

Sentimen dan Klasifikasi Aspek Ulasan Mobile Legends dan Honor of Kings Menggunakan Support Vector Machine

¹Alfin Nur Fadli, ²Danang Wahyu Widodo, ³Patmi Kasih

¹Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹0070.alfinnur@gmail.com, ²danangwahyuwidodo@unpkediri.ac.id, ³patmikasih@unpkediri.ac.id

Penulis Korespondens : Alfin Nur Fadli

Abstrak—Pertumbuhan pesat game bergenre *Multiplayer Online Battle Arena (MOBA)* memicu tingginya volume ulasan pengguna yang memuat informasi penting bagi pengembang. Penelitian ini bertujuan membangun sistem analisis sentimen komparatif dan pemodelan aspek berbasis *Flask* menggunakan 4.000 data ulasan dari game *Mobile Legends* dan *Honor of Kings*. Guna mengatasi ketimpangan kelas, diterapkan metode *downsampling* yang dikombinasikan dengan *Synthetic Minority Oversampling Technique (SMOTE)*. Fitur teks diekstraksi menggunakan *TF-IDF Bigram* dan diklasifikasikan menggunakan algoritma *Support Vector Machine (SVM) Linier*. Hasil penelitian menunjukkan bahwa pendekatan berantai pada *preprocessing* berhasil mereduksi noise teks ulasan secara optimal. Selain itu, sistem berhasil mengelompokkan ulasan ke dalam tujuh aspek spesifik industri game melalui aturan *keyword matching*. Pengujian manual dengan pendekatan hibrida juga terbukti andal menyempurnakan klasifikasi sentimen netral sekaligus mengukur kedekatan karakteristik bahasa ulasan baru menggunakan *Cosine Similarity*.

Kata Kunci— Analisis Sentimen, *Cosine Similarity*, Klasifikasi Aspek, *MOBA*, *Support Vector Machine*.

Abstract— The rapid growth of *Multiplayer Online Battle Arena (MOBA)* games has triggered a high volume of user reviews containing vital information for developers. This study aims to build a comparative sentiment analysis and aspect modeling system based on *Flask* using 4,000 review data from *Mobile Legends* and *Honor of Kings*. To address class imbalance, a downsampling method combined with the *Synthetic Minority Oversampling Technique (SMOTE)* was applied. Text features were extracted using *TF-IDF Bigram* and classified using the *Linear Support Vector Machine (SVM)* algorithm. The results show that the chained approach in preprocessing successfully reduced text review noise optimally. Furthermore, the system successfully grouped reviews into seven game-specific aspects through keyword matching rules. Manual testing with a hybrid approach also proved reliable in refining neutral sentiment classification while measuring the proximity of new review language characteristics using *Cosine Similarity*.

Keywords— Aspect Classification, *Cosine Similarity*, *MOBA*, Sentiment Analysis, *Support Vector Machine*.

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



I. PENDAHULUAN

Industri *game mobile* dalam beberapa tahun terakhir mengalami pertumbuhan yang sangat cepat. Salah satu jenis permainan yang banyak diminati adalah *Multiplayer Online Battle*

Arena (MOBA), di mana pemain dapat berkompetisi secara *daring* dengan jutaan pengguna aktif di seluruh dunia. *Mobile Legends: Bang Bang* dan *Honor of Kings* merupakan dua game *MOBA* populer yang memiliki basis pengguna besar sehingga menghasilkan banyak ulasan pada platform *Google Play Store*. Ulasan tersebut berisi berbagai pendapat, kritik, saran, dan pengalaman pengguna yang dapat dimanfaatkan untuk mengevaluasi kualitas layanan dan fitur permainan [1].

Google Play Store menyediakan sarana bagi pengguna untuk memberikan penilaian dan ulasan terhadap aplikasi yang digunakan. Data ulasan tersebut dapat menjadi sumber informasi penting bagi pengembang untuk mengetahui tingkat kepuasan pengguna terhadap berbagai aspek aplikasi. Namun, banyaknya jumlah ulasan yang tersedia membuat proses analisis secara manual menjadi tidak efisien dan memerlukan waktu yang lama. Oleh karena itu, diperlukan metode otomatis yang mampu mengolah dan menganalisis data ulasan dalam jumlah besar secara efektif [2].

