

Analisis Tingkat Kepuasan Pada Ulasan Game *Wuthering Waves* Menggunakan Algoritma LSTM (*Long Short-Term Memory*)

Analysis of User Satisfaction Levels in Wuthering Waves Game Reviews Using the Long Short-Term Memory (LSTM) Algorithm

Galva Al Godzali*¹, Muhammad Fatchan², Asep Supriyanto³

^{1,2,3}Program Studi Teknik Informatika, Universitas Pelita Bangsa, Indonesia

Email: ¹galva.312210356@mhs.pelitabangsa.ac.id

Article	Received	Revised	Accepted	Published:
Info:	03 Juni 2026	06 Juni 2026	21 Juni 2026	24 Juni 2026

Abstrak

Wuthering Waves merupakan permainan bergenre *Action Role-Playing Game (RPG)* dunia terbuka yang memperoleh banyak ulasan di *Google Play Store*. Volume ulasan yang besar membuat analisis manual menjadi tidak efektif dan subjektif. Penelitian terdahulu umumnya menggunakan pendekatan klasifikasi dua kelas sentimen dengan algoritma konvensional. Penelitian ini bertujuan mengklasifikasikan ulasan pemain game *Wuthering Waves* dari *Google Play Store* ke dalam tiga kategori sentimen (positif, netral, dan negatif) untuk menganalisis tingkat kepuasan pemain serta mengidentifikasi pola keluhan utama secara sistematis. Metode penelitian yang digunakan adalah pendekatan data mining berbasis *deep learning* dengan menerapkan algoritma *Long Short-Term Memory (LSTM)*. Data penelitian diperoleh melalui proses *crawling* pada periode 1 September 2025 hingga 14 Januari 2026. Dari 3.600 ulasan yang dikumpulkan, proses *preprocessing* menghasilkan 3.205 data unik. Pengujian dan pembagian dataset dilakukan menggunakan metode *Stratified K-Fold Cross Validation*. Evaluasi model menunjukkan bahwa algoritma *LSTM* mencapai rata-rata akurasi sebesar 83,15%, presisi 81,77%, *recall* 78,95%, dan *F1-score* 79,06%, yang menunjukkan performa stabil dalam klasifikasi sentimen multikelas. Hasil visualisasi *WordCloud* mengungkap pola keluhan pengguna, terutama terkait kendala optimasi sistem seperti *lag* dan *force close* serta belum tersedianya lokalisasi Bahasa Indonesia. Penelitian ini membuktikan bahwa algoritma *LSTM* efektif dalam mengenali konteks dan hubungan jangka panjang antar kata. Hasil penelitian memberikan informasi bagi pengembang sebagai bahan evaluasi peningkatan kualitas permainan.

Kata Kunci: Analisis Sentimen, *Long Short-Term Memory (LSTM)*, *Wuthering Waves*, *Text Mining*, *Kepuasan Pengguna*.

Abstract

Wuthering Waves is an open-world *Action Role-Playing Game (RPG)* that has received numerous reviews on the *Google Play Store*. The large volume of reviews makes manual analysis ineffective and subjective. Previous studies generally applied two-class sentiment classification approaches using conventional algorithms. This study aims to classify player reviews of the *Wuthering Waves* game from the *Google Play Store* into three sentiment categories (positive, neutral, and negative) to analyze player satisfaction levels and systematically identify major complaint patterns. The research method employs a data mining approach based on *deep learning* by implementing the *Long Short-Term Memory (LSTM)* algorithm. The research data were collected through a *crawling* process from September 1, 2025, to January 14, 2026. From 3,600 collected reviews, the *preprocessing* stage resulted in 3,205 unique data

entries. Dataset testing and splitting were conducted using the Stratified K-Fold Cross Validation method. Model evaluation showed that the LSTM algorithm achieved an average accuracy of 83.15%, precision of 81.77%, recall of 78.95%, and F1-score of 79.06%, indicating stable performance in multiclass sentiment classification. Word Cloud visualization results revealed specific user complaint patterns, particularly related to system optimization issues such as lag and force close, as well as the unavailability of Indonesian language localization. This study demonstrates that the LSTM algorithm is effective in recognizing context and long-term relationships between words. The results provide valuable information for developers as an evaluation reference to improve game quality.

Keywords : Long Short-Term Memory (LSTM), Wuthering Waves, Text Mining, User Satisfaction

This is an open access article under the CC BY-SA license.



1. PENDAHULUAN

Perkembangan teknologi digital telah menghasilkan data dalam jumlah besar yang berasal dari berbagai aktivitas pengguna di internet, termasuk ulasan pada aplikasi dan permainan digital. Ulasan pengguna merupakan sumber informasi yang penting karena mencerminkan pengalaman, persepsi, dan tingkat kepuasan terhadap suatu produk atau layanan. Data tersebut dapat dimanfaatkan melalui teknik data mining untuk menemukan pola dan informasi yang berguna dalam pengambilan keputusan [1].

Industri game merupakan salah satu sektor industri digital dengan pertumbuhan tercepat di dunia, ditandai oleh meningkatnya jumlah pemain, pendapatan pasar, dan aktivitas interaksi pengguna pada berbagai platform game [2]. Dan salah satu sektor yang mengalami pertumbuhan pesat dalam beberapa tahun terakhir yang banyak diminati adalah *open world action role-playing game (RPG)* yang menawarkan kebebasan eksplorasi, sistem pertarungan yang dinamis, dan pengalaman bermain yang imersif [3]. *Wuthering Waves* merupakan salah satu game bergenre *open world action RPG* yang dikembangkan oleh *Kuro Games* dan tersedia pada platform *Android*, *iOS*, serta *PC*. Sejak dirilis, game ini memperoleh perhatian yang besar dari pemain dan menghasilkan ribuan ulasan pada *Google Play Store*. Sebagai platform distribusi aplikasi resmi milik *Google*, *Google Play Store* memungkinkan pengguna memberikan penilaian dan ulasan terhadap aplikasi maupun permainan yang digunakan. Banyaknya ulasan yang tersedia membuat proses analisis secara manual menjadi kurang efektif untuk mengetahui tingkat kepuasan pemain secara menyeluruh [4]. Salah satu metode *deep learning* yang mampu memahami hubungan antar kata dalam urutan kalimat sehingga efektif digunakan pada analisis sentimen berbasis teks. Penelitian terdahulu menunjukkan bahwa LSTM mampu memberikan hasil yang baik dalam analisis sentimen pada ulasan aplikasi maupun game [5][6].

