

Analisis Sentimen pada *Trailer Deadpool vs Wolverine* Menggunakan Model Machine Learning

Verio Aberossy Renedominick¹, Simon Barus Pranata²

^{1,2} Program Studi Informatika, Fakultas Sains, Komputer dan Matematika, Universitas Matana Matana University Tower, Jl. CBD Barat Kav, RW.1, Curug Sangereng, Kelapa Dua, Tangerang Regency

¹verio.renedominick@student.matanauniversity.ac.id, ²simon.barus@matanauniversity.ac.id

Abstract

Social media has become the primary platform for public response to entertainment content, including movie trailers. This research aims to analyze public sentiment in comments on the Deadpool vs Wolverine trailer on YouTube using four machine learning models. Comment data is collected through YouTube Data API v3 and processed using standard preprocessing techniques. Sentiment labeling using Sentiment Intensity Analyzer produces three categories: positive, negative, and neutral. The analysis results show dominance of neutral (7,072) and positive (6,289) comments over negative ones (2,690). Model performance comparison indicates that Logistic Regression achieves the highest accuracy (86.61%), followed by Decision Tree (85.89%), Naive Bayes (79.08%), and KNN (60.07%). This research identifies challenges related to data distribution imbalance and provides insights into the effectiveness of various machine learning models in sentiment analysis of entertainment content.

Keywords: sentiment analysis, machine learning, text classification, social media, YouTube

Abstrak

Media sosial telah menjadi platform utama bagi masyarakat dalam memberikan tanggapan terhadap konten hiburan, termasuk trailer film. Penelitian ini bertujuan untuk menganalisis sentimen komentar publik terhadap trailer film *Deadpool vs Wolverine* di YouTube menggunakan empat model pembelajaran mesin. Data komentar dikumpulkan melalui YouTube Data API v3 dan diproses menggunakan teknik *preprocessing* standar. Pelabelan sentimen menggunakan *Sentiment Intensity Analyzer* menghasilkan tiga kategori: positif, negatif, dan netral. Hasil analisis menunjukkan dominasi komentar netral (7.072) dan positif (6.289) dibanding negatif (2.690). Perbandingan performa model menunjukkan *Logistic Regression* memiliki akurasi tertinggi (86,61%), diikuti *Decision Tree* (85,89%), *Naive Bayes* (79,08%), dan *KNN* (60,07%). Penelitian ini mengidentifikasi tantangan terkait ketidakseimbangan distribusi data dan memberikan wawasan tentang efektivitas berbagai model pembelajaran mesin dalam analisis sentimen konten hiburan.

Kata kunci: analisis sentimen, pembelajaran mesin, klasifikasi teks, media sosial, YouTube

© 2025 Jurnal Pustaka AI

1. Pendahuluan

Di era digital saat ini, media sosial telah menjadi platform utama bagi masyarakat untuk mengekspresikan pendapat dan berinteraksi, khususnya dalam menanggapi konten hiburan. YouTube, sebagai salah satu platform media sosial terbesar, telah menjadi wadah yang signifikan bagi publik untuk memberikan tanggapan dan berkomentar terhadap berbagai konten, termasuk

trailer film [1]. Analisis sentimen terhadap konten di media sosial telah menjadi metode yang semakin penting untuk memahami respons dan persepsi publik. Di era big data saat ini, analisis sentimen dapat dilakukan dengan lebih akurat berkat kemampuan memproses data dalam jumlah masif berupa komentar dan opini pengguna. Namun, volume data yang sangat besar ini juga menciptakan tantangan tersendiri, karena pembacaan dan analisis manual

menjadi sangat tidak efisien dan memakan waktu. Untuk mengatasi hal tersebut, dibutuhkan suatu sistem yang dapat secara otomatis mengklasifikasikan berbagai opini dan ulasan ke dalam kategori-kategori yang relevan [2].

Dalam konteks industri hiburan, trailer film sering kali memicu berbagai tanggapan dari masyarakat di platform media sosial. Masyarakat memberikan tanggapan yang beragam, mulai dari antusiasme hingga kritik terhadap konten yang disajikan [3]. Fenomena ini khususnya menarik untuk diamati pada trailer film-film yang dinantikan seperti *Deadpool vs Wolverine*, yang memiliki basis penggemar yang besar dan aktif di media sosial.

