

Sentiment Classification Analysis of Tokopedia Reviews Using TF-IDF, SMOTE, and Traditional Machine Learning Models

Herianta Barus ^{1*}, Ika Nur Fajri ^{2*}, Yoga Pristyanto ^{3*}

* Sistem Informasi, Universitas Amikom Yogyakarta

heriantabarus@students.amikom.ac.id ¹, fajri@amikom.ac.id ², yoga.pristyanto@amikom.ac.id ³

Article Info

Article history:

Received 2025-07-28

Revised 2025-09-03

Accepted 2025-09-10

Keyword:

Sentiment Analysis, Tokopedia, E-commerce, Naïve Bayes, Random Forest, Support Vector Machine (SVM), Logistic Regression, Decision Tree, TF-IDF, SMOTE.

ABSTRACT

This study explores sentiment classification on Tokopedia user reviews using TF-IDF for feature extraction and SMOTE to handle class imbalance. From nearly one million raw reviews sourced from Kaggle ("E-Commerce Ratings and Reviews in Bahasa Indonesia"), a final set of 6,477 relevant entries was obtained after rigorous preprocessing, including case folding, noise removal (emojis, URLs, numbers), normalization to KBBI standards, tokenization, stopword removal, and stemming with Sastrawi. The dataset consisted of 5,213 positive and 1,264 negative reviews (80.4% positive). SMOTE balanced the classes to 10,426 reviews with a 1:1 ratio for training. Five traditional machine learning models were evaluated: Naive Bayes, Logistic Regression, Support Vector Machine (SVM), Decision Tree, and Random Forest. Assessments were based on accuracy, precision, recall, F1-score, ROC-AUC, and computational time, using an 80:20 stratified split and 5-fold cross-validation. Random Forest achieved the best overall performance (accuracy: 0.9163, F1-score: 0.9133, ROC-AUC: 0.9784), while tuned SVM (C=10, RBF kernel) attained the highest accuracy of 0.9473 and F1-score of 0.9321. Cross-validation on Naive Bayes showed consistent results with an average accuracy of 88.09%. Further analysis using Logistic Regression coefficients identified influential features: positive sentiment associated with words like "mantap", "mudah", and "sukses", while negative sentiment correlated with "kecewa", "parah", and "lemot". These insights provide practical value for Tokopedia's teams to enhance user experience, such as improving app speed and addressing complaints. The findings demonstrate the effectiveness and efficiency of traditional machine learning techniques for sentiment analysis in Bahasa Indonesia contexts.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Pertumbuhan pesat e-commerce di Indonesia telah mengubah pola konsumsi masyarakat. Marketplace seperti Tokopedia, Shopee, Lazada, Blibli, dan Bukalapak menjadi platform utama untuk transaksi digital [1], [2]. Tokopedia, dengan 157 juta pengunjung bulanan pada 2021 [3], menghadapi tantangan dalam menjaga kepuasan pelanggan. Ulasan pengguna, yang mencerminkan pengalaman, kepuasan, dan keluhan, merupakan sumber penting untuk memahami persepsi pelanggan [4]. Analisis sentimen terhadap ulasan ini penting untuk meningkatkan kualitas layanan dan menyusun strategi bisnis yang akurat [5]. Penelitian sebelumnya banyak menggunakan metode

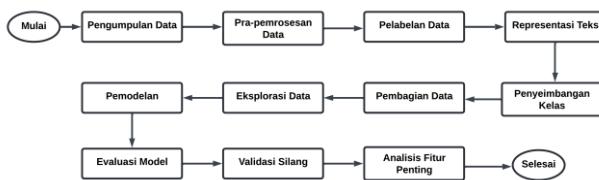
machine learning, seperti Naive Bayes dan SVM, untuk analisis sentimen e-commerce [6], [7]. Namun, pendekatan ini sering mengabaikan ketidakseimbangan kelas, di mana ulasan positif mendominasi (80-90%) dibandingkan ulasan negatif, menyebabkan bias klasifikasi dan akurasi rendah pada kelas minoritas (F1-score < 70% pada kelas negatif dalam [8]). Selain itu, kompleksitas morfologi Bahasa Indonesia, seperti slang (misalnya, "gak" untuk "tidak"), idiom lokal (misalnya, "mantap"), dan struktur kalimat non-standar, mempersulit representasi teks tanpa preprocessing yang memadai, seperti normalisasi dan stemming [9]. Banyak studi juga jarang mengintegrasikan teknik penyeimbangan data seperti SMOTE, yang dapat meningkatkan performa

pada dataset tidak seimbang [10]. Algoritma machine learning tradisional, seperti Naive Bayes, Logistic Regression, Support Vector Machine (SVM), Decision Tree, dan Random Forest, menawarkan efisiensi komputasi dan interpretabilitas tinggi [11]. Model ini efektif untuk dataset menengah (5.000-10.000 ulasan) dan memungkinkan analisis fitur penting, seperti koefisien Logistic Regression yang mengidentifikasi kata-kata seperti "mantap" (positif) atau "lemot" (negatif) untuk strategi bisnis Tokopedia, misalnya, mempercepat aplikasi atau meningkatkan layanan pelanggan [6]. Kombinasi TF-IDF dan SMOTE jarang dieksplorasi untuk ulasan e-commerce Indonesia, menjadikan pendekatan ini relevan untuk konteks lokal dengan sumber daya terbatas [12][13]. Penelitian ini mengisi kesenjangan dengan: (1) menerapkan TF-IDF untuk representasi teks yang menangani kompleksitas morfologi Bahasa Indonesia, (2) menggunakan SMOTE untuk mengatasi ketidakseimbangan kelas, (3) membandingkan kinerja lima algoritma ML tradisional secara komprehensif, dan (4) menyediakan interpretasi fitur untuk mendukung strategi bisnis Tokopedia, seperti peningkatan kecepatan aplikasi berdasarkan keluhan "lemot". Dengan demikian, penelitian ini menegaskan efektivitas ML tradisional dalam analisis sentimen e-commerce di Indonesia, khususnya untuk data tidak seimbang.

II. METODE

Bab ini menguraikan pendekatan metodologis yang digunakan dalam penelitian analisis sentimen terhadap ulasan pengguna aplikasi Tokopedia. Metodologi penelitian disusun berdasarkan praktik terbaik dalam bidang analisis teks dan Natural Language Processing (NLP), yang mencakup serangkaian tahapan mulai dari pengumpulan data, pra-pemrosesan, pelabelan, representasi teks, penyimbangan kelas, hingga pemodelan dan evaluasi. Setiap tahapan didukung dengan teknik yang relevan, seperti case folding, normalisasi, tokenisasi, stopword removal, stemming, representasi teks dengan TF-IDF, penyeimbangan data dengan SMOTE, serta validasi model menggunakan k-fold cross validation. Selain itu, pemilihan algoritma machine learning tradisional (Naïve Bayes, Logistic Regression, Support Vector Machine, Decision Tree, dan Random Forest) didasarkan pada efisiensi komputasi dan interpretabilitas yang lebih tinggi dibandingkan model berbasis deep learning.

