

Transformasi Ulasan Game Menjadi Insight Strategis Menggunakan Aspect-Based Sentiment Analysis dan Business Intelligence

Fahrani Indah Pramita¹ dan Riska Dhenabayu²

^{1,2} Bisnis Digital, Universitas Negeri Surabaya, Indonesia

*) Corresponding Author (Email: fahrani.22092@mhs.unesa.ac.id)

ABSTRACT

Video games were initially developed in military environments as a form of entertainment and have since evolved into a global industry with billions of players. One of the popular RPG games is Elden Ring, which generates a large volume of user reviews, making it difficult for Developers to effectively identify praise and complaints. This study aims to perform ABSA on player reviews, covering gameplay, graphics, performance, and story, while also comparing the performance of LR, BERT, and RF, and visualizing the results through a BI dashboard. The results indicate that all aspects are dominated by positive sentiment. BERT achieved the highest accuracy at 90%, followed by LR (89%) and RF (83%), although LR demonstrated more stable confidence values. The visualization further reveals that positive sentiments are more consistent, whereas negative sentiments tend to be more complex and diverse. For future research, it is recommended to apply model calibration techniques, including softmax output calibration, incorporating uncertainty into the training process, and evaluating calibration using Expected Calibration Error to improve alignment between confidence and actual model performance.

Keywords: Sentiment Analysis; ABSA; BERT; Business Intelligence; Game Review

I. PENDAHULUAN

Video game pertama kali diperkenalkan di lingkungan pangkalan militer sebagai sarana hiburan bagi para rekrutan untuk melepaskan diri dari tekanan pelatihan dasar. Perkembangan industri game sendiri dimulai pada tahun 1950-an dan lahirlah SEGA (Service Games), yang menjadi salah satu pelopor dalam industri permainan elektronik. Langkah tersebut menandai awal mula evolusi industri game modern yang terus berkembang pesat hingga kini, melahirkan berbagai inovasi dalam



teknologi, desain, dan pengalaman bermain interaktif (Lee & Novak, 2022). Sebuah perusahaan riset pasar global yang berperan sebagai penyedia utama data, analisis dan wawasan industri game yang berbasis di Belanda melaporkan bahwa terdapat 3,6 Milyar pemain di seluruh dunia dari total 2,985 Juta pemain mobile, 936 Juta pemain komputer dan 645 Juta pemain konsol (Newzoo, 2025). Tercatat adanya 953,426 ribu pemain memainkan game dengan genre RPG (Role-Playing Game) berjudul Elden Ring yang rilis pada Februari 2022 yang mengunggah ulasan sebanyak 1.096.686 ribu ulasan (SteamDB.info, 2025; ELDEN RING on Steam, 2025).

Dari banyaknya data ulasan yang ada, Developer perlu memastikan kualitas layanan game untuk membangun lingkungan komunitas yang seimbang dan terjaga sehingga diperlukan menganalisis ulasan tersebut secara komprehensif. Karena pada dasarnya ulasan game berfungsi sebagai sumber daya yang berharga untuk mencerminkan preferensi pemain dan menunjukkan masalah yang ada (Yu et al., 2023). Melalui analisis tersebut, Developer dapat memahami kebutuhan dan harapan pemain dengan lebih baik. Oleh karena itu, analisis sentimen diperlukan untuk mengetahui persepsi pemain secara lebih mendalam berdasarkan polaritas dataset dengan mendeteksi polaritas dari kalimat, paragraf atau spam teks (Liu, 2012; Trisna & Jie, 2022). Untuk mengkaji ulasan pemain secara mendalam dapat dilakukan pendekatan lanjutan menggunakan Aspect-Based Sentiment Analysis (ABSA) yang dapat menganalisis pengaruh atau sentimen dalam teks dan diklasifikasikan ke dalam beberapa kategori polaritas positif, netral dan negatif (Hua et al., 2024).

Terdapat penelitian-penelitian lain yang mengkaji klasifikasi sentimen seperti penelitian Bećirović et al. (2024) berjudul *Assessing the Accuracy of Logistic Regression and Bert in Sentiment Analysis and Mental Disorder Classification* yang menunjukkan bahwa algoritma Logistic Regression mencapai tingkat Accuracy sebesar 72%, sedangkan BERT mencapai 84%. Selanjutnya ada penelitian oleh Salsabila et al. (2023) berjudul *Aspect-based Sentiment and Correlation-based Emotion Detection on Tweets for Understanding Public Opinion of Covid-19* yang menunjukkan bahwa algoritma Random Forest 3% lebih unggul daripada BERT pada dataset pertama sedangkan pada dataset kedua BERT lebih unggul hingga 7% dari Random Forest. Terakhir, ada penelitian oleh Purbolaksono et al. (2024) berjudul *Sentiment Analysis of Game Review in Steam Platform using Random Forest* menunjukkan bahwa penerapan Bigram TF-IDF tanpa proses lemmatization pada tahap preprocessing mampu menghasilkan performa model yang cukup tinggi, dengan nilai F1-Score sebesar 62%.

Berdasarkan penelitian terdahulu, masih ditemukan kesenjangan penelitian dimana penelitian menunjukkan fokus yang dominan pada peningkatan Accuracy dan klasifikasi dan perbandingan performa algoritma tanpa mengeksplorasi bagaimana hasil analisis dapat dimanfaatkan untuk mendukung pengambilan keputusan strategis. Penelitian tersebut juga menggunakan pendekatan sentimen tingkat dokumen atau kalimat, sehingga belum mampu mengidentifikasi sentimen berdasarkan aspek-aspek spesifik yang menjadi perhatian pengguna khususnya

pemain dalam konteks industri game, seperti gameplay, graphics, performance dan story. Dapat disimpulkan bahwa masih terdapat ruang untuk pengembangan lebih lanjut, baik dari sisi pemilihan algoritma maupun keragaman dan cakupan dataset, guna memperoleh hasil klasifikasi sentimen yang lebih optimal dan adaptif terhadap konteks data yang berbeda. Di samping itu, hasil analisis pada penelitian-penelitian tersebut belum diintegrasikan ke dalam rangka Business Intelligence yang dapat menghasilkan insight strategis bagi Developer game. Oleh karena itu, penelitian ini mengisi kesenjangan tersebut melalui penerapan ABSA pada ulasan game Elden Ring dan akan menampilkan visualisasi hasilnya dalam dashboard Business Intelligence untuk mendukung pengambilan keputusan berbasis data guna mempertahankan reputasi dan mengatasi kepuasan pengguna.

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan dalam menginterpretasikan hasil yang diperoleh. Dataset yang digunakan hanya berfokus pada ulasan Elden Ring dari platform Steam sehingga karakteristik bahasa, perilaku pengguna, dan pola sentimen yang diperoleh belum tentu mewakili game lain maupun platform distribusi yang berbeda. Selain itu, penelitian ini hanya membandingkan tiga algoritma klasifikasi, yaitu Logistic Regression, BERT, dan Random Forest, sehingga masih terdapat peluang untuk mengevaluasi model lain yang berpotensi memberikan performa lebih baik. Dari sisi analisis bisnis, dashboard hanya dikembangkan menggunakan Tableau sehingga belum mengeksplorasi kemampuan platform Business Intelligence lainnya. Potensi bias juga dapat muncul akibat ketidakseimbangan distribusi sentimen dan ekspresi berbasis emoji yang umum ditemukan dalam ulasan pengguna. Pada model lebih optimal dalam mengenali kelas mayoritas dibandingkan kelas minoritas. Di samping itu, pendekatan ABSA yang digunakan belum mampu mengidentifikasi makna implisit seperti sarkasme, ironi,

II. TINJAUAN PUSTAKA

Pembelajaran Mesin dalam Klasifikasi Sentimen

Dasar awal perkembangan teori klasifikasi diperkenalkan oleh R.A. Fisher (1936) melalui konsep Linear Discriminant Function (LDF), yang kemudian berkembang menjadi metode Linear Discriminant Analysis (LDA). Pendekatan ini memberikan kerangka matematis untuk memisahkan suatu data ke dalam kelompok-kelompok berbeda berdasarkan kombinasi linear dari variabel prediktor. LDA tidak hanya memungkinkan proses klasifikasi dilakukan secara lebih sistematis, tetapi juga menyediakan landasan statistik untuk mengevaluasi tingkat Accuracy, termasuk menghitung probabilitas terjadinya kesalahan klasifikasi. Dengan demikian, metode ini menjadi salah satu fondasi penting dalam perkembangan teknik klasifikasi modern yang banyak diaplikasikan di berbagai bidang, termasuk analisis data teks dan pengenalan pola.

