

Analisa Sentimen Pada Pembatalan Piala Dunia U20 di Indonesia Menggunakan Metode Support Vector Machine

Adelia Damayanti^{1*}, Arifah Khairina Maulidiyah²

^{1,2}Program Studi Matematika, UIN Sunan Ampel Surabaya, Indonesia

¹adeliadamayanti869@gmail.com*; ²arifahmaulidiyah47815@gmail.com

ABSTRAK

Sebelum Piala Dunia U20, perdebatan tentang penolakan kedatangan Timnas Israel dan konflik Israel-Palestina sangat menarik perhatian publik. Hal ini menimbulkan suara pro dan kontra di masyarakat Indonesia. Banyak komentar masyarakat, baik positif maupun negatif, telah disampaikan tentang masalah ini di media sosial, terutama Twitter. Metode Support Vector Machine (SVM) digunakan dalam penelitian ini untuk menganalisis sentimen masyarakat terkait pembatalan Piala Dunia U20 di Indonesia. Data yang digunakan dalam penelitian ini diperoleh dari platform media sosial twitter melalui web hawkeye.info dengan hastag #PialaDuniaU20. Jumlah 2966 tweet yang akan dianalisis dalam penelitian ini dihasilkan dari proses tersebut. Data yang telah diperoleh kemudian diproses untuk menghilangkan noise dan dilakukan ekstraksi fitur menggunakan teknik seperti tokenisasi, stemming, dan penghapusan stop word. Selanjutnya, hasil analisis sentimen akan diklasifikasikan menjadi positif atau negatif berdasarkan probabilitas prediksi yang diberikan oleh model SVM. Analisis klasifikasi menggunakan metode SVM menghasilkan akurasi 74,4% dengan nilai presisi 74%, recall 74%, dan f1-score 73%.

Kata kunci : Analisa Sentimen; Pembatalan Piala Dunia U20, Support Vector Machine (SVM).

This is an open-access article under the [CC-BY-SA](#) license.



Histori Naskah

Naskah di-Kirim : 12 Juli 2023

Naskah di-Terima: 01 September 2023

Naskah di-Publikasi : 07 September 2023

I. PENDAHULUAN

Masyarakat Indonesia sangat menyukai sepak bola. Berdasarkan survei perusahaan riset multinasional Ipsos tahun 2022, Indonesia menjadi negara nomor satu yang memiliki penggemar sepak bola terbanyak di dunia [1]. 69% responden Indonesia adalah penggemar sepak bola [1]. Ini menunjukkan bahwa sepak bola sudah sangat melekat di masyarakat Indonesia, yang menimbulkan ekspektasi besar untuk masa depan sepak bola di Indonesia.

Penelitian ini berfokus pada banyaknya pemberitaan kontroversi tentang Piala Dunia U20 yang akan diadakan di Indonesia. Hal itu disebabkan karena adanya pro-kontra mengenai kedatangan Timnas Israel yang lulus kualifikasi Piala Dunia U20 ke Indonesia. Terdapat beberapa pihak yang menolak datangnya Timnas Israel ke Indonesia dengan berbagai alasan, salah satunya karena tindakan Israel terhadap Palestina yang sampai sekarang masih di bawah jajahan Israel. Polemik tersebut membuat *Federation Internationale de Football Association* (FIFA) menarik Indonesia dari status tuan rumah Piala Dunia U20 yang menyebabkan Timnas Indonesia kehilangan *privilege* untuk tampil di ajang Piala Dunia U20.

Berdasarkan permasalahan tersebut banyak menjadi polemik tanggapan positif ataupun negatif masyarakat. Masyarakat ramai-ramai menyampaikan opininya di berbagai media sosial salah satunya *Twitter*. Pengguna *Twitter* mengunggah opini mereka melalui Tweet dengan hastag #PialaDuniaU20, #PSSI, #FIFA, dan sebagainya berkaitan dengan problematika tersebut. Bagian dari text mining adalah analisis sentimen, yang mengekstrak informasi tentang sentimen positif, netral, atau negatif dari data teks. Peneliti akan menganalisis untuk mengetahui gagasan positif dan negatif apa yang terjadi di masyarakat. Masyarakat yang menggunakan media sosial menggunakan analisis sentimen untuk membuat keputusan pribadi [2]. Support Vector Machines (SVM) adalah salah satu metode klasifikasi supervised learning yang paling umum digunakan dalam penelitian ini [3]. SVM memungkinkan kita untuk secara akurat menempatkan tanggapan orang ke dalam kelas positif atau negatif [4].

