



Sentiment Analysis of Using TikTok as a Learning Media Using the Naïve Bayes Classifiers Algorithm

Analisis Sentimen Penggunaan TikTok Sebagai Media Pembelajaran Menggunakan Algoritma Naïve Bayes Classifier

**Elsa Apriani^{1*}, Isa Faqihuddin Hanif², Falah Oktavianalisti³,
La Denna Hasri Monasari⁴, Indah Winarni⁵**

^{1,2,3,4,5}Program Studi Teknik Informatika, Fakultas Teknologi Industri dan Informatika,
Universitas Muhammadiyah Prof. Dr. Hamka, Indonesia

E-Mail: ¹elsaapriani14@gmail.com, ²isa@uhamka.ac.id, ³vianalisti@gmail.com,
⁴ladennahasrimonasari@gmail.com, ⁵indahwinar29@gmail.com

*Received Jun 22th 2024; Revised Jul 23th 2024; Accepted Jul 28th 2024
Corresponding Author: Elsa Apriani*

Abstract

Social media platforms such as TikTok have become popular among internet users, especially among the younger generation. On the other hand, social media also has the potential to be an effective learning medium. This study aims to analyze user sentiment regarding the use of TikTok as a learning medium by using the Naïve Bayes algorithm to analyze the sentiment data. A total of 176 data points were collected through a questionnaire distributed via Google Form. The obtained data were then analyzed using the Naïve Bayes Classifiers algorithm to categorize sentiments into positive, negative, and neutral. The analysis results showed that the model accuracy was 52.27%, with a precision value of 19.60% and a recall value of 50.83%. Although the accuracy and precision obtained are relatively low, these results provide an initial overview of users' perceptions of TikTok as a learning tool. This research is expected to provide further insights for developers of social media-based learning applications and help improve the quality of educational content on TikTok.

Keyword: Learning Media, Naïve Bayes Classifier, Sentiment Analysis, TikTok

Abstrak

Media sosial seperti TikTok telah menjadi platform yang populer di kalangan pengguna internet, terutama di kalangan generasi muda. Di sisi lain, media sosial juga memiliki potensi sebagai media pembelajaran yang efektif. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terkait penggunaan TikTok sebagai media pembelajaran yang menggunakan algoritma Naïve Bayes untuk menganalisis data sentimen tersebut. Sebanyak 176 data dikumpulkan melalui kuesioner yang disebar menggunakan Google Form. Data yang diperoleh kemudian dianalisis menggunakan algoritma Naïve Bayes Classifiers untuk mengkategorikan sentimen positif, negatif, dan netral. Hasil analisis menunjukkan bahwa akurasi model adalah 52.27%, dengan nilai precision sebesar 19.60%, dan recall sebesar 50.83%. Meskipun akurasi dan precision yang diperoleh relatif rendah, hasil ini memberikan gambaran awal mengenai persepsi pengguna terhadap TikTok sebagai alat pembelajaran. Penelitian ini diharapkan dapat memberikan wawasan lebih lanjut bagi pengembangan aplikasi pembelajaran berbasis media sosial dan membantu dalam meningkatkan kualitas konten edukatif di TikTok.

Kata Kunci: Analisis Sentimen, Media Pembelajaran, Naïve Bayes Classifier, TikTok

1. PENDAHULUAN

Pada tahun 2021 telah tersebar virus Covid-19 yang menyebabkan seluruh dunia melakukan pembatasan sosial tanpa terkecuali. Di Indonesia sendiri Covid-19 masuk pada maret 2020, yang menyebabkan pemerintah Indonesia mengeluarkan pembatasan sosial berskala besar. Di mana masyarakat dipaksa melakukan pembatasan sosial dan melakukan semua aktivitas di dalam rumah, hal ini membuat mereka tidak siap secara fisik dan mental untuk menghadapinya [1]

Media sosial merupakan sebuah sarana komunikasi dan pemasaran di mana orang dapat menuangkan pemikiran, ide dan pengalaman mereka dengan orang lain. Dengan menggunakan internet dan teknologi web, media sosial ini bertujuan untuk meningkatkan interaksi sosial manusia. Menurut laporan terbaru *We Are*

