
KLASIFIKASI SENTIMEN BERBASIS MACHINE LEARNING TERHADAP PENONTON “WONDERLAND INDONESIA” PADA PLATFORM YOUTUBE

Surya Darma¹, Linda Wahyuni², Nauval Alfarizi³

^{1,2}Universitas Potensi Utama, Medan

³Universitas Pembangunan Panca Budi, Medan

Email: ¹suryadarma090693@gmail.com, ²lindawahyuni391@gmail.com,
³nauvalalfarizi026@gmail.com

Abstract: *The purpose of this study is to analyze and classify audience sentiment toward the Wonderland Indonesia video on the YouTube platform using the Naive Bayes and Support Vector Machine (SVM) algorithms. The dataset consists of 10,000 comments collected through a scraping process. The data were then further processed using the Sastrawi library, including data cleaning, case folding, tokenization, stopword removal, and stemming. The TF-IDF method with a dimensional size of (9048, 5000) was used for feature representation. The data distribution indicates a class imbalance, where neutral (4,550) and positive (4,297) sentiments dominate compared to negative (201). The experimental results show that SVM outperforms Naive Bayes, achieving an accuracy of 93% compared to 86%. However, both models encounter difficulties in identifying the negative class. The use of imbalanced data handling techniques shows that SVM with class weighting improves the detection of minority classes without reducing overall accuracy. In contrast, Naive Bayes with SMOTE improves recall but reduces overall performance. The findings indicate that SVM with class weighting is the best-performing model for sentiment classification. Additionally, the results show that the majority of comments are positive, suggesting that audiences respond favorably to the Wonderland Indonesia content.*

Keywords: *Sentiment Analysis, Naive Bayes, Support Vector Machine, YouTube, TF-IDF*

Abstrak: Tujuan penelitian ini adalah untuk menganalisis dan mengklasifikasikan sentimen penonton terhadap video Wonderland Indonesia yang ditampilkan di platform YouTube menggunakan algoritma Naive Bayes dan Support Vector Machine (SVM). Data yang digunakan terdiri dari 10.000 komentar yang dikumpulkan melalui proses scraping. Kemudian, data diproses lebih lanjut dengan menggunakan pustaka Sastrawi untuk membersihkan, menggabungkan case, tokenize, menghilangkan stopword, dan stemming. Metode TF-IDF dengan dimensi (9048, 5000) digunakan untuk mempresentasikan fitur. Distribusi data menunjukkan ketidakseimbangan kelas; sentimen netral (4.550) dan positif (4.297) mendominasi dibandingkan dengan negatif (201). Hasil pengujian menunjukkan bahwa SVM unggul dengan akurasi 93% dibandingkan Naive Bayes sebesar 86%. Namun, kedua model kesulitan menemukan kelas negatif. SVM dengan berat kelas dapat meningkatkan deteksi kelas minoritas tanpa mengurangi akurasi, seperti yang ditunjukkan oleh penggunaan teknik penanganan data tidak seimbang. Di sisi lain, Naive Bayes dengan SMOTE meningkatkan recall tetapi mengurangi performa keseluruhan. Hasil penelitian menunjukkan bahwa model terbaik untuk klasifikasi sentimen adalah SVM dengan berat kelas. Hasil penelitian juga menunjukkan bahwa sebagian besar komentar bersentimen positif, yang menunjukkan bahwa audiens menyukai konten Wonderland Indonesia.

Kata Kunci: Analisis Sentimen, Naive Bayes, SVM, YouTube, TF-IDF

PENDAHULUAN

Dengan berkembangnya teknologi informasi dan komunikasi, pola konsumsi media telah berubah secara signifikan. Ini terutama berlaku untuk platform berbasis video seperti YouTube, yang selain berfungsi sebagai tempat untuk mendistribusikan konten audiovisual, juga memungkinkan pengguna berinteraksi dengan orang lain secara langsung melalui fitur komentarnya. Data tekstual yang sangat besar dihasilkan dari interaksi ini menunjukkan pendapat, persepsi, dan perasaan audiens terhadap konten tertentu. (Saifullah et al., 2021), (Hoiles et al., 2019), (Subha dan Bharathi, 2024).

