ANALISIS SENTIMEN PENGGUNA APLIKASI FORE MELALUI KOMENTAR DI GOOGLE PLAY STORE MENGGUNAKAN METODE SVM

Mohamad Barkah[1], Edi Tohidi[2], Tati Suprati[3], Arif Rinaldi Dikananda[4].

Program Studi Teknik Informatika[1 3]
Program Studi Komputerisasi Akuntansi[2]
Program Studi Rekayasa Perangkat Lunak[4]

STMIK IKMI CIREBON
https://ikmi.ac.id/page/18/?lang=de
muhamadbarkah0@gmail.com
(*) Corresponding Author : muhamadbarkah0@gmail.com
Published : 16 Juli 2026

Abstract— *The development of consumer needs for the convenience of ordering and delivering drinks via the internet, the FORE mobile-based food and beverage (F&B) application has become very popular and is available on the Google Play Store. Therefore, user reviews can help understand user perceptions, satisfaction levels, and problems that arise during the application usage experience. The purpose of this study is to analyze FORE application user sentiment using the Support Vector Machine (SVM) algorithm method and build a classification model to obtain objective user satisfaction levels. This study has 6,600 user review data with web scraping techniques on the Google Play Store in the period December 2018 - October 2025. Data was collected by including text preprocessing (data cleaning, normalization, tokenization, stopword removal, and stemming using the Sastrawi library) which was then carried out using the Term Frequency–Inverse Document Frequency (TF-IDF) feature extraction technique by producing 5,180 features, then the dataset was divided into training data of 80% and test data of 20%. SVM model training was used by classifying reviews into positive & negative categories and the results showed that the majority of positive sentiments were 4,854 data, followed by negative sentiments of 1,631 data. The SVM model produced an accuracy of 81.79% with very good performance in the positive class (Precision 0.85 and recall 0.91). The results of this study are expected to help the FORE application improve service quality and enhance features that suit user needs, and a more optimal data approach is needed to improve model performance and more accurate classification.*

Keywords: *Sentiment Analysis, Support Vector Machine (SVM), TF-IDF, User Reviews, Google Play Store.*

Abstrak — Perkembangan kebutuhan konsumen akan kemudahan pemesanan dan pengiriman minuman melalui internet, aplikasi makanan dan minuman (F&B) berbasis ponsel FORE menjadi sangat populer dan tersedia di Google Play Store. Oleh karena itu, ulasan pengguna dapat membantu memahami persepsi pengguna, tingkat kepuasan, dan masalah yang muncul selama pengalaman penggunaan aplikasi Tujuan dari penelitian ini adalah untuk menganalisis sentimen pengguna aplikasi FORE dengan metode algoritma Support Vector Machine (SVM) serta membangun model klasifikasi guna memperoleh tikat kepuasan pengguna secara objektif. Penelitian ini memiliki 3.966 data ulasan pengguna dengan teknik web scrapping di Google Play Store pada periode Desember 2018 – Oktober 2025. Data dikumpulkan dengan meliputi preprocessing teks (pembersihan data, normalisasi, tokenisasi, penghapusan stopword, dan stemming menggunakan pustaka sastrawi) yang selanjutnya dilakukan menggunakan Teknik ekstraksi fitur Term Frequency–Inverse Document Frequency (TF-IDF) dengan menghasilkan 3.733 fitur, kemudian Dataset dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%. Pelatihan model SVM digunakan dengan mengklasifikasikan ulasan ke dalam kategori positif, negatif dan netral dengan hasil menunjukan bahwa mayoritas sentimen positif sebanyak 1.923 data, diikuti sentimen netral sebanyak 1.572 data dan sentimen

negatif sebanyak 223 data`. Model SVM menghasilkan akurasi sebesar 84.36% dengan performa sangat baik pada kelas positif (Precision 0.88 dan recall 0.95) namun relatif rendah pada kelas negatif (precision 0.72 dan recall 0.71) Hasil penelitian ini diharapkan dapat membantu aplikasi FORE meningkatkan kualitas layanan dan meningkatkan fitur yang sesuai dengan kebutuhan pengguna, dan diperlukan pendekatan data yang lebih optimal untuk meningkatkan kinerja model serta klasifikasi yang lebih akurat.
Kata kunci: Analisis Sentimen, Support Vector Machine (SVM), TF-IDF, Ulasan Pengguna, Google Play Store.

1. Pendahuluan

Perubahan perilaku konsumen yang menginginkan kemudahan, kecepatan, dan kenyamanan dalam memenuhi kebutuhan sehari-hari mendorong pertumbuhan layanan aplikasi makanan dan minuman (F&B) di Indonesia. Studi terbaru menunjukkan bahwa pengalaman pengguna, atau user experience, sangat penting untuk keberhasilan aplikasi F&B. Ini termasuk bagaimana pengalaman pengguna mempengaruhi ulasan aplikasi.

Menurut Technology Acceptance Model (TAM), persepsi kegunaan, kepuasan, kepercayaan, dan pengaruh sosial adalah faktor utama yang mendorong pengguna untuk menggunakan aplikasi pengantaran makanan (Kim et al., 2022). Hasil ini diperkuat oleh (Macias et al., 2025), yang menekankan bahwa kepuasan pengguna dan kesetiaan merek secara langsung dipengaruhi oleh kualitas layanan digital (e-service quality), interaksi dengan kurir, dan persepsi tentang kualitas makanan.

Dalam hal pengambilan keputusan konsumen, ditemukan (Zhai et al., 2022) bahwa tiga faktor utama yang memengaruhi keinginan konsumen untuk menggunakan layanan F&B digital adalah harga, kemudahan penggunaan aplikasi, dan pilihan menu. Dengan demikian, model respons stimulus-organism (SOR) menekankan bahwa kualitas produk, kurir, dan platform sangat memengaruhi keinginan pengguna untuk terus menggunakan aplikasi yang relevan (Williams et al., 2024). Sementara itu, motivasi hedonis, kenyamanan, dan persepsi kebermanfaatan setelah penggunaan adalah faktor penting dalam perilaku penggunaan berkelanjutan (Soesanto, 2023).

Di sisi teknis, pengembang dapat memperoleh informasi penting dari analisis ulasan aplikasi. Studi software engineering, seperti yang ditunjukkan oleh Motger et al. (2025), Sun et al. (2025), dan Dąbrowski et al. (2021), menunjukkan bahwa ulasan pengguna dapat menunjukkan masalah sistem, peluang untuk inovasi aplikasi, dan kebutuhan untuk perbaikan fitur.

