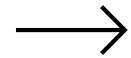




# Bank Marketing Campaign



MACHINE LEARNING MODEL  
PREDICTION

By : JCDS 0808 - 007 -Yanfa Anandika

# Business Understanding

- Sebuah perusahaan perbankan melakukan kampanye pemasaran untuk menawarkan deposito
- Untuk meningkatkan penjualan produk tersebut Bank melakukan kampanye pemasaran
- Dengan menghubungi calon nasabah, diharapkan nasabah yang ditelepon dapat membuka rekening deposito



# Goals

Maka tujuan dari kampanye pemasaran ini diantaranya :

- Meningkatkan penggunaan produk Bank yaitu deposito
- Menambah nasabah baru untuk membuka rekening deposito
- Meningkatkan pemasukan Bank

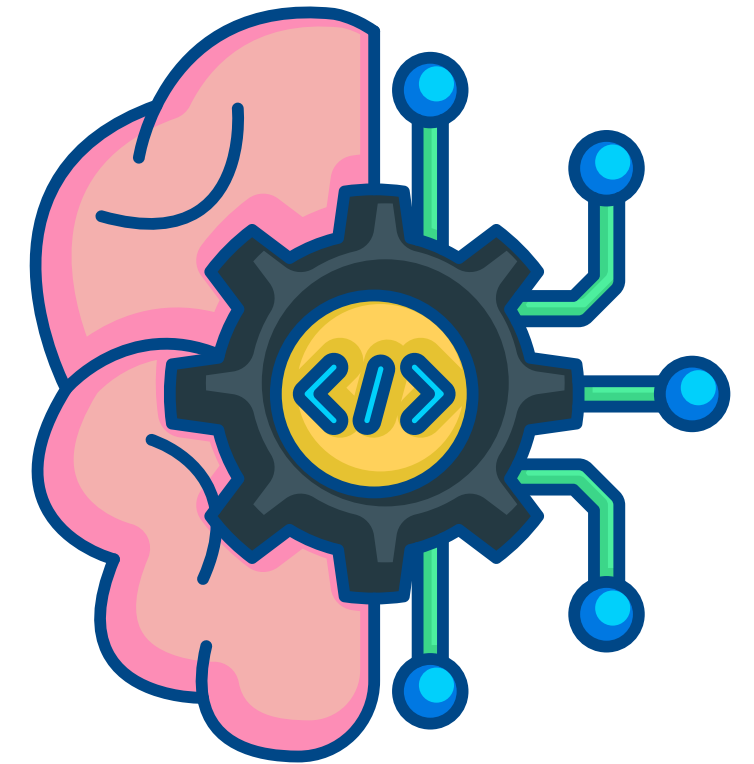
Namun perlu diperhatikan untuk cost dan benefitnya, agar lebih efisien dan tepat sasaran

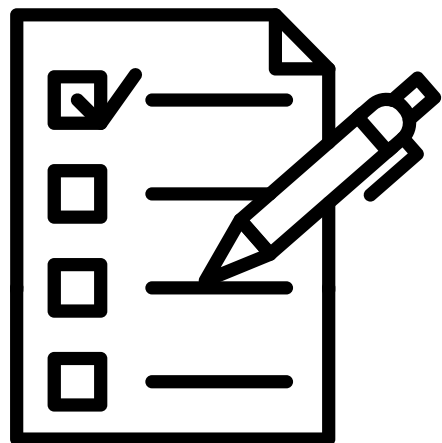
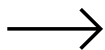


# Analytical Approach

Salah satu solusinya yaitu dengan menggunakan Machine Learning Model untuk memprediksi nasabah potensial, dengan tujuan :

- Campaign Program yang lebih cepat dan efisien
- Meminimalisir biaya operasional
- Mengurangi jumlah panggilan yang tidak tepat sasaran





# PREPARATION



# Cost Estimation

Asset Tier	Mean Marketing Budget	Median Marketing Budget	Est Marketing Budget as a % of Average Assets
Less than 150MM	\$35,000	\$35,000	0.05%
\$150MM-\$299MM •	\$153,594	\$175,000	0.07%
\$300MM-\$499MM	\$197,766	\$205,000	0.05%
\$500MM-\$699MM	\$262,289	\$250,000	0.04%
\$700MM-\$999MM	\$446,672	\$440,000	0.06%
\$1B-\$2.9B	\$927,879	\$800,000	0.05%
\$3B-\$9.9B	\$2,010,046	\$1,987,350	0.04%
\$10B+	\$5,400,000	\$4,000,000	0.05%

Budget Item	Average % of Budget (2024)	Average % of Budget (2023)	Average % of Budget (2022)
New Customer Acquisition	34%	27%	28%
Customer Retention	23%	21%	25%
Advertising	40%	38%	38%
Outside Agency Support	12%	17%	18%
Sponsorships	Not asked	Not asked	27%

Sumber : American Banking Association - Banking Journal

Jika Bank tersebut memiliki Total Asset kurang dari \$ 299M

- Marketing Budget : +- \$ 175.000 atau € 150.000
- Budget Akuisisi Customer Baru : +- € 50.000
- Akuisisi 10.000 Nasabah Baru = +- € 5 / Nasabah

# Cost Estimation

## Source Fund

- Data menunjukkan rata rata saldo nasabah : € 550
- Asumsi Nasabah mendepositokan 50% total saldo : **€ 275**

## Biaya

- Biaya deposito 3% / bulan  $\rightarrow 275\% \times 3\% = \text{€ } 8,25$  (**Biaya Deposito**) / Nasabah
- Biaya operasional / Nasabah  $\rightarrow \text{€ } 5$

## Source Profit

- Bank menyalurkan fund ke kredit dengan bunga 10%
- $\text{€ } 275 \times 10\% = \text{€ } 27,5$  (**Pendapatan Bank**) / Nasabah



# Metrics Evaluation

	Tidak Deposit (0)	Deposit (1)
Tidak Deposit (0)	<b>True Negative (TN)</b> Prediksi & Aktual Tidak Deposit	<b>False Positive (FP)</b> Prediksi Deposit Aktual Tidak Deposit
Deposit (1)	<b>False Negative (FN)</b> Prediksi Tidak Deposit Aktual Deposit	<b>True Positive (TP)</b> Prediksi & Aktual Deposit

Pendapatan Bank : **€ 27,5**

Biaya Deposito : **€ 8,25**

Biaya Operasional : **€ 5**

Perhitungan :

