

Analisis Sentimen Komentar Pengguna Aplikasi AlloFresh Di Google Playstore Menggunakan Metode Support Vector Machine (SVM)

Muchamad Sandy¹, Eri Dariato²

Prodi Teknik Informatikas, Universitas Dian Nusantara, Jakarta Barat, Indonesia

ABSTRACT (Nunito 9 pt)

Keywords:

Analisis Sentimen,
Aplikasi Allofresh,
Google Playstore,
SVM.

Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi AlloFresh dengan memanfaatkan teknik Text Mining dan pembelajaran mesin. Data yang digunakan berupa ulasan konsumen yang dikumpulkan dari platform digital, kemudian melalui serangkaian tahap pra-pemrosesan, meliputi case folding, cleaning, tokenizing, stopword removal, normalisasi, dan stemming. Selanjutnya dilakukan pelabelan sentimen ke dalam kategori positif dan negatif, sementara ulasan netral dikecualikan agar fokus analisis lebih terarah. Representasi teks kemudian dikonversi ke bentuk numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) untuk memberikan bobot pada setiap kata berdasarkan frekuensi dan tingkat kepentingannya dalam keseluruhan korpus. Dalam tahap klasifikasi, algoritma Support Vector Machine (SVM) digunakan untuk mengukur performa model. Hasil pengujian menunjukkan bahwa SVM mampu memberikan tingkat akurasi yang cukup tinggi dengan capaian mendekati 95%. Hal ini menunjukkan bahwa SVM efektif dalam mengklasifikasikan ulasan pengguna, khususnya dalam mengidentifikasi sentimen positif maupun negatif.

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



Corresponding Author:

Muchamad Sandy
Universitas Dian Nusantara
Jakarta Barat
muchamad.sandy@dosen.undira.ac.id

INTRODUCTION

Kemajuan teknologi digital dalam beberapa tahun terakhir telah memicu perubahan besar dalam pola hidup masyarakat, khususnya dalam mengakses informasi dan memanfaatkan layanan berbasis internet (Ghani Ardiansyah et al., 2025). Aplikasi digital kini menjadi sarana utama dalam memenuhi kebutuhan sehari-hari, termasuk dalam aktivitas pemesanan bahan makanan serta barang keperluan rumah tangga (Melisa Nur Aini et al., 2024). Salah satu platform yang menawarkan kemudahan ini adalah AlloFresh, yang hadir sebagai solusi pemesanan instan melalui perangkat mobile dengan sistem yang efisien dan praktis (Tarigan et al., 2025a).

Seiring meningkatnya adopsi aplikasi digital oleh masyarakat, ulasan dan tanggapan dari pengguna menjadi sumber informasi yang signifikan (Rahman et al., 2024). Umpan balik berupa komentar yang ditinggalkan pada platform seperti Google Play Store tidak hanya menunjukkan kepuasan atau ketidakpuasan individu terhadap layanan, tetapi juga menggambarkan persepsi kolektif pengguna terhadap kualitas aplikasi (Wafi et al., 2025). Oleh karena itu, mengolah dan memanfaatkan ulasan pengguna secara optimal dapat memberikan masukan strategis bagi pengembang untuk meningkatkan mutu layanan (Tarigan et al., 2025b).

Namun demikian, besarnya volume komentar yang masuk setiap hari menimbulkan tantangan tersendiri dalam proses analisis (Data et al., 2025). Pendekatan manual dinilai tidak efisien dan membutuhkan waktu yang cukup lama, terutama ketika diperlukan identifikasi pola opini atau

kecenderungan sentimen dalam skala besar (Ahmed & Wang, 2023). Untuk mengatasi hal tersebut, pendekatan berbasis teknologi seperti analisis sentimen menjadi pilihan yang relevan dan efektif.

Analisis sentimen, yang merupakan bagian dari bidang Natural Language Processing (NLP), bertujuan untuk mengidentifikasi serta menginterpretasi makna emosional dari suatu teks (Sentimen et al., 2025). Dengan menggunakan metode ini, data dalam jumlah besar dapat dianalisis dalam waktu yang relatif singkat. Dalam konteks evaluasi aplikasi, teknik ini mampu memberikan gambaran umum mengenai persepsi publik serta mengidentifikasi aspek yang perlu ditingkatkan (Giovani et al., 2020).

Salah satu metode klasifikasi yang banyak digunakan dalam analisis sentimen adalah Support Vector Machine (SVM) (Syamia & Furqan, 2025). Algoritma ini bekerja dengan menentukan garis batas (hyperplane) terbaik yang dapat memisahkan data ke dalam beberapa kategori. SVM dikenal memiliki performa yang baik dalam menangani data berdimensi tinggi dan distribusi yang kompleks, menjadikannya sangat cocok untuk pengolahan teks dan klasifikasi opini (Dewi, 2022).

