

Analisis Sentimen Ulasan Pengguna Aplikasi Fintech menggunakan *Framework CRISP-DM* dalam Penentuan Prioritas Pengembangan Produk

Sentiment Analysis of Fintech Application User Reviews Using the CRISP-DM Framework for Product Development Prioritization

¹Muhammad Rizky Amalsyah, ²Dedy Kurniawan*, ³Ahmad Rifai, ⁴Purwita Sari

^{1,2,3,4}Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya
^{1,2,3,4}Jl. Masjid Al Gazali, Bukit Lama, Kec. Ilir Bar. I, Kota Palembang, Sumatera Selatan 30128, Indonesia

*e-mail: dedykurniawan@unsri.ac.id

(received: 5 February 2025, revised: 6 February 2025, accepted: 12 February 2025)

Abstrak

Pertumbuhan pesat aplikasi fintech telah meningkatkan kebutuhan akan analisis sentimen guna memahami persepsi pengguna terhadap produk yang ditawarkan. Penelitian ini berfokus pada analisis sentimen terhadap ulasan pengguna aplikasi Flip pada Google Play Store dengan menerapkan algoritma *Support Vector Machine* (SVM) dalam proses analisis berbasis kerangka kerja CRISP-DM. Proses analisis melibatkan *preprocessing* teks, pemberian label sentimen menggunakan model *pretrained* BERT, serta klasifikasi menggunakan SVM dengan ekstraksi fitur TF-IDF. Hasil penelitian menunjukkan bahwa mayoritas pengguna memberikan sentimen positif (56,9%), terutama terkait efisiensi biaya, kemudahan transaksi, dan kecepatan produk. Namun, terdapat pula sentimen negatif (43,1%), yang mencakup masalah biaya tambahan, keterlambatan transaksi, serta kendala teknis dalam penggunaan aplikasi. Analisis pemodelan topik dengan metode *Latent Dirichlet Allocation* (LDA) mengidentifikasi topik utama yang menjadi keunggulan serta tantangan produk Flip. Kesimpulan dari penelitian ini menegaskan bahwa Flip memiliki potensi besar dalam memenuhi kebutuhan pengguna, namun perlu dilakukan peningkatan pada aspek produk, transparansi biaya, serta optimalisasi performa aplikasi. Hasil penelitian ini diharapkan dapat menjadi landasan strategis bagi pengembang aplikasi fintech dalam meningkatkan kualitas produk berbasis data, guna meningkatkan kepuasan dan loyalitas pengguna.

Kata kunci: analisis sentimen, fintech, *support vector machine*, CRISP-DM, pemodelan topik

Abstract

The rapid growth of fintech applications has increased the need for sentiment analysis to understand user perceptions of the offered products. This study focuses on sentiment analysis of user reviews for the Flip application on Google Play Store by applying the Support Vector Machine (SVM) algorithm within the CRISP-DM framework. The analysis process involves text preprocessing, sentiment labeling using a pretrained BERT model, and classification using SVM with TF-IDF feature extraction. The results indicate that the majority of users express positive sentiment (56.9%), primarily regarding cost efficiency, transaction ease, and product speed. However, negative sentiment (43.1%) is also present, mainly concerning additional fees, transaction delays, and technical issues in app usage. A topic modeling analysis using the Latent Dirichlet Allocation (LDA) method identifies key topics that highlight both Flip's strengths and challenges. The findings suggest that while Flip holds significant potential in meeting user needs, improvements are needed in product aspects, cost transparency, and app performance optimization. This study is expected to serve as a strategic foundation for fintech app developers to enhance data-driven product quality, ultimately increasing user satisfaction and loyalty.

Keywords: sentiment analysis, fintech, support vector machine, CRISP-DM, topic modeling

1 Pendahuluan

Perkembangan pesat teknologi informasi dan komunikasi telah mempercepat digitalisasi di berbagai sektor, termasuk industri keuangan. Salah satu kemajuan teknologi yang muncul yaitu aplikasi keuangan digital yang menawarkan kemudahan dan efisiensi dalam melakukan transaksi [1]. Aplikasi fintech, seperti Flip, Dana, OVO, dan lainnya, telah menjadi solusi populer untuk kebutuhan transaksi keuangan sehari-hari. Aplikasi-aplikasi ini menyediakan berbagai layanan, seperti transfer antar bank tanpa biaya administrasi, pembayaran tagihan, hingga pengelolaan keuangan pribadi [2]. Seiring dengan meningkatnya jumlah pengguna smartphone, penggunaan aplikasi fintech juga terus mengalami pertumbuhan yang signifikan.

Meningkatnya jumlah pengguna aplikasi fintech menyebabkan meningkatnya tuntutan terhadap kualitas produk yang diberikan [3]. Pengguna memiliki kesempatan untuk memberikan ulasan dan penilaian terhadap aplikasi yang mereka gunakan melalui platform seperti Google Play Store. Ulasan dan rating tersebut berperan sebagai indikator utama bagi pengembang dalam memahami pengalaman pengguna serta mengoptimalkan kualitas produk mereka. Beberapa aplikasi fintech telah diunduh oleh jutaan pengguna dan mendapatkan ribuan ulasan. Misalnya, Flip memiliki lebih dari 10 juta unduhan dengan rating rata-rata 4.8 dan 651 ribu ulasan pengguna [4]. Data ini menunjukkan bahwa ulasan pengguna merupakan sumber informasi yang berharga dalam mengevaluasi kepuasan dan persepsi pengguna terhadap produk fintech.

Namun, jumlah ulasan yang sangat besar menjadi tantangan dalam hal pengelolaan dan analisis persepsi pengguna secara manual. Menganalisis ribuan ulasan secara manual tidak hanya membutuhkan waktu yang lama, tetapi juga berpotensi menimbulkan kesalahan dalam pemahaman [5]. Dengan demikian, dibutuhkan pendekatan atau metode yang lebih efisien dan akurat untuk menganalisis ulasan pengguna. Di antara metode yang dapat diterapkan adalah analisis sentimen melalui penerapan teknik *text mining*. Analisis sentimen memungkinkan pengelompokan ulasan berdasarkan sentimen positif atau negatif, yang membantu dalam memahami persepsi pengguna secara lebih sistematis dan objektif [6].

