

ANALISIS SENTIMEN PADA KOMENTAR YOUTUBE DALAM TURNAMEN MPL SEASON 13 DENGAN METODE ENSEMBLE MACHINE LEARNING

Oleh :

Zendhi Yuna Burnama

Mochamad Alfian Rosid

Program Studi Informatika

Universitas Muhammadiyah Sidoarjo

December, 2024



Latar Belakang

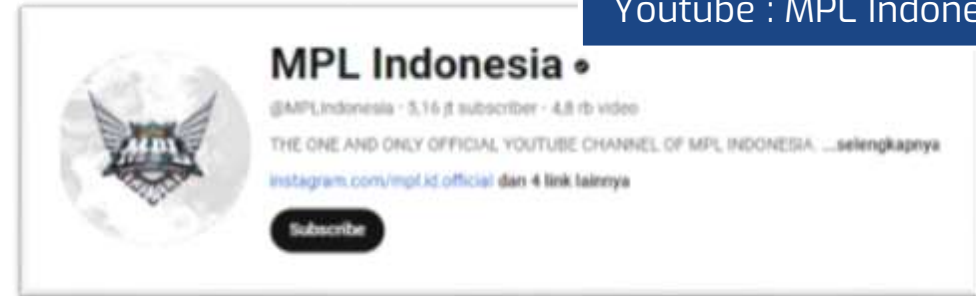
- ✓ Perkembangan teknologi telah menjadikan YouTube sebagai video library terbesar di internet, memungkinkan pengguna memberikan opini dan kritik melalui fitur komentar pada video dan siaran langsung.



Latar Belakang

- ✓ Video dan siaran langsung bertema game, seperti yang diunggah oleh channel YouTube **MPL Indonesia**, semakin populer karena menayangkan turnamen bergengsi **Mobile Legends**, yaitu **MPL Season 13**. Sejak diluncurkan oleh Moonton pada 2016, Mobile Legends telah menjadi salah satu game MOBA yang paling diminati di Indonesia, terutama di kalangan remaja.

Youtube : MPL Indonesia



dorangadget.com



Latar Belakang

- ✓ dukungan investor sangat penting bagi tim e-sports di MPL Season 13, dengan popularitas tim menjadi faktor utama yang menarik minat dan kepercayaan investor.
- ✓ Penelitian ini menganalisis sentimen komentar YouTube terkait MPL Season 13 menggunakan teknik **text mining** dan **berbagai metode klasifikasi**, termasuk **pendekatan Ensemble Machine Learning** untuk meningkatkan akurasi.
- ✓ **Tujuan penelitian ini adalah** mengukur popularitas tim e-sports dari sentimen komentar dan menganalisis pengaruhnya terhadap keputusan investasi dalam industri e-sports.

Rumusan Masalah

Rumusan masalah yang diangkat dalam penelitian ini adalah:
“Bagaimana analisis sentimen terhadap komentar YouTube dapat mempengaruhi keputusan investor dalam mendukung tim e-sports di MPL Season 13?”

