

Absolute – Kelompok 7

Dokumen
Laporan Final
Project –
Stage 4



Background

Deposito berjangka :

- Salah satu sumber pendapatan utama bank
- Investasi tunai
- Tingkat bunga disepakati dengan nasabah dalam waktu tertentu

Pemasaran Deposito Berjangka :

- Menggunakan telepon masih cara paling efektif
- Biayanya mahal
- Penting untuk identifikasi dan pemetaan prioritas pelanggan

Goals dan Objective

Goal :

- Meningkatkan jumlah nasabah yang membeli deposito

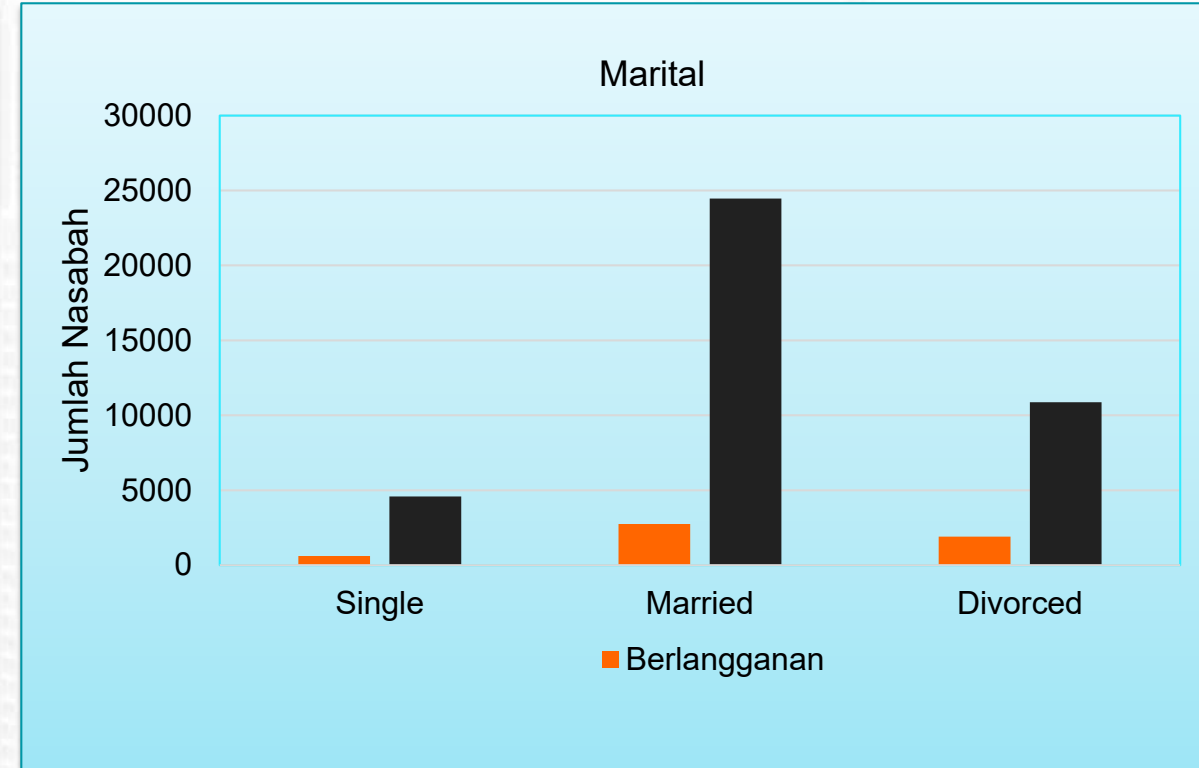
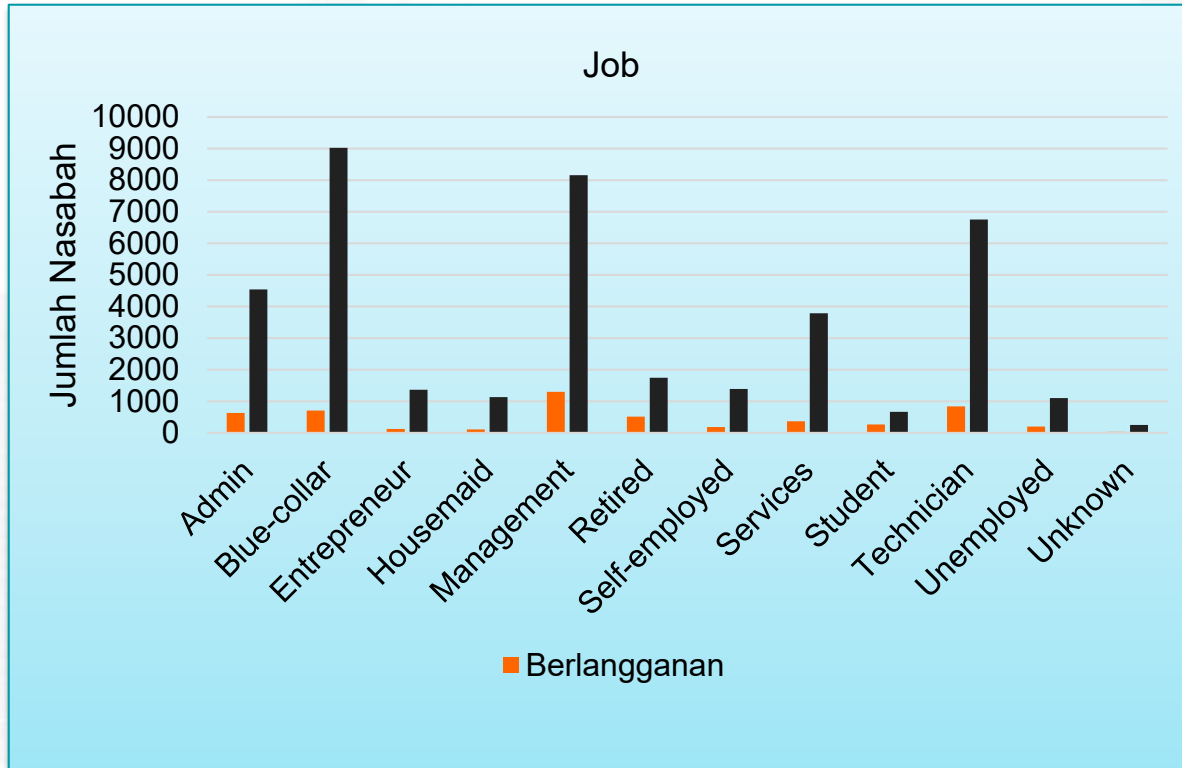
Objectives:

- Membuat machine learning model untuk memprediksi nasabah yang berpotensi membeli deposito
- Membuat sistem pengelompokan nasabah berdasarkan insight data
- Membuat penawaran khusus deposito (conversion rate) sesuai dengan data calon customer

Business Metrics :

- Client subscribed deposit conversion rate

EDA & Insight



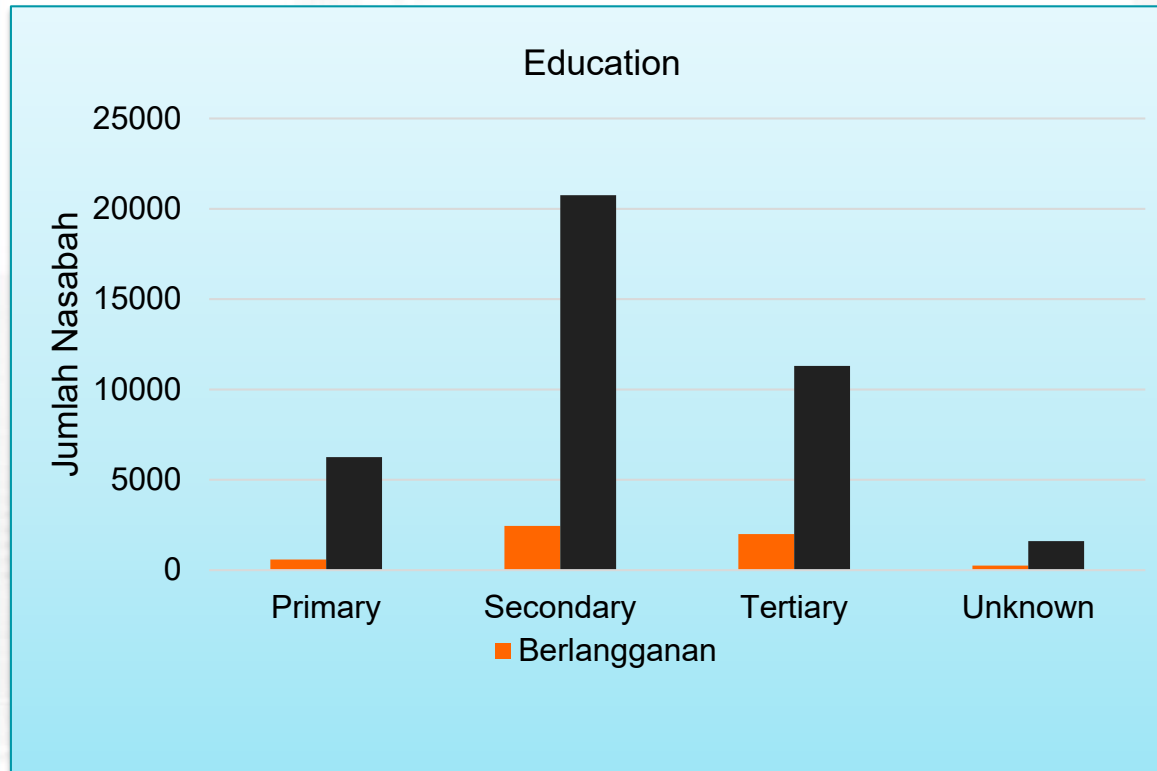
Job yang berpotensi berlangganan deposito tertinggi dengan total yaitu :

