ANALISIS SENTIMEN TERHADAP BERHENTINYA TIKTOKSHOP PADA MEDIA SOSIAL *TWITTER* MENGGUNAKAN *ALGORITMA NAÏVE BAYES*

Riki Sanjaya, Edi Tohidi, Edi Wahyudi, Kaslan
Teknik Informatika, STMIK IKMI CIREBON
Jalan Perjuangan No. 10B Majasem Kec. Kesambi Kota Cirebon
rikisan1004@gmail.com

ABSTRAK

Perkembangan teknologi informasi telah membawa perubahan besar dalam dunia perdagangan, salah satunya adalah munculnya perdagangan dengan sistem elektronik (PSE) yang memungkinkan transaksi jual beli dilakukan secara online. Tiktokshop adalah salah satu platform PSE yang populer di Indonesia [1], yang merupakan bagian dari aplikasi Tiktok yang menyediakan fitur jual beli produk. Namun, pada tanggal 04 Oktober 2023, pemerintah Indonesia mengeluarkan kebijakan untuk menutup Tiktokshop berdasarkan surat Menteri perdagangan no 31 tahun 2023. Alasan pemerintah adalah untuk melindungi usaha mikro, kecil, dan menengah (UMKM) tradisional dari persaingan tidak sehat yang ditimbulkan oleh Tiktokshop. Kebijakan ini menuai banyak reaksi di media sosial, khususnya di Twitter, yang menjadi salah satu media untuk menyuarakan pendapat dan aspirasi masyarakat. Penelitian ini bertujuan untuk menganalisis sentimen terkait penutupan Tiktokshop di Twitter menggunakan algoritma Naive Bayes, yang merupakan salah satu metode klasifikasi teks berdasarkan kemunculan kata-kata tertentu. Hasil analisis menunjukkan bahwa sebagian besar netizen yang berkomentar terkait penutupan Tiktokshop menunjukkan bahwa respon positif mendominasi dengan nilai prediksi positif 965, sedangkan respon negatif hanya 535 nilai akurasi 85.00%, margin error +-2.92% Precision 99,33% recall 77,20 %.

Kata kunci : *Naive Bayes, Sentimen, Tiktokshop, Twitter*

1. PENDAHULUAN

Media sosial adalah teknologi yang populer dan penting bagi banyak orang [2]. Media sosial menyediakan informasi terbaru dan sumber pendapatan bagi beberapa orang [3]. Tiktok adalah media sosial yang memiliki fitur TiktokShop, yang memungkinkan pengguna berbelanja di aplikasi. Penjual Tiktok menggunakan live streaming untuk menarik dan berkomunikasi dengan pelanggan. [4]

PSE (Perdagangan dengan sistem elektronik) adalah cara berdagang menggunakan teknologi digital. Tiktokshop, platform PSE populer di Indonesia, menawarkan produk dengan harga murah dan pengiriman cepat. Tiktokshop ditutup sesuai Surat Mentri Perdagangan No. 31 Tahun 2023 pada 4 Oktober 2023 pukul 17.00 WIB. Pada tahun 2023, kebijakan ini melindungi UMKM tradisional dari persaingan tidak adil oleh Tiktokshop. Surat tersebut mengatur izin usaha, kualitas produk, dan pengaduan konsumen PSE. Tiktokshop tak memenuhi persyaratan karena tak berizin dan tak transparan. Penutupan Tiktokshop mendapat banyak kritik di media sosial. Penelitian ini mengkaji sentimen di Twitter terkait penutupan tersebut, fokus pada reaksi dan pendapat positif dan negatif. Penelitian ini akan memberikan pemahaman tentang dampaknya pada penjual, pembeli, dan masyarakat Twitter secara keseluruhan. [5]

Penelitian sebelumnya oleh Friska Aditia Indriyani, Ahmad Fauzi dan Sutan Faisal berjudul "Analisis sentimen aplikasi Tiktok menggunakan algoritma Naïve Bayes dan Support Vector Machine" menggunakan 2000 data ulasan Tiktok dari Google Play Store dengan algoritma Naïve Bayes dan SVM. Hasilnya menunjukkan 76.7% opini positif dan 23.3% opini negatif, dengan SVM memiliki akurasi 84% dibandingkan dengan Naïve Bayes yang memiliki akurasi 79% [6]. Penelitian sekanjutnya oleh Laurensz B, Sentimen A, dan Sediyono berjudul ‘‘Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19‘‘ menggunakan metode klasifikasi teks (SVM dan Naïve Bayes).Untuk kata kunci ‘‘vaksinsinovac,‘‘ Naïve Bayes menghasilkan sentimen positif 66%, sementara SVM menghasilkan 96% positif.Untuk kata kunci ‘‘vaksinmerahputih,‘‘ Naïve Bayes menghasilkan 89% positif, sedangkan SVM menghasilkan 98% positif [7]. Penelitian selanjutnya dilakukan oleh Muljono A, Putri Artanti D Dkk berjudul "Analisis Sentimen untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes" menggunakan 1200 dokumen dengan label data, 610 positif dan 590 negatif. Model klasifikasi Naïve Bayes menghasilkan akurasi 93.33%, presisi 93.33%, sensitivitas 93.33%, dan f-measure 93.65% setelah diuji dengan data uji yang berbeda menggunakan validasi silang 10 kali lipat[8].

