

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian yang Relevan

Penelitian terdahulu dipelajari untuk mengetahui perkembangan penelitian pada bidang yang sama sekaligus mengidentifikasi pendekatan yang telah digunakan oleh peneliti sebelumnya. Melalui kajian tersebut dapat diketahui perbedaan maupun persamaan dari sisi objek penelitian, metode yang diterapkan, serta hasil yang diperoleh. Informasi tersebut menjadi acuan dalam menentukan posisi penelitian sekaligus menunjukkan kebaruan penelitian yang dilakukan. Beberapa penelitian yang memiliki keterkaitan dengan analisis sentimen menggunakan metode *Support Vector Machine* (SVM) disajikan sebagai berikut.

Styawati, Hendrastuty, dan Isnain (2021) mengkaji analisis sentimen masyarakat terhadap Program Kartu Prakerja menggunakan algoritma *Support Vector Machine*. Penelitian tersebut memperlihatkan bahwa tahapan *preprocessing*, seperti normalisasi kata, tokenisasi, dan *stemming*, mampu meningkatkan kualitas data teks berbahasa Indonesia sehingga berdampak pada peningkatan performa model klasifikasi. Selain itu, pembobotan TF-IDF dinilai mampu merepresentasikan fitur teks secara lebih baik sebelum proses klasifikasi dilakukan (Styawati et al., 2021).

Kurniawati, Fa'rifah, dan Witarsyah (2023) memanfaatkan metode *Support Vector Machine* untuk mengklasifikasikan sentimen ulasan aplikasi Maxim yang diperoleh dari Google Play Store. Data yang dianalisis berjumlah 1.076 ulasan hasil *web scraping* dan diuji menggunakan tiga skenario pembagian data, yaitu 60:40, 70:30, dan 80:20. Melalui optimasi *Grid search* pada *kernel* Linear, konfigurasi 70:30 memberikan hasil terbaik dengan akurasi 89,82%, presisi 92,66%, *recall* 94,09%, serta *F1-Score* 93,38% (Kurniawati et al., 2023).

Akbar, Rusydi, Hasrul, Ramadhanti, dan Erfiana (2022) mengangkat ulasan pengguna aplikasi Maxim sebagai objek penelitian. Sebanyak 1.200 ulasan dikumpulkan dari Google Play Store melalui teknik *web scraping*. Tahap *preprocessing* dilakukan menggunakan tokenisasi dan *label encoding*, sedangkan penentuan label sentimen didasarkan pada nilai *rating* pengguna. Model SVM yang dibangun mengelompokkan data ke dalam dua kelas

sentimen dan menghasilkan akurasi sebesar 79% (Akbar et al., 2022).

Fremmuzar & Baita, 2023 mengkaji pengaruh penggunaan beberapa jenis *kernel* terhadap performa algoritma *Support Vector Machine* (SVM). Penelitian tersebut membandingkan *kernel* Linear, RBF, dan Polynomial dalam klasifikasi sentimen layanan Telkomsel di Twitter. Hasil pengujian menunjukkan bahwa pemilihan *kernel* yang sesuai berpengaruh terhadap performa model sehingga perlu disesuaikan dengan karakteristik data yang digunakan Fremmuzar & Baita, 2023.

Rabbani et al., 2023 mengevaluasi beberapa jenis *kernel* SVM menggunakan TF-IDF sebagai representasi fitur. Penelitian tersebut menemukan bahwa *kernel* Linear lebih sesuai untuk data teks berdimensi tinggi, sedangkan *kernel* RBF menunjukkan hasil yang lebih baik pada data dengan pola nonlinier. Temuan tersebut menjadi dasar penting dalam proses optimasi parameter menggunakan *Grid search*.

Putri Husain et al., 2024 menggabungkan TF-IDF dan SVM untuk mengklasifikasikan sentimen ulasan aplikasi TikTok di Google Play Store. Berdasarkan hasil evaluasi, model yang dibangun memperoleh akurasi 85%, presisi 79%, *recall* 75%, dan *F1-Score* 77%, sehingga kombinasi kedua metode dinilai cukup efektif dalam pengolahan ulasan pengguna.

(Farhan Wiguna et al., 2023) menganalisis sentimen ulasan pengguna aplikasi Gojek menggunakan tahapan *text preprocessing*, pembobotan TF-IDF, serta algoritma Support Vector Machine (SVM) dan Naïve Bayes. Penelitian tersebut menunjukkan bahwa TF-IDF mampu merepresentasikan fitur teks secara efektif, sedangkan SVM memberikan kinerja yang baik dalam proses klasifikasi sentimen ulasan aplikasi berbahasa Indonesia.

Nouval et al., 2026 mengembangkan model analisis sentimen terhadap ulasan aplikasi Maxim menggunakan pendekatan yang paling mendekati penelitian ini. Proses yang diterapkan meliputi *cleaning*, *case folding*, normalisasi kata, *tokenizing*, *stopword removal*, *stemming*, pelabelan *lexicon-based*, serta ekstraksi fitur TF-IDF. Berdasarkan beberapa skenario pengujian, SVM memperlihatkan performa yang lebih stabil dibandingkan Naïve Bayes dengan tingkat akurasi antara 0,77 hingga 0,81.

Wahyudi et al., 2021 menerapkan SVM untuk mengklasifikasikan ulasan aplikasi Grab pada Google Play Store ke dalam kategori positif dan negatif. Kesamaan sumber data serta objek penelitian menjadikan studi tersebut relevan sebagai referensi dalam penerapan SVM pada analisis sentimen aplikasi transportasi.

Dometian & Jatmiko, 2025 membandingkan performa SVM, Naïve Bayes, dan *Logistic Regression* dalam klasifikasi sentimen ulasan Google Play Store. Dengan dukungan TF-

IDF dan optimasi *hyperparameter*, SVM memberikan hasil terbaik pada *dataset* GoPay sekaligus mempertahankan performa yang konsisten pada *dataset* LinkAja. Temuan tersebut memperkuat penggunaan SVM sebagai salah satu algoritma yang andal untuk klasifikasi sentimen berbasis teks.

Berdasarkan hasil telaah terhadap penelitian terdahulu, terlihat bahwa Support Vector Machine (SVM) merupakan salah satu metode yang paling banyak digunakan dalam analisis sentimen berbasis teks. Penelitian oleh Nouval et al., 2026, Akbar et al., 2022, Kurniawati et al., 2023, dan Wahyudi et al., 2021 memiliki keterkaitan yang paling dekat dengan penelitian ini karena sama-sama menggunakan ulasan aplikasi transportasi pada Google Play Store sebagai sumber data. Sementara itu, penelitian Putri Husain et al., 2024 menunjukkan kesamaan dalam penggunaan kombinasi TF-IDF dan SVM untuk mengolah ulasan aplikasi. Dari sisi pengembangan metode, Rabbani et al., 2023 menegaskan pentingnya pemilihan *kernel* yang sesuai untuk meningkatkan performa klasifikasi, sedangkan Dometian & Jatmiko, 2025 memperlihatkan bahwa SVM mampu menghasilkan performa yang lebih baik dibandingkan Naïve Bayes dan *Logistic Regression*. Berbagai temuan tersebut menjadi dasar dalam pemilihan SVM sebagai metode klasifikasi pada penelitian ini.

Meskipun SVM telah banyak diterapkan pada penelitian analisis sentimen berbasis Google Play Store, penelitian ini memiliki sejumlah perbedaan mendasar dibandingkan penelitian-penelitian tersebut. Dari sisi metode pelabelan, sebagian besar penelitian terdahulu menggunakan pelabelan berbasis rating atau lexicon-based generik tanpa validasi tambahan; (Dometian & Jatmiko, 2025) bahkan menemukan bahwa pelabelan lexicon-based murni menghasilkan performa model yang lebih rendah (rentang akurasi 74–81%) dibandingkan pelabelan berbasis rating (89–90%) pada objek penelitian mereka. Penelitian ini merespons temuan tersebut dengan menerapkan pelabelan lexicon-based menggunakan kamus sentimen kustom yang dikembangkan khusus untuk konteks Green SM, disertai validasi manual menyeluruh terhadap seluruh data berlabel netral, sehingga sensitivitas konteks yang menjadi kelebihan pendekatan lexicon-based tetap dipertahankan sekaligus kelemahan akurasinya diminimalkan. Dari sisi optimasi model, penelitian ini menerapkan Grid search CV dengan skema Stratified K-Fold Cross-Validation secara sistematis, berbeda dengan sebagian penelitian terdahulu yang tidak menerapkan tuning parameter maupun menggunakan pendekatan pencarian acak (Random Search) seperti pada (Dometian & Jatmiko, 2025). Dari sisi objek penelitian, mayoritas penelitian terdahulu berfokus pada aplikasi transportasi maupun dompet digital

konvensional yang sudah banyak dikaji, sedangkan penelitian ini secara khusus membahas aplikasi Green SM sebagai layanan transportasi berbasis kendaraan listrik yang masih relatif terbatas diteliti. Kombinasi ketiga perbedaan tersebut — metode pelabelan, strategi optimasi parameter, dan objek penelitian — menjadikan penelitian ini berbeda dari penelitian-penelitian SVM sejenis yang telah ada.

