

# Analisis Sentimen dan Topik Ulasan Aplikasi Komunitas Digital Gamer dengan SVM-LDA dan CRISP-DM

## *Sentiment and Topic Analysis of Digital Community Application Gamer Reviews using SVM-LDA and CRISP-DM*

<sup>1</sup>Muhammad Mayda Ary Pratama, <sup>2</sup>Dedy Kurniawan\*, <sup>3</sup>Ahmad Rifai, <sup>4</sup>Ken Ditha Tania

<sup>1,2,3,4</sup>Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya  
<sup>1,2,3,4</sup>Jl. Raya Palembang-Prabumulih KM. 32 Indralaya, Ogan Ilir, Sumatera Selatan, 30662, Indonesia

\*e-mail: [dedykurniawan@unsri.ac.id](mailto:dedykurniawan@unsri.ac.id)

(received: 16 September 2025, revised: 9 October 2025, accepted: 10 October 2025)

### Abstrak

Perilaku tidak sabar dari para gamer kerap tercermin melalui ulasan digital yang tajam, terutama dalam penggunaan aplikasi komunitas seperti Discord. Penelitian ini mengeksplorasi ekspresi ketidaksabaran tersebut melalui pendekatan analisis sentimen dan topik. Dengan mengadopsi kerangka kerja CRISP-DM, sebanyak 10.000 ulasan berbahasa Indonesia dari Google Play Store dikumpulkan dan dianalisis. Proses analisis dimulai dengan pelabelan sentimen menggunakan IndoBERT, klasifikasi polaritas dengan algoritma Support Vector Machine (SVM), dan eksplorasi topik melalui metode Latent Dirichlet Allocation (LDA). Hasil menunjukkan bahwa 57,4% ulasan bersentimen positif, terutama terkait kualitas komunikasi suara dan interaksi komunitas. Sebaliknya, 42,6% ulasan negatif umumnya menyampaikan frustrasi terhadap kendala login dan proses verifikasi. Model SVM yang dioptimasi dengan Bayesian Optimization mencatat akurasi sebesar 90,46%. Studi ini menegaskan bahwa Discord tidak hanya menjadi sarana komunikasi, tetapi juga representasi dari ekspektasi tinggi pengguna terhadap kecepatan dan stabilitas sistem. Kontribusi utama penelitian ini adalah integrasi metode SVM-LDA dalam kerangka CRISP-DM untuk memahami perilaku digital gamer Indonesia. Implikasi praktis dari temuan ini dapat menjadi masukan strategis bagi pengembang dalam meningkatkan keandalan proses autentikasi dan fitur komunitas sesuai karakteristik pengguna.

**Kata kunci:** CRISP-DM, *latent dirichlet allocation*, analisis Sentimen, gamer, *support vector machine*

### Abstract

*Impatient behavior among gamers is often reflected in sharp and emotionally charged digital reviews, particularly in the use of community applications such as Discord. This study explores expressions of impatience through sentiment and topic analysis. By adopting the CRISP-DM framework, a total of 10,000 Indonesian-language reviews collected from the Google Play Store were analyzed. The analytical process begins with sentiment labeling using IndoBERT, followed by polarity classification using the Support Vector Machine (SVM) algorithm, and topic exploration through the Latent Dirichlet Allocation (LDA) method. The results indicate that 57.4% of the reviews express positive sentiment, primarily related to voice communication quality and community interaction features. In contrast, 42.6% of the negative reviews commonly convey frustration regarding login issues and verification processes. The SVM model optimized using Bayesian Optimization achieved an accuracy of 90.46%. This study highlights that Discord serves not only as a communication platform but also as a reflection of users' high expectations for system speed and stability. The main contribution of this research lies in the integration of SVM-LDA methods within the CRISP-DM framework to better understand the digital behavior of Indonesian gamers. The practical implications of these findings provide strategic insights for developers to improve authentication reliability and community features in alignment with user characteristics.*

**Keywords:** CRISP-DM, *latent dirichlet allocation*, sentiment analysis, gamer, *support vector machine*

## 1 Pendahuluan

Industri permainan digital mengalami perkembangan pesat seiring munculnya berbagai genre dan sistem yang dirancang untuk mempertahankan keterlibatan pemain [1]. Permainan daring yang menghubungkan pemain dalam ruang virtual mendorong terbentuknya budaya komunikasi cepat dan respons instan di kalangan gamer [2]. Dalam situasi kompetitif, pemain cenderung mengambil keputusan tergesa-gesa, yang dapat memicu perilaku tidak sabar [3]. Sistem pencapaian dan penghargaan dalam permainan turut memperkuat ekspektasi terhadap hasil instan [4].

Paparan media seperti ulasan video dan teks secara cepat membentuk persepsi pemain [5]. Selain itu, desain permainan yang berorientasi pada pencapaian memperkuat kecenderungan pemain untuk fokus pada hasil instan, bukan pada proses bermain [6]. Fasilitas komunikasi dalam permainan mempercepat koordinasi dan reaksi terhadap kondisi permainan [7], sementara ruang diskusi daring memungkinkan pertukaran pengalaman secara real-time yang memperkuat kebiasaan ingin hasil segera [8]. Dengan pemanfaatan teknologi analisis data, pengembang dapat lebih memahami perilaku pengguna dalam konteks permainan maupun aplikasi pendukung [9].

Salah satu aplikasi komunitas digital yang populer di kalangan gamer adalah Discord. Aplikasi ini bukan hanya digunakan sebagai alat bantu dalam bermain, tetapi juga sebagai ruang interaksi sosial yang aktif. Discord berbeda dari permainan itu sendiri, karena ulasan terhadap aplikasi ini mencerminkan pengalaman teknis dan sosial pengguna di luar gameplay. Ulasan tersebut sering kali bersifat emosional, dan dapat mencerminkan ekspresi ketidaksabaran pengguna terhadap fitur atau kendala teknis [10].

Penelitian ini berfokus pada eksplorasi sentimen dan topik dalam ulasan aplikasi komunitas digital Discord untuk mengidentifikasi bagaimana ketidaksabaran gamer tercermin dalam persepsi mereka. Sebagian besar studi sebelumnya hanya menganalisis kepuasan pengguna secara umum, tanpa mengaitkan sentimen dengan karakteristik perilaku spesifik seperti ketidaksabaran, terutama dalam konteks komunitas digital gamer di Indonesia. Selain itu, pendekatan yang digunakan pada riset terdahulu umumnya masih terpisah antara klasifikasi sentimen dan pemodelan topik.

