



Analisis Sentimen Produk Berdasarkan Review Pelanggan Shopee Menggunakan KNN

Product Sentiment Analysis Based on Shopee Customer Reviews Using KNN

Fira Irwannia & Andre Hasudungan Lubis*

Teknik Informatika, Fakultas Teknik, Universitas Medan Area, Indonesia

Diterima: 21 April 2025; Direview: 22 April 2025; Disetujui: 24 April 2025

*Corresponding Email: andrehasudunganlubis@uma.ac.id

Abstrak

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan pelanggan mengenai produk mukena yang tersedia di aplikasi Shopee menggunakan algoritma K-Nearest Neighbors (KNN). Data yang digunakan merupakan data primer sebanyak 200 ulasan yang dikumpulkan secara manual. Proses analisis dimulai dari preprocessing data berupa case folding, tokenisasi, penghapusan stopwords, dan stemming, kemudian dilanjutkan dengan ekstraksi fitur menggunakan metode TF-IDF, dan klasifikasi menggunakan algoritma KNN. Penelitian ini juga melakukan evaluasi terhadap performa model dengan akurasi. Hasil pengujian menunjukkan bahwa proporsi data training dan nilai parameter $n_neighbors$ sangat mempengaruhi akurasi model. Proporsi data training sebesar 90% dan testing 10% menghasilkan akurasi tertinggi sebesar 90%. Namun, ketika nilai $n_neighbors = 3$, proporsi 70:30 menghasilkan performa terbaik sebesar 81,67%. Penelitian ini menunjukkan bahwa KNN mampu digunakan sebagai metode yang efektif dalam analisis sentimen terhadap ulasan produk.

Kata Kunci: Analisis Sentimen; Ulasan Produk; KNN; Shopee; TF-IDF.

Abstract

This study aims to conduct sentiment analysis on customer reviews of mukena products available on the Shopee application using the K-Nearest Neighbors (KNN) algorithm. The data used is primary data consisting of 200 reviews collected manually. The analysis process begins with data preprocessing such as case folding, tokenization, stopword removal, and stemming, followed by feature extraction using the TF-IDF method, and classification using the KNN algorithm. The model's performance is evaluated using a confusion matrix. The results show that the proportion of training data and the $n_neighbors$ parameter significantly affect the model's accuracy. A 90% training and 10% testing proportion produced the highest accuracy of 90%. However, with $n_neighbors = 3$, the best performance was achieved with a 70:30 data split, reaching 81.67% accuracy. This study demonstrates that KNN is an effective method for sentiment analysis on product reviews.

Keywords: Sentiment Analysis; Product Reviews; KNN; Shopee; TF-IDF.

PENDAHULUAN

Dalam beberapa tahun terakhir, belanja online melalui platform e-commerce telah menjadi bagian yang tidak terpisahkan dari kehidupan masyarakat modern. Kemudahan akses, efisiensi waktu, dan berbagai pilihan produk yang ditawarkan menjadi alasan utama meningkatnya popularitas platform e-commerce di Indonesia. Salah satu platform yang berhasil menarik perhatian masyarakat adalah Shopee, dengan jumlah pengguna aktif bulanan mencapai 117 juta pada kuartal pertama tahun 2023 [1]. Shopee menempati posisi kedua sebagai situs e-commerce dengan jumlah kunjungan bulanan terbanyak di Indonesia pada kuartal pertama tahun 2021 [2].

Pengguna dapat menilai dan mengulas aplikasi di Google Playstore [3]. Ulasan produk online mengurangi kemungkinan inkonsistensi pembelian dan membantu calon pembeli dalam mengambil keputusan [4]. Playstore merupakan salah satu layanan milik Google yang menyediakan berbagai jenis konten digital, seperti game, aplikasi, film, musik, dan buku [5].