Dalam mendukung pengembangan game dan proses pengambilan keputusan oleh pihak Developer, penerapan teknik Machine Learning dan Deep Learning menjadi semakin krusial, terutama dalam klasifikasi sentimen ulasan pengguna. Analisis sentimen ini memungkinkan Developer untuk memahami opini dan preferensi pemain secara lebih sistematis, sehingga dapat mengarahkan pengembangan produk dan strategi dengan lebih tepat. Beberapa algoritma yang umum digunakan dalam klasifikasi sentimen mencakup Support Vector Machine (SVM), Naive Bayes, Decision Tree (DT), masing-masing menawarkan keunggulan tersendiri dalam memproses data teks dan menghasilkan prediksi sentimen yang akurat (Talaat, 2023).

Logistic Regression

Logistic Regression merupakan sebuah kerangka analisis yang bersifat umum, relatif sederhana, namun sangat tangguh dalam menangani data biner (D.R. Cox, 1958). Tidak seperti Linear Regression yang menghasilkan output kontinu, Logistic Regression memprediksi probabilitas suatu pengamatan ke dalam kelas tertentu, sehingga sangat efektif digunakan pada kasus klasifikasi. Model ini bekerja dengan mentransformasikan kombinasi linear dari variabel prediktor melalui fungsi logit atau sigmoid agar output berada pada rentang 0–1. Logistic Regression banyak diterapkan untuk berbagai tipe klasifikasi, mulai dari binomial (dua kelas), multinomial (lebih dari dua kelas tanpa urutan), hingga ordinal (lebih dari dua kelas dengan peringkat tertentu). Selain itu, metode ini memiliki interpretabilitas yang baik melalui koefisien model yang dapat menunjukkan pengaruh setiap variabel prediktor terhadap peluang terjadinya sebuah kelas, sehingga menjadikannya alat analitik yang mudah dipahami (Starbuck, 2023).

Bidirectional Encoder Representations for Transformers (BERT)

BERT pertama kali diperkenalkan oleh Devlin et al. (2018) menggunakan mekanisme utama bernama Self-Attention yang memungkinkan model untuk menangkap hubungan antar kata dalam sebuah kalimat secara lebih menyeluruh, tanpa bergantung pada urutan kata secara linear, sehingga pemahaman konteks menjadi lebih akurat dan komprehensif. Model ini memeriksa teks di kedua arah secara bersamaan dan menyediakan penyematan kontekstual untuk setiap kata yang ada di dalam teks, sehingga dapat memahami bahasa lebih baik (Khurana et al., 2023). Selain itu, BERT memberikan peningkatan yang signifikan pada berbagai tugas NLP yaitu Sentence Boundary Detection, Tokenization, Dependency Parsing, Part of Speech Tagging, Machine Translation, Question Answering System, Named Entity Recognition, serta Grammatical Error Detection and Correction (Gardazi et al., 2025).

Random Forest

Salah satu metode ensemble learning yang banyak digunakan adalah Random Forest (RF), yang pertama kali diperkenalkan oleh Leo Breiman (2001)

sebagai penyempurnaan dari algoritma pohon keputusan. Prinsip dasar RF adalah membangun sejumlah pohon keputusan secara acak dengan memanfaatkan teknik bootstrap sampling pada data serta melakukan pemilihan subset fitur secara acak pada setiap percabangan pohon yang kemudian digabungkan melalui mekanisme voting mayoritas untuk kasus klasifikasi, atau melalui proses averaging untuk kasus regresi. Breiman mengembangkan teori matematis yang menunjukkan bahwa generalization error bergantung pada rasio c/s^2 (correlation/strength squared), di mana semakin kecil rasio ini maka semakin baik performa model. Teori ini memberikan upper bound untuk generalization error: $PE^* \leq \bar{\rho}(1-s^2)/s^2$, yang menunjukkan bahwa untuk mendapatkan akurasi tinggi, diperlukan tree individual yang kuat (high strength) namun dengan korelasi rendah antar pohon (low correlation).

Aspect Based Sentiment Analysis (ABSA)

Analisis sentimen merupakan cabang ilmu yang berfokus pada pengkajian opini, perasaan, serta sikap seseorang terhadap suatu objek tertentu. Objek yang dimaksud dapat berupa produk, layanan, individu, organisasi, peristiwa, atau topik yang sedang dibahas. Melalui analisis ini, penilaian dan evaluasi dari masyarakat dapat diidentifikasi untuk memahami persepsi mereka secara lebih mendalam. Dengan demikian, analisis sentimen membantu memberikan gambaran tentang emosi dan pandangan publik terhadap suatu entitas (Liu, 2012).

Business Intelligence

Business Intelligence mencakup banyak teknologi, sistem, dan alat yang digunakan untuk menyusun, menganalisis, dan menampilkan laporan untuk memfasilitasi pengambilan keputusan perusahaan. Hal tersebut berpotensi untuk meningkatkan efektivitas standar operasional serta pengaruhnya di tingkat perusahaan dimana dapat menawarkan opsi strategis yang lebih baik dalam lingkungan bisnis yang dinamis (Zhao & Zhang, 2024). Dalam konteks analisis sentimen, data yang digunakan dapat diproses sehingga memberikan bagi perusahaan untuk menafsirkan persepsi pengguna dan merespon dengan cara yang tepat terhadap pujian atau masalah yang diungkapkan (Mirea & Grădinaru, 2025). Tujuannya adalah untuk meminimalkan waktu analisis semua jenis operasi bisnis dan manajemen tingkat atas menggunakan Business Intelligence untuk mengambil keputusan dalam waktu singkat (Mantena & Murugan, 2025).

III. METODE

Penelitian ini menggunakan jenis penelitian kuantitatif yang bertujuan untuk membandingkan kinerja algoritma Logistic Regression, Bidirectional Encoder Representations for Transformers (BERT) dan Random Forest (RF) dalam klasifikasi sentimen ulasan game Elden Ring berdasarkan aspek performa klasifikasi yaitu Confusion Matrix, Accuracy, Precision, Recall dan F1-Score. Penelitian ini diharapkan

dapat memberikan pandangan komprehensif tentang aspek performa polaritas sentimen dari ulasan serta mengidentifikasi faktor teknologi dan manajerial hasil penerapan algoritma dalam Business Intelligence (Fauzan, 2025; Shuvo et al., 2025).

Data dan Subjek Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diambil dari Dataset Steam Reviews of Elden Ring (Lorenzo L., 2025). Dataset ini berisi total 564.971 data hasil web scraping dari platform Steamworks menggunakan Steam API yang memuat berbagai ulasan pengguna terhadap game Elden Ring. Dari jumlah tersebut, populasi yang akan digunakan dalam penelitian ini hanya mencakup ulasan dalam bahasa Inggris sebanyak 474.561 data. Setelah dilakukan pre-processing, yaitu menghilangkan data yang terduplikat, menghapus data yang tidak lengkap atau kosong, maka didapatkan sampel sejumlah 257.122 data dan akan dibagi menjadi dua subset: 20% (51.516 sampel) akan digunakan sebagai data uji kinerja model, 80% (205.606 sampel) akan digunakan sebagai data pelatihan model. Pembagian dengan rasio 80:20 ini dipilih karena telah terbukti memberikan performa optimal pada berbagai kasus klasifikasi (Bichri et al., 2024).

