

# ANALISIS SENTIMEN BERBASIS TOPIK ULASAN PENGGUNA APLIKASI ROBLOX MENGGUNAKAN INTEGRASI LDA DAN SVM

Galih Rafianto<sup>1\*</sup>, Hannie<sup>2</sup>, Aziz Ma'sum<sup>3</sup>

<sup>1,2,3</sup>Universitas Singaperbangsa Karawang; Jl. HS. Ronggo Waluyo, Puserjaya, Telukjambe Timur, Karawang, Jawa Barat, 41361; (0267) 641177

## Keywords:

Analisis Sentimen Berbasis Topik;  
Roblox;  
*Knowledge Discovery in Databases*;  
*Latent Dirichlet Allocation*;  
*Support Vector Machine*.

## Correspondent Email:

galihrafianto179@gmail.com

**Abstrak.** Perkembangan platform permainan virtual komunal seperti Roblox diiringi oleh tingginya volume ulasan pengguna di Google Play Store. Ulasan tersebut mengandung informasi penting mengenai kepuasan pengguna dan berbagai kendala teknis, namun jumlahnya yang besar menyulitkan proses evaluasi secara manual. Penelitian ini bertujuan menganalisis opini pengguna melalui pendekatan analisis sentimen berbasis topik menggunakan metodologi *Knowledge Discovery in Databases* (KDD). *Dataset* sebanyak 40.089 ulasan berbahasa Indonesia dianalisis menggunakan *Latent Dirichlet Allocation* (LDA) untuk mengekstraksi topik utama dan *Support Vector Machine* (SVM) untuk mengklasifikasikan polaritas sentimen. Ketidakseimbangan kelas sentimen diatasi melalui penerapan penyesuaian bobot kelas (*class weight*) pada proses pelatihan model SVM. Hasil pemodelan LDA mengidentifikasi empat topik utama dengan nilai *Coherence Score* sebesar 0,5547 dan *Perplexity* 191,5, yaitu masalah akun dan gangguan teknis, pengalaman bermain dan interaksi sosial, fitur *item* serta monetisasi, serta performa aplikasi dan koneksi jaringan. Model SVM memberikan performa terbaik pada pembagian data 70:30 dengan akurasi 85,22%, presisi 89%, *recall* 85%, dan *F1-score* 86%. Hasil penelitian menunjukkan bahwa integrasi LDA dan SVM efektif dalam mengungkap pola opini pengguna dan dapat menjadi dasar bagi pengembang dalam memprioritaskan perbaikan aplikasi.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

**Abstract.** The rapid growth of communal virtual gaming platforms such as Roblox has been accompanied by a large volume of user reviews on the Google Play Store. These reviews contain valuable information regarding user satisfaction and technical issues; however, the large quantity of data makes manual evaluation difficult. This study aims to analyze user opinions through a topic-based sentiment analysis approach using the *Knowledge Discovery in Databases* (KDD) methodology. A dataset consisting of 40,089 Indonesian-language reviews was analyzed using *Latent Dirichlet Allocation* (LDA) to extract discussion topics and *Support Vector Machine* (SVM) to classify sentiment polarity. The imbalance in sentiment classes was addressed by applying class weight adjustment during the SVM training process. The LDA modeling identified four main topics with a *Coherence Score* of 0.5547 and *Perplexity* of 191.5, including account issues and technical disruptions, gameplay experience and social interaction, item features and monetization, and application performance and network connectivity. The SVM model achieved the best performance with a 70:30 data split, resulting in an accuracy of 85.22%, precision of 89%, recall of 85%, and *F1-score* of 86%. The findings demonstrate that the integration of LDA and SVM is effective in

*uncovering user opinion patterns and can provide strategic insights for developers to prioritize application improvements.*

## 1. PENDAHULUAN

Perkembangan teknologi digital telah mengubah cara pengguna berinteraksi dan menghasilkan informasi di internet. Pengguna tidak lagi hanya berperan sebagai konsumen, tetapi juga sebagai produsen konten melalui fenomena *User-Generated Content* (UGC) yang mencakup ulasan, komentar, maupun berbagai bentuk ekspresi digital lainnya [1]. Dalam konteks industri game online, kontribusi pengguna melalui UGC memiliki peran penting karena dapat memberikan masukan langsung terhadap kualitas layanan dan pengalaman bermain yang disediakan oleh pengembang [2].