Hasil penelusuran terhadap penelitian-penelitian tersebut juga memperlihatkan bahwa masih terdapat ruang untuk pengembangan. Sebagian besar penelitian memanfaatkan data yang berasal dari media sosial, terutama Twitter, sedangkan penelitian yang menggunakan ulasan pengguna dari Google Play Store masih belum banyak dilakukan. Selain itu, objek yang dikaji umumnya berfokus pada aplikasi transportasi konvensional, seperti Grab dan Maxim, ataupun aplikasi digital lainnya. Penelitian yang secara khusus membahas sentimen pengguna terhadap layanan transportasi berbasis kendaraan listrik, terutama Green SM, masih relatif terbatas. Padahal, ulasan yang tersedia di Google Play Store umumnya berasal dari pengguna yang telah menggunakan aplikasi sehingga mampu memberikan gambaran yang lebih rinci mengenai kualitas layanan, performa aplikasi, pengalaman penggunaan, serta tingkat kepuasan pengguna (Fitrana et al., 2024). Perbedaan karakteristik antar sumber data tersebut turut memberikan implikasi terhadap kualitas data yang digunakan dalam proses klasifikasi sentimen. Ulasan pada Google Play Store, mengingat ditulis oleh pengguna yang telah benar-benar mengunduh dan menggunakan aplikasi, cenderung memuat opini yang lebih spesifik terhadap fitur maupun performa aplikasi yang bersangkutan, sehingga proses pelabelan sentimen berbasis lexicon dapat dilaksanakan dengan tingkat akurasi yang lebih baik karena konteks pembicaraan telah terfokus pada objek penelitian (Wahyudi et al., 2021). Sebaliknya, data yang bersumber dari media sosial seperti Twitter/X berpotensi memuat opini yang lebih beragam dan tidak selalu berkaitan langsung dengan pengalaman penggunaan aplikasi tertentu, sehingga berisiko menimbulkan derau (noise) pada proses pelabelan maupun ekstraksi fitur (Fitrana et al., 2024). Dengan demikian, pemilihan sumber data tidak semata-mata bergantung pada ketersediaannya, melainkan turut memengaruhi relevansi dan akurasi label sentimen yang dihasilkan, yang pada akhirnya berdampak terhadap performa model klasifikasi secara keseluruhan.

Berdasarkan kondisi tersebut, penelitian ini menawarkan pendekatan yang berbeda dengan memanfaatkan *Support Vector Machine* (SVM) dan pembobotan fitur *Term Frequency–Inverse Document Frequency* (TF-IDF) untuk mengklasifikasikan sentimen

ulasan pengguna Green SM yang diperoleh dari Google Play Store. Perbedaan utama penelitian ini terletak pada pemilihan objek, yaitu aplikasi transportasi berbasis kendaraan listrik yang masih jarang menjadi fokus penelitian. Selain itu, pelabelan sentimen dilakukan menggunakan kamus sentimen kustom yang disusun khusus untuk konteks layanan Green SM, yang kemudian divalidasi secara manual agar hasil pelabelan lebih akurat dan sesuai dengan karakteristik ulasan pada domain transportasi berbasis kendaraan listrik. Penelitian ini juga menerapkan optimasi *hyperparameter* melalui *Grid search Cross Validation* untuk memperoleh konfigurasi model yang memberikan performa terbaik. Dengan pendekatan tersebut, penelitian ini diharapkan dapat memberikan tambahan referensi pada bidang analisis sentimen sekaligus memperluas kajian mengenai ulasan aplikasi transportasi berbasis kendaraan listrik di Indonesia.

Tabel 2. 1 Ringkasan Penelitian Yang Relevan

No.	Nama Peneliti	Judul Penelitian	Metode Yang Digunakan	Hasil
1.	Styawati et al. (2021)	Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode <i>Support Vector Machine</i>	<i>SVM, TF-IDF, Preprocessing</i>	<i>Preprocessing</i> (normalisasi, <i>stemming</i>) meningkatkan kualitas data dan performa klasifikasi sentimen teks berbahasa Indonesia secara signifikan.
2.	Kurniawati et al. (2023)	<i>Sentiment Analysis of Maxim Online Transportation App Reviews using Support Vector Machine (SVM) Algorithm</i>	<i>SVM, Grid search</i>	<i>SVM</i> efektif mengidentifikasi persepsi pengguna; kamus sentimen berbasis domain lebih akurat dibandingkan kamus umum.
3.	Akbar et al. (2022)	<i>Sentiment Analysis Terhadap Review Aplikasi Maxim di Google Play Store Menggunakan SVM</i>	<i>SVM</i>	<i>SVM</i> berhasil mengklasifikasikan 1.200 ulasan transportasi <i>online Maxim di Google Play Store</i> dengan akurasi 79%.

No.	Nama Peneliti	Judul Penelitian	Metode Yang Digunakan	Hasil
4.	Fremmuzar & Baita (2023)	Uji <i>Kernel SVM</i> dalam Analisis Sentimen Terhadap Layanan <i>Telkomsel</i> di <i>Twitter</i>	<i>Multi-Kernel SVM (Linear, RBF, Poly)</i>	Pemilihan <i>kernel</i> berpengaruh signifikan terhadap performa <i>SVM</i> ; menjadi dasar penerapan <i>Grid search</i> untuk optimasi <i>kernel</i> dalam penelitian ini.
5.	Rabbani et al. (2023)	Perbandingan Evaluasi <i>Kernel SVM</i> untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM	<i>SVM, TF-IDF</i>	<i>Kernel</i> linear kompetitif pada data <i>TF-IDF</i> berdimensi tinggi; <i>kernel RBF</i> unggul pada pola nonlinier. Mendukung penggunaan <i>Grid search</i> .
6.	Putri Husain et al. (2024)	Analisis Sentimen Ulasan Pengguna <i>TikTok</i> pada <i>Google Play Store</i> Berbasis <i>TF-IDF</i> dan <i>SVM</i>	<i>TF-IDF, SVM</i>	<i>TF-IDF + SVM</i> pada ulasan <i>Google Play Store</i> berbahasa Indonesia: akurasi 85%, presisi 79%, <i>recall</i> 75%, <i>F1-Score</i> 77%.
7.	Farhan Wiguna et al. (2023)	<i>Sentiment Analysis of On-Demand Ride-Hailing Systems using Support Vector Machine and Naïve Bayes</i>	<i>TF-IDF, SVM, Naïve Bayes</i>	<i>TF-IDF</i> menghasilkan representasi fitur yang efektif, sedangkan <i>SVM</i> memberikan kinerja yang baik dalam klasifikasi sentimen ulasan pengguna aplikasi.
8.	Nouval et al. (2026)	Analisis Sentimen Ulasan Aplikasi Maxim di <i>Google Play Store</i> Menggunakan Algoritma <i>SVM</i>	<i>TF-IDF, SVM, Naive Bayes</i>	Data 400 ulasan Maxim melalui pipeline <i>preprocessing</i> 6 tahap dan pelabelan <i>lexicon based</i> + validasi manual, diuji pada 3 skenario data dan 5-fold CV. <i>SVM</i> lebih stabil dari NB (akurasi 0,77-0,81).
9.	Wahyudi et al., 2021	Analisis Sentimen pada Aplikasi <i>Grab</i> di <i>Google Play Store</i> Menggunakan <i>SVM</i>	<i>SVM</i>	<i>SVM</i> efektif mengklasifikasikan sentimen ulasan aplikasi Grab pada <i>Google Play Store</i> ; mendukung penerapan <i>SVM</i> pada domain aplikasi transportasi berbasis ulasan pengguna.

