

Sentiment Analysis of Quizizz Application User Reviews Using Logistic Regression Algorithm

Ari Aditya^{1*}, Shandy Tresnawati²

¹ Program Studi Teknik Informatika, Politeknik TEDC Bandung, Cimahi, 40513, Indonesia

² Program Studi Teknik Komputer, Politeknik TEDC Bandung, Cimahi, 40513, Indonesia

Informasi Artikel

Diterima : 1 Juli 2025
Revisi : 7 Juli 2025
Publikasi : 30 September 2025

Kata Kunci:

Analisis Sentimen
Logistic Regregssion
Quizizz
Google Play Store
Data mining SEMMA

ABSTRAK

Aplikasi pembelajaran digital seperti Quizizz semakin diminati karena menyajikan proses belajar yang interaktif. Seiring bertambahnya pengguna, ulasan di Google Play Store juga meningkat dan mencerminkan persepsi pengguna terhadap kualitas aplikasi. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap aplikasi Quizizz menggunakan algoritma *Logistic Regression*. Data berupa ulasan berbahasa Indonesia dikumpulkan dari Google Play Store selama Maret hingga Desember 2024. Proses analisis mencakup pra-pemrosesan teks menggunakan pustaka Sastrawi, pelabelan sentimen berbasis leksikon, pembobotan *TF-IDF*, serta klasifikasi menggunakan *Logistic Regression*. Evaluasi dilakukan menggunakan akurasi, *precision*, *recall*, dan *F1-score*. Hasil menunjukkan bahwa mayoritas ulasan bersentimen positif, dan model dapat mengklasifikasikan sentimen dengan performa yang baik. Temuan ini memberikan masukan bagi pengembang dalam meningkatkan kualitas dan pengalaman pengguna aplikasi.

ABSTRACT

Digital learning applications like Quizizz are increasingly popular for offering interactive learning experiences. As user numbers grow, so do the reviews on platforms like Google Play Store, reflecting user perceptions of app quality. This study aims to analyze user review sentiment toward the Quizizz application using the Logistic Regression algorithm. The data consists of Indonesian-language reviews collected from March to December 2024. The analysis process includes text Preprocessing using the Sastrawi library, lexicon-based sentiment labeling, TF-IDF weighting, and classification using Logistic Regression. The model is evaluated using accuracy, precision, recall, and F1-score. The results show that most reviews are positive, and the model performs well in sentiment classification. These findings offer insights for developers to improve the app's quality and user experience.

This is an open-access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



*Penulis Koresponden

Email: ariadityaa28@gmail.com

Cara sitasi IEEE

A.Aditya dan S. Tresnawati, "Sentiment Analysis of Quizizz Application User Reviews Using Logistic Regression Algorithm," *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 3, pp. 983-997, September 2025, doi: 10.30811/jaise.v5i3.7292

1. PENDAHULUAN

Salah satu bentuk inovasi yang lahir dari kemajuan teknologi adalah hadirnya aplikasi pembelajaran berbasis permainan, seperti Quizizz, yang bertujuan untuk meningkatkan motivasi belajar siswa melalui kuis-kuis interaktif. Aplikasi Quizizz memadukan unsur hiburan dengan proses pembelajaran melalui fitur-fitur seperti *leaderboard*, waktu pengerjaan, dan tampilan visual menarik, sehingga mampu membangkitkan semangat kompetitif dan kolaboratif siswa dalam suasana belajar yang menyenangkan. Kemudahan akses melalui perangkat digital seperti *smartphone* dan *laptop* juga memungkinkan penggunaannya di berbagai situasi, sementara guru dapat dengan mudah menyusun soal, memantau hasil, serta memberikan umpan balik secara langsung guna meningkatkan efektivitas pembelajaran yang inklusif dan adaptif[1].

Meski penggunaannya semakin meluas, pada tahun 2024 muncul berbagai kendala teknis yang menghambat efektivitas pemanfaatannya. Permasalahan seperti gangguan sistem, keterbatasan akses, serta ketidaksesuaian perangkat dengan aplikasi pembelajaran menjadi tantangan tersendiri bagi para pengguna. Selain itu, perubahan kebijakan pada layanan pihak ketiga juga berdampak pada berkurangnya fitur-fitur yang sebelumnya mempermudah penggunaannya. Hambatan-hambatan ini berpotensi menurunkan motivasi dan partisipasi siswa dalam pembelajaran dengan menggunakan aplikasi ini.

Pesatnya pertumbuhan aplikasi digital juga menyebabkan melonjaknya ulasan pengguna, yang dapat menyulitkan identifikasi ulasan bermasalah. Kritik atau keluhan sering tertutupi oleh banyaknya ulasan lain, baik netral maupun positif, sehingga pengguna kesulitan mendapatkan gambaran akurat tentang kualitas aplikasi. Selain itu, praktik pemasaran seperti jasa ulasan positif semakin mempersulit perbedaan antara ulasan autentik dan manipulatif, menghambat pengguna baru dalam menilai aplikasi secara objektif. Hal ini menyulitkan calon pengguna mendapatkan gambaran objektif, terutama dengan banyaknya ulasan yang tersebar dan tidak terstruktur. *Data science* menjadi solusi efektif melalui teknik seperti *Text mining* dan analisis sentimen untuk mengidentifikasi kelemahan aplikasi dari ulasan pengguna. Pendekatan ini membantu memberikan transparansi bagi pengguna tanpa perlu mencoba aplikasinya terlebih dahulu. *Text mining* merupakan metode yang digunakan untuk mengekstraksi informasi dari data berbentuk teks yang bersifat tidak terstruktur, seperti ulasan digital atau dokumen. Proses ini dilakukan dengan menganalisis serta mengelompokkan kata-kata guna mengidentifikasi pola atau hubungan tertentu dalam data. Teknik ini banyak diterapkan dalam berbagai bidang, seperti analisis sentimen, pemetaan kecenderungan pengguna, deteksi *spam*, serta berbagai aplikasi lain dalam ranah pemrosesan bahasa alami[2].

Analisis terhadap sentimen ulasan pengguna aplikasi pendidikan seperti Quizizz menjadi semakin relevan, karena meningkatnya konten secara online yang dihasilkan pengguna, analisis sentimen menjadi alat yang penting untuk memahami opini publik. Dalam konteks pembelajaran, wawasan dari analisis tersebut dapat membantu pengembang dan pendidik dalam meningkatkan kualitas serta efektivitas aplikasi[3]. Menurut Yusuf et al, algoritma *Logistic Regression* menghasilkan performa tinggi dalam mengklasifikasikan ulasan pengguna terhadap aplikasi pembelajaran daring[4]. Untuk mendukung analisis sentimen secara efektif, algoritma *Logistic Regression* menjadi salah satu pendekatan yang sering digunakan. Algoritma ini sangat relevan karena kemampuannya dalam memprediksi sentimen positif, negatif dan netral, berdasarkan data yang telah diolah. Keunggulan algoritma ini terletak pada kesederhanaan dan akurasi yang cukup tinggi, terutama dalam analisis sentimen yang memiliki data yang terstruktur dan bersih.

Melalui analisis sentimen berbasis *Logistic Regression*, pengembang dapat mengidentifikasi dan mengatasi berbagai masalah, seperti kendala teknis, desain antarmuka, konten, pengalaman pengguna, permintaan fitur baru, kualitas pembelajaran, hingga keberadaan iklan yang mungkin mengurangi kenyamanan pengguna. Dengan menggabungkan hasil analisis ini, pengembang dapat mengambil langkah yang tepat untuk memperbaiki aplikasi dan meningkatkan pengalaman pengguna.

Secara keseluruhan, penggunaan algoritma *Logistic Regression* pada analisis sentimen memiliki peran penting dalam memahami ulasan pengguna terhadap aplikasi pendidikan seperti Quizizz. Sebagaimana dijelaskan oleh Ainunnisa dan Sulastri, *Logistic Regression* memiliki keunggulan dalam klasifikasi sentimen karena sifatnya yang mudah digunakan serta interpretatif, dan terbukti mampu menghasilkan tingkat akurasi yang tinggi dalam penelitian mereka terhadap ulasan pengguna aplikasi yang menunjukkan efektivitas algoritma ini dalam menganalisis sentimen pengguna[5]. Literatur yang tersedia umumnya masih berfokus pada aspek penerapan Quizizz dalam proses pembelajaran, seperti efektivitas penggunaan, persepsi siswa, atau pengaruhnya terhadap motivasi belajar. Namun, kajian yang secara khusus menganalisis sentimen pengguna terhadap aplikasi Quizizz berdasarkan ulasan berbahasa Indonesia masih sangat terbatas. Hasil penelusuran pustaka menunjukkan bahwa meskipun analisis sentimen terhadap aplikasi mobile dalam Bahasa Indonesia telah dilakukan pada beberapa domain lain, studi serupa pada konteks aplikasi edukasi berbasis permainan seperti Quizizz belum banyak ditemukan. Dengan memanfaatkan teknik analisis sentimen yang canggih, pengembang dapat mendapatkan wawasan strategis untuk meningkatkan kualitas aplikasi, sekaligus menciptakan pengalaman belajar yang lebih memuaskan bagi pengguna.

