

Penerapan SVM Sebagai Algoritma Machine Learning Dalam Analisis Sentimen Terhadap Telemedicine di Indonesia

Shafa Nadhifah ¹⁾, Aslam Fatkhudin ^{*2)}, Fenilinas Adi Artanto ³⁾

1. Sarjana Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pekajangan Pekalongan, Indonesia
2. Sarjana Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pekajangan Pekalongan, Indonesia
3. Sarjana Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pekajangan Pekalongan, Indonesia

Article Info

Kata Kunci: Media sosial, Sentimen, SVM, Telemedicine, Text Mining

Keywords: Social media, Sentiment, Support Vector Machine, Telemedicine, Text Mining

Article history:

Received: 28 Agustus 2025

Revised: 01 September 2025

Accepted: 04 September 2025

Available online: 01 November 2025

DOI :

[10.48144/suryainformatika.v15i2.2201](https://doi.org/10.48144/suryainformatika.v15i2.2201)

* Corresponding author.

Aslam Fatkhudin

E-mail address:

fathudin@gmail.com

ABSTRAK

Telemedicine di Indonesia telah berkembang sejak diperkenalkan pada tahun 2012, namun masih menghadapi tantangan terkait pemahaman teknologi dan keterbatasan akses di daerah terpencil. Mengingat pentingnya *telemedicine* dalam meningkatkan akses layanan kesehatan, pemahaman terhadap persepsi publik menjadi krusial untuk mengevaluasi penerimaan dan kendala yang dihadapi. Analisis sentimen dapat membantu mengidentifikasi opini masyarakat terhadap layanan ini, baik dari segi manfaat maupun hambatan yang dialami. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap layanan *telemedicine* menggunakan data dari media sosial X. Data dikumpulkan melalui teknik scraping dengan kata kunci "*Telemedicine*" dan dianalisis menggunakan algoritma Support Vector Machine (SVM). Hasil penelitian menunjukkan bahwa dari 525 *tweet* yang dianalisis, 468 *tweet* memiliki sentimen positif, 48 *tweet* negatif, dan 9 *tweet* netral. Model SVM menunjukkan tingkat akurasi sebesar 84%, dengan *precision* 28%, *recall* 33%, dan *F1-score* 30%.

ABSTRACT

Telemedicine in Indonesia has grown since its introduction in 2012, but it still faces challenges related to technological understanding and limited access in remote areas. Given the importance of telemedicine in improving access to health services, understanding public perception is crucial to evaluating acceptance and the obstacles faced. Sentiment analysis can help identify public opinion about this service, both in terms of benefits and obstacles experienced. This study aims to analyze public sentiment towards telemedicine services using data from social media X. Data was collected through scraping techniques with the keyword "Telemedicine" and analyzed using the Support Vector Machine (SVM) algorithm. The results showed that of the 525 tweets analyzed, 468 had positive sentiment, 48 were negative, and 9 were neutral. The SVM model showed an accuracy rate of 84%, with a precision of 28%, recall of 33%, and F1-score of 30%.

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan signifikan dalam berbagai sektor, termasuk bidang kesehatan. Salah satu inovasi yang muncul dari perkembangan ini adalah *telemedicine*, yaitu layanan kesehatan jarak jauh yang

memungkinkan pasien untuk berkonsultasi dengan tenaga medis tanpa harus datang ke fasilitas kesehatan. World Health Organization [1] mendefinisikan *telemedicine* sebagai "menyembuhkan dari jarak jauh" dengan memanfaatkan teknologi komunikasi untuk meningkatkan akses pasien terhadap layanan

kesehatan. Di Indonesia, *telemedicine* mulai dikembangkan sejak tahun 2012 oleh Kementerian Kesehatan melalui program Telemedicine Indonesia (TEMENIN) [2]. Namun, penerapannya masih menghadapi berbagai kendala, seperti keterbatasan infrastruktur internet, rendahnya literasi digital, serta belum adanya regulasi yang matang terkait layanan kesehatan berbasis digital. Pada tahun 2017, pemanfaatan *telemedicine* di Indonesia mengalami hambatan akibat rendahnya pemahaman masyarakat terhadap teknologi dan kurangnya kesadaran akan manfaat layanan ini [3]