Penelitian ini menerapkan algoritma SVM untuk menganalisis sentimen dari komentar pengguna aplikasi AlloFresh yang diperoleh dari Google Play Store. Proses analisis dilakukan melalui beberapa tahap, yaitu pra-proses data, ekstraksi fitur, serta pelatihan model klasifikasi. Tujuan utama penelitian ini adalah untuk mengevaluasi kecenderungan opini pengguna terhadap aplikasi AlloFresh serta menguji efektivitas SVM dalam mengklasifikasikan komentar ke dalam kategori sentimen yang relevan, yaitu positif, negatif, atau netral.

METHODS

Analisis dan Perancangan Sistem

Penelitian ini berfokus pada perancangan dan implementasi sistem analisis sentimen berbasis algoritma Support Vector Machine (SVM) yang ditujukan untuk mengklasifikasikan opini pengguna berbahasa Indonesia pada ulasan aplikasi di Google Play Store. Sistem ini bertujuan untuk menilai efektivitas SVM dalam mengidentifikasi kecenderungan sentimen pengguna terhadap aplikasi tertentu. Tahapan pengembangan sistem diawali dengan proses pelatihan model klasifikasi yang memanfaatkan data yang telah melalui serangkaian proses pra-pemrosesan. Model yang dihasilkan kemudian digunakan untuk mengelompokkan ulasan ke dalam kategori sentimen positif dan negatif. Sebelum proses klasifikasi dilakukan, data ulasan menjalani beberapa tahap pra-pemrosesan, termasuk normalisasi teks (case folding), penghapusan kata umum yang tidak membawa makna khusus (stopwords), penghilangan tanda baca, serta penandaan khusus pada kalimat yang mengandung indikasi sentimen negatif.

Metodologi Pengumpulan Data

Sumber data dalam penelitian ini berasal dari ulasan pengguna aplikasi Duolingo yang diperoleh melalui platform Google Play Store. Pengumpulan data dilakukan secara otomatis menggunakan teknik web scraping untuk meningkatkan efisiensi proses ekstraksi. Data yang terkumpul kemudian diproses dalam lingkungan Google Colaboratory dengan bantuan pustaka TensorFlow. Dataset disimpan dalam format spreadsheet di media penyimpanan lokal atau sistem basis data sebelum dilanjutkan ke tahap pra-pemrosesan lanjutan.

Tahapan Pra-Pemrosesan Data

Tahap ini dilakukan untuk memastikan data berada dalam kondisi bersih, seragam, serta siap diproses oleh algoritma pembelajaran mesin. Prosedur yang diterapkan meliputi:

- a) Case folding: menyeragamkan teks dengan mengubah seluruh huruf menjadi huruf kecil (Darwis et al., 2021).

- b) **Cleaning:** menghapus elemen yang tidak relevan seperti URL, angka, tanda baca, maupun karakter khusus (Petiwi et al., 2022).
- c) **Tokenizing:** memecah kalimat menjadi potongan kata (token) agar mudah dianalisis.
- d) **Stopword removal:** menghilangkan kata-kata umum yang tidak memberi pengaruh signifikan terhadap makna, seperti kata hubung atau kata bantu (Tanggraeni & Sitokdana, 2022).
- e) **Normalisasi:** mengganti kata tidak baku atau slang dengan bentuk standar sesuai kaidah bahasa (Herlinawati et al., 2020).
- f) **Stemming:** mengembalikan kata ke bentuk dasar dengan cara menghapus imbuhan (awalan atau akhiran) untuk menyederhanakan variasi kata (Ilmawan & Mude, 2020).

Pelabelan Sentimen

Dalam proses pelatihan model, hanya ulasan dengan sentimen positif dan negatif yang digunakan, sedangkan komentar netral dikeluarkan setelah melalui tahap identifikasi polaritas.

TF-IDF (Term Frequency–Inverse Document Frequency)

Tahap selanjutnya adalah pembobotan istilah yang berfungsi mengubah teks menjadi representasi numerik agar dapat diproses lebih efektif oleh algoritma machine learning. Pada penelitian ini digunakan metode TF-IDF Vectorizer dari pustaka scikit-learn, yang memberikan bobot pada kata berdasarkan frekuensi kemunculannya dalam dokumen sekaligus mempertimbangkan tingkat kepentingannya dibandingkan dengan seluruh korpus. Dengan pendekatan ini, kata-kata yang jarang muncul namun memiliki makna penting memperoleh bobot lebih tinggi.