Salah satu pendekatan yang banyak digunakan untuk mengekstraksi informasi dari data ulasan adalah analisis sentimen. Analisis sentimen merupakan bagian dari *text mining* yang bertujuan untuk mengidentifikasi dan mengelompokkan opini pengguna ke dalam kategori positif, netral, atau negatif. Berbagai penelitian menunjukkan bahwa algoritma *Support Vector Machine (SVM)* memiliki kinerja yang baik dalam tugas klasifikasi teks karena mampu menangani data berdimensi tinggi dan menghasilkan tingkat akurasi yang kompetitif. Selain itu, metode *Term Frequency-Inverse Document Frequency (TF-IDF)* banyak digunakan untuk merepresentasikan teks ke dalam bentuk numerik sehingga dapat diproses oleh algoritma pembelajaran mesin [3].

Dalam proses klasifikasi sentimen, salah satu permasalahan yang sering ditemui adalah ketidakseimbangan jumlah data pada setiap kelas sentimen. Kondisi ini dapat menyebabkan model cenderung memprediksi kelas mayoritas dan menurunkan kemampuan dalam mengenali kelas minoritas. Untuk mengatasi permasalahan tersebut, dapat diterapkan teknik *Synthetic Minority Over-sampling Technique (SMOTE)* yang bertujuan menyeimbangkan distribusi data pelatihan sehingga model dapat mempelajari karakteristik setiap kelas secara lebih baik [4].

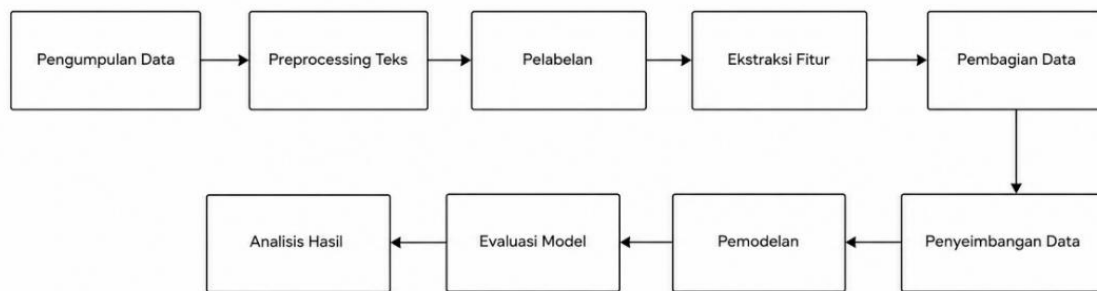
Selain menganalisis sentimen, identifikasi aspek dalam ulasan juga penting karena mampu memberikan informasi yang lebih rinci mengenai bagian aplikasi atau permainan yang memperoleh tanggapan positif maupun negatif dari pengguna. Pendekatan *aspect-based sentiment analysis* telah digunakan untuk mengidentifikasi aspek-aspek tertentu yang memengaruhi tingkat kepuasan pengguna terhadap suatu layanan atau aplikasi [5]. Dengan mengelompokkan ulasan berdasarkan aspek tertentu, pengembang dapat mengetahui faktor-faktor yang perlu dipertahankan maupun ditingkatkan, seperti kualitas permainan, tampilan grafis, performa aplikasi, dan fitur-fitur lainnya.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen dan mengklasifikasikan aspek ulasan pengguna *Mobile Legends* dan *Honor of Kings* yang diperoleh dari *Google Play Store*. Penelitian ini menggunakan metode *TF-IDF* untuk ekstraksi fitur, algoritma *Support Vector Machine (SVM)* untuk klasifikasi sentimen, serta

teknik *SMOTE* untuk mengatasi ketidakseimbangan data. Hasil penelitian diharapkan dapat memberikan gambaran mengenai persepsi pengguna terhadap kedua game tersebut serta menjadi bahan pertimbangan bagi pengembang dalam meningkatkan kualitas layanan dan pengalaman bermain pengguna [6].