Berdasarkan hasilnya, penelitian analisis sentimen terhadap ulasan aplikasi FORE menjadi relevan karena faktor-faktor pengalaman pengguna seperti kemudahan, kepercayaan, kepuasan, kualitas layanan, interaksi antar karyawan, dan kemudahan sangat mungkin memengaruhi sentimen dalam komentar di Google Play Store. Oleh karena itu, pemodelan analisis sentimen berbasis pembelajaran mesin yang menggunakan metode Support Vector Machine (SVM) diperlukan untuk memahami persepsi pengguna secara lebih akurat dan tidak bias.

Jumlah aplikasi yang beredar di berbagai platform seperti Google Play Store meningkat sebagai akibat dari perkembangan teknologi mobile yang sangat pesat. Keberhasilan sebuah aplikasi sangat bergantung pada kualitas layanan dan kemampuan pengembang untuk memenuhi kebutuhan pengguna dalam ekosistem digital yang kompetitif ini. Ulasan pengguna, atau user reviews, adalah salah satu sumber informasi paling penting bagi pengembang untuk memahami pengalaman dan harapan pengguna. Studi terbaru menunjukkan bahwa ulasan pengguna sangat penting untuk meningkatkan aplikasi, kualitas layanan, dan evolusi proses perangkat lunak. Ketika organisasi secara aktif menanggapi ulasan, terutama yang negatif atau memiliki deviasi sentimen tinggi, mereka dapat meningkatkan kualitas ulasan di masa mendatang (Su et al., 2024). Respon yang lebih panjang dan lebih personal telah terbukti mendorong pengguna untuk memberikan umpan balik yang lebih menyeluruh, yang membantu pengembang memahami kebutuhan yang lebih dalam.

Selain itu, analisis sistematis terhadap 182 penelitian (Dąbrowski et al., 2021) menunjukkan bahwa ulasan pengguna sangat penting bagi para engineer untuk memahami kebutuhan pengguna, merencanakan perbaikan bug, dan merencanakan perkembangan aplikasi. Ini diperkuat oleh penelitian tambahan (Dąbrowski et al., 2021), yang menunjukkan bahwa teknik penambangan opini pada ulasan dapat membantu perencanaan rilis dan pemeliharaan perangkat lunak serta mengekstrak informasi terkait fitur.

Selain itu, dinamika ulasan positif dan negatif juga memiliki dampak yang berbeda terhadap inovasi aplikasi. Ulasan negatif memiliki kecenderungan yang lebih besar untuk memicu inovasi dibandingkan dengan ulasan positif (Sun et al., 2025). Ini karena ulasan negatif mengandung lebih banyak informasi dan sinyal emosional. Dengan kata lain, keluhan pengguna mendorong

pengembang untuk meningkatkan fungsi dan kinerja aplikasi.

Sebaliknya, kemajuan AI telah membawa metode baru untuk menilai ulasan pengguna. T-FREX adalah metode berbasis transformer yang dapat mengekstraksi fitur fungsional dari ulasan secara otomatis (Motger et al., 2024). Teknik ini memungkinkan pengembang menilai prioritas perbaikan berdasarkan kebutuhan nyata pengguna, menghubungkan konten ulasan dengan tindakan perbaikan yang nyata.

Serangkaian hasil menunjukkan bahwa ulasan pengguna tidak hanya dapat digunakan untuk menunjukkan kepuasan atau ketidakpuasan pengguna, tetapi juga dapat digunakan sebagai sumber data yang berharga untuk mengoptimalkan kualitas layanan, membuat fitur baru, dan terus meningkatkan kinerja aplikasi. Oleh karena itu, analisis ulasan pengguna sangat penting untuk pengembangan perangkat lunak kontemporer, terutama melalui teknik text mining dan analisis perasaan.

2. Metode Penelitian

Metode penelitian ini menjelaskan tahapan pengumpulan data, EDA (Ekspolarasi Data Awal), *pre-processing* data, pelabelan, pembagian data, Vektorisasi Teks Menggunakan TF-IDF, pelatihan model SVM, evaluasi model.

2.1 Pengumpulan Data

Tahap ini bertujuan untuk mengumpulkan data ulasan pengguna dari Google Play Store. Dengan menggunakan google-play-scraper, peneliti mengambil ulasan pengguna Fore sebanyak 3.001 data. Data yang diambil mencakup kolom teks ulasan dan nilai rating. Nilai rating digunakan untuk membantu pelabelan awal sentimen.

2.3 EDA (Ekspolarasi Data Awal)

Langkah ini dilakukan untuk memastikan data bersih dan siap diolah. Proses meliputi:

1. Filter Kolom: mengambil hanya kolom yang relevan, yaitu kolom rating dan content (teks ulasan)
2. Hapus Data Duplikat: menghapus ulasan yang sama agar tidak terjadi bias pada hasil klasifikasi
3. Visualisasi Data Awal: menampilkan distribusi rating dan jumlah ulasan untuk memahami persebaran data sebelum diproses lebih lanjut.

Visualisasi awal dilakukan menggunakan matplotlib dan word cloud untuk mengidentifikasi frekuensi kata yang sering muncul.

2.4 *Pre-Processing* Data

Tahapan ini merupakan inti dari pipeline NLP. Prosesnya meliputi:

1. Cleaning Data: Menghapus karakter tak relevan seperti URL, tag HTML, simbol, angka, dan emoji menggunakan modul regex (re) dan string.
2. Case Folding: Mengonversi seluruh teks menjadi huruf kecil agar seragam.
3. Normalisasi Data: Mengubah kata tidak baku menjadi bentuk baku dengan menggunakan kamus dari dataset kamus-slang Indonesia.
4. Tokenisasi: Memecah kalimat menjadi unit kata (tokens) menggunakan pustaka NLTK.
5. Stopword Removal: Menghapus kata umum yang tidak berpengaruh terhadap makna sentimen seperti “yang”, “dan”, “itu”.
6. Stemming Data: Mengubah kata ke bentuk dasarnya menggunakan pustaka Sastrawi.
7. Visualisasi Data Setelah Preprocessing: Membuat word cloud untuk melihat perubahan distribusi kata.
8. Pelabelan data dilakukan dengan memberikan kategori “positif” atau “negatif” berdasarkan hasil penilaian (rating).
9. Visualisasi Sentimen: Membuat visualisasi word cloud terpisah untuk kata positif dan negatif guna memastikan distribusi label seimbang.

