

Perbandingan Kinerja SVM dan Naive Bayes dalam Klasifikasi Data Evaluasi Pembelajaran

Muhammad Rizwan Hakimi^{1,a,*}, Nur Aulia Ramadhani², Daniel Pratama Wijaya³

^a Program Studi Teknik Informatika, Universitas Islam Al-Azhar, Mataram, Indonesia

^b Program Studi Sistem Informasi, Universitas Bumigora, Mataram, Indonesia

^c Program Studi Ilmu Komputer, Universitas Teknologi Mataram, Mataram, Indonesia

¹muhammadrizwanhakimi@gmail.com, ²nurauliaramadhani@gmail.com, ³danielpratamawijaya@gmail.com

*Corresponding author

ARTICLE INFO	ABSTRACT
Article history Received July 17, 2026 Revised July 18, 2026 Accepted July 30, 2026 Published July 31, 2026 Keywords Support Vector Machine Naive Bayes Classification Learning Evaluation Educational Data Mining  License by CC-BY-SA Copyright © 2026, The Author(s).	The application of machine learning in educational data analysis has become increasingly important in improving learning evaluation systems. Classification methods are widely used to identify student performance patterns and support academic decision-making. Among the most commonly applied algorithms are Support Vector Machine (SVM) and Naive Bayes, both of which have shown competitive performance in various classification tasks. However, their effectiveness in learning evaluation data classification still requires deeper empirical investigation. This study aims to compare the performance of SVM and Naive Bayes in classifying learning evaluation data based on accuracy, precision, recall, and F1-score. The study used a quantitative experimental approach with a dataset of 1,250 student learning evaluation records consisting of attendance, assignments, participation, midterm, and final examination scores. Data preprocessing included cleaning, normalization, and transformation before model training. The dataset was divided into 80% training data and 20% testing data, with 10-fold cross-validation for validation. The results indicate that SVM achieved an accuracy of 89.6%, precision of 88.9%, recall of 90.2%, and F1-score of 89.5%, outperforming Naive Bayes which obtained 84.3%, 83.7%, 85.1%, and 84.4% respectively. The findings confirm that SVM provides better performance for complex educational datasets and can be recommended for data-driven learning evaluation systems.
How to cite: Hakimi, M. R., Ramadhani, N. A., & Wijaya, D. P. (2026). Perbandingan Kinerja SVM dan Naive Bayes dalam Klasifikasi Data Evaluasi Pembelajaran. <i>Journal of Science and Technology: Alpha</i>, 2(3), 123-129. doi: https://doi.org/10.70716/alpha.v2i3.760	

PENDAHULUAN

Perkembangan teknologi digital telah mengubah cara institusi pendidikan mengelola proses pembelajaran dan evaluasi akademik. Sistem evaluasi pembelajaran yang sebelumnya dilakukan secara manual kini menghasilkan data dalam jumlah besar melalui platform digital, mulai dari nilai tugas, ujian, absensi, hingga aktivitas kelas. Data tersebut memiliki potensi besar untuk dianalisis sebagai dasar pengambilan keputusan akademik. Dalam konteks ini, pendekatan Machine Learning menjadi salah satu solusi utama untuk mengolah data pendidikan secara lebih sistematis dan objektif. Penggunaan machine learning dalam pendidikan mampu meningkatkan efektivitas analisis data, terutama dalam mengidentifikasi pola performa siswa. Akanbi (2023) menunjukkan bahwa algoritma Naive Bayes efektif dalam mengklasifikasikan performa siswa berdasarkan data akademik historis, sehingga dapat membantu institusi pendidikan dalam memprediksi hasil belajar dengan lebih cepat dan akurat.

Klasifikasi merupakan salah satu teknik utama dalam machine learning yang digunakan untuk mengelompokkan data berdasarkan karakteristik tertentu. Dalam konteks evaluasi pembelajaran, klasifikasi dapat digunakan untuk mengidentifikasi kategori capaian belajar siswa, seperti sangat baik, baik, cukup, atau

kurang. Penelitian Azhari, Situmorang, dan Rosnelly (2021) membandingkan beberapa algoritma klasifikasi, yaitu C4.5, Random Forest, SVM, dan Naive Bayes, dan menemukan bahwa setiap algoritma memiliki performa berbeda tergantung karakteristik data. Temuan ini menunjukkan bahwa pemilihan model klasifikasi tidak dapat digeneralisasi, tetapi harus disesuaikan dengan struktur dan distribusi dataset yang digunakan.

