ANALISIS SENTIMEN TERHADAP GAME PALWORLD DI STEAM MENGGUNAKAN ALGORITMA BIDIRECTIONAL ENCODER REPRESENTATION FROM TRANSFORMERS

Muhammad Fiqri Faturrian, Jajam Haerul Jaman, Iqbal Maulana
Program Studi Informatika, Universitas Singaperbangsa Karawang
Jl. HS. Ronggo Waluyo, Telukjambe Timur, Karawang, Indonesia
2010631170096@student.unsika.ac.id

ABSTRAK

Industri video *game* terus berkembang pesat, dengan lebih dari 3,09 miliar pemain aktif secara global. Salah satu *game* yang menarik perhatian adalah *Palworld*, yang mencapai 15 juta unit penjualan pada Februari 2024 namun menghadapi kontroversi plagiarisme. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna *Palworld* di platform *Steam* menggunakan algoritma BERT, guna menganalisis efek dari kontroversi tersebut terhadap sikap dan kepuasan pemain terhadap *game*. Data sebanyak 10.020 ulasan dikumpulkan dengan menggunakan teknik *web scraping* melalui *Steam API* dari bulan Januari hingga November 2024. Pemodelan dilakukan dengan pembagian data training, data testing, dan data validation sebesar 80:10:10. *Hyperparameter* yang digunakan mencakup *batch size* 16, *learning rate* 0,00002, dan *epoch* 10. Evaluasi model menggunakan *confusion matrix* dengan metrik *accuracy, precision, recall*, dan *f1-score*. Hasil menunjukkan pembagian data 80:10:10 menghasilkan model dengan akurasi 86%, *precision* 86%, *recall* 86%, dan *f1-score* 86%.

Kata kunci : *Analisis Sentimen, Palworld, Steam, BERT, Wordcloud*

1. PENDAHULUAN

Industri *video game* telah berkembang pesat dan menjadi salah satu segmen paling menguntungkan dalam industri hiburan, didukung oleh kemajuan teknologi yang meningkatkan aksesibilitas dan popularitas[1]. Pada tahun 2024, jumlah pemain *game* global diperkirakan mencapai 3,32 miliar[2]. Transformasi industri ini mencakup evolusi dari mesin arkade hingga platform digital modern seperti *Steam*, yang memudahkan distribusi dan promosi *game* secara *online*[3].

Palworld, *game survival* yang dirilis melalui akses awal di *Steam* pada Januari 2024, mencatat kesuksesan besar dengan penjualan delapan juta unit dalam enam hari pertama dan lebih dari dua juta pemain aktif bersamaan. Pada Februari 2024, total penjualannya melampaui 15 juta unit, menjadikannya salah satu game PC terlaris. Namun, keberhasilan ini diwarnai kontroversi terkait tuduhan plagiarisme desain *Pokémon*, yang memicu diskusi di komunitas gaming. Meski menuai pujian atas mekanika inovatif dan grafis menarik, tuduhan tersebut menimbulkan reaksi negatif dari sebagian pemain. Dengan memeriksa berbagai tanggapan dan pendapat dari komunitas gaming online, kita dapat menelaah dampak kontroversi tersebut terhadap persepsi dan penerimaan terhadap *Palworld*. Analisis sentimen memberikan informasi berharga bagi pengembang game untuk memahami respons pasar dan mengambil tindakan terkait isu-isu yang muncul. Ini krusial untuk memahami pandangan pelanggan terhadap pengalaman bermain. Membaca ulasan dari berbagai platform membantu menentukan arah pengembangan dan perbaikan iterasi game selanjutnya[4].

Berdasarkan penelitian yang dilakukan oleh Carvajal dan Garrido menyebutkan bahwa Salah satu algoritma analisis sentimen yang populer dan memiliki kinerja unggul saat ini adalah *Bidirectional Encoder Representations from Transformers* (BERT). Teknik ini mengatasi keterbatasan metode tradisional dengan meningkatkan pemahaman kontekstual bahasa. BERT menyediakan model pra-latih yang dapat diadaptasi untuk berbagai aplikasi dengan mudah, tersedia dalam berbagai bahasa, dan mampu memproses teks secara efektif meskipun membutuhkan waktu pelatihan yang lebih lama. Keakuratan dan kemampuannya menangkap nuansa bahasa mirip manusia menjadikannya salah satu model *Natural Language Processing* (NLP) yang paling banyak digunakan. BERT hanya memerlukan data minimal dan sedikit penyesuaian untuk memberikan hasil unggul di berbagai tugas NLP[5].

Pada penelitian sebelumnya yang dilakukan oleh Kang, analisis sentimen digunakan untuk menilai sentimen masyarakat terhadap *Malaysian Airlines* menggunakan enam model yang berbeda, termasuk *Linear Support Vector Classifier, BERT Model, Ensemble Method, Multinomial Naïve Bayes, Bi-LSTM,* dan *Random Forest.* Hasil penelitian menunjukkan bahwa pendekatan *deep learning* menunjukkan kinerja yang lebih baik dibandingkan dengan metode pembelajaran mesin tradisional. Secara khusus *Bi-LSTM* mencapai tingkat akurasi sebesar 77%, sementara model BERT melampaui semua model lain dengan akurasi 86%[6].

