



LEMBAR PENGESAHAN
Nomor: 40/LP-LPBK/III/2026

Judul : ANALISIS SENTIMEN KOMENTAR PENGGUNA YOUTUBE TERHADAP
REVIEW FILM "BATTLE OF SURABAYA" PADA CHANNEL N2 STORY
MENGUNAKAN METODE NAIVE BAYES

Nama : Ibnu Yonanda Leksono

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

Pekalongan, 10 Maret 2026

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



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

ABSTRAK

Ibnu Yonanda Leksono¹, Hadwitya Handayani², Ahmad Khambali³

ANALISIS SENTIMEN KOMENTAR PENGGUNA YOUTUBE TERHADAP REVIEW FILM “BATTLE OF SURABAYA” PADA CHANNEL N2 STORY MENGUNAKAN METODE *NAIVE BAYES*

Media sosial YouTube merupakan salah satu *platform* yang memungkinkan pengguna untuk saling berinteraksi menyampaikan opini atau pandangan secara bebas melalui fitur yang disediakan yaitu komentar, dari fitur komentar ini dapat dimanfaatkan sebagai sumber data untuk analisis sentimen. Banyaknya tingkat interaksi pengguna dalam penyampaian opini publik pada konten *review* film menunjukkan adanya beragam pandangan publik yang perlu dianalisis secara sistematis untuk memperoleh hasil dari respon penonton pada konten tersebut. Pada video konten *review* film “*battle of Surabaya*” pada kanal “N2 STORY” menampilkan komentar dengan banyak reaksi yang bisa dikategorikan sentimen positif, negatif, dan netral. Penelitian ini bertujuan untuk menganalisis sentimen komentar pada video tersebut dengan menggunakan algoritma *Naïve Bayes* dalam mengklasifikasikan tiga kelas sentimen. Penelitian ini menggunakan pendekatan kuantitatif dengan data sebanyak 859 komentar yang dikumpulkan melalui YouTube Data API melalui teknik pengambilan data. Pelabelan dilakukan menggunakan metode *lexicon-based* untuk menentukan polaritas sentimen. Tahapan pengolahan data meliputi *preprocessing* yang terdiri dari *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 menggunakan metode *Naïve Bayes* menunjukkan nilai akurasi sebesar 52,32%, Nilai *precision* sebesar 50,36%, Nilai *recall* sebesar 52,32%, dan *F1-score* sebesar 48,63. Temuan ini menunjukkan bahwa algoritma *Naïve Bayes* dapat digunakan sebagai salah satu metode klasifikasi sentimen, namun masih diperlukan pengembangan lebih lanjut agar diperoleh hasil yang lebih optimal.

Kata Kunci: *YouTube, analisis sentimen, lexicon-based, TF-IDF, Naïve Bayes*

**Undergraduate Program in Informatics
Faculty of Engineering and Computer Sciences
University of Muhammadiyah Pekajangan Pekalongan
February, 2026**

ABSTRACT

Ibnu Yonanda Leksono¹, Hadwitya Handayani², Ahmad Khambali³

**THE SENTIMENT ANALYSIS OF YOUTUBE USER COMMENTS ON THE
“BATTLE OF SURABAYA” FILM REVIEW ON THE N2 STORY
CHANNEL USING THE NAIVE BAYES METHOD**

YouTube social media is one of the platforms that enables users to interact and freely express opinions or viewpoints through the available feature, namely the comment section. These comments can be utilized as a data source for sentiment analysis. The high level of user interaction in expressing public opinions on film review content indicates the presence of diverse public perspectives that need to be systematically analyzed in order to obtain insights into audience responses to the content. In the video review of the film “Battle of Surabaya” on the N2 STORY channel, numerous comments appear with varying reactions that can be categorized into positive, negative, and neutral sentiments. This study aims to analyze the sentiment of comments on the video using the Naive Bayes algorithm to classify the three sentiment classes. This research employs a quantitative approach using a dataset of 859 comments collected through the YouTube Data API using a data collection technique. Sentiment labeling was carried out using a lexicon-based method to determine sentiment polarity. The data processing stages include preprocessing consisting of cleaning, case folding, normalization, tokenizing, stopword removal, and stemming, in order to improve data quality and consistency. Text features were constructed using Term Frequency–Inverse Document Frequency (TF-IDF) weighting before performing the classification process using the algorithm. Model evaluation was conducted using a confusion matrix with the metrics of accuracy, precision, recall, and F1-score to measure performance comprehensively. The results of the study using the Naive Bayes method show an accuracy value of 52.32%, a precision value of 50.36%, a recall value of 52.32%, and an F1-score of 48.63. These findings indicate that the Naive Bayes algorithm can be used as one of the sentiment classification methods; however, further development is still needed to achieve more optimal results.

Keywords: *YouTube, Sentiment Analysis, Lexicon-Based, TF-IDF, Naive Bayes*