- TP :  $€ 27,5 - € 8,25 - € 5 = + € 14,25$  (net benefit)
- TN : **+ € 5** (hemat biaya operasional)
- FP : **- € 5** (rugi biaya operasional)
- FN : **- € 14,25** (kehilangan peluang net benefit)

Fokus pada Recall dan memperhatikan Precision menggunakan F1 Score



# Data Understanding

Dataset yang digunakan adalah (data\_bank\_marketing\_campaign.csv) dari UCI Machine Learning Repository

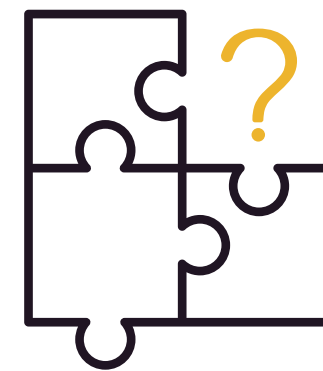
Feature	Deskripsi
age	Usia nasabah
job	Pekerjaan nasabah
balance	Saldo tabungan rata-rata nasabah
housing	Apakah nasabah memiliki kredit rumah
loan	Apakah nasabah memiliki pinjaman pribadi
contact	Jenis komunikasi kontak yang digunakan
month	Bulan terakhir nasabah dihubungi pada tahun tersebut
campaign	Jumlah kontak yang dilakukan selama kampanye ini untuk nasabah tersebut
pdays	Jumlah hari sejak nasabah terakhir dihubungi dalam kampanye sebelumnya
poutcome	Hasil dari kampanye pemasaran se

# Data Cleaning

Pembersihan data dilakukan sebelum proses pembangunan model

Total Data : 7813

	Column	Unique_Count	Unique_Sample	Null	Unknown
0	age	75	[55, 51, 36, 37, 35]	0	0
1	job	12	[admin., housemaid, technician, management, st...]	0	53
2	balance	2797	[1662, 3025, 205, 4803, 911]	0	0
3	housing	2	[no, yes]	0	0
4	loan	2	[no, yes]	0	0
5	contact	3	[cellular, telephone, unknown]	0	1460
6	month	12	[jun, may, nov, jan, sep]	0	0
7	campaign	29	[2, 1, 4, 5, 3]	0	0
8	pdays	372	[-1, 352, 21, 91, 186]	0	0
9	poutcome	4	[unknown, other, failure, success]	0	5385
10	deposit	2	[1, 0]	0	0



## Missing Value

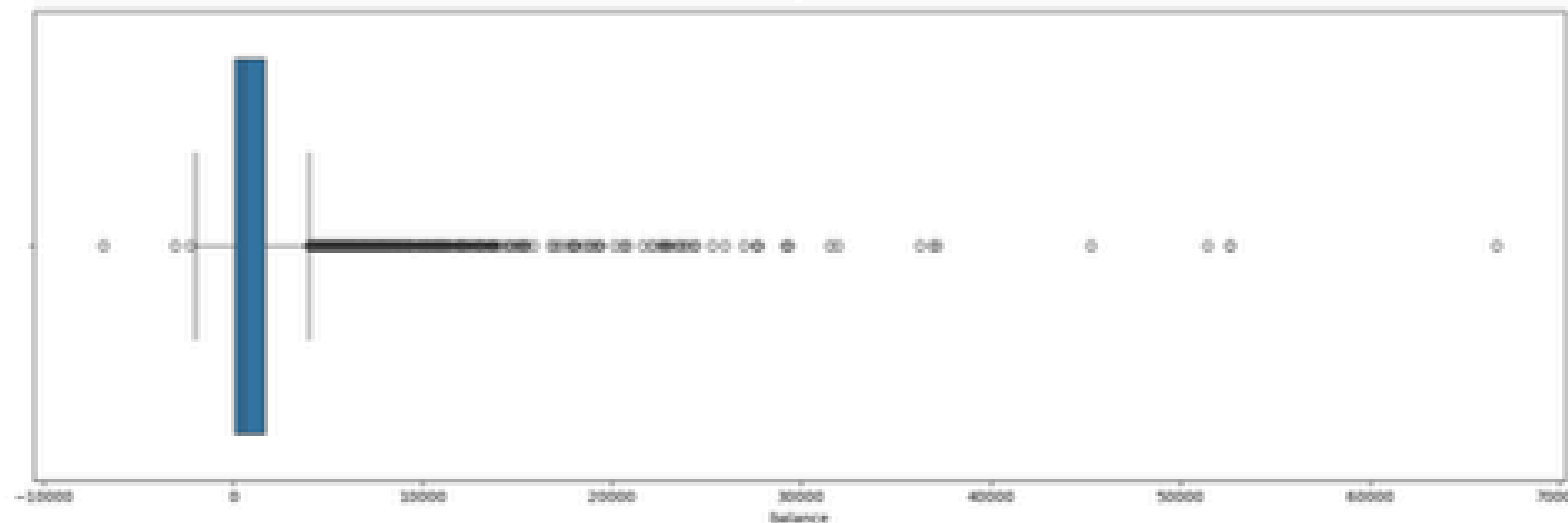
- job 0,6% → Imputasi
- contact 18%
- poutcome 69% → Tidak digunakan