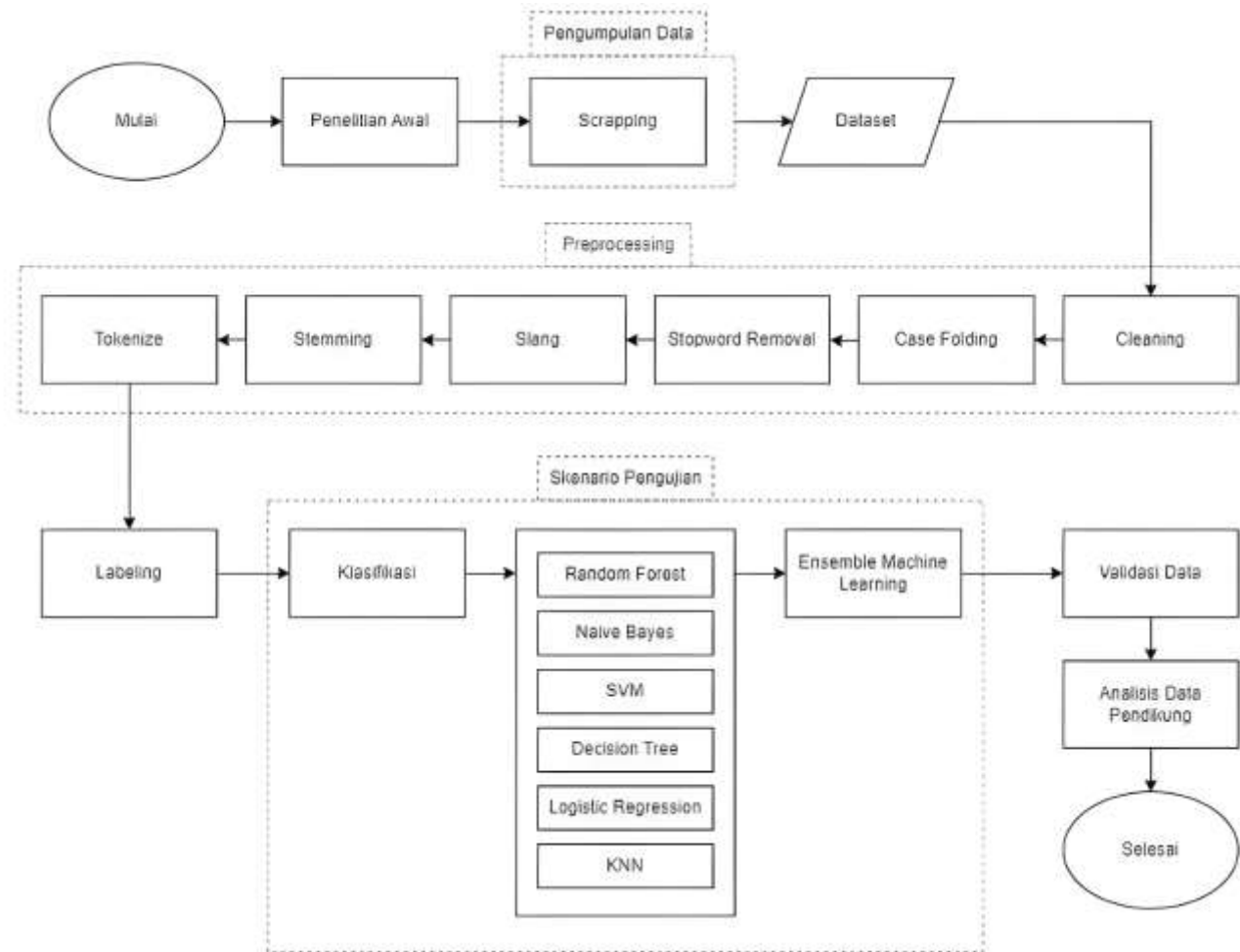
Tujuan

- ✓ Untuk mengetahui bagaimana analisis sentimen pada komentar-komentar YouTube dapat memberikan indikasi terhadap minat dan kepercayaan investor terhadap tim e-sports.

Manfaat Penelitian

- ✓ Dapat mengidentifikasi bagaimana popularitas tim e-sports dalam MPL Season 13 berpengaruh
- ✓ Hasil analisis dapat digunakan oleh investor untuk membuat keputusan yang lebih baik dalam memilih tim e-sports yang berpotensi dan menjanjikan.
- ✓ Menyediakan informasi penting tentang dinamika antara publik, tim, dan investor di industri e-sports, membantu memperkuat dukungan dan strategi investasi dalam turnamen serupa di masa depan.

