



Kelompok: 2 (Analytic Adventurers)
Stage: 0 - Preparation
Mentor: Ali Murtadho
Pukul/ Tanggal: 19.00/23-Agustus-2024

Pembagian tugas di stage ini:

Nama: Nofri Rhamawati

Nama : Wikudhara Shofrina Wijaya

Nama: Wahyu Ismi Azizah

Nama : Reynaldi Silalalahi

Nama: Muhamad Rajwa AthorIQ

Nama : Muhammad Hamzah

Nama : Danu Satria Wiratama

Poin pembahasan:

1. Apa problem yang mau diselesaikan dari dataset tsb?
2. Sebagai siapa kalian pada dataset tersebut?
3. Apa goal yang mau dicapai?
4. Apa Objective yang sesuai dengan goal tsb?
5. Apa business metrics yang cocok untuk mengukur ketercapaian Objective tsb?

Hasil Diskusi:

- pada problem, fokus langsung ke tujuannya mau apa. tidak usah ditulis kembali apa yang ada di keterangan dataset.
- goal-nya juga harus lebih spesifik, bisa langsung fokus ke permasalahannya. misalnya ingin mendapatkan customer yang lebih banyak atau penjualan meningkat. intinya isinya adalah harapan dari si company tsb.
- bisa dianalisis fitur product pitched, apa yang mempengaruhi customer dalam membeli paket-paket tertentu?
- cari segmentasi customer yang paling sesuai dengan deskripsi produk yang akan launch. pada dataset dijelaskan bahwa: Customers monthly income in range of 15000- 25000, and age range 15-30, prefer 5 star properties also have higher chances of taking new package based on EDA.

Tindak Lanjut:

- Memperbaiki penentuan problem pada dataset tersebut agar fokus terhadap tujuan utama
- Menspesifikasikan goal yang ingin dicapai agar mempermudah pengerjaan kedepannya
- Menentukan objective apa saja untuk mencapai goal
- Menentukan matriks yang digunakan pada objective yang ingin dilakukan



Kelompok: Analytic Adventurers
Stage: 1 - Data Preprocessing
Mentor: Ali Murtadho
Pukul/ Tanggal: 19.00/5-September-2024

Pembagian tugas di stage ini:

Nama : Danu Satria Wiratama
: Wikudhara Shofrina Wijaya

Poin pembahasan:

1. Melakukan data cleansing yang mencakup handling missing value, handling duplicated data, handling outliers, feature transformation, feature encoding, dan handling class imbalance.
2. Melakukan feature engineering diantaranya penghapusan feature yang redundant berdasarkan visualisasi heatmap, feature extraction dengan membuat klasifikasi umur.
3. Membuat kemungkinan feature tambahan
4. Pembuatan repository pada git

Hasil Diskusi:

1. Disarankan untuk menambahkan metodologi data science untuk mencari/menggali permasalahan pada dataset.
2. Sebelum masuk ke data cleansing, kategorikan data yang penting dan tidak penting kemudian baru dilakukan data cleansing pada kolom yang penting saja.
3. Pada feature transformasi, lakukan pengecekan pada x train dan y train apakah nilai sudah cukup baik atau belum.

Tindak Lanjut:

Setelah dilakukan EDA kami memilih beberapa data yang penting dan menggunakan data tersebut untuk di proses lebih lanjut, seperti pada tahap handling outliers kami hanya membuang pada fitur yang akan kami fokus untuk gunakan ('Age', 'MonthlyIncome', 'NumberOfFollowups', 'DurationOfPitch'). Selanjutnya kami lakukan oversampling dengan teknik SMOTE untuk mengatasi ketidakseimbangan pada data train.