Social, WhatsApp mencapai 90,9% penggunaan media sosial di Indonesia, Instagram mencapai 85,3%, Facebook mencapai 81,6%, dan TikTok mencapai 73,5% [2]

Perusahaan asal Tiongkok, ByteDance pertama kali meluncurkan aplikasi berdurasi pendek yang bernama Douyin. Hanya dalam waktu satu tahun Douyin memiliki 100 juta pengguna dan 1 miliar tayangan video setiap hari [3]. TikTok, yang awalnya dikenal sebagai platform hiburan, telah berkembang menjadi alat yang dapat digunakan dalam berbagai bidang, termasuk pendidikan. Aplikasi ini memungkinkan pengguna untuk membuat dan berbagi video pendek yang dapat digunakan sebagai konten edukatif. TikTok awalnya dikenal sebagai platform hiburan, tetapi peneliti mulai melihat bahwa itu bisa menjadi alat pembelajaran yang bagus [4].

TikTok adalah media pembelajaran yang menarik dan menyenangkan, dan penggunaan aplikasi ini sebagai media pembelajaran yang dapat membantu para pengguna dalam belajar [5]. Salah satu cara untuk mendapatkan hasil belajar yang bagus adalah dengan menggunakan media pembelajaran. Keseluruhan kegiatan pengumpulan data atau informasi, pengolahan, penafsiran, dan pertimbangan untuk membuat keputusan tentang tingkatan hasil belajar yang dicapai setelah melakukan kegiatan belajar untuk mencapai tujuan pembelajaran. Dengan demikian, TikTok memiliki potensi untuk menjadi alat pembelajaran yang efektif, terutama bagi generasi muda yang lebih akrab dengan teknologi digital dan media sosial.

Dengan latar belakang yang ada, peneliti melakukan Analisis Penggunaan Media Sosial TikTok sebagai Media Pembelajaran Dengan Menggunakan Algoritma Naïve Bayes Classifiers. Dengan memanfaatkan komentar, like, dislike, dan interaksi pengguna lainnya, algoritma ini dapat membantu dalam memahami sentimen pengguna terhadap berbagai konten pembelajaran yang disajikan.

Analisis sentimen adalah proses otomatis yang berfokus pada pemahaman, ekstraksi, dan pengolahan data teks untuk menemukan sentiment dalam pernyataan atau pernyataan pendapat. Analisis sentiment dapat digunakan untuk menentukan apakah informasi dalam teks termasuk pendapat atau perspektif tentang masalah atau peristiwa yang bersifat positif atau negatif [6]. Analisis sentimen juga disebut sebagai penambangan opini yang bertujuan untuk mengetahui pendapat masyarakat terkait penggunaan TikTok sebagai media pembelajaran. Subyek dalam kumpulan data dapat berupa individu, organisasi, atau produk [7].

Algoritma Naïve Bayes adalah salah satu teknik statistik yang digunakan untuk klasifikasi data [8]. Algoritma ini berdasarkan pada Teorema Bayes dengan asumsi 'Naïve' bahwa semua fitur dalam dataset adalah independen. Keuntungan dari algoritma Naïve Bayes adalah efisiensinya dalam mengklasifikasikan data teks dan kemampuannya untuk bekerja dengan baik bahkan dengan jumlah data latihan yang relatif kecil.

