

Analisis Sentimen Ulasan Pemain *Genshin Impact* Menggunakan Kombinasi TF-IDF, Lexicon, dan *Support Vector Machine*

Mochammad Fiqi Fahrudillah¹, Danang Arbian Sulisty¹

¹Teknik Informatika, Fakultas Teknologi dan Desain, Institut Teknologi dan Bisnis Asia Malang

¹danangarbian@gmail.com *

Abstract

The rapid growth of the digital gaming industry in Indonesia has been accompanied by a significant increase in user-generated reviews on distribution platforms such as Google Play Store. This condition necessitates automated methods capable of efficiently interpreting player perceptions on a scale. This study conducts sentiment analysis on player reviews of *Genshin Impact* by developing a seven-stage analytical pipeline consisting of data preparation, lexicon-based labeling, TF-IDF feature extraction, Support Vector Machine (SVM) training, multi-metric evaluation, rule-based post-processing, and automated summarization using a Large Language Model. A total of 40,000 reviews from 2023 until 2025 were collected through web scraping and processed through text cleaning, slang normalization, tokenization, stopword removal, and stemming. Initial labels were generated using an updated domain-specific sentiment lexicon and subsequently refined through a rule-patch mechanism that handles negation, contrastive expressions, and domain-specific technical cues such as lag, bug, and crash. The SVM model was trained using a TF-IDF configuration (1–3 grams) and evaluated across 10 runs with different random seeds, producing an average accuracy of 0.945, a macro-F1 of 0.900, and stable performance across iterations. Visualization of sentiment distribution and WordClouds highlights prominent thematic patterns within each class, while automated summarization using IBM Granite provides qualitative insights into player appreciation of visual and character design, alongside complaints related to performance issues and the game's gacha system. Overall, the integration of statistical, rule-based, and LLM-driven approaches demonstrates an effective and contextually robust framework for sentiment analysis in game analytics.

Keywords: Sentiment Analysis, TF-IDF, Lexicon, SVM, *Genshin Impact*

Abstrak

Perkembangan industri permainan digital di Indonesia diikuti oleh meningkatnya jumlah ulasan pengguna pada platform seperti *Google Play Store*, sehingga diperlukan metode analisis otomatis untuk memahami persepsi pemain secara efisien. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pemain terhadap permainan *Genshin Impact* dengan membangun pipeline tujuh tahap yang mencakup data preparation, lexicon-based labeling, TF-IDF feature extraction, pelatihan model Support Vector Machine (SVM), evaluasi multi metrik, rule-based post-processing, serta automated summarization menggunakan Large Language Model. Dataset berisi 40.000 ulasan dari tahun 2023 sampai 2025 yang diperoleh melalui web scraping, kemudian diproses melalui pembersihan teks, normalisasi slang, tokenisasi, stopword removal, dan stemming. Label awal dihasilkan menggunakan lexicon domain spesifik yang diperbarui berdasarkan konteks permainan, kemudian disempurnakan melalui mekanisme rule patch untuk menangani negasi, frasa kontras, dan istilah teknis seperti lag, bug, dan crash. Model SVM dilatih menggunakan konfigurasi TF-IDF (1–3 grams) dan diuji sebanyak 10 kali dengan random seed berbeda. Penelitian ini menghasilkan akurasi rata-rata 0.94, macro-F1 0.90. Visualisasi distribusi sentimen dan WordCloud memperlihatkan tema-tema utama tiap kelas, sementara ringkasan otomatis menggunakan IBM Granite memberikan insight kualitatif mengenai apresiasi pemain terhadap elemen visual dan karakter, serta keluhan terhadap performa teknis dan sistem gacha. Hasil penelitian menunjukkan bahwa kombinasi metode statistik, rule-based, dan LLM mampu menghasilkan analisis sentimen yang akurat, stabil, dan relevan secara konteks untuk kebutuhan game analytics.

Kata kunci: Analisis Sentimen, TF-IDF, Lexicon, SVM, *Genshin Impact*

©This work is licensed under a Creative Commons Attribution -ShareAlike 4.0 International License

1. Pendahuluan

Perkembangan teknologi informasi telah mendorong peningkatan signifikan dalam jumlah ulasan pengguna pada berbagai platform digital [1]. Salah satu platform yang memiliki volume ulasan besar adalah *Google Play Store*, di mana pengguna secara aktif menyampaikan opini terhadap aplikasi maupun permainan digital yang mereka gunakan. Ulasan-ulasan tersebut mencerminkan persepsi, kepuasan, maupun keluhan pengguna terhadap suatu produk, sehingga dapat dimanfaatkan sebagai sumber informasi dalam pengambilan keputusan, peningkatan kualitas layanan,

maupun pemahaman perilaku pengguna secara lebih luas [1].

Namun, proses analisis opini dalam jumlah besar secara manual menjadi tidak efisien dan sulit dilakukan secara konsisten. Oleh karena itu, diperlukan pendekatan berbasis Natural Language Processing (NLP) dan Machine Learning (ML) untuk mengotomatisasi proses klasifikasi sentimen [2]. Salah satu algoritma yang banyak digunakan dalam klasifikasi teks adalah Support Vector Machine (SVM), yang dikenal memiliki kemampuan yang baik dalam menangani data berdimensi tinggi serta menghasilkan

performa yang relatif stabil pada berbagai domain penelitian [3]. Selain itu, metode *lexicon-based* seperti kamus *InSet* dan *VADER* juga banyak dimanfaatkan untuk mempercepat proses pelabelan otomatis, khususnya pada teks berbahasa Indonesia [4], [5].

Kombinasi antara *lexicon-based labeling* dan algoritma SVM telah menunjukkan hasil yang menjanjikan dalam berbagai penelitian terdahulu. Penelitian yang dilakukan oleh Siregar et al. [1] menunjukkan bahwa SVM dengan pembobotan TF-IDF mampu mengelompokkan ulasan permainan digital dengan akurasi di atas 79%, meskipun masih menghadapi kesulitan dalam mengklasifikasikan sentimen netral. Studi lain oleh Khotimah et al. [6] menerapkan algoritma Naïve Bayes untuk analisis sentimen ulasan aplikasi investasi di *Google Play Store* dan memperoleh akurasi tinggi, namun dengan menghilangkan kelas sentimen netral serta bergantung pada translasi bahasa dan teknik *oversampling* untuk menangani ketidakseimbangan data. Sementara itu, penelitian oleh Aprinastya et al. [7] membandingkan performa Naïve Bayes dan SVM pada ulasan *Genshin Impact* dalam konteks multibahasa, dan menemukan bahwa SVM memiliki performa yang lebih rendah pada dataset berbahasa Indonesia akibat kompleksitas morfologi dan struktur linguistik yang khas.

Temuan-temuan tersebut menunjukkan bahwa meskipun pendekatan berbasis *Support Vector Machine* (SVM) dan *lexicon-based* cukup efektif dalam analisis sentimen, masih terdapat sejumlah tantangan utama, khususnya pada konteks ulasan permainan digital berbahasa Indonesia. Tantangan tersebut tidak hanya berkaitan dengan ketidakseimbangan distribusi kelas sentimen, tetapi juga dengan kompleksitas bahasa informal yang digunakan oleh pemain, seperti singkatan, slang, serta struktur kalimat yang tidak baku. Kondisi ini menyebabkan model statistik dan kamus sentimen umum belum sepenuhnya mampu menangkap makna sentimen secara utuh.

Pendekatan *lexicon* konvensional juga memiliki keterbatasan dalam menangkap konteks linguistik yang bersifat domain-spesifik. Pada ulasan permainan digital, pemain sering menggunakan istilah teknis, frasa komunitas, serta ekspresi campuran antara pujian dan keluhan dalam satu kalimat, yang tidak tercakup dalam kamus umum. Akibatnya, proses pelabelan otomatis berpotensi menghasilkan representasi sentimen yang kurang akurat secara semantik, terutama pada kasus sentimen netral atau ulasan dengan makna implisit. Hal ini menunjukkan bahwa analisis sentimen pada domain permainan memerlukan pendekatan yang lebih adaptif terhadap konteks dan karakteristik bahasa pengguna.

Selain itu, sebagian besar penelitian terdahulu masih berfokus pada evaluasi performa model klasifikasi berdasarkan metrik kuantitatif, seperti *accuracy* dan *F1-score*, tanpa memberikan penjelasan kualitatif yang

mendalam mengenai isi dan kecenderungan opini pengguna. Padahal, dalam konteks pengambilan keputusan dan evaluasi pengalaman pengguna, pemahaman naratif mengenai alasan di balik sentimen positif maupun negatif menjadi aspek yang sama pentingnya dengan nilai performa numerik. Keterbatasan ini membuka peluang untuk mengintegrasikan pendekatan analisis sentimen dengan metode interpretatif yang mampu merangkum dan menjelaskan pola opini pengguna secara lebih komprehensif.

Oleh karena itu, pengembangan dan penyesuaian kamus *lexicon* berbasis konteks domain menjadi penting agar proses pelabelan lebih relevan terhadap karakteristik data yang dianalisis. Dalam penelitian ini, penulis mengembangkan *domain-specific lexicon* untuk permainan *Genshin Impact*, kemudian mengombinasikannya dengan representasi fitur *Term Frequency-Inverse Document Frequency* (TF-IDF) dan algoritma SVM. Untuk mengatasi keterbatasan model statistik dalam memahami konteks linguistik tertentu, penelitian ini juga menambahkan tahap *rule-based post-processing* guna menangani negasi, frasa kontras, serta istilah teknis permainan yang sering muncul dalam ulasan pemain.

Selain analisis kuantitatif, penelitian ini turut memperkaya interpretasi hasil dengan menambahkan proses pasca-prediksi berupa *automated summarization* menggunakan *Large Language Model* (LLM) IBM Granite. Tahap ini bertujuan untuk merangkum kecenderungan opini positif, negatif, dan netral secara naratif tanpa memengaruhi hasil klasifikasi numerik. Selain itu hal ini juga bisa menjadi solusi untuk perhitungan manual dibanding perhitungan secara otomatis [8]. Pendekatan ini sejalan dengan tren penelitian terkini yang mulai memanfaatkan model bahasa besar sebagai alat *post-hoc interpretation* untuk meningkatkan pemahaman kualitatif terhadap hasil analisis sentimen [9], [10].

