



PERBANDINGAN PERFORMA ALGORITMA *NAIVE BAYES* DAN *SVM* UNTUK ANALISIS SENTIMEN KOMENTAR *YOUTUBE* TERHADAP INDUSTRI *ESPORTS* DI INDONESIA

Tito Dian Permana

Universitas Muhammadiyah Bangka Belitung

Yudistira Bagus Pratama

Universitas Muhammadiyah Bangka Belitung

Zikri Wahyuzi

Universitas Muhammadiyah Bangka Belitung

Eka Altiarika

Universitas Muhammadiyah Bangka Belitung

Arvi Pramudyantoro

Universitas Muhammadiyah Bangka Belitung

Alamat: Pangkalan Baru

Korespondensi penulis: titodian768@gmail.com

Abstract The esports industry in Indonesia is rapidly growing and gaining significant attention on social media, particularly YouTube, where comments reflect public perceptions. This study compares the performance of Naive Bayes and Support Vector Machine (SVM) in classifying sentiments from YouTube comments and explores key themes using Latent Dirichlet Allocation (LDA). Data were collected via the YouTube Data API v3, labeled with TextBlob and manually verified into positive, negative, and neutral categories. After preprocessing and TF-IDF representation, class imbalance was handled with SMOTE, and models were trained and evaluated using accuracy, precision, recall, F1-score, and confusion matrix. Results indicate that Naive Bayes achieved 73.85% accuracy with an F1-score of 0.71, while SVM slightly outperformed with 73.97% accuracy and the same F1-score. SVM showed better consistency in classifying negative and neutral comments, whereas Naive Bayes was more effective for positive ones. LDA revealed dominant discussion topics such as appreciation, enthusiasm, community interaction, criticism, and support for esports development. These findings highlight SVM's superior overall performance and the value of LDA in uncovering public discourse, providing both academic contribution and practical insights for the esports industry in understanding public sentiment.

Keywords: Esports, Latent Dirichlet Allocation, Naive Bayes, Sentiment Analysis, Support Vector Machine

Abstrak Industri esports di Indonesia berkembang pesat dan mendapat perhatian besar melalui media sosial, khususnya YouTube, di mana komentar mencerminkan opini publik. Penelitian ini bertujuan membandingkan performa algoritma Naive Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen komentar YouTube, serta mengidentifikasi tema utama menggunakan Latent Dirichlet Allocation (LDA). Data diperoleh melalui YouTube Data API v3, dilabeli dengan TextBlob, dan diverifikasi manual menjadi kategori positif, negatif, dan netral. Tahapan preprocessing mencakup pembersihan teks, case folding, tokenisasi, stopword removal, dan stemming. Representasi teks menggunakan TF-IDF, sedangkan ketidakseimbangan kelas ditangani dengan SMOTE. Model dilatih dengan Naive Bayes dan SVM, kemudian dievaluasi menggunakan akurasi, presisi, recall, F1-score, dan confusion matrix. Hasil menunjukkan Naive Bayes meraih akurasi 73,85% dengan F1-score 0,71, sementara SVM sedikit unggul dengan akurasi 73,97% dan F1-score 0,71. SVM lebih konsisten dalam mengklasifikasikan komentar negatif dan netral, sedangkan Naive Bayes lebih baik pada komentar positif. Analisis LDA mengungkap topik dominan seperti apresiasi, antusiasme, interaksi komunitas, kritik, dan dukungan terhadap esports. Temuan ini menegaskan keunggulan SVM secara keseluruhan sekaligus memberikan wawasan praktis bagi industri esports dalam memahami persepsi publik.

Kata kunci: Esports, Latent Dirichlet Allocation, Naive Bayes, Analisis Sentimen, Support Vector Machine

LATAR BELAKANG

Industri *esports* kini menjadi salah satu sektor yang paling berkembang dan populer di dunia, termasuk di Indonesia. *Esports* adalah kompetisi video game profesional yang mendukung pemain yang kompetitif dalam permainan tertentu[1]. Karena pendapatan dan popularitasnya yang tinggi, *esports* telah menjadi fenomena global. Meningkatnya kompetisi *esports* dan prestasi tim-tim Indonesia di kompetisi internasional menjadi alasan di balik hal ini.

