

Analisis Sentimen Masyarakat Berdasarkan Komentar Kerja Sama Tiktok Shop dan Tokopedia di Instagram Menggunakan Metode *Naïve Bayes Classifier*

Zaenab Kurnia¹, Amalina Maryam Zakiiyyah², Nur Qodariyah Fitriyah³, Agus Milu Susetyo⁴

^{1,2,3,4}Universitas Muhammadiyah Jember

Alamat: Jl. Karimata No.49 Jember

Korespondensi penulis: zaenabkurnia1@gmail.com¹

Abstract. *The reopening of the Tiktok Shop and collaboration with Tokopedia has caught the public's attention. That's proven by a post on Instagram about their collaboration that received various responses. This research aims to conduct sentiment analysis to ascertain whether the public approves of the two's partnership. The Naïve Bayes Classifier method was used to analyze 641 comment data from the period 11 December 2023 to 11 February 2024. The results show the composition of positive, negative, neutral sentiment and unclassified data, as well as accuracy, precision and recall. Of the 641 data recorded, there were 269 neutral sentiment data, 194 negative sentiment data, 176 positive sentiment data, and 437 unclassified data. The Naïve Bayes Classifier model with oversampling techniques obtained an accuracy of 83%, precision of 81%, and recall of 93%.*

Keywords: *sentiment analysis, Tiktok Shop, Naïve Bayes Classifier.*

Abstrak. Kembalinya Tiktok Shop dan berkolaborasi dengan Tokopedia menyita perhatian masyarakat. Hal ini ditandai dengan, postingan kerjasama antar keduanya di instagram yang dibanjiri berbagai komentar. Penelitian ini bertujuan untuk melakukan analisis sentimen untuk mengetahui penerimaan masyarakat dan sebagai strategi pengembangan kerjasama keduanya. Metode *Naïve Bayes Classifier* digunakan untuk menganalisis 641 data komentar dari periode 11 Desember 2023 hingga 11 Februari 2024. Hasilnya menunjukkan komposisi sentimen positif, negatif, netral, dan data tidak terklasifikasi serta akurasi, presisi, dan recall model. Ditemukan bahwa dari 641 data, 269 data sentimen netral, 194 data sentimen negatif, 176 data sentimen positif, dan 437 data tidak terklasifikasi. Model *Naïve Bayes Classifier* dengan teknik *oversampling* mencapai akurasi 83%, presisi 81%, dan recall 93%.

Kata Kunci: Analisis sentimen, Tiktok Shop, *Naïve Bayes Classifier*.

LATAR BELAKANG

TikTok merupakan salah satu platform media sosial yang paling cepat berkembang di dunia. Aplikasi ini memungkinkan penggunaannya untuk membuat, membagikan, dan menonton video pendek yang biasanya disertai dengan musik, efek suara, dan berbagai filter. Sejak diluncurkan pada tahun 2016 oleh perusahaan Tiongkok *ByteDance*, TikTok telah menarik perhatian miliaran pengguna di seluruh dunia. Di Indonesia, popularitas TikTok melesat dengan cepat, terutama di kalangan remaja dan anak muda. Menurut data terbaru, jumlah pengguna TikTok di Indonesia mencapai 105 juta pada tahun 2023, meningkat sebesar 188% dibandingkan tahun sebelumnya (Sarina & Tanniewa, 2023). Seiring dengan meningkatnya jumlah pengguna, TikTok juga mengembangkan fitur *e-commerce* yang dikenal sebagai TikTok Shop. Fitur ini memungkinkan pengguna untuk membeli produk langsung melalui aplikasi TikTok, yang menarik perhatian banyak pelaku bisnis dan konsumen. TikTok Shop

Received Mei 10, 2024; Accepted Juni 06, 2024; Published Juni 30, 2024

*Zaenab Kurnia, zaenabkurnia1@gmail.com

menjual berbagai produk dan jasa dengan harga yang relatif lebih rendah karena sering memberikan diskon dan voucher yang menarik minat pengguna untuk berbelanja.