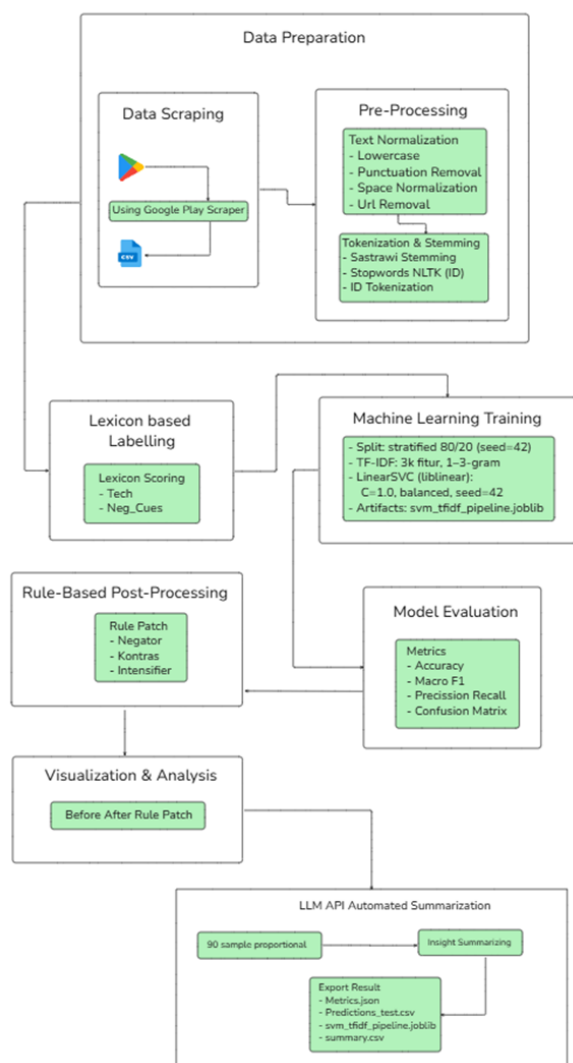
Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen ulasan pemain *Genshin Impact* di *Google Play Store* dengan mengombinasikan pendekatan *lexicon-based labeling*, representasi fitur TF-IDF, dan algoritma *Support Vector Machine* (SVM), serta dilengkapi dengan *rule-based post-processing* dan *automated summarization berbasis Large Language Model*. Pendekatan ini diharapkan mampu menghasilkan analisis sentimen yang lebih akurat, stabil, dan kontekstual pada domain permainan digital berbahasa Indonesia.

2. Metode Penelitian

2.1. Pipeline Penelitian

Penelitian ini menggunakan pendekatan terstruktur berbasis *pipeline* untuk memastikan setiap tahapan analisis sentimen dilakukan secara sistematis dan saling terintegrasi. Pendekatan *pipeline* dipilih karena mampu menggambarkan alur pemrosesan data secara

menyeluruh, mulai dari tahap pengumpulan dan persiapan data hingga interpretasi hasil analisis. Setiap tahapan dirancang agar keluaran dari satu proses menjadi masukan bagi proses berikutnya, sehingga keseluruhan metode penelitian dapat direplikasi dan dievaluasi secara konsisten. Alur metodologi yang digunakan dalam penelitian ini divisualisasikan secara ringkas dan komprehensif pada Gambar 1, yang menampilkan hubungan antar tahap dalam proses analisis sentimen ulasan pemain *Genshin Impact*.



Gambar 1. Pipeline Architecture

Dapat dilihat pada Gambar 1 bahwa alur penelitian yang digunakan dalam penelitian ini direpresentasikan dalam bentuk *pipeline architecture* yang tersusun secara berurutan dan saling terhubung. *Pipeline* tersebut terdiri dari tujuh fase utama, yaitu *Data Preparation*, *Lexicon-Based Labeling*, *Machine Learning Training*, *Model Evaluation*, *Rule-Based Post-Processing*, *Visualization & Analysis*, dan *Automated Summarization*. Setiap fase dirancang untuk menghasilkan keluaran yang menjadi masukan bagi tahap berikutnya, sehingga membentuk alur analisis sentimen yang terstruktur dan sistematis. Melalui

desain *pipeline* ini, proses pengolahan data tidak hanya mengandalkan pendekatan statistik berbasis *machine learning*, tetapi juga diperkaya dengan aturan linguistik dan peringkasan berbasis *Large Language Model*, sehingga keseluruhan tahapan analisis dapat dilakukan secara terpadu dan mudah direplikasi pada penelitian sejenis.

2.2. Data Preparation

Tahap *data preparation* dilakukan untuk memperoleh data ulasan pemain *Genshin Impact* yang siap digunakan dalam proses analisis sentimen. Data penelitian ini diperoleh dari platform *Google Play Store* melalui teknik *web scraping* menggunakan pustaka *google-play-scraper*, yang memungkinkan pengambilan ulasan pengguna aplikasi secara otomatis berdasarkan kriteria tertentu. Pengumpulan data dilakukan pada rentang waktu tahun 2023 hingga 2025, sehingga data yang digunakan mencerminkan opini pemain yang relatif mutakhir terhadap perkembangan permainan. Penggunaan *Google Play Store* sebagai sumber data ulasan telah banyak diterapkan pada penelitian analisis sentimen aplikasi digital karena ketersediaan data yang besar dan beragam [6].

Dataset awal yang diperoleh terdiri dari 40.000 ulasan pemain, yang masing-masing memuat beberapa atribut utama, antara lain *userName*, *content* (teks ulasan), *score* (rating bintang), *reviewCreatedVersion*, dan *timestamp* (at). Ulasan yang dikumpulkan berupa teks bebas berbahasa Indonesia dengan karakteristik bahasa informal, termasuk penggunaan singkatan, kata slang, dan ekspresi nonformal yang umum digunakan oleh pemain gim.

Data hasil *scraping* selanjutnya melalui tahap validasi untuk memastikan konsistensi struktur dan kelayakan isinya. Proses validasi meliputi pengecekan kesesuaian kolom, penghapusan data duplikat, serta pembesihan entri yang tidak lengkap atau memiliki nilai kosong. Setelah proses validasi, diperoleh sebanyak 39.842 ulasan valid yang digunakan sebagai *dataset* akhir dalam penelitian ini. Tahap validasi awal ini penting untuk menjamin kualitas data sebelum masuk ke tahap pemrosesan lanjutan, sebagaimana diterapkan pada penelitian analisis sentimen berbahasa Indonesia [11].

Teks ulasan yang telah divalidasi kemudian diproses melalui serangkaian tahapan *text preprocessing* untuk mengurangi *noise* dan menyeragamkan struktur bahasa. Tahapan tersebut mencakup normalisasi huruf menjadi *lowercase*, pembersihan karakter *non-alphanumeric* menggunakan *regular expression*, normalisasi kata slang berdasarkan kamus yang diperbarui, tokenisasi, penghapusan *stopword* bahasa Indonesia, serta proses *stemming* menggunakan pustaka *Sastrawi Stemmer* [12]. Serangkaian langkah *preprocessing* ini diterapkan untuk menyesuaikan karakteristik bahasa informal dalam ulasan gim dan meningkatkan kualitas representasi fitur pada model berbasis TF-IDF dan SVM, sebagaimana direkomendasikan pada penelitian terkait [13], [14].

2.3. *Lexicon Based Labeling*

Tahap pelabelan sentimen dilakukan dengan pendekatan *lexicon-based*, yang digunakan untuk menghasilkan *pseudo ground-truth* sebelum proses pelatihan model. Pendekatan ini dipilih karena dataset ulasan tidak memiliki anotasi sentimen bawaan, sehingga diperlukan metode otomatis untuk memberikan label awal yang konsisten. Studi yang memanfaatkan *lexicon* sebagai dasar pelabelan otomatis telah banyak digunakan dalam analisis sentimen berbahasa Indonesia [15], [16].

Proses pelabelan dilakukan dengan memetakan setiap token dalam teks ke dalam dua kamus sentimen utama, yaitu kosakata positif dan kosakata negatif. Penelitian ini menggunakan *base lexicon* dari InSet, kemudian memperluasnya menjadi *lexicon domain-spesifik* yang mencakup kosakata khas permainan digital, termasuk istilah teknis dan istilah gacha yang sering muncul dalam ulasan pemain. Penambahan kosakata ini mengikuti praktik umum pada penelitian analisis sentimen bertema hiburan dan permainan, sebagaimana ditunjukkan pada studi mengenai ulasan gim *mobile* [17], yang menekankan perlunya adaptasi *lexicon* terhadap domain yang dianalisis.

Setiap ulasan diberi skor sentimen berdasarkan akumulasi bobot token yang sesuai dengan kosakata positif dan negatif. Skor ini kemudian dikonversi menjadi label sentimen melalui aturan berbasis nilai total, sehingga setiap ulasan dapat diberi label positif, negatif, atau netral. Penggunaan mekanisme scoring seperti ini juga diterapkan pada metode *lexicon-based* lain dalam beberapa publikasi di bidang analisis opini, termasuk penelitian mengenai persepsi publik terhadap layanan digital dan isu sosial. Dengan cara ini, tahap pelabelan dapat menghasilkan label awal yang seragam dan siap digunakan pada proses pelatihan model pada tahap berikutnya.

2.4. *Machine Learning Training*

Proses pelatihan model dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dengan pendekatan linear yang banyak digunakan dalam analisis sentimen teks berbahasa Indonesia. Pemilihan algoritma ini didasarkan pada stabilitasnya dalam menangani data berdimensi tinggi yang dihasilkan oleh representasi TF-IDF, serta performanya yang terbukti unggul dalam berbagai studi sebelumnya. Sejumlah penelitian terdahulu [1], [18]. Menunjukkan bahwa SVM mampu memberikan hasil yang konsisten pada berbagai jenis data opini, termasuk data tidak terstruktur dari media sosial dan platform ulasan aplikasi.