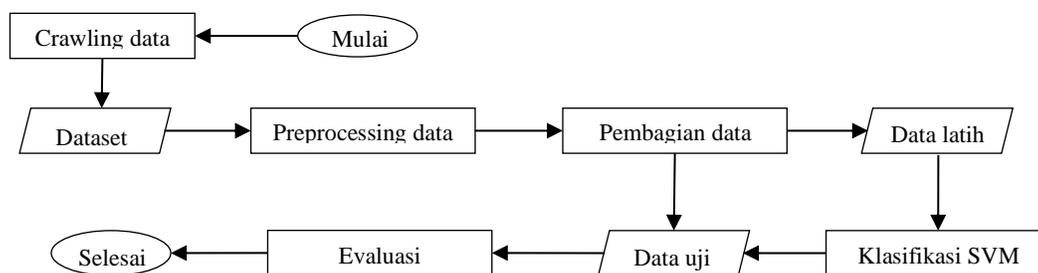
Penelitian sebelumnya telah membuktikan kelebihan metode SVM dalam analisis sentimen. Analisis sentimen mengenai kinerja PSSI dengan menggunakan perbandingan Metode Decision Tree dan SVM dengan TF-IDF pada Instagram mampu

mendapatkan nilai akurasi 94.36% [5]. Hasil penelitian [6] mengulas penggunaan metode SVM menghasilkan akurasi yang sangat baik yaitu 98,34% untuk analisa sentimen oleh Masyarakat pengguna kartu prakerja dengan dataset pada Twitter. Peneliti [7] menggunakan metode SVM untuk melakukan analisis sentimen pada layanan Indihome dengan dataset Twitter; metode SVM berhasil mencapai akurasi 87%. Sementara peneliti [8] membandingkan metode Naive Bayes Classifier (NBC) dan SVM untuk melakukan analisis sentimen opini pembangunan infrastruktur di media sosial Twitter, dengan nilai akurasi NBC 85,6 persen dan SVM 86,2 persen. Perbandingan metode yang sama juga dilakukan peneliti [9] dengan menghasilkan nilai akurasi 84% dengan metode SVM dan 82,8% dengan metode NBC pada analisis sentimen ulasan produk Amazhone. Berdasarkan studi literatur tersebut dapat disimpulkan bahwa SVM menghasilkan akurasi yang paling tinggi sehingga sangat baik digunakan dalam klasifikasi.

Tujuan penelitian ini dilakukan untuk melakukan analisis sentimen terhadap tanggapan masyarakat di Twitter mengenai polemik Piala Dunia U20 dengan metode SVM dengan tujuan untuk mengetahui tingkat akurasi metode SVM dan sebagai acuan untuk pengambilan keputusan dan pemahaman terhadap persepsi publik terkait isu-isu sensitif di masa depan.

II. METODE

Dengan menggunakan data yang dikumpulkan dari respon in-text, analisis sentimen mencoba untuk mendeteksi kecenderungan umum dari opini seseorang terhadap sesuatu, misalnya apakah opini tersebut cenderung positif atau negatif [10]. Berdasarkan informasi yang disajikan pada Gambar 1, langkah pertama dalam proses penelitian adalah data crawling. Selain itu, data mentah akan diproses terlebih dahulu untuk mengurangi jumlah noise yang ada pada data. Pada tahap ini, ada empat alur proses yang berbeda yang berperan: pembersihan, pelipatan huruf besar-kecil, tokenisasi, penyaringan, dan stopwords. Pada tahap ini, data yang telah dibersihkan akan dipisahkan menjadi data uji dan data latih sebagai persiapan untuk pemeriksaan lebih lanjut. Matriks kebingungan digunakan pada langkah terakhir dari proses ini untuk mengevaluasi model dan memperkirakan tingkat akurasi yang dicapai oleh model.