Namun, pada tanggal 26 September 2023, Menteri Perdagangan Indonesia mengesahkan Peraturan Nomor 31 tahun 2023 yang mengatur perizinan dan pengawasan perdagangan melalui sistem elektronik, yang berdampak pada pemberhentian operasi TikTok Shop di Indonesia. Alasan utama pemberhentian ini adalah karena TikTok Shop belum memiliki surat perizinan untuk melakukan perdagangan niaga elektronik dan kombinasi antara *e-commerce* dan media sosial dalam satu aplikasi. Meski demikian, pada tanggal 11 Desember 2023, TikTok Shop kembali beroperasi dengan syarat harus bekerja sama dengan Tokopedia, salah satu marketplace terbesar di Indonesia yang telah memiliki perizinan perdagangan niaga elektronik yang sah. Tokopedia beroperasi dengan model bisnis marketplace dan memiliki ketentuan yang berbeda dalam hal diskon, voucher, dan harga dibandingkan dengan TikTok Shop. Kerja sama ini menarik perhatian masyarakat, terutama di media sosial Instagram, di mana banyak komentar terkait kolaborasi tersebut. Oleh karena itu, penting untuk melakukan analisis sentimen terhadap komentar masyarakat untuk memahami penerimaan mereka terhadap kerja sama ini. Misalnya, penelitian oleh Nandaresta & Warman (2023) yang menyoroti pentingnya analisis sentimen dalam memahami penerimaan masyarakat terhadap produk atau layanan tertentu.

Analisis Sentimen merupakan proses untuk mencari makna dari sebuah pandangan, pendapat atau emosi yang disampaikan melalui sebuah text, ucapan, *post* (aktivitas di Internet), atau basis data dengan pendekatan *Natural Language Processing* (NLP) (Flores et al., 2020). Penelitian ini bertujuan untuk mengidentifikasi sentimen masyarakat terhadap kerja sama antara TikTok Shop dan Tokopedia dengan menggunakan metode *Naïve Bayes Classifier*. Beberapa pertanyaan yang menjadi fokus penelitian ini adalah Bagaimana komposisi sentimen positif, negatif, dan netral dari komentar masyarakat terhadap kerja sama TikTok Shop dan Tokopedia? Apakah terdapat perbedaan signifikan dalam sentimen masyarakat sebelum dan sesudah kerja sama tersebut. Penelitian ini mengacu pada berbagai literatur yang membahas tentang analisis sentimen, *machine learning*, dan implementasi metode *Naïve Bayes Classifier* dalam berbagai konteks. *Naïve Bayes Classifier* ditemukan oleh Thomas Bayes pada abad ke-18, Pengklasifikasi Bayes didasari oleh *teorema bayes* (Siregar et al., 2020). *Naïve Bayes Classifier* merupakan sebuah metode sederhana namun memiliki akurasi dan performansi yang tinggi dalam pengklasifikasian text (Maulidina, 2020).

KAJIAN TEORITIS

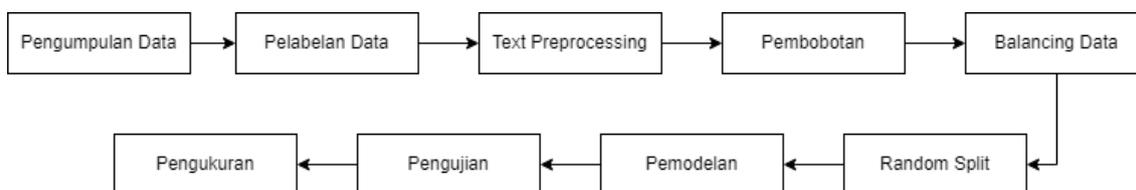
Machine Learning adalah aplikasi komputer dan algoritma matematika yang menghasilkan prediksi di masa yang akan datang dan diadopsi dengan cara pembelajaran yang berasal dari data (Yordanova dkk., 2013). *Text Mining* adalah proses menambang data dengan tujuan mencari kata-kata yang dapat mewakili isi dari dokumen berupa teks, di mana sumber data biasanya didapatkan dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen (Putri & Setiadi, 2014). *Sentiment analysis* adalah kombinasi dari data *mining* dan *text mining* (Akbar et al., 2023). Analisis Sentimen merupakan proses untuk mencari makna dari sebuah pandangan, pendapat, atau emosi yang disampaikan melalui sebuah teks, ucapan, post (aktivitas di Internet), atau basis data dengan pendekatan *Natural Language Processing* (NLP) (Flores dkk., 2020).

