STORAGE – Jurnal Ilmiah Teknik dan Ilmu Komputer, Vol. 5 No. 3, Agustus 2026, 252 – 260
e-ISSN: 2828-5344 | DOI: 10.55123

ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI OJEK INDRALAYA MENGGUNAKAN PERBANDINGAN *SUPPORT VECTOR MACHINE* DAN *NAIVE BAYES*

[1)]Muslimin, [2)]Herri Setiawan, [3)]Dwi Asa Verano
[1,2),3)] Prodi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Indo Global Mandiri
[1)]202211059@student.uigm.ac.id, [2)]herri@uigm.ac.id, [3)]dwi.verano@gmail.com

INFO ARTIKEL

Riwayat Artikel :
Diterima : 28 April 2026
Disetujui : 23 Juli 2026

Kata Kunci :
Analisis Sentimen Berbasis Aspek, Google Playstore, Naive Bayes., Ojek Indralaya, Support Vector Machine.

ABSTRAK

Kehadiran Ojek Indralaya sebagai transportasi lokal sangat dipengaruhi oleh kepuasan pengguna di Google Playstore, namun banyaknya ulasan membuat analisis manual menjadi tidak efisien dan rentan subjektivitas. Penelitian ini menerapkan *Aspect-Based Sentiment Analysis* (ABSA) untuk memetakan sentimen pengguna (positif dan negatif) pada lima aspek layanan: harga, keamanan, kenyamanan, fitur aplikasi, dan pelayanan driver. Menggunakan metode *text mining*, sebanyak 1.023 data ulasan dikumpulkan melalui teknik *scraping* dan diolah melalui tahap *preprocessing* (termasuk mempertahankan *slang word* bernilai emosi), lalu dibobotkan menggunakan TF-IDF. Model klasifikasi dibangun dengan membandingkan performa algoritma *Support Vector Machine* (SVM) dan Naive Bayes menggunakan pembagian data 80:20. Kinerja model dievaluasi melalui metrik *Accuracy, Precision, Recall, F1-score*, dan *Confusion Matrix*. Hasil pengujian menunjukkan bahwa algoritma SVM memiliki performa yang lebih stabil dan unggul pada seluruh aspek dibandingkan Naive Bayes, dengan performa tertinggi dicapai oleh SVM pada aspek harga (*Accuracy* 95,65%). Sebaliknya, performa terendah ditemukan pada algoritma Naive Bayes pada aspek keamanan yang dipengaruhi oleh adanya *imbalanced dataset*. Penelitian ini memberikan kontribusi berupa pemetaan sentimen berbasis aspek yang spesifik sebagai rekomendasi terstruktur bagi pengembang untuk menentukan prioritas perbaikan layanan transportasi lokal.

ARTICLE INFO

Article History :
Received : Apr 28, 2026
Accepted : Jul 23, 2026

Keywords:
Aspect-Based Sentiment Analysis, Google Playstore, Naive Bayes, Ojek Indralaya, Support Vector Machine.

ABSTRACT

The popularity of Ojek Indralaya as a local transportation service is heavily influenced by user satisfaction ratings on the Google Play Store; however, the large volume of reviews makes manual analysis inefficient and prone to subjectivity. This study applies Aspect-Based Sentiment Analysis (ABSA) to map user sentiment (positive and negative) across five service aspects: price, safety, comfort, app features, and driver service. Using text mining methods, a total of 1,023 review data points were collected via scraping techniques and processed through preprocessing stages (including preserving slang words with emotional value), then weighted using TF-IDF. A classification model was built by comparing the performance of the Support Vector Machine (SVM) and Naive Bayes algorithms using an 80:20 data split. Model performance was evaluated using the metrics Accuracy, Precision, Recall, F1-score, and Confusion

Matrix. The test results showed that the SVM algorithm had more stable and superior performance across all aspects compared to Naive Bayes, with the highest performance achieved by SVM in the price aspect (Accuracy 95.65%). Conversely, the lowest performance was found in the Naive Bayes algorithm regarding the safety aspect, which was influenced by the presence of an imbalanced dataset. This study contributes aspect-specific sentiment mapping as structured recommendations for developers to determine priorities for improving local transportation services.