Gambar 1. Alur Proses Penelitian

A. Preprocessing Data

Preprocessing data adalah proses menghilangkan dan mengurangi *noisy* dari kumpulan data untuk mendapatkan hasil yang paling akurat. Item ini harus dihapus karena tidak memiliki nilai informasi untuk klasifikasi sentimen [11]. *Cleansing, case folding, tokenizing, filtering, dan stopword removal* adalah proses *pre-processing*. Proses membersihkan melibatkan penghapusan simbol seperti nama pengguna, tautan URL, *emoticon, hastage, dan retweet* [6]. *Case folding* adalah proses menggabungkan semua huruf dalam teks menjadi huruf kecil, sedangkan *tokenizing* adalah proses membagi dokumen menjadi bagian-bagian kecil seperti paragraf, kata, subbab, dan bab. Kedua proses ini masing-masing disebut sebagai *folding* dan *tokenizing*. Salah satu cara untuk menghilangkan kata-kata yang sering digunakan tetapi tidak memiliki arti yang sebenarnya adalah dengan menghilangkan kata-kata pengisi yang tidak berguna seperti "atau", "dan", "ke", dan sebagainya. Penghapusan kata-kata yang tidak berarti dapat dilakukan dengan proses yang disebut "penghapusan kata terakhir" [12].

B. Pembobotan

Nilai kata dihitung dengan menggunakan pembobotan dan berdasarkan isi dari paper atau tweet [13]. Term Frequency (TF) adalah jumlah term yang ada pada sebuah tweet, sedangkan Inverse Document Frequency (IDF) menentukan seberapa baik kata-kata yang dicari sesuai dengan kata kunci yang dibutuhkan [13]. Inverse Document Frequency disingkat menjadi IDF. Algoritma TF-IDF direpresentasikan oleh persamaan (1).

$$W_{j,i} = \frac{n_{j,i}}{\sum_k n_{k,i}} \cdot \log_2 \frac{D}{d_j} \quad (1)$$

Dimana variabel $W_{j,i}$ merepresentasikan pembobotan TF-IDF untuk term ke- j pada dokumen ke- i , variabel $n_{j,i}$ merepresentasikan jumlah kemunculan term ke- j pada dokumen ke- i , variabel $n_{k,i}$ merepresentasikan jumlah kemunculan seluruh term pada dokumen ke- i , variabel D merepresentasikan jumlah dokumen yang dihasilkan, dan variabel d_j merepresentasikan jumlah dokumen yang mengandung term ke- j .

C. Support Vector Machine (SVM)

Metode klasifikasi SVM memilih fungsi pemisah yang paling efektif dari semua fungsi pemisah yang tersedia untuk membedakan dua kelas [6]. Metode SVM menggunakan ide utama dalam klasifikasi, yaitu memilih hyperplane terbaik untuk memisahkan dua kelas [14]. Dia melakukannya dengan menemukan hyperplane optimal, atau hyperplane dengan margin terbesar. Margin adalah istilah yang digunakan untuk menggambarkan jarak antara hyperplane dan titik data terkecil [15], dengan tahapan untuk mencapai hasil yaitu [7]:

- a. Lakukan analisa setiap tweet untuk menentukan istilah yang paling sering muncul.
- b. Tetapkan parameter berikut sebagai titik awal analisis, $\alpha = 0.5$, $C = 1$, $\lambda = 0.5$, $\gamma = 0.5$, and $\epsilon = 0.001$.
- c. Gunakan persamaan (2) untuk melakukan perhitungan pada matriks. Gunakan persamaan (3), (4), dan (5) untuk mendapatkan data untuk $n = 1, 2, 3, \dots n$. Di mana variabel E_i merepresentasikan nilai error dari data ke- i , variabel γ merupakan singkatan dari learning rate, dan variabel $\max_{(i)} D_{ij}$ menunjukkan nilai maksimum dari diagonal larik Hessian.