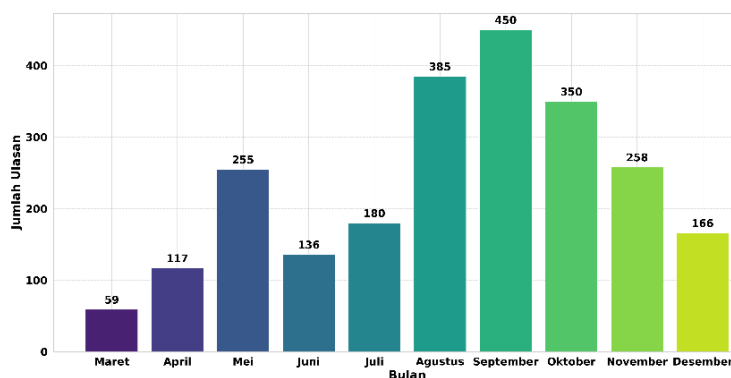
2. METODE

2.1 Pengumpulan data

Pengumpulan data merupakan tahap penting dalam penelitian karena menentukan validitas hasil yang diperoleh. Metode yang digunakan akan sangat memengaruhi kualitas dan keakuratan data. Dalam penelitian ini, peneliti menerapkan metode pengumpulan data sebagai berikut:

a) Data primer

Data primer merupakan data yang diperoleh langsung dari sumber data yang relevan dengan penelitian yang dilakukan. Menurut Muslimin, data primer adalah data yang diperoleh secara langsung oleh peneliti dari sumber asli atau responden tanpa melalui perantara. Jenis data ini dikumpulkan menggunakan berbagai metode seperti observasi, wawancara, survei, atau eksperimen, dan dianggap sebagai data mentah karena belum diolah oleh pihak lain[6]. Data primer yang digunakan, diperoleh langsung dari ulasan pengguna aplikasi Quizizz di Google Play Store. Pengumpulan data dilakukan dengan metode *web scraping*, yaitu teknik otomatis untuk mengekstraksi dan mengumpulkan informasi dari situs *web* secara sistematis tanpa interaksi manual [7]. Teknik ini memungkinkan peneliti memperoleh data dalam jumlah besar, seperti teks ulasan, tanggal unggah, dan *rating* pengguna, yang selanjutnya digunakan dalam proses analisis sentimen. Data ulasan yang berhasil diperoleh bisa dilihat pada gambar dibawah ini.



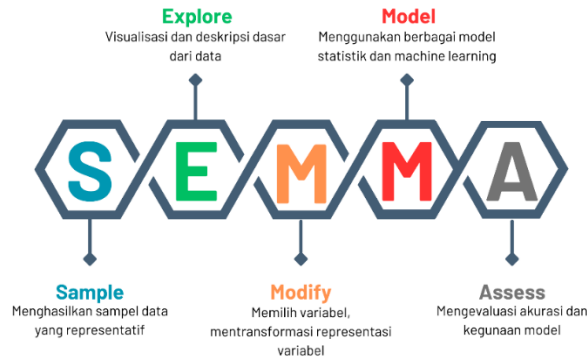
Gambar 1. Data Yang Diperoleh Pada Tahun 2024

b) Data Sekunder

Data sekunder adalah data yang diperoleh secara tidak langsung, biasanya melalui dokumentasi, literatur, buku, jurnal, maupun informasi lain yang relevan dengan rumusan masalah dan kebutuhan penelitian. Menurut Rahman, data sekunder merupakan data yang didapat melalui pihak lain atau bukan hasil pengumpulan langsung oleh peneliti. Umumnya, data ini diperoleh dari publikasi resmi, seperti buku, laporan, atau rilis institusi tertentu. Sebagai contoh, data sosial ekonomi yang diterbitkan melalui publikasi SUSENAS dapat dikategorikan sebagai data sekunder[8].

2.2 Metode Penelitian

Penelitian ini mengacu pada metode SEMMA (*Sample, Explore, Modify, Model, Assess*) sebagai pendekatan dalam proses *data mining*. Fokus utama dalam metode SEMMA adalah pada modifikasi data, pemodelan, dan evaluasi hasil penambangan data. Menurut Alizah et al., SEMMA merupakan singkatan dari *Sample, Explore, Modify, Model, dan Assess*. Sementara itu, dalam penelitian yang dilakukan oleh Shafique, disebutkan bahwa SEMMA merupakan metode *data mining* yang dikembangkan oleh *SAS Institute*, dan digunakan untuk memberikan solusi terhadap berbagai tujuan serta permasalahan dalam bidang bisnis[9]. Tahapan metode SEMMA bisa dilihat pada gambar dibawah ini.



Gambar 2. Tahapan Metode SEMMA

2.1 Sample

Pada tahap ini dilakukan pengumpulan dataset yang berasal dari ulasan pengguna aplikasi Quizizz pada tahun 2024. Pengambilan data dilakukan dengan teknik *web scraping* menggunakan *API* untuk memperoleh ulasan dari Google Play Store. Hasil *scraping* menghasilkan sebanyak 2.354 ulasan pengguna, yang kemudian disimpan dalam format CSV atau Excel untuk keperluan analisis lebih lanjut. *Sample* merupakan tahap dalam proses *data mining* yang berfungsi untuk mengumpulkan sampel data yang cukup besar dan relevan, sehingga dapat membentuk informasi yang penting dan signifikan. Namun demikian, tahap *sample* ini bersifat opsional, sehingga tidak selalu wajib dilakukan dalam setiap proses *data mining*[10]. Contoh *sample* ulasan bisa dilihat pada gambar dibawah ini.

	userName	rati	tangga	bulan	
2	zenal muttaqin	5	31-12-2024	12	mantap
3	Yuli Priyanto	5	31-12-2024	12	Asik banget, cuman k
4	Kesantrian PPISB	5	26-12-2024	12	Bagus
5	ria arummanah	5	25-12-2024	12	Kereeen pokoknya
6	Rizqi Asaduddin	5	24-12-2024	12	karena sangatlah me
7	M daffa Adly r	5	23-12-2024	12	Sangat mendidik
8	Luthfiyah Akram12	3	23-12-2024	12	Bagus. Tapi aku ngga
9	Kenichi Satya Manggala	5	23-12-2024	12	Bagus
10	Endis Maoludin	5	23-12-2024	12	Sangat membantu da
11	saka sukena	5	22-12-2024	12	Bagus buat asesmer
12	Aulia Mutia Allya	5	22-12-2024	12	bagusss banget
13	Yohana Ika Saprina	5	22-12-2024	12	Bagus buat belajar

Gambar 3. Sample Ulasan

2.2 Explore

Pada tahap ini peneliti mendeskripsikan data yang didapat dari hasil *scraping* data ulasan pengguna aplikasi Quizizz. Kemudian dilakukan tahap visualisasi data yaitu data yang didapat dibuat grafik berdasarkan waktu untuk menampilkan informasi dari data secara visual. *Explore* merupakan sebuah proses *data mining* yang dapat digunakan untuk mencari kumpulan data dan menjadi informasi yang terkait dengan tren dan *anomaly* yang tidak terduga yang dapat digunakan untuk mendapatkan pengertian dan ide. Tahapan ini berfungsi untuk mengeksplorasi data guna mengidentifikasi adanya kesalahan penulisan atau nilai kosong (null). Melalui proses eksplorasi ini, peneliti dapat menentukan langkah-langkah pengolahan lanjutan terhadap data yang tersedia[11].