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah mendorong perubahan besar pada berbagai sektor kehidupan masyarakat. Hadirnya transportasi daring memberikan manfaat berupa proses pemesanan yang praktis dan transparansi tarif bagi pengguna. Di tingkat lokal, Ojek Indralaya (OJIN) hadir sebagai solusi inovatif untuk menjawab kebutuhan mobilitas masyarakat Indralaya (Anam, Saifuddin and Fitriah, 2020). Keberhasilan aplikasi ini sangat dipengaruhi oleh faktor *User Experience* (UX) yang menentukan kenyamanan dalam penggunaan sistem. UX yang optimal berkontribusi langsung terhadap peningkatan kepercayaan serta loyalitas pelanggan terhadap suatu layanan (Ningrum *et al.*, 2025).

Tingkat kepuasan tersebut dapat dievaluasi secara komprehensif melalui ulasan pengguna di platform Google Playstore (Astuti, Firmansyah and Riyadi, 2024). Volume ulasan yang sangat besar membuat proses analisis secara konvensional menjadi tidak efisien (Agustina, Subanti and Zukhronah, 2020). Oleh karena itu, diperlukan pendekatan *text mining* untuk mengekstraksi informasi penting dari data teks tidak terstruktur. Proses pengolahan teks ini melibatkan tahapan *Natural Language Processing* (NLP) agar mesin mampu mengenali konteks bahasa manusia (Rivaldi and Wismarini, 2024). Penelitian ini menerapkan *Aspect-Based Sentiment Analysis* (ABSA) untuk memetakan opini pada kategori layanan secara spesifik (Rahmajati, Saputra and Herlambang, 2025). Tinjauan sistematis menunjukkan bahwa ABSA telah banyak digunakan untuk menggali wawasan dari teks opini digital pada berbagai domain aplikasi (Hua *et al.*, 2024). Perkembangan ABSA juga terus berlanjut dengan pemanfaatan pemodelan bahasa tingkat lanjut seperti IndoBERT yang terbukti efektif dalam menangani kompleksitas bahasa Indonesia (Aryanti, Luthfiarta and Soeroso, 2025). Selain itu, pendekatan generatif berbasis *multitask learning* juga telah dikembangkan untuk menyelesaikan berbagai tugas analisis sentimen berbasis aspek secara spesifik dalam bahasa Indonesia (Suchrady and Purwarianti, 2023). ABSA pada akhirnya membantu pengembang dalam merumuskan strategi perbaikan kualitas layanan berdasarkan persepsi.

Untuk klasifikasi sentimen, algoritma *Support Vector Machine* (SVM) dipilih karena keunggulannya dalam menangani data teks berdimensi tinggi (Arifin, Enri and Sulistiyowati, 2021). Penelitian sebelumnya membuktikan bahwa algoritma SVM mampu menghasilkan kinerja yang mengesankan dalam mengklasifikasikan sentimen ulasan aplikasi di Google Playstore berdasarkan aspek-aspek tertentu (Helmayanti, Hamami and Fa'rifah, 2023). Algoritma SVM bekerja dengan mencari *hyperplane* optimal untuk memisahkan kelas data dengan tingkat akurasi tinggi (Yolanda and Mulya, 2024). Sebagai pembanding, algoritma Naive Bayes diterapkan karena efisiensi komputasinya pada pengklasifikasian teks yang cepat (Darwis, Siskawati and Abidin, 2020). Algoritma *Naive Bayes* juga terbukti sangat efektif dan berguna dalam mengklasifikasikan data ulasan aplikasi berbasis aspek dengan tingkat akurasi yang tinggi (Helmayanti, Hamami and Fa'rifah, 2023). Evaluasi performa kedua algoritma dilakukan guna menentukan model klasifikasi yang paling akurat untuk mendukung perbaikan sistem (Ernawati and Wati, 2024).

Meskipun penelitian mengenai analisis sentimen transportasi daring sudah banyak dilakukan, mayoritas literatur terdahulu masih berfokus pada klasifikasi sentimen secara umum (global), sehingga kurang mampu