$$D_{ij} = y_i y_j (K(\vec{x}_i \cdot \vec{x}_j) + \lambda^2) \tag{2}$$

$$E_i = \sum_{j=1}^n \alpha_j D_{ij} \tag{3}$$

$$\delta \alpha_i = \min\{\max[\gamma(1 - E_i) - \alpha_i], C - \alpha_i\} \tag{4}$$

$$\alpha_i = \alpha_i + \delta \alpha_i \tag{5}$$

- d. Gunakan persamaan (6) untuk menemukan nilai bias (b).

$$b = -\frac{1}{2}[w \cdot x^+ + w \cdot x^-] \tag{6}$$

- e. Tweet yang sedang dinilai akan melalui proses, dan persamaan (7) digunakan dalam proses mencapai penilaian. Jika hasil perhitungan lebih dari atau sama dengan nol, maka nilai $h(x)$ adalah +1, yang mengindikasikan bahwa tweet tersebut termasuk ke dalam kelas positif. Sebaliknya, jika hasil perhitungan kurang dari atau sama dengan nol, maka nilai $h(x)$ adalah -1, yang menandakan bahwa tweet tersebut termasuk ke dalam kelas negatif.

$$h(x) = \begin{cases} +1, & \text{if } w \cdot x + b \geq 0 \\ -1, & \text{if } w \cdot x + b < 0 \end{cases} \tag{7}$$

$$h(x) = w \cdot x + b$$

$$h(x) = \sum_{i=1}^m \alpha_i y_i K(x, x_i) + b$$

D. Confusion Matrix

Pendekatan Confusion Matrix, seperti yang diuraikan pada Tabel I, digunakan dalam investigasi ini sehingga hasil kategorisasi dapat dievaluasi. Perbandingan akan dilakukan antara matriks prediksi dan kelas asli, yang terdiri dari data asli dan juga nilai klasifikasi yang diprediksi [6]. Nilai akurasi ditentukan dengan menentukan berapa proporsi dari total jumlah prediksi yang benar yang diperoleh. Persamaan (8) [12] adalah persamaan yang digunakan untuk menghitung akurasi. Dengan menggunakan persamaan (9), seseorang dapat menentukan nilai recall, yang didefinisikan sebagai persentase contoh positif yang diidentifikasi dengan benar [12]. Persamaan (10) digunakan untuk menentukan nilai presisi, yang merupakan persentase dari kasus yang diantisipasi secara positif yang diprediksi secara akurat [12]. Nilai F1-score diperoleh dari komputasi yang mengevaluasi seberapa baik nilai precision dan recall bekerja sama (11) [5].

TABEL I
 CONFUSION MatrikX

	Prediction True	Prediction Negative
Actual True	TP	FN
Actual Negative	FP	TN

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

$$Presisi = \frac{TP}{TP + FP} \tag{10}$$

$$F1 - Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

E. Dataset

Data sentimen Twitter digunakan dalam penelitian ini adalah data yang diperoleh melalui crawling data menggunakan web hawksy.info, yang merupakan basis Tabel II. Data yang diambil adalah data dengan hastag #PialaDuniaU20 sejak tanggal 30 Maret 2023 sampai 1 April 2023. Dari proses tersebut dihasilkan 2966 jumlah tweet yang akan dianalisis.