Metode Analisis Data

Sebelum proses pembersihan data dilakukan, tahap Exploratory Data Analysis (EDA) perlu dilaksanakan terlebih dahulu untuk memahami struktur, kualitas, serta distribusi data. Pada tahap ini, digunakan library Pandas untuk meninjau kolom yang tersedia, tipe data pada setiap kolom, serta gambaran awal mengenai bentuk dan karakteristik dataset secara keseluruhan. Selanjutnya dilakukan preprocessing data untuk mengubah data mentah yang tidak terstruktur menjadi format yang lebih bersih sehingga proses analisis menggunakan model algoritma menghasilkan hasil yang akurat (Jonnala et al., 2025).

ABSA Transformation

ABSA Transformation merupakan proses mengubah teks ke dalam bentuk yang tidak hanya merepresentasikan sentimen secara keseluruhan, tetapi juga mampu memisahkan opini berdasarkan aspek yang dibahas dalam suatu ulasan dilakukan Aspect Extraction, yaitu proses mengidentifikasi entitas, fitur, atau elemen spesifik yang menjadi fokus opini dalam teks. Proses anotasi dilakukan berdasarkan pedoman anotasi yang telah disusun dengan mengacu pada pendekatan ABSA dan definisi operasional setiap aspek, yaitu gameplay, graphics, performance, dan story. Untuk memastikan kualitas anotasi, hasil anotasi akan dibandingkan menggunakan inner-annotator agreement menggunakan koefisien Cohen's Kappa melalui proses diskusi dan peninjauan ulang hingga diperoleh kesepakatan akhir. Tahapan ini dilakukan untuk meningkatkan konsistensi dan reliabilitas data sebelum digunakan pada proses pelatihan model. Nilai kesepakatan yang tinggi menunjukkan bahwa pedoman anotasi dapat diterapkan secara konsisten oleh annotator. Selanjutnya, seluruh perbedaan hasil anotasi didiskusikan dan direkonsiliasi hingga diperoleh label akhir yang digunakan sebagai dataset pelatihan (Pontiki et al., 2015; Imron et al., 2023).

Split Validation

Split Validation akan digunakan pada penelitian ini di mana Dataset akan dibagi menjadi rasio 80:20. Data Latih sebanyak 80% dari total data untuk melatih model algoritma dan Data Uji sebanyak 20% dari total data untuk mengukur kinerja model algoritma. Parameter `random_state` atau `random seed=42` akan diterapkan untuk memastikan reproduktivitas dari semua proses pelatihan dan validasi model (Sharma et al., 2025).

Evaluasi Model Algoritma

Setiap model yang dihasilkan dari algoritma yang digunakan akan dievaluasi menggunakan data uji dengan beberapa metrik evaluasi, yaitu Confusion Matrix, Accuracy, Precision, Recall, dan F1-Score. Penggunaan beberapa metrik ini bertujuan untuk memberikan gambaran yang lebih komprehensif mengenai kemampuan model dalam mengklasifikasikan sentimen secara tepat. Confusion Matrix digunakan untuk melihat distribusi prediksi yang benar dan salah pada setiap kelas, sedangkan Accuracy mengukur tingkat ketepatan prediksi secara keseluruhan. Selain itu, Precision, Recall, dan F1-Score digunakan untuk mengevaluasi performa model secara lebih mendalam, khususnya dalam menangani ketidakseimbangan kelas yang sering ditemukan pada data sentimen. Hasil evaluasi dari setiap model kemudian akan dibandingkan secara sistematis untuk menentukan algoritma yang memiliki performa terbaik dan paling sesuai dengan karakteristik dataset yang digunakan dalam penelitian ini.

Business Intelligence

Tableau akan digunakan sebagai sarana Business Intelligence (BI) untuk mengolah dan memvisualisasikan hasil analisis sentimen menjadi informasi yang lebih mudah dipahami oleh berbagai pemangku kepentingan. Melalui dashboard interaktif, hasil pemodelan sentimen dan Aspect-Based Sentiment Analysis (ABSA) dapat disajikan dalam bentuk grafik, indikator, serta ringkasan informasi yang memudahkan identifikasi pola, tren, dan persepsi pengguna terhadap berbagai aspek game. Visualisasi ini memungkinkan data yang kompleks diinterpretasikan secara lebih cepat dan intuitif dibandingkan penyajian dalam bentuk tabel atau angka semata.

Tujuan utama dari visualisasi hasil analisis sentimen ini adalah untuk memberikan nilai praktis dalam mendukung pengambilan keputusan strategis di industri game. Informasi yang dihasilkan dapat membantu Developer mengidentifikasi aspek yang paling banyak mendapat apresiasi maupun kritik dari pemain, seperti gameplay, graphics, performance, dan story. Selain itu, dashboard BI juga dapat digunakan untuk memantau perubahan sentimen pengguna dari waktu ke waktu, mengevaluasi dampak pembaruan game, serta menentukan prioritas perbaikan dan pengembangan fitur. Dengan demikian, hasil analisis tidak hanya berhenti pada proses klasifikasi sentimen, tetapi juga dapat dikonversi menjadi wawasan bisnis (business insights) yang mendukung peningkatan kualitas produk dan kepuasan pemain.

IV. HASIL DAN PEMBAHASAN

Hasil Distribusi Aspek dan Sentimen

Sebelum proses pemodelan dilakukan, karakteristik dataset dianalisis melalui Exploratory Data Analysis (EDA) untuk memahami distribusi sentimen dan aspek yang terdapat dalam ulasan. Hasil eksplorasi kemudian digunakan sebagai dasar pada tahap preprocessing yang meliputi penanganan missing values, penghapusan duplikasi, tokenisasi, normalisasi teks, serta penghapusan stopwords menggunakan NLTK dan daftar kata tambahan yang relevan dengan domain game. Selanjutnya, pendekatan Aspect-Based Sentiment Analysis (ABSA) diterapkan untuk mengidentifikasi aspek yang dibahas pemain dan menentukan polaritas sentimen pada setiap aspek. Proses ini dilanjutkan dengan stemming untuk mengubah kata ke bentuk dasar dan pembobotan TF-IDF untuk menghasilkan representasi fitur yang lebih informatif bagi model klasifikasi.

Tabel 1. Distribusi Aspek Ulasan

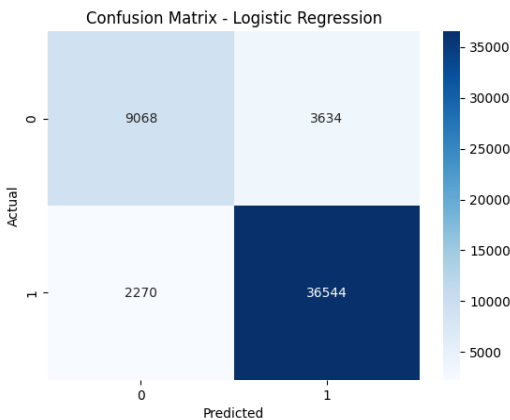
Aspek	Sentimen		Total per Aspek
	Positif	Negatif	
gameplay	63.776	19.227	83.003
graphics	50.821	15.158	65.979
performance	38.113	16.213	54.326
story	38.129	10.277	48.406
Total per Sentimen	190.839	60.875	

Sumber: Data Sekunder diolah oleh Peneliti (2026)

Distribusi data hasil anotasi ditunjukkan pada Tabel 1. Secara keseluruhan, dataset didominasi oleh sentimen positif dengan jumlah 190.839 ulasan (75,81%), sedangkan sentimen negatif berjumlah 60.875 ulasan (24,19%), sehingga total data yang dianalisis mencapai 251.714 pasangan aspek-sentimen. Dari sisi distribusi aspek, gameplay merupakan aspek yang paling banyak dibahas oleh pemain dengan total 83.003 ulasan, diikuti graphics sebanyak 65.979 ulasan, performance sebanyak 54.326 ulasan, dan story sebanyak 48.406 ulasan. Temuan ini menunjukkan bahwa mayoritas pemain memberikan apresiasi positif terhadap berbagai aspek dalam Elden Ring, sekaligus mengindikasikan bahwa gameplay menjadi aspek yang paling banyak mendapatkan perhatian dalam ulasan pengguna. Penyajian distribusi sentimen dan aspek ini penting untuk mendukung interpretasi performa model dan hasil ABSA karena memberikan gambaran mengenai karakteristik data yang digunakan selama proses pelatihan dan evaluasi model.

