

Marketing Campaign

Kelompok 2 LASKAR ONLINE

Dokumen
Laporan Final
Project



Nama Anggota Kelompok

Angelus Felix Sihombing

Aisyah Raudhatuzzahra

Saip Ardo Pratama

Edhita Kristasari

Richard Noel

Marha Nur Amalina

Andri Lau

NO 1

1. Data Cleansing

Lakukan pembersihan data, sesuai yang diajarkan di kelas, seperti:

A. Handle missing values

Menghapus 24 missing values pada feature '**Income**' dengan metode penggunaan nilai median

B. Handle duplicated data

Tidak ada feature dengan data duplicate

C. Handle outliers

Handle outliers dilakukan pada features '**MntWines**', '**MntFruits**', '**MntMeatProducts**', '**MntFishProducts**', '**MntSweetProducts**', '**MntGoldProds**', '**NumDealsPurchases**', '**NumWebPurchases**', '**NumCatalogPurchases**', '**NumStorePurchases**', '**Income**', '**Year_Birth**', '**Total_Prod**', '**Total_Purchases**', dan '**Age**' dengan metode menghapus baris data outliers karena jumlah outliers tiap feature tidak banyak dan total data dihapus <20%

D. Feature transformation

- Melakukan feature transformation pada features **'MntWines'**, **'MntFruits'**, **'MntMeatProducts'**, **'MntFishProducts'**, **'MntSweetProducts'**, **'MntGoldProds'**, **'NumDealsPurchases'**, **'NumWebPurchases'**, **'NumCatalogPurchases'**, **'NumStorePurchases'**, **'Total_Prod'**, dan **'Total_Purchases'** dengan metode log transformation
- Feature **'Income'**, **'Year_Birth'**, **'Age'** dengan metode normalisasi

E. Feature encoding

- Feature **'Marital_Status'** dikelompokkan menjadi **'In Pairs'** dan **'Single'** kemudian dilakukan label encoding pada feature tersebut
- Feature **'Education'** dikelompokkan menjadi **'S1'**, **'S2'**, **'S3'** kemudian dilakukan label encoding pada feature tersebut
- Melakukan one hot encoding pada feature **'Segmentation'**

F. Handle class imbalance

Melakukan oversampling pada target **'Response'** karena ketimpangan data

2. Feature Engineering

A. Feature selection (membuang feature yang kurang relevan atau redundan)

Membuang features **'Response', 'ID', 'Recency', 'AcceptedCmp1', 'AcceptedCmp2', 'AcceptedCmp3', 'AcceptedCmp4', 'AcceptedCmp5', 'Complain', 'Year_Birth', 'Education', 'Marital_Status', 'Kidhome', 'Teenhome', 'Dt_Customer', 'Z_CostContact', 'Z_Revenue', 'Total_Purchases', 'Total_Prod', 'Education2', 'Year_Birth', 'Education', 'Marital_Status', 'Age', 'Segmentation'**

B. Feature extraction (membuat feature baru dari feature yang sudah ada)

- Membuat feature **'dependents'** yang berisi data total anak tiap user yang berasal dari features **'Kidhome'** dan **'Teenhome'**
- Membuat feature **'Age'** untuk menunjukkan usia user dengan mengurangi tahun 2022 dengan feature **'Year_Birth'**
- Membuat feature **'Segmentation'** yang berisi pengelompokan usia user dari feature **'Age'** yang berasal dari feature **'Year_Birth'**
- Membuat feature **'Total_Prod'** dengan data jumlah keseluruhan produk yang pernah dibeli oleh user, berasal dari feature **'MntWines'**, **'MntFruits'**, **'MntMeatProducts'**, **'MntFishProducts'**, **'MntSweetProducts'**, **'MntGoldProds'**
- Membuat feature **'Total_Purchases'** untuk menunjukkan jumlah pembelian yang dilakukan user melalui toko, katalog, dan website. Berasal dari feature **'NumWebPurchases'**, **'NumCatalogPurchases'**, **'NumStorePurchases'**

C. Tuliskan minimal 4 feature tambahan (selain yang sudah tersedia di dataset) yang mungkin akan sangat membantu membuat performansi model semakin bagus (ini hanya ide saja, untuk menguji kreativitas teman-teman, tidak perlu benar-benar dicari datanya dan tidak perlu diimplementasikan)

- **'Distance'** - Jarak antara tempat tinggal dengan toko. Untuk melihat kemungkinan user untuk melakukan kunjungan toko dengan frekuensi tinggi
- **'AvgMntPurchase'** - Jumlah rata-rata pembayaran yang dilakukan. Untuk melihat seberapa besar pengeluaran user untuk berbelanja
- **'Gender'** - Jenis kelamin. Untuk melihat trend berbelanja berdasarkan jenis kelamin
- **'Satisfaction'** - Tingkat kepuasan atas produk dan layanan. Untuk melihat user yang berpotensi untuk memiliki preferensi pada penjualan yang ditawarkan

3. Git

Upload project teman-teman di sebuah repository git. Berkolaborasi di Git jika ada perubahan version dari waktu ke waktu.

A. Buat Repository Git

<https://github.com/marhanura/finprokel2/tree/stage2>

B. Upload file notebook atau file pengerjaan lainnya pada repository tersebut

Link Google Drive

<https://drive.google.com/drive/u/2/folders/1ruagtNg9BZzWYbfDZmJsWwi9jqTERnlp>