Dengan demikian, penelitian ini mengonfirmasi bahwa model BERT, mampu memberikan performa yang sangat baik dalam analisis sentimen, sejalan dengan temuan penelitian sebelumnya yang

menunjukkan superioritas model BERT dibandingkan dengan metode tradisional dan beberapa model *deep learning* lainnya. Peneliti melakukan penyesuaian *hyperparameters* untuk mencapai hasil optimal. *Hyperparameters* yang digunakan dalam penetian ini yaitu *learning rate* 0,00002, *batch size* 16 dan *epoch* 10. Dataset ulasan game *Palworld* dikumpulkan dari *Steam API* berdasarkan ulasan terbaru. Karena dataset tidak memiliki label, Dalam penelitian ini data ulasan akan dilabeli menjadi positif dan negatif, Dengan asumsi hasil penilaian pengguna yang mengatakan "Recommended" akan dilabeli positif, sedangkan yang "Not Recommended" akan dilabeli negatif. Model BERT dipilih karena penelitian sebelumnya menunjukkan bahwa BERT memiliki tingkat akurasi yang tinggi. Tidak seperti algoritma pemrosesan bahasa lainnya, BERT mempelajari konteks kata berdasarkan kata-kata di sekitarnya, memproses konteks penuh dengan melihat pola sebelum dan sesudah kata tersebut.

2. TINJAUAN PUSTAKA

2.1. *Text Mining*

Text mining adalah proses ekstraksi pengetahuan implisit dari data tekstual yang tidak tersimpan secara eksplisit [7]. Berbeda dengan informasi yang diperoleh dari penyimpanan data, pengetahuan ini ditemukan melalui tugas-tugas seperti klasifikasi teks, pengelompokan, dan asosiasi. *Text mining* merupakan varian dari *data mining*, dengan fokus pada teks sebagai data tak terstruktur yang terdiri dari string atau kata-kata [8]. Dalam konteks ini, teks merujuk pada artikel yang terdiri dari paragraf dan ditulis dalam bahasa alami, tidak termasuk kode sumber atau persamaan matematika.

Teknik *text mining* menganalisis jumlah besar teks dalam bahasa alami untuk mendeteksi pola leksikal dan mengekstrak informasi yang berguna. Ada beberapa tahap dalam text mining, yang meliputi pengumpulan dokumen, pra-pemrosesan, transformasi teks, seleksi fitur, seleksi pola, dan evaluasi [9].

2.2. Analisis Sentimen

Analisis sentimen atau opinion mining adalah salah satu cabang *Natural Language Processing* (NLP) yang memfokuskan diri pada evaluasi pendapat, perilaku, penilaian, dan emosi individu terkait dengan berbagai entitas seperti produk, topik, layanan, organisasi, individu, atau kegiatan lainnya yang diekspresikan dalam bentuk teks [10]. Dalam proses analisis sentimen, teks subjektif dapat dikelompokkan ke dalam berbagai klasifikasi, termasuk positif, negatif, netral, atau dijelaskan dalam bentuk emosi seperti kegembiraan, kemarahan, atau kesedihan [11].

Berbagai tugas dapat dilakukan dalam analisis sentimen, antara lain klasifikasi subjektivitas, klasifikasi sentimen, deteksi spam opini, deteksi bahasa implisit, dan ekstraksi aspek. Terdapat tiga pendekatan utama dalam analisis sentimen, yaitu *machine learning* (*unsupervised dan semi-supervised*), berbasis lexicon (berbasis korpus dan berbasis kamus), dan pendekatan *hybrid* yang menggabungkan keduanya [12].

2.3. *Bidirectional Encoder Representation from Tranformers* (BERT)

Bidirectional Encoder Representation from Transformers (BERT) merupakan sebuah algoritma baru dalam bidang *Natural Language Processing* (NLP) yang dikembangkan oleh Google. Pertama kali diperkenalkan oleh para peneliti di Google AI pada tahun 2018, BERT telah terbukti efektif dalam menangani berbagai tugas NLP seperti menjawab pertanyaan, menyimpulkan bahasa alami, klasifikasi, dan evaluasi pemahaman bahasa secara umum. BERT memiliki dua versi utama, yakni BERT-base dan BERT-large. BERT-base memiliki 12 *layers* (L=12), 768 ukuran *embedding* (H=768), 12 *attention head* (A=12), dengan total parameter sebanyak 110 juta. Sedangkan BERT-large memiliki 24 lapisan (L=24), 1024 ukuran *embedding* (H=1024), 16 *attention head* (A=16), dengan total parameter sebanyak 340 juta [13].

2.4. *BERT-base*

Model BERT dapat dijalankan melalui dua tahap. Tahap pertama melibatkan *pre-training*, di mana model memahami data teks *input* dan konteksnya, sedangkan tahap kedua melibatkan *fine-tuning*, dimana model menyesuaikan diri dan mengidentifikasi solusi. *Fine-tuning* model BERT yang sudah dilatih dapat dilakukan dengan menambahkan satu lapisan untuk menghasilkan hasil yang lebih mutakhir.

2.5. *Hyperparameter*

Hyperparameter adalah parameter yang tidak teratur oleh mesin dan tidak memerlukan pengujian. *Hyperparameter* diatur secara manual sesuai dengan kebutuhan peneliti untuk membantu memperkirakan parameter model. *Hyperparameter* yang digunakan dalam penelitian ini meliputi:

a. *Batch Size*
   Batch size adalah *hyperparameter* yang menentukan jumlah sampel yang diproses sebelum bobot model diperbaharui. Proses satu batch akan memakan waktu lebih lama jika ukuran batch lebih besar [14].
b. *Epoch*
   Epoch adalah *hyperparameter* yang menentukan berapa kali jaringan neural memproses data pelatihan. Satu *epoch* menunjukan bahwa algoritma *deep learning* telah mempelajari seluruh dataset pelatihan [15].
c. *Learning Rate*
   Learning rate adalah *hyperparameter* yang digunakan untuk mengatur nilai pembaruan bobot selama proses pelatihan. Semakin besar nilai learning rate, semakin cepat proses pelatihan berlangsung.

2.6. *Confusion Matrix*

Confusion matrix adalah matriks dengan ukuran M x N yang digunakan untuk mengevaluasi performa model klasifikasi, di mana N merepresentasikan jumlah kelas target [16]. Matriks ini membandingkan nilai target aktual dengan nilai yang diprediksi oleh model machine learning [17]. Hal ini memberikan penjelasan holistik tentang seberapa baik kinerja model klasifikasi yang dibangun dan jenis-jenis kesalahan.