Hasil Evaluasi Model Logistic Regression

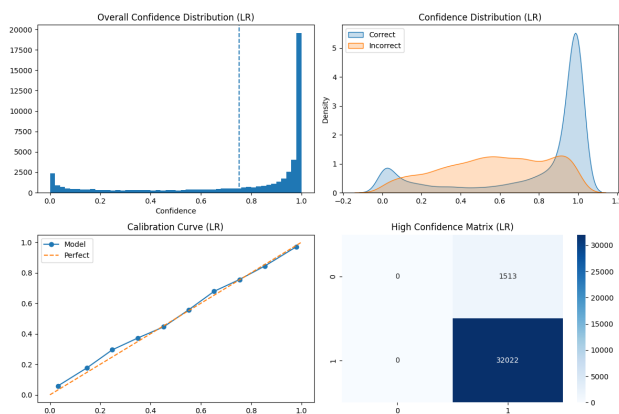
Model dibangun menggunakan Pipeline yang terdiri dari 2 komponen utama, yaitu TF-IDF Vectorizer dan Logistic Regression. TF-IDF Vectorizer digunakan untuk mengubah data teks menjadi bentuk numerik dengan parameter max_features=5000 agar model tetap efisien dan tidak terlalu kompleks. Penggunaan ngram_range=(1,2) memungkinkan model menangkap unigram dan bigram dengan tujuan konteks sederhana dapat dipertimbangkan. Parameter min_df=2 dan max_df=0.9 digunakan karena dapat menyaring kata-kata umum maupun yang jarang muncul. Untuk mencegah overfitting maka eksperimen ini menggunakan regularisasi L2. Parameter C=10 menunjukkan bahwa regularisasi yang diterapkan relatif lemah sehingga model akan memiliki fleksibilitas lebih dalam menyesuaikan data. Solver lbfgs dipilih karena efisien untuk dataset berukuran besar, sedangkan max_iter=100 untuk menentukan batas maksimum iterasi pelatihan. Parameter random_state=42 digunakan untuk memastikan hasil yang konsisten.



Sumber: Data Sekunder diolah oleh Peneliti (2026)

Gambar 1. Confusion Matrix - Logistic Regression

Pada Gambar 1 diatas, model ini mampu mengklasifikasikan data dengan benar dengan nilai Accuracy 89%, terutama pada kelas positif dimana nilai TP tinggi yaitu sebanyak 36.544 data. Namun, masih terdapat kesalahan klasifikasi pada FP dengan nilai yang cukup besar yaitu sebanyak 2.270 data.



Sumber: Data Sekunder diolah oleh Peneliti (2026)

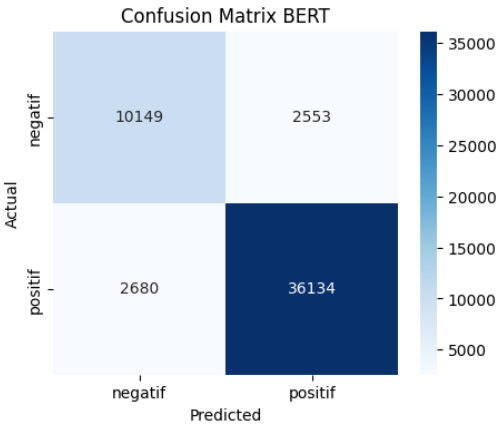
Gambar 2. Distribusi nilai Confidence - Logistic Regression

Berdasarkan hasil distribusi confidence pada Gambar 2, sebagian besar prediksi model terkonsentrasi pada nilai yang tinggi, yaitu mendekati 1, dengan nilai rata-rata sebesar 0,75. Hal ini menunjukkan bahwa model memiliki tingkat keyakinan yang cukup tinggi dalam mengklasifikasikan sentimen ulasan. Temuan tersebut didukung dengan prediksi yang benar sebagian besar berada pada rentang confidence 0,9–1,0, sedangkan prediksi yang salah memiliki distribusi yang lebih menyebar. Meskipun demikian, masih terdapat beberapa kasus di mana model memberikan tingkat keyakinan yang tinggi terhadap prediksi yang keliru. Selanjutnya, kurva kalibrasi menunjukkan bahwa performa model cukup mendekati garis diagonal yang merepresentasikan kalibrasi sempurna, sehingga probabilitas yang dihasilkan model relatif sesuai dengan probabilitas aktual pada data. Adapun hasil High Confidence Matrix yang menunjukkan bahwa mayoritas prediksi dengan tingkat keyakinan tertinggi berada pada kelas positif sebanyak 32.022 prediksi, sedangkan kelas negatif hanya berjumlah 1.513 prediksi. Perbedaan yang cukup besar ini mengindikasikan dominasi sentimen positif dalam dataset sekaligus menunjukkan bahwa sebagian besar pengguna memberikan persepsi yang baik terhadap game yang diteliti, meskipun sentimen negatif yang muncul tetap dapat menjadi sumber informasi penting bagi Developer untuk mengidentifikasi aspek-aspek yang masih memerlukan perbaikan.

Hasil Evaluasi Model BERT

Model BERT yang digunakan dalam penelitian ini memanfaatkan arsitektur bert-base-uncased yang telah dilatih sebelumnya sehingga memiliki pemahaman dasar terhadap struktur dan konteks bahasa. Representasi teks diperoleh melalui tokenizer BERT yang mengubah ulasan menjadi token-token numerik yang dapat diproses oleh model. Panjang token diseragamkan menggunakan parameter `max_length=128` untuk menjaga efisiensi komputasi sekaligus memastikan informasi penting dalam ulasan tetap dapat dipresentasikan. Label sentimen dikonversi ke dalam bentuk numerik dan dikelola menggunakan kelas `SentimentDataset` yang

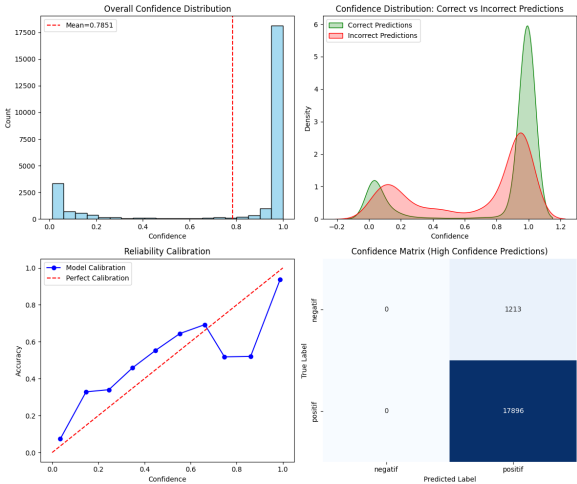
terintegrasi dengan framework PyTorch. Konfigurasi pelatihan menggunakan num_train_epoch=5 dan batch_size=16, yang dipilih untuk memberikan keseimbangan antara kualitas pembelajaran model dan efisiensi waktu komputasi.



Sumber: Data Sekunder diolah oleh Peneliti (2026)

Gambar 3. Confusion Matrix - BERT

Pada Gambar 3 diatas, performa yang dihasilkan oleh BERT cenderung seimbang dengan nilai Accuracy sebesar 90%. Nilai TN pada BERT cenderung tinggi sebesar 10.149 data. Disisi lain, terdapat kesalahan prediksi yang tinggi pada FP dengan nilai sebesar 2.680 data dimana lebih tinggi daripada LR.