Pemodelan Data

Setelah tahap pra-pemrosesan, langkah selanjutnya adalah membangun model klasifikasi sentimen menggunakan dua algoritma pembelajaran mesin, yakni Naive Bayes dan Support Vector Machine (SVM). Pelatihan dan pengujian model dilakukan dengan metode 10-fold cross-validation, di mana dalam setiap iterasi, satu bagian data digunakan sebagai data uji, sementara sembilan bagian sisanya digunakan sebagai data latih. Tujuan utama dari proses ini adalah untuk mengidentifikasi model dengan performa terbaik dalam mengklasifikasikan ulasan ke dalam kategori sentimen positif atau negatif. Seluruh proses pemodelan dilakukan dalam lingkungan Google Colaboratory dengan bahasa pemrograman Python dan dukungan pustaka TensorFlow.

Evaluasi Kinerja Model

Evaluasi model bertujuan untuk mengukur kinerja serta kemampuan generalisasi dari model klasifikasi yang telah dikembangkan. Dua metrik utama yang digunakan dalam evaluasi adalah tingkat akurasi (accuracy) dan nilai Area Under the Curve (AUC). Evaluasi dilakukan dengan pendekatan 10-fold cross-validation untuk menghasilkan estimasi performa yang lebih objektif dan representatif. Fokus utama evaluasi adalah menilai seberapa baik model dapat membedakan antara kategori sentimen positif dan negatif serta memberikan prediksi yang akurat terhadap data ulasan yang belum pernah dilihat sebelumnya.

RESULTS AND DISCUSSION

Pengumpulan dan Pengolahan Data Awal

Dalam penelitian ini, digunakan kumpulan data yang terdiri atas 50.000 komentar pengguna, yang diambil dari ulasan terhadap aplikasi dompet digital Allofresh di platform Google Play Store. Pengambilan data dilakukan secara otomatis menggunakan teknik web crawling yang dibangun dengan pemrograman Python. Setelah proses akuisisi selesai, dilakukan tahapan prapemrosesan

Table 2. Hasil Case Folding

Sebelum	Sesudah
produknya banyak pilihan, kualitasnya terjamin dan pengirimannya satset øÿ'øÿ'øÿ'	produknya banyak pilihan, kualitasnya terjamin dan pengirimannya satset øÿ'øÿ'øÿ'
Kecewa padahal mau beli barang tetapi alamat saya diluar jangkauan padahal di keterangan ada area Jabodetabek dan saya berada di wilayah tersebut, tolong perluas area pengiriman biar bisa pesan barang di mana saja seperti aplikasi belanja pada umumnya	kecewa padahal mau beli barang tetapi alamat saya diluar jangkauan padahal di keterangan ada area jabodetabek dan saya berada di wilayah tersebut, tolong perluas area pengiriman biar bisa pesan barang di mana saja seperti aplikasi belanja pada umumnya
saya suruh menginstal apk ini , nanti di sawer dangan mengasih kan no dana.. tapi itu bohong, apk ini telah bekarja sama dengan orang penipuan	saya suruh menginstal apk ini , nanti di sawer dangan mengasih kan no dana.. tapi itu bohong, apk ini telah bekarja sama dengan orang penipuan
bagus bngt	bagus bngt
Sudah gk bisa dipake	sudah gk bisa dipake

Pembersihan Teks (Cleaning)

Tahapan ini mencakup penghapusan elemen-elemen yang tidak berkontribusi terhadap analisis, seperti tanda baca, karakter non-alfabetik, simbol, serta kata-kata yang tidak relevan dengan konteks sentimen. Hasil Sampel Cleaning Text dapat dilihat pada Table 3.

Table 3. Hasil Pembersihan Text (Cleaning)

Sebelum	Sesudah
produknya banyak pilihan, kualitasnya terjamin dan pengirimannya satset øÿ'øÿ'øÿ'	produknya banyak pilihan kualitasnya terjamin dan pengirimannya satset
kecewa padahal mau beli barang tetapi alamat saya diluar jangkauan padahal di keterangan ada area jabodetabek dan saya berada di wilayah tersebut, tolong perluas area pengiriman biar bisa pesan barang di mana saja seperti aplikasi belanja pada umumnya	kecewa padahal mau beli barang tetapi alamat saya diluar jangkauan padahal di keterangan ada area jabodetabek dan saya berada di wilayah tersebut tolong perluas area pengiriman biar bisa pesan barang di mana saja seperti aplikasi belanja pada umumnya
saya suruh menginstal apk ini , nanti di sawer dangan mengasih kan no dana.. tapi itu bohong, apk ini telah bekarja sama dengan orang penipuan	saya suruh menginstal apk ini nanti di sawer dangan mengasih kan no dana tapi itu bohong apk ini telah bekarja sama dengan orang penipuan

bagus bngt	bagus bngt
sudah gk bisa dipake	sudah gk bisa dipake

Tokenisasi

Tokenisasi merupakan proses segmentasi teks menjadi unit-unit kata (tokens), yang mempermudah proses ekstraksi fitur dan analisis lanjutan pada tahap klasifikasi. Hasil sampel tokenizing dapat dilihat pada Table 4.