Sebelum proses pelatihan dilakukan, teks ulasan direpresentasikan dalam bentuk vektor melalui metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Representasi ini digunakan untuk menangkap bobot kata berdasarkan frekuensi kemunculannya serta signifikansinya dalam keseluruhan korpus. Model

kemudian dilatih menggunakan data yang telah melalui proses pelabelan otomatis, sehingga SVM berfungsi sebagai model *supervised learning* yang mempelajari hubungan antara fitur tekstual dan label sentimen awal. Pendekatan pelatihan ini sejalan dengan metodologi yang digunakan pada penelitian terkait analisis sentimen aplikasi *Play Store* [1], yang menunjukkan bahwa TF-IDF dan SVM merupakan kombinasi yang efektif untuk memetakan pola sentimen dalam ulasan berbahasa Indonesia.

Untuk memastikan kinerja model yang stabil dan dapat direplikasi, proses pelatihan dilakukan secara berulang dengan beberapa konfigurasi yang tetap, termasuk penggunaan pembagian data yang seimbang antara data latih dan data uji. Pendekatan multi percobaan ini mengikuti praktik umum dalam penelitian *machine learning* modern, yang menekankan pentingnya evaluasi berulang untuk mengurangi pengaruh variabilitas data acak. Dengan cara ini, model yang dihasilkan dapat diverifikasi konsistensinya dan dipastikan memiliki kemampuan generalisasi yang baik terhadap data ulasan yang lebih luas.

2.5. *Model Evaluation*

Tahap evaluasi model dilakukan untuk menilai kemampuan algoritma dalam mengklasifikasikan ulasan berdasarkan label sentimen yang telah ditetapkan pada tahap sebelumnya. Proses evaluasi menggunakan skema pengujian yang mempertimbangkan performa model dari berbagai aspek, sehingga dapat memberikan gambaran menyeluruh mengenai efektivitas pendekatan yang digunakan. Evaluasi berbasis *multi-metric* diterapkan karena model sentimen umumnya menghadapi variasi panjang teks, penggunaan bahasa informal, serta ketidakseimbangan distribusi kelas yang dapat memengaruhi kualitas prediksi. Pendekatan ini sejalan dengan praktik evaluasi yang dilakukan dalam penelitian analisis sentimen berbahasa Indonesia di platform digital [19], yang memanfaatkan beberapa metrik untuk memperoleh penilaian yang lebih komprehensif.

Proses evaluasi memanfaatkan metrik utama dalam klasifikasi teks, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*, yang masing-masing memberikan perspektif berbeda mengenai kinerja model. Pemilihan metrik ini didasarkan pada sifat data ulasan gim yang cenderung memiliki distribusi kelas yang tidak merata dan variasi intensitas emosi antarulasan. Dalam konteks seperti ini, satu metrik saja tidak cukup untuk menilai performa model secara utuh [11], [20].

Proses evaluasi dilakukan pada data uji yang dipisahkan dari data latih untuk memastikan bahwa performa model tidak hanya baik pada data yang dikenal tetapi juga pada data baru. Pendekatan evaluasi berulang digunakan untuk memeriksa konsistensi performa model terhadap variasi pembagian data. Prosedur ini memungkinkan analisis terhadap stabilitas hasil dan mendeteksi potensi bias yang mungkin

muncul selama proses pelatihan. Dengan strategi ini, evaluasi dapat memberikan landasan yang kuat untuk menentukan keandalan model sebelum diterapkan pada analisis sentimen skala besar pada tahap berikutnya.

2.6. Rule Based Post Processing

Tahap *rule-based post-processing* diterapkan untuk menyempurnakan label sentimen yang dihasilkan oleh model SVM dengan mempertimbangkan konteks linguistik yang tidak dapat ditangani secara optimal melalui pendekatan statistik berbasis fitur. Pendekatan ini diperlukan karena ulasan permainan digital sering kali mengandung ekspresi kebahasaan yang kompleks, seperti negasi, frasa kontras, dan istilah teknis yang mengindikasikan keluhan, tetapi tidak selalu terdeteksi oleh model berbasis TF-IDF. Beberapa penelitian sebelumnya menunjukkan bahwa penggabungan model statistik dengan aturan linguistik dapat meningkatkan koherensi makna dan ketepatan label, terutama pada teks berbahasa Indonesia yang cenderung lebih variatif dan informal. Hal ini terlihat pada penelitian yang menggabungkan pendekatan *lexicon* dan SVM [21].

Proses *rule patch* mencakup analisis terhadap struktur token dalam ulasan untuk mendeteksi pola kebahasaan yang relevan. Aturan negasi digunakan untuk mengidentifikasi frasa yang membalikkan makna sentimen, seperti pada kalimat dengan penggunaan kata “tidak”, “nggak”, atau “bukan” yang diikuti kata positif. Aturan frasa kontras diterapkan pada ulasan yang menyertakan kata hubung seperti “tapi”, “namun”, atau “meski”, karena frasa setelah konjungsi tersebut sering kali mengandung sentimen utama yang berbeda dari bagian sebelumnya [22]. Selain itu, istilah teknis seperti lag, bug, crash, server, ping, dan berbagai keluhan terkait performa permainan dianalisis sebagai indikator kuat sentimen negatif. Pendekatan berbasis aturan seperti ini juga digunakan pada penelitian analisis sentimen domain permainan dan aplikasi daring, di mana istilah teknis sering memiliki bobot semantik yang lebih signifikan daripada kata sifat umum.

Aturan-aturan tersebut diimplementasikan dalam bentuk fungsi yang memeriksa token hasil pra-pemrosesan dan kemudian menyesuaikan label berdasarkan konteks kemunculan pola linguistik tertentu. Pendekatan ini memungkinkan refinemen label yang lebih sensitif terhadap struktur bahasa dan kondisi domain ulasan permainan digital. Dengan memadukan informasi linguistik dan keluaran model statistik, proses *rule patch* berfungsi sebagai lapisan koreksi tambahan untuk meningkatkan ketepatan makna tanpa menggantikan proses klasifikasi utama. Strategi ini sejalan dengan praktik *hybrid sentiment classification* yang digunakan dalam beberapa penelitian yang menggabungkan model berbasis aturan dan machine learning pada teks berbahasa Indonesia.

2.7. Visualisasi & Analisis

Tahap visualisasi digunakan untuk menyajikan representasi grafis dari data sentimen yang telah diproses, sehingga pola-pola utama dalam teks ulasan dapat terlihat secara lebih intuitif sebelum dilakukan interpretasi pada tahap analisis. Visualisasi berfungsi sebagai langkah pendukung dalam mengidentifikasi kecenderungan tematik, distribusi sentimen, serta karakteristik linguistik dalam dataset. Pendekatan ini umum digunakan dalam penelitian analisis sentimen karena mampu memberikan gambaran awal mengenai komposisi data, keberadaan noise, maupun indikasi ketidakseimbangan kelas. Dalam beberapa penelitian terdahulu [1], visualisasi berbasis grafik dan peta kata digunakan sebagai bagian dari tahap analisis awal untuk memahami struktur opini dalam ulasan digital.

Representasi visual yang digunakan meliputi diagram distribusi sentimen, grafik pola kemunculan kata, dan *WordCloud* untuk menunjukkan kata-kata yang paling sering muncul pada setiap kategori sentimen. Visualisasi ini dibuat setelah seluruh proses preprocessing, pelabelan, dan penerapan aturan linguistik selesai dilakukan, sehingga data yang divisualisasikan telah berada pada kondisi yang paling bersih dan stabil. Teknik visualisasi ini memanfaatkan pustaka seperti *matplotlib*, *seaborn*, dan *wordcloud*, yang umum digunakan dalam penelitian berbasis teks. Penggunaan *WordCloud* sebagai salah satu alat eksplorasi linguistik juga sejalan dengan praktik pada penelitian analisis sentimen aplikasi dan media sosial [10], yang menekankan bahwa peta kata dapat membantu mengidentifikasi topik dominan pada dataset berbahasa Indonesia.

2.8. LLM API Summarization

Tahap LLM API Summarization digunakan untuk menghasilkan ringkasan otomatis dari kumpulan ulasan yang telah melalui seluruh proses klasifikasi dan penyempurnaan berbasis aturan. Selain itu, LLM juga terbukti sangat mudah dalam adaptasi teknologi untuk kedepan [23]. Ringkasan ini berfungsi sebagai pelengkap analisis kuantitatif dengan memberikan gambaran kualitatif mengenai tema-tema utama yang muncul dalam ulasan pemain. Pada penelitian ini digunakan model bahasa besar berbasis API, yaitu IBM Granite, yang dipilih karena kemampuannya dalam memahami konteks teks berbahasa Indonesia serta menghasilkan ringkasan yang padat dan relevan. Pendekatan serupa digunakan pada penelitian yang menerapkan LLM untuk summarization dalam domain kesehatan dan opini public [9], yang menunjukkan bahwa model bahasa besar mampu menangkap struktur tematik dalam dokumen panjang secara efektif [24].

Proses summarization dilakukan dengan memilih sampel ulasan yang mewakili distribusi sentimen yang dihasilkan pada tahap sebelumnya. Pemilihan sampel yang proporsional terhadap komposisi kelas bertujuan memastikan bahwa model menerima masukan yang mencerminkan keragaman pola ekspresi dalam ulasan

pemain, sehingga ringkasan yang dihasilkan dapat memberikan gambaran umum yang akurat. Ulasan tersebut kemudian dikirimkan ke layanan LLM API Granite dalam format teks terstruktur. Model memproses kumpulan ulasan ini dengan mempertimbangkan konteks linguistik serta pola semantik yang terkandung di dalamnya, kemudian menghasilkan ringkasan tematik yang bersifat deskriptif dan sesuai dengan karakteristik domain permainan digital.

Penggunaan LLM sebagai bagian dari metodologi penelitian memungkinkan analisis yang lebih komprehensif karena menggabungkan teknik *machine learning* klasik dan pemrosesan bahasa alami berbasis model berskala besar. Pendekatan hybrid seperti ini semakin banyak digunakan dalam penelitian analisis sentimen modern, terutama ketika diperlukan pemahaman naratif terhadap opini pengguna pada skala besar. Studi-studi terbaru mengenai integrasi LLM, termasuk penelitian summarization pada domain medis dan aplikasi digital, menegaskan bahwa penggunaan model berskala besar dapat membantu mengungkap tema-tema yang tidak selalu terlihat dari analisis kuantitatif semata. Dengan demikian, tahap summarization memberikan nilai tambah dalam interpretasi hasil dan melengkapi proses analisis yang dilakukan pada tahapan sebelumnya, sekaligus membuka peluang untuk pengembangan sistem analisis sentimen otomatis yang lebih adaptif di masa mendatang.

