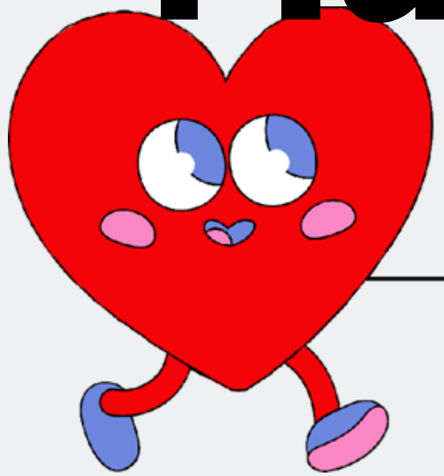
Welcome

Rakamin - Data Science - Batch 22

Final Project:

Marketing Campaign



Bulletproof Scouts (Group 5)

Welcome

Hello!

Team Members

Group 5

idol!

Adam

Mentor

Radifan

Leader / EDA

Afrian

Data
Pre-Processing

Naufal

Modelling &
Evaluation

Humairah

Background &
Model Impact

April

Business Insight

Hesti

Modelling &
Evaluation

Welcome Hello!

Table of Contents

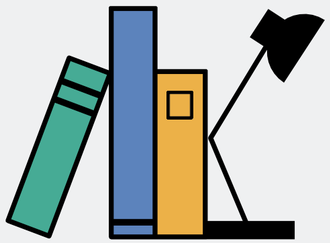


Table Of Contents

Background

EDA (Exploratory Data Analysis)

Data Pre-Processing

Modelling and Evaluation

Business Recommendation

Welcome | Hello! | Background

External Resources

Background

- Rata-rata perusahaan menghabiskan 4 - 11,2% dari *revenue* untuk marketing.
- Marketing yang tidak efektif dapat menyebabkan kerugian bagi perusahaan.
- PT. BTS hanya memiliki 15% *acceptance rate*.

(Source: Hook Agency; Brafton/Gardner CMO)

Background

Goal

Meningkatkan *revenue* perusahaan melalui peningkatan jumlah respon pada campaign.

Objective

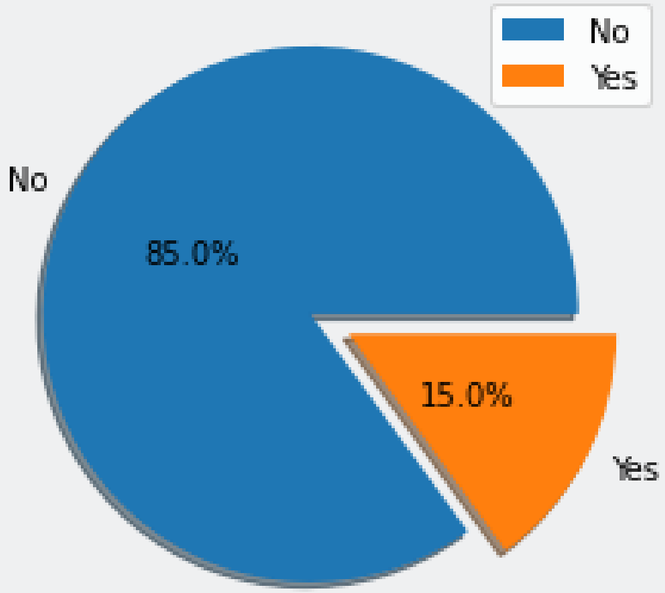
- Memprediksi respon customer terhadap *campaign* selanjutnya.
- Mencari faktor yang berkaitan dengan kesuksesan *campaign*.