2.5 Pelabelan

Metode ini memanfaatkan dua kamus sentimen, yaitu positive.tsv dan negative.tsv, yang berisi daftar kata berbahasa Indonesia beserta bobot (weight) yang mencerminkan kekuatan sentimen positif atau negatif. Setiap ulasan yang telah dipreproses dipecah menjadi token kata, kemudian setiap token dicocokkan dengan kedua kamus tersebut. Jika kata ditemukan dalam kamus positif, bobotnya ditambahkan pada skor positif ulasan; jika ditemukan dalam kamus negatif, bobotnya menambah skor negatif. Setelah seluruh token diperiksa, ulasan diberi label positif jika skor akhir lebih besar dari nol, dan negatif jika lebih kecil dari nol.

2.6 Pembagian Data

Setelah teks siap digunakan, langkah berikutnya adalah mempersiapkan data untuk pelatihan model aplikasi FORE, Splitting Data: Dataset dibagi menjadi 80% data latih dan 20% data uji menggunakan fungsi train_test_split dari scikit-learn. Pembobotan TF-IDF mengubah teks menjadi representasi numerik berbasis TF-IDF untuk merepresentasikan kepentingan kata dalam dokumen. Tahap ini penting karena kualitas pembagian data akan berpengaruh langsung terhadap tingkat generalisasi model dalam memprediksi sentimen pada data baru.

2.7 Vektorisasi Teks Menggunakan TF-IDF

Sebelum memasuki tahap pemodelan, data teks harus diubah menjadi bentuk numerik agar dapat diproses oleh algoritma machine learning.

Pada penelitian ini digunakan teknik Term Frequency–Inverse Document Frequency (TF-IDF), yang mampu memberikan bobot pada kata berdasarkan tingkat kepentingannya dalam dokumen. Vektorisasi TF-IDF menghasilkan matriks berdimensi tinggi yang berisi representasi numerik setiap ulasan. Tahap ini merupakan proses krusial karena kualitas vektorisasi akan memengaruhi kemampuan model dalam mengenali pola sentimen. Model klasifikasi teks yang telah melalui tahap prapemrosesan diubah menjadi representasi numerik menggunakan teknik Term Frequency–Inverse Document Frequency (TF-IDF) (Akhir & Komaran, 2025)

Metrik Akurasi mengukur frekuensi keakuratan prediksi model di setiap kelas. Metrik ini ditentukan melalui pembagian jumlah prediksi akurat (termasuk Positif Benar dan Negatif Benar) dengan jumlah keseluruhan prediksi yang dihasilkan. Rumus evaluasi mengikuti formula umum oleh (Arifin et al., 2025):

$$accuracy) = \frac{TP + TN}{Total\ Prediction}$$

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

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

$$F1\ Score\ = 2 \times \frac{recall \times precision}{recall \times precision}$$

Evaluasi model ini untuk menghasilkan prediksi berdasarkan data uji. Ini bukan sekadar detail teknis, melainkan faktor krusial yang menentukan kelayakan penerapan model kami. dengan menampilkan confusion matrix dan laporan klasifikasi menggunakan fungsi classification_report() dari scikit-learn. Hasil terbaik dipilih berdasarkan kombinasi tertinggi dari accuracy dan F1-scorebotan TF-IDF mengubah teks menjadi representasi numerik berbasis TF-IDF untuk merepresentasikan kepentingan kata dalam dokumen. Tahap ini penting karena kualitas pembagian data akan berpengaruh langsung terhadap tingkat generalisasi model dalam memprediksi sentimen pada data baru.

2.8 Pelatihan Model SVM

Tahap ini bertujuan menguji model pada data baru untuk menilai kemampuan generalisasi model. Model SVM yang telah dilatih diujikan pada ulasan baru dari pengguna iPusnas untuk melihat kesesuaian label prediksi dengan opini aktual pengguna.

2.9 Evaluasi Model

Setelah pelatihan, kami beralih ke fase kritis: menguji kualitas model dengan data yang belum pernah dilihat sebelumnya (*data uji*). Evaluasi ini bukan hanya tentang angka, melainkan tentang validasi kemampuan model dalam membedakan sentimen secara benar dan menggunakan alat-alat evaluasi utama yaitu Confusion Matrix, Akurasi, Precision, Recall, dan F1-score, untuk mendapatkan pandangan menyeluruh.

3. Hasil Pembahasan

A. Dataset

Subjek dalam penelitian ini berasal dari ulasan pengguna (*user reviews*) pada aplikasi Fore yang diperoleh melalui pendekatan kuantitatif eksperimental dengan *supervised machine learning* menggunakan metode Support Vector Machine (SVM). Tahap penelitian meliputi pengumpulan data (*scraping*), pra-pemrosesan data (*preprocessing*), pembentukan fitur (*feature engineering*), pelatihan model (*training*), dan evaluasi model (*testing*).

Data diambil dari Google Play Store pada Pengumpulan data dilakukan dengan teknik web scraping menggunakan pustaka google-play-scraper pada bahasa pemrograman Python. Jumlah ulasan pengguna aplikasi Fore yang berhasil dikumpulkan adalah 3.966

Setiap data ulasan mencakup atribut utama seperti teks ulasan (content) dan nilai *rating* (1–5 bintang). Nilai *rating* ini kemudian digunakan sebagai basis untuk pelabelan awal sentiment seperti

TABEL 1
HASIL PRA-PEMROSESAN DATA
(*DATA SCRAPPING*)

No	ReviewId	Rating	Review Text	Date
0	6086d279-db51-425a-982b-ff1d6868ed7e	5	kopi ter the best saat ini matcha nya jg mantuullll	2018-12-03
1	84fa1126-ba1d-4b37-abc2-4967cb4a93e8	5	nice	2018-12-03
2	85c00ca1-7676-4ff7-9792-11916b21b4f8	1	tambahkan fitur "batalkan pembayaran", karena beberapa outlet fore yang ramai baristanya cenderung mengutamakan pesanan offline, pesanan via apl bisa mencapai 1-2 jam, dan sangat disayangkan tidak bisan dibatalkan	2019-01-16
3	f650c9ad-bfaa-4820-b772-24aedb77ec46	5	suka mampir di sini	2020-03-10
4	6fbe7bf9-23b3-4b2b-a0db-c1eb4e9c56c8	5	mantap	2024-02-15

