Analisis Sentimen Berbasis Aspek Pada Ulasan Produk Fashion *Shopee* Dengan Normalisasi Bahasa *Slang*

Wahyuni Nareswari[1], Herdiesel Santoso*[2]
[1] Program Studi Informatika, STMIK El Rahma Yogyakarta
[2] Program Studi Sistem Informasi STMIK El Rahma Yogyakarta
*e-mail: [1]wahyuninareswari798@gmail.com, *[2]herdiesel.santoso@stmikelrahma.ac.id*
*Correspondence author email: **

Abstrak

Pesatnya pertumbuhan e-commerce di Indonesia, khususnya pada platform Shopee kategori fashion, menghasilkan ulasan pelanggan dalam jumlah besar yang berpotensi dimanfaatkan sebagai evaluasi bisnis. Namun, ulasan tersebut menghadirkan tantangan dalam klasifikasi teks karena didominasi oleh bahasa tidak baku (slang) serta ketidakseimbangan distribusi data sentimen. Penelitian ini bertujuan mengembangkan model Analisis Sentimen Berbasis Aspek (Aspect-Based Sentiment Analysis/ABSA) dengan membandingkan algoritma Multinomial Naïve Bayes (MNB) dan Support Vector Machine (SVM), serta mengintegrasikan normalisasi bahasa slang dan teknik Synthetic Minority Oversampling Technique (SMOTE). Data yang digunakan sebanyak 5.877 ulasan pelanggan, yang dibagi menggunakan skema train-test split 80:20 dan dievaluasi menggunakan confusion matrix dengan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa tanpa preprocessing, SVM memperoleh akurasi sebesar 0,852, lebih tinggi dibandingkan MNB sebesar 0,722. Setelah penerapan normalisasi bahasa slang dan SMOTE, performa kedua model meningkat, di mana MNB mencapai akurasi 0,853 dan SVM mencapai akurasi 0,938 dengan presisi 0,939, recall 0,938, dan F1-score 0,937. Berdasarkan dataset dan skenario pengujian yang digunakan, integrasi normalisasi bahasa slang dan SMOTE terbukti meningkatkan performa klasifikasi sentimen berbasis aspek, dengan SVM menunjukkan kinerja terbaik pada data ulasan produk fashion Shopee.

Kata kunci*— Analisis Sentimen, Berbasis Aspek, SVM, SMOTE, Normalisasi Bahasa Slang*

1. PENDAHULUAN

Transformasi digital mendorong peningkatan aktivitas *e-commerce* di Indonesia, khususnya pada kategori fashion di platform *Shopee*. Tingginya aktivitas transaksi di sektor ini secara otomatis menghasilkan volume data ulasan pelanggan yang sangat besar dan dinamis [1]. Ulasan pelanggan mengandung informasi penting untuk evaluasi bisnis. Namun, karakteristik teks yang tidak terstruktur dan didominasi bahasa tidak baku menjadi tantangan dalam proses analisis. Ulasan tersebut memuat informasi krusial mengenai pengalaman konsumen yang sangat berharga untuk evaluasi performa bisnis. Namun, dalam pemanfaatannya, data ini menghadirkan tantangan teknis yang signifikan karena karakteristik teks ulasan produk fashion yang cenderung tidak terstruktur, memuat banyak singkatan, dan didominasi oleh bahasa *slang* khas pengguna media sosial di Indonesia [2]. Ulasan pelanggan didominasi oleh penggunaan bahasa tidak baku seperti ‘bgs bgt’, ‘ga nyesel’, dan ‘krmnya lma’ yang menyulitkan proses komputasi teks [3].

Dalam mengevaluasi ulasan pelanggan, pendekatan analisis sentimen pada tingkat dokumen dinilai kurang presisi. Rating bintang secara global sering kali bias; seorang pembeli mungkin memberikan rating bintang 4 karena kualitas produknya bagus, namun di dalam teks ulasannya ia mengeluhkan pengiriman yang lambat. Oleh karena itu, diperlukan Analisis Sentimen Berbasis Aspek (ABSA) yang mampu mengekstraksi sentimen secara spesifik pada elemen-elemen krusial seperti kualitas produk, pengiriman, dan layanan [4]. Melalui ABSA, hasil analisis menjadi lebih tepat sasaran pada setiap dimensi bisnis.

Dalam ranah Pemrosesan Bahasa Alami (NLP), algoritma *Multinomial Naïve Bayes* (MNB) dipilih sebagai metode klasifikasi utama untuk ABSA karena keandalannya dalam menangani data teks berdimensi tinggi secara efisien [5]. Sebagai pembanding komprehensif, algoritma *Support Vector Machine* (SVM) juga diimplementasikan mengingat kemampuannya

dalam membentuk batas keputusan yang optimal pada klasifikasi teks [6]. Secara operasional, algoritma probabilistik seperti MNB sangat bergantung pada distribusi frekuensi kata [7]. Tantangan teknis muncul akibat masifnya penggunaan bahasa tidak baku pada ulasan. Jika data tidak melalui tahapan normalisasi bahasa slang, algoritma akan memperlakukan variasi penulisan dari kata yang sama (misalnya "bgs" dan "bagus") sebagai entitas fitur yang sepenuhnya berbeda [8]. Kondisi ini memicu masalah ketersebaran fitur (feature sparsity) yang memecah bobot probabilitas kata, sehingga secara langsung akan menurunkan akurasi model. Oleh karena itu, integrasi antara normalisasi bahasa slang dan pemodelan ABSA menjadi sangat krusial agar algoritma dapat membaca data secara konsisten.

Di samping kelemahan pada level representasi teks, performa algoritma klasifikasi (baik MNB maupun SVM) juga sangat rentan dipengaruhi oleh distribusi kelas sentimen yang tidak seimbang (*data imbalance*) [9], [10]. Pada data ulasan *e-commerce*, kelas sentimen tertentu sering kali jauh lebih dominan dibandingkan kelas lainnya. Ketimpangan ini menyebabkan model cenderung bias dalam memprediksi kelas mayoritas dan mengabaikan kelas minoritas. Oleh karena itu, penelitian ini mengintegrasikan teknik penyeimbang data *Synthetic Minority Over-sampling Technique* (SMOTE) yang beroperasi dengan menghasilkan data sintetis pada kelas minoritas, sehingga kinerja model klasifikasi dapat meningkat secara signifikan [6].