Meskipun demikian, industri *esports* di Indonesia masih menghadapi sejumlah tantangan. Salah satu masalah utamanya adalah stigma negatif yang sering melekat, terutama di kalangan generasi yang lebih tua, yang menganggapnya sebagai aktivitas yang kurang produktif. Selain itu, minimnya regulasi yang jelas dan belum meratanya infrastruktur teknologi di berbagai daerah turut menjadi penghambat bagi pertumbuhan yang lebih inklusif dan berkelanjutan[2].

Di era internet, media sosial seperti *YouTube* menjadi tempat utama bagi orang-orang untuk berbagi pendapat mereka tentang *esports*[3]. Komentar yang diunggah oleh pengguna menunjukkan persepsi publik terhadap industri ini, baik positif maupun negatif. Oleh karena itu, analisis sentimen terhadap komentar *YouTube* dapat menjadi alat yang berguna untuk mengetahui persepsi masyarakat tentang *esports* di Indonesia.

Analisis sentimen telah banyak digunakan dalam berbagai penelitian untuk memahami opini publik terhadap suatu isu, seperti evaluasi produk, ulasan film, serta respons masyarakat terhadap kebijakan tertentu. Untuk melakukan analisis sentimen, diperlukan sebuah algoritma pembelajaran mesin yang dapat mengklasifikasikan teks berdasarkan polaritasnya. Di mana algoritma *Naive Bayes* dan *SVM* banyak digunakan dalam analisis sentimen untuk mengklasifikasikan teks berdasarkan polaritasnya[4]. Penelitian ini difokuskan pada analisis sentimen terhadap komentar di *YouTube* tahun 2024 terkait industri *esports* di Indonesia menggunakan algoritma *Naive Bayes* dan *SVM*. Penelitian ini memiliki tujuan untuk memberikan wawasan mengenai persepsi publik terhadap industri *esports* di Indonesia, sehingga dapat mendukung pengembangan dan penerimaannya di masyarakat. Selain itu, penelitian ini juga menerapkan metode *Latent Dirichlet Allocation (LDA)* untuk mengidentifikasi dan mengkategorikan topik-topik tersembunyi dalam komentar *YouTube*, sehingga dapat memberikan wawasan lebih mendalam mengenai isu-isu yang sering dibahas dalam komunitas *esports*. Melalui penelitian ini, diharapkan bahwa hasil yang didapatkan dapat menyajikan pemahaman yang lebih mendalam mengenai pendapat publik terhadap industri *esports* serta menghasilkan sistem analisis sentimen yang dapat diterapkan untuk memahami tren dan respons masyarakat terhadap perkembangan *esports* di Indonesia.

KAJIAN TEORITIS

Analisis Sentimen

Analisis sentimen adalah proses menemukan, mengkategorikan, dan mengevaluasi emosi yang terkandung dalam sebuah teks. Emosi ini dapat mencerminkan perspektif, perasaan, atau sikap penulis terhadap subjek tertentu. Studi ini menganalisis sentimen komentar pada video *YouTube* yang berhubungan dengan industri *esport* di Indonesia. Tiga kategori utama sentimen dapat ditemukan dalam komentar: sentimen positif, negatif, dan netral[5].

Esports

Esports adalah suatu tempat kegiatan olahraga untuk membangun dan melatih mental serta kemampuan fisik dengan menggunakan teknologi dan komunikasi[6]. seiring meningkatnya pengguna internet dan popularitas game online, *esports* di Indonesia berkembang pesat. Di

Indonesia, esports telah menjadi bagian penting dari industri hiburan digital, seperti yang ditunjukkan oleh turnamen esports nasional seperti Piala Presiden Esports, yang pertama kali diadakan pada tahun 2019[7].

