



Kelompok: UNO

Stage: 2

Mentor: Hendra

Pukul/ Tanggal: 18:45/16 Desember 2023

Pembagian tugas di stage ini:

Nama: Nur Rahman, Febby Maghfirani (Pre-Processing)

Nama: Putri Vina, Asmiyeni (Pre-Processing)

Nama : Daviro Yota Nagasan Wahyudi, Amodya Subagio, Daniel Rowin Pardamean (Feature Engineering)

Nama: Fakhry Abdurrohman (Compiler, Notulen, Git)

Poin pembahasan:

1. Pre-Processing
2. Feature Engineering
3. Git

Hasil Diskusi:

1. **Handle Missing value** : Fungsi unknown menjaga data agar data tidak bias, jobs sebaiknya tetap diunknown. Cek unknown ada patternnya atau tidak? Mungkin unknown adalah orang-orang yang belum pernah dikontak, dengan memeriksa unknown mungkin akan menemukan value baru.
2. Contact masih terlalu tinggi unknownnya yaitu 40-50% sehingga bias.
3. Outcome harus dicari polanya karena sangat tinggi sekali unknownnya.
4. Handle outliers saran jangan menggunakan IQR, coba menggunakan Zscore, Zscore karena menghapus outliers extreme tapi syarat datanya harus terdistribusi normal.

Tindak Lanjut:

1. Cari lagi pattern pada data unknown
2. Handle outliers saran jangan menggunakan IQR
3. Others didrop saja yang 5 data
4. Unknownnya tetap dikeep dan tetap pada tempatnya tidak usah direplace.



Kelompok: UNO

Stage: 2

Mentor: Hendra

Pukul/ Tanggal: 18:45/16 Desember 2023

Pembagian tugas di stage ini:

Nama: Nur Rahman, Febby Maghfirani (Pre-Processing)

Nama: Putri Vina, Asmiyeni (Pre-Processing)

Nama : Daviro Yota Nagasan Wahyudi, Amodya Subagio, Daniel Rowin Pardamean (Feature Engineering)

Nama: Fakhry Abdurrohman (Compiler, Notulen, Git)

Poin pembahasan:

1. Pre-Processing
2. Feature Engineering
3. Git

Hasil Diskusi:

1. Jika mau hapus outliers jangan dipilih pilih, sebaiknya dihapus saja karena ada algoritma yang sensitif terhadap outliers.

2. jika left skewed menggunakan kuadrat, jika right skewed menggunakan log.

Untuk feature scaling, wajib split data train dan data test baru lakukan scaling karena apabila langsung standarisasi sebelum split, scaler ini akan ngefit keseluruhan data dan merubah seluruh data test sehingga akan terjadi data leakage.

Tindak Lanjut:

- 1.. Pisahkan scaler.fit ke data train

Scaler.transform x_ data train

Scaler.fit ke data test



Kelompok: UN0

Stage: 2

Mentor: Hendra

Pukul/ Tanggal: 18:45/16 Desember 2023

Pembagian tugas di stage ini:

Nama: Nur Rahman, Febby Maghfirani (Pre-Processing)

Nama: Putri Vina, Asmiyeni (Pre-Processing)

Nama : Daviro Yota Nagasan Wahyudi, Amodya Subagio, Daniel Rowin Pardamean (Feature Engineering)

Nama: Fakhry Abdurrohman (Compiler, Notulen, Git)

Poin pembahasan:

1. Pre-Processing
2. Feature Engineering
3. Git

Hasil Diskusi:

1. Pisahkan antara numeric dan kategori, Numeric menggunakan pearson, category menggunakan kendall.
2. Cek pada month, karena bisa jadi ada insight di sana karena secara bisnis masih masuk. Apabila jelek boleh di drop, tapi apabila ada insight maka tetap dikeep

Tindak Lanjut:

- 1..