

Analisis Sentimen Wacana Penerapan Jalan Berbayar di Jakarta Menggunakan Algoritma SVM

Ahmad Fauzi*, Eko Harli, Tria Hadi Kusmanto

Fakultas Teknik dan Ilmu Komputer, Teknik Informatika, Universitas Indraprasta PGRI, Jakarta, Indonesia

Email: ¹ahmadfauzi.udzi@gmail.com, ²eko.harli@gmail.com, ³triahadi226@gmail.com

Email Penulis Korespondensi: ahmadfauzi.udzi@gmail.com

Abstrak—Kemacetan lalu lintas merupakan salah satu masalah serius yang dihadapi oleh kota-kota besar di seluruh dunia, termasuk Jakarta. Salah satu upaya dalam menangani kemacetan di Jakarta adalah dengan menerapkan Jalan Berbayar (ERP) pada sejumlah ruas jalan. Usulan penerapan jalan berbayar merupakan kebijakan yang kontroversial di Jakarta, dan pemahaman sentimen masyarakat terkait kebijakan ini memiliki nilai penting dalam pengambilan keputusan. Twitter sebagai media sosial terbesar menjadi wadah aspirasi masyarakat dalam menyuarakan pendapat mengenai kebijakan tersebut. Penelitian ini bertujuan untuk menganalisis sentimen wacana terkait penerapan jalan berbayar di Jakarta menggunakan algoritma Support Vector Machine (SVM). Data sentimen yang digunakan terdiri dari sentimen positif dan sentimen negatif. Hasil penelitian menunjukkan bahwa algoritma SVM dengan nilai akurasi sebesar 81%, presisi sebesar 83%, dan recall sebesar 85%.

Kata Kunci: Sentimen; Twitter; ERP; Jakarta; SVM; Jalan Berbayar

Abstract—Traffic congestion is a serious problem faced by major cities worldwide, including Jakarta. One effort to address traffic congestion in Jakarta is by implementing Electronic Road Pricing (ERP) on certain road sections. The proposal for implementing ERP is a controversial policy in Jakarta, and understanding the sentiment of the community regarding this policy is crucial for decision-making. Twitter, as the largest social media platform, serves as a platform for the public to voice their opinions regarding this policy. This research aims to analyze the discourse sentiment related to the implementation of ERP in Jakarta using the Support Vector Machine (SVM) algorithm. The sentiment data used consists of positive and negative sentiments. The research results show that the SVM algorithm achieved an accuracy rate of 81%, precision of 83%, and recall of 85%.

Keywords: Sentiment; Twitter; ERP; Jakarta; SVM; Road Pricing

1. PENDAHULUAN

Kemacetan lalu lintas merupakan salah satu masalah serius yang dihadapi oleh kota-kota besar di seluruh dunia, termasuk Jakarta. Tingginya volume kendaraan, pertumbuhan populasi, dan infrastruktur yang terbatas telah menyebabkan kemacetan yang parah di ibu kota Indonesia. Pada tahun 2017 kemacetan menyebabkan peningkatan waktu tempuh secara rata-rata adalah 72 menit bagi para pekerja untuk berangkat ke kantor. Hal ini juga turut menimbulkan kerugian dari sisi ekonomi Jakarta hampir 65 Triliun di tahun 2023.

Untuk mengatasi masalah ini, pemerintah Jakarta telah menerapkan berbagai kebijakan, dimulai dari mengganti kebijakan three in one, menjadi sistem ganjil genap. Hingga adanya usulan penerapan sistem jalan berbayar (Electronic Road Pricing – ERP). Usulan jalan berbayar ini terus timbul-tenggelam 5 tahun belakangan ini, banyaknya faktor-faktor pendukung yang harus dipersiapkan oleh pemerintah DKI Jakarta menjadi salah satu penyebab belum dapat diterapkannya usulan tersebut.

Sistem ganjil genap melibatkan pembatasan kendaraan berdasarkan nomor plat, di mana kendaraan dengan nomor plat ganjil dan genap diperbolehkan untuk beroperasi pada hari-hari tertentu. Kebijakan ganjil genap sendiri sudah berjalan dengan baik [1]. Dimana 2017 indeks kemacetan 61% dan pada tahun 2020 turun menjadi 36%, namun ditahun 2023 ini indeks kemacetan kembali meningkat seiring dengan sudah diberlakukannya kembali kehidupan normal setelah pandemi Covid-19. Ini dapat terlihat dimana kota Jakarta dari peringkat ke-46 menjadi posisi ke-29 kota termacet di dunia (TomTom Traffic Index).