Pada penelitian Hidayatullah (2021), analisis sentimen produk Eiger dilakukan dengan menggunakan 700 data komentar dan diperoleh 62,71% (439) data dengan sentimen positif dan 37,29% (261) data dengan sentimen negatif. Hasilnya menunjukkan tingkat akurasi 85,71%, presisi 96,62%, dan *recall* 83,49% dengan skenario uji menggunakan data split. Penelitian lain yaitu Taufiq (2018) menggunakan 803 data tweet dengan 451 data latih, 58 data uji, dan 294 data tidak terklasifikasi. Pada data latih, dihasilkan 50,96% data sentimen positif dan 49,03% data sentimen negatif. Pada data uji, dihasilkan akurasi 65,51%, presisi 51,35%, dan *recall* 90,47%. Selain itu, penelitian Kadir(2021) memperoleh nilai pada jenis dataset binary seimbang (*cat vs dog*) dengan akurasi sebesar 89,20%, sedangkan untuk *multiclass* seimbang (*African wildlife*) menghasilkan akurasi sebesar 77,10%. Pada dataset *multiclass* tidak seimbang (*garbage classification*) menghasilkan akurasi 80,70%, dan untuk *binary* tidak seimbang (*painting vs photograph*), metode split data yang cocok diterapkan yakni *fixed train split* menghasilkan akurasi sebesar 77,10%. *Naïve Bayes Classifier*, ditemukan oleh Thomas Bayes pada abad ke-18, adalah pengklasifikasi yang didasari oleh teorema Bayes (Siregar dkk., 2020) dan merujuk pada implementasi praktis dari teori *Naive Bayes* dalam bentuk pengklasifikasian ke dalam kategori tertentu. Metode ini memiliki kelebihan seperti dapat digunakan untuk data kuantitatif maupun kualitatif, tidak memerlukan jumlah data yang banyak, tidak perlu melakukan data training yang banyak, dapat mengabaikan nilai yang hilang dalam perhitungan, cepat dan efisien, mudah dipahami dan diimplementasikan, baik untuk klasifikasi teks, sederhana dalam kode pemrograman, dan dapat digunakan untuk klasifikasi masalah biner ataupun *multiclass* (Rayuwati dkk., 2022). Namun, metode ini juga memiliki kekurangan seperti jika probabilitas kondisionalnya bernilai nol, maka probabilitas prediksi juga akan bernilai nol, keakuratannya tidak bisa diukur menggunakan satu probabilitas saja

sehingga butuh bukti-bukti lain, memerlukan pengetahuan awal atau pengetahuan mengenai masa sebelumnya untuk membuat keputusan, keberhasilannya sangat bergantung pada pengetahuan awal tersebut, dan dirancang hanya untuk mendeteksi kata-kata saja, tidak bisa berupa gambar.

METODE PENELITIAN

Penelitian ini terdiri dari beberapa tahap. Tahapan – tahapan tersebut yaitu pengumpulan data, *preprocessing*, implementasi *Naïve Bayes*, pengujian. Tahapan-tahapan tersebut digambarkan dalam bentuk diagram sebagai berikut :



Gambar 1. Tahapan Penelitian

Pengumpulan data merupakan proses pengambilan dataset dari akun Instagram @Tiktokofficialindonesia dan @Tokopedia pada postingan kerjasama antara keduanya yang diunduh melalui *plugin web browser google chrome scrapper* yaitu *IGCommentExport*. Kemudian data komentar yang didapatkan dari 2 postingan tersebut akan digabungkan menjadi satu. Pelabelan data dilakukan oleh seorang pakar berdasarkan kategori sentimen analisis. Pada proses *Text Preprocessing* terdapat beberapa proses seperti *case folding*, *tokenizing*, *stopword*, *stemming*. Kemudian dilakukan pembobotan kata menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* untuk memberi bobot pada setiap kata dalam setiap komentar. Setelah dibobotkan, selanjutnya yaitu jumlah data dalam kelas yang tidak seimbang akan diproses menggunakan *oversampling*. Kemudian data dibagi menjadi 80% data *train* dan 20% data *test*. selanjutnya yaitu metode *Naive Bayes Classifier* diimplementasikan terhadap data *train* dan data *test* untuk mengklasifikasikan komentar-komentar berdasarkan sentimennya. Terakhir dilakukan pengukuran hasil analisis menggunakan metrik-metrik seperti akurasi, presisi, dan *recall* untuk mengevaluasi performa model *Naive Bayes Classifier* dalam memprediksi sentimen dari komentar-komentar tersebut.

