

Analisis Sentimen Opini Publik Terhadap Chatgpt di Twitter Menggunakan Metode Naive Bayes

Donny Panca Wilie

Program Studi Informatika, Universitas Multi Data Palembang
Jl. Rajawali No.14, Kec. Ilir Timur II, Kota Palembang, Sumatra Selatan, 30113
e-mail: donnypanca@mhs.mdp.ac.id

Informasi Artikel:

Received: 14 September 2023, Accepted: 27 September 2023, Published: November 2023

Abstract

The presence of ChatGPT, which is based on Artificial Intelligence (AI), has attracted public attention and become a polemic. This application is beneficial for users in finding information, translating languages, and generating text similar to human speech. OpenAI's decision to launch the ChatGPT Plus service is related to its monetization strategy and maintaining user interest. In this regard, a study was conducted using the Naïve Bayes method to classify the sentiment of the Twitter community towards ChatGPT in the period 30/11/2022 to 31/12/2022. This study took 5001 tweet data from the available dataset and used TF-IDF and BOW feature extraction, dividing data into 70% training and 30% test data. This research gets the highest accuracy value of 59% using BOW feature extraction.

Keywords: ChatGPT, Twitter, Naïve Bayes, TF- IDF, BOW.

1. Pendahuluan

ChatGPT adalah sebuah aplikasi yang dikembangkan oleh OpenAI, yang menggunakan teknologi *deep learning* dengan *Natural Language Processing* (NLP) yang berguna dalam membantu pengguna dalam menemukan informasi, menerjemahkan bahasa, atau menghasilkan teks baru yang mirip dengan ucapan manusia (Nursaid, 2023). Saat ini juga, media sosial telah menjadi sarana bagi masyarakat Indonesia untuk berkomunikasi dan menyatakan pendapat mereka. *Twitter* merupakan salah satu media sosial yang sangat populer di Indonesia sebagai platform untuk menyuarakan aspirasi melalui *tweet* atau cuitan (Sabily dkk., 2019). Salah satu topik yang banyak dibicarakan oleh masyarakat di *Twitter* saat ini adalah ChatGPT. ChatGPT dapat dijadikan pendukung dalam pembuatan keputusan tertentu untuk menunjang keberhasilan layanan monetisasi dan mempertahankan minat pengguna terhadap ChatGPT.

Untuk menentukan nilai frekuensi sebuah kata dalam sebuah dokumen dapat menggunakan metode TF-IDF. Metode ini menentukan seberapa relevan sebuah kata di dalam sebuah dokumen. Penggunaan metode ini berguna untuk memberikan peringkat pada dokumen yang dipelajari secara ekstensif, untuk model ruang vektor. Mengikuti pertimbangan dari model

Published by:

diskriminasi kata, mereka berpendapat bahwa kata yang muncul dalam dokumen harus diberi nilai sebanding dengan frekuensi kata dan terbalik sebanding dengan frekuensi dokumen. Pembobotan skema yang mengikuti pendekatan ini disebut TF x IDF (Qasem dkk., 2015).

Selain metode TF-IDF, metode *Bag of Words* (BOW) adalah salah satu metode yang digunakan untuk mengambil fitur dari kalimat atau dokumen. Model BOW merubah data teks menjadi vektor dengan panjang yang telah ditentukan melalui penghitungan kemunculan kata-kata. Penggunaan BOW melibatkan dua tahap, yaitu menentukan kata-kata yang ada dalam seluruh kalimat atau dokumen, dan menentukan metode penilaian untuk setiap kemunculan kata yang telah diidentifikasi sebelumnya. Metode penilaian dapat berupa metode sederhana seperti menggunakan nilai biner untuk menunjukkan keberadaan kata atau menggunakan nilai non-biner untuk menunjukkan frekuensi kemunculan kata tersebut (Beheshti dkk., 2022). Penerapan metode BOW telah dilakukan pada penelitian mengenai analisis sentimen yang berfokus pada pendekatan yang digunakan untuk menganalisis sentimen pada *tweet* berbahasa Azerbaijan. Penelitian ini menggunakan metode *Logistic Regression*, *Naïve Bayes*, dan *Support Vector Machine* dengan menerapkan teknik *vectorizer Bag of Words* dan TF-IDF pada teks. Hasil penelitian ini menunjukkan bahwa *Bag of Words* memberikan tingkat akurasi yang lebih tinggi dibandingkan dengan TF-IDF dalam ketiga metode analisis sentimen tersebut (Hasanli & Rustamov, 2019).