Wonderland Indonesia adalah produk audiovisual yang menarik perhatian publik karena menggabungkan sinematografi, musik, dan kekayaan budaya Indonesia secara inovatif. Banyak komentar yang dibuat menunjukkan keterlibatan tinggi audiens terhadap konten ini. Ini memungkinkan untuk melakukan analisis lebih lanjut tentang tanggapan penonton. Analisis komentar menjadi penting untuk memahami bagaimana audiens memaknai konten. Ini juga berfungsi sebagai pengukur seberapa baik penyampaian pesan dalam media digital.

Dalam situasi seperti ini, analisis sentimen menjadi metode yang relevan untuk menemukan dan mengkategorikan opini publik ke dalam kategori tertentu, seperti sentimen positif, netral, dan negatif. Analisis sentimen juga merupakan bagian dari text mining, yang memanfaatkan teknik pembelajaran mesin untuk mengekstrak informasi penting dari data teks tidak terstruktur. Namun, menggunakan analisis sentimen pada data media sosial menghadapi beberapa tantangan, salah satunya adalah ketidakkonsistenan. (Saifullah et al., 2021), (Subha dan Bharathi, 2024), (Puspasari dan Subarkah, 2022), (Poornima, 2024), (Pokharel dan Bhatta, 2021), (Sangeetha dan Nimala, 2024), (Fadhilah dan Utomo, 2024).

Untuk menyelesaikan masalah ini, metode klasifikasi yang memiliki kemampuan generalisasi yang baik terhadap data teks diperlukan. Dua teknik yang banyak digunakan dalam klasifikasi sentimen adalah Algoritma Naive Bayes dan Support Vector Machine (SVM). Naive Bayes dikenal karena sederhana dan efektif dalam memproses data teks berbasis probabilistik, sedangkan SVM memiliki keunggulan dalam menciptakan batas keputusan yang optimal melalui konsep maximum margin, yang memungkinkannya menangani data yang sangat besar dengan lebih baik.

Penelitian ini bertujuan untuk menerapkan dan membandingkan kinerja algoritma Naive Bayes dan SVM saat mengklasifikasikan sentimen penonton terhadap video Wonderland Indonesia di YouTube. Selain itu, penelitian ini juga melihat bagaimana ketidakseimbangan data mempengaruhi kinerja model dan bagaimana metode penanganan dapat meningkatkan akurasi klasifikasi, terutama untuk kelas minoritas. Diharapkan hasil penelitian ini akan membantu mengembangkan teknik analisis sentimen yang menggunakan pembelajaran mesin dan membantu memahami bagaimana audiens Indonesia merespons konten digital berbasis budaya.

METODE

Untuk mengklasifikasikan sentimen, penelitian ini menggunakan pendekatan kuantitatif yang menggunakan teknik *text mining* dan *machine learning*. Penelitian mencakup beberapa proses utama berikut:

Pengumpulan Data (*Scraping Youtube*).

Data dikumpulkan melalui proses *scraping* komentar pada video Wonderland Indonesia di YouTube sebanyak 10.000 data. Setiap komentar diasumsikan sebagai pendapat pribadi karena struktur tanggapan tidak dipertimbangkan. Setelah itu, proses *preprocessing* dilakukan untuk

meningkatkan kualitas data, yang mencakup:

1. *Data cleaning* (menghapus emoji, simbol, dan bahasa asing)
 2. *Case folding*
 3. *Tokenizing*
 4. *Stopword removal*
 5. *Stemming* menggunakan Sastrawi
- Setelah proses ini, data berkurang menjadi 9.048 komentar.

Pelabelan Data (*Labeling*)

Data dilabeli dengan tiga kategori emosi: positif, netral, dan negatif.

Ekstraksi Fitur

Metode TF-IDF digunakan untuk mengubah data teks menjadi bentuk numerik dengan dimensi (9048, 5000).

Pembagian Data

Data dibagi menjadi data latih dan data uji untuk proses klasifikasi.