2.3 Modify

Pada tahap *Modify* di penelitian ini adalah tahapan untuk persiapan data atau *Pre-processing*. Tahapan yang akan dilakukan oleh peneliti terdiri dari *Case folding*, *cleaning*, *Tokenize*, dan *normalize*. Menurut Suwitono & Kaunang, *modify* adalah sebuah proses untuk memodifikasi data dan mengelompokkan variabel variabel. Data-data yang tidak relevan dan tidak dibutuhkan untuk pembelajaran mesin akan dibuang beserta suatu kelas yang memiliki data berlebih dan mengganggu keseimbangan akan dikurangi[12].

2.3.1 Case folding

Langkah pertama pada tahap *Modify* pada penelitian ini adalah *Case folding*. Proses *Case folding* adalah langkah dalam *text Preprocessing* yang bertujuan untuk mengubah semua huruf dalam teks ulasan menjadi huruf kecil (*lowercase*). Langkah ini penting untuk memastikan konsistensi dalam analisis teks, terutama dalam menghindari perbedaan antara kata yang sebetulnya identik[13]. Tabel berikut ini adalah hasil dari *Case folding*.

Tabel 1. Hasil Dari *Case folding*

No	Sebelum <i>Case folding</i>	Setelah <i>Case folding</i>
----	-----------------------------	-----------------------------

1	<i>Lag</i> paraaaaah padahal sering banget drill soal tapi <i>lag</i> bangeet, gak ada perubahan, dari soal ke soal nge <i>lag</i> nya lama ngabisin waktu	<i>lag</i> paraaaaah padahal sering banget drill soal tapi <i>lag</i> bangeet, gak ada perubahan, dari soal ke soal nge <i>lag</i> nya lama ngabisin waktu
2	Sangat mudah digunakan dan dipahami 🍌🍌	sangat mudah digunakan dan dipahami 🍌🍌

2.3.2 Cleaning

Selanjutnya adalah membersihkan data dari komponen yang tidak relevan dan tidak memiliki makna seperti ASCII, angka, link, hastag, url, tanda baca dan whitespace. *Cleaning* Dilakukan untuk menghilangkan URL atau tautan link website. Data *cleaning* berfungsi untuk menghapus karakter –karakter yang tidak diperlukan atau noise pada data dengan menghilangkan tanda baca, emoji, angka dan link. Proses ini dilakukan karena keberadaan elemen-elemen tersebut dapat membuat data menjadi tidak efektif [14]. Tabel berikut ini adalah hasil *Cleaning* pada dataset dalam penelitian.

Tabel 2. Hasil *Cleaning*

Sebelum <i>Cleaning</i>	Setelah <i>Cleaning</i>
Sangat mudah digunakan dan dipahami 🍌🍌	sangat mudah digunakan dan dipahami

2.3.3 Tokenize

Tahap selanjutnya adalah *Tokenize* atau tokenisasi dimana suatu kalimat dari ulasan atau dataset akan dipecah menjadi potongan kata atau token. Hal ini dilakukan untuk mengetahui kemunculan dari kata tersebut. Tokenisasi berfungsi untuk memecah komentar menjadi satuan kata. Proses tokenisasi dilakukan dengan melihat setiap spasi yang ada dalam komentar maka berdasarkan spasi tersebut kata-kata dapat dipecah [15]. Tabel berikut ini adalah hasil *Tokenize* pada dataset dalam penelitian.

Tabel 3. Hasil *Tokenize*

Sebelum <i>Tokenize</i>	Setelah <i>Tokenize</i>
sangat mudah digunakan dan dipahami	['sangat', 'mudah', 'digunakan', 'dan', 'dipahami']

2.3.4 Normalize

Tahap berikutnya adalah normalisasi atau *normalize*. Didalam penelitian ini, tahap normalisasi dilakukan untuk menstandarisasi kata yang memiliki makna yang sama dengan melakukan perubahan penulisan kata yang disingkat dan atau tidak baku. Normalisasi ini memperbaiki huruf yang ganda pada suatu kalimat serta memperbaiki kalimat yang salah ketik. Tujuannya agar kata bisa dikenali dan dideteksi makna dari kata-kata tersebut [16]. Tabel berikut ini adalah hasil dari *Normalize*

Tabel 4. Hasil *Normalize*

Sebelum <i>Normalize</i>	Setelah <i>Normalize</i>
['bagus', 'banget', 'pokok', 'kayak', 'anak', 'yg', 'belajar', 'disini', 'tu', 'seru', 'gitu', 'kayak', 'game', 'tapi', 'belajar', 'keren deh']	['bagus', 'banget', 'anak', 'yang', 'belajar', 'di sini', 'itu', 'seru', 'seperti', 'game', 'tapi', 'belajar', 'keren']

2.4 Model

Pada tahap ini, *dataset* bersih dibagi menjadi kelas positif, negatif dan netral dengan metode *Lexicon based*. Menurut Fathullah, Sari & Adikara, *lexicon based* merupakan metode seleksi fitur yang bertujuan untuk pemilihan kata yang mengandung sentimen. Proses pemilihan kata dilakukan dengan mencocokkan setiap kata dengan kamus *Lexicon* Bahasa Indonesia [17]. Kelebihan metode ini adalah kemampuan untuk secara otomatis memberi label pada kalimat, yang dapat menghemat waktu, terutama saat mengolah dataset berukuran besar [18]. Metode *Lexicon based* pada penelitian ini digunakan untuk melakukan klasifikasi kelas pada *dataset*. Suatu komentar atau dokumen pada *dataset* diklasifikasikan ke dalam tiga kelas, apakah komentar tersebut termasuk pada kelas positif, negatif atau netral. Kata-kata yang ada dalam komentar terdapat di dalam kamus *lexicon*, maka kata tersebut akan diberikan nilai *score*. Jumlah nilai *score* pada suatu komentar yang menentukan label positif atau negatif. Table berikut ini adalah hasil dari *dataset* yang sudah melalui pelabelan metode *Lexicon based*.

Tabel 5. Hasil dari Metode *Lexicon based*

No	Ulasan	Label
1	<i>Lag</i> paraaaaah padahal sering banget <i>drill</i> soal tapi <i>lag</i> bangeet, gak ada perubahan, dari soal ke soal nge <i>lag</i> nya lama ngabisin waktu	Negatif
2	Sangat mudah digunakan dan dipahami 🍌🍌	Positif

2.5 Asses

Asses merupakan sebuah proses *data mining* yang dapat digunakan untuk menilai data dengan mengevaluasi kegunaan dan keandalan penemuan dari data proses *data mining* dan memperkirakan seberapa baik kinerja tersebut. Menurut Putri et al, pada tahap ini dilakukan evaluasi model menggunakan *confusion matrix*. *Confusion matrix* merepresentasikan data yang diprediksi benar (sesuai label/data aktual) dan salah (tidak sesuai label/data aktual). Pada tahap ini dilakukan evaluasi dari model[19]. Pada algoritma *Logistic Regression*, digunakan 10-fold *Cross-validation* untuk mengetahui akurasi, *precision*, *recall*, dan *F1-score* dari model yang digunakan. Evaluasi ini bertujuan untuk mengukur performa model secara keseluruhan dengan mempertimbangkan keseimbangan antara bias dan *varians*.

Penelitian ini menggunakan algoritma *Logistic Regression*. Dalam penelitian Alfyando et al., pada tahun 2024 *Logistic Regression* merupakan metode statistik yang digunakan untuk memprediksi probabilitas dari suatu kejadian dengan dua kemungkinan hasil, seperti “ya” atau “tidak”. Dalam konteks *machine learning*, metode ini termasuk dalam klasifikasi biner, di mana variabel target memiliki dua kategori. *Logistic Regression* bekerja dengan memanfaatkan fungsi logistik (sigmoid) untuk mengubah nilai output menjadi probabilitas dalam rentang 0 hingga 1[23].

2.6 Proses dan Hasil

Penelitian ini menggunakan algoritma *Logistic Regression*. Peneliti menggunakan algoritma tersebut untuk mengetahui performa dari algoritma tersebut terhadap *dataset* yang sudah didapatkan. Memperhatikan nilai akurasi, *precision*, *recall*, dan *F1-score* untuk mendapatkan pemahaman yang lebih mengenai algoritma tersebut saat digunakan dalam analisis sentimen.