Metode



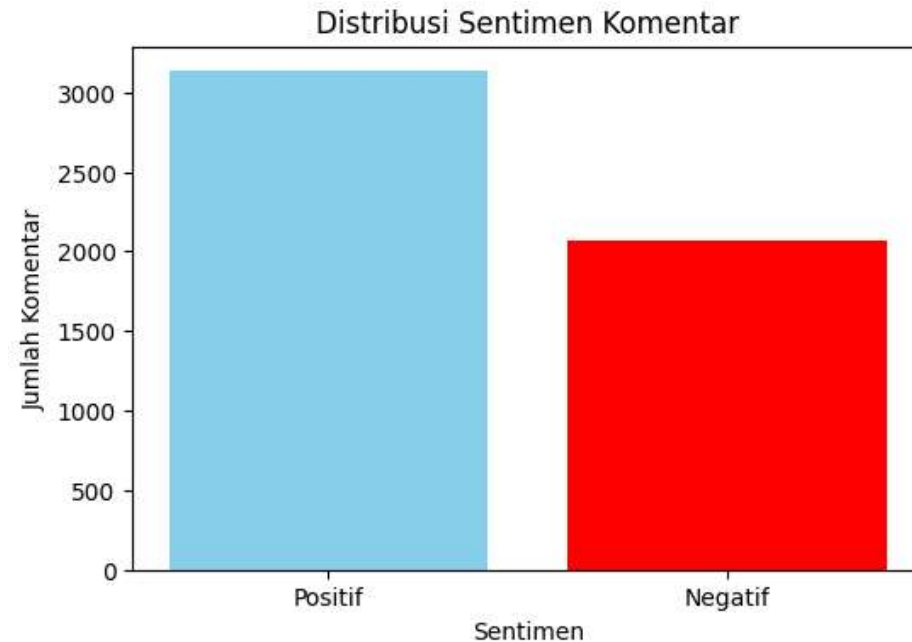
Validasi Data

Pada tahap validasi data, hasil klasifikasi algoritma digabungkan dengan teknik Ensemble Machine Learning dan divalidasi menggunakan Confusion Matrix. Evaluasi dilakukan dengan menghitung F1-Score dan akurasi, serta menganalisis TP, TN, FP, dan FN.

Actual Values	
	Positive (1) Negative (0)
Positive (1)	TP FP
Negative (0)	FN TN

Hasil Sentimen

Tahapan ini menjelaskan hasil pelabelan dengan Flair, mencakup metode otomatis dan manual untuk meningkatkan akurasi. Dari 6.424 komentar, 5.185 relevan, dengan 3.131 positif dan 2.064 negatif. Kombinasi metode ini bertujuan menghasilkan pelabelan yang lebih akurat.



Hasil Wrodcloud

Wordcloud digunakan sebagai visualisasi frekuensi kata dalam teks, memudahkan analisis sentimen dengan cepat mengidentifikasi tema dalam komentar. Kata yang lebih besar menunjukkan frekuensi tinggi, membantu memahami sentimen positif dan negatif. Contoh hasil wordcloud dapat dilihat.



Positif



Negatif

Hasil Pengujian

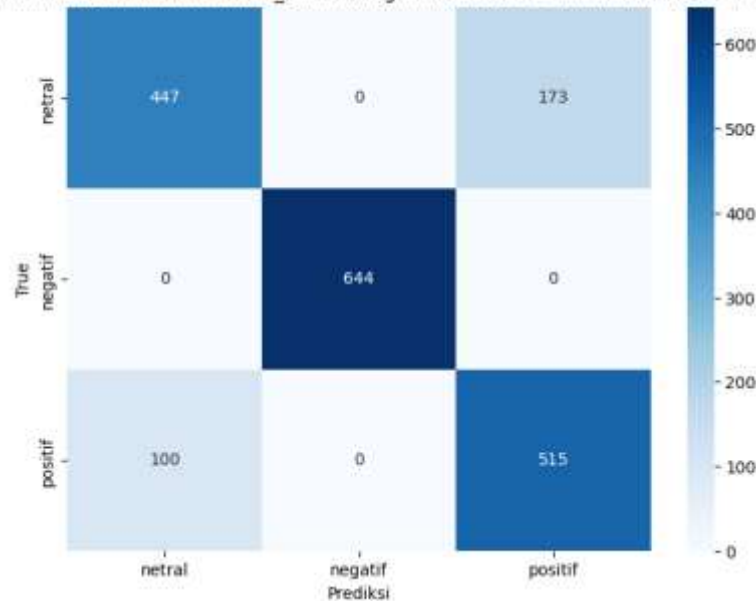
Tahapan ini melibatkan penerapan berbagai metode klasifikasi yang digabungkan menggunakan teknik ensemble machine learning dengan pendekatan majority voting, meliputi hard voting dan soft voting. Sebelum klasifikasi, data diolah menggunakan SMOTE untuk mengatasi ketidakseimbangan kelas, lalu dibagi menjadi data latih dan uji dengan variasi proporsi. Hasil pengujian menunjukkan bahwa performa ensemble machine learning lebih unggul dibanding metode individual, dengan variasi akurasi seperti yang tercantum dalam tabel.