Penelitian ini mengusulkan integrasi metode Support Vector Machine (SVM) dan Latent Dirichlet Allocation (LDA) dalam kerangka kerja CRISP-DM untuk menganalisis ulasan pengguna. Tujuan dari penelitian ini adalah untuk: (1) mengidentifikasi pola ketidaksabaran dalam ulasan pengguna Discord, (2) memetakan distribusi sentimen dan topik secara simultan, dan (3) memberikan rekomendasi yang dapat dimanfaatkan oleh pengembang aplikasi komunitas digital untuk meningkatkan pengalaman pengguna.

## 2 Tinjauan Literatur

Penelitian mengenai analisis sentimen telah berkembang pesat dalam lima tahun terakhir. Analisis ini dipandang sebagai pendekatan penting untuk memahami persepsi publik melalui teks digital yang bersumber dari media sosial maupun ulasan aplikasi. Metode ini telah terbukti mampu memetakan opini masyarakat secara real-time [11]. Proses analisis sentimen memerlukan tahapan sistematis mulai dari pembersihan data, tokenisasi, normalisasi, hingga ekstraksi fitur agar model klasifikasi dapat bekerja secara optimal [12]. Kualitas preprocessing menjadi faktor krusial karena data ulasan digital sering kali mengandung noise, kata tidak baku, dan simbol yang berpotensi menurunkan akurasi model.

Support Vector Machine (SVM) merupakan salah satu algoritma yang paling banyak diterapkan dalam analisis sentimen karena terbukti menghasilkan akurasi tinggi. Pada penelitian aplikasi kependudukan digital, SVM mampu mencapai akurasi lebih dari 80% dalam mengklasifikasikan ulasan positif, negatif, dan netral [13]. Hal ini sejalan dengan temuan lain yang menegaskan bahwa keunggulan SVM terletak pada kemampuannya mencari hyperplane optimal yang memaksimalkan margin antar kelas, sehingga tetap efektif meskipun digunakan pada data berdimensi tinggi [14]. Fleksibilitas ini membuat SVM tidak hanya dipakai untuk klasifikasi berbasis sentimen, tetapi juga dapat dimanfaatkan untuk analisis kualitas perangkat lunak.

Selain klasifikasi, pendekatan *topic modeling* berbasis Latent Dirichlet Allocation (LDA) juga banyak digunakan dalam penelitian ulasan pengguna. Sebuah studi memperkenalkan metode *LDA Filter* yang dapat meningkatkan efisiensi proses klasifikasi dengan mengurangi dimensi fitur tanpa mengorbankan akurasi [15]. Aplikasi LDA pada ulasan perbankan di Indonesia mengungkapkan bahwa

ulasan negatif didominasi oleh kendala login, error sistem, dan kegagalan top-up, sedangkan ulasan positif lebih banyak menyoroti kemudahan transaksi dan keamanan aplikasi [16]. Dengan demikian, LDA terbukti relevan untuk memetakan ulasan pengguna ke dalam tema-tema strategis yang bermanfaat bagi pengembangan aplikasi.

Kerangka kerja CRISP-DM menjadi metodologi yang banyak dipakai dalam proyek analisis data karena bersifat terstruktur dan fleksibel. Sebuah penelitian menerapkan enam fase CRISP-DM untuk mengidentifikasi faktor keterlibatan karyawan yang berdampak terhadap kinerja bisnis, dan hasilnya menunjukkan efektivitas yang tinggi [17]. Studi lain juga menegaskan bahwa setiap fase CRISP-DM memiliki kontribusi penting, khususnya pada tahap *Business Understanding* dan *Data Understanding*, karena kegagalan pada dua fase awal sering kali membuat hasil analisis tidak relevan atau sulit diimplementasikan [18]. Oleh karena itu, CRISP-DM tidak hanya dipandang sebagai kerangka teknis, tetapi juga sebagai panduan agar analisis data dapat menghasilkan nilai nyata bagi organisasi.

Ulasan pengguna merupakan sumber data yang semakin diakui dalam penelitian. Ulasan dianggap sebagai ekspresi digital yang mencerminkan kepuasan maupun ketidakpuasan pengguna secara langsung [19]. Berbeda dengan survei tradisional yang bersifat terstruktur, ulasan ditulis secara spontan sehingga memberikan informasi yang lebih alami. Analisis terhadap ulasan memungkinkan pengembang untuk mengidentifikasi pola berulang dari masalah teknis maupun apresiasi pengguna. Dengan bantuan algoritma klasifikasi, ulasan dapat dipetakan ke dalam kategori tertentu, sehingga lebih mudah dimanfaatkan dalam pengambilan keputusan [20].

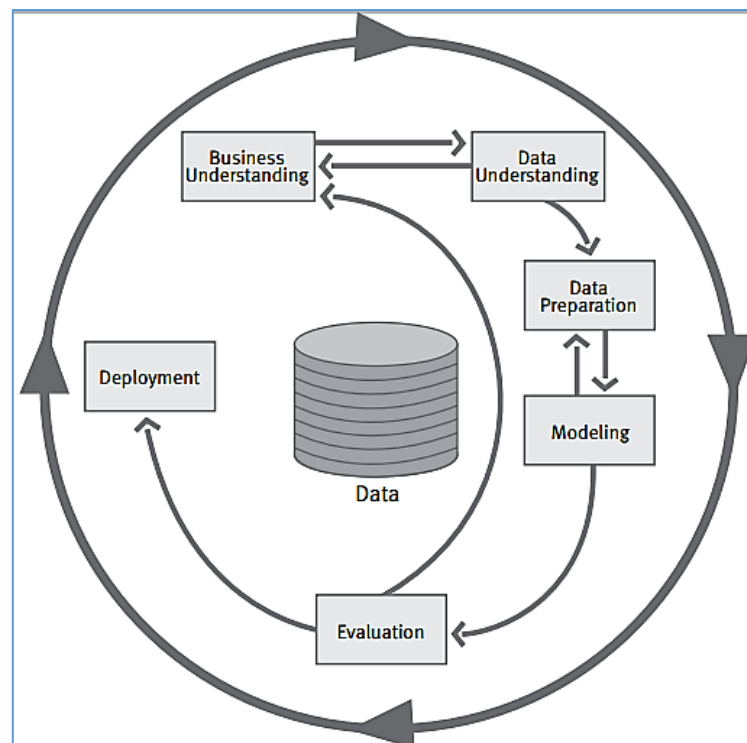
Dalam konteks aplikasi komunitas digital, beberapa penelitian menunjukkan bahwa platform ini memiliki peran yang signifikan. Aplikasi komunitas digital berfungsi untuk meningkatkan interaksi masyarakat, memperkuat literasi digital, serta mendorong kolaborasi berbasis teknologi [21]. Selain itu, aplikasi semacam ini juga memiliki fungsi manajerial yang mendukung koordinasi organisasi secara lebih efektif [22]. Discord, misalnya, berkembang menjadi aplikasi komunikasi multifungsi yang digunakan tidak hanya oleh komunitas gamer, tetapi juga oleh organisasi, institusi pendidikan, dan tim kerja lintas wilayah [23]. Hal ini menunjukkan potensi besar aplikasi komunitas digital dalam mengakomodasi berbagai kebutuhan komunikasi.

