

MODEL PREDIKSI SENTIMEN ULASAN MOBILE LEGENDS DI GOOGLE PLAY STORE DAN YOUTUBE BERBASIS PELABELAN OTOMATIS ROBERTA DAN KLASIFIKASI RANDOM FOREST

Muhammad Rafid Pratama¹, Maya Rini Handayani², Wenty Dwi Yuniarti³, Khothibul Umam⁴

^{1,2,3,4} Teknologi Informasi, Fakultas Sains dan Teknologi, UIN Walisongo Semarang, Indonesia
Jl. Prof. Dr. Hamka, Semarang 50185, Jawa Tengah, Indonesia

¹ 2208096024@student.walisongo.ac.id

² maya@walisongo.ac.id

³ wenty@walisongo.ac.id

⁴ Khothibul_umam@walisongo.ac.id*

Abstract

The rapid development of the mobile gaming industry has led to an increase in users and reviews for various popular titles, including Mobile Legends: Bang Bang. This research aims to analyze user perceptions of the Mobile Legends application through reviews obtained from Google Play Store and YouTube. The methods employed include data crawling, automatic labeling using the RoBERTa model for sentiment classification (positive, negative, and neutral), and modeling with the Random Forest algorithm. The dataset comprises 1,400 data points from Google Play Store and hundreds of data points from YouTube, all of which have undergone preprocessing. Model evaluation uses precision, recall, and f1-score metrics. The test results show that the model can classify reviews quite well, achieving an accuracy of 80% on Google Play Store data and 82% on YouTube data. The model demonstrates high performance in detecting negative and positive reviews, although the accuracy for the neutral class remains low. Overall, the Random Forest-based model is sufficiently reliable in processing user review data, providing insights into public perception of Mobile Legends across different platforms.

Keywords: Mobile Legends, Random Forest, Sentiment Analysis, Google Play Store, YouTube

I. PENDAHULUAN

Industri game mobile telah berkembang pesat dalam beberapa tahun terakhir, dengan jutaan pengguna aktif di seluruh dunia. Salah satu game yang mendominasi pasar adalah Mobile Legends: Bang Bang (MLBB), sebuah permainan bergenre Multiplayer Online Battle Arena (MOBA) yang dikembangkan oleh Moonton. Sejak dirilis pada tahun 2016, MLBB telah mendapatkan popularitas yang luar biasa, terutama di kawasan Asia Tenggara, termasuk Indonesia [1].

Popularitasnya tidak hanya terlihat dari angka unduhan yang tinggi di Google Play Store[2], tetapi juga dari tingginya interaksi publik yang terjadi di berbagai platform media sosial seperti YouTube. Para pemain aktif sering kali membagikan pengalaman, kritik, maupun apresiasi mereka terhadap permainan ini melalui berbagai kanal digital tersebut. Oleh karena itu, ulasan pengguna yang tersedia di platform-platform tersebut menyimpan informasi berharga mengenai persepsi, kepuasan, dan harapan pengguna terhadap game ini.

Sebagai game dengan basis pengguna yang luas, MLBB menerima banyak ulasan dari pemainnya di berbagai platform media sosial. Dengan meningkatnya jumlah ulasan yang tersedia, analisis sentimen menjadi metode yang efektif untuk memahami bagaimana persepsi

gamer terhadap MLBB, baik dalam aspek positif, negatif, maupun netral.[3] Dalam volume data yang sangat besar yang dihasilkan setiap hari oleh para pengguna, analisis manual terhadap ulasan menjadi sangat tidak efisien. Untuk itu, pendekatan berbasis pemrosesan bahasa alami (Natural Language Processing/NLP) semakin relevan dalam penelitian kontemporer. Salah satu strategi yang digunakan dalam penelitian ini adalah pelabelan data secara otomatis menggunakan model deep learning berbasis Transformer, yaitu model w11wo/indonesian-roberta-base-sentiment-classifier yang telah dilatih khusus untuk bahasa Indonesia[4]. Model ini memungkinkan proses anotasi sentimen (positif, netral, negatif) secara cepat dan konsisten, menghindari bias subjektif yang kerap terjadi pada pelabelan manual. Pelabelan otomatis ini merupakan tahap penting sebelum dilakukan analisis lanjutan menggunakan algoritma klasifikasi yang lebih sederhana namun kuat.

Setelah data diberi label secara otomatis, proses klasifikasi dilakukan dengan menggunakan algoritma Random Forest. Algoritma ini dipilih karena kemampuannya dalam menangani dataset besar dengan fitur kompleks serta kestabilannya dalam menghasilkan akurasi yang baik. Random Forest menggabungkan banyak pohon keputusan (decision tree) untuk memberikan prediksi yang lebih akurat dan tahan terhadap overfitting.