Sumber: Data Sekunder diolah oleh Peneliti (2026)

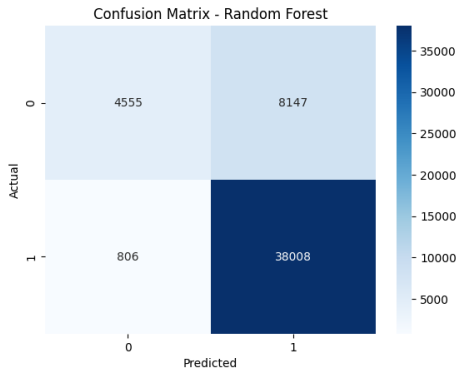
Gambar 4. Distribusi nilai Confidence - BERT

Berdasarkan hasil distribusi confidence pada Gambar 4, sebagian besar prediksi model terkonsentrasi pada nilai yang tinggi dengan rata-rata sebesar 0,78. Hal ini menunjukkan bahwa model memiliki tingkat keyakinan yang cukup kuat

dalam mengklasifikasikan sentimen ulasan. Temuan tersebut didukung oleh distribusi prediksi di mana nilai prediksi yang benar memiliki puncak kurva densitas mencapai angka 6 pada nilai confidence mendekati 1, sedangkan prediksi yang salah cenderung lebih menyebar pada rentang confidence rendah hingga menengah, meskipun masih terdapat beberapa prediksi keliru dengan tingkat keyakinan tinggi pada kisaran 0,8–1,0. Selanjutnya, pada kurva kalibrasi menunjukkan bahwa akurasi yang dicapai cenderung lebih rendah dibandingkan probabilitas yang diprediksi, sehingga model belum sepenuhnya terkalibrasi dengan baik dan masih menunjukkan kecenderungan terlalu percaya diri atau kurang percaya diri pada beberapa prediksi. Adapun hasil High Confidence Matrix menunjukkan bahwa mayoritas prediksi dengan tingkat keyakinan tertinggi berada pada kelas positif sebanyak 17.896 prediksi, sedangkan kelas negatif hanya berjumlah 1.213 prediksi. Perbedaan tersebut mengindikasikan dominasi sentimen positif dalam dataset sekaligus menunjukkan bahwa sebagian besar pengguna memberikan persepsi yang baik terhadap game yang diteliti, meskipun sentimen negatif yang muncul tetap dapat menjadi masukan penting bagi Developer untuk melakukan perbaikan pada aspek tertentu.

Hasil Evaluasi Model Random Forest

Pada eksperimen ketiga, digunakan algoritma Random Forest sebagai metode klasifikasi untuk membandingkan performanya dengan model sebelumnya. Model dibangun menggunakan pendekatan Pipeline yang mengintegrasikan proses ekstraksi fitur dan klasifikasi dalam satu alur dengan menggunakan dua komponen yaitu TF-IDF Vectorizer dan Random Forest Classifier. Parameter max_features=5000 membatasi jumlah fitur yang digunakan sehingga model tetap efisien dan tidak mengalami kompleksitas berlebih. Parameter n_estimators=100 menunjukkan bahwa model menggunakan 100 pohon keputusan untuk melakukan klasifikasi. Sementara itu, max_depth=665 menunjukkan bahwa setiap pohon dapat tumbuh sangat dalam, yang memungkinkan model menangkap pola kompleks dalam data.

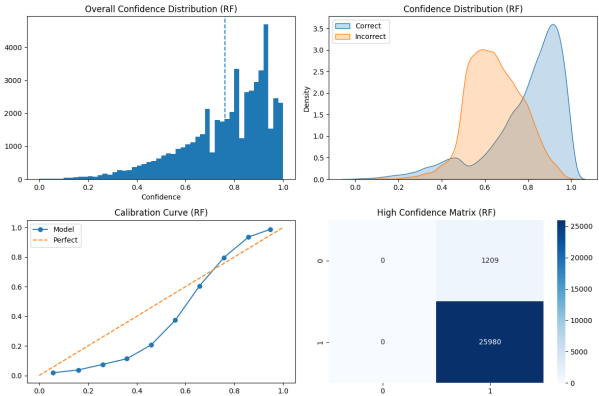


Sumber: Data Sekunder diolah oleh Peneliti (2026)

Gambar 5. Confusion Matrix - Random Forest

Model Random Forest menunjukkan performa yang baik dengan nilai Accuracy 83%. Model memiliki nilai FP yang bagus sebesar 38.008 data. Namun,

model juga menghasilkan FN yang cukup besar sebanyak 8.147 data. Hal ini mengindikasikan bahwa model cukup baik dalam mengenali data positif, tetapi kurang optimal dalam membedakan data negatif.



Sumber: Data Sekunder diolah oleh Peneliti (2026)

Gambar 6. Distribusi nilai Confidence - Random Forest

Berdasarkan hasil distribusi confidence pada Gambar 6, sebagian besar prediksi model berada pada rentang 0,7 hingga 1,0, yang menunjukkan bahwa Random Forest memiliki tingkat keyakinan yang cukup tinggi dalam melakukan klasifikasi sentimen. Namun, distribusi nilai confidence terlihat meningkat secara bertahap dan tidak terkonsentrasi secara ekstrem pada nilai 0 dan 1 seperti yang ditemukan pada model Logistic Regression. Temuan tersebut didukung dengan adanya tumpang tindih yang cukup besar antara distribusi prediksi benar dan salah, terutama pada rentang confidence menengah sekitar 0,5–0,8. Kondisi ini mengindikasikan bahwa nilai confidence yang dihasilkan model belum sepenuhnya mampu membedakan prediksi yang benar dan salah secara konsisten. Selanjutnya, kurva kalibrasi menunjukkan bahwa performa model cukup jauh dari garis diagonal yang merepresentasikan kalibrasi sempurna, sehingga probabilitas yang dihasilkan Random Forest kurang dapat merefleksikan peluang sebenarnya. Adapun hasil High Confidence Matrix menunjukkan bahwa mayoritas prediksi dengan tingkat keyakinan tertinggi berada pada kelas positif sebanyak 25.980 prediksi. Dominasi tersebut mengindikasikan kecenderungan model untuk lebih sering memprediksi kelas positif, yang kemungkinan dipengaruhi oleh ketidakseimbangan distribusi data. Meskipun demikian, tingginya jumlah prediksi positif juga menunjukkan bahwa sebagian besar pengguna memberikan penilaian yang baik terhadap game yang diteliti, sementara sentimen negatif tetap dapat dimanfaatkan sebagai sumber informasi untuk mengidentifikasi aspek yang perlu ditingkatkan oleh Developer.

Hasil Komparasi Algoritma

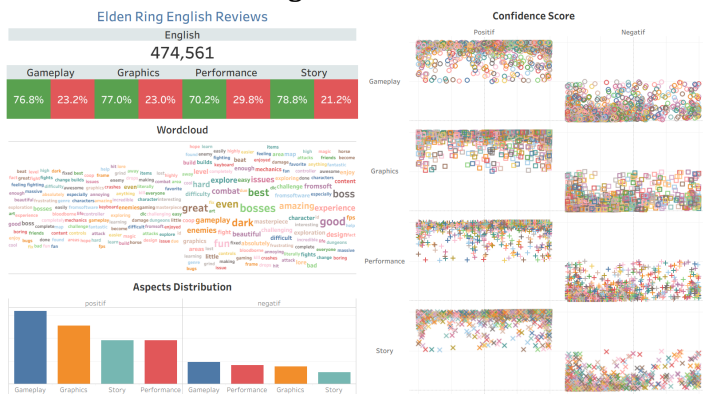
Tabel 2. Komparasi Algoritma

Model	Precision	Recall	F1-Score	Accuracy
LR	0.85	0.83	0.89	0.89
BERT	0.88	0.85	0.86	0.90
RF	0.84	0.67	0.70	0.83

Sumber: Data Sekunder diolah oleh Peneliti (2026)

Tabel 2 menunjukkan perbandingan kinerja tiga model klasifikasi berdasarkan nilai Macro Average, yaitu LR, BERT, dan RF berdasarkan evaluasi Confusion Matrix seperti Precision, Recall, F1-score, dan Accuracy. Pada metrik Precision dan Recall, BERT juga menunjukkan nilai yang relatif lebih tinggi dibandingkan model lainnya, sementara Logistic Regression memiliki nilai F1-score tertinggi, yang menunjukkan keseimbangan antara precision dan recall.

