

Analisis Sentimen Perbedaan Pendapat Netizen Indonesia Terhadap Penutupan Tiktok Shop Menggunakan Algoritma Naïve Bayes

Eko Kurnianto, Dimas Febriawan*

²Fakultas Teknologi Industri dan Informatika, Teknik Informatika, Universitas Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia

Email: ¹ekok203@email.com, ^{2,*}dimas.febriawan@uhamka.ac.id

Email Penulis Korespondensi: dimas.febriawan@uhamka.ac.id

Submitted: 15/12/2023; Accepted: 28/12/2023; Published: 29/12/2023

Abstrak—Penelitian ini menggunakan algoritma Naïve Bayes untuk menganalisis sentimen netizen Indonesia terkait penutupan TikTok Shop. Penelitian ini berfokus pada analisis perbedaan pendapat yang tersebar di platform media sosial. Data yang didapatkan dari social media seperti Youtube, Tiktok, dan Threads. Terdapat data yang nantinya akan digunakan pada penelitian ini dengan jumlah 1366 data. Kemudian, terdapat 987 data positif dan 379 data negatif. Setelah dilakukannya riset, maka akan diperoleh hasil dengan akurasi 86.97% pada percobaan pertama yang tidak menggunakan operator Split Data, dan akurasi 89.23% pada percobaan kedua yang telah menggunakan operator Split Data. Kemudian dari pada hasil analisis ini menyatakan variasi sentimen yang signifikan di antara netizen Indonesia terkait penutupan TikTok Shop. Beberapa kelompok netizen mungkin mengekspresikan kekecewaan atau ketidaksetujuan sementara yang lain mungkin menunjukkan dukungan terhadap keputusan tersebut. Analisis juga mengidentifikasi faktor-faktor kunci yang memengaruhi perbedaan pendapat, seperti pengalaman pengguna, ekspektasi terhadap platform, dan dampak ekonomi. Dari akan itu, dalam penelitian ini berkontribusi didalam bidang analisis sentimen dan pengolahan bahasa alami yang menerapkan prosedur pemecahan sehingga dapat di klasifikasi pada data komentar netizen pada platform.

Kata Kunci: Analisis Sentimen; Algoritma Naïve Bayes; Tiktok Shop; Rapid Miner; Identifikasi

Abstract—This research uses the Naïve Bayes algorithm to analyze the sentiments of Indonesian netizens regarding the closure of the TikTok Shop. This research focuses on analyzing differences of opinion spread on social media platforms. Data obtained from social media such as Youtube, Tiktok, and Threads. There is data that will later be used in this research with a total of 1366 data. Then, there were 987 positive data and 379 negative data. After conducting research, results will be obtained with an accuracy of 86.97% in the first experiment which does not use the Split Data operator, and an accuracy of 89.23% in the second experiment which uses the Split Data operator. Then the results of this analysis reveal significant variations in sentiment among Indonesian netizens regarding the closure of the TikTok Shop. Some groups of netizens may express disappointment or disapproval while others may show support for the decision. The analysis also identified key factors influencing dissent, such as user experience, expectations of the platform and economic impact. Due to this, this research contributes to the field of sentiment analysis and natural language processing which applies splitting procedures so that netizen comment data on the platform can be classified.

Keywords: Sentiment Analysis; Naïve Bayes Algorithm; TikTok Shop; Fast Miner; Identification

1. PENDAHULUAN

Pada era digital ini, media sosial telah menjadi salah satu wadah utama bagi masyarakat untuk berbagi pendapat, informasi, dan interaksi secara online. Dalam beberapa tahun terakhir, perkembangan media sosial di Indonesia menjadi sangat signifikan. Netizen Indonesia aktif menggunakan berbagai platform media sosial untuk menyampaikan pendapat, berbagi informasi, dan terlibat dalam berbagai diskusi online. Tiktok, sebagai salah satu platform media sosial yang paling populer, memiliki pengguna aktif yang signifikan di Indonesia. Salah satu platform yang sangat populer di Indonesia adalah Tiktok, sebuah aplikasi berbagi video pendek yang telah meraih popularitas besar di kalangan netizen Indonesia [1].

Dalam beberapa waktu terakhir, Tiktok membuat sebuah fitur yang memungkinkan pengguna dapat menjual dan membeli produk secara langsung kepada pengikut mereka melalui aplikasi. Setelah beberapa hari fitur tersebut rilis di aplikasi Tiktok dan diberi nama Tiktok Shop. Tiktok Shop ini menjadi topik pembicaraan hangat di kalangan netizen Indonesia. Karena banyak sekali benefit yang di dapat jika menggunakan Tiktok Shop. Oleh, karena itu banyak netizen Indonesia menjadikan fitur Tiktok Shop ini sebagai tempat utama untuk melakukan transaksi barang jika dibandingkan dengan e-commerce yang sudah ada seperti Shoppe, Tokopedia, Bukalapak atau yang lainnya [2].

Tiktok merupakan platform media sosial yang mayoritasnya menyuguhkan konten hiburan. Dan Tiktok diciptakan sebagai media sosial untuk hiburan sekaligus belanja. Belanja dari menonton konten hiburan di Tiktok lebih mempunyai daya tarik tersendiri dibandingkan dengan e-commerce lain [3]. Harga lebih murah juga menjadi salah satu alasan kenapa masyarakat Indonesia lebih memilih Tiktok Shop dibandingkan e-commerce lainnya. Hal itu dikarenakan di dalam Tiktok Shop terdapat fitur canggih yang dimiliki oleh Tiktok Shop yaitu berupa live streaming yang menjadikan customer dapat membeli barang yang diinginkan dengan harga murah. Tiktok merupakan platform media sosial yang mayoritas menyuguhkan konten hiburan. Berbeda dengan Shopee yang sengaja diciptakan khusus untuk menjadi sebuah e-commerce untuk jual beli barang yang dibutuhkan penggunanya [4].

Salah satu peristiwa yang memicu perbincangan di kalangan netizen Indonesia adalah penutupan TikTok Shop. TikTok Shop awalnya menjadi platform yang memungkinkan pengguna TikTok untuk menjual produk langsung melalui aplikasi. Keputusan untuk menutup TikTok Shop dapat menciptakan perbedaan pendapat di antara netizen. Beberapa netizen mungkin mendukung keputusan ini atas alasan tertentu, seperti keamanan atau privasi, sementara yang lain mungkin mengecamkannya karena melihatnya sebagai pembatasan peluang bisnis.

Terdapat beberapa algoritma yang digunakan untuk melakukan analisis sentimen, seperti Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbors, Decision Trees [5]. Dalam penelitian ini, peneliti memilih menggunakan algoritma Naïve Bayes karena algoritma ini memiliki beberapa keunggulan yang memudahkan implementasi dan mengurangi kesulitan dalam perhitungan. Salah satu keunggulan Naïve Bayes adalah kemampuannya untuk bekerja dengan baik pada data yang terdistribusi tidak merata atau tidak seimbang. Ini menjadikannya efektif dalam situasi di mana jumlah sampel dalam kelas tertentu jauh lebih banyak atau lebih sedikit dibandingkan dengan kelas lain. Selain itu, Naïve Bayes dikenal memiliki kemampuan untuk menangani data besar dengan efisien.