HASIL DAN PEMBAHASAN

Pengumpulan Data

Jumlah data yang diperoleh dari postingan akun *@Tiktokofficialindonesia* sebanyak 107 komentar dan dari postingan akun *@Tokopedia* sebanyak 971 komentar.

Pelabelan Data

Data yang telah digabung akan diproses ke tahap selanjutnya yaitu Pelabelan data. Data akan dibagi menjadi sentimen Positif, Negatif dan Netral yang dilakukan oleh seorang pakar. Berikut data yang telah di labeli:

Tabel 1. Pelabelan Data

No	Comment	Label
1	Semoga payment gopay nya segera hadir!	1
2	Apa pembayaran bisa pakai go-pay min	0
3	Wkwkw masih social commerce , kukira tokonya ngelink ke tokped kegiatan transaksi view toko semua ditokped eh sama kek dlu ..	-1
4	@arisa_takahashii saya belanja di tokped alhamdulillah selalu dpt cashback banyak dan potongan banyak , ongkir emg gk ada subsidi jd lebih hemat	1
5	Cara ikut campaign belilokal gmn di tiktokshop?	0
....
1074	@acengdis tokopedia punyahobi	X
1075	Bismillah semoga jualan di Tokopedia bisa serame dulu ☐ jangan lupa mampir Kak ke toko Cahayagrosirbandung ya♥☐	0
1076	kenapa sahamnya turun min? ☐	0
1077	♥☐♥☐♥☐♥☐	X
1078	Ngeri, congrats Tokopedia and Tiktok! ☐	1

Keterangan:

- 1 = Sentimen Positif
- 1 = Sentimen Negatif
- 0 = Sentimen Netral
- X = Sentimen Tidak Terklasifikasi

Terdapat 437 data tidak terklasifikasi karena tidak termasuk ke dalam kriteria label, positif, negatif, dan netral. Dalam penelitian ini data yang akan digunakan sebanyak 641 data.

Text Preprocessing

Tahapan *Text Preprocessing* terbagi menjadi beberapa tahapan yaitu *case folding*, *tokenizing*, *stopword*, *stemming*.

a. Case folding

Komponen-komponen tersebut tidak memiliki pengaruh terhadap sentimen maka dibuang nantinya. Berikut ini hasil dari proses *case folding*:

	TextBaku	Regular Exspresion
0	semoga payment gopay nya segera hadir!	semoga payment gopay nya segera hadir!
1	apa pembayaran bisa pakai gopay min	apa pembayaran bisa pakai gopay min
2	wkwkw masih social commerce , kukira tokonya ngelink ke tokped kegiatan transaksi view toko semu...	wkwkw masih social commerce kukira tokonya ngelink ke tokped kegiatan transaksi view toko semu...
3	saya belanja di tokped alhamdulillah selalu dapat cashback banyak dan potongan banyak , ongkir m...	saya belanja di tokped alhamdulillah selalu dapat cashback banyak dan potongan banyak ongkir m...
4	cara ikut campaign bellokal gimana di tiktokshop?	cara ikut campaign bellokal gimana di tiktokshop
...
636	tokopedia mobilbutut	tokopedia mobilbutut
637	mantap duolingo x tiktok	mantap duolingo x tiktok
638	bismillah semoga jualan di tokopedia bisa serame dulu jangan lupa mampir kak ke toko cahayagrosi...	bismillah semoga jualan di tokopedia bisa serame dulu jangan lupa mampir kak ke toko cahayagrosi...
639	kenapa saham nya turun min? 🙄	kenapa saham nya turun min
640	ngeri, congrats tokopedia and tiktok! 🙄	ngeri congrats tokopedia and tiktok

Gambar 2. Hasil Case folding

b. Tokenizing

Pada tahap ini yaitu membagi text menjadi perkata. Berikut merupakan hasil data yang telah melalui proses *Tokenizing*.