2.7 Analisis

Pada tahap ini, dilakukan analisis terhadap hasil evaluasi model yang telah diterapkan menggunakan algoritma *Logistic Regression*. Analisis ini bertujuan untuk memahami bagaimana performa model dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Quizizz serta mengidentifikasi faktor-faktor yang mempengaruhi akurasi model.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

Tujuan dari penelitian ini adalah untuk menganalisis sentimen pengguna dalam ulasan pengguna terhadap aplikasi Quizizz di Google Play Store menggunakan algoritma *Logistic Regression*. Data ulasan yang digunakan adalah teks yang dikumpulkan melalui web scraping antara tanggal 21 Maret hingga 31 Desember 2024. Pendekatan analisis mengikuti tahapan SEMMA (*Sample, Explore, Modify, Model, Assess*) yang mencakup pengumpulan data, eksplorasi data, *Preprocessing*, pelabelan, visualisasi, ekstraksi fitur, pelatihan model, serta evaluasi performa.

3.1.1 Deskripsi Dataset

Data yang digunakan dalam penelitian ini adalah data ulasan pengguna aplikasi Quizizz di Google Play Store tahun 2024. Setelah dilakukan pengumpulan data melalui *web scraping*, berhasil terkumpul 2.354 ulasan keseluruhan data yang didapata dianalisis. Data yang digunakan meliputi beberapa atribut penting seperti *username*, tanggal ulasan, *rating*, dan isi teks ulasan itu sendiri. Proses seleksi dan validasi data meliputi penghapusan nilai kosong (null) dan *filter* bahasa menggunakan *library* “*langdetect*” untuk memastikan hanya ulasan berbahasa Indonesia yang dianalisis. Menurut Wyawhare, *langdetect* adalah model deteksi bahasa berbasis aturan (*rule-based*) yang menggunakan aturan yang telah ditentukan sebelumnya untuk mengidentifikasi bahasa dari suatu teks[20]. Setelah *difilter* data yang tersedia menjadi 1.403. Hal ini berfungsi untuk memastikan keakuratan hasil *Preprocessing* dan analisis sentimen berbasis leksikon yang digunakan dalam penelitian ini berikut gambar proses *filtering* bahasa.

```
[ ] # 4. Deteksi bahasa setiap ulasan
from langdetect import detect, DetectorFactory
DetectorFactory.seed = 0 # agar hasil deteksi stabil

def detect_language(text):
    try:
        return detect(str(text))
    except:
        return "unknown"

df["bahasa"] = df["ulasan"].apply(detect_language)

[ ] # 5. Filter hanya yang berbahasa Indonesia
df_indonesia = df[df["bahasa"] == "id"]
print(f"Jumlah ulasan berbahasa Indonesia: {len(df_indonesia)}")
df_indonesia.head()
```

Gambar 4. Proses *Filtering* Bahasa

3.1.2 Eksplorasi Data

Eksplorasi data dilakukan untuk memahami pola dan frekuensi kata yang muncul dalam ulasan pengguna terhadap aplikasi Quizizz. Analisis ini berfokus pada isi teks ulasan tanpa mempertimbangkan rating, sesuai pendekatan *lexicon-based* yang digunakan dalam penelitian ini.

Ulasan-ulasan dianalisis secara tekstual untuk mengidentifikasi tema atau kata kunci yang dominan. Visualisasi menggunakan *Wordcloud* dibuat untuk memperlihatkan kata-kata yang paling sering muncul dalam ulasan dan dikelompokkan berdasarkan tiga kategori sentimen:

- Sentimen Positif: Kata-kata yang sering muncul antara lain “*mudah*”, “*bagus*”, “*menarik*”, “*membantu*”, dan “*interaktif*”. Kata-kata ini mencerminkan kepuasan pengguna terhadap fitur dan kemudahan penggunaan aplikasi.
- Sentimen Netral: Kata-kata seperti “*kuis*”, “*soal*”, “*fitur*”, dan “*pengguna*” lebih bersifat informatif atau deskriptif, tanpa menunjukkan emosi yang kuat.
- Sentimen Negatif: Kata-kata yang dominan dalam ulasan negatif meliputi “*lag*”, “*error*”, “*susah*”, “*lemot*”, dan “*keluar sendiri*”. Kata-kata ini mencerminkan ketidakpuasan terhadap performa teknis aplikasi.

3.1.3 Modify

Tahap pra-pemrosesan (*Preprocessing*) dilakukan untuk membersihkan data teks sebelum digunakan dalam proses klasifikasi sentimen. Tahap ini bertujuan untuk menghasilkan data teks yang bersih, konsisten, dan siap diolah lebih lanjut oleh algoritma klasifikasi.

Proses *Preprocessing* dimulai dengan *Case folding*, yaitu mengubah seluruh huruf dalam teks ulasan menjadi huruf kecil guna menyamakan representasi kata. Setelah itu, dilakukan tahap *Cleaning* untuk menghapus elemen-elemen yang tidak relevan seperti URL, angka, simbol, emotikon, serta karakter non-alfabet lainnya yang dapat mengganggu analisis. Selanjutnya, teks dipecah menjadi bagian-bagian kecil berupa kata-kata individual melalui proses tokenisasi, dan kemudian dilakukan normalisasi terhadap kata-kata tidak baku, misalnya “*gk*” diubah menjadi “*tidak*” dan “*bgt*” menjadi “*banget*”, agar seluruh kata konsisten dalam bentuk bakunya.

Tahap berikutnya adalah penghapusan *Stopword*, yaitu proses menghilangkan kata-kata umum yang tidak memiliki kontribusi makna terhadap analisis sentimen, seperti “*dan*”, “*itu*”, atau “*atau*”. Setelah *Stopword* dihapus, dilakukan *Stemming* menggunakan pustaka Sastrawi, yakni proses mengembalikan setiap kata ke bentuk dasarnya. Misalnya, kata “*belajarnya*” akan distem menjadi “*ajar*”.

Keenam proses ini secara keseluruhan menghasilkan representasi teks yang lebih sederhana namun tetap memuat informasi penting untuk keperluan klasifikasi sentimen. Berikut merupakan Gambar yang menunjukkan hasil dari proses *Preprocessing* terhadap salah satu contoh ulasan pengguna.

	ulasan	clean_ulasan
0	Asik banget, cuman kalo bisa di perbanyak lagi...	asik banget cuman kalo banyak yaa jangan suka ...
1	Kereeen pokoknya	kereeen pokok
2	karena sangatlah menyenangkan serta menambah w...	sangat senang tambah wawas
3	Sangat mendidik	sangat didik
4	Bagus. Tapi aku nggak tau cara liat siapa saja...	bagus aku tau cara liat siapa yang mainin kuis...
...
1398	Sya menggunakan aplikasi ini untuk belajar	sya aplikasi untuk ajar
1399	dapat meningkatkan pengalaman pembelajaran, me...	tingkat alam ajar didik motivasi sangat untuk ...
1400	seruu bangettt bisa main sambil belajar apalag...	seruu bangettt main ajar ada karakter karakter...
1401	Menambah motivasi belajar	tambah motivasi ajar
1402	Bagus banget buat yang mau ulangan dan yang ma...	bagus banget buat mau ulang yang mau ajar

Gambar 5. Hasil Dari *Preprocessing* Teks

3.1.4 Model

a) Pelabelan Menggunakan *Lexicon based*

Pelabelan sentimen dilakukan menggunakan pendekatan *lexicon-based*, yaitu dengan mencocokkan setiap kata dalam teks yang telah dipra-proses dengan daftar kosakata yang telah diberi skor sentimen. Dalam penelitian ini, digunakan custom *lexicon* berbasis Bahasa Indonesia yang disusun dengan mengadaptasi dari beberapa sumber, seperti *SentiStrength* Bahasa Indonesia, daftar kata dari penelitian terdahulu, serta disesuaikan secara manual berdasarkan konteks ulasan edukasi. Kosakata positif mencakup kata-kata seperti mudah, menarik, menyenangkan, membantu, bagus, sementara kosakata negatif meliputi *lag*, *error*, lemot, keluar, lambat. Setiap kata diberi bobot skor (+1 untuk positif, -1 untuk negatif), dan akumulasi skor dari setiap ulasan menentukan kelas sentimennya (positif, negatif, atau netral jika skor = 0). Pendekatan ini dipilih karena efektif dalam menangani ulasan pendek dan cocok untuk teks berbahasa Indonesia tanpa perlu pelabelan manual dalam skala besar. Berikut gambar hasil pelabelan.

	clean_ulasan	label_sentimen
0	asik banget cuman kalo banyak yaa jangan suka nge lag di tengahtengah	netral
1	kereeen pokok	netral
2	sangat senang tambah wawas	netral
3	sangat didik	netral
4	bagus aku tau cara liat siapa yang mainin kuis aku sma berapa skornyamohon maaf	positif