Banyaknya ulasan yang tersedia memberikan peluang bagi pengembang untuk memahami tingkat kepuasan pemain. Namun, volume data yang besar serta beragamnya opini yang disampaikan membuat proses analisis secara manual menjadi kurang efektif [7]. Ulasan pemain tidak hanya berisi apresiasi terhadap aspek tertentu seperti *grafis*, *gameplay*, dan desain karakter, tetapi juga memuat kritik terkait bug, optimasi, serta performa permainan. Oleh karena itu, diperlukan metode otomatis yang mampu mengelompokkan ulasan berdasarkan tingkat kepuasan pemain secara objektif dan terukur [8]. Selain itu, LSTM memiliki kemampuan untuk mempertahankan informasi jangka panjang sehingga lebih efektif dalam menangani data teks dibandingkan metode konvensional yang hanya mempertimbangkan frekuensi kata [9].

Sejumlah penelitian terdahulu menunjukkan bahwa algoritma tradisional seperti *Naive Bayes* memiliki kinerja yang cukup baik dalam analisis sentimen pada ranah permainan digital, namun performanya sangat bergantung pada karakteristik data dan metode yang digunakan. Pada data ulasan permainan *Wuthering Waves* di *Google Play Store*, pendekatan *Naive Bayes* yang dikombinasikan dengan pembobotan TF-IDF dan teknik *undersampling* mampu menghasilkan akurasi sebesar 78%. Namun, penelitian tersebut masih terbatas pada klasifikasi dua kelas sentimen (positif dan negatif)

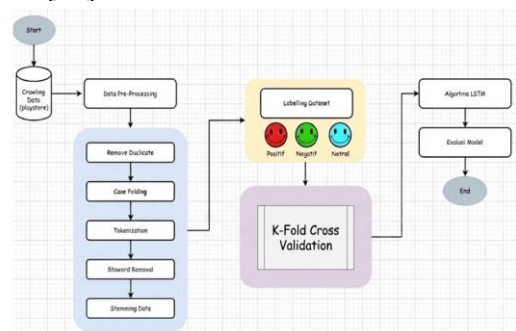
sehingga mengabaikan ulasan pengguna yang bersifat netral atau bimbang, serta menggunakan algoritma *machine learning* konvensional yang mengasumsikan setiap kata bersifat independen dan mengabaikan urutan kata dalam kalimat. Semenara algoritma LSTM terbukti lebih baik dalam menangkap pola data yang kompleks dan mempertahankan hubungan jangka panjang antar kata. Akan tetapi, kajian-kajian tersebut masih terbatas pada kondisi data tertentu dan belum mengeksplorasi pengaruh distribusi kelas sentimen yang lebih rinci secara mendalam. Selain itu, penerapan klasifikasi tiga kelas terbukti dapat meningkatkan pemetaan tingkat kepuasan secara lebih komprehensif, namun implementasinya belum banyak dikaji pada konteks ulasan permainan berbahasa Indonesia yang kompleks. Sementara itu, pada dataset ulasan Google Play Store berbahasa Indonesia, tantangan utama terletak pada penggunaan bahasa informal, tingginya variasi kosakata slang, serta adanya indikasi *overfitting* yang memengaruhi hasil analisis model. [10].

Berdasarkan kajian literatur di atas, terdapat beberapa research gap yang masih ditemukan pada penelitian terdahulu. Pertama, sebagian besar penelitian analisis sentimen ulasan permainan masih berfokus pada klasifikasi dua kelas sentimen (positif dan negatif), sehingga belum mampu menggambarkan tingkat kepuasan pengguna secara lebih rinci melalui keberadaan sentimen netral. Kedua, penelitian sebelumnya umumnya menggunakan algoritma *machine learning* konvensional yang memiliki keterbatasan dalam menangkap hubungan kontekstual dan urutan kata pada teks ulasan. Ketiga, penelitian mengenai analisis sentimen multi-kelas pada ulasan permainan berbahasa Indonesia, khususnya pada permainan *Wuthering Waves*, masih relatif terbatas sehingga belum memberikan gambaran yang komprehensif mengenai tingkat kepuasan pemain.

Berdasarkan research gap tersebut, penelitian ini bertujuan untuk menganalisis tingkat kepuasan pemain berdasarkan ulasan permainan *Wuthering Waves* di Google Play Store menggunakan algoritma Long Short-Term Memory (LSTM). Novelty penelitian ini terletak pada penerapan klasifikasi sentimen tiga kelas, yaitu positif, netral, dan negatif, pada ulasan permainan *Wuthering Waves* berbahasa Indonesia menggunakan algoritma LSTM yang dievaluasi melalui Stratified K-Fold Cross Validation. Pendekatan ini memungkinkan pemetaan tingkat kepuasan pemain yang lebih komprehensif dibandingkan pendekatan dua kelas yang banyak digunakan pada penelitian sebelumnya. Selain itu, penggunaan LSTM dipilih karena kemampuannya dalam menangkap hubungan kontekstual antar kata sehingga diharapkan mampu meningkatkan kualitas klasifikasi sentimen pada data ulasan yang memiliki karakteristik bahasa informal dan beragam [11]. Kontribusi penelitian ini adalah menyediakan model analisis sentimen multi-kelas yang mampu mengidentifikasi tingkat kepuasan pemain secara lebih rinci serta memberikan informasi yang dapat digunakan oleh pengembang permainan sebagai bahan evaluasi dalam meningkatkan kualitas layanan dan pengalaman bermain pengguna.