No.	Nama Peneliti	Judul Penelitian	Metode Yang Digunakan	Hasil
10.	Dometian & Jatmiko (2025)	Analisis Sentimen Ulasan Google Play Store: Studi Komparatif Algoritma SVM, Naïve Bayes, dan Logistic Regression	<i>TF-IDF, Support Vector Machine (SVM), Naïve Bayes, Logistic Regression</i>	SVM memberikan performa terbaik dibandingkan Naive Bayes dan Logistic Regression dalam klasifikasi sentimen ulasan Google Play Store.

Penelitian-penelitian terdahulu menunjukkan bahwa metode Support Vector Machine (SVM) merupakan salah satu algoritma yang banyak digunakan dalam analisis sentimen karena memiliki kemampuan klasifikasi yang baik, khususnya pada data teks berbahasa Indonesia. Selain itu, metode *Term Frequency–Inverse Document Frequency* (TF-IDF) juga sering digunakan sebagai teknik pembobotan kata karena mampu merepresentasikan data teks ke dalam bentuk numerik sehingga memudahkan proses pengolahan data oleh sistem komputasi.

Sebagian besar penelitian sebelumnya menggunakan data yang bersumber dari media sosial, seperti Twitter dan berbagai platform digital lainnya. Sementara itu, penelitian yang memanfaatkan ulasan pengguna aplikasi pada Google Play Store masih relatif terbatas. Di sisi lain, penelitian yang mengombinasikan pelabelan sentimen berbasis lexicon, validasi manual, TF-IDF, dan SVM pada domain aplikasi transportasi berbasis kendaraan listrik juga masih jarang ditemukan.

Berdasarkan kondisi tersebut, penelitian ini dilakukan untuk menganalisis sentimen ulasan pengguna aplikasi Green SM pada Google Play Store dengan menggunakan metode Support Vector Machine (SVM) dan pembobotan TF-IDF. Penelitian ini diharapkan dapat memberikan gambaran mengenai persepsi pengguna terhadap layanan Green SM, sekaligus menjadi referensi bagi pengembangan penelitian analisis sentimen pada bidang transportasi berbasis kendaraan listrik di Indonesia.

2.2 Landasan Teori

Bagian ini memuat landasan teori yang digunakan sebagai dasar dalam penelitian. Penyusunannya dilakukan melalui kajian terhadap berbagai sumber ilmiah, seperti buku, jurnal, dan penelitian terdahulu yang relevan dengan topik yang diteliti.

Pada bagian ini, dibahas beberapa konsep utama yang mendukung penelitian, meliputi Google Play Store, *text mining*, *Natural Language Processing* (NLP), analisis sentimen,

serta metode klasifikasi yang digunakan. Keseluruhan teori tersebut menjadi dasar konseptual untuk menjelaskan setiap tahapan penelitian secara terstruktur, mulai dari proses awal hingga diperolehnya hasil penelitian.

2.2.1 Ulasan Pengguna

Ulasan pengguna adalah bentuk umpan balik yang diberikan setelah seseorang menggunakan suatu aplikasi atau layanan. Umpan balik tersebut biasanya berupa pemberian nilai (*rating*) dan komentar yang berisi pengalaman, pendapat, kritik, maupun saran terhadap layanan yang telah digunakan.

Menurut (Wahyudi et al., 2021), ulasan pada Google Play Store terdiri atas rating dan komentar tekstual sebagai bentuk umpan balik pengguna terhadap aplikasi. Rating menunjukkan penilaian pengguna dalam bentuk numerik, sedangkan komentar tekstual memuat informasi yang lebih rinci mengenai pengalaman pengguna, seperti kritik, saran, maupun keluhan terhadap aplikasi. Oleh karena itu, komentar tekstual menjadi sumber informasi yang penting karena mampu menjelaskan alasan yang mendasari pemberian rating. (Kurniawati et al., 2023) juga menjelaskan bahwa analisis terhadap ulasan pengguna memberikan informasi yang lebih mendalam mengenai pengalaman dan persepsi pengguna dibandingkan hanya melihat nilai rating, sehingga hasil analisis dapat dimanfaatkan sebagai bahan evaluasi untuk meningkatkan kualitas aplikasi.

Pada Google Play Store, ulasan pengguna menjadi salah satu sumber data yang sering dimanfaatkan dalam penelitian analisis sentimen. Melalui ulasan tersebut, peneliti dapat memperoleh informasi mengenai pengalaman pengguna saat menggunakan aplikasi, mulai dari kualitas layanan, performa aplikasi, hingga tingkat kepuasan pengguna. Informasi tersebut dapat dimanfaatkan untuk mengidentifikasi aspek yang masih perlu diperbaiki sekaligus menjadi bahan evaluasi bagi pengembang aplikasi (Dometian & Jatmiko, 2025; Fitriana et al., 2024).

Dalam penelitian ini, ulasan pengguna aplikasi Green SM dijadikan sebagai sumber data utama untuk mengidentifikasi kecenderungan sentimen pengguna terhadap layanan transportasi berbasis kendaraan listrik. Hasil analisis diharapkan mampu memberikan gambaran mengenai pengalaman pengguna sekaligus menunjukkan aspek layanan yang masih memerlukan perhatian.

2.2.2 Aplikasi Green SM dan Google Play Store

Green SM adalah layanan transportasi berbasis kendaraan listrik yang beroperasi di Indonesia sebagai salah satu pilihan transportasi yang lebih ramah lingkungan. Penggunaan

armada kendaraan listrik pada layanan ini diharapkan dapat membantu mengurangi emisi dari sektor transportasi sekaligus mendukung terciptanya sistem transportasi yang lebih berkelanjutan (Regina & Ulmi, 2023).

Green SM merupakan aplikasi yang tersedia pada Google Play Store, yaitu platform distribusi aplikasi berbasis Android yang dikembangkan oleh Google. Melalui platform ini, pengguna dapat mengunduh, menginstal, serta memperbarui aplikasi sesuai dengan kebutuhan mereka. Keberadaan Green SM pada Google Play Store memudahkan pengguna dalam mengakses layanan aplikasi secara praktis dan luas.

Selain berfungsi sebagai sarana distribusi aplikasi, Google Play Store juga menyediakan fitur rating dan review yang memungkinkan pengguna memberikan penilaian serta menyampaikan pengalaman mereka setelah menggunakan aplikasi. Fitur ini memiliki peran penting sebagai media umpan balik langsung dari pengguna terhadap kualitas dan kinerja aplikasi yang digunakan.

Dalam konteks penelitian, ulasan pengguna yang terdapat pada Google Play Store dapat dimanfaatkan sebagai sumber data yang relevan untuk menganalisis persepsi, penilaian, dan pengalaman pengguna terhadap aplikasi Green SM. Data ulasan tersebut mencerminkan berbagai aspek penggunaan aplikasi, seperti kualitas layanan, kemudahan penggunaan, performa sistem, serta pengalaman pengguna dalam memanfaatkan layanan transportasi yang disediakan.

Melalui analisis terhadap ulasan pengguna, peneliti dapat memperoleh gambaran yang lebih sistematis mengenai persepsi pengguna terhadap aplikasi Green SM. Temuan tersebut dapat menjadi dasar untuk mengidentifikasi kelebihan dan kekurangan aplikasi, sekaligus memberikan masukan bagi pengembangan dan peningkatan kualitas layanan di masa mendatang (Dometian & Jatmiko, 2025; Fitriana et al., 2024).

Oleh karena itu, dalam penelitian ini, ulasan pengguna yang diperoleh dari Google Play Store digunakan sebagai sumber data utama karena ditulis secara langsung oleh pengguna aplikasi Green SM. Dengan demikian, data tersebut dinilai mampu merepresentasikan pengalaman nyata pengguna dalam menggunakan layanan yang tersedia pada aplikasi Green SM

2.2.3 Web scraping

Data penelitian dikumpulkan dengan menggunakan teknik web scraping melalui bantuan pustaka *google-play-scraper* pada bahasa pemrograman *Python*. Teknik ini digunakan untuk memperoleh data ulasan pengguna secara otomatis dari halaman

aplikasi Green SM yang tersedia di *Google Play Store*. Melalui proses tersebut, data yang berhasil dikumpulkan mencakup beberapa informasi penting, yaitu nama pengguna, isi ulasan, nilai rating, serta tanggal publikasi ulasan. Data tersebut selanjutnya digunakan sebagai bahan penelitian untuk menganalisis tanggapan dan penilaian pengguna terhadap aplikasi Green SM.

2.2.4 Text mining

Text mining merupakan proses mengekstraksi informasi dari data berbentuk teks sehingga pola, hubungan, atau pengetahuan yang sebelumnya tidak tampak dapat diidentifikasi secara lebih sistematis. Pendekatan ini mengintegrasikan beberapa bidang, seperti temu kembali informasi, *machine learning*, statistika, dan *Natural Language Processing* (NLP), untuk mengolah data tidak terstruktur menjadi informasi yang lebih terorganisasi dan mudah dianalisis. (Mao et al., 2024).