II. METODE

Penelitian ini adalah penelitian eksperimental yang menggunakan pendekatan komputasi, khususnya dengan teknik *Natural Language Processing (NLP)* dan pembelajaran terawasi [4], [6], untuk menganalisis sentimen dari data ulasan pengguna. Subjek penelitian adalah pengguna di Indonesia yang memberikan penilaian dan ulasan melalui platform *Google Play Store* [2], [3], sebagai sarana untuk mendistribusikan aplikasi digital. Objek penelitian ini adalah data teks dari ulasan pengguna yang berasal dari dua permainan *mobile* bergenre *Multiplayer Online Battle Arena (MOBA)*, yaitu *Mobile Legends: Bang Bang (ML)* dan *Honor of Kings (HoK)*. Data ulasan tersebut digunakan sebagai referensi untuk mengetahui perasaan pengguna dan mengelompokkan hal-hal yang menjadi sorotan dalam pengalaman bermain game [7].



Gambar 1. Alur Penelitian

A. Pengumpulan Data

Pengumpulan data ulasan pengguna *Mobile Legends* dan *Honor of Kings* dari *Google Play Store*[1]. Data diperoleh menggunakan teknik *scraping* sehingga diperoleh kumpulan ulasan yang akan digunakan sebagai dataset penelitian[6], hasil dari *scraping* ditunjukkan pada Tabel 1.

Tabel 1. Dataset

REVIEW	SCORE	GAME
GAME SEPI	1	<i>Honor of Kings</i>
game seru dan banyak skin gratis	5	<i>Honor of Kings</i>
saya barusan top up wdp 2 kali dalam waktu bersamaan kenapa masuk 1 saja?	3	<i>Mobile Legends</i>
game bgus	2	<i>Mobile Legends</i>
Total		4000 Dataset

Format Kolom Utama Dataset:

1. *content* (atau review): Berisi teks opini/komentar asli dari pengguna.
2. *score* (atau rating): Skala penilaian bintang 1 sampai 5 yang diberikan pengguna.
3. *game*: Atribut penanda kategori data (*ML* untuk *Mobile Legends* atau *HoK* untuk *Honor of Kings*).

B. Preprocessing

Data ulasan yang telah dikumpulkan kemudian diproses melalui tahap preprocessing untuk meningkatkan kualitas data. Tahapan yang dilakukan meliputi *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan *stemming* menggunakan *library Sastrawi*. Hasil dari tahap ini adalah teks yang lebih bersih dan siap untuk dianalisis[6].

C. Pelabelan (Aspect Labeling)

Pelabelan kelas polaritas sentimen dilakukan secara otomatis dengan memakai skor berupa nilai angka dari rating bintang yang diberikan oleh pengguna. Pendekatan ini sering digunakan dalam penelitian analisis sentimen karena mampu memberikan label sentimen seperti positif, netral, dan negatif [4], [6], secara efektif dari data ulasan pengguna. Selain itu, teks ulasan dikelompokkan ke dalam topik yang lebih spesifik dengan metode pencocokan kata kunci yang sudah ditentukan dalam variabel *labels_dict*. Pemetaan kata kunci tersebut dibagi ke dalam beberapa aspek kualitas aplikasi untuk mengetahui hal-hal yang menjadi perhatian pengguna, sehingga informasi yang didapat tidak hanya menunjukkan perasaan positif atau negatif, tetapi juga faktor-faktor yang mempengaruhi kepuasan pengguna [8].

$$SENTIMEN \begin{cases} \text{Positif jika } score \geq 4 \\ \text{Netral jika } score = 3 \\ \text{negatif jika } score < 3 \end{cases} \quad (1)$$