Dalam penelitian ini, data ulasan yang berasal dari dua platform digital Google Play Store dan YouTube dianalisis untuk melihat bagaimana persepsi pengguna terhadap permainan ini dapat berbeda tergantung pada konteks platform. Hasil klasifikasi kemudian dievaluasi melalui metrik seperti precision, recall, f1-score, dan akurasi untuk mengetahui efektivitas pendekatan yang digunakan[5].

Penelitian serupa sebelumnya telah dilakukan oleh Andika Prasetyo et al. dari Universitas Singaperbangsa Karawang[6]. Dalam penelitian tersebut peneliti mengklasifikasikan sentimen terhadap aplikasi GBWhatsApp menggunakan dua metode yaitu Naïve Bayes Classifier (NBC) dan Random Forest Classifier (RFC). Data yang digunakan berasal dari komentar pengguna di YouTube dan Twitter dengan total 1.686 data, yang setelah dibersihkan menjadi 1.545 data. Data tersebut dibagi menjadi 90% data latih dan 10% data uji. Hasil penelitian menunjukkan bahwa metode NBC mencapai akurasi sebesar 71,43%, precision sebesar 75,81%, dan recall sebesar 61,84%. Sedangkan metode RFC memperoleh akurasi sebesar 64,94%, precision sebesar 80,56%, dan recall sebesar 38,16%. Hal ini menunjukkan bahwa NBC memiliki performa lebih baik dibandingkan RFC dalam mengklasifikasikan sentimen terhadap aplikasi GBWhatsApp.

Penelitian yang dilakukan oleh peneliti Fanka Angelina Larasati, Dian Eka Ratnawati, dan Buce Trias Hanggara di Universitas Brawijaya[7], metodologi Random Forest digunakan untuk menilai sentimen pengguna mengenai aplikasi dompet digital Dana. Pendekatan inovatif ini melibatkan analisis cermat dari kumpulan data substansial yang terdiri dari 1.354 ulasan, yang secara sistematis diklasifikasikan ke dalam tiga kategori sentimen yang berbeda: positif, negatif, dan netral. Penerapan algoritma Random Forest terbukti sangat efektif, yang berpuncak pada tingkat akurasi yang mengesankan sebesar 84%. Tingkat presisi ini menggarisbawahi potensi teknik pembelajaran mesin dalam bidang analisis sentimen, khususnya dalam konteks layanan keuangan digital. Temuan penelitian ini tidak hanya meningkatkan pemahaman kita tentang persepsi pengguna mengenai platform pembayaran seluler tetapi juga berkontribusi pada wacana yang lebih luas tentang integrasi metode analisis tingkat lanjut dalam mengevaluasi sikap konsumen dalam ekonomi digital yang berkembang pesat.

Pada penelitian yang dilakukan oleh Muhammad Yusril Aldean dan Novanda Alim Setya Nugraha dari Institut Teknologi Telkom Purwokerto, digunakan algoritma Random Forest untuk analisis sentimen terhadap 1.500 tweet mengenai vaksinasi Sinovac. Dengan TF-IDF dan SMOTE sebagai teknik praproses, model divalidasi menggunakan K-Fold dan Confusion Matrix, menghasilkan akurasi 79%, precision 85%, recall 90%, dan F1-score 88%[8].

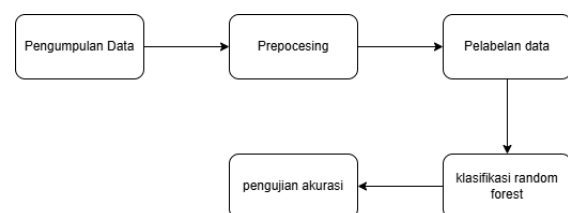
Pada penelitian yang dilakukan oleh Thifal Fadiyah Basar, Dian Eka Ratnawati, dan Issa Arwani dari Universitas Brawijaya, dilakukan klasifikasi opini pengguna Twitter terhadap platform ShopeePay menggunakan algoritma Random Forest. Dengan parameter jumlah tree 300 dan kedalaman 55, diperoleh

akurasi sebesar 95%, precision 95%, recall 94%, dan F1-Score 95%[9].