Hasil Pengujian

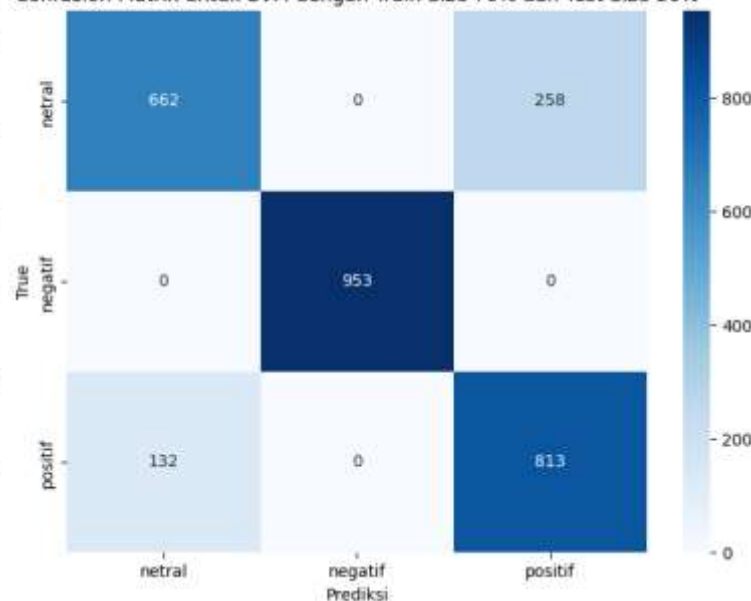
<i>Data train</i>	<i>Data test</i>	<i>Random forest</i>	<i>Support Vector Machine</i>	<i>Naïve bayes</i>	<i>Decision tree</i>	<i>Logistic regression</i>	<i>K-Nearest Neighbors</i>	<i>Voting hard</i>	<i>Voting soft</i>
90	10	83.94%	84.47%	83.62%	81.38%	85.74%	80.85%	86.70%	86.17%
80	20	85.47%	85.42%	83.29%	81.27%	84.83%	78.98%	85.63%	85.90%
70	30	85.34%	86.16%	83.11%	81.72%	85.02%	77.86%	86.48%	85.63%
60	40	83.61%	85.71%	81.80%	79.80%	84.19%	77.86%	85.02%	84.73%
50	50	82.69%	85.03%	81.29%	78.94%	83.18%	77.16%	84.69%	84.27%
40	60	81.92%	84.30%	80.57%	78.57%	83.22%	76.33%	84.14%	83.64%
30	70	79.90%	83.21%	79.32%	76.98%	81.17%	74.67%	82.22%	81.42%
20	80	78.83%	82.16%	78.03%	77.76%	79.49%	73.12%	81.30%	81.29%
10	90	74.88%	79.43%	73.50%	73.49%	75.73%	70.23%	77.70%	77.03%

Validasi Hasil (Confusion Matrix)

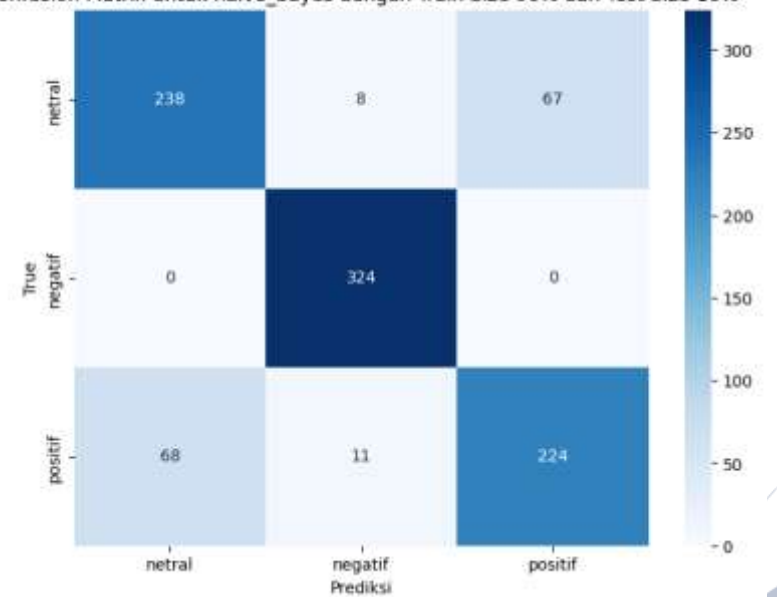
Confusion Matrix untuk random_forest dengan Train Size 80% dan Test Size 20%



