



Banking Marketing Precision

Data Science batch 41

Presented by
Kelompok 4 - DataWhiz



About Us

DataWhiz adalah salah satu kelompok Final Project pada **Rakamin Academy** Data Science - Batch 41 yang terdiri dari 7 (tujuh) orang. Nama kelompok **DataWhiz** mengandung makna yang cukup jelas terkait dengan keahlian dalam data. Nama ini menggambarkan bahwa kelompok ini ahli dalam ilmu data, siap untuk menghadapi tantangan dan menciptakan solusi yang inovatif menggunakan data.

Pada Final Project ini, kami membuat Machine Learning dengan dataset "Banking Market Precision" untuk mengidentifikasi pola dan faktor-faktor yang berkontribusi terhadap keberhasilan *campaign* pemasaran perbankan. Kami melakukan analisis mendalam terhadap data pelanggan terhadap *campaign* pemasaran sebelumnya.

Our Team.



Achmad
Fathony



Aldi
Maulana H



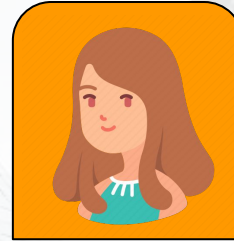
Galih J.
Rudianto



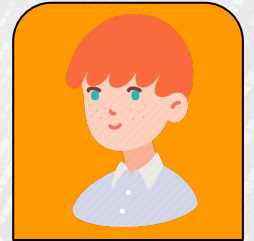
M.
Rayyan



Sutrisni



Wika
Rabila P



Yoga
Bagas P



Table of Content

- ❑ **Problem Statement**
- ❑ **Exploratory Data Analysis & Insight**
- ❑ **Pre-Processing**
- ❑ **Machine Learning Model**
- ❑ **Model Simulation**
- ❑ **Recommendation**





Problem Statement



Current Problem.

Data ini terkait dengan kampanye pemasaran langsung dari sebuah lembaga perbankan di Portugal. Data mencakup informasi tentang klien bank dan hasil panggilan telepon selama kampanye pemasaran.

1. **Menghadapi tantangan dalam meningkatkan langganan produk deposito berjangka**, padahal sudah berupaya untuk mempromosikan.
2. **Conversion rate** nasabah yang menyetujui deposito berjangka tidak maksimal, yaitu hanya **9.1%**. Sehingga campaign yang dilakukan belum efektif dalam menarik nasabah.

Goals.



Meningkatkan efisiensi dan efektivitas dalam kampanye pemasaran produk



Meningkatkan profitabilitas perusahaan dengan meningkatkan pendapatan dari langganan deposito dan mengoptimalkan pengeluaran pemasaran

Objective.



Mengembangkan Model Prediktif, dapat memprediksi kemungkinan langganan deposito oleh nasabah berdasarkan atribut dan perilaku mereka.



Identifikasi Fitur yang Signifikan, sehingga perusahaan dapat fokus pada faktor-faktor kunci dalam strategi pemasaran.



Pengembangan Model Interpretable, model yang dapat diinterpretasikan dengan baik sehingga pemangku kepentingan dapat memahami faktor-faktor yang mempengaruhi prediksi langganan deposito.



Business Metrics.

1.



Conversion Rate

2.



Meningkatkan profit





Exploratory Data Analysis & **Insight**

Data Overview.



1. Data terdiri dari **45.663 baris** dan **17 kolom**
2. Data terdiri dari **10 data kategorikal** dan **7 data numerik**
3. **Terdapat data null** atau **kosong** sebanyak 346 pada kolom *duration*, 382 pada kolom *campaign*, 217 pada kolom *pdays*, 268 pada kolom *previous*, dan 308 pada kolom *y*

Setelah melakukan **cleaning**, data berkurang sebanyak **9.534 baris**

Baseline.

9.15%

90.85%

-  Nasabah **menyetujui** berlangganan deposito berjangka
-  Nasabah **menolak** berlangganan deposito berjangka

Target Feature.



"Yes"

Nasabah **setuju** untuk berlangganan deposito berjangka



"No"

Nasabah **menolak** untuk berlangganan deposito berjangka

Data Insight.

Distribusi Customer Subscription

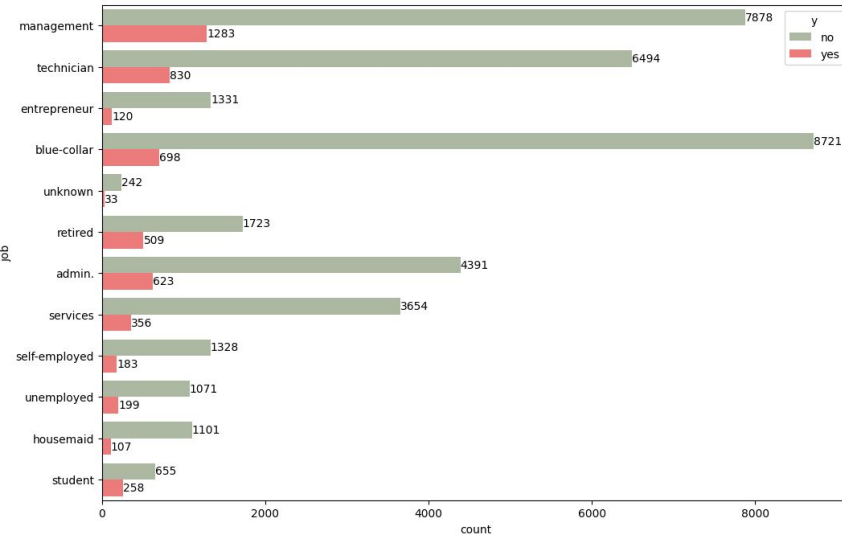


Dari grafik disamping sebaran data imbalance dimana customer yang melakukan subscription sebesar 11,87% dan 88,13 % tidak melakukan subscription