Dari keempat penelitian yang telah disebutkan, terdapat kesamaan dalam penggunaan algoritma Random Forest sebagai metode klasifikasi sentimen, khususnya terhadap opini publik di platform digital seperti Twitter dan YouTube. Selain itu, rata-rata akurasi yang diperoleh cukup tinggi, berkisar antara 79% hingga 95%, yang menunjukkan efektivitas Random Forest dalam tugas analisis sentimen. Berdasarkan hasil tersebut, kami tertarik untuk menerapkan algoritma Random Forest dalam menganalisis ulasan pengguna terhadap game Mobile Legends: Bang Bang untuk mengetahui persepsi gamer berdasarkan data ulasan dari Google Play Store.

II. METODOLOGI PENELITIAN

Bagian ini menjelaskan secara rinci tentang penelitian yang dilakukan[10].



Gambar 1. Alur Penelitian

A. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan secara daring melalui platform Google Play Store dan YouTube, dengan fokus pada ulasan pengguna terhadap aplikasi Mobile Legends: Bang Bang. Untuk memperoleh data secara sistematis dan efisien, peneliti menggunakan sebuah library Python bernama google_play_scraper[11]. Tools ini memungkinkan penarikan data ulasan dari aplikasi Android secara langsung tanpa harus melakukan scraping manual.

Sebanyak 7000 ulasan berhasil dikumpulkan dari halaman aplikasi Mobile Legends: Bang Bang yang tersedia di Google Play Store. Data yang diambil meliputi teks ulasan dan rating bintang dari pengguna. Selain itu, peneliti juga mengumpulkan 1000 komentar dari video YouTube yang berkaitan dengan Mobile Legends: Bang Bang menggunakan YouTube Data API.[12] Komentar-komentar ini ditujukan untuk memperluas cakupan persepsi gamer dari platform lain yang banyak digunakan untuk berdiskusi dan memberikan tanggapan terhadap permainan tersebut.

Proses pengambilan data dilakukan secara otomatis menggunakan script Python yang mengakses masing-masing endpoint dari Google Play dan YouTube, kemudian menyimpan hasilnya dalam format CSV untuk keperluan analisis lebih lanjut. Data yang diperoleh bersifat publik karena ditampilkan secara terbuka oleh pengguna di masing-masing platform, dan telah melalui proses penyaringan untuk memastikan hanya ulasan atau komentar berbahasa Indonesia yang disertakan dalam penelitian ini.

B. Preprocessing

Setelah data ulasan berhasil dikumpulkan dari Google Play, tahap selanjutnya yang dilakukan adalah proses

preprocessing atau pra-pemrosesan data. Pra-pemrosesan ini bertujuan untuk membersihkan data teks dari elemen-elemen yang tidak relevan serta menyiapkannya agar dapat dianalisis secara lebih akurat oleh algoritma klasifikasi[13]. Tahapan pre-processing ini mencakup case folding (mengubah semua huruf menjadi huruf kecil), tokenizing (memisahkan kalimat menjadi kata-kata), filtering (menghapus kata-kata tidak penting atau stopwords seperti "yang", "dengan", dan "dari"), serta stemming (mengembalikan kata ke bentuk dasarnya, seperti "bermain", "bermainlah", menjadi "main"). Dalam proses ini, digunakan berbagai library Python seperti nltk untuk menghapus stopwords, emoji untuk menghilangkan simbol dan emotikon, serta Sastrawi sebagai stemmer bahasa Indonesia. Penghapusan tanda baca juga dilakukan menggunakan modul string.punctuation. Seluruh proses ini bertujuan untuk menyederhanakan teks, mengurangi keragaman bentuk kata, dan meningkatkan akurasi model dalam menganalisis sentimen dari data ulasan yang diperoleh.

C. Pelabelan

Proses pelabelan data dalam penelitian ini bertujuan untuk memberikan anotasi sentimen pada ulasan pengguna Mobile Legends yang dikumpulkan dari Google Play Store dan YouTube. Pelabelan dilakukan secara otomatis untuk mendukung efisiensi dan konsistensi dalam tahap klasifikasi sentimen. Peneliti menggunakan pendekatan deep learning berbasis model pra-latih bernama `w1lwo/indonesian-roberta-base-sentiment-classifier`, yang merupakan model RoBERTa khusus untuk klasifikasi sentimen dalam bahasa Indonesia. Model ini berperan penting dalam penelitian karena memungkinkan sistem klasifikasi yang akurat tanpa perlu pelabelan manual yang subjektif dan memakan waktu.[14].