Dalam studi lain, teknik analisis statistik, contohnya SPSS (*Statistical Package for the Social Sciences*), dimanfaatkan untuk menilai validitas dan kelayakan suatu produk. Sebagai contoh, penelitian tentang Pengembangan Aplikasi Evaluasi Pembelajaran Wizer.Me pada Mata Pelajaran IPS di MTs Darussalam Aryojeding memanfaatkan metode statistik untuk menguji keandalan media pembelajaran yang telah dikembangkan. [7]. Tidak seperti penelitian tersebut, Penelitian ini mengadopsi metode CRISP-DM (*Cross-Industry Standard Process for Data Mining*) sebagai acuan *framework* dalam mengelola seluruh tahapan analisis data, dimulai dari tahap pemahaman terhadap bisnis hingga tahap evaluasi hasil yang diperoleh [8]. Selain itu, algoritma *Support Vector Machine* (SVM) akan dimanfaatkan dalam mengklasifikasi sentimen dari ulasan pengguna. SVM dipilih karena keunggulannya dalam klasifikasi teks, termasuk analisis sentimen, serta kemampuannya dalam memberikan hasil yang akurat. [9]. Dengan pendekatan ini, penelitian ini tidak hanya berfokus pada validasi dan evaluasi produk tetapi juga pada pemahaman yang lebih mendalam mengenai opini pengguna aplikasi fintech. Hasil yang diperoleh diharapkan dapat membantu pemangku kepentingan dalam pengambilan keputusan berbasis data serta memberikan wawasan strategis untuk peningkatan kualitas produk aplikasi fintech secara keseluruhan.

Penelitian ini bertujuan untuk mengkaji penerapan kerangka kerja CRISP-DM serta pemanfaatan model *machine learning* dalam menganalisis sentimen pengguna terhadap aplikasi fintech di Google Play Store guna menentukan prioritas pengembangan produk. Selain itu, studi ini juga menekankan pada identifikasi dan analisis sentimen pengguna terhadap aplikasi fintech untuk memahami persepsi mereka serta mengevaluasi potensi dan manfaat adopsi metode CRISP-DM. Penelitian ini diharapkan dapat memberikan pemahaman yang bermanfaat bagi pengembang aplikasi fintech dalam mengoptimalkan kualitas produk serta menjadi referensi dalam penerapan metode analisis sentimen berbasis *machine learning*.

2 Tinjauan Literatur

Analisis sentimen dalam ulasan pengguna aplikasi fintech menjadi fokus utama dalam berbagai penelitian yang bertujuan untuk memahami pengalaman pengguna dan meningkatkan kualitas produk

aplikasi. Beberapa penelitian terdahulu telah mengimplementasikan berbagai metode *machine learning* untuk mengklasifikasikan sentimen pengguna terhadap aplikasi dan fenomena tertentu.

Penelitian yang dilakukan oleh Mardiana meneliti sentimen ulasan pengguna terhadap aplikasi Alfagift dengan menerapkan algoritma *Support Vector Machine* (SVM). Hasilnya menunjukkan bahwa sebagian besar ulasan pengguna bersifat netral, yang mengindikasikan adanya ketidakjelasan dalam cara kerja aplikasi akibat kurang optimalnya pengembangan sistem dan sumber daya. Namun, penelitian ini lebih berfokus pada identifikasi kategori sentimen tanpa mengeksplorasi bagaimana hasil analisis tersebut dapat digunakan untuk mendukung pengambilan keputusan pengembangan produk lebih lanjut [10].

Sementara itu, pada penelitian Fudholi dkk, mengaplikasikan metode *Naïve Bayes* dalam analisis sentimen perilaku penggemar *band* Coldplay di Twitter. Melalui nilai akurasi model sebesar 80,25%, penelitian ini mampu mengidentifikasi perilaku positif dan negatif dalam interaksi digital penggemar. Meskipun demikian, penelitian ini menghadapi tantangan dalam ketidakseimbangan data yang menyebabkan bias pada prediksi sentimen negatif. Dari hasil ini, tampak bahwa penggunaan algoritma yang lebih kompleks atau strategi pemrosesan data tambahan dapat meningkatkan akurasi model dalam analisis sentimen [11].

Pada studi lain, Saputra dkk, menganalisis sentimen pengguna terhadap aplikasi Tokocrypto menggunakan metode algoritma *Naïve Bayes*. Penelitian ini menggunakan data dari Google Play Store dan menemukan bahwa mayoritas ulasan bersifat negatif, mengindikasikan adanya ketidakpuasan pengguna terhadap produk aplikasi. Akurasi model mencapai 74,22%, yang mengindikasikan bahwa algoritma *Naïve Bayes* bekerja secara efektif dalam menganalisis dan mengkategorikan sentimen pengguna. Namun, penelitian ini tidak membahas bagaimana hasil analisis tersebut dapat digunakan oleh pengembang aplikasi untuk menetapkan prioritas dalam pengembangan fitur baru [12].

Lebih lanjut, pada penelitian Purnama dan Wahyudi meneliti analisis sentimen masyarakat terhadap pembatalan Indonesia sebagai tuan rumah Piala Dunia U-20 dengan menggunakan metode algoritma *K-Nearest Neighbor* (KNN). Dengan tingkat akurasi 99,83%, penelitian ini mengungkapkan bahwa algoritma KNN memiliki kinerja optimal dalam mengklasifikasikan sentimen. Akan tetapi, penelitian ini lebih berfokus pada klasifikasi sentimen tanpa membahas secara mendalam bagaimana opini publik yang teridentifikasi dapat digunakan dalam pengambilan kebijakan strategis [13].

Terakhir, penelitian oleh Ridwansyah menerapkan metode *Naïve Bayes Classifier* dengan *K-Fold Cross Validation* untuk menganalisis sentimen masyarakat dunia terhadap Kota Medan di Twitter. Dengan hasil yang menunjukkan bahwa mayoritas sentimen bersifat positif (80,2%), penelitian ini menyoroti potensi *text mining* dalam memahami opini publik. Namun, studi ini lebih menekankan pada aspek akurasi model tanpa membahas implikasi hasilnya terhadap pengambilan keputusan atau strategi pengembangan kota berbasis data [14].

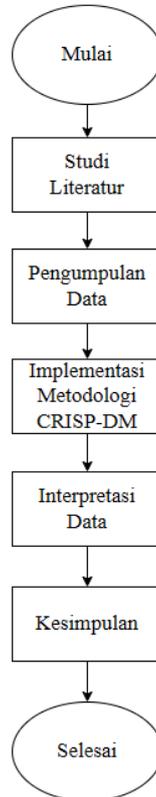
Dari berbagai penelitian terdahulu, dapat disimpulkan bahwa sebagian besar penelitian analisis sentimen hanya berfokus pada klasifikasi sentimen dan evaluasi model *machine learning* yang digunakan, tanpa mengaitkannya secara langsung dengan strategi pengembangan produk atau layanan berdasarkan temuan tersebut. Selain itu, pendekatan *framework* yang sistematis seperti CRISP-DM (*Cross-Industry Standard Process for Data Mining*) masih jarang digunakan untuk penelitian analisis sentimen aplikasi fintech. Penelitian ini berusaha mengisi celah tersebut dengan menerapkan metode CRISP-DM dalam analisis sentimen ulasan pengguna aplikasi fintech, menggunakan algoritma SVM untuk klasifikasi sentimen. Dengan pendekatan ini, tidak terbatas pada klasifikasi sentimen pengguna, melainkan juga mengeksplorasi bagaimana hasil analisis dapat digunakan untuk menetapkan prioritas pengembangan produk, sehingga pengembang aplikasi dapat memperoleh wawasan yang berguna untuk meningkatkan kualitas produk dan pengalaman pengguna secara keseluruhan.