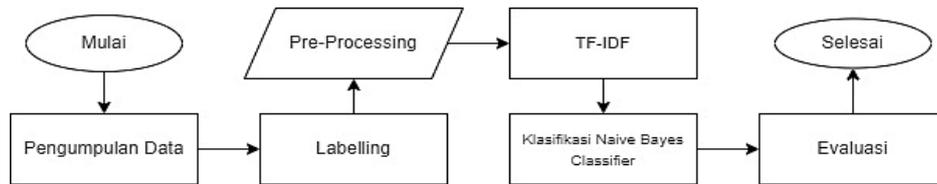
Terdapat beberapa penelitian sebelumnya terkait analisis sentimen menggunakan Naïve Bayes, di antaranya yaitu penelitian pertama yang ditulis oleh Havadz Faradian, Albar Rubhasy, dan Yunan Fauzi Wijaya dengan algoritma Naïve Bayes Classifier mengklasifikasikan sentimen terhadap penutupan TikTok Shop, mencapai akurasi 77%, precision 78%, recall 90%, dan F1-Score 83% [9]. Penelitian kedua yang ditulis oleh Nanda Try Luchia, Sephia Nazwa Auliani, Hani Handayani, Nilam Wahdiaz Azani, dan Rimelda Adha menggunakan algoritma yang sama untuk menganalisis 2110 tweet tentang TikTok di Twitter, menemukan mayoritas sentimen netral (95,78%) dengan akurasi 88% [10]. Penelitian ketiga yang ditulis oleh Friska Aditia Indriyani, Ahmad Fauzi, dan Sutan Faisa menganalisis 2000 ulasan di Google Play Store dengan Naïve Bayes dan SVM, di mana SVM lebih akurat (84%) dibanding Naïve Bayes (79%) [11]. Secara keseluruhan, penelitian ini menunjukkan bahwa Naïve Bayes efektif dalam analisis sentimen, namun SVM bisa memberikan akurasi lebih tinggi. Perbedaan hasil disebabkan oleh variasi data, platform, dan tujuan penelitian yang mempengaruhi performa algoritma.

Dalam penelitian ini menggunakan algoritma Naïve Bayes Classifier untuk menganalisis sentimen pengguna terkait penggunaan TikTok sebagai media pembelajaran. Algoritma ini dipilih karena kemampuannya dalam mengklasifikasikan data teks dengan efisien, serta memberikan wawasan mengenai persepsi pengguna terhadap konten edukatif di TikTok. Melalui analisis sentimen ini, diharapkan dapat diperoleh gambaran yang lebih jelas mengenai efektivitas TikTok sebagai alat bantu pembelajaran, serta bagaimana platform ini dapat dioptimalkan untuk mendukung pendidikan di masa depan.

Hasil analisis ini diharapkan dapat memberikan arahan yang berguna bagi pembuat konten, pendidik, dan peneliti untuk meningkatkan kualitas materi pembelajaran di TikTok, serta memperluas pemahaman tentang peran media sosial dalam pembelajaran informal.

2. METODE PENELITIAN

Penelitian ini berfokus pada metode yang digunakan dalam penelitian untuk menganalisis penggunaan TikTok sebagai media pembelajaran menggunakan Algoritma Naïve Bayes. Penelitian ini dilakukan melalui beberapa tahapan utama yaitu pengumpulan data, pelabelan, praproses data, ekstraksi fitur dengan menggunakan TF-IDF, klasifikasi menggunakan Naïve Bayes, dan evaluasi hasil. Alur penelitian dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Proses pengumpulan data pada penelitian ini memakai data yang diperoleh eksklusif tidak melalui perantara. Berdasarkan hal tersebut peneliti mengumpulkan secara khusus untuk memecahkan konflik penelitian. Pengumpulan data tersebut menggunakan kuesioner yang berupa susunan daftar pertanyaan secara tertulis yang bertujuan untuk memperoleh data yang diungkapkan oleh responden.

2.2. Labelling

Labelling adalah proses untuk menentukan kelas positif atau negatif dari ulasan pada dataset. Labelling dapat dilakukan secara manual atau otomatis. Tujuan dari labelling adalah untuk mengetahui representasi dari data dan mengenali sentimen data [12].

2.3. Pre-Processing

Tahap ini merupakan langkah penting yang mana data mentah diolah kemudian diubah menjadi bentuk yang lebih terstruktur, atau yang disebut *Pre-Processing* [13].

1. Case Folding (Transform Cases)

Case Folding (Transform Cases) adalah proses mengubah semua huruf dalam teks menjadi huruf kecil (lowercase) [14]. Langkah ini penting untuk memastikan konsistensi dalam pengolahan data teks, menghindari perbedaan yang disebabkan oleh huruf besar dan kecil.

2. Tokenization (Tokenize)

Tokenization adalah proses memecah teks menjadi unit-unit yang lebih kecil yang disebut token [15]. Setiap token bisa berupa kata, frasa, atau simbol. Proses ini penting karena memungkinkan analisis lebih detail terhadap setiap elemen dalam teks. Tokenization biasanya dilakukan dengan menggunakan pemisah kata seperti spasi atau tanda baca.