Selain aspek teknologi, karakteristik gamer juga menjadi faktor penting dalam memahami ulasan aplikasi komunitas digital. Sebuah penelitian menemukan bahwa gamer memiliki kecenderungan bersikap kompulsif dan impulsif, yang erat kaitannya dengan perilaku tidak sabar [24]. Kondisi ini menjelaskan mengapa ulasan gamer sering kali bersifat ekstrem, baik sangat positif maupun sangat negatif, tergantung pengalaman mereka. Pemahaman mengenai karakteristik gamer ini penting untuk mengaitkan hasil analisis sentimen dan pemodelan topik dengan kebutuhan nyata pengguna.

Berdasarkan tinjauan literatur tersebut, dapat disimpulkan bahwa penelitian terdahulu telah mengkaji analisis sentimen menggunakan SVM maupun pemodelan topik dengan LDA secara terpisah. Namun, penelitian yang menggabungkan keduanya dalam konteks aplikasi komunitas digital masih terbatas. Oleh karena itu, penelitian ini mengisi kesenjangan tersebut dengan mengintegrasikan SVM dan LDA untuk menganalisis ulasan pengguna Discord, sehingga dapat memberikan pemahaman lebih komprehensif mengenai persepsi dan perilaku gamer.

### 3 Metode Penelitian

Metode penelitian ini menjelaskan langkah-langkah yang digunakan untuk menganalisis ulasan pengguna aplikasi Discord dari Google Play Store, mencakup proses sistematis untuk memperoleh hasil yang akurat dan relevan. Penelitian ini menggunakan pendekatan **CRISP-DM (Cross Industry Standard Process for Data Mining)** sebagai kerangka kerja utama, yang terdiri atas enam tahapan: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment* [25]. Alur lengkap tahapan CRISP-DM dalam penelitian ini ditampilkan pada **Gambar 1**.



Gambar 1 Kerangka kerja CRISP-DM

### 3.1 Tahapan Penelitian

Tahapan penelitian dirancang mengikuti alur CRISP-DM secara terstruktur. Pada tahap *Business Understanding*, peneliti mengidentifikasi permasalahan utama, yaitu bagaimana perilaku ketidaksabaran gamer tercermin dalam ulasan Discord, serta menetapkan tujuan analisis berupa pemetaan sentimen dan topik. Tahap *Data Understanding* dilakukan dengan mengumpulkan dan mengeksplorasi data ulasan, termasuk distribusi rating dan panjang teks ulasan.

Tahap *Data Preparation* melibatkan pembersihan dan normalisasi teks, tokenisasi, penghapusan stopword, serta perbaikan kata tidak baku menggunakan kamus slang bahasa Indonesia. *Labeling* awal sentimen dilakukan dengan model IndoBERT untuk mempersiapkan dataset terstruktur. Selanjutnya pada tahap *Modeling*, data dianalisis menggunakan algoritma **Support Vector Machine (SVM)** untuk klasifikasi polaritas sentimen, dan **Latent Dirichlet Allocation (LDA)** untuk pemodelan topik.

Hasil pemodelan kemudian diuji pada tahap *Evaluation* dengan menggunakan metrik akurasi, precision, recall, dan F1-score. Jika hasil evaluasi tidak memadai, dilakukan iterasi ulang dengan menyesuaikan parameter atau preprocessing. Pada tahap akhir, *Deployment*, hasil analisis ditransformasikan menjadi informasi praktis berupa rekomendasi pengembangan fitur aplikasi, serta divisualisasikan dalam bentuk grafik distribusi sentimen, wordcloud, dan tabel topik dominan.

### 3.2 Metode Pengumpulan Data

Data yang digunakan adalah ulasan pengguna aplikasi Discord yang diperoleh langsung dari Google Play Store melalui teknik **web scraping** dengan bantuan Python pada Google Colaboratory. Informasi yang dikumpulkan mencakup teks ulasan, rating bintang, dan tanggal publikasi. Untuk menjaga kualitas penelitian, data dibatasi pada ulasan berbahasa Indonesia dengan rentang waktu tertentu, sehingga diperoleh dataset representatif sesuai konteks penelitian.

### 3.3 Metode Pengelolaan Data

Data ulasan yang terkumpul selanjutnya dikelola melalui tahapan preprocessing untuk menghilangkan noise, seperti emoji, angka, tanda baca, atau kata tidak baku. Setelah itu dilakukan normalisasi dan tokenisasi agar data lebih seragam. Proses *labeling* sentimen menggunakan model **IndoBERT** menghasilkan dataset dengan kategori positif dan negatif. Dataset kemudian dibagi menjadi data latih dan data uji untuk memastikan model dapat digeneralisasi dengan baik.

Model **SVM** dilatih menggunakan data berlabel untuk memprediksi polaritas sentimen secara otomatis, sedangkan **LDA** digunakan untuk mengelompokkan ulasan ke dalam topik-topik utama.

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

Hasil prediksi divisualisasikan dalam bentuk grafik distribusi sentimen, wordcloud kata kunci dominan, serta tabel topik untuk memberikan gambaran yang jelas mengenai persepsi pengguna. Seluruh tahapan pengelolaan data dilakukan secara berulang sampai diperoleh hasil yang stabil dan relevan dengan tujuan penelitian.

## 4 Hasil dan Pembahasan

Pada bagian ini disajikan hasil penelitian serta analisis yang dilakukan berdasarkan implementasi kerangka kerja CRISP-DM dalam mengolah data ulasan pengguna aplikasi Discord. Hasil penelitian mencakup setiap tahapan CRISP-DM, mulai dari pemahaman bisnis hingga evaluasi model. Analisis difokuskan pada bagaimana algoritma Support Vector Machine (SVM) digunakan untuk mengklasifikasikan polaritas sentimen, serta bagaimana metode Latent Dirichlet Allocation (LDA) mengungkap topik dominan dalam ulasan pengguna. Seluruh temuan kemudian dibahas dalam kaitannya dengan tujuan penelitian, yaitu mengidentifikasi ekspresi ketidaksabaran gamer dalam ulasan aplikasi komunitas digital serta memberikan rekomendasi praktis bagi pengembang dalam meningkatkan kualitas layanan Discord.