Table 4. Hasil Tokenisasi

Sebelum	Sesudah
produknya banyak pilihan kualitasnya terjamin dan pengirimannya satset	produknya', 'banyak', 'pilihan', 'kualitasnya', 'terjamin', 'dan', 'pengirimannya', 'satset'
kecewa padahal mau beli barang tetapi alamat saya diluar jangkauan padahal di keterangan ada area jabodetabek dan saya berada di wilayah tersebut tolong perluas area pengiriman biar bisa pesan barang di mana saja seperti aplikasi belanja pada umumnya	'kecewa', 'padahal', 'mau', 'beli', 'barang', 'tetapi', 'alamat', 'saya', 'diluar', 'jangkauan', 'padahal', 'di', 'keterangan', 'ada', 'area', 'jabodetabek', 'dan', 'saya', 'berada', 'di', 'wilayah', 'tersebut', 'tolong', 'perluas', 'area', 'pengiriman', 'biar', 'bisa', 'pesan', 'barang', 'di', 'mana', 'saja', 'seperti', 'aplikasi', 'belanja', 'pada', 'umumnya'
saya suruh menginstal apk ini nanti di sawer dangan mengasih kan no dana tapi itu bohong apk ini telah bekarja sama dengan orang penipuan	'saya', 'suruh', 'menginstal', 'apk', 'ini', 'nanti', 'di', 'sawer', 'dangan', 'mengasih', 'kan', 'no', 'dana', 'tapi', 'itu', 'bohong', 'apk', 'ini', 'telah', 'bekarja', 'sama', 'dengan', 'orang', 'penipuan'
bagus bngt	'bagus', 'bngt'
sudah gk bisa dipake	'sudah', 'gk', 'bisa', 'dipake'

Penghapusan Stopword

Pada tahap ini, dilakukan eliminasi terhadap kata-kata umum seperti konjungsi atau preposisi yang frekuensinya tinggi namun tidak memiliki makna signifikan dalam proses analisis sentimen. Hasil sampel stopwords removal dapat dilihat pada Table 5.

Table 5. Hasil Stopword

Sebelum	Sesudah
produknya banyak pilihan kualitasnya terjamin dan pengirimannya satset	produknya banyak pilihan kualitasnya terjamin pengirimannya satset
kecewa padahal mau beli barang tetapi alamat saya diluar jangkauan padahal di keterangan ada area jabodetabek dan saya berada di wilayah tersebut tolong perluas area umumnya	kecewa mau beli barang alamat diluar jangkauan keterangan area jabodetabek berada wilayah perluas area pengiriman biar bisa pesan barang aplikasi belanja umumnya

pengiriman biar bisa pesan barang di mana saja seperti aplikasi belanja pada umumnya

saya suruh menginstal apk ini nanti di sawer dengan mengasih kan no dana tapi itu bohong apk ini telah bekarja sama dengan orang penipuan

bagus bngt

sudah gk bisa dipake

suruh menginstal apk nanti sawer dengan mengasih no dana bohong apk bekarja sama orang penipuan

bagus bngt

sudah gk bisa dipake

Normalisasi

Normalisasi bertujuan untuk menyeragamkan kata-kata yang tidak baku atau memiliki ejaan tidak standar, agar konsistensi data tetap terjaga dan representasi kata menjadi seragam. Hasil sampel normalisasi dapat dilihat pada Table 6.

Table 6. Hasil Normalisasi

Sebelum	Sesudah
produknya banyak pilihan kualitasnya terjamin pengirimannya satset	produk banyak pilihan kualitas terjamin pengiriman satset
kecewa mau beli barang alamat diluar jangkauan keterangan area jabodetabek berada wilayah perluas area pengiriman biar bisa pesan barang aplikasi belanja umumnya	kecewa mau beli barang alamat diluar jangkauan keterangan area jabodetabek berada wilayah perluas area pengiriman biar bisa pesan barang aplikasi belanja umum
suruh menginstal apk nanti sawer dengan mengasih no dana bohong apk bekarja sama orang penipuan	suruh instal aplikasi nanti sawer dengan memberi nomor dana bohong aplikasi bekerja sama orang penipuan
bagus bngt	bagus banget
sudah gk bisa dipake	sudah tidak bisa dipakai

Stemming

Langkah ini bertujuan mengembalikan setiap kata ke bentuk dasarnya dengan menghapus imbuhan, baik awalan, sisipan, maupun akhiran. Proses ini penting untuk mengurangi keragaman morfologis kata dan meningkatkan efisiensi dalam representasi teks saat proses klasifikasi dilakukan. Hasil sampel stemming dapat dilihat pada Table 7.