Penelitian ini bertujuan menganalisis sentimen Twitter terkait penutupan Tiktokshop dengan algoritma Naive Bayes. Analisis akan mengidentifikasi dampak positif dan negatif pada penjual dan pembeli, serta mencari hubungan dengan Tanah Abang. Algoritma Naive Bayes digunakan untuk mengkategorikan tweet tentang penutupan Tiktokshop menjadi sentimen positif dan negatif. Hal ini memberikan pemahaman tentang dampak

penutupan Tiktokshop pada penjual, pembeli, dan konsekuensi terkait Tanah Abang sebagai pusat perdagangan utama. Analisis ini mengulas pengaruh penutupan Tiktokshop pada berbagai pihak.

Metode data mining dengan Naive Bayes menganalisis sentimen terkait penutupan Tiktokshop. Data Twitter digunakan untuk kategorisasi opini pengguna sebagai positif atau negatif berdasarkan kamus sentimen. Algoritma Naive Bayes memperhitungkan probabilitas sentimen tweet, dalam konteks penutupan Tiktokshop di media sosial. Memberikan wawasan tentang dampaknya, mengidentifikasi emosi dan pandangan terkait keputusan tersebut. Berfokus pada data Twitter sebagai sumber utama opini dan sentimen pengguna. Teksa tersebut tidak tersedia, jadi tidak dapat mempersingkatnya.

Penelitian ini berpotensi memberikan kontribusi pada pemahaman sentimen publik terhadap penutupan Tiktokshop oleh pemerintah dengan Surat Mentri Perdagangan No. 31 Tahun 2023. Penelitian ini diharapkan memberikan masukan evaluatif kepada pemerintah dan bermanfaat bagi praktisi dan peneliti yang tertarik pada analisis sentimen terkait topik serupa, sekaligus berkontribusi pada pengembangan teknologi informasi sesuai kebutuhan dan aspirasi masyarakat.

2. TINJAUAN PUSTAKA

2.1. Penelitian terdahulu

Menurut hasil kajian sebelumnya yang telah dijalankan oleh Ikbar Wibowo Dkk melakukan penelitian perbandingan algoritma klasifikasi sentimen Twitter terhadap insiden kebocoran data Tokopedia. Tujuannya adalah mengembangkan model klasifikasi sentimen untuk mengidentifikasi apakah twit berkaitan dengan insiden tersebut bersifat positif, negatif, atau netral. Dengan menggunakan tiga jenis classifier (Random Forest, Support-Vector Machine, dan Logistic Regression), hasil penelitian menunjukkan bahwa Support Vector Machine memiliki performa terbaik dengan f1-score 0.503583, menjadikannya pilihan optimal untuk analisis sentimen terkait insiden kebocoran data Tokopedia [9].

Penelitian yang dilakukan Fendyputra Pratama berjudul Analisis Sentimen Twitter Debat Calon Presiden Indonesia menggunakan metode Fined-Grained Sentiment Analysis mengeksplorasi opini publik di Twitter tentang debat capres pertama di Indonesia dengan hashtag paslon dan menilai sikap masyarakat, positif, negatif, atau netral. Metode analisis sentimen yang digunakan adalah Fined-grained Sentiment Analysis pada data tweet yang diambil menggunakan Twitter API dan R. Metode ini menghitung nilai sentimen tweet berdasarkan jumlah kalimat positif dan negatif. Penelitian menunjukkan bahwa tweet dengan hashtag paslon memiliki sentimen positif lebih banyak daripada negatif dan netral (+1) [10] .

Penelitian yang dilakukan oleh Irvandi, Bambang Irawan, Odi Nurdiawan menggunakan Algoritma Naive Bayes untuk menganalisis sentimen wisatawan terhadap wisata halal di Pulau Lombok. Data ulasan diambil dari Google Maps melalui web scraping. Data diolah dengan RapidMiner menggunakan algoritma Naive Bayes dan dipakai wordcloud untuk visualisasi kata-kata umum dalam ulasan wisatawan. Penelitian ini bertujuan memeriksa kecocokan wisata halal di Lombok dengan ekspektasi pengunjung. Hasilnya menunjukkan bahwa model algoritma Naive Bayes memiliki akurasi 74,75%. Teks tersebut dapat dipersingkat menjadi:Kata-kata yang paling banyak digunakan oleh wisatawan dalam wordcloud adalah ‘‘indah‘‘, ‘‘wisata‘‘, ‘‘pantai‘‘, ‘‘alam‘‘, ‘‘gunung‘‘, dan ‘‘masjid‘‘. [11].

2.2. Naïve Bayes

Naive Bayes merupakan metode yang mudah dan efektif yang menerapkan konsep-konsep teori peluang untuk menghitung probabilitas tertinggi berdasarkan frekuensi atau jumlah kemunculan dari setiap kelas dalam data latih. Dalam pengolahan basis data, Naive Bayes masuk dalam pembelajaran terbimbing jenis pembelajaran mesin yang membutuhkan contoh sebagai data latih berlabel. dibagi menjadi dua jenis, yaitu. Klasifikasi dan regresi. Klasifikasi ketika variabel menjadi kategori, panas dan dingin, sakit atau tidak sakit dll. Variabel berupa nilai nyata seperti berat, nilai uang dll contoh lainnya yaitu Support Vector Machine (SVM), K-Nearst Neighbor (KNN), Artificial Neural Network (ANN). Naive Bayes juga merupakan metode yang menggunakan teknik probabilistik, dimana setiap fitur dan fitur lainnya dalam data yang sama diasumsikan tidak saling berhubungan [12].