Alur metodologi penelitian ini divisualisasikan pada Gambar 1, yang menggambarkan hubungan antar tahapan secara sistematis mulai dari input data mentah hingga analisis fitur penting untuk menghasilkan insight bisnis yang relevan.



Gambar 1. gambar tahapan penelitian

A. Pengumpulan Data

Pengumpulan data dilakukan dengan memanfaatkan dataset publik yang tersedia di platform Kaggle. Dataset tersebut memuat ulasan pengguna aplikasi Tokopedia dalam bentuk teks beserta skor penilaian atau rating. Jumlah entri pada dataset ini mencapai sekitar satu juta baris, yang kemudian digunakan sebagai sumber utama dalam proses pemodelan dan analisis. Pemilihan Kaggle sebagai sumber data didasarkan pada kredibilitas serta kemudahan akses terhadap data yang telah terstruktur dan relevan dengan topik penelitian.

B. Pra-Pemrosesan Data

Pemrosesan data dilakukan untuk memastikan ulasan Tokopedia dalam kondisi bersih, konsisten, dan siap digunakan pada tahap representasi fitur serta pemodelan. Rangkaian langkah ini bertujuan mengurangi noise, menyamakan format teks, serta menyesuaikan perbedaan morfologi Bahasa Indonesia. Dengan adanya pemrosesan data, kualitas input ke algoritma machine learning dapat ditingkatkan sehingga hasil klasifikasi lebih akurat [2], [3].

1) Data Cleaning

Tahap awal adalah pembersihan data dari elemen yang tidak relevan. Proses ini meliputi penghapusan data duplikat, entri kosong (NaN), serta ulasan dengan skor netral (3). Selain itu dilakukan case folding untuk menyeragamkan seluruh teks ke huruf kecil. Elemen non-teks seperti angka, URL, tanda baca, emoji, dan simbol khusus dihapus menggunakan regular expressions (regex). Dengan demikian, teks menjadi lebih sederhana dan siap diproses pada tahap berikutnya [2], [3].

2) Normalisasi

Normalisasi dilakukan untuk menyamakan kata tidak baku atau slang menjadi kata baku sesuai Kamus Besar Bahasa Indonesia (KBBI). Contoh: kata "gak" atau "nggak" diubah menjadi "tidak", sedangkan "btw" diubah menjadi "ngomong-ngomong". Tahap ini penting karena ulasan pengguna e-commerce sering menggunakan bahasa sehari-hari yang beragam. Penelitian sebelumnya menunjukkan bahwa normalisasi dapat meningkatkan konsistensi data ulasan dan mempermudah proses representasi teks [4], [6].

3) Tokenisasi

Tokenisasi adalah proses memecah teks ulasan menjadi unit kata (token). Dengan tokenisasi, setiap kata dapat diperlakukan sebagai fitur individual dalam model. Pada penelitian ini digunakan pustaka Sastrawi untuk membagi kalimat ke dalam token. Contoh: kalimat "barang cepat sampai" diubah menjadi [barang, cepat, sampai]. Tokenisasi merupakan tahap penting dalam membangun vektor kata untuk representasi numerik, dan telah digunakan pada penelitian serupa di domain e-commerce [2], [7].

4) Stopword Removal

Stopword removal digunakan untuk menghapus kata-kata umum yang tidak memiliki kontribusi signifikan terhadap penentuan sentimen. Kata seperti "dan", "yang", atau "di" dikategorikan sebagai stopword dan dihapus dari dataset. Penghapusan stopword membantu mengurangi noise,

sehingga kata-kata yang tersisa lebih relevan terhadap analisis sentimen. Studi-studi sebelumnya juga menegaskan bahwa tahapan ini penting untuk meningkatkan kualitas representasi teks [2], [5].

5) Stemming

Stemming bertujuan mengubah kata berimbuhan ke bentuk dasar menggunakan algoritma stemming Sastrawi. Contoh: "berbelanja" → "belanja", "membeli" → "beli". Dengan stemming, variasi kata turunan dapat disatukan sehingga model tidak perlu menganggapnya sebagai entitas yang berbeda. Penelitian terdahulu menunjukkan bahwa stemming efektif dalam mengurangi kompleksitas morfologis Bahasa Indonesia dan meningkatkan akurasi representasi fitur [3], [7].

Hasil dari seluruh tahapan pra-pemrosesan ini adalah teks yang bersih, konsisten, dan sudah dalam bentuk kata dasar. Data yang telah diproses kemudian direpresentasikan menggunakan metode TF-IDF (Term Frequency–Inverse Document Frequency) sebelum masuk ke tahap pembelajaran mesin. Tahapan pra-pemrosesan serupa (cleaning → normalisasi → tokenisasi → stopword → stemming/Sastrawi → TF-IDF) telah terbukti efektif dalam analisis sentimen e-commerce berbahasa Indonesia [2], [4].

C. Pelabelan Data

Pelabelan data dilakukan berdasarkan nilai rating yang diberikan pengguna pada ulasan Tokopedia. Dalam penelitian ini, ulasan dengan rating ≥ 4 diberi label positif (1), sedangkan ulasan dengan rating ≤ 2 diberi label negatif (0). Sementara itu, ulasan dengan skor 3 dianggap netral dan dikeluarkan dari dataset karena cenderung mengandung sentimen campuran yang dapat menimbulkan ambiguitas dalam klasifikasi. Skema ini dipilih untuk menyederhanakan permasalahan klasifikasi menjadi dua kelas utama, yaitu positif dan negatif. Pendekatan serupa juga telah diterapkan pada penelitian terdahulu di domain e-commerce maupun ulasan aplikasi Google Play, di mana data netral diabaikan untuk meningkatkan akurasi model dan memudahkan interpretasi hasil [2], [4], [7].

Dengan demikian, dataset akhir yang digunakan dalam penelitian ini lebih fokus pada dua kelas utama sehingga mendukung proses analisis sentimen yang lebih optimal.