Business Metrics

ROI

Exploratory Data Analysis

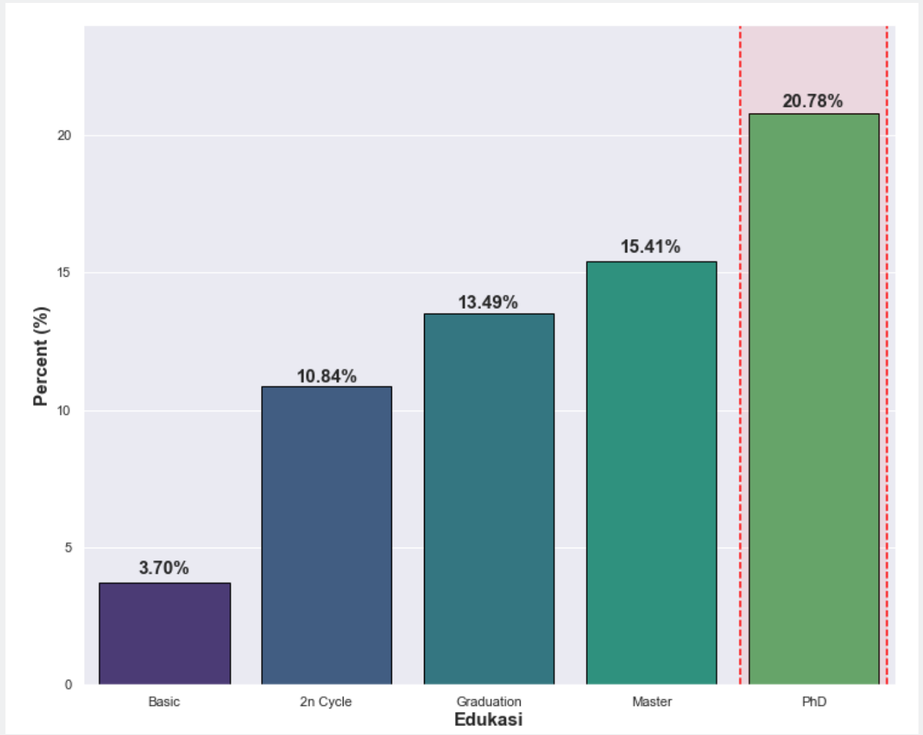
EDA (Exploratory Data Analysis)

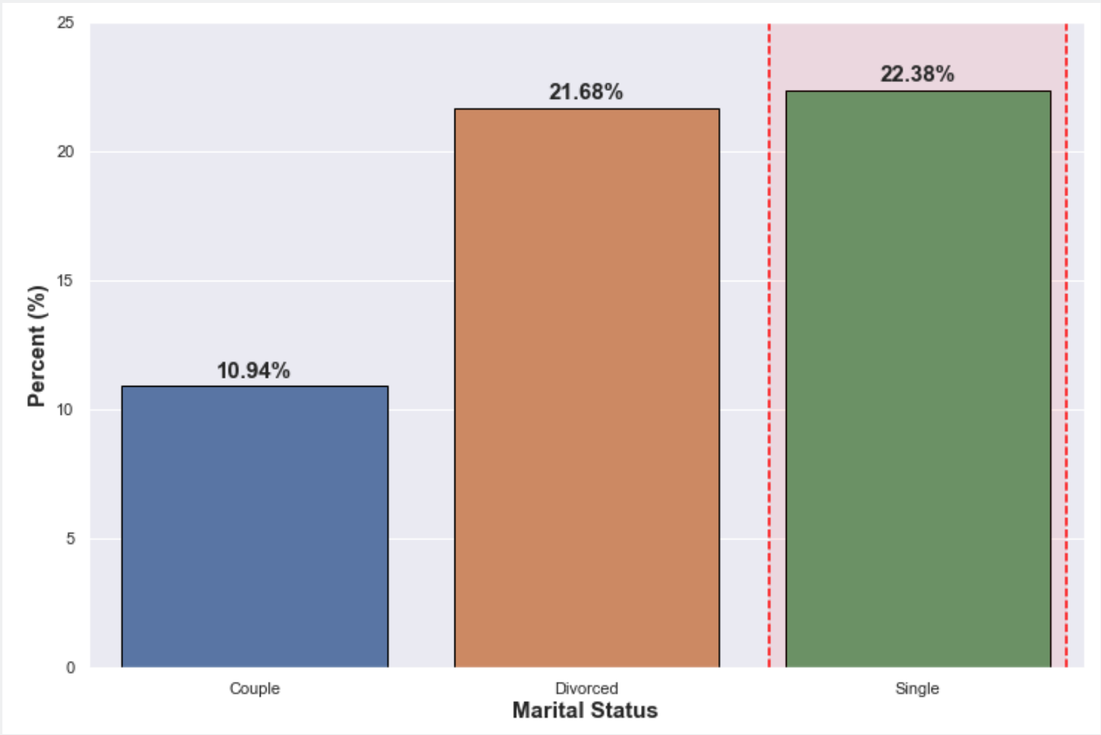


Data Understanding

- Sebaran data target 'Response', 85% yang 'No' dan hanya 15% untuk 'Yes'.
- Akan dilakukan *handling class imbalance* ketika *split data train* dan *test*.