Proses kerja pelabelan dimulai dengan membaca data ulasan menggunakan library `pandas`, kemudian setiap komentar pada kolom teks diproses melalui tokenizer dan diklasifikasikan menggunakan model RoBERTa yang diunduh melalui library `transformers` dari Hugging Face. Model dijalankan menggunakan TensorFlow untuk efisiensi pemrosesan. Hasil klasifikasi berupa label sentimen (Positif, Netral, atau Negatif) ditambahkan ke dalam dataset sebagai kolom baru. Dataset yang telah dilabeli secara otomatis kemudian disimpan dalam format CSV dan digunakan untuk tahap pelatihan model klasifikasi menggunakan algoritma Random Forest. Dengan demikian, proses pelabelan otomatis ini menjadi bagian penting yang menghubungkan tahap pengumpulan data dengan tahap pemodelan dalam penelitian ini.

D. Klasifikasi Random forest

Pemodelan Random Forest digunakan dalam penelitian ini untuk melakukan klasifikasi sentimen berdasarkan ulasan pengguna terhadap aplikasi Mobile Legends: Bang Bang. Random Forest adalah teknik ensemble learning yang memanfaatkan gabungan beberapa pohon keputusan (decision tree) guna menghasilkan prediksi yang lebih akurat dan stabil. Setiap pohon dalam Random Forest dilatih pada subset acak dari data pelatihan dan menghasilkan prediksi yang kemudian digabungkan

melalui voting mayoritas untuk menentukan hasil akhir klasifikasi.[15]

Dalam penelitian ini, dataset yang telah dibagi menjadi dua bagian, yaitu 80% data untuk proses pelatihan dan 20% untuk pengujian, sesuai dengan praktik standar dalam evaluasi model klasifikasi.

Setelah proses pelabelan selesai, data vektor teks diubah menjadi representasi numerik menggunakan teknik TF-IDF (Term Frequency-Inverse Document Frequency). TF-IDF digunakan untuk mengekstraksi fitur penting dari data teks dengan memberikan bobot pada kata-kata yang paling relevan dalam dokumen. Proses ekstraksi fitur dilakukan menggunakan library `sklearn.feature_extraction.text.TfidfVectorizer` dengan pengaturan `max_features=5000` untuk membatasi jumlah fitur agar tetap optimal.

Pada proses pelatihan Random Forest, dilakukan tuning parameter menggunakan `GridSearchCV` dari `scikit-learn`. Parameter yang disetel meliputi jumlah estimators (`n_estimators`) yang diuji dengan nilai 100, 200, dan 300; serta kedalaman pohon (`max_depth`) dengan opsi None, 10, dan 30. Kombinasi parameter terbaik dipilih berdasarkan hasil validasi silang (cross-validation) dengan skor `f1` tertinggi.

E. Evaluasi

Penilaian yang diterapkan dalam penelitian ini menggunakan kerangka kerja Confusion Matrix untuk mengevaluasi metrik kinerja sistem yang dikembangkan secara kuantitatif, sebagaimana yang diperoleh dari hasil klasifikasinya. Pendekatan metodologis ini memungkinkan interpretasi sistematis dari akurasi prediktif dan distribusi kesalahan dalam hasil operasional model.

Untuk menilai kinerja model klasifikasi dalam penelitian ini, akurasi, presisi, recall, dan `f1-score` dihitung. Persamaan berikutnya (1-4) merinci rumus yang digunakan untuk evaluasi kinerja ini.[10]

$$Akurasi = \frac{Tp + Tn}{(Tp + Fp + Fn + Tn)} \times 100\% \quad (1)$$

$$presisi = \frac{Tp}{(Tp + Fp)} \times 100\% \quad (2)$$

$$recall = \frac{Tp}{(Tp + Fn)} \times 100\% \quad (3)$$

$$f1 - score = \frac{2 \times presisi \times recall}{(recall + presisi)} \times 100\% \quad (4)$$

III. HASIL DAN PEMBAHASAN

Dalam penelitian ini, penulis mengumpulkan data dari platform Google Play Store dan YouTube. Komentar pengguna diperoleh melalui proses crawling dan disimpan dalam file berformat .csv. Total data yang terkumpul terdiri dari ratusan ulasan dari youtube dan ribuan ulasan dari google play yang membahas persepsi pengguna terhadap game Mobile Legends: Bang Bang. Data tersebut kemudian dilabeli secara otomatis

menggunakan model bahasa Roberta untuk mengklasifikasikan sentimen menjadi kategori positif dan negatif. Selanjutnya, klasifikasi dilakukan dengan memanfaatkan algoritma Random Forest. untuk mengukur performa model berdasarkan nilai akurasi, precision, recall, dan F1-Score.

A. Pengumpulan Data

Pengumpulan data menggunakan rentan waktu 12 maret hingga 9 april 2025 dengan hasil data pada platform google play store sebanyak 7000 data dan pada youtube sebanyak 1000 data.

