



Tugas#3 Mini Project Part 1

➤ course	<u>Text mining</u>
📅 Date	@June 22, 2024
⚙️ progress	not started



Instructions

Deadline : Sabtu, 22 Juli 2024 pukul 23:59 WITA

Keterlambatan nilai = 0 (tidak menerima alasan apapun)

- Bentuk tim yang terdiri dari **3 orang** kemudian isi gsheets berikut dengan memilih topik mini project pada link berikut:
https://docs.google.com/spreadsheets/d/1RHTPtklwMoGZ0_H2WQ_Wh-b09rKM7Z3XiwHloaGo-jM/edit#gid=0
(perhatikan tab sesuaikan dengan kelas)
- Buat dataset dengan cara **crawling** (WAJIB) dari portal berita Indonesia untuk topik Klasifikasi Berita; Atau **crawling** (WAJIB) dari google playstore/twitter berbahasa Indonesia untuk topik Analisis Sentimen Komentar. Crawling wajib menggunakan python pada Google Collab.
- Minimal data terdiri dari **50 record** untuk topik Klasifikasi Berita; Dan **300 record** untuk Analisis Sentimen Komentar
- 1 record berisi konten isi berita untuk topik Klasifikasi Berita; Atau 1 record berisi 1 komentar netizen pada google playstore/twitter
- Kelas cukup terdiri dari 2 saja misal untuk kasus berita "Teknologi" dan "Ekonomi", sedangkan untuk analisis sentimen "Positif" dan "Negatif". Kelas dapat dibuat secara manual atau otomatis by crawling dengan mengambil

kategori berita untuk topik Klasifikasi Berita; Dan kelas dibuat secara manual untuk Analisis Sentimen Komentar

- Sehingga dataset hanya terdiri dari 2 kolom: content dan class
- Buat laporan dalam bentuk jurnal sesuai template, setidaknya hingga bagian **2. METODE**. Template terlampir
- Pada bagian 2 WAJIB terdapat sub bagian **2.1 Data, 2.2 Proses Klasifikasi** yang berisi flowchart proses, + sub bab lain yang dianggap perlu

Pengumpulan:

- File .zip/.rar dengan format: Kelas_No Kelompok.zip/.rar yang berisi1. Laporan **.docx**2. **.txt** yang berisi link colab crawling data3. **.csv** hasil crawling
- Dikumpulkan oleh ketua kelompok saja

Ket :

- Ada nilai bonus +10 bagi yang memilih topik Klasifikasi Berita
- Pemilihan topik terbatas sesuai gsheets yang tersedia, siapa cepat dia dapat
- Template terlampir



Execution



Files