Table 7. Hasil Stemming

Sebelum	Sesudah
produk banyak pilihan kualitas terjamin pengiriman satset	produk banyak pilih kualitas jamin kirim satset

kecewa mau beli barang alamat diluar jangkauan keterangan area jabodetabek berada wilayah perluas area pengiriman biar bisa pesan barang aplikasi belanja umum	kecewa mau beli barang alamat luar jangkau terang area jabodetabek ada wilayah luas area kirim biar bisa pesan barang aplikasi belanja umum
suruh instal aplikasi nanti sawer dengan memberi nomor dana bohong aplikasi bekerja sama orang penipuan	suruh instal aplikasi nanti sawer beri nomor dana bohong aplikasi kerja sama orang tipu
bagus banget	bagus banget
sudah tidak bisa dipakai	sudah tidak bisa pakai

Hasil Pelabelan Sentimen

Setelah tahapan pra pemrosesan selesai, langkah selanjutnya adalah memberi label pada pola data. Pada studi ini, pelabelan dilakukan dengan analisis sentimen terhadap seluruh komentar dalam dataset, yang diklasifikasikan ke tiga kategori: positif, netral, dan negatif. Sampel hasil pelabelan ditampilkan pada Figure 1.

	content	sentimen
0	produk banyak pilih kualitas jamin kirim satset	Netral
1	kecewa mau beli barang alamat luar jangkau ter...	Negatif
2	suruh instal aplikasi nanti sawer beri nomor d...	Netral
3	bagus banget	Positif
4	sudah tidak bisa pakai	Netral

Figure 1. Hasil Sampel Pelabelan Data

Hasil TF-IDF (Term Frequency–Inverse Document Frequency)

Hasil perhitungan TF-IDF (Term Frequency–Inverse Document Frequency) pada Figure menunjukkan bobot kepentingan masing-masing kata dalam kumpulan dokumen. Kata “allofresh” menempati posisi teratas dengan nilai tf-idf sebesar 0.126325 frekuensinya 616 tidak jauh berbeda dari kata lain. Selanjutnya, kata “allofriends” dan “mohon” memiliki nilai tf-idf yang mirip sekitar 0.107. Kata “hai” sedikit lebih rendah dengan tf-idf 0.106943 dengan jumlah kemunculannya 479 hampir sama dengan “mohon”. Kata “maaf” memiliki frekuensi 365 dan idf 3.359302, sehingga bobot tf-idf-nya turun menjadi 0.088665.

	term	frequency	idf	tf	tfidf
0	allofresh	616	2.835952	0.044544	0.126325
1	allofriends	482	3.081255	0.034854	0.107395
2	mohon	481	3.083332	0.034782	0.107244
3	hai	479	3.087499	0.034637	0.106943
4	maaf	365	3.359302	0.026394	0.088665

Figure 2. Proses Perhitungan TF-IDF

Hasil Permodelan Data

Figure 3. menunjukkan hasil evaluasi model SVM (Support Vector Machine) pada klasifikasi sentimen menunjukkan performa yang cukup baik secara keseluruhan, dengan tingkat akurasi 95%. Pada kelas negative, model memiliki performa yang sangat tinggi dengan precision 0.95, recall 0.99, dan f1-score 0.97. Untuk kelas positive precision (0.97), recall 0.78 dan f1-score 0.86. Kelas neutral memiliki nilai precision, recall, dan f1-score 0.00. Nilai macro average f1-score sebesar 0.61 dan nilai weighted average f1-score 0.94. Hasil permodelan data dapat dilihat pada Figure 4.

```
SVM - Classification Report:
              precision    recall  f1-score   support

negative      0.95         0.99         0.97         497
neutral       0.00         0.00         0.00           4
positive      0.97         0.78         0.86         109

accuracy              0.95         610
macro avg             0.64         0.59         0.61         610
weighted avg          0.94         0.95         0.94         610
```

