

Implementasi Algoritma *K-Nearest Neighbors* untuk Klasifikasi Sentimen pada Ulasan Pakaian Anak di Platform TikTok Shop

Murtiwiayati¹, Rischa Fitriawhalia Arini², Leli Safitri³, Santi Widiyanti⁴, Iwan Setiadi⁵
^{1,2,3,4,5}Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Gunadarma, Jakarta, Indonesia

¹murtiwiayati@staff.gunadarma.ac.id, ²rischa20122003@gmail.com, ³leli.s@staff.gunadarma.ac.id,
⁴santiw@staff.gunadarma.ac.id, ⁵iwangsa@staff.gunadarma.ac.id

Submit : 04 Nov 25 | Diterima : 12 Nov 2025 | Terbit : 15 Nov 2025

ABSTRAK

Di tengah pesatnya perkembangan *e-commerce*, ulasan konsumen di platform digital telah menjadi indikator penting dalam menilai persepsi terhadap suatu produk. Sayangnya, potensi informasi yang terkandung dalam volume besar ulasan tersebut kerap belum dimanfaatkan secara maksimal oleh pelaku usaha. Penelitian ini mengkaji penerapan algoritma *K-Nearest Neighbors* (KNN) dalam menganalisis sentimen ulasan pelanggan terhadap produk pakaian anak yang dijual melalui toko Aluna Dearest di TikTok Shop. Data ulasan dikumpulkan dari periode 2022 hingga 2025, kemudian melalui serangkaian tahap *preprocessing* meliputi *case folding*, eliminasi *stopword*, dan *stemming* untuk membersihkan dan menstandarkan teks. Representasi numerik dari data teks diperoleh menggunakan pendekatan *Term Frequency-Inverse Document Frequency* (TF-IDF), sebelum diklasifikasikan ke dalam tiga kategori sentimen: positif, negatif, dan netral. Model KNN dilatih dengan pembagian data 80% untuk pelatihan dan 20% untuk pengujian, serta dilakukan eksplorasi nilai parameter *K* guna mengoptimalkan kinerja model. Hasil evaluasi menunjukkan bahwa model mencapai akurasi sebesar 87,7%, presisi 85,8%, *recall* 87,7%, dan *F1-score* 82,8%. Temuan ini mengindikasikan bahwa algoritma KNN mampu memberikan performa klasifikasi sentimen yang andal pada data ulasan *e-commerce*, sehingga dapat menjadi alat bantu strategis bagi penjual dalam memahami kepuasan pelanggan, meningkatkan kualitas produk, serta merancang pendekatan pemasaran yang lebih responsif terhadap kebutuhan pasar.

Kata kunci : Analisis Sentimen, *K-Nearest Neighbors*, TF-IDF, *R-commerce*, TikTok Shop, *Social Commerce*, *Text Classification*

PENDAHULUAN

Kemajuan teknologi digital menjadikan aktivitas jual beli melalui *e-commerce* sebagai salah satu pola konsumsi yang melekat dalam kehidupan masyarakat modern. Salah satu komponen penting dalam aktivitas perdagangan daring adalah sistem ulasan dan peringkat pelanggan yang merepresentasikan pengalaman mereka terhadap produk yang dibeli. Platform *e-commerce* seperti TikTok Shop, dengan jumlah pengguna serta ulasan yang sangat besar, menyimpan berbagai informasi yang bernilai tinggi. Namun, meskipun volume ulasan yang tersedia sangat banyak, sebagian penjual dan produsen masih mengalami kesulitan dalam menganalisis serta memanfaatkan data tersebut secara optimal untuk peningkatan kualitas produk.

Dengan menerapkan algoritma *K-Nearest Neighbors* (KNN), penelitian ini diharapkan dapat mengklasifikasikan ulasan ke dalam tiga kategori, yaitu positif, negatif, dan netral, sehingga mampu memberikan gambaran umum mengenai persepsi pelanggan terhadap kualitas produk. Hasil analisis diharapkan dapat mengungkap keluhan yang paling sering muncul, pujian yang paling banyak diterima, serta aspek-aspek kualitas produk yang perlu ditingkatkan oleh penjual. Dalam konteks ini, pemanfaatan teknologi kecerdasan buatan (*Artificial Intelligence*) dan pembelajaran mesin (*machine learning*) memiliki potensi besar untuk mengoptimalkan pengelolaan serta interpretasi ulasan pelanggan.