3. Hasil dan Pembahasan

3.1. Pipeline Penelitian

Alur penelitian yang diimplementasikan dapat dilihat pada Gambar 1. *Pipeline* ini terdiri atas tujuh fase utama, yaitu: data *preparation*, *lexicon-based*, *labeling*, *machine learning training*, *model evaluation*, *rule-based post-processing*, *visualization & analysis*, dan *automated summarization*.

3.2. Data Preparation

Pada tahap *data preparation*, dilakukan eksplorasi awal terhadap *dataset* ulasan pemain yang telah dikumpulkan untuk memberikan gambaran mengenai bentuk, isi, dan karakteristik data yang digunakan dalam penelitian. Eksplorasi ini penting untuk menunjukkan contoh konkret data mentah yang menjadi dasar proses analisis sentimen, serta untuk memastikan bahwa data yang digunakan sesuai dengan tujuan penelitian. Data dibawah didapatkan pada bulan Juni 2025 di *platform Google Play Store* dengan metode *Scraping* menggunakan pustaka *Python* yang terkenal, yaitu *Google_play_Scraper*. Contoh ulasan pemain beserta informasi rating yang diperoleh dari *Google Play Store* disajikan pada Tabel 1 sebagai representasi data sebelum dilakukan pemrosesan lanjutan.

Tabel 1. *Data Review Example*

No	Content	Rating
1	Game ini seperti mewujudkan ...	4
2	jangan jadi game power ...	3
3	thank you for making ...	4
4	my istri	5
5	Why did you fix ...	3
6	agak pelit yah!	1
7	ok	5
8	kalo di bilang bagus ...	3
9	gut	5
10	Can Genshin Impact Make ...	5
39998	Sedikit memburuk	2
39999	Game yang bagus	5
40000	Entahlah mungkin hp gue ...	5

Dataset berupa 40.000 *review* pemain yang diambil dari *Google Play Store* melalui metode *web scraping* menggunakan pustaka *google play scraper*, dengan hasil data disimpan dalam format CSV. Setiap entri ulasan memiliki beberapa atribut utama, yaitu *userName*, *content*, *score*, *reviewCreatedVersion*, dan *at*. Contoh isi data dapat dilihat pada Tabel 1, yang memperlihatkan variasi teks ulasan dengan panjang dan gaya bahasa yang beragam. Setelah dilakukan validasi awal, ditemukan sebanyak 158 baris duplikat dan sekitar 0,4% data dengan nilai kosong, yang kemudian dihapus. *Dataset* akhir yang digunakan dalam penelitian ini terdiri dari 39.842 ulasan valid. Proses pra-pemrosesan selanjutnya mencakup pembersihan teks, normalisasi kata slang (misalnya “gk” menjadi “tidak” dan “bgt” menjadi “banget”), tokenisasi, penghapusan *stopword*, serta *stemming* menggunakan pustaka Sastrawi [12].

Selain meninjau contoh isi ulasan, analisis karakteristik panjang teks juga dilakukan untuk memahami distribusi jumlah karakter pada setiap ulasan pemain. Analisis ini bertujuan untuk menggambarkan kecenderungan panjang ulasan yang ditulis oleh pemain serta mengidentifikasi dominasi ulasan singkat atau panjang dalam *dataset*. Distribusi panjang teks ulasan berdasarkan jumlah karakter disajikan pada Tabel 2.

Tabel 2. *Text Length Distribution*

Karakter	Baris Konten Review
0 -20	17,968
21 - 50	11,137
51 - 100	6,428
101 - 200	3,280
201 - 500	1,187
501 ++	0

Berdasarkan hasil distribusi panjang teks pada Tabel 2, terbentuk kolom baru *clean text* yang merepresentasikan teks ulasan setelah tahap pra-pemrosesan. Distribusi menunjukkan bahwa mayoritas ulasan memiliki panjang 0–20 karakter, dengan jumlah 17.968 baris, yang mengindikasikan bahwa sebagian besar pemain menyampaikan opini dalam bentuk komentar singkat. Jumlah ulasan menurun seiring bertambahnya panjang teks, dan tidak ditemukan ulasan dengan panjang lebih dari 500 karakter. Pola ini mencerminkan karakteristik umum ulasan gim pada

platform *Google Play Store*, yang cenderung bersifat ringkas dan informal.

3.3. Lexicon Based Labeling

Pelabelan dilakukan menggunakan *lexicon domain-specific* yang dikembangkan dari *InSet Lexicon* yang dibuat oleh Koto pada tahun 2017 dengan penambahan kurang lebih 1.500 kata baru terkait konteks permainan seperti *gacha*, *resin*, *artifact*, dan *event*. *Lexicon* disimpan dalam format TSV di GitHub dan dimuat menggunakan *pandas.read_csv()*.

Skor sentimen dihitung dengan formula pada Rumus 1:

$$S = \sum_{i=1}^n (P_i - N_i) \quad (1)$$

Dengan P_i dan N_i masing-masing bobot positif dan negatif dari token ke- i . Yang nantinya kemudian label akhir ditentukan berdasarkan aturan; jika skor > 0 = Positif, jika skor = 0 = Netral, dan jika skor < 0 = negatif. Kemudian diimplementasikan kedalam *pseudocode* berikut;

```

Program Jurnal
Input: token_list T, positive_lexicon PLEX,
negative_lexicon NLEX
Output: sentiment_score S
Initialization:
i ← 1
n ← length(T)
S ← 0
for i = 1 to n do
    token ← T[i]
    if token ∈ PLEX then
        Pi ← PLEX[token]
    else
        Pi ← 0
    end if
    if token ∈ NLEX then
        Ni ← NLEX[token]
    else
        Ni ← 0
    end if
    S ← S + (Pi - Ni)
end for
    
```

Distribusi hasil pelabelan diperlihatkan pada Tabel 3. Tabel tersebut dibagi menjadi dua bagian yaitu distribusi hasil yang menjelaskan presentasi distribusi sentimen dan bagian contoh isi konten didalamnya. Sebagai contohnya, bisa dilihat bahwa presentase distribusi terbanyak adalah sentimen positif dengan hasil 63,33% dari review keseluruhan.

Tabel 3. Label Distribution

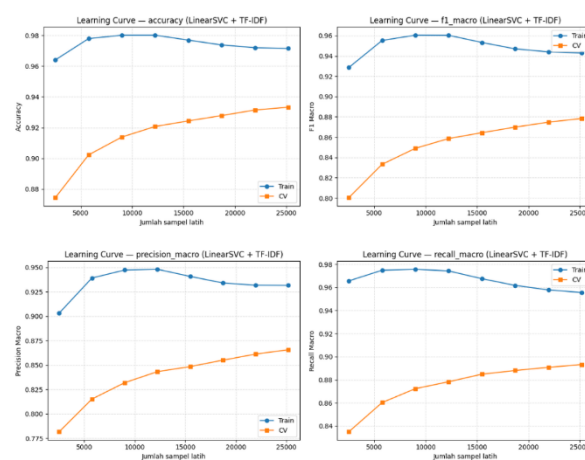
Distribusi hasil pelabelan		
Label	Jumlah	Presentase
Positif	24.928	63,33%
Negatif	11.288	28,68%
Netral	3147	7,99%
Konten hasil pelabelan lexicon		
Clean text	skor	label
game bagus banget	10	Positif
gacha susah dan mahal	-12	Negatif
ok	0	Netral

3.4. Machine Learning Training

Dataset berlabel kemudian dibagi menjadi 80% data latih dan 20% data uji menggunakan *stratified split* untuk menjaga proporsi tiap kelas. Fitur diekstraksi

menggunakan TF-IDF *vectorizer* dengan konfigurasi *ngram_range=(1,3)* dan *max_features=3000*. Model utama adalah LinearSVC dari *scikit-learn* dengan parameter; $C = 1.0$, *class_weight = 'balanced'*, dan *max_iter = 5000*. Proses pelatihan diulang 10 kali pengujian dengan nilai *random seed* berbeda untuk menilai konsistensi performa model.

Untuk mengevaluasi perilaku pembelajaran model serta hubungan antara jumlah data latih dan kinerja klasifikasi, dilakukan analisis menggunakan *learning curve*. Visualisasi ini bertujuan untuk mengamati stabilitas performa model, mendeteksi potensi *overfitting* atau *underfitting*, serta menilai kemampuan generalisasi model terhadap data yang belum pernah dilihat. Hasil evaluasi performa model menggunakan *learning curve* disajikan pada Gambar 2.



Gambar 1. Learning Curve SVM

Pada Gambar 2 diperlihatkan hasil pengujian performa model yang divisualisasikan melalui *learning curve*, yang menggambarkan hubungan antara jumlah sampel latih dan stabilitas performa model pada empat metrik utama, yaitu *accuracy*, *macro F1-score*, *macro precision*, dan *macro recall*. Kurva pelatihan (*train*) menunjukkan performa tinggi yang relatif stabil pada seluruh metrik, sementara kurva *cross-validation* (CV) memperlihatkan peningkatan bertahap seiring bertambahnya jumlah data latih. Pola ini menunjukkan bahwa model semakin mampu melakukan generalisasi dengan baik pada data baru tanpa indikasi *overfitting*. Kinerja model yang konsisten hingga pada jumlah sampel terbesar (± 25.000 data latih) mengonfirmasi bahwa representasi TF-IDF dengan konfigurasi 1–3 grams dan algoritma LinearSVC mampu mengolah variasi teks ulasan pemain secara efektif.

Selain evaluasi visual menggunakan *learning curve*, dilakukan pula pengujian numerik berulang untuk menilai konsistensi performa model terhadap variasi pembagian data latih dan data uji. Pengujian ini dilakukan melalui sepuluh kali pelatihan model dengan nilai *random seed* yang berbeda, sehingga dapat diamati kestabilan metrik evaluasi yang dihasilkan.

Ringkasan hasil pengujian berulang tersebut disajikan pada Tabel 4.

