

PENUGASAN DATA SCIENCE AND ARTIFICIAL INTELLIGENCE 2023/2024

I. Deskripsi Penugasan

Penugasan yang diberikan terdiri dari dua jenis, yaitu:

1. Praktek

Pada penugasan ini, peserta diminta mengikuti kompetisi di *platform* Kaggle. *Output* penugasan berupa submisi jawaban dan skor di platform Kaggle, submisi prediksi dengan skor terbaik di Kaggle (csv), *notebook* yang mencakup keseluruhan proses analisis data (ipynb), dan file *display name* Kaggle (txt). Penjelasan lebih lanjut mengenai kompetisi dan *output* dapat diakses di [sini](#).

2. Teori

Pada penugasan ini, akan diberikan 5 (lima) buah pertanyaan umum mengenai analisis data yang juga berhubungan dengan penugasan Praktek. *Output* penugasan berupa dokumen berisi jawaban pertanyaan dengan format pdf. Dokumen pertanyaan tersebut dapat diakses di [sini](#).

II. Ketentuan Penugasan

Ketentuan khusus untuk Penugasan Praktek, meliputi:

1. Peserta mengikuti kompetisi dengan sebelumnya mengganti *Display Name* dan *Team Name* dengan nama karakter favorit, tidak boleh dengan nama asli. (Boleh karakter di anime, movie, dsb). Panduan dapat dilihat di [sini](#).
2. Peserta dibebaskan memilih dan menggunakan *platform* dan lingkungan *programming* seperti Kaggle, Google Colaboratory, Jupyter Notebook, dll.
3. Peserta dibebaskan menggunakan *library* yang tersedia seperti Pandas, Numpy, Scikit Learn, dll. Namun, diharuskan menggunakan bahasa Python.
4. Penggunaan data eksternal **tidak diperbolehkan** dalam proses pemrosesan, modelling, dan prediksi. Namun, masih diperbolehkan apabila digunakan sebagai pendukung pada proses eksplorasi dan analisis data.
5. Peserta dibebaskan menggunakan teknik *machine learning* maupun *deep learning* apapun, tetapi dilarang melakukan prediksi secara manual (contoh: menentukan nilai target secara manual, mengambil data target dari sumber lain, dan sebagainya).
6. Peserta mendapatkan kesempatan men-*submit* prediksi di Kaggle sebanyak 5 (lima) kali dalam 1 (satu) hari.
7. Batas waktu pengerjaan penugasan praktek di Kaggle akan mengikuti batas waktu penugasan *Open Recruitment*.
8. Notebook yang dikirimkan sebagai output dapat dijalankan kembali dengan baik (tidak *error*).

Secara umum, ketentuan keseluruhan penugasan meliputi:

1. Peserta mengerjakan keseluruhan penugasan secara individu, sesuai dengan kemampuan masing-masing.
2. Peserta mengirim file txt berisi *Display Name* Kaggle (nama karakter favorit) bersamaan dengan keseluruhan penugasan.
3. Apabila terdapat penjelasan tambahan mengenai jawaban penugasan (link eksternal, notes, dsb) dapat dituliskan sekaligus pada file txt *Display Name* Kaggle. Contoh file dapat dilihat di [sini](#).
4. Keseluruhan jawaban penugasan dikumpulkan dalam folder zip, yang berisi:
 - Submisi prediksi terbaik (csv).
 - *Notebook* (ipynb).
 - Jawaban Penugasan Teori (pdf).
 - *Display Name* Kaggle (txt).

III. Matriks Penilaian

Aspek Penilaian	Bobot
Skor <i>Leaderboard</i> Kaggle*	15%
<i>Exploratory Data Analysis</i> (EDA)**	25%
<i>Data Preprocessing</i> **	15%
<i>Modelling</i> **	20%
Kerapian <i>Notebook</i> **	10%
Jawaban Penugasan Teori	15%

* Skor setelah kompetisi berakhir dengan 100% *public score*.

** Berdasarkan *notebook*