



Penerapan Naive Bayes untuk Klasifikasi Opini Fans Manchester United pada Media Sosial

Muhammad Luthfi Lubis*, Aidil Halim Lubis

Fakultas Sains dan Teknologi, Program Studi Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email: ^{1,*}luthfilubis19@gmail.com, ²aidilhalimlubis@uinsu.ac.id

Email Penulis Korespondensi: luthfilubis19@gmail.com

Abstrak—Penelitian ini membahas penerapan algoritma Naïve Bayes untuk mengklasifikasikan opini penggemar Manchester United pada komentar berbahasa Indonesia di media sosial YouTube. Beragamnya bentuk bahasa dalam komentar, seperti slang, singkatan, candaan, dan sarkasme, membuat analisis manual menjadi kurang efisien dan berpotensi subjektif. Data dikumpulkan dari delapan video YouTube menggunakan YouTube Data API v3 dan disimpan dalam basis data MySQL. Dari 4.886 komentar yang diperoleh, sebanyak 1.962 komentar terdeteksi berbahasa Indonesia. Sebanyak 1.000 komentar digunakan sebagai ground truth, terdiri atas 400 komentar positif, 300 negatif, dan 300 netral. Data dibagi menjadi 80% data latih dan 20% data uji menggunakan stratified split. Tahapan pengolahan meliputi pra-pemrosesan teks, pembobotan TF-IDF, klasifikasi Naïve Bayes, dan evaluasi menggunakan confusion matrix, accuracy, precision, recall, serta F1-score. Hasil pengujian memperoleh accuracy sebesar 92,50%, macro precision 92,81%, macro recall 92,36%, dan macro F1-score 92,57%. Dari 1.944 komentar yang berhasil diklasifikasikan, sentimen positif mendominasi sebesar 47,58%, diikuti negatif 28,34% dan netral 24,07%. Sistem berbasis web yang dibangun menggunakan Laravel, PHP, dan MySQL mampu mengintegrasikan proses pengambilan data, pelabelan, klasifikasi, evaluasi, dan visualisasi hasil.

Kata Kunci: Naïve Bayes; Analisis Sentimen; Manchester United; YouTube; TF-IDF

Abstract—This study examines the application of the Naïve Bayes algorithm to classify the opinions of Manchester United fans in Indonesian-language comments on YouTube. The diverse linguistic forms found in the comments such as slang, abbreviations, jokes, and sarcasm make manual analysis inefficient and potentially subjective. Data was collected from eight YouTube videos using the YouTube Data API v3 and stored in a MySQL database. Of the 4,886 comments obtained, 1,962 were identified as being in Indonesian. A total of 1,000 comments were used as ground truth, consisting of 400 positive, 300 negative, and 300 neutral comments. The data was divided into 80% training data and 20% test data using a stratified split. The processing stages included text preprocessing, TF-IDF weighting, Naïve Bayes classification, and evaluation using a confusion matrix, accuracy, precision, recall, and F1-score. The test results yielded an accuracy of 92.50%, a macro precision of 92.81%, a macro recall of 92.36%, and a macro F1-score of 92.57%. Of the 1,944 comments successfully classified, positive sentiment dominated at 47.58%, followed by negative at 28.34% and neutral at 24.07%. The web-based system, built using Laravel, PHP, and MySQL, is capable of integrating the processes of data extraction, labeling, classification, evaluation, and result visualization.

Keywords: Naïve Bayes; Sentiment Analysis; Manchester United; YouTube; TF-IDF

1. PENDAHULUAN

Perkembangan teknologi informasi telah membawa perubahan pada cara masyarakat berkomunikasi, mencari informasi, dan menyampaikan pendapat mengenai berbagai peristiwa (Agustiranti et al., 2024). Media sosial tidak hanya digunakan sebagai sarana hiburan dan komunikasi, tetapi juga menjadi tempat bagi masyarakat untuk memberikan dukungan, kritik, dan penilaian terhadap suatu topik (Aprilia & Isnain, 2024). Salah satu media sosial yang memiliki tingkat interaksi pengguna cukup tinggi adalah YouTube (Yunus et al., 2021). Selain menampilkan konten dalam bentuk video, YouTube menyediakan kolom komentar yang dapat digunakan pengguna untuk menyampaikan opini dan tanggapan terhadap isi video (Adhianti et al., 2025).

Interaksi pengguna di YouTube banyak ditemukan pada konten olahraga, terutama sepak bola. Manchester United merupakan salah satu klub sepak bola yang memiliki banyak penggemar dan sering menjadi pembahasan di media sosial. Hasil pertandingan, penampilan pemain, keputusan pelatih, kegiatan transfer, dan kondisi klub sering menimbulkan berbagai tanggapan dari para penggemar. Komentar yang diberikan dapat berupa dukungan, pujian, kritik, kekecewaan, sindiran, maupun pendapat yang tidak memiliki kecenderungan tertentu. Banyaknya komentar membuat proses analisis secara manual menjadi kurang efektif, membutuhkan waktu yang lama, dan dapat menghasilkan penilaian yang bersifat subjektif.

Analisis terhadap opini penggemar perlu dilakukan karena komentar di media sosial dapat menggambarkan pandangan publik terhadap performa dan keadaan suatu klub sepak bola (Lubis & Kurniawan, 2025). Informasi tersebut dapat digunakan oleh media olahraga, komunitas penggemar, pengelola konten, dan akademisi untuk mengetahui kecenderungan tanggapan pengguna secara lebih terukur. Oleh karena itu, diperlukan pendekatan komputasional yang mampu mengolah komentar dalam jumlah besar dan mengelompokkannya secara otomatis berdasarkan jenis sentimen. Pendekatan yang dapat digunakan adalah analisis sentimen, yaitu proses untuk mengenali dan mengklasifikasikan opini dalam teks menjadi sentimen positif, negatif, dan netral (Mauliza & Sipayung, 2024).