-	-	-	-	-
-	-	-	-	-
6597	f86c63bd-fce5-4758-bdfd-2f40bd561945	5	TAMBAHIN PEMBAYARAN QRIS	2025-11-21
6598	890301e1-b29f-46f3-b03b-359d786b6066	5	Coffee dengan harga merakyat dan kualitas rasa yang istimewa . Hampir semua golongan masyarakat bisa mencoba kualitas rasa coffee yg di sediakan	2025-11-23
6599	59225e1c-419d-418a-ad9c-6692447fb265	5	Good	2025-11-25
6600	7aaeaf05-9ce1-4ec8-b1a9-40f8414ed082	2	pdhal br download tp ga bs atur negara nya.. udh di klik2 tetep ga bs.. apa lg ber mslh ya?	2025-11-25

Tahap awal analisis dilakukan dengan melakukan preprocessing data untuk memastikan bahwa dataset yang digunakan dalam eksperimen memiliki kualitas yang baik, masukan mengenai fitur , hingga ulasan positif tentang layanan aplikasi

B. Filter Kolom

Langkah pertama dilakukan dengan memilih kolom yang relevan dari dataset hasil scraping Google Play Store. Dataset awal memiliki sejumlah atribut bawaan seperti nama pengguna, waktu unggah ulasan, serta informasi tambahan lainnya. Namun, penelitian ini hanya membutuhkan dua variabel, yaitu Rating dan Ulasan, sehingga kolom lain dieliminasi untuk menjaga efisiensi pemrosesan. Penyajian hasil penyaringan kolom ditampilkan pada

TABEL II
HASIL FILTER KOLOM

Rating	Content
1	gabisa di install, mau ngerjain tugas tpi mau kontribusi minum atau nongkrong dstu. mau pilih menu ga bisa nampil ui app ny. loading lama koneksi wifi aman bgt
2	pesanan sudah lebih dari 1 jam statusnya masih dalam pembuatan. mau dibatalkan syarat refundnya ribett
4	selalu suka ama fore, rekomendasi deh untuk semua minumannya. tempatnya juga nyaman.
1	Gak jelass banget. ni app log out" sendiri setiap mau pake vochernya. tiba" blank. pas udh beli gak melalui app tiba" app nya normall lagi. wkwkwk gak hanya sekali dua kali kek gini. sering wkwk
1	fore bikin promo lewat diskon pengguna baru, giliran dipake connect ke ovo tapi ga berhasil melakukan pembayaran, ditungguin 15menit ga berhasil juga, hampir telat masuk kantor gara2 ini
3	pakai aplikasinya sama ojol,lebih cepat ojol di handlenya.wasting time nunggu kopi sampai setengah jam-an
5	pleaseee bgt balikin lg watermelon & starycano 🙏 itu enak bgt oiii , bingung pen coba rasa itu dmna lg 😭😭 udh lope2 sama fore yg hrganya pun masih ramah di kantong
5	fore Cofee mak nyuuus👍👍 mantap 👍👍

Pada tahap ini, data hasil scraping kemudian dipetakan menjadi DataFrame dengan dua kolom inti. Kolom Rating berisi penilaian numerik pengguna terhadap aplikasi Fore, sedangkan kolom Ulasan berisi teks opini yang menjadi objek analisis sentimen. Informasi struktur data menunjukkan bahwa total terdapat 3.966 entri ulasan pelanggan Fore dan seluruh kolom tidak memiliki nilai kosong, menandakan dataset berada dalam kondisi siap untuk diproses lebih lanjut.

C. Hapus Data Duplicate

Tahap berikutnya adalah eliminasi data duplikat dari kolom Ulasan. Data ganda (duplikat) dapat timbul dari aksi pengguna yang mengunggah ulasan yang sama berulang kali atau dari proses *scraping* yang tidak efisien. Penghapusan duplikat sangat krusial untuk menjaga objektivitas model, sebab duplikasi dapat mendistorsi distribusi sentimen yang sebenarnya.

Analisis hasil menunjukkan bahwa dari total 3.966 entri data awal, teridentifikasi ulasan ganda. Setelah proses penyingkiran, *dataset* berkurang menjadi 3.733 entri. Angka ini merepresentasikan *dataset* yang telah bersih dari entri berulang dan siap untuk tahapan *preprocessing* selanjutnya:

```
<class 'pandas.core.frame.DataFrame'>
Index: 3733 entries, 0 to 3965
Data columns (total 8 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   userName  3733 non-null   object
 1   score     3733 non-null   int64
 2   at        3733 non-null   datetime64[ns, UTC+07:00]
 3   content   3733 non-null   object
 4   tanggal   3733 non-null   int64
 5   bulan     3733 non-null   int64
 6   tahun     3733 non-null   int64
 7   jam       3733 non-null   object
dtypes: datetime64[ns, UTC+07:00](1), int64(4), object(3)
memory usage: 262.5+ KB
```

Gbr.1 Hasil Data Duplicate

D. Eksplorasi Data Awal

Tahap Analisis Data Eksploratif (EDA) diimplementasikan sebelum *preprocessing* data ulasan. EDA berfungsi sebagai alat diagnostik untuk memahami struktur *dataset* dan mengidentifikasi anomali yang perlu ditangani.

Fokus utama EDA dalam studi ini meliputi:

1. Pemeriksaan Distribusi Skor: Menganalisis penyebaran skor *rating* yang diberikan pengguna.

2. Identifikasi Token Frekuensi Tinggi: Menemukan kata-kata yang paling sering muncul dalam data ulasan mentah.
3. Pencarian Pola Komentar: Mendeteksi bentuk dan gaya penulisan umum yang digunakan oleh pengguna.
4. Deteksi Elemen Non-Tekstual**:** Mengidentifikasi *noise* berupa *emoji*, karakter berulang, singkatan, dan kata-kata tidak baku.

Data yang disajikan sebagai dari EDA meliputi:

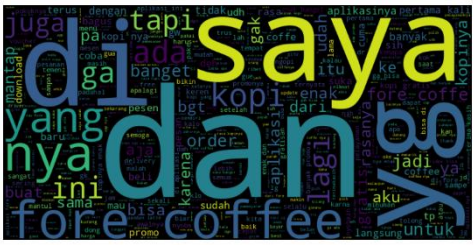
A. Wordcloud Ulasan Mentah (Sebelum *Preprocessing*)

Wordcloud dibuat sebagai representasi visual awal berdasarkan keseluruhan korpus data ulasan sebelum dilakukan proses pembersihan apapun. Visualisasi ini berfungsi sebagai alat diagnostik untuk memahami distribusi dan frekuensi kata-kata yang paling dominan digunakan oleh pengguna aplikasi Fore. Word Cloud yang dihasilkan menggarisbawahi beberapa temuan penting mengenai karakteristik data mentah:

Kosakata Dominan: Visualisasi menampilkan kata-kata dengan ukuran terbesar, yang mengindikasikan kemunculan frekuensi tertinggi. Kata-kata seperti "error", "tidak bisa", "akses", "susah", dan "memuaskan" menonjol. Dominasi kata-kata ini memberikan petunjuk awal mengenai isu fungsional (keluhan teknis) maupun sentimen positif utama yang dirasakan pengguna.

