

Analisis Sentimen Pengguna pada Ulasan Game Honkai Star Rail Menggunakan Algoritma *Support Vector Machine* (SVM)

Ichsani Mursidah^{1,*}, Remi Sanjaya², Bambang Yulianto³, Dhian Sweetania⁴, Puji Sularsih⁵
^{1,4,5}Prodi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi

Universitas Gunadarma, Jakarta, Indonesia

²Prodi Informatika, Fakultas Teknologi Industri
Universitas Gunadarma, Jakarta, Indonesia

³Prodi Teknik Industri, Fakultas Teknologi Industri
Universitas Gunadarma, Jakarta, Indonesia

^{1,*}Mursidah@staff.gunadarma.ac.id, ²remi@staff.gunadarma.ac.id,

³bambang_yulianto@staff.gunadarma.ac.id, ⁴dhian_sweetania@staff.gunadarma.ac.id,

⁵puji@staff.gunadarma.ac.id

Submit : 26 Jun 2025 | Diterima : 04 Jul 2025 | Terbit : 06 Jul 2025

ABSTRAK

Pesatnya perkembangan industri game digital membawa dampak pada tingginya volume ulasan pengguna di berbagai platform aplikasi. Salah satu game yang mendapat perhatian besar adalah *Honkai Star Rail*. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap game tersebut dengan memanfaatkan algoritma *Support Vector Machine* (SVM). Data yang digunakan berupa 6.193 ulasan pengguna yang diperoleh melalui proses web scraping dari platform *Google Play Store*. Sebelum dilakukan proses klasifikasi, data diproses melalui tahapan *preprocessing* yang meliputi *cleansing*, normalisasi, *tokenizing*, *stopword removal*, dan *stemming*. Kemudian dilakukan pembobotan menggunakan metode TF-IDF. Proses klasifikasi menghasilkan akurasi sebesar 80%, dengan kecenderungan sentimen pengguna lebih dominan pada kategori positif. Hasil penelitian ini dapat menjadi masukan bagi pengembang game dalam meningkatkan kualitas aplikasi berdasarkan persepsi pengguna.

Kata kunci : Analisis Sentimen, Honkai Star Rail, *Support Vector Machine*, *Text Mining*

PENDAHULUAN

Perkembangan teknologi informasi membawa dampak signifikan dalam berbagai aspek kehidupan, salah satunya dalam industri hiburan digital. Game menjadi salah satu bentuk hiburan yang paling banyak diminati masyarakat global. Berdasarkan laporan yang dirilis oleh Jasmine Katatikarn (2024), tercatat sekitar 3,03 miliar penduduk dunia memainkan video game, dan sebanyak 2,82 miliar di antaranya merupakan pemain game daring. Dengan tingginya angka pengguna tersebut, ulasan atau *review* dari para pemain menjadi salah satu indikator utama dalam menilai kualitas sebuah game.

Salah satu game yang menonjol dan mendapatkan perhatian besar dari komunitas global adalah *Honkai Star Rail*. Game ini merupakan salah satu produk unggulan dari HoYoverse yang berhasil menarik perhatian jutaan pemain dalam waktu singkat sejak peluncurannya. Ulasan pengguna terhadap game ini berperan penting dalam memberikan masukan kepada pengembang untuk melakukan perbaikan dan pengembangan lebih lanjut. Mengingat volume ulasan yang sangat besar, proses analisis manual terhadap ulasan pengguna menjadi tidak efisien dan memakan waktu.

Untuk mengatasi permasalahan tersebut, digunakan metode analisis sentimen berbasis algoritma *Support Vector Machine* (SVM). Algoritma SVM dipilih karena memiliki tingkat akurasi yang baik dalam mengklasifikasikan data teks menjadi dua kategori sentimen, yaitu positif dan

negatif. Dengan bantuan analisis sentimen, pengembang dapat memperoleh gambaran persepsi pengguna secara cepat dan akurat. Melalui penelitian ini, diharapkan dapat memberikan kontribusi nyata dalam mendukung pengembangan aplikasi berbasis data ulasan pengguna.

TINJAUAN PUSTAKA

Pada penulisan ini, digunakan berbagai referensi terkait penelitian yang terdiri dari buku, jurnal nasional dan internasional berisikan penjelasan teori-teori yang menjadi sumber dasar penelitian terkait dengan penelitian yang dilakukan serta referensi mengenai pemakaian *deep learning* pada citra *x-ray* untuk penyakit tuberkulosis menggunakan metode *Convolutional Neural Network* (CNN).

