



Penerapan *Data Mining* Dalam Menentukan Pelajaran yang Diminati Dengan Metode *Support Vector Mechine (SVM)*

Iis Sulistilawati ¹, * Ahmad Musyafa ², Rafi Mahmud Zain ³

^{1,2,3}Teknik Informatika, Program Pascasarjana, Universitas Pamulang, Tangerang Selatan, Banten

Email: ¹ iissuull77@gmail.com, ² dosen00668@unpam.ac.id, ³ rafizain777@gmail.com

ABSTRACT

This research focuses on predicting students' grades and learning interests using the Support Vector Machine (SVM) method in education. It is crucial for students to determine the subjects they are most interested in to achieve optimal results. Data mining, particularly through the SVM method, can help identify students' learning interests based on historical data. SVM analyzes various variables related to students' academic performance, such as test scores, assignments, attendance, and class participation. By analyzing this data, SVM builds a model that predicts students' interest in specific subjects. This model provides insights into students' potential academic performance. The predictions generated by this SVM model can be used by educators to design more personalized learning strategies. By understanding students' interests and potential, educators can offer lesson recommendations that align with individual needs, enhancing students' motivation and academic performance. Moreover, these predictions assist students in making more informed decisions regarding subject selection, optimizing their potential and achieving better educational outcomes. This research aims to provide a tool that helps direct students' learning interests effectively, thereby improving the overall quality of education and academic results.

Keywords: Prediction, Data Mining, Support Vector Machine (SVM).

ABSTRAK

Penelitian ini berfokus pada prediksi nilai dan minat belajar siswa menggunakan metode Support Vector Machine (SVM) dalam dunia pendidikan. Penting bagi siswa untuk menentukan pelajaran yang paling diminati agar dapat mencapai hasil yang optimal. Data mining, khususnya melalui metode SVM, dapat membantu mengidentifikasi minat belajar siswa berdasarkan data historis. SVM menganalisis berbagai variabel terkait performa akademik siswa, seperti nilai ujian, hasil tugas, kehadiran, dan partisipasi dalam kelas. Dengan menganalisis data ini, SVM membangun model yang memprediksi minat siswa terhadap mata pelajaran tertentu. Model ini memberikan wawasan tentang potensi performa akademik siswa. Prediksi yang dihasilkan oleh model SVM ini dapat digunakan oleh pendidik untuk merancang strategi pembelajaran yang lebih personal. Dengan memahami minat dan potensi siswa, pendidik dapat menawarkan rekomendasi pelajaran yang sesuai dengan kebutuhan individu, sehingga meningkatkan motivasi belajar dan prestasi akademik siswa. Selain itu, prediksi ini membantu siswa dalam membuat keputusan yang lebih tepat terkait pemilihan mata pelajaran, mengoptimalkan potensi diri, dan mencapai hasil pendidikan yang lebih baik. Penelitian ini bertujuan untuk menyediakan alat yang membantu mengarahkan minat belajar siswa secara efektif, sehingga meningkatkan kualitas pendidikan dan hasil akademik secara keseluruhan.

Kata kunci: Prediksi, *Data Mining*, *Support Vector Mechine (SVM)*

1. PENDAHULUAN

Pendidikan adalah elemen krusial dalam meningkatkan peradaban, memajukan dunia kerja, dan membentuk perilaku serta karakter individu. Setiap orang memiliki sikap dan karakter yang unik, yang sering kali dipengaruhi oleh lingkungan masyarakat sekitar [1]. Namun, banyak pelajar yang belum menemukan minat mereka dalam proses belajar mengajar. Hal ini menimbulkan tantangan dalam memetakan minat belajar yang bervariasi di antara pelajar.

Untuk mengatasi masalah ini, teknologi machine learning dapat dimanfaatkan, khususnya melalui *Data Mining* [2]. *Data Mining* memungkinkan untuk memprediksi minat atau ketertarikan pelajar terhadap berbagai materi pembelajaran. Saat ini, sistem input nilai yang dilakukan secara manual sering kali memperumit proses penentuan minat belajar yang paling diminati oleh setiap pelajar. Oleh karena itu, diperlukan solusi yang lebih efisien dan akurat [3].