Tabel 2. Aspect Labeling

No	Aspek	Kata Kunci (Keyword)
1	Gameplay	<i>gameplay, game, hero, skill, karakter, battle, rank, ranked, mode, pertandingan</i>
2	Grafik	<i>grafik, grafis, visual, animasi, desain, efek, tampilan, texture, hd</i>
3	Kontrol	<i>kontrol, control, joystick, analog, tombol, button, sensitivitas, aim, kamera</i>
4	Matchmaking	<i>matchmaking, match, lawan, tim, teammate, musuh, afk, feeder, rank match</i>
5	Monetisasi	<i>monetisasi, diamond, skin, topup, top up, voucher, harga, beli, premium</i>
6	Performa	<i>lag, fps, frame drop, crash, bug, patah-patah, loading, koneksi, server, ping</i>
7	Umum	<i>bagus, jelek, menarik, seru, membosankan, rekomendasi, game terbaik</i>
8	Tidak	Ulasan yang tidak mengandung kata kunci aspek

Terlabel manapun

D. Ekstraksi Fitur

Data teks yang telah diberi label diubah ke dalam bentuk numerik menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)*. Metode ini digunakan untuk merepresentasikan kata-kata dalam bentuk vektor sehingga dapat diproses oleh algoritma *machine learning*[4],[6]

$$TFIDF(t, d) = TF(t, d) \times IDF(t) \quad (2)$$

Dengan:

$$IDF(t) = \log \frac{N}{df(t)} \quad (3)$$

E. Pembagian Data

Dataset kemudian dibagi menjadi data latih (training data) dan data uji (testing data)[6]. Pembagian ini bertujuan untuk melatih model dan mengukur kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dilihat sebelumnya[4], [6].

F. Penyeimbang Data

Pada data latih diterapkan metode *Synthetic Minority Over-sampling Technique (SMOTE)* untuk mengatasi ketidakseimbangan jumlah data pada setiap kelas sentimen. Teknik ini membantu meningkatkan kemampuan model dalam mengenali kelas minoritas Rumus pembentukan data sintesis[4]:

$$x_{baru} = x_i + \lambda(x_{nn} - x_i) \quad (4)$$

Keterangan:

1. x_i = data minoritas asal
2. x_{nn} = tetangga terdekat (*nearest neighbor*)
3. λ = nilai acak antara 0 dan 1

G. Pemodelan

Proses klasifikasi dilakukan menggunakan algoritma *Support Vector Machine (SVM)*. Algoritma ini dipilih karena memiliki performa yang baik dalam klasifikasi teks dan mampu menangani data berdimensi tinggi yang dihasilkan dari proses *TF-IDF*[9]

H. Evaluasi Model

Evaluasi model dilakukan untuk mengetahui sejauh mana algoritma *Support Vector Machine (SVM)* mengelompokkan sentimen dari ulasan pengguna menjadi tiga kategori yaitu positif, netral, dan negatif. Berdasarkan nilai yang didapat dari *confusion matrix*, selanjutnya dihitung beberapa ukuran penilaian, seperti akurasi, presisi, *recall*, dan *F1-score*[6].

Akurasi digunakan untuk mengetahui seberapa tepat model bekerja secara keseluruhan, presisi menunjukkan kemampuan model dalam memberikan prediksi yang benar untuk kelas tertentu, *recall* mengukur seberapa baik model dalam mengenali semua data yang seharusnya termasuk ke dalam kelas tersebut [10]

$$Akurasi = \frac{\text{jumlah prediksi benar}}{\text{total data}} \times 100\% \quad (5)$$

Sedangkan *F1-score* memberikan nilai yang seimbang antara presisi dan *recall* [10], [11], hasil evaluasi yang didapat digunakan untuk mengetahui sejauh mana model berhasil mengklasifikasikan sentimen dari ulasan pengguna memastikan model yang dibuat memiliki performa yang bagus dalam membantu analisis sentimen terhadap ulasan game *Mobile*

Legends dan Honor of Kings[1].