	Regular Exspresion	Tokenize
0	semoga payment gopay nya segera hadir	[semoga, payment, gopay, nya, segera, hadir]
1	apa pembayaran bisa pakai gopay min	[apa, pembayaran, bisa, pakai, gopay, min]
2	wkwkw masih social commerce kukira tokonya ngelink ke tokped kegiatan transaksi view toko semu...	[wkwkw, masih, social, commerce, kukira, tokonya, ngelink, ke, tokped, kegiatan, transaksi, view, toko, semu, ...]
3	saya belanja di tokped alhamdulillah selalu dapat cashback banyak dan potongan banyak , ongkir m...	[saya, belanja, di, tokped, alhamdulillah, selalu, dapat, cashback, banyak, dan, potongan, banyak, , ongkir, m, ...]
4	cara ikut campaign bellokal gimana di tiktokshop	[cara, ikut, campaign, bellokal, gimana, di, tiktokshop]
...
636	tokopedia mobilbutut	[tokopedia, mobilbutut]
637	mantap duolingo x tiktok	[mantap, duolingo, x, tiktok]
638	bismillah semoga jualan di tokopedia bisa serame dulu jangan lupa mampir kak ke toko cahayagrosi...	[bismillah, semoga, jualan, di, tokopedia, bisa, serame, dulu, jangan, lupa, mampir, kak, ke, to, ...]
639	kenapa saham nya turun min	[kenapa, saham, nya, turun, min]
640	ngeri congrats tokopedia and tiktok	[ngeri, congrats, tokopedia, and, tiktok]

Gambar 3. Hasil Tokenizing

c. Stopword

Pada tahap ini kata yang dianggap tidak penting akan dihilangkan. Contohnya kata sambung, kata depan, dan kata ganti. Berikut merupakan contoh data yang telah melewati proses *Stopword*:

	Tokenize	Stopwords
0	[semoga, payment, gopay, nya, segera, hadir]	[semoga, payment, gopay, hadir]
1	[apa, pembayaran, bisa, pakai, gopay, min]	[pembayaran, pakai, gopay, min]
2	[wkwkw, masih, social, commerce, kukira, tokonya, ngelink, ke, tokped, kegiatan, transaksi, view, toko, semu, ...]	[social, commerce, kukira, tokonya, ngelink, tokped, kegiatan, transaksi, view, ditokped]
3	[saya, belanja, di, tokped, alhamdulillah, selalu, dapat, cashback, banyak, dan, potongan, banyak, , ongkir, m, ...]	[belanja, tokped, alhamdulillah, cashback, potongan, ongkir, subsidi, hemat]
4	[cara, ikut, campaign, bellokal, gimana, di, tiktokshop]	[campaign, bellokal, gimana, tiktokshop]
...
636	[tokopedia, mobilbutut]	[tokopedia, mobilbutut]
637	[mantap, duolingo, x, tiktok]	[mantap, duolingo, x, tiktok]
638	[bismillah, semoga, jualan, di, tokopedia, bisa, serame, dulu, jangan, lupa, mampir, kak, ke, to, ...]	[bismillah, semoga, jualan, tokopedia, serame, lupa, mampir, kak, ya]
639	[kenapa, saham, nya, turun, min]	[saham, turun, min]
640	[ngeri, congrats, tokopedia, and, tiktok]	[ngeri, congrats, tokopedia, and, tiktok]

Gambar 4. Hasil Stopword

d. *Stemming*

Pada tahap ini yaitu mengubah kata imbuhan menjadi sebuah kata dasar. Berikut merupakan hasil data yang telah melalui proses *Stemming*.

	Stopwords	Stemming
0	[semoga, payment, gopay, hadir]	[moga, payment, gopay, hadir]
1	[pembayaran, pakai, gopay, min]	[bayar, pakai, gopay, min]
2	[social, commerce, kukira, tokonya, ngelink, tokped, kegiatan, transaksi, view, ditokped]	[social, commerce, kira, toko, ngelink, tokped, giat, transaksi, view, ditokped]
3	[belanja, tokped, alhamdulillah, cashback, potongan, ongkir, subsidi, hemat]	[belanja, tokped, alhamdulillah, cashback, potong, ongkir, subsidi, hemat]
4	[campaign, belilokal, gimana, tiktoshop]	[campaign, bilokal, gimana, tiktoshop]
...
636	[tokopedia, mobilbutut]	[tokopedia, mobilbutut]
637	[mantap, duolingo, x, tiktok]	[mantap, duolingo, x, tiktok]
638	[bismillah, semoga, jualan, tokopedia, serame, lupa, mampir, kak, ya]	[bismillah, moga, jual, tokopedia, serame, lupa, mampir, kak, ya]
639	[saham, turun, min]	[saham, turun, min]
640	[ngeri, congrats, tokopedia, and, tiktok]	[ngeri, congrats, tokopedia, and, tiktok]

