

Analisis Sentimen Ulasan Pengguna Mobile Legends di Google Play Store menggunakan Naïve Bayes dan KNN

Dwi Wijiyanto*, Arif Setiawan, Muhammad Arifin

Fakultas Teknik, Program Studi Sistem Informasi, Universitas Muria Kudus, Kudus, Indonesia

Email: ^{1,*}202153035@std.umk.ac.id, ²arif.setiawan@umk.ac.id, ³arifin.m@umk.ac.id

Email Penulis Korespondensi: 202153035@std.umk.ac.id*

Submitted: 04/10/2025; Accepted: 07/11/2025; Published: 31/12/2025

Abstrak– Mobile Legends merupakan salah satu game mobile populer dengan jumlah unduhan tinggi dan ribuan ulasan pengguna di Google Play Store. Penelitian ini bertujuan menganalisis sentimen ulasan pengguna berbahasa Indonesia menggunakan algoritma Naïve Bayes dan K-Nearest Neighbors (KNN). Sebanyak 4.000 data ulasan dikumpulkan melalui teknik *web scraping* dan diproses melalui tahapan *cleaning*, *tokenization*, *stopword removal*, serta *stemming*. Data kemudian diklasifikasikan ke dalam dua sentimen, yaitu positif dan negatif. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, presisi, *recall*, F1-score, dan Cross Validation. Hasil menunjukkan bahwa algoritma Naïve Bayes memiliki performa yang lebih stabil, sedangkan KNN lebih unggul dalam mengenali kemiripan pola kata. Penelitian ini memberikan wawasan terhadap persepsi pengguna serta dapat menjadi acuan bagi pengembang dalam meningkatkan kualitas layanan game Mobile Legends.

Kata Kunci: Google Play Store; K-Nearest Neighbors; Mobile Legends; Naïve Bayes; Analisis Sentimen

Abstract– Mobile Legends is one of the most popular mobile games with a high number of downloads and thousands of user reviews on the Google Play Store. This study aims to analyze the sentiment of Indonesian-language user reviews using the *Naïve Bayes* and *K-Nearest Neighbors* (KNN) algorithms. A total of 4,000 reviews were collected through *web scraping* and processed through several preprocessing stages, including cleaning, tokenization, stopwords removal, and stemming. The data were then classified into two sentiment categories: positive and negative. Model performance was evaluated using accuracy, precision, recall, F1-score, and cross-validation metrics. The results indicate that the Naïve Bayes algorithm provides more stable performance, while KNN excels in identifying word pattern similarities. This study offers insights into user perceptions and serves as a reference for developers to improve the quality of Mobile Legends' features and services.

Keywords: Google Play Store; K-Nearest Neighbors; Mobile Legends; Naïve Bayes; Sentiment Analysis

1. PENDAHULUAN

Penelitian ini berfokus pada analisis sentimen ulasan pengguna game Mobile Legends di Google Play Store dengan memanfaatkan pendekatan berbasis komputasi. Mobile Legends merupakan salah satu game bergenre *Multiplayer Online Battle Arena* (MOBA) yang paling populer di Indonesia bahkan di dunia. Popularitasnya tidak hanya tercermin dari jumlah unduhan yang telah mencapai ratusan juta, tetapi juga dari banyaknya ulasan pengguna yang terus bertambah setiap hari. Ulasan-ulasan tersebut mengandung berbagai persepsi pengguna, baik berupa apresiasi maupun kritik yang dapat menjadi masukan penting bagi pengembang dalam meningkatkan kualitas layanan.

Namun, jumlah ulasan yang sangat besar dan beragam menimbulkan tantangan tersendiri. Proses analisis secara manual tidak efisien karena memerlukan waktu yang lama dan berpotensi menyebabkan banyak masukan berharga terabaikan. Kondisi ini dapat menghambat pengembang dalam mengidentifikasi permasalahan pengguna secara cepat, yang pada akhirnya berdampak pada penurunan kepuasan dan loyalitas. Oleh karena itu, dibutuhkan pendekatan otomatis berbasis komputasi, seperti analisis sentimen, yang mampu mengolah data teks dalam jumlah besar secara cepat dan akurat untuk mengekstraksi informasi bermakna dari ulasan pengguna.

Sejalan dengan kebutuhan tersebut, dalam lima tahun terakhir telah banyak penelitian yang menerapkan analisis sentimen terhadap ulasan aplikasi menggunakan berbagai algoritma klasifikasi. Prayogi dan Masitoh, misalnya, menganalisis ulasan aplikasi Alfagift dengan algoritma *Random Forest*. Dari 4.379 ulasan yang dikumpulkan melalui Google Play Store, data dikategorikan menjadi sentimen positif dan negatif. Hasil penelitian menunjukkan akurasi sebesar 97,6% dengan AUC 0,98, yang termasuk dalam kategori sangat baik[1]. Temuan ini menunjukkan bahwa metode *ensemble learning* dapat bekerja optimal dalam klasifikasi teks.

Sementara itu, Triputra dan Rubhasy membandingkan algoritma *Naïve Bayes* dan *K-Nearest Neighbors* (KNN) pada ulasan pengguna aplikasi *Facebook* yang dikumpulkan melalui teknik *web scraping* dan diolah dengan *Natural Language Processing* (NLP). Hasil pengujian menunjukkan bahwa Naïve Bayes menghasilkan akurasi 73%, sedangkan KNN 72%[2]. Meskipun selisihnya kecil, Naïve Bayes terbukti lebih stabil dalam mengklasifikasikan sentimen positif dan negatif.

Penelitian lain oleh Mualfah dkk. membandingkan algoritma Naïve Bayes dan Random Forest dalam menganalisis sentimen publik di *Twitter* terkait kasus pembobolan data Bank Syariah Indonesia (BSI) pada tahun 2023. Hasilnya menunjukkan bahwa *Naïve Bayes* memperoleh akurasi 81%, lebih tinggi dibandingkan *Random Forest* yang mencapai 78%[3]. Adapun Palepa dkk. menerapkan algoritma KNN untuk menganalisis komentar *YouTube* mengenai politik identitas pada Pemilu 2024 dan memperoleh akurasi 65%[4].

Dari berbagai penelitian tersebut dapat disimpulkan bahwa performa algoritma analisis sentimen sangat dipengaruhi oleh karakteristik data, teknik praproses, serta metode evaluasi yang digunakan. *Random Forest* terbukti unggul pada kasus aplikasi belanja daring dengan data yang relatif bersih[1], sedangkan *Naïve Bayes* lebih stabil pada data media sosial yang beragam dan singkat[3]. KNN, meskipun sederhana, tetap relevan karena mampu mendeteksi pola kedekatan antar kata dalam dataset tertentu[2].