Klasifikasi Model

Model yang digunakan:

1. *Naive Bayes*
2. *Support Vector Machine (SVM)*

Selain itu, diterapkan teknik penanganan *imbalanced data*:

1. *Class Weight* pada SVM
2. *SMOTE* pada Naive Bayes

Evaluasi Model

Evaluasi dilakukan dengan metrik:

1. *Accuracy*
2. *Precision*
3. *Recall*
4. *F1-Score*



Gambar 1 Diagram Metode Penelitian

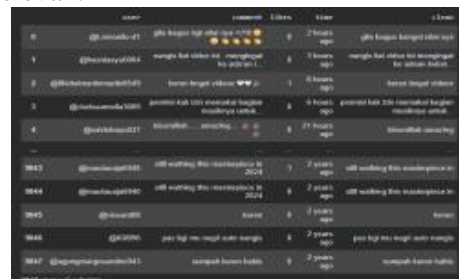
HASIL DAN PEMBAHASAN

Penelitian ini berfokus pada pengolahan data menjadi analisis sentimen terhadap video promosi. Fokus utama penelitian adalah mengidentifikasi pola sentimen pada kolom komentar yang dapat dimanfaatkan sebagai dasar evaluasi performa dalam membandingkan dua algoritma klasifikasi, yaitu *Naive Bayes* dan *Support Vector Machine (SVM)*.



Gambar 2 Data Collecting Dataset

Setiap komentar dianggap sebagai pendapat pribadi karena struktur tanggapan tidak dipertimbangkan. Setelah *preprocessing*, 9.048 data dari 10.000 data yang dihapus dihapus, termasuk emoji, bahasa asing, dan karakter tidak relevan.



Gambar 3 Data Cleaning

Ketika *preprocessing* data selesai, hal yang perlu dilakukan berikutnya adalah *tf-idf*, seperti yang ditunjukkan pada gambar di bawah ini.



Gambar 4 Stemming

Untuk menguji kemampuan model untuk memahami dan mengklasifikasikan bahasa secara otomatis, teknik *lexicon* manual digunakan untuk melabelkan data, dengan masing-masing 150 data untuk setiap kelas sentimen.

```
data['label'].value_counts()

count
label
netral    4550
positif   4297
negatif    201
dtype: int64
```

Gambar 5 Labeling Kelas Pada Dataset

Penelitian ini menggunakan pustaka *sastrawi* untuk melakukan tahapan *preprocessing*, yang mencakup membersihkan data, memotong *casefold*, *tokenizing*, menghapus *stopword*, dan *stemming*.



Gambar 6 Alur preprocessing

Dataset berbahasa Indonesia, jadi penggunaan *Sastrawi* dipilih. Untuk mengubah kata berimbuhan menjadi bentuk dasar, seperti "pembaca" menjadi "baca", buku ini menggunakan algoritma *stemming* Nazief dan Adriani. (Iskandar et al., 2023), (Hidayat et al., 2021).

Dataset dianggap tidak seimbang atau tidak seimbang karena kelas netral (4.550) dan positif (4.297) mendominasi dibandingkan kelas negatif (201) berdasarkan distribusi data. Kondisi ini dapat mengurangi kemampuan model untuk mengidentifikasi kelas minoritas.

Menggunakan TF-IDF untuk mengekstrak fitur, dimensi (9048, 5000) menghasilkan 9.048 dokumen dan 5.000 fitur. Pelatihan model *Naive Bayes* dan *Support Vector Machine (SVM)* didasarkan pada representasi ini.

Hasil perbandingan performa kedua algoritma ditunjukkan pada Tabel 1 berikut.

Tabel 1 Perbandingan Performa Sebelum Tindakan preprocessing

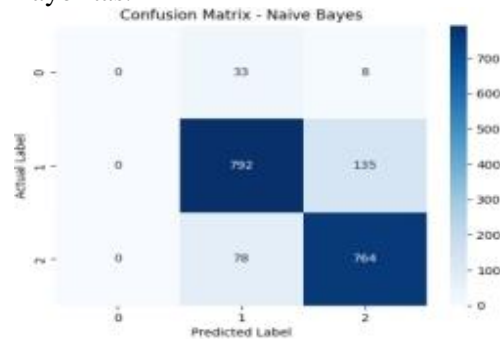
Mod el	Kelas	Preci sion	Reca ll	F1- Score	Sup port
Naive Bayes	Negatif	0.00	0.00	0.00	41
	Netral	0.88	0.85	0.87	927
	Positif	0.84	0.91	0.87	842
	Accurac y			0.86	1810
	Macro Avg	0.57	0.59	0.58	1810
	Weighted Avg	0.84	0.86	0.85	1810