Hasil Dashboard Business Intelligence



Sumber: <http://bit.ly/497ywIE>

Gambar 7. Dashboard Business Intelligence

Hasil tersebut menunjukkan bagaimana pendekatan Business Intelligence dapat mengubah data ulasan pengguna menjadi informasi yang bernilai bagi pengambilan keputusan berbasis data. Melalui visualisasi dashboard, Developer dapat mengidentifikasi aspek-aspek yang menjadi prioritas perbaikan berdasarkan sentimen negatif yang dominan, sekaligus mempertahankan aspek yang telah memperoleh respons positif dari pengguna. Informasi ini memungkinkan perusahaan menyusun strategi pengembangan produk yang lebih terarah dan sesuai dengan kebutuhan pemain. Selain itu, aspek-aspek yang memperoleh sentimen positif tinggi dapat dimanfaatkan sebagai value proposition dalam aktivitas pemasaran dan promosi untuk memperkuat citra produk di pasar. Dashboard juga berperan sebagai sarana knowledge discovery yang membantu perusahaan memahami pola persepsi

pengguna secara lebih mendalam, sehingga proses evaluasi dan peningkatan kualitas produk digital dapat dilakukan secara berkelanjutan dan berbasis bukti empiris.

Pembahasan Sentimen Pemain Terhadap Setiap Aspek

Berdasarkan distribusi aspek ulasan dari penelitian yang telah dilakukan, aspek gameplay memperoleh jumlah sentimen positif tertinggi diikuti oleh aspek graphics. Dominasi sentimen positif pada kedua aspek tersebut menunjukkan bahwa mekanisme game, eksplorasi dunia, sistem pertarungan, serta kualitas visual menjadi faktor utama yang membentuk kepuasan pemain terhadap game. Temuan ini diperkuat oleh visualisasi wordcloud sentimen positif yang menampilkan kata-kata seperti good, great, best, dan amazing, yang menggambarkan apresiasi pemain terhadap pengalaman bermain yang dianggap menarik dan imersif.

Di sisi lain, aspek gameplay dan performance juga menjadi aspek dengan dominasi sentimen negatif tertinggi. Hal tersebut menunjukkan bahwa walaupun gameplay menjadi kekuatan utama permainan, masih terdapat keluhan terkait keseimbangan permainan, tingkat kesulitan, bug, stuttering, serta optimasi performa yang memengaruhi pengalaman pengguna. Visualisasi wordcloud sentimen negatif memperlihatkan kemunculan kata seperti issues, combat, dan enemies, yang mengindikasikan bahwa masalah teknis dan mekanisme pertarungan menjadi sumber utama ketidakpuasan pemain.

Informasi tersebut menjelaskan bahwa analisis sentimen berbasis aspek mampu membantu Developer untuk mengidentifikasi aspek pengalaman pengguna yang paling memengaruhi kepuasan maupun ketidakpuasan pemain (Mujilawati & Ubaydillah, 2024). Jumlah sentimen positif yang lebih banyak dibandingkan sentimen ulasan negatif menunjukkan bahwa elemen-elemen utama permainan berhasil membangun user engagement dan loyalitas pemain. Aspek graphics dan gameplay menunjukkan bahwa kualitas visual serta mekanisme permainan menjadi faktor dominan dalam membentuk pengalaman bermain yang memuaskan. Aspek story yang memperoleh persentase sentimen positif tertinggi mengindikasikan bahwa pengalaman naratif memiliki pengaruh besar terhadap keterikatan emosional pemain (Febrianta et al., 2021).

Di sisi lain, keberadaan sentimen negatif terutama pada aspek performance tetap menjadi perhatian penting bagi Developer. Persentase sentimen negatif yang relatif tinggi ini menunjukkan bahwa pemain masih menghadapi permasalahan teknis seperti lag, stuttering, frame drop, dan bugs yang dapat memengaruhi kenyamanan bermain. Hal tersebut dapat digunakan sebagai dasar pengambilan keputusan strategis untuk menentukan prioritas perbaikan sistem, optimasi perangkat lunak, dan peningkatan kualitas layanan berbasis kebutuhan pengguna (Fadhlurrahman et al., 2023).

Pembahasan Keandalan Metode ABSA-BERT dibandingkan Baseline Logistic Regression dan Random Forest

Berdasarkan hasil evaluasi yang telah dilakukan, terlihat adanya perbedaan performa yang cukup signifikan antara model Logistic Regression, BERT, dan Random Forest dalam melakukan klasifikasi sentimen. Secara umum, model BERT

menunjukkan performa terbaik dari segi Accuracy, diikuti oleh Logistic Regression dan Random Forest. Hal ini menunjukkan bahwa pendekatan berbasis deep learning seperti BERT lebih unggul dalam memahami konteks teks. Dari sisi Precision dan Recall, model BERT juga memiliki nilai tertinggi yang menunjukkan bahwa prediksi positif yang dihasilkan memiliki tingkat ketepatan yang lebih baik dibandingkan model lainnya. Logistic Regression berada di posisi kedua dengan Precision, sedangkan Random Forest memiliki nilai yang lebih rendah. Hal ini mengindikasikan bahwa BERT lebih mampu meminimalkan kesalahan dalam memprediksi kelas positif (Devlin et al., 2018).

Menariknya, pada metrik F1-score, Logistic Regression justru memiliki nilai tertinggi yang mengungguli BERT. Hal ini menunjukkan bahwa Logistic Regression memiliki keseimbangan yang lebih baik antara precision dan recall dibandingkan BERT. Sementara itu, Random Forest kembali menunjukkan performa terendah dengan nilai F1-score, yang mengindikasikan bahwa model ini kurang optimal secara keseluruhan.

Selain itu, perbedaan jumlah support pada model Logistic Regression dan Random Forest menunjukkan bahwa kedua model tersebut menggunakan jumlah data latih yang lebih besar dibandingkan model BERT. Hasil tersebut menunjukkan bahwa BERT memiliki kemampuan yang lebih unggul dalam memahami konteks semantik, hubungan antar kata, dan arsitektur transformer. Dengan kemampuan tersebut, BERT dapat menghasilkan representasi bahasa yang lebih kompleks dan kontekstual sehingga mampu meningkatkan akurasi klasifikasi sentimen.

Secara keseluruhan, dapat disimpulkan bahwa model BERT merupakan model dengan performa terbaik dalam penelitian ini, terutama dari segi Accuracy, Precision, dan Recall. Hasil tersebut sejalan dengan berbagai penelitian terdahulu yang menunjukkan bahwa model berbasis Transformer, khususnya BERT, mampu menangkap konteks semantik dan hubungan antar kata secara lebih baik dibandingkan pendekatan machine learning tradisional sehingga menghasilkan performa klasifikasi sentimen yang lebih tinggi pada berbagai domain teks (Bećirović, F., et al., 2024). Namun, Logistic Regression tetap menjadi alternatif yang kompetitif karena memiliki keseimbangan metrik yang sangat baik serta kalibrasi probabilitas yang lebih stabil, sehingga disarankan untuk menerapkan Expected Calibration Error (ECE), agar dapat diketahui perbedaan antara tingkat kepercayaan model dan akurasi aktual pada berbagai rentang probabilitas prediksi (Filho T. S. et al., 2023). Sementara itu, Random Forest menunjukkan performa yang paling rendah di antara ketiga model, terutama dalam hal Recall dan F1-score, sehingga kurang direkomendasikan untuk digunakan pada kasus klasifikasi ini.

Pembahasan dalam Perspektif Business Intelligence

Penelitian ini menunjukkan implementasi konsep Business Intelligence melalui proses transformasi data ulasan pengguna menjadi informasi strategis yang mudah dianalisis. Dalam konteks industri game, informasi ini sangat penting karena Developer dapat mengetahui aspek mana yang menjadi kekuatan utama produk dan aspek mana yang memerlukan optimalisasi lebih lanjut. Aspek gameplay dan graphics

yang memperoleh dominasi sentimen positif menunjukkan bahwa kualitas permainan dan visual berhasil memenuhi ekspektasi pemain, sedangkan tingginya sentimen negatif pada aspek performance mengindikasikan perlunya peningkatan optimasi sistem dan stabilitas permainan (Shukla et al., 2026).