Algoritma ini merupakan algoritma machine learning yang digunakan untuk masalah klasifikasi, berdasarkan teorema probabilitas Bayes. Keberhasilannya terbukti dalam berbagai aplikasi seperti penyaringan spam, analisis sentimen, dan klasifikasi artikel berita, terutama pada dataset dengan dimensi tinggi seperti pada penelitian ini. Dengan menerapkan algoritma ini pada data yang diambil dari platform-platform media sosial yang mencerminkan keragaman pandangan, penelitian ini bertujuan untuk memberikan wawasan yang lebih mendalam mengenai pendapat netizen Indonesia tentang penutupan TikTok Shop. Dengan hasil analisis sentimen ini, diharapkan dapat memberikan informasi berharga kepada pihak berkepentingan, termasuk retailer, pemerintah, dan konsumen, sehingga mereka dapat membuat keputusan yang lebih baik terkait penutupan TikTok Shop. Melalui kontribusinya, penelitian ini diharapkan dapat memberikan pemahaman yang lebih baik tentang peran media sosial dalam membentuk opini publik serta kebijakan terkait isu-isu lingkungan dan Kesehatan.

Penelitian dengan cakupan yang serupa dilakukan oleh Desy Oktaheriyani [6] Dalam penelitian ini, peneliti menggunakan algoritma Naïve Bayes dan menerapkan dua operator yang berbeda, yaitu split data dan cross-validation data, untuk menganalisis sentimen. Pada tahap penelitian, data yang dikumpulkan terdiri dari 359 komentar yang diklasifikasikan ke dalam dua kelas, yaitu kelas positif dan kelas negatif. Hasil akurasi dari penggunaan algoritma Naïve Bayes dengan operator split data mencapai 86.29%, sedangkan dengan operator cross-validation data mencapai 86.23%.

Penelitian dengan cakupan yang serupa dilakukan oleh Artanti Inez Tanggraeni [7] dengan memanfaatkan algoritma Naïve Bayes, dilakukan analisis sentimen terhadap 65 data ulasan pengguna yang dikumpulkan melalui teknik Web Scraping. Hasil analisis sentimen ini menunjukkan tingkat akurasi yang sangat baik. Dengan tingkat akurasi yang mendekati 100%, mencapai 87%, hal ini menegaskan tingkat keberhasilan yang signifikan dalam analisis sentimen menggunakan klasifikasi Naïve Bayes dengan penerapan pembobotan TF-IDF. Hasil precision, yang merupakan rasio antara data ulasan positif yang diprediksi dengan tepat dan keseluruhan data yang diprediksi, juga menunjukkan kinerja yang sangat baik, mencapai 83%. Selain itu, hasil recall sebesar 87% menunjukkan sensitivitas prediksi terhadap data ulasan positif berdasarkan data yang sebenarnya.

Penelitian dengan cakupan yang serupa dilakukan oleh Sevira Chairunnisa Nandaresta [8] dengan mencermati pandangan masyarakat terhadap Tiktok Shop dan Shopee serta menilai sejauh mana kedua e-commerce tersebut memberikan manfaat kepada masyarakat melalui penerapan metode klasifikasi Algoritma Naïve Bayes. Melalui analisis data yang terdiri dari 800 ulasan, di mana 400 di antaranya berasal dari ulasan Tiktok Shop dan 400 dari ulasan Shopee. Dari total 800 ulasan yang dianalisis, terungkap bahwa hanya 300 ulasan yang menyoroti tanggapan positif masyarakat terhadap Tiktok Shop dan Shopee. Hasil analisis ini dapat memberikan pemahaman lebih mendalam mengenai persepsi dan pengalaman pengguna terhadap Tiktok Shop dan Shopee dalam konteks layanan e-commerce.

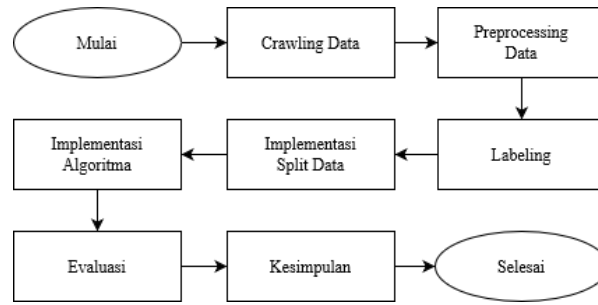
Penelitian dengan cakupan yang serupa dilakukan oleh Soleh Jafar Sidiq [9] yang menyelidiki jumlah ulasan yang mencerminkan tanggapan positif dan negatif terhadap layanan Tiktok Shop. Sebanyak 64,4% dari data tersebut menunjukkan sentimen positif, sementara 14,4% menunjukkan sentimen negatif. Dalam kategori sentimen positif, tingkat precision mencapai 0,88%, recall sebesar 0,74%, dan f1-score sebesar 0,80%. Untuk sentimen yang merujuk pada aspek negatif, akurasi mencapai 0,88%, tingkat penyadaran (recall) mencapai 0,96%, dan skor f1 mencapai 0,91%. Dengan demikian, penelitian ini memberikan gambaran rinci mengenai evaluasi sentimen pelanggan terhadap layanan Tiktok Shop berdasarkan output analisis data yang telah dilakukan.

2. METODOLOGI PENELITIAN

2.1. Tahapan Penelitian

Langkah-langkah yang akan diterapkan dalam penelitian ini mencakup: Pengumpulan Data (Crawling Data), Pra-pemrosesan Data (Preprocessing Data), Pelabelan (Labeling), Implementasi Split Data, Implementasi Algoritma Naïve Bayes, dan Evaluasi Hasil. Evaluasi melibatkan perbandingan antara percobaan pertama tanpa menggunakan operator Pembagian Data dan percobaan kedua yang melibatkan operator Pembagian Data.