Salah satu platform yang berkembang pesat dalam ekosistem tersebut adalah Roblox, sebuah platform permainan berbasis metaverse yang memungkinkan pengguna untuk membuat, membagikan, dan memainkan pengalaman virtual yang dikembangkan oleh komunitas pengguna [3]. Model ekosistem ini menjadikan pemain tidak hanya sebagai konsumen tetapi juga sebagai kreator konten digital dalam lingkungan yang dinamis [4]. Dengan jumlah pengguna aktif harian yang mencapai lebih dari 82,9 juta pada tahun 2024 [5], Roblox menghasilkan volume ulasan pengguna yang sangat besar pada berbagai platform distribusi aplikasi seperti Google Play Store. Ulasan tersebut memuat berbagai opini terkait pengalaman bermain, kendala teknis, maupun saran pengembangan aplikasi. Namun, jumlah ulasan yang sangat besar menyebabkan proses analisis secara manual menjadi tidak efisien sehingga diperlukan pendekatan analisis berbasis komputasi.

Salah satu pendekatan yang umum digunakan untuk menganalisis opini pengguna adalah analisis sentimen, yaitu teknik komputasi yang digunakan untuk mengidentifikasi polaritas opini dalam teks, seperti sentimen positif, negatif, atau netral [6]. Meskipun demikian, analisis sentimen konvensional hanya memberikan gambaran umum mengenai distribusi opini pengguna dan belum mampu menunjukkan aspek atau topik spesifik yang menjadi sumber kepuasan maupun keluhan pengguna [7]. Oleh karena itu,

diperlukan pendekatan yang lebih komprehensif melalui analisis sentimen berbasis topik (*topic-based sentiment analysis*), yang memungkinkan identifikasi sentimen pengguna berdasarkan topik pembahasan tertentu dalam ulasan [8].

Dalam implementasinya, pemodelan topik sering dilakukan menggunakan algoritma *Latent Dirichlet Allocation* (LDA), yaitu metode *unsupervised learning* yang mampu mengidentifikasi pola kemunculan kata untuk membentuk topik tertentu dalam kumpulan dokumen teks [9]. Setelah topik diperoleh, klasifikasi sentimen dapat dilakukan menggunakan algoritma *Support Vector Machine* (SVM) yang dikenal memiliki performa baik dalam pengolahan data teks berdimensi tinggi [10]. Beberapa penelitian sebelumnya menunjukkan bahwa integrasi kedua metode tersebut mampu menghasilkan analisis opini yang lebih komprehensif. Penelitian Kustiyangsih dan Permana (2024) pada aplikasi EdLink menunjukkan bahwa kombinasi LDA dan SVM mampu mengidentifikasi aspek penting seperti *usability* dan *efficiency* dengan tingkat akurasi klasifikasi sebesar 90% [11]. Penelitian Sa'dul Asyhar et al. (2024) pada aplikasi Jenius juga berhasil mengidentifikasi lima topik utama menggunakan LDA dan mencapai akurasi sentimen sebesar 94,03% dengan SVM [12]. Selain itu, Prastyo et al. (2024) menunjukkan bahwa pendekatan serupa efektif dalam memetakan keluhan pengguna pada aplikasi myIM3 [13].

Meskipun berbagai penelitian telah menerapkan integrasi LDA dan SVM pada analisis ulasan aplikasi, kajian yang secara khusus menganalisis ulasan pengguna Roblox masih terbatas, terutama pada ulasan berbahasa Indonesia di Google Play Store. Padahal, ulasan pengguna pada platform tersebut memiliki karakteristik yang sangat informatif karena sering memuat laporan bug, keluhan teknis, serta permintaan fitur yang penting bagi pengembangan perangkat lunak [14]. Oleh karena itu, diperlukan penelitian yang mampu mengidentifikasi topik utama sekaligus

memetakan sentimen pengguna secara lebih mendalam pada ulasan aplikasi Roblox.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengidentifikasi topik utama dalam ulasan pengguna Roblox menggunakan metode *Latent Dirichlet Allocation* (LDA), membangun model klasifikasi sentimen menggunakan *Support Vector Machine* (SVM), serta memetakan distribusi sentimen pada setiap topik yang ditemukan. Melalui integrasi kedua metode tersebut, penelitian ini diharapkan dapat memberikan pemahaman yang lebih komprehensif mengenai persepsi pengguna serta memberikan wawasan yang bermanfaat bagi pengembang dalam meningkatkan kualitas aplikasi Roblox.