Honkai Star Rail

Honkai Star Rail merupakan salah satu game bergenre *turn-based role-playing game* (RPG) yang dikembangkan oleh HoYoverse, perusahaan pengembang game yang juga dikenal melalui judul populer seperti *Honkai Impact 3rd* dan *Genshin Impact*. Game ini dirilis secara global pada 26 April 2023 dan berhasil memperoleh respons positif dari para pemain. Dalam *gameplay*-nya, pemain dituntut untuk mengatur strategi dalam setiap serangan maupun pertahanan secara bergiliran. Kombinasi antara alur cerita yang menarik dengan mekanisme pertarungan yang strategis menjadikan *Honkai Star Rail* memiliki basis pemain yang kuat dan loyal.

Google Play Store

Google Play Store merupakan salah satu platform distribusi aplikasi digital terbesar yang dikembangkan oleh *Google*. Platform ini menyediakan berbagai macam aplikasi, game, dan konten digital lainnya bagi pengguna perangkat berbasis Android. Diluncurkan pertama kali pada tahun 2008, *Google Play Store* kini memiliki lebih dari dua miliar pengguna aktif bulanan di seluruh dunia. Salah satu fitur penting yang tersedia adalah kolom ulasan pengguna, yang menjadi salah satu sumber informasi utama bagi pengembang untuk mengetahui tanggapan terhadap aplikasi atau game yang dirilis.

Web Scraping

Web scraping merupakan metode pengambilan data secara otomatis dari sebuah halaman website. Dalam penelitian ini, teknik web scraping digunakan untuk mengumpulkan data ulasan pengguna dari laman aplikasi *Honkai Star Rail* di *Google Play Store*. Teknik ini umum digunakan dalam penelitian berbasis data, terutama dalam lingkup analisis sentimen, karena kemampuannya dalam memperoleh data dalam jumlah besar secara sistematis dan efisien (Sirisuriya, 2023).

Analisis Sentimen

Analisis sentimen merupakan cabang dari *text mining* yang bertujuan untuk mengidentifikasi dan mengklasifikasikan opini atau perasaan dalam suatu teks menjadi kategori tertentu, seperti positif atau negatif. Teknik ini memanfaatkan pendekatan *Natural Language Processing* (NLP) dan algoritma pembelajaran mesin. Analisis sentimen sering digunakan oleh perusahaan untuk mengetahui persepsi publik terhadap produk atau layanan yang mereka tawarkan (Prakash & Aloysius, 2021). Tantangan dalam analisis sentimen meliputi pengolahan bahasa ambigu, deteksi sarkasme, serta kualitas data yang bervariasi.

Machine Learning

Machine learning merupakan salah satu cabang dari kecerdasan buatan (AI) yang memungkinkan komputer untuk mempelajari pola dari data dan membuat prediksi tanpa diprogram secara eksplisit. Salah satu metode yang umum digunakan dalam analisis data teks adalah *supervised learning*, di mana algoritma belajar dari data yang telah diberi label sebelumnya (Cholissodin et al., 2020).

Support Vector Machine (SVM)

Support Vector Machine merupakan salah satu algoritma pembelajaran mesin berbasis *supervised learning* yang digunakan untuk menyelesaikan masalah klasifikasi. SVM bekerja dengan memisahkan data ke dalam dua kelas berbeda melalui pembuatan *hyperplane* optimal. Keunggulan utama dari SVM adalah kemampuannya dalam menangani data berdimensi tinggi dan performanya yang baik dalam memproses data teks (Sangeetha & Kumaran, 2022).

Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF merupakan metode pembobotan kata yang digunakan untuk mengukur pentingnya suatu kata terhadap sebuah dokumen dalam kumpulan dokumen (Putri Ratna et al., 2019). Pembobotan ini penting dalam proses ekstraksi fitur pada teks agar algoritma klasifikasi seperti SVM dapat bekerja lebih optimal.

Confusion Matrix

Confusion matrix adalah alat evaluasi yang digunakan untuk mengukur kinerja model klasifikasi. Matriks ini menampilkan perbandingan antara hasil prediksi model dengan data sebenarnya, sehingga memudahkan dalam menghitung metrik evaluasi seperti akurasi, presisi, *recall*, dan *f1-score* (Valero-Carreras et al., 2023).

METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan *text mining* dengan memanfaatkan algoritma *Support Vector Machine* (SVM) untuk melakukan analisis sentimen pada ulasan pengguna game *Honkai Star Rail*. Adapun tahapan-tahapan dalam penelitian ini terdiri dari:

Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh melalui proses web scraping pada halaman ulasan aplikasi *Honkai Star Rail* di *Google Play Store*. Proses scraping dilakukan secara otomatis untuk memperoleh ulasan dengan rentang rating tertentu, yaitu 1, 2, 4, dan 5. Total data yang berhasil dikumpulkan berjumlah 6.193 ulasan yang selanjutnya akan diproses pada tahapan berikutnya.

Preprocessing Data

Tahapan preprocessing dilakukan untuk membersihkan data dari unsur-unsur yang tidak relevan serta mempersiapkannya agar sesuai untuk diproses oleh algoritma SVM. Langkah-langkah preprocessing meliputi:

1. **Case Folding:** Mengubah seluruh huruf dalam teks menjadi huruf kecil (lowercase).
2. **Cleansing:** Menghapus karakter yang tidak diperlukan seperti emotikon dan simbol khusus.
3. **Normalisasi:** Mengubah kata tidak baku atau kata singkatan menjadi bentuk yang sesuai dengan kaidah bahasa Indonesia.
4. **Tokenizing:** Memecah kalimat menjadi potongan-potongan kata.
5. **Stopword Removal:** Menghilangkan kata-kata yang dianggap tidak memiliki makna signifikan.
6. **Stemming:** Mengubah kata menjadi bentuk dasarnya dengan menghapus imbuhan.

Pembobotan

Proses pembobotan dilakukan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF digunakan untuk mengukur seberapa penting suatu kata terhadap sebuah dokumen dalam kumpulan dokumen. Hasil dari pembobotan ini akan digunakan sebagai input bagi algoritma klasifikasi. Metode ini memiliki rumus sendiri pada berikut ini.

$$TF(t, d) = \frac{\text{Jumlah kemunculan } t \text{ pada dokumen}}{\text{Total jumlah kata dalam } d}$$

$$IDF(t, D) = \log\left(\frac{N}{|\{d \in D: t \in d\}|}\right)$$

$$TF-IDF(t, d, D) = TF(t, d) \times IDF(t, D)$$

Klasifikasi

Tahapan klasifikasi dilakukan menggunakan algoritma *Support Vector Machine (SVM)*. SVM bekerja dengan memisahkan data ulasan menjadi dua kelas sentimen, yaitu positif dan negatif, dengan memanfaatkan hyperplane optimal. Setiap data ulasan akan dipetakan ke dalam ruang fitur berdasarkan bobot TF-IDF, kemudian diprediksi kategorinya oleh model SVM. SVM memiliki rumus tersendiri yang terlampir pada berikut ini.

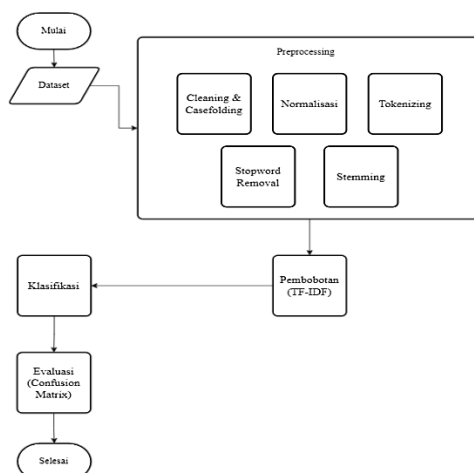
$$f(x) = w^T x + b$$

Evaluasi

Tahap evaluasi dilakukan untuk mengukur kinerja model klasifikasi. Evaluasi dilakukan dengan menggunakan confusion matrix untuk menghasilkan metrik evaluasi berupa precision, recall, *f1-score*, dan akurasi. Selain itu, untuk memberikan visualisasi terhadap hasil analisis, digunakan word cloud untuk menampilkan kata-kata yang paling sering muncul dalam ulasan positif maupun negatif.

HASIL DAN PEMBAHASAN

Analisis pada *game Honkai Star Rail* pada penulisan ilmiah ini, terdapat beberapa tahapan alur yang digunakan, seperti yang ditunjukkan pada Gambar 1.



Gambar 1. *Flowchart* Tahapan Penelitian

Dari Gambar 1 dapat terlihat beberapa tahapan yang dimulai dari *Dataset/Pengumpulan Data*, *Preprocessing* (terdiri dari *cleansing & casefolding*, *normalisasi*, *tokenizing*, *stopword removal*, dan *stemming*), *Pembobotan (TF-IDF)*, *Klasifikasi* (menggunakan SVM), dan berakhir pada *Evaluasi (confusion matrix)*.