2. METODE PENELITIAN

Penelitian ini menggunakan metode data mining dengan pendekatan analisis sentimen berbasis deep learning menggunakan algoritma *Long Short-Term Memory (LSTM)*. Metode ini dipilih karena LSTM memiliki kemampuan untuk mempelajari hubungan antar kata dalam urutan kalimat serta mempertahankan informasi jangka panjang pada data teks, sehingga efektif digunakan untuk klasifikasi sentimen pada ulasan pengguna [12].



Gambar 1. Alur Penelitian

Pada alur penelitian ini bisa dilihat melalui alur penelitian pada *Gambar 1.* yang dilakukan melalui beberapa tahapan, yaitu pengumpulan data, *preprocessing*, pelabelan data, pembagian dataset, pelatihan model, dan evaluasi model. Data penelitian berupa ulasan pengguna game *Wuthering Waves* yang diperoleh dari Google Play Store menggunakan teknik *web crawling* dengan Google Play Scraper [13]. Data yang terkumpul kemudian melalui tahap *preprocessing* yang meliputi penghapusan data duplikat (*remove duplicate*), *case folding*, *tokenization*, *stopword removal*, dan *stemming*. Tahap ini bertujuan untuk mengurangi *noise* serta meningkatkan kualitas data sebelum digunakan pada proses klasifikasi [14]. Setelah proses pelabelan sentimen ke dalam tiga kategori (positif, netral, dan negatif), Dataset yang telah berlabel ini kemudian diproses menggunakan metode *K-Fold Cross Validation* untuk pembagian dan validasi data secara terstruktur. Klasifikasi sentimen kemudian dilakukan menggunakan algoritma LSTM, di mana model dilatih pada data training dan diuji pada data testing untuk menilai kemampuannya dalam mengenali sentimen ulasan. Tahap akhir berupa evaluasi model untuk mengukur kinerja klasifikasi sentimen pada ulasan pengguna game *Wuthering Waves* [15].

2.1 Pengumpulan Dataset

Data penelitian diperoleh langsung dari ulasan pengguna game *Wuthering Waves* di *Google Play Store* menggunakan *Google Colab* dan *library google_play_scraper*. Pengambilan data dilakukan berdasarkan App ID game, sehingga diperoleh informasi berupa teks ulasan, rating, tanggal ulasan, dan identitas pengguna yang bersifat anonim. Seluruh data kemudian disimpan dalam format CSV untuk memudahkan proses pengolahan dan analisis sentimen pada tahap berikutnya. [16].

2.2 Preprocessing

Tahap *preprocessing* dilakukan untuk membersihkan data ulasan dari *noise* guna meningkatkan kualitas input sebelum proses klasifikasi [17]. Langkah-langkah yang diterapkan pada dataset ulasan *Wuthering Waves* meliputi: (a) *Cleaning* Menghapus elemen non-teks pada ulasan Google Play Store seperti emoji (misalnya 🙏🏻), angka (seperti jumlah gacha '150'), tanda baca, tautan, hashtag, dan mention (@) [18]. (b) *Case folding* adalah tahapan dalam proses *preprocessing* teks yang bertujuan menyeragamkan bentuk huruf pada komentar pengguna di ulasan Google Play Store [19]. (c) *Tokenization* adalah tahap dalam proses *preprocessing* teks yang bertujuan untuk memecah sebuah kalimat atau dokumen menjadi bagian-bagian kecil yang disebut token [20]. (d) *Stopword removal* adalah tahap dalam proses *preprocessing* teks yang bertujuan untuk menghapus kata-kata yang dianggap tidak memiliki kontribusi penting terhadap makna kalimat [21]. (e) *Stemming* adalah tahap dalam proses *preprocessing* teks yang bertujuan untuk mengubah setiap kata menjadi bentuk dasarnya [22].

2.3 Labeling Data

Data hasil *preprocessing* yang belum memiliki kategori kelas memerlukan pelabelan (*labeling*) sebelum digunakan dalam pelatihan model. Pelabelan dilakukan secara otomatis menggunakan skrip Python berdasarkan nilai rating 1–5 dari Google Play Store yang mencerminkan tingkat kepuasan pengguna. Penggunaan rating bintang sebagai label sentimen ini didasarkan pada asumsi akademik bahwa nilai rating berbanding lurus dengan polaritas emosi dan kepuasan pengguna, di mana nilai ekstrem ujung mencerminkan sentimen yang jelas [23]. Data ulasan dipetakan ke dalam tiga kategori sentimen rating 4 dan 5 sebagai positif, rating 3 sebagai netral, serta rating 1 dan 2 sebagai negatif. Contohnya bisa dilihat pada tabel 1.

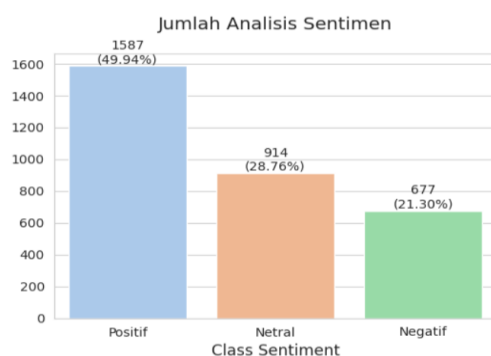
Tabel 1. Pelabelan data

Komentar	Score	Label
Tolong tambahkan subtitle indonesia	3	Netral
Sedang menikmati permainan ini	5	Positif
Optimalisasi mobile masih sampah, mohon dibenerin. Habis update baru force close terus.	1	Negatif

Metode otomatis ini memiliki keterbatasan laten berupa inkonsistensi antara nilai rating dan makna kontekstual teks, pengguna sering kali memberikan rating 3 namun menulis komentar sepenuhnya negatif, atau sebaliknya [24]. Untuk memvalidasi akurasi metode ini, dilakukan pengujian sampel manual terhadap 300 sampel data secara acak. Hasil pengujian menunjukkan tingkat kesesuaian sebesar 84.3% antara rating dan isi teks. Ketidaksesuaian sebesar 15.7% sebagian besar ditemukan pada rating 3 (netral) karena adanya muatan kritik spesifik di dalam teks ulasan serta rating ekstrem yang salah input oleh pengguna. Meski demikian, tingkat kesalahan ini masih berada dalam batas toleransi yang dapat diterima untuk proses pelatihan model.