D. Representasi Teks dan Penyeimbangan Kelas

Setelah melalui tahap pra-pemrosesan, data ulasan direpresentasikan ke dalam bentuk numerik menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF). Metode ini menghitung bobot setiap kata berdasarkan frekuensi kemunculannya dalam sebuah dokumen serta seberapa jarang kata tersebut muncul di seluruh koleksi dokumen. Dengan demikian, kata yang sering muncul dalam sebuah ulasan namun jarang muncul di ulasan lain akan memiliki bobot yang lebih tinggi. Secara matematis, bobot TF-IDF dirumuskan sebagai berikut:

$$w_{t,d} = tf_{t,d} \times \log \left(\frac{N}{df_t} \right)$$

dengan $w_{t,d}$ adalah bobot kata t pada dokumen d , $tf_{t,d}$ adalah frekuensi kemunculan kata t dalam dokumen d , N adalah jumlah total dokumen, dan df_t adalah jumlah dokumen yang mengandung kata t . Representasi berbasis TF-IDF dipilih karena mampu menangkap informasi penting dari kata tanpa dipengaruhi oleh kata umum yang sering muncul, sehingga relevan untuk analisis sentimen berbahasa Indonesia [2], [4]. Selain representasi fitur, penelitian ini juga memperhatikan permasalahan ketidakseimbangan kelas yang umum terjadi pada data ulasan e-commerce. Sebagian besar ulasan cenderung positif, sedangkan ulasan negatif berjumlah lebih sedikit. Ketidakseimbangan ini dapat menyebabkan bias model terhadap kelas mayoritas. Untuk mengatasi hal tersebut digunakan metode *Synthetic Minority Oversampling Technique* (SMOTE) yang menghasilkan sampel sintetis pada kelas minoritas dengan cara interpolasi terhadap tetangga terdekat. Rumus SMOTE dituliskan sebagai berikut:

$$x_{new} = x_i + \lambda \times (x_{nn} - x_i)$$

dengan x_i adalah sample asli dari kelas minoritas, x_{nn} adalah salah satu tetangga terdekatnya, dan λ adalah bilangan acak dalam interval $[0,1]$. Dengan cara ini, data minoritas diperluas tanpa harus mengurangi jumlah data mayoritas, sehingga distribusi kelas menjadi lebih seimbang.

Pendekatan TF-IDF + SMOTE ini dipilih karena masih jarang dieksplorasi dalam penelitian analisis sentimen berbahasa Indonesia, terutama pada domain ulasan e-commerce. Sebagian besar penelitian terdahulu berfokus pada representasi berbasis embedding atau hanya menggunakan TF-IDF tanpa menangani masalah imbalance [2], [4], [7]. Dengan menggabungkan TF-IDF untuk ekstraksi fitur teks dan SMOTE untuk penyeimbangan data, penelitian ini menegaskan kebaruan dalam mengeksplorasi efektivitas metode machine learning tradisional pada dataset ulasan e-commerce yang kompleks dan tidak seimbang.

E. Pembagian Data

Dataset yang telah melalui tahap representasi fitur dengan TF-IDF dan penyeimbangan kelas menggunakan SMOTE selanjutnya dibagi menjadi data latih (training set) dan data uji (testing set). Proses pembagian dilakukan dengan rasio 80:20 menggunakan metode stratifikasi untuk memastikan distribusi kelas positif dan negatif tetap proporsional pada kedua subset data. Dengan teknik ini, baik data latih maupun data uji memiliki representasi kelas yang seimbang, sehingga mengurangi risiko bias akibat distribusi yang tidak merata. Data latih digunakan untuk membangun model klasifikasi dengan lima algoritma machine learning tradisional, sedangkan data uji dipakai untuk mengevaluasi performa model pada data yang belum pernah dilihat sebelumnya. Pembagian data secara stratifikasi 80:20 ini merupakan praktik umum dalam analisis sentimen e-commerce,

sebagaimana juga diterapkan pada penelitian terdahulu [2], [4].

Selain pembagian data, penelitian ini juga menerapkan validasi silang (cross-validation) untuk memperkuat hasil evaluasi. Validasi silang dengan skema 5-fold memastikan setiap data berkesempatan menjadi data uji sekaligus data latih pada iterasi yang berbeda, sehingga performa model tidak hanya bergantung pada satu kali pembagian data. Kombinasi antara split stratifikasi 80:20 dan validasi silang 5-fold memberikan evaluasi yang lebih komprehensif terhadap generalisasi model [3], [6].

F. Eksplorasi Data

Eksplorasi data dilakukan untuk memahami karakteristik dataset sebelum masuk ke tahap pemodelan. Analisis ini mencakup distribusi jumlah ulasan berdasarkan label sentimen positif dan negatif, panjang rata-rata teks ulasan, serta identifikasi kata-kata yang paling sering muncul. Distribusi label dianalisis kembali setelah penerapan SMOTE untuk memastikan keseimbangan antara kelas positif dan negatif. Selain itu, dilakukan visualisasi kata menggunakan word cloud guna memperlihatkan kata-kata yang dominan pada masing-masing kelas. Hasil eksplorasi menunjukkan bahwa ulasan positif cenderung didominasi kata seperti “mantap”, “mudah”, dan “cepat”, sedangkan ulasan negatif lebih sering memuat kata “kecewa”, “lemot”, dan “parah”. Informasi ini membantu memberikan gambaran awal mengenai kecenderungan sentimen pada ulasan pengguna Tokopedia sekaligus memvalidasi relevansi data dengan tujuan penelitian.

G. Pemodelan dan Evaluasi

Tahap pemodelan dilakukan dengan membandingkan lima algoritma machine learning tradisional, yaitu Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), Decision Tree, dan Random Forest. Pemilihan kelima algoritma ini didasarkan pada efektivitasnya dalam analisis sentimen teks berbahasa Indonesia, efisiensi komputasi, serta kemampuan interpretasi hasil klasifikasi yang lebih mudah dibandingkan dengan model deep learning [1], [2], [7].

1) Logistic Regression

Logistic Regression adalah algoritma klasifikasi biner yang menggunakan fungsi logit (sigmoid) untuk memetakan keluaran ke rentang 0–1. Persamaan dasarnya:

$$P(y = 1 | x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

Logistic Regression tidak hanya efektif dalam klasifikasi, tetapi juga memungkinkan analisis koefisien untuk mengidentifikasi kata yang berpengaruh terhadap prediksi sentimen. Studi terdahulu menunjukkan algoritma ini banyak digunakan dalam analisis sentimen e-commerce untuk mengungkap kata kunci yang relevan dengan pengalaman pengguna [1], [4], [14], [15].

2) Decision Tree

Decision Tree membangun struktur pohon keputusan dengan memecah data berdasarkan fitur yang memberikan pemisahan terbaik pada setiap node. Salah satu ukuran yang digunakan adalah Gini Index, yang dirumuskan sebagai:

$$Gini(t) = 1 - \sum_{i=1}^c p(i | t)^2$$

dengan $p(i | t)^2$ adalah probabilitas kelas i pada node t . Decision Tree mudah diinterpretasikan karena menghasilkan aturan klasifikasi yang sederhana, namun rentan terhadap *overfitting*. Penelitian sebelumnya menunjukkan algoritma ini sering dipakai sebagai baseline dalam klasifikasi sentimen [6][12].

3) Support Vector Machine (SVM)

Support Vector Machine bekerja dengan mencari *hyperplane* terbaik yang dapat memisahkan data dari dua kelas dengan margin maksimum. Fungsi kernel digunakan untuk memetakan data non-linear ke ruang berdimensi lebih tinggi. Model ini dioptimasi dengan Grid Search untuk menentukan kombinasi hiperparameter terbaik. Parameter yang diuji meliputi $C = \{1, 5, 10, 100\}$, $\text{kernel} = \{\text{linear}, \text{polynomial}, \text{rbf}\}$, dan gamma dengan nilai default “scale”. Kombinasi optimal diperoleh pada $C = 10$ dengan kernel RBF, yang menghasilkan akurasi 94,73% dan F1-score 93,21%, dengan rata-rata waktu komputasi sekitar 16,3 detik. SVM juga terbukti efektif pada berbagai penelitian analisis sentimen e-commerce di Indonesia [2], [9], [11].