Dengan hasil pengumpulan data dari google play store dan youtube sebagai berikut

Tabel 1. Data screaping dari google play

No	komentar
1	kgak bisa login
2	apasi min mau pindah akun gabisa"" padahal terhubung ke ml akun ku satu nya, susah pengen gonta ganti min tolong min plissdðŸ™ðŸŸ»
3	mantap

Pengumpulan data dari Google Play Store menggunakan library google_play_scraper dan dan YouTube melalui YouTube Data API, dengan proses otomatis menggunakan script Python yang menyimpan data dalam format CSV, dimana semua data telah difilter hanya mencakup konten berbahasa Indonesia yang tersedia secara publik.

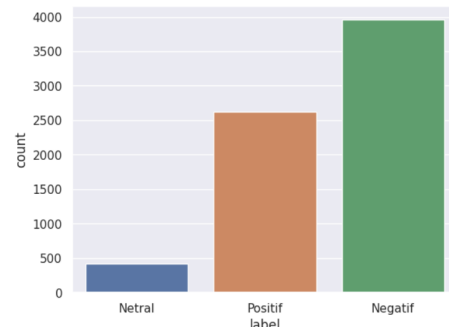
B. Preprocessing

Tahap kedua dari pra-pemrosesan data, yang dijalankan melalui Natural Language Toolkit (NLTK) dalam Python, melibatkan normalisasi teks melalui konversi huruf besar/kecil. Proses ini memastikan keseragaman dengan mengubah semua karakter menjadi huruf besar/kecil yang konsisten, sehingga meningkatkan akurasi dan efisiensi analisis teks dan tugas pemrosesan bahasa alami selanjutnya.

Tabel 3. Data setelah melauai preproceasing

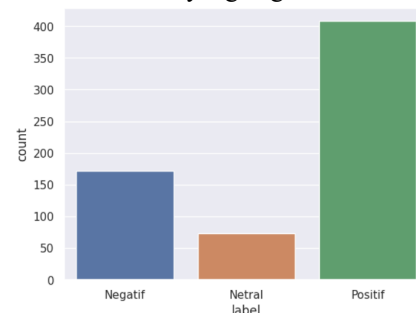
no	komentar	Setelah pre-proceasing
1	Min cara dapetin skinnya secara gratis gimana	min cara dapetin skinnya gratis gimana
2	Serasa perang shinobi main mL sekarang, seru	serasa perang shinobi main sekarang seru
3	Musik nya harus lebih bagus dari musik ff ga sih min	Positif,musik nya lebih bagus musik sih min
4	Tolong buat aku hoki pas ngegacha ya ml, semoga dpt skin nya lgsg di gacha. Mau nuker semua nya ðŸŸð	buat aku hoki pas ngegacha moga dpt skin nya lgsg gacha mau nuker semua nya
5	bang pertamanya deg-degan mau defeat :-D	bang pertama degdegan mau defeat nyata victory

C. Pelabelan



Gambar 2. hasil pelabelan data dari google play

Gambar diatas menunjukan pelabelan dari data yang didapatkan melalui google play dengan 423 netral 2619 data positif dan 3958 data negative maka di playstore memiliki ulasan yang negatif.



Gambar 3. hasil pelabelan Youtube

Data diatas menunjukan 72 data netral 408 data positif,dan 172 data negative dan di youtube mebile legend memiliki ulasan yang positif.

D. Klasifikasi Random Forest

```

PROCEDURE SplitData:
  INPUT:
    df: DataFrame dengan kolom 'content' dan 'label'
    test_size: rasio data test (default: 0.2)
    random_state: seed untuk reproduktibilitas (default: 42)

  OUTPUT:
    x_train: data training untuk fitur
    x_test: data test untuk fitur
    y_train: data training untuk label
    y_test: data test untuk label

  BEGIN
    // Ekstrak fitur dan label dari dataframe
    SET x TO df.content
    SET y TO df.label

    // Bagi data menjadi training set dan test set
    // dengan stratifikasi berdasarkan label y
    SPLIT (x, y) INTO (x_train, x_test, y_train, y_test)
      WITH test_size = test_size
      STRATIFIED BY y
      RANDOM SEED = random_state

    // Kembalikan bentuk/dimensi dari setiap subset
    RETURN (SHAPE_OF(x_train), SHAPE_OF(x_test),
            SHAPE_OF(y_train), SHAPE_OF(y_test))
  END
  
```

Gambar 4. Pembagian data training dan testing

Membagi data menjadi 80% untuk training dan 20% untuk testing yang berarti dalam data yang berasal dari google play sekitar 5600 data untuk pelatihan dan 1400

data untuk pengujian dan untuk data yang berasal dari youtube sebanyak 640 untuk pelatihan dan 160 untuk pengujian. kemudian klasifikasi random forest digunakan untuk mengetahui apakah pendapat bersifat positif,netral atau negative.