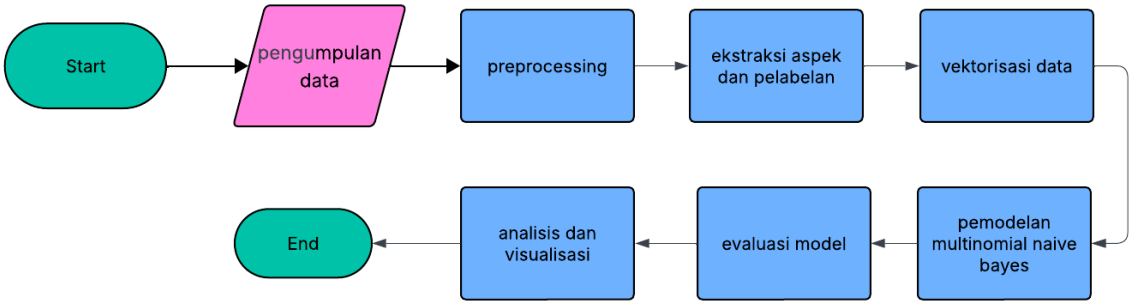
Beberapa penelitian terkait telah berupaya melakukan analisis sentimen pada domain *e-commerce*, namun masih meninggalkan *research gap*. Penelitian yang dilakukan oleh [11] dan [10] menggunakan algoritma Naïve Bayes untuk mengklasifikasikan ulasan aplikasi *Shopee*, namun fokusnya masih berada pada sentimen umum aplikasi dan belum menyentuh level produk maupun menerapkan normalisasi bahasa secara spesifik. Pendekatan analisis sentimen berbasis aspek mulai dikembangkan oleh [9], akan tetapi objek penelitiannya difokuskan pada aspek teknis aplikasi di luar *e-commerce*. Lebih lanjut, penelitian oleh [12] telah menerapkan analisis sentimen berbasis aspek pada ulasan *Shopee* menggunakan *Naïve Bayes*, namun interaksi bahasanya murni merujuk pada fungsionalitas sistem aplikasi. Sementara itu, [13] menerapkan MNB pada produk Bukalapak, namun masih terbatas pada kamus leksikon standar. Pada ranah *fashion*, [14] menggunakan *Support Vector Machine* (SVM) untuk melihat tren *fashion* di media sosial, namun pendekatannya belum membedah aspek ulasan secara spesifik. Bahkan, penelitian-penelitian terbaru seperti yang dilakukan oleh [15] masih berfokus pada klasifikasi sentimen ulasan produk menggunakan pendekatan machine learning dan Naïve Bayes dengan tahapan pra-pengolahan (*preprocessing*) standar. Penelitian tersebut belum secara spesifik menggunakan leksikon *slang* khusus fashion untuk diintegrasikan pada ekstraksi aspek.

Sebagian besar penelitian sebelumnya masih berfokus pada analisis sentimen tingkat dokumen dan belum mengintegrasikan pendekatan berbasis aspek dengan normalisasi bahasa slang secara spesifik. Selain itu, penelitian yang mengombinasikan Analisis Sentimen Berbasis Aspek (ABSA), normalisasi bahasa slang, dan teknik penyeimbangan data seperti *Synthetic Minority Oversampling Technique* (SMOTE) dalam satu kerangka terpadu masih terbatas. Di sisi lain, studi yang membandingkan kinerja algoritma *Multinomial Naïve Bayes* (MNB) dan *Support Vector Machine* (SVM) pada data ulasan dengan karakteristik bahasa tidak baku serta distribusi kelas yang tidak seimbang, khususnya dengan evaluasi sebelum dan sesudah proses normalisasi dan SMOTE, masih jarang dilakukan secara komprehensif. Oleh karena itu, penelitian ini bertujuan untuk menganalisis pengaruh normalisasi bahasa slang dan penyeimbangan data (SMOTE) terhadap kinerja algoritma MNB dan SVM dalam mengklasifikasikan ulasan non-formal. Selain itu, penelitian ini juga mengkaji hasil analisis sentimen secara spesifik pada aspek kualitas produk, pelayanan, dan pengiriman pada ulasan produk fashion di platform Shopee. Pendekatan ini diharapkan mampu menghasilkan model analisis sentimen yang lebih akurat, terstruktur, dan relevan untuk mendukung pengambilan keputusan bisnis.

2. METODE PENELITIAN

Penelitian ini dilakukan melalui tahapan pemrosesan teks yang terstruktur, mulai dari pengumpulan data hingga evaluasi model. Keseluruhan alur metodologi yang diterapkan, mulai

dari proses pengumpulan data mentah hingga tahap evaluasi kinerja model klasifikasi, divisualisasikan secara komprehensif pada Gambar 1.



Gambar 1. Diagram alur metode penelitian

Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data primer berupa ulasan pelanggan pada produksi gamis wanita di *Shopee*. Data diperoleh menggunakan teknik *web scraping* dengan bantuan *Instant Data Scraper*, yaitu ekstensi pada *browser* Google Chrome yang digunakan untuk mengekstraksi data secara otomatis dari halaman web *Shopee*. Data yang dikumpulkan dalam penelitian ini berjumlah 5.877 ulasan yang terdiri dari teks ulasan, rating bintang, serta waktu ulasan. Data yang diperoleh bersifat tidak terstruktur dan mengandung variasi bahasa, termasuk slang, singkatan, dan typo, sehingga perlu dilakukan tahap *preprocessing* sebelum digunakan dalam analisis lebih lanjut.

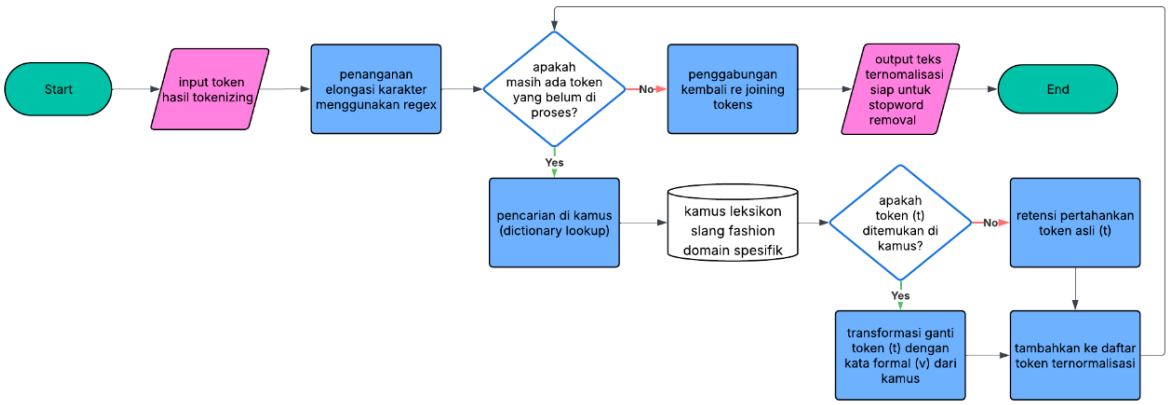
Pra-pengolahan Teks (Text Preprocessing)