Untuk menganalisis sentimen dari komentar-komentar tersebut, penelitian ini menggunakan empat model *machine learning* yang berbeda: Naïve Bayes (NB), KNN, Logistic Regression (LR), dan Decision Tree (DT). *Machine learning*, yang merupakan subdisiplin dari *artificial intelligence* (kecerdasan buatan), memungkinkan komputer untuk memahami pola data melalui pemrosesan data historis berskala besar dengan memanfaatkan bahasa pemrograman [4]. Analisis sentimen sendiri, sebagai bagian dari disiplin text mining, berperan dalam mengidentifikasi dan menginterpretasikan berbagai nuansa opini, perasaan, serta sikap yang terkandung dalam sebuah teks atau kumpulan teks [5]. Sebelum diproses oleh model *machine learning*, data berbentuk teks perlu dikonversi ke format numerik menggunakan teknik seperti TF-IDF, yang menghitung bobot setiap kata berdasarkan kombinasi nilai *term frequency* dan *inverse document frequency* [5].

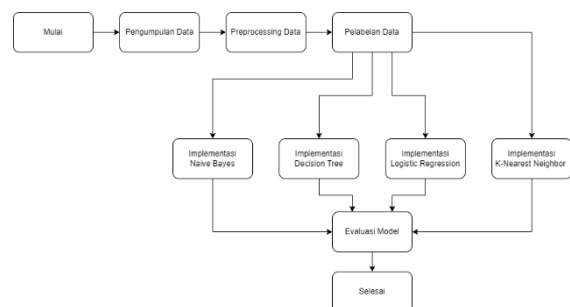
Dalam penelitian ini juga digunakan *Valence Aware Dictionary and sEntiment Reasoner* (VADER), sebuah alat analisis sentimen berbasis leksikon dan aturan yang dirancang khusus untuk mengenali sentimen yang diungkapkan dalam media sosial [6]. Kombinasi VADER dengan berbagai model *machine learning* seperti Naïve Bayes yang menunjukkan keunggulan dalam mengklasifikasikan dokumen [7], Logistic Regression yang efektif untuk menentukan sentimen dari teks [8], K-Nearest Neighbor yang mampu melakukan klasifikasi berbasis perhitungan matematis [9], dan Decision Tree yang mengadopsi struktur pohon untuk pengambilan keputusan [10], diharapkan dapat menghasilkan analisis sentimen yang lebih akurat dan komprehensif.

Penelitian ini bertujuan untuk mengembangkan dan membandingkan efektivitas keempat model *machine learning* tersebut dalam menganalisis sentimen komentar YouTube pada trailer film *Deadpool vs Wolverine*. Pengumpulan data dilakukan melalui Google API Key untuk mengekstrak komentar-komentar dari YouTube secara efisien dan terstruktur. Diharapkan penelitian ini mampu memberikan kontribusi berupa *insight* yang bermanfaat tentang bagaimana publik merespons dan mempersepsikan

trailer *Deadpool vs Wolverine*, serta memberikan pemahaman yang lebih dalam tentang perbandingan efektivitas berbagai model *machine learning* dalam analisis sentimen. Melalui perbandingan akurasi dari keempat model tersebut, penelitian ini juga berkontribusi pada pengembangan metodologi analisis sentimen yang lebih efektif untuk konten hiburan digital di masa depan.

2. Metode Penelitian

Metode penelitian adalah serangkaian langkah atau proses yang disusun secara sistematis dan logis untuk melaksanakan penelitian demi mencapai tujuan yang telah ditetapkan [5]. Proses-proses yang diterapkan dalam penelitian ini disajikan pada Gambar 1.



Gambar 1 Metode Penelitian

2.1. Pengumpulan Data

Penelitian ini menggunakan metode pengumpulan data berbasis web scraping untuk memperoleh komentar dari video YouTube berjudul "*Deadpool & Wolverine | Official Trailer | In Theaters July 26*". Data berupa komentar pengguna diambil secara langsung melalui YouTube Data API v3 dengan menggunakan *script Python*. Data yang berhasil diambil dari video ditampilkan dalam Tabel 2 untuk memberikan gambaran isi komentar yang dianalisis:

Tabel 1. Hasil pengumpulan data

Komentar
I wanna see more of Cassandra Nova. Maybe a 1v1 fight with Scarlet Witch or Jean Grey
This movie was horrible. Why are the marvels' movies getting worse and worse?
Who here is still watching this trailer after seeing the movie a dozen times? 🤔
Movie Name
Never been excited for the marvel's movies logo since guardians three witch was not that long ago