E. Evaluasi

Actual	Negatif	748	1	43
	Netral	78	3	3
	Positif	150	1	373
		Negatif	Netral	Positif

Prediction

Gambar 5. hasil prediksi data googl play

Gambar data diatas adalah hasil uji dari confusion matrix dari data google play dengan perhitungan akurasi.

$$\begin{aligned}
 \text{akurasi} &= \frac{Tp + Tn}{Tp + Tn + fp + fn} \\
 &= \frac{748 + 3 + 373}{748 + 1 + 43 + 78 + 3 + 3 + 150 + 1 + 373} \\
 &= \frac{1124}{1400} = 0.8029
 \end{aligned}
 \tag{1}$$

Berdasarkan perhitungan diatas didapatkan hasil akurasi dari data google play adalah 80,29%

	precision	recall	f1-score	support
Negatif	0.77	0.94	0.85	792
Netral	0.60	0.04	0.07	84
Positif	0.89	0.71	0.79	524
accuracy			0.80	1400
macro avg	0.75	0.56	0.57	1400
weighted avg	0.80	0.80	0.78	1400

Gambar 6.Hasil perhitungan confusion matrix data google play

Data dari Google Play Store dievaluasi menggunakan metrik precision, recall, dan f1-score untuk menilai kinerja model. Dari hasil pengujian terhadap 1400 data ulasan, secara keseluruhan model mendapatkan performa yang cukup baik, dengan nilai *accuracy* mencapai 80%. Untuk kelas *Negatif*, model berhasil memperoleh *precision* sebesar 0.77 dan *recall* yang tinggi sebesar 0.94, menghasilkan *f1-score* sebesar 0.85. Hal ini menunjukkan

bahwa model sangat baik dalam mendeteksi ulasan negatif. Pada kelas *Positif*, model mencatat *precision* sebesar 0.89 dan *recall* sebesar 0.71, dengan *f1-score* sebesar 0.79, yang juga mencerminkan performa yang stabil dan akurat dalam mengenali ulasan positif. Namun, performa model pada kelas *Netral* relatif rendah, dengan *precision* hanya 0.60 dan *recall* sebesar 0.04, menghasilkan *f1-score* yang rendah yaitu 0.07. Rendahnya hasil pada kelas netral kemungkinan disebabkan oleh distribusi data yang tidak seimbang serta kesulitan model dalam membedakan ekspresi netral yang sering tumpang tindih dengan sentimen positif atau negatif secara implisit. Selain distribusi data yang timpang, tingginya kemiripan semantik antara komentar netral dengan komentar positif atau negatif juga menjadi tantangan. Komentar seperti “lumayan”, “boleh lah”, atau “biasa aja” sering kali ditafsirkan berbeda oleh model, tergantung konteks dan kosakata sekitarnya. Hal ini menunjukkan bahwa klasifikasi netral membutuhkan pendekatan berbasis konteks yang lebih kuat, yang mungkin lebih cocok diselesaikan dengan model berbasis deep learning kontekstual seperti BERT atau LSTM. Penggunaan teknik aspect-based sentiment analysis juga dapat membantu mengidentifikasi sentimen yang hanya muncul pada aspek tertentu dari game, tanpa menyimpulkan keseluruhan sentimen sebagai positif atau negatif. Meskipun demikian, nilai *weighted average f1-score* sebesar 0.78 menunjukkan bahwa model ini cukup andal untuk klasifikasi sentimen dalam konteks aplikasi Mobile Legends di Google Play Store, khususnya untuk mendeteksi sentimen positif dan negatif.

	precision	recall	f1-score	support
Negatif	0.68	0.79	0.73	34
Netral	0.75	0.60	0.67	15
Positif	0.90	0.87	0.88	82
accuracy			0.82	131
macro avg	0.77	0.75	0.76	131
weighted avg	0.82	0.82	0.82	131