4) Naïve Bayes

Naïve Bayes merupakan algoritma probabilistik yang sederhana namun efektif untuk klasifikasi teks. Prinsip utamanya adalah menerapkan Teorema Bayes dengan asumsi independensi antar fitur. Persamaan umum Naïve Bayes adalah:

$$P(c | d) = \frac{P(c) \prod_i P(w_i | c)}{P(d)}$$

Algoritma ini dipilih karena mampu memberikan hasil yang konsisten dengan waktu komputasi sangat cepat, meskipun performanya bisa menurun bila asumsi independensi antar fitur tidak terpenuhi. Beberapa penelitian e-commerce menunjukkan Naïve Bayes mampu mencapai akurasi baik pada dataset berukuran menengah [2], [7], [8].

5) Random Forest

Random Forest adalah metode ansambel yang terdiri dari sejumlah Decision Tree yang dibangun secara acak. Setiap pohon dilatih dengan subset data dan subset fitur yang berbeda melalui teknik *bagging*. Hasil klasifikasi diperoleh dari voting mayoritas semua pohon. Dengan cara ini, Random Forest mampu mengurangi risiko *overfitting* yang sering terjadi pada Decision Tree tunggal dan menghasilkan prediksi yang lebih stabil. Sejumlah penelitian e-commerce di Indonesia melaporkan bahwa Random Forest memberikan performa lebih tinggi dibandingkan model tunggal [5], [10], [16].

Evaluasi kelima model dilakukan menggunakan data uji hasil pembagian stratifikasi 80:20 serta validasi silang 5-fold untuk memastikan stabilitas dan generalisasi model. Metrik

yang digunakan meliputi akurasi, presisi, recall, F1-score, ROC-AUC, serta waktu komputasi. Penggunaan berbagai metrik evaluasi memberikan gambaran yang lebih komprehensif mengenai performa model. Analisis waktu komputasi juga ditambahkan untuk menegaskan efisiensi model, yang merupakan salah satu alasan utama pemilihan algoritma machine learning tradisional pada penelitian ini [1], [2], [9].

H. Validasi Silang

Validasi model dilakukan dengan menggunakan metode *stratified split* 80:20, di mana 80% data digunakan sebagai data latih dan 20% data digunakan sebagai data uji. Untuk memastikan hasil evaluasi tidak bias terhadap distribusi data tertentu, seluruh model juga divalidasi menggunakan teknik *k-fold cross-validation* dengan nilai $k = 5$. Teknik ini membagi data latih menjadi lima lipatan (*fold*) yang seimbang, kemudian secara bergantian empat lipatan digunakan untuk pelatihan dan satu lipatan untuk pengujian. Hasil evaluasi diperoleh dari rata-rata performa kelima lipatan tersebut.

Penerapan validasi silang ini memastikan model tidak hanya menyesuaikan diri dengan subset data tertentu (*overfitting*), tetapi mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat. Pada penelitian ini, validasi silang diterapkan pada semua algoritma tradisional yang diuji, yaitu Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), Decision Tree, dan Random Forest. Pendekatan validasi silang juga banyak digunakan dalam penelitian analisis sentimen e-commerce berbahasa Indonesia untuk menjaga konsistensi hasil, misalnya pada studi oleh Ramadhan [11] dan Hasanah [6].

I. Analisis Fitur Penting

Selain mengevaluasi performa model berdasarkan metrik akurasi, presisi, recall, F1-score, dan ROC-AUC, penelitian ini juga melakukan analisis terhadap fitur penting (feature importance) untuk memperoleh wawasan praktis dari ulasan pengguna. Analisis ini dilakukan menggunakan koefisien dari model Logistic Regression, yang dapat menunjukkan kontribusi setiap kata terhadap prediksi sentimen. Nilai koefisien positif mengindikasikan kata tersebut lebih sering muncul pada ulasan dengan sentimen positif, sedangkan nilai koefisien negatif menunjukkan keterkaitan dengan sentimen negatif.

Dengan pendekatan ini, penelitian tidak hanya menghasilkan model klasifikasi yang andal, tetapi juga mengidentifikasi kata-kata yang berperan penting dalam persepsi pelanggan terhadap Tokopedia. Sebagai contoh, kata “mantap”, “mudah”, dan “sukses” berkontribusi pada sentimen positif, sedangkan kata “kecewa”, “parah”, dan “lemot” lebih dominan pada sentimen negatif. Informasi ini dapat dimanfaatkan oleh pihak Tokopedia untuk memperbaiki kualitas layanan, misalnya dengan meningkatkan kecepatan aplikasi atau merespons keluhan yang paling sering muncul.

Pendekatan serupa juga telah digunakan pada penelitian analisis sentimen e-commerce di Indonesia, di mana pemetaan kata kunci terbukti membantu menghubungkan hasil klasifikasi dengan strategi bisnis yang lebih terarah [3], [6], [8].

J. Hyperparameter Tuning SVM

Untuk memperoleh performa terbaik pada algoritma *Support Vector Machine* (SVM), dilakukan pencarian hiperparameter menggunakan metode *Grid Search*. Parameter yang diuji meliputi nilai $C = \{0.1, 1, 10, 100\}$, $\text{kernel} = \{\text{linear}, \text{polynomial}, \text{rbf}\}$, serta fungsi gamma dengan nilai default scale . Proses ini dilakukan untuk menemukan kombinasi parameter yang menghasilkan akurasi dan F1-score paling tinggi pada data uji.

Hasil tuning menunjukkan bahwa kombinasi terbaik diperoleh pada kernel RBF dengan $C = 10$ dan $\text{gamma} = \text{scale}$. Kombinasi ini meningkatkan akurasi SVM secara signifikan dari baseline 0.8849 menjadi 0.9473, dengan F1-score 0.9321. Proses tuning dilakukan dengan validasi silang 5-fold untuk memastikan hasil yang diperoleh stabil dan tidak bias terhadap subset data tertentu.

Pendekatan tuning hiperparameter dengan *Grid Search* telah banyak digunakan pada penelitian analisis sentimen berbasis SVM, karena mampu meningkatkan kinerja klasifikasi dibandingkan penggunaan parameter default [6], [9], [10].

III. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi analisis sentimen terhadap ulasan pengguna Tokopedia dengan menggunakan algoritma machine learning tradisional berbasis representasi TF-IDF dan penyeimbangan data SMOTE. Pembahasan difokuskan pada dua aspek, yaitu evaluasi kuantitatif melalui metrik kinerja model (akurasi, presisi, recall, F1-score, ROC-AUC, serta waktu komputasi) dan validasi silang 5-fold, serta evaluasi kualitatif melalui analisis fitur penting, visualisasi word cloud, distribusi kata dominan, dan interpretasi kesalahan klasifikasi. Hasil penelitian dijelaskan secara terstruktur, dimulai dari deskripsi dan eksplorasi dataset, representasi teks menggunakan TF-IDF, proses penyeimbangan kelas dengan SMOTE, hingga pelatihan dan evaluasi lima algoritma klasifikasi: Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), Decision Tree, dan Random Forest. Selanjutnya, ditampilkan analisis perbandingan kinerja model, validasi silang, serta interpretasi fitur dan kesalahan prediksi. Dengan struktur ini, Tidak hanya menekankan performa numerik antar model, tetapi juga memberikan wawasan praktis mengenai persepsi pengguna yang dapat dimanfaatkan Tokopedia untuk meningkatkan kualitas layanan. Selain itu, temuan penelitian ini memperkuat relevansi metode tradisional dalam analisis sentimen berbahasa Indonesia, terutama pada dataset dengan ketidakseimbangan kelas.

A. Deskripsi Data dan Distribusi Label

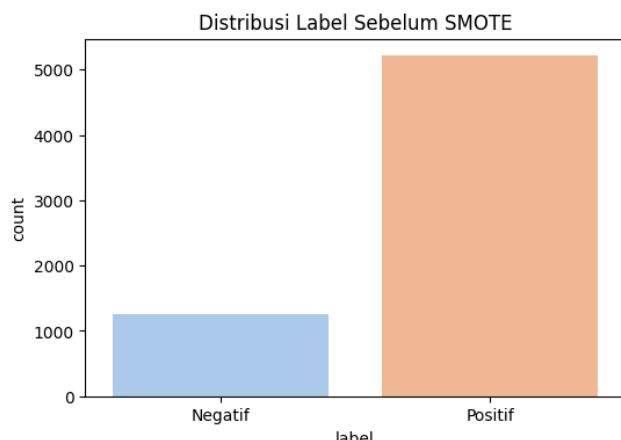
Dataset yang digunakan dalam penelitian ini berasal dari Kaggle dengan judul “E-Commerce Ratings and Reviews in Bahasa Indonesia”. Dataset awal berjumlah hampir satu juta ulasan pengguna Tokopedia. Setelah dilakukan preprocessing (cleaning, normalisasi, tokenisasi, stopword removal, dan stemming), jumlah ulasan yang relevan diperoleh sebanyak 6.477, terdiri dari 5.213 ulasan positif (80,5%) dan 1.264 ulasan negatif (19,5%). Ulasan dengan skor netral (rating = 3) dikeluarkan karena berpotensi menimbulkan ambiguitas.

Distribusi data menunjukkan adanya ketidakseimbangan kelas, di mana ulasan positif mendominasi. Untuk mengatasi hal ini, digunakan metode *Synthetic Minority Oversampling Technique* (SMOTE). Setelah dilakukan penyeimbangan, diperoleh 10.426 ulasan dengan distribusi seimbang (50% positif dan 50% negatif).

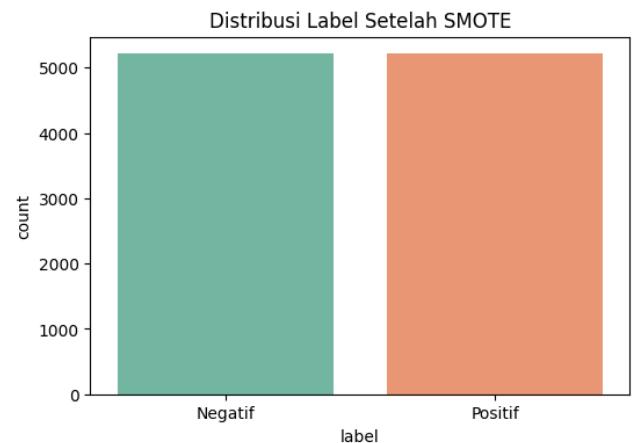
Distribusi dataset sebelum dan sesudah SMOTE dapat dilihat pada Tabel 1, sedangkan perbandingan visualnya ditampilkan pada Gambar 2 (sebelum SMOTE) dan Gambar 3 (sesudah SMOTE).

TABEL I
DISTRIBUSI DATASET SEBELUM DAN SESUDAH SMOTE

Kelas Sentimen	Jumlah (sebelum SMOTE)	Proporsi (%) (sebelum SMOTE)	Jumlah (sesudah SMOTE)	Proporsi (%) (sesudah SMOTE)
Positif	5213	80,5%	5213	50,0%
Negatif	1264	19,5%	5213	50,0%
Total	6477	100%	10426	100%



Gambar 2. gambar Data sebelum terapkan SMOTE



Gambar 3. gambar Data setelah terapkan SMOTE

B. Eksplorasi Data

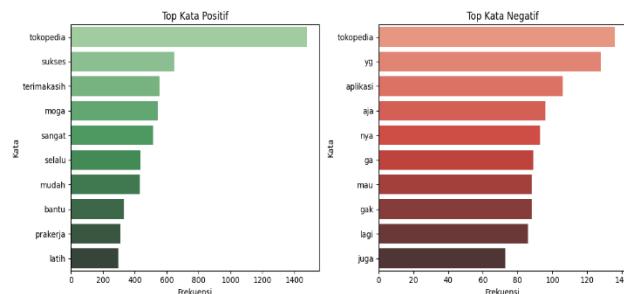
Eksplorasi data dilakukan untuk memahami karakteristik ulasan Tokopedia sebelum tahap pemodelan. Analisis distribusi panjang ulasan menunjukkan bahwa sebagian besar ulasan berada pada rentang 10–25 kata, yang menunjukkan bahwa mayoritas pengguna menuliskan ulasan singkat namun padat informasi.

1 Visualisasi word cloud digunakan untuk menggambarkan kata-kata yang paling dominan pada masing-masing kelas sentimen. Pada ulasan positif, kata-kata yang menonjol antara lain “mantap”, “mudah”, dan “sukses”. Sementara itu, ulasan negatif banyak didominasi kata “kecewa”, “lemot”, dan “parah”. Perbedaan kosakata ini mengindikasikan adanya kecenderungan ekspresi yang jelas antara ulasan positif dan negatif.

Selain itu, analisis frekuensi kata memperlihatkan 10 kata teratas pada masing-masing kelas. Pada ulasan positif, kata-kata seperti “tokopedia”, “sukses”, dan “terimakasih” memiliki frekuensi paling tinggi. Sebaliknya, ulasan negatif lebih sering memuat kata seperti “gagal”, “lemot”, dan “tidak”. Analisis ini memberikan gambaran awal bahwa kosakata yang digunakan konsisten dengan sentimen yang dinyatakan.