2.2. Pengumpulan Data

Pada tahap *preprocessing*, berbagai langkah dilakukan untuk mempersiapkan data mentah agar siap digunakan dalam proses analisis [1]. Proses *preprocessing* dilakukan dengan tahapan sebagai berikut:

a. Penanganan Data Kosong

Data komentar yang kosong atau memiliki nilai *NaN* diubah menjadi *string* kosong ("") agar dapat diproses dengan lancar tanpa menyebabkan kesalahan.

b. Konversi ke Huruf Kecil

Semua teks komentar dikonversi menjadi huruf kecil (*lowercase*) untuk menghilangkan perbedaan akibat kapitalisasi yang tidak relevan dalam analisis.

c. Penghapusan Tanda Baca dan Angka

Teks dibersihkan dari tanda baca dan angka menggunakan pola ekspresi reguler (*regular expressions*). Hanya huruf dan spasi yang dipertahankan agar analisis terfokus pada kata-kata utama.

d. Penghapusan *Stopwords*

Kata-kata yang bersifat umum dan tidak memiliki makna penting dalam analisis, seperti "*the*," "*is*," "*and*," dihapus dari teks. Proses ini menggunakan daftar *stopwords* yang telah didefinisikan oleh library NLTK (*Natural Language Toolkit*).

e. Hasil Akhir Teks Bersih

Setelah tahap *preprocessing* selesai, setiap komentar akan menghasilkan teks yang lebih bersih dan hanya mengandung kata-kata penting yang relevan untuk analisis sentimen atau analisis lainnya.

Data yang telah melalui tahap *preprocessing* disajikan dalam Tabel 2, yang memuat hasil transformasi berupa teks komentar yang telah dibersihkan dan siap untuk dianalisis lebih lanjut.

Tabel 2. Hasil *preprocessing* data

Komentar
wanna see cassandra nova maybe v fight scarlet witch jean grey
movie horrible marvels movies getting worse worse
still watching trailer seeing movie dozen times
movie name
never excited marvels movies logo since guardians three witch long ago

2.3. Pelabelan Data

Setelah tahap prapemrosesan selesai, data komentar diklasifikasikan berdasarkan sentimen menggunakan pendekatan berbasis *Sentiment Intensity Analyzer* (SIA) dari library NLTK. Proses pelabelan dilakukan dengan menghitung skor sentimen dari setiap komentar menggunakan metode *polarity scoring*. Skor ini terdiri dari tiga kategori utama, yaitu:

a. **Positive:**

Jika nilai *compound* lebih besar dari 0,05.

b. **Negative:**

Jika nilai *compound* lebih kecil dari -0,05.

c. **Neutral:**

Jika nilai *compound* berada di antara -0,05 dan 0,05.

Hasil dari proses pelabelan ini memberikan informasi tambahan yang memetakan setiap komentar ke dalam salah satu kategori sentimen tersebut. Data yang telah dilabeli akan digunakan untuk analisis lebih lanjut, dan hasilnya disajikan dalam Tabel 3.

Tabel 3. Hasil pelabelan data

Komentar	Label
wanna see cassandra nova maybe v fight scarlet witch jean grey	Negative
movie horrible marvels movies getting worse worse	Negative
still watching trailer seeing movie dozen times	Neutral
movie name	Neutral
never excited marvels movies logo since guardians three witch long ago	Negative

2.4. Implementasi Model Machine Learning

Setelah tahap prapemrosesan selesai, data Dalam tahap implementasi *model machine learning*, penelitian ini memanfaatkan empat algoritma berbeda untuk melakukan klasifikasi sentimen berdasarkan data komentar yang telah melalui proses pembersihan. Langkah awal melibatkan ekstraksi fitur menggunakan metode *TF-IDF Vectorization* (*Term Frequency-Inverse Document Frequency*). Setelah fitur diekstrak, data dibagi menjadi dua set, yaitu data pelatihan (80%) dan data pengujian (20%), dengan menggunakan metode *train_test_split* serta menetapkan nilai *random_state* sebesar 42 guna menjamin hasil yang dapat direproduksi.

Empat algoritma yang diterapkan pada data tersebut adalah:

1. *Naïve Bayes*, menggunakan distribusi Multinomial untuk klasifikasi berbasis probabilitas,
2. *K-Nearest Neighbors (KNN)*, yang menentukan kategori berdasarkan kedekatan jarak antar data,
3. *Logistic Regression*, untuk memodelkan hubungan linier antara fitur dan label, dan
4. *Decision Tree*, yang menggunakan pendekatan berbasis pohon keputusan untuk membagi data ke dalam kategori.