## Duplicated

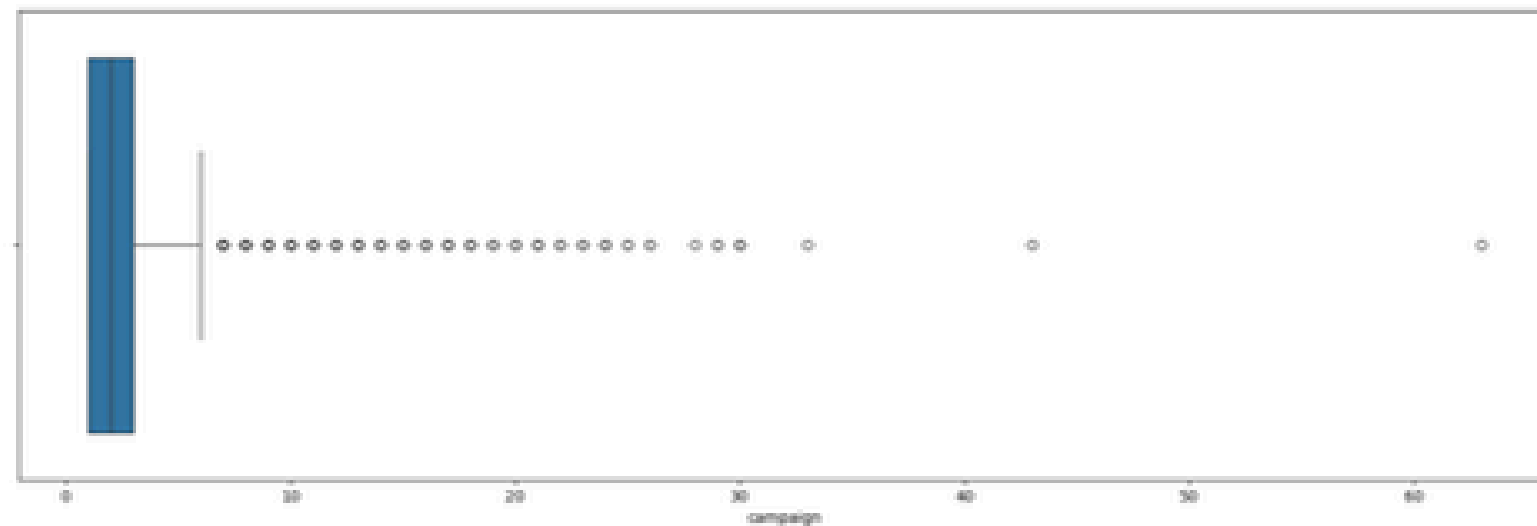
- 8 Data → Delete duplicate

# Outlier Handling



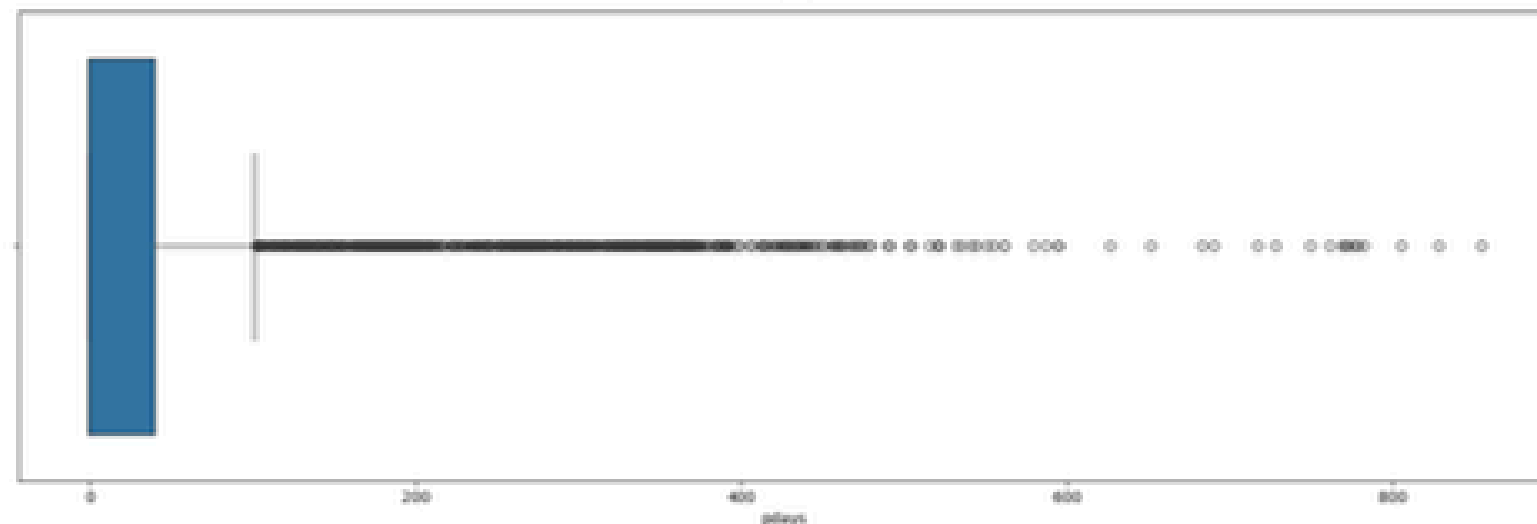
**balance** : menghapus nilai

- Kurang dari 0
- Lebih dari 30000



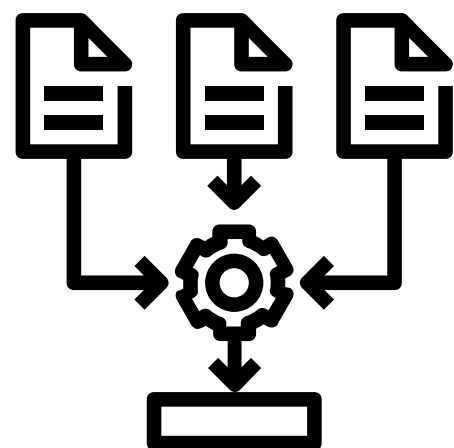
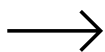
**campaign** : menghapus nilai

