

Abstraksi

Raihan Suwardhana Arisyi. 55419267

IMPLEMENTASI NLP DENGAN PEMODELAN NLTK NAÏVE BAYES UNTUK KLASIFIKASI SENTIMEN ANALISIS PADA TEKS BERBAHASA INGGRIS BER-BASIS WEB.

Skripsi, Fakultas Teknologi Industri, Jurusan Informatika, Universitas Gunadarma, 2023.

Kata Kunci : NLP, Analisis Sentimen, Klasifikasi, Python, Artificial Intelligence, Library, Social media, Streamlit

(XIV+ 102+ lampiran)

Analisis sentimen telah menjadi pendekatan penting dalam memahami respons dan opini publik terhadap berbagai hal, seperti produk, layanan, atau isu-isu tertentu. Dalam era digital dan media sosial, data teks yang dihasilkan oleh pengguna media sosial memiliki potensi besar sebagai sumber informasi untuk analisis sentimen. Penelitian ini bertujuan untuk mengembangkan model analisis sentimen menggunakan metode *Naïve Bayes* pada data teks dari media sosial. Pada tahap pengembangan, data teks diambil dari dataset *twitter_samples* yang disediakan oleh *NLTK* (*Natural Language Toolkit*). Penerapan *NLP* berupa teks *processing* dilakukan, termasuk *cleaning* (menghilangkan tanda baca, angka, simbol), *tokenization* (pemisahan kata pada kalimat), *stopword removal* (menghilangkan kata yang tidak memiliki kata terlalu berarti), dan *stemming* (perubahan kata menjadi kata dasar bahasa inggris). Setelah itu, model klasifikasi *Naïve Bayes* dilatih menggunakan data latih yang telah diproses. Hasil dari penelitian ini mencakup pengembangan model analisis sentimen yang dapat mengklasifikasikan teks komentar berbahasa inggris dari media sosial sebagai sentimen positif atau negatif . Hasil penelitian ini menunjukkan bahwa algoritma *Naïve Bayes* dapat digunakan untuk mengklasifikasi sentimen pada teks bahasa inggris dengan menggunakan rasio pembagian data latih 80% dan data tes 20% memiliki akurasi sebesar \pm 79%-80% pada dataset yang digunakan. Hasil penelitian ini adalah berhasil mengimplementasikan *NLP* pada aplikasi analisis sentimen pada teks berbahasa inggris dengan menggunakan algoritma *naïve bayes*.

Daftar Pustaka (2018-2023)