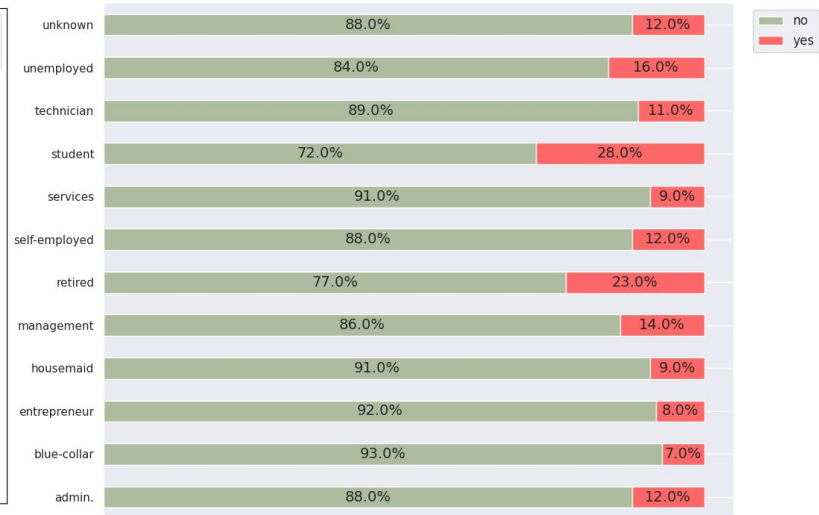
Data Insight.

Pengaruh job nasabah terhadap Subscription

Subscription by job



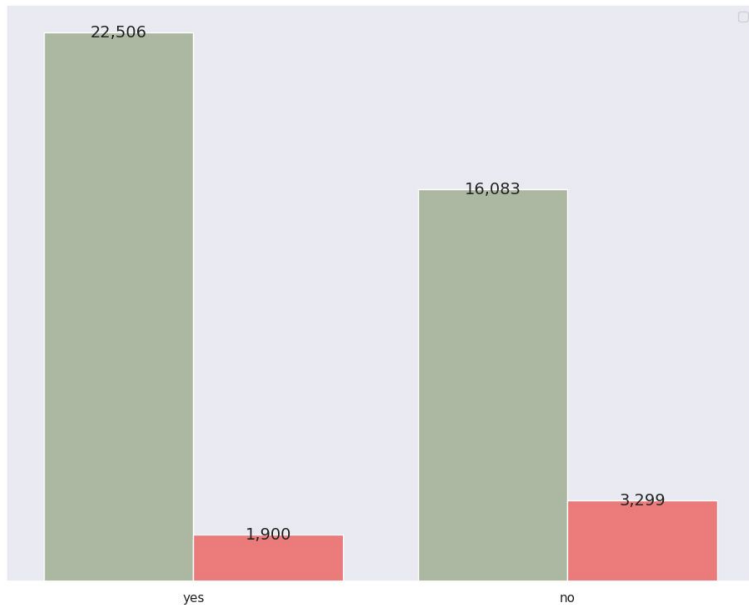
Ratio of Subscription by job



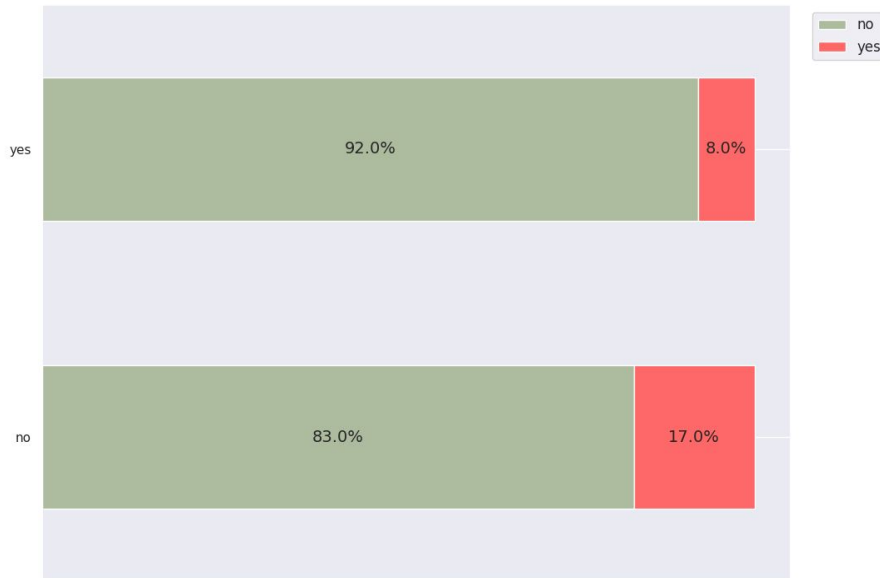
Dapat dilihat bahwa customer dengan pekerjaan management memiliki total Subscription terbanyak sebanyak 1283 orang dan pekerjaan technician melakukan Subscription sebanyak 830 orang. Dari rasio Subscription berdasarkan jenis job student dan retired memiliki perbandingan lebih banyak melakukan Subscription deposit yaitu 28% dan 23%.