2.4 Distribusi Kelas dataset

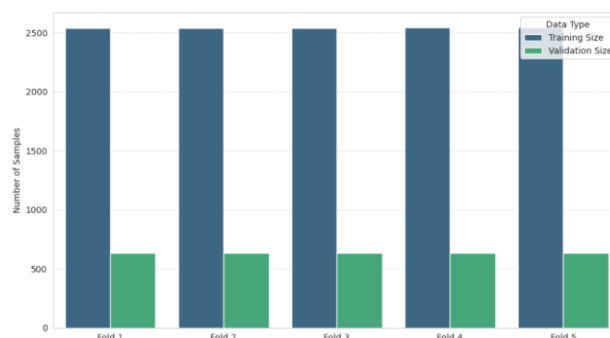
Pada Penelitian ini diperoleh distribusi dataset yang terdiri atas 1.587 data kategori positif (49,94%), 914 data kategori netral (28,76%), dan 677 data kategori negatif (21,30%). Hasil tersebut menunjukkan bahwa sebagian besar pemain memberikan ulasan yang mencerminkan tingkat kepuasan terhadap permainan Wuthering Waves. Distribusi dataset bisa dilihat pada gambar dibawah ini.



Gambar 2. Distribusi Kelas Dataset

2.5 Split Dataset

Penelitian ini menggunakan Stratified 5-Fold Cross Validation untuk mempertahankan proporsi kelas positif, netral, dan negatif pada setiap fold. Pada penelitian ini, distribusi jumlah data pada setiap fold dalam proses K-Fold Cross Validation. Setiap iterasi menggunakan sekitar 80% data sebagai data pelatihan (2543 sampel) dan 20% data sebagai data validasi (635 sampel). Distribusi yang konsisten pada setiap fold menunjukkan bahwa pembagian data dilakukan secara seimbang sehingga proses evaluasi model menjadi lebih representatif dan mengurangi bias akibat pemilihan data pelatihan dan validasi [25].



Gambar 3. Pembagian data dengan K-fold Cross Validation

2.6 Representasi Teks Menggunakan Embedding

Sebelum masuk ke tahap pemodelan, data teks hasil preprocessing ditransformasikan ke dalam bentuk numerik menggunakan Tokenizer dari Keras. Proses tokenisasi ini memetakan setiap kata menjadi indeks bilangan bulat, dengan batas maksimum 20.000 kosakata (`num_words = 20.000`) serta menerapkan token Out-of-Vocabulary (OOV) untuk menangani kata baru di luar kamus pelatihan. Selanjutnya, dilakukan sequence padding dengan panjang maksimum 150 token guna menghasilkan ukuran input yang seragam. Karakteristik teks tersebut kemudian direpresentasikan ke dalam ruang

vektor kontinu melalui embedding layer berdimensi 128. Teknik word embedding ini dipilih agar hubungan semantik antar-kata dapat dipelajari secara lebih efektif dibandingkan metode berbasis frekuensi kata [26], sebelum akhirnya digunakan sebagai masukan (input) utama proses klasifikasi.

2.7. Konfigurasi Model LSTM

Model klasifikasi sentimen dibangun menggunakan arsitektur *Long Short-Term Memory*. Arsitektur model terdiri atas Embedding Layer berdimensi 128, LSTM Layer dengan 64 unit neuron, GlobalMaxPooling1D, Batch Normalization, Dense Layer dengan 64 neuron dan fungsi aktivasi ReLU, serta Dropout sebesar 0,3 untuk mengurangi risiko overfitting [27]. Pada lapisan keluaran digunakan fungsi aktivasi Softmax untuk mengklasifikasikan ulasan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Model dikompilasi menggunakan optimizer Adam dan fungsi loss sparse categorical crossentropy.

Tabel 2. Konfigurasi Model

Parameter	Nilai
Jumlah Kosakata Maksimum	20.000
Panjang Sequence Maksimum	150
Embedding Dimension	128
LSTM Layer	64
Dropout	0,3
Optimizer	Adam
Loss Function	Sparse Categorical Crossentropy
Output Activation	Softmax
Jumlah Kelas Sentiment	3

Berdasarkan pengaturan pada Tabel 2, model LSTM akan membaca teks ulasan secara urut satu arah untuk mengenali pola hubungan antar kata. Lapisan Dropout sengaja dipasang sebesar 0,3 agar model tidak hanya menghafal data latihan, sehingga kemampuan model dalam menebak data ulasan baru yang belum pernah dilihat sebelumnya tetap terjaga dengan baik.

2.8 Eksperimen Hyperparameter

Tabel 3. Eksperimen Hyperparameter

Eksperimen Embedding LSTM Dropout Akurasi				
1	64	32	0.2	78%
2	64	64	0.3	80%
3	128	64	0.3	83%
4	128	128	0.5	81%

Proses pencarian kombinasi parameter optimal dalam penelitian ini dilakukan menggunakan metode pencarian eksploratif terstruktur (*structured manual tuning*) melalui eksperimen komparatif. Penentuan nilai konfigurasi tidak dilakukan secara acak, melainkan melalui pengujian *trial-and-error* sistematis yang didasarkan pada karakteristik dataset teks informal. Eksperimen ini dirancang secara bertahap dengan mengisolasi dan memvariasikan komponen kunci model—meliputi dimensi *embedding*, jumlah unit tersembunyi LSTM, serta regulasi tingkat *dropout*—untuk mengamati perilaku konvergensi model serta meminimalkan nilai *loss* fungsi *sparse categorical crossentropy*. Langkah penataan parameter secara empiris ini bertujuan untuk menemukan arsitektur paling efisien yang mampu menghasilkan akurasi tinggi sekaligus mencegah generalisasi yang buruk (*overfitting*). Detail dari skenario eksperimen komparatif tersebut disajikan pada Tabel 3.

