

# Customer Revenue Uplift

No. \_\_\_\_\_

Date: \_\_\_\_\_

Import libraries like alat/pertakal.

class Project Config buku resep. (takaran nya berapa).  
instansi

Data Path

Random state → nomor urutan. → elo diacak

penggunaan harian/  
bulan

← Time Windows

Pola jangka

Outlier → data menyimpang

Pendek / Jangka  
Panjang

lain dr yg lain

seperti misalnya.

← Random Forest

Protected Columns

output yg sama  
setiap dimana.

↓

Hasil & evaluasi  
Model ga benar

nilai FCM  
di rata

← Network FP I

Nest = jumlah temen yg diajak  
diskusi (biar adil)

Max D = bgt kasi temen boleh  
nanya (biar ga muter x)

rf min = jumlah org dalam 1 tel  
sebelum dipercabig

(biar ga kecil)

leaf = jml yg ts sejauh

ditepungin akhir  
(biar ga dr 1 org)

biaya wh/sms

← Cost / cust : 5000

perlangsiran

Hour per month ← Multiplier : 1.2

naik setelah promo

ketika yg tinggi → High value Thold : 100 k.

→ 100k dianggap penting &  
dipindin rusak.

fire Plot. ← <sup>verakan teman</sup>  
<sup>seberapa banyak hasil</sup>  
Pilihan warna

protected column

↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓

daftar nama kolom i

yg dilindungi saat proses  
outlier handling.

↓ file path Many2x CSV → kode lain gampang  
ambil lokasi

ga diubah /

diberikan tipe

data pun hrs ini

terp acara.



bahan <sup>xx</sup> di cuci, potong <sup>xx</sup> No.  
belum dimasak chef  
<sup>↑</sup>  
(model)

Date:

## class data loader.

Int → lalu untuk nyimpen

load data set CSV (error handle)

data, bahan ada & bagus → ambil bahan & chef nya.

else ~~ada~~, ~~ada~~ → ~~ambil~~ ~~chef~~ → ~~x~~

↓  
load all  
load data set → Call → ambil semua data/  
bahan <sup>xx</sup> dr kelas,  
cek satu x success/gagal

Quality check → Cek tangggal kadaluwarsa bahan,  
uang yg dibutuhkan  
list yg hilang  
list yg salah

Missing value ↙ num = median  
cat = unknown

dates kolom → drop yg salah

Outlier : t - score

summary

dataset. data loader. load all data()

Untuk mensortir semua bahan diatas

list semua data set. ↗  
having a  
feature ada  
di data set.



(1)

## Customer profile (5)

ID

Age → 16 - 69 (Avg : 47.5)

Gender → M: 15811, F: 1437

City → Bangalore (10.2%)

MNR

Plan type → OM6 30 6B (25.1%)

APPU → 5.468 - 203.934 → High

Debt → And. mid (34%) 3.912

SAC ID

(2)

## Campaign history (6)

ID → 21801

Campaign → Topup cashback

Offer date

Treatment AB → 0 12k. / 1: 22k

✓ Conversion

✓ Uplift → Total ₹139.1t.

(3)

## Usage (7)

ID

date

data → Avg. 4.7 GB / month

voice. → Avg : 53 min / month

SMS → Avg : 5. SMS / month

APP social → 3.8 Jam

All gaming → 0.9 Jam



for men + Heavy data  
 Gaming ✓ kdom.  
 Social media.

low usage → Risk. churn

(4) Network (15)

site ID

date

Avail

PL

TMR

VL

VSUPR

KSPP

KSPQ

PRB

Max user

Active user

CSSR

Excellent = ~~5~~ 3.08 SD

Poor = ~~10~~ 7.6.9 SD → different

Campaign

MSDS

feature engineering → menyiapkan & merakit dataset supaya  
dapat diproses oleh model RF

(1) Usage + Mean, Max, Std, Sum

Nombat filter baru pada usage

Usage - features

(2) Rolling Usage 7 & 30

filter remarcian terakhir dalam 7 & 30 hari

(3) Data = Medium / Heavy / Low / Poor

Visual = Regular / Minimal / Active / Addict

Gaming = Casual / Non-gamer / Regular

segmentasi pelaku Hard core

Pelanggan

(4) Network Quality Score

be peral pelanggan  
menyertai  
berapa batas area yg  
merent keingin di loek' itam

combine all KPI → ~~all~~

(5) Campaign + Count, Sum, Mean,  
Max, Last.

pelanggan dalam mengikuti kampanye

Complain + Count, Sum, Mean,

max, last.