Machine Learning

Pembelajaran mesin, sebuah bidang dalam ilmu komputer dan kecerdasan buatan, menggunakan data dan algoritma untuk meniru proses pembelajaran manusia. Proses machine learning mempelajari sesuatu disebut pembelajaran. Proses pembelajaran berkaitan dengan input seperti data latih, data uji, dan algoritma untuk berperan dalam pengambilan keputusan dan prediksi. Pembelajaran secara iteratif akan meningkatkan akurasi secara bertahap, sehingga menghasilkan prediksi atau pengambilan keputusan yang lebih akurat[8]. Dua algoritma yang paling sering digunakan adalah Naive Bayes dan Support Vector Machine (SVM). Naive Bayes merupakan algoritma berbasis probabilistik yang sederhana namun efektif dalam klasifikasi teks. Kelebihannya adalah efisiensi komputasi dan kemudahan implementasi, meskipun keterbatasan muncul ketika data tidak memenuhi asumsi independensi antar fitur. Sebaliknya, SVM merupakan algoritma yang memisahkan data ke dalam kelas berbeda dengan menentukan hyperplane terbaik. Algoritma ini dikenal lebih konsisten dalam menangani data berdimensi tinggi dan cenderung menghasilkan akurasi lebih baik, meskipun membutuhkan tuning parameter yang cermat.

Youtube

YouTube adalah media sosial yang menyediakan video online yang dapat diakses dari lokasi mana pun, apabila terhubung dengan internet[9]. Informasi yang diperoleh dari YouTube diambil melalui YouTube API dengan mencari kata kunci yang relevan, sehingga menghasilkan daftar video yang sesuai.

Google Colab

Google Colab adalah salah satu alat ilmu data dan AI yang paling populer karena mudah digunakan. kemampuan untuk bekerja sama dalam tim dan akses ke sumber daya komputer yang kuat. Google Colab telah berkembang menjadi alat yang disukai oleh pengembang dan praktisi ilmu data dalam beberapa tahun terakhir. Karena kemudahan akses, sumber daya gratis, Berkat kemampuan kerja timnya, Google Colab telah mempermudah pembuatan dan pengujian ide dalam ilmu data dan pembelajaran mesin[10].

Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) adalah salah satu metode topik modeling yang digunakan untuk menemukan topik-topik tersembunyi dalam kumpulan dokumen teks secara otomatis. LDA bekerja dengan mengasumsikan bahwa setiap dokumen merupakan campuran dari sejumlah topik, dan setiap topik terdiri dari sekumpulan kata yang memiliki probabilitas kemunculan tertentu. Dalam prosesnya, LDA menggunakan pendekatan probabilistik untuk menentukan distribusi kata dalam topik serta distribusi topik dalam dokumen[11]

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif Secara umum, tahapan penelitian adalah kumpulan tindakan yang diambil peneliti untuk memastikan bahwa proses penelitian dilakukan secara sistematis, terstruktur, dan metodis untuk meraih tujuan yang telah ditentukan sebelumnya[12]. Tujuan dari penelitian ini adalah untuk mengevaluasi kinerja algoritma *Naive Bayes* dan *SVM* dalam mengidentifikasi serta mengklasifikasikan sentimen pada komentar *YouTube* yang membahas industri *esports* di Indonesia. Pra-pemrosesan data, pelatihan model,

dan evaluasi kinerja model merupakan langkah awal dalam proses analisis data[13]. Berikut ini penjelasan mengenai tahapan analisis data

a) Pra-pemrosesan data

Sebelum menggunakan data untuk model pelatihan, tahap pemrosesan awal sangat penting untuk menjamin kualitas data[14]. Pra-pemrosesan data dilakukan untuk memastikan kualitas data sebelum pelatihan model. Tahapan meliputi *cleaning* (menghapus simbol, URL, emoji, dan teks tidak relevan), *normalization* (mengubah kata informal menjadi bentuk baku), *stopword removal* (menghapus kata umum), *stemming* (mengubah kata menjadi bentuk dasar), *tokenization* (memecah teks menjadi kata), serta *labelling* sentimen positif, negatif, atau netral menggunakan *TextBlob*.