2.3. *Data mining*

Data mining merupakan metode analisis data besar yang disimpan dalam basis data. Data mining memanfaatkan pengalaman atau kesalahan sebelumnya sebagai sumber belajar untuk meningkatkan kualitas model dan hasil analisis, salah satu metodenya adalah dengan menerapkan teknik klasifikasi data mining [13].

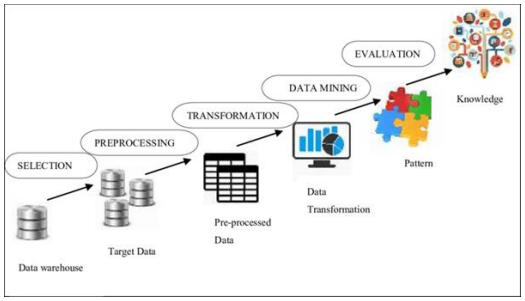
2.4. *RapidMiner*

RapidMiner adalah aplikasi yang dapat mengolah data dengan metode data mining. RapidMiner dapat menemukan pola-pola tersembunyi dari kumpulan data yang besar dengan menggunakan teknik statistika, kecerdasan buatan dan database [14]

3. METODE PENELITIAN

Penelitian ini menerapkan metode Knowledge Discovery in Database (KDD) yang berguna untuk mengeksplorasi teknik dari database yang sudah ada [15]. Penggunaan istilah Knowledge Discovery in database (KDD) seringkali saling dipertukarkan untuk menggambarkan proses mengungkap informasi yang

tersembunyi dalam suatu basis data yang besar[16].Tujuan dari KDD dan data mining adalah menggali informasi tersembunyi dari basis data yang besar Metode KDD terdiri dari tahapan *Selection, Preprocessing, Transformation, data mining,* dan *Evaluation.*



Gambar 1. Alur Penelitian KDD

3.1. *Selection*

Tahap selection melibatkan pengumpulan data, di mana dataset yang digunakan terdiri dari 946 data yang diperoleh dari platform Twitter. Data ini dikumpulkan dengan menggunakan teknik crawling pada Googe Colab Untuk crawling Twitter, peneliti perlu mendapatkan auth-token terlebih dahulu. Auth-token adalah kode unik yang digunakan untuk mengidentifikasi dan mengotentikasi pengguna Twitter. Peneliti dapat memanfaatkan Google Colab, sebuah platform cloud yang memungkinkan eksekusi kode Python secara online, untuk memperoleh auth-token. Proses pengambilan data dari Twitter dapat dilakukan menggunakan Google Colab.

3.2. *Preprocessing*

Preprocessing merupakan fase di mana data dibersihkan dan diperbaiki. Saat mengumpulkan data, informasi yang diperoleh biasanya bersifat tidak terstruktur dan mengandung sejumlah besar karakter. Tujuan dari proses preprocessing adalah untuk menghilangkan noise tersebut. Pada tahap ini, terdapat beberapa langkah yang perlu dilakukan.

1. *Cleaning* merupakan proses menghilangkan data yang tidak relevan, seperti emoticon, hashtag, username, url, atau karakter lain yang tidak berguna.
2. *Case Folding* merupakan tahap mengubah teks menjadi huruf kecil
3. *Tokenize* merupakan tahap memecah kalimat jadi kata per kata.
4. *Filter by* length merupakan Salah satu metode untuk menyeleksi data berdasarkan kriteria tertentu.
5. *Stopword* merupakan tahap menghapus kata yang tidak memiliki nilai pada proses data mining, seperti kata sambung.
6. *Stemming* merupakan tahap menghapus imbuhan pada kata.
7. Pelabelan merupakan tahap memberikan label pada kalimat agar memiliki nilai sentimen

3.3. *Transformation*

Transformasi data adalah langkah yang dilakukan untuk mengubah skala pengukuran data awal ke format lain, sehingga data tersebut sesuai dengan asumsi-asumsi yang menjadi dasar analisis berbagai teknik data mining[17]. Pada fase ini, akan dilakukan pembobotan kata dengan menerapkan metode TF-IDF.

3.4. *Data mining*

Data mining adalah proses untuk menemukan pengetahuan dari data besar. Data mining menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstrak informasi yang bermanfaat dari kumpulan data yang besar. Dalam penelitian ini, kami menggunakan algoritma Naïve Bayes untuk mengklasifikasikan emosi. Sebelum melakukan data mining, data dibagi menjadi dua bagian. Bagian pertama adalah data pelatihan dan bagian kedua adalah data pengujian dengan rasio 70% pelatihan dan 30% pengujian. Pada tahap ini, kami melakukan analisis sentimen menggunakan algoritma Naive Bayes Classifier yang digunakan untuk menghitung nilai pengklasifikasi presisi, presisi, recall, dan f1 score.[18]

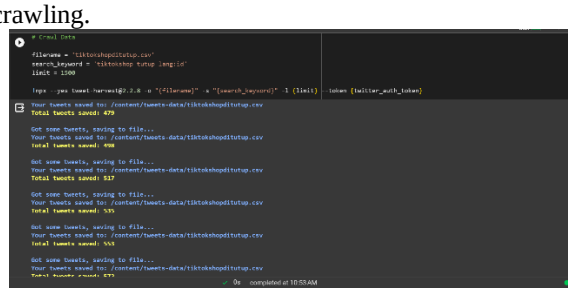
3.5. *Evaluation*

Peneliti menggunakan metode pengujian untuk menilai kualitas model klasifikasi yang dikembangkan. Model klasifikasi diuji dengan menghitung akurasi. Akurasi adalah rasio antara jumlah rekaman yang diklasifikasikan dengan benar dan jumlah total rekaman. Peneliti menyajikan hasil pengujian dalam bentuk tabel matrix konfusi yang berisi empat nilai, yaitu true positif (TP), false negatif (FN), true negatif (TN), dan false positif (PF). Peneliti memanfaatkan confusion matrix sebagai alat evaluasi untuk mengukur seberapa baik model dapat membedakan antara klasifikasi yang benar atau salah.