Telemedicine menjadi solusi utama dalam bidang kesehatan di era digital selama pandemi covid-19, ketika akses masyarakat terhadap layanan kesehatan konvensional menjadi terbatas akibat pembatasan sosial dan risiko penularan penyakit. Layanan ini memungkinkan pasien untuk tetap mendapatkan konsultasi medis tanpa harus mengunjungi rumah sakit atau klinik secara langsung. Menurut [4], jumlah pengguna *telemedicine* di Indonesia meningkat sebesar 41% selama pandemi, menunjukkan bahwa *telemedicine* memiliki potensi besar dalam meningkatkan akses layanan kesehatan bagi masyarakat. Namun, meskipun memiliki manfaat yang signifikan, *telemedicine* masih menghadapi tantangan, terutama dalam hal penerimaan masyarakat. Beberapa faktor yang menjadi penghambat utama adalah keraguan terhadap efektivitas konsultasi jarak jauh, kekhawatiran terhadap keamanan data pasien, serta kurangnya interaksi langsung antara dokter dan pasien. Oleh karena itu, diperlukan pendekatan berbasis data untuk memahami persepsi masyarakat terhadap *telemedicine* dan faktor-faktor yang memengaruhi penerimaannya [5].

Di era digital, media sosial menjadi platform utama bagi masyarakat untuk berbagi pengalaman dan opini tentang berbagai topik, termasuk layanan kesehatan. Media sosial X merupakan salah satu platform yang banyak digunakan untuk mendiskusikan inovasi dalam bidang kesehatan, termasuk *telemedicine*. Dengan banyaknya diskusi dan opini yang tersedia di media sosial X, analisis sentimen dapat digunakan untuk menganalisis persepsi masyarakat secara luas dan real-time. Analisis sentimen atau opinion mining bertujuan untuk mengidentifikasi dan mengklasifikasikan opini seseorang terhadap suatu objek berdasarkan teks yang ditulisnya [6]. Dalam penelitian ini, analisis sentimen digunakan untuk memahami bagaimana masyarakat merespon layanan *telemedicine* melalui opini yang mereka bagikan di media sosial X. Sentimen dalam teks dapat diklasifikasikan ke dalam tiga kategori utama, yaitu positif, negatif, dan netral.

Dalam melakukan klasifikasi sentimen, penelitian ini menggunakan pendekatan Machine Learning dengan algoritma Support Vector Machine. SVM adalah salah satu algoritma klasifikasi yang dikenal memiliki performa tinggi dalam analisis teks, terutama dalam

klasifikasi sentimen. Keunggulan utama dari SVM adalah kemampuannya dalam menangani data berdimensi tinggi serta menghasilkan klasifikasi yang akurat bahkan dengan jumlah data yang terbatas [7]. Berdasarkan penelitian sebelumnya, SVM memiliki keunggulan dibandingkan metode lain seperti Naïve Bayes (NB). Studi yang membandingkan kedua algoritma menunjukkan bahwa SVM memiliki tingkat akurasi 92,91%, lebih tinggi dibandingkan NB yang hanya mencapai 85,93% [8]. Selain itu, penelitian oleh [9] menemukan bahwa kernel linear pada SVM memiliki performa terbaik dalam klasifikasi sentimen dengan akurasi mencapai 81%, precision 82%, dan recall 81%. Oleh karena itu, dalam penelitian ini digunakan algoritma SVM dengan kernel linear untuk mengklasifikasikan sentimen data dari media sosial X terkait layanan *telemedicine*.

Untuk meningkatkan akurasi klasifikasi, penelitian ini juga menerapkan metode Term Frequency - Inverse Document Frequency (TF-IDF) dalam proses ekstraksi fitur. TF-IDF merupakan teknik pembobotan kata dalam dokumen yang membantu algoritma Machine Learning dalam memahami kata-kata yang memiliki signifikansi lebih tinggi dalam analisis teks. Dengan menerapkan TF-IDF, model SVM dapat mengklasifikasikan data dengan lebih akurat, karena kata-kata yang lebih relevan akan memiliki bobot lebih tinggi dalam analisis sentimen.[10]