Pengaruh housing nasabah terhadap Subscription

Subscription by housing

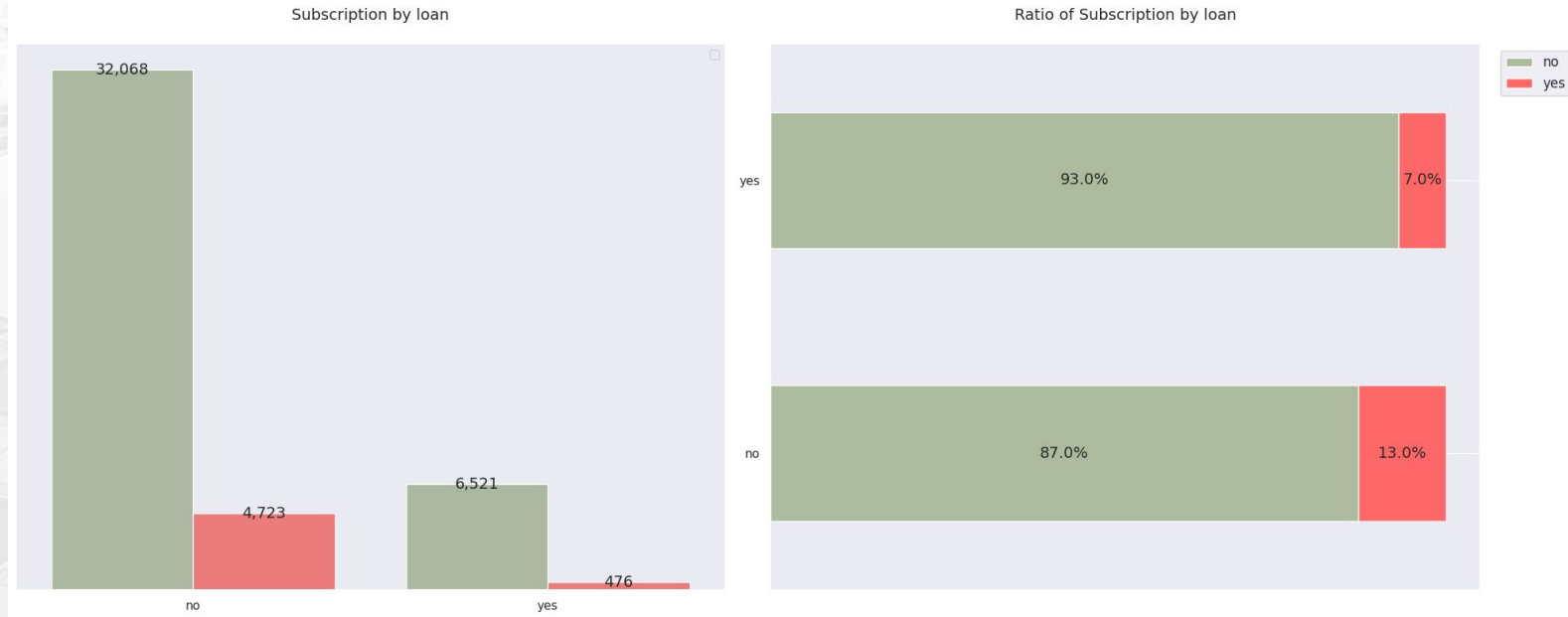


Ratio of Subscription by housing



Dapat dilihat bahwa customer yang tidak memiliki pinjaman rumah total Subscription sebanyak 3299 orang dan customer yang memiliki pinjaman rumah total Subscription sebanyak 1900 orang. Dari rasio Subscription berdasarkan jenis housing memiliki perbandingan lebih banyak melakukan Subscription customer yang tidak memiliki pinjaman rumah atau housing 17%