Salah satu algoritma yang sering digunakan untuk klasifikasi teks adalah Naïve Bayes (Ni'mah & Archaniga, 2026). Algoritma ini menggunakan perhitungan probabilitas berdasarkan Teorema Bayes untuk menentukan kemungkinan sebuah komentar termasuk ke dalam kelas sentimen tertentu (Prasetyo, 2025). Naïve Bayes memiliki kelebihan karena proses perhitungannya relatif sederhana, efisien, dan dapat memberikan hasil yang baik dalam pengolahan data teks. Dalam penelitian ini, komentar penggemar Manchester United diubah menjadi data numerik

menggunakan pembobotan Term Frequency–Inverse Document Frequency atau TF-IDF. Pembobotan tersebut digunakan untuk mengetahui tingkat kepentingan setiap kata sebelum diproses oleh model klasifikasi.

Beberapa penelitian sebelumnya menunjukkan bahwa Naïve Bayes dapat digunakan untuk analisis sentimen dalam konteks olahraga. Melakukan analisis sentimen terhadap performa Manchester United menggunakan data dari Twitter (Muludi et al., 2023). Penelitian tersebut menunjukkan bahwa Naïve Bayes memperoleh akurasi sebesar 79% pada data berbahasa Inggris dan 75% pada data berbahasa Indonesia. Menganalisis sentimen masyarakat terhadap Tim Nasional Indonesia pada turnamen Piala Dunia U-17. Penelitian tersebut menghasilkan akurasi sebesar 78,34%, precision sebesar 82,96%, recall sebesar 78,34%, dan F1-score sebesar 74,95%. Hasil tersebut menunjukkan bahwa Naïve Bayes cukup baik dalam mengklasifikasikan opini pengguna media sosial pada bidang sepak bola (Juniardi & Sugianto, 2024).

Walaupun memiliki kesamaan dalam penggunaan Naïve Bayes dan pembahasan sepak bola, penelitian terdahulu masih menggunakan Twitter atau X sebagai sumber data utama (Ghozali et al., 2023). Sementara itu, penggunaan komentar YouTube untuk menganalisis opini penggemar Manchester United, khususnya komentar berbahasa Indonesia, masih dapat dikembangkan lebih lanjut. Komentar YouTube memiliki bentuk yang beragam, mulai dari kalimat pendek, kalimat panjang, bahasa tidak baku, singkatan, sindiran, hingga campuran istilah asing. Selain itu, penelitian ini tidak hanya menguji algoritma, tetapi juga menggabungkan proses pengambilan komentar, penyimpanan dataset, pelabelan, pra-pemrosesan teks, klasifikasi, evaluasi, dan visualisasi dalam satu sistem berbasis web.

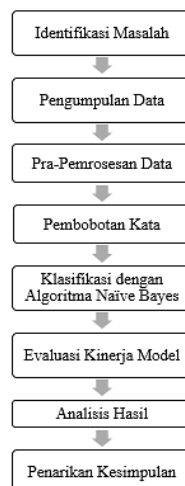
Berdasarkan kesenjangan tersebut, penelitian ini menerapkan Naïve Bayes untuk mengklasifikasikan opini penggemar Manchester United pada komentar YouTube menjadi sentimen positif, negatif, dan netral. Data dikumpulkan melalui YouTube Data API v3 dari beberapa video yang membahas Manchester United, kemudian disimpan dalam basis data MySQL (Hartati, 2022). Komentar berbahasa Indonesia yang telah diseleksi diberi label secara manual. Selanjutnya, komentar diproses melalui tahap cleaning, case folding, normalisasi, tokenizing, stopword removal, dan stemming (Hasibuan et al., 2024). Teks yang telah diproses kemudian diberi bobot menggunakan TF-IDF serta dibagi menjadi data latih dan data uji. Kinerja model diukur menggunakan confusion matrix, accuracy, precision, recall, dan F1-score (Normawati & Prayogi, 2021).

Penelitian ini bertujuan menerapkan Naïve Bayes untuk mengklasifikasikan opini penggemar Manchester United, mengukur kinerja model, dan menampilkan hasil klasifikasi melalui sistem berbasis web. Sistem dikembangkan menggunakan Laravel 12, PHP 8.2, dan MySQL serta dihubungkan dengan YouTube Data API v3 (Hendra & Riti, 2023). Kontribusi penelitian ini terletak pada penerapan proses analisis sentimen yang terintegrasi, mulai dari pengambilan komentar, pelabelan dataset, pra-pemrosesan, pembobotan TF-IDF, klasifikasi tiga kelas menggunakan Naïve Bayes, hingga visualisasi hasil. Dari sisi akademis, penelitian ini diharapkan dapat menambah kajian mengenai text mining dan analisis sentimen berbahasa Indonesia dalam bidang olahraga (Fathonah & Herliana, 2021). Dari sisi praktis, sistem yang dikembangkan dapat menjadi referensi bagi media olahraga, komunitas penggemar, pengelola konten, dan pengembang sistem untuk memahami kecenderungan opini publik di media digital.

2. METODOLOGI PENELITIAN

2.1 Jenis dan Tahapan Penelitian

Penelitian ini menggunakan metode kuantitatif dengan pendekatan eksperimen komputasional. Pendekatan ini dipilih karena penelitian mencakup proses pengumpulan komentar dari media sosial YouTube, pengolahan data teks secara teratur, pembuatan model klasifikasi, serta pengukuran kinerja model menggunakan nilai angka. Penelitian ini bertujuan untuk menerapkan algoritma Naïve Bayes dalam mengelompokkan opini penggemar Manchester United ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.



Gambar 1. Tahapan Penelitian



Pada Gambar 1, tahapan penelitian terdiri dari identifikasi masalah, pengumpulan dan penyimpanan komentar, penyaringan komentar berbahasa Indonesia, pelabelan data secara manual, pra-pemrosesan teks, pembobotan kata menggunakan TF-IDF, pembagian data menjadi data latih dan data uji, proses klasifikasi menggunakan Naïve Bayes, evaluasi kinerja model, serta penerapan dan penyajian hasil melalui sistem berbasis web. Seluruh tahapan tersebut dilakukan secara berurutan agar proses penelitian berjalan dengan teratur dan dapat dilakukan kembali menggunakan dataset yang sama. Alur penelitian dapat dilihat pada Gambar 1.

Identifikasi masalah dilakukan dengan melihat banyaknya komentar pengguna pada video YouTube yang membahas Manchester United. Komentar-komentar tersebut berisi berbagai macam opini, seperti dukungan, pujian, kritik, sindiran, kekecewaan, serta pendapat yang tidak menunjukkan sentimen tertentu. Jumlah komentar yang banyak serta penggunaan bahasa tidak baku, singkatan, sarkasme, dan istilah asing membuat proses analisis secara manual menjadi kurang efisien dan dapat menimbulkan penilaian yang subjektif. Oleh karena itu, diperlukan pendekatan komputasional yang mampu mengolah dan mengklasifikasikan komentar secara otomatis.