### 4.1 Business Understanding

Discord merupakan aplikasi komunikasi digital yang banyak digunakan oleh komunitas gamer. Dengan fitur *voice channel*, *text channel*, *screen sharing*, dan server komunitas, Discord menyediakan ruang interaksi intensif di luar permainan utama. Namun, popularitas ini juga memunculkan beragam ulasan pengguna, baik berupa apresiasi maupun keluhan. Penelitian ini berangkat dari pertanyaan bagaimana karakteristik ketidaksabaran gamer tercermin dalam ulasan, serta bagaimana analisis sentimen dan topik dapat digunakan untuk memahami persepsi mereka.

### 4.2 Data Understanding

Dataset penelitian ini terdiri dari **10.000 ulasan** pengguna aplikasi Discord yang diperoleh dari Google Play Store melalui proses *web scraping*. Setiap entri mencakup atribut id, username, rating, tanggal publikasi, dan teks ulasan. Eksplorasi awal memperlihatkan pola distribusi rating ekstrem: rating 5 bintang banyak memuat apresiasi terkait kualitas suara dan fitur komunitas, sementara rating 1 bintang umumnya berisi keluhan teknis seperti gagal login, bug, atau aplikasi crash. Temuan ini sejalan dengan pola yang juga dijelaskan pada penelitian-penelitian terdahulu mengenai ulasan aplikasi digital, di mana mayoritas ulasan cenderung berada pada kutub sangat positif atau sangat negatif.

Selain itu, untuk kebutuhan analisis topik digunakan metode **Latent Dirichlet Allocation (LDA)**. LDA bekerja dengan ketentuan bahwa setiap dokumen (ulasan) dipandang sebagai campuran dari beberapa topik, sedangkan setiap topik direpresentasikan oleh distribusi kata-kata tertentu. Dengan demikian, satu ulasan dapat memuat lebih dari satu topik dengan proporsi yang berbeda. Dalam penelitian ini jumlah topik ditentukan berdasarkan *coherence score* untuk memastikan interpretasi yang konsisten, dan kata-kata dengan probabilitas tertinggi pada tiap topik dipilih sebagai kata kunci. Ketentuan tersebut memungkinkan LDA mengelompokkan ulasan ke dalam tema-tema yang bermakna, misalnya topik mengenai kualitas komunikasi, keluhan login, maupun bug aplikasi, sehingga analisis tidak hanya terbatas pada polaritas sentimen, tetapi juga menggali isu spesifik yang relevan dengan pengalaman pengguna. Rincian hasil scraping data ulasan pengguna aplikasi Discord yang diperoleh dari Google Play Store ditampilkan pada **Tabel 1**.

**Tabel 1. Hasil scraping data ulasan discord dari google playstore**

id	user	rating	tanggal	ulasan
d128fbc9-e5a4-420b-84cd-fef8f199f7e8	miswan fuad	5	2025-07-14 14:59:37	very good to be grow community
80f8b69d-5519-4249-b3ba-8da9854e15c8	Dea Indah Sari	5	2025-07-14 13:06:40	sangat bagus untuk saya terima kasih 😊😊😊😊



a6dc5ba6-aa26-4605-8eaa-ad3a38c9732e	rafael tmz	3	2025-07-14 12:43:48	semoga nanti update selanjutnya bisa ada sistem keamanan profile picture ☺
a32d8c4f-4ecc-45a2-8b6c-f24b1effee18	Sheva Dhavinvi rendra	1	2025-07-14 11:58:19	Saya tidak sengaja berlangganan tiba tiba duit 100rb saya hilang ketarik sama nitro discord tanggung jawab dong itu duit buat bayar di Tokopedia
4734876a-fbf4-41ea-aeab-9857e9df3bee	Muhamm ad Alvin	1	2025-07-14 08:21:00	kenapa saat saya mencoba membuat akun menggunakan nomor hp,kode verifikasi nya tidak di kirim sama sekali jadi tolong perbaiki masalah ini.
64db03e6-86cf-41bc-950a-8ea9d6c9ac0a	Farizky	5	2025-07-14 07:46:05	bagus
ebddd6bc-d57d-48a9-b131-ccfa3c11db5d	Arfan	1	2025-07-14 07:43:15	apa sih discord ini minta verify phone harus nya pake email jangan nomor hp karena pas aku masukin nomor hp malah invalid phone number, aku udah perhatiin itu sama banget sama nomor hp aku, tapi tetep aja invalid phone number gak jelas banget padahal ada yang memang lagi kubutuhin banget malah disuruh verify phone, kalo gk dibenerin bug ini aku hapus aja discord, jelek banget
7e3eb460-b5ae-4db2-acf7-f1fc9dfe7528	YT MKF	5	2025-07-14 06:45:53	bagus
a16cf525-5da1-42a3-a489-aa6fd6259f23	Hanifah Qia	5	2025-07-14 04:50:47	apk bagus tapi kadang suka saling hujat 🤬
bec34b07-cc5e-4c0c-87f6-591a41437a8f	anton rorympan dey	5	2025-07-14 01:23:34	siipp

### 4.3 Data Preparation

Proses *preprocessing* dilakukan agar teks siap dianalisis. Langkah yang ditempuh meliputi *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Hasilnya, teks ulasan menjadi lebih ringkas dan konsisten. Contohnya, ulasan “Aplikasi ini bagus banget buat ngobrol sama teman, fiturnya keren!” diproses menjadi “aplikasi bagus ngobrol teman fitur keren”. Dengan teks yang sudah bersih, model analisis dapat lebih akurat dalam mengidentifikasi kata kunci yang bermakna. Hasil penerapan tahapan preprocessing terhadap teks ulasan ditunjukkan pada **Tabel 2**.