Tabel 1 Kriteria Confusion Matrix

Actual Class	***Predicted Class***	
	Positive	***Negative***
Positive	*True Positive (TP)*	*False Negative (FN)*
Negative	*False Positive (FP)*	*True Negative (TN)*

Keterangan:

a. *True Positive* (TP) mengacu pada data yang memiliki kondisi aktual positif dan diprediksi dengan benar.
b. *True Negative* (TN) merujuk pada data negatif yang diprediksi secara akurat.
c. *False Positive* (FP) terjadi ketika data yang seharusnya negatif diprediksi sebagai positif.
d. *False Negative (FN) terjadi ketika data yang seharusnya positif diprediksi sebagai negatif.*

Terdapat beberapa metrik evaluasi pada *confusion matrix* yaitu:

a. Accuracy

$$Accuracy = \frac{TN + TP}{TN + FN + TP + FP} \times 100\% \quad (1)$$

b. Precision

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (2)$$

c. Recall

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

d. F1-score

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% \quad (4)$$

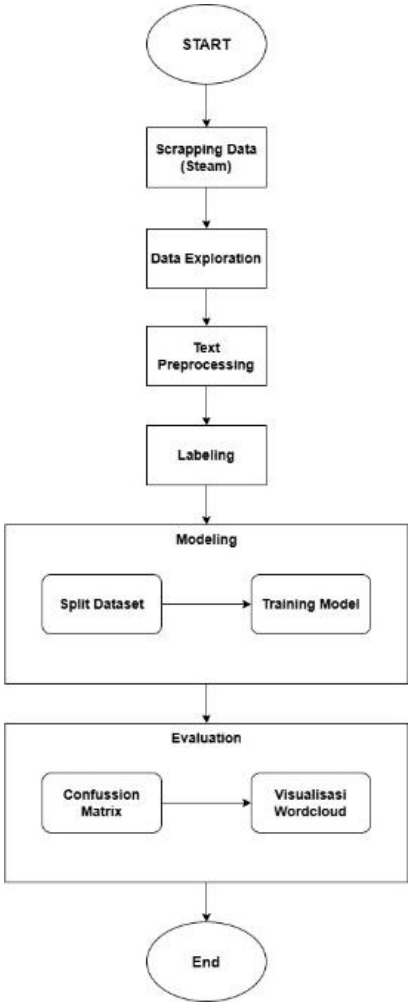
2.2. *Wordcloud*

Pendekatan visualisasi data yang dikenal sebagai *wordcloud* sangat populer dan sudah umum digunakan. Dalam visualisasi *wordcloud*, ukuran *font* yang digunakan untuk kata akan berbanding lurus dengan berapa kali kata tersebut muncul dalam sekelompok dokumen. Artinya, semakin sering suatu subjek disebutkan, semakin besar ukuran yang akan muncul dalam visualisasi [18].

Wordcloud merupakan bagian penting dari data mining, yang belakangan ini semakin mendapatkan perhatian dan menyediakan lebih banyak opsi aplikasi. Saat ini, terdapat berbagai generator *wordcloud online* yang dapat diakses oleh pengguna dengan kebutuhan dasar, seperti mengulang frasa tertentu atau mengumpulkan data teks dari situs web. Pengguna dapat memanfaatkan generator *wordcloud* ini [19].

3. METODE PENELITIAN

Adapun langkah-langkah dalam penelitian ini mengikuti diagram alur yang menjelaskan setiap tahap yang dilakukan selama penelitian, seperti yang ditunjukan pada Gambar 1 berikut.



Gambar 1. Tahapan Penelitian

3.1. Pengumpulan Data

Dalam penelitian ini, data yang akan diambil berasal dari ulasan pengguna *game Palworld* di platform *Steam*. Data diambil menggunakan teknik *scraping* melalui *Steam API*. yang memungkinkan peneliti untuk mengumpulkan ulasan secara otomatis dari platform tersebut.

3.2. *Data Exploration*

Tahap selanjutnya yaitu memahami setiap kolom/*feature* yang ada pada dataset.

3.3. *Text Preprocessing*

Tahap awal *preprocessing* teks bertujuan untuk membersihkan dataset dari gangguan agar lebih terstruktur. Proses ini menggunakan modul NLTK di *Python*. Langkah-langkah yang dilakukan meliputi:

a. *Remove Duplicate:* Menghapus data ulasan yang duplikat berdasarkan *review_id* agar tidak ada ulasan ganda dari pengguna yang sama.
b. *Cleaning:* Menghilangkan karakter tidak valid seperti tanda baca, angka, simbol, dan emoji untuk memudahkan pengolahan data.
c. *Case folding:* Menyeragamkan semua teks menjadi huruf kecil.
d. *Tokenizing:*Memisahkan kata-kata dalam kalimat.
e. *Remove stopwords:* Menghapus kata-kata yang sering muncul tetapi tidak penting, seperti kata penghubung atau kata ganti.
f. *Stemming:* Mengubah kata-kata menjadi bentuk dasar dengan menghapus imbuhan.
g. *Remove missing value:* Menghapus ulasan yang kosong atau hanya berisi karakter khusus setelah proses pembersihan, karena ulasan tersebut akan bernilai *NaN*.

3.4. *Labeling*

Dalam penelitian ini, proses pelabelan ditentukan berdasarkan hasil ulasan Pengguna. Dengan asumsi hasil "Not Recommended" adalah negatif, dan "Recommended" adalah positif.