Kesimpulan akan diambil berdasarkan perbandingan hasil kedua percobaan tersebut. Rincian tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

- a) **Crawling Data**
Pada fase ini, Proses mencakup penghimpunan data, meskipun metode ini belum umum digunakan dalam penelitian lain. Penelitian ini fokus pada akuisisi komentar secara daring melalui platform sosial media seperti Youtube, Tiktok, dan Threads. [10].
- b) **Preprocessing Data**
Preprocessing merupakan tahapan awal dalam analisis data yang bertujuan untuk menyiapkan dan membersihkan data mentah sehingga dapat dimanfaatkan secara efisien dalam proses analisis berikutnya. Fokus utama dari preprocessing adalah meningkatkan kualitas data, menghilangkan gangguan, dan mengoptimalkan data agar sesuai dengan persyaratan analisis yang akan dilakukan [11].
- c) **Labeling**
Data yang akan digunakan merupakan data yang diperoleh dari platform media sosial online seperti Youtube, Tiktok, dan Threads. Selanjutnya, dilakukan proses pelabelan secara manual dengan mengkategorikan ulasan ke dalam dua klasifikasi, yaitu sentimen positif dan sentimen negatif [12].
- d) **Implementasi Split Data**
Pada langkah ini, Pembagian Data dapat dijalankan melalui RapidMiner Studio dan berkaitan dengan hasil labeling yang dibagi menjadi dua kategori, yaitu data latih dan data uji. Dalam kerangka penelitian ini, data dipisahkan dengan memisahkan dataset latih, yang nantinya akan diuji menggunakan Algoritma Naïve Bayes pada dataset uji [13].
- e) **Implementasi Algoritma Naïve Bayes**
Pada langkah ini, algoritma klasifikasi yang diterapkan yaitu Naïve Bayes pada data latih [14]. Dengan menggunakan algoritma ini, peneliti dapat memprediksi apakah sentimen ulasan pelanggan itu positif atau negatif berdasarkan kata-kata yang terdapat dalam ulasan tersebut.
- f) **Evaluasi**
Pada langkah evaluasi ini, dilakukan perhitungan nilai akhir dari analisis klasifikasi menggunakan metrik-metrik seperti akurasi, presisi, dan recall [15]. Dengan menggunakan matriks kebingungan (confusion matrix) untuk memisahkan berdasarkan pengelompokan dataset, dimana eksperimen pertama dilakukan tanpa menggunakan operator Split Data, sementara eksperimen kedua melibatkan penggunaan operator Split Data. Tabel confusion matrix dapat digunakan untuk menyajikan hasil perhitungan tersebut. [16].
- g) **Kesimpulan**
Dalam rangkuman ini, diperoleh hasil perbandingan antara percobaan pertama dan percobaan kedua untuk menilai apakah hasilnya dapat diinterpretasikan dalam analisis sentimen ulasan, yang mencakup aspek positif dan negatif dari tingkat kepuasan.

2.2 Preprocessing

Preprocessing merupakan sebuah teknik yang digunakan untuk mengubah suatu data mentah kedalam format yang efisien. Karena data mentah seringkali tidak lengkap dan memiliki format yang tidak konsisten. Dalam preprocessing ada beberapa tahapan kerja yang digunakan antara lain:

- a. **Data Cleaning (Cleansing):** tahapan ini data dibersihkan melalui beberapa proses seperti menambahkan nilai yang hilang.
- b. **Letter Casing (Case Folding):** tahapan ini untuk mengkonvers seluruh teks menjadi bentuk standar (huruf kecil atau lowercase)
- c. **Tokenizing:** tahapan ini untuk mengubah kata-kata yang dipisahkan oleh spasi dalam teks.
- d. **Stemming:** tahapan ini untuk pengurangan bentuk menjadi kata dasarnya.
- e. **Filter Stopword Removal:** tahapan ini untuk mengambil kata-kata yang penting dan membuang kata-kata yang tidak berpengaruh pada proses klasifikasi.
- f. **Filter Token (by length):** tahapan ini untuk memfilter tokenize berdasarkan panjangnya.

2.3 Algoritma Naïve Bayes

Algoritma Naive Bayes adalah metode klasifikasi yang berdasarkan teorema Bayes dengan asumsi bahwa setiap fitur yang digunakan untuk klasifikasi adalah independen. Secara sederhana, algoritma ini memprediksi kelas dari suatu data berdasarkan probabilitas dari fitur-fiturnya.

Penting untuk diingat bahwa algoritma ini "naive" karena mengasumsikan independensi antara setiap pasangan fitur, meskipun kenyataannya fitur mungkin saling terkait. Meskipun demikian, Naive Bayes sering kali memberikan hasil yang baik, terutama dalam konteks klasifikasi teks dan dokumen. Salah satu teknik yang diterapkan dalam melakukan klasifikasi ini adalah metode Naïve Bayes. Pendekatan ini beroperasi dengan memproyeksikan probabilitas kejadian di masa yang akan datang berdasarkan pengalaman yang telah terjadi di masa sebelumnya [17]. Untuk rumus Algoritma Naïve Bayes [18] dapat dilihat pada (i).

Rumus Naïve Bayes :

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

Penjelasan

P(A) : Probabilitas kejadian A.

P(B) : Probabilitas kejadian B.

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

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

2.4 Confusion Matrix

Confusion Matriks dipakai untuk mengukur sejauh mana model klasifikasi dapat memprediksi dengan benar terhadap suatu kelas. Dan biasanya digunakan dalam konteks masalah klasifikasi biner (dua kelas), walaupun dapat disesuaikan untuk masalah klasifikasi multikelas. Rumus confusion terlihat pada tabel 1 berikut.

Tabel 1. Rumus Confusion

Prediction Class	Actual Class	
	Pred. Positif	Pred. Negatif
Negatif	TP	FN
Positif	FP	TN

Dalam penilaian kualitas klasifikasi dengan melakukan perhitungan, terdapat beberapa metode yang dapat digunakan, namun salah satu yang umum digunakan adalah menggunakan akurasi, presisi, dan recall [19]. Akurasi adalah persentase dari prediksi model yang telah dibuat. Perhitungan nilai akurasi dilakukan dengan menggunakan persamaan rumus sebagai berikut (2).

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FN+FP)} \times 100\% \tag{2}$$

Selain mengukur akurasi, metode untuk menilai kualitas klasifikasi melibatkan perhitungan precision dan recall. Precision mengindikasikan tingkat keakuratan informasi yang digunakan untuk menghasilkan jawaban oleh sistem. Penghitungan nilai precision dapat dilakukan dengan menggunakan persamaan rumus(3).

$$\text{Precision} = \frac{TP}{(TP+FN)} \times 100\% \tag{3}$$

Recall adalah perbandingan antara jumlah materi yang relevan dengan total materi yang relevan. Penghitungan nilai recall dapat dilakukan dengan menggunakan persamaan rumus(4).

$$\text{Recall} = \frac{TP}{(TP+FN)} \times 100\% \tag{4}$$

3. HASIL DAN PEMBAHASAN

3.1 Hasil Crawling Data

Pada fase permulaan ini, berhasil menghimpun 1366 data dari berbagai platform media sosial, termasuk Youtube, Tiktok, dan Threads. Proses pengambilan data dilakukan secara manual dari setiap platform untuk menjamin kualitas dan ketepatan data yang dihimpun. Dalam analisis ini, studi menggunakan dua istilah kunci, yakni Positif dan Negatif, sebagai fokus utama untuk mengevaluasi sentimen atau tanggapan netizen terhadap fitur Tik Tok Shop [20]. Data diambil menggunakan Dengan menggunakan Bahasa Indonesia, analisis dapat memberikan penafsiran yang lebih akurat mengenai respons dari pengguna fitur Tik Tok Shop, sebagaimana yang terlihat pada gambar 2.