mengidentifikasi titik permasalahan spesifik pada fitur layanan. Pada sektor transportasi daring berskala besar, pengujian menunjukkan bahwa algoritma SVM memiliki akurasi yang lebih tinggi daripada Naive Bayes dalam menganalisis sentimen ulasan aplikasi Grab (Nurfauzi *et al.*, 2025). Studi komparatif lain pada ulasan aplikasi InDrive juga menegaskan bahwa SVM lebih unggul dibandingkan Naive Bayes dalam menangani klasifikasi data teks ulasan pengguna yang kompleks (Romadhoni, Rachmadany and Prasojo, 2025). Terdapat *research gap* yang nyata di mana evaluasi sentimentil berbasis aspek (*Aspect-Based*) untuk skala transportasi lokal masih sangat terbatas jika dibandingkan dengan aplikasi besar berskala nasional. Oleh karena itu, penelitian ini hadir untuk mengisi celah tersebut dengan memetakan ulasan pengguna ke dalam lima aspek spesifik, yaitu harga, keamanan, kenyamanan, fitur aplikasi, dan pelayanan driver. Melalui pendekatan komparasi algoritma SVM dan Naive Bayes pada level aspek ini, hasil penelitian diharapkan dapat memberikan kontribusi empiris yang lebih presisi dan aplikatif bagi pengembang Ojek Indralaya dalam menentukan prioritas perbaikan sistem. Hasil pemetaan sentimen yang terperinci pada kelima aspek krusial tersebut tidak hanya akan menjadi landasan evaluasi performa aplikasi saat ini, tetapi juga dapat diintegrasikan sebagai bagian dari sistem pemantauan berkelanjutan guna memastikan bahwa setiap pembaruan perangkat lunak maupun kebijakan operasional yang diterapkan oleh pihak manajemen di masa mendatang benar-benar sejalan dengan ekspektasi dan kebutuhan riil masyarakat pengguna di tingkat lokal.

2. METODE

Penelitian ini dijalankan menggunakan bahasa pemrograman Python di lingkungan *cloud-based* Google Colaboratory (Google Colab). Tahapan metodologi dirancang secara sistematis untuk menjawab kebutuhan *Aspect-Based Sentiment Analysis* (ABSA) yang mencakup langkah-langkah berikut:

2.1 Pengumpulan Data (*Web Scraping*)

Data ulasan dikumpulkan dari platform Google Playstore menggunakan teknik *web scraping* dengan bantuan pustaka google-play-scraper. Parameter penarikan *App_ID* ditargetkan secara spesifik pada aplikasi transportasi lokal yaitu "com.myOjekIndralaya.OjekIndralaya".

Proses pengumpulan data dilakukan pada 12 Desember 2025 dengan menarik ulasan pengguna berbahasa Indonesia yang ditulis dalam rentang waktu 2019 - 2025. Penarikan data ini berhasil mengekstrak 1.023 sampel dataset ulasan yang meliputi atribut tanggal, nama pengguna, jumlah bintang (*rating*), dan isi teks ulasan.

Sebelum masuk ke tahap berikutnya, dilakukan *data cleaning* awal untuk mengeliminasi data yang tidak layak, dengan kriteria eksklusi meliputi:

1. Penghapusan ulasan duplikat (ulasan ganda dari pengguna yang sama).
2. Penghapusan ulasan kosong atau ulasan yang hanya berisi karakter non-teks (seperti emoji saja). Setelah proses eliminasi tersebut, diperoleh sebanyak [Isi Jumlah Data Bersih Akhir, misal: 1.023] data ulasan bersih yang siap digunakan untuk analisis.

2.2 *Pra-pemrosesan Teks (Pre-processing)*

Data mentah yang diperoleh melalui serangkaian standarisasi agar algoritma machine learning dapat memprosesnya dengan optimal. Tahapan ini meliputi:

1. *Case Folding*: Penyeragaman seluruh teks ulasan menjadi huruf kecil (lowercase).
2. *Cleaning*: Penghapusan elemen tidak relevan (*noise*) seperti angka, tautan URL, simbol, dan karakter tanda baca.
3. *Tokenizing*: Pemisahan struktur kalimat utuh menjadi unit-unit kata tunggal (*token*).
4. *Stopword Removal*: Eliminasi kata hubung dan kata umum bahasa Indonesia yang tidak memiliki bobot informatif (seperti "ini", "dan", "yang") menggunakan bantuan pustaka Sastrawi.
5. *Stemming*: Pemotongan kata berimbuhan kembali ke bentuk dasar linguistiknya menggunakan algoritma Sastrawi (misalnya kata "membantu" ditransformasikan menjadi "bantu"). Pada tahap ini, kata-kata tidak baku (*slang words*) yang memuat intensitas emosi kuat (seperti "mantaappp" dan "sangaat") sengaja dipertahankan agar tidak kehilangan bobot sentimennya.