3.5. *Modeling*

Pada fase ini merupakan bagian terpenting dari proses penelitian. Sebelum memulai proses pemodelan, pembagian dataset dilakukan dengan memisahkan data menjadi tiga bagian: *data train, data validation, dan data test*. Pembagian ini penting untuk menentukan data yang akan digunakan dalam pelatihan dan pengujian model. Proporsi pembagian data pada penelitian ini yaitu 80:10:10. Dimana pembagiannya adalah 80% untuk *data train*, 10% untuk *data validation*, dan 10% untuk *data test*.

Selanjutnya, dalam pelatihan model *BERT-base*, diperlukan definisi parameter yang berbeda untuk penelitian ini. *Hyperparameter* yang digunakan dalam penelitian ini yaitu *Batch size* 16, *epoch* 10, dan *learning rate* 0,00002.

3.6. *Evaluation*

Proses evaluasi model bertujuan untuk mengukur seberapa baik model dalam melakukan klasifikasi pada berbagai kelas. Salah satu metode evaluasi yang digunakan dalam penelitian ini adalah *confusion matrix*, yang merupakan alat pengukuran kinerja model pada tugas klasifikasi *machine learning* dengan *output* berupa dua kelas atau lebih. Selanjutnya, pada tahap evaluasi ini, hasil prediksi analisis sentimen untuk setiap kalimat dalam dataset dievaluasi. Nilai akurasi tertinggi yang diperoleh selama proses pelatihan model sebelumnya dijadikan sebagai akurasi model. Terdapat empat kinerja evaluasi yang digunakan, yaitu f1-score, precision, recall, dan accuracy yang sering digunakan dalam tugas klasifikasi teks dan analisis sentiment. kemudian dilakukan visualisasi data menggunakan wordcloud yang bertujuan untuk melihat sebaran kata yang terdapat pada ulasan yang berlalel positif dan negatif. Nantinya, akan terlihat sebaran kata yang sering muncul pada setiap ulasan baik ulasan yang berlabel positif dan berlabel negatif.

4. HASIL DAN PEMBAHASAN

4.1. *Scraping Data*

Sumber data yang digunakan dalam penelitian ini berasal dari ulasan pengguna *game Palworld* di platform *Steam*. Data diambil menggunakan teknik scraping melalui *Steam API,* yang memungkinkan peneliti untuk mengumpulkan ulasan secara otomatis dari platform tersebut. Rentang waktu pengambilan data adalah sejak *Palworld* dirilis hingga penelitian ini dilakukan (Januari 2024 – November 2024).

Jumlah dataset ditentukan dengan mengambil masing-masing hasil ulasan pengguna (Positif atau Negatif) sebanyak 5.000 atau dengan jumlah total 10.020 data. Setiap ulasan diberi label berdasarkan hasil ulasan yang diberikan oleh pengguna, yaitu hasil "Not Recommended" adalah negatif, dan "Recommended" adalah positif. Data yang terdiri dari 10.020 baris ini memiliki 6 kolom, yaitu *review_id, Review, Recommendation, User_id, play_hours, dan helpful_votes*.

	A	B	C	D	E	F
1	review_id	Review	Recommendation	User_id	play_hour	helpful_vote
2	156721663	dont get sued please	Recommended	76561198415138439	33,1	11480
3	156816846	It's not P*kemon with guns, it's ARK but Japanese.	Recommended	76561198010997926	31,9	10008
4	156497135	So it's a couple of days after release. Sale numbers have just hit 4 million. I've played nearly 30 hours since its r	Not Recommended	76561197994874894	35,3	9626
5	156717723	most accurate amazon warehouse simulator	Recommended	76561198331631571	40,6	6892
6	156741513	My depresso got depressed. We cured it with medicine. My alpaca developed an eating disorder so we ate him	Recommended	76561198057424435	166	6849
7	156282364	came out in better state than pokemon scarlet/violet	Recommended	76561198202504052	15,2	6233
8	156387115	Game is so fucking good that I actually decided to use the bug report feature.	Recommended	76561197995199561	36,7	5435
9	157099250	Some people compare this to Pokemon, but lets be honest here, Nintendo wished they had a game as good as	Recommended	76561198028513363	238,2	4796
10	156458903	Get stranded on a beach and start punching sheep.This is Pokemon Black and Blue.	Recommended	76561198297946291	27,5	4573
11	156708897	i will eat a spoon full of mayonnaise for every like this gets	Recommended	76561198113075195	34	4300
12	156683353	I punched the first person i saw and he drew a musket and plugged me. best opening 30s of a game i can reme	Recommended	76561198024617563	50,1	3998
13	156805800	Some will call this a Pokemon rip off, I'd argue this game capitalises on Pokemon's lack of innovation and chan	Recommended	76561198336682259	47,3	3754
14	156506334	Beat them up, enslave them, starve them, force them to breed, and then eat them when they get out of line.10	Recommended	76561198275767842	41,8	3751
15	156540092	Pokemon: Gun of the Wild Survival Evolved Edition! Plus it pisses Twitter off, so that's always a net positive.	Recommended	76561198110874467	28,2	3464
16	156294895	[h1] THERE'S NO LAWS AGAINST THE PALS, BATMAN! [/h1]	Recommended	76561198099630511	79	3331
17	156281171	Protagonists in any Pokemon-like game: As long as we stick together, we can do anything.Protagonist Palworld:	Recommended	76561198338419772	0,1	3306
18	158128427	It's like Pokemon, but you shoot them in the face and force them into unpaid labor.	Recommended	76561197997208586	54,9	2997
19	156577399	Like Pokemon and ark, but better than both somehow	Recommended	76561198064852560	47,6	2956
20	156813762	Please don't let Nintendo take this away from us	Recommended	76561198082243396	82,6	2594
21	156157740	It's Pokemon, meets monster hunter, meets Ark. Then add in a bunch of automatic weapons, forced animal lab	Recommended	76561197989917694	23,4	2563
22	159982783	My friend who I play CS2 with said if I post a review on palworld and it gets 50 likes and 25 Awards that he will	Recommended	76561198084723326	13	2406
23	156387052	Pokemon slavery with machine guns	Recommended	76561198061533386	53,7	2050
24	156281223	In short, it's very addictive and overall fun game. Also I will preface that you should not expect to in the game p	Recommended	76561198043639891	283,8	1997
25	167476339	I am a 54 yo father, probably one of the oldest people playing this game. I am a single father to my Son, who is	Recommended	76561198121352316	86,8	1839
26	156281128	It's that new Pokémon: red, white and blue:.	Recommended	76561198028931632	32,8	1818
27	156458848	This was supposed to be a meme game, an absolute rip on Pokemon and Ark. Instead, we got something that r	Recommended	76561198160732774	88,2	1736
28	157699789	War crimes, slavery, endless breeding farms, human trafficking, animal cruelty, animal genocide.would recomm	Recommended	76561198382238889	386,5	1719
29	157097053	So it's a bit more than a week after release. Sale numbers have just hit 8 million. I've played nearly 60 hours sir	Not Recommended	76561198050914196	1,8	1670