Berdasarkan hasil *tuning hyperparameter* tersebut, diperoleh bahwa konfigurasi terbaik adalah *embedding size* 128, jumlah unit LSTM 64, dan *dropout* 0.3 dengan capaian akurasi sebesar 83% (Eksperimen 3). Pemilihan dimensi *embedding* 128 terbukti optimal karena menyediakan ruang representasi vektor yang lebih kaya untuk menangkap hubungan semantik fitur teks dibanding ukuran 64. Sementara itu, penggunaan 64 unit LSTM dipilih sebagai titik tengah yang ideal. Konfigurasi yang

lebih rendah (32 unit) memicu terjadinya *underfitting*, sedangkan konfigurasi yang lebih tinggi (128 unit) justru menurunkan performa akibat kompleksitas model yang berlebihan. Terakhir, nilai *dropout* sebesar 0.3 dipilih karena mampu meregulasi model secara seimbang tanpa gejala *overfitting*. Sebaliknya, peningkatan *dropout* menjadi 0.5 pada Eksperimen 4 terbukti menurunkan akurasi menjadi 81%. Hal ini mengindikasikan bahwa tingkat regularisasi yang terlalu tinggi justru menghambat proses pembelajaran model karena mengeliminasi terlalu banyak informasi penting selama pelatihan [28].

2.9 Evaluasi Model

Evaluasi model dilakukan untuk mengukur performa algoritma LSTM dalam mengklasifikasikan sentimen ulasan pengguna game *Wuthering Waves*. Komponen dalam evaluasi model dibagi menjadi empat istilah, yaitu *TP* (*True Positive*), *TN* (*True Negative*), *FP* (*False Positive*), serta *FN* (*False Negative*) [29]. Berdasarkan komponen-komponen tersebut, evaluasi kinerja model dihitung menggunakan rumus sebagai berikut:

2.1.1 Accuracy

Nilai Akurasi (*Accuracy*) memberikan gambaran model tingkat efektivitas secara keseluruhan dibanding dengan seluruh data yang di uji. Akurasi Diuji dengan perhitungan sebagai berikut:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

2.1.2 Precision

Presisi adalah metrik yang mengukur seberapa baik model memprediksi kelas positif dengan benar dari semua prediksi positif yang dibuat. Rumus untuk menghitung precision adalah sebagai berikut:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

2.1.3 Recall

Sensitivitas (*recall*) adalah metrik evaluasi yang digunakan untuk mengukur seberapa banyak data positif yang berhasil ditemukan oleh model. Formula untuk menghitung recall adalah sebagai berikut:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

2.1.4 F1-Score

F1-Score merupakan metrik evaluasi yang digunakan untuk mengukur keseimbangan antara nilai precision dan recall.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

3. HASIL DAN PEMBAHASAN

3.1 Crawling Data (Pengumpulan Data)

Data yang digunakan dalam penelitian ini berupa ulasan pengguna game *Wuthering Waves* yang diperoleh dari *Google Play Store*. Pengumpulan data dilakukan menggunakan teknik web *crawling* dengan bantuan *Google Play Scraper* pada platform *Google Colab*. Proses pengambilan data dilakukan pada periode September 2025 hingga Januari 2026 dan menghasilkan sebanyak 3.600 ulasan pengguna yang disimpan dalam format CSV.

3.2 Preprocessing

Tahap selanjutnya adalah *preprocessing* data, yaitu proses pembersihan dan pengolahan data teks sebelum digunakan dalam proses analisis sentimen. Data ulasan yang diperoleh dari *Google Play Store* umumnya masih mengandung berbagai bentuk *noise*, seperti simbol, angka, kata tidak baku, maupun karakter yang tidak relevan.

Tabel 4. Hasil Praprocessing

No.	Tahap Preprocessing	Hasil Transformasi Kalimat
1.	Data Awal	wuwa.. kenapa hard banget gachaa nya udah hampir 150 kali gacha lynae ga pulang ² gw belom gacha signature nya pliss lah kuroo. akun lama ini 🙄🙏
2.	Cleaning	wuwa kenapa hard banget gachaa nya udah hampir 150 kali gacha lynae ga pulang gw belom gacha signature nya pliss lah kuroo akun lama ini
3.	Case Folding	wuwa kenapa hard banget gachaa nya udah hampir 150 kali gacha lynae ga pulang gw belom gacha signature nya pliss lah kuroo akun lama ini
4.	Tokenization	['wuwa', 'kenapa', 'hard', 'banget', 'gachaa', 'ya', 'sudah', 'hampir', '150', 'kali', 'gacha', 'lynae', 'tidak', 'pulang', 'gue', 'belum', 'gacha', 'signature', 'ya', 'pliss', 'lah', 'kurooakun', 'lama', 'ini']
5.	Stopword Removal	wuwa kenapa hard banget gachaa ya 150 kali gacha lynae pulang gue undian signature ya pliss kuroo akun
6.	Steaming	wuwa kenapa hard banget gacha ya kali gacha lynae pulang gue undi signature ya pliss kuroo akun

3.3 Wordcloud

WordCloud digunakan untuk memvisualisasikan frekuensi kemunculan kata pada ulasan game *Wuthering Waves* setelah tahap preprocessing. Ukuran kata dalam visualisasi berbanding lurus dengan frekuensi kemunculannya, sehingga memberikan gambaran awal mengenai topik utama yang dibahas pemain. Hasil visualisasi WordCloud dapat dilihat pada Gambar



Gambar 4. Wordcloud

Analisis WordCloud ulasan *Wuthering Waves* mengungkap pola sentimen yang merefleksikan tingkat kepuasan pengguna. Pada sentimen positif, dominasi kata "bagus", "grafik", "gameplay", dan "seru" menunjukkan bahwa kualitas visual yang tinggi dan sistem pertarungan yang dinamis menjadi daya tarik utama permainan. Sebaliknya, analisis pada sentimen negatif dan netral mengungkapkan dua keluhan utama. Pertama, masalah lokalisasi yang ditandai kata "bahasa" dan "Indonesia", di mana absennya opsi bahasa lokal mempersulit pemain memahami alur cerita yang kompleks. Kedua, masalah optimasi perangkat yang diwakili kata "lag" dan "close" (*force close*), khususnya pada perangkat kelas menengah ke bawah. Hal ini mengindikasikan beban komputasi grafis yang berat dan perlunya optimasi sistem pada pembaruan mendatang.