2.3 Pelabelan Aspek dan Sentimen

Proses pelabelan dalam penelitian ABSA ini dibagi menjadi dua tahap utama, yaitu penentuan aspek layanan dan penentuan polaritas sentimen:

1. Mekanisme Penentuan Aspek: Setiap ulasan diklasifikasikan ke dalam 5 kategori aspek layanan spesifik, yaitu: *harga, keamanan, kenyamanan, fitur aplikasi,* dan *pelayanan driver*. Penentuan aspek dilakukan secara manual. Jika sebuah ulasan memuat lebih dari satu aspek, maka ulasan tersebut dianalisis secara terpisah pada masing-masing aspek yang bersangkutan.
2. Mekanisme Penentuan Sentimen: Selaras dengan batasan masalah untuk menghindari ambiguitas, ulasan pada tiap aspek dilabeli ke dalam 2 kelas sentimen, yaitu Positif dan Negatif. Pelabelan sentimen ini ditentukan berdasarkan [Pilih salah satu sesuai eksperimen Anda: konversi skor rating (rating 4-5 sebagai positif, rating 1-2 sebagai negatif, sedangkan rating 3 dieksklusi) ATAU anotasi manual oleh tiga orang *annotator* menggunakan sistem *majority voting*]

2.4 Pembobotan Fitur, Klasifikasi, dan Evaluasi Model

Teks ulasan yang telah bersih dan dilabeli kemudian diekstraksi ke dalam matriks vektor numerik menggunakan algoritma pembobotan kata TfidfVectorizer (TF-IDF). Selanjutnya, matriks fitur tersebut dibagi menjadi 80% data latih (*training data*) dan 20% data uji (*testing data*). Pembagian dataset ini dilakukan menggunakan teknik *stratified split* untuk memastikan proporsi kelas sentimen positif dan negatif tetap seimbang dan representatif baik pada data latih maupun data uji, mengingat adanya kondisi data yang tidak seimbang (*imbalanced dataset*) antar-aspek.

Proses pengklasifikasian sentimen dijalankan dengan membandingkan dua algoritma *machine learning*, yaitu:

1. ***Support Vector Machina* (SVM):** Dijalankan melalui modul SVC (*Support Vector Classification*) menggunakan kernel linear dan parameter regularisasi $C = 1.0$$.
2. ***Naive Bayes*:** Dijalankan menggunakan algoritma Multinomial Naive Bayes dengan parameter *smoothing* $\alpha = 1.0$.

Kedua model tersebut dilatih menggunakan data latih dan diuji menggunakan data uji untuk masing-masing aspek secara terpisah. Evaluasi unjuk kerja dan keandalan model diukur secara komprehensif melalui nilai persentase dari kalkulasi *Confusion Matrix*, yang memuat metrik *Accuracy, Precision, Recall*, dan *F1-Score*.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan dan Pra-pemrosesan Data (*Pre-prosecing*)

Tahapan awal penelitian menghasilkan 1.023 data ulasan pengguna yang ditarik secara *real-time* dari Google Playstore. Atribut data yang diekstraksi mencakup identitas pengguna (*username*), tanggal ulasan (*at*), skor penilaian atau *rating* (*score*), dan isi teks ulasan (*content*).

Sebelum diekstraksi menjadi matriks numerik, ribuan teks ulasan tersebut melalui proses pembersihan (*preprocessing*) untuk mereduksi dimensi linguistik dan membuang elemen yang tidak relevan (*noise*). Proses standarisasi ini meliputi pengubahan ke huruf kecil (*Case Folding*), penghapusan kata hubung yang tidak bermakna seperti "ini", "dan", serta "yang" (*Stopword Removal*), dan pemangkasan kata berimbuhan kembali ke bentuk dasarnya menggunakan pustaka Sastrawi (*Stemming*), misalnya transformasi kata "membantu" menjadi "bantu". Untuk mempermudah pemahaman mengenai perubahan struktur data, Contoh konkret perubahan teks pada setiap tahapan *preprocessing* disajikan pada Tabel 1.

Tabel 1. Contoh Konkret Tahapan *Preprocessing* Data Teks

Tahapan	Output Teks Ulasan	Keterangan
Data Mentah	"Aplikasi ini sangat membantu!! Tapi driver-nya Sering lama bgt menjemput, mantaappp	Teks awal hasil *scraping*.

	tingkatkan lagi."	
Case Folding	"aplikasi ini sangat membantu!! tapi driver-nya sering lama bgt menjemput, mantaappp tingkatkan lagi."	Mengubah seluruh huruf menjadi kecil.
Cleaning	"aplikasi ini sangat membantu tapi driver nya sering lama bgt menjemput mantaappp tingkatkan lagi"	Menghapus tanda baca (!!) dan tanda hubung (-).
Tokenizing	['aplikasi', 'ini', 'sangat', 'membantu', 'tapi', 'driver', 'nya', 'sering', 'lama', 'bgt', 'menjemput', 'mantaappp', 'tingkatkan', 'lagi']	Pemotongan string menjadi token kata tunggal.
Stopword Removal	['aplikasi', 'membantu', 'driver', 'sering', 'lama', 'bgt', 'menjemput', 'mantaappp', 'tingkatkan']	Menghilangkan kata non-potensial ("ini", "sangat", "tapi", "nya", "lagi").
Stemming	['aplikasi', 'bantu', 'driver', 'sering', 'lama', 'bgt', 'jemput', 'mantaappp', 'tingkat']	Mengembalikan kata berimbuhan ke bentuk kata dasar menggunakan Sastrawi.