Sementara itu usulan kebijakan jalan berbayar adalah konsep di mana pengguna jalan harus membayar tarif tertentu untuk mengakses dan menggunakan jalan tersebut. Tujuan utama dari kebijakan ini adalah untuk mengurangi volume kendaraan pribadi, merangsang penggunaan transportasi publik, dan meningkatkan efisiensi lalu lintas. Dari hal tersebut maka perlu adanya peningkatan transportasi public yang nyaman, aman dan mudah dijangkau oleh masyarakat Jakarta.

Dalam konteks ini, analisis sentimen wacana kebijakan jalan berbayar menjadi penting untuk memahami dan mengevaluasi efektivitas kebijakan ini dalam mengatasi masalah kemacetan di Jakarta. Analisis sentimen adalah proses untuk memahami dan mengelompokkan suatu kata yang dibagi menjadi beberapa kelas. Analisis sentimen juga disebut opinion mining yang dapat diartikan menggali opini dan emosi dari data uji. Tujuan dari analisis sentimen itu sendiri adalah untuk mengekstrak atribut dan komponen dari beberapa komentar yang ada di media sosial dan sehingga dapat menentukan beberapa kelas positif, negatif dan netral [2][3].

Media sosial berperan sangat penting terhadap analisis sentimen publik sekarang ini, dengan banyaknya informasi yang didapat, akan lebih memudahkan dan bermanfaat untuk melakukan analisis sentimen terutama mengenai opini publik. Media sosial ialah salah satu sumber yang sangat umum digunakan untuk berkomunikasi, berbagi dokumen serta data dengan jumlah komunitas yang besar. Beragam opini diberikan masyarakat melalui media sosial dan yang disampaikan cenderung terpolarisasi menjadi sentimen yang mendukung dan menolak, salah satu media sosial tersebut

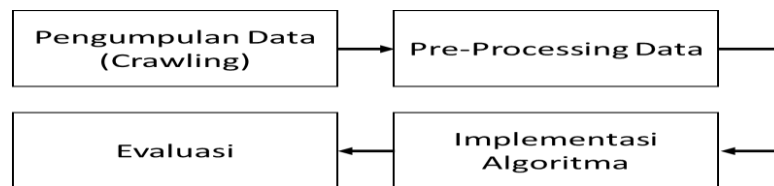
adalah Twiter [4]. Twitter merupakan salah satu platform yang banyak digunakan di hampir seluruh negara di Dunia, termasuk Indonesia. Berdasarkan laporan Wearesocial Hootsuite pengguna media sosial Twitter di Indonesia diawal tahun 2023 mencapai hingga 24 juta pengguna [5]

Salah satu metode yang dapat digunakan untuk melakukan analisis wacana adalah algoritma Support Vector Machine (SVM) [6]. SVM adalah metode pembelajaran mesin yang efektif untuk klasifikasi dan analisis sentimen dalam teks. SVM juga sangat cocok untuk mengatasi masalah dalam dunia nyata yang bersifat linear separable [7]. Akurasi SVM sebesar 82% lebih baik dari Algoritma Naïve Bayes yang hanya sebesar angka 76,56% ditunjukkan pada penelitian dalam mengulas review analisis [8]. Hasil yang sama juga ditunjukkan pada penelitian sentimen analisis dari data twitter terkait gempa yang terjadi di negara Chile, penelitian tersebut menunjukkan nilai akurasi SVM sebesar 81,2% dibandingkan dengan Naïve Bayes (74,2%) dan Random Forest (72,5%) [9]. Penelitian lainnya SVM menunjukkan nilai akurasi sebesar 91% berbanding Clustering I-means dengan nilai akurasi 82%, terhadap data sentimen analisis pada pelaksanaan PILKADA serentak dengan sumber data berasal dari Twitter [10].