b) Pelatihan Model

Pelatihan model dilakukan dengan Naive Bayes dan SVM ber-kernel linear. Naive Bayes mengklasifikasi sentimen dengan pendekatan probabilistik, sedangkan SVM memisahkan kelas menggunakan *hyperplane* optimal pada data berdimensi tinggi. Kedua model dilatih untuk mengelompokkan komentar YouTube menjadi positif, negatif, atau netral, lalu dievaluasi dengan validasi silang menggunakan metrik akurasi, presisi, *recall*, dan F1-score untuk mengukur kinerjanya.

c) Evaluasi Model

dilakukan menggunakan Naive Bayes dan SVM dengan kernel linear. Naive Bayes memanfaatkan pendekatan probabilistik untuk menghitung peluang kata dalam tiap kategori sentimen, sedangkan SVM menentukan *hyperplane* optimal untuk memisahkan kelas pada data berdimensi tinggi. Kedua model dilatih untuk mengklasifikasikan komentar YouTube menjadi positif, negatif, atau netral, serta dievaluasi menggunakan validasi silang dan metrik akurasi, presisi, *recall*, dan F1-score.

HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pada penelitian ini, data yang digunakan diperoleh dari platform *YouTube*, yaitu berupa komentar pada video-video yang membahas topik *esports* di Indonesia. Pengumpulan data dilakukan dalam kurun waktu mulai dari Januari 2024 hingga Desember 2024, jumlah komentar yang berhasil dikumpulkan yakni sebanyak 15.100 komentar dari total 501 video yang dianalisis.

B. Pelabelan Data

Pelabelan data pada penelitian ini dilakukan dengan pendekatan semi-otomatis, yaitu menggabungkan proses pelabelan otomatis menggunakan library *TextBlob* dengan verifikasi manual untuk meningkatkan akurasi hasil.

	text	sentimen
0	Semoga Di Lirik Pemimpin negara	Netral
1	Emang msh ada lg yg bakal bertanding?	Netral
2	Buset masuk berita juga cukk	Netral
3	kereen	Netral
4	Ayo pubg nyusul emas juga	Netral
...	...	...
15095	Keren timnas indo. Moga rrq sama tliid ketu...	Netral
15096	CONGRATS INDONESIA ♥	Netral
15097	3 kali ketemu PH girls dikasih telor sama ID G...	Netral
15098	Montoon harus nya buat M SERIES versi ladies p...	Netral
15099	Pertama min	Netral

Gambar 1. Hasil Pelabelan Otomatis

	clean_text	sentiment
0	moga lirik pemimpin negara	Positif
1	tanding	Netral
2	buset masuk berita cukk	Netral
3	kereen	Positif
4	ayo pubg nyusul emas	Positif
...	...	...
12382	keren timnas indo moga rrq tliid tular bawa med...	Positif
12383	congrats indonesia	Positif
12384	ketemu ph girls kasih telor id girls	Positif
12385	montoon m series versi ladies indonesia juara	Positif
12386	min	Netral

Gambar 2. Hasil Pelabelan Manual

C. Preprocessing Data

Tahap *preprocessing* data merupakan proses penting dalam pengolahan teks, terutama untuk data komentar *YouTube* yang tidak terstruktur dan cenderung mengandung unsur non-baku seperti kata-kata gaul, simbol, serta berbagai elemen kebahasaan informal lainnya. Dalam penelitian ini, proses *preprocessing* dilakukan secara menyeluruh menggunakan berbagai pustaka pemrosesan bahasa alami seperti *nlp-id*, *nlTK*, *BeautifulSoup*, dan modul pemetaan kata tidak baku.

D. Pelatihan Model

Setelah proses pra-pemrosesan data selesai, tahap selanjutnya dalam penelitian ini adalah melakukan pelatihan model klasifikasi sentimen. Tujuan dari pelatihan ini adalah untuk membangun model yang mampu mengklasifikasikan komentar *YouTube* ke dalam tiga kategori sentimen, yaitu *positif*, *negatif*, dan *netral*.