1. Management : 1301
2. Technician : 840
3. Blue-collar : 708

Status yang berpotensi berlangganan deposito tertinggi dengan total yaitu :

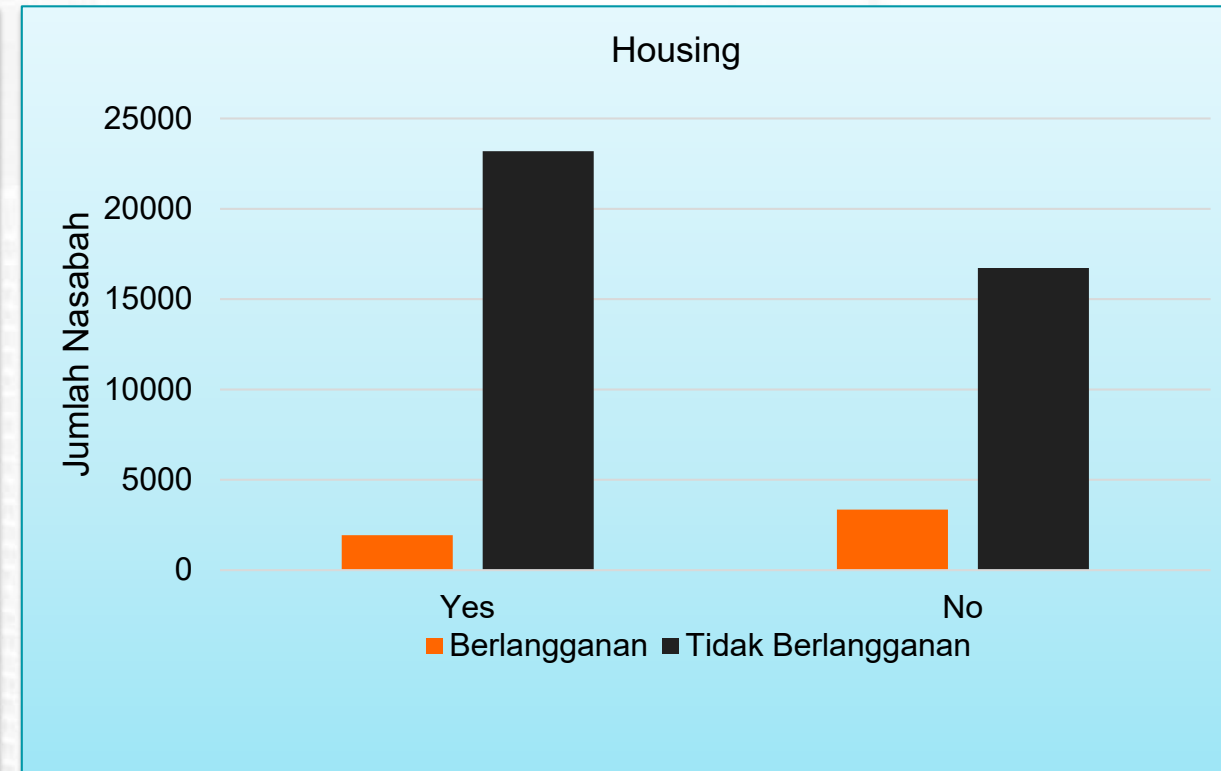
1. Married : 2755
2. Divorced : 1912
3. Single : 622