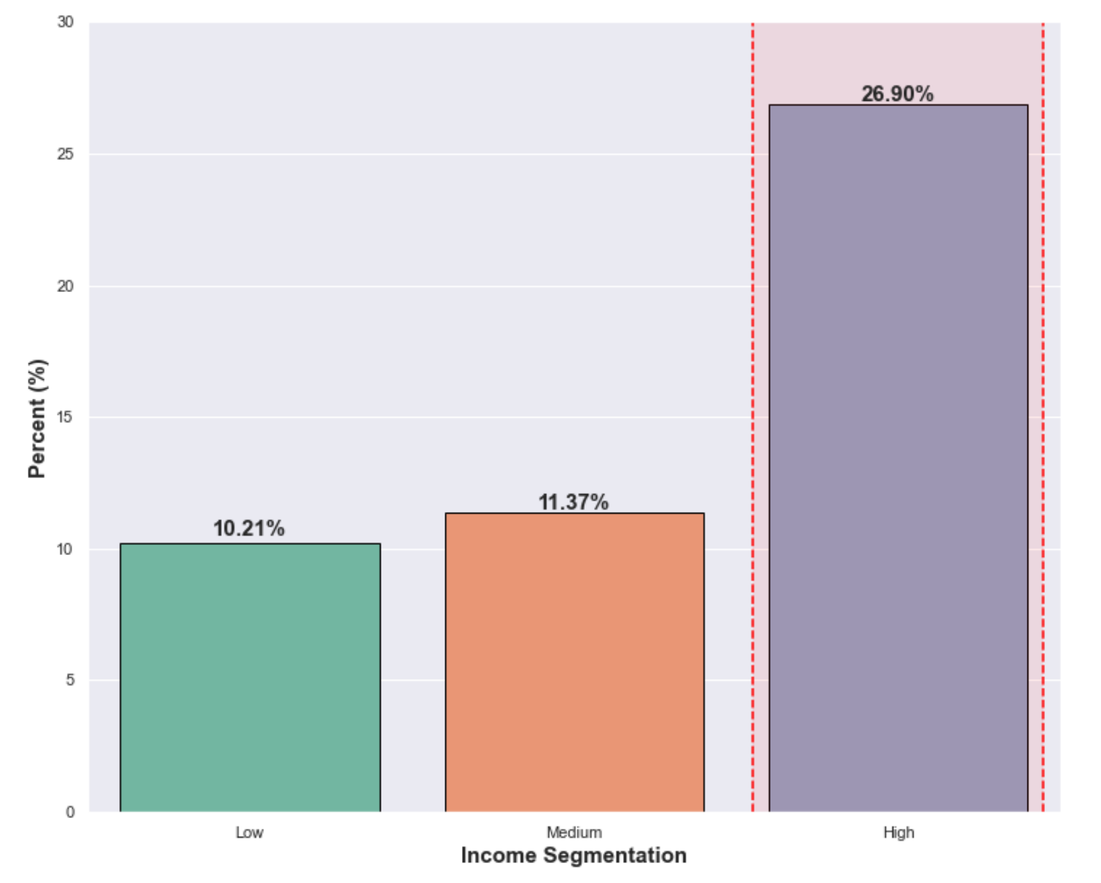
EDA (Exploratory Data Analysis)





Rasio customer yang merespon terhadap campaign berdasarkan kategori `**Education**` dan `**Marital_Status**`