Seiring dengan meningkatnya penggunaan internet, jumlah data teks yang tersedia juga terus bertambah, baik dalam bentuk komentar, ulasan, maupun percakapan pada berbagai platform digital. Kondisi tersebut menjadikan *text mining* sebagai salah satu metode yang relevan dan banyak digunakan untuk mengelola serta menganalisis data teks dalam jumlah besar secara otomatis. (Rodríguez-Ibáñez et al., 2023).

Dalam penelitian ini, *text mining* diterapkan untuk menganalisis ulasan pengguna aplikasi Green SM yang diperoleh dari *Google Play Store*. Proses analisis dilakukan melalui beberapa tahapan, yaitu *preprocessing*, pelabelan sentimen, ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), serta klasifikasi menggunakan algoritma Support Vector Machine (SVM). Melalui tahapan tersebut, data teks yang semula tidak terstruktur dapat diubah menjadi data yang siap diolah dan dianalisis lebih lanjut dalam mendukung tujuan penelitian.

2.2.5 Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang berfokus pada kemampuan komputer untuk memahami, mengolah, dan menginterpretasikan bahasa manusia. Melalui NLP, data berbentuk teks dapat dianalisis sehingga sistem mampu mengenali struktur bahasa, memahami makna, serta mengekstraksi informasi penting yang terkandung di dalamnya. (Jim et al., 2024; Mao et al., 2024).

Penerapan NLP telah berkembang dalam berbagai bidang, seperti *chatbot*, penerjemahan otomatis, mesin pencari, dan analisis sentimen. Dalam konteks analisis

sentimen, NLP berperan penting dalam mengubah data teks yang semula tidak terstruktur menjadi data yang lebih terorganisasi dan siap diproses. Dengan demikian, proses klasifikasi sentimen menggunakan metode *machine learning* dapat dilakukan secara lebih efektif dan menghasilkan tingkat akurasi yang lebih baik. (Jim et al., 2024; Mao et al., 2024).

Pada penelitian ini, penerapan NLP dilakukan melalui serangkaian tahapan *preprocessing* data. Tahapan tersebut meliputi *cleaning*, *case folding*, normalisasi kata, *tokenizing*, *stopword removal*, dan *stemming*. Setiap tahap memiliki fungsi untuk mengurangi *noise* pada data, menyeragamkan bentuk kata, serta meningkatkan kualitas data teks sebelum memasuki proses analisis lebih lanjut. (Nouval et al., 2026; Styawati et al., 2021).

Setelah melalui tahapan *preprocessing*, data teks akan memiliki representasi yang lebih baik sehingga dapat mendukung proses ekstraksi fitur dan klasifikasi secara lebih optimal. Oleh karena itu, penerapan NLP dalam penelitian ini diharapkan mampu meningkatkan kinerja model klasifikasi sentimen serta menghasilkan analisis yang lebih akurat dan relevan.

2.2.6 Analisis Sentimen

Analisis sentimen merupakan salah satu metode yang digunakan untuk mengidentifikasi kecenderungan opini atau penilaian individu berdasarkan data berbentuk teks. Melalui pendekatan ini, komentar atau ulasan dapat diklasifikasikan ke dalam kategori sentimen positif, negatif, atau netral. Dengan demikian, analisis sentimen berperan dalam membantu peneliti memahami persepsi pengguna terhadap suatu produk, layanan, atau topik tertentu secara sistematis. (Du et al., 2024; Mao et al., 2024; Wankhade et al., 2022).

Analisis sentimen merupakan bagian dari penerapan *Natural Language Processing* (NLP), yaitu cabang kecerdasan buatan yang berfokus pada kemampuan komputer untuk memahami, mengolah, dan menganalisis bahasa manusia. Melalui NLP, data teks yang semula tidak terstruktur dapat diproses menjadi bentuk yang lebih terorganisasi, sehingga sistem mampu mengenali pola bahasa, makna, serta kecenderungan opini yang terkandung di dalamnya. Seiring dengan perkembangan teknologi digital, analisis sentimen semakin banyak diterapkan untuk mengolah data teks dalam jumlah besar, seperti komentar dan ulasan pengguna pada berbagai platform digital. Pendekatan ini memberikan efisiensi dalam proses identifikasi opini karena analisis dapat dilakukan secara otomatis tanpa memerlukan pemeriksaan manual terhadap setiap data. Oleh sebab itu,

analisis sentimen menjadi salah satu metode yang relevan dalam mengevaluasi persepsi pengguna secara efektif dan efisien.

Penerapan analisis sentimen pada penelitian sebelumnya telah dilakukan terhadap ulasan pengguna aplikasi PLN Mobile. Data ulasan diperoleh dari pengguna aplikasi, kemudian diolah melalui beberapa tahapan, yaitu *web scraping*, *preprocessing*, pelabelan sentimen menggunakan metode Lexicon VADER, ekstraksi fitur menggunakan TF-IDF, serta klasifikasi dengan algoritma Naïve Bayes. Selanjutnya, performa model dievaluasi menggunakan *confusion matrix* untuk mengukur tingkat akurasi klasifikasi yang dihasilkan. (Asri et al., 2022; Du et al., 2024). Temuan tersebut memperkuat argumentasi mengenai signifikansi analisis sentimen dibandingkan penggunaan rating semata pada Google Play Store. Pada studi terhadap aplikasi PLN Mobile misalnya, teridentifikasi bahwa skor rating yang diberikan pengguna kerap tidak sejalan dengan muatan sentimen pada teks ulasan yang menyertainya, sehingga rating dinilai kurang representatif dalam menggambarkan kualitas layanan aplikasi secara menyeluruh (Asri et al., 2022). Pola serupa turut teramati pada penelitian terhadap ulasan aplikasi transportasi berbasis Google Play Store lainnya, yang menunjukkan bahwa rating hanya mencerminkan penilaian agregat secara numerik, sementara narasi tekstual pada ulasan mampu menguraikan pengalaman pengguna secara lebih komprehensif (Wahyudi et al., 2021). Berlandaskan temuan-temuan tersebut, penelitian ini memposisikan analisis terhadap teks ulasan sebagai pendekatan utama dalam mengevaluasi persepsi pengguna, bukan semata-mata sebagai pelengkap terhadap data rating aplikasi Green SM.

Jim et al. (2024) menjelaskan bahwa analisis sentimen dapat dilakukan menggunakan berbagai pendekatan, seperti *lexicon-based*, *machine learning*, maupun *deep learning*. Masing-masing pendekatan memiliki karakteristik yang berbeda sehingga pemilihannya perlu disesuaikan dengan tujuan penelitian serta karakteristik data yang digunakan (Wankhade et al., 2022).

Pada penelitian ini, analisis sentimen diterapkan pada tingkat dokumen (*document level*), yaitu setiap ulasan diperlakukan sebagai satu dokumen yang mewakili satu opini pengguna. Pendekatan tersebut sesuai untuk memberikan satu label sentimen terhadap keseluruhan isi dokumen (Wankhade et al., 2022). Hasil penelitian tersebut menunjukkan bahwa model yang digunakan mampu mengelompokkan ulasan pengguna ke dalam kategori sentimen positif, negatif, dan netral. Temuan ini menunjukkan bahwa analisis sentimen dapat dimanfaatkan untuk menggambarkan persepsi pengguna terhadap kualitas layanan

aplikasi. Selain itu, metode ini juga berpotensi menjadi alat evaluasi layanan serta mendukung proses pengambilan keputusan berbasis data.

Kategori sentimen yang digunakan pada penelitian ini dijelaskan sebagai berikut.

a. **Sentimen Positif**

Sentimen positif menggambarkan ulasan yang berisi apresiasi atau kepuasan pengguna terhadap layanan Green SM. Ulasan dalam kategori ini umumnya memuat penilaian yang baik mengenai kualitas layanan, kemudahan penggunaan aplikasi, kenyamanan kendaraan, maupun pelayanan yang diberikan oleh pengemudi.

b. **Sentimen Negatif**

Sentimen negatif merupakan ulasan yang mengandung kritik, keluhan, atau ketidakpuasan terhadap aplikasi Green SM. Ulasan pada kategori ini biasanya berkaitan dengan kendala penggunaan aplikasi, proses pemesanan, kualitas layanan, maupun pengalaman pengguna yang tidak sesuai harapan.

c. **Sentimen Netral**

Sentimen netral merupakan ulasan yang tidak menunjukkan kecenderungan positif maupun negatif secara dominan. Ulasan pada kategori ini umumnya berupa penyampaian informasi, pertanyaan, atau komentar yang bersifat deskriptif sehingga tidak mengandung penilaian yang jelas terhadap layanan Green SM.