(2)

## Master feature

1 customer profile

2 usage

3 network

4 campaign

5 complaint

63

$\Sigma + = 77$

(3)

## Interaction feature

menambah pola gabungan yg lebih kompleks.

fill missing value

$\text{Cat} = \text{Unknown}$   
 $\text{Num} = 0$

Kes. Ahmad Rifai



# Modeling

X	Y
80%	with

No. \_\_\_\_\_

Date: \_\_\_\_\_

train data ~~not~~

menyajikan data train, ls cara menambahkan master feature & campaign history. karena kolom target conversion & uplift ada di data campaign

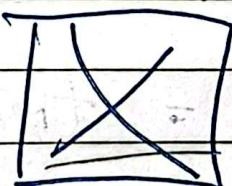
Campaign history + Master feature

$$5 + 77 = 82$$

Target = 82

Drop: cust\_id, campaign\_id, offsite  
conversion, uplift, master feature

↓  
Save as



menyajikan data  
feature untuk training

as exclude kolom identitas, date, target

menyajikan 2 kolom

Target conversion & uplift.

Classification

Migration

Encode Cat 9 kolom kategorikal di master feature  
akan di encode.

bagi data untuk train & test.

Split — conversion = stratify } Drop column  
                  uplift   = No stratify } column

Classification → conversion & feature importance  
Regression → uplift

besaranya peningkatan pendapatan

hanya train data yg res uplift lebih dari nol ( $>0$ )



- Learner.  $\rightarrow$  prediksi uplift lebih akurat dibanding baseline

Nembak data jadi dua  
one terpisah

Model RF sebelum yg a.

Group

Treatment = 1

$\downarrow$

Control = 0

karena fokus ke customer  
yg benar  $\Rightarrow$  benefit dr campaign

Hapus kolom identitas, date, target.

- feature drop off rd, conversion,

campaign, uplift.

offdate, treatment, group.



IMPLEMENTASI BAGIAN 1 Learner

IMPLEMENTASI BAGIAN 2

Save as

$\times$  treatment &

$\times$  Control

Learner

Target  $[Y\_conversion]$

$[Y\_control]$

Predict uplift = treatment, control.

- Maka can bl Master feature.

$$77 + 8 = 80$$

Uplift

theory (Probability)

control. (probability)

Customer 20536 jika diberi campaign probability conversion rate 42%

~~Top 10 customer by uplift ratio of given campaign~~

$$\text{Uplift} = 80.8 - 38.6 = 42.2\%$$

Simulation campaign 3 input penting : persentase target pelanggan, biaya per pelanggan, revenue multiplier

• Target

• Customer

• Expected Conversion

• Revenue

• Cost

• Net revenue

• ROI

• Avg Uplift

• Avg ARPU

Simulation  
Result.

Give max

strategy.

Vitalidari

1. Net revenue & ROI vs Target &

2. Expected Revenue, Cost, Net Revenue per  
Target + ROI

3. ~~Avg~~ Uplift per Customer vs Target &

# Graph ML

- Test 1000 cust by ARPU  
 - Encode

- feature cols.  
 - Scaler

- customer list  $\rightarrow$  Mr. Graph()

↳  
 connections

Avg strength

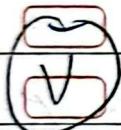
- centrality metrics

↳ Denary

Avg weight

Degree centrality

Influence score



ID

Influence score

Network influence

Degree centrality

Betwenees

Eigenvector

Page rank

Lp norm

ARPU

Connections

City

Plan Type

Age month



- **TOP 10 Influencer**

Campaign Status.  $\leftarrow$  Top 1%

Highest connected

Vizualization.

- • Influential features (same let Lebeum zu)
- • Plan Type Performance
- • Geographic Performance
- • Network & Influence Insight
- • Top 5 Influencers

- **SHAP**.

- • Features importance
- • Individual customer explanations

Top Increasing column

Top Decreasing column

- Top increasing column: Feature 1

- Same thing for decreasing column: Feature 2

Max Feature 1: Feature 1

Max Feature 2: Feature 2

Min Feature 1: Feature 1

Min Feature 2: Feature 2

Top Feature 1: Feature 1

Top Feature 2: Feature 2



No. :

Date :

	bayangan kamu kerja di bid. Marketing. Tiba bos kamu req reservasi. "Mane ya si Damelia ya", haa ini dia. Kita mau naikin revenue nih, tolong kasi PROMO $\% 10$ pelanggannya ingat sejam ya". "a.. iya iya bos. Siap siap" "Saya tunggu ya" "aku jadi ok". "aku di sini buat sejam, yg tener aja bos, emang bangun candi". "Ngomong apa tadi". "man kita coba".
-	Step 1: Maptan library yg kita butuhkan.
-	Step 2: Kumpulkan konfigurasi $\times$ di satu class yaitu project config. seperti folder tempat dataset, random state, test size, rolling penggunaan data harian & bulanan, threshold & score untuk outlier, konfigurasi Model RF, Threshold network KPI, biaya campaign per customer, charapain omset setelah campaian, threshold high value customer, dan daftar nama kolom yang dilindungi saat proses outlier handling.
-	Step 3: Data cleaning. - Mengisi missing value dengan Median untuk type data numerik & unknown untuk type data kategorikal. - Drop kolom tanggal yg tidak - Outlier handling. Now data kita sudah bersih.
-	Step 4: Exploratory Data Analysis - Memahami data customer profile, campaign history, Usage Metrics, & Network KPI