**Tabel 2 Hasil preprocessing teks ulasan**

No	Ulasan	Ulasan Preprocessed
1	very good to be grow comunity	very, good, to, be, grow, comunity
2	sangat bagus untuk saya terima kasih 🥰🥰🥰🥰	bagus, terima, kasih
3	semoga nanti update selanjutnya bisa ada sistem keamanan profile picture ☺	moga, update, sistem, aman, profile, picture
4	saya tidak sengaja berlangganan tiba tiba duit 100rb saya hilang ketarik sama nitro discord tanggung jawab dong itu duit buat bayar di tokopedia	sengaja, langgan, duit, rb, hilang, tarik, nitro, discord, tanggung, duit, bayar, tokopedia
5	kenapa saat saya mencoba membuat akun	coba, akun, nomor, hpkode,

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

	menggunakan nomor hp,kode verifikasi nya tidak di kirim sama sekali jadi tolong perbaiki masalah ini.	verifikasi, nya, kirim, tolong, baik
6	bagus	bagus
7	apa sih discord ini minta verify phone harus nya pake email jangan nomor hp karena pas aku masukin nomor hp malah invalid phone number, aku udah perhatiin itu sama banget sama nomor hp aku, tapi tetep aja invalid phone number gak jelas banget padahal ada yang memang lagi kubutuhin banget malah disuruh verify phone, kalo gk dibenerin bug ini aku hapus aja discord, jelek banget	sih, discord, verify, phone, nya, pakai, email, nomor, hp, pas, masukin, nomor, hp, invalid, phone, number, udah, perhatiin, banget, nomor, hp, tetep, aja, invalid, phone, number, banget, kubutuhin, banget, suruh, verify, phone, kalo, dibenerin, bug, hapus, aja, discord, jelek, banget
8	bagus	bagus
9	apk bagus tapi kadang suka saling hujat 🤔	apk, bagus, kadang, suka, hujat
10	siipp	siipp

#### 4.4 Modeling

Tahap modeling adalah inti dari penelitian ini. Ada tiga model utama yang digunakan: labeling sentimen dengan IndoBERT, klasifikasi dengan SVM, dan pemodelan topik dengan LDA.

##### 4.4.1 Labeling Sentimen dengan IndoBERT

Pelabelan awal dilakukan menggunakan model IndoBERT Sentiment Classifier terhadap data ulasan pengguna aplikasi Discord. Dari 10.000 ulasan yang dikumpulkan, diperoleh hasil bahwa **57,4% diklasifikasikan sebagai sentimen positif** dan **42,6% sebagai sentimen negatif**. Dominasi sentimen positif menunjukkan adanya apresiasi pengguna terhadap fitur-fitur utama aplikasi Discord seperti kemudahan akses, performa stabil, dan komunikasi yang lancar.

Namun demikian, jumlah ulasan negatif yang signifikan mengindikasikan adanya permasalahan umum yang dirasakan pengguna, seperti gangguan teknis dan ketidaksesuaian ekspektasi dalam pengalaman penggunaan. Distribusi ini menjadi dasar untuk eksplorasi lebih lanjut melalui klasifikasi dan pemodelan topik.

##### 4.4.2 Klasifikasi Sentimen dengan SVM

Hasil pengujian menunjukkan bahwa model SVM mampu mencapai **akurasi sebesar 90,46%**, dengan **precision, recall, dan F1-score** berkisar antara **0,89 hingga 0,92**. **Gambar 2** menampilkan classification report yang menggambarkan performa model secara menyeluruh, sementara **Gambar 3 memperlihatkan confusion matrix** yang menunjukkan jumlah prediksi benar dan salah pada masing-masing kelas sentimen.

```

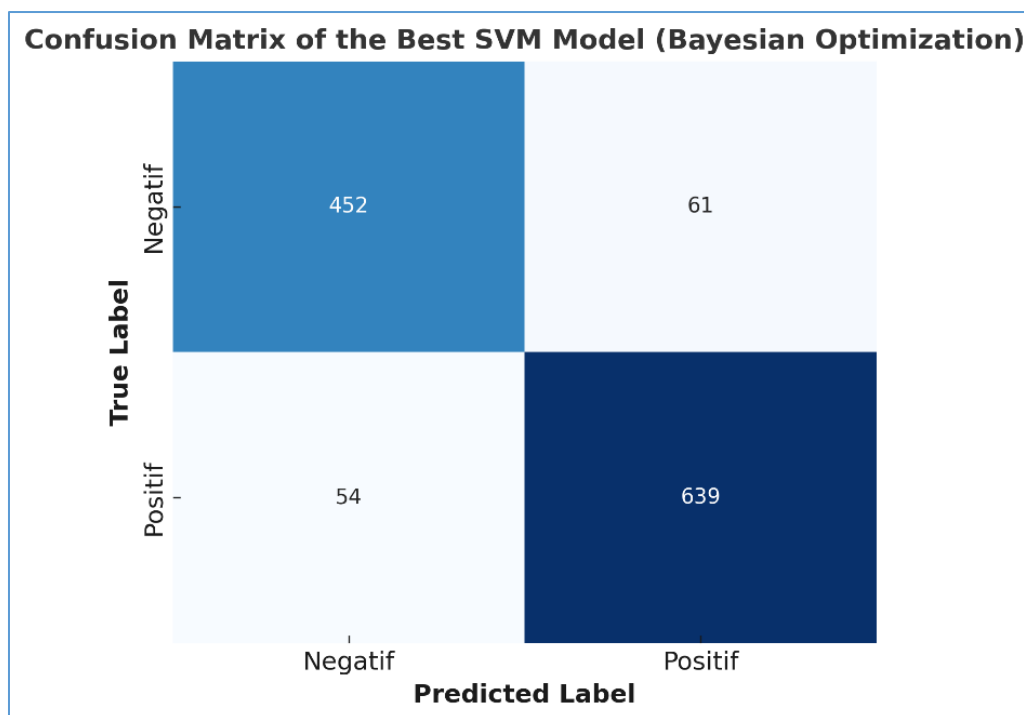
Accuracy of the best SVM model (Bayesian Optimization): 0.9046

Classification Report of the best SVM model (Bayesian Optimization):

```

	precision	recall	f1-score	support
negatif	0.89	0.88	0.89	513
positif	0.91	0.92	0.92	693
accuracy			0.90	1206
macro avg	0.90	0.90	0.90	1206
weighted avg	0.90	0.90	0.90	1206

Gambar 2. Hasil Classification Report Model SVM



**Gambar 3 Hasil confusion matrix model SVM**

Confusion matrix menunjukkan bahwa tingkat kesalahan klasifikasi hanya sebesar 9,54% (115 dari 1.206 ulasan), sebagaimana ditampilkan pada **Gambar 3**. Capaian ini menegaskan bahwa model memiliki tingkat keandalan yang tinggi dalam membedakan polaritas ulasan, khususnya dalam konteks bahasa informal yang umum digunakan oleh komunitas digital.

Model juga menunjukkan kinerja yang seimbang dalam mengklasifikasikan sentimen positif dan negatif, mencerminkan kemampuannya dalam menangani variasi ekspresi yang kompleks. Keseimbangan klasifikasi ini sangat penting karena mendukung tujuan utama analisis, yaitu mengidentifikasi pola perilaku pengguna seperti ketidaksabaran dan frustrasi dalam interaksi digital.

