

ANALISIS SENTIMEN MEDIA

X

BERDASARKAN KATA KASAR

Dalam era digital ini, di mana informasi tersebar luas melalui berbagai platform online, pemahaman terhadap sentimen menjadi semakin krusial. Bisnis, organisasi, dan peneliti sering kali menggunakan analisis sentimen untuk mendapatkan wawasan tentang persepsi masyarakat terhadap merek, produk, atau isu tertentu.

Dengan teknologi pemrosesan bahasa alami (NLP) dan kecerdasan buatan, kita dapat mengimplementasikan algoritma dan model yang dapat menganalisis dan mengklasifikasikan teks berdasarkan emosinya, seperti positif, negatif, atau netral. Hal ini memungkinkan pemahaman yang lebih baik terhadap umpan balik pengguna dan membantu pengambilan keputusan yang lebih tepat.

Rumusan Masalah:

- Apakah kata kasar akan selalu berkaitan dengan sentimen negatif?
- Apakah kata kasar akan selalu berkaitan dengan sentimen positif?
- Apakah kata kasar akan selalu berkaitan dengan sentimen netral?

Tujuan Masalah:

- Mengetahui kata kasar akan selalu berkaitan dengan sentimen negatif
- Mengetahui kata kasar akan selalu berkaitan dengan sentimen positif
- Mengetahui kata kasar akan selalu berkaitan dengan sentimen netral

Dalam menghadapi pembersihan teks atau cleaning teks dari data yang dimiliki ketika dalam pemilihan untuk membersihkan teks harus sesuai dengan yang data dimiliki seperti contoh ketika memiliki data teks dari twitter harus memiliki proses pembersihan data yang mengandung kata rt karena di twitter ada fungsi untuk meretweet dari hasil cuitan orang lain

Dalam pemilihan metode statistika untuk menjawab latar belakang permasalahan metode yang saya gunakan adalah menggunakan package text blob untuk mengetahui sentimen yang muncul di masyarakat terkait cuitan yang ada. TextBlob adalah pustaka Python yang menyediakan antarmuka sederhana untuk melakukan tugas pemrosesan bahasa alami (NLP). Pustaka ini dibangun di atas pustaka Natural Language Toolkit (NLTK) dan menyediakan antarmuka yang mudah digunakan untuk melakukan berbagai tugas seperti analisis sentimen, pemrosesan teks, dan ekstraksi informasi.

Dalam pemilihan metode statistika untuk menjawab latar belakang permasalahan metode yang saya gunakan adalah menggunakan package text blob untuk mengetahui sentimen yang muncul di masyarakat terkait cuitan yang ada. TextBlob adalah pustaka Python yang menyediakan antarmuka sederhana untuk melakukan tugas pemrosesan bahasa alami (NLP). Pustaka ini dibangun di atas pustaka Natural Language Toolkit (NLTK) dan menyediakan antarmuka yang mudah digunakan untuk melakukan berbagai tugas seperti analisis sentimen, pemrosesan teks, dan ekstraksi informasi.

Dalam menggunakan metode untuk mencari hubungan antar dua data kategorikal yaitu data sentimen dengan data kata kasar, maka metode yang digunakan adalah dengan uji chi square.

Uji Chi-Square, atau Chi-Square Test, adalah metode statistik yang digunakan untuk menentukan apakah ada hubungan signifikan antara dua variabel kategorikal. Ini dapat digunakan untuk menguji apakah distribusi frekuensi observasi antar kategori variabel berbeda dari yang diharapkan secara acak

Berdasarkan hasil dan analisis yang sudah dilakukan terdapat 3 sentimen dasar yang didapatkan yaitu netral, negatif, dan juga positif didapatkan hasil berupa 647 sentimen positif, 12061 sentimen netral, 461 sentimen negatif dengan cuitan yang mengandung kata kasar berjumlah 4977 dan cuitan yang mengandung kata tidak kasar berjumlah 8192

Berdasarkan uji korelasi menggunakan metode chi square didapatkan hasil P value untuk uji chi square sebanyak $2.655708595468085 \times 10^{-16}$, yang dimana berarti berdasarkan hasil uji chi square dengan menggunakan taraf signifikansi sebanyak 95% atau dengan alpha sebanyak 5% maka didapatkan hasil yaitu terdapat hubungan antara sentimen dan kata kasar