Kelompok: Analytic Adventurers
Stage: 2 - Exploratory Data Analysis
Mentor: Ali Murtadho
Pukul/ Tanggal: 19.00/19-September-2024

Pembagian tugas di stage ini:

Nama: Wahyu Ismi Azizah

Nama: Reynaldi Silalalahi

Nama: Nofri Rhamawati

Poin pembahasan:

1. Cara menggali lebih dalam insight dari visualisasi
2. Buat distribusi perbandingan dan dipisahkan berdasarkan kolom target
3. Melakukan correlation plot untuk melihat feature yg mengalami redundant
4. Pemilihan feature berdasarkan chi square

Hasil Diskusi:

1. Disarankan fokus terhadap beberapa feature yg dianggap penting untuk digali lebih lanjut
2. Membuat plot pendistribusian dari setiap feature yang ada
3. Disarankan membuat correlation plot untuk melihat korelasi antara feature
4. Berdasarkan hasil chi square dicocokkan juga dengan heatmap agar tidak ada feature redundat

Tindak Lanjut:

Berdasarkan hasil diskusi, langkah tindak lanjut yang disarankan adalah fokus pada fitur-fitur penting untuk analisis lebih lanjut, membuat visualisasi distribusi setiap fitur, dan menerapkan correlation plot untuk mengevaluasi hubungan antar fitur. Selain itu, hasil pemilihan fitur menggunakan Chi-Square sebaiknya dicocokkan dengan heatmap untuk menghindari redundansi, sehingga hanya mempertahankan fitur yang memberikan informasi baru



Kelompok: Analytic Adventurers
Stage: 3 - Machine Learning Modeling & Evaluation
Mentor: Ali Murtadho
Pukul/ Tanggal: 19.00/3-Oktober-2024

Pembagian tugas di stage ini:

Nama: Muhamad Rajwa

Poin pembahasan:

1. Melakukan pemilihan model terbaik yang akan digunakan
2. Melakukan Hyperparameter tuning untuk mendapatkan parameter terbaik dari model yang dipilih
3. Menentukan matriks evaluasi berdasarkan tujuan bisnis yang ingin dicapai

Hasil Diskusi:

1. Disarankan untuk mencoba lebih dari satu atau dua algoritma yang ada untuk dilakukannya pemodelan
2. Setelah mendapatkan model dengan algoritma yang kita anggap terbaik, lakukan hyperparameter untuk mendapatkan hasil yang maksimal
3. Dikarenakan tujuan utama dari permasalahan ini adalah ingin meningkatkan conversion rate pembeli, maka diputuskan menggunakan matriks recall

Tindak Lanjut:

Kami mencoba menggunakan beberapa algoritma klasifikasi, yaitu Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, AdaBoost, dan XGBoost. Berdasarkan evaluasi, model dengan performa terbaik adalah XGBoost. Namun, model tersebut menunjukkan gejala overfitting, yang ditandai dengan adanya perbedaan signifikan antara skor pada data training dan data testing. Untuk mengatasi hal ini, kami melakukan hyperparameter tuning guna mengurangi kesenjangan tersebut. Setelah dilakukan tuning, kami dapat menarik kesimpulan mengenai performa dari model machine learning yang telah dibangun.



Kelompok: Analytic Adventurers
Stage: 4 – Final Preparation
Mentor: Ali Murtadho
Pukul/ Tanggal: 19.00/17-Oktober-2024

Pembagian tugas di stage ini:

Nama: Nofri Rhamawati

Nama : Wikudhara Shofrina Wijaya

Nama: Wahyu Ismi Azizah

Nama : Reynaldi Silalalahi

Nama: Muhamad Rajwa Athoriq

Nama : Danu Satria Wiratama

Poin pembahasan:

1. Pembagian tugas
2. Penentuan moderator
3. Latihan presentasi
4. Manajemen waktu presentasi

Hasil Diskusi:

1. Dibagikan pembacaan slide presentasi berdasarkan pengerjaan Stage
2. Wahyu Ismi Azizah terpilih menjadi moderator
3. Melakukan presentasi diawasi oleh mentor
4. Diberikannya saran agar tidak banyak terbuangnya waktu saat presentasi

Tindak Lanjut:

berlatih presentasi dengan setiap anggota membacakan slide yang sudah di tentukan sebelumnya, setelah itu dilatih dengan ditanya oleh mentor agar terbiasa saat presentasi didepan juri, mengevaluasi hasil latihan presentasi tersebut