



Tersedia Online : <http://e-journals.unmul.ac.id/>

ADOPSI TEKNOLOGI DAN SISTEM INFORMASI (ATASI)

Alamat Jurnal : <http://e-journals2.unmul.ac.id/index.php/atasi/index>



Analisis Sentimen SEA Games 2023 di Twitter Metode dengan Machine Learning

Akhmad Irsyad ^{1)*}, Raihan Daiva Geralda ²⁾, Reza Wardhana ³⁾

Sistem Informasi, Universitas Mulawarman

E-Mail : akhmadirsyad@ft.unmul.ac.id ¹⁾; daivageralda831@gmail.com ²⁾; rezawardhana@fkti.unmul.ac.id ³⁾;

ARTICLE INFO

Article history:

Received : 24 January 2022

Revised : 02 February 2022

Accepted : 22 February 2022

Available online : xxxxx

Keywords:

Sea games 2023

SVM

Naïve Bayes

Random forest

Sentimen analisis

ABSTRACT

Analisis sentimen adalah sebuah metode yang digunakan untuk menganalisis dan mengidentifikasi polaritas (positif, negatif, atau netral) dari teks atau data yang berkaitan dengan pendapat, opini, atau emosi pengguna. Metode ini banyak digunakan dalam berbagai bidang, termasuk dalam analisis sentimen pada data sosial media. Salah satu platform media sosial yang sangat populer adalah Twitter. Salah satu ajang olahraga terbesar di Asia adalah SEA Games, yang diadakan setiap dua tahun sekali. SEA Games, banyak warga Indonesia yang mengungkapkan pendapat, dukungan, dan emosi mereka terkait SEA Games 2023 di Twitter. Dengan menggunakan metode supervised learning dapat memberikan wawasan yang berharga tentang bagaimana masyarakat Indonesia merespons dan bereaksi terhadap perhelatan olahraga yang penting ini. Hasil analisis ini dapat membantu pihak terkait, termasuk penyelenggara dan sponsor SEA Games 2023, dalam memahami sentimen publik, mengevaluasi kinerja, dan membuat keputusan yang lebih baik dalam rangka menyelenggarakan ajang olahraga yang sukses. Hasil dari uji coba yang dilakukan metode SVM memiliki kinerja terbaik dengan F-1 score 61,53 %.

ABSTRAK

Sentiment analysis is a method used to analyze and identify the polarity (positive, negative, or neutral) of text or data related to a user's thoughts, opinions, or emotions. This method is widely used in various fields, including sentiment analysis on social media data. One very popular social media platform is Twitter. One of the biggest sporting events in Asia is the SEA Games, which are held every two years. SEA Games, many Indonesians expressed their opinions, support and emotions regarding the 2023 SEA Games on Twitter. Using supervised learning methods can provide valuable insight into how Indonesian society responds and reacts to this important sporting event. The results of this analysis can help related parties, including organizers and sponsors of the 2023 SEA Games, in understanding public sentiment, evaluating performance, and making better decisions in order to organize a successful sporting event. The results of the trials carried out by the SVM method had the best performance with an F-1 score of 61.53%.

Kata Kunci :

Maksimal

Sea games 2023

SVM

Naïve Bayes

Random forest

Sentiment analysis

APA style in citing this article:

Irsyad, A., & Geralda, R. D. (2023).

Analisis Sentimen SEA Games 2023 di

Twitter Metode dengan Machine

Learning. *Adopsi Teknologi Dan Sistem*

Informasi (ATASI), 2(2), 126–131.

<https://doi.org/10.30872/atasi.v2i2.1138>

2022 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

*) Correspondenting Author

<https://doi.org/10.30872/atasi.v2i2.1138>

2023 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

1. PENDAHULUAN

Analisis sentimen adalah sebuah metode yang digunakan untuk menganalisis dan mengidentifikasi polaritas (positif, negatif, atau netral) dari teks atau data yang berkaitan dengan pendapat, opini, atau emosi pengguna. Metode ini banyak digunakan dalam berbagai bidang, termasuk dalam analisis sentimen pada data sosial media (Winahyu & Suharjo, 2021). Salah satu platform media sosial yang sangat populer adalah Twitter. Twitter memberikan akses luas kepada pengguna untuk berbagi pendapat dan pengalaman mereka dengan cepat dan dalam jumlah besar (Normah et al., 2022). Pengguna Twitter secara aktif mengungkapkan perasaan mereka tentang berbagai topik, termasuk peristiwa dan ajang olahraga.