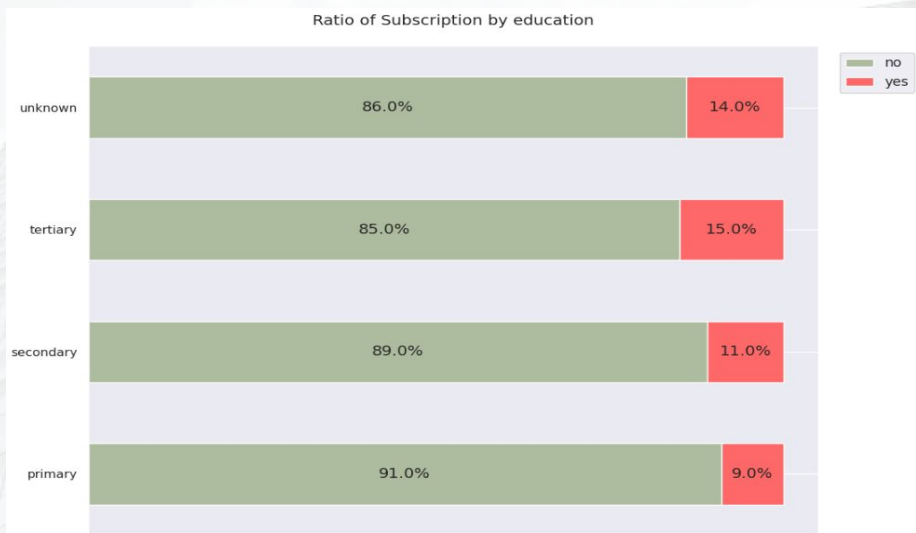
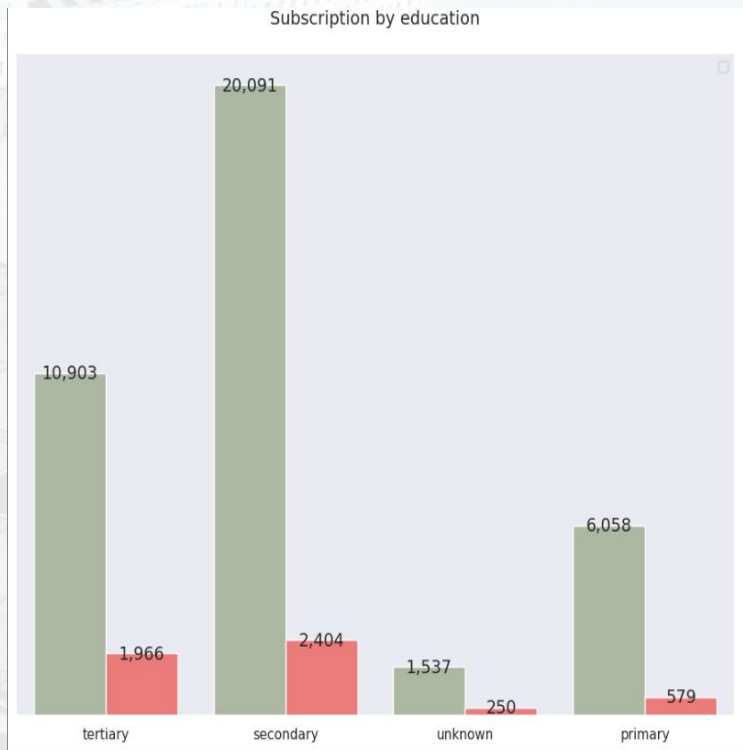
Pengaruh loan nasabah terhadap Subscription



Dapat dilihat bahwa customer yang tidak memiliki pinjaman pribadi total Subscription sebanyak 4723 orang dan customer yang memiliki pinjaman pribadi total Subscription sebanyak 476 orang. Dari rasio Subscription berdasarkan jenis loan memiliki perbandingan lebih banyak melakukan Subscription customer yang tidak memiliki pinjaman pribadi atau loan 13%

Data Insight.

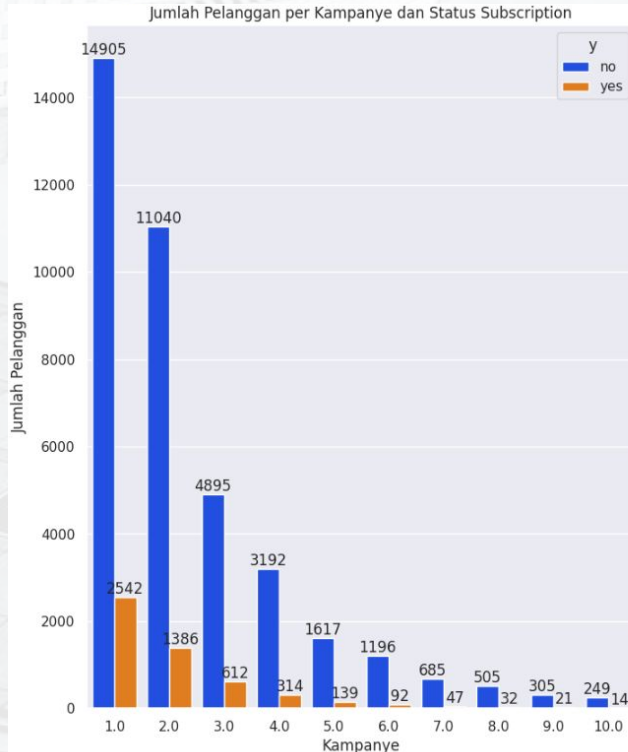
Pengaruh education nasabah terhadap Subscription



Dapat dilihat bahwa customer yang melakukan Subscription lebih banyak pada tingkat pendidikan secondary dengan jumlah 2404 orang, lalu pada posisi kedua pada tingkat pendidikan tertiary berjumlah 1966. Hal tersebut didukung dengan rasio Subscription by education dimana customer melakukan Subscription pada tingkat pendidikan tertiary 15% dan paling sedikit dari tingkat pendidikan primary 9%