Pengumpulan Data

Tahapan awal dalam penelitian ini diawali dengan proses pengumpulan data berupa ulasan pengguna terhadap game *Honkai Star Rail* yang tersedia pada platform distribusi digital *Google*

Play Store. Pengambilan data dilakukan secara otomatis melalui teknik *web scraping* yang dijalankan menggunakan lingkungan pemrograman *Google Colaboratory (Google Colab)*. Total data yang berhasil dikumpulkan berjumlah 6.193 ulasan, yang kemudian diekstraksi secara sistematis dan disusun dalam bentuk tabular. Data yang diperoleh mencakup 12 atribut utama, yaitu: *reviewId* (identitas pengguna), *userName* (nama pengguna), *userImage* (foto profil), *content* (isi ulasan), *score* (rating pengguna terhadap aplikasi), *thumbsUpCount* (jumlah pengguna lain yang menyukai ulasan tersebut), *reviewCreatedVersion* (versi aplikasi saat ulasan dibuat), *at* (tanggal dan waktu ulasan diberikan), *replyContent* (isi balasan dari pengembang), *repliedAt* (waktu balasan diberikan), *appVersion* (versi aplikasi saat ulasan dibuat), dan *appId* (kode identifikasi aplikasi).

Tabel 1. Hasil Pengambilan Data Pengguna ID, Nama, Gambar

<i>reviewId</i>	<i>userName</i>	<i>userImage</i>
41751f51-c72a-49c4-9fdf-2b8b8f5c1ddc	Kikyo97	https://play-lh.googleusercontent.com/a-/ALV-UjVhM7p3P8cKhhh8t4pbRdJrEDIth5mHEIaWJ1EiKxxxq1TjEFfb
1474ec7f-030d-4fde-9770-2c9a802dfd9f	Adelia Janeeta	https://play-lh.googleusercontent.com/a-/ALV-UjWPBEilnFy4_g7KCNaTvrZ4TLg49SU7tAPO9sIQfj173GG2Y1U

Tabel 2. Hasil Pengambilan Data Isi Ulasan, *Rating* dan Jumlah *Like*

<i>content</i>	<i>score</i>	<i>thumbsUpCount</i>
Sebenarnya dah lama pengen ngasih bintang 1 perihal kenapa setiap pra unduh selalu saja size nya tidak normal, seakan nge scam banget asli, fitur yang seharusnya memudahkan pemain malah jadi merepotkan pemain	1	0

Tabel 3. Hasil Pengambilan Data Versi Ulasan, Tanggal dan Waktu Ulasan, Balasan *Developer*

<i>reviewCreatedVersion</i>	<i>at</i>	<i>replyContent</i>
2.3.0	19/05/2025 01:46:32	
2.2.0	14/03/2025 01:10:10	Halo Trailblazer! Mohon kunjungi Web Official Honkai: Star Rail untuk mengecek informasi

Tabel 4. Hasil Pengambilan Data Tanggal Balasan *Developer*, Versi Aplikasi, dan ID Aplikasi

<i>repliedAt</i>	<i>appVersion</i>	<i>appId</i>
	2.3.0	com.HoYoverse.hkrpgoversea
19/03/2025 07:42:34	2.0.0	com.HoYoverse.hkrpgoversea

Pemilihan data dilakukan secara acak namun dengan parameter terarah, yaitu berdasarkan skor atau *rating* yang diberikan oleh pengguna. Rentang skor yang dijadikan acuan terdiri atas nilai 1, 2, 4, dan 5, untuk memastikan representasi dari sentimen negatif dan positif dapat terklasifikasi secara berimbang.

Preprocessing

Dilakukan tahapan *preprocessing* guna mengubah data mentah menjadi bentuk yang terstruktur dan siap diolah secara komputasional. Proses ini bertujuan untuk meningkatkan kualitas data teks agar lebih representatif dan relevan terhadap kebutuhan klasifikasi sentimen yang akan dilakukan pada tahap berikutnya.