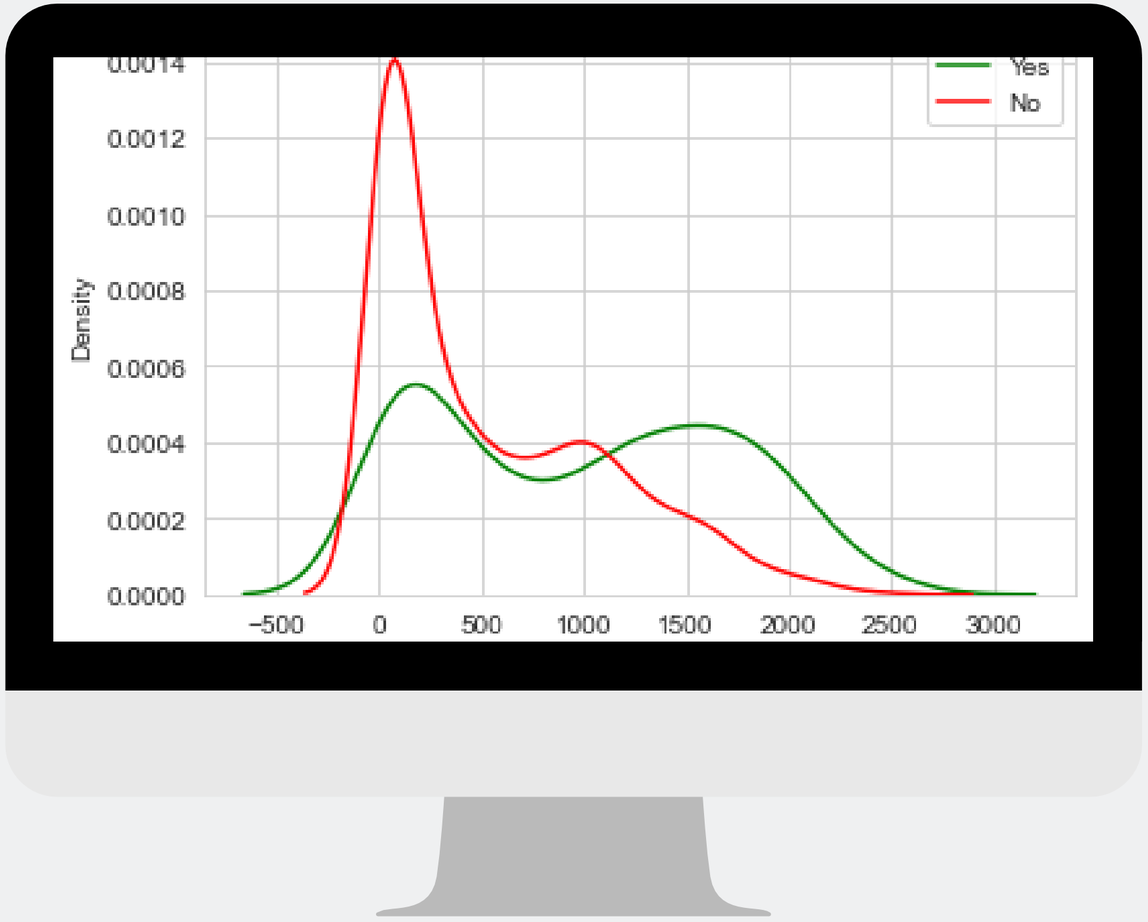
EDA (Exploratory Data Analysis)