Setiap model dilatih menggunakan data pelatihan dengan parameter default yang sesuai dengan implementasi pustaka *scikit-learn*. Proses ini bertujuan untuk membangun model yang mampu mengklasifikasikan sentimen secara akurat berdasarkan fitur yang diekstraksi dari data komentar.

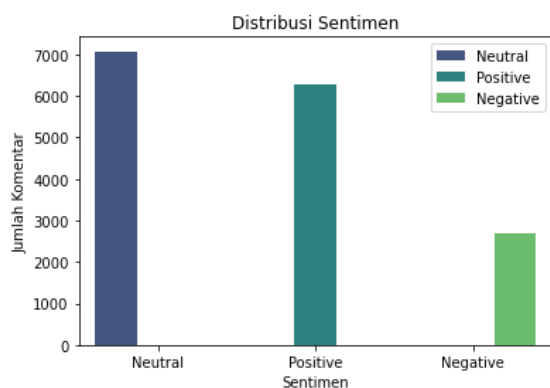
2.5. Evaluasi Model

Setelah Untuk mengevaluasi performa model, penelitian ini menggunakan metrik evaluasi, yaitu akurasi, *recall*, *precision*, dan *F1-score*. Metrik evaluasi dipilih untuk memberikan gambaran yang komprehensif mengenai kemampuan setiap model dalam melakukan klasifikasi sentimen. Selain itu, matriks kebingungan (*confusion matrix*) juga digunakan untuk menganalisis lebih lanjut hasil prediksi model, dengan memberikan visualisasi yang menunjukkan distribusi prediksi yang benar maupun salah pada setiap kelas.

Perbandingan hasil evaluasi antara model dilakukan dengan menyusun tabel kinerja yang mencakup nilai dari setiap metrik untuk masing-masing model. Hasil ini kemudian divisualisasikan dalam bentuk grafik batang untuk mempermudah analisis perbandingan performa. Visualisasi ini memberikan pandangan yang lebih jelas mengenai kelebihan dan kekurangan setiap algoritma dalam menyelesaikan tugas klasifikasi sentimen.

3. Hasil dan Pembahasan

Rangkaian Setelah dilakukan pengumpulan dan pengolahan data, distribusi sentimen pada dataset menunjukkan bahwa sebagian besar data memiliki sentimen *Neutral* dengan jumlah 7.072 data, diikuti oleh sentimen *Positive* sebanyak 6.289 data, dan sentimen *Negative* yang relatif lebih sedikit, yaitu 2.690 data sebagaimana divisualisasikan pada Gambar 2.



Gambar 2. Distribusi Sentimen

Distribusi ini menunjukkan adanya ketidakseimbangan jumlah data antar kategori sentimen, dengan dominasi kategori *Neutral*. Hal ini perlu diperhatikan karena ketidakseimbangan data dapat memengaruhi performa model dalam memprediksi kelas dengan jumlah data yang lebih kecil.

Dalam implementasi model machine learning untuk analisis sentimen, empat model yang berbeda telah digunakan dan menghasilkan performa yang

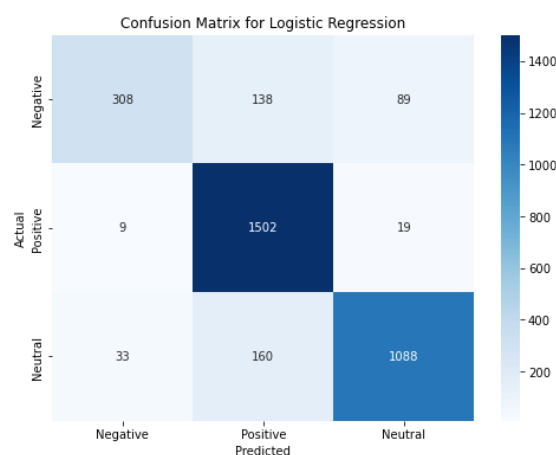
bervariasi. Model *Logistic Regression* menunjukkan performa terbaik dengan akurasi 86.61%, diikuti oleh *Decision Tree* dengan akurasi 85.89%, *Naive Bayes* dengan akurasi 79.08%, dan *KNN* dengan akurasi 60.07%. Performa masing-masing model dapat dilihat lebih detail pada Tabel 4.