## 2. TINJAUAN PUSTAKA

### 2.1. Text Mining

*Text mining* merupakan teknik komputasi yang digunakan untuk mengekstraksi informasi dan pola dari data teks yang tidak terstruktur [15]. Proses ini menggabungkan berbagai bidang seperti *Natural Language Processing* (NLP), *machine learning*, dan *information retrieval* untuk mengubah data teks mentah menjadi informasi yang dapat dianalisis [16]. Dalam penelitian terkait ulasan pengguna aplikasi, *text mining* sering dimanfaatkan untuk memahami opini pengguna secara otomatis dalam jumlah besar.

### 2.2. Analisis Sentimen

Analisis sentimen merupakan metode komputasi yang digunakan untuk mengidentifikasi dan mengklasifikasikan polaritas opini dalam teks ke dalam kategori positif, negatif, atau netral [17]. Teknik ini banyak dimanfaatkan untuk menangkap *voice of customer* dari ulasan pengguna aplikasi sehingga pengembang dapat memahami tingkat kepuasan serta permasalahan yang dihadapi pengguna [18].

### 2.3. Representasi Fitur Teks

Agar data teks dapat diproses oleh algoritma *machine learning*, teks perlu diubah ke dalam bentuk numerik melalui proses ekstraksi fitur. Salah satu metode yang umum digunakan adalah *Bag of Words* (BoW), yaitu teknik representasi dokumen berdasarkan frekuensi

kemunculan kata dalam dokumen tanpa memperhatikan urutan kata [19]. Metode ini banyak digunakan pada pemodelan topik karena mampu merepresentasikan distribusi kata dalam dokumen secara sederhana.

Selain BoW, metode lain yang sering digunakan adalah *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF memberikan bobot pada setiap kata berdasarkan tingkat kemunculan kata dalam dokumen dan tingkat kelangkaannya dalam seluruh korpus dokumen [20]. Representasi ini mampu menyoroti kata-kata yang lebih informatif dalam dokumen sehingga sering digunakan sebagai fitur input pada algoritma klasifikasi teks seperti *Support Vector Machine*.

### 2.4. Latent Dirichlet Allocation (LDA)

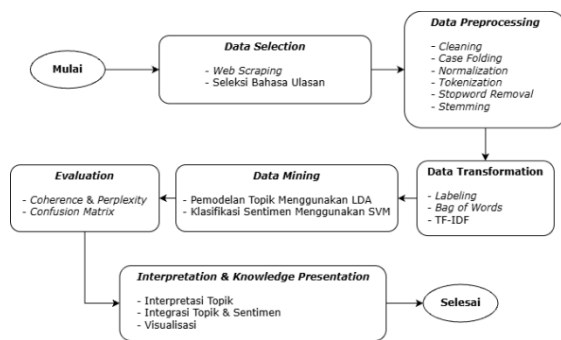
*Latent Dirichlet Allocation* (LDA) merupakan salah satu algoritma pemodelan topik yang banyak digunakan untuk mengidentifikasi tema tersembunyi dalam kumpulan dokumen teks. LDA bekerja dengan mengasumsikan bahwa setiap dokumen merupakan kombinasi dari beberapa topik, sedangkan setiap topik direpresentasikan sebagai distribusi probabilitas atas kata-kata tertentu [21]. Dengan pendekatan ini, LDA mampu mengidentifikasi kelompok kata yang sering muncul bersama dan membentuk topik tertentu dalam korpus teks.

### 2.5. Support Vector Machine (SVM)

*Support Vector Machine* (SVM) merupakan algoritma *supervised learning* yang banyak digunakan dalam tugas klasifikasi teks. Algoritma ini bekerja dengan mencari *hyperplane* optimal yang mampu memisahkan data dari dua kelas atau lebih dengan margin maksimum [22]. Dengan pendekatan ini, SVM memiliki kemampuan generalisasi yang baik terhadap data baru.

## 3. METODE PENELITIAN

Penelitian ini akan mengacu pada metodologi *Knowledge Discovery in Databases* (KDD). Adapun susunan tahapannya sebagai berikut:



Gambar 1 Rancangan Alur Penelitian

### 3.1. Data Selection

Objek penelitian dalam studi ini adalah aplikasi permainan Roblox yang tersedia pada platform Google Play Store. Data yang digunakan berupa data sekunder berupa ulasan pengguna yang diperoleh melalui proses web scraping menggunakan pustaka google-play-scraper.