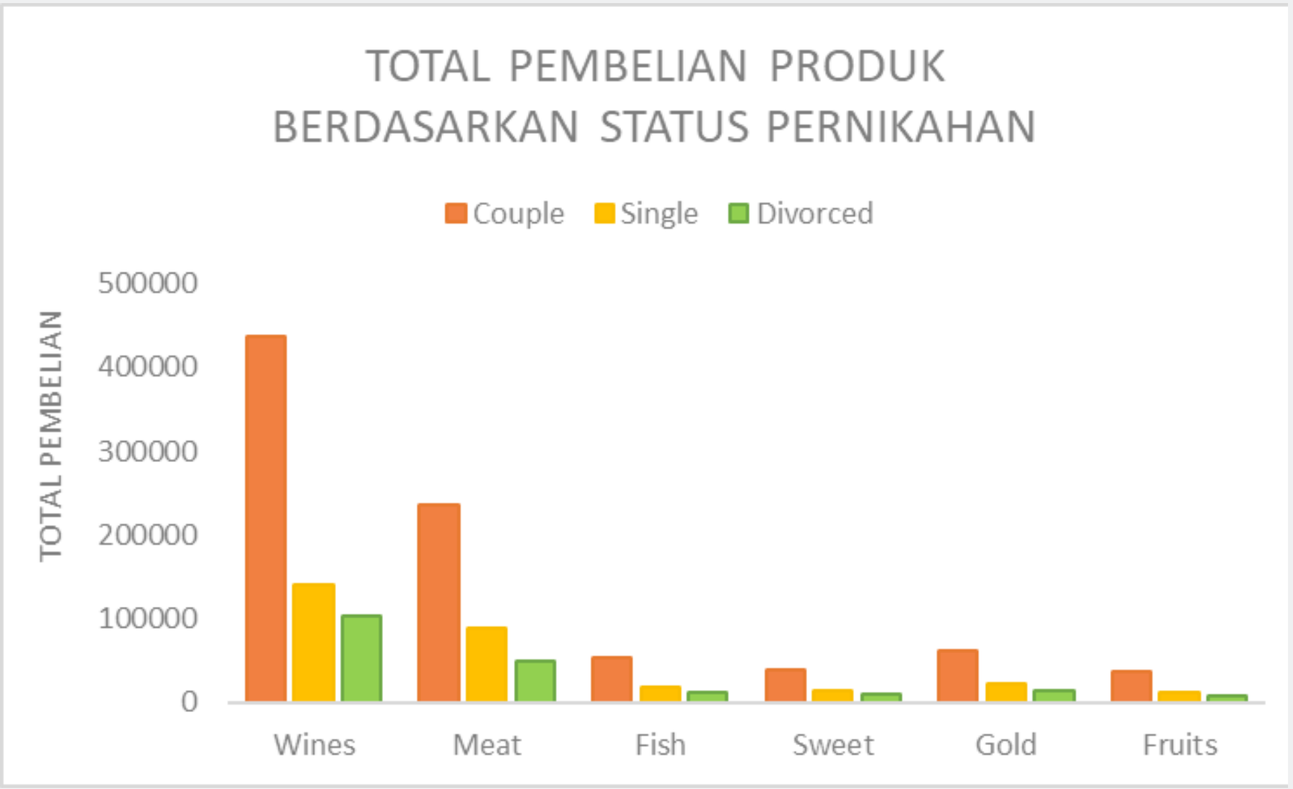
Pembagian segmentasi income berdasarkan Q3, apabila lebih dari Q3 maka dia 'High', dan kurang dari Q1 maka dia 'Low'

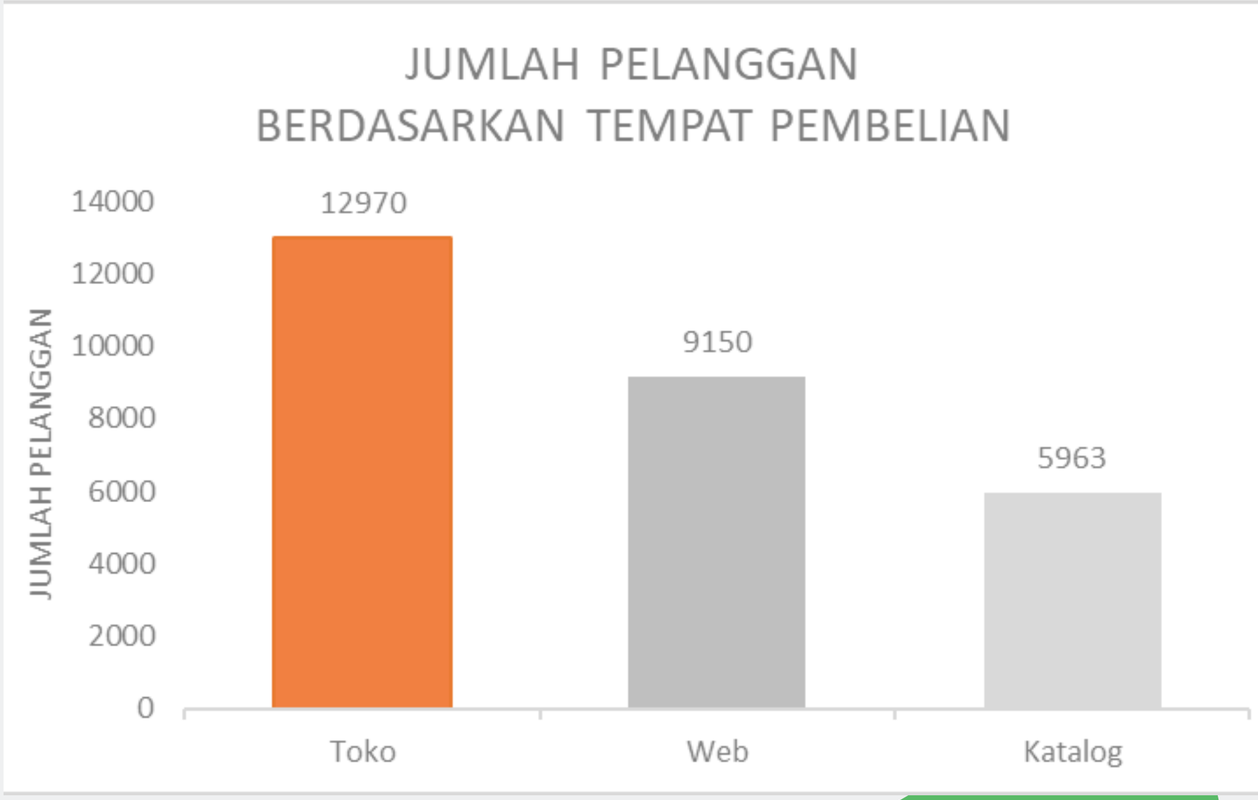
EDA (Exploratory Data Analysis)