2.2 Pengumpulan dan Pelabelan Data

Data penelitian dikumpulkan dari delapan video YouTube yang membahas Manchester United. Video-video tersebut memuat topik tentang hasil pertandingan, penampilan tim, pemain, pelatih, kegiatan transfer, dan keadaan klub. Pemilihan video dilakukan berdasarkan kesesuaian topik penelitian serta tingkat aktivitas pengguna pada bagian komentar.

Komentar dikumpulkan secara otomatis menggunakan YouTube Data API v3 dengan ID video sebagai parameter utama. Data yang diperoleh berbentuk JSON dan berisi ID video, nama pengguna, isi komentar, waktu komentar dipublikasikan, serta informasi tambahan mengenai video. Proses pengambilan data tersebut dikembangkan menggunakan Laravel 12 dengan lingkungan PHP 8.2.

Selanjutnya, data yang telah diperoleh disimpan ke dalam basis data MySQL. Penyimpanan secara lokal dilakukan agar proses penyaringan, pelabelan, pra-pemrosesan, pelatihan, pengujian, dan klasifikasi dapat dilakukan berkali-kali tanpa perlu mengambil ulang komentar dari YouTube Data API v3. Penggunaan dataset lokal juga dapat mengurangi pemakaian kuota API dan menjaga agar data tetap stabil selama proses eksperimen berlangsung.

Hasil pengumpulan data memperoleh sebanyak 4.886 komentar yang tersimpan dalam basis data lokal. Dari jumlah tersebut, terdapat 1.962 komentar yang terdeteksi menggunakan bahasa Indonesia. Sementara itu, sebanyak 2.924 komentar tidak digunakan karena menggunakan bahasa selain bahasa Indonesia atau tidak lolos tahap penyaringan. Dari komentar berbahasa Indonesia, terdapat 18 komentar yang tidak dapat dianalisis karena tidak menghasilkan token akhir atau token yang dihasilkan tidak terdapat dalam kosakata model. Oleh karena itu, jumlah komentar yang berhasil diproses pada tahap klasifikasi akhir adalah sebanyak 1.944 komentar, Sesuai pada Tabel 1.

Tabel 1. Komposisi Dataset Penelitian

Komponen Dataset	Jumlah
Video YouTube	8 video
Total komentar dalam basis data	4.886 komentar
Komentar berbahasa Indonesia	1.962 komentar
Komentar di luar penyaringan	2.924 komentar
Komentar berlabel manual	1.000 komentar
Data latih	800 komentar
Data uji	200 komentar
Komentar yang berhasil diklasifikasikan	1.944 komentar
Komentar dengan token akhir kosong	18 komentar

Sebanyak 1.000 komentar berbahasa Indonesia dipilih untuk diberi label secara manual dan digunakan sebagai data acuan atau ground truth. Proses pelabelan dilakukan melalui halaman Ground Truth Builder yang tersedia pada sistem. Melalui halaman tersebut, peneliti dapat membaca komentar asli, melihat hasil pra-pemrosesan, menentukan kategori sentimen, menyimpan hasil pelabelan, serta memeriksa dan memperbaiki label jika terdapat kesalahan atau ketidaksesuaian. Pelabelan dibagi ke dalam tiga kategori sentimen. Sentimen positif diberikan pada komentar yang mengandung dukungan, pujian, rasa optimis, atau penghargaan terhadap Manchester United, pemain, maupun pelatih. Sentimen negatif diberikan pada komentar yang memuat kritik, hinaan, sindiran, ejekan, rasa tidak puas, atau kekecewaan. Sementara itu, sentimen netral diberikan pada komentar yang berisi informasi, pertanyaan, candaan tanpa makna sentimen yang jelas, atau komentar yang tidak menunjukkan kecenderungan sentimen tertentu. Identitas akun pengguna disamarkan dalam artikel untuk menjaga privasi. Sementara itu, isi komentar tetap digunakan sebagai bahan analisis karena komentar tersebut berasal dari kolom komentar YouTube yang dapat diakses oleh masyarakat umum.

Tabel 2. Distribusi dan Pembagian Data Berlabel

Kelas Sentimen	Data Latih	Data Uji	Total
Positif	320	80	400
Negatif	240	60	300
Netral	240	60	300



Kelas Sentimen	Data Latih	Data Uji	Total
Total	800	200	1.000

Data dibagi dengan perbandingan 80:20 menggunakan teknik stratified split. Teknik ini digunakan agar proporsi kelas positif, negatif, dan netral tetap seimbang pada data latih maupun data uji, seperti yang ditunjukkan pada Tabel 2.

2.3 Pra-pemrosesan Teks dan Pembobotan TF-IDF

Komentar yang dikumpulkan dari YouTube masih mengandung berbagai unsur yang dapat memengaruhi proses klasifikasi, seperti tautan, nama akun, emoji, simbol, tanda baca, huruf kapital, kata tidak baku, singkatan, dan kata berimbuhan. Oleh karena itu, komentar harus melalui tahap pra-pemrosesan yang meliputi cleaning, case folding, normalisasi, tokenizing, stopword removal, stemming, dan deteksi bahasa Indonesia (Yunus et al., 2021).

Tahap cleaning dilakukan untuk menghapus URL, nama akun, simbol, serta karakter yang tidak diperlukan. Selanjutnya, case folding digunakan untuk mengubah semua huruf menjadi huruf kecil. Tujuannya agar kata yang sama tetap dianggap sebagai satu fitur meskipun ditulis dengan penggunaan huruf kapital yang berbeda.

Tahap normalisasi dilakukan untuk mengubah kata tidak baku dan singkatan menjadi bentuk yang lebih seragam, seperti “ga” menjadi “tidak”, “udh” menjadi “sudah”, dan “utd” menjadi “united”. Normalisasi hanya diterapkan pada token yang berdiri sebagai kata utuh. Hal ini dilakukan agar proses penggantian tidak mengubah bagian dari kata lain yang memiliki susunan karakter yang hampir sama.