Penggunaan tiga kategori sentimen tersebut disesuaikan dengan karakteristik data ulasan yang diperoleh dari Google Play Store. Kehadiran kategori netral memungkinkan variasi opini pengguna direpresentasikan secara lebih menyeluruh sehingga hasil analisis sentimen diharapkan dapat menggambarkan kondisi yang lebih mendekati persepsi pengguna sebenarnya.

2.2.7 Text preprocessing

Sebelum dilakukan proses klasifikasi, data teks perlu dipersiapkan agar memiliki kualitas yang sesuai untuk dianalisis. Tahapan tersebut dikenal sebagai *text preprocessing*, yaitu proses membersihkan dan menyeragamkan data teks sehingga lebih mudah diolah pada tahap berikutnya. Jim et al. (2024) menjelaskan bahwa kualitas hasil *preprocessing* berpengaruh terhadap performa model analisis sentimen karena data yang lebih bersih mampu menghasilkan representasi fitur yang lebih baik. Oleh sebab itu, tahapan ini menjadi bagian penting sebelum data diproses menggunakan algoritma *machine learning* (Jim et al., 2024; Septiyanti et al., 2024).

Pentingnya *text preprocessing* juga ditunjukkan oleh penelitian Djunaidi et al. (2024) pada sistem pencarian fungsional obat berbasis TF-IDF dan *Cosine Similarity*. Dalam penelitian tersebut, *preprocessing* dilakukan setelah proses *web scraping* dan sebelum pembobotan TF-IDF maupun perhitungan tingkat kemiripan. Hasil penelitian menunjukkan bahwa data yang telah melalui tahapan *preprocessing* lebih siap digunakan pada proses ekstraksi fitur dan analisis selanjutnya.

Pada penelitian ini, *text preprocessing* dilakukan lewat beberapa tahap seperti berikut.

a) *Cleaning* (Pembersihan Data)

Tahap *cleaning* berfokus buat hapus bagian-bagian yang tidak terlalu penting untuk analisis sentimen. Yang dihapus misalnya URL, mention (@username), hashtag (#), angka, tanda baca, karakter khusus, dan emoji yang tidak mengubah arti ulasan. Dengan noise dikurangi, data jadi lebih bersih dan lebih gampang diproses di tahap selanjutnya.

b) *Case folding*

Case folding dilakukan dengan mengubah semua huruf dalam teks jadi huruf kecil (*lowercase*). Tujuannya supaya tidak ada perbedaan karena huruf besar dan huruf kecil, jadi kata yang maknanya sama tetap dibaca sebagai satu token.

c) *Slangword normalization* (Normalisasi Kata Tidak Baku)

Normalisasi kata dipakai untuk mengubah singkatan, kata tidak baku, dan bahasa gaul jadi bentuk baku sesuai ejaan bahasa Indonesia. Contohnya, kata “gak” diubah jadi “tidak”, lalu “yg” diubah jadi “yang”. Tahap ini penting karena ulasan pengguna biasanya masih banyak memakai bahasa santai.

d) *Stopword removal*

Pada tahap ini, kata-kata yang sering muncul tetapi tidak memberi informasi penting dalam klasifikasi sentimen akan dihapus, misalnya “dan”, “di”, “yang”, dan “ini”. Dengan mengurangi kata-kata seperti itu, jumlah fitur jadi lebih sedikit dan proses komputasi jadi lebih ringan.

e) *Tokenizing*

Tokenizing adalah proses memisahkan kalimat menjadi kumpulan kata atau token. Setelah itu, setiap token dipakai sebagai unit analisis untuk proses ekstraksi fitur dan klasifikasi.

f) *Stemming*

Stemming mengembalikan kata berimbuhan menjadi kata dasar dengan menggunakan algoritma Sastrawi yang dibuat untuk bahasa Indonesia. Contohnya, kata

“menggunakan”, “digunakan”, dan “penggunaan” diubah menjadi “guna”. Penyamaan bentuk kata ini membantu mengurangi variasi kosakata supaya representasi data jadi lebih konsisten.

Melalui rangkaian tahapan tersebut, data ulasan yang semula masih mengandung berbagai variasi penulisan dapat diubah menjadi data yang lebih seragam dan siap digunakan pada proses analisis sentimen.

2.2.8 *Lexicon-based Labeling*

Pendekatan *lexicon-based* merupakan teknik analisis sentimen yang menentukan kategori sentimen berdasarkan kumpulan kata yang telah memiliki nilai sentimen. Setiap kata pada kamus sentimen (*lexicon*) diberikan skor positif maupun negatif sesuai kecenderungan maknanya. Skor sentimen suatu dokumen diperoleh dengan menjumlahkan nilai seluruh kata yang terdapat pada dokumen tersebut (Koto & Rahmaningtyas, 2017).

Untuk teks berbahasa Indonesia telah tersedia beberapa kamus sentimen yang banyak dimanfaatkan dalam penelitian, salah satunya InSet (Indonesia Sentiment Lexicon). Leksikon tersebut berisi kosakata bahasa Indonesia yang telah diberi bobot sentimen sehingga dapat digunakan sebagai acuan dalam proses pelabelan sentimen secara otomatis (Koto & Rahmaningtyas, 2017).

Penelitian ini tidak menggunakan InSet secara langsung, tetapi memanfaatkan kamus sentimen yang disusun khusus sesuai karakteristik ulasan Green SM. Kamus tersebut memuat 70 entri yang terdiri atas 32 kata bersentimen positif dan 38 kata bersentimen negatif. Pendekatan ini dipilih karena berbagai istilah yang sering muncul pada ulasan Green SM belum seluruhnya tercakup dalam kamus sentimen umum. Ulasan dengan skor sentimen nol kemudian diberi label netral dan divalidasi secara manual agar hasil pelabelan sesuai dengan isi ulasan.

2.2.9 TF-IDF (*term Frequency-Inverse Document Frequency*)

TF-IDF (*Term Frequency-Inverse Document Frequency*) merupakan teknik pembobotan kata yang banyak digunakan pada *text mining* maupun *information retrieval*. Metode ini mengukur tingkat kepentingan suatu kata dalam sebuah dokumen dengan mempertimbangkan frekuensi kemunculannya pada dokumen tersebut serta penyebarannya pada seluruh kumpulan dokumen. Menurut Salton dan Buckley (1988), kata yang sering muncul pada satu dokumen tetapi jarang ditemukan pada dokumen lain akan memperoleh bobot yang lebih tinggi sehingga mampu merepresentasikan isi dokumen dengan lebih baik. TF-IDF juga banyak diterapkan sebagai teknik ekstraksi fitur dalam analisis sentimen

karena mampu menghasilkan representasi numerik yang efektif untuk proses klasifikasi (Farhan Wiguna et al., 2023).

a) *Term Frequency (TF)*

Term Frequency (TF) digunakan untuk menghitung seberapa sering suatu kata muncul dalam sebuah dokumen. Semakin tinggi frekuensi kemunculan suatu kata, semakin besar pula nilai TF yang diperoleh. Nilai tersebut menunjukkan bahwa kata tersebut memiliki keterkaitan yang lebih kuat dengan isi dokumen.

$$TF(t, d) = \frac{f(t, d)}{\sum f(t', d)} \quad (2.1)$$

Keterangan:

$TF(t, d)$ menyatakan nilai *Term Frequency* dari kata t pada dokumen d . Nilai $f(t, d)$ menunjukkan jumlah kemunculan kata t , sedangkan penyebutnya merupakan jumlah seluruh kata yang terdapat pada dokumen tersebut.

b) *Inverse Document Frequency (IDF)*

Berbeda dengan TF yang hanya memperhatikan frekuensi dalam satu dokumen, *Inverse Document Frequency (IDF)* digunakan untuk melihat tingkat keunikan suatu kata pada seluruh koleksi dokumen. Kata yang muncul hampir di semua dokumen akan memiliki nilai IDF yang rendah, sedangkan kata yang lebih jarang muncul tetapi mampu membedakan isi dokumen akan memperoleh nilai IDF yang lebih tinggi.

$$IDF(t) = \log \left(\frac{N}{df(t)} \right) \quad (2.2)$$

Keterangan:

N merupakan jumlah keseluruhan dokumen, sedangkan $df(t)$ menunjukkan banyaknya dokumen yang memuat kata t .

c) Nilai Akhir *TF-IDF*

Bobot akhir *TF-IDF* diperoleh dengan mengalikan nilai TF dan *IDF*.

$$TF-IDF(t, d) = TF(t, d) \times IDF(t) \quad (2.3)$$

Nilai *TF-IDF* yang besar menunjukkan bahwa suatu kata sering muncul pada dokumen tertentu, tetapi tidak banyak ditemukan pada dokumen lainnya. Kondisi tersebut menjadikan kata tersebut lebih representatif dalam menggambarkan isi dokumen.