3. Stopwords Removal (Filter Stopwords)

Stopwords Removal adalah proses menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis teks, seperti "dan", "yang", "di", "ke", dan "dari" [16]. Menghapus stopwords meningkatkan kualitas analisis dengan mengurangi kebisingan dalam data, menggunakan daftar stopwords dari library NLP seperti Kaggle.

4. Filter Token (by Length)

Filter Token (by Length) menghilangkan token berdasarkan panjang karakternya. Token yang terlalu pendek atau panjang dihapus karena kurang informatif. Biasanya, token dengan panjang kurang dari 4 karakter atau lebih dari 25 karakter dihapus, membantu mengeliminasi kesalahan pengetikan atau kata-kata yang tidak relevan [17].

2.4. TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) merupakan tahapan pre-processing dalam ekstraksi dataset. Selanjutnya TF-IDF akan melakukan pembobotan pada sebuah kata (term) dalam dokumen [18]. Oleh karena itu hal ini menekankan bahwa nilai TF-IDF sebuah kata akan meningkat seiring dengan frekuensi kemunculan dalam sebuah dokumen berita. Selain itu TF-IDF digunakan untuk menyeimbangkan bobot kata-kata dalam berita, sehingga kata yang spesifik memiliki arti yang lebih besar pada sebuah dokumen [19].

2.5. Klasifikasi Naïve Bayes Classifier

Konsep dasar dari klasifikasi adalah semacam struktur data dengan jenis analisis data dasar yang melibatkan teknik penentuan kelompok berdasarkan data sebelumnya. Naïve Bayes adalah teknik statistik untuk penalaran induktif untuk masalah klasifikasi. Keuntungan dari teknik ini adalah hanya memerlukan sejumlah kecil data latihan untuk menentukan perkiraan parameter yang diperlukan untuk proses klasifikasi [20] Persamaan umum Naïve Bayes Classifier dapat dilihat pada persamaan 1.

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \tag{1}$$

Keterangan:

- P (A|B) : Probabilitas bersyarat A yang diberikan oleh B
- P (B|A) : Probabilitas bersyarat B yang diberikan oleh A
- P (A) : Probabilitas kejadian A
- P (B) : Probabilitas kejadian B

2.6. Evaluasi

Proses ilmiah yang sistematis yang dikenal sebagai evaluasi bertujuan untuk menentukan apakah hasil, atau efektifitas, suatu program atau proyek, secara objektif konsisten dengan tujuan yang direncanakan untuk mencapainya Setelah itu, membuat dan menetapkan kebijakan dengan mempertimbangkan manfaat dan keuntungan program terlebih dahulu [21].

1. *Accuracy*

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \tag{2}$$

2. *Precision*

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

3. *Recall*

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

3. HASIL DAN PEMBAHASAN

Proses pengumpulan data penelitian dilakukan dengan cara menyebarkan kuesioner kepada pengguna aplikasi TikTok. Kuesioner disebarkan kepada pengguna media sosial TikTok agar mereka mengisi berdasarkan experience. Tujuan dikumpulkannya data tersebut yaitu untuk mengetahui seberapa efektif atau tidak aplikasi TikTok sebagai media pembelajaran. Proses pengumpulan data dapat dilihat pada gambar 2.



Gambar 2. Proses Pengumpulan Data

Pengumpulan data melalui kuesioner terkumpul sebanyak 176 data. Dataset tersebut kemudian akan disimpan ke dalam file dengan format CSV. Hasil tampilan data dapat dilihat pada tabel 1.

Tabel 1. Pengumpulan Data

Apakah Anda pernah menggunakan TikTok untuk tujuan pembelajaran?	Seberapa efektif menurut Anda TikTok sebagai media pembelajaran?	Apakah Anda lebih suka pembelajaran melalui TikTok daripada metode tradisional?	Bagaimana pendapat Anda tentang kualitas konten pembelajaran di TikTok?	Apakah Anda merasa konten pembelajaran di TikTok mudah dipahami?
Pernah	Cukup Efektif	Lebih suka metode tradisional	Cukup baik dan menarik	Sangat Mudah
Tidak Pernah	Sangat Efektif	Tradisional	Sangat bagus tetapi sangat sedikit	Iya
Belum pernah	Tidak terlalu efektif	Lebih bagus metode tradisional	Tidak terlalu buruk	Ya
Pernah	Cukup efektif	Tidak juga	Cukup membantu	Cukup mudah dipahami
Pernah	Lumayan efektif	Lumayan	Singkat, jelas, dan tidak membosankan	Iya