Tahap tokenizing digunakan untuk memisahkan komentar menjadi beberapa kata atau token. Setelah itu, stopword removal dilakukan untuk menghapus kata-kata umum yang kurang memberikan informasi terhadap sentimen, seperti “dan”, “yang”, “di”, dan “itu”. Selanjutnya, proses stemming dilakukan menggunakan pustaka Sastrawi untuk mengubah kata berimbuhan menjadi bentuk kata dasar. Deteksi bahasa Indonesia kemudian digunakan untuk memilih komentar yang memiliki ciri kosakata bahasa Indonesia (Siregar et al., 2023).

Setelah melalui seluruh tahap pra-pemrosesan, komentar diubah ke dalam bentuk angka menggunakan metode Term Frequency–Inverse Document Frequency atau TF-IDF. Nilai Term Frequency dihitung menggunakan Persamaan (1).

$$TF(t, d) = \frac{f(t, d)}{\sum_k f(k, d)} \quad (1)$$

Pada perhitungan Term Frequency (TF), $TF(t, d)$ menunjukkan frekuensi kemunculan kata t dalam dokumen d . Sementara itu, $f(t, d)$ merupakan jumlah kemunculan kata t pada dokumen d , sedangkan $\sum_k f(k, d)$ menyatakan jumlah keseluruhan kata atau token yang terdapat dalam dokumen d . Nilai Inverse Document Frequency dihitung menggunakan Persamaan (2).

$$IDF(t) = \ln \left(\frac{1+N}{1+df(t)} \right) + 1 \quad (2)$$

Pada perhitungan Inverse Document Frequency (IDF), $IDF(t)$ menunjukkan nilai bobot kelangkaan kata t pada seluruh dokumen. Variabel N menyatakan jumlah keseluruhan dokumen yang terdapat dalam korpus, sedangkan $df(t)$ merupakan jumlah dokumen yang mengandung kata t . Adapun $\ln$ menunjukkan fungsi logaritma natural yang digunakan dalam proses perhitungan nilai IDF. Bobot akhir TF-IDF dihitung menggunakan Persamaan (3).

$$TFIDF(t, d) = TF(t, d) \times IDF(t) \quad (3)$$

Pada perhitungan bobot TF-IDF, $TFIDF(t, d)$ menunjukkan bobot akhir kata t pada dokumen d . Nilai $TF(t, d)$ menyatakan frekuensi kemunculan kata t dalam dokumen d , sedangkan $IDF(t)$ menunjukkan tingkat keunikan atau kelangkaan kata t pada keseluruhan dokumen dalam korpus. Hasil pembobotan TF-IDF membentuk vektor numerik untuk setiap komentar. Kata yang sering muncul dalam suatu komentar, tetapi jarang ditemukan pada komentar lainnya, akan memperoleh bobot lebih tinggi. Vektor TF-IDF tersebut digunakan sebagai masukan dalam proses pembentukan model Naïve Bayes.

2.4 Klasifikasi Menggunakan Naïve Bayes

Data latih digunakan untuk membangun model agar mampu mengenali hubungan antara pola kata dalam komentar dan kategori sentimennya (Normawati & Prayogi, 2021). Model Naïve Bayes kemudian menghitung kemungkinan suatu komentar termasuk ke dalam kelas positif, negatif, atau netral. Nilai probabilitas awal dari setiap kelas dihitung menggunakan Persamaan (4).

$$P(c) = \frac{N_c}{N} \quad (4)$$

Pada perhitungan probabilitas awal kelas, $P(c)$ menunjukkan probabilitas awal dari kelas c . Nilai N_c menyatakan jumlah dokumen latih yang termasuk dalam kelas c , sedangkan N merupakan jumlah keseluruhan dokumen yang digunakan dalam proses pelatihan. Implementasi Naïve Bayes dalam penelitian ini menggunakan akumulasi bobot TF-IDF pada setiap kelas. Laplace smoothing diterapkan untuk menghindari probabilitas bernilai nol apabila suatu kata tidak ditemukan pada kelas tertentu. Probabilitas suatu kata terhadap kelas dihitung menggunakan Persamaan (5).



$$P(w|c) = \frac{W(w,c)+1}{W_{total(c)}+|V|} \quad (5)$$

Pada perhitungan probabilitas kata, $P(w|c)$ menunjukkan probabilitas kemunculan kata w pada kelas c . $W(w,c)$ merupakan jumlah bobot TF-IDF kata w pada kelas c , sedangkan $W_{total(c)}$ menyatakan jumlah keseluruhan bobot fitur yang terdapat pada kelas c . Selanjutnya, $|V|$ menunjukkan jumlah kata unik dalam kosakata, sementara nilai 1 merupakan nilai Laplace smoothing yang digunakan untuk menghindari probabilitas bernilai nol. Skor suatu komentar terhadap setiap kelas dihitung dalam bentuk logaritma menggunakan Persamaan (6).

$$Score(c|d) = \ln(P(c)) + \sum_{w \in d} TFIDF(w,d) \times \ln(P(w|c)) \quad (6)$$

Pada perhitungan skor klasifikasi, $Score(c|d)$ menunjukkan skor dokumen d terhadap kelas c . $P(c)$ merupakan probabilitas awal dari kelas c , sedangkan $TFIDF(w,d)$ menyatakan bobot TF-IDF kata w pada dokumen d . Selanjutnya, $P(w|c)$ menunjukkan probabilitas kata w pada kelas c , sementara $\ln$ merupakan fungsi logaritma natural yang digunakan dalam proses perhitungan skor.

Model menghitung skor komentar terhadap kelas positif, negatif, dan netral. Kelas dengan skor tertinggi atau paling mendekati nol ditetapkan sebagai hasil prediksi akhir. Penggunaan logaritma dilakukan untuk mencegah nilai probabilitas menjadi terlalu kecil ketika banyak fitur dikalikan.