Tahapan ini bertujuan untuk mengubah data teks yang tidak terstruktur menjadi data yang siap diolah oleh algoritma [16]. Tahapannya meliputi:

1. *Case Folding*: Mengubah semua huruf menjadi huruf kecil (*lowercase*).
2. *Filtering/Cleansing*: Menghapus angka, simbol, tanda baca, URL, dan karakter non-alfabet lainnya.
3. *Tokenizing*: Memecah kalimat menjadi potongan kata tunggal (*token*).
4. Normalisasi Bahasa Slang: Tahap krusial di mana kata-kata tidak baku atau singkatan (misal: "bgs", "gercep", "pw") dipetakan ke dalam kamus leksikon fashion buatan peneliti untuk diubah menjadi kata baku ("bagus", "gerak cepat", "nyaman").
5. *Stopword Removal*: Menghapus kata-kata umum yang tidak memiliki pengaruh signifikan terhadap sentimen (seperti "dan", "yang", "di").

Mekanisme Normalisasi Bahasa Slang

Normalisasi bahasa slang bertujuan menyelaraskan distribusi frekuensi kata agar representasi fitur menjadi lebih konsisten. Proses ini dilakukan dengan mentransformasikan kata-kata tidak baku, singkatan, dan istilah *slang* spesifik dunia fashion ke dalam bentuk formal sesuai dengan Kamus Besar Bahasa Indonesia (KBBI) atau istilah teknis yang disepakati [16]. Alur kerja dari proses normalisasi bahasa slang yang diusulkan pada penelitian ini dapat dilihat secara rinci pada Gambar 2. Mekanisme normalisasi beroperasi dengan mengkomparasi setiap token hasil tahapan *tokenizing* dengan sebuah kamus leksikon (*slang dictionary*) yang dikembangkan secara spesifik untuk domain produk *fashion*. Alur dimulai ketika sistem menerima *input* berupa token tunggal. Selanjutnya, sistem melakukan proses pencarian (*lookup*) untuk mengecek apakah token tersebut terdaftar sebagai lema tidak baku di dalam kamus [17]. Jika ditemukan kecocokan (*match*), sistem akan melakukan penggantian (*replace*) menjadi bentuk kata bakunya. Jika tidak, token dipertahankan dalam bentuk aslinya (*keep original*).



Gambar 2. Alur kerja dari proses normalisasi bahasa *slang*

Pendekatan Analisis Sentimen Berbasis Aspek (*Aspect-Based Sentiment Analysis/ABSA*) dalam penelitian ini dilakukan melalui mekanisme ekstraksi aspek berbasis leksikon (*lexicon-based approach*) yang dikombinasikan dengan proses pencocokan kata kunci (*keyword matching*). Tiga aspek utama yang dianalisis, yaitu kualitas produk, pelayanan, dan pengiriman, ditentukan berdasarkan studi literatur serta karakteristik umum ulasan pada platform e-commerce. Dalam pendekatan ini, satu ulasan dapat diklasifikasikan ke dalam lebih dari satu aspek sesuai dengan konteks yang terkandung di dalamnya.

Secara operasional, setiap ulasan diproses melalui tahap *keyword matching* untuk masing-masing aspek secara independen, sehingga satu dokumen dapat menghasilkan beberapa pasangan *(aspek, sentimen)*. Untuk setiap aspek yang terdeteksi, dilakukan proses klasifikasi sentimen secara terpisah menggunakan representasi fitur yang sama (TF-IDF), sehingga setiap aspek menghasilkan label sentimen yang spesifik. Pendekatan ini termasuk dalam skema *multi-label classification* pada level aspek, dengan proses klasifikasi sentimen yang bersifat *multi-class* pada masing-masing aspek.

Konstruksi Kamus Leksikon Slang Fashion

Penelitian ini membangun sebuah kamus leksikon khusus (*Domain-Specific Lexicon*) yang dikumpulkan secara manual melalui observasi pada ulasan Shopee. Kamus ini mencakup tiga kategori utama:

1. Singkatan Umum: Contoh: "bgs" → "bagus", "bgt" → "banget", "krm" → "kirim".
2. Istilah Slang Fashion/E-commerce: Contoh: "gercep" → "gerak cepat", "pw" → "nyaman", "nerawang" → "tembus pandang", "recsell" → "rekomendasi penjual".
3. Typo & Elongasi Karakter: Menangani penekanan emosi seperti "baguuuuusss" menjadi "bagus" menggunakan teknik *regular expression* (*Regex*) sebelum dilakukan *matching* ke kamus leksikon.

Algoritma Pencocokan String (String Matching)

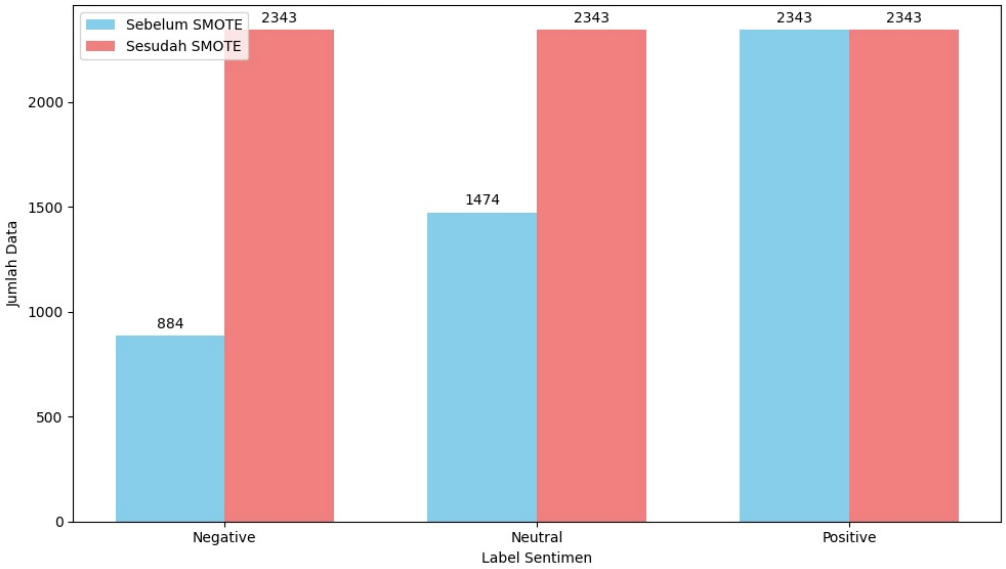
Proses normalisasi dilakukan dengan metode *Dictionary-based Lookup*. Setiap token (t) yang dihasilkan dari tahap tokenizing akan diperiksa keberadaannya di dalam kamus leksikon (L) [18]. Jika $t \in L$, maka t akan digantikan dengan nilai formal v yang berkorespondensi dalam kamus. Secara logis, proses ini mengikuti fungsi persamaan (1).