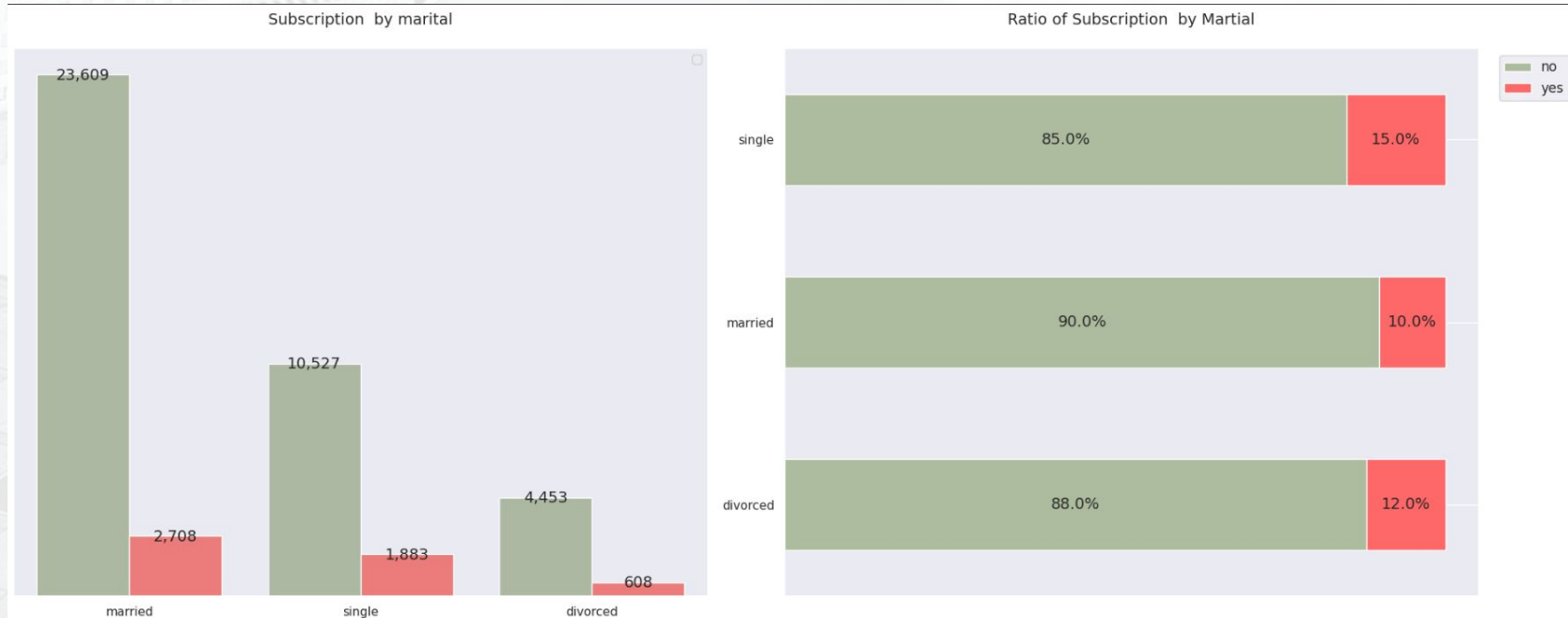
Data Insight.

Pengaruh jumlah campaign terhadap Subscription



Dapat dilihat kampanye 1 tidak terlalu dapat menarik customer untuk berlangganan deposit. Namun, pada kampanye 1 lebih banyak menarik customer daripada kampanye lainnya. Menaikkan jumlah kontak customer tidak terbukti menaikkan customer yang ingin berlangganan.

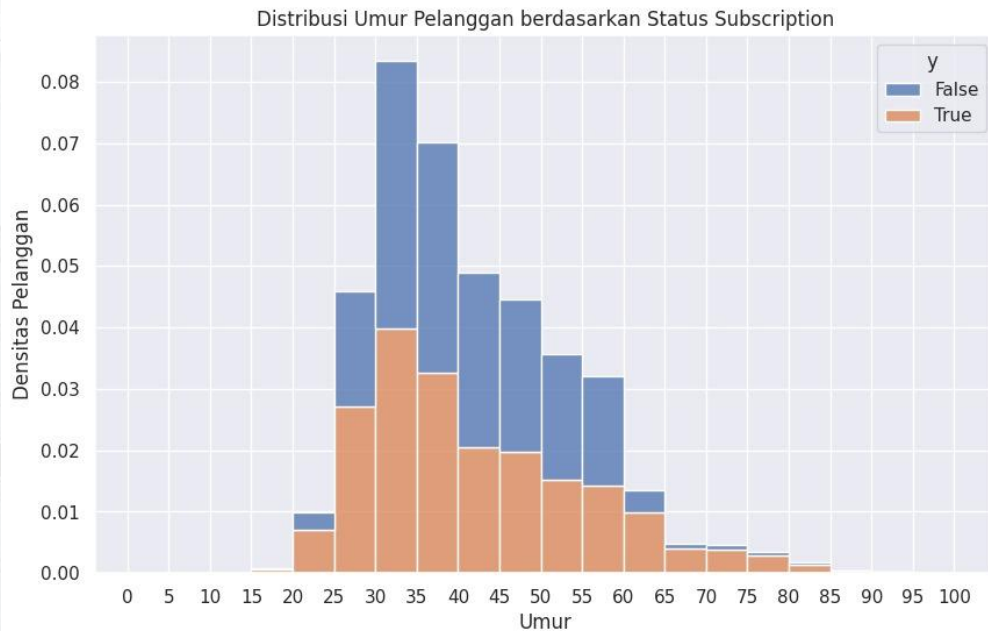
Pengaruh marital nasabah terhadap Subscription