Data Mining menawarkan metode yang dapat membantu dalam pengumpulan dan analisis data nilai pelajar. Data ini kemudian diakumulasikan dan dianalisis untuk menentukan pelajaran yang memiliki nilai rata-rata tertinggi di antara semua pelajaran lainnya. Proses ini bertujuan untuk mengekstrak informasi penting yang nantinya disampaikan kembali kepada pelajar dan orang tua/wali. Dengan demikian, mereka dapat lebih memahami perkembangan atau perubahan yang terjadi pada setiap pelajar.

Salah satu metode yang dapat digunakan dalam *Data Mining* untuk tujuan ini adalah *Support Vector Machine (SVM)* [4]. SVM merupakan metode klasifikasi yang bertujuan menemukan batas optimal dengan margin maksimal, yang membantu dalam menentukan nilai maksimal yang diperoleh pelajar. Metode ini memungkinkan penentuan pelajaran yang paling diminati secara otomatis dan lebih akurat dibandingkan dengan metode manual. Selain itu, penggunaan machine learning dapat mengurangi kesalahan manusia (*human error*), sehingga hasil yang diperoleh lebih andal dan dapat diandalkan. Dengan memanfaatkan teknologi seperti *Data Mining* dan SVM, kita dapat lebih memahami dan mendukung minat belajar pelajar, membantu mereka mencapai potensi maksimal, dan memberikan informasi yang bermanfaat bagi orang tua/wali. Hal ini akan berkontribusi pada peningkatan kualitas pendidikan dan pengembangan sumber daya manusia yang lebih baik di masa depan.

2. TUJUAN PENELITIAN

Tujuan penelitian ini adalah untuk mendapatkan pemahaman mendalam mengenai minat dan ketertarikan pelajar terhadap berbagai mata pelajaran di Pondok Pesantren Nafidatunnajah. Penelitian ini berupaya mengeksplorasi bagaimana dinamika pembelajaran berlangsung di pesantren tersebut dan mengidentifikasi faktor-faktor yang mempengaruhi minat belajar para pelajar. Dengan menggunakan metode kualitatif dan analisis data akademis yang dikumpulkan setelah ujian akhir semester, penelitian ini bertujuan untuk menemukan pola dan tren yang dapat memberikan wawasan lebih dalam tentang interaksi pelajar dengan materi pembelajaran serta menginformasikan pengembangan metode pembelajaran yang lebih efektif di masa depan.

3. LANDASAN TEORI

3.1. *Data Mining*

Data Mining adalah proses menemukan pola-pola yang bermanfaat dan informasi dari sejumlah besar data. Dalam konteks penelitian ini, *Data Mining* digunakan untuk menganalisis nilai akademis pelajar dengan tujuan untuk mengidentifikasi minat mereka terhadap berbagai mata pelajaran. Metode ini memungkinkan peneliti untuk mengekstrak informasi berharga dari data yang dikumpulkan dan memberikan wawasan yang dapat digunakan untuk meningkatkan proses pembelajaran [5].

3.2. *Support Vector Machine (SVM)*

Support Vector Machine (SVM) merupakan salah satu algoritma yang digunakan dalam data mining untuk klasifikasi dan regresi [6]. Dalam konteks penelitian ini, SVM digunakan untuk mengklasifikasikan data nilai akademis pelajar guna menentukan minat mereka terhadap berbagai mata pelajaran [7]. Algoritma ini sangat efisien dalam menangani masalah klasifikasi, terutama ketika data memiliki banyak dimensi. SVM bekerja dengan cara mencari *hyperplane* di ruang multidimensi yang dapat memisahkan data ke dalam kelas-kelas yang berbeda. *Hyperplane* ini dipilih sedemikian rupa sehingga memiliki margin maksimal, yaitu jarak terdekat antara data dari masing-masing kelas ke *hyperplane* tersebut [8]. Dengan kata lain, SVM berusaha menemukan *hyperplane* yang paling optimal yang memisahkan kelas-kelas data dengan margin terbesar. Ada dua jenis SVM yang dapat digunakan, yaitu [9]:

1. Linear SVM: Digunakan ketika data dapat dipisahkan dengan *hyperplane* linier.
2. Non-Linear SVM: Digunakan ketika data tidak dapat dipisahkan dengan *hyperplane* linier. Dalam hal ini, SVM menggunakan *kernel trick* untuk memetakan data ke dalam ruang dimensi yang lebih tinggi di mana data dapat dipisahkan secara linier.