Untuk mengklasifikasikan banyaknya cuitan pada *Twitter*, diperlukan waktu yang lama. Oleh karena itu, diperlukan metode untuk mengklasifikasikan data *Twitter* tersebut agar dapat dianalisis dengan dikelompokkan ke dalam suatu kelas. Pada penelitian ini digunakan metode klasifikasi *Naïve Bayes*. Metode *Naïve Bayes* sering digunakan karena dianggap cukup efektif dalam mendapatkan hasil yang tepat dengan akurasi yang tinggi (Kusumadewi, 2009).

Salah satu kelebihan dari menggunakan algoritma *Naïve Bayes* adalah kemampuannya untuk bekerja dengan baik dalam kasus-kasus di mana jumlah fitur yang digunakan untuk klasifikasi sangat besar. Selain itu, *Naïve Bayes* juga relatif cepat dan mudah diimplementasikan, sehingga cocok untuk digunakan dalam aplikasi real-time. Dalam sebuah studi yang dilakukan oleh Domingos dan Pazzani pada tahun 1997, mereka menemukan bahwa *Naïve Bayes* sering menghasilkan performa yang lebih baik daripada algoritma klasifikasi lainnya pada data yang kompleks dan dengan jumlah fitur yang besar (P. Domingos & Pazzani, 1997). Metode *naïve bayes* memiliki tingkat akurasi yang baik dalam melakukan analisis sentimen. Oleh karena itu, pada penelitian ini dilakukan analisis sentimen opini publik terhadap ChatGPT di twitter menggunakan metode *naïve bayes* diikuti dengan pembobotan data menggunakan metode TF-IDF, dan metode BOW untuk meningkatkan perfoma algoritma dalam melakukan analisis sentimen.

2. Tinjauan Literatur

Penelitian Sebelumnya dilakukan oleh Hennie Tuhuteru pada tahun 2020 dengan judul Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine “ dengan hasil analisis sentimen yang dilakukan terhadap 725 dokumen data testing dengan metode klasifikasi SVM dan 9-fold cross validation menghasilkan sentimen positif sebesar 28%, sentimen negatif 27% dan sentimen netral sebesar 45%. Jika dilihat dari jumlah data dan sumbernya, data terbanyak berasal dari platform *Twitter*, yaitu 825 *tweets* sedangkan Facebook hanya 250 komentar karena hasil pencarian yang ditampilkan terbatas.

Penelitian yang dilakukan oleh Akhmad Deviyanto, Muhammad Didik Rohmad Wahyudi pada

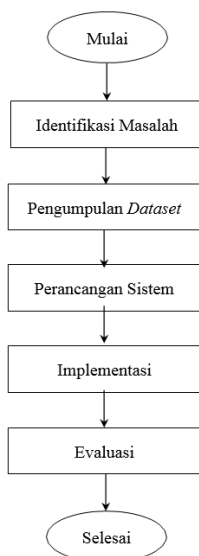
Published by:

tahun 2018 dengan judul “Penerapan Analisis Sentimen pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor” Hasil analisis pengujian menghasilkan nilai akurasi terbesar adalah 67,2% ketika $k=5$, presisi tertinggi 56,94% ketika $k=5$, dan recall 78,24% dengan $k=15$