SVM	Negatif	0.33	0.05	0.09	41
	Netral	0.90	0.98	0.94	927
	Positif	0.98	0.92	0.95	842
	Accurac y			0.93	1810
	Macro Avg	0.74	0.65	0.66	1810
	Weighted Avg	0.92	0.93	0.92	1810

Hasil pengujian menunjukkan bahwa SVM lebih baik daripada Naive Bayes, dengan akurasi 93% versus 86%. Kedua model bekerja dengan baik di kelas netral dan positif, tetapi Naive Bayes gagal menemukan kelas negatif dengan skor f1 0.00, sementara SVM hanya mendapatkan skor f1 0,09.

Keterbatasan ini terjadi karena ketidakseimbangan data; jumlah kelas negatif lebih sedikit daripada jumlah kelas lainnya. Ini menunjukkan bahwa model cenderung bias terhadap kelas mayoritas,

yang juga ditunjukkan oleh nilai macro average yang rendah. Meskipun dataset memiliki nilai rata-rata yang berat, nilai ini dipengaruhi oleh dominasi kelas mayoritas.

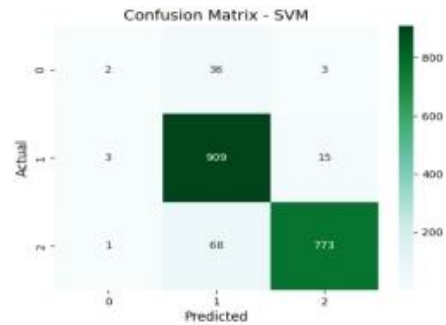


Gambar 7 Confusion Matrix Naïve Bayes

Hasil Naive Bayes menunjukkan kinerja yang buruk pada kelas negatif. Model tidak dapat mengklasifikasikan seluruh 41 data negatif dengan benar, dengan 33 diprediksi sebagai netral dan 8 diprediksi sebagai positif. Ketidakmampuan model untuk mengidentifikasi karakteristik sentimen negatif adalah buktinya.

Confusion matrix menunjukkan bahwa *Naive Bayes* cenderung bias terhadap kelas mayoritas (netral dan positif) akibat ketidakseimbangan data, di mana kelas negatif memiliki jumlah yang jauh lebih sedikit. Kondisi ini menyebabkan model lebih sering memprediksi ke kelas dominan.

Akibatnya, untuk meningkatkan kinerja klasifikasi, terutama untuk kelas negatif, metode penanganan data yang tidak seimbang diperlukan. Untuk membandingkan peningkatan kinerja model dalam penelitian ini, metode class weight dan SMOTE digunakan.



Gambar 8 Confusion Matrix SVM

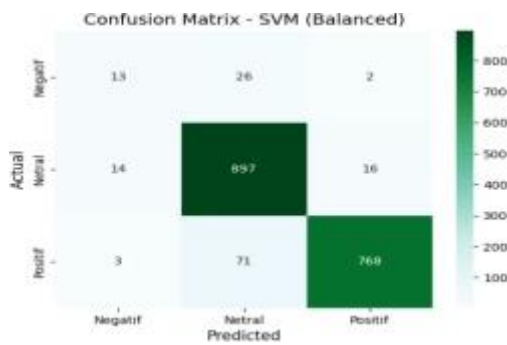
Penelitian ini menerapkan algoritma SVM dengan penyesuaian *class weight*, sebagaimana ditunjukkan pada Tabel 2

Tabel 2 Hasil SVM dengan Penanganan Imbalance (Class Weight)

Kelas	Precision	Recall	F1-Score	Support
Negatif	0.43	0.32	0.37	41
Netral	0.90	0.97	0.93	927
Positif	0.98	0.91	0.94	842
Accuracy			0.93	1810
Macro Avg	0.77	0.73	0.75	1810
Weighted Avg	0.93	0.93	0.93	1810

Kemampuan model untuk mengatasi bias kelas, terutama pada kelas negatif, ditingkatkan dengan menggunakan berat kelas pada SVM. Nilai akurasi adalah 0.43, dan performa pada kelas positif tetap tinggi, dengan akurasi 0.98 dan skor f1-nya 0.94, sementara akurasi keseluruhan tetap stabil pada 93%.