Identifikasi *Noise*: Selain kata kunci bermakna, Word Cloud juga memperlihatkan tingkat kontaminasi data (*noise*) yang tinggi. Hal ini ditunjukkan oleh banyaknya karakter non-tekstual, seperti *emoji*, singkatan, dan variasi ejaan yang tidak standar.

Wawasan Awal: Visualisasi ini membantu peneliti dalam memetakan pola keluhan yang sering diungkapkan dan mengonfirmasi bahwa data memerlukan serangkaian proses *text cleaning* yang komprehensif (seperti *stopword removal* dan *stemming*) untuk mengurangi *noise* dan meningkatkan akurasi model klasifikasi.

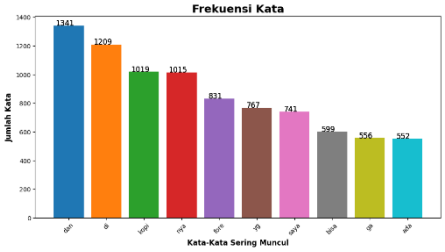


Gbr.2 Hasil *Wordcloud* (Sebelum Preprocessing)

B. Analisis Frekuensi Kata (Grafik Batang)

Selain visualisasi *Word Cloud*, penelitian ini menyajikan grafik batang (*bar chart*) untuk mendemonstrasikan frekuensi kemunculan kata-kata tunggal yang paling dominan dalam *dataset* ulasan mentah. Grafik ini memberikan gambaran kuantitatif yang lebih presisi mengenai sebaran token dibandingkan dengan representasi visual *Word Cloud*. Penggunaan grafik batang ini memiliki tiga tujuan analitis utama:

1. Mengidentifikasi Kosakata Utama Secara Kuantitatif: Grafik ini secara definitif menunjukkan 20 kata teratas (atau jumlah tertentu) yang paling sering digunakan oleh pengguna, memberikan data frekuensi numerik yang jelas.
2. Menganalisis Kecenderungan Sentimen Awal: Dengan melihat kata-kata yang menempati posisi teratas, peneliti dapat menilai kecenderungan polaritas sentimen. Misalnya, dominasi kata seperti "kopi", "order", atau "mantap" menunjukkan sentimen positif yang kuat, sementara kata seperti "error", "gagal", atau "tidak bisa" mengindikasikan dominasi isu negatif.
3. Mengukur Tingkat *Noise*: Grafik ini secara eksplisit mengungkapkan seberapa banyak kata-kata yang tidak memiliki nilai sentimen (misalnya, *stopword* seperti "di", "yang", "dan", atau "ke") muncul dalam data. Tingginya frekuensi kata-kata *noise* ini sebelum dibersihkan membuktikan urgensi dilakukannya proses Stopword Removal pada tahap *preprocessing* selanjutnya.



Gbr. 3 Diagram Frekuensi Data Ulasan Mentah

Grafik batang (*bar chart*) tidak hanya mengidentifikasi *stopword*, tetapi juga memperkuat pemahaman bahwa tahap pra-pemrosesan (*preprocessing*) sangatlah penting

dan wajib dilakukan. Data ulasan mentah terbukti mengandung banyak elemen yang tidak relevan, yang jika tidak dibersihkan akan menurunkan kinerja dan validitas model klasifikasi. Kata-kata yang tidak relevan dan *noise* yang teridentifikasi secara jelas meliputi:

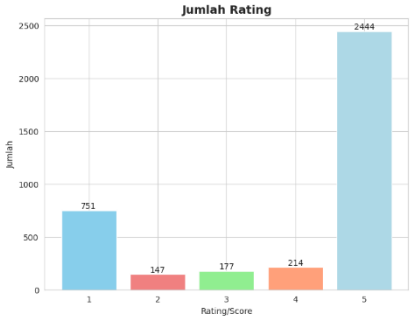
1. *Emoji*: Simbol piktogram yang tidak dapat diolah oleh metode Bag-of-Words seperti TF-IDF dan SVM.
2. Pengulangan Huruf (*Character Repetition*): Contohnya seperti penulisan "gimanaaa" atau "apaaaa", yang menunjukkan ekspresi emosi tetapi memecah token menjadi bentuk yang tidak standar.
3. Singkatan Non-Baku (*Slang*): Penggunaan kata tidak baku seperti "gk" (tidak) atau "tdk" (tidak), yang memerlukan tahap normalisasi.
4. Campuran Kata Baku dan Non-Baku: Ketidakseragaman leksikal yang mengharuskan penggunaan *stemming* atau normalisasi untuk menyatukan akar kata.

Hasil EDA ini menjadi dasar yang kuat dan empiris untuk menentukan urutan dan jenis langkah-langkah pembersihan teks yang akan diterapkan pada tahap selanjutnya. Secara keseluruhan, visualisasi *Word Cloud* dan grafik frekuensi kata telah berhasil memperjelas pola ulasan pengguna dan mengonfirmasi perlunya transformasi data yang komprehensif sebelum data dapat digunakan untuk pelatihan model *machine learning*.