Confusion Matrix untuk SVM dengan Train Size 70% dan Test Size 30%

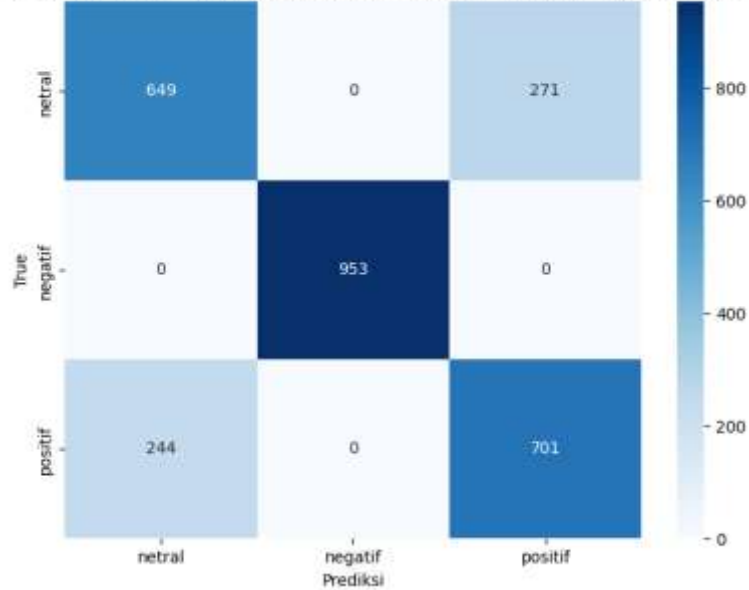


Confusion Matrix untuk naive_bayes dengan Train Size 90% dan Test Size 10%

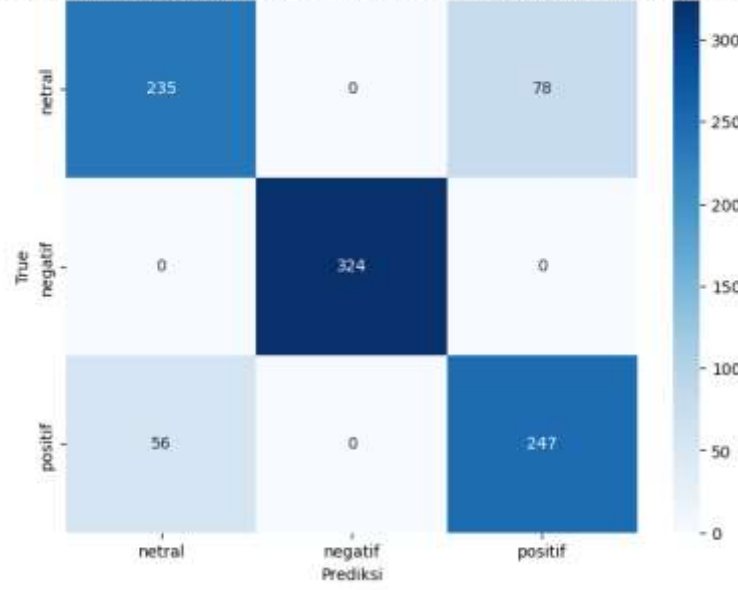


Validasi Hasil (Confusion Matrix)