2.5 Evaluasi Model dan Implementasi Sistem

Kinerja model dinilai menggunakan confusion matrix tiga kelas dengan membandingkan label sebenarnya dan label hasil prediksi. Metrik evaluasi yang digunakan meliputi accuracy, precision, recall, dan F1-score. Nilai accuracy dihitung dengan membandingkan jumlah prediksi yang benar pada seluruh kelas dengan jumlah keseluruhan data uji.

$$Accuracy = \frac{TP_{positif} + TP_{negatif} + TP_{netral}}{N} \times 100\% \quad (7)$$

Nilai precision pada setiap kelas dihitung menggunakan Persamaan (8).

$$Precision_c = \frac{TP_c}{TP_c + FP_c} \quad (8)$$

Nilai recall pada setiap kelas dihitung menggunakan Persamaan (9).

$$Recall_c = \frac{TP_c}{TP_c + FN_c} \quad (9)$$

Nilai F1-score dihitung menggunakan Persamaan (10).

$$F1Score_c = 2 \times \left(\frac{Precision_c \times Recall_c}{Precision_c + Recall_c} \right) \quad (10)$$

Dalam perhitungan metrik evaluasi pada kelas c , $TP(c)$ merupakan jumlah data yang berasal dari kelas c dan berhasil diprediksi dengan benar sebagai kelas c . $FP(c)$ menunjukkan jumlah data dari kelas lain yang keliru diprediksi sebagai kelas c , sedangkan $FN(c)$ merupakan jumlah data pada kelas c yang salah diprediksi sebagai kelas lain. Nilai precision, recall, dan F1-score dihitung untuk kelas positif, negatif, dan netral. Penelitian ini juga menggunakan macro average agar setiap kelas memperoleh bobot penilaian yang sama.

$$MacroPrecision = \frac{Precision_{positif} + Precision_{negatif} + Precision_{netral}}{3} \quad (11)$$

$$MacroRecall = \frac{Recall_{positif} + Recall_{negatif} + Recall_{netral}}{3} \quad (12)$$

$$MacroF1Score = \frac{F1_{positif} + F1_{negatif} + F1_{netral}}{3} \quad (13)$$

Seluruh proses penelitian diterapkan dalam sistem berbasis web menggunakan Laravel 12, PHP 8.2, dan MySQL yang terhubung dengan YouTube Data API v3 untuk mengambil informasi video, komentar utama, dan balasan komentar. Data disimpan dalam tabel “videos”, “video_comments”, “analysis_runs”, dan “comments” untuk menyimpan metadata video, komentar mentah, hasil pra-pemrosesan, token akhir, status bahasa, label manual, riwayat analisis, konfigurasi model, serta hasil klasifikasi (Sinlae et al., 2024).

Sistem juga menggunakan beberapa layanan, yaitu “YouTubeService”, “TextPreprocessingService”, “IndonesianLanguageService”, “TfidfService”, “NaiveBayesService”, dan “EvaluationService” untuk menjalankan proses pengambilan data, pra-pemrosesan, penyaringan bahasa, pembobotan TF-IDF, klasifikasi, pembagian data, serta evaluasi model (Narulita et al., 2024). Penyimpanan dataset dalam basis data memungkinkan eksperimen dilakukan kembali tanpa mengambil komentar baru, sehingga penggunaan kuota API dapat dikurangi dan sumber data tetap sama selama dataset belum diperbarui.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan dan Pengolahan Dataset

Pengumpulan data dari delapan video YouTube menghasilkan 4.886 komentar yang disimpan dalam basis data lokal. Setelah proses deteksi bahasa, ditemukan 1.962 komentar berbahasa Indonesia, dengan 1.000 komentar digunakan sebagai ground truth dan 1.944 komentar berhasil diproses pada klasifikasi akhir. Sementara itu, 18 komentar tidak dapat dianalisis karena tidak menghasilkan token akhir atau tokennya tidak terdapat dalam kosakata model. Data ground truth terdiri atas 400 komentar positif, 300 komentar negatif, dan 300 komentar netral, sehingga setiap kelas memiliki jumlah data yang cukup untuk proses pelatihan dan pengujian. Pembagian data secara stratified juga menjaga agar seluruh kelas tetap terwakili pada data uji. Hasil pra-pemrosesan menunjukkan bahwa sistem mampu membersihkan dan menyeragamkan teks komentar sebelum dilakukan pembobotan TF-IDF, seperti yang ditunjukkan pada Tabel 3.

Tabel 3. Contoh Hasil Pra-pemrosesan Teks

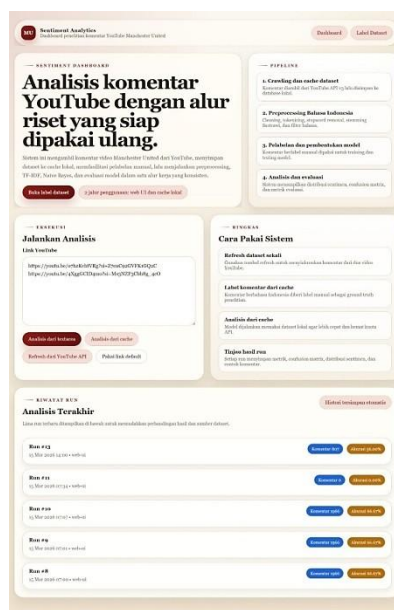
No.	Komentar Asli	Hasil Akhir Pra-pemrosesan	Label
1	min, berita terkini hari ini dtmpt gua listrik hidup mati tampaknya PLN sudah tidak sanggup menyaingi cahaya dan energi KING MUFC yang teramat luar biasa	min berita kini hari dtmpt gua listrik hidup mati tampak pln sanggup saing cahaya energi king mufc yang amat luar biasa	Positif
2	Hari ini terlihat ada fans Mencretan United ngalaMUn sendirian dikamar setelah kalah dari Galatasaray	hari lihat fans mencret united ngalamun sendiri kamar telah kalah galatasaray	Negatif
3	01:12 nama edit nya apa min?	01 12 nama edit apa min	Netral

Hasil pada Tabel 3 menunjukkan bahwa tahap pra-pemrosesan tidak hanya menghapus karakter yang tidak diperlukan, tetapi juga menyamakan bentuk kata. Normalisasi yang diterapkan hanya pada token atau kata utuh dapat mencegah perubahan yang merusak bagian tengah kata. Tahap ini membantu menghasilkan fitur yang lebih konsisten sebelum diproses menggunakan TF-IDF dan Naïve Bayes.

3.2 Hasil Implementasi Sistem

Sistem berbasis web yang dikembangkan berhasil menggabungkan seluruh tahapan penelitian dalam satu aplikasi (Kadim et al., 2023). Pengguna dapat memasukkan tautan video, memperbarui komentar melalui YouTube Data API v3, menyimpan komentar ke dalam basis data, memberikan label secara manual, menjalankan proses analisis, serta melihat hasil evaluasi model (Hendra & Riti, 2023).