Dalam proses *stemming* dan pembobotan, algoritma sengaja dikondisikan untuk mempertahankan *slang word* atau kata dengan pengetikan tidak baku yang sengaja ditekankan pengguna, seperti kata "mantaappp" dan "sangaat", karena kata-kata tersebut memuat bobot emosi yang kuat. Agar tidak mengganggu kinerja *stemming* Sastrawi, kata *slang* diisolasi sebelum proses pemotongan imbuhan dilakukan. Karena pustaka Sastrawi hanya mengenali kata dasar baku Indonesia, kata *slang* seperti "mantaappp" akan dilewati (*skipped*) oleh fungsi *stemmer* tanpa diubah menjadi rusak, sehingga karakteristik unik leksikon emosi tersebut tetap terjaga hingga masuk ke tahap pembobotan TF-IDF.

3.2 Distraksi dan Proporsi Aspek Ulasan

Berdasarkan pemetaan data, proporsi pembahasan pengguna menunjukkan distribusi yang sangat bervariasi. Aspek Fitur Aplikasi terkonfirmasi sebagai topik yang paling dominan dibahas oleh pengguna dengan jumlah mencapai 446 ulasan, mencakup hampir separuh data yaitu sebesar 49,2%. Tingginya persentase ini membuktikan bahwa stabilitas *software* dan antarmuka (UI/UX) merupakan fokus utama evaluasi pelanggan. Proporsi pembahasan terbesar selanjutnya diikuti oleh aspek Kenyamanan dengan 162 ulasan (17,9%), aspek Pelayanan dengan 147 ulasan (16,2%), dan aspek Harga dengan 112 ulasan (12,4%). Sebaliknya, aspek Keamanan menempati irisan paling marjinal dengan hanya memuat 39 ulasan (4,3%) dari total keseluruhan data.

3.3 Hasil Pengujian Model Klasifikasi Sentimen

Kinerja algoritma *Support Vector Machine* (SVM) bervarian kernel Linear dan Multinomial Naive Bayes (NB) dievaluasi berdasarkan empat metrik utama: Akurasi, *Precision*, *Recall*, dan *F1-Score*. Guna meningkatkan keterbacaan data secara terstruktur, hasil rekapitulasi performa komparasi disajikan pada Tabel 2.

Tabel 2. Rekapitulasi Perbandingan Kinerja SVM dan Naive Bayes

Aspek	Algoritma	Akurasi	Precision	Recall	F1-*Score*
Harga	SVM	95.65%	95.88%	95.65%	95.47%
	Naive Bayes	78.26%	61.25%	78.26%	68.72%
Keamanan	SVM	75.00%	82.14%	75.00%	70.83%
	Naive Bayes	62.50%	39.06%	62.50%	48.08%
Kenyamanan	SVM	90.91%	90.52%	90.91%	90.58%
	Naive Bayes	81.82%	66.94%	81.82%	73.64%
Fitur	SVM	90.00%	89.91%	90.00%	89.84%
	Naive Bayes	83.33%	85.07%	83.33%	81.60%
Pelayanan	SVM	86.67%	86.58%	86.67%	86.14%
	Naive Bayes	70.00%	9.00%	70.00%	57.65%

Untuk memastikan detail keandalan klasifikasi, matriks kebingungan (*Confusion Matrix*) model SVM dibedah secara empiris. Representasi angka persebaran prediksi aktual *Confusion Matrix* SVM dapat dilihat pada Tabel 3.