Di antara berbagai algoritma klasifikasi, Support Vector Machine (SVM) dan Naive Bayes merupakan dua metode yang paling sering digunakan karena memiliki karakteristik yang berbeda. SVM dikenal memiliki kemampuan tinggi dalam membangun hyperplane optimal untuk memisahkan data dengan kompleksitas tinggi. Eligo et al. (2022) menjelaskan bahwa SVM menunjukkan efisiensi klasifikasi yang baik pada masalah multiclass classification dengan akurasi yang konsisten. Penelitian Chrisinta dan Simarmata (2023) juga membuktikan bahwa SVM menghasilkan performa lebih baik dibanding Naive Bayes pada analisis sentimen evaluasi dosen. Temuan ini diperkuat oleh Ilmawan dan Mude (2020) yang menemukan bahwa SVM lebih unggul dalam klasifikasi ulasan tekstual pada platform digital.

Di sisi lain, Naive Bayes memiliki kelebihan pada efisiensi komputasi dan kesederhanaan model. Algoritma ini bekerja berdasarkan probabilitas posterior dengan asumsi independensi antaratribut. Dalam banyak kasus, metode ini mampu memberikan hasil klasifikasi yang cepat dengan biaya komputasi rendah. Howay dan Suhirman (2023) menemukan bahwa Naive Bayes mampu memberikan hasil yang stabil dalam penentuan jurusan siswa di sekolah menengah kejuruan. Penelitian Lestari et al. (2022) juga menunjukkan bahwa Naive Bayes cukup efektif dalam mengekstraksi opini mahasiswa terhadap layanan pendidikan tinggi. Namun, keterbatasan utama metode ini terletak pada asumsi independensi atribut yang sering kali tidak sesuai dengan kondisi data nyata.

Beberapa penelitian terdahulu telah membandingkan SVM dan Naive Bayes dalam berbagai domain, termasuk analisis sentimen, kesehatan, dan pendidikan. Khaira, Aryani, dan Hardian (2023) menemukan bahwa SVM memiliki akurasi lebih tinggi dibanding Naive Bayes dalam analisis sentimen kebijakan pendidikan selama pandemi. Pridiptama, Wasono, dan Amijaya (2024) juga melaporkan bahwa SVM lebih unggul dalam klasifikasi penyakit tekanan darah tinggi. Ramadhani dan Widyaningrum (2022) membuktikan bahwa SVM lebih efektif dalam klasifikasi predikat kelulusan mahasiswa. Hasil serupa ditemukan oleh Nasrulloh et al. (2025) yang membandingkan Naive Bayes, Decision Trees, dan SVM pada klasifikasi sentimen aplikasi digital. Secara umum, hasil-hasil tersebut menunjukkan kecenderungan bahwa SVM memiliki performa lebih stabil pada data dengan pola yang kompleks.

Meskipun demikian, tantangan utama dalam klasifikasi data pendidikan terletak pada distribusi data yang sering tidak seimbang. Zhang, Sadaoui, dan Mouhoub (2015) menjelaskan bahwa ketidakseimbangan data dapat menyebabkan bias klasifikasi dan menurunkan akurasi model. Penelitian Gunawan, Devianto, dan Sari (2024) menggunakan metode NearMiss untuk mengatasi masalah data tidak seimbang dan menunjukkan adanya peningkatan performa klasifikasi pada SVM maupun Naive Bayes. Selain itu, Khaldi et al. (2025) menekankan bahwa kemampuan adaptasi model terhadap kebutuhan praktis sangat menentukan keberhasilan implementasi dalam sistem nyata. Dalam konteks evaluasi pembelajaran, data seperti nilai tugas, kehadiran, dan partisipasi kelas sering memiliki distribusi yang tidak merata, sehingga pemilihan algoritma menjadi semakin penting.

