



LEMBAR PENGESAHAN
Nomor: 33/LP-LPBK/III/2026

Judul : ANALISIS PERBANDINGAN METODE SVM DAN NAIVE BAYES DALAM
KLASIFIKASI SENTIMEN PUBLIK TERHADAP VIDEO YOUTUBE ZOALS
PRASETYO "BEAT VERSI TERBAIK"
Nama : Muhammad Ali Yafi

Menerangkan bahwa abstrak dengan judul di atas telah diterjemahkan ke dalam Bahasa Inggris oleh Lembaga Pengembangan Bahasa dan Kerja Sama (LPBK), Universitas Muhammadiyah Pekajangan Pekalongan.

Pekalongan, 04 Maret 2026

Disahkan oleh,
Kepala Lembaga Pengembangan Bahasa dan Kerja Sama (LPBK)



Aida Rusmariana, S.Kep., Ns., MAN

ABSTRAK

Muhammad Ali Yafi¹, Fenilinas Adi Artanto², Ahmad Khambali³

ANALISIS PERBANDINGAN METODE SVM DAN NAIVE BAYES DALAM KLASIFIKASI SENTIMEN PUBLIK TERHADAP VIDEO YOUTUBE ZOALS PRASETYO “BEAT VERSI TERBAIK”

YouTube merupakan media sosial yang memungkinkan pengguna menyampaikan opini secara bebas melalui kolom komentar sehingga dapat dimanfaatkan sebagai sumber data analisis sentimen. Tingginya tingkat interaksi dan partisipasi pengguna pada konten otomotif menunjukkan adanya beragam persepsi publik yang perlu dianalisis secara sistematis untuk memperoleh gambaran objektif terhadap respons penonton. Video “Beat Versi Terbaik” pada akun Zoals Prasetyo menampilkan komentar dengan kecenderungan sentimen positif, negatif, dan netral yang mencerminkan pandangan masyarakat terhadap topik yang dibahas. Penelitian ini bertujuan untuk menganalisis sentimen komentar pada video tersebut serta membandingkan kinerja algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) dalam mengklasifikasikan tiga kelas sentimen. Penelitian ini menggunakan pendekatan kuantitatif dengan data sebanyak 1.500 komentar yang dikumpulkan melalui YouTube Data API melalui teknik pengambilan data. Pelabelan dilakukan menggunakan metode *lexicon* untuk menentukan polaritas sentimen. Tahapan pengolahan data meliputi *preprocessing* teks berupa *cleaning*, *case folding*, normalisasi, *tokenizing*, *stopword removal*, dan *stemming* guna meningkatkan kualitas dan konsistensi data. Fitur teks dibentuk menggunakan pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF) sebelum dilakukan proses klasifikasi menggunakan kedua algoritma. Evaluasi model dilakukan menggunakan *confusion matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur performa secara komprehensif. Hasil penelitian menunjukkan bahwa SVM memiliki kinerja lebih baik dibandingkan *Naïve Bayes* dengan nilai akurasi dan F1-score sebesar 0,70, sedangkan *Naïve Bayes* memperoleh akurasi 0,60 dan F1-score 0,53. Temuan ini menunjukkan bahwa pemilihan algoritma berpengaruh terhadap kualitas klasifikasi sentimen berbasis TF-IDF. SVM lebih optimal digunakan pada konteks penelitian ini.

Kata Kunci: *YouTube, analisis sentimen, TF-IDF, Naïve Bayes, Support Vector Machine*

ABSTRACT

Muhammad Ali Yafi¹, Fenilinas Adi Artanto², Ahmad Khambali³

THE COMPARATIVE ANALYSIS OF SVM AND NAIVE BAYES METHODS IN PUBLIC SENTIMENT CLASSIFICATION OF THE YOUTUBE VIDEO BY ZOALS PRASETYO “BEAT VERSI TERBAIK”

YouTube is a social media platform that allows users to freely express their opinions through the comment section, making it a valuable data source for sentiment analysis. The high level of interaction and user participation in automotive-related content indicates the presence of diverse public perceptions that need to be systematically analyzed to obtain an objective overview of viewers' responses. The video “Beat Versi Terbaik” on the Zoals Prasetyo channel contains comments with tendencies toward positive, negative, and neutral sentiments that reflect public views on the topic discussed. This study aims to analyze the sentiment of comments on the video and compare the performance of the Naive Bayes and Support Vector Machine (SVM) algorithms in classifying three sentiment classes. The research uses a quantitative approach with a dataset of 1,500 comments collected through the YouTube Data API using a data retrieval technique. Labeling was carried out using the lexicon-based method to determine sentiment polarity. The data processing stages include text preprocessing consisting of cleaning, case folding, normalization, tokenizing, stopword removal, and stemming to improve data quality and consistency. Text features were generated using Term Frequency–Inverse Document Frequency (TF-IDF) weighting before the classification process was conducted using both algorithms. Model evaluation was performed using a confusion matrix with the metrics of accuracy, precision, recall, and F1-score to measure performance comprehensively. The results show that SVM performs better than Naive Bayes, achieving an accuracy and F1-score of 0.70, while Naive Bayes obtains an accuracy of 0.60 and an F1-score of 0.53. These findings indicate that the selection of algorithms influences the quality of TF-IDF–based sentiment classification. SVM is more optimal for use in this research context.

Keywords: *YouTube, sentiment analysis, TF-IDF, Naive Bayes, Support Vector Machine*