Salah satu ajang olahraga terbesar di Asia adalah SEA Games, yang diadakan setiap dua tahun sekali. SEA Games 2023 merupakan perhelatan yang sangat dinantikan oleh warga Indonesia, karena ajang ini akan diselenggarakan di Indonesia. Dalam menyikapi perhelatan ini, banyak warga Indonesia yang akan mengungkapkan pendapat, dukungan, dan emosi mereka terkait SEA Games 2023 di Twitter. Dalam konteks tersebut, metode supervised learning dapat digunakan untuk menganalisis sentimen warga Indonesia terhadap SEA Games 2023 di Twitter. Dengan menggunakan dataset yang berisi teks-teks tweet yang dikategorikan sebagai positif, negatif, atau netral, model supervised learning dapat dilatih untuk mengklasifikasikan sentimen dari tweet-tweet tersebut.

Analisis sentimen warga Indonesia terhadap SEA Games 2023 di Twitter menggunakan metode supervised learning dapat memberikan wawasan yang berharga tentang bagaimana masyarakat Indonesia merespons dan bereaksi terhadap perhelatan olahraga yang penting ini. Hasil analisis ini dapat membantu pihak terkait, termasuk penyelenggara dan sponsor SEA Games 2023, dalam memahami sentimen publik, mengevaluasi kinerja, dan membuat keputusan yang lebih baik dalam rangka menyelenggarakan ajang olahraga yang sukses.

2. TINJAUAN PUSAKA

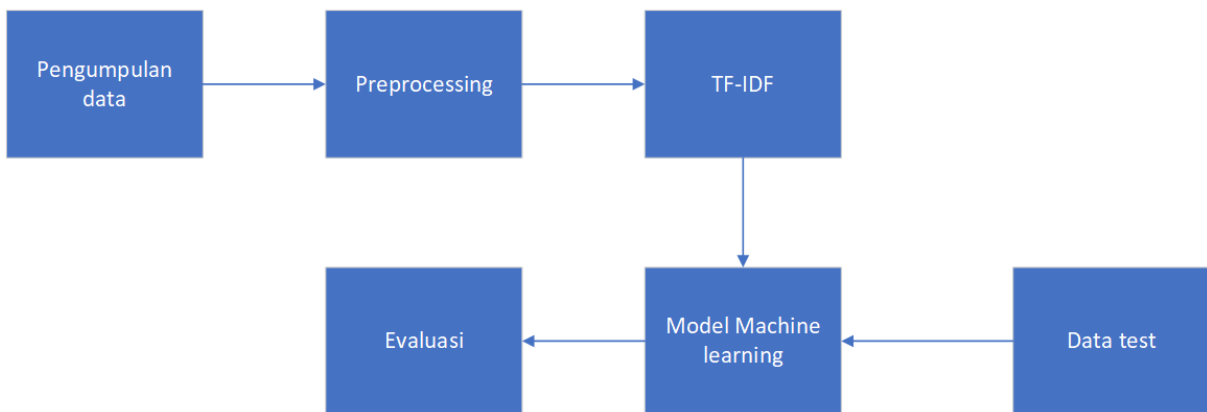
Saat ini telah banyak penelitian sentimen analisis dengan berbagai tema, mulai dari masalah politik, ekonomi, dan bahkan masalah olahraga, salah satunya adalah penelitian yang dilakukan oleh Winahyu (2021), penelitian yang dilakukan dengan menggunakan metode *text classification* yang telah banyak secara umum digunakan adalah algoritma Multinomial Naive Bayes. Metode ini memanfaatkan teorema probabilitas yang mana teorema bayes dan fungsionalitas data mining yaitu klasifikasi naive bayes. Penelitian ini berfokus pada opini pengguna jejaring sosial twitter dengan kata kunci aplikasi web data ini di ambil dari API twitter dengan jumlah item data 1000 selanjutnya klasifikasi opini pengguna dan preproses data. Dari hasil pengujian Dalam mengevaluasi performance algoritma kita menggunakan acuan Confusion Matrix dan merepresentasikan prediksi dan kondisi sebenarnya (aktual) dari data yang dihasilkan oleh algoritma kita bisa menentukan Accuracy 69% , Precision merupakan rasio jumlah yang benar diprediksi pada label kelas positif 0,67 , netral 0, negatif 1 dan recall merupakan rasio jumlah yang diprediksi benar terhadap keseluruhan data yang benar yaitu kelas positif 1, netral 0, negatif 0,14 dan f1-score merupakan rata-rata dari *precision* dan *recall* yaitu kelas positif 0,81, netral 0, negatif 0,25 (Winahyu & Suharjo, 2021).