Figure 3. Hasil Pengujian SVM

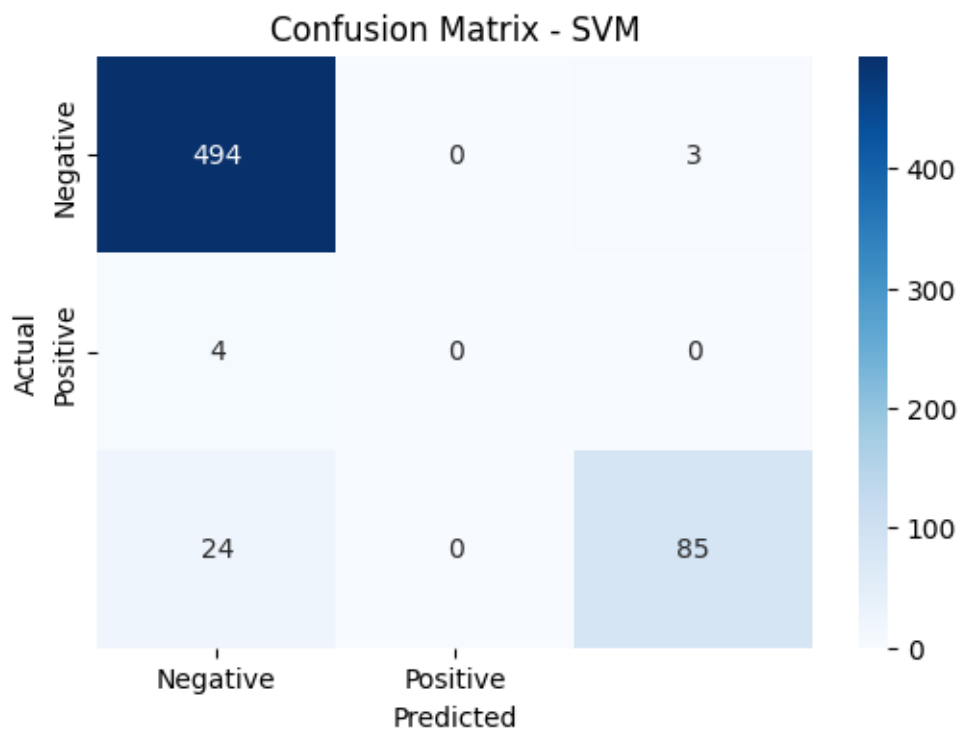


Figure 4. Confusion Matrix SVM

Hasil Evaluasi Kinerja Model

Hasil evaluasi akurasi cross-validation pada model SVM dengan 10 fold menunjukkan bahwa performa model relatif konsisten dengan kisaran nilai antara 0.910 hingga 0.938. Pada fold pertama, akurasi berada di titik terendah yaitu sekitar 0.910, kemudian meningkat tajam pada fold kedua hingga

mencapai 0.938 yang merupakan nilai tertinggi. Setelah itu, akurasi menurun pada fold ketiga (0.922) dan keempat (0.914), lalu kembali stabil pada fold kelima hingga ketujuh dengan nilai yang sama yaitu 0.922. Fold kedelapan memperlihatkan peningkatan signifikan dengan akurasi 0.934, namun menurun kembali pada fold kesembilan (0.914), sebelum akhirnya naik lagi pada fold kesepuluh dengan nilai 0.926. Secara keseluruhan, model SVM mencapai rata-rata akurasi sebesar 92.24% dengan standar deviasi sekitar 0.83%, yang menunjukkan bahwa performa model cukup stabil dengan fluktuasi yang relatif kecil antar fold.