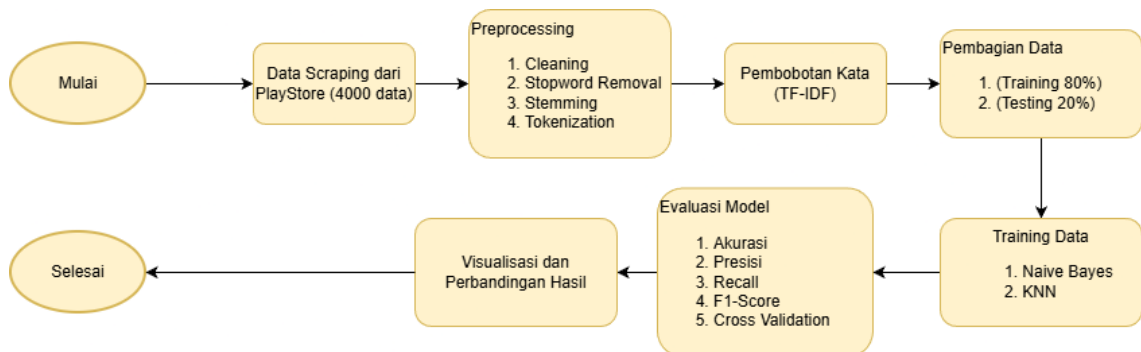
Namun, hingga saat ini masih terbatas penelitian yang secara khusus membahas analisis sentimen pada ulasan pengguna game *Mobile Legends* di *Google Play Store*. Padahal, dengan basis pengguna yang sangat besar, ulasan-ulasan game ini berpotensi memberikan wawasan berharga bagi pengembang untuk memahami persepsi pengguna dan meningkatkan pengalaman bermain. Kesenjangan penelitian (gap analysis) ini menjadi dasar bagi penelitian ini, mengingat studi sebelumnya lebih banyak berfokus pada aplikasi belanja daring, layanan perbankan, maupun isu politik, sementara domain game online populer masih jarang dikaji.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna *Mobile Legends* di *Google Play Store* dengan membandingkan efektivitas algoritma *Naïve Bayes* dan *K-Nearest Neighbors*. Secara praktis, hasil penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi pengembang dalam memahami persepsi positif maupun negatif dari pengguna, sehingga dapat merancang strategi peningkatan fitur yang lebih tepat sasaran. Dari sisi akademis, penelitian ini juga diharapkan dapat memperkaya literatur di bidang analisis sentimen dan klasifikasi teks, khususnya dalam konteks aplikasi game berbahasa Indonesia. Dengan demikian, penelitian ini diharapkan mampu menjembatani kesenjangan antara teori dan praktik serta memberikan kontribusi nyata bagi industri maupun dunia akademik.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pemilihan algoritma *Naïve Bayes* dan *K-Nearest Neighbors* (KNN) dalam penelitian ini didasarkan pada karakteristik keduanya yang sesuai untuk analisis sentimen berbasis teks. *Naïve Bayes* dipilih karena memiliki kemampuan yang baik dalam menangani data berdimensi tinggi seperti representasi teks hasil transformasi TF-IDF. Algoritma ini bekerja dengan prinsip probabilistik dan asumsi independensi antar fitur, sehingga mampu memberikan hasil klasifikasi yang cepat dan efisien meskipun dengan jumlah data yang besar. Selain itu, *Naïve Bayes* terbukti efektif dalam mengidentifikasi pola kata yang merepresentasikan sentimen positif atau negatif tanpa memerlukan proses pelatihan yang kompleks. Sementara itu, *K-Nearest Neighbors* (KNN) digunakan sebagai pembanding karena sifatnya yang non-parametrik dan berbasis jarak, di mana prediksi dilakukan berdasarkan kedekatan antar data. KNN mampu menangkap pola lokal pada dataset dan memberikan gambaran tentang bagaimana hubungan antarulasan memengaruhi hasil klasifikasi sentimen. Kombinasi kedua algoritma ini memberikan perbandingan yang komprehensif antara pendekatan probabilistik dan berbasis jarak, sehingga dapat dievaluasi secara mendalam keunggulan masing-masing metode dalam konteks analisis sentimen teks pengguna aplikasi *Mobile Legends*. Rangkaian tahapan penelitian pada Gambar 1 disusun guna memberikan kemudahan dalam pelaksanaan penelitian. Alur ini menjadi pedoman terstruktur yang menguraikan langkah demi langkah proses penelitian.



Gambar 1 Tahapan Penelitian

2.2 Pengumpulan Data

Proses pengumpulan (*web scraping*) dalam penelitian ini menargetkan sebanyak 4.000 ulasan pengguna dari aplikasi '*com.mobile.legend*s' di *Google Play Store* [5]. Proses ini dijalankan secara otomatis menggunakan bahasa pemrograman *Python* [6] di lingkungan *Google Colab*. Sesuai kode yang digunakan, library utama *google-play-scraper* dimanfaatkan untuk mengambil data ulasan terbaru (diurutkan berdasarkan *Sort.NEWEST*) secara iteratif dalam bahasa Indonesia. Untuk memastikan data yang diambil bersih dan terstruktur, library *pandas* digunakan untuk mengelola hasil *scraping* ke dalam struktur *DataFrame*, memformat tanggal, dan akhirnya menyimpan data dalam format file CSV untuk dianalisis lebih lanjut.

2.3 Preprocessing Data

Sesudah data berhasil diperoleh, dilakukan tahap preprocessing guna memastikan ulasan berada dalam kondisi yang bersih serta siap dipakai pada proses analisis selanjutnya. Tahapan ini dikerjakan menggunakan Google Colab dengan bahasa pemrograman Python dan meliputi sejumlah langkah penting yang ditujukan untuk meningkatkan kualitas data. Tahap preprocessing menjadi bagian krusial dalam penyiapan data karena berfungsi menyederhanakan analisis dengan melakukan standarisasi sekaligus pemurnian terhadap data awal[7].

