

EKSTRAKSI ARAH SENTIMEN TWITTER MENGGUNAKAN PENDEKATAN KUALITATIF DESKRIPTIF

Adi Kurniawan DCS Wave 13



Pendahuluan

Twitter telah menjadi platform utama di mana pengguna dapat berbagi pemikiran, pendapat, dan perasaan mereka secara real-time. Oleh karena itu, menganalisis sentimen dalam data Twitter dapat memberikan wawasan berharga tentang bagaimana masyarakat merespons berbagai kondisi.

Pendekatan kualitatif dalam ekstraksi sentimen bertujuan untuk menemukan kata-kata dan kalimat yang digunakan oleh pengguna Twitter. Dengan menggunakan metode kualitatif deskriptif untuk menggambarkan berbagai jenis sentimen yang muncul dalam data Twitter, mulai dari positif, negatif, hingga netral.

Pada analisis data twitter kali ini akan mengetahui sentimen dengan cara mengekstraksi banyaknya kata *abusive* yang muncul dan seberapa banyak kata yang mengandung arti tidak baku (alay). Setelah mengetahui data (kata abusive dan kata alay) tersebut akan di ganti dengan kata yang baku dan mensensor kata *abusive* tersebut. Setelah mengganti kata-kata tersebut akan dibandingkan dengan data yang sudah dibersihkan, sehingga dapat mengetahui arah sentiment dari data twitter tersebut.



Metodologi Penelitian

01

Proses Code Function

Proses Code Function Cleansing yang dapat menjawab latar belakang permasalahan

02

Metode Statistika dan EDA

Metoda statistika dan EDA yang dapat menjawab latar belakang permasalahan

03

Visualisasi

Proses visualisasi yang dapat menjawab latar belakang permasalahan





01

Process Code Function Cleansing

- **Import Library yang dibutuhkan:** Library dibutuhkan agar fungsi atau metode yang digunakan berjalan dengan semestinya
- **Memebangun FLASK dan SWAGGER:** Merupakan frame unutk API menampilkan proses dan fusngi pemanggilan dan menamilkan hasil prosessing celansing data selain tersimpan di Data base.
- **Import data:** Mengimport data Abusive.csv dan new_kamusalay ke data frame Pandas
- **Cleansing data:** Menghilangkan duplikasi Tweet dan mengecek apakah pada kolom Tweet ada yang kosong (NaN)

01

Process Code Function Cleansing

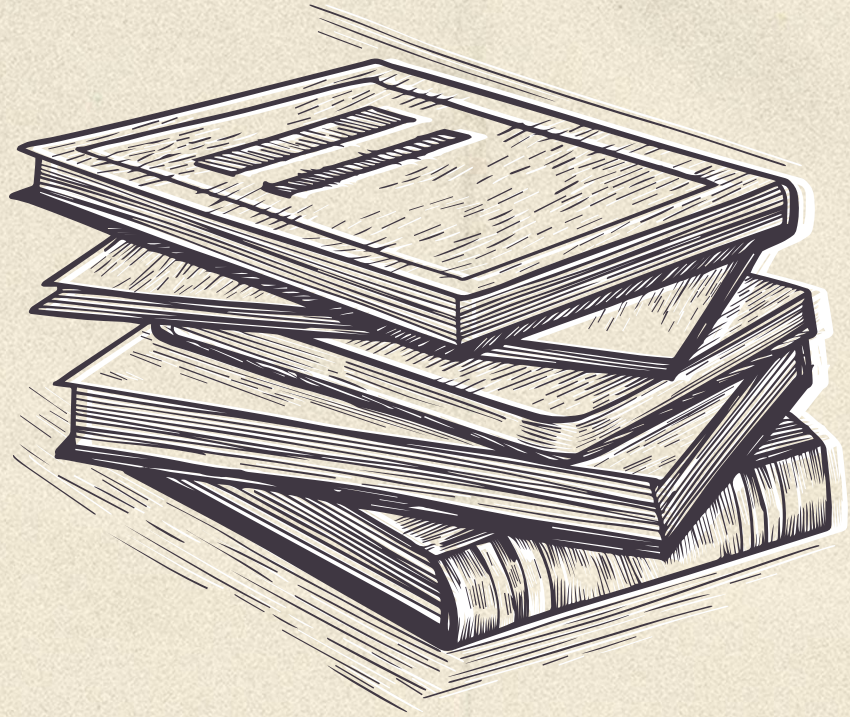
- **Membuat End Point:** Membangun 3 endpoint : **Basic GET, Masukan Text, dan Masukan File.** **Basic Get** digunakan untuk mengecek **Swagger UI** berjalan dengan baik, **Masukan Text** sebagai API yang digunakan untuk meng-*cleansing* data berupa masukan Text. **Masukan File** sebagai API yang akan digunakan untuk meng-*cleansing* data berupa file.
- **Membuat Fungsi:** Fungsi yang dibangun adalah fungsi *Cleansing data*, fungsi *normalize kata* , fungsi *sensor kata-kata*, dan fungsi *processing kata-kata*
- **Membangun database:** Data base dibangun untuk menampilkan perbandingan hasil *processing* data twitter yang telah diproses dengan data mentahan. Di dalam database akan dibuat kolom sesuai dengan yang dibutuhkan.
- **Visualisasi:** berfungsi untuk mendeskripsikan hasil dari proses *cleansing* secara mudah, sehingga dapat mengetahui kata yang sering dipakai untuk menarik kesimpulan sentimen

01

Process Code Function Cleansing

02

Metode Statistika



[illegible]

03

Visualisasi

Kesimpulan



Dari hasil proses yang telah dilakukan dapat disimpulkan bahwa arah sentimen dari data twitter ini adalah cenderung **negative**, dibuktikan dengan kata **disensor** yang mendominasi visualisasi dari wordcloud. Seperti yang telah dijelaskan bahwa kata disensor adalah hasil dari daftar kata *abusive* yang telah di-*cleansing*