$$f(t) = \begin{cases} L(t), jika\ t\ \in\ Kamus\ Slang \\ t,\ jika\ t\ \notin\ Kamus\ Slang \end{cases} \quad (1)$$

Keterangan: $f(t)$ = Token hasil normalisasi (kata baku); t = Token input dari hasil tokenizing; L = Himpunan kamus leksikon bahasa slang; v = Nilai kata baku yang berpasangan dengan token t di dalam kamus L.

Penyeimbangan Data Dengan SMOTE

Data ulasan e-commerce umumnya mengalami ketidakseimbangan kelas sentimen. Berdasarkan hasil pengumpulan data awal yang divisualisasikan pada Gambar 3 sebelum SMOTE, distribusi kelas sentimen menunjukkan kondisi ketidakseimbangan (*imbalanced data*) yang cukup ekstrem. Jumlah ulasan dengan sentimen Positif (sebagai kelas mayoritas) mendominasi secara signifikan dibandingkan dengan ulasan bersentimen Negatif (sebagai kelas minoritas). Kondisi tersebut dapat menyebabkan model klasifikasi seperti *Multinomial Naïve Bayes* (MNB) dan *Support Vector Machine* (SVM) mengalami bias terhadap kelas mayoritas sehingga kemampuan model dalam mengenali kelas minoritas menjadi kurang optimal [5]. Oleh karena itu, digunakan metode *Synthetic Minority Oversampling Technique* (SMOTE) pada data latih setelah proses *train-test split* dengan rasio 80:20 guna menghasilkan distribusi data yang lebih seimbang serta menghindari data *leakage* [7]. Penerapan SMOTE dilakukan menggunakan parameter bawaan (*default parameter*) yaitu k=5 sebagai jumlah *nearest neighbor* dari *library sklearn* yang digunakan pada proses *oversampling* data minoritas.



Gambar 3. Perbandingan label sentimen sebelum dan sesudah SMOTE

Guna mengatasi permasalahan bias komputasi tersebut, penelitian ini menerapkan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) pada data latih (*training data*). Berbeda dengan teknik *random over-sampling* tradisional yang hanya menduplikasi data minoritas dan berisiko memicu *overfitting*, algoritma SMOTE bekerja dengan cara menginterpolasi titik-titik data minoritas terdekat (*k-nearest neighbors*) untuk menghasilkan data sintetis baru yang unik namun tetap merepresentasikan karakteristik kelas aslinya [6]. Teknik ini telah berhasil menyeimbangkan distribusi data dengan menaikkan jumlah sampel pada sentimen Negatif (secara sintetis) hingga proporsinya setara dengan sentimen Positif. Distribusi kelas yang seimbang ini sangat krusial dalam tahapan pemodelan. Bagi algoritma MNB, data yang seimbang akan mencegah dominasi probabilitas *prior* dari kelas mayoritas. Sementara bagi algoritma SVM, keseimbangan ini akan membantu algoritma dalam membentuk batas keputusan (*margin hyperplane*) yang lebih adil dan optimal dalam memisahkan ruang fitur antar sentimen [5].

Klasifikasi Dengan Multinomial Naïve Bayes

Penelitian ini menggunakan varian *Multinomial Naïve Bayes* (MNB) karena kemampuannya dalam menangani distribusi frekuensi kata dalam dokumen teks [7]. Proses klasifikasi dilakukan dengan menghitung probabilitas posterior suatu kelas c berdasarkan dokumen d menggunakan persamaan (2).

$$P(c|d) = \frac{P(c)\ \prod_{i=1}^{n} P\ (t_i|c)}{P(d)} \quad (2)$$

Keterangan: $P(c \mid d)$= probabilitas kelas c pada dokumen d; $P(c)$= probabilitas prior kelas; $P(t_i \mid c)$= probabilitas kemunculan term t_i pada kelas c; n= jumlah kata dalam dokumen.
Untuk menghindari nilai probabilitas nol, digunakan *Laplace Smoothing* dalam perhitungan probabilitas *term* menggunakan persamaan (3)[13].

$$P(t|c) = \frac{T_{ct} + 1}{\sum_{t' \in V} (T_{ct'} + B)} \quad (3)$$

Keterangan: T_{ct}= frekuensi kemunculan term t pada kelas c; V= kosakata (vocabulary) seluruh dokumen; B= jumlah total kata unik dalam kosakata.

Proses normalisasi bahasa *slang* memiliki pengaruh langsung terhadap perhitungan *likelihood* pada model MNB. Tanpa normalisasi, kata seperti "bagus" dan "bgs" akan diperlakukan sebagai fitur yang berbeda sehingga menghasilkan *sparsity* pada data. Dengan adanya normalisasi seperti pada menggunakan persamaan (4), kedua token tersebut digabung menjadi satu entitas sehingga frekuensi kemunculan term meningkat dan distribusi probabilitas menjadi lebih stabil.

$$P(t_{formal}|c) = \frac{(T_{c,bagus} + T_{c,bgs}) + 1}{\sum_{t' \in V} (T_{ct'} + B)} \quad (4)$$

Penggabungan ini memperkuat bobot probabilitas pada kelas sentimen yang sesuai serta mengurangi dimensi kosakata, sehingga model menjadi lebih efisien dan lebih akurat dalam mengenali pola sentimen pada aspek kualitas, pelayanan, maupun pengiriman. Proses klasifikasi dilakukan menggunakan algoritma MNB yang bekerja berdasarkan probabilitas kemunculan kata dalam suatu kelas sentimen [8]. Model dilatih menggunakan data hasil TF-IDF yang telah diseimbangkan dengan SMOTE, kemudian digunakan untuk mengklasifikasikan ulasan ke dalam kelas positif dan negatif berdasarkan nilai probabilitas tertinggi dari setiap kelas.