EDA & Insight



Pendidikan yang berpotensi berlangganan deposito tertinggi dengan total yaitu :

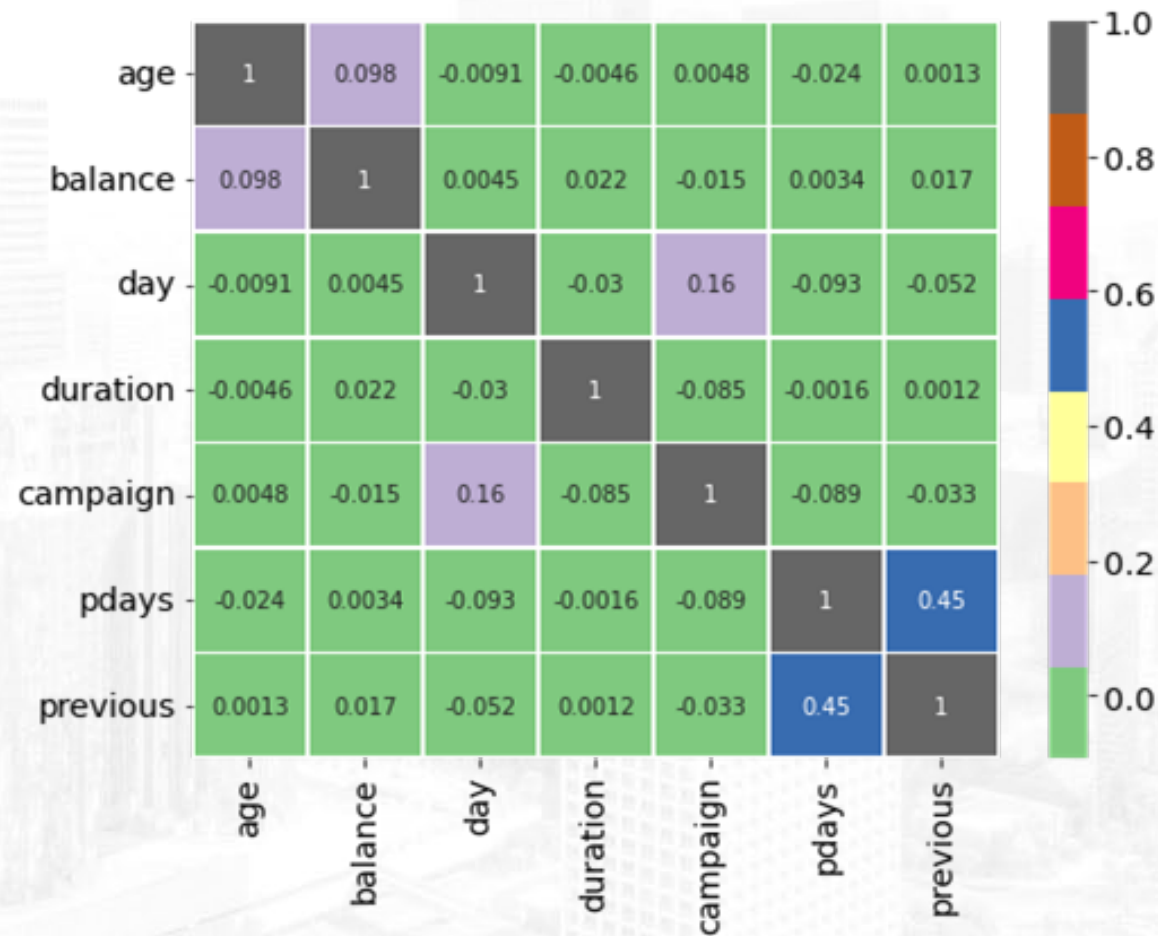
1. Secondary : 2450
2. Tertiary : 1996
3. Primary : 591



Nasabah yang tidak menyewa rumah atau memiliki tempat tinggal sendiri berpotensi berlangganan deposito dengan total yaitu :