E. Evaluasi Model

1. Evaluasi Model Naive Bayes

Evaluasi dilakukan untuk menilai kinerja algoritma *Naive Bayes* dalam mengklasifikasikan sentimen komentar *YouTube* terkait industri esports di Indonesia. Penilaian performa model menggunakan metrik akurasi, *precision*, *recall*, dan *f1-score* guna memberikan gambaran menyeluruh mengenai kemampuan algoritma dalam mengenali pola teks serta memprediksi kategori sentimen positif, negatif, dan netral.

```

Hasil Evaluasi Naive Bayes
Parameter Terbaik: {'clf_alpha': 1.0}
Akurasi Terbaik (Cross-Validation): 0.7610
Akurasi pada Data Uji: 0.7385
Laporan Klasifikasi Naive Bayes:

```

	precision	recall	f1-score	support
Negatif	0.78	0.83	0.81	1279
Netral	0.70	0.57	0.63	566
Positif	0.67	0.70	0.68	633
accuracy			0.74	2478
macro avg	0.72	0.70	0.71	2478
weighted avg	0.74	0.74	0.74	2478

Gambar 3. Hasil Evaluasi Naive Bayes

Hasil evaluasi algoritma *Naive Bayes* menunjukkan bahwa parameter optimal yang diperoleh adalah $\alpha = 1.0$, dengan akurasi terbaik pada cross-validation sebesar 0.7610. 40 Pada saat pengujian menggunakan data uji, model menghasilkan akurasi sebesar 0.7385. Berdasarkan laporan klasifikasi, kinerja terbaik terdapat pada kelas negatif dengan *precision* 0.78, *recall* 0.83, dan *f1-score* 0.81. Pada kelas positif, model menunjukkan *precision* 0.67, *recall* 0.70, dan *f1 score* 0.68, sedangkan pada kelas netral nilai *precision* mencapai 0.70 namun *recall* lebih rendah, yaitu 0.57, dengan *f1-score* sebesar 0.63. Nilai macro average untuk *precision*, *recall*, dan *f1-score* masing-masing sebesar 0.72, 0.70, dan 0.71, sedangkan *weighted average* menunjukkan konsistensi pada angka 0.74.

2. Evaluasi Model Support Vector Machine (SVM)

Pada tahap ini dilakukan evaluasi terhadap algoritma *Support Vector Machine (SVM)* dalam mengklasifikasikan sentimen komentar *YouTube* mengenai industri esports di Indonesia. Proses evaluasi memanfaatkan metrik akurasi, *precision*, *recall*, dan *f1-score*

```

--- Hasil Evaluasi SVM ---
Parameter Terbaik: {'clf_C': 1, 'clf_gamma': 'scale',
'clf_kernel': 'rbf'}
Akurasi Terbaik (Cross-Validation): 0.7459
Akurasi pada Data Uji: 0.7397
Laporan Klasifikasi SVM:

```

	precision	recall	f1-score	support
Negatif	0.75	0.86	0.80	1279
Netral	0.69	0.60	0.64	566
Positif	0.76	0.61	0.68	633
accuracy			0.74	2478
macro avg	0.73	0.69	0.71	2478
weighted avg	0.74	0.74	0.73	2478

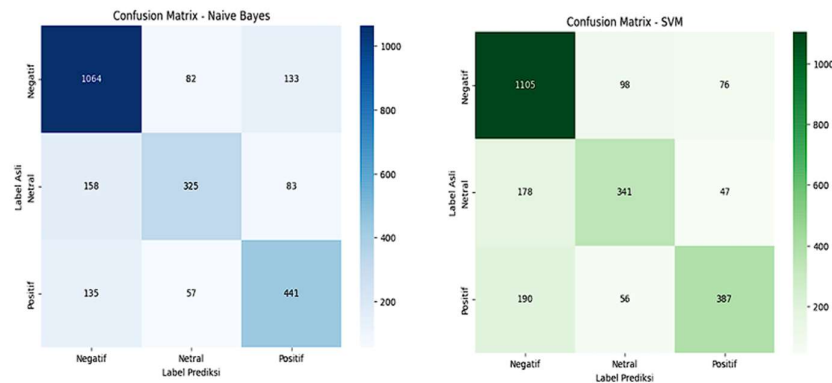
untuk menilai sejauh mana *SVM* mampu membedakan pola teks serta menghasilkan prediksi yang sesuai dengan kategori sentimen positif, negatif, maupun netral.