Gambar 6. Hasil Pelabelan

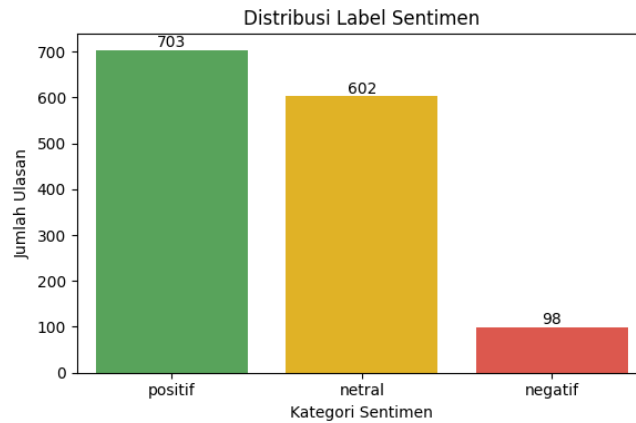
b) Visualisasi *Wordcloud* untuk Masing-Masing Sentimen

Selama fase eksplorasi data, visualisasi data teks dalam bentuk *wordcloud* dilakukan untuk setiap label sentimen (positif, negatif, dan netral). *Wordcloud* secara visual menggambarkan frekuensi kata dalam ulasan pengguna. Semakin sering suatu kata digunakan, semakin besar kata tersebut muncul dalam visualisasi. Visualisasi ini bertujuan untuk mengidentifikasi kata-kata dominan yang sering digunakan oleh pengguna di setiap jenis sentimen dan memberikan gambaran umum tentang fokus dan konteks ulasan. Menurut Wibowo et al, *Wordcloud* sendiri merupakan salah satu teknik visualisasi yang menggambarkan representasi visual dari frekuensi kata. *Wordcloud* sendiri cukup populer digunakan pada *text mining* karena dapat memberikan tampilan visual pada kata-kata yang akan dianalisis namun tetap informatif[21]. Berikut merupakan gambar hasil dari masing masing visualisasi *wordcloud*.



Gambar 7. Visualisasi *Worldcloud* (a) *Worldcloud* Sentimen Positif (b) *Worldcloud* Sentimen Negatif (c) *Worldcloud* Sentimen Netral

Selain visualisasi wordcloud, dilakukan juga analisis kuantitatif berupa grafik batang (*bar chart*) untuk menampilkan distribusi jumlah ulasan dalam masing-masing kategori sentimen: positif, netral, dan negatif. Grafik ini bertujuan untuk memberikan gambaran proporsi dominasi sentimen dalam korpus data. Hasilnya menunjukkan bahwa kategori sentimen positif merupakan yang paling dominan, diikuti oleh netral dan negatif. Distribusi ini penting untuk mengidentifikasi seberapa besar kecenderungan pengguna terhadap aplikasi serta menjadi dasar dalam mengevaluasi ketidakseimbangan kelas pada saat pelatihan model klasifikasi. Berikut gambar *bar chart* distribusi label sentimen.



Gambar 8. Bar Chart Distribusi Label Sentimen

c) Ekstraksi Fitur dengan *TF-IDF*

Proses ekstraksi fitur dilakukan menggunakan metode *TF-IDF* (*Term Frequency – Inverse Document Frequency*). Teknik ini bertujuan untuk mengubah teks ulasan menjadi representasi numerik yang dapat diproses oleh model klasifikasi. *TF-IDF* memberikan bobot yang lebih tinggi pada kata-kata yang sering muncul dalam sebuah dokumen tetapi jarang muncul di keseluruhan korpus, sehingga lebih mencerminkan konteks spesifik dari masing-masing ulasan. Dalam penelitian Andayani & Ryansyah tahun 2022 dikatakan juga bahwa *Term Frequency – Inverse Document Frequency (TF-IDF)*. Algoritma ini bekerja dengan bobot tiap kata di dalam dokumen yang dibandingkan. Proses pembobotan dengan algoritma ini membutuhkan sekumpulan dokumen agar bobot yang dihasilkan tidak hanya bergantung pada jumlah kemunculan kata. Setelah proses pembobotan, bobot tersebut disusun ke dalam vektor yang akan diukur tingkat kesamaannya dengan menggunakan cosine similarity [22]. Berikut gambar hasil ekstraksi ekstraksi fitur dengan *TF-IDF*.

	abiss	ad	ada	adapun	adek	adil	aduhai	ae	agak	agam	...	xixi	\
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
	ya	yaa	yaaa	yah	yang	yant	yes	yg	you				
0	0.0	0.357711	0.0	0.0	0.000000	0.0	0.0	0.0	0.0				
1	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	0.0	0.0				
2	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	0.0	0.0				
3	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	0.0	0.0				
4	0.0	0.000000	0.0	0.0	0.277445	0.0	0.0	0.0	0.0				

Gambar 9. Hasil Ekstraksi Fitur Dengan *TF-IDF*

d) Penerapan Algoritma

Pada penerapan analisis sentimen, algoritma *Logistic Regression* dimanfaatkan untuk memperkirakan kemungkinan sebuah teks tergolong dalam kelas positif atau negatif. Model ini dinilai efektif dalam menangani klasifikasi biner sederhana serta memberikan interpretasi yang transparan terhadap pengaruh setiap fitur terhadap prediksi akhir [24].

Logistic Regression merupakan metode analisis yang digunakan untuk memodelkan hubungan antara variabel respon dengan variabel prediktor, baik yang bersifat kategorik maupun kontinu. Model ini dapat digunakan untuk respon biner, ordinal, maupun multinomial. Pada regresi logistik biner, variabel respon hanya memiliki dua kemungkinan, yaitu sukses ($y=1$) dengan probabilitas $p(x)$ dan gagal ($y=0$) dengan probabilitas $1-p(x)$, di mana distribusi respon mengikuti distribusi Bernoulli pada setiap observasi. [25].

Data teks diubah ke dalam bentuk numerik dengan teknik *TF-IDF* (*Term Frequency-Inverse Document Frequency*) agar dapat dibaca oleh model. Representasi ini membantu model memahami kata-kata mana yang paling penting dalam membedakan sentimen antar dokumen.

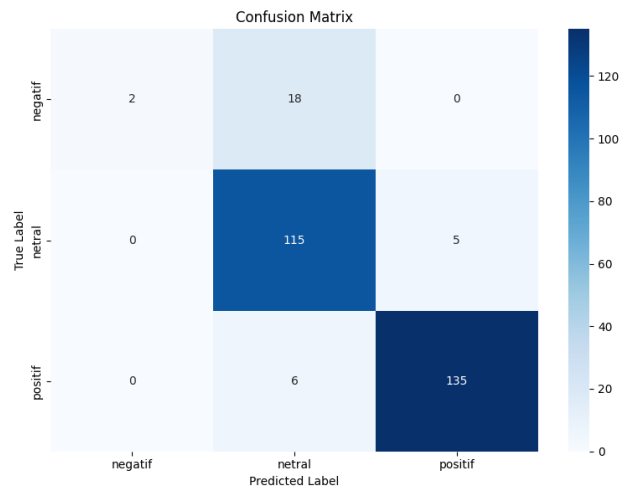
Dataset dibagi menjadi dua bagian, yaitu data latih sebesar 80% dan data uji sebesar 20%. Model dilatih pada data latih untuk membentuk pola klasifikasi, kemudian diuji menggunakan data uji untuk mengevaluasi kemampuannya dalam memprediksi label sentimen berdasarkan ulasan teks. Berikut tabel contoh hasil prediksi model pada beberapa data uji.

Tabel 6. Hasil Prediksi Model

Teks Ulasan	Label Asli	Prediksi
“sangat bagus ajar”	Positif	Positif
“seru banget harus ajar wajib siswa”	Positif	Positif
“asik banget cuman kalo banyak yaa jangan suka nge <i>lag</i> di tengahtengah”	Netral	Netral
“susah <i>login</i> ”	Negatif	Netral
“ <i>lag</i> dikit ga ngaruh”	Negatif	Netral

3.1.5 Evaluasi Model

Setelah model *Logistic Regression* dilatih, dilakukan evaluasi untuk mengukur seberapa baik model dalam memprediksi sentimen ulasan. Evaluasi dilakukan dengan menghitung akurasi, yaitu persentase prediksi yang benar dibandingkan seluruh data pengujian. Selain itu, *classification report* digunakan untuk melihat performa model dalam memprediksi masing-masing label sentimen, mencakup nilai *precision*, *recall*, dan *F1-score*. Terakhir, *confusion matrix* ditampilkan sebagai visualisasi yang menunjukkan jumlah prediksi benar dan salah dari tiap label, sehingga kita bisa mengetahui jenis kesalahan yang sering dilakukan oleh model, seperti apakah ulasan netral sering diklasifikasikan sebagai positif atau negatif. Dari hasil evaluasi ini, dapat disimpulkan apakah model sudah cukup baik atau masih perlu diperbaiki. Berikut gambar *confusion matrix* hasil pengujian model.

