

CREDIT CARD PAYMENT DEFAULT PREDICTION

STAGE 2



Alaska Napitupulu



Rafsanjani Lestari Negara



Ike Setyawati



Ukhti Aprillia Vanny

Pre-Processing

1. Data Cleansing

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

- A. Handle missing values
- B. Handle duplicated data
- C. Handle outliers
- D. Feature transformation
- E. Feature encoding
- F. Handle class imbalance

JAWABAN

- A. Tidak terdapat missing values
- B. Tidak terdapat duplicated data
- C. Terdapat outliers, kami handle menggunakan z-score
- D. Feature transformation
 - Scaling = Standardization dan Robust (Akan kita bandingkan hasil algoritma)
 - Transformation = yeo-jhonson (Karena bisa handle skew positif maupun negatif)
- E. Feature encoding = SEX dan EDUCATION (label encod) sedangkan MARRIAGE dan PAY_0 – PAY_6 (OneHotEncod)
- F. Handle class imbalance (Z-score tetapi ada kemungkinan kita tidak handle imbalance dan focus ke F1-score karena ternyata oversampling smote cukup mempengaruhi hasil (menurunkan nilai dari algo tersebut))

STAGE 2

**CREDIT CARD
PAYMENT DEFAULT
PREDICTION**

2. Feature Engineering

Cek feature yang ada sekarang, lalu lakukan:

- A. Feature selection (membuang feature yang kurang relevan atau redundan)
- B. Feature extraction (membuat feature baru dari feature yang sudah ada)
- 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)

JAWABAN

- A. Tidak membuang features (akan mencoba semua features kemudian akan melakukan features importants sebagai perbandingan nantinya)
- B. Tidak ada
- C. Penghasilan, Pekerjaan, Outstanding (Total Tagihan Pokok), Cycle (tanggal cetak), DPD (hari keterlambatan), Current balance.

STAGE 2

**CREDIT CARD
PAYMENT DEFAULT
PREDICTION**