Gambar 4. Hasil Evaluasi *SVM*

Hasil evaluasi algoritma Support Vector Machine (*SVM*) menunjukkan bahwa parameter terbaik diperoleh pada kombinasi kernel RBF dengan nilai $C = 1$ dan $\gamma = \text{scale}$. Model ini menghasilkan akurasi cross-validation sebesar 0,7459, sedangkan akurasi pada data uji mencapai 0,7397. Berdasarkan laporan klasifikasi, kelas Negatif memperoleh precision sebesar 0,75, recall 0,86, dan f1-score 0,80, yang menandakan performa tinggi dalam mendeteksi komentar negatif. Pada kelas Netral, nilai precision mencapai 0,69 dan recall 0,60, dengan f1-score 0,64. Sementara untuk kelas 41 Positif, precision tercatat sebesar 0,76, recall 0,61, dan f1 score 0,68.

F. Confusion Matrix Naive Bayes dan SVM

Confusion matrix digunakan untuk memberikan gambaran yang lebih detail mengenai hasil prediksi model dengan membandingkan nilai aktual dan nilai prediksi. Melalui confusion matrix, dapat diketahui jumlah prediksi yang benar maupun salah pada setiap kelas sentimen, baik positif, negatif, maupun netral. Pada penelitian ini, confusion matrix ditampilkan untuk



Gambar 5. Hasil Confusion Matrix

kedua algoritma, yaitu Naive Bayes dan Support Vector Machine (*SVM*), sehingga dapat terlihat distribusi kesalahan klasifikasi yang terjadi pada masing masing model. Confusion matrix Naive Bayes memperlihatkan bahwa model mampu mengklasifikasikan komentar Negatif dengan cukup baik, yaitu sebanyak 1.064 prediksi benar, sementara untuk kelas Netral dan Positif masing-masing tercatat 325 dan 441 prediksi benar. Kesalahan klasifikasi masih cukup dominan, khususnya pada komentar Netral yang kerap terdeteksi sebagai Negatif. Sementara itu, confusion matrix *SVM* menunjukkan hasil yang sedikit berbeda, dengan kinerja lebih baik pada kelas Negatif (1.105 prediksi benar) serta peningkatan pada kelas Netral dengan 341 prediksi benar. Namun, untuk kelas Positif, jumlah prediksi benar lebih rendah dibanding Naive Bayes, yaitu sebanyak 387 komentar. Dengan demikian, dapat dilihat bahwa *SVM* lebih konsisten dalam mendeteksi komentar 42 Negatif dan Netral, sedangkan Naive Bayes lebih unggul pada komentar Positif.

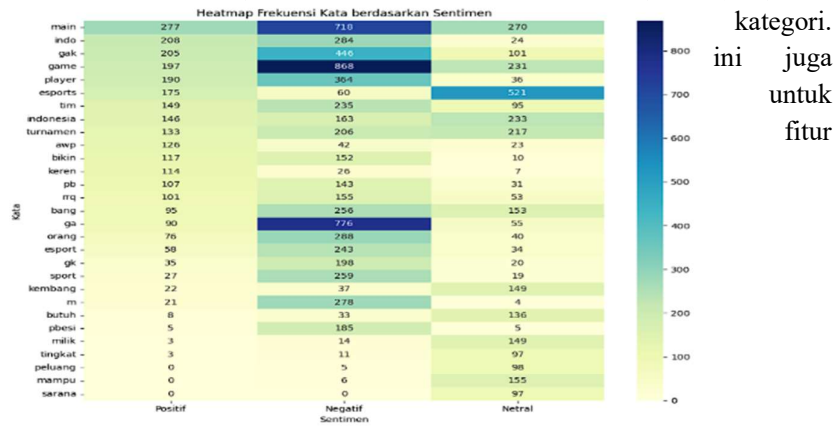
G. Visualisasi Hasil

Visualisasi hasil merupakan salah satu metode yang digunakan untuk menggambarkan data secara grafis guna memperoleh pemahaman yang lebih komprehensif terhadap pola dan karakteristik data sentimen yang dianalisis. Pada penelitian ini, visualisasi dilakukan menggunakan data komentar yang telah melalui proses praproses dan pelabelan secara manual. Hal ini dimaksudkan agar visualisasi yang dihasilkan mampu merepresentasikan kondisi aktual dari opini pengguna secara objektif.