Dataset yang dikumpulkan terdiri dari 40.089 ulasan pengguna yang dipublikasikan pada Google Play Store dalam rentang waktu 7 Oktober 2025 hingga 7 November 2025. Untuk menjaga konsistensi analisis linguistik, hanya ulasan berbahasa Indonesia yang dipertahankan dalam dataset penelitian.

### 3.2. Data Preprocessing

Data ulasan yang diperoleh dari Google Play Store merupakan data teks tidak terstruktur yang mengandung berbagai bentuk *noise*. Oleh karena itu dilakukan proses *data preprocessing* untuk meningkatkan kualitas data sebelum dianalisis. Tahapan pra-pemrosesan meliputi *cleaning*, *case folding*, *normalization*, *tokenization*, *stopword removal*, *stemming*.

### 3.3. Data Transformation

Setelah proses pra-pemrosesan, data teks diubah menjadi representasi numerik agar dapat diproses oleh algoritma *machine learning*. Penelitian ini menggunakan dua metode representasi fitur, yaitu *Bag of Words* (BoW) yang digunakan untuk merepresentasikan frekuensi kemunculan kata dalam dokumen yang kemudian digunakan sebagai input bagi algoritma pemodelan topik *Latent Dirichlet Allocation* (LDA). Yang kedua adalah *Term Frequency–Inverse Document Frequency* (TF-IDF) yang digunakan untuk menghasilkan bobot kata yang mencerminkan tingkat kepentingannya dalam dokumen. Representasi

ini digunakan sebagai fitur input untuk algoritma klasifikasi sentimen *Support Vector Machine* (SVM). Selain itu, proses pelabelan sentimen otomatis dilakukan berdasarkan rating pengguna pada Google Play Store, di mana ulasan dengan rating 1–2 dikategorikan sebagai sentimen negatif, sedangkan rating 4–5 dikategorikan sebagai sentimen positif.

### 3.4. Data Mining

Proses *data mining* dalam penelitian ini terdiri dari dua tahap utama, yaitu pemodelan topik dan klasifikasi sentimen.

Algoritma *Latent Dirichlet Allocation* (LDA) digunakan untuk mengidentifikasi topik utama yang muncul dalam ulasan pengguna. Model dilatih menggunakan representasi fitur *Bag of Words* dengan variasi jumlah topik (K) dari 2 hingga 10 topik.

Sedangkan proses klasifikasi sentimen dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dengan *kernel linear*. *Dataset* dibagi menjadi data latih dan data uji menggunakan tiga skenario pembagian data, yaitu 70:30, 80:20, dan 90:10. Untuk mengatasi ketidakseimbangan distribusi kelas sentimen, digunakan parameter *class\_weight = balanced* selama proses pelatihan model.

### 3.5. Evaluation

Evaluasi model dilakukan untuk mengukur kualitas pemodelan topik dan performa klasifikasi sentimen. Evaluasi model LDA dilakukan menggunakan metrik *coherence score* dan *perplexity*.

Sementara itu, evaluasi model SVM dilakukan menggunakan confusion matrix untuk menghitung metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik tersebut digunakan untuk menilai kemampuan model dalam mengklasifikasikan sentimen positif dan negatif pada ulasan pengguna.

### 3.6. Interpretation dan Knowledge Presentation

Tahap akhir penelitian adalah interpretasi hasil pemodelan serta penyajian pengetahuan yang dihasilkan. Kata-kata kunci yang dihasilkan dari model LDA dianalisis untuk menentukan label topik yang merepresentasikan tema utama ulasan pengguna. Selanjutnya, hasil klasifikasi sentimen dari model SVM diintegrasikan

dengan hasil pemodelan topik untuk memetakan distribusi sentimen pada setiap topik.

Hasil analisis kemudian divisualisasikan menggunakan *word cloud* untuk menampilkan kata-kata dominan dalam setiap topik serta diagram batang untuk menunjukkan distribusi sentimen positif dan negatif pada masing-masing topik.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Data Selection

Penelitian ini menganalisis ulasan pengguna aplikasi Roblox pada Google Play Store menggunakan pendekatan *topic-based sentiment analysis*. Dataset diperoleh melalui proses web scraping pada periode 7 Oktober 2025 hingga 7 November 2025 dengan total 43.875 ulasan mentah. Setelah dilakukan proses seleksi bahasa dan pembersihan data, jumlah dataset yang digunakan dalam penelitian menjadi 40.089 ulasan berbahasa Indonesia.