Dari perspektif Business Intelligence, dashboard yang dihasilkan dalam penelitian ini berperan sebagai alat pendukung keputusan yang mampu menyederhanakan data tidak terstruktur menjadi indikator performa yang mudah dipahami. Visualisasi distribusi aspek membantu proses identifikasi pola opini pengguna secara cepat, sedangkan wordcloud memberikan gambaran mengenai kata kunci dominan yang membentuk persepsi pengguna terhadap game. Selain itu, grafik confidence score menunjukkan bahwa model memiliki tingkat kepercayaan prediksi yang tinggi, sehingga hasil analisis dapat digunakan sebagai dasar evaluasi yang reliabel. Implementasi dashboard BI pada penelitian ini sejalan dengan penelitian terdahulu yang menyatakan bahwa kombinasi ABSA dan visual analytics mampu meningkatkan efektivitas interpretasi data pelanggan dalam mendukung strategi peningkatan kualitas produk. Dengan demikian, dashboard yang dihasilkan tidak hanya memberikan informasi deskriptif mengenai sentimen pengguna, tetapi juga menghasilkan wawasan analitis yang dapat digunakan sebagai dasar penyusunan strategi pengembangan game secara berkelanjutan (Suhdi & Siswoyo, 2025).

V. KESIMPULAN DAN REKOMENDASI

Berdasarkan hasil penelitian pada Dataset Steam Reviews of Elden Ring, dapat disimpulkan bahwa analisis sentimen berbasis aspek mampu memberikan gambaran menyeluruh mengenai pengalaman pemain terhadap permainan. Aspek gameplay, graphics, performance, dan story secara umum didominasi oleh sentimen positif, dengan aspek gameplay menjadi faktor utama yang paling memengaruhi kepuasan pengguna, diikuti oleh kualitas visual dan pengalaman naratif permainan. Dominasi sentimen positif tersebut diperkuat melalui visualisasi wordcloud yang menampilkan kata-kata bernuansa apresiatif seperti good, great, best, dan amazing, yang menunjukkan bahwa pemain menilai mekanisme permainan, eksplorasi dunia, sistem pertarungan, dan kualitas grafis mampu memberikan pengalaman bermain yang menarik serta imersif. Selain itu, aspek story juga menunjukkan pengaruh yang kuat terhadap keterikatan emosional pemain sehingga turut meningkatkan user engagement dan loyalitas pengguna. Meskipun demikian, aspek gameplay dan performance juga menjadi sumber sentimen negatif yang cukup dominan karena masih ditemukan berbagai kendala teknis seperti lag, stuttering, frame drop, bug, serta ketidakseimbangan mekanisme pertarungan yang memengaruhi kenyamanan bermain. Temuan ini diperkuat oleh kemunculan kata-kata seperti issues, combat, dan enemies pada wordcloud sentimen negatif yang mengindikasikan adanya permasalahan teknis dan mekanisme permainan yang masih perlu diperbaiki.

Evaluasi kinerja klasifikasi pada algoritma BERT menunjukkan keunggulan dalam Accuracy jika dibandingkan dengan algoritma Logistic Regression dan Random Forest. BERT mencapai Accuracy 90%, sedangkan Logistic Regression mencapai

Accuracy 89%, dan Random Forest mencapai Accuracy 83%. Oleh karena itu, BERT lebih unggul dalam mengklasifikasikan data ulasan game terutama dalam hal prediksi. Meskipun dari segi nilai confidence, Logistic Regression menunjukkan nilai yang lebih stabil dibandingkan BERT. Namun, nilai confidence yang tinggi tidak selalu menunjukkan kualitas pemahaman konteks yang lebih baik, terutama pada data ulasan panjang yang memiliki variasi makna dan struktur kalimat yang kompleks.

Dashboard hasil analisis ABSA-BERT pada penelitian ini menunjukkan penerapan konsep Business Intelligence melalui proses pengolahan data ulasan pengguna menjadi informasi strategis yang mudah dipahami dan dianalisis oleh Developer. Ini memungkinkan Developer untuk mengidentifikasi aspek yang menjadi kekuatan utama produk maupun aspek yang masih memerlukan optimalisasi lebih lanjut. Dominasi sentimen positif pada aspek gameplay dan graphics menunjukkan bahwa mekanisme permainan serta kualitas visual telah mampu memenuhi ekspektasi pemain dan menciptakan pengalaman bermain yang memuaskan. Sebaliknya, tingginya sentimen negatif pada aspek performance mengindikasikan masih adanya kendala teknis yang berkaitan dengan optimasi sistem dan stabilitas permainan sehingga memerlukan perhatian lebih lanjut. Dari perspektif Business Intelligence, dashboard ini berfungsi sebagai alat pendukung keputusan yang mampu mentransformasikan data tidak terstruktur menjadi indikator performa yang lebih informatif dan mudah diinterpretasikan. Visualisasi distribusi aspek membantu proses identifikasi pola opini pengguna secara cepat, sementara wordcloud memberikan gambaran mengenai kata-kata dominan yang membentuk persepsi pemain terhadap permainan. Selain itu, visualisasi confidence score menunjukkan bahwa model memiliki tingkat keyakinan prediksi yang tinggi sehingga hasil analisis dapat digunakan sebagai dasar evaluasi yang reliabel. Implementasi dashboard ini juga memperlihatkan bahwa kombinasi metode ABSA dan visual analytics mampu meningkatkan efektivitas interpretasi data pelanggan dalam mendukung strategi peningkatan kualitas layanan maupun pengembangan produk digital secara berkelanjutan. Dengan demikian, dashboard yang dihasilkan tidak hanya menyajikan informasi deskriptif mengenai sentimen pengguna, tetapi juga menghasilkan wawasan analitis yang dapat dimanfaatkan sebagai dasar pengambilan keputusan strategis dalam pengembangan game di masa mendatang.

Penelitian berikutnya disarankan untuk menerapkan teknik model kalibrasi agar probabilitas yang dihasilkan lebih merepresentasikan tingkat akurasi sebenarnya. Selain itu, perlu dilakukan proses kalibrasi pada output softmax, misalnya dengan metode seperti temperature scaling atau pendekatan lainnya, sehingga distribusi probabilitas menjadi lebih seimbang. Di samping itu, proses training juga dapat dikembangkan dengan mempertimbangkan aspek uncertainty, sehingga model tidak hanya fokus pada akurasi prediksi tetapi juga mampu mengukur tingkat ketidakpastian secara lebih baik. Terakhir, evaluasi terhadap kualitas kalibrasi model perlu dilakukan secara khusus menggunakan metrik seperti

Expected Calibration Error (ECE), sehingga dapat diketahui sejauh mana kesesuaian antara confidence dan performa aktual model.

REFERENSI

- Bećirović, F., Merdović, N., Hiyari, M. H., Smajlhodžić-Deljo, M., Salkić, I., & Softić, A. (2024, September 1). Assessing the Accuracy of Logistic Regression and Bert in Sentiment Analysis and Mental Disorder Classification. *Artificial Intelligence in Industry 4.0: The future that comes true*, 317-326. <https://doi.org/10.5644/PI2024.215.19>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Cox, D. R. (1958). THE REGRESSION ANALYSIS OF BINARY SEQUENCES. *Journal of the Royal Statistical Society SERIES B (METHODOLOGICAL)*, 20(2), 215–232. <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. <https://doi.org/10.48550/arXiv.1810.04805>
- ELDEN RING on Steam. (2025). Steam. Retrieved November 29, 2025, from https://store.steampowered.com/app/1245620/ELDEN_RING/
- Fadhurrahman, J. A. M., Herawati, N. A., Aulya, H. R. W., Puspasari, I., & Utama, N. P. (2023). Sentiment Analysis of Game Reviews on STEAM using BERT, BiLSTM, and CRF. *International Conference on Electrical Engineering and Informatics (ICEEI)*. <https://doi.org/10.1109/ICEEI59426.2023.10346219>
- Fauzan, N. A. (2025, June 30). A Zero-Shot Aspect-Based Sentiment Analysis of Public Perception Toward AI Chatbots. *TIERS Information Technology Journal*, 6(1), 97-111. <https://doi.org/10.38043/tiers.v6i1.6488>
- Febrianta, M. Y., Widiyanesti, S., & Ramadhan, S. R. (2021, Oktober). Analisis Ulasan Indie Video Game Lokal pada Steam Menggunakan Analisis Sentimen dan Pemodelan Topik Berbasis Latent Dirichlet Allocation. *Journal of Animation & Games Studies*, 7(2). <https://doi.org/10.24821/jags.v7i2.5162>
- Gardazi, N. M., Daud, A., Malik, M. K., Bukhari, A., Alsahfi, T., & Alshemaimri, B. (2025). BERT applications in natural language processing: a review. *Artificial Intelligence Review*, 58(166). <https://doi.org/10.1007/s10462-025-11162-5>
- Hua, Y. C., Denny, P., Wicker, J., & Taskova, K. (2024). A systematic review of aspect-based sentiment analysis: domains, methods, and trends. *Artificial Intelligence Review*, 57, 296. <https://doi.org/10.1007/s10462-024-10906-z>