3.4 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja arsitektur dalam mengklasifikasikan sentimen ulasan pengguna permainan *Wuthering Waves* secara objektif.

Tabel 5. Hasil Evaluasi Model Dengan *K-Fold Cross Validation*

Kategori Sentimen (Aktual)	Prediksi Negatif	Prediksi Netral	Prediksi Positif	Total Aktual
Negatif	98	26	1	125
Netral	28	141	18	187
Positif	9	26	289	324
Total Prediksi	135	193	308	636

Berdasarkan data pada Tabel 5, model menunjukkan performa yang sangat baik dalam mengenali pola bahasa pada ulasan bersentimen positif. Dari total 324 data aktual positif, model berhasil memprediksi dengan benar sebanyak 289 ulasan, sedangkan 26 ulasan salah diprediksi sebagai netral dan 9 ulasan sebagai negatif. Tingginya akurasi pada kelas ini didorong oleh karakteristik ulasan positif yang cenderung menggunakan kosakata apresiatif yang konsisten dan eksplisit. Sebaliknya, tantangan klasifikasi terbesar masih ditemukan pada kategori netral. Model berhasil mengidentifikasi 141 ulasan netral dengan benar dari total 187 data aktual. Sebanyak 28 data netral meleset diprediksi sebagai negatif dan 18 data diprediksi sebagai positif. Hal ini terjadi karena ulasan netral di Google Play Store sering kali memiliki struktur kalimat yang ambigu, di mana pengguna mencampurkan kritik dan pujian dalam satu ulasan sekaligus, sehingga membuat batas antarkelas menjadi tumpang tindih. Sementara itu, untuk kategori negatif, model mampu memprediksi 98 ulasan dengan benar dari total 125 data aktual. Kesalahan klasifikasi pada kelas ini mayoritas bergeser ke arah sentimen netral (26 ulasan), yang disebabkan oleh kemiripan penggunaan kata kunci keluhan ringan (seperti masalah optimasi atau permintaan lokalisasi) yang juga sering muncul pada ulasan bersentimen netral. Secara keseluruhan, distribusi hasil ini membuktikan bahwa model memiliki tingkat efektivitas yang tinggi, namun tetap memerlukan perhatian khusus pada penanganan teks yang bersifat ambigu

3.5 Hasil Klasifikasi

Hasil Klasifikasi menggunakan metode *5-Fold Cross Validation* menunjukkan bahwa model LSTM memperoleh rata-rata akurasi sebesar 83,15%, rata-rata precision 81,77%, rata-rata recall 78,95%, dan rata-rata F1-score 79,06%. Nilai tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengklasifikasikan tingkat kepuasan pemain pada ulasan permainan *Wuthering Waves*. Perbedaan hasil pada setiap *fold* relatif kecil, dengan akurasi tertinggi sebesar 85,71% pada *fold* ke-4 dan akurasi terendah sebesar 80,66% pada *fold* ke-2. Hal ini menunjukkan bahwa model mampu mempertahankan performa yang cukup stabil pada berbagai pembagian data. Nilai *precision* yang lebih tinggi dibandingkan recall mengindikasikan bahwa model cukup baik dalam menghasilkan prediksi yang tepat, namun masih terdapat beberapa data yang belum berhasil dikenali pada kelas yang sebenarnya. Sementara itu, rata-rata F1-score sebesar 79,06% menunjukkan keseimbangan yang cukup baik antara precision dan recall dalam proses klasifikasi. Variasi nilai precision, recall, dan F1-score antar *fold* juga mengindikasikan bahwa karakteristik data pada setiap pembagian *fold* dapat memengaruhi kemampuan model dalam mengenali pola sentimen. Ekperimen secara keseluruhan, hasil *5-Fold Cross Validation* menunjukkan bahwa model LSTM memiliki performa yang konsisten dan mampu melakukan klasifikasi sentimen ulasan pemain dengan baik. Penggunaan *cross validation* juga memberikan evaluasi yang lebih robust dibandingkan pembagian data tunggal karena seluruh data memperoleh kesempatan menjadi data pelatihan maupun data validasi, sehingga hasil yang diperoleh lebih representatif terhadap kemampuan model dalam menggeneralisasi data baru.

Tabel 6. Hasil Klasifikasi

Fold	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
1	83.99	81.56	77.99	75.79
2	80.66	80.31	80.66	79.78
3	83.49	86.43	83.49	83.91
4	85.71	76.22	70.71	74.30
5	81.89	84.31	81.89	81.54
Rata-rata	83.15	81.77	78.95	79.06

3.6 Hasil Prediksi

Berdasarkan hasil prediksi model LSTM, model LSTM mampu memprediksi sebagian besar data dengan baik. Namun, masih terdapat beberapa kesalahan klasifikasi pada data berlabel Netral yang diprediksi sebagai Positif. Hal ini menunjukkan bahwa model cukup baik dalam mengenali sentimen, meskipun masih terdapat kesalahan pada beberapa kasus tertentu.

