

ABSTRAK

Karel Fadhillah. 10120567

ANALISIS SENTIMEN TRANSPORTASI WHOOSH MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE

Skripsi Jurusan Sistem Informasi. Fakultas Ilmu Komputer dan Teknologi Informasi. Universitas Gunadarma. 2024

Kata kunci : Analisis Sentimen, Naïve Bayes Classifier, Support Vector Machine, Twitter, WHOOSH.

(xiv + 80 + Lampiran)

Transportasi merupakan elemen krusial bagi kota besar, dan kereta api adalah salah satu solusi penting untuk negara berkembang seperti Indonesia, yang mengandalkan kapasitas besar untuk mempermudah mobilitas orang dan barang. Proyek pembangunan kereta cepat (*High Speed Train, HST*) di Indonesia kini dioperasikan dengan nama WHOOSH sejak 2 Oktober 2023. Kontroversi mengenai kebutuhan, harga tiket, dan pelayanan WHOOSH memunculkan pentingnya analisis sentimen di media sosial, khususnya Twitter, untuk memahami persepsi masyarakat. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap kereta cepat Jakarta-Bandung menggunakan dua algoritma *machine learning*, yaitu *Naïve Bayes Classifier* dan *Support Vector Machine* (SVM), serta membandingkan tingkat akurasinya. Metode penelitian meliputi pengumpulan data, *preprocessing*, dan pemrosesan data yang mencakup *labeling*, *balancing* (*oversampling SMOTE*), pembagian data (*training* dan *testing*), serta inisialisasi algoritma. Hasil analisis menunjukkan bahwa sentimen masyarakat bervariasi dengan mayoritas sentimen negatif dan netral. Model *Naïve Bayes* mengklasifikasikan data uji sebagai negatif (41,7% atau 489 data), netral (32,8% atau 385 data), dan positif (25,4% atau 298 data). Sebaliknya, SVM memberikan distribusi yang lebih seimbang dengan negatif (36,2% atau 424 data), netral (35,0% atau 410 data), dan positif (28,8% atau 338 data). SVM menunjukkan performa yang superior dengan akurasi 88% dan *F1-score* tinggi pada ketiga kelas sentimen, sedangkan *Naïve Bayes* mencapai akurasi 81% dengan *F1-score* lebih rendah pada kelas positif. Visualisasi *wordcloud* mengungkapkan dominasi sentimen negatif, dengan kata-kata seperti "mahal" dan "biaya" sering muncul, mengindikasikan keluhan masyarakat terkait harga tiket dan pengeluaran. Sentimen positif didominasi oleh kata-kata seperti "cepat" dan "nyaman". Hasil ini memberikan wawasan penting bagi pengelola dalam menangani kekhawatiran publik dan meningkatkan kualitas layanan serta efisiensi biaya.

Daftar Pustaka (2019-2023)