1. *TP (True Positive)* : jumlah data positif yang berhasil diklasifikasikan dengan benar sebagai positif.
2. *FP (False Positive)* : jumlah data negatif yang salah diklasifikasikan sebagai positif.
3. *FN (False Negative)* : jumlah data positif yang salah diklasifikasikan sebagai negatif.
4. *Precision* : tingkat ketepatan model dalam memprediksi suatu kelas positif.
5. *Recall* : kemampuan model dalam menemukan seluruh data yang sebenarnya termasuk ke dalam kelas positif.
6. *F1-Score* : nilai rata-rata antara precision dan recall yang digunakan untuk mengukur keseimbangan kedua metrik tersebut.

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

$$F1 = 2 \times \frac{Precision \times recall}{Precision + recall} \quad (8)$$

I. Analisis Hasil

Tahap terakhir adalah menganalisis hasil klasifikasi sentimen dan aspek ulasan. Analisis dilakukan untuk mengetahui persepsi pengguna terhadap *Mobile Legends dan Honor of Kings* serta membandingkan karakteristik sentimen yang muncul pada kedua game tersebut[1], [6]

III. HASIL DAN PEMBAHASAN

A. Hasil Pengumpulan data

Proses pengumpulan data pada penelitian ini berhasil menghimpun total 4.000 baris data ulasan yang berasal dari dua game *Multiplayer Online Battle Arena (MOBA)* seluler terkemuka. Data tersebut diunggah ke dalam sistem melalui mekanisme *multi-file upload* dengan rincian alokasi sebagai berikut:

1. Dataset 1 : Berisi 2.000 ulasan pengguna untuk game, yang diberi penanda sumber (*dataset_source*) sebagai "Dataset 1" di dalam sistem.
2. Dataset 2 Berisi 2.000 ulasan pengguna untuk game, yang diberi penanda sumber (*dataset_source*) sebagai "Dataset 2" di dalam sistem.

Setelah kedua dataset digabungkan ke dalam satu struktur *data frame (df_combined)*, sistem melakukan tahapan standardisasi dan pembersihan awal sebagai berikut:

1. *Penyelarasan Kolom Ulasan*: Berbagai variasi nama kolom teks asli dari kedua dataset (seperti *review*, *text*, *ulasan*, atau *content*) disatukan otomatis ke dalam satu atribut standar, yaitu *content*.
2. *Penyelarasan Kolom Rating*: Variasi penamaan angka bintang (seperti *rating*, *skor*, atau *score*) disatukan ke dalam atribut standar *score*.
3. *Eliminasi Data Kosong (Data Cleaning)*: Sistem mendeteksi dan menghapus baris data yang cacat atau tidak memiliki nilai (*missing values*) pada kolom teks ulasan maupun skor rating menggunakan fungsi *dropna()* selanjutnya.

4. Pelabelan Sentimen Awal (Ekstraksi Kelas): Nilai pada atribut *score* (skala 1–5) dikonversi secara otomatis oleh sistem menjadi 3 kelas sentimen awal, dengan rincian aturan:
 - a. Sentimen Positif: Ditujukan untuk ulasan Mobile Legends maupun Honor of Kings yang memiliki rating bintang 4 dan 5 (≥ 4).
 - b. Sentimen Netral: Ditujukan untuk ulasan yang memiliki rating bintang 3 ($=3$).
 - c. Sentimen Negatif: Ditujukan untuk ulasan yang memiliki rating bintang 1 dan 2 (≤ 2).

Total 4.000 data ulasan yang telah terintegrasi dan terstandardisasi ini disimpan sementara ke dalam objek `DATA_STORE['df']` dan siap dialirkan menuju tahapan pengolahan awal teks (*preprocessing*).