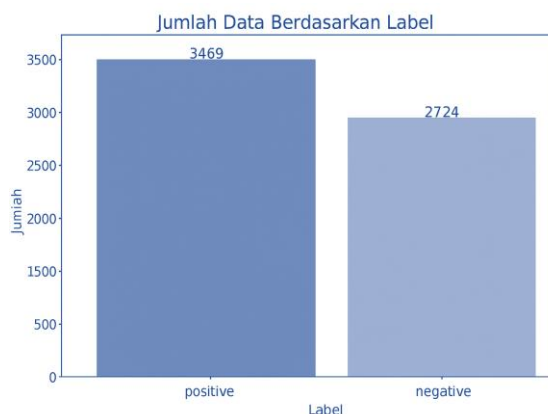
1. Pelabelan

Setiap data ulasan diberi label sentimen berdasarkan rating yang diberikan pengguna. Rating 1 dan 2 diberi label *negative*, sedangkan rating 4 dan 5 diberi label *positive*.

Tabel 5. Hasil Pelabelan

<i>score</i>	<i>label</i>
1	<i>negative</i>
2	<i>negative</i>
3	<i>positive</i>

Dari proses pelabelan diperoleh 2.724 data dengan label *negative* dan 3.469 data dengan label *positive*.



Gambar 2. Bar Pelabelan

2. Cleansing & Case Folding

Cleansing akan menghapus *noise* berupa *emoticon* dan katakter seperti tanda baca. Yang secara bersamaan juga akan melakukan *Casefolding* yang akan mengubah keseluruhan teks menjadi huruf kecil atau *lowercase*. Normalisasi: Mengubah kata tidak baku menjadi kata baku dengan bantuan kamus normalisasi.

Tabel 6. Hasil *Cleansing* dan *Case Folding*

<i>content</i>	<i>cleaning</i>
Gamenya bagus banget. Dari segi cerita, gameplay, dll.	gamenya bagus banget dari segi cerita gameplay dll

3. Normalisasi

Setelah tahap pembersihan awal (*cleansing*), data teks selanjutnya diproses melalui tahap normalisasi. Tahapan ini bertujuan untuk menggantikan kata-kata tidak baku, termasuk istilah gaul maupun singkatan, menjadi bentuk kata yang sesuai dengan kaidah bahasa Indonesia. Proses normalisasi ini dibantu oleh kamus padanan kata yang bersumber dari repositori *open-source* di GitHub, yang digunakan sebagai acuan konversi kata. Teks yang menjadi objek normalisasi adalah hasil keluaran dari proses *cleansing*, sehingga input pada tahap ini sudah terbebas dari simbol dan karakter non-alfabet.

Tabel 7. Hasil Normalisasi

cleaning	normalisasi
sering keluar dari game sejak update tolong aku ingin main dengan tenang	sering keluar dari game sejak update tolong aku ingin main dengan tenang

4. Tokenizing

Langkah selanjutnya adalah tokenisasi (*tokenizing*), yaitu proses pemisahan teks menjadi unit-unit kata (token) berdasarkan tanda pemisah berupa spasi. Setiap kalimat dalam ulasan pengguna dipecah menjadi potongan kata tunggal yang kemudian disimpan dalam format struktur data berupa larik (*array*). Proses ini dilakukan untuk memudahkan sistem dalam memahami setiap elemen teks secara individu sebelum dilakukan analisis lanjutan.

Tabel 8. Hasil *Tokenizing*

normalisasi	tokenize
sering keluar dari game sejak update tolong aku ingin main dengan tenang	['sering', 'keluar', 'dari', 'game', 'sejak', 'update', 'tolong', 'aku', 'ingin', 'main', 'dengan', 'tenang']
overall game ini sudah bagus terutama mekanik dan gameplay	['overall', 'game', 'ini', 'sudah', 'bagus', 'terutama', 'mekanik', 'dan', 'gameplay']

5. Stopword Removal

Setelah proses tokenisasi, dilakukan tahap penghilangan *stopword*. *Stopword* adalah kata-kata umum dalam bahasa yang frekuensi kemunculannya tinggi, tetapi memiliki kontribusi semantik yang rendah dalam konteks analisis. Contoh kata semacam ini antara lain “ini”, “tapi”, “sudah”, “setiap”, dan sejenisnya. Proses *stopword removal* juga turut menghapus kata-kata tidak penting yang belum tersaring pada tahap normalisasi sebelumnya, termasuk singkatan-singkatan yang tidak dikenali.