Penelitian yang dilakukan oleh Samsir, Ambiyar, Unung Verawardina, Firman Edi, Ronal Watrianthos pada tahun 2021 dengan judul “Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode *Naïve Bayes*” Hasil dari pengujian ini adalah Penelitian pada periode tersebut menunjukkan 30% sentimen positif, 69% sentimen negatif, dan 1% netral. Persepsi negatif dihasilkan karena ketidakpuasan masyarakat terhadap pembelajaran daring. Beberapa tweet menunjukkan kekecewaan dengan kata ‘stres’ dan ‘malas’ merupakan kata yang memiliki frekuensi tinggi dalam percakapan pada periode tersebut. Berdasarkan hasil ini pembelajaran daring belum maksimal diterapkan di Indonesia pada masa pandemi yang terlihat dari tingginya kekecewaan public pada awal November 2020. Penelitian lebih lanjut diharapkan menerapkan algoritma yang berbeda untuk mendapatkan hasil yang lebih akurat dalam analisis opini.

3. Metode Penelitian

Berikut adalah beberapa tahapan yang dilakukan dapat dilihat pada Gambar 1 untuk melakukan sentimen analisis terhadap Chat-GPT menggunakan fitur TF-IDF dan *Bag of Words*, serta metode *Naïve Bayes*.



Gambar 1. Kerangka Kerja Penelitian

3.2 Identifikasi Masalah

Tahap ini dimulai dengan mencari informasi dari jurnal, buku, artikel dan teori yang terkait dengan ChatGPT. Selanjutnya mengumpulkan literatur sebagai referensi untuk memberikan kontribusi penelitian yang baru.

3.2 Pengumpulan Dataset

Pada tahap ini, *dataset* yang dikumpulkan berupa *public dataset* yang bersumber dari *kaggle* dengan alamat url: <https://www.kaggle.com/datasets/charunisa/ChatGPT-sentiment-analysis>. *Dataset* berisi data cuitan (*tweet*) sebanyak 219.294 baris yang dimulai dari tanggal 30/11/2022 hingga 31/12/2022, dengan kata kunci #ChatGPT, dan #chatgpt. *Dataset* terbagi menjadi 3

kelas, yakni bad, good, dan neutral. Dengan masing-masing total label tersendiriya ialah:

- Untuk kelas bad memiliki total 107.796 data *tweet*
- Untuk kelas good memiliki total 56.011 data *tweet*
- Untuk kelas neutral memiliki total 55.487 data *tweet*

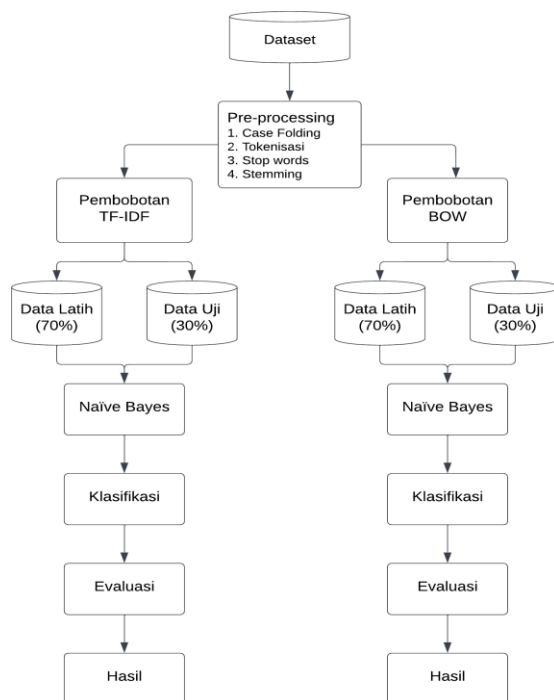
Berikut contoh data *tweet* dari setiap kelas pada *dataset* dapat dilihat pada Tabel 1

Tabel 1. *Dataset* ChatGPT

No	Kelas	Contoh Sentimen
1	<i>Good</i>	ChatGPT is coming out strong, refusing to help me stalk someone but agreeing, providing that someone is W...
2	<i>Neutral</i>	ChatGPT has this curious behaviour where it often starts out by saying it doesn't know something, the...
3	<i>Bad</i>	hmmMmmMM ChatGPT anywhere, anytime delivered by Sentien Audio. It's late but hackathon? https://t.co/...