Total spent yang *customer* keluarkan untuk pembelian semua produk.



EDA (Exploratory Data Analysis)





Data Pre-Processing

Data Pre-Processing

Variable	Data Type
Response	int64
ID	int64
Year_Birth	int64
Education	Object
Marital_Status	Object
Income	Float64
Kidhome	int64
Teenhome	int64
Dt_Customer	Object
Recency	int64
MntWines	int64

Variable	Data Type
MntFruits	int64
MntMeatProducts	int64
MntFishProducts	int64
MntSweetProducts	int64
MntGoldProducts	int64
NumDealPurchases	int64
NumWebPurchases	int64
NumCatalogPurchases	int64
NumStorePurchases	int64
NumWebVisitsMonth	int64
AcceptedCmp1	int64

Variable	Data Type
AcceptedCmp2	int64
AcceptedCmp3	int64
AcceptedCmp4	int64
AcceptedCmp5	int64
Complain	int64
Z_CostContact	int64
Z_Revenue	int64

Total Columns : 29

Total Rows : 2240

Data Pre-Processing

HANDLING MISSING VALUE	Drop 26 Rows
HANDLING DUPLICATED DATA	0 Duplicated Data
FEATURE ENCODING	a. Label Encoding : [Education], [Has Child], [Segmentasi] b. One Hot Encoding : [Marital Status]
TRAIN - TEST SPLIT	Data train : 70% (1551 rows) Data test : 30% (665 rows)
FEATURE SELECTION	a. Drop Irrelevant Features b. Feature Importances
HANDLING IMBALANCE TARGET	SMOTE 50%

Modelling and Evaluation

Feature

- Education
- Marital_Status
- Income
- Recency
- MntWines
- MntFruits
- MntMeatProducts
- MntFishProducts
- MntGoldProds
- NumDealsPurchases
- NumCatalogPurchases
- NumStorePurchases
- NumWebVisitsMonth
- AcceptedCmp1
- AcceptedCmp2
- AcceptedCmp3
- AcceptedCmp4
- AcceptedCmp5
- Complain
- month_Customer
- Spending
- Total_Cmp
- Has_child
- Segmentasi
- MS_couple
- MS_Divorced
- MS_single

Model

- KNN
- Logistic
- Regression
- Decision Tree
- Random Forest
- AdaBoost
- XGBoost

