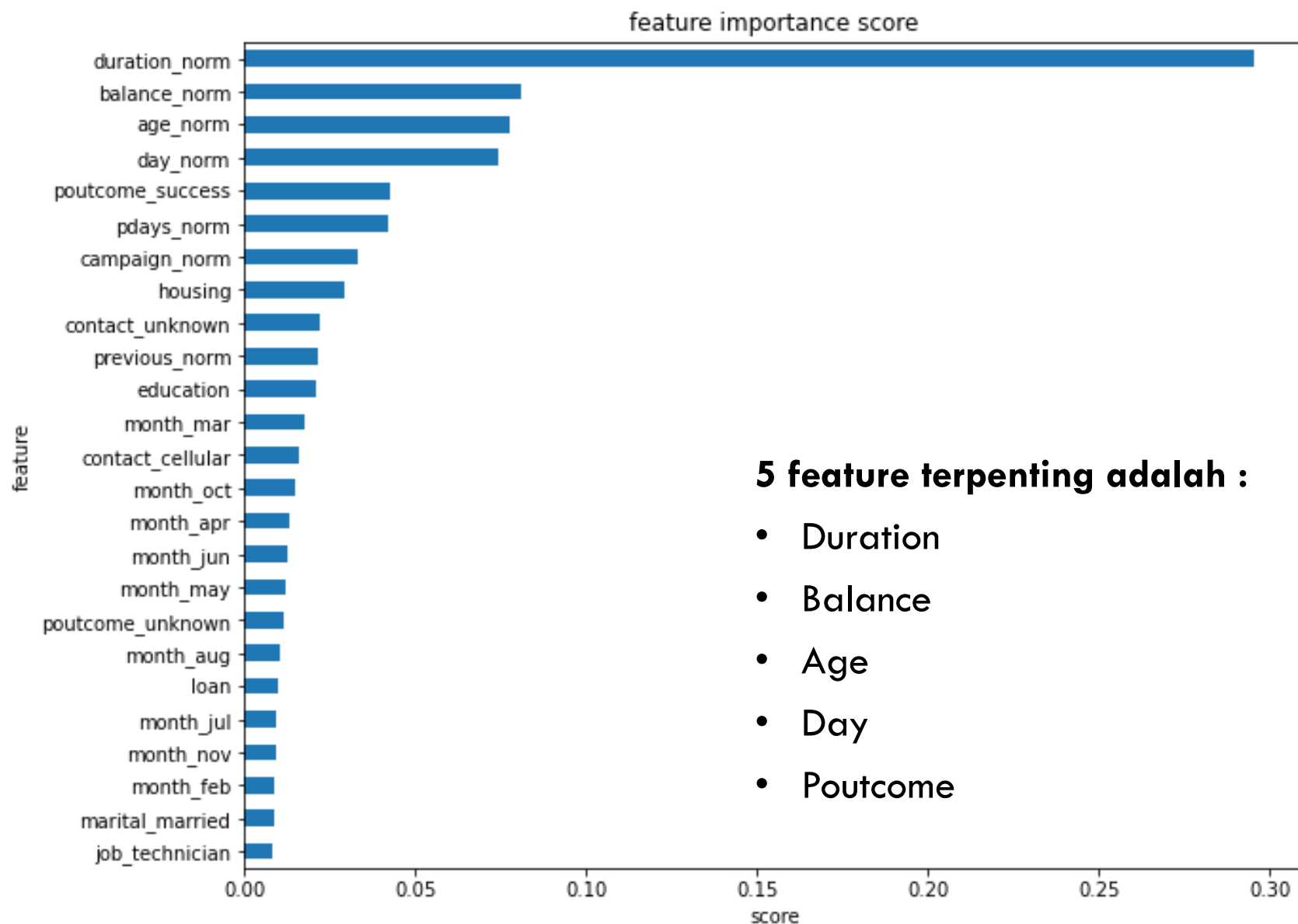
```
Correlation contact_cellular & contact_unknown: -0.874  
Correlation pdays_norm & poutcome_unknown: 0.986
```

Kita akan menghapus fitur-fitur yang redundant dan berpotensi mengakibatkan multikolinieritas.

Kolom yang dihapus: `contact_cellular` & `poutcome_unknown`



# Feature Importance



## Business Calculation

## CvR and Marketing Cost

	Before Using Model	After Using Model
<b>Sample</b>	45211	45211
<b>Target customer</b>	45211	15547
<b>Converted customer</b>	5289	14848
<b>Conversion rate</b> (converted customer / target customer x 100%)	11,7 %	95,3 %
<b>Marketing cost</b> (converted customer x telemarketing cost x duration x frequency)	€ 90.422	€ 77.836

- Terjadi kenaikan conversion rate sebesar **83.6%**
- Terjadi penurunan marketing cost sebesar **€ 12.586** atau sebesar **13.92%**

[\\*https://blog.useproof.com/calculate-conversion-rate](https://blog.useproof.com/calculate-conversion-rate)



Business Insights  
and Recommendations

## Business Calculation

## CPA and Return on Investment (ROI)

Assumption :

Net interest income = 1,5%

Average deposits amount = € 630 / Rp. 10 mio

	Before Using Model	After Using Model
<b>Marketing cost</b>	€ 124.782	€ 77.836
<b>Cost per Acquisition</b> (marketing cost / converted customer)	€ 23,59	€ 5,24
<b>Potential Revenue</b> (net interest income x average deposits amount x converted user)	€ 49.981	€ 140.313
<b>Return on Investment (ROI)</b> ( (potential revenue – marketing cost) / marketing cost x 100%)	-44,72%	80,27%

- Terjadi penurunan cost per acquisition sebesar € 18,35 atau sebesar 77.78%
- Terjadi kenaikan ROI menjadi 80,27%

[bi.go.id/suku-bunga-lending-deposits](https://bi.go.id/suku-bunga-lending-deposits)

[mountain.com/cost-per-acquisition-calculation](https://mountain.com/cost-per-acquisition-calculation)

<https://wartadana.com/dari-mana-bank-dapat-untung>

<https://www.investopedia.com/terms/r/returnoninvestment.asp>



Business Insights  
and Recommendations



## Executive Summary

**5** menit  
awal  
call

adalah durasi  
penentuan untuk  
membuat customer  
tertarik dan subscribe

**3** Total Call  
per  
customer

adalah jumlah optimum  
yang menentukan  
ketertarikan customer  
kepada deposito

**Customer dengan  
Campaign Sukses**

pada periode sebelumnya  
cenderung akan sukses juga  
jika diberikan campaign di  
masa mendatang



Business Insights  
and Recommendations

	before modelling	after modelling	
<b>Conversion Rate</b>	11,7%	95,3%	↑ 83,6%
<b>Marketing Cost</b>	€ 124.782	€ 77.836	↓ 13,9%
<b>Cost per Acquisition</b>	€ 23,59	€ 5,24	↓ 77,7%
<b>Return on Investment</b>	-44,72%	80,27%	↑ 124,9%

# THANK YOU !



The End

Business Insights  
and Recommendations

Modeling and Evaluation

Data Preparation

EDA

Background

Introduction