Gambar 5. Hasil *Stemming*

TF-IDF

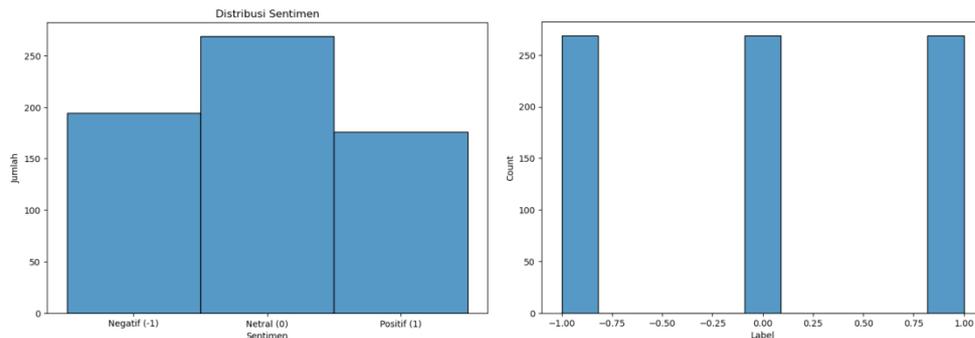
Selanjutnya yaitu menghitung nilai TF-IDF (*Term Frequency – Inverse Document Frequency*). Berikut hasil perhitungan TF-IDF:

Tabel 2. Hasil TF-IDF

No	Comment	Label
1.	['moga', 'payment', 'gopay', 'hadir']	['0,366404', '0,561323084', '0,485367', '0,561323084']
2.	['bayar', 'pakai', 'gopay', 'min']	['0,501675', '0,55132', '0,569039', '0,347223']
3.	['social', 'commerce', 'kira', 'toko', 'ngelink', 'tokped', 'giat', 'transaksi', 'view', 'ditokped']	['0,305794887', '0,246573989', '0,343159114', '0,327651559', '0,343159114', '0,188361', '0,365015786', '0,305794887', '0,365015786', '0,327651559']
4.	['belanja', 'tokped', 'alhamdulillah', 'cashback', 'potong', 'ongkir', 'subsidi', 'hemat']	['0,288751746', '0,225781314', '0,340346716', '0,392744358', '0,411332717', '0,256394531', '0,411332717', '0,437531538']
5.	['campaign', 'bilokal', 'gimana', 'tiktoshop']	['0,594731343', '0,559119602', '0,423823299', '0,392496767']
...
6.	['tokopedia', 'mobilbutut']	['0,482179576', '0,876072404']
7.	['mantap', 'duolingo', 'tiktok']	['0,579191594', '0,740432833', '0,341022457']
8.	['bismillah', 'moga', 'jual', 'tokopedia', 'serame', 'lupa', 'mampir', 'kak', 'ya']	['0,414308823', '0,267594239', '0,270440554', '0,228030528', '0,414308823', '0,389500549', '0,414308823', '0,299783105', '0,216151144']
9.	['saham', 'turun', 'min']	['0,498352825', '0,733604587', '0,462026808']
10.	['ngeri', 'congrats', 'tokopedia', 'and', 'tiktok']	['0,574144889', '0,511630216', '0,33612934', '0,475061476', '0,264434925']

Balancing data Oversampling

Data yang digunakan sebelum dilakukan tahap balancing data sebanyak 641 data terdiri dari 269 data sentimen netral, 194 data sentimen negatif, 176 data sentimen positif. Setelah dilakukan teknik *Balancing Oversampling*, jumlah data minoritas menjadi meningkat. Terdapat 269 data sentimen netral, 269 data sentimen negatif, 269 data sentimen positif, dengan jumlah keseluruhan data sampel yaitu 807 data. Berikut merupakan gambaran data sebelum dan setelah dilakukan teknik *balancing* data menggunakan *oversampling*:



Gambar 6. Diagram Data Sebelum (Kiri) Dan Setelah (Kanan) *Oversampling*

Implementasi Metode

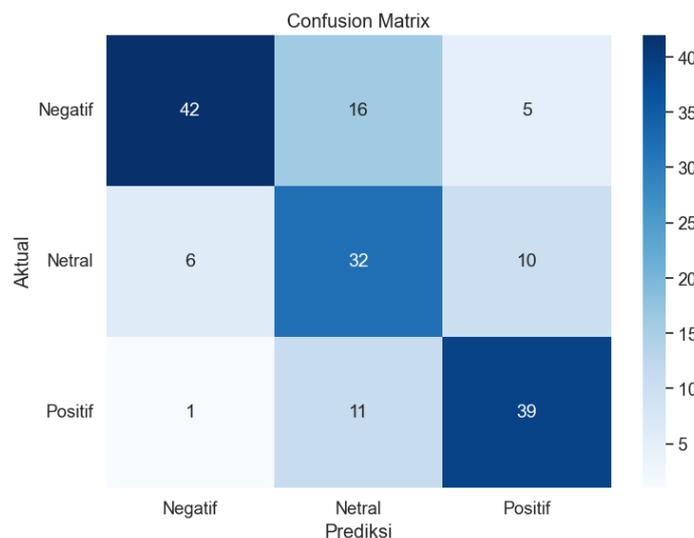
Proses pertama yaitu membagi data menjadi data *train* dan data *test* menggunakan *Random Split*. Pada skenario data yang telah melalui teknik *balancing* data *oversampling*, terdapat 807 data yang di acak terlebih dahulu kemudian di bagi menjadi 80% data latih dan 20% data *test* yang menghasilkan 645 data *train* dan 162 data *test*. Proses berikutnya yaitu implementasi *Naive Bayes Classifier* dengan *balancing* data *Oversampling* menghasilkan nilai prediksi yang diperoleh dari hasil implementasi *Naive Bayes Classifier*. Berikut merupakan hasil dari nilai aktual dan nilai prediksi:

Tabel 3. Nilai Aktual dan Prediksi

Comment	Nilai Aktual	Nilai Prediksi
Sekarang keranjang kuning atau hijau min?	0	0
Sdh ketebak, dibalik teriak2 di tanah abang, ada ijo yg teriak jg krn market lbh pilih oren	-1	0
Green card, biar langgeng persunting gadis lokal dulu ☐	0	1
@afaisholin pengiriman just ekspedisi ka, sicepat jnt jne ninja. jnt kebagian pengiriman 50%	0	0
@meicamelisa nah semoga jualanya ditiktok dan klo mau co langsung direct ke tokped☐	1	1
Pernah jadi Kebanggaan Indonesia ☐	1	1
Kenapa ngak sekalian yaa bisa bayar nya pake gopaypaylater	-1	1
Minto.. Kok di TikTokShop belum bisa bayar pake GoPayLater ya?? Kira-kira kapan nih bisa co pake GoPayLaternya? @tokopedia	0	0
semoga barangnya tetep bagus bagus	0	1
Sekarang keranjang kuning atau hijau min?	0	0
Sdh ketebak, dibalik teriak2 di tanah abang, ada ijo yg teriak jg krn market lbh pilih oren	-1	0

Confusion Matrix

Pada proses ini dilakukan perbandingan terhadap data aktual dan data prediksi untuk menentukan nilai elemen TP(*True Positive*), FP(*False Positive*), TN(*True Negative*), dan FN(*False Negative*). Berikut hasil perhitungan *Confusion Matrix Oversampling*:

**Gambar 7. Diagram Confusion Matrixs**

Dari hasil gambar diatas didapatkan akurasi 83%, presisi 81%, dan *recall* 93%. Dari hasil gambar diatas didapatkan akurasi 83%, presisi 81%, dan *recall* 93%. Penelitian ini menganalisis dataset sebanyak 641 data sentimen yang terdiri dari 269 data netral, 194 data

negatif, dan 176 data positif. Mayoritas data memiliki sentimen netral, hal tersebut berarti bahwa lebih banyak orang yang beropini lemah terhadap kerjasama tiktok shop dan tokopedia. Sedangkan perbedaan sentimen negatif dan netral tidak terlalu jauh, yaitu lebih tinggi sentimen negatif. Dari hasil komposisi tersebut menunjukkan persepsi masyarakat terhadap kerjasama keduanya mengandung opini yang tidak kuat namun cenderung ke arah negatif.