C. Analisis Distribusi Jumlah *Rating*

Pemeriksaan sebaran nilai *rating* pengguna merupakan bagian integral dari Analisis Data Eksploratif (EDA) yang berfungsi memberikan gambaran awal mengenai kecenderungan penilaian terhadap aplikasi FORE di Google Play Store. Pada Gambar 4.3, disajikan diagram yang memperlihatkan frekuensi kemunculan setiap nilai *rating* (bintang 1 hingga 5). Pola distribusi menunjukkan adanya fenomena polaritas ekstrem dalam penilaian pengguna:

1. Kepuasan Puncak: *Rating* "5" mendominasi secara signifikan dengan jumlah mencapai 2.444 ulasan. Ini menegaskan bahwa mayoritas pengguna memiliki pengalaman yang sangat positif dan memuaskan.
2. Ketidakpuasan Mayor: Di sisi lain, *rating* "1" juga terekam dalam jumlah yang cukup substansial, yakni 751 ulasan. Angka ini mengindikasikan adanya kelompok pengguna yang menghadapi masalah kritis atau merasa sangat kecewa.
3. Netralitas Rendah: Sementara itu, *rating* tengah, yaitu "2" (147 ulasan), "3" (177 ulasan), dan "4" (214 ulasan), memiliki frekuensi yang jauh lebih rendah.Pola ini secara tegas menunjukkan adanya dua kutub penilaian yang kuat—pengguna cenderung berada di posisi sangat puas (bintang 5) atau sangat tidak puas (bintang 1). Sentimen ulasan bersifat sangat kontras dan beragam.



Gbr.4 Analisis Distribusi Jumlah *Rating*

D. *Preprocessing* Data

Pada penelitian *Analisis Sentimen Pengguna Aplikasi FORE melalui Komentar di Google Play Store menggunakan metode SVM*, tahap *text preprocessing* merupakan bagian yang sangat penting sebelum data teks digunakan pada proses pelatihan model. Komentar pengguna yang diperoleh melalui Google Play Store umumnya memiliki karakteristik tidak terstruktur, mengandung singkatan, emotikon, tanda baca berlebih, hingga *noise* seperti URL dan simbol. Agar data tersebut dapat diolah menggunakan algoritma Support Vector Machine (SVM), maka teks harus dibersihkan dan diubah ke bentuk representasi numerik yang sesuai dengan Gambar berikut:

Gbr. 5 Hasil *Praprocessing* Data Ulasan

Setelah seluruh tahapan *preprocessing* diterapkan, jumlah data yang semula berjumlah 3.966 ulasan berkurang menjadi 3.733 ulasan yang valid dan dapat digunakan untuk pemodelan. Proses *preprocessing* meliputi :

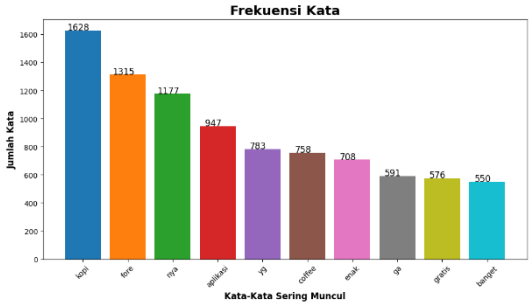
1. Pembersihan teks,
2. Normalisasi kata tidak baku,
3. Penghapusan *stopword*,
4. Tokenisasi,
5. Dan *stemming*.

Sebagai bagian dari EDA sederhana pada tahap *preprocessing*, ditampilkan dua visualisasi berikut :

1. *Wordcloud* setelah *preprocessing*, yang menunjukkan kata-kata dominan seperti *coffe*, aplikasi, kopi, mantap, saldo, dan banget.

Gbr.6 *Wordcloud* Setelah *Preprocessing*

2. Grafik frekuensi kata, yang menampilkan kata-kata dengan kemunculan terbanyak setelah teks dibersihkan dan di*stemming*.



Gbr.7 Diagram Frekuensi (Setelah Preprocessing)

Secara kolektif, kedua visualisasi ini (Word Cloud dan Grafik Frekuensi) berhasil menyediakan wawasan awal yang berharga mengenai topik utama dan isu-isu yang paling sering dibicarakan dalam ulasan pengguna.

E. Pelabelan Data

Proses pelabelan data pada penelitian ini dilakukan menggunakan metode *lexicon-based*, di mana setiap kata pada hasil *preprocessing* dibandingkan dengan daftar kata positif dan negatif pada kamus sentimen.

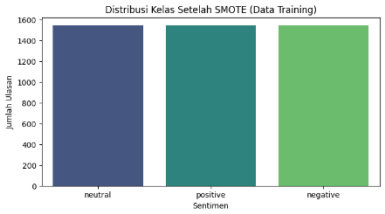
Dari total 3.966 ulasan awal, dilakukan penyaringan menggunakan metode *dropna* untuk menghapus data kosong atau tidak valid. Setelah tahap ini, tersisa 3.733 data yang digunakan untuk proses pelabelan.

Berikut merupakan contoh hasil pelabelan data pada lima ulasan pertama:

Gbr.8 Hasil Pelabelan Data Ulasan Aplikasi

Hasil pelabelan divisualisasikan menggunakan diagram batang untuk menunjukkan distribusi sentimen secara keseluruhan. Diagram tersebut menampilkan bahwa :

1. Sentimen Positif: 1.547 data
2. Sentimen Netral 1.547 data
3. Sentimen Negatif 1.547 data



Gbr. 9 Diagram Kelas Sentimen Ulasan

Visualisasi ini memperlihatkan bahwa persebaran sentimen hampir seimbang, dengan sedikit dominasi ulasan positif. Untuk memahami pola kata pada masing-masing kategori, dibuat *wordcloud* terpisah untuk sentimen positif dan negatif.

1. *Wordcloud* Sentimen Positif

Wordcloud kategori positif didominasi kata-kata seperti: memuaskan, berhasil, mudah, cepat, bagus, terbantu, dan baik. Hal ini menunjukkan sebagian pengguna merasa aplikasi sangat puas dengan pelayanan yang diberikan Fore

Gbr.10 *Wordcloud* Sentimen Positif

2. Wordcloud Sentimen Negatif Wordcloud kategori negatif menampilkan kata dominan seperti: error, akses, gagal, tidak bisa, email, verifikasi, login, dan lemot. Kata-kata tersebut menegaskan bahwa sebagian pengguna mengalami kendala

teknis terutama terkait akses dan autentikasi

Gbr.10 *Wordcloud* Sentimen Negatif

F. Pembagian Data

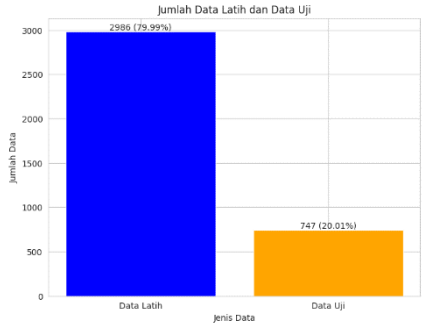
Setelah melalui tahap *preprocessing* dan pelabelan, data kemudian dibagi menjadi data latih dan

data uji untuk proses pemodelan menggunakan algoritma *Support Vector Machine* (SVM). Pembagian dilakukan dengan rasio 80% data latih dan 20% data uji. Pemilihan rasio 80:20 didasarkan pada pertimbangan bahwa proporsi tersebut merupakan standar yang umum digunakan dalam penelitian *machine learning*, karena memberikan jumlah data latih yang cukup besar untuk mempelajari pola sentimen secara optimal, tanpa mengurangi proporsi data uji yang diperlukan untuk mengukur performa model secara objektif.