Halaman Ground Truth Builder digunakan untuk memberikan label positif, negatif, atau netral pada setiap komentar. Sistem juga menyediakan fitur untuk menampilkan komentar yang belum diberi label, komentar yang sudah diberi label, hasil pra-pemrosesan, serta fasilitas untuk memperbaiki label yang dinilai kurang sesuai. Setelah data berlabel tersedia, pengguna dapat menjalankan analisis menggunakan dataset yang tersimpan dalam basis data. Sistem membagi data dengan teknik stratified split, membentuk fitur TF-IDF, melatih model Naïve Bayes, memprediksi data uji, serta mengklasifikasikan seluruh komentar berbahasa Indonesia yang dapat diproses. Hasil klasifikasi ditampilkan dalam bentuk grafik distribusi sentimen, tabel metrik setiap kelas, confusion matrix, metadata video, serta contoh komentar yang dilengkapi label aktual dan label prediksi. Sistem juga menyimpan riwayat pengujian sehingga hasil analisis dapat diperiksa kembali tanpa harus mengulangi seluruh proses dari awal.



Gambar 2. Halaman Dashboard Sistem Analisis Sentimen

Video	Komentar Asli	Preprocessing	Aktual	Prediksi
4XggGCID4mo	Giliran Arsenal gak kalah,dah ada rekor di UCL,gak upload kau nol,gk konsisten kau nol jadi youtuber 🤔	gilir arsenal kalah dah rekor ucl upload kau nol gk konsisten kau nol youtuber	-	Positif
4XggGCID4mo	Yaaaa tim mu kalah yaaa 😭😭	yaaaa tim mu kalah yaaa	negatif	Negatif
4XggGCID4mo	kalo MU menang dia dimunculin tetep krg seru	kalo mu menang dimunculin tetep krg seru	positif	Positif
4XggGCID4mo	Kalo king Mu menang , baru aku nonton cinol... Gak kuat kalok king Mu kala	kalo king mu menang nonton cinol kuat kalok king mu kala	netral	Positif
o7hzKvhtVRg	Ngomong apa biawak satu ini 🤔	ngomong apa biawak satu	-	Positif
o7hzKvhtVRg	@hafidbailhaqiz601king loserpoo 🤔 king yg di akui itu king M.U stiap club kecil macam loserpoo aja suka termotivasi untuk menang lawan king m.u sama ke club kecil lainnya suka termotivasi kala lawan king biar dibawa media haha ups	loserpool king aku king m u stiap club kecil macam loserpoo suka motivasi menang lawan king m u club kecil lai suka motivasi kala lawan king biar bawa media haha ups	-	Positif

Gambar 3. Halaman Hasil Klasifikasi dan Evaluasi Model

Tampilan halaman pelabelan, komentar yang belum diberi label, komentar yang sudah diberi label, dan riwayat pengujian dapat disatukan dalam satu gambar komposit apabila diperlukan. Penyajian tersebut lebih ringkas dibandingkan menampilkan setiap halaman dalam gambar yang berbeda.

3.3 Hasil Klasifikasi Sentimen

Sebanyak 1.944 komentar berhasil dikelompokkan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Hasil klasifikasi menunjukkan bahwa sentimen positif menjadi kategori terbanyak dengan 925 komentar atau 47,58%. Sentimen negatif berjumlah 551 komentar atau 28,34%, sedangkan sentimen netral berjumlah 468 komentar atau 24,07%.

Tabel 4. Distribusi Hasil Klasifikasi Sentimen

Sentimen	Jumlah Komentar	Persentase
Positif	925	47,58%
Negatif	551	28,34%
Netral	468	24,07%
Total	1.944	100%

Jumlah sentimen positif yang paling besar menunjukkan bahwa sebagian besar komentar pada delapan video yang diteliti berisi dukungan, penghargaan, rasa optimis, atau tanggapan positif terhadap Manchester United. Meskipun demikian, persentase sentimen negatif sebesar 28,34% menunjukkan bahwa pembahasan mengenai klub tersebut juga menimbulkan cukup banyak kritik, sindiran, ejekan, dan rasa kecewa.

Komentar netral terdiri atas pertanyaan, informasi, candaan yang tidak menunjukkan sentimen secara jelas, serta komentar yang tidak secara langsung berisi dukungan maupun kritik. Adanya ketiga kategori tersebut menunjukkan bahwa percakapan tentang Manchester United di YouTube memiliki beragam bentuk opini, seperti yang ditunjukkan pada Tabel 4.

3.4 Hasil Evaluasi Model

Pengujian dilakukan terhadap 200 data uji yang terdiri atas 80 komentar positif, 60 komentar negatif, dan 60 komentar netral. Hasil evaluasi menunjukkan akurasi sebesar 92,50%, macro precision sebesar 92,81%, macro recall sebesar 92,36%, dan macro F1-score sebesar 92,57%, seperti yang ditunjukkan pada Tabel 5.

Tabel 5. Ringkasan Evaluasi Model

Metrik Evaluasi	Nilai
Accuracy	92,50%
Macro Precision	92,81%
Macro Recall	92,36%
Macro F1-score	92,57%

Hasil evaluasi untuk setiap kelas ditunjukkan pada Tabel 6.

Tabel 6. Metrik Evaluasi Setiap Kelas

Kelas	Precision	Recall	F1-score	Support
Positif	96,55%	93,33%	94,92%	60
Negatif	91,53%	90,00%	90,76%	60
Netral	90,36%	93,75%	92,02%	80



Kelas netral memperoleh precision tertinggi, yaitu 96,55%. Hasil tersebut menunjukkan bahwa sebagian besar komentar yang diprediksi sebagai netral memang memiliki label aktual netral. Kelas positif memperoleh recall tertinggi sebesar 93,75%, yang berarti model mampu mengenali sebagian besar komentar positif pada data uji. Kelas negatif memperoleh precision sebesar 91,53%, recall sebesar 90,00%, dan F1-score sebesar 90,76%. Nilai tersebut menunjukkan bahwa model tetap memiliki kemampuan yang baik dalam mengenali komentar negatif meskipun bahasa kritik dan sindiran pada komentar YouTube dapat memiliki bentuk yang beragam. Hasil confusion matrix ditunjukkan pada Tabel 7.