Tabel 9. Hasil *Stopword Removal*

tokenize	stopword
['sering', 'keluar', 'dari', 'game', 'sejak', 'update', 'tolong', 'aku', 'ingin', 'main', 'dengan', 'tenang']	['sering', 'keluar', 'game', 'sejak', 'update', 'aku', 'main', 'tenang']

6. Stemming

Tahapan akhir dari preprocessing adalah *stemming*, yaitu proses transformasi setiap kata menjadi bentuk dasarnya (akar kata). Proses ini dilakukan dengan cara menghapus awalan, akhiran, dan sisipan pada kata-kata berimbuhan. Tujuannya adalah untuk menyamakan representasi kata-kata turunan dengan bentuk aslinya, sehingga dapat meningkatkan efisiensi dan akurasi dalam tahap klasifikasi.

Tabel 10. Hasil *Stemming*

stopword	stemming
['sering', 'keluar', 'game', 'sejak', 'update', 'aku', 'main', 'tenang']	['sering', 'keluar', 'game', 'sejak', 'update', 'aku', 'main', 'tenang']
['gamenya', 'bagus', 'banget', 'segi', 'cerita', 'gameplay']	['gamenya', 'bagus', 'banget', 'segi', 'cerita', 'gameplay']

Setelah seluruh data melalui proses *stemming*, setiap token kata yang telah dibersihkan dikembalikan ke dalam bentuk representasi kalimat utuh. Proses ini dilakukan dengan menyatukan kembali elemen-elemen token dalam satu baris teks, serta menghilangkan simbol-simbol teknis seperti tanda kurung siku ([]) dan kutip tunggal (') yang semula terbentuk dari struktur data bertipe larik. Langkah ini bertujuan untuk mempersiapkan data dalam format teks yang konsisten dan dapat digunakan sebagai input dalam proses klasifikasi sentimen.

Tabel 11. Hasil Pengubahan Menjadi Teks

stemming	teks
['sering', 'keluar', 'game', 'sejak', 'update', 'aku', 'main', 'tenang']	sering keluar game sejak update aku main tenang
['gamenya', 'bagus', 'banget', 'segi', 'cerita', 'gameplay']	gamenya bagus banget segi cerita gameplay

Pembobotan (TF-IDF)

Tahap pembobotan dilakukan untuk memberikan nilai numerik pada setiap kata dalam dokumen menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF). Metode ini digunakan untuk mengukur tingkat kepentingan suatu kata terhadap dokumen secara relatif terhadap keseluruhan korpus. Dengan TF-IDF, fitur-fitur penting dalam teks dapat diidentifikasi, diprioritaskan, dan digunakan sebagai input representatif untuk algoritma klasifikasi. Pembobotan ini berperan penting dalam meningkatkan akurasi model dalam memahami relevansi kata terhadap konteks ulasan.

Tabel 12. Hasil Pembobotan (TF-IDF)

kata	bobot
overall	4.7074558396868715
game	7.373508095732813
utama	5.427271267328979
gameplay	3.867705184935051
unik	5.799510727808823
hoyoverse	4.365360345511297
puzzle	6.958747638293367
musuh	10.354918937865214
hasil	6.265600457733422
strategi	6.512460535664947

Pengujian Data

Tahap pengujian dimulai dengan proses pembagian dataset ke dalam dua kelompok utama, yaitu data latih (*training data*) dan data uji (*testing data*). Proporsi pembagian yang digunakan dalam penelitian ini adalah 80% untuk data latih dan 20% untuk data uji. Tujuan dari pembagian ini adalah agar model *Support Vector Machine* (SVM) dapat dilatih terlebih dahulu menggunakan data latih, kemudian dievaluasi kinerjanya berdasarkan hasil klasifikasi pada data uji.

Dalam implementasinya, atribut teks hasil *preprocessing* digunakan sebagai variabel independen (*fitur*) yang diberi notasi X , sedangkan label sentimen digunakan sebagai variabel dependen atau target klasifikasi yang dilambangkan dengan y . Kedua variabel ini selanjutnya dipisahkan kembali ke dalam subset *training* dan *testing* dengan rasio yang sama, yaitu 80:20, guna memastikan validitas proses pelatihan dan pengujian model.