Klasifikasi Dengan Support Vector Machine (SVM)

Selain menggunakan *Multinomial Naïve Bayes*, penelitian ini juga menggunakan *Support Vector Machine* (SVM) sebagai metode pembanding untuk mengevaluasi performa klasifikasi. SVM digunakan untuk membentuk *hyperplane* optimal yang memisahkan data antar kelas dalam ruang fitur berdasarkan hasil representasi TF-IDF[14]. Proses klasifikasi pada SVM dapat direpresentasikan menggunakan fungsi keputusan pada persamaan (5).

$$\boldsymbol{f(x) = w \cdot x + b} \quad (5)$$

Keterangan: w merupakan vektor bobot (*weight vector*), x merupakan vektor fitur data, dan b merupakan bias atau konstanta pemisah *hyperplane*. Hasil dari SVM digunakan sebagai pembanding terhadap MNB untuk melihat model mana yang memberikan performa terbaik dalam klasifikasi sentimen ulasan produk.

Evaluasi Model

Model klasifikasi digunakan untuk mengelompokkan ulasan ke dalam tiga kategori sentimen yaitu positif, negatif, dan netral. Sentimen netral mempresentasikan ulasan yang tidak menunjukkan kecenderungan opini yang dominan terhadap suatu aspek tertentu. Oleh karena itu, kelas netral tetap dipertahankan pada proses pelatihan dan pengujian model sehingga pendekatan yang digunakan termasuk ke dalam *multi-class classification*. Selain itu, kelas netral juga merepresentasikan opini objektif/*non-opinionated*, sehingga penting untuk menjaga representasi distribusi sentimen secara realistis.

Selanjutnya, performa model diukur menggunakan *Confusion Matrix* untuk membandingkan hasil prediksi sistem dengan label aktual [8]. *Confusion Matrix* digunakan untuk menghitung kinerja klasifikasi berdasarkan empat kategori utama yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Berdasarkan nilai tersebut,

digunakan beberapa metrik evaluasi untuk mengukur performa model, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian guna mengevaluasi performa model secara objektif.
Accuracy digunakan untuk mengukur tingkat kedekatan hasil prediksi dengan nilai sebenarnya, dihitung menggunakan persamaan (6).

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

Precision digunakan untuk mengukur ketepatan prediksi positif dibandingkan seluruh hasil prediksi positif, dihitung menggunakan persamaan (7).

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

Recall digunakan untuk mengukur kemampuan model dalam menemukan data positif yang sebenarnya, yang dihitung menggunakan persamaan (8).

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

F1-Score merupakan rata-rata harmonik dari precision dan recall yang digunakan untuk menyeimbangkan kedua metrik tersebut, yang dihitung menggunakan persamaan (9).

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

Keterangan: *True Positive* = Data positif terprediksi positif; *True Negative* = Data negatif terprediksi negatif; *False Positive*: Data negatif terprediksi positif; *False Negative*: Data positif terprediksi negatif.

3. HASIL DAN PEMBAHASAN

Hasil Pengumpulan Data

Penelitian ini mengumpulkan 5.877 ulasan pelanggan pada produk fashion wanita di platform Shopee menggunakan teknik web scraping dengan Instant Data Scraper. Data yang diperoleh terdiri dari teks ulasan, rating, dan waktu ulasan. Selanjutnya, ulasan diklasifikasikan ke dalam tiga aspek utama, yaitu kualitas produk, pelayanan, dan pengiriman menggunakan metode *keyword matching*. Setiap ulasan diperiksa berdasarkan kemunculan kata kunci yang merepresentasikan masing-masing aspek, seperti “bagus”, “jahitan”, dan “bahan” pada aspek kualitas produk, “ramah” dan “respon cepat” pada aspek pelayanan, serta “pengiriman cepat” dan “packing” pada aspek pengiriman. Contoh *keyword* aspek ditunjukkan pada Tabel 1.

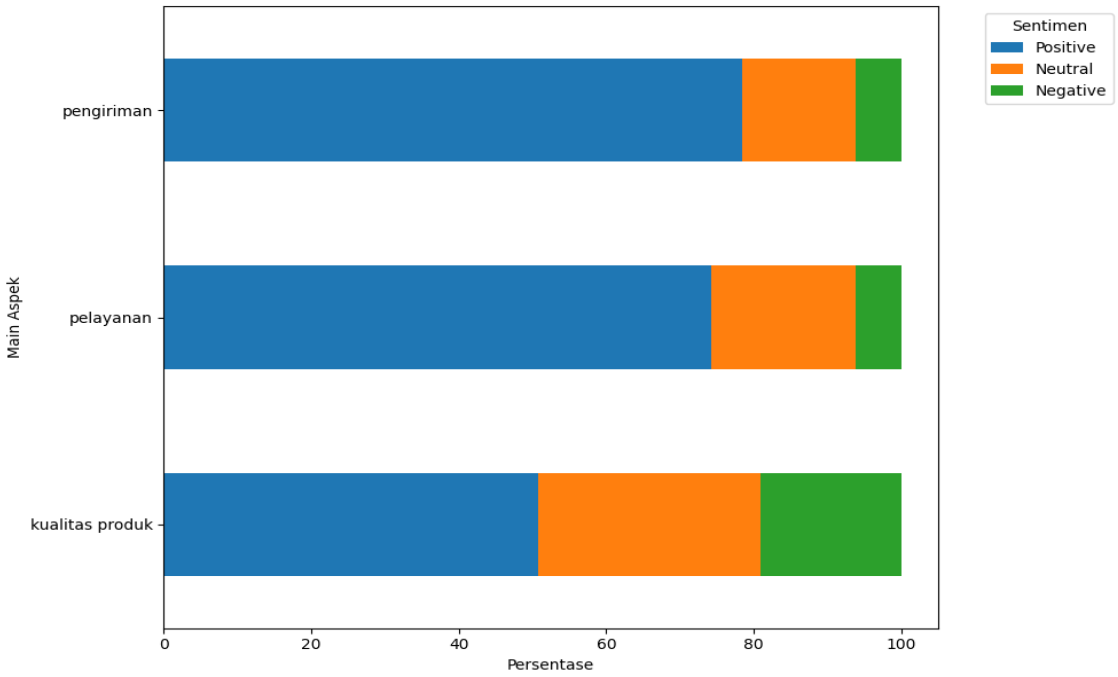
Tabel 1. Kamus *Keyword* Aspek

Aspek	Contoh *Keyword*
Kualitas Produk	bagus, bahan, jahitan, tebal, adem
Pelayanan	ramah, respon, cepat, seller
Pengiriman	kirim, lama, cepat, packing

Dalam pendekatan Analisis Sentimen Berbasis Aspek (ABSA), satu ulasan dapat mengandung lebih dari satu aspek. Oleh karena itu, satu ulasan dapat diklasifikasikan ke dalam beberapa aspek sekaligus sesuai dengan konteks yang terkandung di dalamnya. Hal ini menyebabkan jumlah data pada tingkat aspek lebih besar dibandingkan jumlah ulasan awal. Distribusi data pada masing-masing aspek ditunjukkan pada Tabel 2.