TABEL II
 DATASET TWITTER

User	Tweet	Tanggal
reviel08	RT @detikcom: FIFA membatalkan Indonesia sebagai tuan rumah Piala Dunia U-20. Jalan panjang sudah ditempuh, asa itu pun pupus. Simak kronologinya. #PSSI #FIFA #TimnasU20 #PialaDuniaU20	1/4/2023 13:42
sportstarsdotid	Usai bertemu Presiden Jokowi, Shin Tae-yong bubar skuad Timnas Indonesia U-20 #pialadunia20	1/4/2023 13:35
Diseminasi_Info	#batal jd tuan rumah #pialadunia20 udah gak usah ngeles jelas itu karena jaminan keamanan kurang memadai #jokowi	1/4/2023 13:33
suaralebay	beginilah nasib petugas partai, ... kena semprit si mbok terus. #PialaDuniaU20	1/4/2023 13:27
BambangJonn	Udh kehabisan kata2, satu hal yang hanya bisa terucap, Semangat garuda muda! Semoga dengan adanya musibah ini, tidak melunturkan semangat kalian utk mengharumkan bangsa Indonesia. Jangan menyerah ya! #pialadunia #pialadunia20 #timnas #fifa #sepakbola #indonesia #pildun #pssi https://t.co/tkzjeMcsRZ	1/4/2023 13:07
⋮	⋮	⋮
heraldindo	Bukan saat ini, tapi nanti kita akan buktikan kalau kita bisa. #timnasu20 #FIFA #PSSI #PialaDuniaU20 #PakErick	30/03/2023 04:04:12

III. HASIL DAN PEMBAHASAN

Crawling tweet dengan menggunakan tagar "#PialaDuniaU20" di Twitter menghasilkan sebanyak 2.966 data yang berkaitan dengan Piala Dunia U-20. Data-data ini dikumpulkan dari Twitter. Data yang telah dikumpulkan akan melalui tahap yang disebut preprocessing. Data yang telah terkumpul akan dilakukan preprocessing data. Proses preprocessing data mencakup pembersihan (menghapus mention, emoticon, url, hashtag, dan retweet), case folding (mengubah semua huruf menjadi huruf kecil), tokenizing (menghapus stopword), dan penghapusan tweet yang sama. Jumlah tweet menjadi 1169 setelah proses preprocessing selesai. Hasil data setelah proses preprocessing dapat dilihat pada Tabel III.

TABEL III
 DATA SETELAH PREPROCESSING

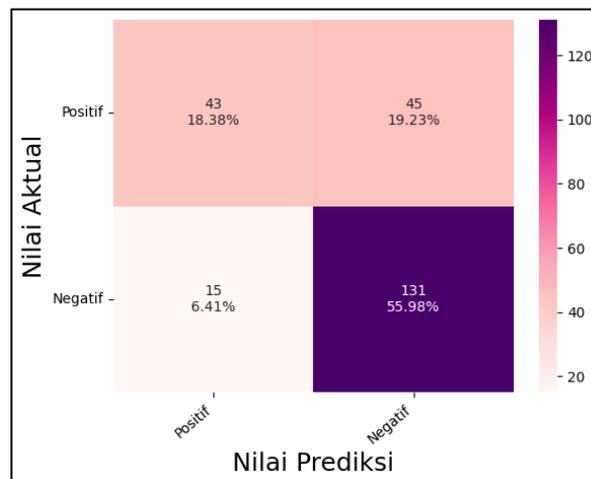
No	Text	Processed Text
1	RT @detikcom: FIFA membatalkan Indonesia sebagai tuan rumah Piala Dunia U-20. Jalan panjang sudah ditempuh, asa itu pun pupus. Simak kronologinya. #PSSI #FIFA #TimnasU20 #PialaDuniaU20	fifa batal indonesia tuan rumah piala dunia u 20 jalan tempuh asa pupus simak kronologi

No	Text	Processed Text
2	Usai bertemu Presiden Jokowi, Shin Tae-yong bubar skuad Timnas Indonesia U-20 #pialadunia20 https://t.co/9ODECHVM07	temu presiden jokowi shin tae-yong bubar skuad timnas indonesia u 20
3	1. batal jd tuan rumah #pialadunia20 udah gak usah ngeles jelas itu karena jaminan keamanan kurang memadai #jokowi https://t.co/jFxA6C4zIR	batal jd tuan rumah udah gak ngeles jamin aman pada
:	:	:
1168	RT @my_supersoccer: Segala macam usaha udah dilakuin Erick Thohir biar FIFA bisa tergerak hatinya dikit. Sayang Piala Dunia U-20 tetap gak bisa digelar di Indonesia 🇮🇩🇮🇩🇮🇩 🇮🇩 : PSSI #SUPERLIVE #SUPERSOCCER #NIRASANYASUPER #PIALADUNIAU20 https://t.co/voF3YW1VIS	usaha udah dilakuin erick thohir biar fifa gerak hati dikit sayang piala dunia u 20 gak gelar indonesia pssi
1169	Bukan saat ini, tapi nanti kita akan buktikan kalau kita bisa. #timnasu20 #FIFA #PSSI #PialaDuniaU20 #PakErick https://t.co/rH7LKwEw2u	bukti bisa