4. HASIL DAN PEMBAHASAN

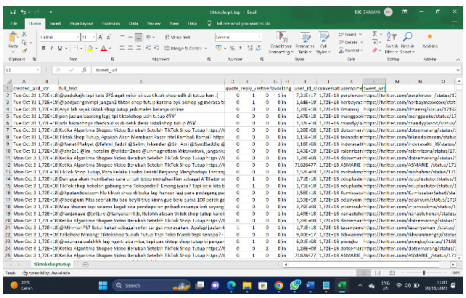
4.1. Selection

Tahap selection adalah metode yang digunakan untuk memilih data yang relevan dari *Twitter* berdasarkan kata kunci tertentu. Dalam penelitian ini, kata kunci yang dipilih adalah "tiktokshop tutup" yang berkaitan dengan isu yang sedang ramai di media sosial. Untuk mengambil data dari *Twitter*, peneliti menggunakan Google Colab sebagai platform untuk menjalankan proses crawling. Gambar 2 menunjukkan langkah-langkah yang dilakukan dalam proses crawling.



Gambar 2. Proses *Crawling*

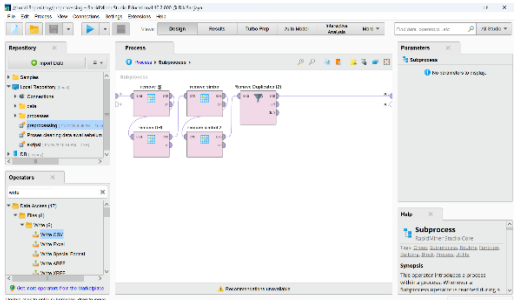
Hasil dari proses crawling yang dilakukan untuk mengumpulkan data tweet yang berkaitan dengan isu pemberhentianya tiktokshop di Indonesia. Dari 951 tweet yang berhasil diperoleh, sekitar 81% mengekspresikan rasa kecewa, marah, atau bingung atas keputusan tersebut. Gambar 3 menunjukan hasil dari crawling



Gambar 3. Hasil *Crawling*

4.2. Cleaning

Cleaning adalah proses menghilangkan elemen-elemen yang tidak relevan atau mengganggu dari data, seperti tanda baca koma, tanda tanya, tanda seru, dan *emoticon*, *URL*, dan *noisy*. Preprocessing adalah proses memperbaiki kualitas data dengan mengatasi masalah missing value atau nilai kosong dan data yang tidak konsisten. Proses ini sangat penting karena dapat mempengaruhi hasil dari proses data.



Gambar 4. *Cleaning*

Hasil dari proses *cleaning* dapat dilihat pada tabel 1 berikut :

Tabel 1 Tahap proses *cleaning*

Sebelum	Sesudah
@padukajh tapi kalo ðŸ□Š agak mikir sii cuz tiktok shop udh di tutup kan :(	padukajh tapi kalo agak mikir sii cuz tiktok shop udh di tutup kan

4.3. Case Folding

Case folding merupakan proses mengubah huruf besar menjadi huruf kecil atau sebaliknya dalam sebuah teks

Tabel 2 Tahap *case folding*

Sebelum	Sesudah
padukajh tapi kalo agak mikir sii cuz tiktok shop udh di tutup kan	padukajh tapi kalo agak mikir sii cuz tiktok shop udh di tutup kan

4.4. Tokenize

Tokenize merupakan merupakan proses membagi kalimat atau teks menjadi data menjadi kata atau token.

Tabel 3 Hasil proses *tokenize*

Sebelum	Sesudah
padukajh tapi kalo agak mikir sii cuz tiktok shop udh di tutup kan	‘padukajh’,‘tapi’, ‘kalo’, ‘agak’, ‘mikir’, ‘sii’, ‘cuz’, ‘tiktok’, ‘shop’, ‘udh’, ‘di’, ‘tutup’, ‘kan’

4.5. Stopword

Stopword bertujuan untuk menghilangkan semua kata yang masih sering muncul dalam kumpulan data dan dianggap tidak relevan serta tidak mempengaruhi mood kalimat yang dihilangkan. Kata-kata yang dianggap tidak berhubungan di sini antara lain adalah kata sambung seperti di, ke, ini, dan dari. Tujuan menghilangkan konjungsi adalah untuk membuat teks ulasan menjadi lebih kecil dan ringkas.

Tabel 4 Hasil proses *stopword.*

Sebelum	Sesudah
padukajh tapi kalo agak mikir sii cuz tiktok shop udh di tutup kan	Tapi agak mikir tiktok shop tutup kan

4.6. Filter by length

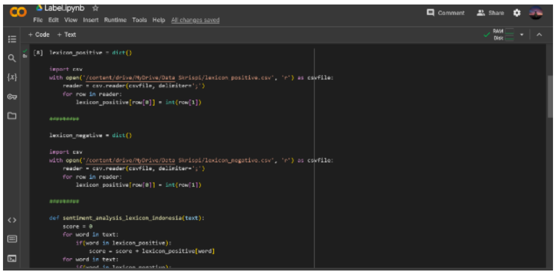
Filter by length merupakan Salah satu metode untuk menyeleksi data berdasarkan kriteria tertentu. Metode ini dapat diterapkan pada berbagai tipe data, seperti string, list, atau array. Dalam penelitian ini, peneliti menggunakan kriteria panjang minimal 4 karakter dan maksimal 25 karakter, hasil *filter by length* dapat dilihat pada tabel 5 berikut.