Arsi (2021) melakukan penelitian dengan metode Support Vector Machine (SVM) untuk diterapkan pada tweets topik pemindahan ibu kota Indonesia untuk tujuan klasifikasi kelas sentimen pada media sosial twitter. Teknis klasifikasi dilakukan dengan cara mengklasifikasikan menjadi 2 kelas yakni positif dan negatif. Berdasarkan hasil pengujian yang dilakukan terhadap tweets sentimen pemindahan ibu kota dari media sosial twitter sebanyak 1.236 tweets (404 positif dan 832 negatif) menggunakan SVM diperoleh akurasi =96,68%, precision=95.82%, recall=94.04% dan AUC = 0,979 (Arsi & Waluyo, 2021).

Pane (2022) dalam penelitian ini menggunakan data yang bersumber dari cuitan para pengguna media sosial Twitter di wilayah Jakarta terhadap kebijakan pemerintah yaitu PPKM dalam menghadapi pandemi COVID-19. Dalam penelitian ini kami menggunakan pendekatan Machine Learning yaitu LSTM. Pemodelan ini menghasilkan klasifikasi sentimen positif dan negatif. Dataset yang digunakan sebanyak 3000 tweet dengan periode waktu September - November 2021. Pada tahap preprocessing data yang siap digunakan untuk pemodelan sebanyak 2176. Hasil penelitian ini mendapatkan akurasi sebesar 0,943. Sehingga model yang kami usulkan yaitu LSTM berhasil mengklasifikasikan sentimen yang memuaskan dengan angka positif sebesar 92% dan negatif 8% dari 2176 sentimen, sehingga dapat disimpulkan bahwa kebijakan PPKM yang diterapkan oleh pemerintah Indonesia di wilayah DKI Jakarta daerah tersebut dikatakan efektif atau positif (Pane & Ramdan, 2022).

Hikmawan et al., (2020) melakukan penelitian dengan menggunakan kata kunci “Jokowi” dan “Covid” untuk mengumpulkan data di twitter, hal ini memungkinkan kita menjadikan tweet didalamnya untuk dijadikan dataset. Dalam data mining untuk sentimen analisis, dilakukan teknik seperti transform, tokenize, stemming, classification, dan lain-lain sangat berpengaruh pada akurasi. Gata Framework digunakan untuk preprocessing, dan Rapidminer juga digunakan untuk menganalisa dan membandingkan tiga metode klasifikasi yaitu Naive Bayes, Support Vector Machine, dan k-NN. Dan dihasilkan nilai terbaik yaitu Support Vector Machine dengan accuracy 84.58%, precision 82.14% dan recall 85.82% (Hikmawan et al., 2020). Dari beberapa penelitian sebelumnya mengenai sentimen analisis diketahui bahwa dengan menggunakan machine learning dapat diperoleh hasil klasifikasi yang baik, dan diketahui bahwa SVM adalah metode yang sering digunakan oleh karena itu pada penelitian ini dilakukan dengan SVM dan beberapa metode machine learning yang lain untuk dibandingkan.

3. METODE PENELITIAN



Gambar 1. Alur Penelitian

Penelitian ini dimulai dengan mengumpulkan data lalu dilakukan preprocessing, preprocessing dilakukan dengan membersihkan data seperti melakukan lowercase, menghilangkan spesial karakter, menghilangkan emotikon, dari data yang telah dibersihkan lalu dilakukan pembagian data menjadi data *train* dan data *test*. Data *train* lalu dilakukan ekstraksi fitur dengan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Data *test* digunakan untuk melatih model machine learning, hasil dari pelatihan lalu dilakukan uji coba dengan data *test* untuk mengetahui kinerja yang dibuat apakah sudah cukup baik.