Tabel 4. 10 Loop Test

run	loop	train	train	train	test	test	test
n	point	precision	recall	f1-score	precision	recall	f1-score
1	1.0	0.93	0.95	0.94	0.87	0.89	0.88
1	2.0	0.93	0.95	0.94	0.87	0.89	0.88
1	3.0	0.93	0.95	0.94	0.87	0.89	0.88
1	4.0	0.93	0.95	0.94	0.87	0.89	0.88
1	5.0	0.93	0.95	0.94	0.87	0.89	0.88
1	6.0	0.93	0.95	0.94	0.87	0.89	0.88
1	7.0	0.93	0.95	0.94	0.87	0.89	0.88
1	8.0	0.93	0.95	0.94	0.87	0.89	0.88
2	1.0	0.93	0.95	0.94	0.87	0.9	0.89
2	2.0	0.93	0.95	0.94	0.87	0.9	0.89
2	3.0	0.93	0.95	0.94	0.87	0.9	0.89
2	4.0	0.93	0.95	0.94	0.87	0.9	0.89
.....							
10	7.0	0.93	0.95	0.94	0.87	0.9	0.88
10	8.0	0.93	0.95	0.94	0.87	0.9	0.88

Pada Tabe 4 ditunjukkan detail numerik dari sepuluh kali pengujian, yang memperlihatkan nilai metrik yang relatif stabil pada setiap run dan setiap titik learning curve. Nilai *precision*, *recall*, dan *F1-score* pada data latih konsisten berada pada rentang 0,93–0,95, sementara performa pada data uji stabil di rentang 0,87–0,90 untuk seluruh metrik inti. Konsistensi ini menunjukkan bahwa perubahan nilai *random seed* tidak memberikan variasi signifikan terhadap performa model, menandakan *robust*-nya model terhadap perbedaan komposisi data pada setiap *split* [25]. Selain itu, tren performa yang homogen antar-*run* memperlihatkan bahwa model memiliki stabilitas tinggi dan dapat direproduksi dengan hasil yang konsisten, yang merupakan karakteristik penting dalam penelitian berbasis pembelajaran mesin.

3.5. Model Evaluation

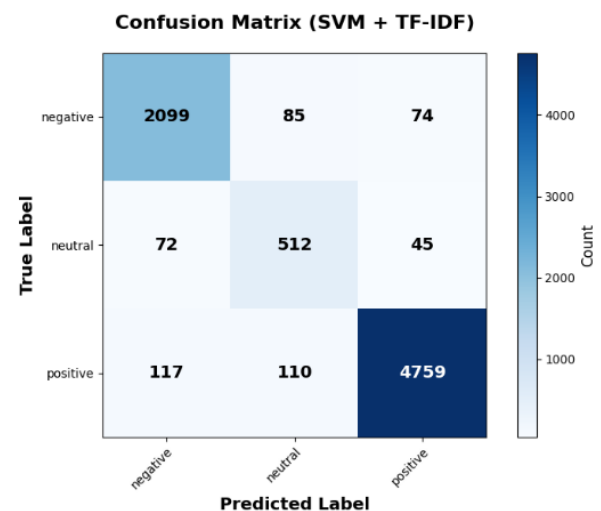
Tahap ini dilakukan untuk menilai kinerja model klasifikasi sentimen yang telah dilatih pada tahap sebelumnya. Evaluasi dilakukan dengan menggunakan beberapa metrik klasifikasi yang umum digunakan dalam analisis sentimen, yaitu *precision*, *recall*, dan *F1-score*, baik untuk setiap kelas sentimen maupun secara agregat. Ringkasan hasil evaluasi performa model disajikan pada Tabel 5.

Tabel 5. Evaluation Summary

	precision	recall	f1-score
negative	0.933509	0.938884	0.936189
neutral	0.814262	0.780604	0.797078
positive	0.967193	0.969715	0.968453
accuracy	0.945764	0.945764	0.945764
macro avg	0.904988	0.896401	0.900573
weighted avg	0.945315	0.945764	0.945508

Hasil evaluasi pada Tabel 5 mencakup nilai *precision*, *recall*, dan *F1-score* untuk setiap kelas sentimen. Model menghasilkan performa terbaik pada kelas positif dengan nilai *F1-score* sebesar 0,968, diikuti oleh kelas negatif dengan *F1-score* 0,936, sedangkan kelas netral memiliki *F1-score* paling rendah, yaitu 0,797. Perbedaan performa ini sejalan dengan karakteristik dataset, di mana ulasan netral cenderung lebih pendek dan tidak mengandung ekspresi emosional yang jelas, sehingga lebih sulit dipetakan secara konsisten oleh model. Nilai *macro-average* sebesar 0,900 menunjukkan bahwa performa model relatif merata di seluruh kelas meskipun distribusi data tidak seimbang. Sementara itu, nilai *weighted-average* dan *accuracy* yang mencapai 0,945 mengindikasikan bahwa model mampu bekerja dengan baik pada distribusi data yang didominasi oleh ulasan positif dan negatif.

Selain evaluasi berbasis metrik numerik, analisis performa model juga dilakukan melalui visualisasi *confusion matrix* untuk memberikan gambaran yang lebih rinci mengenai pola prediksi dan kesalahan klasifikasi antar kelas sentimen. Visualisasi ini memungkinkan identifikasi kelas yang paling sering tertukar serta membantu memahami karakteristik kesalahan yang dihasilkan oleh model. Hasil visualisasi *confusion matrix* ditampilkan pada Gambar 3.



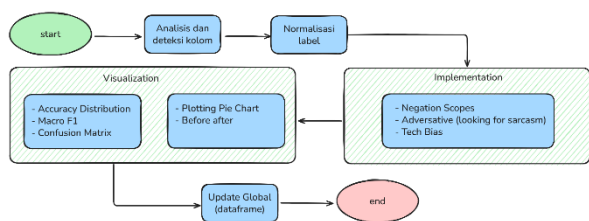
Gambar 2. Confusion Matrix Output

Visualisasi *confusion matrix* pada Gambar 3 memberikan gambaran yang lebih detail mengenai pola prediksi model. Untuk kelas negatif, model berhasil mengklasifikasikan 2.099 ulasan dengan benar, dengan jumlah kesalahan yang relatif kecil pada prediksi netral sebanyak 85 ulasan dan positif sebanyak 74 ulasan. Pola serupa terlihat pada kelas positif, di mana 4.759 ulasan terklasifikasi dengan benar, sementara kesalahan ke kelas netral (110 ulasan) dan negatif (117 ulasan) masih berada dalam batas wajar. Tantangan terbesar terlihat pada kelas netral; dari total 629 ulasan, hanya 512 ulasan yang terprediksi dengan benar, sementara sebagian lainnya salah diklasifikasikan sebagai negatif (72 ulasan) atau positif (45 ulasan). Pola ini mengonfirmasi bahwa kelas netral memiliki

ambiguitas leksikal yang lebih tinggi dibandingkan kelas lainnya. Meskipun demikian, distribusi kesalahan yang muncul tetap konsisten dengan hasil evaluasi numerik sebelumnya dan mendukung kesimpulan bahwa model memiliki kemampuan generalisasi yang baik.

3.6. Rule Based Post Processing

Tahap *rule-based post-processing* diterapkan sebagai langkah lanjutan setelah proses klasifikasi sentimen menggunakan model SVM untuk menyempurnakan hasil prediksi berbasis statistik. Tahap ini bertujuan untuk menangkap konteks linguistik dan pola kebahasaan yang tidak sepenuhnya terakomodasi oleh representasi fitur TF-IDF, khususnya pada ulasan permainan digital yang sering mengandung negasi, frasa kontras, serta istilah teknis domain-spesifik. Alur penerapan *rule patch* dalam penelitian ini divisualisasikan secara ringkas melalui *flowchart* pada Gambar 4.



Gambar 4. Rule Patch Flowchart

Penerapan *rule patch* dilakukan setelah model SVM menghasilkan label prediksi awal, dengan tujuan mengoreksi konteks linguistik yang tidak sepenuhnya tertangkap oleh pendekatan berbasis TF-IDF. Aturan yang diterapkan bersifat domain-spesifik untuk konteks permainan digital dan mencakup tiga mekanisme utama, yaitu: (1) negation handling, misalnya frasa “tidak jelek” atau “nggak buruk” yang secara semantik bermakna positif; (2) contrastive phrase handling, seperti “bagus tapi mahal” yang menunjukkan dominasi sentimen negatif; serta (3) technical bias correction, yaitu pengenalan istilah teknis seperti *crash*, *lag*, atau resin habis sebagai indikator kuat sentimen negatif. Alur mekanisme ini, sebagaimana ditunjukkan pada Gambar 4, dimulai dari analisis struktur kolom dan normalisasi label, dilanjutkan dengan penerapan aturan linguistik, hingga pembaruan *dataframe global*. *Pseudocode rule patch* yang diimplementasikan dalam notebook penelitian memperlihatkan bagaimana sistem mendeteksi negator, istilah teknis, dan pola frasa kontras untuk melakukan *override* label secara terkontrol.

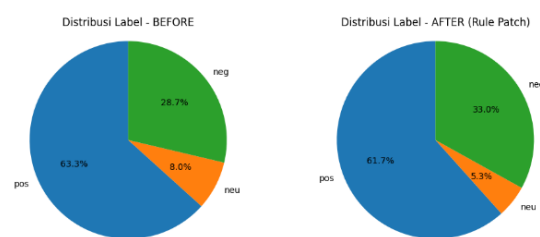
Untuk menilai dampak penerapan *rule-based post-processing* terhadap hasil klasifikasi sentimen, dilakukan perbandingan performa model sebelum dan sesudah *rule patch*. Evaluasi ini bertujuan untuk mengamati perubahan metrik kinerja utama sebagai konsekuensi dari penyesuaian label berbasis aturan linguistik. Ringkasan hasil evaluasi tersebut disajikan pada Tabel 6.