Tabel 7. Confusion Matrix Hasil Pengujian

Aktual / Prediksi	Netral	Negatif	Positif
Netral	56	2	2
Negatif	0	54	6
Positif	2	3	75

Berdasarkan confusion matrix, jumlah prediksi yang benar adalah 56 komentar netral, 54 komentar negatif, dan 75 komentar positif. Dengan demikian, perhitungan akurasi dapat dituliskan sebagai berikut:

$$Accuracy = \frac{56+54+75}{200} \times 100\%$$

$$Accuracy = \frac{185}{200} \times 100\%$$

$$Accuracy = 92,50\%$$

Kesalahan klasifikasi terbesar terjadi pada enam komentar negatif yang diprediksi sebagai positif. Kesalahan tersebut dapat terjadi karena komentar negatif mengandung kata yang secara umum memiliki hubungan dengan sentimen positif, tetapi digunakan dalam bentuk sindiran atau sarkasme. Selain itu, terdapat dua komentar netral yang diprediksi negatif, dua komentar netral yang diprediksi positif, dua komentar positif yang diprediksi netral, dan tiga komentar positif yang diprediksi negatif. Contoh keluaran hasil klasifikasi ditunjukkan pada Tabel 8.

Tabel 8. Contoh Hasil Klasifikasi Komentar

Video ID	Hasil Pra-pemrosesan	Aktual	Prediksi
o7hzKvhtVRg	king emyu cuma kalah lawan tim kecil	Negatif	Negatif
gpn4NjOn50s	lawak habis habis	Negatif	Negatif
Krf9fH6yUfE	mu makan	Netral	Netral
OVvaV6TxoF8	latih house morienho dapat 3 bel kecil2an dithn 2017 ingat g musim depan epl seru mu cit liverpool chelsea arsenal tottenham	Positif	Positif

3.5 Pembahasan

Nilai akurasi sebesar 92,50% dan macro F1-score sebesar 92,57% menunjukkan bahwa kombinasi pra-pemrosesan bahasa Indonesia, TF-IDF, dan Naïve Bayes mampu mengklasifikasikan opini penggemar Manchester United dengan baik. Nilai macro precision, macro recall, dan macro F1-score yang hampir sama juga menunjukkan bahwa performa model cukup seimbang pada kelas positif, negatif, dan netral. Hasil tersebut didukung oleh penggunaan 1.000 komentar sebagai ground truth, pembagian data secara stratified, serta perbaikan tahap normalisasi yang membuat token lebih konsisten. Performa penelitian ini lebih tinggi dibandingkan (Muludi et al., 2023) yang memperoleh akurasi 79% pada data bahasa Inggris dan 75% pada data bahasa Indonesia, serta (Juniardi & Sugianto, 2024) dengan akurasi 78,34% dan F1-score 74,95%. Namun, perbandingan tersebut tetap perlu dilakukan secara hati-hati karena setiap penelitian memiliki jumlah data, sumber data, proses pelabelan, karakteristik komentar, dan konfigurasi model yang berbeda.

Kesalahan prediksi masih terjadi karena komentar YouTube sering mengandung sarkasme, candaan, ejekan, singkatan, julukan klub, konteks pertandingan, dan campuran bahasa yang sulit dipahami hanya dari arti kata. Hasil klasifikasi menunjukkan sentimen positif sebesar 47,58%, sehingga opini pada delapan video cenderung berisi dukungan atau respons positif, meskipun sentimen negatif sebesar 28,34% menunjukkan bahwa kritik dan ketidakpuasan tetap banyak ditemukan. Sistem berbasis Laravel juga berhasil menggabungkan proses pengambilan komentar, penyimpanan data, deteksi bahasa, pelabelan, pra-pemrosesan, TF-IDF, klasifikasi, evaluasi, dan visualisasi dalam satu aplikasi. Meskipun demikian, penelitian ini masih memiliki keterbatasan berupa kemungkinan subjektivitas pelabelan, belum mampu menganalisis sentimen berdasarkan aspek tertentu, serta penggunaan dataset yang hanya membahas Manchester United. Secara umum, TF-IDF dan Naïve Bayes dapat digunakan untuk mengklasifikasikan komentar YouTube berbahasa Indonesia ke dalam tiga kelas sentimen dan diterapkan dalam sistem berbasis web.

4. KESIMPULAN

Kombinasi pra-pemrosesan teks bahasa Indonesia, pembobotan TF-IDF, dan algoritma Naïve Bayes mampu mengklasifikasikan komentar penggemar Manchester United di YouTube ke dalam sentimen positif, negatif, dan netral



dengan baik. Dari 1.944 komentar yang berhasil diproses, sentimen positif menjadi kategori terbanyak dengan 925 komentar atau 47,58%, diikuti sentimen negatif sebanyak 551 komentar atau 28,34% dan sentimen netral sebanyak 468 komentar atau 24,07%. Pengujian terhadap 200 data uji menghasilkan akurasi sebesar 92,50%, macro precision sebesar 92,81%, macro recall sebesar 92,36%, dan macro F1-score sebesar 92,57%, yang menunjukkan bahwa performa model cukup tinggi dan seimbang pada ketiga kelas sentimen. Sistem berbasis web yang dikembangkan juga berhasil menggabungkan proses pengambilan komentar, penyimpanan data, pelabelan, pra-pemrosesan, klasifikasi, evaluasi, dan visualisasi hasil dalam satu aplikasi. Meskipun masih terdapat kesalahan prediksi akibat sarkasme, candaan, singkatan, dan campuran bahasa, hasil penelitian menunjukkan bahwa TF-IDF dan Naïve Bayes dapat digunakan untuk menganalisis opini penggemar Manchester United pada komentar YouTube berbahasa Indonesia secara otomatis.