Selain dua label utama (positif dan negatif), proses labeling awal juga mengungkap keberadaan ulasan dengan nada netral atau ambigu, misalnya “lumayan”, “cukup oke”, atau “ya begitulah”. Karena keterbatasan pada model IndoBERT yang hanya mendukung klasifikasi biner, ulasan dengan nada netral tersebut dikecualikan dari proses pelatihan dan evaluasi model. Keterbatasan ini menjadi catatan penting dalam interpretasi hasil, dan membuka peluang untuk pendekatan klasifikasi tiga kelas pada penelitian selanjutnya.

#### 4.4.3 Pemodelan Topik dengan LDA

Analisis topik dilakukan dengan **Latent Dirichlet Allocation (LDA)** secara terpisah pada ulasan positif dan negatif. Untuk sentimen positif, **Tabel 3**. menampilkan hasil LDA yang menghasilkan 10 topik utama.

**Tabel 3. Hasil Latent Dirichlet Allocation (LDA) sentimen positif**

No	Nama Topik	Kata Kunci	Jumlah	Persentase
1	Aplikasi yang baik dan menarik meskipun ada sedikit bug	bagus, aplikasi, bug, komunikasi, orang	1245	35,95%
2	Apresiasi kualitas dan fungsi aplikasi	good, app, great, perfect, job	387	11,18%
3	Pengalaman login dan penggunaan aplikasi secara umum positif	mantap, nice, login, manfaat	379	10,94%
4	Pengalaman bermain dan berinteraksi bersama teman	main, teman, ngobrol, discord	278	8,03%
5	Manfaat aplikasi Discord dan	fitur, tambah, aplikasi, update	244	7,05%



	penggunaan fitur tambahan			
6	Ekspresi kekaguman dalam bahasa Inggris	the, i, this, like, app	239	6,90%
7	Ungkapan kepuasan dan ucapan terima kasih	oke, login, email, aman	235	6,79%
8	Penggunaan Discord untuk komunikasi suara dan bermain gim	keren, suka, suara, seru	220	6,35%
9	Permintaan fitur tambahan serta perbaikan	fitur, update, tolong	177	5,11%
10	Pengalaman bersosialisasi dan bermain gim bersama teman	teman, bareng, ngobrol, seru	61	1,76%

**Ulasan positif** menghasilkan 10 topik utama. Topik terbesar adalah “Aplikasi yang baik dan menarik meskipun ada sedikit bug” (35,95%), diikuti apresiasi kualitas aplikasi (11,18%) dan pengalaman login yang positif (10,94%). Tema lain yang muncul adalah interaksi dengan teman, manfaat fitur tambahan, kualitas suara, serta permintaan fitur baru. Mayoritas ulasan positif menekankan kenyamanan penggunaan, kualitas komunikasi, dan kepuasan terhadap fungsi utama aplikasi.

Sementara itu, pada kelompok sentimen negatif, LDA mengidentifikasi tema seperti **keterlambatan kode verifikasi, gangguan login, bug teknis setelah pembaruan, dan ketergantungan koneksi jaringan**. Topik-topik ini secara lebih rinci dapat dilihat pada **Tabel 4**, yang menampilkan distribusi tema dominan dalam ulasan negatif beserta bobotnya.

**Tabel 4. Hasil Latent Dirichlet Allocation (LDA) sentimen negatif**

No	Nama Topik	Kata Kunci	Jumlah	Persentase
1	Kendala login dan akses akun	login, akun, discord, gagal	464	18,09%
2	Masalah verifikasi nomor HP atau email	verifikasi, kode, email, nomor	379	14,78%
3	Isu teknis bug pada suara, video, update aplikasi	bug, suara, server, update	319	12,44%
4	Permintaan bantuan pengguna disertai kekecewaan	tolong, gagal, susah, hapus	294	11,46%
5	Masalah login dan password pengguna	login, password, akun	261	10,18%
6	Hambatan terkait akun, login, verifikasi	akun, login, verifikasi	257	10,02%
7	Kesulitan login dan akses fitur	login, akses, fitur	240	9,36%
8	Permintaan penambahan/pengembalian fitur	fitur, tambah, update	197	7,68%
9	Topik spam atau ulasan tidak jelas	spam, kosong, ulasan	106	4,13%
10	Pengalaman negatif saat bermain gim atau bersosialisasi	gim, main, susah	47	1,83%

**Ulasan negatif** juga terbagi ke dalam 10 topik utama. Keluhan terbesar adalah “Kendala login dan akses akun” (18,09%), “Masalah verifikasi nomor HP atau email” (14,78%), dan “Bug teknis pada suara, video, dan update aplikasi” (12,44%). Topik lain mencakup permintaan bantuan, masalah password, hambatan autentikasi, serta permintaan penambahan fitur. Sebagian kecil ulasan negatif berupa spam atau keluhan emosional, namun tetap menunjukkan tingkat frustrasi pengguna terhadap gangguan teknis.

Temuan ini menunjukkan bahwa hambatan kecil sekalipun dalam interaksi teknis dapat memicu frustrasi dan respons negatif dari pengguna pola yang konsisten dengan perilaku impulsif dalam komunitas digital. Topik-topik negatif tersebut mencerminkan ekspresi langsung terhadap

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

Untuk mendukung temuan ini secara visual, disajikan **wordcloud** dari masing-masing **kelompok sentimen**. Pada **Gambar 4**, kata-kata dominan dalam ulasan positif meliputi *bagus*, *mantap*, *keren*, *teman*, dan *suara*. Sementara itu, **Gambar 5** menunjukkan bahwa kata-kata yang sering muncul dalam ulasan negatif adalah *login*, *akun*, *verifikasi*, *tolong*, dan *email*.



## 4.5 Evaluation

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

Dari sisi kuantitatif, Model SVM menunjukkan **akurasi sebesar 90,46%**, dengan **F1-score antara 0,89 hingga 0,92**. Nilai precision dan recall yang tinggi menandakan bahwa model mampu membedakan ulasan positif dan negatif secara akurat. Confusion matrix menunjukkan bahwa proporsi kesalahan klasifikasi berada di bawah 10%, memperkuat keandalan model dalam menangani variasi data nyata dari komunitas pengguna digital.

Model LDA juga menghasilkan koherensi topik yang baik dengan distribusi topik yang stabil. Hasil ini menunjukkan bahwa pendekatan integratif SVM-LDA mampu menangkap baik dimensi emosional maupun tematik dari ulasan.

Sementara itu, evaluasi kualitatif melalui topik-topik hasil LDA sangat merepresentasikan pengalaman pengguna yang sesungguhnya. Ulasan positif banyak berkaitan dengan pengalaman komunitas yang menyenangkan, fitur suara yang stabil, dan desain antarmuka yang intuitif. Ulasan negatif mencerminkan ekspresi ketidaksabaran terhadap masalah teknis kecil, seperti bug minor atau gangguan login.