Tabel 5 Hasil proses *filter by length*.

Sebelum	Sesudah
padukajh tapi kalo agak mikir sii cuz tiktok shop udh di tutup kan	Tapi agak mikir tiktok shop tutup

4.7. Labeling

Setelah data ulasan selesai di-steaming, langkah berikutnya adalah memberi label data menggunakan lexicon Vader. Lexicon Vader menghitung nilai compound sebagai rata-rata tertimbang dari bobot sentimen yang ada di dalam ulasan (Asri et al., 2022). Ulasan yang memiliki nilai compound > 0 diklasifikasikan sebagai sentimen positif, sedangkan ulasan yang memiliki nilai compound < 0 diklasifikasikan sebagai sentimen negatif. Untuk memberi label data ulasan, peneliti masih memanfaatkan software Python. Dari 921 data ulasan yang diberi label, terdapat 171 data ulasan berlabel positif dan 750 data ulasan berlabel negatif.



Gambar 5 pelebelan *vader lexicon*

Gamber 5 menampilkan cara menghitung skor sentimen teks dengan kamus vader lexicon. Teks dibagi dan dicocokkan dengan entri kamus yang punya nilai sentimen. Skor sentimen total diubah sesuai konteks token dan dibagi dengan jumlah token untuk dapat skor sentimen rata-rata.

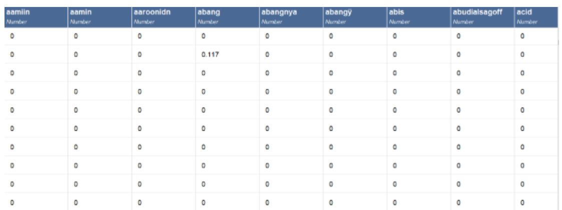
Tabel 6 Hasil label *vader lexcion*

Text	polarity_score	polarity
['padukajh', 'kalo', 'mikir', 'tiktok', 'shop', 'tutup'	-5	negative
anjir tiktok shop tutup nonton live produk	2	positive

Tabel 6 menampilkan nilai emosi positif atau negatif dari beberapa teks, yang diberikan oleh kamus vader lexicon. Kamus ini berisi kata dan frasa dengan nilai emosi. Nilai emosi ini bisa mengukur perasaan atau pandangan dari teks.

4.8. Transformation

Data yang sudah diolah pada langkah sebelumnya diberikan label positif atau negatif dengan proses validasi menggunakan kamus vader lexcion, lalu dataset dan label yang awalnya berbentuk data teks dikonversi ke format angka dengan algoritma TF–IDF (Term Frequency–Inverse Document Frequency) sebagai berikut:



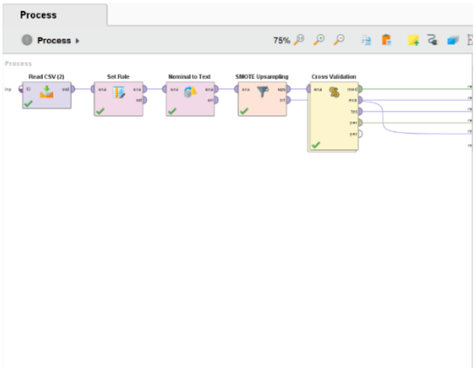
aamiin Number	aamin Number	aaroonidn Number	abang Number	abangnya Number	abangy Number	abis Number	abudialsagoff Number	acid Number
0	0	0	0	0	0	0	0	0
0	0	0	0.117	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

Gambar 6 hasil proses *tf-idf*

Gambar 6 menunjukkan hasil perhitungan TF-IDF untuk mengukur frekuensi kemunculan kata dalam data twitter. Teknik ini memberikan bobot untuk setiap kata berdasarkan seberapa sering kata tersebut muncul dalam dokumen. Sebagai contoh, kata "abis" memiliki bobot 0 karena tidak terdapat dalam dokumen pertama. Kata "Abang" memiliki bobot 0,117 karena terdapat dalam dokumen kedua dan dokumen lainnya.

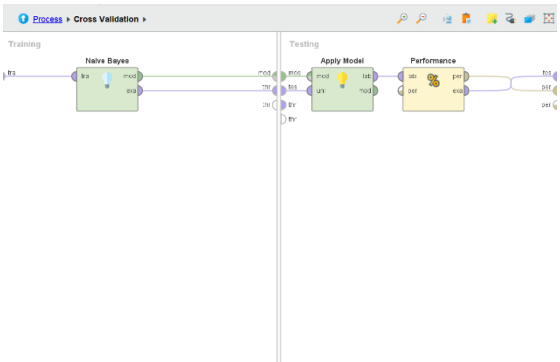
4.9. Data Mining

Langkah ini adalah proses mendapatkan informasi dari data sebelumnya yang telah dikumpulkan dan diproses. Dalam penelitian ini, metode klasifikasi *Naive Bayes* digunakan ini adalah salah satu pendekatan pembelajaran mesin yang didasarkan pada *teorema Bayes*.