Dari total 3.733 data, diperoleh hasil pembagian sebagai berikut:

1. Jumlah Data Latih : 2.986
2. Jumlah Data Uji : 747

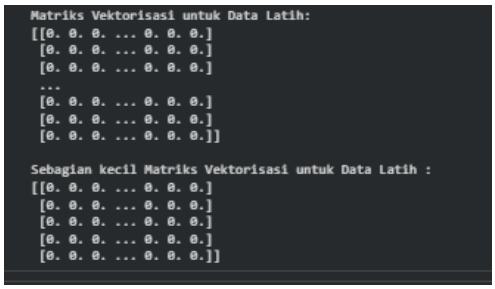
Perbandingan jumlah tersebut juga divisualisasikan menggunakan diagram batang (bar chart) untuk menunjukkan proporsi data yang digunakan pada masing-masing tahap pemodelan.



Gbr.11 Diagram Data Latih dan Data Uji

G. Vektorisasi Teks Menggunakan TF-IDF

Tahap berikutnya adalah proses vektorisasi teks menggunakan TF-IDF Vectorizer. Metode ini mengubah setiap ulasan menjadi representasi numerik berupa bobot TF-IDF untuk setiap kata dalam korpus. Hasil vektorisasi berupa matriks dengan ukuran (9142 × n_features) untuk data latih, di mana n_features merupakan jumlah total kata unik setelah *preprocessing*. Secara umum, nilai matriks didominasi angka nol karena sebagian besar kata tidak muncul pada setiap dokumen (sparse matrix)Berikut merupakan tampilan bentuk Matriks Vektorisasi untuk Data Latih:



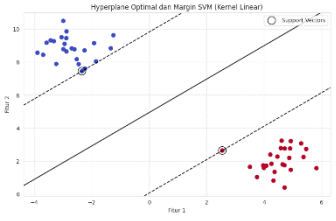
```
Matriks Vektorisasi untuk Data Latih:
[[0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 ...
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]]

Sebagian kecil Matriks Vektorisasi untuk Data Latih :
[[0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]]
```

Gbr.13 Matriks Vektorisasi Data Latih

H. Pemodelan SVM

Setelah proses vektorisasi selesai, dilakukan pelatihan model menggunakan *Support Vector Machine* (SVM) dengan kernel linear. Pemilihan kernel linear digunakan karena data teks umumnya memiliki dimensi tinggi dan bersifat linearly separable, sehingga kernel linear lebih efisien serta memberikan performa yang baik pada tugas klasifikasi sentimen. Model kemudian dilatih menggunakan data latih yang telah direpresentasikan dalam bentuk vektor TF-IDF.



Gbr.14 Visualisasi *Hyperlane* SVM

Visualisasi *hyperplane* pada pemodelan *Support Vector Machine* (SVM) menggunakan Kernel Linear menunjukkan bahwa algoritma mampu mengidentifikasi batas keputusan yang optimal untuk memisahkan sentimen ulasan aplikasi Fore menjadi dua kelas (positif dan negatif). Garis pemisah (hyperplane) utama berhasil memisahkan sebagian besar data secara linier, dengan margin yang jelas dan terdefinisi baik oleh dua garis putus-putus. Titik-titik data yang berperan krusial dalam menentukan batas keputusan ini adalah *Support Vectors* (ditandai dengan lingkaran tebal), yang menegaskan posisi hyperplane dan margin. Keterpisahan linier yang terlihat dalam plot ini mengindikasikan bahwa fitur-fitur yang digunakan (Fitur 1 dan Fitur 2) telah efektif

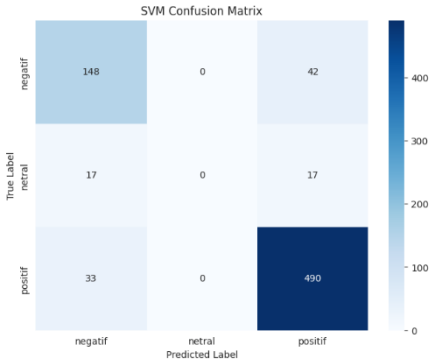
merepresentasikan karakteristik ulasan, memungkinkan model SVM Kernel Linear mencapai kinerja klasifikasi yang kuat dan mudah diinterpretasikan.

I. Evaluasi Model SVM

Evaluasi performa dilakukan setelah model *Support Vector Machine* (SVM) dengan kernel linear selesai dilatih menggunakan data latih dan diuji menggunakan data uji sebanyak 3.733 data. Evaluasi dilakukan menggunakan confusion matrix, akurasi, serta metrik *precision*, *recall*, dan *F1-score*. Hasil confusion matrix menunjukkan distribusi prediksi model terhadap kelas Positif dan Negatif sebagai berikut :

Gbr.15 Confussion Matrix SVM

Berdasarkan matriks tersebut, model mampu mengklasifikasikan 490 data Positf



dengan benar dan salah mengklasifikasikan 148 data sebagai Negatif. Pada kelas Positif, model berhasil mengklasifikasikan 490 data dengan benar dan melakukan kesalahan prediksi sebanyak 148 data. Secara umum, model menunjukkan kinerja yang cukup baik dalam membedakan kedua kelas.

Akurasi keseluruhan model adalah:

1. SVM Accuracy : 0.8540829986613119
2. SVM Accuracy : 85.41%

Nilai akurasi tersebut menunjukkan bahwa model mampu memberikan prediksi yang benar pada sekitar 80% data uji. Untuk melihat performa yang lebih detail pada masing-masing kelas, berikut adalah hasil *classification report*:

```
SVM Accuracy: 0.8540829986613119

SVM Accuracy: 85.41%

SVM Classification Report:
              precision    recall  f1-score   support

    negative       0.75      0.78      0.76       190
     neutral       0.00      0.00      0.00        34
    positive       0.89      0.94      0.91       523

    accuracy                           0.85       747
   macro avg       0.55      0.57      0.56       747
weighted avg       0.82      0.85      0.83       747
```

Gbr.16 Hasil *classification report*

Dari hasil tersebut, terlihat bahwa nilai *precision* dan *recall* untuk kedua kelas berada pada kisaran 0.78–0.94, yang menunjukkan perbedaa performa antara kemampuan model mengenali kelas Negatif dan Positif. Nilai *F1-score* sebesar 0.76-0.91 pada kedua kelas mengindikasikan bahwa model memiliki kemampuan klasifikasi yang berbeda. Secara keseluruhan, hasil evaluasi ini menunjukkan bahwa model SVM dengan kernel linear mampu melakukan klasifikasi sentimen dengan performa yang sangat baik dan dapat diandalkan untuk menganalisis sentimen pada ulasan pengguna aplikasi FORE.