Tabel 3. Distribusi *Confusion Matrix* Model SVM Per Aspek

Aspek Layanan	TP	TN	FP	FN
Harga	18	4	1	0
Fitur Aplikasi	59	22	6	3
Pelayanan *Driver*	20	6	3	1
Kenyamanan	26	4	2	1
Kemanan	5	1	2	0

Keterangan: TP (*True Positve*), TN (*True Negative*), FP (*False Positive*), FN (*False Negative*)

1. Aspek Harga: SVM menebak dengan presisi tinggi melalui 18 ulasan *True Positive* (TP) dan 4 ulasan *True Negative* (TN). Kesalahan klasifikasi hanya terjadi pada 1 ulasan keluhan yang ditebak positif (*False Positive* / FP), dengan angka *False Negative* (FN) bernilai 0.
2. Aspek Fitur: Sebagai aspek dengan sampel terbanyak, model berhasil mendeteksi 59 TP dan 22 TN. Kesalahan prediksi sangat minim secara proporsional, yakni 6 FP dan 3 FN, membuktikan konsistensi algoritma membaca terminologi UI/UX.
3. Aspek Pelayanan: Evaluasi mencatatkan 20 TP dan 6 TN, diwarnai 3 FP di mana model salah mengenali keluhan respons pengemudi sebagai pujian, serta 1 FN.
4. Aspek Kenyamanan: Model memprediksi 26 TP, 4 TN, 2 FP, dan 1 FN, membuktikan pembobotan TF-IDF sangat tangguh membedakan leksikon fasilitas.
5. Aspek Keamanan: Model berhasil mendeteksi 5 TP dan 1 TN, namun dibayangi oleh 2 FP dan 0 FN akibat keterbatasan distribusi data uji.

3.4 Analisis Komparasi Kinerja Algoritma

Penelitian ini menemukan adanya kesenjangan (*gap*) keandalan komputasional yang tajam antara SVM dan Naive Bayes.

1. Dominasi SVM pada Data Berdimensi Tinggi:

SVM menunjukkan stabilitas prediksi yang konsisten di kelima kategori aspek. Keunggulan ini dibuktikan secara empiris melalui penggunaan kernel Linear, yang terbukti optimal menarik batas *hyperplane* memisahkan sentimen kalimat. SVM juga memperlihatkan tingkat ketahanan (*robustness*) yang baik menghadapi *slang word* dan bahasa non-standar, mempertahankan skor di atas 86% pada aspek padat data seperti Fitur dan Kenyamanan.

2. Penurunan Performa Naive Bayes pada Data Tak Seimbang (*Imbalanced Dataset*):

Naive Bayes mengalami penurunan performa yang cukup signifikan ketika memproses data ulasan yang asimetris. Kelemahan mendasar ini terekspos pada rendahnya nilai *Precision* di aspek Keamanan (39,06%) dan Pelayanan (49,00%). Kondisi ini menandakan tingginya akumulasi kesalahan *False Positive*; sistem terlalu agresif melabeli rentetan kalimat keluhan seakan-akan merupakan apresiasi. Kegagalan probabilitas ini

berpangkal dari kelemahan asumsi independensi fitur (*feature independence*); Naive Bayes menghitung setiap kata secara terisolasi, sehingga gagal memahami tautan semantik kalimat utuh serta luput menangkap frasa negasi.

3. Kompleksitas Terminologi Keamanan

Aspek Keamanan teridentifikasi sebagai parameter yang paling menyulitkan mesin, di mana akurasi SVM menurun ke angka 75,00% dan Naive Bayes berada di rentang 62,50%. Kompleksitas ini bukan hanya disebabkan oleh terbatasnya jumlah sampel data, melainkan tingginya ambiguitas leksikon. Kata "ngebut" dapat dinilai sangat positif oleh penumpang yang mengejar waktu, sekaligus bermakna negatif bila disandingkan dengan parameter keselamatan fisik.

3.5 Analisis Kualitatif Berdasarkan Visualisasi *Word Cloud*

Untuk membedah motivasi di balik polaritas sentimen pengguna, diksi dominan yang diekstraksi lewat teknik visualisasi *Word Cloud* diuraikan pada Tabel 4 sebagai representasi tekstual awan kata.

Tabel 4. Daftar Kluster Kata Dominan Hasil Visualisai *Word Cloud*

Aspek Layanan	Kata Kunci Dominan (Sentimen Positif)	Kata Kunci Dominan (Sentimen Negatif)
Harga	"murah", "harga", "jangkau", "sahabat"	"hujan", "mahal", "biaya", "tambah"
Kenyamanan dan Kemanan	"aman", "cepat", "ramah", "bantu", "mudah"	"pesan", "lama", "order", "update", "gak"
Fitur Aplikasi	"mudah", "bantu"	"update", "sering", "pesan", "buka", "apk", "gak"
Pelayanan Driver	"ramah", "sopan", "layan", "baik"	"driver", "sering", "lama", "minta"