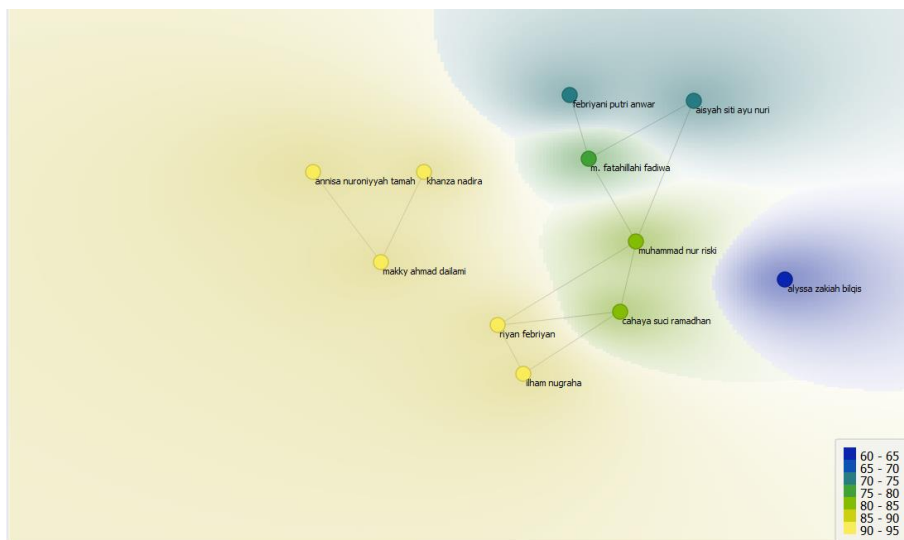
Proses implementasi SVM dalam penelitian ini melibatkan beberapa langkah sebagai berikut [10]:

1. *Preprocessing* Data: Data nilai akademis pelajar yang telah dikumpulkan diolah terlebih dahulu. Tahap ini meliputi normalisasi data, pengisian nilai yang hilang, dan penghapusan *outliers*. Normalisasi dilakukan untuk memastikan bahwa data berada dalam skala yang sama, sehingga dapat meningkatkan kinerja algoritma SVM.
2. Pemilihan Fitur: Fitur-fitur yang relevan dipilih untuk digunakan dalam model SVM. Dalam konteks penelitian ini, fitur-fitur yang dipilih meliputi nilai akademis dari berbagai mata pelajaran, usia, jenis kelamin, dan latar belakang keluarga pelajar. Pemilihan fitur yang tepat sangat penting untuk meningkatkan akurasi model.
3. Pemilihan Kernel: Dalam SVM, pemilihan kernel sangat berpengaruh terhadap kinerja model. Beberapa kernel yang umum digunakan adalah *linear*, *polynomial*, dan *radial basis function* (RBF). Peneliti akan melakukan percobaan dengan berbagai kernel untuk menentukan kernel yang memberikan hasil terbaik.
4. Pelatihan Model: Data yang telah *preprocessing* dan fitur yang telah dipilih kemudian digunakan untuk melatih model SVM. Data dibagi menjadi data pelatihan dan data uji untuk mengukur kinerja model. Proses pelatihan melibatkan optimasi parameter SVM untuk mendapatkan *hyperplane* yang paling optimal.
5. Evaluasi Model: Model SVM yang telah dilatih dievaluasi menggunakan data uji. Beberapa metrik evaluasi yang digunakan meliputi akurasi, *precision*, *recall*, dan *F1-score*. Metrik-metrik ini memberikan gambaran mengenai kinerja model dalam mengklasifikasikan data dengan benar.