Penelitian ini bertujuan untuk menganalisis opini masyarakat terhadap layanan *telemedicine* di Indonesia dengan menggunakan data dari media sosial X. Selain itu, penelitian ini juga mengevaluasi kinerja model SVM dalam mengklasifikasikan sentimen, serta memberikan rekomendasi berbasis data untuk meningkatkan adopsi dan efektivitas layanan *telemedicine* di Indonesia. Dengan memahami opini masyarakat secara lebih mendalam melalui analisis sentimen, penelitian ini diharapkan dapat memberikan wawasan yang dapat digunakan oleh penyedia layanan *telemedicine*, pemerintah, serta tenaga kesehatan untuk meningkatkan kualitas dan penerimaan layanan kesehatan berbasis digital di Indonesia

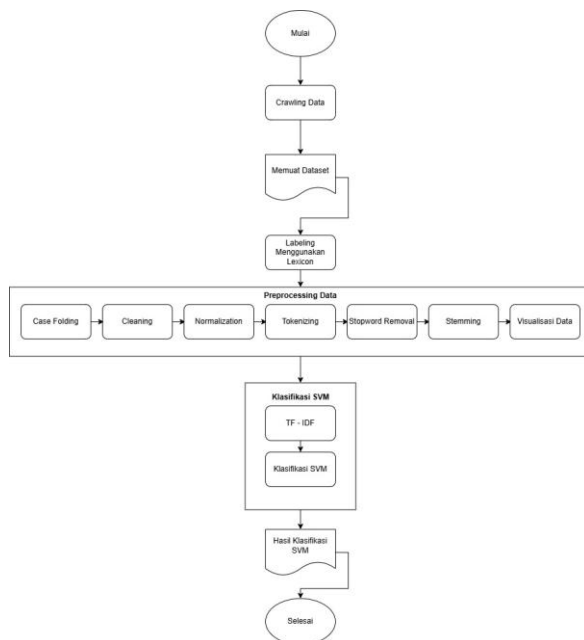
2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif yang bertujuan untuk menganalisis data melalui pengumpulan, mengolah, dan menilai data numerik secara objektif. Pendekatan kuantitatif dipilih karena penelitian ini berfokus pada klasifikasi sentimen masyarakat terhadap *telemedicine* berdasarkan data media sosial Media social X. Proses penelitian menggunakan teknik scraping untuk mengumpulkan data teks dari Media sosial X, yang kemudian diolah menggunakan algoritma Support Vector Machine (SVM). Untuk mengubah data teks mentah menjadi format yang siap untuk analisis, teknik Text Mining

juga digunakan dalam penelitian ini. Berkat penggunaan kata kunci seperti *Telemedicine* di media sosial X, sehingga hasil analisis benar mencerminkan pendapat publik. Selanjutnya, model dievaluasi dengan menggunakan matrik seperti akurasi, ketepatan, recall, dan F1-score untuk menilai kinerja algoritma secara objektif dan kuantitatif..

2.1 Kerangka Penelitian

Penelitian ini melakukan analisis data mendeskripsikan tentang data yang digunakan untuk menganalisis sentimen opini masyarakat di Media sosial X. penelitian ini menggunakan metode SVM, penelitian ini termasuk dalam penelitian text mining karena proses *crawling* data untuk mendapatkan data yang dapat digunakan untuk menganalisis sentimen. Berikut adalah kerangka penelitian:



Gambar 1. Alur Penelitian

Pada Gambar 1 Alur Penelitian menunjukkan bahwa penelitian ini diawali dengan pengumpulan data opini publik dari media sosial X menggunakan *crawling* data *tweet* dengan *library Tweet Harvest* dengan kata kunci *Telemedicine*. Data yang dikumpulkan dimuat ke dalam dataset dan dilabeli menggunakan metode berbasis *lexicon* untuk menentukan sentimen (positif, negatif, atau netral). *Lexicon* yang digunakan adalah *domain-specific lexicon*, sehingga Tingkat kepercayaan hasil pelabelan otomatisnya tinggi. *Lexicon* digunakan karena dapat memproses jutaan data dengan cepat. Jika menggunakan pelabelan manual, proses ini akan memakan waktu yang sangat lama, tidak efisien, dan mahal, selain itu dengan menggunakan *Lexicon* dapat memastikan konsistensi dalam proses pelabelan. Tidak seperti manusia yang bisa memiliki bias atau perbedaan interpretasi, algoritma berbasis leksikon akan selalu melabeli kata yang sama dengan skor yang sama. Selanjutnya, dilakukan

preprocessing data yang meliputi *case folding*, *cleaning*, normalisasi, *tokenizing*, *stopword removal*, *stemming*, dan visualisasi data. Setelah *preprocessing*, bobot kata dihitung menggunakan metode TF-IDF sebagai input untuk klasifikasi sentimen menggunakan Support Vector Machine (SVM). Hasil klasifikasi SVM kemudian dianalisis untuk memahami opini publik terkait topik penelitian. Penelitian diakhiri dengan evaluasi hasil analisis dan penarikan kesimpulan.