Name	Comment	Time
@user-vz3tc9h	Tik tok itu sebenarnya punya Zionis krna sham2 disna bnyk mrka yg pegang, Tiongkok	2023-12-01T01:42:09Z
@adism707	Budayakan beli prudak dalam negri biar rupiah gak anjlok	2023-11-11T05:56:32Z
@DilAbadyi	TikTok shop 90% produkt dari China. Tanah di negara dikuasi asing	2023-11-07T20:12:23Z
@puryantoparr	Yang jelas pajak nya kurang	2023-11-02T12:09:24Z
@ekosatrio520	JGN TAKUT DGN ADANYA PERDAGANGAN PASAR BEBAS ILEGAL, TAPI YANG PERLU DI	2023-10-18T18:16:11Z
@panjihananto	Owh saya paham maksudnya, saya salah satu orang yg baru bekerja di ekspedisi JNT	2023-10-18T08:01:33Z
@user-ld9px7x	PULIHLAH KAU BANGSAKU, TANPA CAMPUR TANGAN TETANGGA LAIN	2023-10-15T23:48:10Z
@yanudinjitafa	Jika semua pasar dikuasai luar percayalah lapangan kerja akan susah. Jika indones	2023-10-15T06:56:47Z
@michaelstefa	kita tst tau sama tau lahh, visi tiktok itu " Poduksi lokal ? atau Produksi China ? &	2023-10-14T07:40:21Z

Gambar 2. Hasil Crawling Data

3.2 Preprocessing

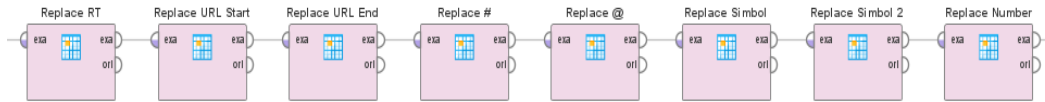
Setelah berhasil mengumpulkan sejumlah 1366 data, langkah berikutnya dalam penelitian ini adalah melakukan preprocessing. Proses preprocessing dilakukan dengan tujuan mengeliminasi kata-kata yang kurang terkait, dalam dataset, meningkatkan mutu data dan menghasilkan teks yang lebih optimal. Tahapan ini melibatkan beberapa langkah, yakni:

- Dalam tahap Pembersihan Data (Cleaning Data) ini, peneliti memulai langkah dengan menggunakan operator penggantian. Selanjutnya, peneliti menentukan mengatur parameter penggantian sesuai dengan data yang perlu disucikan, seperti menghilangkan data yang duplikat, data yang mengandung tautan situs web, tanda baca, simbol-simbol, dan juga data yang berisi emotikon, sebagaimana ditunjukkan dalam Tabel 2.

Tabel 2. Hasil Sample Cleaning Data

Input	Output
JGN TAKUT DGN ADANYA PERDAGANGAN PASAR BEBAS ILEGAL, TAPI YANG PERLU DI TAKUTKAN JIKA PENGAWAS PERDAGANGAN PURA-PURA BUTA. 😊😊😊 pemerintah semestinya bersyukur rakyatnya bisa menjangkau daya beli. krn pemerintah sendiri gk bisa bikin produk murah bagi rakyatnya sendiri 😊😊😊	JGN TAKUT DGN ADANYA PERDAGANGAN PASAR BEBAS ILEGAL TAPI YANG PERLU DI TAKUTKAN JIKA PENGAWAS PERDAGANGAN PURA-PURA BUTA pemerintah semestinya bersyukur rakyatnya bisa menjangkau daya beli. krn pemerintah sendiri gk bisa bikin produk murah bagi rakyatnya sendiri
ADA ONLINE SHOP PAKAYAN DI TIKTOK SHOP BISA MERAUP KEUNTUNGAN 20MILYAR PERBULAN KALO KAGA ADA MUNGKIN BISA DI BAGI2 KE PEDAGANG2 KECIL 😊😊 rakyat miskin bisa beli barang murah pemerintah gak suka, rakyat gak kebeli barang karena mahal, pemerintah senang.... 😊😊😊😊	ADA ONLINE SHOP PAKAYAN DI TIKTOK SHOP BISA MERAUP KEUNTUNGAN PERBULAN KALO KAGA ADA MUNGKIN BISA DI BAGI KE PEDAGANG KECIL rakyat miskin bisa beli barang murah pemerintah gak suka, rakyat gak kebeli barang karena mahal, pemerintah senang
YOUTUBE SEDIKIT LAGI BERHASIL TUMBANGKAN TIKTOK. MEREKA BERSAING SAMPAI YT MENGADAKAN YT SHORT BERBAYAR UNTUK TUMBANGKAN TIKTOK. TIKTOK ORANG ORANG SYNDROOM □□□□	YOUTUBE SEDIKIT LAGI BERHASIL TUMBANGKAN TIKTOK MEREKA BERSAING SAMPAI YT MENGADAKAN YT SHORT BERBAYAR UNTUK TUMBANGKAN TIKTOK. TIKTOK ORANG ORANG SYNDROOM
Harus jika penjualan online tidak dihapus maka pasar tradisional tinggal kenangan Dengan hadir nya tiktok , banyak umkm dan toko ² lokal di indonesia banyak yg tutup dan akhirnya pengangguran merajalela ., tolong buat pemerintah selamatkan lah kami pelaku umkm dan usaha lokal 🙏🙏🙏🙏	Harus jika penjualan online tidak dihapus maka pasar tradisional tinggal kenangan Dengan hadir nya tiktok banyak umkm dan toko lokal di indonesia banyak yg tutup dan akhirnya pengangguran merajalela tolong buat pemerintah selamatkan lah kami pelaku umkm dan usaha lokal

Tahap Pembersihan Data melibatkan eliminasi data yang diperoleh secara manual dari media sosial menggunakan aplikasi RapidMiner Studio. Proses ini mencakup penghapusan data yang mengandung tautan situs web, simbol, tanda baca, dan emotikon, serta identifikasi dan penghapusan data yang memiliki duplikat. Dengan menggunakan operator yang disediakan, data ganda dihapus terlebih dahulu, diikuti oleh eliminasi tautan situs web, emotikon, simbol, dan tanda baca lainnya. Proses Pembersihan Data ini tergambar dalam gambar 3.



Gambar 3. Proses Cleaning Data

- b. Selanjutnya, langkah penggunaan case folding digunakan untuk mengubah kata atau kalimat menjadi huruf kecil dengan memanfaatkan lower case teks. Proses ini bermanfaat untuk mengonversi kata atau kalimat ke dalam format huruf kecil. Hasil dari proses Case Folding dapat diidentifikasi dalam Tabel 3.