A. Preprocessing

Proses mempersiapkan teks dokumen atau dataset mentah disebut juga dengan proses *text preprocessing*. *Text preprocessing* berfungsi untuk mengubah data teks yang tidak terstruktur atau sembarang menjadi data yang terstruktur (Alita & Isnain, 2020). Pada penelitian ini proses yang dilakukan dalam tahapan preprocessing adalah sebagai berikut :

1. Case Folding Case Folding adalah proses penyamaan case dalam sebuah dokumen. Hal ini dilakukan untuk mempermudah pencarian. Tidak semua dokumen teks konsisten dalam penggunaan huruf kapital. Oleh karena itu peran case folding dibutuhkan dalam mengkonversi keseluruhan teks dalam dokumen menjadi suatu bentuk standar (dalam hal ini huruf kecil atau lowercase).
2. Tokenizing Tokenizing adalah proses pemotongan sebuah dokumen menjadi bagian-bagian, yang disebut dengan token. Pada saat bersamaan tokenizing juga berfungsi untuk membuang beberapa karakter tertentu yang dianggap sebagai tanda baca.
3. Stopword Removal Stopword removal adalah proses penghilangan kata-kata yang tidak berkontribusi banyak pada isi dokumen Kata-kata yang termasuk ke dalam stopwords dihilangkan karena memberikan pengaruh yang tidak baik dalam proses text mining seperti kata-kata “and”, “I”, “you”, “with”, “she”, “he”, dan lain-lain.

B. TF-IDF

Term Frequency-Inverse Document Frequency merupakan salah satu metode pembobotan yang menggabungkan antara term frequency (TF) dan inverse document frequency (IDF) atau pembobotan pada setiap kata dalam setiap dokumen teks (Assidyk et al., 2020). Term frequency (TF) adalah kemunculan sebuah kata dalam suatu dokumen sedangkan inverse document frequency (IDF) adalah jumlah seluruh dokumen yang mengandung kata tertentu. Maka metode TF-IDF ini dapat menghitung total bobot dari kata dalam sebuah dokumen dengan persamaan.

$$TF(t, d) = \frac{\text{total jumlah term dalam dokumen } d}{\text{jumlah kemunculan term } t \text{ dalam dokumen } d} \quad (1)$$

Term Frequency adalah rasio jumlah kemunculan term *t* dalam suatu dokumen *d* dibagi dengan total jumlah term dalam dokumen *d*. *Inverse Document Frequency* adalah logaritma dari rasio total jumlah dokumen dalam koleksi *D* dibagi oleh jumlah dokumen yang mengandung term *t*, ditambah satu.

$$IDF(t, D) = \log \frac{\text{total jumlah dokumen dalam koleksi } D}{\text{jumlah dokumen yang mengandung term } t+1} \quad (2)$$

TF-IDF adalah hasil perkalian antara Term Frequency (TF) dan Inverse Document Frequency (IDF). Dimana *t* adalah term atau kata kunci tertentu, *d* adalah dokumen tertentu, *D* adalah koleksi seluruh dokumen

$$TF - IDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (3)$$

C. Random forest

Random forest classifier merupakan metode klasifikasi yang terdiri dari kumpulan pohon keputusan yang nantinya akan dijadikan vote untuk mendapatkan hasil terakhir dari pendeteksian sarkasme dengan pendukung berupa data latih dan fitur acak yang independen dengan fitur yang berbeda-beda (Alita & Isnain, 2020). Pohon keputusan dibuat dengan menentukan node akar dan berakhir dengan beberapa node daun untuk mendapatkan hasil akhir

D. SVM

SVM adalah teknik untuk memprediksi klasifikasi dan regresi. SVM pertama kali diperkenalkan oleh Vapnik pada tahun 1992 sebagai seperangkat konsep unggul yang di bidang pattern recognition (Gifari et al., 2022). SVM adalah algoritma pembelajaran mesin yang bekerja berdasarkan prinsip Structural Risk Minimization (SRM) dengan tujuan untuk menemukan hyperplane optimal yang memisahkan dua kelas dalam ruang input.