4. Kesimpulan

Berdasarkan analisis sentimen pengguna aplikasi Fore melalui komentar di Google Play Store dengan metode SVM, preprocessing data berhasil dilakukan dengan langkah-langkah seperti cleaning, case folding, tokenisasi, removal of stopwords, dan stemming. Dari 6.596 data, 6.472 dapat digunakan setelah pembersihan. Ulasan dikategorikan menjadi positif (60,41%) dan negatif (10,83%), menunjukkan perspektif pengguna yang beragam. Model SVM dengan kernel linear menghasilkan akurasi 81,99%, menunjukkan kemampuan dalam mengklasifikasikan sentimen dengan baik. Kesimpulannya, tahapan preprocessing dan metode lexicon-based labeling dapat mempersiapkan data dengan optimal, sementara SVM memberikan kinerja klasifikasi yang efisien untuk analisis sentimen aplikasi FORE.

Daftar Pustaka

[1] M. N. Akbar, N. Hasanahlmar, and M. H. Hasrul, “Sentiment Analysis Terhadap Review Aplikasi Maxim di Google Play Store Menggunakan Support Vector Machine (SVM),” vol. 2, no. 2, 2022.

[2] T. Akhir and R. M. Komaran, “Model klasifikasi sentimen berita online menggunakan metode tf-idf + svm dan fine-tuned indobert,” 2025.

[3] H. Allam and F. Alrowais, “Text Classification: How Machine Learning Is Transforming the Field,” Information, vol. 16, no. 2, p. 130, 2025, doi: 10.3390/info16020130.

[4] M. N. Arifin, A. Hamzah, M. A. Huda, and N. Hasanah, “Analysis of Google Play Store User Sentiment Towards Application X Using the SVM Algorithm,” vol. 5, no. 1, pp. 249–258, 2025.

[5] M. Bordoloi, S. Misra, and A. Raman, “Sentiment analysis: A survey on design framework and future directions,” PLoS ONE, 2023, doi: 10.1371/journal.pone.10026245.

[6] J. Dąbrowski, E. Letier, and others, "Analysing app reviews for software engineering: a systematic literature review," Empirical Software Engineering, 2021, doi: 10.1007/s10664-021-10065-7.

[7] E. Damayanti, A. V. Vitianingsih, S. Kacung, and D. Cahyono, "Sentiment Analysis of Alfagift Application User Reviews Using Long Short-Term Memory (LSTM) and Support Vector Machine (SVM) Methods," vol. 4, no. 2, pp. 509–521, 2024.

[8] G. Ferdinands et al., "Performance of active learning models for screening prioritization in systematic reviews: a simulation study into the Average Time to Discover relevant records," Systematic Reviews, vol. 12, no. 1, p. 100, 2023, doi: 10.1186/s13643-023-02257-7.

[9] J.-J. Kim, S. Lee, and M. Choi, "A Comparative Study of Term-Weighting Schemes for Text Classification," Journal of Information Processing & Management, 2022.

[10] P. Kurniawati, R. Yanu, and D. Witarsyah, "Sentiment Analysis of Maxim Online Transportation App Reviews using Support Vector Machine (SVM) Algorithm," vol. 5, no. 2, pp. 466–475, 2023, doi: 10.47065/bits.v5i2.4265.

[11] W. Macias, K. Rodriguez, and H. Barriga, "Determinants of satisfaction with online food delivery providers and their impact on restaurant brands," vol. 14, no. 4, pp. 557–578, 2025, doi: 10.1108/JHTT-04-2021-0117.

[12] Q. Motger, M. Oriol, M. Tiessler, X. Franch, and J. Marco, "What About Emotions? Guiding Fine-Grained Emotion Extraction from Mobile App Reviews," Empirical Software Engineering, 2025.

[13] M. J. Prasetyo and I. M. A. Agastya, "Analisis Sentimen Ulasan Aplikasi Perbankan di Google Play Store menggunakan Algoritma Support Vector Machine," vol. 13, pp. 2386–2400, 2024.

[14] S. Roth et al., "Machine Learning Methods for Systematic Reviews: A Rapid Scoping Review," Systematic Reviews, 2023, doi: 10.1186/s13643-023-02257-7.

[15] A. Saeidmehr, P. David, G. Steel, and F. F. Samavati, "Systematic review using a spiral approach with machine learning," Systematic Reviews, pp. 1–17, 2024, doi: 10.1186/s13643-023-02421-z.

[16] G. J. Sagala and Y. T. Samuel, "Sentiment Analysis on ChatGPT App Reviews on Google Play Store Using Random Forest Algorithm, Support Vector Machine and Naïve Bayes," Int. J. Eng. Business and Social Science, vol. 2, no. 04, pp. 1194–1204, 2024.

[17] H. Soesanto, "Behavioral Intention Using Online Food Delivery Service: Information Technology," vol. 24, no. 3, pp. 1186–1196, 2023.

[18] Q. Su, A. Namin, and S. Ketron, "The effect of online company responses on app review quality," Journal of Consumer Marketing, 2024, doi: 10.1108/JCM-06-2023-6098.

[19] L. Sun, Y. He, and F. Fu, "Impact of online negative and positive reviews on app innovation," pp. 1–18, 2025, doi: 10.1371/journal.pone.0323205.

[20] A. M. Williams, N. Ansai, N. Ahluwalia, and D. T. Nguyen, "Anemia Prevalence: United States, August 2021–August 2023," NCHS Data Brief, no. 519, pp. 1–8, 2024.

[21] Y. Zhai, X. Song, Y. Chen, and W. Lu, "A study of mobile medical app user satisfaction incorporating theme analysis and review sentiment tendencies," Int. J. Environ. Res. Public Health, vol. 19, p. 7466, 2022, doi: 10.3390/ijerph19127466.