Namun, perlu dicatat bahwa ulasan netral belum dapat ditangani secara langsung oleh model karena keterbatasan skema klasifikasi biner. Ini menjadi catatan penting, karena komentar-komentar ambigu cukup umum dalam komunitas digital, dan ke depannya pendekatan klasifikasi tiga kelas (positif-netral-negatif) dapat lebih mewakili spektrum sentimen yang lebih realistis.

Visualisasi data turut memperkuat interpretasi hasil evaluasi. Grafik distribusi sentimen memperlihatkan dominasi ulasan positif, meskipun ulasan negatif memiliki proporsi yang cukup signifikan untuk diperhatikan. Wordcloud yang dihasilkan menampilkan perbedaan kata kunci yang jelas antara kedua jenis ulasan. Pada wordcloud ulasan positif pada **Gambar 4**, kata-kata seperti *bagus*, *mantap*, *teman*, dan *suara* menegaskan kepuasan pengguna terhadap fitur komunikasi. Sebaliknya, pada wordcloud ulasan negatif pada **Gambar 5**, kata-kata seperti *login*, *kode*, *bug*, dan *error* menekankan masalah teknis yang menjadi hambatan utama dalam pengalaman penggunaan aplikasi. Keterkaitan antara distribusi statistik, topik dominan, dan representasi visual memperlihatkan konsistensi hasil, sekaligus memastikan bahwa interpretasi yang diberikan valid secara metodologis maupun bermakna secara praktis.

Jika dikaitkan dengan tujuan penelitian, evaluasi ini menunjukkan bahwa metode yang digunakan telah berhasil mengungkap pola ketidaksabaran gamer secara menyeluruh. SVM memungkinkan pemetaan polaritas sentimen dengan akurasi tinggi, sehingga dapat mengidentifikasi secara jelas perbedaan antara ulasan positif dan negatif. LDA memberikan konteks yang lebih dalam mengenai tema-tema yang melatarbelakangi sentimen tersebut, terutama keluhan teknis yang sering kali memicu respons negatif dari pengguna. Dengan demikian, kombinasi IndoBERT, SVM, dan LDA tidak hanya terbukti efektif dari sisi teknis, tetapi juga relevan secara konseptual dengan fenomena perilaku pengguna yang menjadi fokus penelitian.

Secara keseluruhan, evaluasi membuktikan bahwa penelitian ini berhasil mencapai tujuan yang telah ditetapkan. Model klasifikasi mampu memberikan gambaran kuantitatif mengenai distribusi sentimen, sementara analisis topik dan visualisasi melengkapi hasil tersebut dengan pemahaman kualitatif yang lebih mendalam. Hasil evaluasi ini pada akhirnya memberikan kontribusi baik secara akademis, melalui penerapan metode analisis teks yang terbukti efektif, maupun secara praktis, melalui rekomendasi yang dapat dimanfaatkan pengembang Discord untuk meningkatkan kualitas layanan aplikasi.

## 5 Kesimpulan

Penelitian ini berhasil mencapai tujuannya dalam menganalisis sentimen dan topik dari ulasan pengguna aplikasi Discord dengan memanfaatkan kerangka kerja CRISP-DM. Algoritma Support Vector Machine (SVM) yang dioptimasi menggunakan Bayesian Optimization mampu mengklasifikasikan polaritas sentimen secara akurat, dengan akurasi mencapai 90,46%. Sementara itu, pemodelan topik dengan Latent Dirichlet Allocation (LDA) mengungkap bahwa ulasan positif umumnya berkaitan dengan apresiasi terhadap kualitas komunikasi suara, fitur komunitas, dan antarmuka pengguna. Sebaliknya, ulasan negatif didominasi oleh keluhan terkait login, verifikasi, bug, dan gangguan jaringan. Temuan ini menegaskan bahwa perilaku ketidaksabaran gamer berdampak signifikan terhadap pola ulasan yang muncul. Hasil penelitian memberikan rekomendasi



praktis bagi pengembang Discord untuk meningkatkan aspek teknis aplikasi serta memperkuat fitur komunitas agar lebih selaras dengan ekspektasi pengguna.