Model	Akurasi	Precision	Recall	F1-Score
<i>Logistic Regression</i>	86.61%	87.05%	86.61%	86.01%
<i>Decision Tree</i>	85.89%	85.53%	85.89%	85.59%
<i>Naive Bayes</i>	79.08%	82.09%	79.08%	77.22%
<i>KNN</i>	60.07%	73.11%	60.07%	55.08%

Tabel 4 menunjukkan bahwa *Logistic Regression* tidak hanya unggul dalam akurasi, tetapi juga memiliki nilai *precision* tertinggi (87.05%) dan *F1-Score* terbaik (86.01%). Hal ini mengindikasikan bahwa model ini memiliki keseimbangan yang baik antara *precision* dan *recall* dalam mengklasifikasikan sentimen.

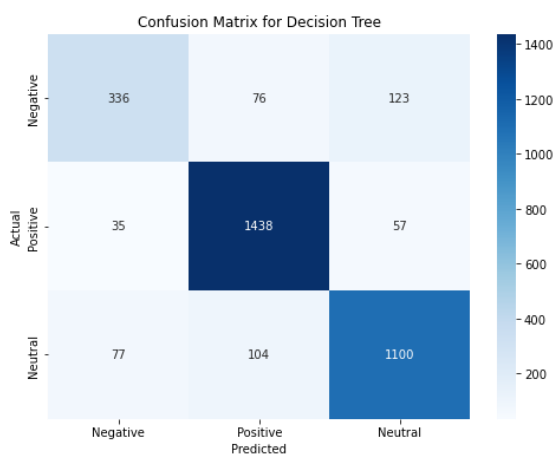
Berdasarkan *confusion matrix* pada Tabel 2 hingga Tabel 5, dapat diobservasi bahwa:

a. Model *Logistic Regression* menunjukkan kinerja yang stabil dalam mengklasifikasikan ketiga kategori sentimen. Model ini menghasilkan prediksi yang akurat, terutama pada kelas positif (1.502 prediksi benar) dan netral (1.088 prediksi benar). Selain itu, model ini cukup andal dalam mengidentifikasi sentimen negatif dengan 308 prediksi benar. *Confusion matrix* untuk model ini divisualisasikan pada Gambar 3.



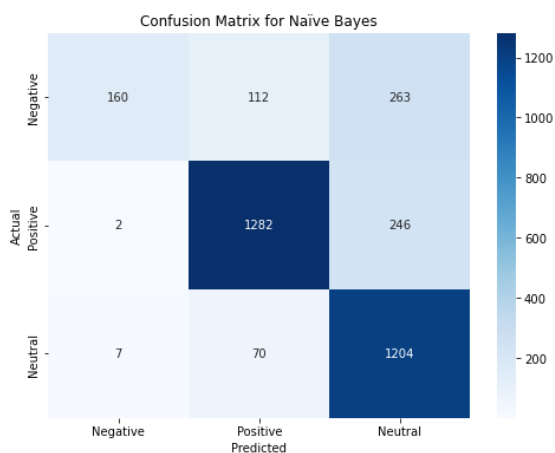
Gambar 3. *Confusion matrix* untuk Logistik Regression

b. *Decision Tree* memiliki performa yang hampir setara dengan *Logistic Regression*, dengan kemampuan prediksi yang baik untuk ketiga kelas. Model ini berhasil mengidentifikasi 336 kasus negatif, 1.438 kasus positif, dan 1.100 kasus netral dengan tepat. *Confusion matrix* untuk model ini divisualisasikan pada Gambar 4.



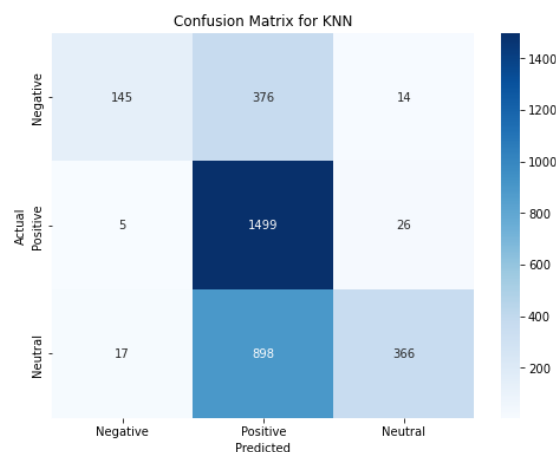
Gambar 4. *Confusion matrix* untuk *Decision Tree*

c. *Naive Bayes* menunjukkan performa yang kurang memadai, khususnya dalam mengklasifikasikan sentimen negatif. Meskipun demikian, model ini cukup baik dalam mengidentifikasi sentimen positif dengan 1.282 prediksi benar. *Confusion matrix* untuk model ini divisualisasikan pada Gambar 5.