Gambar 4. Wordcloud Sentimen Positif dan Negatif



Gambar 5. Top 10 Kata pada Ulasan Positif dan Negatif

C. Hasil Pemodelan

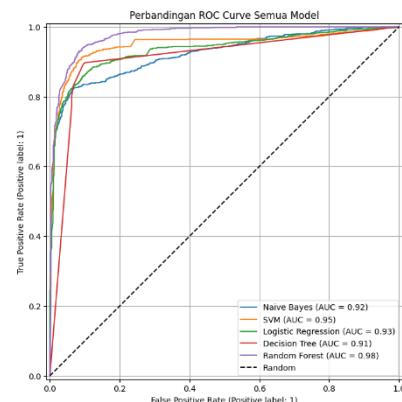
Lima algoritma machine learning tradisional diuji pada dataset ulasan Tokopedia, yaitu Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), Decision Tree, dan Random Forest. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, ROC-AUC, serta waktu komputasi. Hasil pengujian menunjukkan bahwa SVM dengan kernel RBF ($C=10$) memberikan akurasi tertinggi sebesar 94,73%, dengan nilai presisi 0,9578 dan F1-score 0,8750. Random Forest menempati posisi kedua dengan akurasi 91,28% dan ROC-AUC 0,9759, yang menegaskan kemampuannya dalam membedakan sentimen positif dan negatif. Sementara itu, Decision Tree memiliki performa yang cukup seimbang dengan akurasi 89,84% dan F1-score 0,8977. Model yang lebih sederhana, seperti Naïve Bayes dan Logistic Regression, tetap relevan sebagai baseline karena memiliki waktu eksekusi sangat cepat meskipun akurasinya lebih rendah dibandingkan SVM dan Random Forest.

TABEL 2

PERBANDINGAN KINERJA MODEL MACHINE LEARNING TRADISIONAL

Model	Akurasi	Precision	Recall	F1-Score	ROC-AUC	Waktu (s)
Naïve Bayes	0.8749	0.9335	0.8073	0.8658	0.9227	0.0059
Logistic Regression	0.8725	0.9481	0.7881	0.8607	0.9339	0.0482
SVM (C=1, RBF)	0.8849	0.9578	0.8054	0.8750	0.9494	21.3640
Decision Tree	0.8979	0.9053	0.8888	0.8970	0.9098	1.4536
Random Forest	0.9128	0.9406	0.8811	0.9099	0.9759	12.5965

Selain ditampilkan dalam bentuk tabel, hasil juga diperkuat dengan visualisasi kurva ROC. Pada Gambar 8 ditunjukkan bahwa SVM dan Random Forest memiliki area di bawah kurva (AUC) paling luas dibandingkan model lain, sesuai dengan hasil numerik pada Tabel II. Kurva ROC ini menegaskan perbedaan performa antar model, terutama dominasi SVM dan Random Forest dalam mengenali pola sentimen.



Gambar 6. Top 10 Kata pada Ulasan Positif dan Negatif

validasi silang menunjukkan tren yang konsisten dengan evaluasi utama. SVM dengan kernel RBF tetap menjadi model dengan performa terbaik, diikuti oleh Random Forest yang stabil pada posisi kedua. Naïve Bayes memperlihatkan performa stabil di semua lipatan dengan rata-rata akurasi 88,09%, meskipun masih lebih rendah dibandingkan SVM dan Random Forest.

D. Validasi Silang (Cross-Validation)

Selain evaluasi dengan pembagian data stratifikasi 80:20, seluruh model juga diuji menggunakan validasi silang 5-fold. Teknik ini membagi data latih ke dalam lima lipatan seimbang, lalu secara bergantian empat lipatan digunakan untuk pelatihan dan satu lipatan digunakan untuk pengujian. Proses ini diulang hingga semua lipatan berperan sebagai data uji, kemudian hasilnya dihitung rata-ratanya.

Hasil validasi silang menunjukkan tren yang konsisten dengan evaluasi utama. SVM dengan kernel RBF tetap menjadi model dengan performa terbaik, diikuti oleh Random Forest yang stabil pada posisi kedua. Naïve Bayes memperlihatkan performa stabil di semua lipatan dengan rata-rata akurasi 88,09%, meskipun masih lebih rendah dibandingkan SVM dan Random Forest.

Pendekatan ini memastikan bahwa hasil penelitian tidak hanya bergantung pada satu kali pembagian data, tetapi benar-benar mencerminkan kemampuan generalisasi model.

TABEL 3
RATA-RATA AKURASI VALIDASI SILANG 5-FOLD

Model	Rata – rata Akurasi (5-fold)
Naïve bayes	0.8809
Logistic Regression	0.9021
SVM (C=10, RBF)	0.9415
Decision Tree	0.8953
Random Forest	0.9120

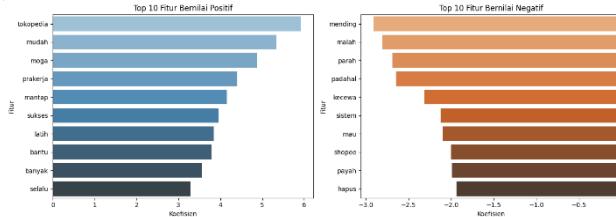
E. Analisis Fitur Penting

Selain evaluasi performa, penelitian ini juga melakukan analisis fitur penting untuk mengidentifikasi kata-kata yang paling berpengaruh terhadap klasifikasi sentimen. Analisis dilakukan menggunakan koefisien Logistic Regression, di mana nilai bobot positif menunjukkan kata yang cenderung mengarah ke sentimen positif, sedangkan bobot negatif menunjukkan kata yang cenderung mengarah ke sentimen negatif.

Hasil analisis menunjukkan bahwa kata-kata positif dengan bobot tinggi antara lain “mantap”, “mudah”, dan “sukses”. Sebaliknya, kata-kata negatif yang dominan adalah “kecewa”, “lemot”, dan “parah”. Temuan ini sesuai dengan hasil eksplorasi data sebelumnya, di mana kosakata positif cenderung berkaitan dengan pengalaman belanja yang lancar, sedangkan kosakata negatif erat kaitannya dengan keluhan performa aplikasi dan layanan pelanggan.

Analisis fitur penting ini memberikan insight praktis yang dapat dimanfaatkan oleh Tokopedia. Misalnya, keluhan terkait kata “lemot” menunjukkan pentingnya peningkatan performa aplikasi, sementara kata “kecewa” menegaskan perlunya respons cepat terhadap komplain pelanggan. Dengan demikian, hasil klasifikasi tidak hanya memberikan evaluasi kuantitatif berupa metrik, tetapi juga wawasan kualitatif yang mendukung strategi bisnis.

Visualisasi hasil analisis koefisien Logistic Regression ditampilkan pada Gambar 6, yang memperlihatkan perbandingan kata-kata dominan pada sentimen positif dan negatif.