Setiap aplikasi memiliki keunggulan dan kelemahan tersendiri yang dapat memicu tanggapan berbeda dari para pengguna, seperti rasa puas maupun tidak puas [6]. Analisis sentimen terhadap ulasan aplikasi media sosial menjadi sarana yang bermanfaat bagi bisnis untuk memahami kebutuhan serta preferensi pelanggan, sekaligus mengidentifikasi permasalahan yang berpotensi merusak citra mereka [7]. Analisis sentimen terhadap produk juga merupakan metode yang efisien untuk mengetahui bagaimana pelanggan menilai suatu produk. Penelitian ini menunjukkan bahwa ulasan dari pengguna memiliki pengaruh terhadap minat beli dan menjadi faktor penting dalam peningkatan penjualan [8].

Ulasan produk pada platform marketplace memberikan informasi yang berguna bagi calon pembeli dalam mengambil keputusan pembelian dan mengurangi risiko ketidaksesuaian produk. Marketplace yang dijadikan objek kajian dalam penelitian ini adalah Shopee, yang merupakan salah satu platform e-commerce populer di Indonesia maupun mancanegara, karena menawarkan kemudahan dalam transaksi jual beli berbagai kebutuhan sehari-hari [9]. Menurut Lubis dkk., [10], perilaku belanja online milenial dipengaruhi oleh kepercayaan, sikap, norma subjektif, dan kontrol perilaku yang dirasakan.



Melakukan analisis sentimen secara manual merupakan pekerjaan yang menantang karena melibatkan volume data yang besar, adanya ulasan pelanggan yang tidak seragam, serta keterbatasan kemampuan manusia dalam melakukan analisis secara manual.

Dalam penelitian berjudul "Analisis Sentimen Review Pelanggan Marketplace Shopee Indonesia Menggunakan Metode Algoritma K-Nearest Neighbors" oleh Susanti dan Akrom [11], melakukan analisis sentimen terhadap 150 ulasan produk Shopee yang dibagi rata antara komentar positif dan negatif. Mereka menerapkan algoritma K-Nearest Neighbors (KNN) untuk mengklasifikasikan sentimen ulasan. Hasil penelitian menunjukkan bahwa model KNN mencapai akurasi sebesar 66,67%, dengan precision positif 64,71%, precision negatif 69,23%, recall positif 73,33%, dan recall negatif 60%. Penelitian ini mengindikasikan bahwa meskipun KNN efektif dalam klasifikasi sentimen, terdapat ruang untuk perbaikan dalam hal akurasi dan keseimbangan antara precision dan recall.

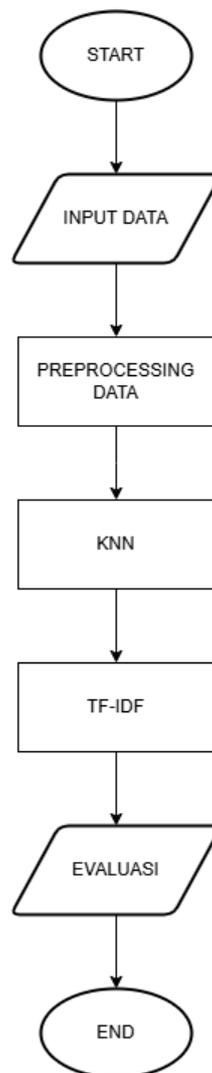
Rais, Said, dan Ruliana [12] dalam penelitian mereka yang berjudul "Text Classification on Sentiment Analysis of Marketplace SHOPEE Reviews On Twitter Using K-Nearest Neighbor (KNN) Method", melakukan analisis sentimen terhadap 150 tweet pengguna Twitter mengenai Shopee. Mereka menggunakan algoritma KNN dengan pembobotan TF-IDF dan menghitung jarak menggunakan Euclidean. Hasil penelitian menunjukkan bahwa model KNN mencapai akurasi tertinggi sebesar 90% dengan perbandingan data training dan testing 80%:20%. Penelitian ini menyoroti efektivitas KNN dalam menganalisis sentimen dari data teks yang diperoleh melalui media sosial.