3 Metode Penelitian

Metode penelitian ini menjelaskan tentang metode yang digunakan dalam penelitian, mencakup tahapan yang dilakukan untuk mencapai tujuan serta memastikan keakuratan dan relevansi dalam analisis sentimen pengguna aplikasi Fintech.

3.1 Tahapan Penelitian

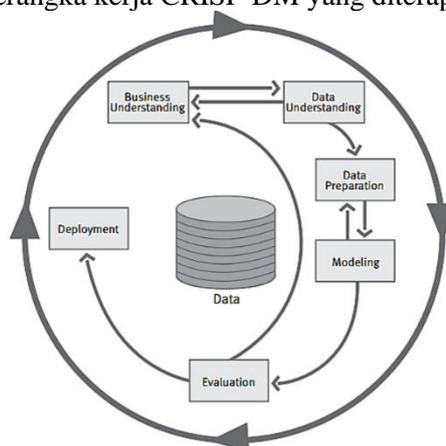
Tahapan penelitian merupakan rangkaian proses yang diimplementasikan secara sistematis, terstruktur, dan terorganisir guna mencapai tujuan yang telah ditetapkan [15]. Bagian ini menguraikan proses tahapan yang akan dijalankan oleh penulis untuk menyelesaikan penelitian ini. Adapun diagram alir metodologi yang digunakan dalam penelitian ini adalah sebagaimana Gambar 1 berikut, yang dimulai dari studi literatur hingga kesimpulan.



Gambar 1. Tahapan penelitian

3.2 Pendekatan Metode

Penelitian ini menerapkan *Cross-Industry Standard Process for Data Mining* (CRISP-DM) sebagai acuan kerangka kerja utama. Metode ini mencakup enam tahapan sistematis dalam menganalisis sentimen pengguna terhadap aplikasi Flip berdasarkan ulasan di Google Play Store, dengan menggunakan algoritma *Support Vector Machine* (SVM). Gambar 2 di bawah ini menunjukkan tahapan dalam kerangka kerja CRISP-DM yang diterapkan pada penelitian ini.



Gambar 2. Kerangka kerja CRISP-DM [16]

Business Understanding (Pemahaman Bisnis)

Tahapan ini bertujuan untuk mengidentifikasi permasalahan dan menentukan tujuan penelitian, yaitu menganalisis sentimen pengguna aplikasi Flip untuk memahami persepsi mereka terhadap

<http://sistemasi.ftik.unisi.ac.id>

produk aplikasi. Hasil analisis diharapkan dapat memberikan wawasan kepada pengembang dalam meningkatkan kualitas aplikasi.

Data Understanding (Pemahaman Data)

Penelitian ini memanfaatkan data berupa ulasan pengguna aplikasi Flip yang diperoleh secara *real-time* dari Google Play Store melalui teknik *web scraping*. Data tersebut meliputi teks ulasan, rating, dan tanggal publikasi, dengan batas maksimal 2.000 ulasan. Pada tahap ini juga dilakukan eksplorasi awal terhadap pola sentimen dan karakteristik data.

Data Preparation (Persiapan Data)

Sebelum dianalisis, data yang diperoleh harus melalui tahapan *preprocessing*. Tahapan *preprocessing* meliputi **pembersihan data**, yang mencakup penghapusan karakter yang tidak relevan (emoji, tanda baca, angka), serta mengonversi teks menjadi huruf kecil. Selanjutnya dilakukan **normalisasi teks** seperti menghapus *stopwords*, memperbaiki kata tidak baku menggunakan kamus slang. Setelah itu, proses **penyaringan data** yang dilakukan untuk mengeliminasi ulasan kosong atau tidak relevan. Lalu, **Labeling sentiment** dilakukan menggunakan *pretrained* model BERT untuk memberi kategori sentimen (positif atau negatif).

Modeling (Pemodelan)

Pada langkah pemodelan, model Support Vector Machine (SVM) dilatih menggunakan data yang telah diberi label guna memprediksi sentimen dalam ulasan pengguna. Dilakukan parameter tuning guna mengoptimalkan performa model, serta evaluasi awal terhadap data pelatihan sebelum diuji pada dataset yang baru.

Model Evaluation (Evaluasi Model)

Berdasarkan penelitian [17], evaluasi model dilakukan untuk mengukur kinerja model dalam mengklasifikasikan sentimen. Metode evaluasi yang digunakan meliputi akurasi, presisi, *recall*, *F1-score* menggunakan data *testing* yang sudah ada. Hal ini bertujuan untuk menilai kinerja model dalam melakukan prediksi secara tepat dan mengidentifikasi pola dengan efektif [18]. Evaluasi ini memastikan model dapat mengidentifikasi sentimen secara efektif.

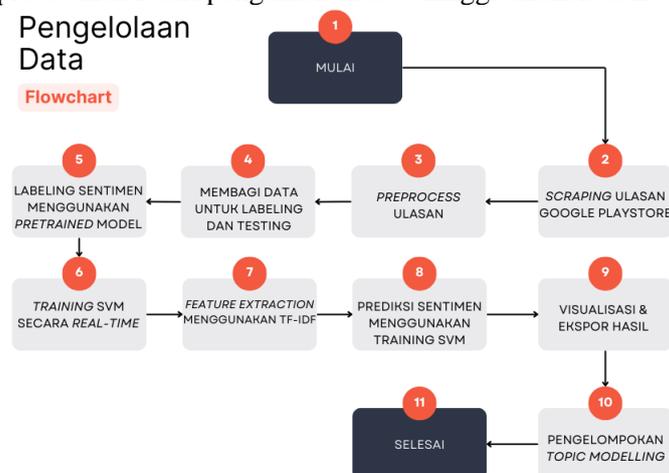
Deployment (Penerapan Model)

Model yang telah dikembangkan diimplementasikan dalam *website* penelitian untuk prediksi sentimen secara *real-time*. Hasil analisis ditampilkan dalam bentuk grafik visualisasi dan dapat diekspor dalam format seperti Excel untuk mendukung pengambilan keputusan atau pengembangan fitur aplikasi lebih lanjut.