Gambar 10. *Confusion matrix* Hasil Pengujian Model

Berdasarkan *confusion matrix* di atas, diketahui bahwa dari total 20 ulasan dengan label negatif, hanya 2 ulasan yang berhasil diklasifikasikan dengan benar oleh model, sementara 18 ulasan negatif lainnya justru diprediksi sebagai netral, dan tidak ada satupun yang diklasifikasikan sebagai positif. Untuk ulasan dengan label netral, mayoritas berhasil diprediksi dengan tepat, yaitu sebanyak 115 data, meskipun masih terdapat 5 ulasan netral yang salah diklasifikasikan sebagai positif, dan tidak ada yang diklasifikasikan sebagai negatif. Sementara itu, pada ulasan dengan label positif, model menunjukkan kinerja yang sangat baik dengan jumlah prediksi benar sebanyak 135 data, dan hanya terjadi kesalahan klasifikasi terhadap 6 data yang diprediksi sebagai netral, serta tidak ada yang diklasifikasikan sebagai negatif. Secara keseluruhan, hasil ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang baik terutama untuk label positif dan netral, namun masih perlu perbaikan dalam mendeteksi ulasan negatif yang sering kali keliru diklasifikasikan sebagai netral. Untuk memperjelas performa model secara kuantitatif, berikut adalah *classification report* yang menampilkan metrik evaluasi secara lengkap.

Tabel 7. *Classification Report* Hasil Evaluasi Model

Sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Negatif	1.00	0.10	0.18	20
Netral	0.83	0.96	0.89	120
Positif	0.96	0.96	0.96	141
<i>Accuracy</i>			0.90	281
<i>Macro Avg</i>	0.82	0.88	0.68	281
<i>Weighted Avg</i>	0.93	0.91	0.87	281

Tabel 8. Akurasi Keseluruhan Model

Metrik	Nilai
Akurasi (desimal)	0.90
Akurasi (persentase)	89.68%

Dari tabel di atas dapat disimpulkan bahwa model bekerja sangat baik pada kelas positif dan netral, dengan nilai *precision* dan *recall* yang tinggi. Untuk kelas positif, *precision* dan *recall* masing-masing mencapai 0.96, yang berarti sebagian besar prediksi terhadap kelas ini benar dan hampir seluruh ulasan positif berhasil dikenali oleh model. Begitu pula pada kelas netral, model menunjukkan performa yang solid dengan *precision* sebesar 0.83 dan *recall* 0.96, mengindikasikan bahwa model sangat efektif dalam mengklasifikasikan ulasan netral. Namun demikian, performa model pada kelas negatif yang menjadi fokus utama dalam penelitian ini masih tergolong lemah. Hal ini terlihat dari nilai *recall* yang sangat rendah, yaitu hanya 0.10, meskipun *precision*-nya sempurna (1.00). Artinya, model hanya mampu mengenali sebagian kecil dari total ulasan negatif yang ada, sementara sisanya cenderung salah diklasifikasikan ke dalam kelas lain, khususnya netral. Nilai *F1-score* untuk kelas negatif pun hanya sebesar 0.18, mencerminkan ketidakseimbangan performa yang signifikan dan rendahnya sensitivitas model terhadap ulasan dengan sentimen negatif. Secara keseluruhan, akurasi model mencapai 0.90 (atau 89.68%), yang berarti sekitar 9 dari 10 prediksi yang dilakukan model adalah benar. Meskipun demikian, nilai *macro average F1-score* sebesar 0.68 menegaskan bahwa performa antar kelas masih belum merata. Sementara itu, nilai *weighted average F1-score* sebesar 0.87 cenderung dipengaruhi oleh jumlah data yang lebih besar pada kelas netral dan positif.

3.2 Pembahasan

Pembahasan pada bagian ini dilakukan untuk menginterpretasikan hasil penelitian yang telah disajikan pada subbab sebelumnya. Penelitian ini bertujuan untuk mengetahui pola sentimen pengguna terhadap aplikasi Quizizz, mengukur akurasi model *Logistic Regression* dalam klasifikasi sentimen, serta mengidentifikasi faktor-faktor utama yang memengaruhi persepsi pengguna melalui ulasan yang diambil dari Google Play Store.

3.2.1 Pola Sentimen Ulasan Pengguna

Hasil klasifikasi sentimen terhadap 1.403 ulasan pengguna aplikasi Quizizz menunjukkan bahwa mayoritas ulasan termasuk dalam kategori sentimen positif, disusul oleh sentimen netral, dan yang paling sedikit adalah sentimen negatif. Temuan ini menggambarkan bahwa secara umum, pengguna memiliki persepsi yang baik terhadap aplikasi, meskipun masih terdapat sebagian pengguna yang menyampaikan kritik atau komentar netral.

a) Sentimen Positif

Ulasan positif menunjukkan bahwa pengguna merasa puas terhadap pengalaman menggunakan Quizizz. Mereka menilai aplikasi ini sangat membantu dalam proses pembelajaran, terutama karena kemudahan akses, tampilan antarmuka yang user-friendly, dan elemen gamifikasi yang membuat belajar terasa lebih menyenangkan. Kata-kata dominan yang muncul dalam *Wordcloud* kategori positif meliputi: “mudah”, “bagus”, “membantu”, “interaktif”, dan “seru”. Kata-kata ini mencerminkan antusiasme pengguna terhadap fitur-fitur yang disediakan oleh Quizizz, seperti kuis yang menarik, sistem skor, dan fleksibilitas penggunaan tanpa harus membayar.

b) Sentimen Negatif

Meskipun jumlahnya lebih sedikit, ulasan negatif memberikan masukan yang krusial terkait aspek teknis aplikasi. Mayoritas komentar negatif menyoroti permasalahan seperti *lag* saat menjalankan kuis, aplikasi tiba-tiba keluar, waktu *loading* yang lama, dan kompatibilitas perangkat yang kurang optimal. Hal ini terlihat dari dominasi kata-kata seperti “lag”, “lemot”, “error”, dan “keluar” dalam *Wordcloud* kategori negatif. Namun demikian, model mengalami kesulitan dalam mengenali ulasan negatif. Dari 20 data uji berlabel negatif, hanya 2 yang berhasil diklasifikasikan dengan benar, sedangkan sisanya sebagian besar salah diklasifikasikan sebagai netral. Contohnya seperti ulasan “susah login” atau “lag dikit ga ngaruh” yang secara kontekstual bernada negatif, tetapi model menganggapnya netral karena struktur kalimat pendek atau penggunaan negasi yang melemahkan bobot kata negatif. Kinerja buruk ini tercermin dari nilai *recall* kelas negatif yang sangat rendah (0,10), meskipun *precision*-nya sempurna (1,00). Hal ini mengindikasikan bahwa model hanya mengenali sebagian kecil ulasan negatif secara akurat. Beberapa faktor penyebabnya adalah ketidakseimbangan distribusi data antar kelas serta keterbatasan *lexicon* dalam menangkap konteks negatif yang implisit. Perbaikan ke depan dapat dilakukan dengan menyeimbangkan data, menerapkan class *weighting*, serta menyempurnakan daftar kosakata agar lebih sensitif terhadap ekspresi negatif dalam bahasa Indonesia.

c) Sentimen Netral

Ulasan netral umumnya bersifat informatif atau deskriptif tanpa menyatakan opini yang kuat, baik positif maupun negatif. Pengguna dalam kategori ini cenderung memberikan penilaian yang datar atau hanya menjelaskan fitur tanpa disertai emosi. Kata-kata seperti “fitur”, “kuis”, “soal”, “pengguna”, dan “nilai” sering muncul dalam

Wordcloud sentimen netral. Beberapa komentar dalam kategori ini juga berisi saran, masukan, atau pertanyaan teknis dari pengguna, seperti cara melihat skor atau mengakses soal tertentu.