Tabel 6. Evaluation Output After Rule Patch

Metric	After	After (Rule Patch)
Accuracy	0.946	0.873
Macro-F1	0.901	0.790

Evaluasi hasil *rule-patch* dirangkum pada tabel 6 yang memperlihatkan perubahan metrik setelah aturan diterapkan. Secara umum, akurasi dan macro-F1 mengalami penurunan dari 0.946 menjadi 0.873 (accuracy) dan dari 0.901 menjadi 0.790 (macro-F1). Penurunan ini sejalan dengan tujuan *rule-patch*, yaitu bukan untuk meningkatkan performa model secara statistik, melainkan untuk menyesuaikan label agar lebih representatif dengan konteks domain permainan. Banyak ulasan yang sebelumnya diklasifikasikan sebagai positif atau netral oleh SVM dialihkan menjadi negatif karena mengandung keluhan teknis yang tidak terbaca sebagai sentimen negatif oleh fitur TF-IDF. Dengan demikian, *rule-patch* berfungsi sebagai mekanisme koreksi semantik, bukan sebagai optimisasi metrik evaluasi. Selain itu, evaluasi manual terhadap sebagian sampel menunjukkan bahwa aturan linguistik ini mampu menangkap nuansa keluhan yang muncul pada kalimat-kalimat bercampur, misalnya ulasan yang memuji visual tetapi menyoroti performa buruk sehingga menghasilkan label akhir yang lebih akurat secara makna.

Selain evaluasi berbasis metrik numerik, dampak penerapan *rule-based post-processing* juga dianalisis melalui perubahan distribusi label sentimen. Analisis distribusi ini bertujuan untuk melihat bagaimana penerapan aturan linguistik memengaruhi proporsi masing-masing kelas sentimen secara keseluruhan, sehingga dapat memberikan gambaran makro mengenai pergeseran persepsi yang dihasilkan oleh proses *rule patch*. Perbandingan distribusi label sentimen sebelum dan sesudah penerapan *rule patch* divisualisasikan dalam bentuk diagram lingkaran pada Gambar 4.



Gambar 5. Rule Patch Distribution

Perubahan distribusi sentimen tampak jelas pada Gambar 4. Sebelum penerapan *rule patch*, proporsi label positif mendominasi sebesar 63,3%, diikuti oleh negatif sebesar 28,7%, dan netral sebesar 8,0%. Setelah aturan linguistik diterapkan, proporsi sentimen negatif meningkat menjadi 33,0%, sementara sentimen netral menurun menjadi 5,3% dan sentimen positif berkurang menjadi 61,7%. Pergeseran ini menunjukkan bahwa *rule patch* berhasil menangkap keluhan teknis yang sebelumnya tersembunyi dalam kelas positif atau netral. Pola tersebut juga mencerminkan karakteristik

khas ulasan permainan digital, di mana pemain sering memadukan pujian dan keluhan dalam satu kalimat, sehingga memerlukan pendekatan linguistik untuk memahami konteks sentimen secara lebih utuh. Dengan demikian, meskipun nilai akurasi model mengalami penurunan, representasi sentimen setelah penerapan *rule patch* menjadi lebih realistis dan selaras dengan wacana ulasan pemain Genshin Impact yang banyak menyoroti isu performa, *lag*, dan *bug*. Hasil ini menegaskan bahwa *rule patch* berfungsi sebagai tahap penyesuaian domain-specific, bukan sebagai bagian dari pemodelan *supervised* utama, serta memberikan dasar yang lebih kuat bagi analisis kualitatif pada tahap peringkasan menggunakan LLM. Untuk logika *rule patch* bisa dilihat pada pseudocode dibawah;

```

Rule Patch Logic


---


Input: text, base_label, base_score
Output: refined_label

Initialization:
NEGATORS ← {...}
CONTRAST ← {...}
TECH ← {...}
NEG_CUES ← {...}
POS_CUES ← {...}
tokens ← NormalizeAndTokenize(text)
window_neg ← 3
window_contrast_start ← 5
window_contrast_end ← 12
weak_pos_margin ← 2

# 1. Negation Handling
for i = 1 to length(tokens) do
  if tokens[i] ∈ NEGATORS then
    scope ← tokens[i+1 : i+window_neg]
    if any(word ∈ POS_CUES for word in scope) then
      return "neg"
    end if
  end if
end for

# 2. Contrastive Phrase Handling
for i = 1 to length(tokens) do
  if tokens[i] ∈ CONTRAST then
    tail ← tokens[i+window_contrast_start : i+window_contrast_end]
    if any(word ∈ TECH or word ∈ NEG_CUES for word in tail) then
      return "neg"
    end if
  end if
end for

# 3. Technical Bias Adjustment
has_tech ← any(token ∈ TECH for token in tokens)
has_neg_cue ← any(token ∈ NEG_CUES for token in tokens)

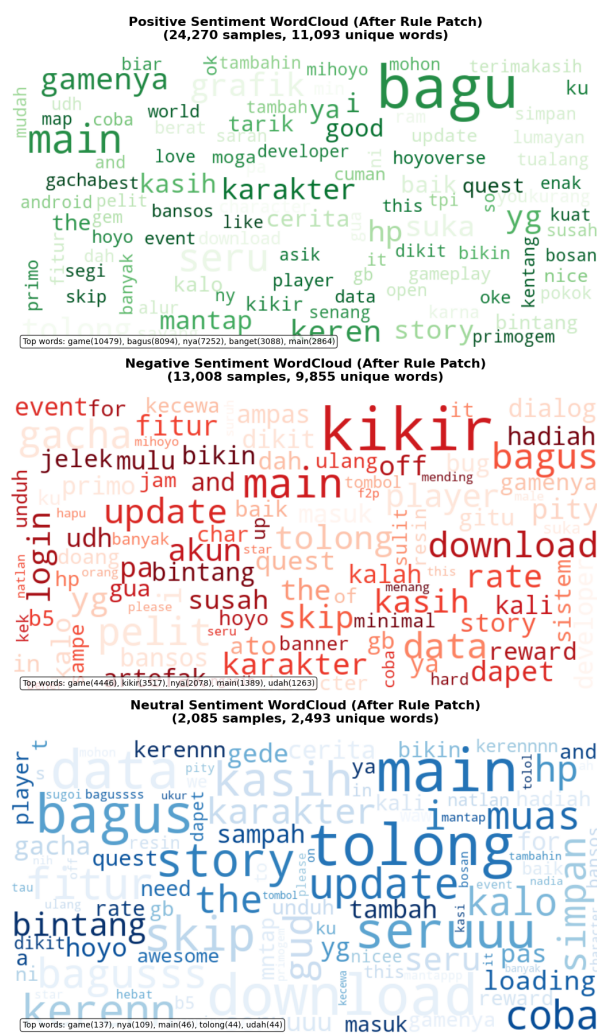
# 3a. Neutral → Negative
if base_label = "neu" AND (has_tech OR has_neg_cue) then
  return "neg"
end if

# 3b. Weak Positive → Negative
if base_label = "pos" AND base_score ≤ weak_pos_margin AND (has_tech OR has_neg_cue) then
  return "neg"
end if
return base_label

```

3.7. Visualization and Analysis

Tahap visualisasi dilakukan untuk mengeksplorasi pola linguistik dan tema dominan yang muncul pada ulasan pemain setelah seluruh proses klasifikasi dan *rule-based post-processing* diterapkan. Visualisasi ini bertujuan untuk memberikan gambaran intuitif mengenai kata-kata yang paling sering digunakan oleh pemain dalam menyampaikan opini mereka pada masing-masing kategori sentimen. Salah satu teknik visualisasi yang digunakan adalah *WordCloud*, yang menampilkan frekuensi kemunculan kata secara visual berdasarkan ukuran dan intensitasnya. Hasil visualisasi *WordCloud* untuk setiap kelas sentimen disajikan pada Gambar 6.



Gambar 6. Sentiment Wordcloud

Visualisasi *WordCloud* pada Gambar 6 memberikan gambaran mengenai kata-kata yang paling sering muncul pada masing-masing kategori sentimen setelah proses *rule patch* diterapkan. Pada kelas positif, kata-kata seperti “bagus”, “main”, “grafik”, “karakter”, dan “cerita” muncul secara dominan, yang mencerminkan bahwa pemain banyak memberikan apresiasi terhadap desain visual, karakter, serta aspek alur cerita permainan. Selain itu, kemunculan kata-kata seperti “mantap”, “keren”, serta variasi kata slang seperti

“banget” dan “gamenya” menunjukkan bentuk ekspresi positif yang konsisten dengan pola bahasa informal pengguna. Jumlah total sampel positif setelah *rule patch* mencapai 24.270 ulasan dengan 11.093 kata unik, yang menunjukkan keragaman ekspresi yang cukup luas meskipun konteks tematiknya relatif stabil.

Pada kelas negatif, WordCloud memperlihatkan dominasi kata-kata seperti “kikir”, “update”, “gacha”, “login”, “ampas”, “bug”, dan “susah”, yang selaras dengan keluhan pemain mengenai isu teknis serta sistem ekonomi permainan. Kata-kata seperti “download”, “akun”, dan “skip” juga muncul dengan frekuensi tinggi, mengindikasikan adanya hambatan teknis maupun mekanik permainan yang secara konsisten menjadi sumber frustrasi pemain. Sementara itu, pada kelas netral, beberapa kata seperti “story”, “tolong”, “main”, dan “update” sering muncul tanpa konteks emosional yang jelas. Pola ini sesuai dengan karakteristik ulasan netral yang umumnya berupa permintaan, informasi, atau komentar deskriptif. Dengan jumlah 2.085 sampel dan 2.493 kata unik, WordCloud netral menunjukkan keragaman kosakata yang jauh lebih kecil dibandingkan dua kategori lainnya, yang memperkuat temuan bahwa kelas netral memiliki ambiguitas yang lebih tinggi dalam pemetaan sentimen.

3.8. Automated Summarization

Selain analisis kuantitatif dan visualisasi berbasis statistik, penelitian ini juga menerapkan *automated summarization* untuk memperoleh pemahaman kualitatif terhadap pola opini pemain. Tahap ini bertujuan untuk merangkum kecenderungan tema dan narasi utama yang muncul dalam ulasan pemain pada setiap kategori sentimen, sehingga hasil analisis menjadi lebih interpretatif. Untuk keperluan tersebut, digunakan *Large Language Model (LLM) IBM Granite*, yang menerima sampel ulasan terpilih sebagai masukan untuk menghasilkan ringkasan naratif. Contoh sampel ulasan yang digunakan sebagai input dalam proses *automated summarization* disajikan pada Tabel 7.