Data yang telah di preprocessing akan dilakukan proses pelabelan, proses pelabelan pada penelitian ini menggunakan metode Lexicon-based, hasil text yang telah dilakukan pelabelan dapat dilihat pada Tabel IV.

TABEL IV
 PELABELAN DATA

No	Processed Text	Label
1	fifa batal indonesia tuan rumah piala dunia u 20 jalan tempuh asa pupus simak kronologi	positif
2	temu presiden jokowi shin tae-yong bubar skuad timnas indonesia u 20	positif
3	batal jd tuan rumah udah gak ngeles jamin aman pada	negatif
:	:	:
1168	usaha udah dilakuin erick thohir biar fifa gerak hati dikit sayang piala dunia u 20 gak gelar indonesia pssi	negatif
1169	bukti bisa	positif



Gambar 1. Confusion Matrix

Data akan dibagi menjadi dua kategori: data pelatihan dan data pengujian. Setelah itu, TF-IDF akan diterapkan untuk memberikan nilai pada kata tersebut. Setelah data dilatih dan diuji, selanjutnya akan diklasifikasikan dengan menggunakan pendekatan SVM, yang akan menggunakan bobot kata yang telah diberikan pada data. Selain itu, prosedur evaluasi akan dilakukan dengan menggunakan confusion matrix, yang dapat ditunjukkan pada Gambar 1. Hal ini akan memungkinkan penentuan tingkat akurasi. Temuan dari penelitian ini mampu memberikan analisis mengenai perasaan yang dimiliki oleh anggota komunitas Twitter terhadap Piala Dunia U-20. Metode SVM dalam melakukan analisis sentimen menghasilkan kategori klasifikasi dengan kategori cukup [16], dimana nilai akurasi sebesar 74.4%, nilai precision sebesar 74%, nilai recall sebesar 74%, dan nilai f1-score sebesar 73%. Kesimpulan yang diambil dari penerapan pendekatan SVM pada sebuah evaluasi dapat menjadi saran untuk penelitian lebih lanjut [17].

IV. CONCLUSION

Dapat ditarik kesimpulan bahwa jumlah data awal adalah 2966 tweet, namun setelah prosedur preprocessing, jumlah tersebut berkurang menjadi 1196 tweet. Hal ini dapat disimpulkan dari penjelasan sebelumnya. Pendekatan Support Vector Machine (SVM) menghasilkan nilai akurasi sebesar 74.4% untuk analisis klasifikasi sentimen, dengan nilai precision sebesar 74%, nilai recall sebesar 74%, dan nilai f1 sebesar 73%. Selain itu, metode SVM memiliki nilai recall sebesar 74%. Temuan ini menunjukkan bahwa metode SVM cukup kompeten untuk tujuan klasifikasi dalam konteks analisis sentimen. Metode klasifikasi lain, seperti Decision Tree, Naive Bayes, dan KNN, harus digunakan dalam penelitian selanjutnya sebagai pengganti SVM sehingga informasi baru dan bahan pembelajaran yang lebih luas dapat diperoleh.

