



Proyek Akhir KASDAD

Asisten Dosen KASDAD Ganjil 24/25

Topik



1. Company Financial Data (CF)
2. Insurance Customer Data (IC)
3. Global Terrorism Data (GT)
4. Earthquakes in US (EUS)
5. Indonesian Schools (IS)
6. College Student Performance (CSP)
7. Taxi Trip (TT)

Pembagian Topik: [Disini](#)

Dataset



1. Dataset yang digunakan dapat dilihat [disini](#).
2. Setiap topik terdapat minimal 3 dataset.
3. 2 dataset (dataset test) akan digunakan dalam pengerjaan kaggle competition.
4. Sisanya akan digunakan sebagai pengerjaan Problem Guideline yang dijelaskan pada slide [Topik Proyek Akhir](#).

*notes: masih ada beberapa topik yang belum di share datasetnya ya

Pengerjaan



1. Setiap anggota harus berkontribusi untuk setiap bagian dari proyek akhirnya.
2. Penilaian tidak hanya dari aspek pengerjaan kelompok tapi dari kontribusi pribadi juga
3. Setiap anggota wajib melakukan pengerjaan pada bagian modeling (regresi, klasifikasi dan clustering). Contoh dapat dilihat pada slide berikutnya.
4. Setiap kelompok wajib melakukan pengerjaan pada kaggle competition. Link untuk setiap kaggle competition akan diberikan kemudian.
5. Secara garis besar, setiap kelompok mengerjakan sesuai dengan Problem Guideline (EDA, Classification, Regresi, dan Clustering) dengan dataset training.
6. Model classification dan regresi yang sudah didapatkan dalam pengerjaan sebelumnya akan digunakan untuk kaggle competition dengan cara memprediksi dengan dataset testing.
7. Pada kaggle competition, setiap kelompok memungkinkan untuk bersaing dengan kelompok lain untuk mendapatkan prediksi yang lebih baik

Contoh Pengerjaan Modeling



- Contoh salah dan hanya orang yang mengerjakan yang mendapatkan nilai:
Pengerjaan dilakukan sebagai berikut
 - Anggota A mengerjakan klasifikasi
 - Anggota B mengerjakan regresi

Penilaian untuk pengerjaan seperti ini:

- Anggota A mendapatkan nilai klasifikasi dan tidak mendapatkan nilai regresi
- Anggota B mendapatkan nilai regresi dan tidak mendapatkan nilai klasifikasi

- Contoh benar dan seluruh anggota akan mendapatkan nilai:
Pengerjaan dilakukan sebagai berikut
 - Anggota A mengerjakan klasifikasi dengan Model A dan pendekatan A
 - Anggota B mengerjakan klasifikasi dengan model A dan pendekatan B
 - Anggota C mengerjakan klasifikasi dengan model B
 - Dari model yang dilakukan, didapatkan model yang terbaik untuk klasifikasi
 - Begitu pula untuk regresi dan clustering