### 4.2. Data Preprocessing

Pada tahap ini menghasilkan data teks yang bersih dan siap digunakan pada tahapan selanjutnya. Ditemukan sebanyak 907 ulasan kosong yang kemudian dihapus dari *dataset*. Jumlah data sebelum melewati tahap preprocessing yaitu 42.951 ulasan dan jumlah data setelah melewati tahap preprocessing menjadi 42.044 ulasan.

### 4.3. Data Transformation

Setelah dilakukan pelabelan, ditemukan bahwa distribusi sentimen pada dataset menunjukkan dominasi sentimen positif dengan jumlah 33.189 ulasan (82,79%), sedangkan 6.900 ulasan (17,21%) dikategorikan sebagai sentimen negatif. Ketidakseimbangan distribusi kelas ini kemudian ditangani melalui penerapan parameter *class weight* pada model *Support Vector Machine* (SVM) agar model mampu mengenali kedua kelas secara lebih proporsional.

Kemudian dilakukan ekstraksi fitur *Bag of Words* yang menghasilkan matriks berdimensi ( $40.089 \times 1.000$ ), yang selanjutnya digunakan sebagai input pada proses pemodelan topik menggunakan algoritma LDA. Sedangkan untuk *input* SVM dilakukan TF-IDF yang

menghasilkan matriks dua dimensi berukuran ( $40.089 \times 5.000$ ).

## 4.4. Data Mining

### 4.4.1. Pemodelan Topik Menggunakan LDA

Pemodelan topik dilakukan menggunakan *Latent Dirichlet Allocation* (LDA) dengan variasi jumlah topik dari  $K=2$  hingga  $K=10$ . Hasil luaran dari tahapan ini adalah terbentuknya sembilan kandidat model topik LDA yang berbeda. Seluruh kandidat model ini kemudian diteruskan ke tahap evaluasi untuk diukur kualitas distribusinya guna menentukan satu model dengan jumlah topik ( $K$ ) yang paling optimal.

### 4.4.2. Klasifikasi Sentimen Menggunakan SVM

Bersamaan dengan pemodelan topik, sistem juga mengeksekusi algoritma *Support Vector Machine* (SVM) dengan *kernel linear*. Proses pelatihan diawali dengan pembagian dataset menggunakan fungsi *train\_test\_split* ke dalam tiga skenario rasio distribusi data, yaitu 70:30, 80:20, dan 90:10. Rincian hasil jumlah data latih dan data uji pada masing-masing skenario disajikan pada Tabel 1

**Tabel 1** Distribusi Pembagian *Dataset* untuk Skenario SVM

Skenario Rasio	Jumlah Data Latih	Jumlah Data Uji	Total Data
70 : 30	28.062 ulasan	12.027 ulasan	40.089 ulasan
80 : 20	32.071 ulasan	8.018 ulasan	40.089 ulasan
90 : 10	36.080 ulasan	4.009 ulasan	40.089 ulasan

## 4.5. Evaluation

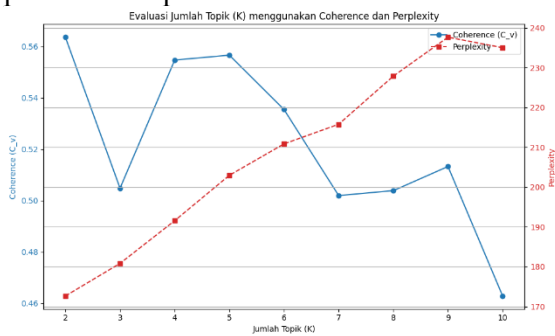
### 4.5.1. Coherence & Perplexity

Evaluasi model LDA dilakukan menggunakan metrik *coherence score* dan *perplexity* untuk menentukan konfigurasi topik yang paling optimal. Hasil evaluasi untuk rentang jumlah topik  $K=2$  hingga  $K=10$  disajikan pada Tabel 2

**Tabel 2** Hasil Evaluasi Performa Model LDA

No.	Jumlah Topik (K)	Coherence Score	Perplexity
1	2	0.563619	172.653802
2	3	0.504822	180.844871
3	4	0.554710	191.576468
4	5	0.556636	202.973413
5	6	0.535453	210.870920
6	7	0.501963	215.803261
7	8	0.503875	227.893453
8	9	0.513273	237.675435
9	10	0.462858	235.031460

Berikut adalah grafik hasil evaluasi pemodelan topik.