DAFTAR PUSTAKA

- [1] Ipsos, "Attitudes Towards the Fifa World Cup 2022 in Qatar," 2022. [Online]. Available: [https://www.ipsos.com/sites/default/files/ct/news/documents/2022-11/Ipsos 2022 FIFA World Cup Global Advisor Survey Report - Public Version.pdf](https://www.ipsos.com/sites/default/files/ct/news/documents/2022-11/Ipsos%202022%20FIFA%20World%20Cup%20Global%20Advisor%20Survey%20Report%20-%20Public%20Version.pdf)
- [2] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [3] R. W. Mayasari, K. Fithriasari, N. Iriawan, and W. S. Winahju, "Surabaya Government Performance Evaluation Using Tweet Analysis," *Matematika*, vol. 36, no. 1, pp. 31–42, 2020, doi: 10.11113/matematika.v36.n1.1176.
- [4] S. Naz, A. Sharan, and N. Malik, "Sentiment Classification on Twitter Data Using Support Vector Machine," *Proc. - 2018 IEEE/WIC/ACM Int. Conf. Web Intell. WI 2018*, pp. 676–679, 2019, doi: 10.1109/WI.2018.00-13.
- [5] M. F. Asshiddiqi and K. M. Laksmana, "Perbandingan Metode Decision Tree dan Support Vector Machine untuk Analisis Sentimen pada Instagram Mengenai Kinerja PSSI," *e-Proceeding Eng.*, vol. 7, no. 3, 2020.
- [6] N. Hendrastuty *et al.*, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021, [Online]. Available: <http://situs.com>
- [7] R. Tineges, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 650, 2020, doi: 10.30865/mib.v4i3.2181.
- [8] D. Maulina and I. Frans, "Studi Komprehensif Algoritma Naive Bayes Classifier Dengan Support Vector Machine Pada Sentiment Analysis Opini Pembangunan Infrastruktur di Media Sosial Twitter," *Inspir. J. Teknol. Inf. dan Komun.*, vol. 12, no. 1, p. 41, 2022, doi: 10.35585/inspir.v12i1.2640.
- [9] S. Dey, S. Wasif, D. S. Tonmoy, S. Sultana, J. Sarkar, and M. Dey, "A Comparative Study of Support Vector Machine and Naive Bayes Classifier for Sentiment Analysis on Amazon Product Reviews," *2020 Int. Conf. Contemp. Comput. Appl. IC3A 2020*, pp. 217–220, 2020, doi: 10.1109/IC3A48958.2020.233300.
- [10] I. Rozi, S. Pramono, and E. Dahlan, "Implementasi Opinion Mining (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik Pada Perguruan Tinggi," *J. EECCIS*, vol. 6, no. 1, pp. 37–43, 2012.
- [11] J. Faret and J. Reitan, "Twitter Sentiment Analysis-Exploring the Effects of Linguistic Negation," Norwegian University, 2015. [Online]. Available: <http://brage.bibsys.no/xmlui/handle/11250/2353488>
- [12] G. T. Santoso, "Analisis Sentimen Pada Tweet Dengan Tagar #Bpjsrasarentenir Menggunakan Metode Support Vectore Machine (Svm)," Universitas Islam Riau Pekanbaru, 2021.
- [13] M. A. F. Umi Rofiqoh1, Rizal Setya Perdana2 and Program, "Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter dengan Metode Support Vector Machine dan Lexicon Based Features Twitter event detection View project Human Detection and Tracking View project," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1(12), no. October, pp. 1725–1732, 2017, [Online]. Available: <https://www.researchgate.net/publication/320234928>
- [14] A. Kowalczyk, *Support Vector Machines Succinctly*. 2017.
- [15] S. Rani and J. Singh, "Sentiment Analysis of Tweets Using Support Vector Machine," *Int. J. Comput. Sci. Mob. Appl.*, vol. 5, pp. 83–91, 2017, [Online]. Available: www.xyz.com
- [16] Florin Gorunescu, *Data Mining: Concept, Models, Techniques*. Springer, 2011.
- [17] A. V. Vitianingsih, Z. Othman, S. S. K. Baharin, A. Suraji, and A. L. Maukar, "Application of the Synthetic Over-Sampling

Method to Increase the Sensitivity of Algorithm Classification for Class Imbalance in Small Spatial Datasets,” Int. J. Intell. Eng. Syst., vol. 15, no. 5, pp. 676–690, 2022.