Tabel 7. Hasil prediksi

Hasil Prediksi	Actual	Predicted
grafik mantap edan	Netral	Positif
Game bagus suka	Positif	Positif
bagus minus teks indonesia moga depan	Netral	Positif
tolong tambah fitur bahasa indonesia	Netral	Netral

4. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma *Long Short-Term Memory* (LSTM) untuk menganalisis tingkat kepuasan pemain berdasarkan ulasan game *Wuthering Waves* di Google Play Store. Model yang dibangun mampu mengklasifikasikan ulasan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Pengujian model yang dilakukan menggunakan metode *Stratified 5-Fold Cross Validation* membuktikan bahwa model memiliki performa yang stabil dan konsisten pada berbagai pembagian data, dengan capaian rata-rata nilai akurasi sebesar 83,15%, precision 81,77%, recall 78,95%, dan F1-score sebesar 79,06%. Hasil tersebut menunjukkan bahwa arsitektur LSTM efektif dan tangguh dalam memahami konteks serta ketergantungan jangka panjang pada data teks ulasan pengguna yang bersifat informal

Meskipun demikian, hasil analisis menunjukkan bahwa ketidakpuasan pemain masih didominasi oleh masalah performa dan bug, seperti *force close*, lag, serta optimasi permainan yang belum stabil pada beberapa perangkat. Selain itu, sebagian pemain juga menyoroti belum tersedianya dukungan Bahasa Indonesia. Temuan ini memberikan implikasi praktis bagi pengembang game untuk menetapkan prioritas pembaruan (*update*), khususnya pada peningkatan stabilitas sistem, perbaikan bug, optimalisasi performa lintas perangkat, dan penambahan fitur lokalisasi Bahasa Indonesia. Dengan demikian, hasil penelitian ini dapat dimanfaatkan sebagai dasar pengambilan keputusan yang lebih terarah dalam pengembangan dan evaluasi game agar sesuai dengan kebutuhan serta harapan pemain.

Penelitian ini memiliki keterbatasan pada cakupan data yang hanya bersumber dari Google Play Store dan terbatas pada ulasan berbahasa Indonesia. Selain itu, basis pelabelan data didasarkan pada nilai rating bintang pengguna. Dari sisi pemodelan, eksperimen ini masih berfokus pada arsitektur LSTM konvensional tanpa melibatkan model berbasis *Transformer* yang lebih mutakhir seperti BERT atau IndoBERT. Faktor-faktor tersebut membuka peluang bagi penelitian mendatang untuk memperluas variasi dataset lintas platform serta menerapkan arsitektur *Transformer* demi menangkap konteks semantik secara lebih mendalam.

KONFLIK KEPENTINGAN

Penulis menyatakan bahwa tidak terdapat konflik kepentingan dalam penyusunan artikel ini, dan seluruh data yang disajikan dalam artikel ini merupakan karya ilmiah orisinal penulis.

DAFTAR PUSTAKA

- [1] S. Mujilawati and M. A. Ubaydillah, "Analisis Sentimen Berbasis Aspek pada Ulasan Game Mobile Legends dengan Pendekatan Decision Tree untuk Evaluasi Pengalaman Pengguna Teknik Informatika, Fakultas Teknik, Universitas Islam Lamongan, Indonesia Aspect-Based Sentiment Analysis on Mobile Leg," *J. Pendidik. dan Teknol. Indones.*, vol. 4, no. 11, pp. 325–333, 2024.

- [2] A. Prof, E. BİL, A. Prof, and S. Ergun, "NEW CONSUMERS OF THE DIGITAL AGE : GAME PLAYERS," pp. 9–22, 2021.
- [3] G. Hartono and D. Priharsari, "Analisis User Experience Konten Spiral Abyss Challenge dari Video Game Genshin Impact menggunakan Game Experience Questionnaire (GEQ)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 8, pp. 3838–3846, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [4] K. C. Astuti, A. Firmansyah, dan A. Riyadi, "Implementasi Text Mining untuk Analisis Sentimen Masyarakat terhadap Ulasan Aplikasi Digital Korlantas Polri pada Google Play Store," *Remik: Riset dan E-Jurnal Manajemen Informatika Komputer*, vol. 8, no. 1, hlm. 383-394, Jan. 2024, doi: 10.33395/remik.v8i1.13421
- [5] R. Noveandini, M. S. Wulandari, and F. Rasyad, "Penerapan Model LSTM pada Analisis Sentimen Ulasan Pengguna Aplikasi Shopee Google Play Store," *Fasilkom*, vol. 15, no. 2, pp. 290–296, 2025, [Online]. Available: <https://ejournal.umri.ac.id/index.php/JIK/article/view/9150>
- [6] A. Pangestu, Y. Tajul Arifin, and R. Ade Safitri, "Analisis Sentimen Review Publik Pengguna Game Online Pada Platform Steam Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 6, pp. 3106–3113, 2024, doi: 10.36040/jati.v7i6.8829.
- [7] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan4, "Analysis Data Mining Netflix Data Using The Rapid Miner Application," *J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021.
- [8] Z. N. -, B. I. -, and N. A. R. -, "Analisis Sentimen Ulasan Pengguna Terhadap Game Zenless Zone Zero Menggunakan Metode Bi-Directional Lstm," *J. Inform. dan Tek. Elektro Terap.*, vol. 14, no. 1, 2026, doi: 10.23960/jitet.v14i1.8722..
- [9] M. A. Tamlica and B. Irawan, "Analisis Sentimen Ulasan Aplikasi Grok di Google Play Store Menggunakan Algoritma Long Short-Term Memory (LSTM)," *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, vol. 14, no. 1, pp. 1168–1175, 2026.
- [10] M. H. Shiddiq, B. Prasetyo, and R. A. Ulayya, "Penerapan Naïve Bayes dalam Analisis Sentimen Pengguna Game Wuthering Waves pada Google Play Store," *JSI: Jurnal Sistem Informasi (E-Journal)*, vol. 17, no. 2, pp. 392–403, Oct. 2025.
- [11] P. Aditya, A. Azzahra, and A. Wijaya, "ANALISIS SENTIMEN PEMAIN SUBWAY SURF MELALUI METODE," vol. 3, no. 2, pp. 267–275, 2023.
- [12] R. Putra Kurniawan, I. Istiadi, and S. Wahyu Iriananda, "Analisis Sentimen Ulasan Aplikasi LinkedIn Berbasis Lexicon Dan Long Short-Term Memory (Lstm)," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 2, pp. 2315–2324, 2025, doi: 10.36040/jati.v9i2.13042.
- [13] S. Nurhikmah, R. Ramadani, and G. Triyono, "Analisis Sentimen pada Ulasan Aplikasi Wondr di Play Store dengan Metode Naïve Bayes," pp. 1919–1930, 2025, doi: 10.33364/algoritma/v.22-2.2507.
- [14] A. Kho and F. F. Tampinongkol, "Analisis Sentimen Pengguna Aplikasi STEAM dengan Algoritma Naive Bayes," *Ranah Res. J. Multidiscip. Res. Dev.*, vol. 7, no. 6, pp. 4620–4627, 2025, doi: 10.38035/rj.v7i6.1803
- [15] D. Putri dan A. Suryanto, "Penerapan LSTM untuk Menganalisis Sentimen Review Pengguna Aplikasi Zoom pada Play Store," *IMTechno: Journal of Industrial Management and Technology*, vol. 6, no. 2, hlm. 100-106, Jul. 2025
- [16] D. S. Nurrochmah, N. Rahaningsih, R. D. Dana, dan C. L. Rohmat, "Penerapan Algoritma Naive Bayes dalam Analisis Sentimen Ulasan Aplikasi Kitalulus di Google Play Store," *Jurnal Informatika Terpadu*, vol. 11, no. 1, hlm. 1-11, Mar. 2025.
- [17] U. Farauk, E. Putra, D. Ayu, N. Ramhan, and L. N. Alifah, "Analisis Sentimen Pada Tweet Buzzer Politik Indonesia Menggunakan Metode SVM Sentiment Analysis On Indonesian Political Buzzer Tweets Using Svm Method," vol. 4, no. 1, pp. 311–326, 2025.
- [18] N. Hidayah, L. Ode, and D. Pramono, "Evaluasi IndoBERTweet untuk Klasifikasi Sentimen Ulasan Wisata melalui Inferensi Langsung Lintas Domain : Studi Kasus Kota Baubau Evaluation of IndoBERTweet for Tourism Review Sentiment Classification via Cross-Domain Direct Inference : A Case Study in Baubau City," vol. 15, no. 1, pp. 60–72, 2026.