Berdasarkan telaah literatur, sebagian besar penelitian sebelumnya masih berfokus pada analisis sentimen, klasifikasi teks, atau prediksi performa umum siswa. Penelitian yang secara spesifik membandingkan SVM dan Naive Bayes pada data evaluasi pembelajaran multidimensional masih terbatas. Hafidh et al. (2025) menunjukkan bahwa lingkungan pembelajaran berpengaruh terhadap pola klasifikasi gaya belajar siswa, sedangkan Hidayatunnisa'i, Kusriani, dan Kusnawi (2023) menekankan pentingnya evaluasi kualitas butir soal melalui metode klasifikasi yang tepat. Herijanto dan Wahyuningsih (2024) juga menunjukkan bahwa validasi silang mampu meningkatkan reliabilitas hasil klasifikasi. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja SVM dan Naive Bayes dalam klasifikasi data evaluasi pembelajaran menggunakan parameter accuracy, precision, recall, dan F1-score. Kontribusi utama penelitian

ini terletak pada penyediaan bukti empiris yang dapat menjadi dasar dalam pengembangan sistem evaluasi pembelajaran berbasis data di lingkungan pendidikan.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen komparatif untuk membandingkan performa algoritma Support Vector Machine (SVM) dan Naive Bayes dalam klasifikasi data evaluasi pembelajaran. Desain eksperimen dipilih karena mampu memberikan pengukuran objektif terhadap kinerja kedua model berdasarkan indikator evaluasi yang terukur. Pendekatan ini sejalan dengan penelitian Simarmata, Gerd-Wolfgang Weber, dan Chrisinta (2024) yang menggunakan evaluasi performa algoritma klasifikasi untuk mengukur efektivitas model pada data berskala besar.

Penelitian dilaksanakan pada periode Januari sampai April 2026 di beberapa sekolah menengah atas di Kota Mataram, Indonesia. Objek penelitian berupa data evaluasi pembelajaran siswa semester ganjil tahun ajaran 2025/2026. Populasi penelitian terdiri dari seluruh data akademik siswa sebanyak 1.250 record. Setiap record memuat lima variabel utama, yaitu nilai tugas, nilai ujian tengah semester, nilai ujian akhir semester, tingkat kehadiran, dan partisipasi kelas. Variabel target berupa kategori hasil evaluasi pembelajaran yang dibagi menjadi empat kelas, yaitu sangat baik, baik, cukup, dan kurang.

Teknik pengumpulan data dilakukan melalui dokumentasi dari sistem informasi akademik sekolah. Data yang diperoleh kemudian melalui tahap preprocessing untuk meningkatkan kualitas data sebelum proses klasifikasi. Tahapan preprocessing terdiri atas data cleaning, penghapusan duplikasi, penanganan missing value, transformasi atribut kategorikal menjadi numerik, serta normalisasi data menggunakan metode Min-Max Scaling. Tahapan ini penting karena kualitas data sangat memengaruhi hasil klasifikasi, sebagaimana dijelaskan oleh Azhari et al. (2021) dan Eligo et al. (2022).

Setelah preprocessing selesai, data dibagi menggunakan metode hold-out dengan komposisi 80% data training dan 20% data testing. Pembagian ini dipilih untuk menjaga keseimbangan antara proses pelatihan model dan pengujian akurasi. Selain itu, validasi model dilakukan menggunakan 10-Fold Cross Validation untuk memastikan konsistensi performa model. Teknik ini digunakan karena terbukti meningkatkan reliabilitas evaluasi model sebagaimana diterapkan oleh Herijanto dan Wahyuningsih (2024). Untuk mengatasi potensi ketidakseimbangan data, penelitian ini juga menerapkan pendekatan balancing sederhana berdasarkan distribusi kelas, merujuk pada pendekatan NearMiss yang digunakan oleh Gunawan et al. (2024).

Implementasi algoritma SVM dilakukan menggunakan kernel radial basis function (RBF) karena kernel ini efektif dalam menangani data non-linear. Pemilihan kernel RBF mengacu pada temuan Turgay et al. (2024) yang menunjukkan bahwa kernel trick meningkatkan efisiensi prediksi. Sementara itu, algoritma Naive Bayes menggunakan model Gaussian Naive Bayes karena sesuai untuk data numerik kontinu. Kedua model diimplementasikan menggunakan bahasa pemrograman Python dengan library Scikit-learn.