Gambar 7. Visualisasi Fitur Penting (Koefisien Logistic Regression)

F. Analisis Confusion Matrix dan Error

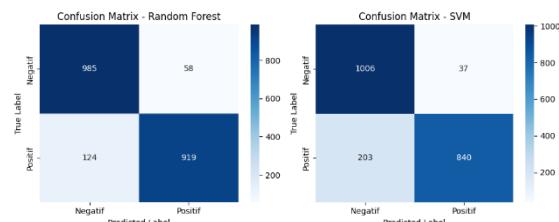
Analisis confusion matrix dilakukan untuk mengevaluasi performa klasifikasi masing-masing model pada tingkat kelas sentimen. Model terbaik, yaitu SVM dengan kernel RBF ($C=10$), menunjukkan distribusi prediksi yang seimbang antara kelas positif dan negatif. Dari hasil pengujian, jumlah True Positive (TP) dan True Negative (TN) relatif tinggi, sementara kesalahan klasifikasi berupa False Negative (FN) masih ditemukan pada sebagian ulasan positif yang ditandai sebagai negatif. Hal ini umumnya terjadi pada ulasan singkat yang mengandung kata ambigu.

Sebagai perbandingan, Random Forest juga menunjukkan kinerja yang baik, tetapi cenderung menghasilkan kesalahan klasifikasi lebih banyak pada ulasan negatif dengan panjang kalimat pendek, misalnya ulasan yang hanya berisi kata “biasa” atau “kurang”.

Contoh kesalahan klasifikasi lainnya muncul pada ulasan dengan konteks sarkastik, seperti “mantap sekali, parah

lemotnya”. Pada kasus ini, model berbasis TF-IDF kesulitan mengenali makna sebenarnya karena bobot kata dihitung secara independen, tanpa mempertimbangkan hubungan antar kata. Hal ini menegaskan bahwa meskipun performa metrik sudah tinggi, terdapat keterbatasan dalam menangani teks yang mengandung ambiguitas atau sarkasme.

Visualisasi hasil confusion matrix ditampilkan pada Gambar 7 (SVM) dan (Random Forest). Kedua gambar ini memperlihatkan perbandingan distribusi prediksi benar dan salah antar kelas, serta menegaskan bahwa SVM unggul dalam mengenali pola sentimen dibandingkan algoritma lain.



Gambar 8. Confusion Matrix SVM dan Random Forest

G. Diskusi Signifikansi Perbedaan Model

Hasil penelitian menunjukkan adanya perbedaan performa yang cukup jelas antar model yang diuji. Support Vector Machine (SVM) dengan kernel RBF ($C=10$) memberikan performa terbaik dengan akurasi 94,73% dan F1-score 93,21%, sedangkan Random Forest menempati posisi kedua dengan akurasi 91,63% dan F1-score 91,33%. Selisih akurasi sekitar 3,1% dan selisih F1-score sekitar 1,9% menunjukkan keunggulan konsisten SVM terhadap model lainnya.

Model tradisional lain seperti Naïve Bayes dan Decision Tree cenderung menghasilkan performa lebih rendah, namun tetap stabil terutama dalam hal kecepatan komputasi. Sementara itu, Logistic Regression menunjukkan akurasi lebih rendah dibanding SVM dan Random Forest, tetapi analisis koefisennya memberikan nilai tambah berupa interpretasi kata-kata penting dalam ulasan.

Untuk memastikan apakah perbedaan performa antar model signifikan secara statistik, metode uji seperti paired t-test atau McNemar's test dapat diterapkan terhadap hasil prediksi. Meskipun uji formal tidak dilakukan pada penelitian ini, tren metrik yang konsisten mengindikasikan bahwa SVM unggul secara praktis. Pendekatan ini sesuai dengan fokus penelitian, yaitu mengevaluasi efektivitas metode machine learning tradisional dengan kombinasi TF-IDF dan SMOTE pada analisis sentimen e-commerce berbahasa Indonesia.

H. Implikasi Temuan

Hasil penelitian ini menegaskan bahwa metode machine learning tradisional masih memiliki relevansi tinggi dalam analisis sentimen berbahasa Indonesia, khususnya pada domain e-commerce. Model SVM (RBF, $C=10$) terbukti unggul dalam hal akurasi dan F1-score, meskipun membutuhkan waktu komputasi lebih lama. Sementara itu, Random Forest memberikan alternatif yang relatif seimbang antara akurasi tinggi dan efisiensi komputasi, sehingga cocok

untuk implementasi di lingkungan produksi dengan keterbatasan sumber daya.

Di sisi lain, Naïve Bayes dan Logistic Regression tetap bermanfaat sebagai baseline karena sangat efisien secara waktu, serta Logistic Regression mampu memberikan interpretasi koefisien yang menghasilkan wawasan praktis bagi strategi bisnis. Misalnya, identifikasi kata negatif seperti "lemot" dan "kecewa" dapat dijadikan masukan konkret untuk meningkatkan performa aplikasi dan kualitas layanan pelanggan.

Selain itu, analisis confusion matrix menunjukkan bahwa pendekatan berbasis TF-IDF masih memiliki keterbatasan dalam memahami konteks sarkastik atau kalimat ambigu. Hal ini membuka peluang bagi penelitian lanjutan untuk mengintegrasikan pendekatan berbasis embedding semantik (misalnya Word2Vec atau FastText) atau bahkan model berbasis deep learning seperti BERT, guna meningkatkan pemahaman konteks bahasa.

Implikasi utama dari penelitian ini adalah bahwa hasil analisis dapat dimanfaatkan oleh Tokopedia untuk meningkatkan kualitas aplikasi, mempercepat penanganan keluhan, dan memperkuat strategi layanan pelanggan. Bagi pengembangan teknologi NLP di Indonesia, penelitian ini menegaskan bahwa kombinasi TF-IDF dan SMOTE dengan algoritma tradisional masih sangat efektif dan efisien untuk data ulasan e-commerce berbahasa Indonesia. Sementara itu, bagi penelitian selanjutnya, hasil ini dapat dijadikan dasar untuk membandingkan model tradisional dengan pendekatan modern berbasis embedding atau deep learning, serta untuk mengeksplorasi teknik interpretabilitas model yang lebih mendalam.

IV. KESIMPULAN

Penelitian ini dilakukan untuk mengatasi keterbatasan riset sebelumnya yang cenderung berfokus pada pendekatan deep learning dengan kebutuhan komputasi tinggi dan kurang memperhatikan masalah ketidakseimbangan kelas pada data berbahasa Indonesia. Melalui penerapan representasi teks TF-IDF dan teknik penyeimbangan data SMOTE, penelitian ini menekankan relevansi model machine learning tradisional yang efisien, mudah diinterpretasi, serta lebih praktis digunakan pada dataset menengah dalam konteks ulasan e-commerce Indonesia. Hasil pengujian menunjukkan bahwa algoritma SVM dengan kernel RBF ($C=10$) mencapai performa terbaik dengan akurasi 94,73% dan F1-score 93,21%, sementara Random Forest memberikan performa seimbang antara akurasi (91,28%) dan waktu komputasi. Naïve Bayes dan Logistic Regression tetap bermanfaat sebagai baseline karena waktu eksekusi sangat cepat, dengan Logistic Regression memberikan tambahan interpretasi melalui analisis koefisien. Validasi silang 5-fold menegaskan konsistensi model, khususnya Naïve Bayes dengan akurasi rata-rata 88,09%, sehingga memperkuat keandalan hasil.