E. Naïve Bayes

Naïve Bayes Classifier merupakan model probabilistik yang digunakan untuk proses klasifikasi berdasarkan teorema Bayes. Naïve Bayes Classifier dikenal sederhana namun sangat efisien (Yutika et al., 2021), dan memiliki asumsi yg sangat kuat terhadap independensi dari masing-masing kondisi, terutama saat memiliki jumlah data train yang sedikit. Pada penelitian ini menggunakan algoritma Complement Naïve Bayes yang merupakan adaptasi dari algoritma Multinomial Naïve Bayes dan dirancang untuk mengatasi dataset yang tidak seimbang (Yutika et al., 2021). Data yang tidak seimbang adalah ketika jumlah suatu kelas lebih tinggi dibandingkan kelas lainnya, sehingga dapat mempengaruhi perhitungan probabilitas kelas lain.

F. Evaluasi

Evaluasi performa dilakukan pada hasil eksperimen sistem analisis sentimen dan pada hasil analisis sentimen responden. Hasil dari Precision, Recall, F1-Score menghasilkan nilai dari 0 sampai 1 dan hasil yang lebih baik adalah nilai yang mendekati 1 (Gifari et al., 2022). Evaluasi bertujuan untuk menilai hasil uji coba sistem yang dibuat apakah telah sesuai antara hasil sistem analisis sentimen dengan hasil sebenarnya. Confusion matrix adalah tabel yang digunakan untuk menganalisis seberapa baik keakuratan sebuah metode klasifikasi untuk memprediksi kelas suatu data. Ada beberapaparameter yang bisa digunakan untuk mengevaluasi kinerja klasifikasi. Penelitian ini menggunakan Accuracy Precision, Recall, F1-Score sebagai parameter evaluasi untuk mengetahui seberapa baik sistem yang dibangun dan hasil yang diberikan oleh sistem.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{FN} + \text{TN}) \quad (4)$$

$$\text{Precision(P)} = \text{TP} / (\text{TP} + \text{FP}) \quad (5)$$

$$\text{Recall(R)} = \text{TP} / (\text{TP} + \text{FN}) \quad (6)$$

$$\text{F-1 Score} = 2 * (\text{R} * \text{P}) / (\text{R} + \text{P}) \quad (7)$$

4. HASIL DAN PEMBAHASAN

A. Dataset

Dalam penelitian ini, saya menggunakan teknik crawling data untuk mengumpulkan data dari Twitter yang berkaitan dengan SEA Games 2023. crawling data tweet yang berbahasa Indonesia dan memiliki rentang waktu dari 1 April 2023 hingga 19 Mei 2023. Dengan memilih bahasa tertentu sehingga penelitian dapat fokus pada tweet-tweet yang relevan dengan SEA Games 2023 dan mengatasi data yang dikumpulkan. Dalam proses labelling dilakukan dengan menggunakan metode terjemahan menggunakan Google Translate dan library TextBlob.

Pertama-tama menterjemahkan tweet-tweet yang telah kumpulan dalam bahasa Indonesia ke dalam bahasa Inggris. Hal ini dilakukan menggunakan Google Translate API atau library seperti googletrans. Setelah tweet-tweet diterjemahkan ke dalam bahasa Inggris, saya menggunakan library TextBlob untuk melakukan labelling. Data lalu dilakukan pembagian menjadi data pelatihan dan data uji. Proporsi pembagian data pelatihan dan data uji ditentukan dengan test_size=0.1, yang berarti 10% data digunakan sebagai data uji.

B. Preprocessing

Pre-processing merupakan tahap pengolahan data untuk membersihkan kata-kata yang tidak perlu atau kata-kata yang tidak memiliki makna. Adapun urutan dari tahapan *pre-processing* yaitu sebagai berikut :

1. Case Folding

Pada tahap ini, semua huruf akan menjadi huruf kecil (*lowercase*). Contoh penerapan pada tahap *case folding* dapat dilihat pada tabel 3.1.