Gambar 7 proses model *naïve bayes*

Gambar 7 menggambarkan metode *Naive Bayes* yang digunakan untuk menganalisis data. Metode ini terdiri dari beberapa langkah. Langkah pertama adalah mengimpor data latih dari file CSV. Langkah kedua adalah menetapkan variabel target sebagai label dengan *operator Set Role*. Langkah ketiga adalah mengubah variabel nominal menjadi teks dengan operator Nominal to Text. Langkah keempat adalah melakukan upsampling untuk menyeimbangkan kelas data. Langkah kelima adalah membuat dan menguji model *Naive Bayes* dengan operator cross-validation. Metode ini membagi data menjadi 10 lipatan dan menguji model pada setiap lipatan.



Gambar 8 Proses *Validation Naive Bayes*

Gambar 8 menunjukkan proses cross-validation yang menggunakan operator Naive Bayes, Apply Model, dan Performance. Operator Naive Bayes membuat model klasifikasi berdasarkan probabilitas dengan menganggap fitur-fitur bersifat saling bebas, menggunakan data pelatihan sebagai input. *Operator Apply Model* menerapkan model ini pada data pengujian untuk mendapatkan prediksi kelas. *Operator Performance* mengukur kinerja model dengan metrik-metrik seperti akurasi dan presisi.

Berikut ini adalah hasil *Performance* yang didapatkan dengan menggunakan metode *Naive bayes*:

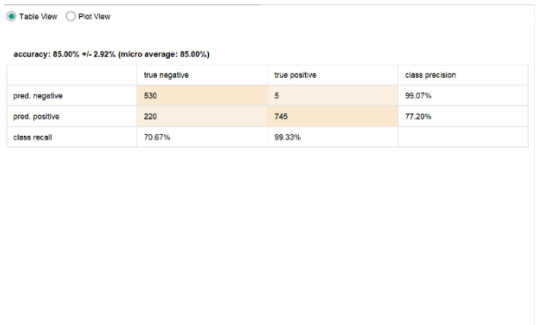


Table View Plot View

accuracy: 85.00% +/- 2.92% (micro average: 85.00%)

	true negative	true positive	class precision
pred. negative	530	5	99.07%
pred. positive	220	745	77.20%
class recall	70.67%	99.33%	

Gambar 9 Hasil *Performance Accuracy Naïve Bayes*

Gambar 9 menunjukkan bahwa algoritma Naïve Bayes memiliki akurasi 85,00%, yang berarti model ini mampu mengklasifikasikan data dengan baik sebanyak 85,00%. Margin error atau tingkat kesalahan dalam pengambilan sampel adalah +- 2,92%. Jumlah sentimen positif adalah 745 dan jumlah sentimen negatif adalah 530. Presisi adalah rasio antara prediksi yang tepat dengan data yang diharapkan. Presisi untuk prediksi positif adalah 100% dan untuk prediksi negatif adalah 70,67%. *Recall* adalah rasio antara prediksi benar positif dengan seluruh data positif. Recall untuk data positif adalah 77,20% dan untuk data negatif adalah 99,07%.

4.10. Evaluasi

Untuk mengevaluasi model klasifikasi, kita dapat menggunakan *Confusion Matrix*. Confusion Matrix adalah tabel yang menampilkan kinerja model dalam mengklasifikasikan objek yang sebenarnya (True) dan yang salah (False). Dari *Confusion Matrix*, kita dapat menghitung berbagai metrik evaluasi, seperti *acuracy*, *precision*, *recall* dan *f-measure*. Berikut adalah rumus untuk menghitung Confusion Matrix.

Rumus dan perhitungan yang digunakan untuk menghitung akurasi adalah :

$$Acuracy = \frac{TP+TN}{\text{TP+TN+FP+FN}} = \frac{745+530}{745+530+5+220} = \frac{1275}{1500} \text{ x } 100\% = 85{,}00\%$$

Rumus dan perhitungan yang digunakan untuk menghitung nilai Precision :

$$Precision = \frac{TP}{\text{TP+FP}} = \frac{745}{745+5} \frac{745}{750} \text{ x } 100\% = 99{,}33\%$$

Rumus dan perhitungan yang digunakan dalam menghitung nilai Recall adalah :

$$Recall = \frac{TP}{\text{TP+FN}} = \frac{745}{745+220} \frac{745}{965} \text{ x } 100\% = 77{,}20\%$$

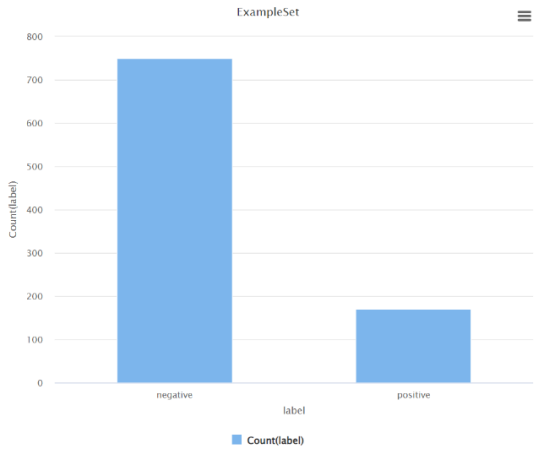
Rumus dan perhitungan yang digunakan dalam menghitung nilai f-measure adalah

$$\text{F} - \text{measure} = 2 * \frac{99{,}33*77{,}20}{99{,}33+77{,}20} = 86{,}94\%.$$

Keterangan :
TP : True Positive
TN : True Negatif
FP : False Positif
FN : False Negatif

Berdasarkan hasil perhitungan yang telah dilakukan, dapat disimpulkan bahwa metode yang digunakan memiliki performa yang baik dalam mengklasifikasikan data. Nilai akurasi yang diperoleh adalah 85,00%, yang menunjukkan bahwa sebagian besar data terklasifikasikan dengan benar. Nilai presisi yang diperoleh adalah 99,33%, yang menunjukkan bahwa hampir semua data yang terklasifikasikan sebagai positif memang benar-benar positif. Nilai recall yang diperoleh adalah 77,20%, yang menunjukkan bahwa sebagian besar data yang benar-benar positif berhasil terdeteksi. Nilai F-measure yang diperoleh adalah 86,94%, yang menunjukkan bahwa metode yang digunakan memiliki keseimbangan antara presisi dan recall.