Tabel 3. Hasil Case Folding

Input	Output
JGN TAKUT DGN ADANYA PERDAGANGAN PASAR BEBAS ILEGAL TAPI YANG PERLU DI TAKUTKAN JIKA PENGAWAS PERDAGANGAN PURA-PURA BUTA pemerintah semestinya bersyukur rakyatnya bisa menjangkau daya beli.krn pemerintah sendiri gk bisa bikin produk murah bagi rakyatnya sendiri	jgn takut dgn adanya perdagangan pasar bebas ilegal,tapi yang perlu di takutkan jika pengawas perdagangan buta pemerintah semestinya bersyukur rakyatnya bisa menjangkau daya beli.krn pemerintah sendiri gk bisa bikin produk murah bagi rakyatnya sendiri
ADA ONLINE SHOP PAKAYAN DI TIKTOK SHOP BISA MERAUP KEUNTUNGAN PERBULAN KALO KAGA ADA MUNGKIN BISA DI BAGI KE PEDAGANG KECIL	ada online shop pakayan di tiktok shop bisa meraup keuntungan perbulan kalo kaga ada mungkin bisa di bagi ke pedagang kecil
rakyat miskin bisa beli barang murah pemerintah gak suka,rakyat gak kebeli barang karena mahal,pemerintah senang	rakyat miskin bisa beli barang murah pemerintah gak suka,rakyat gak kebeli barang karena mahal pemerintah senang
YOUTUBE SEDIKIT LAGI BERHASIL TUMBANGKAN TIKTOK MEREKA BERSAING SAMPAI YT MENGADAKAN YT SHORT BERBAYAR UNTUK TUMBANGKAN TIKTOK. TIKTOK ORANG ORANG SYNDROOM	youtube sedikit lagi berhasil tumbangkan tiktok mereka bersaing sampai yt mengadakan yt short berbayar untuk tumbangkan tiktok tiktok orang orang syndroom
Harus jika penjualan online tidak dihapus maka pasar tradisional tinggal kenangan	harus jika penjualan online tidak dihapus maka pasar tradisional tinggal kenangan
Dengan hadir nya tiktok banyak umkm dan toko lokal di indonesia banyak yg tutup dan akhirnya pengangguran merajalela tolong buat pemerintah selamatkan lah kami pelaku umkm dan usaha lokal	dengan hadir nya tiktok banyak umkm dan toko lokal di indonesia banyak yg tutup dan akhirnya pengangguran merajalela tolong buat pemerintah selamatkan lah kami pelaku umkm dan usaha lokal

- c. Berikutnya, dalam langkah Tokenisasi, tujuannya adalah untuk memisahkan setiap kata dalam sebuah kalimat. Proses ini menggunakan operator Tokenisasi, dengan menggunakan tabel rumus yang tersedia di web.kaggle.com [21] diperlukan untuk melakukan proses secara otomatis pada seluruh data. Hasil dari Tokenisasi dapat ditemukan dalam Tabel 4.

Tabel 4. Hasil Tokennizing

Input	Output
jgn takut dgn adanya perdagangan pasar bebas ilegal,tapi yang perlu di takutkan jika pengawas perdagangan buta pemerintah semestinya bersyukur rakyatnya bisa menjangkau daya beli.krn pemerintah sendiri gk bisa bikin produk murah bagi rakyatnya sendiri	jgn, takut, dgn, adanya, perdagangan, pasar, bebas, ilegal, tapi, yang, perlu, di, takutkan, jika, pengawas, perdagangan, buta, pemerintah, semestinya, bersyukur, rakyatnya, bisa, menjangkau, daya, beli, krn, pemerintah, sendiri, gk, bisa, bikin, produk, murah, bagi, rakyatnya, sendiri,
ada online shop pakayan di tiktok shop bisa meraup keuntungan perbulan kalo kaga ada mungkin bisa di bagi ke pedagang kecil	ada, online, shop, pakayan, di, tiktok, shop, bisa, meraup, keuntungan, perbulan, kalo, kaga, ada, mungkin, bisa, di, bagi, ke, pedagang, kecil,
rakyat miskin bisa beli barang murah pemerintah gak suka,rakyat gak kebeli barang karena mahal pemerintah senang	rakyat, miskin, bisa, beli, barang, murah, pemerintah, gak, suka, rakyat, gak, kebele, barang, karena, mahal, pemerintah, senang,
youtube sedikit lagi berhasil tumbangkan tiktok mereka bersaing sampai yt mengadakan yt short berbayar untuk tumbangkan tiktok tiktok orang orang syndroom	youtube, sedikit, lagi, berhasil, tumbangkan, tiktok, mereka, bersaing, sampai, yt, mengadakan, yt, short, berbayar, untuk, tumbangkan, tiktok, tiktok, orang, orang, syndroom,

Input	Output
harus jika penjualan online tidak dihapus maka pasar tradisional tinggal kenangan dengan hadir nya tiktok banyak umkm dan toko lokal di indonesia banyak yg tutup dan akhirnya pengangguran merajalela tolong buat pemerintah selamatkan lah kami pelaku umkm dan usaha lokal	harus, jika, penjualan, online, tidak, dihapus, maka, pasar, tradisional, tinggal, kenangan, dengan, hadir, nya, tiktok, banyak, umkm, dan, toko, lokal, di, indonesia, banyak, yg, tutup, dan, akhirnya, pengangguran, merajalela, tolong, buat, pemerintah, selamatkan, lah, kami, pelaku, umkm, dan, usaha, lokal,

- d. Langkah berikutnya adalah tahap stemming yang bertujuan untuk merubah kata-kata dengan imbuhan, kata sambung, dan kata-kata modern menjadi bentuk dasarnya supaya bisa dimanfaatkan dalam tahapan stemming. Hal ini tergambar pada Tabel 5.

Tabel 5. Kata Untuk Stemming

Input	Output
jgn	jangan
dgn	dengan
krn	karena
gk, gak, kaga	enggak
yt	youtube
yg	yang
krn	karena

- e. Setelah itu, dari hasil stemming, diperoleh hasil proses pengubahan kata dari bentuk kekinian menjadi bentuk dasar, serta pemisahan berdasarkan parameter yang telah ditetapkan. Hasil stemming dapat ditemukan dalam Tabel 6.