Hasil ini menunjukkan bahwa penanganan data tidak seimbang dapat meningkatkan deteksi kelas minoritas tanpa mengurangi performa model secara keseluruhan.



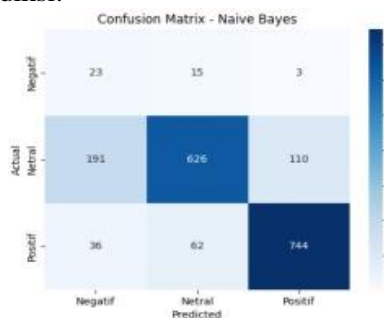
Gambar 9 Confusion Matrix SVM dengan Balanced

Di kelas negatif, 13 dari 41 data diklasifikasikan dengan benar, 26 diprediksi netral, dan 2 positif. Ini menunjukkan kemampuan model untuk mendeteksi sentimen negatif, meskipun masih cenderung bias ke kelas netral.

Tabel 3 Hasil Naïve Bayes dengan Penanganan Smote

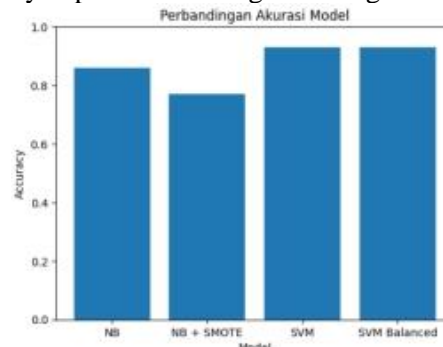
Kelas	Precision	Recall	F1-Score	Support
Negatif	0.09	0.56	0.16	41
Netral	0.89	0.68	0.77	927
Positif	0.87	0.88	0.88	842
Accuracy			0.77	1810
Macro Avg	0.62	0.71	0.60	1810
Weighted Avg	0.86	0.77	0.80	1810

Peningkatan *recall* sebesar 0.56 pada kelas negatif menunjukkan bahwa lebih dari 50% data negatif terdeteksi dengan sukses. Nilai *f1-score* (0,16) masih belum optimal karena *precision* yang rendah (0,09) menunjukkan banyaknya kesalahan prediksi.



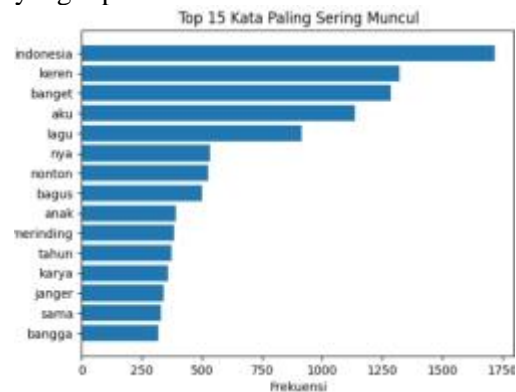
Gambar 10. Confusion Matrix NB dengan Smote

Secara keseluruhan, penerapan SMOTE meningkatkan kemampuan model untuk menemukan kelas negatif, tetapi menurunkan kinerja pada kelas mayoritas, terutama netral, karena proses resampling mendorong model untuk membuat lebih banyak prediksi tentang kelas negatif.



Gambar 11 Perbandingan Performa Algoritma

Secara keseluruhan, SVM dan SVM dengan *class weight* paling baik, tetapi Naive Bayes dengan SMOTE tidak lebih akurat dan mungkin lebih buruk. Hal ini menunjukkan bahwa hasil klasifikasi sangat dipengaruhi oleh metode penanganan data dan algoritma yang dipilih.



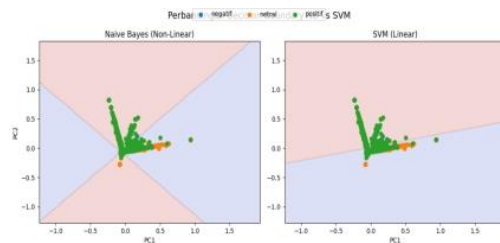
Gambar 12 Distribusi Kata dengan Frekuensi Terbanyak

Hasil visualisasi menunjukkan bahwa kata "indonesia" adalah yang paling banyak digunakan, diikuti oleh kata-kata bernuansa positif seperti "keren", "banget", dan "bagus". Ini menunjukkan bahwa sebagian besar komentar bersifat positif dan menghargai konten yang dibahas.



