



Kelompok

: 1

Stage

: 0

Mentor

: Fadilah Nur Imani

Pukul/ Tanggal

: 20.00 WIB / January 11th 2023

Pembagian tugas di stage ini:

- | | |
|--|--|
| - Devriansyah Sya'ban : management tim, business insight & rekomendasi, management GIT | - Kun Anggiar : multivariate analysis |
| - Vira Diana : statistic descriptive, compile report & PPT | - Laurenzius Julio : multivariate analysis |
| - Farih Afdhalul : statistic descriptive | - Adinda Dita : business insight |
| - Dignu Akbar : univariate analysis | - Ramadhani A : business insight |

Poin pembahasan:

1. Point yang harus dilakukan pada stage 1
2. Melakukan univariate analysis
3. Melakukan multivariate analysis
4. Menentukan business insight

Hasil Diskusi:

1. Point yang harus dilakukan di stage 1
 - Pada stage 1 tidak perlu melakukan data cleansing dan feature engineering
 - Pada stage ini cukup melakukan data exploration dan visualisasi data mentah yang ada untuk mendapatkan insightnya. Fokus pelajari data tanpa modifikasi data.
 - Jika data sudah clean, tetap ditunjukkan saja hasil eksplorasi yang sudah dilakukan
2. Univariate Analysis
 - Sebaiknya ada 5 kolom yang dianalisis
 - Data numerik bisa dibuat boxplot atau distribution plot
 - Data kategorik bisa dibuat bar chart atau pie chart



Kelompok

: 1

Stage

: 0

Mentor

: Fadilah Nur Imani

Pukul/ Tanggal

: 20.00 WIB / January 4th 2023

3. Multivariate Analysis

- Baiknya dibuat korelasi 2 fitur
- Data kategorikal bisa dibuat scatter plot agar terlihat polanya
- Lakukan Analisa fitur vs target dan fitur vs fitur serta berikan penjelasan

4. Business Insight

- Business insight diperoleh setelah melakukan visualisasi data
- Minimal harus memberikan 3 business insight serta rekomendasinya
- Business insight harus bisa memberikan impact ke bisnis, dilihat dari performa ML nya. Minimal 50% atau standardnya 70% .
Cek juga di kaggle performa model orang lain seperti apa
- Baca artikel terkait, think out of the box, dan lakukan brainstorm. Pahami core nya dulu
- Pada business insight bisa menambahkan penawaran paket asuransi dilihat dari segi usia. Sehingga jenis penawaran yang dilakukan bisa disesuaikan dengan golongan usia
Selain itu, juga bisa diberikan insight meningkatkan revenue atau mengurangi cost

5. Tambahan Notes:

Cek duplikat dilakukan untuk mengurangi data yang sama persis, meskipun data tersebut milik user yang berbeda. Data yang sama persis dalam satu baris dengan baris lainnya hanya akan menambah waktu training machine learning dan tidak menambah performa machine learning yang dilakukan



Kelompok

: 1

Stage

: 0

Mentor

: Fadilah Nur Imani

Pukul/ Tanggal

: 20.00 WIB / January 4th 2023

Tindak Lanjut:

1. Data duplikat tidak perlu di drop pada stage ini, namun cukup ditunjukkan saja pada saat pengumpulan HW
2. Menambahkan penjelasan dari statistic deskriptif data
3. Menambahkan penjelasan dari setiap univariate dan multivariate analysis yang dilakukan