Tabel 7. Example Data for Granite

Label	Teks
negative	puzzle sulit tantang sulit
positive	awesome
Negative	game hp aing nge lag parah andro 13 ngotak anjg
neutral	enakan game nya
positive	senang seru
positive	bagus kasih bintang 4 tapa momiji
negative	oke sih cuman fenti lakik kayak cewek ya
negative	lebihin bansos
positive	nice

Tabel 7 menampilkan contoh sampel ulasan yang dikirim ke IBM Granite sebagai bagian dari proses *automated summarization*. Sampel ini dipilih secara proporsional berdasarkan distribusi sentimen hasil *rule patch*, mencakup ulasan positif, negatif, dan netral. Variasi gaya bahasa terlihat jelas, mulai dari komentar teknis seperti “hp aing nge lag parah”, ungkapan

apresiatif seperti “awesome”, hingga kalimat netral seperti “enakan game nya”. Keberagaman bentuk ekspresi ini penting agar Granite dapat menangkap pola tematik dalam ulasan pemain yang tidak bersifat datar, melainkan mencerminkan penggunaan bahasa sehari-hari, slang, serta kritik teknis khas komunitas gim.

Setelah proses *automated summarization* dilakukan menggunakan IBM Granite terhadap sampel ulasan yang telah dipilih, tahap selanjutnya adalah mengevaluasi hasil ringkasan yang dihasilkan. Evaluasi ini bertujuan untuk melihat kesesuaian antara ringkasan naratif yang dihasilkan oleh model LLM dengan distribusi sentimen serta performa klasifikasi yang diperoleh pada tahap sebelumnya. Ringkasan hasil *automated summarization* beserta metrik pendukung dan distribusi sentimen pada data sampel disajikan pada Tabel 8.

Tabel 8. Granite Summarize

accur acy	macr o p	macr o r	macr o fl	dist ne gative	dist n eutral	dist po sitive
0.93	0.87	0.89	0.88	32.9%	5.4%	61.97%

Sum mary Dari analisis sampel teks ulasan tentang game Genshin Impact, dapat disimpulkan bahwa performa klasifikasi SVM sangat tinggi dengan akurasi 0.937, precision 0.873, recall 0.899, dan F1-score 0.885. Ini menunjukkan bahwa metode klasifikasi SVM berhasil dengan baik mengklasifikasikan ulasan pemain sebagai positif, negatif, atau netral.

Dalam kategori ulasan positif, pemain menyampaikan kepuasan dan kegembiraan dengan game. Ulasan mengenai karakter yang menarik, grafik yang keren, dan dunia yang luas menjadi tema utama. Pemain juga mengungkapkan kepuasan mendapatkan hadiah setelah melakukan quest dan menginginkan lebih banyak hadiah, seperti resin dan item-item lainnya. Ulasan-ulasan ini menunjukkan bahwa aspek seperti desain karakter, visual, dan sistem quest dianggap positif.

Sementara itu, kategori ulasan negatif menciptakan kritikan terhadap beberapa aspek game. Kritikus menyebut game sebagai "lebih bansos" dan "kikir," yang mungkin menunjukkan kekecewaan terhadap sistem gacha atau mekanisme monetisasi dalam game. Selain itu, ada kritik terhadap masalah kemacetan saat dimainkan di perangkat pintar Android, serta masalah grafik yang mencurigakan pada perangkat tertentu.

Kategori netral, sebagai namanya, mengandung ulasan yang tidak terlalu positif maupun negatif. Ulasan-ulasan ini biasanya menunjukkan kekebalan atau tidak ada perbedaan yang signifikan, seperti "oke" atau "alhamdulillah."

Dari analisis ini, dapat disimpulkan bahwa Genshin Impact memiliki persepsi yang positif dari bagi sebagian besar pemain, tetapi ada kritikan terhadap beberapa aspek game. Pemain menghargai desain karakter, grafik, dan sistem quest, tetapi ada kekecewaan terhadap sistem monetisasi dan masalah performa pada perangkat tertentu. Perusahaan perlu memperhatikan kritikan negatif untuk meningkatkan pengalaman pemain secara keseluruhan.

Hasil ringkasan yang ditunjukkan pada Tabel 8 memperlihatkan bahwa performa klasifikasi pada data sampel tetap konsisten dengan hasil evaluasi sebelumnya, dengan nilai akurasi 0.937, precision 0.873, recall 0.899, dan *macro-F1* 0.885. Distribusi sentimen pada sampel menunjukkan proporsi negatif

32,9%, netral 5,4%, dan positif 61,97%, yang selaras dengan tren distribusi pada *dataset* utama setelah penerapan *rule patch*. Berdasarkan sampel tersebut, Granite menghasilkan ringkasan naratif yang merangkum pola dominan pada setiap kategori sentimen, baik berupa apresiasi maupun kritik.

Secara lebih mendalam, ringkasan Granite mengidentifikasi bahwa ulasan positif didominasi oleh apresiasi terhadap karakter, grafik, dunia permainan, dan hadiah in-game. Sebaliknya, ulasan negatif banyak menyoroti isu seperti *lag*, sistem *gacha* yang dianggap “kikir”, masalah performa perangkat, serta ketidakstabilan grafik. Ulasan netral cenderung bersifat informatif atau menyampaikan pengalaman tanpa muatan emosional yang kuat. Secara keseluruhan, proses *automated summarization* ini memperkuat analisis kuantitatif sebelumnya dengan memberikan konteks naratif yang lebih kaya, sehingga membantu menggambarkan persepsi pemain secara lebih komprehensif. Pendekatan ini menunjukkan bahwa kombinasi model klasifikasi berbasis SVM dan Large Language Model seperti Granite efektif dalam menghasilkan insight kualitatif pada skala besar.

3.9. Pembahasan

Bagian ini akan membahas hasil penelitian yang telah dipaparkan pada subbab sebelumnya dengan mengaitkan performa model, karakteristik data, serta implikasi metodologis terhadap analisis sentimen ulasan permainan digital. Pembahasan difokuskan pada efektivitas pendekatan yang diusulkan, keterbatasan yang ditemukan, serta kontribusi penelitian terhadap pemahaman sentimen pemain Genshin Impact berbahasa Indonesia.

Hasil evaluasi menunjukkan bahwa kombinasi *lexicon-based labeling*, representasi fitur TF-IDF, dan algoritma *Support Vector Machine* (SVM) menghasilkan performa klasifikasi yang tinggi dan stabil. Berdasarkan Tabel 5, model mencapai nilai *accuracy* sebesar 0,945, dengan *macro-F1* sebesar 0,900, yang mengindikasikan bahwa model mampu melakukan generalisasi dengan baik pada seluruh kelas sentimen meskipun distribusi data tidak seimbang. Nilai *weighted F1-score* sebesar 0,945 juga menunjukkan bahwa model bekerja efektif pada kelas dominan, yaitu sentimen positif dan negatif. Temuan ini sejalan dengan hasil *learning curve* dan pengujian berulang pada Tabel 4, yang memperlihatkan performa data uji yang stabil pada rentang 0,87–0,90 untuk seluruh metrik utama di sepuluh kali pengujian dengan random seed berbeda.

Analisis per kelas menunjukkan bahwa performa terbaik dicapai pada sentimen positif, dengan nilai *precision* 0,968, *recall* 0,969, dan *F1-score* 0,968, diikuti oleh sentimen negatif dengan *F1-score* 0,936. Sebaliknya, sentimen netral memiliki performa paling rendah dengan *F1-score* 0,797, sebagaimana terlihat pada Tabel 5. Pola ini diperkuat oleh hasil *confusion matrix*, di mana dari 629 ulasan netral, hanya 512

ulasan yang terklasifikasi dengan benar, sementara sisanya salah dipetakan ke kelas negatif (72 ulasan) dan positif (45 ulasan). Temuan ini mengonfirmasi bahwa sentimen netral memiliki tingkat ambiguitas linguistik yang lebih tinggi, terutama karena ulasan netral cenderung singkat dan minim ekspresi emosional eksplisit.

Penerapan *rule-based post-processing* memberikan sudut pandang penting terhadap keterbatasan pendekatan statistik murni. Setelah *rule patch* diterapkan, nilai *accuracy* menurun dari 0,946 menjadi 0,873, dan *macro-F1* menurun dari 0,901 menjadi 0,790, sebagaimana ditunjukkan pada Tabel 6. Penurunan ini tidak menunjukkan degradasi kualitas analisis, melainkan mencerminkan perubahan tujuan evaluasi dari optimasi metrik ke penyesuaian semantik berbasis domain. Distribusi sentimen pasca-*rule patch* memperlihatkan peningkatan proporsi sentimen negatif dari 28,7% menjadi 33,0%, serta penurunan sentimen netral dari 8,0% menjadi 5,3%, yang menunjukkan bahwa sejumlah keluhan teknis berhasil diidentifikasi secara lebih akurat sebagai sentimen negatif.

Hasil visualisasi *WordCloud* mendukung temuan kuantitatif dan *rule-based* tersebut. Pada kelas positif, dominasi kata seperti “bagus”, “grafik”, “karakter”, dan “cerita” mencerminkan apresiasi pemain terhadap aspek visual dan desain permainan. Sebaliknya, kelas negatif didominasi kata-kata seperti “kikir”, “lag”, “bug”, dan “gacha”, yang selaras dengan peningkatan proporsi sentimen negatif setelah *rule patch*. Kelas netral, dengan hanya 2.085 sampel dan 2.493 kata unik, menunjukkan keragaman kosakata yang jauh lebih kecil dibandingkan kelas positif (24.270 sampel; 11.093 kata unik), memperkuat temuan bahwa kelas ini memiliki ambiguitas tinggi dan keterbatasan konteks emosional.

Lebih lanjut, penerapan *automated summarization* menggunakan *Large Language Model* IBM Granite memberikan penguatan kualitatif terhadap hasil kuantitatif sebelumnya. Berdasarkan Tabel 8, performa klasifikasi pada data sampel tetap konsisten dengan *accuracy* 0,937, *precision* 0,873, *recall* 0,899, dan *macro-F1* 0,885, dengan distribusi sentimen 61,97% positif, 32,9% negatif, dan 5,4% netral. Ringkasan naratif yang dihasilkan Granite secara eksplisit menyoroti apresiasi terhadap desain visual dan karakter pada ulasan positif, serta kritik terhadap performa teknis, sistem *gacha*, dan stabilitas perangkat pada ulasan negatif. Dengan demikian, *automated summarization* berperan sebagai lapisan interpretatif yang memperkaya analisis sentimen berbasis statistik.