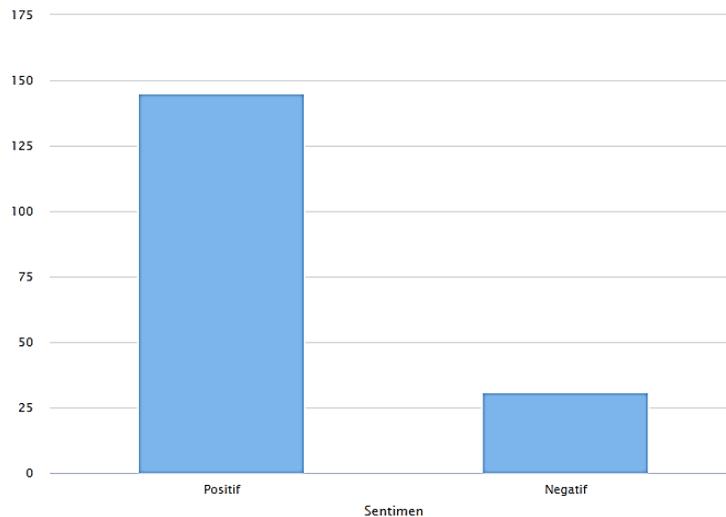
3.1. Labelling

Proses pelabelan pada data kuesioner dilakukan secara manual. Pelabelan dibagi menjadi positif dan negatif. Hasil labelling dapat dilihat pada tabel 2.

Tabel 2. Hasil *Labelling*

No	Text	Sentimen
1	Pernah cukup efektif Lebih suka melalui TikTok Lumayan, semua tergantung dari konten kreator Iya	Positif
2	Belum pernah Tidak terlalu efektif Lebih bagus metode tradisional Tidak terlalu buruk Ya	Negatif

Jumlah data yang didapat untuk labelling yaitu sebanyak 176 data. Jumlah untuk sentimen positif yaitu sebanyak 145 data, sedangkan untuk sentimen negatif yaitu 31 data. Di bawah ini terdapat diagram yang menunjukkan hasil perbandingan antara sentimen positif dan sentimen negatif. Perbandingan sentimen positif dan negatif dapat dilihat pada gambar 3.



Gambar 3. Perbandingan Sentimen Positif dan Negatif

3.2. Pre-Processing

Tahap *pre-processing* diawali dengan operator *tokenize*, *transforms cases*, *filter stopwords (dictionary)*, dan *filter tokens (by length)*. *Tokenize* digunakan untuk memisahkan kata per kata. *Transform Cases* digunakan untuk mengubah huruf besar menjadi huruf kecil. Kemudian, *filter stopwords* digunakan untuk menghilangkan kata-kata yang tidak penting, dan *filter tokens* digunakan untuk menghilangkan kata-kata dengan jumlah huruf terbatas atau terlalu banyak huruf.

3.2.1. Case Folding (Transform Cases)

Proses case folding dilakukan untuk mengubah kalimat pada data menjadi huruf kecil atau lower case. Misalnya, kata "TikTok" dan "TikTok" dianggap sama setelah proses case folding. Hal ini membantu dalam mengurangi jumlah fitur unik dalam dataset dan meningkatkan efisiensi analisis. Hasil case folding dapat dilihat pada tabel 3.

Tabel 3. Hasil *Transform Cases*

No	Text	<i>Transform Cases</i>
1	Pernah cukup efektif Lebih suka melalui TikTok Lumayan, semua tergantung dari konten kreator Iya	pernah cukup efektif lebih suka melalui tiktok lumayan, semua tergantung dari konten kreator iya
2	Belum pernah Tidak terlalu efektif Lebih bagus metode tradisional Tidak terlalu buruk Ya	belum pernah tidak terlalu efektif lebih bagus metode tradisional tidak terlalu buruk ya

3.2.2. Tokenization (Tokenize)

Proses tokenization merupakan proses untuk memotong kalimat menjadi beberapa kata yang lebih kecil. Hasil tokenize dapat dilihat pada tabel 4.