1. Heatmap

PERBANDINGAN PERFORMA ALGORITMA NAIVE BAYES DAN SVM UNTUK ANALISIS SENTIMEN KOMENTAR YOUTUBE TERHADAP INDUSTRI ESPORTS DI INDONESIA

Heatmap ini memberikan gambaran visual mengenai kata-kata yang dominan dalam setiap kelas sentimen, serta memungkinkan identifikasi karakteristik linguistik yang khas untuk tiap kategori. Visualisasi ini juga bermanfaat untuk eksplorasi fitur dalam



pembangunan model klasifikasi berbasis teks.

Gambar 6. Visualisasi Heatmap

2. WordCloud

Visualisasi dilakukan dengan membagi area gambar menjadi tiga bagian (tiga subplot sejajar), di mana setiap bagian menampilkan *word cloud* untuk satu jenis sentimen. Masing-masing *word cloud* diberi warna berbeda, hijau untuk positif, biru untuk netral, dan merah untuk negatif



Gambar 7. WordCloud Setiap Sentimen

3. Ekplorasi LDA (Latent Dirichlet Allocation)

PERBANDINGAN PERFORMA ALGORITMA NAIVE BAYES DAN SVM UNTUK ANALISIS SENTIMEN KOMENTAR YOUTUBE TERHADAP INDUSTRI ESPORTS DI INDONESIA

Eksplorasi topik dengan metode LDA dilakukan untuk menemukan tema utama dalam komentar *YouTube* tentang esports Indonesia. Hasilnya mengungkap topik seperti performa pemain, kualitas turnamen, kebijakan PBESI, serta dukungan atau kritik terhadap tim ternama. Temuan ini memberi wawasan tentang persepsi publik dan dapat dimanfaatkan untuk strategi komunikasi, evaluasi kebijakan, dan peningkatan kualitas kompetisi *esports* nasional



Gambar 8. WordCloud Setiap Topik

KESIMPULAN DAN SARAN

Hasil evaluasi model menunjukkan bahwa kedua algoritma memiliki performa yang cukup seimbang dengan akurasi sekitar 74%. Algoritma Naive Bayes menghasilkan akurasi 0,7385 dengan performa terbaik pada kelas negatif (f1-score 0,81) serta cukup baik pada kelas positif (0,68), namun lemah pada kelas netral (0,63). Sementara itu, SVM memperoleh akurasi 0,7397 dengan keunggulan pada deteksi komentar negatif (f1-score 0,80, recall 0,86), meskipun kinerjanya lebih rendah pada kelas positif (0,68) dan netral (0,64). Dengan demikian, dapat disimpulkan bahwa SVM lebih unggul secara keseluruhan, terutama karena akurasinya sedikit lebih tinggi dan lebih baik dalam mengidentifikasi komentar negatif yang mendominasi dataset. Penelitian selanjutnya disarankan memperluas objek analisis ke sektor digital lain seperti aplikasi, forum, atau berita online, serta mempertimbangkan algoritma lebih kompleks seperti Random Forest, XGBoost, LSTM, atau BERT untuk meningkatkan akurasi. Selain itu, pengumpulan data lintas platform dan analisis berbasis waktu penting dilakukan guna menangkap dinamika sentimen publik terhadap esports secara lebih komprehensif.