**Gambar 2** Grafik Hasil Evaluasi LDA

Hasil evaluasi menunjukkan bahwa model dengan empat topik (K=4) memberikan keseimbangan terbaik antara kualitas koherensi dan stabilitas model dengan coherence score sebesar 0,5547 dan perplexity sebesar 191,5.

**4.5.2. Confusion Matrix**

Setelah dilakukan pengujian model SVM menggunakan tiga skenario pembagian data latih dan data uji, dilakukan proses evaluasi model menggunakan confusion matrix. Adapun hasil performa model disajikan pada tabel berikut.

**Tabel 3** Hasil Evaluasi Performa Model SVM

Rasio Data	Akurasi	Presisi	Recall	F1-score
70:30	85,22%	89%	85%	86%
80:20	84,67%	89%	85%	86%
90:10	83,34%	89%	83%	85%

Berdasarkan hasil tersebut, model dengan rasio 70:30 memberikan performa terbaik dengan akurasi sebesar 85,22%.

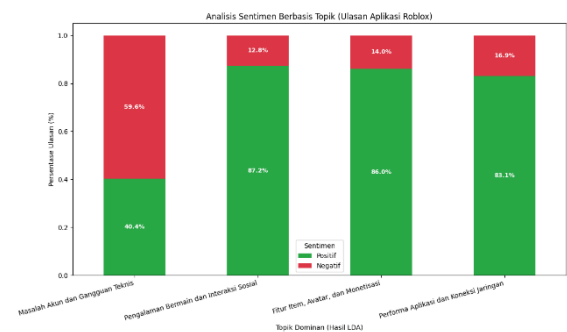
**4.6. Interpretation dan Knowledge Presentation**

Berdasarkan hasil pemodelan LDA dengan K=4, diperoleh empat topik utama yang merepresentasikan isu yang paling sering dibahas oleh pengguna Roblox, yang dapat dilihat pada tabel berikut:

**Tabel 4** Hasil Interpretasi Kata Dominan per Topik

Topik	Kata Kunci Dominan	Interpretasi
1	bug, akun, masuk, hilang	Masalah akun dan gangguan teknis
2	seru, teman, map, mabar	Pengalaman bermain dan interaksi sosial
3	item, beli, baju, avatar	Fitur item dan monetisasi
4	lambat, jaringan, wifi	Performa aplikasi dan koneksi jaringan

Selanjutnya dilakukan integrasi antara pemodelan topik LDA dan klasifikasi sentimen SVM yang menghasilkan pemahaman yang lebih komprehensif mengenai persepsi pengguna terhadap aplikasi Roblox.



**Gambar 3** Distribusi Perbandingan Sentimen per Topik

Analisis distribusi sentimen menunjukkan bahwa topik mengenai pengalaman bermain dan interaksi sosial didominasi oleh sentimen positif, yang menunjukkan bahwa aspek sosial seperti bermain bersama (mabar) dan variasi permainan menjadi faktor utama yang meningkatkan kepuasan pengguna.

Sebaliknya, topik mengenai masalah akun dan gangguan teknis menunjukkan proporsi sentimen negatif tertinggi, yaitu sebesar 59,6%. Temuan ini mengindikasikan bahwa permasalahan autentikasi akun, bug aplikasi, dan gangguan sistem masih menjadi sumber utama keluhan pengguna.

Temuan penelitian ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa laporan bug dan kendala teknis sering menjadi topik utama dalam ulasan pengguna aplikasi digital [14]. Selain itu, hasil penelitian ini juga mendukung temuan Kustiyaningsih dan Permana (2024) yang menunjukkan bahwa integrasi algoritma LDA dan SVM mampu mengidentifikasi aspek penting dalam ulasan pengguna aplikasi secara efektif [11].