Tabel 2. Distribusi sentimen (Positif, Negatif, Netral) pada masing-masing aspek

Aspek	Sentimen Positif	Sentimen Negatif	Sentimen Netral
kualitas produk	2906	1088	1735
pelayanan	239	63	20
pengiriman	515	101	40



Gambar 4. Distribusi sentimen per tiga aspek utama (%)

Berdasarkan Tabel 2 dan Gambar 4, aspek kualitas produk memiliki jumlah ulasan paling dominan dibandingkan aspek lainnya. Selain itu, pada aspek pengiriman ditemukan proporsi sentimen negatif yang relatif lebih tinggi. Pendekatan berbasis aspek ini memiliki beberapa kelebihan. Pertama, mampu menangkap informasi sentimen secara lebih spesifik pada setiap aspek. Kedua, meningkatkan jumlah data pada level aspek sehingga memperkaya distribusi fitur dalam proses pelatihan model. Ketiga, memungkinkan identifikasi permasalahan secara lebih terarah, seperti membedakan keluhan terkait kualitas produk dan pengiriman.

Hasil Preprocessing Data

Tahap preprocessing menghasilkan data teks yang lebih bersih dan terstruktur sehingga siap digunakan dalam proses klasifikasi. Salah satu tahapan kunci dalam penelitian ini adalah normalisasi bahasa slang menggunakan kamus *leksikon* yang disusun secara manual sesuai dengan karakteristik ulasan produk fashion. Contoh hasil normalisasi ditunjukkan pada Tabel 3.

Tabel 3. Contoh Pemrosesan Normalisasi Bahasa *Slang* pada Teks Ulasan

Teks Ulasan Asli (Raw Data)	Token *Slang* Terdeteksi	Hasil Normalisasi Teks Baku
"bgs bgt, bhn adem ga nyesel beli dsn"	bgs, bgt, bhn, ga, nyesel, dsn	bagus banget bahan adem tidak menyesal beli disini
"krmnya lma tpi gpp lah cucok meong"	krmnya, lma, tpi, gpp, cucok meong	kirimnya lama tetapi tidak apa apa lah cocok sekali
"zong bgt pesen wrna htm dtg pth kcewa"	zong, bgt, pesen, wrna, htm, dtg, pth, kcewa	buruk banget pesan warna hitam datang putih kecewa
"seler ramah pol, realpict bgt thx"	seler, pol, realpict, bgt, thx	penjual ramah sekali sesuai gambar banget terima kasih

Berdasarkan Tabel 3, proses normalisasi berhasil mengonversi kata tidak baku seperti “bgs”, “bgt”, dan “gpp” menjadi bentuk baku yang sesuai. Proses ini meningkatkan konsistensi representasi teks dengan menyatukan variasi penulisan ke dalam satu bentuk kata yang seragam. Secara metodologis, normalisasi bahasa slang berperan penting dalam mengurangi masalah *feature sparsity* akibat variasi kata yang tidak konsisten. Dengan penyatuan token ke dalam bentuk baku, frekuensi kemunculan kata menjadi lebih representatif, sehingga distribusi fitur pada

tahap pembobotan (TF-IDF) menjadi lebih stabil. Kondisi ini berdampak langsung pada peningkatan kualitas input model dan berpotensi meningkatkan kinerja algoritma klasifikasi.

Hasil Pembentukan Model dan Evaluasi

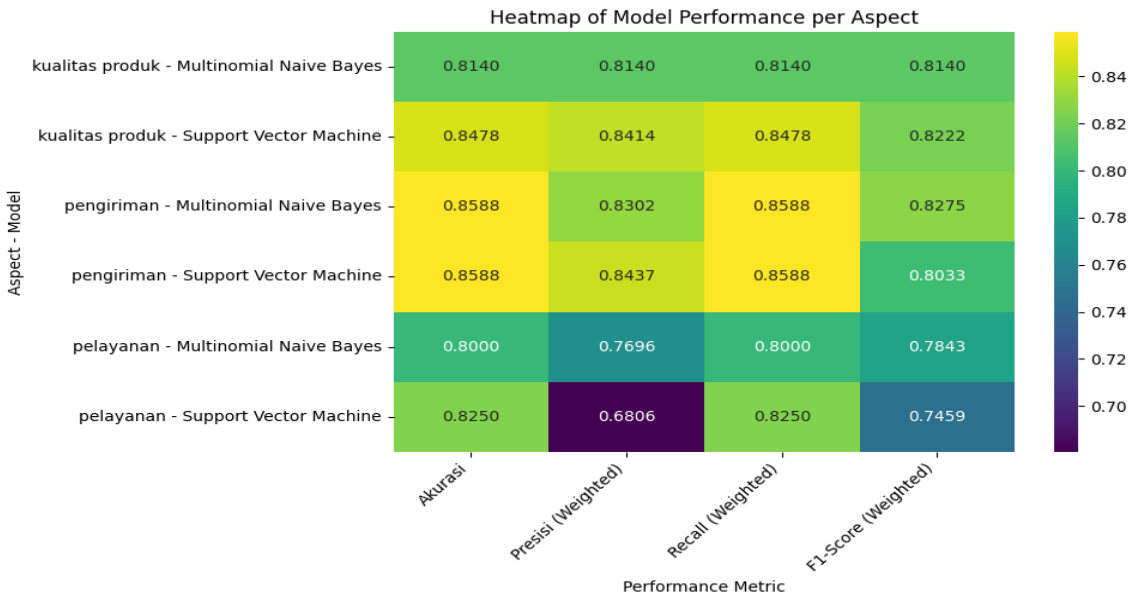
Pengujian model dilakukan menggunakan *Multinomial Naïve Bayes* (MNB) dan *Support Vector Machine* (SVM) dengan data yang telah melalui tahap TF-IDF dan SMOTE. Evaluasi model dilakukan menggunakan confusion matrix dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian ditunjukkan pada Tabel 4.