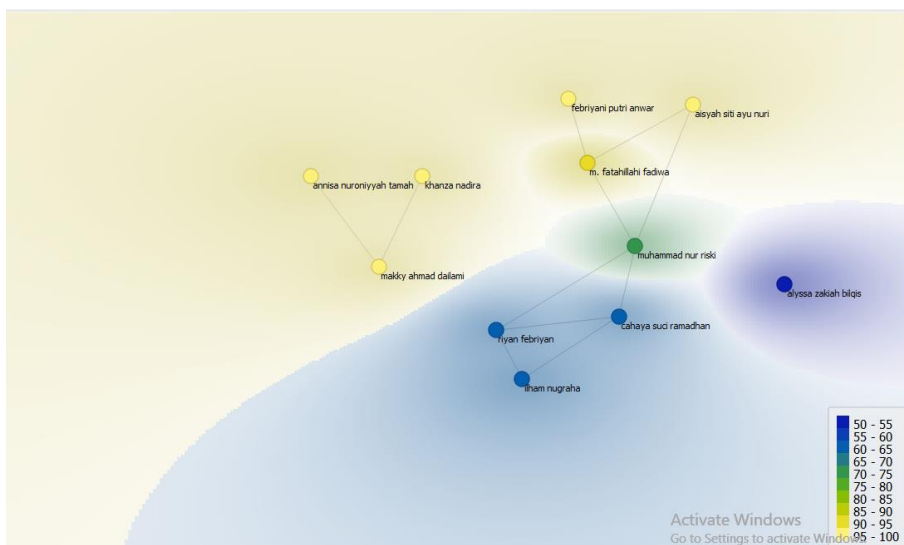
6. Interpretasi Hasil: Hasil evaluasi model SVM kemudian diinterpretasikan untuk menarik kesimpulan mengenai minat pelajar terhadap berbagai mata pelajaran. Hasil ini akan digunakan untuk memberikan rekomendasi kepada pelajar dan orang tua/wali mengenai mata pelajaran yang paling diminati dan area yang memerlukan perhatian lebih.

4. HASIL DAN PEMBAHASAN

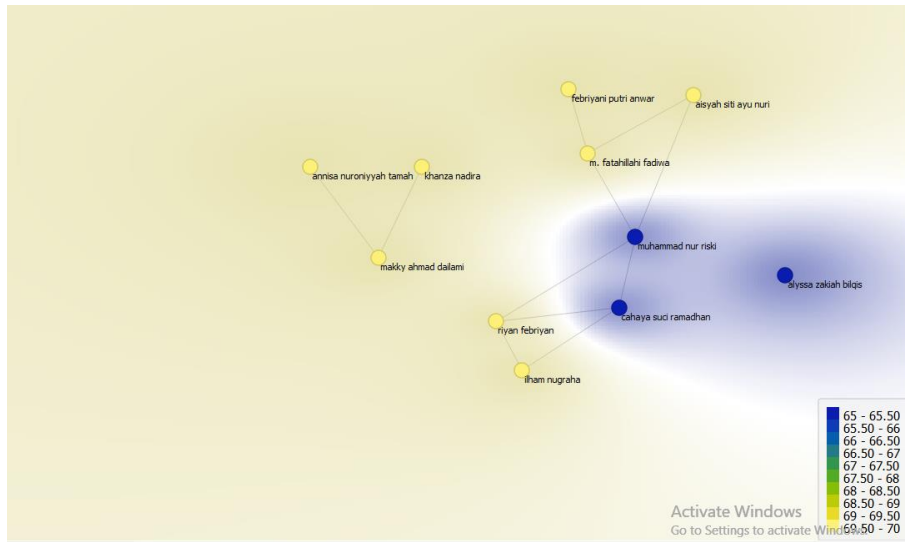
Berdasarkan dari pengumpulan data yang diperoleh pada nilai ujian akhir tahun ajaran 2023-2024 dari siswa kelas 2 Pondok Pesantren Nafidatunnajah pada seluruh mata pelajaran diperoleh sebagai berikut:



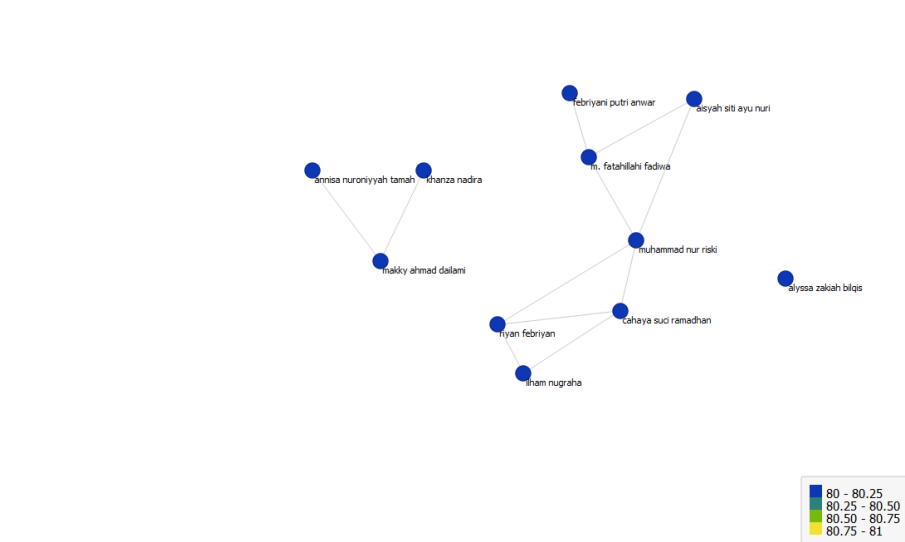
Gambar 1. Persebaran Data Mata Pelajaran Tajwid



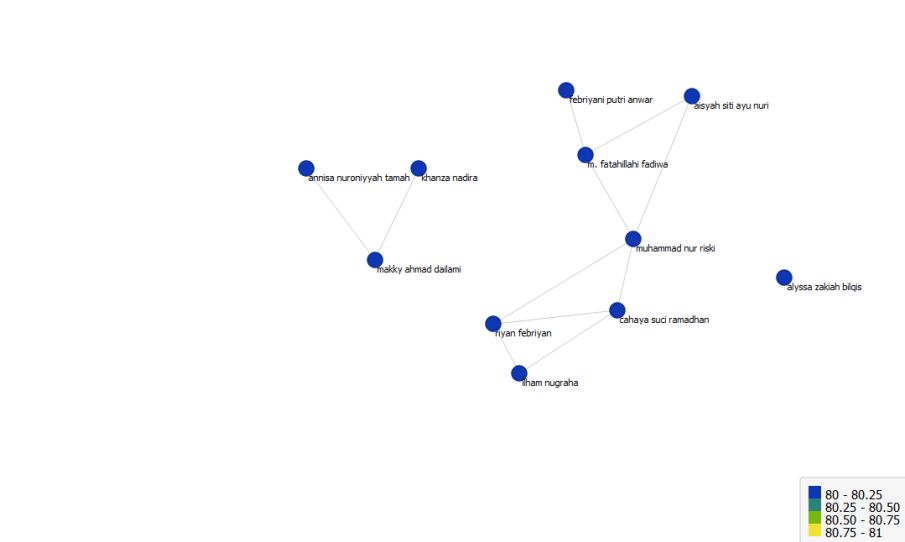
Gambar 2. Persebaran Data Mata Pelajaran Bahasa Inggris