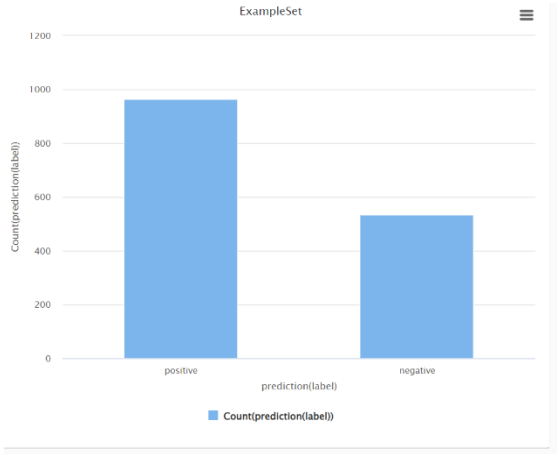
Dari hasil analisis yang dilakukan dengan proses pengumpulan data melalui teknik crawling, didapatkan 945 tweet yang berisi ulasan tentang suatu produk. Untuk mengetahui sentimen dari ulasan tersebut, dilakukan pelabelan dengan menggunakan vader lexcion, sebuah metode analisis sentimen berbasis leksikon. Metode ini sesuai dengan penelitian sebelumnya yang juga menggunakan vader lexcion. Ulasan diklasifikasikan sebagai sentimen positif jika compound > 0, dan sebagai sentimen negatif jika compound < 0. Hasil pelabelan dengan menggunakan vader lexcion adalah sebagai berikut.



Gambar 10 grafik sentimen prediksi awal

Pada gambar 10 dapat dilihat hasil pelabelan sentimen menggunakan *vader lexcion*. Dari hasil tersebut, kita dapat mengetahui bahwa terdapat 750 sentimen yang berlabel negatif dan 171 sentimen yang berlabel positif. Namun, jika kita menggunakan metode *naive bayes* untuk menganalisa sentimen, kita akan mendapatkan hasil yang berbeda. Berikut adalah

jumlah sentimen positif dan negatif yang didapatkan dengan menggunakan *naive bayes.*



Gambar 11 Grafik *Prediction Sentimen naïve bayes*

Berdasarkan gambar 11, metode naive bayes dapat memprediksi sentimen 745 dengan kategori positif dan sentimen 530 dengan kategori negatif. Hasil ini menunjukkan kinerja naive bayes dalam mengklasifikasikan sentiment.

Tabel 7 Perbandingan data awal dan prediksi data

No	Sentimen	Prediction (Sentimen naïve bayes)	Text
1.	*negative*	*positive*	bexlgium pliss seenak bayiii trus kmrn belu perak gara gara tiktokshop tutup
2.	*negative*	*negative*	cilembu fresh from oven dikasih driver ngasihnya nanas makannya sedih tiktok shop tutup drivernya rumahkan mobilnya masuk garasi sehatã driver
3.	*negative*	*negative*	yordanagil tanyakanrl kejadian sampe tutup cabang tutup cabang kemarin jasa kurir garaã tiktokshop tutup

Tabel 7 menampilkan beberapa contoh perbedaan antara label sentimen awal yang diberikan oleh vader lexcion dan label sentimen hasil prediksi naive bayes. Dari contoh tersebut, terlihat bahwa ada perubahan label sentimen dari negatif menjadi positif atau sebaliknya. Jika dilihat secara keseluruhan, vader lexcion memberikan label negatif untuk 750 sentimen dan label positif untuk 171 sentimen. Sementara itu, naive bayes memberikan label positif untuk 745 sentimen dan label negatif untuk 530 sentimen.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil analisa yang dilakukan terkait dengan sentimen terhadap berhentinya tiktokshop pada media sosial twitter menggunakan metode Naïve Bayes dapat disimpulkan Masyarakat Indonesia mempunyai pendapat beragam mengenai penangguhan TikTok. Ada yang mendukung keputusan ini dan ada pula yang menentang.

Namun, berdasarkan hasil penelitian yang telah dilakukan, sebagian besar masyarakat Indonesia memberikan opini positif tentang berhentinya tiktokshop dan Algoritma ini mengklasifikasikan teks berdasarkan kemunculan kata-kata tertentu yang memiliki nilai positif atau negatif. Dalam kasus ini, algoritma naïve bayes digunakan untuk menganalisis respon masyarakat terhadap kebijakan pemerintah yang menutup tiktokshop. Hasil analisis menunjukkan bahwa respon positif mendominasi dengan nilai prediksi positif 965, sedangkan respon negatif hanya 535 nilai akurasi 85.00%, margin error +-2.92% Precision 99,33% recall 77,20 %.