Gambar 2. Hasil Scraping Data

4.2. *Data Exploration*

Terdapat banyak atribut hasil dari *scraping* dalam *dataframe*. adapun terdapat 6 kolom/*feature* yang ada pada dataset diantaranya, *review_id, Review, Recommendation, User_id, play_hours, dan helpful_votes*. Adapun dalam penelitian ini, fitur yang akan digunakan adalah *review_id, Review, Recommendation*. Kolom *review_id* akan digunakan untuk mencari nilai duplikat pada ulasan, *Review* akan digunakan untuk pemodelan sentimen, dan *Recommendation* akan menentukan label sentimen (positif atau negatif).

Data ini kemudian diproses dan dibersihkan untuk memastikan kualitas data yang akan digunakan dalam analisis sentimen. Data ini digunakan sebagai input untuk model *Bert-base* dalam proses klasifikasi sentimen.

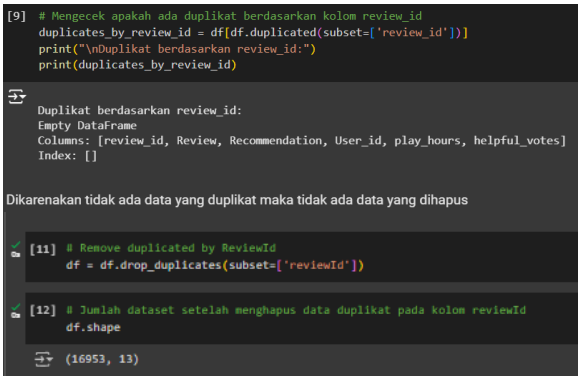
4.3. *Text Preprocessing*

Text Preprocessing adalah proses untuk membersihkan dan memformat data teks agar siap digunakan dalam model pembelajaran mesin. Tahap yang dilakukan yaitu r*emove duplicated, cleaning,*

case folding, *tokenizing*, *remove stopword*, *stemming*, dan *Remove Missing Value*. Berikut adalah penjelasan langkah-langkah yang dilakukan:

4.3.1. *Remove Duplicated*

Pada tahap ini, Dataset dibersihkan dari ulasan duplikat berdasarkan kolom "review_id", yang merupakan pengenal unik untuk setiap ulasan di Steam. Ulasan dengan review_id yang sama dianggap duplikat dan dihapus untuk menghindari bias dalam model.



Gambar 3. Proses Remove Duplicated

Pada Gambar 3 menampilkan bahwa tidak ada data yang duplikat maka tidak ada data yang dihapus.

4.3.2. *Cleaning*

Langkah ini mencakup penghapusan karakter unik seperti tanda baca, angka, emoji, dan karakter non-alfabet lainnya. Hal ini bertujuan untuk memastikan bahwa teks hanya berisi kata-kata yang bermakna.

Tabel 2. Tahap Cleaning Pada Ulasan

Sebelum *Cleaning*	**Setelah *Cleaning***
Get the game before Nintendo KILLS IT. 😠 I love you Nintendo but I hope you lose. This is disappointing.	Get the game before Nintendo KILLS IT I love you Nintendo but I hope you lose This is disappointing

Proses *cleaning* adalah langkah penting dalam preprocessing teks yang membantu dalam meningkatkan kualitas dan kebersihan data teks.

4.3.3. *Case Folding*

Pada tahap ini, semua teks ulasan dikonversi menjadi huruf kecil.

Tabel 3. Tahap *Case Folding*

Sebelum *Case Folding*	**Setelah *Case Folding***
Get the game before **Nintendo KILLS IT I** love you **Nintendo** but I hope you lose **This** is disappointing	**get** the game before **nintendo kills it i** love you **nintendo** but **i** hope you lose **this** is disappointing

Hal ini dilakukan untuk memastikan konsistensi dan menghindari perlakuan berbeda terhadap huruf besar dan huruf kecil.

4.3.4. *Tokenizing*

Tokenisasi adalah proses memecah teks menjadi unit-unit yang lebih kecil seperti kata atau frasa. Tokenisasi memudahkan proses analisis teks deng an memisahkan setiap kata dalam teks.

Tabel 4. Tahap *Tokenizing*

Sebelum *Tokenizing*	**Setelah *Tokenizing***
get the game before nintendo kills it i love you nintendo but i hope you lose this is disappointing	["get", "the", "game", "before", "nintendo", "kills", "it", "i", "love", "you", "nintendo", "but", "i", "hope", "you", "lose", "this", "is", "disappointing"]

Proses *tokenizing* membantu dalam mempersiapkan teks untuk analisis lebih lanjut atau untuk dimasukkan ke dalam model *machine learning*.