Dalam konteks analisis wacana, SVM dapat digunakan untuk mengklasifikasikan teks menjadi kategori yang berbeda, seperti pendukung kebijakan jalan berbayar, penentang kebijakan, atau netral. Dengan memanfaatkan SVM, kita dapat mengidentifikasi dan menganalisis sentimen yang terkandung dalam teks terkait dengan kebijakan jalan berbayar, baik itu dalam bentuk ulasan masyarakat, artikel berita, atau twit media sosial. Penelitian terkait analisis sentimen terhadap bidang transportasi sebelumnya sudah dilakukan pada penelitian Analisa sentimen pengguna transportasi transjakarta menggunakan metode Naives Bayes dan K-Nearest Neighbor [11]. Penelitian lain juga mengangkat ERP dalam pengendalian mobil pribadi di salah satu jalan protokol di Jakarta, namun penelitian tersebut menggunakan model analisis sensitifitas hubungan utilitas antara ERP dan variable lain [12]. Dari penelitian sebelumnya dapat dilihat belum adanya penelitian yang berfokus pada Analisa sentimen terhadap wacana kebijakan ERP di Jakarta menggunakan metode SVM. Melalui pendekatan ini, kami bertujuan untuk memberikan pemahaman yang lebih baik tentang efektivitas kebijakan jalan berbayar dalam mengatasi masalah kemacetan di Jakarta. Dengan demikian, hasil analisis sentimen wacana ini diharapkan dapat memberikan wawasan yang berharga bagi para pengambil keputusan dan pemerintah daerah dalam perencanaan, pengembangan, dan evaluasi kebijakan transportasi di Jakarta

2. METODOLOGI PENELITIAN

Penelitian ini dilakukan dengan tahapan yang tersaji pada gambar 1, terdapat 4 proses tahapan secara umum dalam penelitian ini yaitu; Pengumpulan Data, Preprocessing Data, Implementasi Algoritma dan Evaluasi.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Tahapan ini merupakan tahapan awal untuk pengolahan data, data yang digunakan adalah data yang berasal dari twitter dengan proses crawling. Crawling adalah proses menjelajahi web dan mengunduh halaman web secara otomatis untuk mengumpulkan informasi. Program yang khusus bertugas melakukan crawling disebut Crawler [13]. Data yang diolah sebanyak 300 tweet hasil dari crawling pada twitter dengan rentang waktu 1 tahun terakhir.

2.2 Preprocessing Data

Preprocessing data merupakan tahapan awal sebelum data diolah, tujuan dari tahapan ini adalah untuk memastikan data yang digunakan bersih. Ada beberapa tahapan preprocessing yang dilakukan pada penelitian ini:

- Cleansing Data
Cleansing data membuang kata-kata yang tidak ada maknanya seperti URL, hastag, mention dan karakter yang tidak diperlukan.
- Labeling Data
Pelabelan data pada penelitian ini dilakukan untuk memberikan nilai sentiment positif ataupun negatif pada sebuah tweet.
- Case-folding
Case-folding mengubah semua karakter menjadi sama, misalnya kata-kata diubah dari huruf besar ke huruf kecil [14]

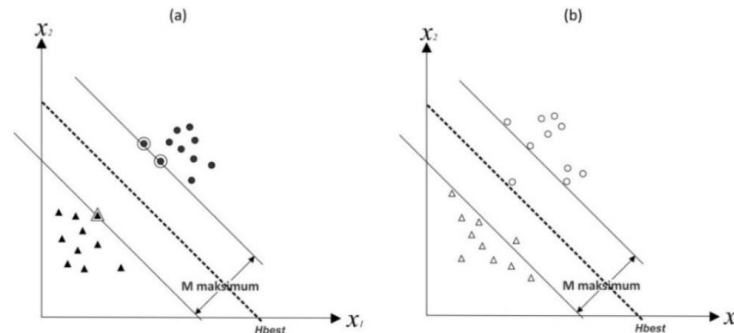
2.3 Implementasi Algoritma

2.3.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah sistem pembelajaran yang pengklasifikasiannya menggunakan ruang hipotesis berupa fungsi-fungsi linear dalam sebuah ruang fitur (feature space) berdimensi tinggi, dilatih dengan algoritma

pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan learning bias yang berasal dari teori pembelajaran statistik [15].