Tahap awal preprocessing adalah *cleaning*, yaitu proses pembersihan teks dengan cara mengubah seluruh huruf menjadi format lowercase, menghapus karakter angka, serta menghilangkan simbol atau tanda yang tidak relevan seperti ':', '_', '=', '+', dan sejenisnya[8]. Setelah melalui tahap *cleaning*, proses berikutnya adalah *stopword removal*, yaitu menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis, seperti kata sambung atau kata keterangan [9]. Tahapan ini dilakukan agar hanya kata-kata yang relevan dan memiliki bobot informasi signifikan yang dipertahankan, sehingga data menjadi lebih ringkas dan siap digunakan pada proses analisis berikutnya.

Selanjutnya masuk ke tahap *stemming*, yaitu proses mengubah kata berimbuhan menjadi bentuk dasarnya agar memiliki makna yang lebih konsisten[10]. Pada penelitian ini, *stemming* dilakukan dengan bantuan library pemrosesan teks berbahasa Indonesia sehingga setiap kata yang memiliki variasi imbuhan dapat diseragamkan ke dalam bentuk kata dasar untuk mempermudah proses analisis sentimen. Setelah proses *stemming* selanjutnya adalah tahap *tokenizing*, yaitu proses memecah teks menjadi unit yang lebih kecil dengan membagi kalimat ke dalam kata per kata[11]. Pada tahap ini, kalimat dipisahkan menjadi potongan kata atau token[12], sementara tanda baca yang masih diperlukan dapat tetap dipertahankan untuk menjaga konteks teks.

2.4 Pembobotan TF-IDF

Setelah tahap preprocessing selesai, langkah berikutnya adalah pembobotan dengan metode TF-IDF. Pembobotan TF-IDF (Term Frequency-Inverse Document Frequency) merupakan metode perhitungan pada tiap bobot kata yang terdapat di dalam data dokumen[13], sehingga setiap kata memperoleh bobot sesuai tingkat kepentingannya. Hasil transformasi ini memungkinkan data teks diproses lebih lanjut dan digunakan secara efektif dalam proses klasifikasi sentimen

Nilai TF-IDF dihitung berdasarkan formula matematis sebagaimana ditunjukkan pada persamaan (1) :

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (1)$$

Term Frequency (TF)

Menunjukkan frekuensi kemunculan suatu kata pada dokumen tertentu, yang dihitung menggunakan persamaan (2) berikut:

$$TF(t, d) = \frac{f_{t,d}}{\sum_k f_{k,d}} \quad (2)$$

Inverse Document Frequency (IDF)

Mengukur keunikan kata di seluruh koleksi dokumen. Dihitung dengan persamaan (3) berikut:

$$IDF(t) = \log\left(\frac{N}{df_t}\right) \quad (3)$$

IDF dihitung dengan N sebagai total seluruh teks pada kumpulan data dan DF(t) sebagai jumlah dokumen yang memuat kata t. Semakin jarang sebuah kata muncul, semakin besar pula nilai IDF yang diberikan, sehingga kata tersebut dinilai lebih bermakna.

2.5 Pembagian Data

Tahap selanjutnya adalah membagi data menjadi dua porsi, yaitu 80% untuk keperluan training dan 20% untuk testing[14]. Data latih digunakan untuk membangun serta mengoptimalkan kinerja algoritma klasifikasi sehingga mampu mengenali pola dari data yang tersedia, sedangkan data uji berfungsi untuk mengevaluasi performa model dengan data yang belum pernah dipelajari sebelumnya. Pembagian ini dilakukan agar hasil analisis sentimen dapat lebih objektif sekaligus mengurangi risiko terjadinya *overfitting* pada model.

2.6 Training Data

Langkah berikutnya adalah menggunakan data tersebut untuk melatih model klasifikasi dengan menerapkan metode *Naïve Bayes* dan *K-Nearest Neighbors* (KNN). Algoritma *Naïve Bayes* melakukan identifikasi sentimen dengan memanfaatkan berbagai elemen tekstual, seperti pilihan kata dan pola kalimat yang merepresentasikan kecenderungan positif maupun negatif[15]. Setiap ulasan kemudian dianalisis dengan menghitung probabilitas keanggotaan ke dalam masing-masing kelas sentimen, sehingga dapat ditentukan apakah suatu teks bersifat positif maupun negatif. Proses perhitungan probabilitas tersebut dilakukan menggunakan varian Multinomial *Naïve Bayes*, yang sesuai untuk data teks hasil transformasi TF-IDF, dengan parameter utama α sebagai nilai

smoothing Laplace untuk menghindari probabilitas nol pada kata yang jarang muncul. Nilai α biasanya disesuaikan melalui proses validasi untuk mencapai performa optimal.

Sebagai algoritma klasifikasi yang berfungsi untuk mencari nilai probabilitas yang baik, Naïve Bayes mampu memberikan hasil klasifikasi yang sederhana namun cukup akurat, terutama dalam pengolahan data teks berukuran besar[16]. Naïve Bayes menggunakan Teorema Bayes sebagai landasannya, dengan perhitungan yang dirumuskan pada persamaan (4) berikut :

$$P(C_i|X) = \frac{P(X|C_i) \cdot P(C_i)}{P(X)} \quad (4)$$

Dimana :

$P(C_i|X)$: robabilitas posterior, yaitu kemungkinan data X termasuk ke dalam kelas C_i

$P(X|C_i)$: likelihood, probabilitas kemunculan data X jika diketahui kelasnya C_i

$P(C_i)$: prior probability, probabilitas awal suatu data masuk ke kelas C_i

$P(X)$: evidence, probabilitas keseluruhan dari data X .

K-Nearest Neighbors (KNN) merupakan algoritma klasifikasi yang sederhana tetapi cukup handal, dengan prinsip bahwa data yang memiliki kemiripan atribut kemungkinan besar berada pada kelas yang sama[17]. Proses klasifikasi pada KNN diawali dengan pemilihan sebuah data uji untuk dibandingkan dengan data latih yang sudah diberi label. Selanjutnya, algoritma akan mencari sejumlah tetangga terdekat sesuai nilai k yang telah ditentukan sebelumnya oleh pengguna, dan kelas mayoritas dari tetangga tersebut akan menjadi hasil prediksi. Dengan demikian, nilai k tersebut menjadi parameter utama dalam model KNN yang digunakan dalam proses voting untuk menentukan kelas prediksi. Selain itu, pemilihan nilai k dan jenis fungsi jarak (distance metric) seperti *cosine similarity* atau *Manhattan distance* juga memengaruhi kinerja model dalam mengenali pola sentimen pada data teks.