Gambar 3. Persebaran Data Mata Pelajaran Matematika



Gambar 4. Persebaran Data Mata Pelajaran Imla

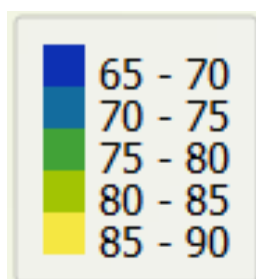


Gambar 5. Persebaran Data Mata Pelajaran Hadist



Gambar 6. Persebaran Data Mata Pelajaran Fiqih

Dengan hasil yang telah diperoleh maka dapat disimpulkan bahwa pada perbedaan warna pada tingkat minat didapatkan 5 warna, yaitu *blue*, *medium blue*, *green*, *medium green*, dan *yellow*. Sebagai berikut:



Dan telah dikelompokkan berdasarkan warna, apabila terdapat warna biru atau biru muda yang cukup banyak maka daya tarik pelajar terhadap pelajaran tersebut itu rendah, dan apabila warna yang terdapat pada hasil penelitian itu hijau atau hijau muda maka daya tarik pelajar terhadap pelajaran tersebut itu normal, dan apabila hasil penelitian menunjukkan bahwa warna kuning muda atau kuning pekat maka daya tariknya tinggi dan nilai yang diperoleh itu besar. Dan telah dikelompokkan sebagai berikut:

Mata Pelajaran	Blue	Medium Blue	Green	Medium Green	Yellow
Mutholaah	1	3	2	3	2
IPA	7		3		1
B. Indonesia	3	4	2	1	1
Imla	11				
Inggris	1	3	1		6
Nahwu	3		4	1	3
Hadist	11				
Matematika	3				8
T. Lugoh	2	1	4	3	1

Mata Pelajaran	Blue	Medium Blue	Green	Medium Green	Yellow
Tajwid	1	2	1	3	5
Tafsir	2	2	2	6	1
Fikih	11				
Shorof	7	1	2		1
Khot	1		4	3	3
Insya	3	4		2	2
Mahfudzot	2	3	1	1	4

Maka dapat disimpulkan bahwa mata pelajaran yang paling diminati adalah:

1. Matematika
2. Inggris
3. Tajwid

Dan berdasarkan data yang diperoleh maka pelajaran yang kurang diminati adalah:

1. Imla
2. Hadist
3. Fiqih

5. KESIMPULAN

Penelitian ini menunjukkan bahwa daya tarik pelajar terhadap mata pelajaran dapat dikategorikan berdasarkan warna, dengan warna biru atau biru muda menunjukkan daya tarik yang rendah, hijau atau hijau muda menunjukkan daya tarik yang normal, dan kuning muda atau kuning pekat menunjukkan daya tarik yang tinggi. Dari hasil analisis, ditemukan bahwa mata pelajaran Matematika, Bahasa Inggris, dan Tajwid adalah yang paling diminati oleh pelajar, ditandai dengan tingginya jumlah indikator warna kuning yang menunjukkan minat dan nilai yang tinggi. Sebaliknya, mata pelajaran Imla, Nahwu, dan Hadist kurang diminati oleh pelajar, yang ditunjukkan oleh dominasi warna biru pada hasil penelitian.

Temuan ini mengindikasikan bahwa pelajar cenderung lebih tertarik dan berprestasi lebih baik dalam mata pelajaran yang mereka anggap menarik, sementara mata pelajaran yang kurang menarik membutuhkan strategi pengajaran yang lebih inovatif untuk meningkatkan daya tarik dan performa pelajar. Untuk meningkatkan minat pelajar pada mata pelajaran yang kurang diminati seperti Imla, Nahwu, dan Hadist, diperlukan strategi pengajaran yang lebih menarik, seperti metode pembelajaran

interaktif atau integrasi teknologi. Kurikulum dan materi ajar juga perlu diperbarui agar lebih relevan dengan kehidupan sehari-hari siswa.

Sebaliknya, untuk mata pelajaran yang sudah diminati seperti Matematika, Bahasa Inggris, dan Tajwid, perlu dipertahankan dan ditingkatkan kualitas pengajarannya dengan menambahkan proyek-proyek praktis, kompetisi, dan kegiatan ekstrakurikuler terkait. Penyediaan materi tambahan bagi siswa yang menunjukkan minat lebih dalam pada mata pelajaran ini juga penting untuk mendukung perkembangan akademis mereka.