- Lebih dari 30



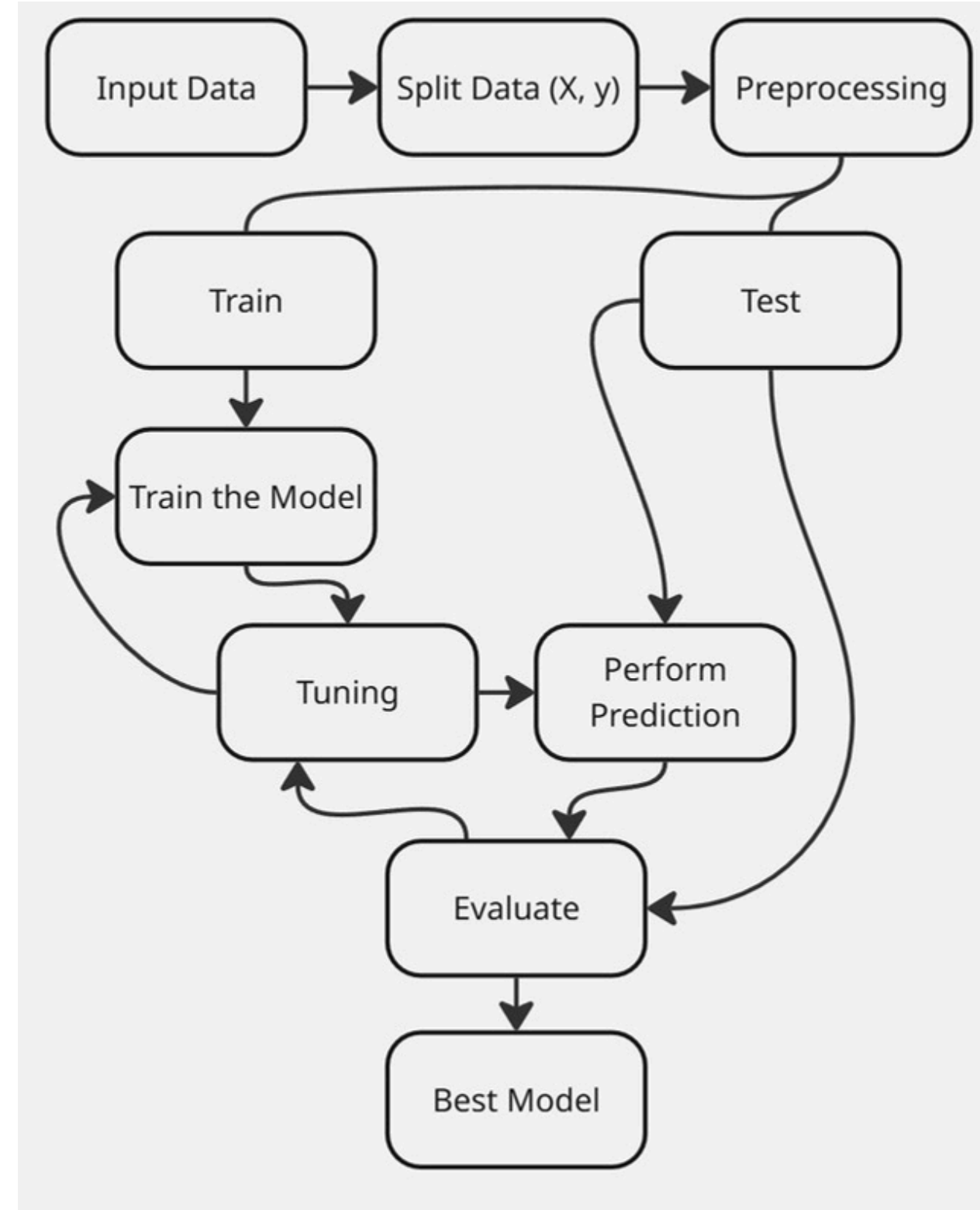
**pdays** : menghapus nilai

- Lebih dari 450

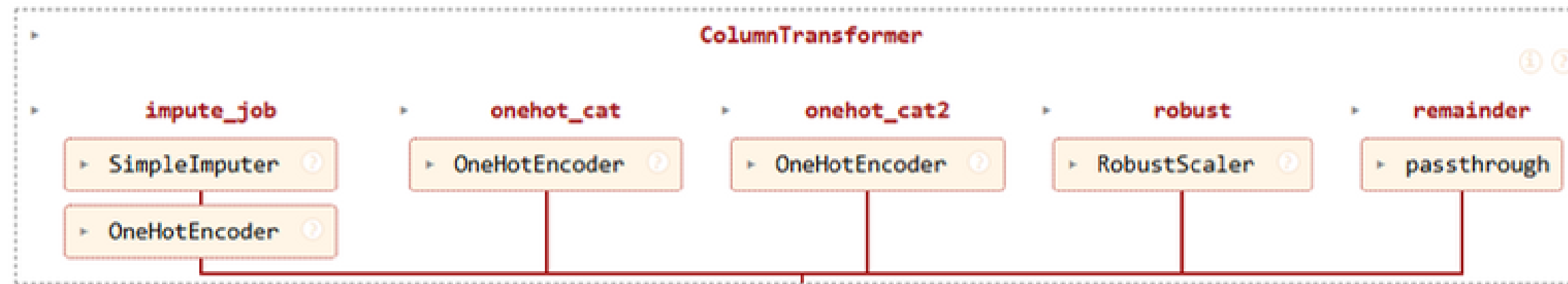


**MODELING**

# Modeling Flow



# Preprocessing



## **impute\_job**

- Imputasi dengan Modus
- Encoding dengan OneHotEncoder

## **onehot\_cat (contact, month)**

- Encoding dengan OneHotEncoder

## **onehot\_cat2 (housing, loan)**

- Encoding dengan OneHotEncoder
- Menggunakan drop='first' agar tidak redundant

## **robust (age, balance, campaign, pdays)**

- Scaling menggunakan RobustScaler

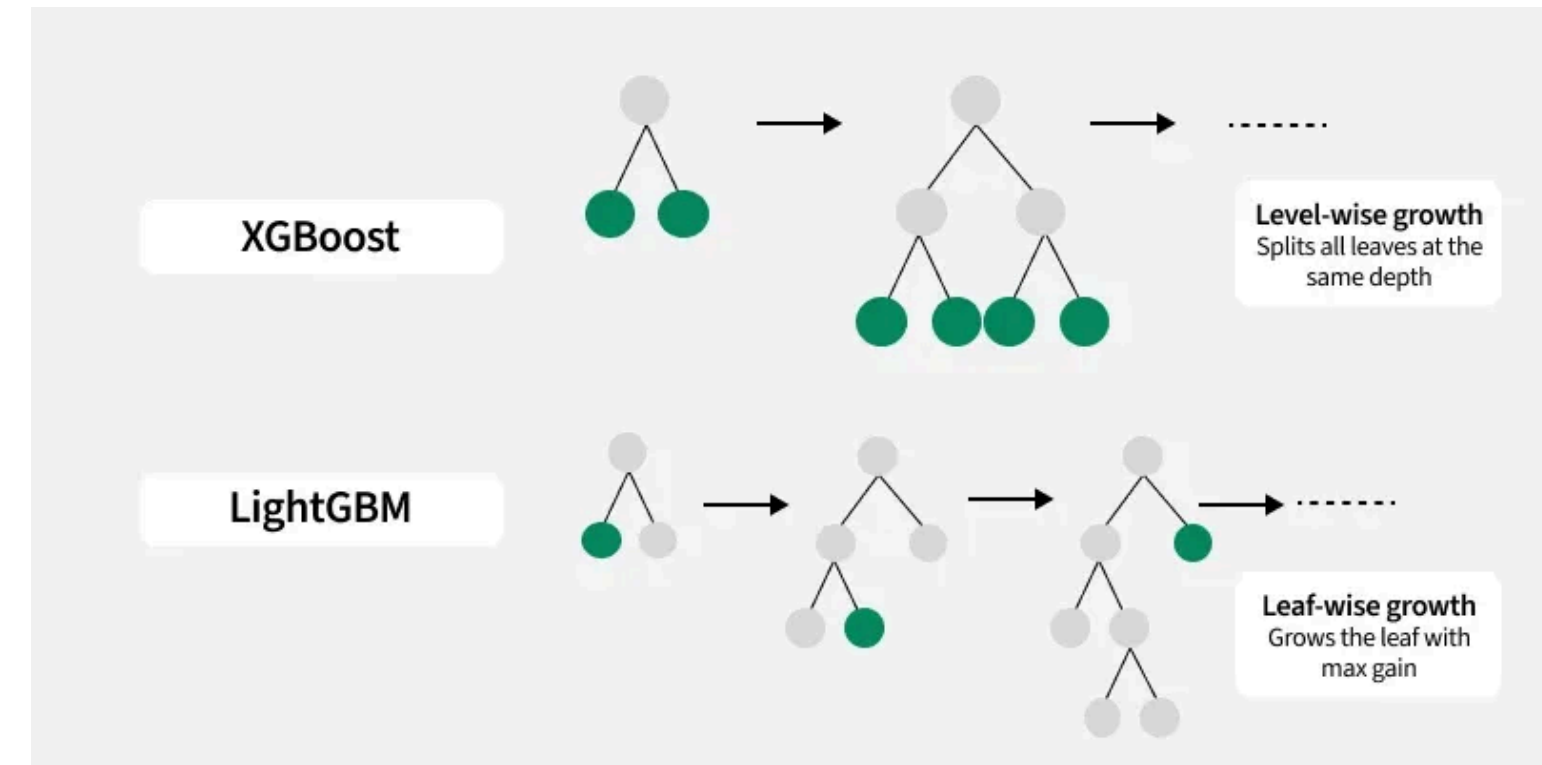
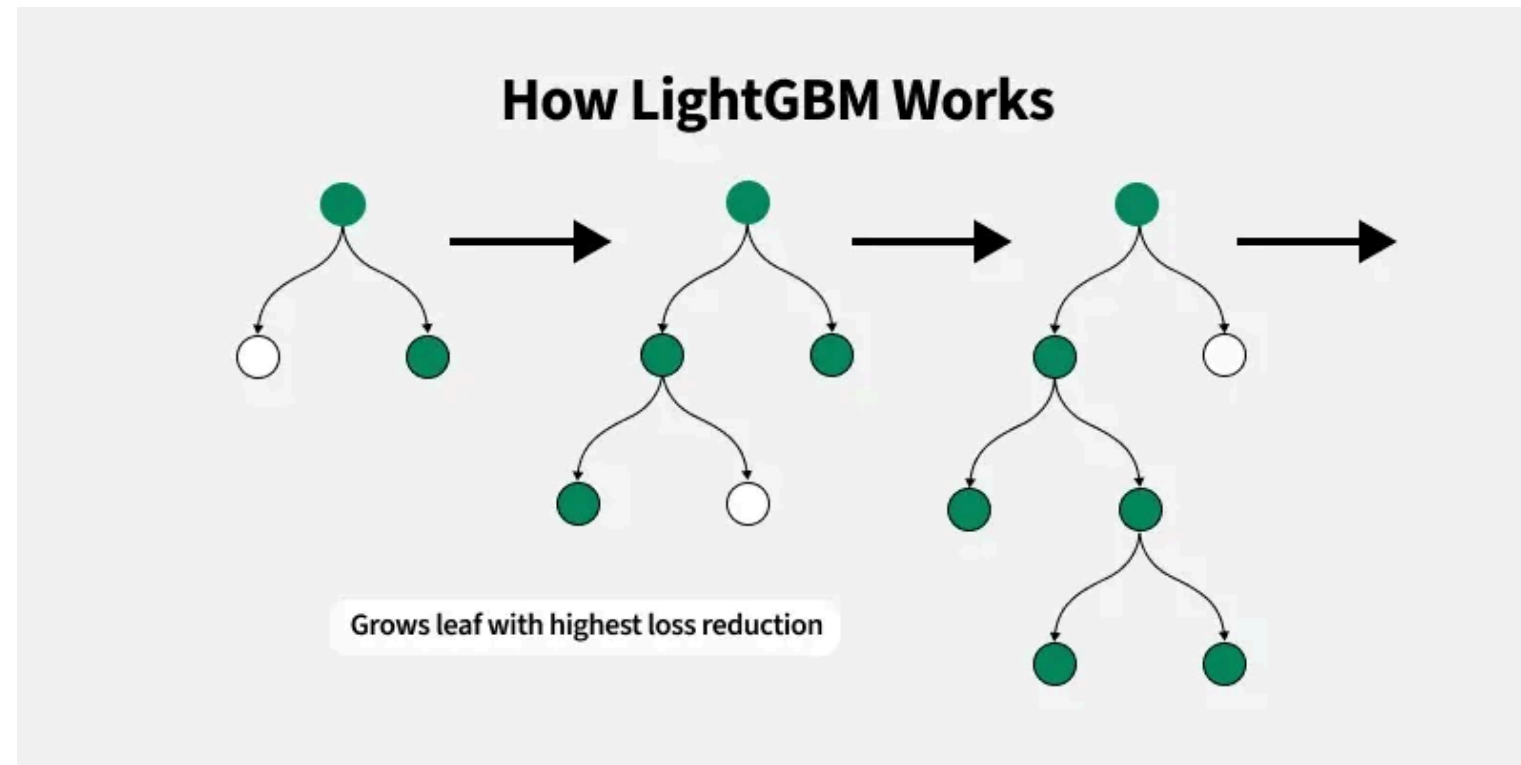
# Model Selection

	model	train_f1_mean	test_f1_mean	train_f1_std	test_f1_std	train_f2_mean	test_f2_mean	train_f2_std	test_f2_std
4	stacking	0.769	0.678	0.011	0.017	0.754	0.665	0.009	0.028
8	lgbm	0.803	0.675	0.004	0.018	0.77	0.647	0.005	0.023
6	rf	1.0	0.673	0.0	0.015	1.0	0.658	0.0	0.022
7	xgb	0.886	0.667	0.004	0.01	0.862	0.646	0.005	0.015
1	logistic	0.67	0.664	0.002	0.011	0.666	0.66	0.002	0.022
3	voting	0.838	0.664	0.003	0.011	0.822	0.651	0.006	0.021
5	bagging	0.976	0.642	0.003	0.009	0.966	0.611	0.004	0.011
0	knn	0.752	0.64	0.004	0.017	0.733	0.624	0.006	0.023
2	dtree	1.0	0.616	0.0	0.015	1.0	0.619	0.0	0.02