Tabel 4. Hasil Komparasi Pengujian Confusion Matrix

Skenario Pengujian Algoritma	Akurasi	Presisi	Recall	F1-Score
Skenario 1 (MNB Tanpa Normalisasi + Tanpa SMOTE)	72,2%	75,9%	72,2%	71,3%
Skenario 2 (MNB Dengan Normalisasi + SMOTE)	85,3%	86,6%	85,3%	85,2%
Skenario 3 (SVM Tanpa Normalisasi + Tanpa SMOTE)	85,2%	85,6%	85,2%	85%
Skenario 4 (SVM Dengan Normalisasi + SMOTE)	93,8%	93,9%	93,8%	93,7%

Hasil Performa Model Pada Masing-Masing Aspek

Selain evaluasi model secara keseluruhan, penelitian ini juga melakukan pengujian performa pada masing-masing aspek untuk melihat kemampuan model dalam mengklasifikasikan sentimen secara lebih spesifik. Evaluasi dilakukan pada aspek kualitas produk, pelayanan, dan pengiriman menggunakan metrik *accuracy, precision, recall dan F1-score.* Hasil pengujian performa pada masing-masing aspek dapat ditunjukkan pada Gambar 5 berikut.



Gambar 5. Heatmap Performa Model pada Masing-Masing Aspek

Berdasarkan Gambar 5, performa klasifikasi menunjukkan variasi pada masing-masing aspek. Pada aspek kualitas produk, algoritma SVM memperoleh performa yang lebih tinggi dibandingkan MNB pada hampir seluruh metrik evaluasi. Sementara itu, pada aspek pengiriman kedua model menunjukkan performa yang relatif seimbang dengan nilai akurasi tertinggi sebesar 85,8%. Pada aspek pelayanan, performa model cenderung lebih rendah dibandingkan aspek lainnya karena jumlah data yang lebih sedikit dan karakteristik ulasan yang lebih beragam. Secara umum, hasil ini menunjukkan bahwa distribusi data dan karakteristik tiap aspek memengaruhi performa klasifikasi model.

Pembahasan

Hasil pengujian menunjukkan bahwa normalisasi bahasa slang dan SMOTE meningkatkan performa model secara signifikan. Normalisasi slang mengurangi sparsity dengan menyatukan variasi kata ke dalam bentuk baku, sehingga probabilitas pada *Multinomial Naïve Bayes* (MNB) menjadi lebih stabil dan akurasinya meningkat dari 72,2% menjadi 85,3%. Lebih lanjut, penerapan teknik penyeimbangan kelas (SMOTE) memastikan model tidak bias terhadap kelas mayoritas. Pengaruh optimasi ini terlihat paling dominan pada algoritma *Support Vector Machine* (SVM), yang berhasil mencapai akurasi tertinggi sebesar 93,8%. Superioritas SVM dibandingkan MNB diduga terjadi karena SVM lebih adaptif dalam membentuk batas keputusan (*hyperplane*) yang optimal pada data teks yang representasi fiturnya (TF-IDF) telah bersih dan seimbang.

Meskipun menghasilkan performa yang tinggi, model masih menunjukkan beberapa kesalahan klasifikasi (*misclassification*) pada ulasan tertentu. Kesalahan tersebut umumnya terjadi pada ulasan yang mengandung lebih dari satu opini dalam satu kalimat serta konteks kalimat yang ambigu. Selain itu, beberapa ulasan memiliki kombinasi sentimen positif dan negatif secara bersamaan pada aspek yang berbeda sehingga model mengalami kesulitan dalam menentukan kelas sentimen yang dominan. Contoh kesalahan klasifikasi ditemukan pada ulasan seperti "barang bagus tapi pengiriman lama", di mana model cenderung mengklasifikasikan keseluruhan sebagai positif karena dominasi kata "bagus", meskipun terdapat sentimen negatif pada aspek pengiriman.

Di luar peningkatan metrik komputasi, model klasifikasi ini juga berhasil mengungkap wawasan bisnis yang krusial melalui pendekatan Analisis Sentimen Berbasis Aspek (ABSA). Berdasarkan hasil ekstraksi aspek, meskipun ulasan pada Kualitas Produk didominasi oleh sentimen positif, temuan yang patut disoroti terdapat pada aspek Pengiriman, di mana rasio sentimen negatif terdeteksi cukup tinggi. Hal ini membuktikan keunggulan metode ABSA dibandingkan analisis level dokumen; penjual dapat mengetahui secara spesifik bahwa ketidakpuasan pelanggan *Shopee* bukan berasal dari kualitas pakaian (*fashion*) yang dijual, melainkan bersumber dari kendala logistik atau keterlambatan kurir pengiriman. Hasil ini memberikan evaluasi bisnis yang lebih terarah dan dapat ditindaklanjuti secara langsung oleh pihak penjual.

4. KESIMPULAN

Penerapan normalisasi bahasa slang dan teknik SMOTE terbukti meningkatkan performa klasifikasi sentimen berbasis aspek pada dataset ulasan produk fashion Shopee yang digunakan dalam penelitian ini. Hasil pengujian menunjukkan bahwa kombinasi preprocessing tersebut memberikan peningkatan performa pada kedua algoritma, baik Multinomial Naïve Bayes (MNB) maupun Support Vector Machine (SVM). Model terbaik diperoleh oleh SVM dengan penerapan normalisasi slang dan SMOTE, yang mencapai nilai akurasi sebesar 93,8%, presisi 93,9%, recall 93,8%, dan F1-score 93,7%, sedangkan MNB pada skenario yang sama memperoleh akurasi sebesar 85,3%. Secara keseluruhan, SVM menunjukkan performa yang lebih unggul dibandingkan MNB dalam menangani data teks berdimensi tinggi dengan distribusi yang tidak seimbang. Selain itu, hasil penelitian ini menunjukkan bahwa tahap preprocessing, khususnya normalisasi bahasa tidak baku dan penyeimbangan data, memiliki peran penting dalam meningkatkan kualitas representasi fitur dan kinerja model klasifikasi sentimen berbasis aspek.