3.3 Perancangan Sistem

Mengumpulkan *dataset*, maka dilakukan preprocessing untuk mendapatkan data untuk diolah. Adapun tahapan dalam *pre-processing* data yaitucase folding untuk mengubah huruf kapital menjadi huruf kecil, selanjutnyadilakukan tokenisasi untuk memisahkan antar kata pada suatu kalimat, dilanjutkan tahapan *stopwords* untuk menghilangkan kata-kata yang sering digunakan dan bermakna imbuhan. Dan yang terakhir adalah *stemming* yaitu mengubah kata yang mengandung imbuhan.



Gambar 2. Skema Perancangan

Setelah *pre-processing*, maka dilakukan pembobotan terhadap seluruh data menggunakan metode TF-IDF, selanjutnya yaitu melakukan pembagian data latih dan data uji dengan perbandingan rasio 70% data latih: 30% data uji, setelah itu melakukan klasifikasi menggunakan Algoritma *Naïve Bayes* hingga mendapatkan hasil pengujian yang selanjutnya digunakan untuk menghitung nilai *precision*, *recall*, *accuracy*, dan *f1-score*.

Jika hasil dari fitur TF-IDF sudah didapatkan, maka data hasil dari proses *pre-processing* digunakan kembali untuk menerapkan fitur *Bag of Words* dengan pembagian *dataset* yang sama seperti TF-IDF. Setelah itu melakukan klasifikasi menggunakan Algoritma *Naïve Bayes* hingga mendapatkan hasil pengujian yang selanjutnya digunakan untuk menghitung nilai *precision*, *recall*, *accuracy*, dan *f1-score*.

4. Hasil dan Pembahasan

Hasil model *naïve bayes* dalam menganalisis data *tweet* yang diambil dari *kaggle*. Pengujian dilakukan sebanyak 2 kali dengan ekstraksi fitur yang berbeda meliputi :

1. Pengujian 1 kelas data yaitu bad, neutral, dan good menggunakan ekstraksi fitur TF-IDF dengan banyak *term* 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, dan 5000.
2. Pengujian 2 kelas data yaitu bad, neutral, dan good menggunakan ekstraksi fitur BOW dengan banyak *term* 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, dan 5000.

Untuk kelas bad diwakilkan dengan “-“, kelas good diwakilkan dengan “+”, dan kelas neutral diwakilkan dengan “0”. Untuk pembagian data, data dibagi menjadi 70% data latih dan 30% data uji. Hasil dari pengujian akan mendapatkan nilai *precision*, *recall*, *f1-score*, dan *accuracy*.

4.1 Hasil Skenario Pengujian Pertama

Pada skenario pertama, dilakukan pengujian terhadap data hasil dari preprocessing menggunakan ekstraksi fitur TF-IDF dengan banyak *term* 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, dan 5000. Hasil dari pengujian skenario pertama ditampilkan pada Tabel 2.

Tabel 2 Hasil Pengujian Skenario Pertama

Banyak <i>term</i>	Precision			Recall			F1-Score			Accuracy
	-	0	+	-	0	+	-	0	+	
500	52%	45%	68%	86%	27%	47%	62%	34%	56%	54%
1000	51%	42%	62%	79%	30%	43%	62%	35%	51%	52%
1500	52%	37%	59%	72%	33%	41%	60%	35%	48%	49%
2000	54%	40%	56%	69%	39%	40%	61%	39%	47%	50%
2500	53%	38%	53%	63%	38%	43%	57%	38%	48%	48%
3000	53%	36%	51%	59%	38%	41%	56%	37%	46%	47%
3500	53%	36%	48%	57%	39%	41%	55%	38%	45%	46%
4000	53%	37%	47%	55%	40%	41%	54%	38%	44%	46%
4500	53%	38%	48%	55%	41%	42%	54%	39%	45%	46%

Berdasarkan hasil pengujian yang dilakukan didapatkan beberapa kesimpulan akhir diantaranya sebagai berikut :