REFERENCES

- Adhianti, P. D., Solikhin, Riyanto, E., Lutfi, S., & Purwanto, A. (2025). Public Sentiment Analysis of the Indonesian National Football Team on Instagram Using Naive Bayes. *Jurnal Teknologi Informasi Dan Komunikasi*, 16(2), 192–201. <https://doi.org/10.51903/jtikp.v16i2.1539>
- Agustiranti, T., Kurdiana, A. K. I., Al Ghiffari, B., Juniar, E. D., & Purnama, D. G. (2024). Penerapan Naive Bayes terhadap sentimen analisis media sosial Twitter pengguna Kereta Cepat Jakarta–Bandung (Whoosh). *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 7(1), 297–305. <https://ejournal.sisfokomtek.org/index.php/jikom>
- Aprilia, N. W. A. S., & Isnain, A. R. (2024). Analisis sentimen terhadap media sosial Twitter dengan kasus kampanye anti-korupsi di Indonesia menggunakan Naive Bayes. *Jurnal Media Informatika Budidarma*, 8(2), 695–703. <https://doi.org/10.30865/mib.v8i2.7582>
- Fathonah, F., & Herliana, A. (2021). Penerapan Text Mining analisis sentimen mengenai vaksin COVID-19 menggunakan metode Naïve Bayes. *Jurnal Sains Dan Informatika*, 7(2), 155–162. <https://doi.org/10.34128/jsi.v7i2.331>
- Ghozali, M. I., Sugiharto, W. H., & Iskandar, A. F. (2023). Analisis sentimen pinjaman online di media sosial Twitter menggunakan metode Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(6), 1340–1348. <https://doi.org/10.30865/klik.v3i6.936>
- Hartati, E. (2022). Sistem Informasi Transaksi Gudang Berbasis Website Pada Cv. Asyura. *Klik - Jurnal Ilmu Komputer*, 3(1), 12–18. <https://doi.org/10.56869/klik.v3i1.323>
- Hasibuan, M. S., Lubis, A. H., & Sari, M. N. (2024). Perbandingan algoritma clustering DBSCAN dan K-Means dalam pengelompokan siswa terbaik. *INFOTECH: Jurnal Informatika & Teknologi*, 5(2), 301–309. <http://jurnal.sttmcleungsi.ac.id/index.php/infotech>
- Hendra, H., & Riti, Y. F. (2023). Perancangan Dan Implementasi Website Dengan Konsep Ui/Ux Untuk Mengoptimalkan Marketing Perusahaan. *Jurnal Informatika Dan Teknik Elektro Terapan*, 11(3s1). <https://doi.org/10.23960/jitet.v11i3s1.3430>
- Juniardi, T., & Sugianto, C. A. (2024). Analisis sentimen tim nasional sepak bola Indonesia di turnamen Piala Dunia U-17 Indonesia pada Twitter (X) menggunakan algoritma Naïve Bayes. *Jurnal Informatika Dan Teknik Elektro Terapan (JITET)*, 12(3S1), 3715–3722. <https://doi.org/10.23960/jitet.v12i3S1.5188>
- Kadim, A. A., Hadjarate, L., & Muthia, M. (2023). Implementasi Framework Laravel Dalam Pembuatan Sistem Pencatatan Notula Berbasis Website. *Jurnal Sistem Informasi Bisnis*, 13(1), 45–51. <https://doi.org/10.21456/vol13iss1pp45-51>
- Lubis, F. Z., & Kurniawan, R. (2025). Sentiment Analysis Towards Naturalization of Indonesian National Team Players on Social Media X Using the Naive Bayes Method. *Jurnal Mandiri IT*, 14(1), 208–218. <https://doi.org/10.35335/mandiri.v14i1.412>
- Mauliza, R. N., & Sipayung, Y. R. (2024). Penerapan Text Mining dalam menganalisis pendapat masyarakat terhadap Pemilu 2024 pada media sosial X menggunakan metode Naive Bayes. *Technomedia Journal (TMJ)*, 9(1). <https://doi.org/10.33050/tmj.v9i1.2212>
- Muludi, K., Humam, M. N., Shofiana, D. A., & Syarif, A. (2023). Perbandingan kinerja CNN dan Naïve Bayes pada analisis sentimen performa Manchester United di Twitter. *Journal Information Engineering and Educational Technology (JIEET)*, 7(2), 83–90.
- Narulita, S., Nugroho, A., & Abdillah, M. Z. (2024). Diagram Unified Modelling Language (UML) untuk Perancangan Sistem Informasi Manajemen Penelitian dan Pengabdian Masyarakat (SIMLITABMAS). *Jurnal Publikasi Sistem Informasi Dan Telekomunikasi*, 3, 244–256.
- Ni'mah, K., & Archaniga, R. A. (2026). Handling Class Imbalance in Fan Sentiment Analysis: Naïve Bayes with TF-IDF on Instagram and Twitter. *Jurnal Teknik Informatika*, 19(1), 22–37. <https://doi.org/10.15408/jti.v19i1.46733>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier dan Confusion Matrix pada analisis sentimen berbasis teks pada Twitter. *Jurnal Sains Komputer Dan Informatika (J-SAKTI)*, 5(2), 697–711. <https://tunasbangsa.ac.id/ejurnal/index.php/jsakti>
- Prasetyo, V. P. (2025). Analisis Sentimen Platform YouTube Mengenai Pro Kontra terhadap Naturalisasi Pemain Sepakbola Timnas Indonesia Menggunakan Metode Naïve Bayes. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(3), 2225–2232. <https://doi.org/10.23960/jitet.v13i3.7235>
- Sinlae, F., Irwanda, E., Maulana, Z., & Syahputra, V. E. (2024). Penggunaan Framework Laravel dalam Membangun



- Aplikasi Website Berbasis PHP. *Jurnal Siber Multi Disiplin (JSMD)*, 2(2), 119–132. <https://creativecommons.org/licenses/by/4.0/>
- Siregar, D., Ladayya, F., Albaqi, N. Z., & Wardana, B. M. (2023). Penerapan metode Support Vector Machines (SVM) dan metode Naïve Bayes Classifier (NBC) dalam analisis sentimen publik terhadap konsep child-free di media sosial Twitter. *Jurnal Statistika Dan Aplikasinya*, 7(1), 93–100. <https://doi.org/10.31284/j.jsa.2023.7.1>
- Yunus, M., Husni, M., & Mufadhdhal, M. M. (2021). Klasifikasi sentimen terhadap Badan Penyelenggara Jaminan Sosial (BPJS) pada media sosial Twitter menggunakan Naive Bayes. *SMATIKA: STIKI Informatika Jurnal*, 11(2), 81–91. <https://doi.org/10.32664/smatika.v11i02.577>