Dalam teknik ini, SVM berusaha untuk menemukan fungsi pemisah (klasifier) yang optimal yang bisa memisahkan dua kelas yang berbeda. Teknik ini berusaha menemukan fungsi pemisah (hyperplane) terbaik diantara fungsi yang tidak terbatas jumlahnya untuk memisahkan dua macam obyek. SVM berusaha menemukan hyperplane dengan memaksimalkan jarak antar kelas [16], hyperplane terbaik adalah hyperplane yang terletak di tengah-tengah antara dua set obyek dari dua kelas. Seperti diilustrasikan pada Gambar 1(a). Dengan cara ini, SVM menjamin kemampuan generalisasi yang tinggi untuk data-data yang akan datang, seperti diilustrasikan pada Gambar 1(b).



Gambar 2. Hyperline Terbaik dengan Margin Maksimum

Dalam SVM, objek-objek data terluar yang paling dekat dengan hyperplane ini disebut support vector. Hanya support vector inilah yang diperhitungkan oleh SVM untuk menemukan hyperplane paling optimal sedangkan objek-objek data yang lain tidak diperhitungkan sama sekali, dengan cara ini SVM dapat bekerja secara lebih efisien [7].

2.3.2 Metode Kernel SVM

Pada dasarnya, kasus-kasus di dunia nyata adalah kasus yang tidak linier. Untuk mengatasi sifat tidak linier adalah dengan memakai metode kernel. Metode kernel pada umumnya terdapat 4 jenis fungsi kernel yang dapat diterapkan pada algoritma SVM, yaitu:

- Kernel Linear
- Kernel Polynomial
- Kernel Gaussian
- Kernel Sigmoid

2.4 Evaluasi

Evaluasi dibutuhkan untuk melihat sejauh mana model yang dibangun terhadap sebuah algoritma bisa diterima atau perlu peningkatan lainnya. Evaluasi pada penelitian ini menggunakan perhitungan Accuracy: ukuran berapa banyak prediksi yang benar dari model yang telah dibuat untuk menguji dataset, Precision: Mengukur seberapa banyak kasus yang diprediksi dengan benar sebenarnya ternyata positif, dan Recall: mengukur kasus nilai actual positif yang dapat diprediksi dengan benar menggunakan model yang telah dibuat. Ketiga perhitungan tersebut merupakan perhitungan berdasarkan total prediksi ulasan yang benar berdasarkan model yang telah dibuat. Penelitian ini menggunakan Confusion Matrix untuk menghitung ketiga nilai tersebut [17].


3. HASIL DAN PEMBAHASAN

3.1 Crawling Data

Terdapat 304 hasil crawling data yang berasal dari twitter, jumlah data yang terbatas ini tidak lepas dari kebijakan twitter yang baru dimana perusahaan tersebut membatasi jumlah tweet yang dapat dilihat seseorang. Kata kunci yang digunakan dalam penelitian ini adalah “jalan berbayar erp Jakarta”. Tabel 1 menunjukkan beberapa hasil tweet yang didapat.

Tabel 1. Contoh Tweet Hasil Crawling Data

No	Tweet
1	Mobil listrik di jakarta sejatinya buat ngehindarin ganjil-genap. trus ada wacana jalan berbayar ERP si mobil listrik pun kena bayar juga kalo mau lewat. lalu temenku yang kemaren beli lonik5 misuh-misuh rak jelas 😞 😞 😞
2	Semoga dicabut Aturan Jalan berbayar/ ERP Di Dki jakarta
3	Makin hari jalanan jakarta makin macet, dan solusinya konon jalan-jalan berbayar (ERP) 🚗 🚘 🚙 🚚 solusi gak sih? #seriusnaya https://t.co/iPIZmet2e2

No	Tweet
4	Sejumlah warga yang tinggal di Ibu Kota mengaku keberatan dengan rencana Pemerintah menerapkan sistem jalan berbayar elektronik di sejumlah ruas jalan Ibu Kota. Transportasi umum di DKI Jakarta dinilai masih belum memadai secara waktu. < #Megapolitan #ERP https://t.co/Y7kKbdQDNc
5	Pemberlakuan jalan berbayar atau Electronic Road Pricing (ERP) di wilayah Jakarta, masih dalam tahap perencanaan. #DishubDKI #ERP #jalanberbayar #pemprovdk Simak selengkapnya klik disini  https://t.co/KYweEWKE0L https://t.co/7KOi5hOstV

3.2 Cleansing Data

Proses ini merupakan proses pembersihan dokumen dari kata- kata yang tidak dibutuhkan dan tweet-tweet yang terlihat duplikat untuk mengurangi noise [16]. Kata yang dihilangkan merupakan suatu karakter HTML, kata kunci, emoticon, hastag (#), mention username (@username), angka dan url [18]. Hasil dari cleansing data terdapat reduksi data dari 304 total tweet menjadi 259. Tabel 2 menunjukkan hasil dari cleansing data.