Salah satu metode yang relevan adalah penerapan algoritma *K-Nearest Neighbors* (KNN), yang mampu mempercepat serta meningkatkan efisiensi proses analisis dalam mengidentifikasi produk dengan kualitas rendah maupun produk yang memperoleh ulasan positif.

METODOLOGI PENELITIAN

Penelitian ini mengimplementasikan algoritma *K-Nearest Neighbors* (KNN) melalui rangkaian proses analitik untuk mengklasifikasikan sentimen ulasan pelanggan terhadap pakaian anak di TikTok Shop, dengan tujuan memperoleh hasil klasifikasi yang akurat dari data ulasan publik.



Gambar 1. Alur Penelitian

Pada gambar 1 dijelaskan bahwa penelitian ini mengikuti alur metodologis yang terstruktur dan terdiri atas beberapa tahapan berurutan untuk menganalisis sentimen ulasan pelanggan terhadap produk pakaian anak-anak di toko *Aluna Dearest* pada platform TikTok Shop. Tahap pertama adalah pengumpulan data, yaitu menghimpun ulasan pelanggan yang bersifat publik dan relevan dengan produk pakaian anak-anak. Data awal berupa teks mentah yang belum melalui proses pembersihan atau transformasi.

Selanjutnya, data teks mengalami preprocessing untuk meningkatkan kualitas dan konsistensi representasi linguistik. Tahapan ini meliputi *case folding* (pengubahan seluruh karakter menjadi huruf kecil), tokenisasi (pemisahan teks menjadi unit kata), eliminasi *stopword* (penghapusan kata umum yang tidak membawa makna sentimen), serta *stemming* (pengembalian kata ke bentuk dasarnya). Proses ini bertujuan menghasilkan teks yang lebih terstruktur dan siap untuk analisis lebih lanjut.

Setelah preprocessing, dilakukan ekstraksi fitur melalui representasi numerik menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF). Pendekatan ini mengubah dokumen teks menjadi vektor numerik berbobot, sehingga dapat diproses oleh algoritma pembelajaran mesin. Model dilatih dan diuji menggunakan skema pembagian data 80% untuk pelatihan (*training*) dan 20% untuk pengujian (*testing*). Data latih digunakan untuk membangun pola klasifikasi berdasarkan label sentimen, sedangkan data uji berfungsi sebagai representasi data baru untuk mengukur generalisasi model.

Tahap terakhir adalah evaluasi kinerja model, yang dilakukan dengan menghitung metrik klasifikasi standar, meliputi akurasi (*accuracy*), presisi (*precision*), *recall*, dan *F1-score*. Evaluasi

ini memberikan gambaran objektif mengenai efektivitas algoritma KNN dalam mengklasifikasikan sentimen ulasan pelanggan secara akurat dan andal. Kinerja model dievaluasi dengan menggunakan beberapa metrik evaluasi, antara lain akurasi, *precision*, *recall*, dan *f1-score*. Perhitungan metrik evaluasi menggunakan persamaan berikut:

Rumus *Precision*:

$$Precision = \frac{TP}{TP+FP} \dots\dots\dots 1$$

Rumus *Recall*:

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots 2$$

Rumus *F1-Score*:

$$F1-Score = 2 \times \frac{(Precision \times Recall)}{Precision+Recall} \dots\dots\dots 3$$

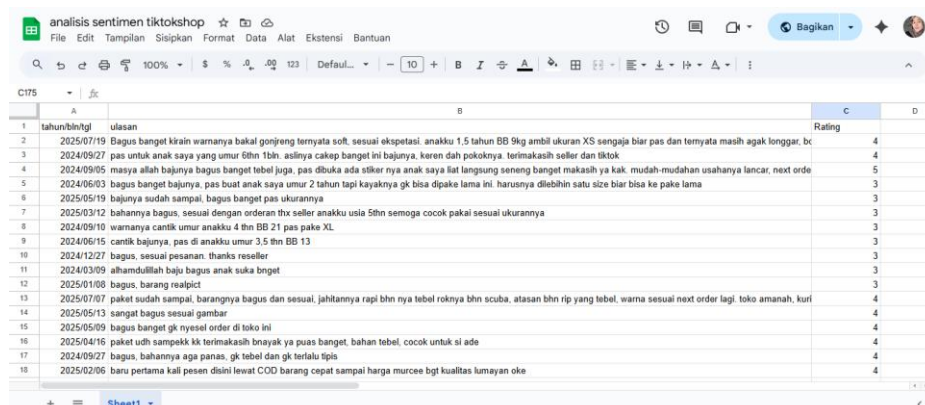
Keterangan:

- **True Positive (TP)** merupakan kondisi ketika model berhasil melakukan prediksi dengan tepat, di mana hasil prediksi sesuai dengan kelas sebenarnya.
- **True Negative (TN)** adalah kondisi ketika model memprediksi sebuah data sebagai bukan bagian dari kelas tertentu, dan hasil tersebut sesuai dengan kondisi sebenarnya.
- **False Positive (FP)** menggambarkan situasi ketika model memberikan prediksi positif, padahal kenyataannya data tersebut bukan termasuk dalam kelas yang dimaksud.
- **False Negative (FN)** terjadi ketika model memprediksi sebuah data sebagai negatif, sedangkan pada kenyataannya data tersebut justru termasuk dalam kelas yang seharusnya diprediksi positif.

HASIL DAN PEMBAHASAN

Pengumpulan Data

Tahap ini merupakan langkah awal dalam penelitian, yaitu proses pengumpulan data yang diperlukan, berupa data ulasan dan *rating* pada *platform TikTok Shop* toko Aruna Dearest. Pada tahap ini, terdapat tiga variabel utama dalam *dataset*, yaitu tahun, tanggal dan bulan, ulasan, serta *rating*.



	A	B	C	D
1	tahun/bulan/tgl	ulasan		Rating
2	2025/07/19	Bagus banget kirain warnanya bakal gonjrong ternyata soft, sesuai ekspektasi. anakku 1,5 tahun BB 9kg ambil ukuran XS sengaja biar pas dan ternyata masih agak longgar. bc		4
3	2024/09/27	pas untuk anak saya yang umur 5thn tbh. aslinya caket banget tpi bajunya. keren dah pokoknya terimakasih seller dan bltk		4
4	2024/09/05	masya allah bajunya bagus banget tebal jaja, pas dibuka ada stiker nya anak saya liat langsung seneng banget makasih ya kak. mudah-mudahan usahanya lancar, next order		5
5	2024/06/03	bagus banget bajunya, pas buat anak saya umur 2 tahun tapi kayaknya gk bisa dipake lama ini. harusnya dilebihin satu size biar bisa ke pake lama		3
6	2025/05/19	bajunya sudah sampai, bagus banget pas ukurannya		3
7	2025/03/12	bahannya bagus, sesuai dengan orderan tnx seller anakku usia 5thn semoga cocok pakai sesuai ukurannya		3
8	2024/09/10	warnanya cantik umur anakku 4 thn BB 21 pas pake XL		3
9	2024/06/15	cantik bajunya, pas di anakku umur 3,5 thn BB 13		3
10	2024/12/27	bagus, sesuai pesanan. thanks reseller		3
11	2024/03/09	alhamdulillah baju bagus anak suka bnget		3
12	2025/01/08	bagus, barang realpict		3
13	2025/07/07	paket sudah sampai, barangnya bagus dan sesuai, jahitannya rapi bhn nya tebal roknya bhn scuba, atasan bhn rip yang tebal, warna sesuai next order lagi. toko amanah, kurl		4
14	2025/05/13	sangat bagus sesuai gambar		4
15	2025/05/09	bagus banget gk nyesel order di toko ini		4
16	2025/04/16	paket udh sampekk kk terimakasih bnayak ya puas banget, bahan tebal, cocok untuk si ade		4
17	2024/09/27	bagus, bahannya aga panas, gk tebal dan gk terlalu tipis		4
18	2025/02/06	baru pertama kali pesen disini lewat COD barang cepat sampai harga murrce bgt kualitas lumayan oke		4

Gambar 2. Hasil Pengumpulan Data

Ekstraksi Fitur

Pada tahap ini dilakukan proses ekstraksi fitur menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)*. Metode *TF-IDF* berfungsi untuk menghitung bobot setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen (ulasan) serta memperhatikan

tingkat kelangkaan kata tersebut di seluruh kumpulan dokumen. Dengan cara ini, kata yang sering muncul namun memiliki makna khusus akan memperoleh bobot yang lebih tinggi, sedangkan kata yang muncul terlalu sering dan dianggap kurang informatif akan memiliki bobot yang lebih rendah. Dalam penelitian ini, jumlah fitur dibatasi hingga 5.000 kata yang dianggap paling informatif.