- Imron, S., Setiawan, E. I., Santoso, J., & Purnomo, M. H. (2023). Aspect Based Sentiment Analysis Marketplace Product Reviews Using BERT, LSTM, and CNN. *Artificial Intelligence*, 7(3). <https://doi.org/10.29207/resti.v7i3.4751>
- Jonnala, N. S., Teja, A.V.S. R., Rajeswari, S. R., Jakeer, S., Dheeraj, A., Bansal, S., Prakash, K., Singh, S., Faruque, M. R. I., & Al-mugren, K.S. (2025, Jan 08). Leveraging hybrid model for accurate sentiment analysis of Twitter data. *Scientific Reports*, 15(24438). <https://doi.org/10.1038/s41598-025-09794-2>
- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural Language Processing: State of The Art, Current Trends and Challenges. *Multimedia Tools and Applications*, 82, 3713-3744. <https://doi.org/10.1007/s11042-022-13428-4>
- Lee, J., & Novak, J. (2022). *Game Development Essentials: An Introduction* (4th Edition) (J. Lee & J. Novak, Eds.). NOVY UNLIMITED, Incorporated d/b/a NOVY PUBLISHING. https://books.google.co.id/books?hl=en&lr=&id=8N0LEQAAQBAJ&oi=fnd&pg=PR2&dq=game&ots=WfBYLSLimj&sig=9lt4mePLeZG_2D2q9ScpsTyxTiA&redir_esc=y#v=onepage&q=game&f=false
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-02145-9>
- Lorenzo, L. (2025). *Steam Reviews of Elden Ring*. Kaggle. <https://www.kaggle.com/datasets/lorenzoshylckl/steam-reviews-of-elden-ring/data>
- Mantena, M. M. K. V., & Murugan, G. (2025). Integration of machine learning techniques in the field of business information system: Study. *AIP Conference Proceedings*. <https://doi.org/10.1063/5.0247131>
- Mirea, B., & Grădinaru, G. (2025). The Business Value of Text Analysis and Topic Modeling: Evaluating Sentiment Shifts in Mobile Game Reviews after Updates. *Proceedings of the International Conference on Business Excellence*, 19(1). <https://doi.org/10.2478/picbe-2025-0082>
- Mujilawati, S., & Ubaydillah, M. A. (2024, November). Analisis Sentimen Berbasis Aspek pada Ulasan Game Mobile Legends dengan Pendekatan Decision Tree untuk Evaluasi Pengalaman Pengguna. *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, 4(11). <https://doi.org/10.52436/1.jpti.465>
- Newzoo. (2025, September 9). *Newzoo's Global Games Market Report 2025 | Free version*. Newzoo. Retrieved October 22, 2025, from <https://newzoo.com/resources/trend-reports/newzoo-global-games-market-report-2025>

- Paramesha, M., Rane, N. L., & Rane, J. (2024). Big Data Analytics, Artificial Intelligence, Machine Learning, Internet of Things, and Blockchain for Enhanced Business Intelligence. *Partners Universal Multidisciplinary Research Journal (PUMRJ)*, 01(02). <https://dx.doi.org/10.2139/ssrn.4855856>
- Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., & Androutsopoulos, I. (2015). SemEval-2015 task 12: Aspect based sentiment analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)* (pp. 486–495). Association for Computational Linguistics. <https://doi.org/10.18653/v1/S15-2082>
- Purbolaksono, M. D. (2024). Sentiment Analysis of Game Review in Steam Platform using Random Forest. *International Journal on Information and Communication Technology (IJoICT)*, 10(2), 161–169. <https://doi.org/10.21108/ijoict.v10i2.1007>
- Salsabila, Tyas, S. M. P., Romadhona, Y., & Purwitasari, D. (2023, April). Aspect-based Sentiment and Correlation-based Emotion Detection on Tweets for Understanding Public Opinion of Covid-19. *Journal of Information Systems Engineering and Business Intelligence*, 9(1). <http://dx.doi.org/10.20473/jisebi.9.1.84-94>
- Sharma, Y., Thompson, C., Mangoni, A. A., Shahi, R., Horwood, C., & Woodman, R. (2025). Performance of Machine Learning Models in Predicting 30-Day General Medicine Readmissions Compared to Traditional Approaches in Australian Hospital Setting. *Healthcare*, 13(11), 1223. <https://doi.org/10.3390/healthcare13111223>
- Shukla, P., Kumar, R., Dwivedi, V. K., & Singh, A. K. (2026). Aspect based sentiment analysis: A systematic review, taxonomy, applications, and future research directions. *Computer Science Review*, 61. <https://doi.org/10.1016/j.cosrev.2026.100924>
- Shuvo, S. A., Tabassum, M., Tafannum, N., & Chadni, S. (2025, June). MACHINE LEARNING IN BUSINESS INTELLIGENCE: FROM DATA MINING TO STRATEGIC INSIGHTS IN MIS. *Green Technologies and Circular Economy Innovations*, 4(02), 339-369. <https://doi.org/10.63125/dr8py41>
- Starbuck, C. (2023). *The Fundamentals of People Analytics: With Applications in R*. Springer International Publishing. https://doi.org/10.1007/978-3-031-28674-2_12
- SteamDB. (2025, Agustus 22). *ELDEN RING Steam Charts · SteamDB*. SteamDB. <https://steamdb.info/app/1245620/charts/#max>
- Filho, T. S., Song, H., Perello-Nieto, M., Santoz-Rodriguez, R., Kull, M., Flach, P. (2023). Classifier calibration: a survey on how to assess and improve predicted class

- probabilities. *Machine Learning* 112, 3211–3260.
<https://doi.org/10.1007/s10994-023-06336-7>
- Suhdi, & Siswoyo, D. A. (2025). Analisis Sentimen Berbasis Aspek dan Visualisasi Business Intelligence Pada Ulasan Pelanggan Produk Minuman Industri F&B Dari Instagram dan Google Review. *Karapan Network Journal*, 02(01).
<https://ejournal.omahtabing.com/knj/en/article/view/94>
- Talaat, A. S. (2023). Sentiment analysis classification system using hybrid BERT models. *Journal of Big Data*, 10(110). <https://doi.org/10.1186/s40537-023-00781-w>
- Trisna, K. W., & Jie, H. J. (2022). Deep Learning Approach for Aspect-Based Sentiment Classification: A Comparative Review. *Applied Artificial Intelligence*, 36, 1193. <https://doi.org/10.1080/08839514.2021.2014186>
- Yu, Y., Dinh, D.-T., Nguyen, B.-H., Yu, F., & Huynh, V.-N. (2023, June 14). Mining Insights From Esports Game Reviews With an Aspect-Based Sentiment Analysis Framework. *IEEE Access*, 11, 61161 - 61172.
<https://doi.org/10.1109/ACCESS.2023.3285864>
- Zhao, L., & Zhang, J. (2024). Machine learning based business intelligence security and privacy analysis with gaming model in training complexity application. *Entertainment Computing*, 50.
<https://doi.org/10.1016/j.entcom.2024.100695>