2.2 Pengumpulan Data (Crawling Data)

Penelitian ini mengumpulkan data dari media sosial X menggunakan perangkat lunak Tweet Harvest untuk menganalisis opini publik tentang *telemedicine*. Proses *crawling* dilakukan dengan memanfaatkan fitur *cookies* untuk mengakses dan mengumpulkan *tweet* yang mengandung kata "*Telemedicine*" dalam rentang waktu 1 Januari 2018 hingga 15 Desember 2024. Setiap *tweet* yang diambil mencakup informasi seperti tanggal, isi teks, dan *username* pengirim, lalu disimpan dalam format CSV. Data yang terkumpul kemudian disaring untuk memastikan relevansi sebelum memasuki tahap *preprocessing*, yaitu pembersihan elemen tidak relevan. Proses pengumpulan data ini bertujuan untuk mendapatkan wawasan komprehensif mengenai sentimen masyarakat terhadap layanan *telemedicine*.

2.3 Preprocessing

Pada tahap *preprocessing* ini, data mentah yang telah terkumpul akan diolah menjadi data yang dapat digunakan pada tahap selanjutnya. Tahap *preprocessing* terdiri dari 6 proses yaitu:

- Case Folding*, merupakan proses untuk mengubah semua huruf yang ada pada dokumen menjadi huruf kecil.
- Cleaning*, merupakan proses untuk menghilangkan atribut yang tidak berpengaruh terhadap klasifikasi yaitu tanda baca, karakter kosong, dan emoji.
- Normalization*, merupakan proses untuk memperbaiki kesalahan yang ada pada kata seperti ejaan yang salah agar kata yang memiliki makna sama menjadi setara.
- Tokenizing*, merupakan proses untuk memisahkan kalimat menjadi beberapa bagian kata.
- Stopword Removal*, merupakan proses untuk menghilangkan katakata yang dianggap tidak berpengaruh terhadap kalimat.
- Stemming*, merupakan proses untuk mengubah kata-kata yang ada menjadi bentuk kata dasar.

2.4 Ekstrasi Fitur

Proses mengubah data yang telah diproses menjadi representasi numerik agar dapat dipahami oleh algoritma klasifikasi. Tahapan ini mencakup *preprocessing*, *feature selection*, dan *feature transformation* untuk menentukan atribut yang paling relevan. Dalam penelitian ini, metode TF-IDF digunakan untuk

mendeskripsikan kata-kata dalam tweet tentang telemedicine. Hasil akhirnya adalah sekumpulan fitur yang mencerminkan pola data dan siap digunakan dalam analisis lebih lanjut.

2.5 Klasifikasi Support Vector Machine (SVM)

Klasifikasi dalam penelitian ini menggunakan algoritma Support Vector Machine (SVM) untuk mengkategorikan sentimen terhadap telemedicine menjadi positif, negatif, dan netral. SVM dipilih karena kemampuannya menangani data besar dan menghasilkan klasifikasi yang akurat. Prosesnya mencakup pelatihan model, penggunaan kernel linear untuk pemetaan data, serta optimasi margin agar model dapat digeneralisasi dengan baik. Setelah pelatihan, data uji digunakan untuk mengukur akurasi dan dianalisis pada tahap evaluasi.

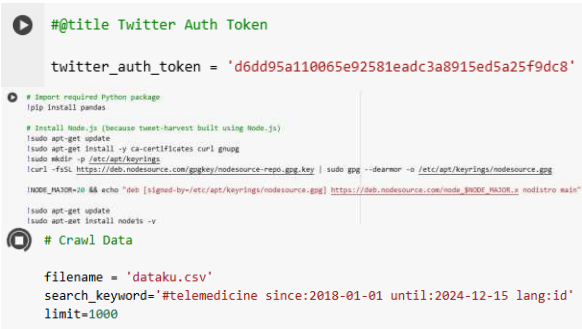
2.6 Evaluasi Model

Tahap evaluasi akan menggunakan confusion matrix untuk mengetahui performansi dari setiap algoritma SVM dengan hasil evaluasi berupa nilai accuracy, precision, recall, dan f1-score.