Tabel 4. Hasil *Tokenize*

No	Text	<i>Tokenize</i>
1	Pernah cukup efektif Lebih suka melalui TikTok Lumayan, semua tergantung dari konten kreator Iya	'Pernah', 'cukup', 'efektif', 'Lebih', 'suka', 'melalui', 'TikTok', 'Lumayan', 'semua', 'tergantung', 'dari', 'konten', 'kreator', 'Iya'
2	Belum pernah Tidak terlalu efektif Lebih bagus metode tradisional Tidak terlalu buruk Ya	'Belum', 'pernah', 'Tidak', 'terlalu', 'efektif', 'Lebih', 'bagus', 'metode', 'tradisional', 'Tidak', 'terlalu', 'buruk', 'Ya'

3.2.3. Stopwords Removal (Filter Stopwords)

Proses stopwords removal yaitu proses untuk menghapus kata umum berdasarkan kaggle. Hasil filter stopwords dapat dilihat pada tabel 5.

Tabel 5. Hasil *Stopwords Removal*

No	Text	<i>Filter Stopwords</i>
1	Pernah cukup efektif Lebih suka melalui TikTok Lumayan, semua tergantung dari konten kreator Iya	cukup efektif Lebih suka melalui TikTok Lumayan, semua tergantung dari konten kreator Iya
2	Belum pernah Tidak terlalu efektif Lebih bagus metode tradisional Tidak terlalu buruk Ya	Belum pernah Tidak terlalu efektif Lebih bagus metode tradisional Tidak terlalu buruk Ya

3.2.4. Filter Token (by Length)

Proses filter token by length merupakan proses untuk membatasi jumlah kata yang digunakan dengan batas minimal 4 karakter dan maksimal 25 karakter. Hasil filter token dapat dilihat pada tabel 6.

Tabel 6. Hasil *Filter Token*

No	Text	<i>Filter Token</i>
1	Pernah cukup efektif Lebih suka melalui TikTok Lumayan, semua tergantung dari konten kreator Iya	Pernah cukup efektif Lebih suka melalui TikTok
2	Belum pernah Tidak terlalu efektif Lebih bagus metode tradisional Tidak terlalu buruk Ya	Belum pernah Tidak terlalu efektif Tidak terlalu buruk

3.3. TF-IDF

Setelah tahap *pre-processing* data selesai, langkah berikutnya adalah pembobotan data menggunakan metode TF-IDF. Metode ini dikenal sebagai frekuensi waktu dokumen invers. Tujuan proses TF-IDF adalah untuk menampilkan berapa banyak kata dalam dataset yang sering muncul. TF-IDF akan membatasi sebuah kata (term) dalam dokumen, menekankan bahwa nilai TF-IDF sebuah kata meningkat seiring dengan frekuensi kemunculannya dalam dokumen berita. Selain itu, TF-IDF digunakan untuk menyeimbangkan bobot kata-kata dalam dokumen berita, sehingga kata-kata tertentu memiliki nilai yang lebih besar dalam dokumen. Proses TF-IDF dapat dilihat pada gambar berikut. Hasil TF-IDF dapat dilihat pada tabel 7.

Tabel 7. Hasil TF-IDF

Word	In Documents	Total	In Class (Positif)	In Class (Negatif)
efektif	137	141	120.0	21.0
mudah	76	93	80.0	13.0
pembelajaran	40	71	62.0	9.0
tiktok	48	68	51.0	17.0
konten	32	61	46.0	15.0
bagus	50	51	39.0	12.0
dipahami	39	45	39.0	6.0
lumayan	32	45	39.0	6.0
tradisional	33	35	29.0	6.0
suka	32	32	28.0	4.0
metode	27	31	26.0	5.0
menarik	25	30	28.0	2.0
tergantung	26	27	25.0	2.0
belajar	18	21	19.0	2.0
informasi	14	19	17.0	2.0
media	13	18	16.0	2.0

