

Tugas Akhir

Buatlah sebuah Analisis Media Sosial dengan topik tertentu (kalian tentukan topiknya) menggunakan metode data sains yang telah dipelajari sebelumnya. Ketentuannya adalah sebagai berikut:

- Kumpulkan dataset dari dua sumber:
 - twitter minimal dari 4 hari kebelakang hingga hari ini. (Jika diumumkan hari Rabu, maka diambil data mulai minggu. (Atau ambil minimal 4 hari)
 - Menggunakan sumber data tambahan dengan cara melakukan webscraping dari situs tertentu (imdb, berita, Wikipedia, dsb).
- Simpan semua data mentah ke dalam csv(misal tweets_buruh.csv atau news_buruh_csv)
- Lakukan Analisis pada data tersebut. (Topic Analysis, Sentiment Analysis dan lainnya)
- Kerjakan semua langkah kerja kalian dalam berkas berformat ipynb. Jalankan berkasnya dan simpan dalam keadaan output muncul.
- Pengerjaan Harus mengikuti kaidah metode untuk data sains, Jadi harus ada 6 komponen utama, yaitu:
 - Business Understanding, buat sel yang menjelaskan kenapa kalian memilih topik tersebut. Apa yang berusaha kalian cari, apa tujuan kalian
 - Data Collection, proses pengumpulan data dan memasukkan ke dalam format yang kalian inginkan agar bisa dilanjutkan ke dalam tahap selanjutnya
 - Data Preparation, Lakukan Text-Preprocessing, EDA, dan langkah pembersihan dan persiapan lainnya.

- Data Processing, Lakukan proses-proses kalian disini, misal sentiment analysis, topic modeling, network analysis dsb.
 - Data Anaysis, Analysis hasil dari proses kalian. Jawab pertanyaan pada poin pertama
 - Share Insight, adalah saat kalian membuat tugas ini dalam bentuk berkas ipynb. Jadi bisa diubah lebih menarik misal warna plotnya, tabelnya, dsb.
- Penilaian akan berdasar pada langkah-langkah pengerjaan yang disebutkan di atas, apakah sudah terlihat alurnya dengan jelas atau belum.
- Komponen penilaian lainnya adalah dari keberjalanan kode, kerapihan kode, dan hasil analisa.
- Masukkan berkas ipynb tersebut kedalam satu folder, dengan ketentuan:
 - Beri nama foldernya, proyek akhir
 - Beri nama file ipynb nya, <email_depan>_proyekakhir.ipynb
 - Masukkan semua berkas yang dipakai untuk menjalankan file kedalam folder tersebut (yang mana bila tidak ada berkas tersebut file tidak akan berjalan sempurna, misalnya dataset, csv, dsb)
 - File token.csv(token twitter) tidak perlu dimasukkan
- Buat serapih mungkin, anggap pemeriksa adalah orang awam yang ingin membaca laporan, dan tidak ada kendala (error) saat menjalankan file ipynbnya.
- Unggah folder ke dalam google drive, salin linknya, dan masukkan ke dalam tempat pengumpulan pada menu quiz
- Pengumpulan dibuka pada hari kamis 00.00 WIB - Minggu 12.00 WIB

- Sistem pengumpulan seperti pada kuis biasa, jadi setelah kuis dijalankan maka peserta punya waktu 24 jam sebelum link ditutup. Maksimal minggu jam 12.00 WIB akan tertutup.
- Jangan lupa untuk memperbaiki pengaturan view dan sharenya.

Beberapa referensi:

<https://pers.droneempit.id/ui-dalam-trending-topik-twitter/>

<https://www.kaggle.com/karnikakapoor/customer-segmentation-clustering>

<https://www.kaggle.com/sonalisingh1411/customer-personality-analysis-segmentation>