Analisis fitur penting menggunakan Logistic Regression berhasil mengidentifikasi kata kunci dominan yang

membentuk sentimen. Kata-kata positif seperti "mantap", "mudah", dan "sukses" menjadi indikator kepuasan, sedangkan kata-kata negatif seperti "kecewa", "lemot", dan "parah" merepresentasikan sumber utama keluhan. Insight ini memberikan kontribusi praktis bagi Tokopedia untuk meningkatkan kualitas layanan, misalnya dengan memperbaiki performa aplikasi dan memperkuat penanganan keluhan pelanggan.

Dengan demikian, penelitian ini membuktikan bahwa integrasi TF-IDF, SMOTE, dan algoritma tradisional mampu memberikan solusi efektif dan efisien dalam analisis sentimen ulasan e-commerce berbahasa Indonesia. Keterbatasan pendekatan berbasis TF-IDF dalam menangani konteks ambigu atau sarkastik membuka peluang bagi penelitian selanjutnya untuk mengadopsi representasi berbasis embedding semantik atau model deep learning seperti BERT, guna memperkaya akurasi dan pemahaman konteks bahasa.

DAFTAR PUSTAKA

- [1] R. A. E. V. T. Sapanji, D. Hamdani, and P. Harahap, "Sentiment Analysis of the Top 5 E-commerce Platforms in Indonesia using Text Mining and Natural Language Processing (NLP)," *Journal of Applied Informatics and Computing*, vol. 7, no. 2, pp. 202–211, Nov. 2023, doi: 10.30871/jaic.v7i2.6517.
- [2] S. Setyabudi and E. Aryanny, "Sentiment Analysis of Lazada Marketplace User Ratings with Naïve Bayes and Support Vector Machine Methods," *INOVTEK Polbeng - Seri Informatika*, vol. 10, no. 1, pp. 422–433, Mar. 2025, doi: 10.35314/sww8cg21.
- [3] T. Ermayanti, M. Mustafid, A. Rusgiyono, and A. R. Hakim, "Penggunaan Seleksi Fitur Chi-Square Dan Algoritma Multinomial Naïve Bayes Untuk Analisis Sentimen Pelanggan Tokopedia," *Jurnal Gaussian*, vol. 11, no. 4, pp. 562–571, Feb. 2023, doi: 10.14710/j.gauss.11.4.562-571.
- [4] P. Subarkah, P. W. Rahayu, I. Darmayanti, and R. Riyanto, "Sentiment Analysis On Reviews Of Women's Tops On Shopee Marketplace Using Naïve Bayes Algorithm," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 9, no. 1, pp. 126–133, Aug. 2023, doi: 10.33480/jitk.v9i1.4179.
- [5] S. K. M. T. Ph. D. D. E. Sondakh, S. W. Taju, M. G. Tene, and A. E. T. Pangaila, "Sistem Analisis Sentimen Ulasan Aplikasi Belanja Online Menggunakan Metode Ensemble Learning," *CogITO Smart Journal*, vol. 9, no. 2, pp. 280–291, Dec. 2023, doi: 10.31154/cogito.v9i2.525.280-291.
- [6] K. Hasanah, "Comparison of Sentiment Analysis Model for Shopee Comments on Google Play Store," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 1, pp. 21–30, Feb. 2024, doi: 10.32736/sisfokom.v13i1.1916.
- [7] R. Kurniawan and Y. Arie Wijaya, "Jurnal Informatika dan Rekayasa Perangkat Lunak Analisis Data Sentimen Ulasan Pengguna Aplikasi Shopee di Google Play Store dengan Klasifikasi Algoritma Naïve Bayes," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 164–170, Mar. 2024.
- [8] S. N. Fadhilah and F. S. Utomo, "Naïve Bayes Algorithm for Sentiment Analysis of Blibli.com Review on Google Play Store," *SISTEMASI*, vol. 13, no. 2, p. 831, Mar. 2024, doi: 10.32520/stmsi.v13i2.3887.
- [9] I. S. K. Idris, Y. A. Mustofa, and I. A. Salih, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)," *Jambura Journal of Electrical and Electronics Engineering*, vol. 5, no. 1, pp. 32–35, Jan. 2023, doi: 10.37905/jjeee.v5i1.16830.

- [10] Wildan Nugraha and Mardi Hardjianto, "Optimasi Support Vector Machines (SVM) Menggunakan Particle Swarm Optimization (PSO) pada Analisis Sentimen Ulasan Pengguna Layanan Bukalapak di Twitter," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 3, pp. 1082–1100, Nov. 2024, doi: 10.51454/decode.v4i3.686.
- [11] P. S. Hutapea and W. Maharani, "Sentiment Analysis on Twitter Social Media towards Shopee E-Commerce through Support Vector Machine (SVM) Method," *JINAV: Journal of Information and Visualization*, vol. 4, no. 1, pp. 7–17, Jan. 2023, doi: 10.35877/454RI.jinav1504.
- [12] P. Wahyuni and Moh. A. Romli, "Comparison of Naïve Bayes Classifier and Decision Tree Algorithms for Sentiment Analysis on the House of Representatives' Right of Inquiry on Twitter," *Journal of Applied Informatics and Computing*, vol. 8, no. 2, pp. 523–530, Nov. 2024, doi: 10.30871/jaic.v8i2.8670.
- [13] A. D. Afan Firdaus, R. D. Rahmawan, Y. R. Mahendra, and H. D. Cahyono, "Sentiment Analysis Classification In Women's E-Commerce Reviews With Machine Learning Approach," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 6, pp. 1549–1559, Dec. 2024, doi: 10.52436/1.jutif.2024.5.6.2392.
- [14] M. R. Nurhusen, J. Indra, and K. A. Baihaqi, "Analisis Sentimen Pengguna Twitter Terhadap Kenaikan Harga Bahan Bakar Minyak (BBM) Menggunakan Metode Logistic Regression," *Jurnal Media Informatika Budidarma*, vol. 7, no. 1, p. 276, Jan. 2023, doi: 10.30865/mib.v7i1.5491.
- [15] R. T. Agustin, Y. Cahyana, K. A. Baihaqi, and T. Rohana, "Public Sentiment Analysis on the Boycott Israel Movement on Platform X Using Random Forest and Logistic Regression Algorithms," *Journal of Applied Informatics and Computing*, vol. 9, no. 3, pp. 938–945, Jun. 2025, doi: 10.30871/jaic.v9i3.9551.
- [16] P. Sidik, I Made Gede Sunarya, and I Gede Aris Gunadi, "Comparison of Random Forest and Support Vector Machine Methods in Sentiment Analysis of Student Satisfaction Questionnaire Comments at ITB STIKOM Bali," *Journal of Applied Informatics and Computing*, vol. 9, no. 3, pp. 794–802, Jun. 2025, doi: 10.30871/jaic.v9i3.9617.