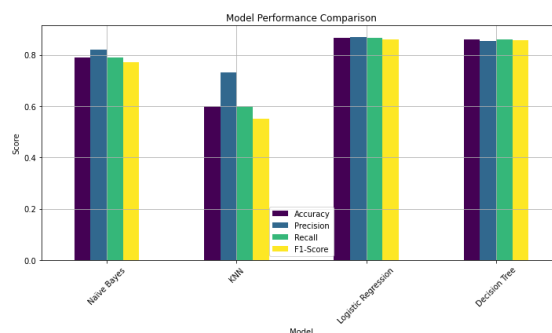
Gambar 5. *Confusion matrix* untuk *Naive Bayes*

d. *KNN* menunjukkan performa paling rendah di antara keempat model, dengan tingkat kesalahan klasifikasi yang tinggi terutama dalam membedakan antara sentimen positif dan netral. Hal ini terlihat dari tingginya jumlah misklasifikasi pada *confusion matrix*, di mana 898 data netral salah diklasifikasikan sebagai positif. *Confusion matrix* untuk model ini divisualisasikan pada Gambar 6.



Gambar 7. *Confusion matrix* untuk *KNN*

Ketidakseimbangan dalam distribusi data sentimen yang teridentifikasi di awal analisis tampaknya memiliki dampak signifikan terhadap performa model, terutama pada model *KNN*. Model *Logistic Regression* dan *Decision Tree* menunjukkan ketahanan yang lebih baik terhadap ketidakseimbangan ini, yang tercermin dari performa mereka yang lebih tinggi dalam mengklasifikasikan ketiga kelas sentimen.



Gambar 8. Grafik perbandingan performa

Berdasarkan hasil visualisasi pada Gambar 8, perbandingan performa keempat model menunjukkan bahwa *Logistic Regression* dan *Decision Tree* memiliki keunggulan dalam semua metrik evaluasi dibandingkan model lainnya. Ini dapat dilihat dari nilai *accuracy*, *precision*, *recall*, dan *F1-score* yang lebih unggul secara konsisten.

Sebaliknya, model *KNN* memiliki performa paling rendah di hampir semua metrik, terutama pada *recall* dan *F1-score*. Sementara itu, model *Naive Bayes* menunjukkan kinerja yang lebih unggul

dibandingkan *KNN* namun tetap lebih rendah dibandingkan *Logistic Regression* dan *Decision Tree*.

Secara keseluruhan, *Logistic Regression* dapat dianggap sebagai model yang paling optimal untuk tugas klasifikasi ini, diikuti oleh *Decision Tree*. Namun, kekuatan dan kelemahan masing-masing model memberikan wawasan penting untuk pengembangan lebih lanjut, termasuk potensi peningkatan melalui optimasi parameter atau metode pembobotan yang lebih kompleks.

4. Kesimpulan

Berdasarkan hasil penelitian analisis sentimen terhadap komentar pada *trailer* *Deadpool* vs *Wolverine* menggunakan empat model *machine learning*, dapat ditarik beberapa kesimpulan penting:

1. Distribusi sentimen dalam dataset menunjukkan dominasi komentar netral (7.072 data) dan positif (6.289 data), sementara komentar negatif memiliki jumlah yang lebih sedikit (2.690 data). Hal ini mengindikasikan bahwa *trailer* film tersebut mendapatkan respon yang cenderung netral hingga positif dari penonton YouTube.

2. Dari keempat model pembelajaran mesin yang diimplementasikan, *Logistic Regression* menunjukkan performa terbaik dengan tingkat akurasi 86.61%, precision 87.05%, recall 86.61%, dan *F1-score* 86.01%. Performa ini diikuti oleh *Decision Tree* yang juga menunjukkan hasil yang kompetitif dengan akurasi 85.89%.