Huda dan Prianto [13] dalam studi mereka yang berjudul "Analisis Algoritma K-Nearest Neighbor terhadap Sentimen Pengguna Aplikasi Shopee", melakukan analisis sentimen terhadap 2.000 ulasan aplikasi Shopee yang diperoleh melalui teknik web scraping. Mereka menggunakan algoritma KNN untuk mengklasifikasikan sentimen ulasan ke dalam kategori positif, netral, dan negatif. Hasil penelitian menunjukkan bahwa model KNN menghasilkan akurasi sebesar 70%, dengan nilai presisi 50,5%, recall 44,8%, dan F1-score 48,3%. Penelitian ini menyarankan penerapan teknik sampling data yang lebih optimal untuk mencapai distribusi kelas yang lebih seimbang pada data latih dan uji



Berdasarkan latarbelakang dan penelitian yang pernah dilakukan sebelumnya, penelitian ini mengusulkan solusi dengan menggunakan metode K-Nearest Neighbors (K-NN) untuk mengukur tingkat akurasi dalam mengklasifikasikan sentimen produk di platform Shopee [14]. Algoritma K-Nearest Neighbors merupakan algoritma populer yang termasuk dalam grup instance-based learning. K-Nearest Neighbors sendiri merupakan metode lazy-learning yang melakukan klasifikasi berdasarkan kedekatan jarak antara data yang dianalisis [15].

METODE PENELITIAN



Gambar 1 Metode Penelitian

Penelitian ini menerapkan algoritma KNN untuk melakukan analisis sentimen terhadap komentar produk mukena. Metode penelitian ini bertujuan untuk memberikan

gambaran yang jelas dan sistematis mengenai tahapan-tahapan utama, mulai dari input data komentar, praproses data, penerapan metode TF-IDF, dan penggunaan algoritma KNN, hingga evaluasi dalam bentuk akurasi.

Pengumpulan Data

Penelitian ini menggunakan data primer yang diperoleh langsung oleh penulis melalui proses pengumpulan secara manual. *Dataset* yang berhasil dihimpun terdiri atas 200 ulasan produk mukena yang tersedia di aplikasi *Shopee*. Proses pengumpulan ini dilakukan untuk memperoleh data ulasan yang autentik dan relevan, yang akan dijadikan dasar analisis dalam penelitian ini. Maka contoh data yang sudah berhasil dikumpulkan dapat dilihat pada Tabel 1.

Tabel 1 Contoh Dataset dari Shopee

No	Nama Pengguna	Komentar
1.	mlaa	pertama kali order alhamdulillah langsung puas, pengiriman jg cepat, packing mantap
2.	Ujibon	Kainnya lembut dan sesuai gambar. Mukena sangat nyaman dan akan membeli kembali
3.	ephy_ephy	Bahan adem sptnya, penjual ramah, cm warna almond milknya mirip grey ya?
4.	esampalupi	Bahannya halus, rendanya ga lebay. Untuk hadiah mama dan ipar semoga sukaa Terima kasih vouchernya
5.	dewi8150	Bahannya bagus adem nyaman dipakai sangat cocok, cuma warna tidak sama dengan dfoto dan asli Warna asli lebih ke coklat muda ato mocca

Preprocessing Data

Preprocessing data merupakan tahap awal dalam pengolahan data yang bertujuan untuk mempermudah proses analisis dan pemrosesan lebih lanjut. Tahapan ini dilakukan untuk membersihkan data dari gangguan (*noise*), menyederhanakan dimensi data, serta menyusun data agar lebih terstruktur [16]. Beberapa teknik yang digunakan dalam preprocessing data antara lain: *case folding*, yang berfungsi menyamakan semua karakter, sehingga kata seperti “The” dan “THE” dianggap identik; *stopword removal*, yang berguna untuk menghapus kata-kata yang dianggap tidak penting seperti kata hubung, biasanya menggunakan pendekatan *bag of words*; serta *stemming*, yaitu proses mengubah kata yang memiliki imbuhan menjadi bentuk dasarnya [17]. Pada penelitian ini, proses *preprocessing* meliputi beberapa tahapan, yaitu *case folding* untuk menyamakan format huruf, *tokenization* untuk Memecah teks menjadi unit-unit yang lebih kecil. Penghapusan



stopwords yang tidak memberikan makna signifikan, serta *stemming* untuk mereduksi kata ke bentuk dasarnya. Hasil dari *preprocessing* data dapat dilihat pada Tabel 2.