## 5. KESIMPULAN

Berdasarkan hasil penelitian analisis sentimen berbasis topik pada ulasan pengguna aplikasi Roblox di Google Play Store menggunakan integrasi metode *Latent Dirichlet Allocation* (LDA) dan *Support Vector Machine* (SVM), diperoleh beberapa kesimpulan sebagai berikut:

- a. Pemodelan topik menggunakan LDA berhasil mengidentifikasi empat topik utama dalam ulasan pengguna dengan konfigurasi optimal  $K=4$ , yang menghasilkan nilai *coherence score* sebesar 0,5547 dan *perplexity* sebesar 191,57. Topik yang terbentuk meliputi masalah akun dan gangguan teknis, pengalaman bermain dan interaksi sosial, fitur item dan monetisasi, serta performa aplikasi dan koneksi jaringan.
- b. Model klasifikasi sentimen menggunakan SVM dengan *kernel linear* menunjukkan performa terbaik pada skenario pembagian data 70:30, dengan akurasi sebesar 85,22%, presisi 89%, *recall* 85%, dan *F1-score* sebesar 86%, yang menunjukkan bahwa model mampu mengklasifikasikan sentimen ulasan pengguna secara efektif.
- c. Integrasi hasil LDA dan SVM menunjukkan bahwa aspek pengalaman bermain dan interaksi sosial menjadi faktor utama kepuasan pengguna dengan dominasi sentimen positif, sementara masalah akun dan gangguan teknis menjadi sumber keluhan terbesar dengan proporsi sentimen negatif mencapai 59,6%, sehingga memberikan wawasan strategis bagi pengembang dalam memprioritaskan perbaikan aplikasi.
- d. Penelitian selanjutnya dapat mengeksplorasi penggunaan metode berbasis *deep learning* seperti IndoBERT untuk klasifikasi sentimen serta BERTopic

untuk pemodelan topik guna membandingkan kualitas hasil analisis. Selain itu, pengembangan sistem pelabelan yang melibatkan anotasi manusia serta penambahan kelas sentimen netral dapat meningkatkan akurasi analisis. Penelitian di masa depan juga dapat mempertimbangkan penerapan teknik deteksi sarkasme dan pemrosesan ulasan multi-bahasa untuk menghasilkan analisis opini pengguna yang lebih komprehensif.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

## DAFTAR PUSTAKA

- [1] T. Fajriah and E. R. Ningsih, "Pengaruh Teknologi Komunikasi Terhadap Interaksi Sosial di Era Digital," *Merdeka Indonesia Journal International* 4, vol. 4, no. 1, Jun. 2024.
- [2] Y. Liu, H. Duan, and W. Cai, "User-Generated Content and Editors in Games: A Comprehensive Survey," *IEEE Transaction On Games*, Jul. 2024, [Online]. Available: <http://arxiv.org/abs/2412.13743>
- [3] Y. Kou and X. Gui, "Harmful Design in the Metaverse and How to Mitigate it: A Case Study of User-Generated Virtual Worlds on Roblox," *Association for Computing Machinery (ACM)*, Jul. 2023, pp. 175–188. doi: 10.1145/3563657.3595960.
- [4] Y. joo Kang, U. jun Lee, and S. Lee, "Who Makes Popular Content? Information Cues from Content Creators for Users' game Choice: Focusing on User-Created Content Platform 'Roblox,'" *Entertain. Comput.*, vol. 50, p. 100697, May 2024, doi: 10.1016/J.ENTCOM.2024.100697.
- [5] Roblox Corporation, "Roblox Reports Fourth Quarter and Full Year 2024 Financial Results," Roblox Corporation – Investor Relations. Accessed: Nov. 10, 2025. [Online]. Available: <https://ir.roblox.com/news/news-details/2025/Roblox-Reports-Fourth-Quarter-and-Full-Year-2024-Financial-Results/default.aspx>
- [6] T. Liu *et al.*, "RoseMatcher: Identifying the impact of user reviews on app updates," *Inf. Softw. Technol.*, vol. 161, p. 107261, Sep. 2023, doi: 10.1016/J.INFSOF.2023.107261.
- [7] G. Brauwiers and F. Frasincar, "A Survey on Aspect-Based Sentiment Classification," *ACM*