1. Ekstraksi fitur TF-IDF, ekstraksi fitur BOW, dan metode *naïve bayes* dapat digunakan untuk analisis sentimen chatgpt pada *twitter*.
2. Penelitian ini menerapkan metode *naïve bayes* untuk analisis sentimen. Dengan menggunakan sebanyak 5001 data *tweet* yang dimulai dari tanggal 30/11/2022 hingga 31/12/2022, dengan kata kunci #ChatGPT dan #chatgpt menunjukkan 49% sentimen negatif (Bad), 26% sentimen positif (Good), dan 25% sentimen netral (Neutral) terhadap ChatGPT.
3. Pada pengujian skenario pertama, dimana dilakukan pengujian terhadap data hasil dari preprocessing menggunakan ekstraksi fitur TF-IDF dengan banyak *term* 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, dan 5000 mendapatkan rata-rata nilai akurasi sebesar 48%.
4. Pada pengujian skenario kedua, dimana dilakukan pengujian terhadap data hasil dari preprocessing menggunakan ekstraksi fitur BOW dengan banyak *term* 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, dan 5000 mendapatkan rata-rata nilai akurasi sebesar 59%.
5. Dari semua skenario pengujian pada penelitian ini, didapatkan bahwa ekstraksi fitur BOW mendapatkan nilai akurasi tertinggi sebesar 59%.
6. Dari semua pengujian pada penelitian ini, model kelas bad selalu memiliki kemampuan klasifikasi yang baik dibandingkan kelas lainnya, yaitu dengan nilai rata-rata *f1-score* tertinggi 64%, dibuktikan dengan kemampuan pada model kelas bad kemampuannya dalam mengenali data yaitu dengan nilai rata-rata *recall* yang dimiliki sebesar 59%. Kemudian, kemampuan dalam membedakan data kelas bad dengan yang lainnya, dengan nilai rata-rata *precision* sebesar 71%.
7. Selanjutnya untuk model kelas Neutral selalu memiliki kemampuan klasifikasi yang paling rendah dibandingkan kelas lain, yaitu dengan nilai rata-rata *f1-score* terkecil 37%. Hal tersebut dibuktikan dengan rendahnya nilai rata-rata *precision* 39% dan nilai rata-rata *recall* 37%.

References

- [1] Al-Ayyoub, M., Khamaiseh, A. A., Jararweh, Y., & Al-Kabi, M. N. (2019). A Comprehensive Survey Of Arabic Sentiment Analysis. *Information Processing And Management*, 56(2), 320–342. <https://doi.org/10.1016/j.ipm.2018.07.006>
- [2] Anugerah, A. P., & Sibaroni, Y. (2022). Sentiment Analysis Of Development Jakarta-Bandung High-Speed Train Using Twitter Social Media With BNN Method. *JURIKOM (Jurnal Riset Komputer)*, 9(3), 536. <https://doi.org/10.30865/Jurikom.V9i3.4143>
- [3] Beheshti, A., Ghodrathnama, S., Elahi, M., & Helia, H. (2022). *Social Data Analytics*. Bhatia, S., Sharma, M., & Bhatia, K. K. (2018). Sentiment Analysis And Mining Of Opinions. *Studies In Big Data*, 30, 503–523. https://doi.org/10.1007/978-3-319-60435-0_20
- [4] Faradhillah, N. Y. A. (2016). Analisis Sentimen Terhadap Kinerja Pelayanan Publik Di Kota Surabaya Berdasarkan Klasifikasi Komentar Di Media Sosial Dengan