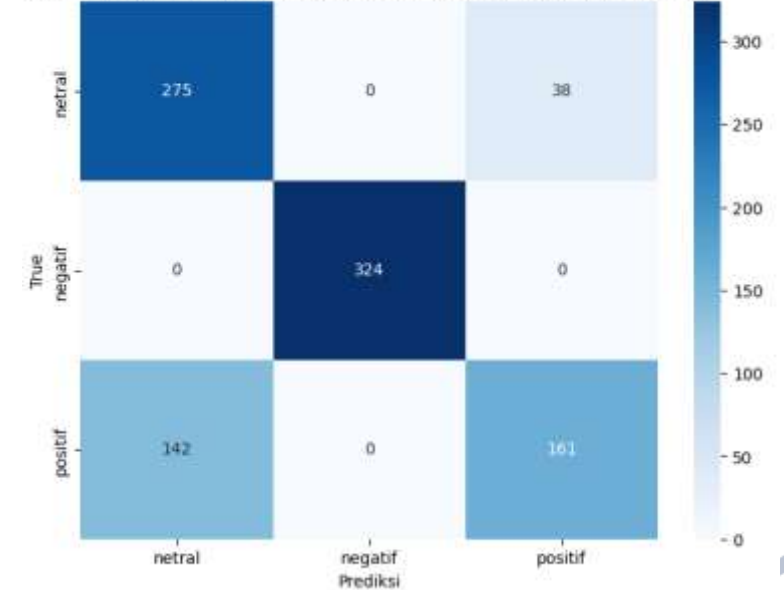
Confusion Matrix untuk decision_tree dengan Train Size 70% dan Test Size 30%



Confusion Matrix untuk logistic_regression dengan Train Size 90% dan Test Size 10%



Confusion Matrix untuk knn dengan Train Size 90% dan Test Size 10%



Referensi

- Agustina, N., & Ihsan, C. N. (2023). Pendekatan Ensemble untuk Analisis Sentimen Covid19 Menggunakan Pengklasifikasi Soft Voting. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 10(2), 263. <https://doi.org/10.25126/jtiik.20231026215>
- Anisa, Y. (2022). Peran Channel Youtube Sebagai Media Alternatif untuk Membantu Proses Pembelajaran Matematika dan Media Informasi pada Tingkat Perguruan Tinggi. In *JPMR* (Vol. 07, Issue 01). <https://ejournal.unib.ac.id/index.php/jpmr>
- Aribowo, A. S., Aribowo, A. S., Basiron, H., Herman, N. S., Khomsah, S., Jaya, H. T., Pembangunan, U., & Veteran, N. (2020). An Evaluation of Preprocessing Steps and Tree-based Ensemble Machine Learning for Analysing Sentiment on Indonesian YouTube Comments. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(5), 7078–7086. <https://doi.org/10.30534/ijatcse/2020/29952020>
- Baskoro, B. B., Susanto, I., & Khomsah, S. (2021). Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR). *INISTA (Journal of Informatics Information System Software Engineering and Applications)*, 3(2), 21–29. <https://doi.org/10.20895/INISTA.V3I2>
- Budiman, J., Limgestu, R., & Tri Sagianto, I. (2023). Perilaku Keputusan Investasi Investor Pasar Saham Indonesia. *Jurnal Ilmiah Akuntansi Dan Keuangan*, 5(9), 3518–3526. <https://journal.ikopin.ac.id/index.php/fairvalue>
- Chely Aulia Misrun, Haerani, E., Fikry, M., & Budianita, E. (2023). Analisis sentimen komentar youtube terhadap Anies Baswedan sebagai bakal calon presiden 2024 menggunakan metode naive bayes classifier. *Jurnal CoSciTech (Computer Science and Information Technology)*, 4(1), 207–215. <https://doi.org/10.37859/coscitech.v4i1.4790>
- Christanti, N., Ariany Mahastanti, L., & Ekonomika dan Bisnis Universitas Kristen Satya Wacana, F. (2011). FAKTOR-FAKTOR YANG DIPERTIMBANGKAN INVESTOR DALAM MELAKUKAN INVESTASI. In *Jurnal Manajemen Teori dan Terapan / Tahun* (Vol. 4, Issue 3).
- Dwi Kurnianto, D., & Waluyo, S. (2023). 3 rd Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) 30 Agustus 2023-Jakarta (Vol. 2, Issue 2).
- Firdaus, R., Al Hariri, R., & Fu'adah Amran, H. (2024). Sentimen Analisis Masyarakat Tentang Penetapan Hari Raya Idul Adha Tahun 2023 Pada Video Youtube Menggunakan Algoritma Random Forest dan Support Vector Machine.
- Firmansyah, D., Sutrisno, F., & Waluyo, S. (2024). ANALISIS SENTIMEN MASYARAKAT MENGENAI KASUS KEKUASAAN NARKOBA PADA KOMENTAR YOUTUBE MENGGUNAKAN METODE K-NEAREST NEIGHBOURS (Vol. 3, Issue 1).