Dengan pendekatan ini, penelitian tidak hanya berfokus pada klasifikasi sentimen tetapi juga mengintegrasikan hasil analisis ke dalam sistem yang dapat diakses untuk mendukung pengembangan produk berbasis data.

3.3 Metode Pengelolaan Data

Tahapan pengelolaan data dalam penelitian ini dapat dilihat pada Gambar 3, yang menggambarkan alur proses mulai dari pengambilan data hingga analisis dan visualisasi hasil.



Gambar 3. Pengelolaan data

Pengambilan Data (Scraping)

Data ulasan pengguna aplikasi Flip dikumpulkan secara otomatis melalui *web scraping* dari Google Play Store. Informasi yang diperoleh mencakup teks ulasan, rating, dan tanggal publikasi, dengan batas maksimal 2.000 data untuk efisiensi analisis.

Preprocessing Data

Data yang diperoleh dibersihkan dari noise seperti simbol khusus, angka, serta kata tidak baku. Proses ini mencakup penghapusan *stopwords*, normalisasi teks, dan *filtering* ulasan yang tidak relevan guna memastikan kualitas data.

Pembagian Data

Data yang telah diproses dibagi menjadi dua subset:

- **Subset Labeling:** Data ini akan diberikan label sentimen (positif atau negatif) menggunakan *pretrained* model BERT.
- **Subset Testing:** Digunakan untuk menguji performa model sentimen yang dikembangkan.

Pelabelan Sentimen dengan Model BERT

Model BERT (*Bidirectional Encoder Representations from Transformers*) digunakan sebagai penentuan kategori sentimen pada data ulasan sebelum diterapkan pada algoritma SVM untuk pelatihan lebih lanjut. Proses pelabelan data dilakukan untuk menandai setiap teks dalam dataset guna mengidentifikasi apakah data tersebut bersifat negatif atau positif [19].

Pelatihan Model SVM

Data berlabel digunakan untuk melatih model SVM secara *real-time*, sehingga model dapat menyesuaikan dengan data terbaru dari hasil *scraping*. Model *Support Vector Machine* (SVM) digunakan untuk mengklasifikasikan data dengan menentukan fungsi keputusan sebagaimana ditunjukkan pada Persamaan (1).

$$f(x) = \text{sign}(W \cdot X + b) \quad (1)$$

Keterangan:

W : vektor bobot.

X : vektor fitur input.

b : adalah bias.

\cdot : menunjukkan hasil skalar antara W dan X .

Prediksi Sentimen

Model SVM yang telah dilatih diterapkan untuk memprediksi sentimen pada data testing dan mengklasifikasikannya ke dalam kategori positif atau negatif.

Visualisasi dan Ekspor Data

Hasil prediksi disajikan dalam bentuk grafik atau tabel guna memudahkan interpretasi. Data ini juga dapat diekspor dalam format tertentu untuk analisis lebih lanjut atau laporan.

Topic modelling

Dengan menggunakan *Topic Modeling*, pola ini dapat dianalisis lebih lanjut untuk menentukan topik utama yang sering muncul dalam ulasan, sehingga dapat memudahkan dalam mengidentifikasi dan mengelompokkan topik yang paling relevan dengan kebutuhan analisis.

Seluruh proses dilakukan secara otomatis melalui *website* yang dibuat oleh peneliti, memungkinkan pengelolaan data yang efisien, akurat, dan *real-time*.

4 Hasil dan Pembahasan

Pada bagian ini menyajikan hasil penelitian serta analisis yang dilakukan berdasarkan implementasi kerangka CRISP-DM dalam mengolah data ulasan pengguna aplikasi Flip. Hasil penelitian mencakup setiap tahapan CRISP-DM, mulai dari pemahaman bisnis hingga evaluasi model. Analisis difokuskan pada bagaimana metode *Support Vector Machine* (SVM) digunakan untuk mengklasifikasikan sentimen pengguna, serta bagaimana temuan dari proses ini dapat memberikan wawasan bagi pengembang dalam meningkatkan kualitas aplikasi.

4.1 Implementasi Kerangka CRISP-DM

Dalam penelitian ini, penulis telah mengimplementasikan setiap tahapan dalam kerangka CRISP-DM sebagai pendekatan sistematis untuk menganalisis sentimen pengguna. Berikut adalah hasil dari setiap tahapan yang telah diterapkan dalam penelitian ini.

Business Understanding (Pemahaman Bisnis)

Transformasi digital di sektor keuangan mendorong pertumbuhan aplikasi fintech seperti Flip, yang memudahkan transaksi tanpa biaya administrasi. Seiring meningkatnya pengguna, kualitas

<http://sistemasi.ftik.unisi.ac.id>

produk menjadi prioritas, dan ulasan di Google Play Store menjadi sumber masukan penting bagi pengembang. Namun, *volume* ulasan yang besar menyulitkan analisis manual, sehingga diperlukan pendekatan *data mining* berbasis *Support Vector Machine* (SVM) untuk mengidentifikasi sentimen pengguna secara otomatis.

Penelitian ini menerapkan CRISP-DM sebagai kerangka sistematis dalam mengolah data ulasan pengguna Flip dari Google Play Store, mencakup teks ulasan. Dengan klasifikasi sentimen menggunakan SVM, penelitian ini bertujuan untuk mengidentifikasi fitur dengan sentimen negatif sebagai prioritas pengembangan, mengevaluasi aspek positif untuk dikembangkan lebih lanjut, serta menyusun strategi peningkatan produk berdasarkan pola sentimen yang ditemukan.

Data Understanding (Pemahaman Data)

Penelitian ini memanfaatkan 2.000 *review* pengguna aplikasi Flip yang dikumpulkan langsung dari Google Play Store melalui aplikasi *web scraping* yang dikembangkan oleh peneliti. Data mencakup tanggal ulasan dan teks ulasan, yang mencerminkan pengalaman pengguna, baik dalam bentuk keluhan, masukan, maupun apresiasi. Ulasan yang diambil merupakan data terbaru, sehingga hasil analisis dapat menyajikan informasi terbaru bagi pengembang untuk meningkatkan mutu produk. Batas pengambilan data ditetapkan pada 2.000 ulasan sesuai dengan kapasitas aplikasi *scraping*, namun tetap mewakili populasi pengguna aktif yang memberikan ulasan. Tabel 1 menyajikan beberapa contoh ulasan yang telah diperoleh, memberikan wawasan awal mengenai data yang akan dianalisis.