- Comput. Surv.*, vol. 55, no. 4, Apr. 2023, doi: 10.1145/3503044.
- [8] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, "A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges," *IEEE Trans. on Knowl. and Data Eng.*, vol. 35, no. 11, pp. 11019–11038, Nov. 2023, doi: 10.1109/TKDE.2022.3230975.
- [9] E. Erniyati, P. Harsani, M. Mulyati, and L. D. Fahriza, "Topic Modeling LDA and SVM in Sentiment Analysis of Hotel Reviews," *Komputasi: Jurnal Ilmiah Ilmu Komputer dan Matematika*, vol. 20, no. 2, pp. 93–100, Jul. 2023, doi: 10.33751/komputasi.v20i2.7604.
- [10] Y. Afrianto Singgalen, "Topic modeling using LDA and performance evaluation of classification algorithm: k-NN, SVM, NBC, and DT," vol. 16, no. 3, pp. 143–157, Jun. 2024.
- [11] Y. Kustiyaningsih and Y. Permana, "Penggunaan Latent Dirichlet Allocation (LDA) dan Support-Vector Machine (SVM) Untuk Menganalisis Sentimen Berdasarkan Aspek Dalam Ulasan Aplikasi EdLink," *Teknika*, vol. 13, no. 1, pp. 127–136, Mar. 2024, doi: 10.34148/teknika.v13i1.746.
- [12] E. Sa'dul Asyhar, S. Hadi Wijoyo, and N. Y. Setiawan, "Analisis Sentimen dan Pemodelan Topik Terhadap Ulasan Aplikasi Jenius Menggunakan Metode Support Vector Machine dan Latent Dirichlet Allocation," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 9, Sep. 2024, [Online]. Available: <http://j->
- [13] P. A. Prastyo, B. Berlilana, and I. Tahyudin, "Analisis Sentimen dan Pemodelan Topik pada Ulasan Pengguna Aplikasi myIM3 Menggunakan Support Vector Machine dan Latent Dirichlet Allocation," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, pp. 1618–1626, Dec. 2024, doi: 10.47065/bits.v6i3.6268.
- [14] Z. Jiang, V. Liu, and M. Erne, "Examining the Usefulness of Customer Reviews for Mobile Applications: The Role of Developer Responsiveness," *Journal of Database Management*, vol. 35, no. 1, pp. 1–23, 2024, doi: 10.4018/JDM.343543.
- [15] H. Hassani, C. Beneki, S. Unger, M. T. Mazinani, and M. R. Yeganegi, "Text mining in big data analytics," *Big Data and Cognitive Computing*, vol. 4, no. 1, pp. 1–34, Mar. 2020, doi: 10.3390/bdcc4010001.
- [16] H. Yan, M. Ma, Y. Wu, H. Fan, and C. Dong, "Overview and analysis of the text mining applications in the construction industry," *Heliyon*, vol. 8, no. 12, p. e12088, Dec. 2022, doi: 10.1016/J.HELIYON.2022.E12088.
- [17] A. Nurian and B. Nurina Sari, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 3, pp. 2830–7062, Sep. 2023, doi: 10.23960/jitet.v11i3%20s1.3348.
- [18] F. Noor Hasan and M. Dwijayanti, "Analisis Sentimen Pada Ulasan Pelanggan Menggunakan Metode Naïve Bayes Classifier (Studi Kasus: Grab Indonesia)," *Jurnal Linguistik Komputasional*, vol. 4, no. 2, pp. 93–99, Sep. 2021, doi: 10.22236/teknoka.v6i1.441.
- [19] P. Cichosz, "BAG OF WORDS AND EMBEDDING TEXT REPRESENTATION METHODS FOR MEDICAL ARTICLE CLASSIFICATION," *International Journal of Applied Mathematics and Computer Science*, vol. 33, no. 4, pp. 603–621, Dec. 2023, doi: 10.34768/amcs-2023-0043.
- [20] M. Das, S. Kamalanathan, and P. Alphonse, "A Comparative Study on TF-IDF feature Weighting Method and its Analysis using Unstructured Dataset," *arXiv preprint arXiv:2308.04037*, 2020.
- [21] L. T. Nguyen *et al.*, "Evaluating the Performance of Topic Modeling Techniques for Bibliometric Analysis Research: An LDA-based Approach," *HighTech and Innovation Journal*, vol. 5, no. 2, pp. 312–330, Jun. 2024, doi: 10.28991/HIJ-2024-05-02-07.
- [22] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, "An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review," *Information (Switzerland)*, vol. 15, no. 4, Apr. 2024, doi: 10.3390/info15040235.