No. :

Date :

- |   |  |
|---|--|
| - | Step 5 : Feature Engineering.  |
| - | Membuat fitur baru pada usage metrics.   |
| - | fitur pemetaan terakhir dalam 7 dan 30 hari.   |
| - | Segmentasi penilaian pelanggan   |
| - | Membuat fitur network quality score untuk mengetahui seberapa baik Network KPI di lokasi utama maupun $\approx$ pelanggan. |
| - | Fitur beberapa sering & surles pelanggan dalam mengikuti campaign  |
| - | Fitur riwayat belanja pelanggan.   |
| - | Gabungkan semua fitur yang dibuat tadi ke dalam Master feature, jadi kolom totalnya ada 63.                                |
| - | Membuat interaction feature dengan mengalihkan fitur satu dg yg lain. Untuk menangkap pola gabungan yg lebih kompleks      |
| - | Interaction feature ada 14, ditambah Master feature 63 sehingga total master feature nya ada 77.                           |
| - | Jangan lupa isi missing value dengan 0 untuk type data numerik & untuk kategori, untuk type data kategorikal.              |
| - | Step 6: Modelling.   |
| - | Memperbaiki data training dg cara trimmings  |
| - | Menggabungkan master feature dg campaign history karena kolom target yaitu "conversion" & "uplift" ada di data campaign    |
| - | menyimpan kolom unit train dg exclude kolom $\approx$ identitas, date, & target simpan sbg X                               |
| - | Menyimpan 2 kolom target yaitu conversion dan uplift.  |
| - | Encode kolom kategorikal sebelum di modeling   |
| - | Membagi data untuk training & testing.   |
| - | Jalankan Model classification untuk memprediksi conversion yaitu apakah pelanggan mau beli atau tidak.                     |

No. :

Date :

- Step 9 : Network Analysis
  - Membuat algoritma untuk koneksi antar pelanggan berdasarkan kesamaan profile mereka untuk mencari pelanggan yg paling berpengaruh & bisa meniru event viral dari campaign.
  - Menghitung influence score setiap pelanggan dalam jangkauan menggunakan 4 metrik : degree centrality, betweenness centrality, eigenvector centrality & pagerank.
  - Menggabungkan 40% network influence, 40% uplift score, dan 20% ARPU untuk mengetahui viral effect.
  - Step 8 mendapatkan top 10 influential customer tertinggi.
    - Jd strateginya bisa menfokuskan di wilayah Yogyakarta + Jabodetabek.
    - Jg target customer yang memiliki  $> 0.33$  influence score.
- Step 8 : Analisis performa segmentasi
  - Halo postpaid adalah plan terbaik dg uplift 3.15% dan customer besar (966 customer)
  - Combo Sakti 5G 6B, uplift nya tinggi 3.25% tapi customer kecil (34)
  - Yogyakarta, tempat paling sukses tlp campaign dg uplift (9.15%) disusul oleh Bandung, Depok, Bekasi, Jakarta.
- Step 9 : SHAP Analysis
  - Shap menjelaskan secara detail bagaimana model memprediksi, jd tta bisa memahami kontribusi real setiap fitur, yg paling berpengaruh dalam keputusan model.

No. :

Date :

- Cek feature yg paling impact dari Model RF.
- Jalankan Model regresi untuk memprediksi uplift yaitu bedanya peningkatan pendapatan.
- 
- Membanding Model T-Learner untuk memprediksi uplift lebih akurat dibanding baseline model RF sebelumnya. Karena informasi customer yang benar<sup>xx</sup> konversi dari campaign. Dengan <sup>xx</sup> cara membagi data jadi dua grup terpisah.
- Implementasi bagian pertama T-Learner, tujuannya mempelajari pola konversi saat menerima treatment.
- Implementasi bagian kedua T-Learner, tujuannya mempelajari pola konversi jika tidak diberi campaign.
- Menghitung skor uplift yaitu delih probabilitas yg dapat campaign tidak dapat campaign.
- 
- Jadi kita dapat top 10 customer dg uplift score tertinggi.
- customer dg id 20536 ini diberi campaign probabilitas conversion nya naik 42%.
- 
- Tadi kita cet top 10 cust, sekarang coba top beberapa persen dr customer berdasarkan uplift score tertinggi.
- Didapatkan total menargetkan top 10 % customer yaitu 3000 cust akan mendapatkan Return Of Investment yg tertinggi. dg net revenue: sekitar 162 juta
- Kalau top 20% customer akan mendapatkan net revenue tertinggi tapi ROI nya turun.
- Jadi kamu mau pilih yang mana ROI tertinggi atau net revenue tertinggi.
-