Tabel 2 Hasil *Preprocessing Data*

No	Komentar
1.	['pertama', 'order', 'alhamdulillah', 'langsung', 'puas', 'kirim', 'packing', 'mantap']
2.	['kain', 'lembut', 'sesuai', 'gambar', 'mukena', 'sangat', 'nyaman', 'beli', 'kembali']
3.	['bahan', 'adem', 'penjual', 'ramah', 'warna', 'almond', 'milk', 'mirip', 'grey']
4.	['bahan', 'halus', 'renda', 'lebay', 'hadiah', 'mama', 'ipar', 'semoga', 'suka', 'terima', 'kasih', 'voucher']
5.	['bahan', 'bagus', 'adem', 'nyaman', 'pakai', 'cocok', 'warna', 'foto', 'asli', 'warna', 'asli', 'coklat', 'muda']

TF-IDF

Setelah tahap *preprocessing* komentar selesai dilakukan, langkah berikutnya adalah memisahkan elemen kata dengan menggunakan teknik *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF merupakan metode pembobotan yang menghitung frekuensi kemunculan kata dalam satu dokumen (TF) serta frekuensi kata tersebut dalam keseluruhan kumpulan dokumen (IDF). Tujuan dari pendekatan ini adalah untuk menyeimbangkan perbedaan panjang dokumen sehingga menghasilkan ukuran bobot yang lebih akurat dan dapat diandalkan (Astiningrum, Batubulan, & Sias, 2020). TF-IDF sendiri merupakan kombinasi dari dua konsep utama, yaitu *term frequency* (TF) dan *inverse document frequency* [18]:

$$TF = \frac{t}{d} \quad (1)$$

$$idf = \log \left(\frac{N}{df(t)} \right) \quad (2)$$

$$TFidf = TF \cdot idf \quad (3)$$

Keterangan:

t = jumlah kemunculan kata tertentu dalam dokumen d

d = total keseluruhan kata pada dokumen

N = total dokumen yang ada

df(t) = jumlah dokumen yang memiliki kata t

K-Nearest Neighbors (KNN)

Algoritma K-Nearest Neighbors (KNN) menjadi salah satu metode populer dalam machine learning karena kemudahan implementasinya dalam menangani permasalahan yang kompleks. KNN juga efektif dalam mengatasi ketidakseimbangan kelas pada dataset, karena proses klasifikasinya didasarkan pada mayoritas tetangga terdekat. Dengan pendekatan ini, KNN tetap mampu menghasilkan prediksi yang baik meskipun terdapat kelas minoritas yang kurang dominan. Selain untuk klasifikasi, algoritma ini juga dapat diterapkan pada masalah regresi, di mana nilai prediksi dihitung dari rata-rata nilai tetangga terdekat. Berikut ini adalah tahapan-tahapan dalam algoritma KNN [4]:

- 1) Menentukan nilai K sebagai parameter utama.
- 2) Menghitung jarak antara data uji dan setiap data pada dataset pelatihan, menggunakan metrik tertentu jika atribut bersifat numerik

$$D(x_i, y_i) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4)$$

Keterangan:

D (xi, yi) = jarak

Xi = data latih

Yi = data uji

i = variabel data

n = Dimensi data

- 3) Mengurutkan jarak yang telah dihitung secara menurun (descending)
- 4) Memilih K data dengan jarak terdekat dari data uji.

Menentukan kelas mayoritas dari K tetangga tersebut dan mengklasifikasikan data uji ke dalam kelas tersebut.

HASIL DAN PEMBAHASAN

Pengaruh Proporsi Data Training terhadap Performa Model KNN

Pada penelitian ini, dilakukan pengujian terhadap performa model K-Nearest Neighbors (KNN) dengan variasi proporsi data training dan testing, sebagaimana terlihat pada Tabel 3.