Preprocessing Data

Pada tahap *preprocessing*, data ulasan yang telah melalui proses ekstraksi fitur dibersihkan agar siap digunakan dalam tahap pemodelan. Proses ini mencakup beberapa langkah penting. Pertama, seluruh teks diubah menjadi huruf kecil (*lowercasing*) untuk menyeragamkan format penulisan. Selanjutnya, dilakukan pembersihan teks dengan menghapus karakter non-huruf seperti angka, tanda baca, dan simbol, sehingga hanya tersisa kata-kata yang relevan untuk analisis lebih lanjut.

1. Missing Value & Duplikat

Tahap pembersihan data dari *missing value* dan data duplikat dilakukan untuk memastikan kualitas data yang digunakan dalam proses analisis dan pemodelan, sehingga hasil yang diperoleh lebih akurat serta dapat diandalkan. Berdasarkan hasil pengecekan data yang ditunjukkan pada Gambar 3, diketahui bahwa tidak terdapat *missing value* pada kolom tahun/bulan/tanggal, ulasan, maupun *rating*, sehingga jumlah data tetap sebanyak 2.702 setelah proses verifikasi nilai kosong. Selain itu, hasil pemeriksaan juga menunjukkan adanya 53 data duplikat. Setelah dilakukan penghapusan data tersebut, jumlah data yang digunakan untuk tahap analisis selanjutnya menjadi 2.649 baris.

```

↩ Jumlah missing value per kolom:
  rating      0
  ulasan    318
  label       0
dtype: int64
Jumlah data setelah drop missing value: 2702
Jumlah data duplikat: 53
Jumlah data setelah hapus duplikat: 2649

```

Gambar 3. Output Missing Value

2. Normalisasi

Pada tahap ini dilakukan proses normalisasi terhadap nilai *rating* yang terdapat dalam *dataset*. Normalisasi bertujuan untuk mengubah skala nilai ke dalam rentang 0 hingga 1, sehingga memudahkan proses pengolahan data serta memastikan bahwa setiap variabel memiliki bobot yang sebanding. Hasil normalisasi ini dimasukkan kembali ke dalam kolom *rating*, di mana nilai *rating* yang semula berada pada skala 1 hingga 5 diubah menjadi skala antara 0 dan 1. Tahap ini penting dilakukan untuk menghindari dominasi variabel dengan skala besar terhadap variabel lain dalam proses analisis dan pemodelan.

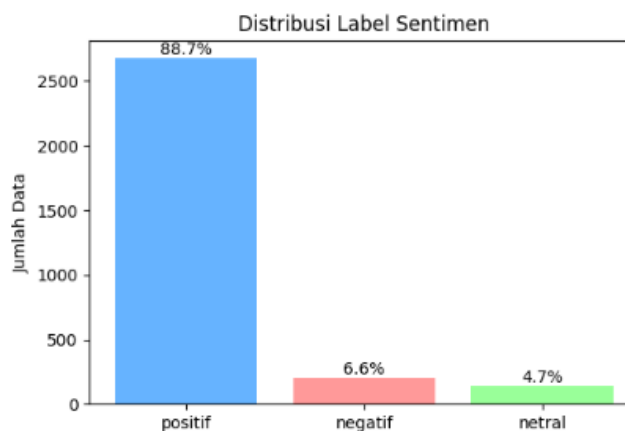
	Rating	Rating_norm
0	4	0.75
1	4	0.75
2	5	1.00
3	3	0.50
4	3	0.50

Gambar 4. Output Normalisasi

3. Visualisasi Distribusi Label

Distribusi label sentimen ditampilkan dalam bentuk diagram lingkaran yang terbagi menjadi beberapa bagian sesuai dengan persentase masing-masing kategori. Berdasarkan diagram tersebut, sentimen positif mendominasi dengan proporsi sebesar 88,7%, diikuti

oleh sentimen negatif sebesar 6,6%, dan sentimen netral sebesar 4,7%. Diagram lingkaran ini digunakan untuk menekankan perbandingan proporsi antar kategori sentimen secara visual sehingga dapat dipahami dengan mudah secara sekilas.