DAFTAR PUSTAKA

[1] Musfiroh, A. Tholib, and Z. Arifin, "Analisis Sentimen Terhadap Ulasan Aplikasi Shopee di Google Play Store Menggunakan Metode TF-IDF dan Long Short-Term Memory (LSTM)," J. Electr. Eng. Comput., vol. 6, no. 2, pp. 371–381, 2024, doi:

https://doi.org/10.33650/jeecom.v6i2.8713.
[2] K. Wau, "Application of Fine-Tuned IndoBERT for Sentiment Classification Local Product Reviews on Tokopedia Marketplace with Limited Dataset," J. Artif. Intell. Eng. Appl., vol. 5, no. 1, pp. 1377–1381, 2025, doi: https://doi.org/10.59934/jaiea.v5i1.1629.
[3] M. Hamka and Tukiran, "Analisis Sentimen Pengguna E-Commerce Dan Marketplace Menggunakan Support Vector Machine," J. Rekayasa Sist. Inf. dan Teknol., vol. 1, no. 4, pp. 273–282, 2024, doi: https://doi.org/10.59407/jrsit.v1i4.555.
[4] A. R. Putra and D. E. Ratnawati, "Analisis Sentimen Berbasis Aspek Pada Aplikasi Mobile Menggunakan Naïve Bayes Berdasarkan Ulasan Pengguna Playstore (Studi Kasus : Jconnect Mobile) Aspect-Based Sentiment Analysis On Mobile Applications Using Naïve Bayes Based On Playstore User Reviews," J. Teknol. Inf. dan Ilmu Komput., vol. 12, no. 2, pp. 293–300, 2025, doi: https://doi.org/10.25126/jtiik.2025127556.
[5] A. A. Asmiran, A. F. Nassa, A. Layinah, and Warto, "Sentiment Analysis of Shopee App User Reviews Based on Naïve Bayes Classifier," J. Surya Inform., vol. 15, no. 2, pp. 91–98, 2025, doi: https://doi.org/10.48144/suryainformatika.v15i2.2181.
[6] N. D. Ramadhanti et al., "Analisis Sentimen Ulasan Pengguna Aplikasi Shopee di Google Play Store Menggunakan Support Vector Machine dengan Teknik SMOTE," JIP (Jurnal Inform. Polinema), vol. 12, no. 2, pp. 313–322, 2026, doi: https://doi.org/10.33795/jip.v12i2.9073.
[7] M. F. O. Widi and D. Yuliawati, "Application of Multinomial Naïve Bayes for Sentiment Classification on Bukalapak Reviews," J. Appl. Informatics Comput., vol. 9, no. 6, pp. 3883–3891, 2025, doi: https://doi.org/10.30871/jaic.v9i6.11671.
[8] Y. A. Prasetyo, F. E. Nastiti, and Nurohman, "Penerapan Multinomial Naive Bayes Untuk Analisis Sentimen Aspek Pada Ulasan Mall Di Kota Solo," J. Inform. Teknol. dan Sains, vol. 7, no. 3, pp. 1093–1102, 2025, doi: https://doi.org/10.51401/jinteks.v7i3.6097.
[9] I. Maulana, W. Apriandari, and A. Pambudi, "Analisis Sentimen Berbasis Aspek Terhadap Ulasan Aplikasi MyPertamina Menggunakan SUPPORT VECTOR MACHINE," Idealis Indones. J. Inf. Syst., vol. 6, no. 2, pp. 172–181, 2023, doi: https://doi.org/10.36080/idealis.v6i2.3022.
[10] N. Agustina, D. H. Citra, and W. Purnama, "The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews on Google Play Store," MALCOM Indones. J. Mach. Learn. Comput. Sci., vol. 2, no. April, pp. 47–54, 2022, doi: https://doi.org/10.57152/malcom.v2i1.195.
[11] J. J. A. Limbong, I. Sembiring, and K. D. Hartomo, "Analisis Klasifikasi Sentimen Ulasan Pada E-Commerce Shopee Berbasis Word Cloud Dengan Metode Naive Bayes Dan K-Nearest Analysis Of Review Sentiment Classification On E-Commerce Shopee Word Cloud Based With Naïve Bayes And K-Nearest Neighbor Methods," J. Teknol. Inf. dan Ilmu Komput., vol. 9, no. 2, pp. 347–356, 2022, doi: https://doi.org/10.25126/jtiik.2022924960.
[12] A. Simanungkalit, J. Panda, P. Naibaho, and A. De Kweldju, "Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Shopee Menggunakan Algoritma Naïve Bayes," Jutisi J. Ilm. Tek. Inform. dan Sist. Inf., vol. 13, no. 1, pp. 659–670, 2021, doi: https://doi.org/10.35889/jutisi.v13i1.1826.
[13] B. A. Bhagaskara, A. Ismail, and H. Ed, "Analisis Sentimen Produk Pada Bukalapak Menggunakan Lexicon-Based dan Multinomial Naïve Bayes," J. Elektron. Ilmu Komput. Udayana, vol. 13, no. 4, pp. 925–936, 2025.
[14] R. Safitri, I. Ali, and N. Rahaningsih, "Analisis Sentimen Terhadap Tren Fashion Di Media Sosial Dengan Metode Support Vector Machine (SVM)," J. Mhs. Tek. Inform., vol. 8, no. 2, pp. 1746–1754, 2024, doi: https://doi.org/10.36040/jati.v8i2.9045.
[15] R. T. Nagur, D. Alamsyah, M. G. Tauran, Y. Nuryamin, and A. Priyatna, "Perbandingan Kinerja Naive Bayes dan Decision Tree untuk Analisis Sentimen Ulasan Produk Online di Shopee dan Tokopedia," JSI (Jurnal Sist. Informasi) Univ. Dirgant. Marsekal Suryadarma, vol. 13, no. 1, pp. 124–132, 2026, doi: https://doi.org/10.35968/jsi.v13i1.1740.
[16] N. F. Putri, M. F. Hidayattullah, and D. I. Af'idah, "Sentimen Analisis Kota Tegal Berbasis

Aspek Menggunakan Algoritma Naïve Bayes," INFOMATEK J. Inform. Manaj. dan Teknol., vol. 26, pp. 45–54, 2024, doi: https://doi.org/10.23969/infomatek.v26i1.11209.
[17] A. S. R. Rufaida, A. E. Permanasari, and N. A. Setiawan, "Lexicon-Based Sentiment Analysis Using Inset Dictionary: A Systematic Literature Review," in Proceedings of the 5th International Conference on Applied Engineering, EAI, 2023. doi: http://dx.doi.org/10.4108/eai.5-10-2022.2327474.
[18] M. Ojeda-Hernández, D. López-Rodríguez, and Á. Mora, "Lexicon-based sentiment analysis in texts using Formal Concept Analysis," Int. J. Approx. Reason., vol. 155, pp. 104–112, 2023, doi: https://doi.org/10.1016/j.ijar.2023.02.001.