- Menggunakan Algoritma Naïve Bayes. *Doctoral Dissertation, Institut Teknologi Sepuluh Nopember*, 1–198.
- [5] Fauzi, A. (2023). *Kominfo Ancam Blokir Chatgpt!? Ini Dia Faktanya*. <https://www.finansialku.com/fakta-di-balik-chatgpt-yang-terancam-diblokir-kominfo/>
 - [6] Hamzah, A. (2012). Klasifikasi Teks Dengan Naïve Bayes Classifier (NBC) Untuk Pengelompokan Teks Berita Dan Abstract Akademis. *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) Periode III, 2011*, 269–277. <https://doi.org/10.1109/AICT47866.2019.8981793>
 - [7] Haryanto, & Rudi, M. (2018). *Twitter Menggunakan Metode Naïve Bayes Dan*. 294.
 - Hasanli, H., & Rustamov, S. (2019). Sentiment Analysis Of Azerbaijani Twits Using Logistic Regression, Naive Bayes And SVM. *13th IEEE International Conference On Application Of Information And Communication Technologies, AICT 2019 - Proceedings*. <https://doi.org/10.1109/AICT47866.2019.8981793>
 - [8] Hermawan, & Yoannita. (2018). Komparasi Metode Evaluasi Pada Credit Scoring Data Mining. *Jtksi*, 01(02), 22–25.
 - [9] Hidayatullah, P. (2017). Pengolahan Citra Digital: Teori Dan Aplikasi Nyata. In *Bandung: Informatika*. Informatika Bandung.
 - U. M. Y. (2023). *Chatgpt Dan Fungsinya Bagi Mahasiswa*.
 - [10] <https://library.umi.ac.id/news/detail/500/chatgpt-dan-fungsinya-bagi-mahasiswa>
 - [11] Kusumadewi, S. (2009). Klasifikasi Status Gizi Menggunakan Naive Bayesian Classification. *Commit (Communication And Information Technology) Journal*, 3(1), 6. <https://doi.org/10.21512/Commit.V3i1.506>
 - [12] Lisangan, E. A., Gormantara, A., & Carolus, R. Y. (2022). Implementasi NaiveBayes Pada Analisis Sentimen Opini Masyarakat Di Twitter Terhadap Kondisi New Normal Di Indonesia. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 2(1), 23–32. <https://doi.org/10.24002/Konstelasi.V2i1.5609>
 - [13] Manning, C. D., Raghavan, P., & Schutze, H. (2009). *Introduction To Modern Information Retrieval*. <https://doi.org/10.1108/00242530410565256>
 - [14] Nursaid, F. A. (2023). *Chatgpt Disinyalir Gantikan Pekerjaan Manusia, Ini Penjelasannya*. <https://www.komite.id/2023/02/21/chatgpt-disinyalir-gantikan-pekerjaan-manusia-ini-penjelasannya/>
 - [15] P. Domingos, & Pazzani, M. (1997). On The Optimality Of The Simple Bayesian Classifier Underzero-One Loss. *Machine Learning*, 29, 103–130. <https://link.springer.com/article/10.1023/A:1007413511361>
 - [16] Pang, B., & Lee, L. (2016). Opinion Mining And Sentiment Analysis. *World Journal Of Gastroenterology*, 22(45), 9898–9908. <https://doi.org/10.3748/WJG.V22.I45.9898>
 - [17] Pietikäinen, M., Hadid, A., Zhao, G., & Ahonen, T. (2011). *Computer Vision Using Local Binary Patterns*. Springer. https://doi.org/10.1007/978-0-85729-748-8_14
 - [18] Prakoso, J. P. (2023). *Kelebihan Dan Kekurangan Aplikasi Yang Hebohkan Dunia, Chatgpt*. <https://bisnisindonesia.id/article/kelebihan-dan-kekurangan-aplikasi-yang-hebohkan-dunia-chatgpt>
 - [19] Prasetyo, E. (2015). *Data Mining: Konsep Dan Aplikasi Menggunakan MATLAB*.
 - [20] Qaiser, S., & Ali, R. (2018). Text Mining: Use Of TF-IDF To Examine The Relevance Of Words To Documents. *International Journal Of Computer Applications*, 181(1), 25–29. <https://doi.org/10.5120/Ijca2018917395>
 - [21] Qasem, M., Thulasiram, R., & Thulasiram, P. (2015). Twitter Sentiment Classification