3. Model *Naive Bayes* menempati posisi ketiga dengan akurasi 79.08%, sementara *KNN* menunjukkan performa yang relatif lebih rendah dengan akurasi 60.07%. Hasil ini mengindikasikan bahwa pendekatan probabilistik dan berbasis pohon keputusan lebih efektif dalam menganalisis sentimen untuk kasus ini dibandingkan dengan pendekatan berbasis jarak seperti *KNN*.

4. Ketidakseimbangan distribusi data sentimen memberikan dampak yang berbeda pada masing-masing model. *Logistic Regression* dan *Decision Tree* menunjukkan ketahanan yang lebih baik terhadap ketidakseimbangan ini, sementara *KNN* paling terpengaruh, terutama dalam membedakan antara sentimen positif dan netral.

Penelitian ini memberikan wawasan berharga tentang efektivitas berbagai model pembelajaran mesin dalam analisis sentimen untuk konten hiburan di media sosial. Untuk pengembangan penelitian selanjutnya, disarankan untuk:

1. Menerapkan teknik penyeimbangan data untuk mengatasi ketidakseimbangan distribusi sentimen.

2. Mengeksplorasi penggunaan model pembelajaran mesin yang lebih canggih seperti deep learning atau ensemble methods.

3. Memperluas cakupan analisis dengan mempertimbangkan aspek temporal dan kontekstual dari komentar.

4. Mengoptimalkan parameter model untuk meningkatkan performa klasifikasi, terutama untuk model dengan performa yang lebih rendah.

Daftar Rujukan

- [1] S. Adelia, E. Milanda, J. Santari, D. T. Kesuma, E. Silvia, and F. Kurniawan, "Analisis Sentimen Belajar Programming Pada Media Sosial Youtube Menggunakan Algoritma Klasifikasi Naive Bayes," 2023. [Online]. Available: <https://journal-computing.org/index.php/journal-ita/index>
- [2] B. Gunawan, H. Sasty, P. #2, E. Esiyudha, and P. #3, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," vol. 4, no. 2, pp. 17–29, 2018, [Online]. Available: www.femaledaily.com
- [3] G. Nusantara Bakry, "ANALISIS SENTIMEN DAN JARINGAN AKTOR TERHADAP KOMERSIALISASI TRAGEDI PADA VIDEO YOUTUBE TRAILER FILM VINA : SEBELUM TUJUH HARI SENTIMENT AND NETWORK ANALYSIS ACTOR OF COMERSIALIZATION OF TRAGEDY ON TRAILERS YOUTUBE FILM VINA : BEFORE SEVEN DAYS," *Jurnal Komunikasi dan Media*, vol. 05, no. 01, pp. 1–13, 2024.
- [4] A. Nursalim and R. Novita, "SENTIMENT ANALYSIS OF COMMENTS ON GOOGLE PLAY STORE, TWITTER AND YOUTUBE TO THE MYPERTAMINA APPLICATION WITH SUPPORT VECTOR MACHINE," *Jurnal Teknik Informatika (JUTIF)*, vol. 4, no. 6, pp. 1305–1312, 2023, doi: 10.52436/1.jutif.2023.4.6.1059.
- [5] I. Afdhal et al., "Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 1, 2022.
- [6] S.-Iong, Ao, *Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment*. International MultiConference of Engineers and Computer Scientists 2019 IMECS 2019, March 13–15, 2019, Hong Kong, 2019.
- [7] A. Karimah and G. Dwilestari, "ANALISIS SENTIMEN KOMENTAR VIDEO MOBIL LISTRIK DI PLATFORM YOUTUBE DENGAN METODE NAIVE BAYES," 2024. [Online]. Available: <https://www.kaggle.com/datasets/billycemerson/anali>
- [8] S. A. H. Bahtiar, C. K. Dewa, and A. Luthfi, "Comparison of Naïve Bayes and Logistic Regression in Sentiment Analysis on Marketplace Reviews Using Rating-Based Labeling," *Journal of Information Systems and Informatics*, vol. 5, no. 3, pp. 915–927, Aug. 2023, doi: 10.51519/journalisi.v5i3.539.
- [9] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal*

KomtekInfo, pp. 1–7,
10.35134/komtekinfo.v10i1.330.

Jan. 2023, doi:

[10] Y. A. Singgalen, “Analisis Sentimen Top 10 Traveler Ranked Hotel di Kota Makassar Menggunakan Algoritma Decision Tree dan Support Vector Machine,” *Media Online*, vol. 4, no. 1, pp. 323–332, 2023, doi: 10.30865/klik.v4i1.1153.