Tabel 13. Hasil Uji Data

<i>Data Training</i> Teks	4947
<i>Data Testing</i> Teks	1237
<i>Data Training</i> Label	4947
<i>Data Testing</i> Label	1237

Berdasarkan pembagian tersebut, dari total 6.193 data yang tersedia, sebanyak 4.947 data dialokasikan sebagai data pelatihan, sementara 1.237 sisanya digunakan sebagai data pengujian. Model SVM kemudian dilatih dengan data tersebut dan menghasilkan laporan klasifikasi (*classification report*) yang mencakup metrik evaluasi seperti akurasi, presisi, *recall*, dan *f1-score* seperti pada gambar 3.

```

Classification Report:
              precision    recall  f1-score   support

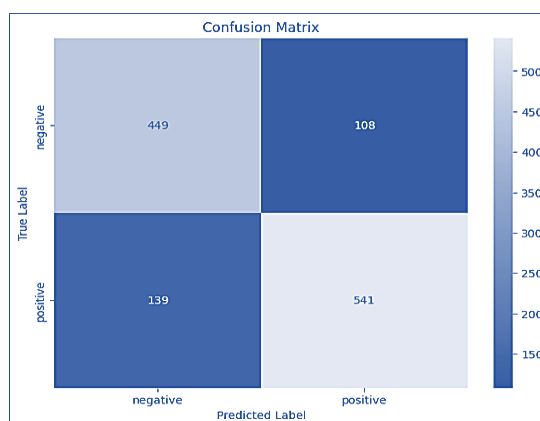
 negative     0.76      0.81      0.78       557
 positive     0.83      0.80      0.81       680

 accuracy                   0.80       1237
 macro avg     0.80      0.80      0.80       1237
 weighted avg  0.80      0.80      0.80       1237
  
```

Gambar 3. Hasil SVM

Evaluasi

Evaluasi terhadap performa model *Support Vector Machine* (SVM) berfungsi untuk mengukur sejauh mana tingkat akurasi dan keandalan model dalam melakukan klasifikasi data. Salah satu metode yang digunakan untuk mengevaluasi hasil klasifikasi adalah *confusion matrix*. Metode ini memberikan representasi visual yang menggambarkan perbandingan antara hasil prediksi model dan nilai aktual pada data uji, sehingga memudahkan identifikasi jenis kesalahan serta tingkat ketepatan klasifikasi.



Gambar 4. *Confusion Matrix* SVM

Berdasarkan hasil evaluasi yang divisualisasikan melalui *confusion matrix*, diperoleh rincian sebagai berikut:

1. **True Negative (TN): 449**
Menunjukkan bahwa terdapat 449 data yang diprediksi sebagai sentimen negatif oleh model, dan memang benar-benar tergolong dalam kelas negatif. Hal ini mencerminkan jumlah prediksi yang tepat untuk kategori negatif.
2. **False Positive (FP): 108**
Merupakan jumlah data yang diprediksi sebagai sentimen positif oleh model, padahal kenyataannya termasuk ke dalam kategori negatif. Kesalahan ini menunjukkan adanya prediksi keliru yang menganggap sentimen negatif sebagai positif.
3. **False Negative (FN): 139**
Mengindikasikan bahwa model memprediksi 139 data sebagai sentimen negatif, meskipun pada kenyataannya data tersebut merupakan sentimen positif. Kesalahan ini dapat menurunkan sensitivitas model terhadap data positif.
4. **True Positive (TP): 541**
Menunjukkan bahwa sebanyak 541 data berhasil diklasifikasikan secara akurat sebagai sentimen positif, sesuai dengan label aslinya. Nilai ini merepresentasikan prediksi yang benar untuk kategori positif.

KESIMPULAN

Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap ulasan game *Honkai Star Rail* dengan menggunakan algoritma *Support Vector Machine* (SVM). Berdasarkan hasil yang diperoleh dari 6.193 data ulasan yang telah melalui proses preprocessing dan pembobotan TF-IDF, model SVM berhasil mencapai tingkat akurasi sebesar **80%**. Evaluasi performa model menunjukkan bahwa SVM memiliki kemampuan yang baik dalam membedakan antara sentimen positif dan negatif, dengan nilai precision, recall, dan f1-score yang relatif seimbang. Visualisasi *word cloud* menunjukkan bahwa ulasan positif didominasi oleh kata-kata yang merepresentasikan kepuasan pengguna terhadap kualitas game, sedangkan ulasan negatif didominasi oleh keluhan teknis seperti bug dan kesulitan login. Dengan demikian, analisis ini tidak hanya mengukur persepsi pengguna secara umum, tetapi juga memberikan informasi yang dapat digunakan oleh pengembang untuk meningkatkan kualitas aplikasi. Berdasarkan hasil penelitian ini, disarankan agar pengembangan analisis sentimen ke depannya mempertimbangkan penggunaan algoritma pembelajaran mesin lainnya, seperti *Random Forest*, *Naive Bayes*, atau pendekatan berbasis *Deep Learning* sebagai pembandingan untuk mengevaluasi performa klasifikasi yang lebih komprehensif. Selain itu, cakupan data dapat diperluas dengan mengambil ulasan dari berbagai platform lain, seperti App Store atau forum komunitas gamer, guna memperoleh persepsi pengguna yang lebih beragam. Untuk meningkatkan akurasi klasifikasi, proses *preprocessing* dapat dikembangkan lebih lanjut dengan menambahkan teknik pengolahan seperti penanganan negasi (*negation handling*) dan deteksi sarkasme (*sarcasm detection*). Ke depannya, sistem analisis sentimen ini juga dapat diintegrasikan ke dalam sistem pemantauan aplikasi secara real-time agar pengembang dapat mengambil keputusan lebih cepat dan responsif terhadap ulasan pengguna.