Tabel 6. Hasil Stemming

Input	Output
jgn, takut, dgn, adanya, perdagangan, pasar, bebas, illegal, tapi, yang, perlu, di, takutkan, jika, pengawas, perdagangan, buta, pemerintah, semestinya, bersyukur, rakyatnya, bisa, menjangkau, daya, beli, krn, pemerintah, sendiri, gk, bisa, bikin, produk, murah, bagi, rakyatnya, sendiri,	jangan, takut, dengan, adanya, perdagangan, pasar, bebas, illegal, tapi, yang, perlu, di, takutkan, jika, pengawas, perdagangan, buta, pemerintah, semestinya, bersyukur, rakyatnya, bisa, menjangkau, daya, beli, karena, pemerintah, sendiri, enggak, bisa, bikin, produk, murah, bagi, rakyatnya, sendiri,
ada, online, shop, pakayan, di, tiktok, shop, bisa, meraup, keuntungan, perbulan, kalo, kaga, ada, mungkin, bisa, di, bagi, ke, pedagang, kecil,	ada, online, shop, pakayan, di, tiktok, shop, bisa, meraup, keuntungan, perbulan, kalo, enggak, ada, mungkin, bisa, di, bagi, ke, pedagang, kecil,
rakyat, miskin, bisa, beli, barang, murah, pemerintah, gak, suka, rakyat, gak, kebele, barang, karena, mahal, pemerintah, senang,	rakyat, miskin, bisa, beli, barang, murah, pemerintah, enggak, suka, rakyat, enggak, kebele, barang, karena, mahal, pemerintah, senang,
youtube, sedikit, lagi, berhasil, tumbangkan, tiktok, mereka, bersaing, sampai, yt, mengadakan, yt, short, berbayar, untuk, tumbangkan, tiktok, tiktok, orang, orang, syndroom,	youtube, sedikit, lagi, berhasil, tumbangkan, tiktok, mereka, bersaing, sampai, youtube, mengadakan, youtube, short, berbayar, untuk, tumbangkan, tiktok, tiktok, orang, orang, syndroom,
harus, jika, penjualan, online, tidak, dihapus, maka, pasar, tradisional, tinggal, kenangan, dengan, hadir, nya, tiktok, banyak, umkm, dan, toko, lokal, di, indonesia, banyak, yg, tutup, dan, akhirnya, pengangguran, merajalela, tolong, buat, pemerintah, selamatkan, lah, kami, pelaku, umkm, dan, usaha, lokal,	harus, jika, penjualan, online, tidak, dihapus, maka, pasar, tradisional, tinggal, kenangan, dengan, hadir, nya, tiktok, banyak, umkm, dan, toko, lokal, di, indonesia, banyak, yg, tutup, dan, akhirnya, pengangguran, merajalela, tolong, buat, pemerintah, selamatkan, lah, kami, pelaku, umkm, dan, usaha, lokal,

- f. Berikutnya, pada langkah ini, merupakan hasil dari tokenisasi yang bertujuan untuk mengambil kata-kata yang signifikan. Oleh karena itu, langkah ini melibatkan penggunaan perpustakaan stopwords (Indonesia) dan formula yang berasal dari web.kaggle.com [22] fungsi ini berguna untuk mengekstrak kata-kata yang memiliki signifikansi dari kamus stopwords yang sudah tersedia dalam Bahasa Indonesia. Hasil dari proses penyaringan stopwords dapat ditemukan dalam Tabel 7.

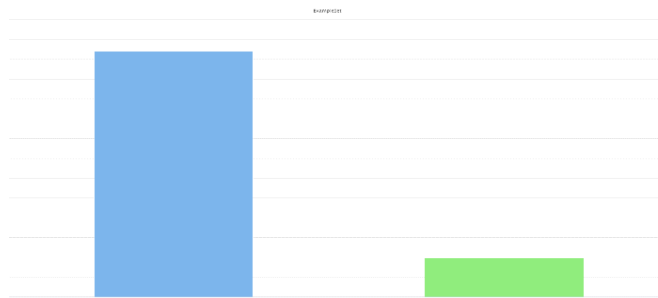
Tabel 7. Hasil Filter Stopword

Input	Output
jangan, takut, dengan, adanya, perdagangan, pasar, bebas, illegal, tapi, yang, perlu, di, takutkan, jika, pengawas, perdagangan, buta, pemerintah, semestinya, bersyukur, rakyatnya, bisa,	jangan, takut, dengan, adanya, perdagangan, pasar, bebas, illegal, tapi, yang, perlu, di, takutkan, jika, pengawas, perdagangan, buta, pemerintah, semestinya, bersyukur, rakyatnya, bisa, menjangkau,

Input	Output
menjangkau, daya, beli, karena, pemerintah, sendiri, enggak, bisa, bikin, produk, murah, bagi, rakyatnya, sendiri, ada, online, shop, pakayan, di, tiktok, shop, bisa, meraup, keuntungan, perbulan, kalo, enggak, ada, mungkin, bisa, di, bagi, ke, pedagang, kecil, rakyat, miskin, bisa, beli, barang, murah, pemerintah, enggak, suka, rakyat, enggak, kebele, barang, karena, mahal, pemerintah, senang, youtube, sedikit, lagi, berhasil, tumbangkan, tiktok, mereka, bersaing, sampai, youtube, mengadakan, youtube, short, berbayar, untuk, tumbangkan, tiktok, tiktok, orang, orang, syndroom, harus, jika, penjualan, online, tidak, dihapus, maka, pasar, tradisional, tinggal, kenangan, dengan, hadir, nya, tiktok, banyak, umkm, dan, toko, lokal, di, indonesia, banyak, yg, tutup, dan, akhirnya, pengangguran, merajalela, tolong, buat, pemerintah, selamatkan, lah, kami, pelaku, umkm, dan, usaha, lokal,	daya, beli, karena, pemerintah, sendiri, enggak, bisa, bikin, produk, murah, bagi, rakyatnya, sendiri, ada, online, shop, pakayan, di, tiktok, shop, bisa, meraup, keuntungan, perbulan, kalo, enggak, ada, mungkin, bisa, di, bagi, ke, pedagang, kecil, rakyat, miskin, bisa, beli, barang, murah, pemerintah, enggak, suka, rakyat, enggak, kebele, barang, karena, mahal, pemerintah, senang, youtube, sedikit, lagi, berhasil, tumbangkan, tiktok, mereka, bersaing, sampai, youtube, mengadakan, youtube, short, berbayar, untuk, tumbangkan, tiktok, tiktok, orang, orang, syndroom, harus, jika, penjualan, online, tidak, dihapus, maka, pasar, tradisional, tinggal, kenangan, dengan, hadir, nya, tiktok, banyak, umkm, dan, toko, lokal, di, indonesia, banyak, yg, tutup, dan, akhirnya, pengangguran, merajalela, tolong, buat, pemerintah, selamatkan, lah, kami, pelaku, umkm, dan, usaha, lokal,

3.3 Labeling Data

Pada langkah ini, terdapat empat proses dalam klasifikasinya. Mulai dari pemberian label (Labeling) pada data yang akan dianalisis, dilanjutkan dengan Pembagian Data (Split Data) Untuk mengelompokkan data menjadi dua sektor, yakni Dataset Pelatihan (Data Latih) dan Dataset Pengujian (Data Uji). Penerapan algoritma dilakukan dengan cara memanfaatkan Algoritma Naïve Bayes [23]. Peneliti ini melibatkan dua percobaan dalam proses evaluasi untuk dianalisis, yakni pertama tidak menggunakan Pembagian Data (Split Data) dan kedua menggunakan operator Pembagian Data (Split Data). Peneliti ini memanfaatkan sampel labeling, dan pada Gambar 4, kemudian terdapat diagram juga yang berfungsi menggambarkan hasil penandaan manual dalam kategori positif dan negatif.