Tabel 1. Ulasan data setelah *scraping*

No	Waktu Ulasan	Ulasan
1	2024-11-03 02:05:29	bagus....
2	2025-01-06 02:05:32	Sangat membantu, terima kasih banyak.
3	2024-11-19 23:56:31	sekarang harus langganan dulu baru bisa ada free untuk trf padahal dulu nggk lho ...
4	2025-01-20 13:07:16	Sekarang pengiriman lama
5	2025-01-06 01:07:34	Sangat membantu, mengurangi biaya Admin
6	2024-12-25 22:09:51	cepat dan lancar dlm transaksi
7	2024-11-05 14:05:11	minimal gratis top up walet 100 kali lahh
8	2024-12-13 10:13:00	Flip sekarang TF aja harus 25 menit nunggu nya lama padahal dulu cepet
9	2024-11-10 11:20:27	mantap, pertahankan
10	2024-12-05 12:45:43	Memang tidak ada biaya admin, tapi lumayan nunggu loading

Data Preparation (Persiapan Data)

Tahap *Data Preparation* bertujuan untuk menyiapkan data ulasan dari Google Play Store agar siap digunakan dalam analisis sentimen dan pemodelan SVM. Proses ini mencakup data *cleaning*, yang melibatkan penghapusan karakter tak relevan, *stopwords*, kata slang, tokenisasi, serta *stemming* untuk meningkatkan konsistensi teks. Proses pembersihan data menghasilkan teks yang lebih terstruktur dan siap untuk analisis, yang ditunjukkan dalam Tabel 2.

Tabel 2. Ulasan data setelah *cleaning*

No	Ulasan Sebelum	Ulasan Sesudah
1	bagus....	bagus
2	Sangat membantu, terima kasih banyak.	sangat membantu terima kasih banyak
3	sekarang harus langganan dulu baru bisa ada free untuk trf padahal dulu nggk lho ...	sekarang langganan dulu baru ada free trf padahal dulu nggk lho
4	Sekarang pengiriman lama	sekarang pengiriman lama
5	Sangat membantu, mengurangi biaya Admin	sangat membantu mengurangi biaya admin
6	cepat dan lancar dlm transaksi	cepat lancar dalam transaksi
7	minimal gratis top up walet 100 kali lahh	minimal gratis top up walet kali lahh
8	Flip sekarang TF aja harus 25 menit	flip sekarang tf aja menit nunggu nya lama

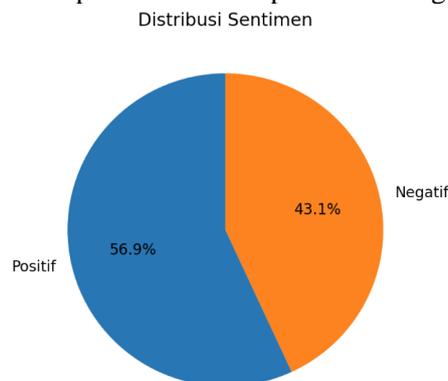
	nunggu nya lama padahal dulu cepet	padahal dulu cepet
9	mantap, pertahankan	mantap pertahankan
10	Memang tidak ada biaya admin, tapi lumayan nunggu loading	memang ada biaya admin lumayan nunggu loading

Setelah itu, dilakukan data *splitting*, Dengan 80% data dialokasikan untuk melatih model, sementara 20% data yang tersisa diterapkan dalam proses pengujian guna mengevaluasi performa model guna memastikan model dapat mengenali pola sentimen dengan baik. Data *integration* menggunakan model *pretrained* BERT untuk memberikan label sentimen pada dataset sebelum digunakan dalam pelatihan SVM. Tabel 3 menyajikan hasil proses labeling data *training* yang dilakukan menggunakan model *pretrained* BERT sebagai tahap awal sebelum pemodelan SVM.

Tabel 3. Labeling data training

No	Ulasan Data Training	Label Sentimen
1	Bagus	Positif
2	sangat membantu terima kasih banyak	Positif
3	sekarang langganan dulu baru ada free trf padahal dulu nggk lho	Negatif
4	sekarang pengiriman lama	Negatif
5	sangat membantu mengurangi biaya admin	Positif
6	cepat lancar dalam transaksi	Positif
7	minimal gratis top up walet kali lahh	Negatif
8	flip sekarang tf aja menit nunggu nya lama padahal dulu cepet	Negatif
9	mantap pertahankan	Positif
10	memang ada biaya admin lumayan nunggu loading	Negatif

Selanjutnya, data *visualization* dilakukan untuk memahami distribusi sentimen, yang menunjukkan 56,9% ulasan positif dan 43,1% negatif, sehingga dapat dievaluasi keseimbangannya. Grafik distribusi sentimen dalam data pelatihan antara positif dan negatif ditampilkan pada Gambar 4.