DAFTAR PUSTAKA

- Prakash, A., & Aloysius, A. (2021). "Sentiment Analysis on Social Media Data Using Machine Learning Algorithms." *Journal of Applied Data Science*, 3(1), 12–23. <https://doi.org/10.1016/j.jads.2021.02.004>
- Sangeetha, K., & Kumaran, A. (2022). "Comparative Study of Support Vector Machine and Naïve Bayes in Sentiment Classification." *International Journal of Intelligent Systems and Applications*, 14(3), 20–27. <https://doi.org/10.5815/ijisa.2022.03.03>
- Sirisuriya, A. (2023). "A Practical Guide to Web Scraping for Natural Language Processing." *Data Science Journal*, 12(4), 44–58.

- <https://doi.org/10.5334/dsj-2023-044>
Valero-Carreras, J., Ramírez, P., & Cazorla, M. (2023). "Evaluating Classification Performance: When and Why to Use Confusion Matrix and Beyond." *IEEE Access*, 11, 77321–77330.
<https://doi.org/10.1109/ACCESS.2023.3273857>
- Akbar, M. N., Rusydi, N. H., H., M. H., Ramadhanti, N. S., & Erfiana. (2022). *Sentiment Analysis of Maxim Application Reviews on Google Play Store Using Support Vector Machine (SVM)*. *AGENTS: Journal of Artificial Intelligence and Data Science*, 2(2), 9–16.
- Tusar, M. T. H. K., & Islam, M. T. (2021, September). A comparative study of sentiment analysis using NLP and different machine learning techniques on US airline Twitter data. In *2021 International conference on electronics, communications and information technology (ICECIT)* (pp. 1-4). IEEE.
<https://doi.org/10.1109/ICECIT54077.2021.9641336>
- Sahbuddin, M., & Agustian, S. (2022). *Support Vector Machine Method with Word2vec for Covid-19 Vaccine Sentiment Classification on Twitter*. *Journal of Informatics and Telecommunication Engineering (JITE)*, 6(1). <https://doi.org/10.31289/jite.v6i1.7534>
- Wicaksono, F. A., Romadhony, A., & Hasmawati. (2022). *Sentiment Analysis of University Social Media Using Support Vector Machine and Logistic Regression Methods*. *Indonesian Journal on Computing (Indo-JC)*, 7(2), 15–24. <https://doi.org/10.34818/INDOJC.2022.7.2.638>
- Tan, J. Y., Chow, A. S. K., & Tan, C. W. (2021, October). Sentiment Analysis on Game Reviews: A Comparative Study of Machine Learning Approaches. In *International Conference on Digital Transformation and Applications (ICDXA2021)* (pp. 209-216).
<https://doi.org/10.56453/icdxa.2021.1023>
- Rizaldi, S. A. R., Alam, S., & Kurniawan, I. (2023). Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes. *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, 2(3), 109-117.
<https://doi.org/10.55123/storage.v2i3.2334>
- Gunawan, A., Gunawan, G., Rohman, R. S., & Kartikasari, R. (2024). ANALISIS SENTIMEN TERHADAP PELAYANAN SAMSAT DI TASIKMALAYA MENGGUNAKAN METODE NAÏVE BAYES. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 2884-2889.
<https://doi.org/10.36040/jati.v8i3.9556>
- Putri, I. P., Terttiaavini, T., & Arminarahmah, N. (2024). Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting pada Anak: Comparative Analysis of Machine Learning Algorithms for Predicting Child Stunting. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(1), 257-265.
<https://doi.org/10.57152/malcom.v4i1.1078>