4.3.5. *Remove Stopword*

Dalam penelitian ini, proses penghapusan *stopwords* dilakukan menggunakan *library* NLTK.

Tabel 5. Tahap *Remove Stopword*

Sebelum *Remove Stopwords*	**Setelah *Remove Stopwords***
["get", "**the**", "game", "**before**", "nintendo", "kills", "**it**", "**i**", "love", "**you**", "nintendo", "**but**", "**i**", "hope", "**you**", "lose", "**this**", "**is**", "disappointing"]	get game nintendo kills love nintendo hope lose disappointing

Penghapusan stopwords membantu dalam mengurangi kebisingan dalam data, sehingga algoritma analisis sentimen dapat lebih efektif dalam mengidentifikasi makna dari teks yang diberikan.

4.3.6. *Stemming*

Pada penelitian ini, proses stemming dilakukan menggunakan modul PorterStemmer, yang merupakan salah satu modul dari library NLTK.

Tabel 7. Tahap Stemming

Sebelum *Stemming*	**Setelah *Stemming***
get game nintendo **kills** love nintendo hope lose **disappointing**	get game nintendo **kill** love nintendo hope lose **disappoint**

Proses *Stemming* membantu mengurangi jumlah kata yang perlu dianalisis dengan menggabungkan variasi kata ke dalam satu bentuk dasar.

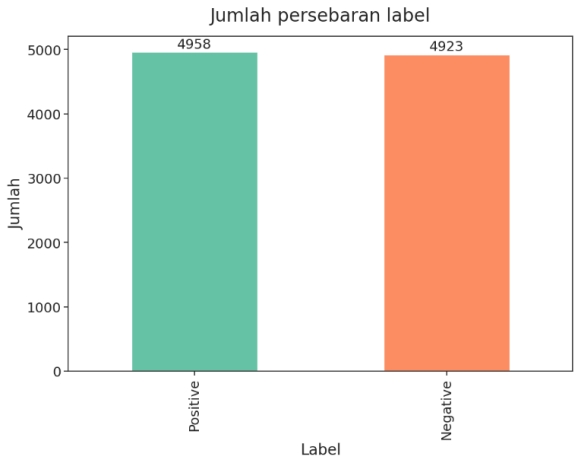
4.3.7. *Remove Missing Value*

Setelah melewati tahapan *stemming* dilakukan kembali proses penghapusan data ulasan yang tidak memiliki nilai atau NaN. Hal ini bisa terjadi karena pada ulasan yang memiliki karakter khusus seperti

nilai angka, tanda baca, dan emoji saja akan otomatis terhapus setelah melewati proses *text preprocessing* sehingga ulasannya menjadi missing value. Adapun setelah melakukan pengecekan terdapat 139 ulasan dengan nilai NaN dan akan dihapus. Dengan demikian jumlah datanya sekarang menjadi 9.881.

4.4. *Labeling*

Pada penelitian ini label yang digunakan ada 2 label yaitu *negative* dan *positive*. Adapun penentuan labelnya berdasarkan hasil ulasan Pengguna. Dengan asumsi hasil "Not Recommended" adalah negatif, dan "Recommended" adalah positif.



Gambar 4. Jumlah Persebaran Label

Gambar 4 menunjukkan jumlah persebaran rating dari ulasan terhadap *game Palworld*. Terdapat 4958 ulasan yang diberi label *positive*. Hal ini menunjukkan bahwa banyak pengguna yang puas atau sangat puas dengan *game* ini. Terdapat 4923 ulasan yang diberi label *negative*. jumlah ulasan negatif menunjukkan adanya sejumlah besar pengguna yang kurang puas atau sangat tidak puas dengan *game* tersebut.

4.5. *Modeling*

Setelah melalui proses pelabelan, langkah selanjutnya adalah melakukan modeling. Pada tahap *modeling* ini, terdapat beberapa langkah yang meliputi *split dataset, dan training model*. Dalam proses pemodelan, akan dilakukan pembagian *data training, data testing, dan data validation* sebesar 80:10:10. Selanjutnya berdasarkan kedua skenario di atas akan dilakukan beberapa proses pemodelan diantaranya yaitu *split dataset, dan training model*.

4.5.1. *Split Dataset*

Proses pembagian data pada penelitian ini terbagi menjadi tiga yaitu data train, data validation, dan data test. Dengan proporsi 80:10:10 dimana 80% untuk *data train*, 10% untuk *data validation*, dan 10% untuk *data test*.

4.5.2. *Training Model*

Penelitian ini menggunakan model bert-base-uncased yang dilatih dengan korpus *Wikipedia* bahasa Inggris dan BookCorpus. Model ini dipilih karena kemampuannya dalam memahami konteks bahasa Inggris, terutama untuk tugas seperti klasifikasi teks, analisis sentimen, dan ekstraksi informasi.

Parameter yang digunakan meliputi *learning rate* 0,00002, *batch size* 16, dan pelatihan selama 10 *epoch*. *Early stopping* diterapkan untuk mencegah *overfitting* dengan memantau performa pada data validasi. *Optimizer* Adam dipilih karena kemampuannya mengadaptasi *learning rate* secara dinamis, sehingga mempercepat konvergensi.