Algoritma KNN menghitung jarak antar dua vektor teks dengan menggunakan rumus yang ditunjukkan pada persamaan (5) :

$$:d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5)$$

Dimana :

$d(x, y)$: jarak antara vektor X dan Y

x_i, y_i : nilai atribut ke- i dari masing-masing vektor,

n : jumlah atribut atau dimensi data.

2.7 Evaluasi Model

Setelah proses pelatihan dan pengujian selesai, dilakukan evaluasi terhadap hasil prediksi dengan menggunakan metrik performa, meliputi akurasi, presisi, recall, serta F1-score, serta Cross Validation guna menilai kemampuan algoritma dalam membedakan kelas sentimen. Untuk memperjelas hasil evaluasi, digunakan diagram perbandingan metrik yang memvisualisasikan sejauh mana efektivitas tiap algoritma dalam melakukan klasifikasi sentimen pengguna. Akurasi (Accuracy) yaitu rasio antara jumlah prediksi yang tepat, baik pada kelas positif maupun negatif, terhadap total keseluruhan data yang diuji[18]. Nilai akurasi diperoleh dengan membandingkan jumlah prediksi yang benar (TP + TN) terhadap keseluruhan jumlah prediksi, yaitu (TP + TN + FP + FN). Dapat dihitung dengan persamaan berikut (6) :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Presisi (Precision) adalah rasio antara jumlah data yang diprediksi benar sebagai positif dengan keseluruhan data yang diprediksi sebagai positif[18]. Presisi dihitung dengan membagi jumlah prediksi positif yang benar (TP) dengan keseluruhan data yang diprediksi sebagai positif, yaitu (TP + FP). Dapat dihitung dengan persamaan (7) berikut :

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

Recall adalah rasio antara jumlah data positif yang berhasil diprediksi dengan benar terhadap keseluruhan data yang sebenarnya termasuk ke dalam kelas positif[18]. Recall dihitung dengan membagi jumlah prediksi positif yang benar (TP) dengan total keseluruhan data positif aktual, yaitu (TP + FN). Dapat dihitung dengan persamaan (8) berikut :

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

F1-Score adalah metrik yang diperoleh dari rata-rata harmonik antara nilai presisi dan recall, sehingga mampu memberikan keseimbangan antara keduanya[18]. dapat dihitung dengan persamaan (9) berikut :

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

Cross Validation adalah salah satu teknik untuk menilai atau memvalidasi keakuratan sebuah model yang dibangun berdasarkan dataset tertentu[19]. *K-fold cross validation*, di mana dataset dibagi menjadi k bagian dengan ukuran yang relatif sama. Pada setiap iterasi, satu fold digunakan sebagai data uji (*testing set*), sedangkan k-1 fold lainnya digunakan sebagai data latih (*training set*), proses ini menggunakan fungsi *cross_val_score* dari *library scikit-learn*. Data fitur yang digunakan adalah teks yang telah ditransformasi menggunakan *TfidfVectorizer* dengan batasan 1.000 fitur teratas (*max_features=1000*).

$$CV = \frac{1}{k} \sum_{i=1}^k M_i \quad (10)$$

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Data penelitian diperoleh dari ulasan pengguna Mobile Legends di Google Play Store menggunakan metode *web scraping* berbasis Python dengan library *google-play-scraper*. Proses pengambilan data dilakukan secara bertahap sebanyak 100 ulasan per permintaan dengan jeda satu detik untuk menghindari deteksi sebagai aktivitas otomatis. Informasi yang dikumpulkan mencakup *username*, tanggal ulasan, rating, dan komentar dengan total 4.000 entri. Dataset disimpan dalam format CSV untuk memudahkan tahap *preprocessing* dan analisis sentimen. Sistem juga dilengkapi mekanisme pemeriksaan duplikasi berdasarkan kombinasi *username* dan isi komentar. Hasil akhir dari proses ini digunakan sebagai data utama pada tahap analisis yang ditampilkan pada Tabel 1.

Tabel 1 Data hasil web scraping

No	Username	Tanggal	Bintang	Komentar
1.	Mofreps	2025-08-10	1	stress gua main di mytic kayak bukan mytic
2.	Levin	2025-08-10	1	Moonton kalo mau di kasih bintang 5 minimal kasih gw tim bagus
3.	Muhammad Dzikri	2025-08-10	5	love you moonton
4.	Beswin Lian	2025-08-10	1	lawak sia sia gw top up game korupsi blokir aja ni game
5.	Komalasari sari22	2025-08-10	1	kikir tim gak baleg tim musuh jago tim gw gak jeng dan lol

3.2 Preprocessing Data

Tahap *preprocessing* dimulai dengan membaca dataset dalam format CSV untuk memastikan kelengkapan dan konsistensi data. Ulasan yang kosong, tidak bermakna, atau hanya berisi simbol dihapus, sementara sentimen ulasan tanpa teks ditentukan berdasarkan kategori rating bintang, di mana rating 1–3 diklasifikasikan sebagai negatif dan 4–5 sebagai positif. Selanjutnya, dilakukan proses *cleaning* dengan mengubah teks menjadi huruf kecil serta menghapus angka dan karakter non-alfabet. Proses tokenisasi menggunakan library NLTK memecah teks menjadi kata tunggal, kemudian disaring dengan menghapus *stopwords*. Tahap akhir adalah *stemming* menggunakan library Sastrawi untuk mengembalikan kata ke bentuk dasarnya. Hasilnya menghasilkan dua atribut baru, yaitu “Komentar_Cleaned” dan “Komentar_Tokenized”, yang selanjutnya digunakan pada tahap analisis klasifikasi sentimen sebagaimana ditunjukkan pada Tabel 2.