Stacking dan LGBM (Light Gradient Boosting Machine)  
adalah 2 model terbaik

- Light Gradient Boosting Machine (LGBM) memiliki test score std lebih stabil
- Stacking memiliki score Train dan Test yang lebih kecil overfitting nya

# Light Gradient Boosting Machine (LGBM)



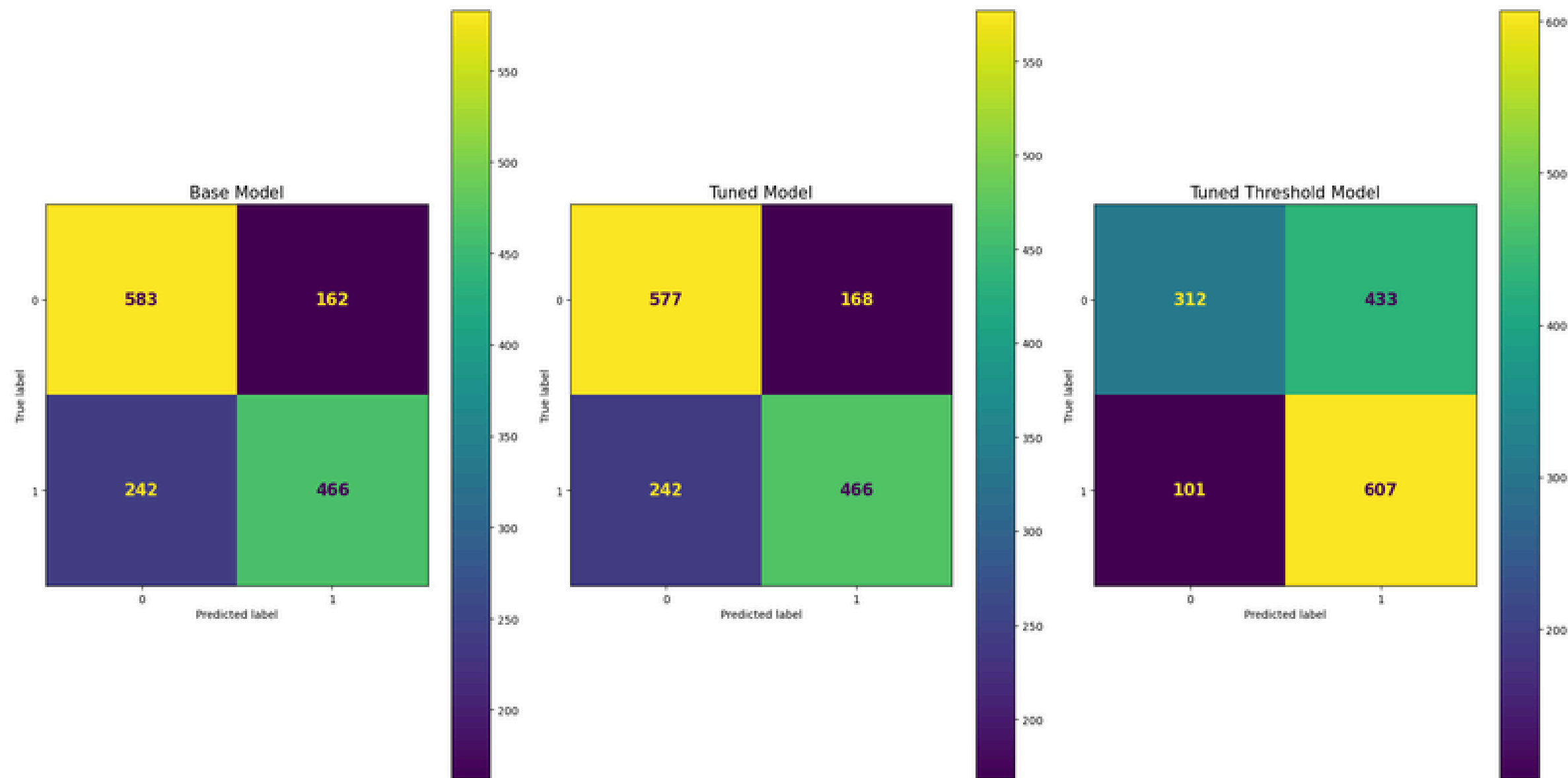
LightGBM (Light Gradient Boosting Machine) adalah metode klasifikasi dan regresi, dengan tujuan membangun model prediktif dengan cara menggabungkan banyak decision tree yang lemah (weak learners) menjadi sebuah model yang kuat (strong learner).

- LGBM dimulai dengan model sederhana
- Model menambah weak learner (tree kecil)
- Memperdalam pohon di bagian terpenting
- Pohon yang lebih dalam dibanding XGBoost

```
param_dist = {  
    'modeling__num_leaves': range(10,21),  
    'modeling__max_depth': range(3,11),  
    'modeling__learning_rate': [0.01, 0.05, 0.1],  
    'modeling__n_estimators': [100, 300, 500],  
    'modeling__objective': ['binary'],  
    'modeling__class_weight': ['balanced'],  
    'modeling__min_child_samples': range(10,21),  
    'modeling__importance_type': ['split','gain'],  
}
```



# Model Tuning + Threshold Optimization



Base Model Score :

	train_base	test_base
precision	0.844	0.742
recall	0.726	0.658
f1	0.781	0.698
f2	0.747	0.673

Tuned Model Score :

	train_tuned	test_tuned
precision	0.762	0.735
recall	0.641	0.658
f1	0.696	0.694
f2	0.662	0.672

Tuned Threshold Model Score :

	train_thres	test_thres
precision	0.607	0.584
recall	0.870	0.857
f1	0.715	0.695
f2	0.801	0.784

# Model Tuning + Threshold Optimization

Base Model Cost :

Total Sample: 1453

Benefits:

- TP Revenue: € 6,640.50 (from 466 successful calls)
- TN Savings: € 2,915.00 (from 583 avoided calls)
- Total Benefit: € 9,555.50

Costs:

- FP Cost: € 810.00 (from 162 wasted calls)
- FN Cost: € 3,448.50 (from 242 missed opportunities)
- Total Cost: € 4,258.50

Net Profit: € 5,297.00

Base model

Tuned Model Cost :

Total Sample: 1453

Benefits:

- TP Revenue: € 6,640.50 (from 466 successful calls)
- TN Savings: € 2,885.00 (from 577 avoided calls)
- Total Benefit: € 9,525.50

Costs:

- FP Cost: € 840.00 (from 168 wasted calls)
- FN Cost: € 3,448.50 (from 242 missed opportunities)
- Total Cost: € 4,288.50