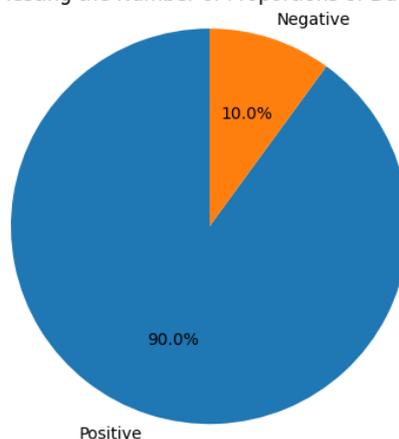
Tabel 3 Pengujian Jumlah Proporsi Data

Pengujian	Data Training	Data Testing	Performa
1	70 %	30 %	73,33 %
2	80 %	20 %	82,5 %
3	90 %	10 %	90 %

Dari hasil pengujian tersebut dapat dilihat bahwa peningkatan proporsi data training dari 70% menjadi 90% diikuti dengan peningkatan akurasi model dari 73,33% menjadi 90%. Hal ini menunjukkan bahwa semakin banyak data yang digunakan untuk pelatihan, model KNN mampu belajar pola sentimen dalam ulasan produk dengan lebih baik, sehingga prediksi sentimen menjadi lebih akurat. Peningkatan akurasi ini juga mengindikasikan bahwa model lebih mampu mengenali fitur-fitur penting dari data ulasan setelah melalui proses pembobotan TF-IDF.

Namun, perlu diingat bahwa proporsi data testing yang terlalu kecil (10%) dapat mempengaruhi validitas evaluasi model, karena data testing menjadi kurang representatif dalam mengukur kemampuan generalisasi model. Oleh karena itu, meskipun akurasi tertinggi didapat pada pengujian dengan 90% data training, penentuan proporsi data yang optimal harus mempertimbangkan keseimbangan antara pelatihan dan pengujian agar hasil dapat digeneralisasi ke data baru.

Testing the Number of Proportions of Data



Gambar 2 Hasil Sentimen Berdasarkan Jumlah Proporsi Data

Gambar 2 memperlihatkan distribusi hasil sentimen pada proporsi data training tertinggi, dengan dominasi sentimen positif sebesar 90% dan sentimen negatif sebesar 10%. Hal ini mengindikasikan bahwa sebagian besar ulasan produk mukena di platform Shopee bersifat positif, yang sejalan dengan konteks produk dan ulasan pengguna.

Pengaruh Parameter $n_neighbors$ dan Proporsi Data terhadap Performa Model

Selain variasi proporsi data training, pengujian juga dilakukan dengan mengatur parameter $n_neighbors$ pada nilai 3 untuk mengetahui pengaruhnya terhadap performa model, sebagaimana disajikan pada Tabel 4.

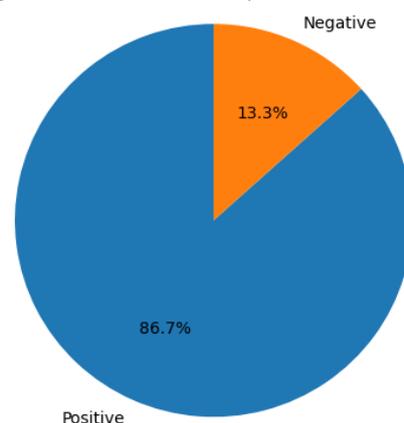
Tabel 4 Pengujian Jumlah Proporsi Data & Parameter

Pengujian	Data Training	Data Testing	$N_neighbors$	Performa
1	70 %	30 %	3	81,67 %
2	80 %	20 %	3	75 %
3	90 %	10 %	3	70 %

Hasil pengujian pada Tabel 4 menunjukkan pola berbeda dibanding Tabel 3. Akurasi terbaik diperoleh pada proporsi data training 70% dan testing 30%, yakni 81,67%. Sedangkan peningkatan proporsi training menjadi 80% dan 90% justru menurunkan performa model dengan akurasi masing-masing 75% dan 70%.

Fenomena ini dapat dijelaskan dari sudut pandang overfitting dan generalisasi model. Proporsi data testing yang cukup besar (30%) memberikan evaluasi yang lebih akurat terhadap kemampuan model memprediksi data yang belum pernah dilihat. Dengan nilai $n_neighbors=3$, model mengandalkan keputusan mayoritas dari tiga tetangga terdekat, yang lebih stabil jika data testing cukup representatif. Sebaliknya, proporsi testing yang terlalu kecil berpotensi menyebabkan model tampak memiliki akurasi rendah saat diuji dengan data yang terbatas, meskipun pada pelatihan model cukup optimal.