1. Aspek Harga: Kepuasan pengguna ditandai dengan kemunculan dominan diksi "murah", "harga", "jangkau", dan "sahabat", yang menegaskan daya saing OJIN dari sisi ekonomis. Sebaliknya, protes terpusat pada frasa "hujan", "mahal", "biaya", dan "tambah", yang menjadi bukti adanya keberatan publik terhadap fluktuasi tarif di musim hujan.
2. Aspek Keamanan dan Kenyamanan: Pada spektrum positif, pengguna memvalidasi layanan melalui diksi "aman", "cepat", "ramah", "bantu", dan "mudah". Namun, visualisasi awan kata negatif justru menonjolkan kata "pesan", "lama", "order", "update", dan "gak". Hal ini mengindikasikan bahwa keresahan psikologis pengguna terkait rasa kurang aman atau kurang nyaman disebabkan oleh kendala teknis, seperti menunggu armada dalam waktu lama atau tidak mendapat kepastian pesanan diproses.
3. Aspek Fitur Aplikasi: Meskipun pujian mengalir lewat kata "mudah" dan "bantu" , kluster kata negatif secara jelas menunjukkan tumpukan kendala sistem operasional melalui keluhan "update", "sering", "pesan", "buka", "apk", dan "gak". Hal ini menandakan tingkat ketidaknyamanan yang cukup tinggi terhadap frekuensi pembaruan aplikasi yang dianggap mengganggu.
4. Aspek Pelayanan *Driver*: eberhasilan layanan didukung oleh kinerja pengemudi yang dinilai "ramah", "sopan", "layan", dan "baik". Namun, sentimen negatif merekam keluhan terhadap oknum pengemudi melalui kemunculan kata "driver", "sering", "lama", dan "minta", merepresentasikan kendala pengemudi yang lamban menjemput atau meminta pembatalan sepihak.

4. PENUTUP

4.1 Kesimpulan

Penerapan *Aspect-Based Sentiment Analysis* (ABSA) berhasil menstrukturisasi opini pelanggan pada aplikasi Ojek Indralaya ke dalam lima kategori aspek layanan. Hasil analisis menunjukkan bahwa "Fitur Aplikasi"

menjadi porsi aspek yang paling banyak dibahas oleh pengguna, yakni sebesar 49,2%.

Berdasarkan analisis perbandingan *machine learning*, algoritma *Support Vector Machine* (SVM) menunjukkan performa yang lebih stabil, kokoh (*robust*), dan presisi dibandingkan model Naive Bayes pada seluruh metrik evaluasi. Performa tertinggi model SVM tercatat pada aspek Harga dengan raihan akurasi mencapai 95,65%. Model SVM juga terbukti memiliki ketahanan yang baik di tengah anomali *imbalanced dataset*, kondisi yang sebaliknya membuat model Naive Bayes mengalami penurunan performa yang signifikan, terutama pada aspek Keamanan dan Pelayanan.

4.2 Saran

Berdasarkan temuan empiris dan analisis eksplorasi teks dalam penelitian ini, saran yang diajukan dibagi menjadi dua bagian sebagai berikut:

1. **Saran Praktis (Bagi Pengembang Aplikasi Ojek Indralaya**)
   - Pengembang sistem OJIN disarankan untuk memprioritaskan perbaikan antarmuka (UI/UX) guna mengatasi kendala *bug error*, tingginya frekuensi pembaruan (*update*) aplikasi yang mengganggu pengguna, serta optimalisasi sistem manajemen *order*.
   - Pihak manajemen disarankan melakukan evaluasi manajerial operasional, khususnya terkait mitigasi pembatalan pesanan (*cancel order*) secara sepihak oleh pengemudi (*driver*) serta pentingnya transparansi skema penambahan harga tarif fluktuatif saat musim penghujan di area Indralaya.
2. **Saran Metodologis (Bagi Penelitian Akademis Selanjutnya)**
   - Untuk mengatasi kendala ketidakseimbangan data (*imbalanced dataset*) yang jamak ditemui pada ulasan lokal, penelitian berikutnya disarankan menerapkan teknik *oversampling* seperti *Synthetic Minority Over-sampling Technique* (SMOTE) sebelum tahap pemodelan.
   - Peneliti selanjutnya disarankan menguji kapabilitas arsitektur *Deep Learning* yang lebih mutakhir, seperti model BERT atau IndoBERT, guna memperkuat performa diagnosis *Natural Language Processing* (NLP) dalam menangani ambiguitas leksikon bahasa Indonesia non-baku
   - Direkomendasikan pula penambahan proses validasi label secara manual melibatkan beberapa *annotator* ahli (*multi-annotator verification*) untuk meminimalkan subjektivitas pelabelan ulasan aspek.