B. Hasil Pengolahan Awal (Preprocessing)

Tahapan *preprocessing* data ulasan mentah dilakukan secara terstruktur melalui lima langkah berantai untuk mereduksi *noise* dan menghasilkan teks bersih yang siap dikomputasi. Proses diawali dengan *case folding* untuk mengubah teks menjadi huruf kecil sekaligus menyatukan frasa majemuk, diikuti dengan *cleaning* berbasis reguler ekspresi, angka, simbol, serta emotikon. Selanjutnya, kata-kata tidak baku atau *slang* disesuaikan melalui tahap normalisasi kata. Teks kemudian disaring menggunakan teknik *stopword removal* untuk menghapus kata umum, namun sistem secara cerdas tetap mempertahankan kata pembalik makna demi menjaga akurasi analisis sentimen. Terakhir, kata berimbuhan dikonversi menjadi kata dasar menggunakan pustaka *Sastrawi.Stemmer*, berikut hasilnya *preprocessing* ditunjukkan pada Gambar 2.

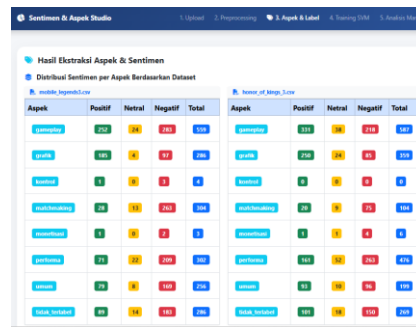
The screenshot shows the 'Sentimen & Aspek Studio' web application. It has a navigation bar with steps: 1. Upload, 2. Preprocessing (active), 3. Aspek & Label, 4. Training SVM, 5. Analisis Manual. Below the navigation bar, there's a section 'Alur Kerja Preprocessing Teks Terurai' with a 'Mulai Preprocessing Berantai' button. A status bar indicates 'STATUS DATASET: Total Berhasil Diupload: 4000 Data'. The main section is 'Pelacakan Proses Perubahan Kalimat (Sampel 5 Data Teratas)' which contains a table showing the transformation of text samples through six steps: 1. Kalimat Asal (Raw), 2. Case Folding, 3. Cleaning Simbol, 4. Normalisasi Kata, 5. Hapus Stopwords, and 6. Stemming (Sastrawi).

Sumber	1. Kalimat Asal (Raw)	2. Case Folding	3. Cleaning Simbol	4. Normalisasi Kata	5. Hapus Stopwords	6. Stemming (Sastrawi)
	game anomali 10% game 90% jualan skin,,ini game bnyk dark system dan bot	game anomali 10% game 90% jualan skin,,ini game bnyk dark system dan bot nva rlan vann	game anomali game jualan skin ini game bnyk dark system dan bot	game anomali game jualan skin ini game bnyk dark system dan bot	anomali jualan skin bnyk dark system bot paling parah bnuk nruar	anomali jual skin bnyk dark system bot paling parah

Gambar 2. Proses *Preprocessing*

C. Hasil Pelabelan Aspek (*Aspect-Based Labeling*)

Sistem mengklasifikasikan ulasan ke dalam beberapa aspek spesifik industri *gaming* menggunakan pendekatan berbasis aturan (*rule-based text matching*). Berdasarkan ekstraksi kata kunci, ulasan dikelompokkan ke dalam 7 kategori aspek, yang ditunjukkan pada Gambar3.



Gambar 3. Aspek & Label

Di dalam sistem, data yang tidak terlabel ini tetap dihitung jumlah dan persentasenya sebagai bahan evaluasi untuk mengetahui seberapa luas cakupan (*coverage*) dari aturan berbasis kata kunci yang Anda bangun terhadap dataset, Ulasan mungkin hanya berisi ekspresi emosi atau kalimat umum yang tidak merujuk pada fitur spesifik game (contoh: "sepi" atau "good"), sehingga lolos penyaringan aturan kata kunci penanda aspek ditunjukkan pada Gambar 4.