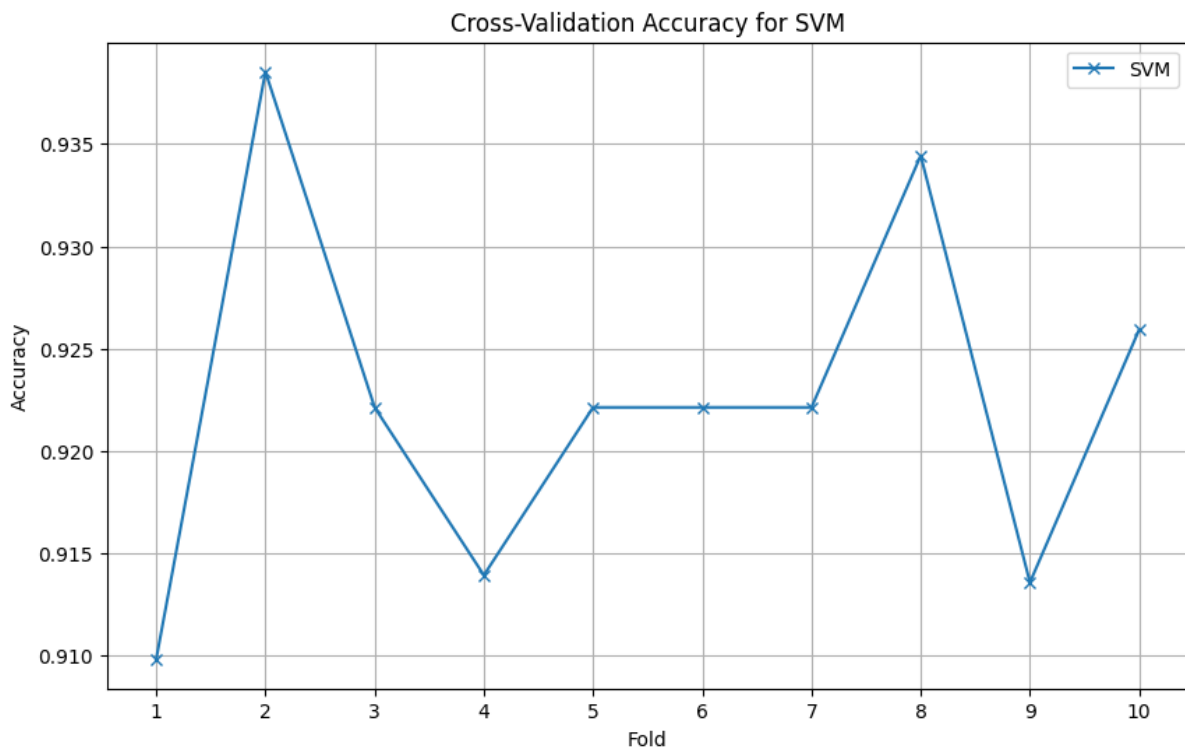


Figure 5. Hasil Akurasi Cross-Validation Model SVM Pada 10 fold.

CONCLUSION

Penelitian ini berhasil menerapkan analisis sentimen pada ulasan pengguna aplikasi AlloFresh dengan menggunakan metode TF-IDF untuk representasi fitur teks dan algoritma Support Vector Machine (SVM) untuk klasifikasi. Hasil TF-IDF menunjukkan bahwa kata “allofresh” menjadi term dengan bobot tertinggi (0.126), diikuti oleh kata “allofriends” dan “mohon” yang memiliki bobot hampir serupa. Hal ini mengindikasikan bahwa kata-kata tersebut merupakan kata kunci penting dalam interaksi pengguna dengan aplikasi.

Model SVM yang dibangun menunjukkan performa yang cukup tinggi dengan akurasi keseluruhan sebesar 95%. Pada kelas negatif, model menunjukkan kinerja terbaik dengan precision 0.95, recall 0.99, dan F1-score 0.97. Untuk kelas positif, nilai precision cukup tinggi (0.97) namun recall lebih rendah (0.78) dengan F1-score 0.86. Sebaliknya, performa pada kelas netral sangat rendah dengan semua metrik evaluasi (precision, recall, F1-score) bernilai 0.00. Evaluasi cross-validation dengan 10 fold menghasilkan rata-rata akurasi 92.24% dengan standar deviasi 0.83%, yang berarti performa model cukup stabil meskipun masih ada fluktuasi antar fold.

Penelitian selanjutnya sebaiknya memperhatikan distribusi data yang tidak seimbang, khususnya pada kelas netral, dengan menerapkan teknik seperti oversampling atau SMOTE.

Perbandingan dengan algoritma lain, termasuk model deep learning, juga dapat dilakukan untuk meningkatkan akurasi klasifikasi. Selain itu, penggunaan teknik lanjutan seperti lemmatization dan word embedding diharapkan mampu memberikan representasi teks yang lebih kontekstual.

