

ABSTRAK

Pembimbing: (1) Moh. Dasuki, M.Kom.; (2) Wiwik Suharso, S.Kom., M.Kom.

Transformasi teknologi informasi mendorong pemanfaatan *Natural Language Processing* (NLP) dalam analisis sentimen untuk menggali opini pengguna terhadap layanan berbasis aplikasi. Penelitian ini bertujuan menganalisis sentimen ulasan pengguna aplikasi ojek daring ACI di *Google Play Store* menggunakan metode *N-gram* dan algoritma *Support Vector Machine* (SVM). Data penelitian terdiri atas 1.419 ulasan yang dikumpulkan pada periode 2023–2025. Setelah melalui tahap *preprocessing*, jumlah data berkurang menjadi 1.285 ulasan dengan menghapus teks yang memiliki kurang dari tiga kata. Pelabelan sentimen dilakukan menggunakan pendekatan *lexicon-based* dengan kamus *sentiwords_id*. Data dengan sentimen netral dihapus, sehingga diperoleh 239 data akhir yang terbagi ke dalam kelas positif dan negatif. Ekstraksi fitur dilakukan menggunakan skema *unigram*, *unigram + bigram*, dan *unigram + trigram* dengan pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF). Proses klasifikasi dilakukan menggunakan SVM linear dengan pembagian data latih dan data uji sebesar 80:20. Hasil pengujian menunjukkan bahwa kombinasi *unigram + bigram* menghasilkan akurasi tertinggi sebesar 96%, diikuti oleh *unigram + trigram* sebesar 94% dan *unigram* sebesar 90%. Seluruh model menghasilkan nilai *precision*, *recall*, dan *F1-score* di atas 88%. Temuan ini menunjukkan bahwa kombinasi *unigram + bigram* mampu merepresentasikan konteks kata secara lebih optimal. Tingkat akurasi yang tinggi mengindikasikan bahwa model yang diusulkan dapat dimanfaatkan sebagai alat evaluasi berbasis data untuk membantu pengembang aplikasi ACI dalam memantau persepsi pengguna dan meningkatkan kualitas layanan.

Kata kunci: Analisis Sentimen, *N-Gram*, *Support Vector Machine*

ABSTRAK

Advisors: (1) Moh. Dasuki, M.Kom.; (2) Wiwik Suharso, S.Kom., M.Kom.

The transformation of information technology has encouraged the application of Natural Language Processing (NLP) in sentiment analysis to extract user opinions on application-based services. This study aims to analyze user sentiment toward the ACI online motorcycle taxi application based on reviews available on the Google Play Store using N-gram features and the Support Vector Machine (SVM) algorithm. The dataset consists of 1,419 user reviews collected between 2023 and 2025. After preprocessing, the dataset was reduced to 1,285 reviews by removing texts containing fewer than three words. Sentiment labeling was performed using a lexicon-based approach with the sentiwords_id dictionary. Reviews with neutral sentiment were excluded, resulting in 239 final samples categorized into positive and negative classes. Feature extraction was conducted using unigram, unigram + bigram, and unigram + trigram schemes with Term Frequency–Inverse Document Frequency (TF-IDF) weighting. Classification was performed using a linear SVM with an 80:20 training–testing split. Experimental results indicate that the unigram + bigram combination achieved the highest accuracy of 96%, followed by unigram + trigram at 94% and unigram at 90%. All models achieved precision, recall, and F1-score values above 88%. These findings demonstrate that the unigram + bigram combination provides a more optimal representation of word context. The high accuracy indicates that the proposed model can serve as a reliable data-driven evaluation tool to assist ACI application developers in monitoring user perceptions and improving service quality.

Keywords: Sentiment Analysis, N-Gram, Support Vector Machine