Secara keseluruhan, pola sentimen dalam ulasan pengguna menunjukkan dominasi opini positif yang memperkuat nilai guna aplikasi Quizizz dalam konteks pembelajaran. Namun demikian, keberadaan sentimen negatif dan netral tetap menjadi aspek penting yang tidak bisa diabaikan. Sentimen negatif menunjukkan adanya masalah performa teknis yang harus segera ditangani, sementara sentimen netral memberikan gambaran objektif terkait penggunaan fitur yang bisa dimanfaatkan sebagai dasar untuk peningkatan layanan. Oleh karena itu, pihak pengembang perlu menindaklanjuti temuan ini secara menyeluruh agar dapat mempertahankan kepuasan pengguna yang sudah tinggi, sekaligus memperbaiki aspek-aspek yang masih menjadi kendala di lapangan.

3.2.2 Kinerja Model Logistic Regression Dalam Analisis Sentimen

Model yang digunakan dalam penelitian ini adalah *Logistic Regression* yang dilatih pada data ulasan pengguna yang telah melalui proses pra-pemrosesan teks dan dikonversi ke dalam bentuk vektor menggunakan metode *TF-IDF*. Evaluasi model dilakukan dengan menggunakan *confusion matrix* serta metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*.

Hasil evaluasi menunjukkan bahwa model memiliki *accuracy* sebesar 89,68%, yang tergolong tinggi untuk tugas klasifikasi sentimen berbasis teks berbahasa Indonesia. Nilai *F1-score* untuk sentimen positif dan netral masing-masing mencapai 0,96 dan 0,89, menunjukkan bahwa model memiliki performa yang sangat baik dalam mengenali dua kategori tersebut.

Namun demikian, performa model terhadap kelas negatif masih jauh dari optimal. Meskipun *precision*-nya mencapai 1,00, nilai *recall* hanya 0,10, dan *F1-score* hanya 0,18, yang menandakan bahwa model sangat rendah dalam mengenali sebagian besar ulasan negatif yang sebenarnya. Hal ini menjadi perhatian utama, mengingat fokus penelitian ini adalah untuk mengevaluasi efektivitas model dalam mengidentifikasi sentimen negatif dalam ulasan pengguna.

Rendahnya performa pada kelas negatif kemungkinan besar disebabkan oleh ketidakseimbangan jumlah data, di mana ulasan negatif jauh lebih sedikit dibandingkan dengan kelas netral dan positif. Selain itu, ekspresi sentimen negatif dalam bahasa Indonesia sering kali lebih bervariasi, *implisit*, dan kontekstual, sehingga menyulitkan model linear seperti *Logistic Regression* dalam menangkap nuansa semantik yang kompleks.

Meski demikian, secara keseluruhan model ini masih menunjukkan performa yang stabil, mudah diinterpretasikan, dan cukup akurat untuk klasifikasi tiga kelas. Namun, untuk mencapai performa yang lebih seimbang, khususnya dalam mendeteksi sentimen negatif, perlu dilakukan perbaikan lebih lanjut, seperti penggunaan *class weighting*, penyeimbangan data (*resampling*), atau eksplorasi terhadap model lain yang lebih kompleks dan sensitif terhadap data minoritas.

3.2.3 Faktor Penentu Sentimen Berdasarkan Ulasan

Berdasarkan hasil *Preprocessing* dan visualisasi kata dominan pada tiap kelas sentimen, dapat diidentifikasi beberapa faktor utama yang memengaruhi persepsi pengguna terhadap aplikasi Quizizz.

Untuk sentimen positif, kata-kata seperti “mudah”, “bagus”, “membantu”, dan “belajar” menunjukkan bahwa pengguna menghargai kesederhanaan antarmuka, efektivitas konten kuis, serta manfaat langsung dalam mendukung kegiatan akademik. Fitur seperti kemudahan akses, gamifikasi, dan evaluasi otomatis memberikan pengalaman belajar yang lebih menyenangkan.

Sedangkan pada sentimen negatif, kata-kata seperti “lag”, “error”, “keluar”, dan “lemot” mengindikasikan bahwa masalah teknis menjadi sumber ketidakpuasan. Keluhan terhadap performa aplikasi seperti koneksi yang tidak stabil atau aplikasi yang menutup sendiri sering muncul dan menjadi penghambat bagi pengguna dalam menjalankan kuis secara optimal.

Faktor-faktor ini penting bagi pengembang karena secara langsung berkaitan dengan *user experience* (UX) dan kualitas layanan. Pengetahuan ini dapat digunakan untuk menyusun prioritas pengembangan fitur, meningkatkan performa teknis, serta mempertahankan elemen yang dinilai sangat positif oleh pengguna.

3.2.4 Implikasi Hasil Penelitian

Hasil analisis ini memberikan implikasi praktis yang dapat dimanfaatkan oleh pengembang aplikasi, pendidik, dan pihak lain yang menggunakan Quizizz sebagai media pembelajaran. Pengembang dapat menggunakan informasi ini untuk meningkatkan performa teknis, memperbaiki bug, dan mempertahankan fitur-fitur yang mendapat tanggapan positif dari pengguna. Pendidik juga dapat mempertimbangkan jenis konten dan pendekatan pembelajaran berbasis kuis yang dinilai paling efektif oleh siswa.

Selain itu, hasil ini menunjukkan bahwa penerapan teknik *text mining* dan algoritma *Logistic Regression* cukup efektif dalam menganalisis ulasan pengguna terhadap aplikasi pendidikan. Kombinasi metode ini bisa diadaptasi pada studi lain dengan konteks serupa untuk memperoleh wawasan pengguna secara lebih luas dan sistematis.

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai analisis sentimen ulasan pengguna aplikasi Quizizz menggunakan algoritma *Logistic Regression*, dapat disimpulkan bahwa :

- a) Mayoritas ulasan pengguna di Google Play Store menunjukkan sentimen positif. Hal ini mencerminkan bahwa secara umum aplikasi Quizizz diterima dengan baik oleh pengguna, terutama karena kemudahan penggunaan, fitur kuis yang interaktif, serta pengalaman belajar yang menyenangkan. Sementara itu, sentimen negatif muncul dalam jumlah yang lebih sedikit dan umumnya disebabkan oleh kendala teknis seperti *lag*, *error*, dan aplikasi yang keluar secara tiba-tiba.
- b) Dari segi performa, algoritma *Logistic Regression* terbukti mampu melakukan klasifikasi sentimen dengan cukup baik, yang ditunjukkan oleh tingkat *accuracy* sebesar 89,68%. Model ini menunjukkan performa optimal pada kategori sentimen positif dan netral, dengan nilai *F1-score* masing-masing sebesar 0,96 dan 0,89. Namun demikian, performa terhadap sentimen negatif masih perlu ditingkatkan, mengingat nilai *recall*-nya yang sangat rendah, yaitu sebesar 0,10, meskipun *precision*-nya mencapai 1,00. Hal ini menunjukkan bahwa model masih memiliki keterbatasan dalam mengenali sebagian besar ulasan negatif secara akurat, sehingga perlu dilakukan perbaikan agar model lebih sensitif terhadap sentimen negatif.
- c) Faktor-faktor yang memengaruhi arah sentimen pengguna umumnya berkaitan dengan performa teknis aplikasi, kualitas antarmuka, dan kelengkapan fitur yang tersedia. Kata-kata seperti “mudah”, “bagus”, “lag”, dan “error” menjadi indikator utama dalam membedakan sentimen pengguna. Secara keseluruhan, kombinasi metode *Preprocessing* teks, perhitungan bobot menggunakan *TF-IDF*, dan penggunaan algoritma *Logistic Regression* terbukti efektif dalam melakukan analisis sentimen berbasis teks dalam bahasa Indonesia, dengan hasil klasifikasi yang akurat dan interpretatif.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Program Studi Teknik Informatika, yang telah memberikan dukungan, arahan, dan fasilitas selama proses penyusunan dan pelaksanaan penelitian ini. Penulis juga menyampaikan apresiasi kepada Politeknik TEDC Bandung atas tersedianya lingkungan akademik yang mendukung serta sarana dan prasarana yang memadai, sehingga penelitian ini dapat diselesaikan dengan baik.

