

## PRAKTIKUM DATA MINING

Nama : Hawa Andini Hadi  
NIM : 23051214238  
Kelas/Angkatan : DTMG1/2023  
Algoritma : Transformers - RoBERTa (w11wo/indonesian-roberta-base-sentiment-classifier)  
Jenis Analisis : Analisis Sentimen Topik "FOMO" di Twitter  
Dataset : Dataset diambil dari Twitter dengan menggunakan kata kunci pencarian:  
fomo since:2024-04-01 until:2025-04-01 lang:id  
Jumlah data yang berhasil diambil adalah 100 tweet berbahasa Indonesia.

### Keterangan Dataset :

Dataset terdiri dari kolom full\_text (isi tweet) dan sentiment (hasil analisis sentimen). Sentimen diklasifikasikan menjadi tiga kategori: negatif, netral, dan positif, menggunakan model "indonesian-roberta-base-sentiment-classifier" dari HuggingFace.

### Pembahasan :

Analisis ini bertujuan untuk mengetahui bagaimana orang-orang di Twitter membicarakan topik FOMO (takut ketinggalan) selama periode 1 April 2024 sampai 1 April 2025. FOMO biasanya terjadi saat seseorang merasa takut tidak ikut dalam sesuatu yang sedang ramai dibicarakan, terutama di media sosial.

Untuk itu, dilakukan analisis sentimen agar bisa tahu apakah tweet-tweet tersebut bernada positif, netral, atau negatif. Analisis ini dilakukan secara otomatis menggunakan alat kecerdasan buatan (model bahasa Indonesia dari RoBERTa).

### Metode Preprocessing :

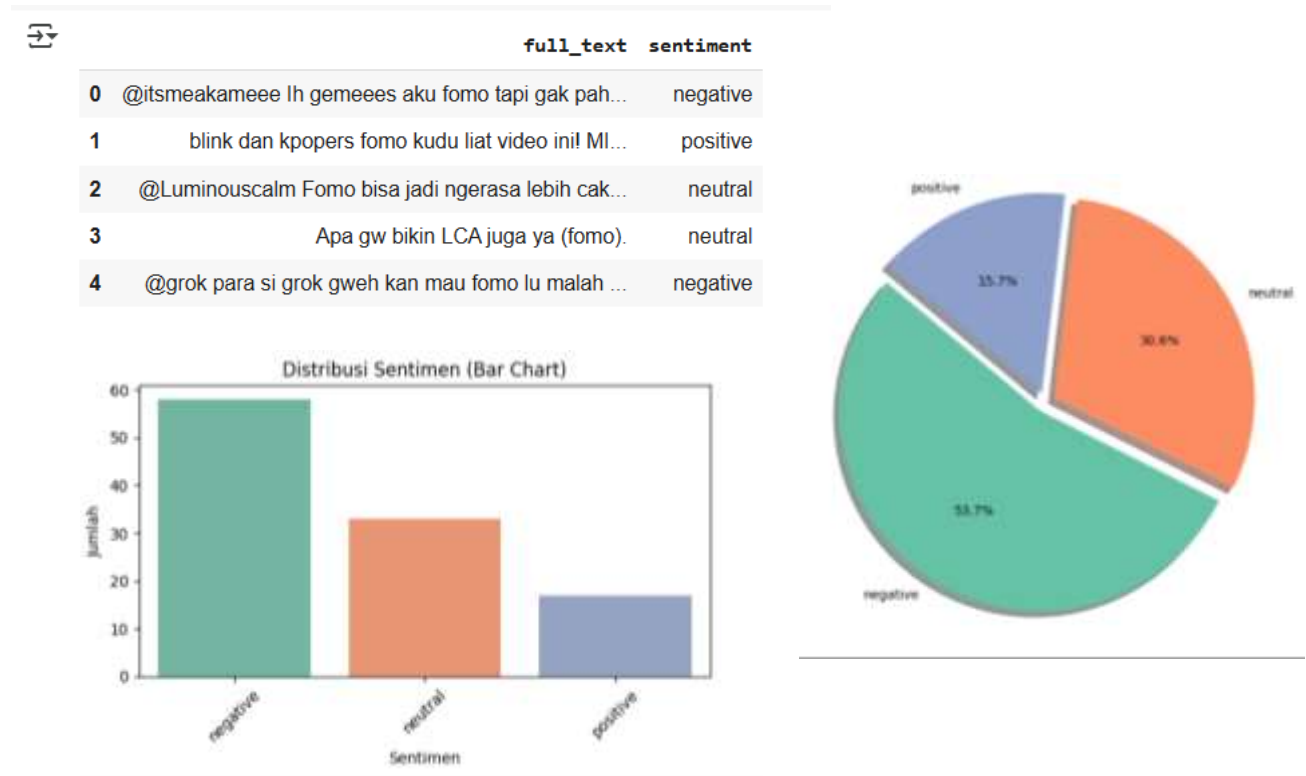
Teks dari tweet tidak dibersihkan atau diproses secara manual (seperti menghapus simbol, mengubah huruf jadi kecil, dll). Hal ini karena pipeline yang digunakan dari Hugging Face sudah otomatis menangani proses ini di dalamnya. Model secara langsung membaca teks asli dari tweet dan memprosesnya dengan cara yang sudah disesuaikan.