5. DAFTAR PUSTAKA

Agustina, D.A., Subanti, S. and Zukhronah (2020) ‘Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine’, 3(2), pp. 109–122.

Anam, M.K., Saifuddin and Fitriah, S.N. (2020) ‘Kontribusi Pengemudi Ojek Online (GRAB) dalam Pelayanan Masyarakat di Kabupaten Malang’, *Jurnal Kajian Ekonomi dan Perbankan*, 4(2), pp. 1–15.

Arifin, N., Enri, U. and Sulistiyowati, N. (2021) ‘Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-GRAM untuk Text Classification’, *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 6(2), pp. 129–136.

Aryanti, F.A.D., Luthfiarta, A. and Soeroso, D.A.I. (2025) ‘Aspect-Based Sentiment Analysis with LDA and IndoBERT Algorithm on Mental Health App : Riliv’, *Journal of Applied Informatics and Computing (JAIC)*, 9(2), pp. 361–375. Available at: http://jurnal.polibatam.ac.id/index.php/JAIC.

Astuti, K.C., Firmansyah, A. and Riyadi, A. (2024) ‘Implementasi Text Mining untuk Analisis Sentimen Masyarakat terhadap Ulasan Aplikasi Digital Korlantas Polri pada Google Play Store’, *Remik: Riset dan E-Jurnal Manajemen Informatika Komputer*, 8(1), pp. 383–394.

Darwis, D., Siskawati, N. and Abidin, Z. (2020) ‘Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional’, *Jurnal TEKNO KOMPAK*, 15(1), pp. 131–145.

Ernawati, S. and Wati, R. (2024) 'Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon', *Jurnal Buana Informatika*, 15(April), pp. 40–49.

Helmayanti, S.A., Hamami, F. and Fa'rifah, R.Y. (2023) 'Penerapan Algoritma TF-IDF dan Naive Bayes untuk Analisis Sentimen Berbasis Aspek Ulasan Flip pada Google Play Store', *Jurnal Indonesia : Manajemen Informatika dan Komunikas*, 4(3), pp. 1822–1834.

Hua, Y.C. *et al.* (2024) 'A Systematic Review of Aspect - Based Sentiment Analysis: Domains, Methods, and Trends', *Artificial Intelligence Review*, 57(11), pp. 1–51. Available at: https://doi.org/10.1007/s10462-024-10906-z.

Ningrum, M.P. *et al.* (2025) 'Evaluasi User Experience pada Aplikasi Maxim menggunakan Metode HEART Metrics', *Sistemasi: Jurnal Sistem Informasi*, 14(4), pp. 1763–1775.

Nurfauzi, O.M. *et al.* (2025) 'Analisis Sentimen Grab Indonesia pada Ulasan Google Play Store Menggunakan Algoritma Naïve Bayes dan Support Vector Machine', *SMARTICS Journal*, 11(1), pp. 8–13.

Rahmajati, H., Saputra, M.C. and Herlambang, A.D. (2025) 'Penerapan Aspect-Based Sentiment Analysis untuk Identifikasi Masalah Kualitas Layanan Aplikasi Mobile iPusnas berdasarkan Persepsi Pengguna', *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 9(7), pp. 1–10.

Rivaldi, R.C. and Wismarini, T.D. (2024) 'Analisis Sentimen Pada Ulasan Produk Dengan Metode Natural Language Processing (NLP) (Studi Kasus Zalika Store 88 Shopee)', *Elkom: Jurnal Elektronika dan Komputer*, 17(1), pp. 120–128.

Romadhoni, A.A., Rachmadany, A. and Prasojo, B.H. (2025) 'Sentiment Analysis of Indrive App Usage Reviews on Google Playstore Using Support Vector Machine (SVM) and Naïve Bayes Algorithm', *International Journal of Artificial Intelligence for Digital Marketing*, 2(10), pp. 102–112.

Suchrady, R.Z. and Purwarianti, A. (2023) 'Indo LEGO-ABSA : A Multitask Generative Aspect Based Sentiment Analysis for Indonesian Language', *Cornell University* [Preprint]. Available at: https://arxiv.org/pdf/2311.01757.

Yolanda, A.M. and Mulya, R.T. (2024) 'Implementasi Metode Support Vector Machine untuk Analisis Sentimen pada Ulasan Aplikasi Sayurbox di Google Play Store', 6(2), pp. 76–83. Available at: https://doi.org/10.35580/variansiunm258.