Gambar 7. Hasil perhitungan confusion matrix data youtube

Evaluasi terhadap model pada data ulasan yang diambil dari YouTube menunjukkan kinerja yang sangat baik dengan nilai *accuracy* sebesar 82%. Model mampu mengklasifikasikan sentimen positif secara akurat, dengan *precision* mencapai 0.90, *recall* sebesar 0.87, dan *f1-score* sebesar 0.88, yang mencerminkan dominasi komentar positif dalam dataset serta kemampuan model dalam mengenali ekspresi positif secara konsisten. Untuk kelas negatif, model mempertahankan kinerja yang konsisten dengan nilai *precision* 0,68 dan *recall* 0,79, yang menghasilkan *f1-score* sebesar 0,73. Sementara itu, meskipun kelas netral memiliki jumlah data yang lebih sedikit, model tetap mampu memberikan hasil yang cukup baik dengan *precision* sebesar 0.75 dan *f1-score* sebesar 0.67. Nilai *macro average f1-score* sebesar 0.76 dan *weighted average f1-score* sebesar 0.82 mengindikasikan bahwa model memiliki kapabilitas yang seimbang dalam mengklasifikasikan berbagai jenis sentimen pada data komentar YouTube. Hal ini menunjukkan bahwa model dapat diandalkan untuk mengolah data dari platform video dengan konteks yang lebih bervariasi dan bahasa yang lebih santai dibandingkan ulasan di Google Play Store.

F. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu dicatat. Pertama, komposisi data tidak seimbang, khususnya pada label sentimen netral yang jumlahnya lebih sedikit dibandingkan label positif dan negatif, sehingga memengaruhi performa model dalam mendeteksi sentimen netral. Kedua, model Random Forest yang digunakan tidak memiliki kemampuan untuk memahami konteks semantik secara mendalam, sehingga tidak mampu menangkap nuansa sentimen dalam kalimat kompleks atau ambigu. Selain itu, teknik praproses standar seperti stemming dan stopword removal dapat menyebabkan hilangnya informasi penting dalam teks ulasan. Terakhir, analisis hanya terbatas pada dua platform (Google Play Store dan YouTube), yang belum sepenuhnya mewakili seluruh persepsi publik terhadap game Mobile Legends di media sosial lainnya. Penelitian lanjutan diharapkan dapat mengatasi keterbatasan ini melalui pendekatan algoritma deep learning, penyeimbangan dataset, dan penambahan data dari platform digital yang lebih luas.

IV. KESIMPULAN

Penelitian ini menganalisis persepsi pengguna terhadap game Mobile Legends: Bang Bang melalui data ulasan dari Google Play Store dan YouTube menggunakan metode klasifikasi Random Forest. Model bahasa Roberta digunakan untuk melakukan pelabelan secara otomatis, yang memungkinkan identifikasi sentimen secara efisien terhadap 1.400 ulasan dari Play Store dan ratusan ulasan dari YouTube. Evaluasi model menunjukkan performa yang baik, dengan akurasi sebesar 80% pada data Google Play Store dan 82% pada data YouTube. Nilai f1-score tertinggi diperoleh pada kelas negatif dan positif, yang menunjukkan bahwa model sangat efektif dalam mendeteksi ekspresi sentimen dominan pengguna. Namun, performa rendah pada kelas netral menjadi catatan penting, yang kemungkinan disebabkan oleh ketidakseimbangan distribusi data dan kompleksitas dalam membedakan ekspresi netral. Implikasi praktis dari penelitian ini adalah bahwa model dapat digunakan untuk memahami opini publik terhadap sebuah game secara lebih akurat, yang berguna bagi pengembang dalam meningkatkan pengalaman pengguna. Penelitian ini juga membuka peluang penerapan metode serupa dalam analisis sentimen terhadap aplikasi digital lainnya. Untuk penelitian lanjutan, disarankan untuk melakukan balancing data dan mengeksplorasi algoritma berbasis deep learning guna meningkatkan ketepatan klasifikasi, khususnya pada kategori netral.

V. SARAN

Dalam penelitian selanjutnya, disarankan untuk meningkatkan performa model terutama pada kelas netral dengan cara menyeimbangkan distribusi data melalui teknik seperti oversampling atau undersampling. Selain itu, pengujian dengan berbagai algoritma pembelajaran mesin lainnya seperti Support Vector Machine (SVM), XGBoost, atau pendekatan berbasis deep learning seperti LSTM dan BERT dapat dilakukan untuk membandingkan performa klasifikasi secara lebih komprehensif. Menambah jumlah data ulasan dari berbagai platform serta mempertimbangkan analisis aspek-aspek sentimen (aspect-based sentiment analysis) juga direkomendasikan untuk memperoleh pemahaman yang lebih

mendalam terhadap persepsi pengguna terhadap fitur-fitur tertentu dari game Mobile Legends.