Target

Response

Modeling

70 % **Train Data**

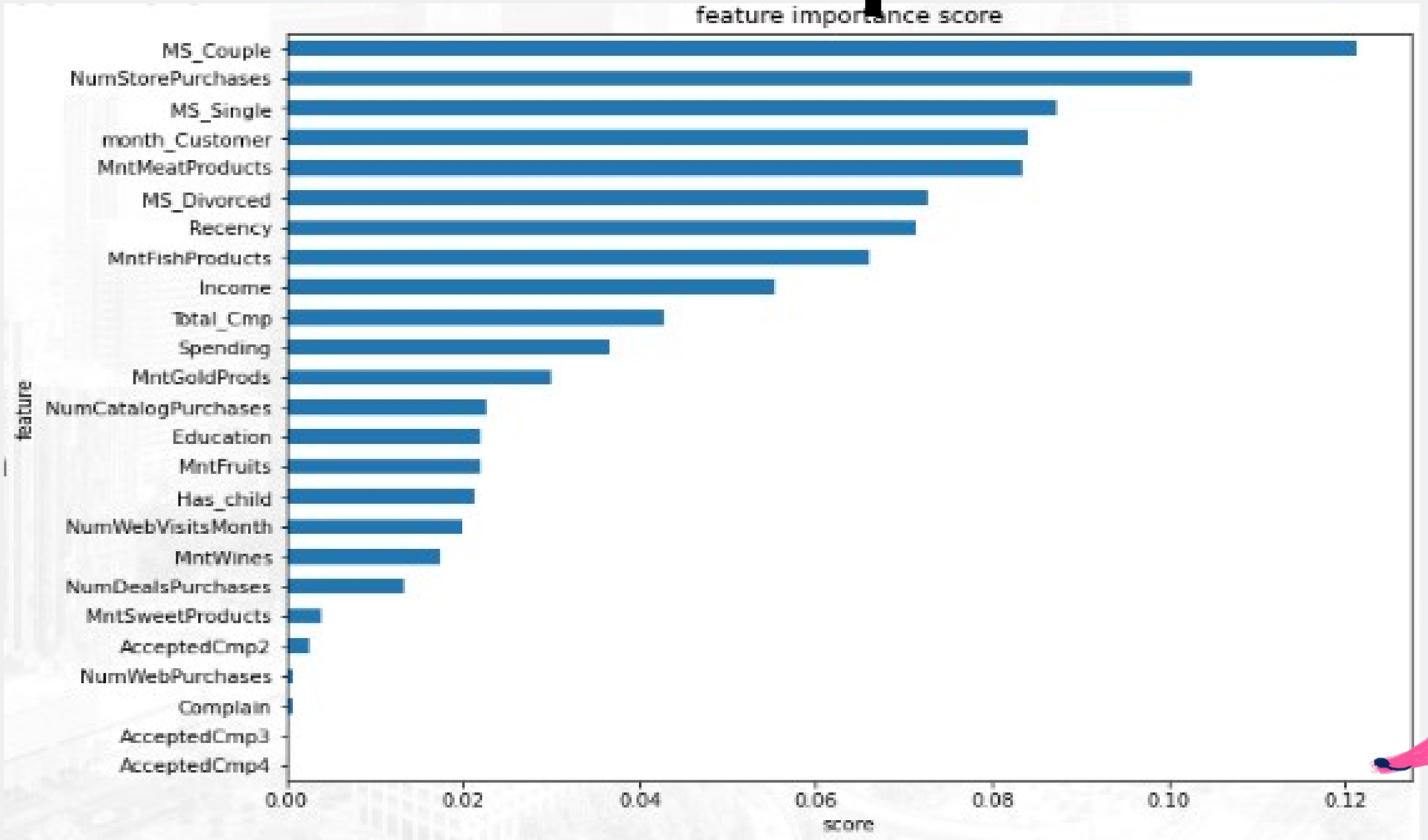
30 % **Test Data**

Hyperparameter Tuning

Best n_estimators: 1500
Best learning_rate: 0.05831
Best algorithm: SAMME.R
Scoring : AUC

JENIS MODEL	Accuracy	Precision	Recall	F1-Score	AUC
KNN	84%	43%	42%	43%	76%
LOGISTIC REGRESSION	83%	40%	36%	38%	77%
DECISION TREE	82%	42%	60%	50%	81%
RANDOM FOREST	90%	68%	59%	63%	92%
ADABOOST	91%	69%	65%	67%	94%
XGBOOST	88%	59%	55%	57%	92%

Feature Importance



Top 5 Features:

- MS_Couple
- NumStorePurchases
- MS_Single
- month_Customer
- MntMeatProducts

Feature Selection

Feature NumDealsPurchases, MntSweetProducts, AcceptedCmp2, NumWebPurchases, Complain, AcceptedCmp3 & AcceptedCmp4 di drop, kemudian iterasi model dengan best parameter model.

Metrics	Tuned AdaBoost (%)
Accuracy	90%
Precision	65%
Recall	65%
F1-Score	65%
AUC	94%

Business Insight & Reccomendations

Bussiness Insight and Recommendation

Model Application Flow

Before Modelling

Data pelanggan PT BTS → Perencanaan pembuatan *campaign* → Pemberian *campaign* ke seluruh pelangan

After Modelling

Data pelanggan PT BTS → Prediksi pelanggan yang akan merespon *campaign* → Perencanaan pembuatan *campaign* → Pemberian *campaign* ke pelangan yang berpotensi merespon