Pada tabel 7, kata-kata seperti "efektif", "mudah", dan "pembelajaran" muncul dengan banyak dalam dokumen dan secara keseluruhan, menunjukkan bahwa kata-kata ini sering muncul saat melakukan analisis. Khususnya, kata "efektif" muncul sebanyak 137 kali dalam dokumen dengan total 141 kali, dan paling sering

		Predicted Values	
		Pred. Positif	Pred. Negatif
Predicted Values	Pred. Positif	76	69
	Pred. Negatif	15	16

Gambar 5. Grafik *Confusion Matrix*

3.5. Evaluasi

Pada penelitian ini, model Naïve Bayes digunakan untuk melakukan klasifikasi sentimen terhadap ulasan aplikasi TikTok. Proses evaluasi menunjukkan bahwa model memiliki tingkat akurasi sebesar 52.27%. Ini berarti model dapat memprediksi dengan benar sekitar 52.27% dari total 176 sampel data yang diuji. Precision untuk kelas positif mencapai 83.52%, yang menunjukkan bahwa dari semua prediksi positif yang dihasilkan model, 83.52% di antaranya benar-benar positif. Namun, precision untuk kelas negatif hanya sebesar 18.82%, menandakan bahwa hanya 18.82% dari prediksi negatif yang benar-benar negatif. Recall untuk kelas positif adalah 52.41%, sementara recall untuk kelas negatif adalah 51.61%. Ini menunjukkan bahwa model memiliki kemampuan yang sedang dalam mengenali data positif dan negatif, tetapi lebih baik dalam mengidentifikasi data positif dibandingkan data negative.

4. KESIMPULAN

Berdasarkan hasil evaluasi, dapat disimpulkan bahwa model Naïve Bayes yang digunakan dalam penelitian ini memiliki performa yang kurang memuaskan dengan akurasi keseluruhan sebesar 52.27%. Meskipun precision untuk kelas positif cukup tinggi, precision untuk kelas negatif sangat rendah. Hal ini mengindikasikan bahwa model cenderung lebih efektif dalam mengidentifikasi ulasan positif dibandingkan ulasan negatif. Recall yang diperoleh menunjukkan bahwa model mampu mengenali lebih dari separuh data positif dan negatif dengan benar. Untuk meningkatkan performa model, diperlukan peningkatan pada metode pre-processing data dan mungkin penggunaan algoritma yang lebih kompleks atau metode ensemble untuk mencapai hasil yang lebih akurat dan seimbang dalam klasifikasi sentimen.

REFERENSI

- [1] B. Yip and V. Perasso, "Asal Covid-19: Apakah kita perlu tahu dari mana asal virus corona ini?," <https://www.bbc.com/indonesia/dunia-57590872>.
- [2] C. M. Annur, "Ini Media Sosial Paling Banyak Digunakan di Indonesia Awal 2024," <https://databoks.katadata.co.id/datapublish/2024/03/01/ini-media-sosial-paling-banyak-digunakan-di-indonesia-awal-2024>.
- [3] A. Wahyu Setiawan and M. B. N. Ariani, "Determinasi Faktor Yang Mempengaruhi Indeks Pembangunan Manusia Provinsi Jawa Barat Tahun 2015 – 2019," *Jurnal Indonesia Sosial Sains*, vol. 3, no. 1, pp. 1–9, Jan. 2022, doi: 10.36418/jiss.v3i1.505.
- [4] W. A. Alvianto, J. Amanullah, and L. Santoso, "Jurnal Kajian Ilmu seni," 2024.
- [5] L. Novia, A. Hikmat, and I. Safi'I, "Pemanfaatan Aplikasi Tiktok Sebagai Media Pembelajaran Pidato," *Jurnal Inovasi Global*, vol. 2, no. 2, pp. 307–311, Mar. 2024, doi: 10.58344/jig.v2i2.64.
- [6] J. Khab Sulaiman Dalam, A. Oktavia Praneswara, N. Cahyono, and U. Amikom Yogyakarta, "Analisis Sentimen Ulasan Aplikasi TikTok Shop Seller Center di Google Playstore Menggunakan Algoritma Naive Bayes," *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 6, p. 3925.
- [7] E. Andrian and A. Rahman Isnain, "JURNAL MEDIA INFORMATIKA BUDIDARMA Analisis Sentimen Masyarakat Terhadap TikTok Shop di Twitter Menggunakan Metode Naive Bayes Classifier," 2024, doi: 10.30865/mib.v8i2.7530.
- [8] A. Komarudin, A. Meutia Hilda, and C. Author, "Analisis Sentimen Ulasan Aplikasi Identitas Kependudukan Digital Pada Play Store Menggunakan Metode Naive Bayes," 2024. [Online]. Available: <http://jurnal.bsi.ac.id/index.php/co-science>
- [9] H. Faradian, A. Rubhasy, and Y. F. Wijaya, "ANALISIS SENTIMEN TERHADAP PENUTUPAN TIKTOK SHOP MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER PADA MEDIA SOSIAL X," *Scientica Jurnal Ilmiah Sain dan Teknologi*, vol. 2, no. 4, pp. 150–163, Mar. 2024.