Dapat dilihat bahwa customer yang melakukan Subscription deposit paling banyak pada status married sebanyak 2708, lalu status single yang melakukan Subscription deposit sebanyak 1883. Rasio customer yang melakukan Subscription deposit paling banyak dilakukan oleh status martial single dengan 15%

Data Insight.

Pengaruh umur nasabah terhadap Subscription

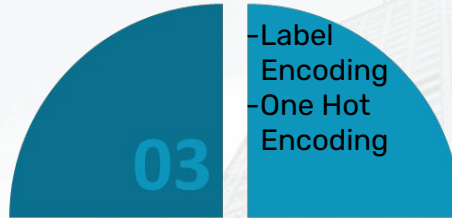


Pre-Processing.

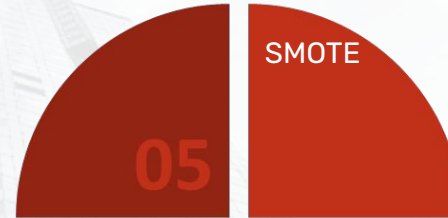
Handling Missing Values & Duplicates



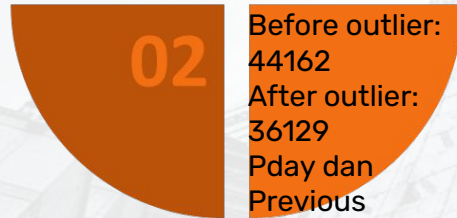
Feature Encoding



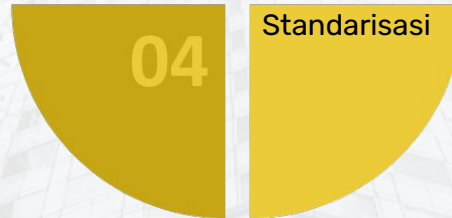
Handling Class Imbalance



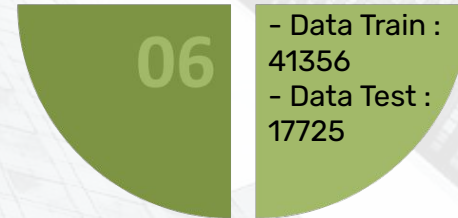
Machine Learning Model



Handling Outlier



Feature Transformation



Split Data Train & Test (70 : 30)

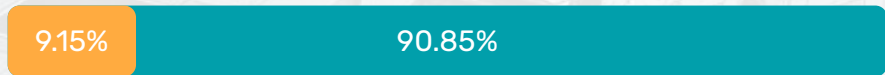




Machine Learning Model

Model Evaluation.

Aktual	Tidak Setuju	Setuju
	Tidak Setuju	Setuju
Tidak Setuju	True Negative	False Positive
Setuju	False Negative	True Positive

Baseline.



-  Nasabah **menyetujui** berlangganan deposito berjangka
-  Nasabah **menolak** berlangganan deposito berjangka

$$\text{Recall} = (\text{TP}) / (\text{TP} + \text{FN})$$

Recall

Recall digunakan untuk menekan nilai **False Negative** agar Prediksi menolak, tetapi berpotensi tinggi untuk setuju

Modeling Result.

Imbalance Data				
	Logistic Reggresion	KNN	Decision Tree	Random Forrest
Accuracy (Test Set)	0,91	0,91	0,82	0,92
Precision (Test Set)	0,57	0,56	0,42	0,61
Recall (Test Set):	0,22	0,17	0,44	0,33
F1-Score (Test Set):	0,31	0,26	0,43	0,43
roc_auc (test-proba)	0,88	0,76	0,69	0,93
roc_auc (train-proba)	0,88	0,96	1	1

Under Sampling				
	Logistic Reggresion	KNN	Decision Tree	Random Forrest
Accuracy (Test Set)	0,8	0,74	0,8	0,85
Precision (Test Set)	0,79	0,75	0,77	0,82
Recall (Test Set):	0,73	0,63	0,77	0,85
F1-Score (Test Set):	0,76	0,69	0,77	0,83
roc_auc (test-proba)	0,88	0,8	0,79	0,92
roc_auc (train-proba)	0,88	0,92	1	1