- Using Machine Learning Techniques For Stock Markets. *2015 International Conference On Advances In Computing, Communications And Informatics, ICACCI 2015*, 834–840. <https://doi.org/10.1109/ICACCI.2015.7275714>
- [22] Rashid, J., Adnan Shah, S. M., Irtaza, A., Mahmood, T., Nisar, M. W., Shafiq, M., & Gardezi, A. (2019). Topic Modeling Technique For Text Mining Over Biomedical Text Corpora Through Hybrid Inverse Documents Frequency And Fuzzy K-Means Clustering. *IEEE Access*, 7, 146070–146080. <https://doi.org/10.1109/ACCESS.2019.2944973>
- [23] Rifai, M. F., Jatnika, H., & Valentino, B. (2019). Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist(MOS). *Petir*, 12(2), 131–144. <https://doi.org/10.33322/Petir.V12i2.471>
- [24] Rizal, A. (2023). *Openai Bakal Uji Coba Chatgpt Versi Berbayar, Ini Keunggulannya*. <https://era.id/tekno/118294/Biaya-Langganan-Chatgpt-Plus>
- [25] Ruhimat. (2018). *CHAT GPT (Teknologi Informasi Dan Komunikasi)*. Universitas Djuanda. <https://www.unida.ac.id/artikel/chat-gpt--teknologi-informasi--dan-komunikasi>
- [26] Sabilly, A. F., Adikara, P. P., & Fauzi, M. A. (2019). Analisis Sentimen Pemilihan Presiden 2019 Pada Twitter Menggunakan Metode Maximum Entropy. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(5), 4204–4209.
- [27] Salton, G., & Buckley, C. (1987). Salton, G. And Buckley, C., 1988. Term-Weighting Approaches In Automatic Text Retrieval. 7896.Pdf. In *National Foundation*.
- [28] Samsir, Ambiyar, Verawardina, U., Edi, F., & Watrianthos, R. (2021). Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM. *Jurnal Media Informatika Budidarma*, 5(1), 149. <https://doi.org/10.30865/Mib.V5i1.2604>
- [29] Solichin. (2017). *Mengukur Kinerja Algoritma Klasifikasi Dengan Confusion Matrix*.
- [30] Suparyati, S., & Fathurahman, A. (2022). Analisis Sentimen Dengan Klasifikasi Naïve Bayes Pada Review Hotel Tripadvisor. *Jurnal Ilmiah Informatika (JIF)*, 10(1).
- [31] Sutoyo, E., & Fadlurrahman, M. A. (2020). Penerapan SMOTE Untuk Mengatasi Imbalance Class Dalam Klasifikasi Television Advertisement Performance Rating Menggunakan Artificial Neural Network. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 6(3), 379. <https://doi.org/10.26418/Jp.V6i3.42896>
- [32] Syarli, & Muin, A. A. (2016). Metode Naïve Bayes Untuk Prediksi Kelulusan. *Jurnal Ilmiah Ilmu Komputer*, 2(1), 22–26. <https://media.neliti.com/media/publications/283828-metode-naive-bayes-untuk-prediksi-kelulu-139fcfea.pdf>
- [33] Xhemali, D., J. Hinde, C., & G. Stone, R. (2009). Naive Bayes Vs. Decision Trees Vs. Neural Networks In The Classification Of Training Web Pages. *International Journal Of Computer Science*, 4(1), 16–23. <http://cogprints.org/6708/>
- [34] Yadollahi, A., Shahraki, A. G., & Zaiane, O. R. (2017). Current State Of Text Sentiment Analysis From Opinion To Emotion Mining. *ACM Computing Surveys*, 50(2). <https://doi.org/10.1145/3057270>
- [35] Zhafira, D. F., Rahayudi, B., & Indriati, I. (2021). Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes Dan Pembobotan TF-IDF Berdasarkan Komentar Pada Youtube. *Jurnal Sistem Informasi, Teknologi Informasi, Dan Edukasi Sistem Informasi*, 2(1), 55–63. <https://doi.org/10.25126/Justsi.V2i1.24>
- [36] Zhao, R., & Mao, K. (2015). Fuzzy Bag-Of-Words Model For Document

Representation. *IEEE Transactions On Fuzzy Systems*, 14, 794–804.
<https://doi.org/10.1109/TFUZZ.2017.2690222>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)