Gambar 5. Output Visualisasi Bar Chart

Penerapan Algoritma KNN

Pada tahap ini dilakukan proses pelabelan sentimen berdasarkan nilai *rating* yang diberikan oleh pengguna. Nilai *rating* ≥ 4 dikategorikan sebagai sentimen positif karena dianggap merepresentasikan tingkat kepuasan pengguna terhadap produk. Nilai *rating* = 3 diklasifikasikan sebagai sentimen netral, karena berada pada posisi tengah dan tidak menunjukkan kecenderungan yang jelas terhadap kepuasan maupun ketidakpuasan. Sementara itu, nilai *rating* ≤ 2 dikategorikan sebagai sentimen negatif, karena dianggap menggambarkan ketidakpuasan pengguna terhadap produk.

```
label
positif    2678
negatif    199
netral     143
Name: count, dtype: int64
```

Gambar 6. Output Pembuatan Label

Output pada Gambar 6 menunjukkan hasil distribusi data setelah dilakukan proses pelabelan berdasarkan nilai *rating*. Data ulasan terbagi ke dalam tiga kategori sentimen, yaitu netral, negatif, dan positif. Dari total 3.752 data ulasan, terdapat 143 ulasan yang termasuk dalam kategori netral, 199 ulasan dalam kategori negatif, dan 2.678 ulasan dalam kategori positif. Hasil ini menunjukkan bahwa mayoritas ulasan pelanggan memiliki kecenderungan sentimen positif, sedangkan ulasan netral dan negatif memiliki proporsi yang lebih kecil.

Pelatihan Model dan Pengujian Model

Berdasarkan penelitian yang dilakukan oleh Saifurridho *et al.* (2024) dalam jurnal *Analisis Algoritma K-Nearest Neighbor terhadap Sentimen Pengguna Aplikasi Shopee*, dilakukan pembagian *dataset* menjadi data latih (*training data*) dan data uji (*testing data*) menggunakan fungsi *train_test_split()*. Proporsi pembagian yang digunakan adalah 80% data untuk pelatihan dan 20% data untuk pengujian, dengan parameter *stratify* = *y* agar distribusi label pada data latih dan data uji tetap seimbang. Selain itu, parameter *random_state* = 42 digunakan untuk memastikan bahwa hasil pembagian data dapat direproduksi secara konsisten.

```

Jumlah data training: 2119
Jumlah data testing: 530
KNeighborsClassifier
KNeighborsClassifier(n_neighbors=3, weights='distance')
  
```

Gambar 7. Output Model

Output pada Gambar 7 menunjukkan bahwa dari total 2.649 data ulasan, sebanyak 2.119 data digunakan sebagai data latih (*training data*) dan 530 data digunakan sebagai data uji (*testing data*) dengan perbandingan 80:20.

Evaluasi Model

Tahap evaluasi dilakukan untuk menilai performa model *K-Nearest Neighbors (KNN)* setelah melalui proses pelatihan. Model yang telah dibangun digunakan untuk melakukan prediksi pada data uji (X_{test}), sehingga diperoleh label prediksi (y_{pred}). Berdasarkan hasil yang ditunjukkan pada Gambar 8, model *K-Nearest Neighbors (KNN)* memiliki performa klasifikasi yang cukup baik secara keseluruhan, dengan nilai akurasi sebesar 87,7%, presisi 85,8%, *recall* 87,7%, dan *F1-score* sebesar 82,8%. Namun, berdasarkan laporan klasifikasi yang dihasilkan, performa model diketahui sangat dipengaruhi oleh distribusi kelas pada data. Kategori sentimen positif yang memiliki jumlah data paling banyak menunjukkan hasil evaluasi tertinggi, dengan nilai *recall* sempurna (1,00) dan *F1-score* sebesar 0,93.

```

Akurasi : 0.8773584905660378
Presisi : 0.8581509433962264
Recall : 0.8773584905660378
F1-Score : 0.8284170348223994

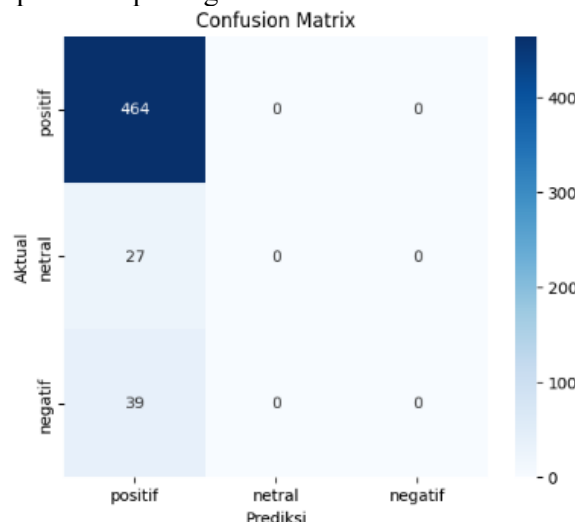
Laporan Klasifikasi:
      precision    recall  f1-score   support

negatif    0.50      0.05      0.09         39
netral     1.00      0.04      0.07         27
positif    0.88      1.00      0.93        464

accuracy    0.88         530
macro avg   0.79         530
weighted avg 0.86         530
  