Djunaidi et al., 2024 menerapkan TF-IDF untuk memberikan bobot pada setiap kata berdasarkan frekuensi kemunculan serta distribusinya pada kumpulan dokumen. Hasil pembobotan tersebut kemudian dimanfaatkan dalam perhitungan *Cosine Similarity* untuk mengukur tingkat kemiripan antar dokumen. Pendekatan tersebut memperlihatkan bahwa TF-IDF mampu menghasilkan representasi fitur yang lebih informatif sebelum dilakukan proses analisis lanjutan.

Pada penelitian ini, TF-IDF digunakan untuk mengubah ulasan yang telah melalui tahap *preprocessing* menjadi representasi numerik berdimensi tinggi. Vektor hasil pembobotan tersebut selanjutnya digunakan sebagai masukan bagi algoritma *Support Vector Machine* (SVM) dalam proses klasifikasi sentimen. Penggunaan TF-IDF dipilih karena mampu merepresentasikan karakteristik setiap ulasan secara efektif sehingga mendukung proses klasifikasi menggunakan SVM (Farhan Wiguna et al., 2023).

2.2.10 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma *machine learning* berbasis supervised learning yang banyak digunakan untuk menyelesaikan permasalahan klasifikasi. Algoritma ini memiliki kemampuan yang baik dalam mengolah data berdimensi tinggi, termasuk data teks yang telah direpresentasikan menggunakan TF-IDF. Karakteristik tersebut menjadikan SVM sebagai salah satu algoritma yang banyak diterapkan pada penelitian analisis sentimen karena mampu menghasilkan performa klasifikasi yang baik (Farhan Wiguna et al., 2023; Jim et al., 2024)

a) Cara Kerja SVM dan Konsep *Hyperplane*

Prinsip utama SVM adalah mencari sebuah *hyperplane* yang mampu memisahkan data dari kelas yang berbeda dengan *margin* maksimum. Semakin lebar *margin* yang dihasilkan, semakin baik kemampuan model dalam melakukan generalisasi terhadap data baru. Konsep ini didasarkan pada prinsip *Structural Risk Minimization* (SRM), yaitu pendekatan yang bertujuan memperoleh model dengan kemampuan prediksi yang baik.

$$w \cdot x + b = 0$$

(2.4)

Pada persamaan tersebut, w menyatakan vektor bobot, x merupakan vektor fitur, sedangkan b adalah nilai *bias*. Untuk membentuk *margin*, digunakan dua buah *hyperplane* pembatas yang dirumuskan sebagai berikut.

$$w \cdot x + b = +1 \text{ dan } w \cdot x + b = -1$$

$$(2.5)$$

Lebar *margin* dapat dihitung menggunakan persamaan berikut.

$$\text{margin} = \frac{2}{\|w\|}$$

$$(2.6)$$

Tujuan proses optimasi pada SVM adalah memperoleh *margin* maksimum. Secara matematis, proses tersebut dinyatakan dengan fungsi berikut.

$$\min \frac{1}{2} \|w\|^2$$

$$(2.7)$$

dengan kendala

$$y_i(w \cdot x_i + b) \geq 1$$

$$(2.8)$$

Data yang berada paling dekat dengan *hyperplane* disebut sebagai *support vector*. Titik-titik inilah yang menentukan posisi akhir *hyperplane* sehingga berperan penting dalam proses pembentukan model.

b) *Soft Margin* dan Parameter C

Pada kondisi tertentu, data dari kelas yang berbeda tidak dapat dipisahkan secara sempurna. Untuk mengatasi kondisi tersebut, SVM menerapkan konsep *soft margin* dengan menambahkan variabel *slack* beserta parameter penalti C ke dalam fungsi optimasi.

$$\min \frac{1}{2} \|w\|^2 + C \sum \xi_i$$

$$(2.9)$$

Parameter C digunakan untuk mengendalikan kompromi antara lebar *margin* dan jumlah kesalahan klasifikasi yang masih dapat ditoleransi. Nilai C yang kecil memberikan toleransi lebih besar terhadap kesalahan sehingga menghasilkan *margin* yang lebih lebar. Sebaliknya, nilai C yang besar mendorong model mengurangi kesalahan klasifikasi semaksimal mungkin, meskipun dapat meningkatkan risiko *overfitting*.

c) Fungsi Kernel pada SVM

SVM menggunakan fungsi *kernel* untuk mengukur tingkat kemiripan antar data sekaligus menentukan apakah data perlu dipetakan ke ruang berdimensi lebih tinggi agar dapat dipisahkan secara optimal. Penelitian ini menguji dua jenis *kernel*, yaitu *kernel* linear dan *kernel Radial Basis Function* (RBF).

$$K(x_i, x_j) = x_i \cdot x_j$$

(2.10)

Kernel linear banyak digunakan pada data teks berdimensi tinggi seperti representasi TF-IDF karena data tersebut umumnya sudah cukup mudah dipisahkan secara linear pada ruang fitur berdimensi tinggi, serta memiliki kebutuhan komputasi yang relatif rendah (Rabbani et al., 2023).

$$K(x, y) = \exp(-\gamma \|x - y\|^2)$$

(2.11)

Selain *kernel* linear, penelitian ini juga menguji *kernel Radial Basis Function* (RBF), yang menggunakan parameter γ (gamma) untuk mengontrol lebar fungsi Gaussian sehingga mampu menangkap hubungan non-linear antar fitur yang tidak selalu dapat dikenali oleh *kernel* linear (Fremmuzar & Baita, 2023; Rabbani et al., 2023). Karena pemilihan jenis *kernel* berpengaruh signifikan terhadap performa SVM (Fremmuzar & Baita, 2023), penelitian ini tidak menetapkan salah satu *kernel* secara langsung, melainkan menentukan *kernel* yang paling sesuai secara empiris melalui proses *Grid search Cross Validation*.

d) SVM untuk Klasifikasi Multi-Kelas

Pada dasarnya, SVM dirancang untuk membedakan dua kelas (*binary classification*). Namun, klasifikasi sentimen umumnya melibatkan lebih dari dua kategori, misalnya positif, negatif, dan netral. Untuk menangani kondisi tersebut, SVM dapat dikembangkan menggunakan pendekatan *One-vs-Rest* (OvR) maupun *One-vs-One* (OvO). Penelitian ini menerapkan pendekatan OvO, sesuai strategi bawaan kelas SVC pada *scikit-learn*, yaitu dibangun satu *classifier* biner untuk setiap pasangan kelas sehingga terbentuk beberapa model klasifikasi yang bekerja secara terpisah.

e) Keunggulan dan Keterbatasan SVM

SVM memiliki beberapa kelebihan dalam analisis sentimen berbasis teks, di antaranya:

- 1) Mampu menangani data berdimensi tinggi dengan baik.

- 2) Memiliki kemampuan generalisasi yang baik karena prinsip *margin maksimum*.
- 3) Relatif tahan terhadap masalah *curse of dimensionality*.
- 4) Memberikan performa yang kompetitif pada berbagai penelitian analisis sentimen berbahasa Indonesia, termasuk penelitian yang memanfaatkan pembobotan TF-IDF sebagai representasi fitur (Farhan Wiguna et al., 2023).

Di samping kelebihan, SVM juga memiliki beberapa keterbatasan.

1. Proses pelatihan memerlukan waktu yang lebih lama ketika diterapkan pada *dataset* berukuran besar.
2. Kinerja model sangat dipengaruhi oleh pemilihan *kernel* dan parameter yang tepat, sehingga memerlukan proses *tuning* yang teliti.