Future works yang dapat dilakukan meliputi penggunaan metode *machine learning* lainnya seperti *Random Forest* atau *Neural Networks* untuk membandingkan hasil dan mendapatkan pemahaman yang lebih mendalam tentang faktor-faktor yang mempengaruhi daya tarik pelajar terhadap mata pelajaran tertentu. Selain itu, pengembangan sistem pendukung keputusan yang menggunakan data ini dapat membantu guru dan pengelola sekolah dalam merancang kurikulum dan strategi pengajaran yang lebih efektif. Studi longitudinal juga perlu dilakukan untuk memantau perubahan minat pelajar terhadap mata pelajaran seiring berjalannya waktu, serta mengintegrasikan feedback langsung dari pelajar untuk memahami persepsi mereka dan menemukan cara untuk meningkatkan minat mereka. Dengan demikian, penelitian ini tidak hanya memberikan wawasan mengenai daya tarik pelajar terhadap mata pelajaran, tetapi juga menawarkan arah yang jelas untuk perbaikan dan pengembangan lebih lanjut.

6. DAFTAR PUSTAKA

- [1] D. Pristiwanti, B. Badariah, S. Hidayat, and R. S. Dewi, "Pengertian Pendidikan," *J. Pendidik. dan Konseling*, vol. 4, no. 6, pp. 337–347, 2022, doi: 10.33387/bioedu.v6i2.7305.
- [2] A. T. Alhasani, H. Alkattan, A. A. Subhi, E.-S. M. El-Kenawy, and M. M. Eid, "A Comparative Analysis of Methods for Detecting and Diagnosing Breast Cancer Based on Data Mining," *J. Artif. Intell. Metaheuristics*, vol. 4, no. 2, pp. 08–17, 2023, doi: 10.54216/jaim.040201.
- [3] R. A. Permana and S. Sahara, "Metode Support Vector Machine Sebagai Penentu Kelulusan Mahasiswa pada Pembelajaran Elektronik," *J. Khatulistiwa Inform.*,

- vol. 7, no. 1, 2019.
- [4] H. Saputra, “Analisis Sentimen Pada Vaksin Booster Menggunakan Algoritma Support Vector Machine Multiclass Di Twitter,” *J. Teknol. Pint.*, vol. 3, no. 10, pp. 1–26, 2023.
- [5] E. Suhartono, “Systematic Literatur Review (SLR): Metode , Manfaat , Dan Tantangan Learning Analytics Dengan Metode Data Mining di Dunia Pendidikan Tinggi,” *J. Ilm. INFOKAM*, vol. 13, no. 1, pp. 73–86, 2017.
- [6] N. Wanaspati, S. Anraeni, and H. Darwis, “Support Vector Machine untuk Analisis Sentimen Masyarakat Terhadap Penggunaan Antibiotik di Indonesia,” *Indones. J. Comput. Sci.*, vol. 12, no. 4, pp. 2196–2206, 2023, doi: 10.33022/ijcs.v12i4.3320.
- [7] F. K. Wattimury and E. Seniwati, “Penentuan Peminatan Mahasiswa Prodi Informatika di Universitas Amikom Yogyakarta Menggunakan SVM,” *INTECHNO Journal-Information Technol. J.*, vol. 1, no. 4, pp. 15–18, 2019.
- [8] M. D. Purbolaksono, M. I. Tantowi, A. I. Hidayat, and A. Adiwijaya, “Perbandingan support vector machine dan modified balanced random forest dalam deteksi pasien penyakit diabetes,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 393–399, 2021.
- [9] E. Indrayuni, “Analisa Sentimen Review Hotel Menggunakan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization,” *EVOLUSI J. Sains dan Manaj.*, vol. 4, no. 2, 2016.
- [10] S. Dwiasnati and Y. Devianto, “Optimasi Prediksi Bencana Banjir menggunakan Algoritma SVM untuk penentuan Daerah Rawan Bencana Banjir,” *Pros. SISFOTEK*, vol. 5, no. 1, pp. 202–207, 2021.