```

Gambar 8. Output Evaluasi Model

Selanjutnya dilakukan pembuatan *Confusion Matrix* untuk mengevaluasi hasil prediksi model *K-Nearest Neighbors (KNN)*. *Confusion Matrix* digunakan untuk membandingkan antara label aktual dan label hasil prediksi, sehingga dapat diketahui jumlah data yang berhasil diprediksi dengan benar maupun yang salah pada setiap kategori sentimen.



Gambar 9. Output Confusion Matrix

Berdasarkan *Confusion Matrix* yang ditampilkan pada Gambar 3.20, model KNN melakukan klasifikasi terhadap tiga kelas sentimen, yaitu positif, netral, dan negatif. Hasil evaluasi menunjukkan bahwa model memiliki kemampuan terbaik dalam mengenali kelas netral, dengan sebanyak 11 data berhasil diprediksi dengan benar, meskipun masih terdapat beberapa kesalahan prediksi ke kelas positif dan negatif. Untuk kelas negatif, sebagian besar data juga dapat diprediksi dengan benar sebanyak 7 data, namun terdapat beberapa yang salah terklasifikasi sebagai netral maupun positif. Sementara itu, performa model dalam mengenali kelas positif tergolong kurang baik, karena tidak ada data positif yang teridentifikasi dengan benar; sebagian besar justru diprediksi sebagai kelas netral.

KESIMPULAN

Berdasarkan penelitian tentang penerapan algoritma *K-Nearest Neighbors* (KNN) dalam analisis sentimen ulasan pelanggan terhadap produk pakaian anak di TikTok Shop, dapat disimpulkan bahwa KNN efektif dalam mengklasifikasikan ulasan ke dalam tiga kategori sentimen: positif, netral, dan negatif. Proses *preprocessing*—meliputi *case folding*, *stemming*, penghapusan *stopword*, normalisasi teks, serta pelabelan data—terbukti signifikan dalam meningkatkan kualitas data sebelum klasifikasi. Representasi teks ke dalam bentuk numerik melalui metode TF-IDF juga terbukti sesuai untuk kebutuhan pemrosesan oleh algoritma KNN. Hasil evaluasi model menunjukkan performa yang cukup baik secara keseluruhan, dengan akurasi mencapai 87,7%, presisi 85,8%, *recall* 87,7%, dan *F1-score* sebesar 82,8%. Analisis lebih lanjut melalui laporan klasifikasi mengungkapkan bahwa kinerja model dipengaruhi oleh ketidakseimbangan distribusi kelas; kategori positif yang dominan menghasilkan nilai *recall* sempurna (1,00) dan *F1-score* tertinggi (0,93). Dengan demikian, penelitian ini menunjukkan bahwa KNN merupakan pendekatan yang layak dan efektif untuk menganalisis persepsi pelanggan melalui ulasan teks. Temuan ini dapat dimanfaatkan oleh penjual untuk memahami preferensi konsumen, meningkatkan kualitas produk, serta merancang strategi pemasaran yang lebih responsif dan tepat sasaran.