Secara keseluruhan, SVM dipilih pada penelitian ini karena mampu menangani karakteristik data teks hasil representasi TF-IDF dengan baik serta memiliki kemampuan klasifikasi yang sesuai untuk analisis sentimen. Pemilihan ini didukung oleh bukti empiris dari penelitian komparatif terhadap ulasan aplikasi berbasis Google Play Store, yang menunjukkan SVM secara konsisten mengungguli Naïve Bayes dan Logistic Regression pada dua dataset independen, dengan akurasi masing-masing 93% berbanding 92% pada kedua algoritma pembandingan (Dometian & Jatmiko, 2025). Keunggulan SVM pada studi tersebut tidak hanya tercermin dari akurasi, tetapi juga dari distribusi recall yang lebih seimbang antar kelas sentimen, berbeda dengan Naïve Bayes yang cenderung bias terhadap kelas mayoritas. Selain itu, pemilihan jenis kernel yang tepat turut terbukti berpengaruh signifikan terhadap performa SVM dalam menangani data teks (Rabbani et al., 2023), sehingga penelitian ini menerapkan optimasi kernel melalui Grid search CV untuk memaksimalkan potensi algoritma ini pada data ulasan Green SM.

2.2.11 Evaluasi Model Klasifikasi

Evaluasi model dilakukan untuk mengetahui kemampuan model klasifikasi dalam menghasilkan prediksi yang sesuai dengan label sebenarnya. Pengujian tidak hanya dilakukan terhadap data pelatihan, tetapi juga menggunakan data yang belum pernah dipelajari model sehingga kemampuan generalisasinya dapat dinilai. Pada penelitian ini, evaluasi dilakukan menggunakan *confusion matrix* sebagai dasar perhitungan berbagai metrik performa. (Rustam et al., 2021) menjelaskan bahwa penggunaan beberapa metrik evaluasi memberikan penilaian yang lebih menyeluruh terhadap kinerja model dibandingkan hanya menggunakan satu metrik.

Penerapan evaluasi menggunakan *confusion matrix* juga ditunjukkan oleh Hendra Jatnika et al., 2025 pada penelitian klasifikasi sentimen ulasan layanan publik menggunakan algoritma Random Forest. Berdasarkan hasil pengujian, model yang dikembangkan memperoleh tingkat akurasi sebesar 88,97%. Temuan tersebut menunjukkan bahwa evaluasi melalui *confusion matrix* mampu memberikan gambaran yang jelas mengenai kemampuan model dalam mengklasifikasikan sentimen sehingga hasilnya dapat dimanfaatkan sebagai bahan evaluasi terhadap kualitas layanan berdasarkan ulasan pengguna.

a) *Confusion matrix*

Confusion matrix merupakan matriks yang digunakan untuk membandingkan hasil prediksi model dengan label sebenarnya. Karena penelitian ini menggunakan tiga kategori sentimen, yaitu positif, negatif, dan netral, maka bentuk *confusion matrix* yang digunakan berukuran 3×3 .

Komponen utama pada *confusion matrix* terdiri atas:

1. **True Positive (TP)**, yaitu data yang berhasil diprediksi sesuai dengan kelas sebenarnya.
2. **True Negative (TN)**, yaitu data yang berhasil dikenali sebagai bukan anggota kelas yang sedang dievaluasi.
3. **False Positive (FP)**, yaitu data yang sebenarnya berasal dari kelas lain, tetapi diprediksi sebagai kelas yang sedang diamati.
4. **False Negative (FN)**, yaitu data yang seharusnya termasuk ke dalam suatu kelas, tetapi diprediksi sebagai kelas yang berbeda.

Tabel 2. 2 Ilustrasi *Confusion matrix* untuk Klasifikasi Tiga Kelas

Aktual \ Prediksi	Positif	Negatif	Netral
Positif	TP (Positif)	FN	FN
Negatif	FP	TP (Negatif)	FN
Netral	FP	FP	TP (Netral)

b) *Accuracy* (Akurasi)

Accuracy menunjukkan persentase prediksi yang benar dibandingkan dengan keseluruhan data yang diuji. Nilai ini memberikan gambaran umum mengenai performa model, meskipun kurang representatif apabila distribusi data antar kelas tidak seimbang.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

(2.12)

c) *Precision*

Precision digunakan untuk mengukur tingkat ketepatan model ketika memberikan prediksi pada suatu kelas tertentu. Nilai *precision* menjadi penting apabila kesalahan berupa *false positive* perlu ditekan seminimal mungkin.

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

(2.13)

d) *Recall*

Recall menggambarkan kemampuan model dalam menemukan seluruh data yang benar-benar termasuk ke dalam suatu kelas. Nilai ini menjadi perhatian ketika kesalahan berupa *false negative* perlu dikurangi.

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

(2.14)

e) *F1-Score*

F1-score merupakan rata-rata harmonik antara nilai *precision* dan *recall*. Metrik ini digunakan untuk memberikan penilaian yang lebih seimbang terhadap performa model, terutama ketika distribusi data pada setiap kelas tidak sama.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

(2.15)

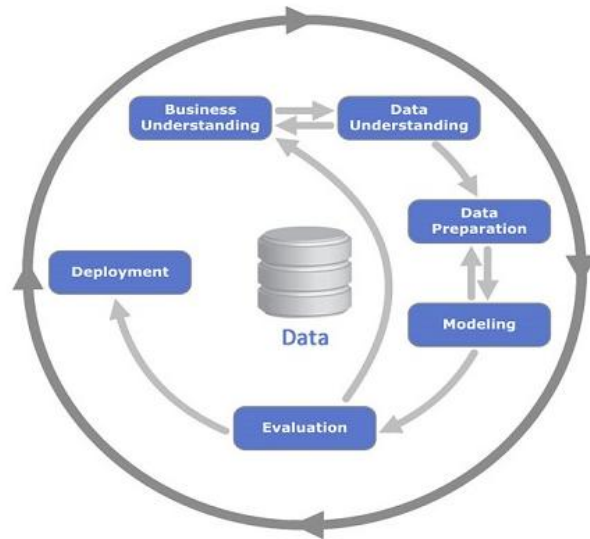
(Wulandari et al., 2025) menggunakan *confusion matrix* untuk membandingkan performa algoritma Naïve Bayes dan Logistic Regression pada tugas klasifikasi sentimen. Evaluasi dilakukan melalui perhitungan *accuracy*, *precision*, *recall*, dan *F1-score* sehingga kelebihan maupun kelemahan masing-masing algoritma dapat diamati secara lebih menyeluruh. Pendekatan tersebut menunjukkan bahwa penggunaan beberapa metrik

evaluasi memberikan gambaran yang lebih lengkap dibandingkan hanya mengandalkan nilai akurasi.

Pada penelitian ini, proses evaluasi dilakukan pada skenario klasifikasi multikelas. Oleh karena itu, nilai *precision*, *recall*, dan *F1-score* dihitung untuk setiap kelas terlebih dahulu, kemudian diringkas menggunakan metode *weighted average*. Pendekatan tersebut mempertimbangkan jumlah data pada masing-masing kelas sehingga hasil evaluasi lebih mencerminkan kondisi sebenarnya, terutama ketika distribusi data tidak seimbang.