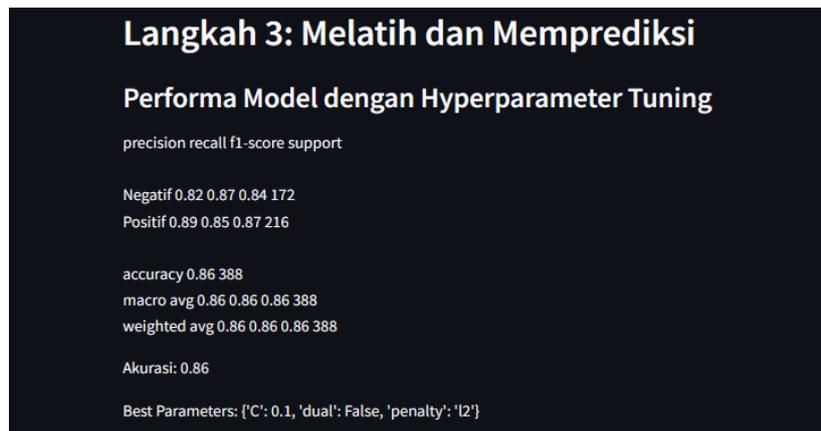


Gambar 4. Disitribusi sentimen data pelatihan

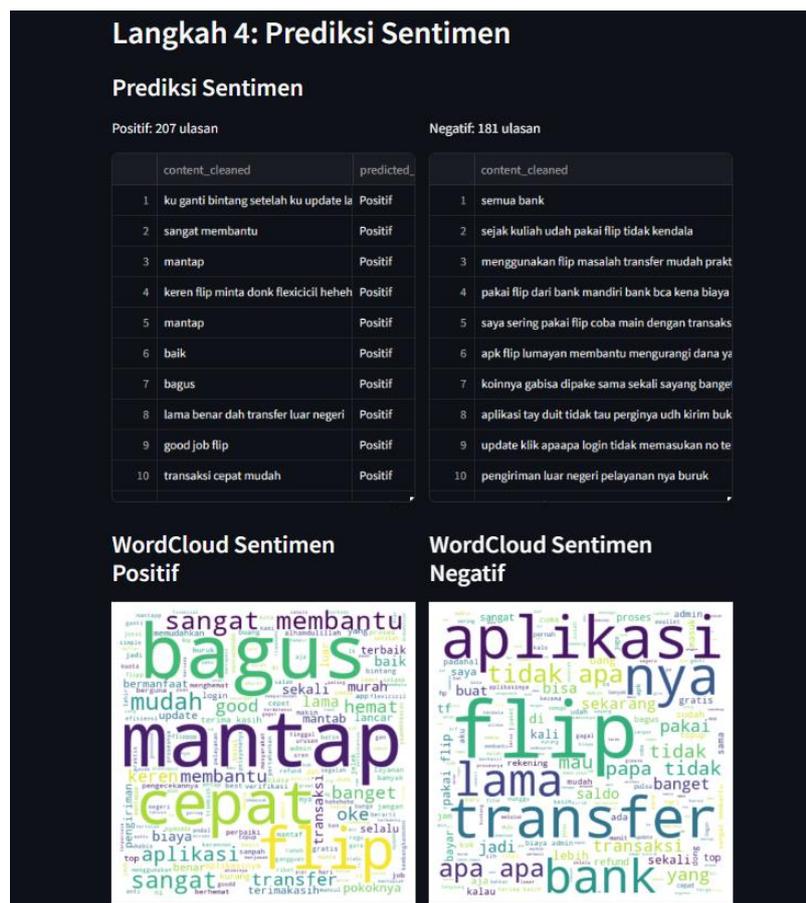
Terakhir, *feature extraction* menggunakan metode TF-IDF diterapkan untuk mengubah teks ulasan menjadi representasi numerik, memastikan model dapat membedakan kata-kata penting dan mengurangi pengaruh kata umum dalam analisis sentimen.

Modeling

Pada tahap *Modeling*, penelitian ini mengimplementasikan *Support Vector Machine* (SVM) sebagai algoritma utama karena keandalannya dalam klasifikasi teks dengan representasi TF-IDF. Model dikembangkan dengan data pelatihan yang telah diolah, dengan identifikasi sentimen positif atau negatif. Untuk mengoptimalkan performa, dilakukan *tuning parameter* menggunakan *GridSearchCV* dengan validasi silang guna menemukan kombinasi parameter terbaik. Setelah pelatihan, model dievaluasi untuk mengukur performanya menggunakan *precision*, *recall*, dan *F1-score*, dengan hasil akurasi keseluruhan sebesar 86%, menunjukkan tingkat keberhasilan yang tinggi dalam mengklasifikasikan sentimen. Model diuji menggunakan data pengujian yang tidak digunakan saat pelatihan, sehingga dapat mengevaluasi kemampuannya dalam mengenali pola sentimen pada ulasan pengguna yang baru. Gambar 5 menampilkan performa dari model SVM dalam melatih dan memprediksi sentimen pengguna, sementara Gambar 6 menampilkan hasil prediksi sentimen pada data uji serta visualisasi *WordCloud* dari kata-kata yang dominan dalam dataset.



Gambar 5. Performa Model SVM



Gambar 6. Prediksi sentimen data uji dan wordcloud

Deployment

Tahap *deployment* tidak termasuk dalam penelitian ini, karena fokus utama adalah analisis sentimen menggunakan kerangka kerja CRISP-DM, bukan implementasi model ke dalam sistem produksi. Namun, hasil prediksi *Support Vector Machine* (SVM) tetap dianalisis untuk memberikan wawasan strategis terkait persepsi pengguna terhadap aplikasi Flip dan prioritas pengembangan produk. Meskipun belum diimplementasikan secara *real-time*, penelitian ini dapat berfungsi sebagai landasan dalam pengembangan sistem otomatis di masa depan untuk analisis ulasan pengguna atau laporan sentimen berkala, sehingga tetap berkontribusi dalam mendukung pengambilan keputusan berbasis data.

4.2 Interpretasi Data Prediksi

Tahap Interpretasi Data bertujuan untuk memahami hasil prediksi model SVM terhadap data uji, yang mencerminkan persepsi pengguna aplikasi Flip.

Hasil Prediksi Data Uji

Dari 388 ulasan uji (20% dari total data), model mengidentifikasi 221 ulasan positif (56,9%) dan 167 ulasan negatif (43,1%). Analisis pola sentimen ini digunakan untuk menghubungkan ulasan dengan fitur aplikasi serta memberikan rekomendasi strategis bagi pengembang dalam meningkatkan mutu produk serta pengalaman pengguna, sebagaimana dicontohkan dalam Tabel 4

Tabel 4. Sentimen hasil prediksi data uji

No	Sentimen Positif	Sentimen Negatif
1	aplikasi fungsional bagus	flip bagus dulu koin ditarik sekarang tidak ditarik makin malas pakai flip koin ditarik kode uniknya lumayan tinggi jadi sama aja kayak bayar tidak bisa ditarik
2	baguss keren cepet proses tanpa admin	terbatas topup flip
3	transaksi mudah murah cepat	sekarang langganan dulu baru ada free trf padahal dulu nggk lho
4	baik memuaskan	awal awal transaksi pakai flip lancar kayak jalan tol sekarang ampun deh loading nya lama ada limit free admin
5	sangat memudahkan masyarakat yang berhemat mantap	kecewa flip sekarang buruk sekali
6	terimakasih banyak tim developer aplikasi luar biasa membantu	ternyata per desember limit transfer topup cuma transfer bank kali topup ewallet kali itu diarahkan beli paket langganan transfer
7	aman amanah cepat	mahal skg ngambil untung nya yang tadinya ratusan perak jadi ribuan
8	aplikasi terbaik buat transfer	kak ya stelah update malah lama banget loading qrisnya
9	transaksi cepat	sekarang proses verifikasi nya jadi lama banget
10	flip terbaik sebelum zaman pandemi pakai flip	mohon maaf admin sekarang flip harganya naik terus buat jualan aja mepet sekali labanya dulu token listrik sekarang jadi naiknya tinggi min coba kalau kembali min terima kasih

Analisis Topic Modelling

Tahap Analisis Topic Modelling diterapkan melalui metode *Latent Dirichlet Allocation* (LDA) dalam mengelompokkan ulasan berdasarkan topik utama yang muncul dalam data, baik untuk sentimen positif maupun negatif. Dengan menganalisis 388 ulasan uji, model mengidentifikasi pola kata dan hubungan antar kata yang sering muncul, sehingga dapat mengungkap aspek yang menjadi keunggulan dan kelemahan aplikasi Flip. Hasil analisis ini menghasilkan kata kunci yang mewakili tiap topik, disajikan pada Tabel 5 dan Tabel 6, yang kemudian diinterpretasikan menggunakan *expert judgment* guna menentukan tema utama dari setiap kelompok topik yang terbentuk.