Tabel 2 Data Hasil Preprocessing

No	Username	Tanggal	Bintang	Komentar	Komentar Cleaned	Komentar Tokenized	Label
1	Mofreps	2025-08-10	1	stress gua main di mytic kayak bukan mytic	stress gua main mytic kayak bukan mytic	['stress', 'gua', 'main', 'mytic', 'kayak', 'bukan', 'mytic']	negatif
2	Levin	2025-08-10	1	Moonton kalo mau di kasih bintang 5	moonton kalo mau kasih bintang minimal	['moonton', 'kalo', 'mau', 'kasih', 'bintang', 'minimal']	negatif

				minimal kasih gw tim bagus	kasih gw tim bagus	'minimal', 'kasih', 'gw', 'tim', 'bagus']	
3	Muhammad Dzikri	2025-08- 10	5	love you moonton	love you moonton	['love', 'you', 'moonton']	positif
4	Beswin Lian	2025-08- 10	1	lawak sia sia gw top up game korupsi blokir aja ni game	lawak sia sia gw top up game korupsi blokir aja ni game	['lawak', 'sia', 'sia', 'gw', 'top', 'up', 'game', 'korupsi', 'blokir', 'aja', 'ni', 'game']	negatif
5	Mofreps	2025-08- 10	1	stress gua main di mytic kayak bukan mytic	stress gua main mytic kayak bukan mytic	['stress', 'gua', 'main', 'mytic', 'kayak', 'bukan', 'mytic']	negatif

3.3 Pembobotan TF-IDF

Setelah melewati tahap preprocessing, data kemudian dikonversi ke dalam bentuk representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Hasil transformasi tersebut dapat dilihat pada Tabel 3 berikut :

Tabel 3 Data Hasil TF-IDF

No	Term	TF	IDF	TF-IDF
1	aaaaaa	30,011,746,882,677,400	860,115,242,829,729	2,581,356,095,773,840
2	aaaaaaaaa	9,003,524,064,803,220	860,115,242,829,729	7,744,068,287,321,530
3	aamiin	3,391,277,476,472,040	860,115,242,829,729	2,916,889,450,178,740
4	aamon	36,690,782,779,008,700	860,115,242,829,729	3,155,830,153,957,990
5	aaaaaa	30,011,746,882,677,400	860,115,242,829,729	2,581,356,095,773,840

Pada tahap implementasi, data hasil *preprocessing* dibaca dari file CSV dan diubah menjadi representasi numerik menggunakan metode TF-IDF melalui fungsi *TfidfVectorizer* dari pustaka *scikit-learn*. Proses ini menghitung frekuensi kemunculan kata dan tingkat kepentingannya dalam keseluruhan korpus, sehingga menghasilkan bobot TF-IDF untuk setiap kata unik. Nilai tersebut kemudian disusun dalam bentuk tabel yang memuat daftar kata beserta bobotnya, dan disimpan kembali dalam file CSV sebagai masukan untuk proses klasifikasi sentimen pada tahap selanjutnya.

3.4 Training Data

Data hasil preprocessing yang telah diberi label digunakan sebagai data latih dalam proses pembelajaran model. Penelitian ini menggunakan dua metode klasifikasi, yaitu *Naïve Bayes* dan *K-Nearest Neighbors* (KNN). Kinerja kedua model tersebut dievaluasi melalui metrik akurasi, presisi, recall, dan F1-score untuk mengukur efektivitas sistem dalam mengenali polaritas sentimen pada ulasan pengguna Mobile Legends..

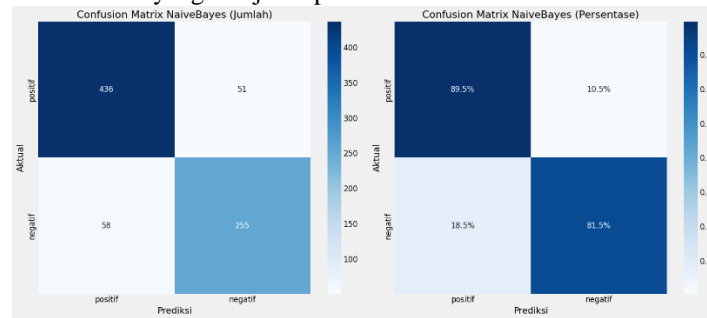
3.4.1 Naïve Bayes

Setelah model *Naïve Bayes* dilatih dengan dataset yang sudah diproses pada tahap preprocessing, dilakukan evaluasi untuk menilai kinerjanya dalam mengklasifikasikan sentimen pengguna terhadap aplikasi Mobile Legends. Ringkasan hasil pengukuran menggunakan metrik evaluasi disajikan pada Tabel 4.

Tabel 4 Metriks Metode Naive Bayes

No	Naïve Bayes Metriks	Jumlah
1.	Akurasi	86.38%
2.	Presisi	85.80%
3.	Recall	85.50%
4.	F1-Score	85.64%
5.	Runtime	0.2102 Detik

Hasil pengujian menunjukkan bahwa model menghasilkan akurasi sebesar 86,38%, dengan nilai presisi 85,80%, recall 85,50%, dan F1-score 85,64%. Waktu komputasi yang dibutuhkan hanya 0,2102 detik, menandakan bahwa algoritma Naïve Bayes mampu melakukan proses klasifikasi secara cepat dan efisien. Rincian kinerja model dapat dilihat pada Confusion Matrix yang disajikan pada Gambar 2.



Gambar 2 Confusion Matrix Naive Bayes

Berdasarkan *Confusion Matrix*, algoritma *Naïve Bayes* menunjukkan kinerja yang seimbang dalam mengenali kedua kelas sentimen. Model mampu mengklasifikasikan ulasan positif dengan tingkat ketepatan 89,5%, sedangkan untuk sentimen negatif mencapai 81,5%. Dari total data, sebanyak 436 ulasan positif dan 255 ulasan negatif berhasil diidentifikasi dengan benar, sementara 51 ulasan positif serta 58 ulasan negatif masih mengalami kesalahan klasifikasi.

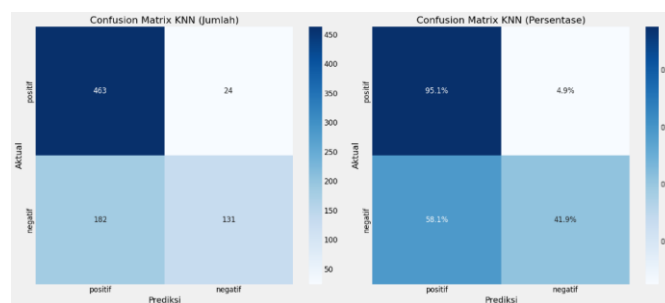
3.4.2 K-Nearest Neighbors (KNN)

Algoritma *K-Nearest Neighbors* (KNN) menghasilkan akurasi sebesar 74,25%, dengan nilai presisi 78,15%, recall 68,46%, dan F1-score 68,89%. Hasil ini menunjukkan bahwa performa KNN masih kurang stabil dibandingkan *Naïve Bayes*. Dari sisi efisiensi, waktu eksekusi mencapai 17,6230 detik, menandakan bahwa proses klasifikasi dengan metode ini membutuhkan waktu yang lebih lama. Ringkasan kinerja KNN ditampilkan pada Tabel 5.