Analisis sentimen berbasis algoritma *K-Nearest Neighbor* (KNN) dalam penelitian ini memiliki sejumlah keterbatasan. Salah satunya adalah sensitivitasnya terhadap volume data yang besar, serta performa yang menurun ketika menghadapi ketidakseimbangan distribusi kelas di mana kelas mayoritas cenderung lebih dominan dalam prediksi, sehingga mengurangi kemampuan model dalam mengenali kelas minoritas. Mengingat hal tersebut, penelitian lanjutan disarankan untuk tidak hanya mengandalkan KNN, melainkan juga mengeksplorasi algoritma pembelajaran mesin alternatif, seperti *Support Vector Machines* (SVM), yang berpotensi memberikan akurasi dan efektivitas yang lebih baik dalam tugas klasifikasi sentimen. Selain itu, pengembangan antarmuka pengguna (*user interface*) dapat menjadi arah pengembangan selanjutnya, sehingga hasil analisis sentimen dapat disajikan secara interaktif, intuitif, dan mudah diakses oleh pengguna akhir.

DAFTAR PUSTAKA

- [1] Sasabone, L., Sudarmanto, E., Yovita, Y., & Adiwijaya, S. (2023). *Pengaruh E-commerce dan Kemudahan Transaksi Terhadap Perubahan Pola Konsumsi Dalam Era Digital Di Indonesia*. *Sanskara Ilmu Sosial dan Humaniora*, 1(01), 32-42. <https://doi.org/10.58812/sish.v1i01.304>
- [2] Vinne, V., Sinurat, D. U., & Prasetianti, Y. (2025). *Penerapan Algoritma K-Nearest Neighbors (KNN) Dalam Menganalisis Sentimen Ulasan Aplikasi Seabank Pada Google Play Store*. *Journal of Information Systems Management and Digital Business*, 2(2), 103-113. <https://doi.org/10.70248/jismdb.v2i2.1508>
- [3] Arrasyid, R. M., Putera, D. E., & Yusuf, A. Y. P. (2024). *Analisis Sentimen Review Pembelian Produk Di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing*. *Jurnal Tekno Kompak*, 18(2), 319-330. <https://doi.org/10.33365/jtk.v18i2.3813>

- [4] Faruk, K. O., Rahman, A., Shusmita, S. A., Awlad, M. S. I., Das, P., Mehedi, M. H. K., ... & Rasel, A. A. (2022, July). *K Nearest Neighbour Collaborative Filtering For Expertise Recommendation Systems*. In *International Symposium on Distributed Computing and Artificial Intelligence* (pp. 187-196). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-20859-1_19
- [5] Nurhidayat, R., & Dewi, K. E. (2023). *Penerapan Algoritma K-Nearest Neighbor Dan Fitur Ekstraksi N-Gram Dalam Analisis Sentimen Berbasis Aspek*. *Komputa: Jurnal Ilmiah Komputer dan Informatika*, 12(1), 91-100. <https://doi.org/10.34010/komputa.v12i1.9458>
- [6] Gifari, O. I., Adha, M., Hendrawan, I. R., & Durrand, F. F. S. (2022). *Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine*. *Journal of Information Technology*, 2(1), 36-40. <https://doi.org/10.46229/jifotech.v2i1.330>
- [7] Yasin, W., Harimoorthy, K., & Vetriveeran, D. (2023, June). *Identification of Consumer Buying Patterns using KNN in E-Commerce Applications*. In *2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)* (pp. 187-192). IEEE. <https://doi.org/10.1109/ICPCSN58827.2023.00036>
- [8] Ramadhansyah, D., Asrofiq, A., & Yunefri, Y. (2024). *Analisis Sentimen Ulasan Penumpang Maskapai Penerbangan Di Indonesia Dengan Algoritma Random Forest Dan KNN*. *Zonasi: Jurnal Sistem Informasi*, 6(2), 287-297. <https://doi.org/10.54914/jit.v10i1.1054>
- [9] Saifurridho, M., Martanto, M., & Hayati, U. (2024). *Analisis Algoritma K-Nearest Neighbor terhadap Sentimen Pengguna Aplikasi Shopee*. *Jurnal Informatika Terpadu*, 10 (1), 21-26. <https://doi.org/10.54914/jit.v10i1.1054>
- [10] Putri, N. L., Warsito, B., & Surarso, B. (2024). *Pengaruh Klasifikasi Sentimen Pada Ulasan Produk Amazon Berbasis Rekayasa Fitur dan K-Nearest Neighbor*. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(1), 65-74. <https://doi.org/10.25126/jtiik.20241117376>