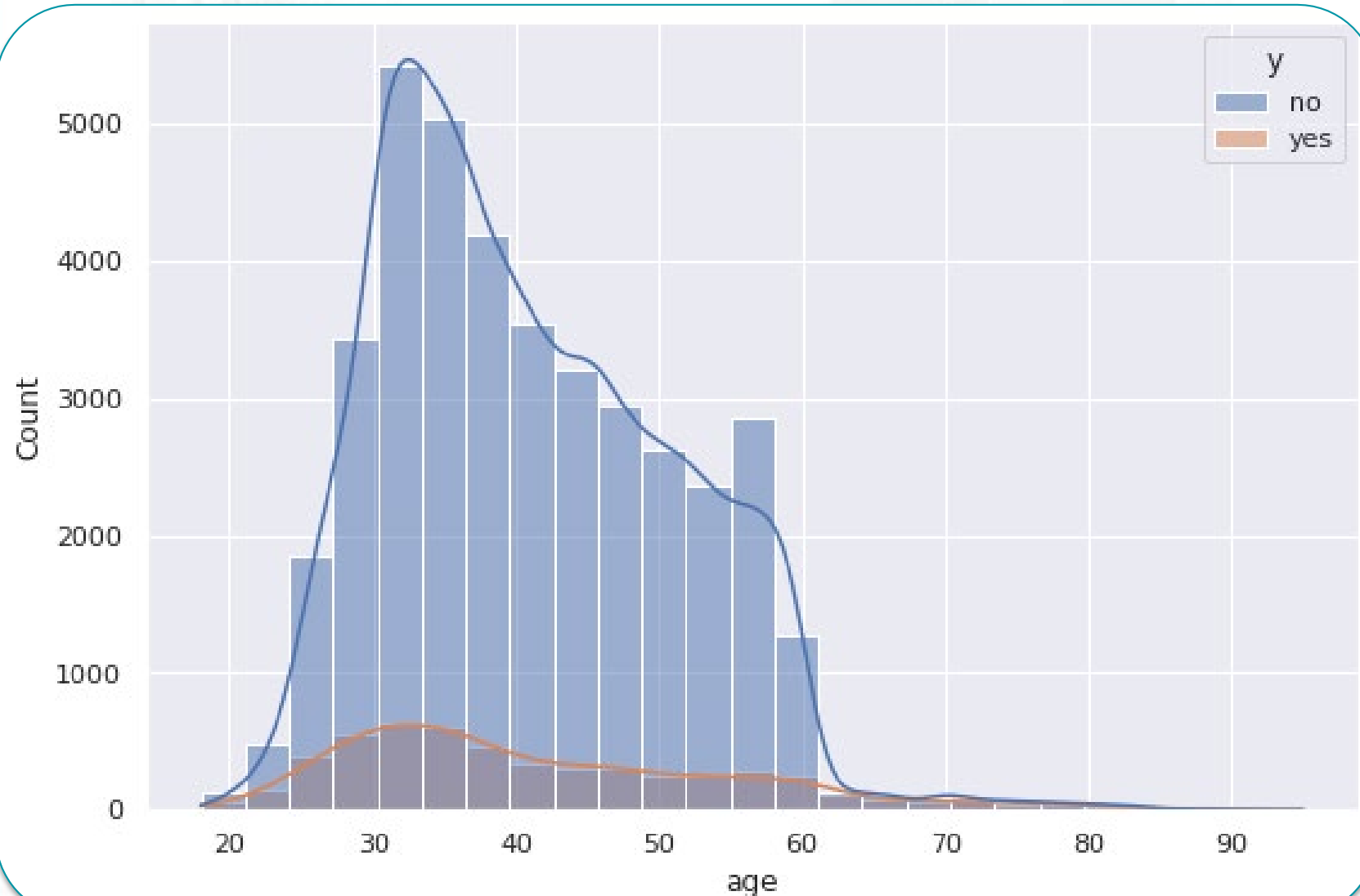
1. No : 3354
2. Yes : 1935

Heatmap



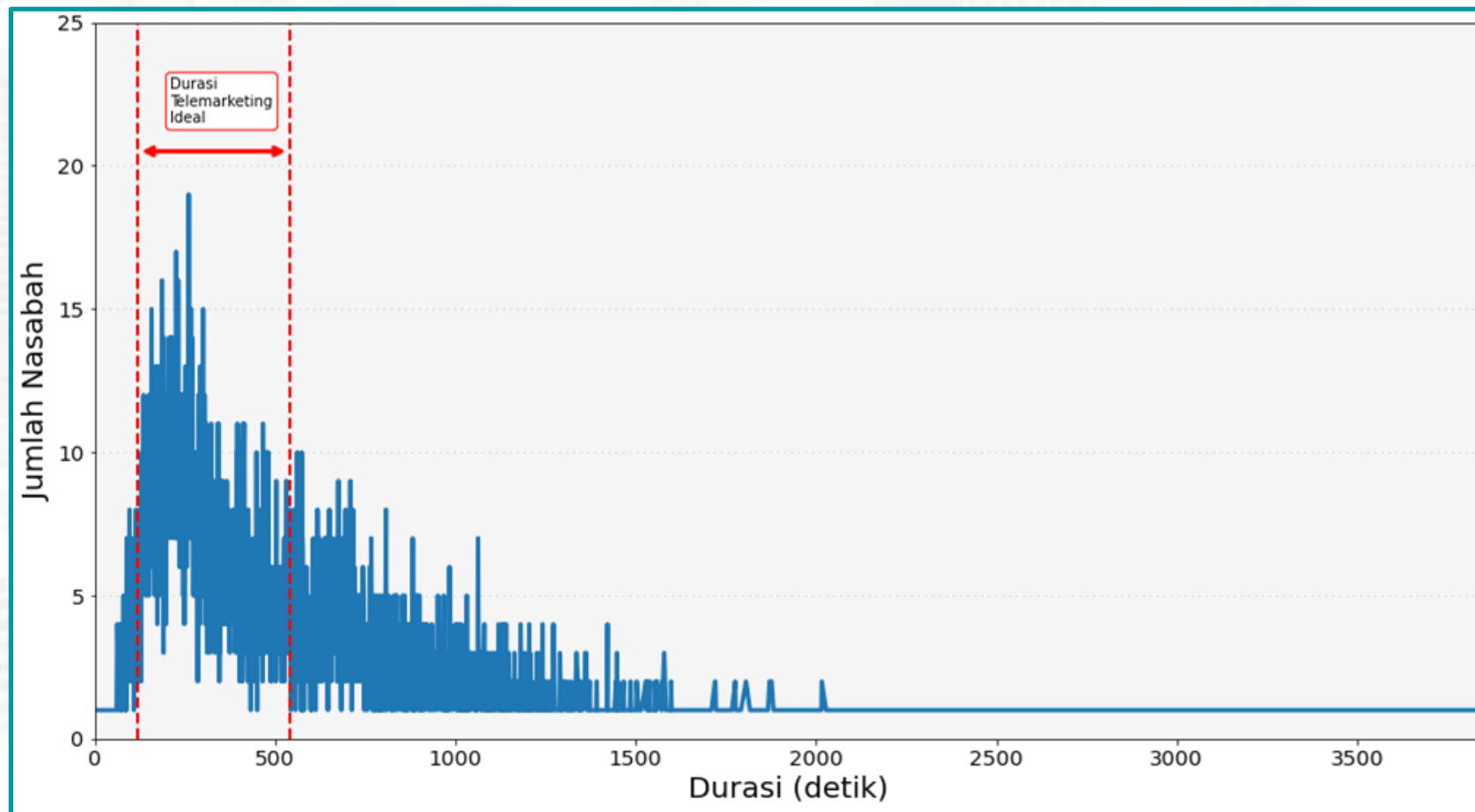
- Berdasarkan correlation heatmap dan category plots semua feature masih relevan untuk digunakan.
- Korelasi previous dan pdays cukup kuat karena bernilai 0.45. Walau begitu, semua feature akan dipertahankan untuk keperluan analisis

Business Insight



Umur nasabah yang berlangganan atau tidak berlangganan deposito menunjukkan distribusi yang hampir identik. Dapat dilihat range umur 22-62 tahun lah yang memiliki potensi untuk berlangganan deposito

Business Insight



Durasi yang cocok untuk menghubungi nasabah adalah sebanyak 2 sampai 9 menit dengan persentase 56,7%.

Dimana client banyak dihubungi melalui celular

Pre-Processing

Duplicated Data

Tidak terdapat data yang duplicate

Outliers

jumlah baris sebelum filter outlier adalah 45211 dan setelah filter outlier jumlah baris menjadi 40211

Feature Transformation

Dilakukan transform pada beberapa kolom dengan normalisasi dan standarisasi

Feature Encoding

Strategi yang kita gunakan untuk dataset banking adalah Label Encoding

Imbalance

Kami menggunakan random oversampling dengan metode SMOTE

Feature Selection

Drop Kolom (default, contact, poutcome, pdays, previous, days, month)