Referensi

- Gunardi, F. (2020). PENGARUH E-SPORTS MARKETING DAN KEPUASAN PELANGGAN TERHADAP BRAND LOYALTY RAZER INC.
- Jimmy Alga, Cindi Wulandari, B. I. (2024). Analisis Sentimen Aplikasi Youtube di Google Play Store Menggunakan Machine Learning. RESOLUSI: Rekayasa Teknik Informatika Dan Informasi - ISSN 2745-7966, 6(1), 197–207. <https://djournals.com/resolusi>
- Khomsah, S. (2021). Sentiment Analysis On YouTube Comments Using Word2Vec and Random Forest Sentimen Analisis pada Opini YouTube Menggunakan Word2Vec dan Random Forest. Jurnal Informatika Dan Teknologi Informasi, 18(1), 61–72. <https://doi.org/10.31515/telematika.v18i1.4493>
- Ningtyas, A. A., Solichin, A., & Pradana, R. (2023). ANALISIS SENTIMEN KOMENTAR YOUTUBE TENTANG PREDIKSI RESESI EKONOMI TAHUN 2023 MENGGUNAKAN ALGORITME NAÏVE BAYES (Vol. 20, Issue 1).
- Puspita Nilamsari, D., & Parma Dewi, I. (2023). Jurnal Vocational Teknik Elektronika dan Informatika. Jurnal Vocational Teknik Elektronika Dan Informatika, 11(1), 96–102. <http://ejournal.unp.ac.id/index.php/voteknika/index>
- Putri, T. H., & Priyatna, C. (2024). Analisis Media Monitoring (Tiara Harjono Putri, dkk.) | 9 Madani. Jurnal Ilmiah Multidisipline, 2(3), 9–21. <https://doi.org/10.5281/zenodo.11099237>
- Rahmani, N. A. (2024). Dinamika Sentimen dalam Mengelola Reputasi ONIC Esports pada MPL ID Season 13. *Jurnal Ilmiah Multidisipline*, 2(5), 336–345. <https://doi.org/10.5281/zenodo.11471752>
- Rosid, M. A., Fitriani, A. S., Astutik, I. R. I., Mulloh, N. I., & Gozali, H. A. (2020). Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi. *IOP Conference Series: Materials Science and Engineering*, 874(1). <https://doi.org/10.1088/1757-899X/874/1/012017>
- Sanjaya, G., & Muslim Lhaksmana, K. (2020). Analisis Sentimen Komentar YouTube tentang Terpilihnya Menteri Kabinet Indonesia Maju Menggunakan Lexicon Based. *E-Proceeding of Engineering*, 7(3), 9698–9710. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/10703>
- Wahid, A., & Saputri, G. (2022). Analisis Sentimen Komentar Youtube Tentang Relawan Patwal Ambulance Menggunakan Algoritma Naïve Bayes dan Decision Tree. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(2), 319. <https://doi.org/10.30865/json.v4i2.4941>
- Wahyuni, D. (2020). TANTANGAN DAN PELUANG ESPORTS DALAM KEOLAHRAGAAN NASIONAL CHALLENGES AND OPPORTUNITIES OF ESPORTS IN NATIONAL SPORTS. *Kajian*, 25(4), 341–353. <https://beritagar.id/artikel/arena/>

Referensi

- Yasir, M., & Suraji, R. (2023). Perbandingan Metode Klasifikasi Naive Bayes, Decision, Tree, Random Forest Terhadap Analisis Sentimen Kenaikan Biaya Haji 2023 pada Media Sosial Youtube. *Jurnal Cahaya Mandalika (JCM)*, 3(2), 180–192.
- Yusrinawati, R. (2023). Analisis Perilaku Konsumtif Pemain Game Mobile Legends Bang Bang di ESports Indonesia Jember يليب. *Nucl. Phys.*, 13(1), 104–116.