-
- [10] N. T. Luchia, S. N. Auliani, H. Handayani, N. W. Azani, and R. Adha, "Analisis Sentimen Pengguna Twitter Terhadap Aplikasi TikTok Menggunakan Algoritma Naïve Bayes Clasifier," *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat*, pp. 100–104, Aug. 2023.
- [11] F. A. Indriyani, A. Fauzi, and S. Faisal, "Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine," *TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika*, vol. 10, no. 2, pp. 176–184, Jul. 2023.
- [12] F. Setya Ananto and F. N. Hasan, "Implementasi Algoritma Naïve Bayes Terhadap Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store," *Jurnal ICT: Information Communication & Technology*, vol. 23, no. 1, pp. 75–80, 2023, [Online]. Available: <https://ejournal.ikmi.ac.id/index.php/jict-ikmi>
- [13] Friska Aditia Indriyani, Ahmad Fauzi, and Sutan Faisal, "Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine," *TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika*, vol. 10, no. 2, pp. 176–184, Jul. 2023, doi: 10.37373/tekno.v10i2.419.
- [14] A. Erfina, M. Fani Al-shufi, C. Kaler, K. Cisaat, and K. Sukabumi, "ANALISIS SENTIMEN APLIKASI JASA KURIR DI PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES," *Jurnal Sistem Informasi dan Informatika (SIMIKA) P-ISSN*, vol. 5, pp. 2622–6901, 2022.
- [15] A. S. Rahayu, A. Fauzi, and R. Rahmat, "Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Spotify," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, p. 349, Dec. 2022, doi: 10.30865/json.v4i2.5398.
- [16] M. Khoirul, U. Hayati, and O. Nurdian, "ANALISIS SENTIMEN APLIKASI BRIMO PADA ULASAN PENGGUNA DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES," 2023.
- [17] R. Faiz Ananda, A. Syahri, and F. N. Hasan, "SENTIMENT ANALYSIS OF CUSTOMER SATISFACTION IN GOJEK AND GRAB APPLICATION REVIEWS USING THE NAIVE BAYES ALGORITHM," *Jurnal Teknik Informatika (JUTIF).1*, vol. 5, no. 1, pp. 233–241, 1680, doi: 10.52436/1.jutif.2024.5.1.1680.
- [18] "Analisis Sentimen Terhadap Pembelajaran Daring dengan Algoritma Naïve Bayes Classifier."
- [19] J. Homepage, N. C. Agustina, D. Herlina Citra, W. Purnama, C. Nisa, and A. Rozi Kurnia, "MALCOM: Indonesian Journal of Machine Learning and Computer Science The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews on Google Play Store Implementasi Algoritma Naïve Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 1, pp. 47–54, 2022.
- [20] E. Apriliyani and Y. Salim, "Analisis performa metode klasifikasi Naïve Bayes Classifier pada Unbalanced Dataset," *Indonesian Journal of Data and Science (IJODAS)*, vol. 3, no. 2, pp. 47–54, 2022.
- [21] R. A. Saputra, D. P. Ray, and F. Irwiensyah, "KLIK: Kajian Ilmiah Informatika dan Komputer Analisis Sentimen Aplikasi Tokocrypto Berdasarkan Ulasan Pada Google Play Store Menggunakan Metode Naïve Bayes," *Media Online*, vol. 4, no. 4, 2024, doi: 10.30865/klik.v4i4.1707.