Machine Learning

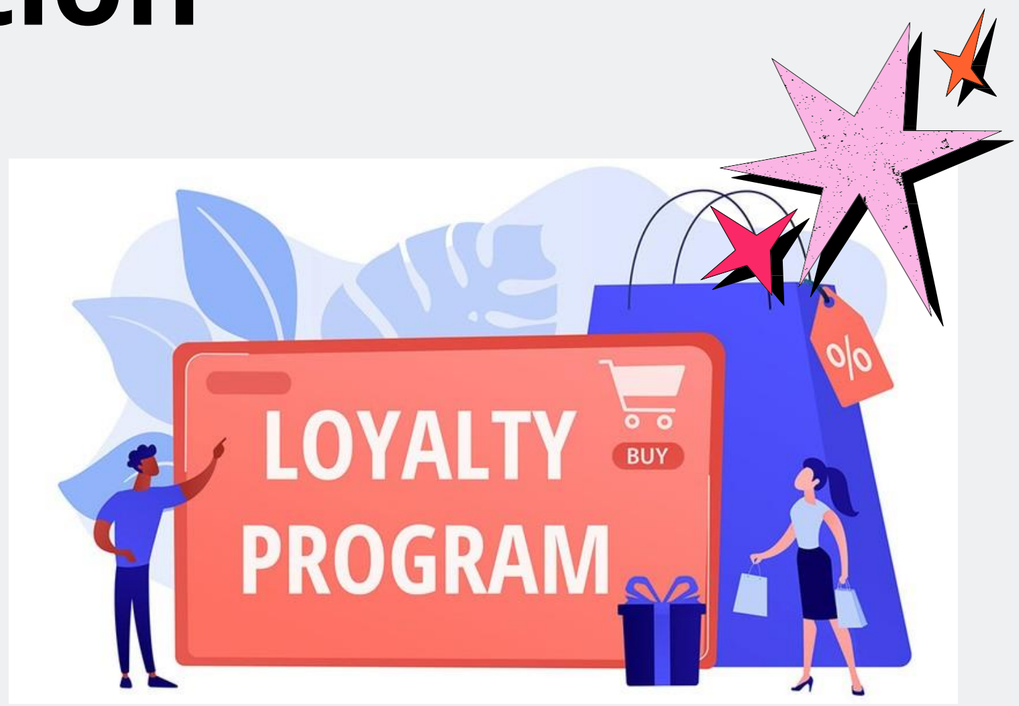
Business Recommendation

Memberikan diskon buah

Memberikan *flyer* berisi penawaran diskon *special wine* kepada pelanggan yang datang ke *store*.

Business Recommendation

Penawaran paket bundling daging dengan produk lain



Pembuatan program 'loyalty reward' berdasarkan lama keanggotaan dan total belanja

*Kurs USD to IDR per 5/8/2022 = 14,948.10

Marketing Cost

Total Pendapatan secara Keseluruhan — $1.356.988,00
Total penjualan produk

Rata-rata Pendapatan per *Campaign* — $226.164,67
Total pendapatan / Jumlah campaign (=6)

Rata-rata Pendapatan per *Customer* — $100,97
Total pendapatan / Jumlah *customer* (=2240)

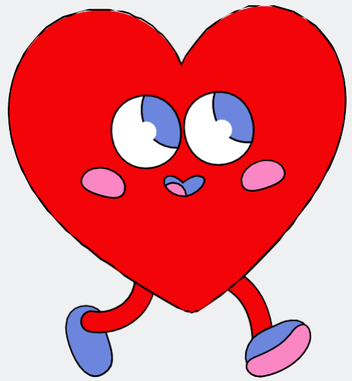
Marketing Cost* per *Customer — **$10,09**
Asumsi 10% dari Rata-rata Pendapatan per *Customer*

Marketing Cost* per *Campaign — **$22.601,60**

*Kurs USD to IDR per 5/8/2022 = 14,948.10

Model Impact

per Campaign

	Before 15% *Acceptance Rate*		After 65% *Acceptance Rate*	
ROI ((Revenue - Marketing Cost) / Marketing Cost) * 100	**50%**	→	**550%**	↑ **1000,0%**
Revenue Pendapatan per Customer * Total Customer	**$ 33,6K**	→	**$ 145,7K**	↑ **333,6%**
Net Profit Pendapatan Campaign - Cost Campaign	**$ 11,2K**	→	**$ 123,4K**	↑ **1001,8%**

Thank You!