3. HASIL DAN PEMBAHASAN

3.1. Crawling Data

Penelitian ini mengumpulkan 525 komentar tentang layanan telemedicine di Indonesia dari media sosial X menggunakan metode *scraping* dengan *Tweet Harvest* di Google Colaboratory (https://colab.research.google.com/drive/1-HrziXJ43VwMRIY5Kbu-PTd-xRdddhoA#scrollTo=i9GI_skUDbDq). Data yang diperoleh diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral, menggunakan metode *modified lexicon*. Selanjutnya, data dianalisis menggunakan algoritma Support Vector Machine (SVM) untuk mengukur akurasi, precision, recall, dan F1-score. Proses pengambilan data dimulai dengan instalasi Tweet Harvest, impor dependensi, dan autentikasi menggunakan X Auth Token. Setelah itu, dilakukan *scraping* data berdasarkan kata kunci "Telemedicine", dengan cakupan waktu mulai 1 Januari 2018 hingga 15 Desember 2024. Hasil *scraping* kemudian disimpan dalam format CSV untuk mempermudah analisis lebih lanjut. Python dipilih dalam penelitian ini karena memiliki dukungan luas untuk analisis data dan pemrosesan teks. Gambar 2 merupakan proses *crawling* data



Gambar 2. Crawling Data

3.2. Preprocessing

Pada tahap ini dilakukan untuk mengubah data agar sesuai dan dapat diolah pada tahapan selanjutnya. Contoh hasil dari 6 proses pada tahap preprocessing yang telah dilakukan dapat dilihat pada table 1

Tabel 1. Preprocessing

Proses	Hasil
Data Ulasan	Telemedicine sedang berkembang pesat terutama di daerah terpencil! Bagaimana teknologi ini membantu pasien tanpa harus bepergian jauh? Temukan lebih lanjut di artikel terbaru IJHA: https://t.co/e044TN0RxO dan https://t.co/0BlgfKgmU5 IJHA #Telemedicine #eHealth https://t.co/3xVTRSEhdg
Case Folding	telemedicine sedang berkembang pesat terutama di daerah terpencil! bagaimana teknologi ini membantu pasien tanpa harus bepergian jauh? temukan lebih lanjut di artikel terbaru ijha: https://t.co/e044tn0rxo dan https://t.co/0blgfkgmu5 ijha #telemedicine #ehealth https://t.co/3xvtrsehdg
Cleaning	telemedicine sedang berkembang pesat terutama di daerah terpencil bagaimana teknologi ini membantu pasien tanpa harus bepergian jauh temukan lebih lanjut di artikel terbaru ijha dan ijha
Normalisasi	telemedicine sedang berkembang pesat terutama di daerah terpencil bagaimana teknologi ini membantu pasien tanpa harus bepergian jauh temukan lebih lanjut di artikel terbaru ijha dan ijha
Tokenizing	['telemedicine', 'sedang', 'berkembang', 'pesat', 'terutama', 'di', 'daerah', 'terpencil',

Scenario	Kernel	Akurasi	Precision	Recall	F1 Score
80:20	Linear	84%	28%	33%	30%

[illegible][illegible][illegible]

Word	IDF
diagnosa	6.471138
resep	5.801338
klinik	5.213745
virtual	4.973397
medis	4.969823
online	4.171674
pasien	3.175208
dokter	2.987729
konsultasi	2.975414
telemedicine	2.121801

Skenario Rasio Perbandingan	Data <i>Training</i>	Data <i>Testing</i>
80 : 20	372	94

Penerapan SVM Sebagai Algoritma Machine Learning Dalam Analisis Sentimen Terhadap Telemedicine di Indonesia

“telemedicine”, dan “sehat”, sedangkan pada ulasan negatif yaitu “pasien”, “covid”, dan “sakit”, dan untuk ulasan netral terdapat kata “aplikasi”, “online”, “konsultasi”.