Tabel 5. Topic modelling sentimen positif

Topik	Kata Kunci	Jumlah	Presentase	Prioritas
Manfaat Aplikasi Flip yang Membantu dan Menghemat Biaya	sangat, membantu, flip, bermanfaat, aplikasi, app, alhamdulillah, berguna, berhemat, nya	54	26.09%	Tinggi
Kepuasan Pengguna terhadap Kemudahan dan Kecepatan Transaksi	mantap, good, pokoknya, mantaf, cepet, lancar, hemat, memudahkan, berhemat, nya	37	17.87%	Tinggi
Pengalaman Positif Pengguna dengan Harga Terjangkau dan Kemudahan Transfer	bagus, mudah, lama, benar, aplikasinya, aplikasi, dah, luar, murah, transfer	33	15.94%	Tinggi
Efisiensi Biaya dan Kecepatan Proses	cepat, hemat, transfer, biaya, murah, sekali, pengiriman,	17	8.21%	Sedang

Transaksi Flip	best, the, memudahkan			
Pelayanan Terbaik dan Dukungan Cepat yang Memuaskan	banget, oke, luar, biasa, betul, admin, terbaik, flip, jangan, cepet	16	7.73%	Sedang
Kemudahan dan Kinerja Lancar dalam Setiap Transaksi	keren, transaksi, cepat, login, lancar, menghemat, perbaiki, update, oke, mudah	13	6.28%	Sedang
Layanan Bebas Biaya yang Membantu dan Mudah Digunakan	baik, terima, kasih, membantu, sangat, mudah, flip, transfer, cepat, gratis	11	5.31%	Sedang
Fitur Lengkap dan Layanan Mudah dari Aplikasi Flip	flip, mantab, good, keren, job, top, pokoknya, banyak, layanan, mudah	10	4.83%	Rendah
Peningkatan Aplikasi untuk Memberikan Pengalaman yang Lebih Baik	aplikasi, selalu, sampah, buruk, aplikasinya, update, flip, app, cepet, alhamdulillah	8	3.86%	Rendah
Perbaikan Verifikasi dan Layanan demi Kepuasan Pengguna	kurang, terimakasih, flip, aplikasi, bintang, ni, verifikasi, update, layanan, jangan	8	3.86%	Rendah

Tabel 6. Topic modelling sentimen negatif

Topik	Kata Kunci	Jumlah	Presentase	Prioritas
Masalah Biaya dan Rekening	aplikasi, bank, transfer, flip, nya, biaya, pakai, rekening, kasih, admin	35	19.34%	Tinggi
Waktu Transaksi yang Lama	flip, pakai, gratis, admin, transaksi, sekarang, lama, top, saldo, nya	34	18.78%	Tinggi
Kebingungan dan Ketidaktepatan Aplikasi	apa, papa, nya, aplikasi, sama, aplikasinya, jadi, kok, bagus, pernah	21	11.6%	Sedang
Gagal Transaksi dan Kecewa	mau, aplikasi, apk, uang, sama, kalau, gagal, aja, semoga, mulu	20	11.05%	Sedang
Masalah pada Proses Transaksi	membantu, sangat, transfer, transaksi, proses, lebih, flip, rekening, cepat, biaya	20	11.05%	Sedang
Keluhan tentang Kecepatan Aplikasi	flip, apa, pakai, banget, lama, aplikasi, bayar, mau, papa, menit	13	7.18%	Sedang
Masalah Pembayaran Token Listrik	tf, flip, listrik, sangat, bank, sekali, mau, login, token, makin	11	6.08%	Sedang
Kesulitan Membayar Pulsa	flip, apa, masuk, bayar, aplikasi, aja, sekarang, pulsa, segera, melalui	10	5.52%	Sedang
Proses yang Lambat dan Kekecewaan	nya, sekali, padahal, makin, proses, lama, sayang, kali, dong, admin	10	5.52%	Sedang
Masalah Top-Up dan Verifikasi	transfer, makin, kali, topup, ewallet, bagaimana, sih, jelas, jadi, konfirmasi	7	3.87%	Rendah

Interpretasi Topik Sentimen Pengguna

Analisis *topic modelling* terhadap ulasan pengguna Flip mengidentifikasi berbagai tema utama, baik untuk sentimen positif maupun negatif, guna memahami persepsi pengguna dan memberikan rekomendasi perbaikan.

Pada sentimen positif, sebagian besar ulasan menyoroti efisiensi biaya, kemudahan transaksi, dan kecepatan layanan, dengan topik utama seperti penghematan biaya (26,09%), kemudahan transaksi (17,87%), dan harga terjangkau (15,94%). Pengguna juga mengapresiasi pelayanan yang responsif dan fitur lengkap, meskipun beberapa mengusulkan peningkatan pada performa aplikasi dan sistem verifikasi.

Sementara itu, sentimen negatif mencerminkan tantangan utama yang dihadapi pengguna, terutama terkait biaya tambahan (19,34%), lamanya waktu transaksi (18,78%), serta kesulitan dalam penggunaan aplikasi (11,6%). Keluhan lainnya mencakup transaksi gagal, performa aplikasi yang lambat, serta kendala dalam pembayaran token listrik dan pulsa. Rekomendasi perbaikan mencakup peningkatan transparansi biaya, optimalisasi infrastruktur transaksi, penyederhanaan antarmuka aplikasi, serta peningkatan respons dukungan pelanggan.

Secara keseluruhan, isu dengan prioritas tinggi seperti biaya transaksi, rekening, dan waktu proses perlu segera diatasi, sementara kendala teknis dan fitur tambahan dapat diperbaiki secara bertahap. Dengan memahami pola sentimen ini, pengembang Flip dapat meningkatkan kualitas produk secara strategis untuk meningkatkan pengalaman pengguna.

5 Kesimpulan

Penelitian ini menunjukkan bahwa implementasi *framework* CRISP-DM efektif dalam analisis sentimen ulasan pengguna aplikasi fintech. Dengan model *Support Vector Machine* (SVM), sentimen pengguna berhasil diklasifikasikan dengan akurasi 86%, terdiri dari 56,9% sentimen positif dan 43,1% negatif. Metode ini terbukti mampu mengidentifikasi persepsi pengguna secara akurat. Keterbatasan penelitian meliputi cakupan data yang hanya diambil dari ulasan terbaru Google Play Store dan penggunaan SVM tanpa perbandingan dengan algoritma lain. Penelitian selanjutnya disarankan untuk memperluas sumber data dari berbagai platform dan mengeksplorasi algoritma lain. Hasil ini menegaskan bahwa CRISP-DM dan *machine learning* dapat menjadi alat strategis bagi pengembang aplikasi fintech dalam menentukan prioritas inovasi berbasis data. Dengan memahami pola sentimen, pengembang dapat mengoptimalkan fitur unggulan, memperbaiki kekurangan, serta meningkatkan kepuasan dan loyalitas pengguna dalam menghadapi persaingan industri.