## Referensi

- [1] O. Luolila, "What Sells Video Games? A Qualitative Study into Customer Preferences in Video Game Design.," 2023, Accessed: Sep. 10, 2025. [Online]. Available: <https://aaltoodoc.aalto.fi/bitstreams/e847fdef-17e5-4cde-ad5c-368575f38d5a/download>
- [2] Y. Yuan, J. Cao, R. Wang, and S. Yarosh, "Tabletop Games in the Age of Remote Collaboration: Design Opportunities for a Socially Connected Game Experience," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–14. Accessed: Sep. 10, 2025. [Online]. Available: <https://par.nsf.gov/servlets/purl/10237337>
- [3] A. Boldi, A. Rapp, and M. V Birk, "Stress in Esports: A Qualitative Study on the Interplay of Player Experiences and Organizational Systems," *Int J Hum Comput Interact*, Vol. 41, No. 15, pp. 9224–9247, 2025, DOI: <https://doi.org/10.1080/10447318.2024.2423343>.
- [4] V. Aalto, "Creating Games to Create Communities," 2024, Accessed: Sep. 10, 2025. [Online]. Available: <https://urn.fi/URN:NBN:fi:amk-2024122037822>
- [5] F. Adigüzel, "The Effect of YouTube Reviews on Video Game Sales," *İşletme Araştırmaları Dergisi*, Vol. 13, No. 3, pp. 2096–2109, 2021, DOI: <https://doi.org/10.20491/isarder.2021.1249>.
- [6] H. E. Park and S.-F. Yap, "A Dualistic Approach to Harmonious and Obsessive Passion: The Impact of Game Mechanics on Mobile Gaming Engagement," *Electronic Commerce Research*, pp. 1–24, 2024, DOI: <https://doi.org/10.1007/s10660-024-09886-7>.
- [7] K. Zheng and R. Farzan, "Understanding Player's Gesture-based Communicative Behavior in MOBA Games," *Proc ACM Hum Comput Interact*, Vol. 7, No. CHI PLAY, pp. 1068–1090, 2023, DOI: <https://doi.org/10.1145/3611061>.
- [8] A. Czauderna, S. von Gillern, and B. Robinson, "Informal Social-Emotional Learning in Gaming Affinity Spaces: Evidence from a Reddit Discussion Thread on Elden Ring," *Simul Gaming*, Vol. 55, No. 1, pp. 30–50, 2024, DOI: <https://doi.org/10.1177/10468781231209697>.
- [9] Y. Liu, Q. Li, B. Wang, Y. Zhang, and D. Song, "A Survey of Quantum-Cognitively Inspired Sentiment Analysis Models," *ACM Comput Surv*, Vol. 56, No. 1, pp. 1–37, 2023.
- [10] J. G. Reitman, C. G. Anderson, and C. Steinkuehler, "Discord Community Challenges, Tools, and Strategies," Irvine, CA: Connected Learning Lab, 2021, Accessed: Sep. 10, 2025. [Online]. Available: <https://connectedlearning.uci.edu/wp-content/uploads/2022/09/2021-Y4-Discord-Community-Challenges-Tools-and-Strategies-Report.pdf>
- [11] F. F. Rachman and S. Pramana, "Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter," *Indonesian of Health Information Management Journal (INOHIM)*, Vol. 8, No. 2, pp. 100–109, 2020, Accessed: Sep. 10, 2025. [Online]. Available: <https://inohim.esaunggul.ac.id/index.php/INO/article/download/223/175>
- [12] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *Jurnal Tekno Kompak*, Vol. 15, No. 1, pp. 131–145, 2021, Accessed: Sep. 10, 2025. [Online]. Available: <https://www.academia.edu/download/97852578/561.pdf>
- [13] M. I. A. G. Wibowo and I. Pratama, "Analisis Sentimen Ulasan Aplikasi Identitas Kependudukan Digital menggunakan Metode Support Vector Machine," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, Vol. 6, No. 4, pp. 715–722, 2024, Accessed: Sep. 10, 2025. [Online]. Available: <http://103.241.192.17/~jurnalunidha/index.php/jteksis/article/download/1552/909>
- [14] R. M. Singh, "Mathematical Concepts of Linear Algebra in AI Tools: A Calculation-based Study," *Journal of Hyperstructures*, 2024, Accessed: Sep. 10, 2025. [Online]. Available: [https://journal.uma.ac.ir/article\\_3548\\_b68e9a95330321ba0344d453052d7431.pdf](https://journal.uma.ac.ir/article_3548_b68e9a95330321ba0344d453052d7431.pdf)
- [15] P. Celard, A. S. Vieira, E. L. Iglesias, and L. Borrajo, "LDA Filter: A Latent Dirichlet Allocation Preprocess Method for Weka," *PLoS One*, Vol. 15, No. 11, p. e0241701, 2020, DOI: <https://doi.org/10.1371/journal.pone.0241701>.
- [16] M. H. Basri, "Pemodelan Topik dan Analisis Sentimen pada Teks Ulasan Pengguna Aplikasi Perbankan Seluler di Indonesia," *The Indonesian Journal of Computer Science*, Vol. 13, No. 4,

- 2024, Accessed: Sep. 10, 2025. [Online]. Available: <http://ijcs.net/ijcs/index.php/ijcs/article/download/4200/725>
- [17] A. N. H. A. K. Rotty, T. S. Dewayana, and A. N. Habyba, “Cross-Industry Standard Process for Data Mining (CRISP-DM) Approach in Determining the Most Significant Employee Engagement Drivers to Sales at X Car Dealership,” in *Proc. 3rd Asia Pacific Int. Conf. Ind. Eng. Oper. Manag. Johor Bahru, Malaysia*, 2022, pp. 3368–3379. Accessed: Sep. 10, 2025. [Online]. Available: <http://ieomsociety.org/proceedings/2022malaysia/552.pdf>
- [18] C. Schröer, F. Kruse, and J. M. Gómez, “A Systematic Literature Review on Applying CRISP-DM Process Model,” *Procedia Comput SCI*, Vol. 181, pp. 526–534, 2021, DOI: <https://doi.org/10.1016/j.procs.2021.01.199>.
- [19] S. Patmanthara *et al.*, “Belantika Pendidikan Pengaruh Integrasi *Google Classroom* dengan Platform Pendukung dalam Pembelajaran Interaktif terhadap Minat Belajar Informatika Siswa Kelas X SMA Negeri 8 Malang,” 2024. Accessed: Sep. 10, 2025. [Online]. Available: <https://core.ac.uk/download/pdf/631296646.pdf>
- [20] M. D. Ikhsan, B. Huda, A. Hananto, and F. Nurapriani, “Analisis Sentimen Ulasan Pengguna Aplikasi Traveloka Pada *Google Play Store* menggunakan *Algoritma Naive Bayes*,” 2025, Accessed: Sep. 10, 2025. [Online]. Available: <https://e-journal.hamzanwadi.ac.id/index.php/infotek/article/download/30444/6909>
- [21] D. A. Ubaidah *et al.*, “Digitalisasi Referensi Klasik Islam sebagai Upaya Modernisasi Pembelajaran di Pesantren Babakan Jamanis, Karangbenda, Kecamatan Parigi, Kabupaten Pangandaran: Pengabdian,” *Jurnal Pengabdian Masyarakat dan Riset Pendidikan*, Vol. 4, No. 1, pp. 3515–3526, 2025, Accessed: Sep. 10, 2025. [Online]. Available: <http://jerk.in.org/index.php/jerk.in/article/download/1815/1681>
- [22] S. Oktaviana, L. Kurniawati, and S. A. Putri, “Pelatihan Trello sebagai Alat Manajemen Kegiatan untuk Organisasi Pemuda: Studi Kasus Ikrar 007 Depok,” *Jurnal Pengabdian Kolaborasi dan Inovasi IPTEKS*, Vol. 3, No. 4, pp. 897–904, 2025, Accessed: Sep. 10, 2025. [Online]. Available: <https://journal.ppmi.web.id/index.php/JPKI2/article/download/2697/1820>
- [23] M. B. Huda, “Pemanfaatan Discord sebagai Alternatif Media Pembelajaran secara Daring,” *Jurnal Ilmu-ilmu Sosial*, Vol. 19, No. 2, pp. 661–666, 2022, Accessed: Sep. 10, 2025. [Online]. Available: <https://core.ac.uk/download/pdf/552629838.pdf>
- [24] M. A. Alghaffaru, A. Soleh, and M. R. Febliansa, “Gaya Hidup Kompulsif *Gamers Online Mobile Legends* Bang Bang di Kota Bengkulu,” in *Seminar Nasional Ekonomi Manajemen, Bisnis dan Akuntansi*, 2023, pp. 11–18. Accessed: Sep. 10, 2025. [Online]. Available: <https://proceeding.unived.ac.id/index.php/sn-emba/article/view/157>
- [25] 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, Accessed: Sep. 10, 2025. [Online]. Available: <https://apic.id/jurnal/index.php/jsc/article/download/162/86>