Gambar 13 Sentimen Wordcloud

Secara visual, *wordcloud* menunjukkan distribusi fitur yang relatif seimbang. Namun, sebagian besar data terdiri dari sentimen positif dan netral, yang mencakup hampir 90% dari dataset.



Gambar 14 Perbandingan Naive Bayes dengan SVM

Sebaliknya, SVM menggunakan optimasi margin antarkelas untuk membentuk batas keputusan linear yang lebih stabil dan konsisten, sementara

Naive Bayes menghasilkan batas keputusan non-linear yang tidak teratur, yang menyebabkan data yang saling tumpang tindih tidak optimal.

Proses klasifikasi menjadi lebih sulit karena distribusi data yang tidak seimbang. Ini terutama berlaku untuk jumlah kelas negatif yang sangat kecil. Namun, SVM masih lebih baik daripada PCA, terutama ketika menangani data yang sangat besar dari hasil ekstraksi TF-IDF yang diproyeksikan.

SIMPULAN

Distribusi sentimen menunjukkan bahwa kelas netral (4.550 data) dan positif (4.297 data) mendominasi, dengan sedikit data kelas negatif. Kondisi ini menyebabkan ketidakseimbangan data, atau kumpulan data tidak seimbang, yang berdampak pada kinerja model, terutama untuk mengidentifikasi kelas minoritas. Dalam pengujian awal, algoritma SVM lebih baik daripada Naive Bayes, dengan akurasi 93% dibandingkan dengan Naive Bayes hanya 86%. Namun, karena ketidakseimbangan data, kedua model menghadapi masalah dalam mengklasifikasikan kelas negatif; Naive Bayes gagal menemukan kelas negatif sama sekali.

Berdasarkan perbandingan tersebut, dapat disimpulkan bahwa SVM dengan penanganan kelas berat adalah model terbaik untuk penelitian ini. Ini karena SVM ini mampu memberikan performa yang seimbang antara kemampuan mendeteksi kelas minoritas dan akurasi tinggi.

Selain itu, hasil visualisasi seperti *wordcloud* dan frekuensi kata menunjukkan bahwa sebagian besar komentar memiliki sentimen positif, dengan kata-kata seperti "indonesia", "keren", dan "bagus" mendominasi. Hal ini menunjukkan bahwa audiens memberikan respons yang baik terhadap konten video yang dianalisis.