REFERENSI

- [1] R. Andreas and D. Arymami, "Volume 8 Nomor 2 Tahun 2021 Hasrat Konsumsi Virtual dalam Permainan Daring Mobile legends: Perspektif Deleuze dan Guattari," *Jurnal Pemikiran Sosiologi*, 2022.
- [2] W. M. Haidar and E. R. Antika, "Prestasi Akademik Siswa Ditinjau Dari Kecanduan Bermain Game Online Mobile Legends," *TERAPUTIK: Jurnal Bimbingan dan Konseling*, vol. 5, no. 3, pp. 319–325, Feb. 2022, doi: 10.26539/teraputik.53872.
- [3] R. H. Nufus and U. Surapati, "Analisis Sentimen Persepsi Masyarakat Terhadap Timnas Indonesia U-23 dalam AFC-23 Asian Cup 2024 Pada Media Sosial X Menggunakan Metode Naïve Bayes Classifier," *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, vol. 5, no. 3, pp. 2647–2657, Sep. 2024, doi: 10.35870/jimik.v5i3.964.
- [4] J. Sirusstara, N. Alexander, A. Alfariy, S. Achmad, and R. Sutoyo, "Clickbait Headline Detection in Indonesian News Sites using Robustly Optimized BERT Pre-training Approach (RoBERTa)," in *2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, IEEE, Sep. 2022, pp. 1–6, doi: 10.1109/AiDAS56890.2022.9918678.
- [5] Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang, and X. Liang, "An improved random forest based on the classification accuracy and correlation measurement of decision trees," *Expert Syst Appl*, vol. 237, p. 121549, Mar. 2024, doi: 10.1016/j.eswa.2023.121549.
- [6] A. Prasetyo, T. Ridwan, and A. Voutama, "Analisis Sentimen Terhadap Aplikasi GBWhatsApp Menggunakan Naive Bayes Classifier dan Random Forest Classifier," *JSiI (Jurnal Sistem Informasi)*, vol. 11, no. 1, pp. 1–9, Mar. 2024, doi: 10.30656/jsii.v11i1.6936.
- [7] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 9, pp. 4305–4313, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [8] M. Y. Aldean, P. Paradise, and N. A. Setya Nugraha, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 di Twitter Menggunakan Metode Random Forest Classifier (Studi Kasus: Vaksin Sinovac)," *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, vol. 4, no. 2, pp. 64–72, Jun. 2022, doi: 10.20895/inista.v4i2.575.
- [9] T. Fadiyah Basar, D. E. Ratnawati, and I. Arwani, "Analisis Sentimen Pengguna Twitter terhadap Pembayaran Cashless menggunakan ShopeePay dengan Algoritma Random Forest," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 3, pp. 1426–1433, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>

- [10] M. H. Aufan, M. R. Handayani, A. B. Nurjanna, N. Cahyo, H. Wibowo, and K. Umam, "The Perceptions of Semarang Five Star Hotel Tourists With Support Vector Machine on Google Reviews," *Jurnal Teknik Informatika (JUTIF)*, vol. x, No. y, pp. x-y, 2023, doi: 10.52436/jutif.
- [11] R. M. Amir Latif, M. Talha Abdullah, S. U. Aslam Shah, M. Farhan, F. Ijaz, and A. Karim, "Data Scraping from Google Play Store and Visualization of its Content for Analytics," in *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, IEEE, Jan. 2019, pp. 1–8. doi: 10.1109/ICOMET.2019.8673523.
- [12] S. Alanazy and J. Tian, "Measuring API Usability and the Environment: A YouTube API Case Study," 2025, pp. 176–186. doi: 10.1007/978-3-031-76273-4_14.
- [13] K. Adib, M. R. Handayani, W. D. Yuniarti, and K. Umam, "Opini Publik Pasca-Pemilihan Presiden: Eksplorasi Analisis Sentimen Media Sosial X Menggunakan SVM," *SINTECH JOURNAL*, 2024, [Online]. Available: <https://doi.org/10.31598>
- [14] Y. O. Sihombing and N. V. Situmorang, "Prediksi Sentimen Pada Teks Media Sosial Corporate University Menggunakan RoBERTa," *Prosiding PITNAS Widyaiswara I*, vol. 12, no. 12, 2024.
- [15] L. Rangga Aditya Tarigan and D. Dahlan, "Optimasi Fitur Dengan Forward Selection pada Estimasi Tingkat Penyakit Paru-paru Menggunakan Algoritma Klasifikasi Random Forest," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 5, pp. 10341–10348, Sep. 2024, doi: 10.36040/jati.v8i5.11064.