Feature Tambahan

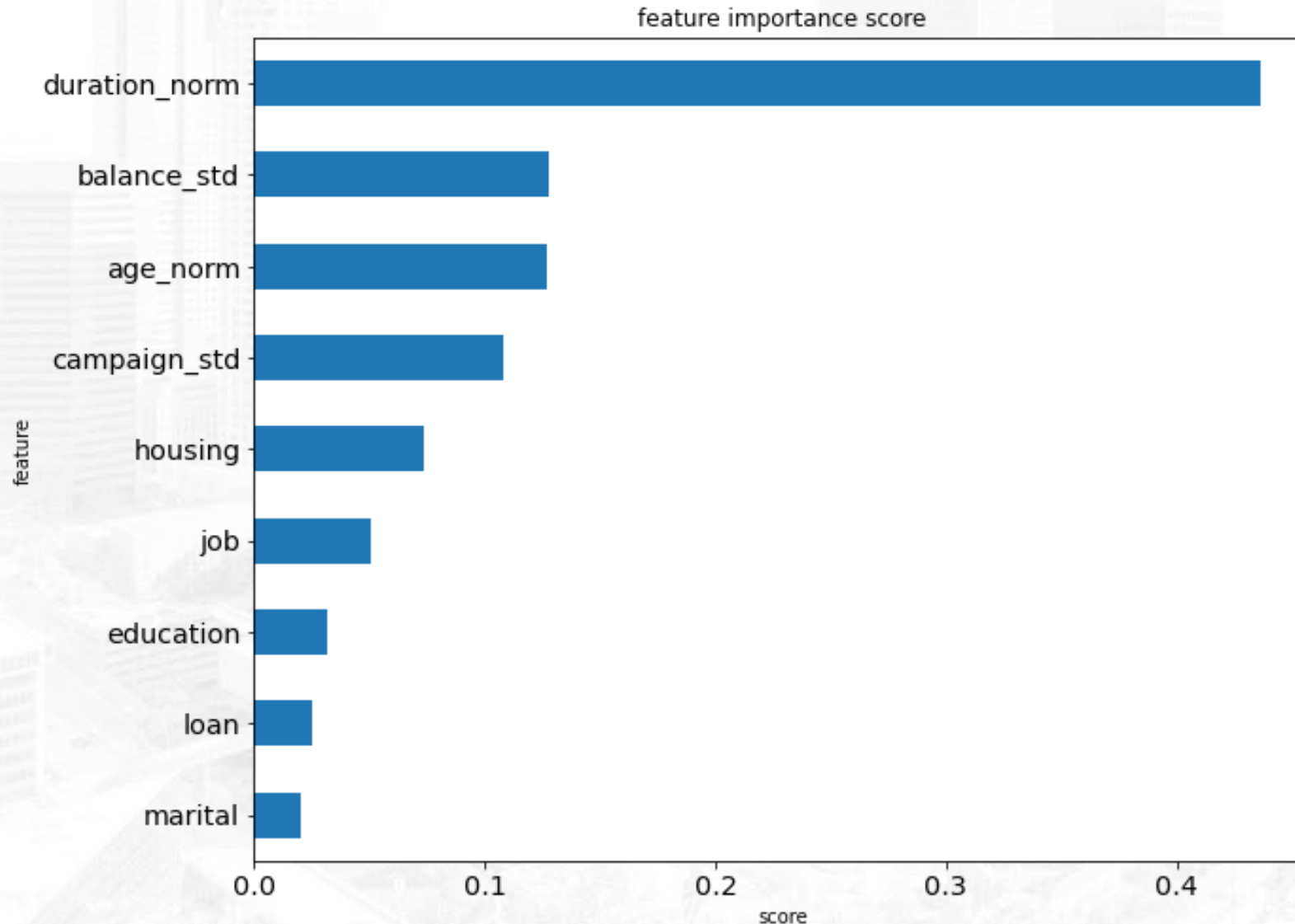
- Feature income nasabah
- Feature nasabah prioritas
- Feature email nasabah
- Feature internet banking (aktivitas nasabah)
- Feature debit / kredit nasabah

Evaluasi Model

XGBOOST		
	Train	Test
Accuracy	0.94	0.94
Precision	0.29	0.29
Recall	0.80	0.72
F1 Score	0.42	0.42
ROC_AUC	0.97	0.95