4. Contoh Data Tidak Terlabel	
TOTAL BERAS DATA KASINDO Makanan yang tidak masuk ke dalam data kurni aspek utama	
	555 Data 11.000
Komentar	Sentimen
kort08888	Positif
pepek udah bintang kash lose streak bati	Negatif
lama lam gua hepiun dapat musuh tidak tara	Negatif
username ku fuadhin not fuadhin ek lucu	Positif
brnyak a baby monoton mending diu brnyak minamar	Negatif
oghey	Positif
barusan top up wdp kali dalam waktu sama kerna masuk saja	Netral

Gambar 4. Data yang Tidak Terlabel

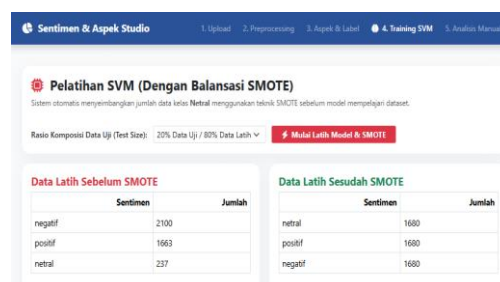
D. Pelatihan Model *SVM*

1. Ekstraksi Fitur *TF-IDF* (*Bigram*)

Teks bersih kemudian diekstraksi menggunakan *TfidfVectorizer* dengan parameter *ngram_range*=(1, 2). Penggunaan kombinasi *Unigram* dan *Bigram* terbukti efektif menangkap konteks kata yang berpasangan, seperti kata "tidak bagus".

2. Sintesis Data dengan SMOTE

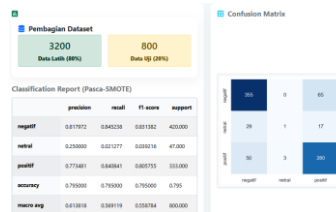
Setelah dataset dibagi menjadi data latih (*training*) dan data uji (*testing*) dengan metode Stratified Split, teknik *SMOTE* (*Synthetic Minority Oversampling Technique*) diterapkan khusus pada data pelatihan. *SMOTE* bertugas menyintesis data baru pada kelas minoritas agar distribusi kelas seimbang secara matematis sebelum diserahkan ke algoritma *SVM*, konfigurasi sebelum & sesudah *smote* ditunjukkan pada Gambar 5.



Gambar 5. Sebelum *Smote* & Sesudah *Smote*

3. Hasil Evaluasi Model

Model dievaluasi menggunakan *Confusion Matrix* dan *Classification Report*. Penggunaan parameter memastikan nilai *precision* atau recall tetap aman secara komputasi meskipun ada kelas yang tidak terprediksi sama sekali, hasilnya ditunjukkan dalam Gambar 7.

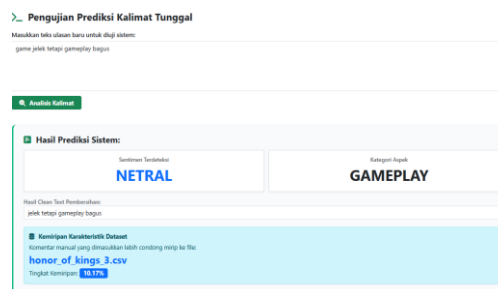


Gambar 7. Confusion Matrix

E. Hasil Pengujian Manual dan Analisis Kemiripan (*Cosine Similarity*)

Pada halaman pengujian manual, sistem menyediakan fungsionalitas untuk mendeteksi ulasan baru yang dimasukkan oleh pengguna:

1. Klasifikasi Sentimen Hibrida: Nilai probabilitas mentah dari *SVM* (*predict_proba*), tetapi juga menambahkan logika berbasis aturan. Jika probabilitas netral $\geq 40\%$ atau terdapat kata konjungsi kontradiktif (“tapi”, “namun”, “tetapi”) dengan selisih probabilitas positif-negatif yang tipis ($\leq 20\%$), sistem secara otomatis mengoreksi prediksi menjadi Netral, analisis manual ditunjukkan pada Gambar 8.
2. Deteksi *Komparasi Dataset* (*Cosine Similarity*): Sistem mampu mengidentifikasi ulasan baru tersebut cenderung lebih mirip ke karakteristik bahasa di Dataset 1 atau Dataset 2. Fitur ini bekerja dengan cara menghitung rata-rata nilai kedekatan kosinus (*mean cosine similarity*)