Testing the Number of Data Proportions and Parameters



Gambar 3 Hasil Sentimen Berdasarkan Jumlah Proporsi Data & Parameter

Gambar 3 mengilustrasikan distribusi sentimen berdasarkan pengujian ini, dengan sentimen positif sebesar 86,7% dan negatif 13,3%. Penurunan proporsi sentimen positif dibanding pengujian sebelumnya menandakan adanya variasi dalam data testing dan

kemungkinan munculnya ulasan dengan sentimen negatif lebih banyak dalam dataset pengujian yang lebih besar.

Pembahasan

Penelitian ini bertujuan untuk mengukur efektivitas algoritma KNN dalam mengklasifikasikan sentimen ulasan produk mukena di Shopee dan mengetahui pengaruh parameter proporsi data dan `n_neighbors` terhadap performa model.

Berdasarkan hasil yang diperoleh:

- Efektivitas KNN: Algoritma KNN terbukti efektif untuk analisis sentimen produk dengan akurasi terbaik mencapai 90% pada proporsi data training 90%. Ini menjawab pertanyaan apakah KNN dapat diaplikasikan untuk analisis sentimen pada dataset ulasan produk lokal.
- Pengaruh Proporsi Data Training dan Testing: Proporsi data training yang lebih besar secara umum meningkatkan performa model, tetapi proporsi testing yang terlalu kecil dapat menyebabkan kurang representatifnya hasil evaluasi, sehingga perlu keseimbangan. Proporsi 70:30 memberikan hasil optimal jika dipadukan dengan pengaturan parameter `n_neighbors` tertentu.
- Pengaruh Parameter `n_neighbors`: Nilai `n_neighbors` sangat berpengaruh dalam performa model. Pada pengujian dengan `n_neighbors=3`, proporsi 70:30 memberikan hasil terbaik, mengindikasikan bahwa jumlah tetangga terdekat yang dipilih harus disesuaikan dengan proporsi data agar akurasi optimal.

Selain itu, dominasi sentimen positif dalam hasil menunjukkan persepsi pelanggan terhadap produk mukena di Shopee sebagian besar positif, yang merupakan indikator positif bagi penjual dan platform e-commerce terkait.

SIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa algoritma K-Nearest Neighbors (KNN) dapat digunakan secara efektif untuk melakukan klasifikasi sentimen terhadap ulasan produk. Proporsi data training dan pemilihan nilai `n_neighbors` berpengaruh signifikan terhadap performa model. Pengujian menunjukkan bahwa proporsi data 90:10 memberikan akurasi tertinggi, sementara nilai `n_neighbors = 3` dengan proporsi data 70:30 menghasilkan performa terbaik. Adapun hasil sentimen



positif sebesar 86,7 % sedangkan sentimen negatif sebesar 13,3%. Oleh karena itu, pemilihan parameter yang tepat sangat penting dalam penerapan algoritma KNN untuk analisis sentimen.