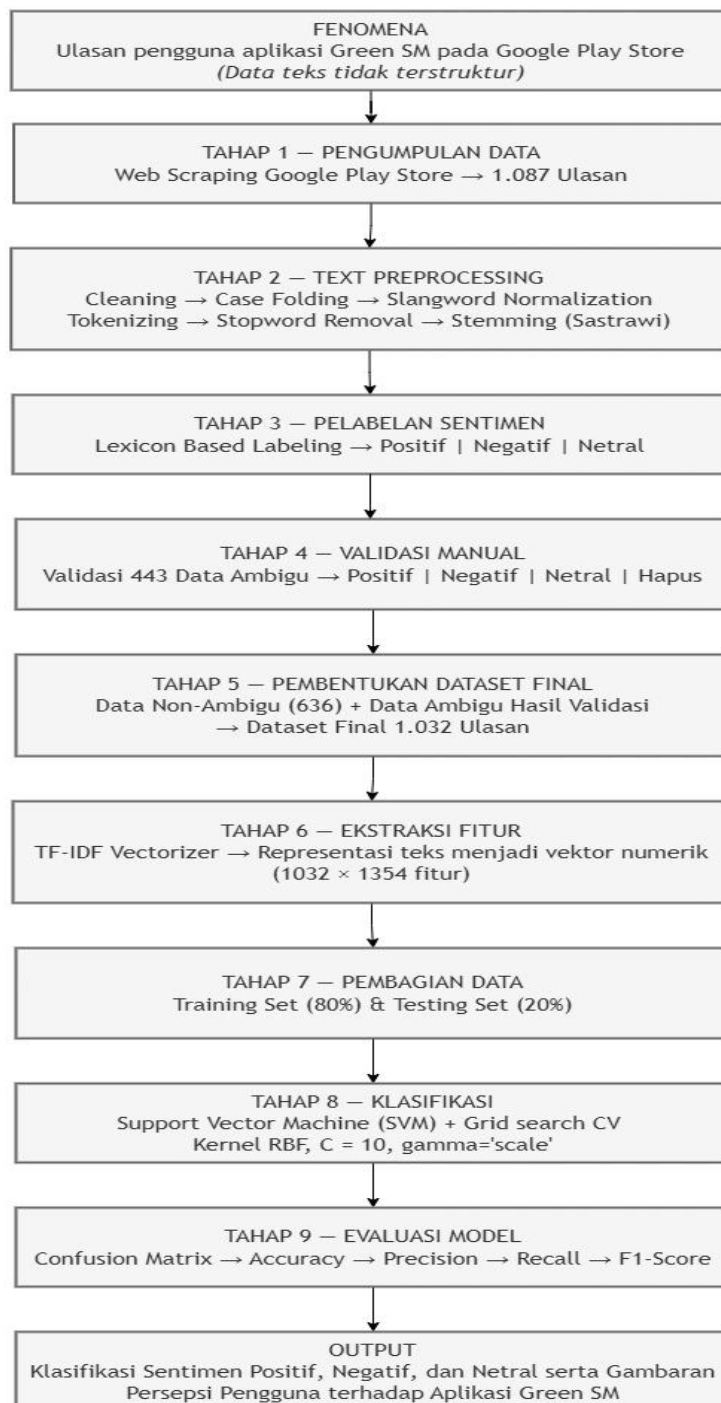
2.2.12 CRISP-DM

Cross-Industry Standard Process for Data Mining (CRISP-DM) merupakan metodologi yang banyak digunakan dalam penelitian data mining karena menyediakan tahapan kerja yang sistematis, terstruktur, dan bersifat iteratif dalam pengembangan model. CRISP-DM terdiri atas enam tahapan, yaitu *Business Understanding*, yaitu memahami permasalahan, tujuan, dan ruang lingkup penelitian; *Data Understanding*, yaitu mengumpulkan serta memahami karakteristik data yang akan dianalisis; *Data Preparation*, yaitu mempersiapkan data melalui proses pembersihan (*cleaning*), transformasi, dan *preprocessing* agar siap digunakan; *Modeling*, yaitu membangun model menggunakan algoritma *machine learning* serta melakukan optimasi parameter; *Evaluation*, yaitu mengevaluasi kinerja model berdasarkan metrik yang digunakan untuk memastikan model telah memenuhi tujuan penelitian; dan *Deployment*, yaitu menyajikan hasil analisis atau model dalam bentuk laporan maupun rekomendasi yang dapat dimanfaatkan oleh pengguna. Karena bersifat iteratif, setiap tahapan dapat diulang apabila ditemukan permasalahan pada tahap berikutnya. Metodologi ini telah banyak diterapkan pada penelitian klasifikasi berbasis *machine learning* (Afandi et al., 2024; Wulandari et al., 2025), sehingga pada penelitian ini CRISP-DM digunakan sebagai pedoman dalam seluruh tahapan analisis sentimen ulasan pengguna aplikasi Green SM.



Gambar 2. 1 *CRISP-DM*
(Sumber: Siswipraptini et al., 2023)

2.3 Kerangka Pemikiran



Gambar 2. 2 Kerangka Pemikiran

Kerangka pemikiran penelitian ini menggambarkan tahapan penelitian yang dilakukan mulai dari proses pengumpulan data hingga menghasilkan model klasifikasi sentimen ulasan pengguna aplikasi Green SM pada Google Play Store. Penyusunan kerangka tersebut didasarkan pada permasalahan berupa banyaknya ulasan pengguna yang tersedia di Google Play Store, tetapi masih berbentuk data teks tidak terstruktur

sehingga sulit dianalisis secara manual, terutama dalam jumlah yang besar.

Penelitian diawali dengan pengumpulan data menggunakan teknik *web scraping* pada halaman aplikasi Green SM di Google Play Store. Melalui proses tersebut diperoleh sebanyak 1.087 ulasan yang digunakan sebagai data awal penelitian. Data yang diperoleh masih berupa teks mentah sehingga perlu melalui beberapa tahapan pengolahan sebelum digunakan pada proses klasifikasi sentimen.

Tahap berikutnya adalah *text preprocessing* untuk meningkatkan kualitas data sebelum dilakukan analisis. Proses ini diawali dengan *cleaning* untuk menghapus karakter, simbol, tanda baca, dan elemen lain yang tidak memberikan informasi penting. Selanjutnya dilakukan *case folding* dengan mengubah seluruh huruf menjadi huruf kecil agar penulisan kata lebih seragam. Setelah itu diterapkan *slangword normalization* untuk mengubah kata tidak baku menjadi bentuk baku. Teks kemudian dipisahkan menjadi token melalui proses *tokenizing*, dilanjutkan dengan *stopword removal* untuk menghilangkan kata-kata yang memiliki kontribusi rendah terhadap proses klasifikasi. Tahap terakhir adalah *stemming* menggunakan algoritma Sastrawi sehingga setiap kata dikembalikan ke bentuk dasarnya. Rangkaian proses tersebut menghasilkan data yang lebih bersih dan konsisten sehingga siap digunakan pada tahapan berikutnya.

Data hasil *preprocessing* kemudian memasuki tahap pelabelan sentimen menggunakan pendekatan *lexicon-based labeling*. Setiap ulasan diberi label berdasarkan skor sentimen yang diperoleh dari kamus sentimen yang digunakan dalam penelitian. Proses pelabelan otomatis menghasilkan tiga kategori sentimen, yaitu positif, negatif, dan netral, dengan komposisi 408 ulasan positif, 228 ulasan negatif, dan 443 ulasan netral.

Banyaknya ulasan yang masih memperoleh label netral menyebabkan perlunya dilakukan validasi secara manual. Sebanyak 443 ulasan tersebut ditinjau kembali untuk menentukan apakah termasuk kategori positif, negatif, atau dihapus apabila tidak memenuhi kriteria penelitian. Tahap ini dilakukan untuk meningkatkan kualitas *dataset* sehingga label sentimen yang digunakan pada proses klasifikasi lebih sesuai dengan isi ulasan.

Setelah proses validasi selesai, data non-netral sebanyak 636 ulasan digabungkan dengan hasil validasi manual yang dinyatakan layak digunakan. Selanjutnya dilakukan pemeriksaan kembali untuk menghapus data yang duplikat sehingga diperoleh 1.032 ulasan sebagai *dataset* akhir. *Dataset* tersebut kemudian digunakan pada tahap pembentukan model klasifikasi.

Tahap berikutnya adalah ekstraksi fitur menggunakan metode *Term Frequency–*

Inverse Document Frequency (TF-IDF). Melalui proses ini, data teks diubah menjadi representasi numerik dalam bentuk vektor sehingga dapat diproses oleh algoritma *machine learning*. Bobot setiap kata dihitung berdasarkan frekuensi kemunculannya pada dokumen dan penyebarannya pada seluruh kumpulan dokumen. Hasil pembobotan menghasilkan matriks TF-IDF berukuran 1.032×1.354 yang digunakan sebagai masukan bagi model klasifikasi.

Dataset selanjutnya dibagi menjadi data latih (*training set*) dan data uji (*testing set*) dengan perbandingan 80% : 20%. Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji dimanfaatkan untuk mengukur kemampuan model dalam mengklasifikasikan data yang belum pernah digunakan selama proses pelatihan.

Proses klasifikasi dilakukan menggunakan algoritma *Support Vector Machine* (SVM), yaitu *Support Vector Classification* (SVC) dengan *kernel* RBF. Parameter terbaik diperoleh melalui proses optimasi *Grid search Cross Validation*, yaitu $C = 10$ dan $\gamma = \text{scale}$. Konfigurasi tersebut dipilih karena mampu memberikan performa yang baik pada data teks berdimensi tinggi dan banyak digunakan dalam penelitian analisis sentimen.

Tahap terakhir adalah mengevaluasi model yang telah dibangun menggunakan *confusion matrix*. Evaluasi dilakukan melalui perhitungan nilai *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui kemampuan model dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Green SM. Hasil penelitian berupa model klasifikasi sentimen beserta informasi mengenai kecenderungan sentimen pengguna terhadap aplikasi Green SM berdasarkan ulasan yang tersedia pada Google Play Store.