Net Profit: € 5,237.00

Tuned model

Tuned Threshold Model Cost :

Total Sample: 1453

Benefits:

- TP Revenue: € 8,649.75 (from 607 successful calls)
- TN Savings: € 1,560.00 (from 312 avoided calls)
- Total Benefit: € 10,209.75

Costs:

- FP Cost: € 2,165.00 (from 433 wasted calls)
- FN Cost: € 1,439.25 (from 101 missed opportunities)
- Total Cost: € 3,604.25

Net Profit: € 6,605.50

Tuned + Threshold Optimization model

Total Sample: 1453

Benefits:

- TP Revenue: € 0.00 (from 0 successful calls)
- TN Savings: € 3,725.00 (from 745 avoided calls)
- Total Benefit: € 3,725.00

Costs:

- FP Cost: € 0.00 (from 0 wasted calls)
- FN Cost: € 10,089.00 (from 708 missed opportunities)
- Total Cost: € 10,089.00

Net Profit: € -6,364.00

Semua Prediksi 0

Total Sample: 1453

Benefits:

- TP Revenue: € 10,089.00 (from 708 successful calls)
- TN Savings: € 0.00 (from 0 avoided calls)
- Total Benefit: € 10,089.00

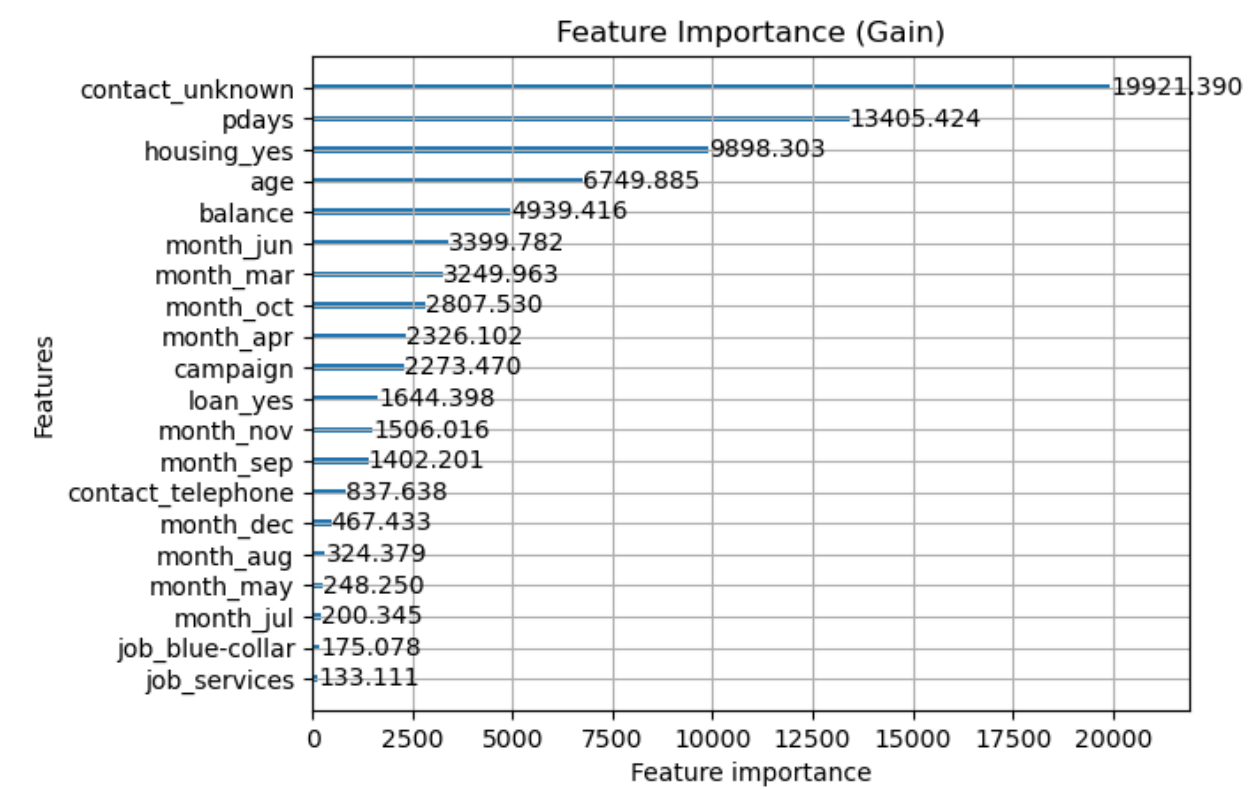
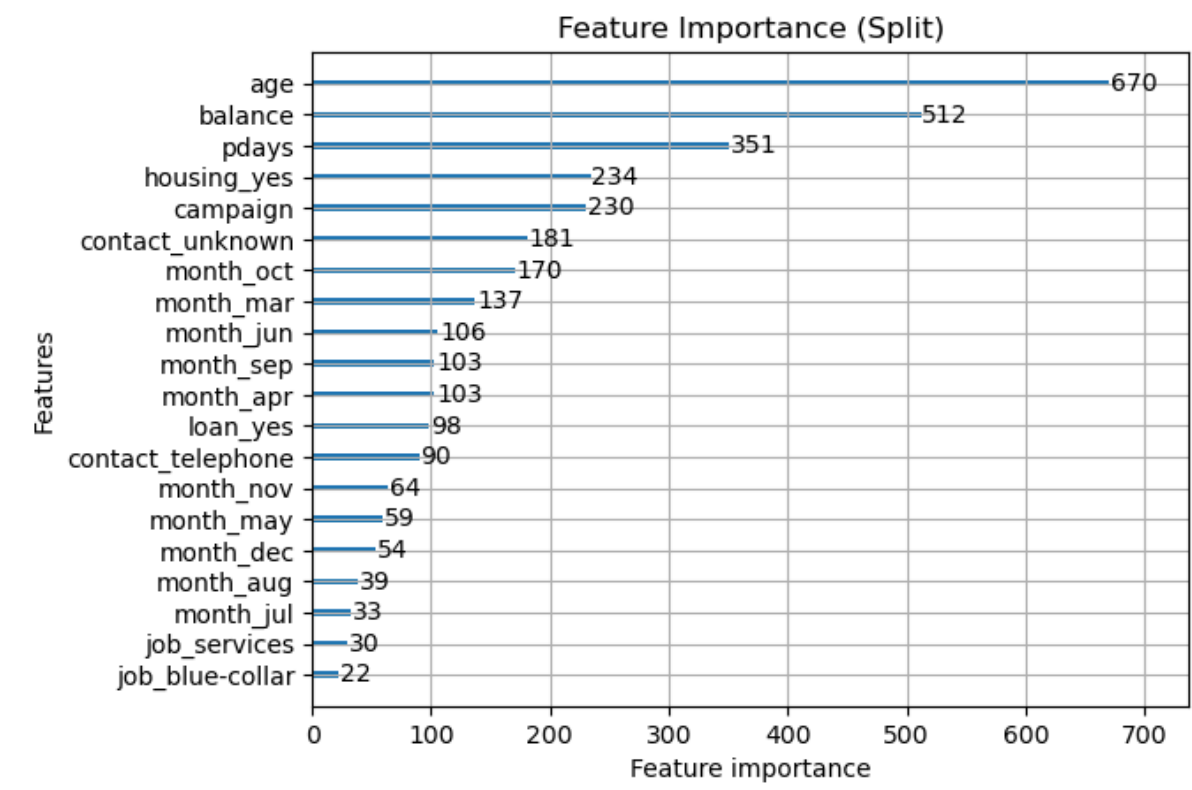
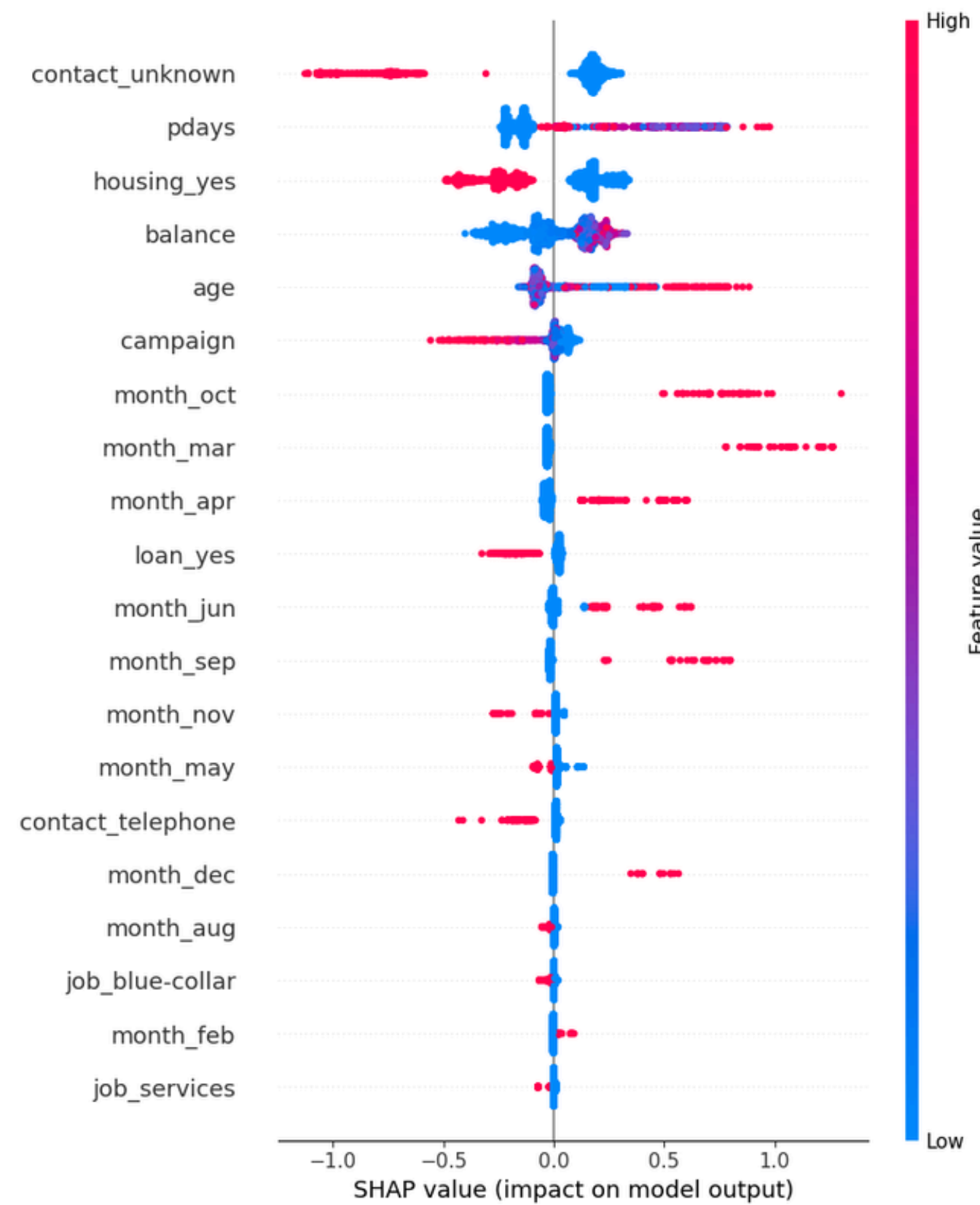
Costs:

- FP Cost: € 3,725.00 (from 745 wasted calls)
- FN Cost: € 0.00 (from 0 missed opportunities)
- Total Cost: € 3,725.00

Net Profit: € 6,364.00

Semua Prediksi 1

# Model Explanation



# Conclusion

## Ringkasan

- Model akhir : LGBM (tuned + threshold optimized)
- Fokus optimasi : Recall kelas 1 (minimalkan nasabah potensial terlewat)
- Konsekuensi : Terima risiko False Positive (hubungi yang tidak tertarik)
- Alasan : Kerugian kehilangan nasabah  $>$  biaya salah kontak

## Manfaat

- Efektivitas karyawan meningkat
- Hemat waktu & biaya
- Profit bertambah
- Lebih efisien dibanding menelpon semua nasabah



# Recommendation

## Rekomendasi Bisnis

- Fokus telepon nasabah dengan probabilitas tinggi
- Segmentasi perlakuan:
  - Tinggi → telepon langsung
  - Sedang → email / WA (lebih murah)
  - Rendah → bisa dipertimbangkan tidak dihubungi
- Evaluasi biaya & sesuaikan threshold secara berkala

## Meningkatkan Performa Model

- Tambah fitur baru (contoh: gaji, jumlah deposit)
- Perjelas nilai unknown untuk treatment yang tepat
- Update model tiap batch campaign agar adaptif



→

# Thank You !

YOUR FEEDBACK AND SUGGESTIONS ARE HIGHLY  
APPRECIATED