Tabel 5 Metriks Metode K-Nearest Neighbors (KNN)

No	KNN Metriks	Jumlah
1	Akurasi	74.25%
2	Presisi	78.15%
3	Recall	68.46%
4	F1-Score	68.89%
5	Runtime	17.6230 detik

Dari aspek efisiensi, algoritma KNN memerlukan waktu komputasi sebesar 17,6230 detik, menunjukkan bahwa proses klasifikasinya relatif lambat dibandingkan *Naïve Bayes*. Kondisi ini terjadi karena KNN harus menghitung jarak antara data uji dan seluruh data pelatihan pada setiap prediksi, sehingga waktu eksekusi meningkat seiring dengan bertambahnya jumlah data. Detail hasil klasifikasi ditampilkan pada Confusion Matrix pada Gambar 3.



Gambar 3 Confusion Matrix KNN

Berdasarkan *Confusion Matrix*, algoritma KNN menunjukkan kemampuan tinggi dalam mengenali sentimen positif dengan tingkat ketepatan 95,1%, namun performanya terhadap sentimen negatif masih rendah, hanya mencapai 4,9%. Dari keseluruhan data, sebanyak 463 ulasan positif terklasifikasi dengan benar, sedangkan

24 salah dikategorikan sebagai negatif. Pada kelas negatif, kesalahan prediksi cukup dominan dengan 182 data keliru teridentifikasi sebagai positif dan hanya 131 yang diprediksi dengan tepat.

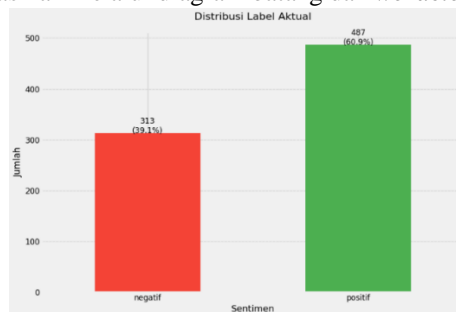
3.4.3 Evaluasi dan Perbandingan

Tujuan dari tahapan ini adalah mengevaluasi serta membandingkan performa algoritma *Naïve Bayes* dan *K-Nearest Neighbors* (KNN) dalam mengklasifikasikan sentimen, dengan hasil yang diperlihatkan pada Tabel 6.

Tabel 6 Perbandingan hasil metrik

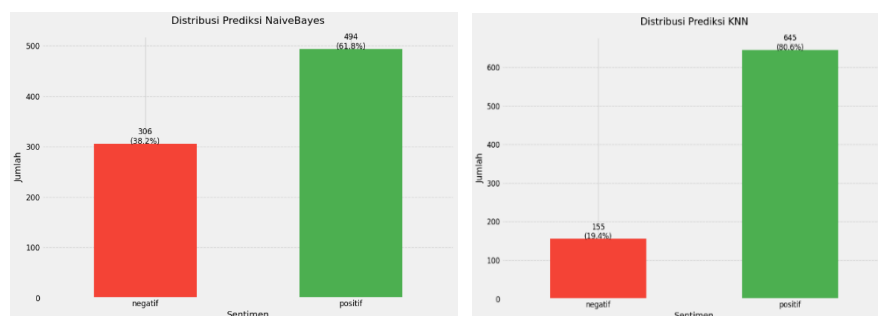
No	Metriks	Naïve Bayes	KNN
1	Akurasi	86.38%	74.25%
2	Presisi	85.80%	78.15%
3	Recall	85.50%	68.46%
4	F1-Score	85.64%	68.89%
5	Runtime	0.2102 detik	17.6230 detik

Hasil evaluasi menunjukkan bahwa algoritma *Naïve Bayes* memiliki performa paling optimal dalam analisis sentimen ulasan pengguna Mobile Legends. Model ini unggul dalam seluruh metrik evaluasi, dengan tingkat akurasi dan konsistensi hasil yang lebih tinggi dibandingkan KNN. Keseimbangan antara presisi dan recall menjadikan *Naïve Bayes* lebih andal dalam mengidentifikasi polaritas sentimen secara keseluruhan. Perbandingan kinerja kedua algoritma divisualisasikan melalui diagram batang dan *wordcloud* pada bagian berikut.



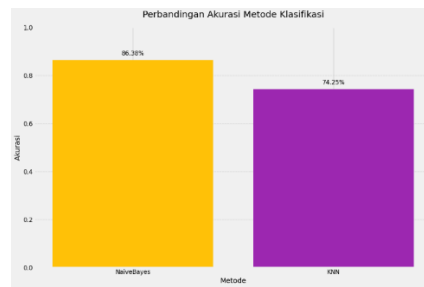
Gambar 4 Diagram batang label aktual

Distribusi label pada Gambar 4 memperlihatkan bahwa ulasan pengguna Mobile Legends di Google Play Store didominasi oleh sentimen positif sebanyak 487 ulasan (60,9%), sedangkan sentimen negatif berjumlah 313 ulasan (39,1%). Kondisi ini menunjukkan kecenderungan pengguna untuk memberikan penilaian yang baik terhadap aplikasi, meskipun masih terdapat sebagian yang menyampaikan pengalaman kurang memuaskan. Ketidakseimbangan proporsi tersebut menjadi pertimbangan penting dalam analisis sentimen agar model mampu mengenali kedua kelas secara proporsional dan menghasilkan prediksi yang lebih akurat.



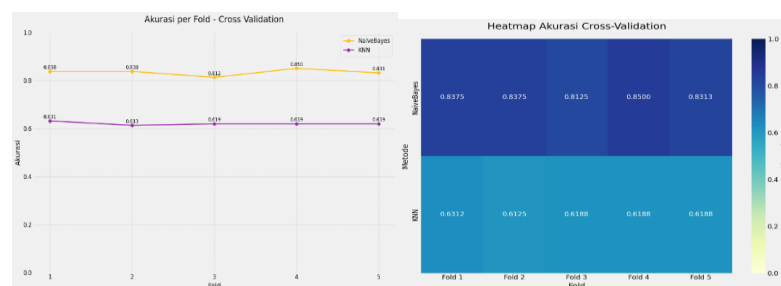
Gambar 5 Tabel batang prediksi Naive Bayes dan KNN

Perbandingan hasil prediksi menunjukkan adanya perbedaan karakteristik antara algoritma *Naïve Bayes* dan *K-Nearest Neighbors* (KNN). Model *Naïve Bayes* menghasilkan distribusi sentimen yang relatif seimbang, dengan proporsi mendekati data aktual, sehingga menunjukkan stabilitas dalam klasifikasi. Sebaliknya, KNN menampilkan kecenderungan kuat terhadap kelas positif dan kesulitan dalam mengidentifikasi sentimen negatif. Ketidakseimbangan hasil prediksi ini mengindikasikan bahwa KNN lebih sensitif terhadap dominasi data positif pada dataset, sehingga akurasinya terhadap kelas minoritas menjadi lebih rendah.