DAFTAR PUSTAKA

- Al Fayed, A. J., Darma, S., Sinabariba, Z., & Pardede, S. M. P. (2025). Comparison of Naïve Bayes, K-Nearest Neighbors, and Decision Tree methods for classifying heart disease risk factors. *Journal of Computer Science and Research (JoCoSiR)*, 3(3), 81–88.
- Al-fraiji, S. S., & Al-Shammary, D. (2021). EEG Signals Classification based on mathematical selection and cosine similarity. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 13(3). <https://doi.org/10.29304/jqcm.2021.13.3.837>
- Darma, S., Al Fayed, A. J., Pardede, S. M. P., Aqsha, M. H., & Novelan, M. S. (2026). Predictive analysis of flood risk factors based on a machine learning approach: Comparative study of SVM and XGBoost algorithms. *Journal of Technology and Computer (JOTECHCOM)*, 3(1), 24–33. <https://journal.technolabs.co.id/index.php/jotechcom/article/view/94>
- Fadhilah, S. N., & Utomo, F. S. (2024). Naïve Bayes Algorithm for Sentiment Analysis of Blibli.com Review on Google Play Store. *Sistemasi*, 13(2), 831. <https://doi.org/10.32520/stmsi.v13i2.3887>
- Fokoué, E. (2018). To Bayes or Not To Bayes? That's no longer the question! <https://doi.org/10.48550/arxiv.1805.11012>
- Gregory, D. E. (2012). Choosing a Graduate School. *Educational Horizons*, 90(3), 5–9. <https://doi.org/10.1177/0013175x1209000302>
- Hidayat, W., Utami, E., & Hartanto, A. D. (2021). Pemilihan Parameter Terbaik pada Algoritma Winnowing dalam Mendeteksi Tingkat Kesamaan Dokumen Bahasa Indonesia. *Creative Information Technology Journal*, 7(2), 119. <https://doi.org/10.24076/citec.2020v7i2.256>
- Hoiles, W., Krishnamurthy, V., & Pattanayak, K. (2019). Rationally Inattentive Inverse Reinforcement Learning Explains YouTube Commenting Behavior. <https://doi.org/10.48550/arxiv.1910.11703>
- Iskandar, A. F., Utami, E., Hidayat, W., Budi, A. P., & Hartanto, A. D. (2023). Modifikasi Fonem Vokal Pada Stemming Kata Tidak Baku. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 10(1), 35. <https://doi.org/10.25126/jtiik.20231015028>
- Mavhemwa, P. M., Zennaro, M., Nsengiyumva, P., & Nzanywayingoma, F. (2024). Weighted naïve bayes multi-user classification for adaptive authentication. *Journal of Physics Communications*, 8(10), 105005. <https://doi.org/10.1088/2399-6528/ad8a16>
- Meyer, M. S., Cranmore, J., Rinn, A. N., & Hodges, J. (2020). College Choice: Considerations for Academically Advanced High School Seniors. *Gifted Child Quarterly*, 65(1), 52–74. <https://doi.org/10.1177/0016986220957258>
- Missaghian, R. (2021). Social Capital and Post-Secondary Decision-Making Alignment for Low-Income Students. *Social Sciences*, 10(3), 83. <https://doi.org/10.3390/socsci10030083>
- Poornima, K. M. (2024). Rating Based on YouTube Comments. *Interantional Journal of Scientific Research in Engineering and Management*, 08(03), 1–5. <https://doi.org/10.55041/ijrsrem29798>
- Pokharel, R., & Bhatta, D. (2021). Classifying YouTube Comments Based on Sentiment and Type of Sentence.

- <https://doi.org/10.48550/arxiv.2111.01908>
- Puspasari, H. M., & Subarkah, P. (2022). Sentiment Analysis for Opinions on the Covid-19 Vaccination Program Using a Naive Bayes Classifier. *Jurnal Borneo Administrator*, 18(3), 213–230.
<https://doi.org/10.24258/jba.v18i3.992>
- Reynolds, J., Elliott, J. L., Castillo, K., Sliwak, R. M., & Halligan, C. S. (2023). I lost my mentor, now what? The experiences of counseling psychology women doctoral students who lost their mentor: Training and program implications. *Qualitative Psychology*, 10(2), 227–244.
<https://doi.org/10.1037/qap0000237>
- Saifullah, S., Fauziyah, Y., & Aribowo, A. S. (2021). Comparison of machine learning for sentiment analysis in detecting anxiety based on social media data. *Jurnal Informatika*, 15(1), 45.
<https://doi.org/10.26555/jifo.v15i1.a2>
- Sangeetha, M., & Nimala, K. (2024). Unravelling Emotional Tones: A Hybrid Optimized Model for Sentiment Analysis in Tamil Regional Languages. *Journal of Machine and Computing*, 114–126.
<https://doi.org/10.53759/7669/jmc202404012>
- Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., Derous, E., Fruyt, F. D., & Duyck, W. (2021). How interest fit relates to STEM study choice: Female students fit their choices better. *Journal of Vocational Behavior*, 129, 103614.
<https://doi.org/10.1016/j.jvb.2021.103614>
- Subha, K., & Bharathi, N. (2024). Leveraging spark-based machine learning algorithm for audience sentiment analysis in youtube content. *Intelligent Data Analysis*, 28(5), 1395–1405.
<https://doi.org/10.3233/ida-240198>
- Taslim, T., Handayani, S., & Fajrizal, F. (2023). Kinerja Komparatif Optimasi Algoritma Naive Bayes dalam Klasifikasi Teks untuk Uji Klinis Kanker. *Eksplora Informatika*, 13(1), 113–123.
<https://doi.org/10.30864/eksplora.v13i1.994>
- Veness, J., Hütter, M., Orseau, L., & Bellemare, M. G. (2014). Online Learning of k-CNF Boolean Functions.
<https://doi.org/10.48550/arxiv.1403.6863>