Penelitian ini menganalisis sentimen ulasan pengguna terhadap *game Palworld* di *Steam*, menggunakan pembagian data 80% untuk pelatihan, 10% untuk pengujian, dan 10% untuk validasi. Evaluasi model mencakup metrik seperti akurasi, presisi, *recall*, dan *f1-score.*

	precision	recall	f1-score	support
Negative	0.89	0.83	0.86	502
Positive	0.83	0.90	0.87	487
accuracy			0.86	989
macro avg	0.86	0.86	0.86	989
weighted avg	0.86	0.86	0.86	989

Gambar 5. Hasil Performa Model

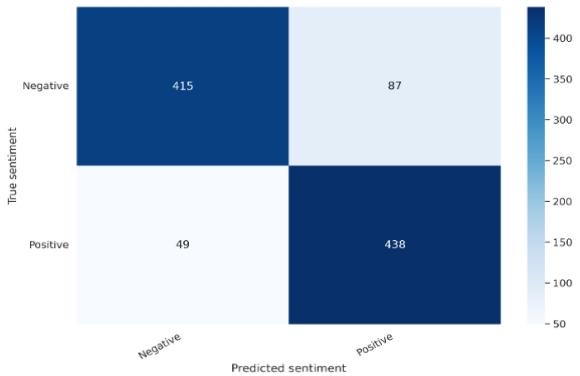
Gambar 5 menunjukan model mencapai akurasi 86%, dengan *macro average dan weighted average* untuk *precision, recall*, serta *F1-score* masing-masing sebesar 0,86. Hasil ini menunjukkan bahwa model memiliki performa tinggi, baik secara seimbang antar kelas maupun dengan mempertimbangkan distribusi kelas.

4.6. *Evaluation*

Pada tahap evaluasi, dilakukan analisis terhadap performa model *BERT-base* menggunakan beberapa metrik seperti *confusion matrix*, dan visualisasi *wordcloud*.

4.6.1. *Confusion Matrix*

Tahap *confusion matrix* digunakan untuk mengukur kinerja model yang telah diterapkan. Hasil *confusion matrix* pada model adalah sebagai berikut.



Gambar 6. Hasil *Confusion Matrix* pada model

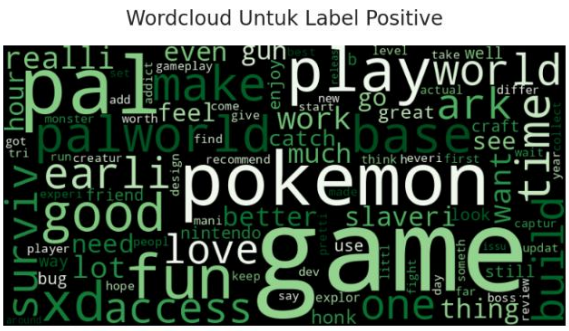
Gambar 6 menunjukkan model berhasil mengklasifikasikan 438 dari 487 ulasan positif dengan benar, namun 49 ulasan positif salah diklasifikasikan sebagai negatif. Dari 502 ulasan negatif, 415 diklasifikasikan dengan benar, sementara 87 ulasan negatif salah diklasifikasikan sebagai positif.

4.6.2. *Visualisasi Wordcloud*

Wordcloud adalah visualisasi kata-kata yang paling sering muncul dalam teks. Dalam analisis sentimen ulasan *game Palworld*, *wordcloud* digunakan untuk menampilkan kata-kata dominan berdasarkan sentimen (positif dan negatif). Proses ini melibatkan pengolahan teks, pemilihan kata kunci, dan visualisasi. Selain mengukur performa model, wordcloud membantu memahami isu-isu utama yang diungkapkan pengguna, menjadikannya alat efektif untuk analisis sentimen dan visualisasi data teks.

Gambar 7. *Wordcloud* Ulasan Negatif *Game Palworld*

Berdasarkan Gambar 7 menunjukkan *wordcloud* untuk ulasan dengan label sentimen negatif, terlihat beberapa kata yang sering muncul dalam ulasan tersebut, mencerminkan berbagai masalah dan keluhan yang dialami pengguna. Kata-kata dominan dalam *wordcloud* tersebut antara lain “game”, “server”, “play”, “time”, “Progress”, “bug”, “want”, “fix”, dan “bored”. Dari hasil analisis ini, dapat disimpulkan bahwa pengguna banyak mengeluhkan kendala *server* yang sering mengganggu pengalaman bermain, masalah terkait waktu dan *progres*, *bug* yang mengganggu performa permainan, serta rasa bosan karena kurangnya variasi atau tantangan dalam *game*. Selain itu, kata-kata seperti "fix" dan "want" mengindikasikan keinginan pemain agar pengembang memperbaiki masalah-masalah ini. Informasi ini dapat membantu pengembang *game* untuk fokus pada perbaikan *server*, peningkatan performa *game*, penambahan variasi dalam *gameplay*, serta perbaikan *bug* demi pengalaman bermain yang lebih baik dan memuaskan bagi pengguna.



Gambar 8. *Wordcloud* Ulasan Positif *Game Palworld*

Berdasarkan Gambar 8 menunjukkan *wordcloud* untuk ulasan dengan label sentimen positif, terlihat beberapa kata yang sering muncul mencerminkan apresiasi dan kesan positif dari pengguna terhadap *game*. Kata-kata yang paling dominan di antaranya adalah “game”, “play”, “fun”, “good”, “love”, “build”, dan “work”. Dari hasil analisis ini, dapat disimpulkan bahwa pengguna memberikan ulasan positif terkait keseruan bermain, yang tercermin dari kata “fun”, “play”, dan “love”, menunjukkan bahwa pengguna menikmati pengalaman bermain yang menyenangkan dan menghargai fitur-fitur dalam *game*. Kata-kata seperti “build”, “base”, dan “work” juga menunjukkan bahwa aspek kreativitas dalam membuat dan membangun sesuatu di dalam *game* sangat disukai. Secara keseluruhan, ulasan positif ini memberikan gambaran bahwa pengguna puas dengan berbagai aspek permainan yang menghadirkan keseruan, tantangan, dan kesempatan berekspresi. Pengembang dapat mempertimbangkan untuk mengembangkan lebih lanjut fitur-fitur ini guna mempertahankan kepuasan pengguna.