Tabel 2. Hasil Cleansing Data

No	Tweet
1	Mobil listrik di jakarta sejatinya buat ngehindarin ganjil-genap. trus ada wacana jalan berbayar ERP si mobil listrik pun kena bayar juga kalo mau lewat lalu temenku yang kemaren beli lonik5 misuh-misuh rak jelas
2	Semoga dicabut Aturan Jalan berbayar ERP Di Dki jakarta
3	Makin hari jalanan jakarta makin macet dan solusinya konon jalan-jalan berbayar ERP solusi gak sih
4	Sejumlah warga yang tinggal di Ibu Kota mengaku keberatan dengan rencana Pemerintah menerapkan sistem jalan berbayar elektronik di sejumlah ruas jalan Ibu Kota Transportasi umum di DKI Jakarta dinilai masih belum memadai secara waktu.
5	Pemberlakuan jalan berbayar atau Electronic Road Pricing ERP di wilayah Jakarta masih dalam tahap perencanaan

3.3 Labeling Data

Data hasil cleansing selanjutnya masuk kedalam tahap pelabelan data, proses pelabelan dilakukan secara manual dengan menentukan tweet tersebut masuk kedalam kelompok sentiment positif atau negatif. Metode pelabelan manual ini sangat cocok untuk menilai opini publik seperti dilakukan penelitian sebelumnya [7][19]. Label negatif diberikan terhadap tweet yang berisi penolakan dan kekecewaan terhadap adanya usulan wacana penerapan jalan berbayar di Jakarta. Sedangkan label positif diberikan terhadap tweet yang menunjukkan antusiasme dan optimistis bahwa dengan adanya penerapan jalan berbayar di Jakarta akan mengurangi kemacetan dan memudahkan semua pihak. Tabel 3 Menunjukkan jumlah sentimen negatif dan positif hasil proses pelabelan data

Tabel 3. Jumlah Pelabelan Data

Label	Jumlah
Positif	114
Negatif	145

3.4 Case Folding

Proses ini digunakan untuk menyamakan seluruh karakter yang ada menjadi huruf kecil seluruhnya [20]. Tabel 4 menunjukkan hasil dari proses case folding

Tabel 4. Hasil Case Folding

No	Sebelum Case Folding	Setelah Case Folding
1	Semoga dicabut Aturan Jalan berbayar ERP Di Dki jakarta	semoga dicabut aturan jalan berbayar erp di dki jakarta
2	Makin hari jalanan jakarta makin macet dan solusinya konon jalan-jalan berbayar ERP solusi gak sih	makin hari jalanan jakarta makin macet dan solusinya konon jalan-jalan berbayar erp solusi gak sih

3.5 Implementasi Algoritma

3.5.1 Data Latih dan Data Uji

Hasil preprocessing data selanjutnya dilakukan data splitting (pembagian data) antara data latih (data train) dengan data uji (data test). Data latih digunakan oleh SVM untuk membentuk model classifier, model yang terbentuk akan digunakan sebagai prediksi kelas data baru yang belum pernah ada sebelumnya. Data latih dan data uji yang digunakan adalah data yang telah memiliki label kelas, dengan perbandingan data latih dan data uji adalah 90% : 10% [7]. Setelah dilakukan data splitting diperoleh 233 data latih dan 26 data uji. Proses ini menggunakan bantuan library Scikit-learn [21]

3.5.2 Klasifikasi Support Vector Machine

Tahapan terakhir pada penelitian adalah menerapkan metode SVM terhadap data yang telah dilatih sebelumnya. Secara umum SVM memiliki 4 fungsi kernel dalam mengelompokkan data. Yaitu:.. Penelitian ini menggunakan fungsi kernel linear. Selanjutnya dilakukan pengujian terhadap nilai-nilai parameter SVM yang optimal untuk proses analisis sentimen. Pada algoritma SVM dengan menggunakan kernel linier, terdapat satu parameter yang dapat diuji yaitu nilai Cost (C), dengan nilai parameter C pada penelitian ini adalah: [0,01; 0,05; 0,25; 0,5; 0,75; 1]. Tabel 5 menunjukkan hasil accuracy score untuk setiap nilai C. Dari tabel terlihat bahwa nilai C terbaik dimulai dari nilai 0.25 dengan akurasi yang didapat adalah 85%