Tabel 3.1. Contoh Proses Penerapan *Case Folding*

| Sebelum <i>Case Folding</i> | Sesudah <i>Case Folding</i> |
|--|--|
| Manajer Thailand Mundur Buntut Kerusakan dengan Indonesia di Final SEA Games 2023 https://t.co/ZDr6yVmycX | manajer thailand mundur buntut kerusakan dengan indonesia di final sea games 2023 https://t.co/zdr6yvmycX |

2. *Cleansing*

Cleansing merupakan tahapan penghapusan karakter selain huruf *alfabet*. Karakter yang dihapus berupa tanda baca seperti hastag (#), koma (,), *username* (@username), *emoticon*, dan *url*. Contoh penerapan tahap *cleansing* dapat dilihat pada tabel 3.2.

Tabel 3.2. Contoh Proses Penerapan *Cleansing*

| Sebelum <i>Cleansing</i> | Sesudah <i>Cleansing</i> |
|--|---|
| manajer thailand mundur buntut kerusuhan dengan indonesia di final sea games 2023 https://t.co/zdr6yvmycx | manajer thailand mundur buntut kerusuhan dengan indonesia di final sea games 2023 |

3. *Tokenizing*

Tokenizing merupakan tahap pemisahan sebuah kalimat menjadi potongan tunggal kata. Contoh penerapan pada tahap *tokenizing* dapat dilihat pada tabel 3.4.

Tabel 3.3. Contoh Proses Penerapan *Tokenizing*

| Sebelum <i>Tokenizing</i> | Sesudah <i>Tokenizing</i> |
|---|--|
| manajer thailand mundur buntut kerusuhan dengan indonesia di final sea games 2023 | manajer,thailand,mundur,buntut,rusuh,indonesia,final,sea,games |

C. TF-IDF

Setelah data bersih dan diberi labelling, tahap selanjutnya adalah pembuatan fitur untuk mempermudah proses klasifikasi. Pada tahap pembuatan fitur yaitu dilakukannya proses mengubah fitur teks menjadi visualisasi vector dan pembobotan kata menggunakan TF-IDF. Proses TF-IDF dilakukan dengan library Sklearn, dan karena data yang imbalance dilakukan oversampling dengan metode SMOTE dan untuk implementasinya dilakukan dengan library imblearn.

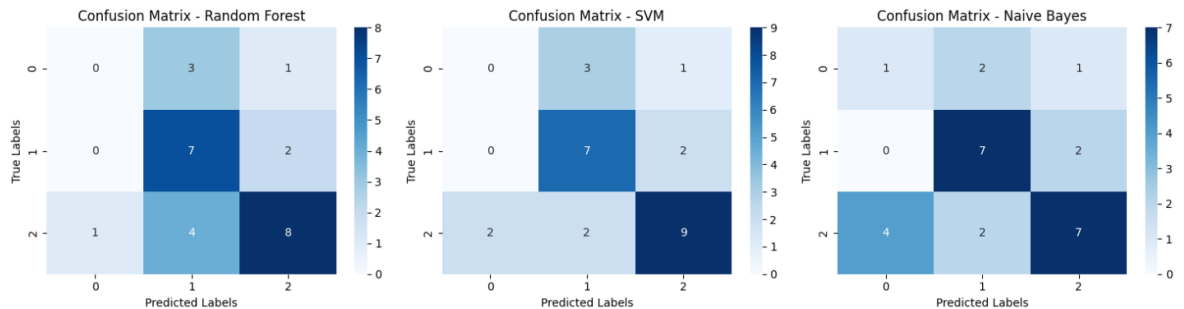
D. Model machine learning

Pada tahap evaluasi dilakukan perhitungan berbagai metrik performa untuk masing-masing model klasifikasi yang telah digunakan. Metrik yang dihitung meliputi akurasi (accuracy), presisi (precision), recall, dan skor F1 (F1 score). Akurasi mengukur sejauh mana model dapat memprediksi dengan benar terhadap seluruh data uji. Presisi mengukur tingkat ketepatan model dalam memprediksi kelas positif, recall mengukur tingkat kelengkapan model dalam mengidentifikasi kelas positif, dan skor F1 merupakan gabungan antara presisi dan recall. Selain itu, penelitian ini menggunakan matriks konfusi (confusion matrix) untuk masing-masing model, yang menunjukkan jumlah prediksi yang benar dan salah untuk setiap kelas sentimen. Matriks konfusi ini kemudian divisualisasikan menggunakan heatmap.