Untuk mengatasi ketidakseimbangan jumlah data di setiap kategori sentimen, teknik *balancing* data dengan metode *oversampling* digunakan dalam penelitian ini. Teknik ini memastikan bahwa jumlah data untuk setiap kategori sentimen seimbang, sehingga model dapat belajar dengan lebih baik dan mengurangi bias terhadap kategori yang dominan. Setelah dilakukan *oversampling*, model ini menghasilkan akurasi 83%, presisi 81% dan recall 93%. Hal ini menunjukkan bahwa model secara umum bekerja dengan baik dalam mengklasifikasi data. Namun masih terdapat beberapa prediksi positif yang salah.

KESIMPULAN DAN SARAN

Penelitian ini menggunakan 641 data dengan 269 data sentimen netral, 194 data sentimen negatif, 176 data sentimen positif. Dalam penelitian *Naive Bayes Classifier* ini dilakukan teknik *balancing* data *Oversampling* yang menghasilkan akurasi akurasi 83%, presisi 81%, dan *recall* 93%. Saran untuk Peneliti selanjutnya dapat menggunakan atau menambah metode klasifikasi lain untuk menemukan tingkat akurasi, presisi, dan *recall* yang lebih baik. Peneliti selanjutnya dapat menambah metode *preprocessing* lain jika diperlukan untuk meningkatkan kualitas proses data. Peneliti selanjutnya diharapkan menggunakan metode pemisah data lain agar mendapatkan nilai akurasi, presisi, dan *recall* yang lebih baik. Peneliti selanjutnya agar mengeksplorasi teknik-teknik *balancing* data yang lebih baru dan inovatif supaya mendapat solusi yang lebih baik untuk masalah ketidakseimbangan data.

DAFTAR REFERENSI

- Akbar, F. N. B., Arifianto, D., Zakiiyyah, M. A., & Susetyo, A. M. (2023). Analisis sentimen terhadap Anies Baswedan menggunakan metode Support Vector Machine: Studi kasus media sosial Twitter. *Jurnal Smart Teknologi*, 4(6), 2774–1702. <http://jurnal.unmuhjember.ac.id/index.php/JST>
- Flores, V. A., Jasa, L., & Linawati. (2020). Analisis sentimen untuk mengetahui kelemahan dan kelebihan pesaing bisnis rumah makan berdasarkan komentar positif dan negatif di Instagram. *Majalah Ilmiah Teknologi Elektro*, 19(1), 49. <https://doi.org/10.24843/mite.2020.v19i01.p07>
- Hidayatullah, R. (2021). Analisis sentimen masyarakat pada komentar produk Eiger di

Instagram menggunakan metode Naïve Bayes Classifier. [Institution or Journal].

Kadir, A. A. W. (2021). Perbandingan kinerja klasifikasi CNN berdasarkan strategi split data pada beragam dataset citra. <http://repository.unhas.ac.id:443/id/eprint/12034>

Maulidina, M. (2020). Analisis sentimen komentar warganet terhadap postingan Instagram menggunakan metode Naive Bayes Classifier dan TF-IDF (Studi kasus: Instagram Gubernur Jawa Barat Ridwan Kamil). University of Technology Yogyakarta. <http://eprints.uty.ac.id/id/eprint/6332>

Sarina, S., & Tanniewa, A. M. (2023). Implementasi algoritma Support Vector Learning terhadap analisis sentimen penggunaan aplikasi Tiktok Shop Seller Center. *Prosiding SISFOTEK*, 7(1), 165–170. <https://www.seminar.iaii.or.id/index.php/SISFOTEK/article/view/404>

Siregar, N. C., Siregar, R. R. A., & Sudirman, M. Y. D. (2020). Implementasi metode Naive Bayes Classifier (NBC) pada komentar warga sekolah mengenai pelaksanaan pembelajaran jarak jauh (PJJ). *Jurnal Teknologia*, 3(1), 102–110. <https://aperti.e-journal.id/teknologia/article/view/67>

Taufiq, R. (2018). Analisis sentimen pada Twitter terhadap kinerja Komisi Pemberantasan Korupsi (KPK) di Indonesia dengan metode Naïve Bayes. [Institution or Journal].