Gambar 6 Tabel batang perbandingan akurasi

Gambar perbandingan akurasi memperlihatkan bahwa algoritma *Naïve Bayes* memiliki performa lebih unggul dibandingkan *K-Nearest Neighbors* (KNN). Pendekatan probabilistik yang diterapkan *Naïve Bayes* terbukti lebih efektif dalam mengenali pola teks ulasan pengguna Mobile Legends, sedangkan metode berbasis jarak pada KNN cenderung menghasilkan tingkat ketepatan yang lebih rendah.



Gambar 7 Cross Validation

Hasil validasi silang, yang divisualisasikan dengan jelas dalam heatmap dan line plot akurasi, menunjukkan perbedaan performa yang sangat signifikan di antara kedua model: *Naive Bayes* (Garis Kuning): Model ini menunjukkan kinerja yang sangat baik dan stabil. Akurasi di setiap fold konsisten tinggi, berkisar antara 0.812 (Fold 3) hingga puncaknya 0.850 (Fold 4). Fluktuasi yang rendah antar fold (garis yang relatif datar) menunjukkan bahwa model ini tidak sensitif terhadap pembagian data latih/uji tertentu dan memiliki kemampuan generalisasi yang baik. Rata-rata akurasi *Naive Bayes* dari kelima fold adalah 83.38%. *K-Nearest Neighbors* (Garis Ungu): Sebaliknya, model KNN (dengan K=5) menunjukkan performa yang jauh lebih rendah. Meskipun juga stabil, akurasinya hanya berkisar antara 0.613 (Fold 2) dan 0.631 (Fold 1). Rata-rata akurasi KNN hanya mencapai 62.02%. Kesenjangan performa yang besar ini (terlihat jelas pada line plot) dengan kuat mengindikasikan bahwa *Naive Bayes* adalah model yang jauh lebih superior dan lebih cocok untuk tugas klasifikasi sentimen pada dataset ini dibandingkan dengan KNN.



Gambar 8 Visualisasi wordcloud label aktual

Gambar wordcloud menampilkan kata-kata yang paling sering muncul pada ulasan pengguna Mobile Legends untuk masing-masing kategori sentimen. Pada sentimen positif, kata seperti “bagus”, “seru”, dan “mantap” merepresentasikan kepuasan pengguna terhadap kualitas dan pengalaman bermain. Sementara itu, pada sentimen negatif, kemunculan kata “tidak”, “gak”, “moonton”, serta istilah seperti “kalah” dan “rusak” mencerminkan keluhan pengguna terhadap performa teknis maupun sistem permainan.



Gambar 9 Visualisasi wordcloud Naive Bayes

Gambar wordcloud hasil prediksi dengan algoritma Naïve Bayes menggambarkan pola kemunculan kata yang merepresentasikan persepsi pengguna Mobile Legends. Pada sentimen positif, istilah seperti “bagus”, “seru”, “mantap”, dan “keren” mencerminkan apresiasi terhadap kualitas permainan dan fitur yang disediakan. Sementara pada sentimen negatif, kemunculan kata “tidak”, “gak”, “tim”, “moonton”, dan “kalah” menunjukkan adanya keluhan terkait performa permainan, ketidakseimbangan sistem, serta kritik terhadap pengembang.



Gambar 10 Visualisasi wordcloud KNN

Gambar wordcloud hasil prediksi dengan algoritma K-Nearest Neighbors (KNN) memperlihatkan sebaran kata yang menggambarkan kecenderungan sentimen pengguna Mobile Legends. Pada sentimen positif, kata seperti “bagus”, “seru”, “mantap”, dan “keren” mendominasi, menunjukkan adanya kepuasan terhadap kualitas permainan dan pengalaman bermain. Sebaliknya, pada sentimen negatif, kemunculan kata “tidak”, “gak”, “kalah”, “sistem”, dan “moonton” mengindikasikan munculnya keluhan terkait performa permainan maupun kritik terhadap pihak pengembang.

Secara kualitatif, kestabilan performa algoritma *Naïve Bayes* dipengaruhi oleh karakteristiknya yang berbasis pada pendekatan probabilistik dengan asumsi independensi antar fitur. Pendekatan ini memungkinkan model mengenali pola umum dalam teks secara efektif, bahkan ketika jumlah data pada tiap kelas tidak seimbang. Karena *Naïve Bayes* berfokus pada probabilitas kemunculan kata yang merepresentasikan suatu sentimen, bukan pada jarak antar data, model ini mampu menghasilkan prediksi yang konsisten terhadap berbagai variasi ulasan. Sementara itu, algoritma *K-Nearest Neighbors* (KNN) cenderung bias terhadap sentimen positif karena proses klasifikasinya ditentukan oleh kedekatan jarak antar data. Dalam dataset yang mayoritas berlabel positif, sebagian besar tetangga terdekat berasal dari kelas tersebut, sehingga KNN lebih sering mengklasifikasikan ulasan sebagai positif meskipun isinya bernada negatif. Kondisi ini menyebabkan performa KNN menjadi kurang stabil dan akurasinya lebih rendah dibandingkan dengan *Naïve Bayes* dalam melakukan analisis sentimen teks.

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai analisis sentimen ulasan pengguna aplikasi Mobile Legends di Google Play Store menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbors*, dapat disimpulkan bahwa *Naïve Bayes* memiliki kinerja yang lebih stabil serta efisien dibandingkan KNN. Hasil pengujian menunjukkan bahwa metode ini menghasilkan klasifikasi dengan tingkat akurasi lebih tinggi sekaligus memerlukan waktu pemrosesan yang lebih singkat. Secara keseluruhan, temuan penelitian ini menegaskan bahwa *Naïve Bayes* lebih tepat diterapkan pada analisis sentimen berbasis teks karena mampu menjaga keseimbangan antara ketepatan prediksi dan efisiensi komputasi. Adapun KNN masih menghadapi kendala dalam mengenali variasi sentimen tertentu serta kurang efektif ketika diterapkan pada dataset berukuran besar. Oleh karena itu, *Naïve Bayes* direkomendasikan sebagai metode utama untuk penelitian sejenis yang berfokus pada klasifikasi sentimen pengguna aplikasi.