Sebagai acuan untuk penelitian selanjutnya, penulis menyarankan untuk menggunakan media sosial lain selain twitter, seperti facebook atau instagram, untuk mendapatkan data yang lebih beragam dan melihat perbedaan pola sentimen antara media sosial yang berbeda.

Penulis juga menyarankan untuk menerapkan metode analisis sentimen yang berbeda dengan menggunakan Bahasa pemrograman lain, seperti python atau bahasa r, untuk membandingkan hasil dan akurasi metode yang digunakan. Selain itu, penulis menyarankan untuk menambahkan kategori sentimen netral selain positif dan negatif, untuk mengidentifikasi opini yang tidak menunjukkan sikap atau emosi tertentu. Terakhir, penulis menyarankan untuk melakukan penentuan label sentimen dengan metode lexicon based, untuk melihat perbedaan hasil dengan metode machine learning.

DAFTAR PUSTAKA

[1] Nurmalasari and Latifah, “Jurnal Ekonomi & Manajemen Universitas Bina Sarana Informatika”, doi: 10.31294/jp.v21i1.

[2] K. A. Nugraha, “Analisis Sentimen Berbasis Emoticon pada Komentar Instagram Bahasa Indonesia Menggunakan Naïve Bayes,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 3, Dec. 2021, doi: 10.28932/jutisi.v7i3.4094.

[3] S. Saifulloh, “Analisis Promethee II Sebagai Pendukung Keputusan Pemilihan Media Sosial,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 3, Dec. 2020, doi: 10.28932/jutisi.v6i3.2956.

[4] W. D. Dinanti and W. Bharata, “Eksplorasi Minat Pembelian Konsumen Live Streaming Tiktok Shop Berdasarkan Framework Stimulus Organism Response (SOR),” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 2,

pp. 254–264, Jul. 2023, doi: 10.32736/sisfokom.v12i2.1658.

[5] Permendag, "PERIZINAN BERUSAHA, PERIKLANAN, PEMBINAAN, DAN PENGAWASAN PELAKU USAHA DALAM PERDAGANGAN MELALUI SISTEM ELEKTRONIK," 2023.

[6] 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.

[7] B. Laurensz, A. Sentimen, and E. Sediyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19 (Analysis of Public Sentiment on Vaccination in Efforts to Overcome the Covid-19 Pandemic)," 2021.

[8] A. Muljono, D. Putri Artanti, A. Syukur, A. Prihandono, and D. I. Rosal Moses Setiadi, "Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes," pp. 8–9, 2018, [Online]. Available: http://twitter.com

[9] N. Ikbar Wibowo, T. Andika Maulana, H. Muhammad, N. Aini Rakhmawati, and P. korespondensi, "Perbandingan Algoritma Klasifikasi Sentimen Twitter Terhadap Insiden Kebocoran Data Tokopedia," *JISKa*, vol. 6, no. 2, pp. 120–129, 2021, [Online]. Available: https://github.com/nadhifikbarw/ep-scrapper

[10] S. Fendyputra Pratama, R. Andrean, and A. Nugroho, "Analisis Sentimen Twitter Debat Calon Presiden Indonesia Menggunakan Metode Fined-Grained Sentiment Analysis," *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 4, no. 2, pp. 2541–3619, 2019, doi: 10.31328/jo.

[11] Irvandi, B. Irawan, and O. Nurdiawan, "NAIVE BAYES DAN WORDCLOUD UNTUK ANALISIS SENTIMEN WISATA HALAL PULAU LOMBOK," *INFOTECH journal*, vol. 9, no. 1, pp. 236–242, May 2023, doi: 10.31949/infotech.v9i1.5322.

[12] M. Khoirul, U. Hayati, and O. Nurdiawan, "ANALISIS SENTIMEN APLIKASI BRIMO PADA ULASAN PENGGUNA DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES," 2023.

[13] H. Putri, A. I. Purnamasari, A. R. Dikananda, O. Nurdiawan, and S. Anwar, "Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode NAÏVE BAYES dan KNN," *Building of Informatics, Technology and Science (BITS)*, vol. 3, no. 3, pp. 331–337, Dec. 2021, doi: 10.47065/bits.v3i3.1093.

[14] Firmansyah and O. Nurdiawan, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA FREQUENT PATTERN-GROWTH UNTUK MENENTUKAN POLA PEMBELIAN PRODUK CHEMICALS," 2023.

[15] U. Kulsum, M. Jajuli, and N. Sulistiyowati, "Analisis Sentimen Aplikasi WETV di Google Play Store Menggunakan Algoritma Support Vector Machine," *Journal of Applied Informatics and Computing (JAIC)*, vol. 6, no. 2, p. 205, 2022, [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[16] T. Prasetya, J. Eka Yanti, A. Irma Purnamasari, A. Rinaldi Dikananda, and S. Anwar, "Analisis Data Transaksi Terhadap Pola Pembelian Konsumen Menggunakan Metode Algoritma Apriori," *INFORMATICS FOR EDUCATORS AND PROFESSIONALS*, vol. 6, no. 1, pp. 43–52, 2021.

[17] O. Nurdiawan, A. Irma Purnamasari, and I. Ali, "Analisa Penjualan Mobil Dengan Menggunakan Algoritma K-Means Di PT. Mulya Putra Kencana," vol. 1, no. 2, pp. 32–35, 2021.

[18] M. Walid and F. Halimiyah, "Klasifikasi Kemandirian Siswa SMA/MA Double Track Menggunakan Metode Naive Bayes," *Jurnal ICT : Information Communication & Technology*, vol. 22, pp. 190–197, 2022.