Referensi

- [1] A. M. Puteri, I. Inanda, R. B. Prasetio, and R. Sanjaya, "Pengaruh Literasi Keuangan dan Literasi Digital terhadap Preferensi Bank Bank Digital di Kalangan Mahasiswa," *Jurnal Publikasi Ilmu Manajemen*, vol. 3, no. 4, pp. 16–25, 2024, doi: 10.55606/jupiman.v3i4.4467.
- [2] N. Kristi, D. F. Shiddieq, and D. Nurhayati, "Analisis Penerimaan Aplikasi Flip menggunakan Model *Unified of Acceptance and use of Technology 3*," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 685–694, Apr. 2024, doi: 10.57152/malcom.v4i2.1316.
- [3] Y. Ngamal and M. A. Perajaka, "Penerapan Model Manajemen Risiko Teknologi Digital di Lembaga Perbankan Berkaca pada Cetak Biru Transformasi Digital Perbankan Indonesia," *Jurnal Manajemen Risiko*, vol. 2, no. 2, pp. 59–74, 2022, doi: 10.33541/mr.v2iIV.4099.
- [4] Google Play Store, "Jumlah Ulasan Aplikasi Flip pada *Google Play Store*."
- [5] A. Fadilla, E. Army, Y. Dwi, P. Rustam, A. Indrijawati, and G. T. Pontoh, "Peran *Artificial Intelligence* dalam meningkatkan Kualitas Audit: Tinjauan Literatur Sistematis," *Jurnal Akuntansi dan Governance*, vol. 5, no. 2, pp. 146–165, 2025, doi: 10.24853/jago.5.2.146-165.
- [6] A. Syafa'aturrohmah, O. Nurdiawan, F. M. Basysyar, and M. Sulaeman, "Naive Bayes meningkatkan Model Analisis Sentimen pada Ulasan Aplikasi DANA di *Playstore* Indonesia," *Information Management for Educators and Professionals*, vol. 9, no. 2, pp. 171–180, 2024, doi: 10.51211/imbi.v9i2.3330.

- [7] Y. Basrina, N. Afryansih, and T. Febriani, "Pengembangan Aplikasi Evaluasi Pembelajaran Wizer.Me pada Mata Pelajaran IPS di MTs Darussalam Aryojeding," *Jurnal Pendidikan dan Ilmu Geografi*, vol. 8, no. 1, pp. 31–38, 2023, doi: 10.21067/jpig.v8i1.7361.
- [8] F. Xaverius, K. Lile, and I. Suharjo, "Optimalisasi Perbandingan Algoritma *K-Nearest Neighbors* dan *Decision Tree* untuk Prediksi Kemenangan di *MPL Season 13 Mobile Legend*," *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no. 4, pp. 7553–7560, 2024, doi: 10.36040/jati.v8i4.10343.
- [9] M. N. Muttaqin and I. Kharisudin, "Analisis Sentimen pada Ulasan Aplikasi Gojek menggunakan Metode *Support Vector Machine* dan *K Nearest Neighbor*," *UNNES Journal of Mathematics*, vol. 10, no. 2, pp. 22–27, 2021, doi: 10.15294/ujm.v10i2.48474.
- [10] M. Mardiana, "Sentimen Analisis terhadap Aplikasi Alfagift pada *Google Play Store* dengan Algoritma *Support Vector Machine (SVM)*," Palembang, 2023.
- [11] L. A. Fudholi, N. Rahaningsih, and R. D. Dana, "Sentimen Analisis Perilaku Penggemar *Coldplay* di Media Sosial *Twitter* menggunakan Metode *Naive Bayes*," *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no. 3, pp. 4150–4159, 2024, doi: 10.36040/jati.v8i3.9827.
- [12] R. A. Saputra, D. P. Ray, and F. Irwiensyah, "Analisis Sentimen Aplikasi Tokocrypto berdasarkan Ulasan pada *Google Play Store* menggunakan Metode *Naive Bayes*," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 4, pp. 2028–2036, 2024, doi: 10.30865/klik.v4i4.1707.
- [13] L. Purnama and T. Wahyudi, "Analisa Sentimen tentang Piala Dunia U-20 Indonesia menggunakan Algoritma *K-Nearest Neighbor*," *Jurnal Sains dan Teknologi*, vol. 6, no. 2, pp. 217–222, 2024, doi: 10.55338/saintek.v6i2.1397.
- [14] T. Ridwansyah, "Implementasi *Text Mining* terhadap Analisis Sentimen Masyarakat Dunia di *Twitter* terhadap Kota Medan menggunakan *K-Fold Cross Validation* dan *Naive Bayes Classifier*," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 5, pp. 178–185, 2022, doi: 10.30865/klik.v2i5.362.
- [15] R. K. Dewi, Q. J. Adrian, H. Sulistiani, and F. Isnaini, "Dashboard Interaktif untuk Sistem Informasi Keuangan pada Pondok Pesantren Mazroatul'ulum," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. 2, no. 2, pp. 116–121, 2021, doi: 10.33365/jtsi.v2i2.883.
- [16] S. Navisa, L. Hakim, and A. Nabilah, "Komparasi Algoritma Klasifikasi Genre Musik pada *Spotify* menggunakan CRISP-DM," *Jurnal Sistem Cerdas*, vol. 4, no. 2, pp. 114–125, 2021, doi: 10.37396/jsc.v4i2.162.
- [17] A. S. Muliana, D. Lestarini, and S. P. Raflesia, "Analysis of Public Sentiment on Election Results using *Naive Bayes* in Social Media X," *Sistemasi: Jurnal Sistem Informasi*, vol. 13, no. 6, pp. 2467–2478, 2024, doi: 10.32520/stmsi.v13i6.4592.
- [18] E. Constancio and K. Tania, "Penerapan Metode *Supervised Learning* dan Teknik Resampling untuk Prediksi Penipuan Transaksi Keuangan," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, pp. 1427–1439, Dec. 2024, doi: 10.47065/bits.v6i3.6110.
- [19] M. Zahirma and R. Rumini, "Sentiment Analysis of Sunscreen Product Reviews using *Naive Bayes Classifier Algorithm*," *Sistemasi: Jurnal Sistem Informasi*, vol. 14, no. 1, pp. 66–75, 2025.