Gambar 4. Grafik Hasil Labeling

Jumlah data yang berhasil diperoleh untuk kelas positif adalah sebanyak 987 data, sementara untuk kelas negatif diperoleh sebanyak 379 data.

3.4. Wordcloud

Penggunaan wordcloud untuk visualisasi bertujuan untuk menampilkan kata-kata yang sering muncul dalam ulasan positif, seperti yang terlihat pada Gambar 5. Beberapa contoh kata-kata tersebut meliputi: 'orang', 'indonesia', 'toko', 'shop', 'barang', 'umkm', 'murah', 'harga', 'sepi', 'pedagang', 'tiktok', 'pemerintah', 'setuju', 'tutup', 'online', 'jual', 'pasar', 'jualan', 'prouduk'.



Gambar 5. Wordcloud Sentimen Positif

Penggunaan wordcloud sebagai visualisasi digunakan untuk mengidentifikasi kata-kata yang sering muncul dalam ulasan negatif, seperti yang ditampilkan dalam Gambar 6. Beberapa contoh kata-kata yang sering muncul adalah berikut:

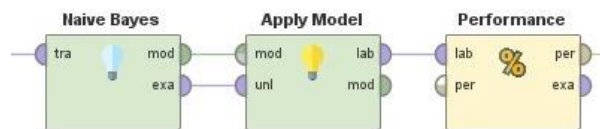
'setuju', 'jualan', 'tutup', 'barang', 'murah', 'beli', 'online', 'tiktok', 'produk', 'orang', 'indonesia', 'umkm', 'toko', 'harga', 'shop', 'pasar', 'jual', 'pemerintah', 'pedagang'.



Gambar 6. Wordcloud Sentimen Negatif

3.5 Implementasi Algoritma Naïve Bayes

Tahapan berikutnya adalah penerapan algoritma Naïve Bayes, yang melibatkan implementasi dan klasifikasi data. Pada langkah ini, data pelatihan akan dianalisis dan diklasifikasikan dengan menggunakan algoritma Naïve Bayes, sementara data pengujian akan dimodelkan. Proses penerapan algoritma Naive Bayes ini tergambar dalam Gambar 7.

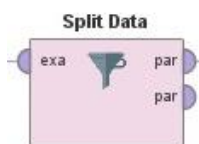


Gambar 7. Processing Algoritma Naïve Bayes

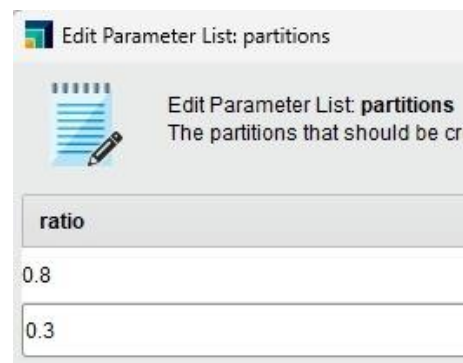
Algoritma Naïve Bayes digunakan untuk klasifikasi data, dengan memisahkan data ke dalam kelas atau jenis tertentu seperti yang telah dijelaskan sebelumnya.

3.6 Implementasi Split Data

Langkah selanjutnya adalah Implementasi Pembagian Data, di mana data dibagi menjadi dua bagian, yaitu data pelatihan (training data) dan data pengujian (testing data). Dalam kerangka penelitian ini, pengelompokan data dilakukan dengan mengalokasikan 80% atau 0.8 dari total data untuk data pelatihan, sementara sisanya, yaitu 30% atau 0.3, digunakan sebagai data pengujian. Proses ini terlihat pada Gambar 8 dan 9.



Gambar 8. Processing Operator Split Data



Gambar 9. Ration Operator Split Data

3.7 Evaluasi

Pada tahap akhir ini, peneliti membandingkan dua percobaan, analisis klasifikasi yang berbeda menggunakan menerapkan algoritma Naïve Bayes tanpa menggunakan operator Pembagian Data dan menggunakan operator Pembagian Data. Hasil perhitungan ini memberikan tampilan dalam bentuk tabel confusion matrix. Berikut

beberapa data yang dihasilkan saat membandingkan kedua percobaan tersebut, dan hasil yang dapat dipresentasikan melalui tabel confusion matrix.

1. Percobaan Pertama

Pada uji coba pertama, diperoleh hasil analisis menggunakan algoritma Naïve Bayes tanpa menerapkan operator Pembagian Data untuk mendapatkan skor dari confusion matrix. Selanjutnya, tabel hasil dari uji coba kedua dapat ditemukan dalam Tabel 8.

Tabel 8. Hasil dari Percobaan Pertama

	Accuracy		86.97%
	true positif	true negatif	class precision
pred.positif	829	0	100.00%
prod.negatif	183	392	68.17%
class recall	81.92%	100.00%	

Diperoleh hasil dari uji coba pertama tanpa menerapkan operator Pembagian Data, dengan tingkat akurasi sebesar 86,97%. Pada kategori positif, diperoleh Precision sebesar 100% dan Recall sebesar 81,92%, sementara pada kategori negatif, Precision mencapai 68,17% dan Recall mencapai 100%.

2. Percobaan Kedua

Dalam hasil uji coba kedua, diperoleh hasil analisis menggunakan algoritma Naïve Bayes dengan penerapan operator Pembagian Data untuk mendapatkan skor dari confusion matrix. Tabel hasil uji coba kedua dapat diidentifikasi pada Tabel 9.

Tabel 9. Hasil dari Percobaan Kedua

	Accuracy		89.23%
	true positif	true negatif	class precision
pred.positif	626	0	100.00%
prod.negatif	110	285	72.15%
class recall	85.05%	100.00%	

Diperoleh hasil dari uji coba kedua dengan menerapkan operator Pembagian Data, dengan tingkat akurasi sebesar 89,23%. Pada kategori positif, diperoleh Precision sebesar 100% dan Recall sebesar 85,05%, sementara pada kategori negatif, Precision mencapai 72,15% dan Recall mencapai 100%.