5. KESIMPULAN DAN SARAN

Memuat Analisis sentimen *game Palworld* menggunakan model *Bert-base* melibatkan beberapa langkah, dimulai dari pengambilan data ulasan menggunakan *Steam API* (10.020 data), pembersihan teks dengan NLTK, hingga pelabelan sentimen berdasarkan ulasan "Recommended" (positif) dan "Not Recommended" (negatif). Data dibagi 80:10:10 untuk pelatihan, validasi, dan pengujian. Model menggunakan *batch size* 16 dan *learning rate* 0,00002, dengan evaluasi berbasis *confusion matrix* menghasilkan akurasi, presisi, *recall, dan F1-score* sebesar 86%. Analisis kata menunjukkan kata-kata dominan seperti “fun” dan “love” pada sentimen positif, serta “bug” dan “server” pada sentimen negatif. Saran perbaikan untuk pengembang meliputi peningkatan kapasitas *server*, pengujian lebih mendalam untuk stabilitas, perbaikan *bug*, pengurangan waktu *grinding*, dan penambahan variasi *gameplay* seperti misi atau tantangan baru. Untuk penelitian selanjutnya, disarankan meningkatkan jumlah data, mencoba teknik pelabelan lain (misalnya TextBlob atau lexicon), mengeksplor model *pre-*

trained lain seperti *BERT-large* atau *RoBERTa*, serta menggunakan variasi *hyperparameter* untuk mengoptimalkan performa.

DAFTAR PUSTAKA

[1] J. Y. Tan, A. S. K. Chow, and C. W. Tan, "A Comparative Study of Machine Learning Algorithms for Sentiment Analysis of Game Reviews", *IEMJ*, vol. 82, no. 3, Nov. 2022.

[2] J. Howarth, "How Many Gamers Are There? (New 2024 Statistics)," *Exploding Topics*, Jan. 29, 2024. [Online]. Available: https://explodingtopics.com/blog/number-of-gamers. [Accessed: May 5, 2024]

[3] CS Agents, "Steam: Everything You Need to Know About the Video Game Distributor," *CS Agents*, May 21, 2019. [Online]. Available: https://cs-agents.com/blog/steam/#:~:text=For%20game%20developers%2C%20Steam%20can,within%20the%20video%20game%20industry. [Accessed: May 5, 2024].

[4] L. Yang, Y. Li, J. Wang and R. S. Sherratt, "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning," in *IEEE Access*, vol. 8, pp. 23522-23530, 2020, doi: 10.1109/ACCESS.2020.2969854

[5] E. C. Garrido-Merchan, R. Gozalo-Brizuela, and S. . Gonzalez-Carvajal, "Comparing BERT Against Traditional Machine Learning Models in Text Classification", *JCCE*, vol. 2, no. 4, pp. 352–356, Apr. 2023, doi: 10.47852/bonviewJCCE3202838.

[6] H. W. Kang, K. K. Chye, Z. Y. Ong, and C. W. Tan, "The Science of Emotion: Malaysian Airlines Sentiment Analysis using BERT Approach," in *International Conference on Digital Transformation and Applications (ICDXA) 2021*, 2021, pp. 129-136, doi: 10.56453/icdxa.2021.1013.

[7] R. Feldman and J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press, 2007.

[8] G. Salton, *Automatic Text Processing: Transformation, Analysis, and Retrieval of Information by Computer*. Addison Wesely, 1988.

[9] V. Mohan, "Text Mining Research: A Survey," *Int. J. Innov. Res. Comput. Commun. Eng.*, 2016, doi: 10.15680/IJIRCCE.2016.

[10] H. Liu, I. Chatterjee, M. Zhou, X. S. Lu, and A. Abusorrah, "Aspect-Based Sentiment Analysis: A Survey of Deep Learning Methods," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 6, pp. 1358–1375, 2020, doi: 10.1109/TCSS.2020.3033302.

[11] Q. T. Ain et al., "Sentiment Analysis Using Deep Learning Techniques: A Review," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 6, 2017.

[12] A. Ligthart, C. Catal, and B. Tekinerdogan, "Systematic Reviews in Sentiment Analysis: A Tertiary Study," *Artif. Intell. Rev.*, vol. 54, no. 7, pp. 4997–5053, 2021, doi: 10.1007/s10462-021-09973-3.

[13] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *CoRR*, abs/1810.04805, 2018. [Online]. Available: http://arxiv.org/abs/1810.04805.

[14] D. Osinga, *Deep Learning Cookbook: Practical Recipes to Get Started Quickly*. O'Reilly Media, Inc., 2018.

[15] M. S. Wibawa, "Pengaruh Fungsi Aktivasi, Optimisasi, dan Jumlah Epoch Terhadap Performa Jaringan Saraf Tiruan," *J. Sist. dan Inform.*, vol. 11, no. 2, 2017.

[16] I. Düntsch and G. Gediga, "Confusion Matrices and Rough Set Data Analysis," *J. Phys. Conf. Ser.*, vol. 1229, no. 1, 2019, doi: 10.1088/1742-6596/1229/1/012055.

[17] A. J. Gallego, J. Calvo-Zaragoza, J. J. Valero-Mas, and J. R. Rico-Juan, "Clustering-based K-Nearest Neighbor Classification for Large-Scale Data with Neural Codes Representation," *Pattern Recognit.*, vol. 74, pp. 531–543, 2018, doi: 10.1016/j.patcog.2017.09.038.

[18] M. R. Maarif, "Content Analysis on Twitter Users Interaction within First 100 Days of Jakarta's New Government by Using Text Mining," *J. Pekommas*, vol. 3, no. 2, pp. 137, 2018, doi: 10.30818/jpkm.2018.2030203.

[19] A. Fahmi, I. Ramadhan, and P. Studi, "Analisis Sentiment Masyarakat Selama Bulan Ramadhan dalam Menghadapi Pandemi Covid-19," *J. Inform. dan Sist. Inform.*, vol. 1, no. 1, pp. 608–617, 2020.