Gambar 8. Analisis Manual

IV. KESIMPULAN

Berdasarkan hasil perancangan, implementasi, dan pengujian sistem Analisis Sentimen dan Aspek yang telah dilakukan, disimpulkan bahwa sistem berbasis Flask ini berhasil mengintegrasikan dan memproses 4.000 data ulasan komparatif dari game *Mobile Legends* dan *Honor of Kings* secara efisien. Melalui lima tahapan pengolahan awal (*preprocessing*) yang terstruktur termasuk penanganan cerdas terhadap frasa majemuk dan pengecualian kata pembalik makna pada tahap *stopword removal* sistem mereduksi noise teks ulasan secara optimal menjadi kata dasar yang siap dikomputasi. Pendekatan berbasis aturan (*rule-based text matching*) pada fungsi pelabelan aspek terbukti dapat mengelompokkan ulasan pengguna ke dalam tujuh kategori aspek spesifik. Selain itu, penerapan strategi pemangkasan data yang dikombinasikan dengan teknik *SMOTE* pada tahap pelatihan terbukti krusial dalam mengatasi

kendala ketimpangan kelas mayoritas dan minoritas, sehingga model *Support Vector Machine* (SVM). Sistem ini juga menunjukkan performa analisis pengujian kalimat tunggal melalui pendekatan *hibrida*, di mana prediksi probabilitas SVM disempurnakan dengan logika konjungsi kontradiktif untuk menghasilkan label sentimen yang lebih kontekstual, sekaligus mampu mengukur tingkat kemiripan karakteristik bahasa ulasan baru terhadap kedua dataset *game* tersebut.

DAFTAR PUSTAKA

- [1] Davin Zainur Robert, Ahmad Bagus Setiawan, Danang Wahyu Widodo. Implementasi Metode Support Vector Machine untuk Klasifikasi Otomatis Pengaduan Publik di Kabupaten Trenggalek. *Prosiding SEMNAS INOTEK*, Vol. 9, 2025
- [2] Hariyadi, Daud Firdo, & Al Rafi, M. H. (2024). Implementasi Algoritma Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Ulasan Aplikasi Canva. *Jurnal Minfo Polgan*, 13(1), 261-269.
- [3] Idris, I. S. K., & Mustofa, Y. A. (2023). Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM). *Jambura Journal of Electrical and Electronics Engineering (JJEED)*, 5(1), 1-6
- [4] Insan, M. K., Hayati, U., & Nurdiawan, O. (2023). Analisis Sentimen Aplikasi Brimo pada Ulasan Pengguna di Google Play Menggunakan Algoritma Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 583-588.
- [5] Ma'arif, W. A., Sarwido, & Tamrin, T. (2026). Analisis Sentimen Terhadap Ulasan Game Mobile Legend di Playstore menggunakan Algoritma Logistic Regression. *JUKI: Jurnal Komputer dan Informatika*, 8(1), 1-6.
- [6] Maulana, R., Voutama, A., & Ridwan, T. (2023). Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store Menggunakan Algoritma NBC. *Jurnal Teknologi Terpadu*, 9(1), 42-48.
- [7] Nurian, A., & Sari, B. N. (2023). Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, 11(3 S1), 828-834.
- [8] Pranata, Bagus Dwi, Mahdiyah, Umi, Kasih, Patmi. Pemodelan Gaya Belajar Siswa dengan Menggunakan Support Vector Machine
- [9] Pratiwi, R. W., Handayani, S. F., Dairoh, Afidah, D. I., A, Q. R., & F, A. G. (2021). Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector Machine (SVM). *Journal of Informatics, Information System, Software Engineering and Applications (J. OF INISTA)*, 4(1), 40-46.
- [10] Rahman, R. A., Pranatawijaya, V. H., & Sari, N. N. K. (2024). Analisis Sentimen BerbasisAspek pada Ulasan Aplikasi Gojek. *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, 4(1), 70-80.
- [11] Tinaliah, & Elizabeth, T. (2022). Analisis Sentimen Ulasan Aplikasi PrimaKu MenggunakanMetode Support Vector Machine. *Jatisi (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(4), 3436-3442.