Tabel 5. Hasil pengujian nilai C

Nilai Cost	0,01	0,05	0,25	0,5	0,75	1
Accuracy Score	0.7307	0.7692	0.8461	0.8076	0.8076	0.8076

Pengujian ini menghasilkan akurasi, recall, percision serta F1- Score/F-Measure seperti yang terlihat pada tabel 6.

Tabel 6. Hasil Akurasi

	Precision	Recall	F1-score	Support
Negative	0,90	0,92	0,86	13
Positive	0,81	0,77	0,83	13
Accuracy			0,85	26
Macro avg	0,85	0,85	0,85	26
Weighted avg	0,85	0,85	0,85	26

Tabel 6. menunjukkan bahwa hasil pengujian untuk metode klasifikasi Support Vector Machine mendapatkan akurasi sebesar 85%.

4. KESIMPULAN

Berdasarkan dari penelitian yang sudah dilakukan, dapat disimpulkan bahwa dengan menggunakan Algoritma SVM dan kernel fungsi linear, menunjukkan akurasi sebesar 85% untuk analisis sentimen terhadap wacana penerapan jalan berbayar di Ibukota Jakarta. Pengembangan yang dapat dilakukan pada penelitian ini antara lain dengan menambah jumlah data yang diolah, menggunakan fungsi kernel yang lain seperti gaussian kernel pada algoritma SVM selain itu juga dapat dengan mengubah metode pemilihan data training. Selanjutnya dapat pula dengan melanjutkan untuk mengimplementasikan kedalam sebuah aplikasi untuk mengklasifikasikan opini negative dan positif.

REFERENCES

- [1] A. H. Rahadian, M. Saputra, and D. Ramadhanty, "Analisis Implementasi Kebijakan Sistem Ganjil Genap Dalam Mengatasi Kemacetan Di Provinsi DKI Jakarta," *J. Ilm. untuk Mewujudkan Masy. Madani*, vol. 50, no. 1, pp. 50–52, 2022, [Online]. Available: <http://ojs.stiami.ac.id>.
- [2] K. A. B. Permana, M. Sudarma, and W. G. Ariastina, "Analisis Rating Sentimen pada Video di Media Sosial Youtube Menggunakan STRUCT-SVM," *Maj. Ilm. Teknol. Elektro*, vol. 18, no. 1, p. 113, 2019, doi: 10.24843/mite.2019.v18i01.p17.
- [3] K. Setiawan, B. Rahmatullah, Burhanuddin, A. B. Paryanti, and F. Fauzi, "Komparasi Metode Naive Bayes Dan Support Vector Machine Menggunakan Particle Swarm Optimization Untuk Analisis Sentimen Mobil Esemka," *JISAMAR (Journal Inf. Syst. Applied, Manag. Account. Res.*, vol. 4, no. 3, pp. 102–111, 2020, [Online]. Available: <http://journal.stmikjayakarta.ac.id/index.php/jisamarTelp.+62-21-3905050>.
- [4] A. Harun and D. Putri Ananda, "Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naive Bayes dan Decision Tree," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 58–64, May 2021, doi: 10.57152/malcom.v1i1.63.
- [5] S. Kemp, "DIGITAL 2023: INDONESIA," 2023. <https://datareportal.com/reports/digital-2023-indonesia> (accessed Jul. 01, 2023).
- [6] P. A. Permatasari, L. Linawati, and L. Jasa, "Survei Tentang Analisis Sentimen Pada Media Sosial," *Maj. Ilm. Teknol. Elektro*, vol. 20, no. 2, p. 177, 2021, doi: 10.24843/mite.2021.v20i02.p01.
- [7] N. Fitriyah, B. Warsito, and D. A. I. Maruddani, "ANALISIS SENTIMEN GOJEK PADA MEDIA SOSIAL TWITTER DENGAN KLASIFIKASI SUPPORT VECTOR MACHINE (SVM)," *J. Gaussian*, vol. 9, no. 3, pp. 376–390, Aug. 2020, doi: 10.14710/j.gauss.v9i3.28932.
- [8] A. M. Rahat, A. Kahir, and A. K. M. Masum, "Comparison of Naive Bayes and SVM Algorithm based on Sentiment Analysis Using Review Dataset," in 2019 8th International Conference System Modeling and Advancement in Research Trends (SMART), Nov. 2019, pp. 266–270, doi: 10.1109/SMART46866.2019.9117512.
- [9] G. A. Ruz, P. A. Henríquez, and A. Mascareño, "Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers," *Futur. Gener. Comput. Syst.*, vol. 106, pp. 92–104, May 2020, doi: 10.1016/j.future.2020.01.005.
- [10] A. Rahmawati, A. Marjuni, and J. Zeniarja, "ANALISIS SENTIMEN PUBLIK PADA MEDIA SOSIAL TWITTER TERHADAP PELAKSANAAN PILKADA SERENTAK MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE," *CCIT J.*, vol. 10, no. 2, pp. 197–206, Aug. 2017, doi: 10.33050/ccit.v10i2.539.
- [11] I. Iwandini, A. Triayudi, and G. Soepriyono, "Analisa Sentimen Pengguna Transportasi Jakarta Terhadap Transjakarta