Tabel 1. Hasil kinerja model machine learning.

| No | Model | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|----|---------------|--------------|---------------|------------|--------------|
| 1 | SVM | 61,53 | 61,53 | 61,53 | 61,53 |
| 2 | Naïve Bayes | 57,69 | 57,69 | 57,69 | 57,69 |
| 3 | Random Forest | 57,69 | 57,69 | 57,69 | 57,69 |

Berdasarkan hasil analisis data yang dilakukan, dapat disimpulkan bahwa algoritma SVM menunjukkan kinerja yang cukup baik dibandingkan dengan algoritma Random Forest dan Multinomial Naive Bayes. Dalam analisis ini, SVM memiliki nilai 0.615385. Hasil ini menunjukkan bahwa SVM dapat melakukan klasifikasi dengan tingkat keakuratan yang relatif tinggi. Dalam perbandingan dengan Random Forest dan Multinomial Naive Bayes, SVM berhasil memberikan performa yang lebih baik dalam hal keseluruhan. Oleh karena itu, dalam konteks analisis data yang dilakukan, SVM dapat dianggap sebagai pilihan yang efektif dan dapat diandalkan untuk tugas klasifikasi



Gambar 2. Hasil Coefficient matrices machine learning

5. KESIMPULAN

Hasil analisis dilakukan untuk mendapatkan tujuan penelitian. Dari hasil analisis data ini juga untuk menjawab pertanyaan penelitian dan yang dilakukan selama penelitian apakah terbukti berdasarkan hasil penelitian yang dilakukan. Berdasarkan tujuan dan pertanyaan penelitian maka hasil analisis membangun model klasifikasi sentimen dengan metode SVM. Model yang telah dibangun dilakukan proses pengujian (testing) dengan menggunakan data testing untuk mengetahui seberapa akurat sistem dalam melakukan prediksi sentimen analisis serta setiap kelas sentimen divisualisasikan berdasarkan kata yang sering muncul sebagai bahan informasi penting.

6. DAFTAR PUSTAKA

- Alita, D., & Isnain, A. R. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi*, 8(2), 50–58. <https://doi.org/10.23960/komputasi.v8i2.2615>
- Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Peminahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(1), 147–156. <https://doi.org/10.25126/JTIK.0813944>
- Assidyk, A. N., Setiawan, E. B., & Kurniawan, I. (2020). Analisis Perbandingan Pembobotan TF-IDF dan TF-RF pada Trending Topic di Twitter dengan Menggunakan Klasifikasi K-Nearest Neighbor. *E-Proceeding of Engineering*, 7(2), 7773–7781.
- Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *Journal of Information Technology*, 2(1), 36–40. <https://doi.org/10.46229/jifotech.v2i1.330>
- Hikmawan, S., Pardamean, A., & Khasanah, S. N. (2020). Sentimen Analisis Publik Terhadap Joko Widodo terhadap wabah Covid-19 menggunakan Metode Machine Learning. *Jurnal Kajian Ilmiah*, 20(2), 167–176. <https://doi.org/10.31599/jki.v20i2.117>
- Normah, Rifai, B., Vambudi, S., & Maulana, R. (2022). Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE. *Jurnal Teknik Komputer AMIK BSI*, 8(2), 174–180. <https://doi.org/10.31294/jtk.v4i2>
- Pane, S. F., & Ramdan, J. (2022). Pemodelan Machine Learning : Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Menggunakan Data Twitter. *Jurnal Sistem Cerdas*, 5(1), 12–20. <https://doi.org/10.37396/jsc.v5i1.191>
- Winahyu, J., & Suharjo, I. (2021). Aplikasi Web Analisis Sentimen Dengan Algoritma Multinomial Naïve Bayes. *Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika (KARMAPATI)*, 10(2), 206. <https://doi.org/10.23887/karmapati.v10i2.36609>
- Yutika, C. H., Adiwijaya, A., & Faraby, S. Al. (2021). Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes. *Jurnal Media Informatika Budidarma*, 5(2), 422. <https://doi.org/10.30865/mib.v5i2.2845>