Secara keseluruhan, pembahasan ini menunjukkan bahwa pendekatan *hybrid* yang menggabungkan *machine learning* berbasis SVM, penyesuaian linguistik melalui *rule-based post-processing*, dan interpretasi kualitatif menggunakan *large language model* mampu menghasilkan analisis sentimen yang tidak hanya kuat secara numerik, tetapi juga relevan

secara semantik. Pendekatan ini sangat sesuai untuk konteks ulasan permainan digital berbahasa Indonesia yang kaya akan bahasa informal, *slang*, dan ekspresi campuran antara pujian dan keluhan.

4. Kesimpulan

Penelitian ini berhasil membangun sebuah *pipeline* analisis sentimen untuk memproses 40.000 ulasan pemain Genshin Impact yang dikumpulkan dari *Google Play Store* dengan mengombinasikan pendekatan *lexicon-based labeling*, representasi fitur TF-IDF, dan algoritma *Support Vector Machine* (SVM). Hasil eksperimen menunjukkan bahwa pendekatan ini menghasilkan performa klasifikasi yang tinggi dan stabil, dengan nilai akurasi sebesar 0,945 dan macro-F1 sebesar 0,900. Temuan ini mengindikasikan bahwa SVM yang dilatih menggunakan TF-IDF efektif dalam menangani variasi bahasa informal, penggunaan *slang*, dan struktur kalimat khas ulasan gim berbahasa Indonesia.

Penerapan *rule-based post-processing* mengungkap keterbatasan pendekatan statistik murni dalam menangani konteks linguistik tertentu, seperti negasi, frasa kontras, dan istilah teknis domain permainan. Meskipun penerapan *rule patch* menurunkan nilai akurasi dan macro-F1 menjadi 0,873 dan 0,790, penyesuaian berbasis aturan ini menghasilkan label sentimen yang lebih representatif secara semantik. Pergeseran distribusi sentimen pasca-*rule patch*, khususnya peningkatan proporsi sentimen negatif yang berkaitan dengan isu *lag*, *bug*, *crash*, dan sistem *gacha*, menunjukkan bahwa pendekatan linguistik mampu mengoreksi ambiguitas yang tidak sepenuhnya tertangkap oleh model statistik.

Tahap *automated summarization* menggunakan *Large Language Model* IBM Granite memberikan lapisan interpretatif yang memperkaya analisis kuantitatif. Ringkasan tematik yang dihasilkan konsisten dengan hasil klasifikasi dan visualisasi sebelumnya, mencakup apresiasi terhadap desain visual dan karakter, serta kritik terhadap performa teknis dan mekanisme permainan. Secara keseluruhan, penelitian ini menegaskan bahwa integrasi *machine learning* klasik, penyesuaian linguistik berbasis aturan, dan *large language model* merupakan pendekatan yang efektif dan komprehensif untuk memahami opini pengguna pada domain permainan digital. Untuk penelitian selanjutnya, disarankan perluasan data lintas platform, pengembangan *domain-specific lexicon* yang lebih adaptif, serta eksplorasi model berbasis *transformer* guna meningkatkan ketepatan analisis sentimen.

Daftar Rujukan

- [1] A. Siregar, K. Alfarizi Siregar, B. Cahyadi, L. Samosir, and A. Azhard, "Sentiment Analysis of Reviews from Google Play: Azur Lane, Genshin Impact, Arknights," 2025, doi: 10.64803/cessmuds.v1i1.33.
- [2] M. Y. Febrianta *et al.*, "Analisis Ulasan Indie Video Game Lokal pada Steam Menggunakan Analisis Sentimen dan

- Pemodelan Topik Berbasis Latent Dirichlet Allocation," 2021.
- [3] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, Feb. 2021, doi: 10.34148/teknika.v10i1.311.
- [4] M. Fernanda and N. Fathoni, "Perbandingan Performa Labeling Lexicon InSet dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM," *Jurnal Informatika dan Sains Teknologi*, vol. 1, no. 3, pp. 62–76, 2024, doi: 10.62951/modem.v1i3.112.
- [5] Eko Arip Winanto, S. M. Z. Ali Difyah, Pareza Alam Jusia, and Sharipuddin, "Analisis Sentimen Terhadap Tagar Kabur Aja Dulu Di Twitter Menggunakan Metode Lexicon-Based," *Jurnal PROCESSOR*, vol. 20, no. 2, Oct. 2025, doi: 10.33998/processor.2025.20.2.2542.
- [6] K. Khotimah, M. Martanto, A. R. Dikananda, and A. Rifa'i, "ANALISIS SENTIMEN ULASAN APLIKASI PINTU DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5789.
- [7] R. Aprinastya, M. Jazman, S. Syaifullah, M. Rahmawita, S. Siregar, and E. Saputra, "Comparative Analysis of Naïve Bayes Classifier and Support Vector Machine for Multilingual Sentiment Analysis: Insights from Genshin Impact User Reviews," *JUSIFO (Jurnal Sistem Informasi)*, vol. 10, no. 2, pp. 117–126, Dec. 2024, doi: 10.19109/jusifo.v10i2.24876.
- [8] V. A. Fitria and L. Widayanti, "Enhancing Accuracy in Stock Price Prediction: The Power of Optimization Algorithms," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 23, no. 2, pp. 405–418, Mar. 2024, doi: 10.30812/matrik.v23i2.3785.
- [9] M. Cao *et al.*, "Using large language models to automate summarization of CT simulation orders in radiation oncology," *J Appl Clin Med Phys*, vol. 26, no. 11, p. e70310, Nov. 2025, doi: 10.1002/acm2.70310.
- [10] Y. Tolla, "Deteksi Stres dan Depresi Unggahan Media Sosial dengan Machine Learning".
- [11] N. Dwi Husna Sadikin, S. Susanti, and A. Reswara Sanjaya, "Analisis Sentimen Publik Terhadap Kampanye Pengurangan Sampah Plastik Menggunakan Algoritma Naïve Bayes," *Agustus*, vol. 15, no. 2, pp. 202–212, 2025.
- [12] M. Ashari, D. Arbian Sulisty, and F. Almu'ini Ahda, "STEMMING IN MADURESE LANGUAGE USING NAZIEF AND ADRIANI ALGORITHM," *J. Tek. Inform. (JUTIF)*, pp. 695–702, 2024.
- [13] A. M. Putra, Candra Saputra, Rahmaddeni, Safril Irsandi, and Vawana Muzaki, "Analisis Sentimen Masyarakat Terhadap Kasus Gas LPG 3 Kg Pada Youtube Kompas Menggunakan Metode Support Vector Machine," *Explore*, vol. 15, no. 2, pp. 163–171, Jul. 2025, doi: 10.35200/ex.v15i2.159.
- [14] B. Gunawan, H. Sasty, P. #2, E. Esyudha, and P. #3, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," vol. 4, no. 2, pp. 17–29, 2018, [Online]. Available: www.femaledaily.com
- [15] E. Sari, L. Afuan, I. Permadi, E. Maryanto, and S. P. Rahayu, "CORRELATION ANALYSIS OF SENTIMENT OF 2024 ELECTION RESULTS AND STOCK MOVEMENTS OF POLITICAL ACTORS IN INDONESIA," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 4, pp. 1213–1227, Aug. 2024, doi: 10.52436/1.jutif.2024.5.4.2701.
- [16] A. S. Rizkia, W. Wufron, and F. F. Roji, "Analisis Sentimen Coretax: Perbandingan Pelabelan Data Manual, Transformers-Based, dan Lexicon-Based pada Performa IndoBERT: Sentiment Analysis of Coretax: A Comparison of Manual, Transformers-Based, and Lexicon-Based Data Labeling on IndoBERT Performance," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 3, pp. 1037–1048, Jul. 2025, doi: 10.57152/malcom.v5i3.2151.

- [17] D. D. Sajmira, K. Umam, and M. R. Handayani, "Enhancing Review Processing in the Video Game Adaptation Domain through VADER and Rating-Based Labeling using SVM," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 14, no. 3, pp. 407–414, Jul. 2025, doi: 10.32736/sisfokom.v14i3.2409.
- [18] O. I. Gifari, M. Adha, I. Rifky Hendrawan, F. Freddy, and S. Durrand, "Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine," *JIFOTECH (JOURNAL OF INFORMATION TECHNOLOGY)*, vol. 2, no. 1, 2022.
- [19] Yusril, W. Fuadi, and Y. Afrillia, "ANALISIS SENTIMEN REVIEW APLIKASI STOCKBIT DI GOOGLE PLAY STORE DAN X(TWITTER) MENGGUNAKAN SUPPORT VECTOR MACHINE," *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 10, no. 2, pp. 1050–1062, Jul. 2025, doi: 10.36341/rabit.v10i2.6446.
- [20] R. Ananta Pratama Ilmu Komputer, "ANALISIS SENTIMEN KONSUMEN DENGAN TEKNIK TEXT MINING," 2024.
- [21] A. Diah Pramesti, K. Umam, and M. R. Handayani, "Identification of Buzzers in Skincare Reviews Using a Lexicon-Based Sentiment Analysis Method," 2025.
- [22] D. Sulisty, F. Ahda, and V. A. Fitria, "Epistomologi dalam Natural Language Processing," *Jurnal Inovasi Teknologi dan Edukasi Teknik*, vol. 1, no. 9, pp. 652–664, Sep. 2021, doi: 10.17977/um068v1i92021p652-664.
- [23] D. A. Sulisty, A. P. Wibawa, D. D. Prasetya, and F. A. Ahda, "An enhanced pivot-based neural machine translation for low-resource languages.," *International Journal of Advances in Intelligent Informatics*, 2025, Vol 11, Issue 2, p258, 2025.
- [24] D. A. Sulisty, D. D. Prasetya, F. A. Ahda, and A. P. Wibawa, "Pivoted Low Resource Multilingual Translation with NER Optimization," *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, vol. 24, no. 5, May 2025, doi: 10.1145/3727876.
- [25] A. R. Habibi, V. A. Fitria, and L. Hakim, "Optimasi Learning Rate Neural Network Backpropagation Dengan Search Direction Conjugate Gradient Pada Electrocardiogram," *NUMERICAL: Jurnal Matematika dan Pendidikan Matematika*, pp. 131–137, Jan. 2020, doi: 10.25217/numerical.v3i2.603.
- [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>