DAFTAR REFERENSI

- [1] F. Febriansyah, V. Octaviani, and S. Sari, "Pengalaman Komunikasi Atlet E-Sport Mobile Legends Kota Bengkulu Dalam Upaya Memenangkan Permainan," *J. Multidisiplin Dehasen*, vol. 3, no. 3, pp. 269–276, 2024, doi: 10.37676/mude.v3i3.6556.
- [2] F. Kurniawan, "E-Sport dalam Fenomena Olahraga Kekinian," *Jorpres (Jurnal Olahraga Prestasi)*, vol. 15, no. 2, pp. 61–66, 2020, doi: 10.21831/jorpres.v15i2.29509.
- [3] J. P. Suharsono and D. Nurahman, "Pemanfaatan Youtube Sebagai Media Peningkatan Pelayanan Dan Informasi," *Ganaya J. Ilmu Sos. dan Hum.*, vol. 7, no. 1, pp. 298–304, 2024, doi: 10.37329/ganaya.v7i1.3157.
- [4] R. Kurniawan and S. Anwar, "Analisis Sentimen Penggemar Treasure Di Karnaval Mandiri Menggunakan Naïve Bayes," vol. 4, pp. 203–213, 2025.

- [5] Ardiyansyah, Ridhwan, Makmur, Ivan, Phai, Sabrina Phalosa, Studi, Program, Informatika, Teknik, Tarumanagara, Universitas, Barat, Jakarta, & Analyzer, Sentiment Intensity. (2024). Sentimen Komentar Youtube Dengan Sentiment Intensity Analyzer Dari NLTK. 31–36.
- [6] Dirgantara, Harya Bima, Lesmana Marselino, Tedi, & Ery Kurniawati, Yulia. (2023). Kajian Literatur Kurikulum E-sport dan Perkembangan Industri Game. *KALBISCIENTIA Jurnal Sains Dan Teknologi*, 10(1), 7–13. 58 <https://doi.org/10.53008/kalbiscientia.v10i1.1799>
- [7] Paramitha, S. T., Hasan, M. F., Iلسya, M. N. F., Anggraeni, L., & Ramadhan, G. (2021). Level of physical activity of Indonesian esport athletes in the Piala Presiden Esport 2019. *Jurnal SPORTIF: Jurnal Penelitian Pembelajaran*, 7(1), 71–83. https://doi.org/10.29407/js_unpk.v7i1.15642
- [8] Fernando, Daud, Huwaidi, Faris, Ananto, Muhammad Hafidz, & Pramadya, Sahrial. (2021). Comparison of Machine Learning Algorithms in the Role of Hepatitis Patient Disease Classification. 4(2), 159–170. Retrieved from <https://doi.org/10.17509/seict.v2i1.34253>
- [9] Firsttama, Risav Arrahman, Arifiyanti, Amalia Anjani, & Kartika, Dhian Satria Yudha. (2024). Analisis Sentimen Komentar Youtube Konferensi Tingkat Tinggi G20 Menggunakan Metode Naive Bayes. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 6(2), 282–285. <https://doi.org/10.47233/jteksis.v6i2.1263>
- [10] Amal, B., Damayanti, S., Khonsa, A. N., Zahra, M. H., Rahmadhani, V. A., Anggraeni, W., & Zendrato, K. D. P. (2025). Penilaian Mahasiswa PBSI Unsika Terhadap Keefektifan Penggunaan Google Colab Dalam Pembelajaran Coding. *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, 6(1), 66–76.
- [11] Sahoo, Satyaajeet, Maiti, J., & Tewari, V. K. (n.d.). Multivariate Gaussian Topic Modelling : A novel approach to discover topics with greater semantic coherence. 1–12.
- [12] A. Widiyanti and D. A. Megawaty, “Perbandingan Algoritma K-Nearest Neighbor dan Support Vector Machine Pada Pengenalan Pola Tulisan Tangan,” *J. Media Inform. Budidarma*, vol. 8, no. 3, p. 1451, 2024, doi: 10.30865/mib.v8i3.7757
- [13] L. Rao, “Sentiment Analysis of English Text with Multilevel Features,” *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/7605125.
- [14] Firdaus, Mutiara Puspita, & Trisnawarman, Dedi. (2025). Public Sentiment Analysis of the Public Housing Savings Program Using the IndoBERT Lite Model on YouTube Comments Analisis Sentimen Publik Terhadap Program Tabungan Perumahan Rakyat Menggunakan Model IndoBERT Lite pada Komentar YouTube. 5(January), 359–368.