4. KESIMPULAN

Berdasarkan hasil penelitian, dari 525 data tweet yang dikumpulkan dari media sosial X, sentimen positif berjumlah 468 data, negatif sebanyak 48 data, dan netral sebanyak 9 data, berdasarkan labeling Modified Lexicon. Klasifikasi menggunakan Support Vector Machine (SVM) dengan kernel linear menunjukkan tingkat akurasi 84%, precision 28%, recall 33%, dan F1-score 30%, meskipun akurasi cukup tinggi, kemampuan membedakan sentimen negatif dan netral masih perlu ditingkatkan. Hasil analisis sentimen dan visualisasi data menunjukkan kata yang dominan dalam sentimen positif adalah “layanan”, “telemedicine”, dan “sehat”, sedangkan dalam sentimen negatif terdapat kata “pasien”, “covid”, dan “sakit”, serta dalam sentimen netral ditemukan kata “aplikasi”, “online”, dan “konsultasi”. Hal ini mengindikasikan bahwa masyarakat merespons positif terhadap kemudahan akses telemedicine, tetapi masih terdapat keluhan terkait masalah teknis dan efektivitas layanan. Penelitian ini memiliki kendala, seperti keterbatasan dataset, yang memengaruhi akurasi dalam membedakan sentimen. Selain itu, penggunaan bahasa tidak baku dalam beberapa tweet juga mempengaruhi hasil klasifikasi karena keterbatasan kamus dalam *Modified Lexicon* yang digunakan

REFERENSI

- [1] 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, 2020, doi: 10.14710/j.gauss.v9i3.28932.
- [2] R. Rakarahayu Putri and N. Cahyono, “Analisis Sentimen Komentar Masyarakat Terhadap Pelayanan Publik Pemerintah Dki Jakarta Dengan Algoritma Super Vector Machine Dan Naive Bayes,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 2, pp. 2363–2371, 2024, doi: 10.36040/jati.v8i2.9472.
- [3] D. Agustina, A. Sufia, H. Shofia, I. Cahyani, J. P. Ralya, and T. Mariani, “Review Article: Efektivitas Penggunaan Telemedicine Pada Masa Pandemi Sebagai Sarana Konsultasi Kesehatan,” *J. Keperawatan dan Kesehat. Masy. Cendekia Utama*, vol. 12, no. 3, p. 257,
- [4] M. Ulfa, “Analisis Sentimen Terhadap Layanan Telemedicine Menggunakan Metode Support Vector Machine (SVM),” Universitas Malikussaleh, 2022.
- [5] R. A. J. Fahmi, W. M. Nur, D. Canawine, M. N. Kusumajaya, A. F. Fadhlillah, and N. A. Rakhmawati, “Analisis Sentimen Masyarakat Terhadap Uu Perlindungan Data Pribadi Pada Media Sosial Twitter Menggunakan Metode Support Vector Machine,” *Method. J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 1, pp. 6–10, 2024, doi: 10.46880/mtk.v10i1.2335.
- [6] S. F. Pane, A. Owen, and C. Prianto, “Analisis Sentimen UU Omnibus Law pada Twitter Menggunakan Metode Support Vector Machine,” *J. Telekomun. dan Komput.*, vol. 11, no. 2, p. 130, 2021, doi: 10.22441/incomtech.v11i2.10874.
- [7] R. Rosdiana, M. Ula, and H. A. K. Aidilof, “Implementasi Pemodelan Citra Model Svm (Support Vector Machine) Dalam Penentuan Pengklasifikasian Jenis Suara Kontes Burung,” *J. Inform. Kaputama*, vol. 5, no. 2, pp. 317–324, 2021, doi: 10.59697/jik.v5i2.264.
- [8] Y. Alkhalifi, W. Gata, A. Prasetya, and I. Budiawan, “Analisis Sentimen Penghapusan Ujian Nasional pada Twitter Menggunakan Support Vector Machine dan Naive Bayes berbasis Particle Swarm Optimization,” *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 6, no. 2, p. 71, 2020, doi: 10.24014/coreit.v6i2.9723.
- [9] A. Liani, “Analisis Perbandingan Kernel Algoritma Support Vector Machine dalam Mengklasifikasikan Skripsi Teknik Informatika berdasarkan Abstrak,” *JOINS (Journal Inf. Syst.*, vol. 5, no. 2, pp. 240–249, 2020, doi: 10.33633/joins.v5i2.3715.
- [10] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, “Comparative Evaluation of SVM Kernels for Sentiment Classification in Fuel Price Increase Analysis,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 153–160, 2023.