REFERENCE

- Ahmed, Z., & Wang, J. (2023). A fine-grained deep learning model using embedded-CNN with BiLSTM for exploiting product sentiments. *Alexandria Engineering Journal*, 65, 731–747. <https://doi.org/10.1016/j.aej.2022.10.037>
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131. <https://doi.org/10.33365/jtk.v15i1.744>
- Data, A., Dengan, M., Bayes, N., Klasifikasi, U., Penerima, K., Iuran, B., Kesehatan, J., Desa, P., El, D.-D., Purba, R., Al, E., El, D., & Siboro, Y. O. (2025). Analisis Kelayakan Penerima Bantuan Iuran Jaminan Kesehatan Menggunakan Teknik Klasifikasi Data Mining dengan Metode Naive Bayes. 07(02), 2721–1800. <https://doi.org/10.54209/jatilima.v7i02.1279>
- Dewi, A. K. (2022). Analisis Sentimen Ekspedisi Sicepat Dari Ulasan Google Play Menggunakan Algoritma Naive Bayes. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 796–805. <https://doi.org/10.35957/jatisi.v9i2.1802>
- Ghani Ardiansyah, A., Hidayati, N., Sari Sinaga, N., Criston Purba, D., Yeng Sinaga, E., Tunas Bangsa Pematangsiantar, S., Ghani Ardiansyah STIKOM Tunas Bangsa Pematang Siantar, A., & Sumatra, N. (2025). Design of Building Shop Information System Abghan22 Web-Based Using The Naive Bayes. *Jatilima: Journal of Multimedia and Information Technology*, 07. <https://doi.org/10.54209/jatilima.v7i01.978>
- Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. *Jurnal Teknoinfo*, 14(2), 115. <https://doi.org/10.33365/jti.v14i2.679>
- Herlinawati, N., Yuliani, Y., Faizah, S., Gata, W., & Samudi, S. (2020). Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naive Bayes dan Support Vector Machine. *CESS (Journal of Computer Engineering, System and Science)*, 5(2), 293. <https://doi.org/10.24114/cess.v5i2.18186>
- Ilmawan, L. B., & Mude, M. A. (2020). Perbandingan Metode Klasifikasi Support Vector Machine dan Naive Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store. *ILKOM Jurnal Ilmiah*, 12(2), 154–161. <https://doi.org/10.33096/ilkom.v12i2.597.154-161>
- Melisa Nur Aini, Rita Yulfani, & Nurul Jariah. (2024). Penerapan Metode Naive Bayes Untuk Analisis Sentimen Ulasan Produk Sunscreen Berdasarkan Female Daily Review. *Jurnal Multimedia Dan Teknologi Informasi (Jatilima)*, 6(01), 24–34. <https://doi.org/10.54209/jatilima.v6i01.421>
- Petiwi, M. I., Triayudi, A., & Sholihati, I. D. (2022). Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naive Bayes dan Support Vector Machine. *Jurnal Media Informatika Budidarma*, 6(1), 542. <https://doi.org/10.30865/mib.v6i1.3530>
- Rahman, Z., Sakinah, P., Hendra, Y., Satria, B., Maulana, F., & Ayun, A. Q. (2024). Sentiment Analysis of Gojek App Reviews on Google Play Store with Natural Language Processing Using Naive Bayes' Algorithm-Zumardi Rahman et.al Sentiment Analysis of Gojek App Reviews on Google Play Store with Natural Language Processing Using Naive Bayes' Algorithm. 06. <https://doi.org/10.54209/jatilima.v6i03.1189>
- Sentimen, A., Pelanggan, U., Hokben, A., Google, D., Menggunakan, P., Support, M., Vikri, R., Al, A. E., Alif, R. V., & Hasugian, A. H. (2025). Analisis Sentimen Ulasan Pelanggan Aplikasi Hokben Di

- Google Playstore Menggunakan Metode Support Vector Machine.* 07(02).
<https://doi.org/10.54209/jatilima.v7i02.1501>
- Syamia, N., & Furqan, M. (2025). *Sentiment Analysis On The Change Of Coach Of The Indonesian National Team Using Support Vector Machine Algorithm : A Case Study Of Twitter Data-Nanda Syamia et.al Sentiment Analysis On The Change Of Coach Of The Indonesian National Team Using Support Vector Machine Algorithm : A Case Study Of Twitter Data.* 07(02).
<https://doi.org/10.54209/jatilima.v7i02.1510>
- Tanggraeni, A. I., & Sitokdana, M. N. N. (2022). Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 785–795. <https://doi.org/10.35957/jatisi.v9i2.1835>
- Tarigan, D. A., Situmorang, Z., & Rosnelly, R. (2025a). *ANALISIS SENTIMEN APLIKASI PLAYSTORE SIREKAP 2024 PASCA PILPRES DENGAN PERBANDINGAN METODE SUPPORT VECTOR MACHINE (SVM), NAÏVE BAYES CLASSIFIER DAN RANDOM FOREST.* 11(3), 661–670.
- Tarigan, D. A., Situmorang, Z., & Rosnelly, R. (2025b). *ANALISIS SENTIMEN APLIKASI PLAYSTORE SIREKAP 2024 PASCA PILPRES DENGAN PERBANDINGAN METODE SUPPORT VECTOR MACHINE (SVM), NAÏVE BAYES CLASSIFIER DAN RANDOM FOREST.* 11(3), 661–670.
- Wafi, A. Z. Al, Rochim, F. P., & Bezaleel, V. (2025). Investigating Liver Disease Machine Learning Prediction Performancethrough Various Feature Selection Methods. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 24(3), 505–518.
<https://doi.org/10.30812/matrik.v24i3.4531>