Model paling bagus adalah model XGBoost, dengan melihat nilai ROC_AUC diatas 90%

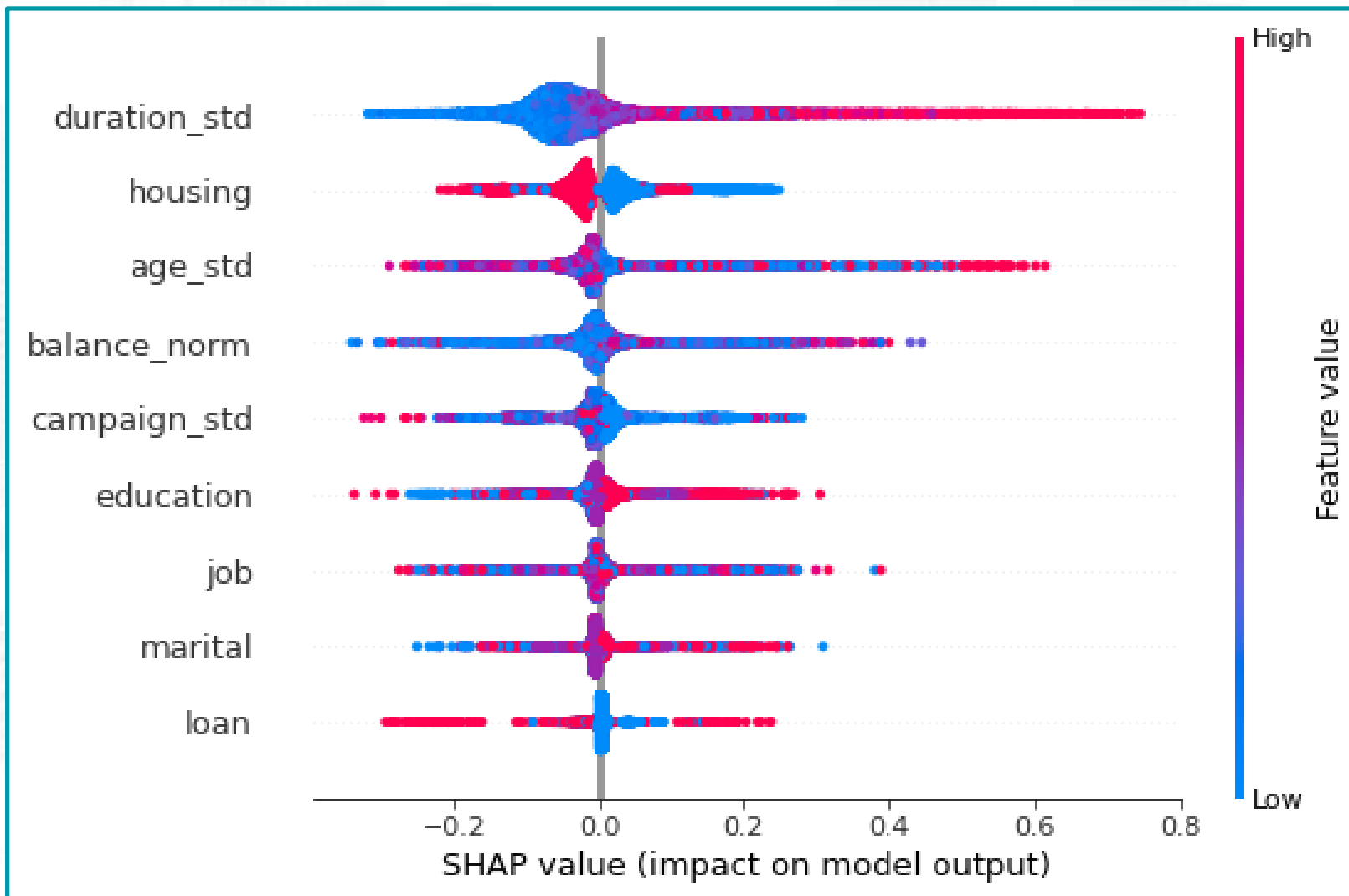
Feature Importance



Dapat disimpulkan bahwa semakin tinggi nilai score, maka semakin penting dalam memprediksi tingkat keinginan nasabah dalam berlangganan deposito. Berikut adalah lima fitur yang paling berpengaruh :

- Duration
- Balance
- Age
- Campaign
- Housing

Shap Value



Dapat disimpulkan dari Shap Value fitur yang digunakan dalam mempengaruhi nasabah berlangganan deposito :

- Duration
- Campaign
- Age
- Balance
- Housing

Recommendation

Berdasarkan Business Insight, berikut rekomendasi yang dapat diberikan:

- Bank perlu melakukan pembiayaan lebih terhadap promosi melalui cellular
- Membuat campaign produk deposito secara masif, baik melalui cellular maupun melalui website perbankan.
- Menargetkan penawaran terhadap nasabah yang berusia 22-62 tahun untuk berlangganan deposito