REFERENCES

- [1] M. B. Prayogi dan G. Masitoh, “Analisis Sentimen Ulasan Pengguna Aplikasi Alfagift Menggunakan Random Forest,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 10, no. 2, hal. 158–170, 2025, doi: 10.14421/jiska.2025.10.2.158-170.
- [2] T. Ramadha Triputra dan A. Rubhasy, “ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI FACEBOOK DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN K-NEAREST NEIGHBOR,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 3, hal. 4607–4615, Mei 2025, doi: 10.36040/jati.v9i3.13403.
- [3] D. Mualfah, A. Prihanti, R. Firdaus, dan Sunanto, “Analisis Sentimen Masyarakat Terhadap Kasus Pembobolan Data Nasabah Bank BSI Pada Twitter Menggunakan Metode Random Forest Dan Naïve Bayes,” *J. Fasilkom*, vol. 13, no. 3, hal. 614–620, 2024, doi: 10.37859/jf.v13i3.6478.
- [4] M. J. Palepa, N. Pratiwi, dan R. Q. Rohmansa, “Analisis Sentimen Masyarakat Tentang Pengaruh Politik Identitas Pada Pemilu 2024 Terhadap Toleransi Beragama Menggunakan Metode K - Nearest Neighbor,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 9, no. 1, hal. 389–401, 2024, doi: 10.29100/jupi.v9i1.4957.
- [5] S. N. S. Muslim, F. Nurdiansyah, dan A. Y. Rahman, “PERBANDINGAN ALGORITMA NAIVE BAYES DAN KNN DALAM ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI CAPCUT,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3S1, hal. 3588–3596, Okt 2024, doi: 10.23960/jitet.v12i3S1.5156.

- [6] S. Aulia dan I. Ismail, "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI OJOL THE GAME MENGGUNAKAN ALGORITMA NAÏVE BAYES," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, Agu 2024, doi: 10.23960/jitet.v12i3.5144.
- [7] R. N. Satrya, "Analisis Sentimen Terhadap Cryptocurrency Pada Platform Twitter Menggunakan Algoritma Support Vector Machine (Svm)," *KONSTELASI Konvergensi Teknol. dan Sist. Informasi*, vol. 4, no. 1, 2023., vol. 33, no. 1, hal. 1–12, 2023, doi: 10.70247/jumistik.v4i1.152.
- [8] A. Baita, Yoga Pristyanto, dan Nuri Cahyono, "Analisis Sentimen Mengenai Vaksin Sinovac Menggunakan Algoritma Support Vector Machine (SVM) dan K-Nearest Neighbor (KNN)," *Inf. Syst. J.*, vol. 4, no. 2, hal. 42–46, 2021, doi: 10.24076/infosjournal.2021v4i2.687.
- [9] S. N. Salombe, S. Alam, dan V. Aris, "Analisis Sentimen Ulasan Netizen Pada Aplikasi Disney + Hotstar Di Aplikasi Google Play Store Menggunakan Machine Learning," vol. 5, hal. 5968–5977, 2025, doi: <https://doi.org/10.31004/innovative.v5i1.15846>.
- [10] R. Damanhuri dan V. A. Husein, "Analisis Sentimen pada Ulasan Aplikasi Access by KAI Berbahasa Indonesia Menggunakan Word-Embedding dan Classical Machine Learning," *J. Masy. Inform.*, vol. 15, no. 2, hal. 97–106, 2024, doi: 10.14710/jmasif.15.2.62383.
- [11] I. D. Nugrayani, M. Hafid, dan D. Irmayanti, "Analisis Sentimen Terhadap Pemindahan Ibu Kota Negara (IKN) Pada Platform Twitter Menggunakan Metode Naive Bayes," *J. Apl. dan Teor. Ilmu Komput.*, vol. 6, no. 2, hal. 91–96, Apr 2025, doi: 10.17509/jatikom.v6i2.49105.
- [12] N. Alvionika, S. Faisal, R. Rahmat, dan A. F. N. Masruriyah, "Analisis Sentimen Pada Komentar Instagram Provider By.U Menggunakan Metode K-Nearest Neighbors (KNN)," *J. Algoritma*, vol. 21, no. 2, hal. 50–63, 2024, doi: 10.33364/algoritma/v.21-2.1672.
- [13] K. Mustaqim, F. A. Amaresti, dan I. N. Dewi, "Analisis Sentimen Ulasan Aplikasi PosPay untuk Meningkatkan Kepuasan Pengguna dengan Metode K-Nearest Neighbor (KNN)," *Edumatic J. Pendidik. Inform.*, vol. 8, no. 1, hal. 11–20, 2024, doi: 10.29408/edumatic.v8i1.24779.
- [14] N. Aditiya, P. Setiaji, dan S. Supriyono, "Sentiment Analysis of Public Satisfaction with the 'INFO BMKG' Application using Naive Bayes, SVM, and KNN," *SISTEMASI*, vol. 14, no. 3, hal. 1418, Mei 2025, doi: 10.32520/stmsi.v14i3.5223.
- [15] D. Audina *dkk.*, "TOMORO COFFEE MENGGUNAKAN ALGORITMA NAÏVE BAYES," vol. 8, no. 1, hal. 112–121, 2025, doi: <https://doi.org/10.36595/jire.v8i1>.
- [16] T. Wiratama Putra, A. Triayudi, dan A. Andrianingsih, "Analisis Sentimen Pembelajaran Daring Menggunakan Metode Naïve Bayes, KNN, dan Decision Tree," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 6, no. 1, hal. 20–26, 2022, doi: 10.35870/jtik.v6i1.368.
- [17] I. H. Kusuma dan N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, hal. 302–307, 2023, doi: 10.30591/jpit.v8i3.5734.
- [18] R. Merdiansah, S. Siska, dan A. Ali Ridha, "Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT," *J. Ilmu Komput. dan Sist. Inf.*, vol. 7, no. 1, hal. 221–228, 2024, doi: 10.55338/jikoms.v7i1.2895.
- [19] A. D. Adhi Putra, "Analisis Sentimen pada Ulasan pengguna Aplikasi Bibit Dan Bareksa dengan Algoritma KNN," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, hal. 636–646, 2021, doi: 10.35957/jatisi.v8i2.962.