- [19] U. Kulsum, M. Jajuli, and N. Sulistiyowati, "Analisis Sentimen Aplikasi WETV di Google Play Store Menggunakan Algoritma Support Vector Machine," *J. Appl. Informatics Comput.*, vol. 6, no. 2, pp. 205–212, 2022, doi: 10.30871/jaic.v6i2.4802.
- [20] D. F. Nawulansih et al., "Analisis Sentimen Ulasan Aplikasi DANA di Google Play Store : Penerapan Support Vector Machine dan Synthetic Minority Over-sampling Technique Sentiment Analysis of DANA Application Reviews on Google Play Store : Implementation of Support Vector Machine and Synthetic Minority Over-sampling Technique," vol. 5, no. 9, pp. 2660–2671, 2025.
- [21] M. Khadapi and V. M. Pakpahan, "Analisis Sentimen Berbasis Jaringan LSTM dan BERT terhadap Diskusi Twitter tentang Pemilu 2024," *JUKI: Jurnal Komputer dan Informatika*, vol. 6, no. 2, pp. 130–137, Nov. 2024.
- [22] S. K. Lubis, M. H. Dar, and F. A. Nasution, "Analisis Sentimen Ulasan Pengguna Aplikasi pada Google Play Store Menggunakan Algoritma Support Vector Machine," *Informatika*, vol. 11, no. 2, pp. 120–128, 2024, doi: 10.36987/informatika.v11i2.5860.
- [23] W. Arisandi, S. Anggai, and Tukiayat, "Analisis Sentimen Ulasan Pengguna Aplikasi Media Sosial X Di PlayStore Menggunakan Algoritma Long Short-Term Memory (LSTM) Dan Gated Recurrent Unit (GRU)," *Jurnal Sistem Komputer dan Kecerdasan Buatan*, vol. IX, no. 1, pp. 63–72, Sep. 2025.
- [24] A. Almansour, R. Alotaibi, and H. Alharbi, "Text - rating review discrepancy (TRRD): an integrative review and implications for research," *Futur. Bus. J.*, 2022, doi: 10.1186/s43093-022-00114-y.
- [25] Z. Sitorus, M. Iqbal, D. Nasution, and R. F. Wijaya, "Penerapan Deep Learning dan Analisis Sentimen terhadap Gap Kompetensi Lulusan Lembaga Pendidikan dan Pelatihan Vokasi terhadap Dunia Kerja dengan Metode Long Short-Term Memory (LSTM)," vol. 6, no. 2, pp. 0–11, 2025, doi: 10.47065/bit.v5i2.2029.
- [26] A. Harris, A. Nugroho, Y. Novianto, J. Jasmir, and D. Fatma, "Fitur Word Embedding untuk meningkatkan Kinerja Machine Learning pada Analisis Sentimen Game Honor of Kings Word Embedding Features to Improve Machine Learning Performance in Sentiment Analysis of the Honor of Kings Game," vol. 15, pp. 522–533, 2026.
- [27] C. Fri, R. Elouahbi, Y. Taki, and A. Remaida, "Enhanced Bidirectional LSTM for Sentiment Analysis of Learners ' Posts in MOOCs," vol. 16, no. 5, pp. 163–172, 2025.
- [28] P. Alkhairi, A. P. Windarto, and M. Efendi, "Optimasi LSTM Mengurangi Overfitting untuk Klasifikasi Teks Menggunakan Kumpulan Data Ulasan Film Kaggle IMDB," *bits*, vol. 6, no. 2, p. 1142–1150, Sep. 2024.
- [29] A. R. Gunawan and R. F. Alfa Aziza, "Sentiment Analysis Using LSTM Algorithm Regarding Grab Application Services in Indonesia," *J. Appl. Informatics Comput.*, vol. 9, no. 2, pp. 322–332, 2025, doi: 10.30871/jaic.v9i2.8696.