DAFTAR PUSTAKA

- [1] Katadata, "Market Size of Indonesia's Hijab Industry Reaches Rp 33 Trillion," 2021.
- [2] C. Cahyaningtyas, Y. Nataliani, and I. R. Widiyanti, "Analisis Sentimen Pada Rating Aplikasi Shopee Menggunakan Metode Decision Tree Berbasis SMOTE," *J. Teknol. Inf.*, pp. 174–184, 2021.
- [3] M. N. Muttaqin and I. Kharisudin, "Analisis Sentimen Aplikasi Gojek Menggunakan Support Vector Machine Dan K-Nearest Neighbor," *J. Math.*, pp. 22–27, 2021.
- [4] S. Alfaris and Kusnawi, "Komparasi Metode KNN dan Naive Bayes Terhadap Analisis Sentimen Pengguna Aplikasi Shopee," *Indones. J. Comput. Sci.*, pp. 2766–2776, 2023.
- [5] M. F. F. Supeli and Setiaji, "Klasifikasi Sentimen Positif dan Negatif Pada Aplikasi Vidio Dengan Algoritma K-Nearest Neighbor," *Infonesian J. Comput. Sci.*, pp. 7–15, 2023.
- [6] A. P. A. Giovani, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru di Twitter Menggunakan Algoritma Klasifikasi," *J. TEKNOINFO*, pp. 116–124, 2020.
- [7] D. Ardiansyah, A. Saepudin, R. Aryanti, E. Fitriani, and Royadi, "Analisis Sentimen Review Pada Aplikasi Media Sosial Tiktok Menggunakan Algoritma K-NN dan SVM Berbasis PSO," *J. Inform. Kaputama*, pp. 233–241, 2023.
- [8] M. Nurul and I. Isnalita, "Pengaruh Jumlah Pengunjung, Ulasan Produk, Reputasi Toko Dan Status Gold Badge Pada Penjualan Dalam Tokopedia," *E-Jurnal Akunt.*, pp. 1855–1865, 2019.
- [9] A. H. Ruger, M. Suyanto, and M. P. Kurniawan, "Sentiment Analisis Pelanggan Shopee di Twitter dengan Algoritma Naive Bayes," *J. Inf. Technol.*, pp. 26–29, 2021.
- [10] A. H. Lubis, W. R. Amelia, S. N. Ramadhani, A. A. Pane, and S. Aryza, "Indonesian millennials' behavior intention to online shopping through instagram," *Int. J. Sci. Technol. Res.*, vol. 8, no. 11, pp. 2466–2471, 2019.
- [11] L. Susanti and A. Akrom, "ANALISIS SENTIMEN REVIEW PELANGGAN MARKETPLACE SHOPEE INDONESIA MENGGUNAKAN METODE ALGORITMA K-NEAREST NEIGHBORS," *J. ESIT (E-Bisnis, Sist. Informasi, Teknol. Informasi)*, vol. 18, no. 2, Jul. 2023, Accessed: May 23, 2025. [Online]. Available: <https://openjournal.unpam.ac.id/index.php/ESIT/article/view/32740>
- [12] Z. Rais, R. N. Said, and R. Ruliana, "Text Classification on Sentiment Analysis of Marketplace SHOPEE Reviews On Twitter Using K-Nearest Neighbor (KNN) Method," *JINAV J. Inf. Vis.*, vol. 3, no. 1, pp. 1–8, Jul. 2022, doi: 10.35877/454RI.JINAV1389.
- [13] M. Saifurridho, M. Martanto, and U. Hayati, "Analisis Algoritma K-Nearest Neighbor terhadap Sentimen Pengguna Aplikasi Shopee," *J. Inform. Terpadu*, vol. 10, no. 1, pp. 21–26, Mar. 2024, doi: 10.54914/JIT.V10I1.1054.
- [14] L. S. W. W. Kean, Indirati., and Marji, "Analisis Sentimen Review Shopee Berbahasa Indonesia Menggunakan Improved K-Nearest Neighbor dan Jaro Winkler Distance," *J. Pengemb. Technol. Inf. dan Ilmu komput.*, pp. 7172–7179, 2019.
- [15] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *J. Inform. Univ. Pamulang*, p. 646, 2021.
- [16] M. B. A. Darmawan, F. Dewanta, and S. Astuti, "Analisis Perbandingan Algoritma Decision Tree, Random Forest, dan Naive Bayes untuk Prediksi Banjir di Desa Dayeuhkolot," *TELKA*, pp. 52–61, 2023.
- [17] A. A. Syam, G. H. M, A. Salim, D. F. Suriyanto, and M. F. B, "Analisis Teknik Preprocessing Pada Sentimen Masyarakat Terkait Konflik Israel-Palestina Menggunakan Support Vector Machine," *J. Ilm. Penelit. dan Pembelajaran Inform.*, pp. 1464–1472, 2024.
- [18] F. Rozi, F. Sukmana, and M. N. Adami, "Pengelompokan Judul Buku dengan Menggunakan Algoritma K-Nearest Neighbor (K-NN) dan Term Frequency - Inverse Document Frequency (TF-IDF)," *J. Inform. Merdeka Pasuruan*, pp. 1–5, 2021.