4. KESIMPULAN

Dari hasil pengujian dalam penelitian ini, dapat disimpulkan bahwa dengan banyaknya data yang terkumpul mencapai 1366 dan melibatkan label manual yang dibantu oleh ahli bahasa, algoritma Naïve Bayes terbukti menjadi pilihan yang tepat untuk menganalisis sentiment untuk tema yang diangkat oleh peneliti. Dalam dua percobaan yang dilakukan, terdapat beberapa poin penting terkait akurasi, presisi, dan recall. Pada percobaan pertama, akurasi Naïve Bayes mencapai 86,97%, dengan presisi sebesar 68,17%, dan recall sebesar 81,92%. Sementara pada percobaan kedua, akurasi mencapai 89,23%, dengan presisi sebesar 72,15%, dan recall sebesar 85,05%. Temuan ini menunjukkan bahwa percobaan kedua menghasilkan performa yang lebih baik daripada percobaan pertama, mengindikasikan bahwa pemisahan data secara manual memberikan hasil yang lebih efisien daripada tidak menggunakan operator Pembagian Data. Sehingga, Hasil analisis sentimen dapat menjadi dasar untuk menyusun rekomendasi kebijakan lebih lanjut atau strategi komunikasi yang dapat membantu dalam memperbaiki persepsi masyarakat atau merespon kekhawatiran yang muncul. Dengan memahami sentimen netizen terkait penutupan TikTok Shop dan implikasinya, kesimpulan ini dapat menjadi dasar untuk menginformasikan kebijakan, komunikasi, dan strategi lanjutan yang lebih baik di masa depan.

REFERENCES

[1] A. S. Widagdo and A. N. W., Bambang Soedijono, "Analisis Tingkat Kepopuleran E-Commerce Di Indonesia Berdasarkan Sentimen Sosial Media Menggunakan Metode Naïve Bayes," *J. Inf. Politek. Indonusa* Surakarta, vol. 6, 2020.

[2] S. Sari, U. Khaira, P. Pradita, and T. S. Tri, "... Beauty Shaming Di Media Sosial Twitter Menggunakan Algoritma SentiStrength: Sentiment Analysis Against Beauty Shaming Comments on Twitter Social Media ...," *Indones. J. ...*, vol. 1, no. 1, pp. 71-78, 2021, [Online]. Available: <https://journal.irpi.or.id/index.php/ijirse/article/view/55%0Ahttps://journal.irpi.or.id/index.php/ijirse/article/download/55/24>.

[3] J. Mantik et al., "Application Of N-Gram On K-Nearest Neighbor Algorithm To Sentiment Analysis Of TikTok Shop Shopping Features," *J. Mantik*, vol. 6, no. 3, pp. 2685-4236, 2022.

[4] C. Maulida, T. Yunanda, M. Hanafi, W. Mega, and P. Dhuhiha, "Sentiment Analysis on TikTok Shop Reviews Using Long Short-Term Memory Method to Find Business Opportunity," *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 9, no.

- 1, pp. 1–7, 2024, [Online]. Available: <https://doi.org/10.25139/inform.v9i1.6524>.
- [5] D. Ardiansyah, A. Saepudin, R. Aryanti, E. Fitriani, and Royadi, “Analisis Sentimen Review Pada Aplikasi Media Sosial Tiktok Menggunakan Algoritma K-NN Dan Svm Berbasis Pso,” *J. Inform. Kaputama*, vol. 7, no. 2, pp. 233–241, 2023, doi: 10.59697/jik.v7i2.148.
- [6] D. Oktaheriyani, M. A. Wafa, and S. Shadiqien, “ANALISIS PERILAKU KOMUNIKASI PENGGUNA MEDIA SOSIAL TIKTOK (Studi Pada Mahasiswa Fakultas Ilmu Sosial dan Ilmu Politik UNISKA MAB Banjarmasin),” *ePRINTS UNISKA*, pp. 1–62, 2020, [Online]. Available: <http://eprints.uniska-bjm.ac.id/id/eprint/3504>.
- [7] A. I. Tanggraeni and M. N. N. Sitokdana, “Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 2, pp. 785–795, 2022, doi: 10.35957/jatisi.v9i2.1835.
- [8] S. C. Nandaresta and C. Warman, “Terhadap Tiktok Shop Dan Shopee Di Twitter Menggunakan Metode Naïve Bayes Dan Knn (K- Nearest Neighbor,” *Sismatik*, vol. 12, no. 1, pp. 1–9, 2023.
- [9] S. Jafar and A. Nur, “Analysis Of Twitter User Sentiment To Tiktok Shop Using Naïve Bayes And Decision Tree Algorithms,” pp. 8–14, 2010.
- [10] D. Normawati and S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021, [Online]. Available: <https://ejournal.tunasbangsa.ac.id/index.php/jsakti/article/view/369/348>.
- [11] A. Surahman, A. F. Octaviansyah, and D. Darwis, “Ekstraksi Data Produk E-Marketplace Sebagai Strategi Pengolahan Segmentasi Pasar Menggunakan Web Crawler,” *Sistemasi*, vol. 9, no. 1, p. 73, 2020, doi: 10.32520/stmsi.v9i1.580.
- [12] A. Aziz, “Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 1, p. 115, 2022.
- [13] 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, pp. 636–646, 2021, doi: 10.35957/jatisi.v8i2.962.
- [14] S. Lestari and S. Saepudin, “Analisis Sentimen Vaksin Sinovac Pada Twitter Menggunakan Algoritma Naive Bayes,” *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform.)*, pp. 163–170, 2021.
- [15] Dedi Darwis, Nery Siskawati, and Zaenal Abidin, “Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional,” *J. TEKNO KOMPAK*, vol. 15, no. 1, pp. 131–145, 2020.
- [16] H. F. Putro, R. T. Vlandari, and W. L. Y. Saptomo, “Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan,” *J. Teknol. Inf. dan Komun.*, vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.
- [17] E. Fitri, “Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine,” *J. Transform.*, vol. 18, no. 1, p. 71, 2020, doi: 10.26623/transformatika.v18i1.2317.
- [18] F. Sidik, I. Suhada, A. H. Anwar, and F. N. Hasan, “Analisis Sentimen Terhadap Pembelajaran Daring Dengan Algoritma Naive Bayes Classifier,” *J. Linguist. Komputasional*, vol. 5, no. 1, p. 34, 2022, doi: 10.26418/jlk.v5i1.79.
- [19] T. Wiratama Putra, A. Triayudi, and 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, pp. 20–26, 2022, doi: 10.35870/jtik.v6i1.368.
- [20] Y. R. Pusvitasari, “Analisis Penggunaan Fitur Tiktok Shop Pada Perilaku Konsumtif Mahasiswa Fakultas Dakwan IAIN Salatiga,” p. 1, 2022, [Online]. Available: <http://e-repository.perpus.iainsalatiga.ac.id/14705/>.
- [21] RAVI SHAH, *Guide to Tokenizers and Preprocessing*. 2022.
- [22] Konrad Banachewicz and Luca Massaron, *Kaggle Workbook*. Packt Publishing, Limited, Packt Publishing, 2022.
- [23] miftahul kahfi al Fath, “Analisis Sentimen Komentar Kebijakan Full Day School Dari Facebook Page Kemendikbud Ri Menggunakan Algoritma Naïve Bayes Classifier,” *Tek. Inform.*, p. 169, 2018, [Online]. Available: <https://repository.uinjkt.ac.id/dspace/handle/123456789/47898>.