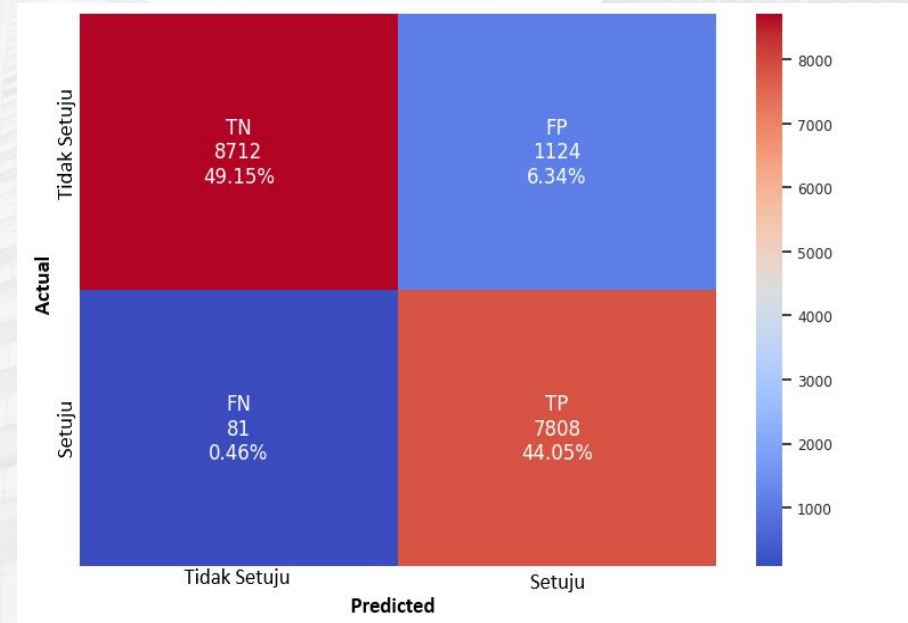
Over Sampling				
	Logistic Reggresion	KNN	Decision Tree	Random Forrest
Accuracy (Test Set)	0,8	0,91	0,96	0,97
Precision (Test Set)	0,81	0,85	0,92	0,93
Recall (Test Set):	0,74	0,97	1	1
F1-Score (Test Set):	0,77	0,91	0,96	0,97
roc_auc (test-proba)	0,89	0,96	0,96	1
roc_auc (train-proba)	0,89	0,99	1	1

SMOTE				
	Logistic Reggresion	KNN	Decision Tree	Random Forrest
Accuracy (Test Set)	0,86	0,92	0,89	0,93
Precision (Test Set)	0,83	0,86	0,87	0,9
Recall (Test Set):	0,84	0,98	0,9	0,95
F1-Score (Test Set):	0,92	0,92	0,88	0,92
roc_auc (test-proba)	0,97	0,97	0,89	0,98
roc_auc (train-proba)	1	1	1	1

Modeling Result.

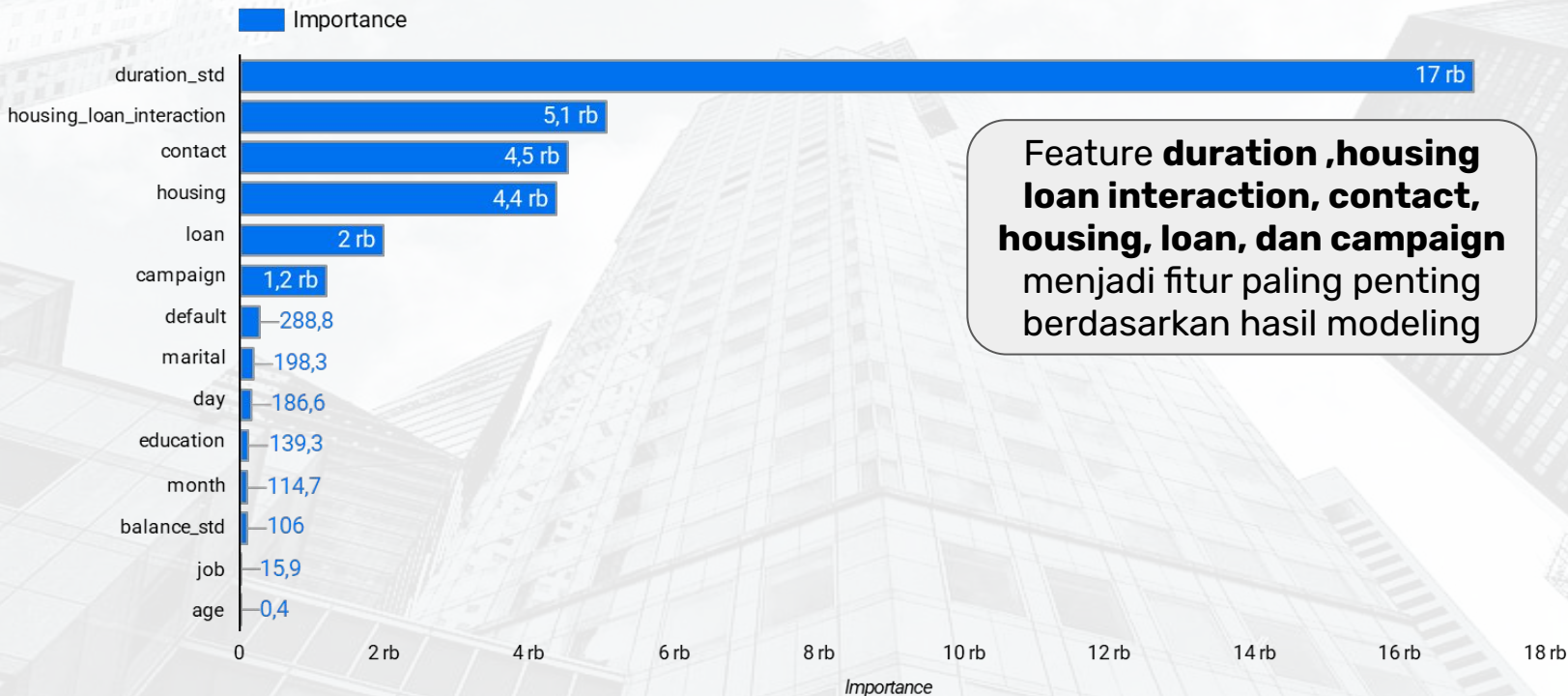
	KNN	KNN_TuningParamater
Accuracy (Test Set)	0,91	0,94
Precision (Test Set)	0,85	0,88
Recall (Test Set):	0,97	0,99
F1-Score (Test Set):	0,91	0,93
roc_auc (test-proba)	0,96	0,97
roc_auc (train-proba)	0,99	0,99