REFERENSI

- [1] S. Krisnasari, D. Suhermah, and I. Latifah, “Pemanfaatan Aplikasi Quizizz dalam Pembelajaran Literasi dan Numerasi di PAUD,” *Jiip - J. Ilm. Ilmu Pendidik.*, vol. 5, no. 6, pp. 1730–1734, 2022, doi: 10.54371/jiip.v5i6.635.
- [2] Y. Findawati and A. Rosid, Muhammad, *BUKU AJAR TEXT MINING*, 1st ed. Sidoarjo: UMSIDA Press, 2020. ISBN: 978-623-6833-19-3.
- [3] F. F. Mailoa and L. Lazuardi, “Analisis sentimen data twitter menggunakan metode text mining tentang masalah obesitas di indonesia,” *J. Inf. Syst. Public Heal.*, vol. 6, no. 1, p. 44, 2021, doi: 10.22146/jisph.44455.
- [4] R. Yusuf, K. Bahumatra, N. Komaria, E. A. Aqma, and L. Cahyani, “Analisis Sentimen Terhadap Aplikasi Google Meet Berdasarkan Komentar Pengguna Menggunakan Metode Logistic Regression,” *J. Ilm. Educic Pendidik. dan Inform.*, vol. 11, no. 1, pp. 53–64, 2024, doi: 10.21107/educic.v11i1.28113.
- [5] I. R. Ainunnisa and S. Sulastri, “Analisis Sentimen Aplikasi Tiktok dengan Metode Support Vector Machine (SVM), Logistic Regression dan Naïve Bayes,” *J. Teknol. Sist. Inf. dan Apl.*, vol. 6, no. 3, pp. 423–430, 2023, doi: 10.32493/jtsi.v6i3.31076.
- [6] D. Nugraha, *Metodologi penelitian: pendekatan kuantitatif, kualitatif, dan campuran*, 1st ed., no. June. Agam, Indonesia: CV LAUK PUYU PRESS, 2024.
- [7] M. R. Fikri, R. T. Handayanto, and D. Irwan, “Web Scraping Situs Berita Menggunakan Bahasa Pemrograman Python,” *J. Students’ Res. Comput. Sci.*, vol. 3, no. 1, pp. 123–136, 2022, doi: 10.31599/jsres.v3i1.1514.
- [8] Rahmawati, B. A. Habsy, and M. Nursalim, “Jenis-Jenis Metode Pengumpulan Data (Qualitative Research),” *J. Pendidik. Tambusai*, vol. 9, no. 1, p. 9935, 2025, doi: 10.47827/jer.v5i4.281.
- [9] M. D. Alizah, A. Nugroho, U. Radiyah, and W. Gata, “Sentimen Analisis Terkait Lockdown pada Sosial Media Twitter,” *Indones. J. Softw. Eng.*, vol. 6, no. 2, pp. 223–229, 2020, doi: 10.31294/ijse.v6i2.8991.
- [10] Omari Firas, “A combination of SEMMA & CRISP-DM models for effectively handling big data using formal concept analysis based knowledge discovery: A data mining approach,” *World J. Adv. Eng. Technol. Sci.*, vol. 8, no. 1, pp. 009–014, 2023, doi: 10.30574/wjaets.2023.8.1.0147.
- [11] A. Salsabila Juwita, A. Rizky Kurniawan, A. Aryaputra Ashari, D. Tyan Putro, V. Nurcahyawati, and H. Artikel, “Implementasi Data Mining untuk Memprediksi Kesehatan Mental Mahasiswa menggunakan Algoritma Naïve Bayes,” *KOMPUTEK J. Tek. Univ. Muhammadiyah Ponorogo*, vol. 8, no. No 1, p. Hal 61-70, 2024, [Online]. Available: <http://studentjournal.umpo.ac.id/index.php/komputek>
- [12] Y. A. Suwitono and F. J. Kaunang, “Implementasi Algoritma Convolutional Neural Network (CNN) Untuk Klasifikasi Daun Dengan Metode Data Mining SEMMA Menggunakan Keras,” *J. Komitika (Komputasi dan Inform.)*, vol. 6, no. 2, pp. 109–121, 2022, doi: 10.31603/komitika.v6i2.8054.
- [13] M. Z. Siregar *et al.*, “ANALISIS SENTIMEN TERHADAP ULASAN APLIKASI MEDIA SOSIAL DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 2, pp. 3373–3381, 2025, doi: 10.36040/jati.v9i2.12841.
- [14] Pande sindu, Agus Aan Jiwa Permana, and I Nyoman Saputra Wahyu Wijaya, “Identifikasi Dan Normalisasi Teks Slang Dengan Fasttext Pada Twitter Dalam Bahasa Indonesia,” *J. Pendidik. Teknol. dan Kejur.*, vol. 21, no. 1, pp. 33–44, 2024, doi: 10.23887/jptkuniksha.v21i1.66381.
- [15] N. Nofiyani and W. Wulandari, “Implementasi Electronic Data Processing Untuk meningkatkan Efektifitas dan Efisiensi Pada Text Mining,” *J. Media Inform. Budidarma*, vol. 6, no. 3, p. 1621, 2022, doi: 10.30865/mib.v6i3.4332.

- [16] D. Yulianto and A. Nugraheni, "DECODE : Jurnal Pendidikan Teknologi Informasi," *Decod. J. Pendidik. Teknol. Inf.*, vol. 1, no. 1, pp. 33–42, 2021, doi: 10.51454/decode.v1i1.5.
- [17] N. S. Fathullah, Y. A. Sari, and P. P. Adikara, "Analisis Sentimen Terhadap Rating dan Ulasan Film dengan menggunakan Metode Klasifikasi Naïve Bayes dengan Fitur Lexicon-Based," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 2, pp. 590–593, 2020, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/6987>
- [18] A. N. Ihsan and S. Tresnawati, "Analisis Sentimen Pengguna X Terhadap Layanan Provider Iconnet Menggunakan Naïve Bayes Dan Support Vector Machine," *JITET (Jurnal Inform. dan Tek. Elektro Ter.)*, vol. 12, no. 3S1, pp. 4105–4113, 2024, doi: 10.23960/jitet.v12i3S1.5264.
- [19] D. S. Putri, N. Sulistiyowati, and A. Voutama, "Analisis Sentimen dan Pemodelan Ulasan Aplikasi AdaKami Menggunakan Algoritma SVM dan KNN," *J. Sensi*, vol. 9, no. 2, pp. 209–225, 2023, doi: 10.33050/sensi.v9i2.2914.
- [20] A. Wyawhare, "Comparative Analysis of Multilingual Text Classification & Identification through Deep Learning and Embedding Visualization," *arXiv Prepr.*, vol. 2312.03789, pp. 1–9, 2023, [Online]. Available: <https://arxiv.org/abs/2312.03789>
- [21] J. A. Wibowo, V. C. Mawardi, and T. Sutrisno, "Visualisasi Word Cloud Hasil Analisis Sentimen Berbasis Fitur Layanan Aplikasi Gojek Dengan Support Vector Machine," *J. Serina Sains, Tek. dan Kedokt.*, vol. 2, no. 1, pp. 61–70, 2024, doi: 10.24912/jsstk.v2i1.32058.
- [22] S. Andayani and A. Ryansyah, "Implementasi Algoritma TF-IDF Pada Pengukuran Kesamaan Dokumen," *JuSiTik J. Sist. dan Teknol. Inf. Komun.*, vol. 1, no. 1, p. 53, 2017, doi: 10.32524/jusitik.v1i1.218.
- [23] M. Alfyanto, F. T. Anggraeny, and A. N. Sihananto, "Perbandingan Algoritma Random Forest dan Logistic Regression Untuk Analisis Sentimen Ulasan Aplikasi Tumbuh Kembang Anak Di Play Store," *J. Sist. Inf. dan Ilmu Komput.*, vol. 2, no. 1, pp. 77–86, 2024, [Online]. Available: <https://doi.org/10.59581/jusiik-widyakarya.v2i1.2262>
- [24] A. Kartika Sari, Akhmad Irsyad, Dinda Nur Aini, Islamiyah, and Stephanie Elfriede Ginting, "Analisis Sentimen Twitter Menggunakan Machine Learning untuk Identifikasi Konten Negatif," *Adopsi Teknol. dan Sist. Inf.*, vol. 3, no. 1, pp. 64–73, 2024, doi: 10.30872/atasi.v3i1.1373.
- [25] A. R. S. Darwanto, Taza Luzia Viarindita, and Yekti Widyaningsih, "Analisis Regresi Logistik Binomial dan Algoritma Random Forest pada Proses Pengklasifikasian Penyakit Ginjal Kronis," *J. Stat. dan Apl.*, vol. 5, no. 1, pp. 1–14, 2021, doi: 10.21009/jsa.05101.