- Menggunakan Metode Naives Bayes dan K-Nearest Neighbor,” *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 543–550, Jan. 2023, doi: 10.47065/josh.v4i2.2937.
- [12] D. Putera and Najid, “PENGENDALIAN PENGGUNAAN MOBIL PRIBADI DENGAN STRATEGI PARKIR Tujuan Penelitian Manfaat / Sasaran Bagi Pengemudi Apabila Diterapkan ERP,” vol. 1, no. 2, pp. 81–89, 2018.
- [13] B. Liu, *Web Data Mining*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.
- [14] I. S. K. Idris and Y. A. Mustofa, “Typo Checking Menggunakan Algoritma Rabin-Karp,” *Jambura J. Electr. Electron. Eng.*, vol. 4, no. 1, pp. 87–91, Jan. 2022, doi: 10.37905/jjee.v4i1.12150.
- [15] M. P. Dwi Cahyo, Widodo, and B. Prasetya Adhi, “Kinerja Algoritma Support Vector Machine dalam Menentukan Kebenaran Informasi Banjir di Twitter,” *PINTER J. Pendidik. Tek. Inform. dan Komput.*, vol. 3, no. 2, pp. 116–121, Dec. 2019, doi: 10.21009/pinter.3.2.5.
- [16] H. Müller and J.-C. Freytag, “Problems, methods, and challenges in comprehensive data cleansing,” Jan. 2003.
- [17] M. I. Mazdadi, R. Ramadhani, T. H. Saragih, and M. Haekal, “KLASIFIKASI TANAMAN JARAK PAGAR MENGGUNAKAN ALGORITME DEEP LEARNING H2O,” *J. Komputasi*, vol. 9, no. 1, Apr. 2021, doi: 10.23960/komputasi.v9i1.2774.
- [18] D. Darwis, E. S. Pratiwi, and A. F. O. Pasaribu, “Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia,” *Eductic - Sci. J. Informatics Educ.*, vol. 7, no. 1, pp. 1–11, 2020, doi: 10.21107/edutic.v7i1.8779.
- [19] Merinda Lestandy, Abdurrahim Abdurrahim, and Lailis Syafa’ah, “Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 802–808, 2021, doi: 10.29207/resti.v5i4.3308.
- [20] M. Syarifuddin, “ANALISIS SENTIMEN OPINI PUBLIK MENGENAI COVID-19 PADA TWITTER MENGGUNAKAN METODE NAÏVE BAYES DAN KNN,” *INTI Nusa Mandiri*, vol. 15, no. 1, pp. 23–28, Aug. 2020, doi: 10.33480/inti.v15i1.1347.
- [21] A. Asro’i and H. Februariyanti, “Analisis Sentimen Pengguna Twitter Terhadap Perpanjangan Ppkm Menggunakan Metode K-Nearest Neighbor,” *J. Khatulistiwa Inform.*, vol. 10, no. 1, pp. 17–24, 2022, doi: 10.31294/jki.v10i1.12624.