Confusion Matrix KNN



Feature Importance.

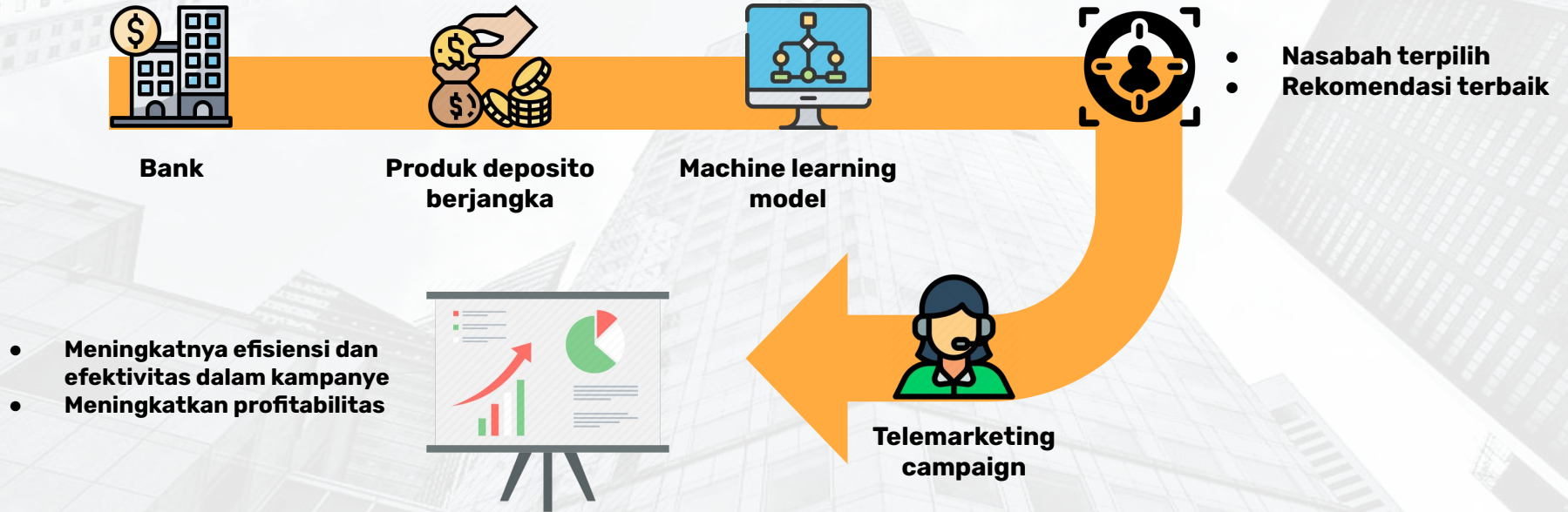
Feature Importance





Business Simulation

Business Simulation.



Business Simulation.

Sebelum machine learning

Profit:

150.000 \$ (3.306 nasabah)

Conversion rate:

9.1 % (3.306 nasabah setuju)



Sesudah machine learning

Profit:

196.700 \$ (4.335 nasabah)

Conversion rate:

11,9 % (4.335 nasabah setuju)



Business Recommendation



Business Recommendation.



- Target segmentasi berdasarkan usia customer yaitu pada usia 25 - 35 , usia 36 - 50, dan usia >50.
- Buat promosi untuk customer yang memiliki level education secondary seperti Cashback & bonus untuk pelajar yang ingin melanjutkan pendidikan tinggi.
- Mencari target customer yang tidak memiliki pinjaman rumah (housing) dan pinjaman pribadi (loan).
- Membuat promosi untuk pasangan yang baru menikah untuk melakukan Subscription deposito.
- Melakukan promosi pada bulan may karena pada bulan tersebut memiliki customer yang banyak.

Thank You



Rakamin
Academy