### Metode Evaluasi Model :

Dalam proses ini tidak dilakukan pengujian seberapa akurat model. Tidak tersedia data pembandingan (label asli) untuk mengetahui apakah hasil klasifikasi model benar atau salah. Karena

itu, tidak dihitung nilai akurasi atau ukuran performa lain seperti precision dan recall. Evaluasi dilakukan hanya secara visual, lewat grafik, untuk melihat distribusi hasil sentimen yang dihasilkan model.

Hasil Evaluasi :



Analisis sentimen terhadap teks terkait "fomo" menunjukkan distribusi yang tidak merata di antara tiga kategori sentimen: negatif, netral, dan positif.

- Sentimen Negatif mendominasi dengan 53.7% dari total data. Ini menunjukkan bahwa mayoritas percakapan terkait topik "fomo" mengandung perasaan negatif seperti kecemasan, iri, atau ketidaknyamanan.
- Sentimen Netral berada di posisi kedua dengan 30.6%, yang berarti sebagian besar teks bersifat informatif atau tidak menunjukkan emosi yang kuat.
- Sentimen Positif merupakan yang paling sedikit, hanya sebesar 15.7%, menunjukkan bahwa ekspresi positif terhadap "fomo" relatif jarang.

Visualisasi berupa bar chart dan pie chart secara konsisten menggambarkan dominasi sentimen negatif dalam data ini.

Kesimpulan :

1. Pengumpulan data tweet dengan topik “FOMO” dari Twitter berhasil dilakukan, dengan total 100 tweet berbahasa Indonesia yang berhasil dikumpulkan.
2. Model RoBERTa berbahasa Indonesia digunakan untuk mengidentifikasi sentimen dari setiap tweet, dan hasilnya diklasifikasikan ke dalam tiga kategori utama: positif, negatif, dan netral.
3. Hasil klasifikasi tersebut kemudian divisualisasikan menggunakan grafik batang dan diagram lingkaran untuk memberikan gambaran yang lebih jelas mengenai persebaran sentimen.
4. Karena tidak tersedia data yang sudah dilabeli sebelumnya, tidak dilakukan pengujian akurasi model, sehingga kinerja model tidak bisa diukur secara kuantitatif.
5. Analisis ini bersifat eksploratif dan memberikan wawasan awal mengenai bagaimana pengguna Twitter merespons isu FOMO.

Lampiran:

1. Link data set dan google colab:  
<https://colab.research.google.com/drive/1ABbEG-ohb2TsfzxG6K9GOziEXE9zVJDp>