Evaluasi performa model dilakukan menggunakan confusion matrix dengan empat indikator utama, yaitu accuracy, precision, recall, dan F1-score. Penggunaan indikator ini mengacu pada standar evaluasi klasifikasi yang digunakan dalam penelitian Khaira et al. (2023), Pridiptama et al. (2024), dan Nasrulloh et al. (2025). Rumus evaluasi yang digunakan adalah sebagai berikut:

Accuracy	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	$Precision = \frac{TP}{TP + FP}$
Recall	$Recall = \frac{TP}{TP + FN}$
F1-score	$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$

Secara sistematis, alur penelitian dimulai dari pengumpulan data, preprocessing, pembagian dataset, pelatihan model, validasi menggunakan cross-validation, pengujian model, dan analisis hasil. Tahapan ini dirancang untuk menghasilkan evaluasi yang objektif dan valid sehingga dapat memberikan gambaran nyata mengenai keunggulan masing-masing algoritma dalam klasifikasi data evaluasi pembelajaran.

HASIL DAN PEMBAHASAN

Tahap pengujian dilakukan setelah seluruh data melalui proses preprocessing dan validasi model. Dataset sebanyak 1.250 record dibagi menjadi 1.000 data training dan 250 data testing. Pengujian menggunakan dua algoritma klasifikasi, yaitu Support Vector Machine (SVM) dan Naive Bayes, dengan skema 10-Fold Cross Validation untuk memastikan konsistensi hasil. Evaluasi dilakukan menggunakan empat parameter utama, yaitu accuracy, precision, recall, dan F1-score. Pendekatan evaluasi ini digunakan karena mampu memberikan gambaran menyeluruh terhadap kualitas model klasifikasi sebagaimana diterapkan oleh Simarmata et al. (2024) dan Herijanto dan Wahyuningsih (2024).

Hasil pengujian menunjukkan bahwa kedua algoritma mampu melakukan klasifikasi data evaluasi pembelajaran dengan tingkat akurasi yang cukup tinggi, namun terdapat perbedaan performa yang signifikan. SVM menunjukkan hasil lebih unggul dibandingkan Naive Bayes pada seluruh indikator evaluasi. Rincian hasil pengujian dapat dilihat pada Tabel 1.

Tabel 1. Perbandingan Hasil Pengujian Model

Algoritma	Accuracy	Precision	Recall	F1-Score
Support Vector Machine	89,60%	88,90%	90,20%	89,50%
Naive Bayes	84,30%	83,70%	85,10%	84,40%

Berdasarkan hasil pada Tabel 1, SVM memperoleh nilai accuracy sebesar 89,6%, lebih tinggi 5,3% dibandingkan Naive Bayes yang memperoleh 84,3%. Selisih ini menunjukkan bahwa SVM memiliki kemampuan lebih baik dalam memetakan pola data evaluasi pembelajaran yang kompleks. Hasil ini konsisten dengan penelitian Chrisinta dan Simarmata (2023) yang menunjukkan bahwa SVM lebih unggul dalam klasifikasi sentimen evaluasi dosen. Temuan serupa juga dilaporkan oleh Ilmawan dan Mude (2020), di mana SVM menghasilkan akurasi lebih tinggi dibanding Naive Bayes pada analisis sentimen ulasan aplikasi.

Pada aspek precision, SVM memperoleh nilai 88,9%, sementara Naive Bayes memperoleh 83,7%. Nilai precision yang lebih tinggi menunjukkan bahwa SVM lebih akurat dalam mengurangi kesalahan klasifikasi positif. Pada parameter recall, SVM juga unggul dengan nilai 90,2%, dibandingkan Naive Bayes sebesar 85,1%. Hal ini menunjukkan bahwa SVM lebih efektif dalam mendeteksi seluruh data yang benar-benar termasuk dalam kategori target. Temuan ini sejalan dengan penelitian Khaira et al. (2023) dan Pridiptama et al. (2024) yang menemukan bahwa SVM memiliki sensitivitas klasifikasi lebih baik pada data dengan banyak atribut.

Keunggulan SVM dalam penelitian ini dapat dijelaskan melalui kemampuannya membentuk hyperplane optimal pada ruang multidimensi. Data evaluasi pembelajaran memiliki hubungan antarvariabel yang cukup kompleks, misalnya antara kehadiran, partisipasi kelas, dan hasil ujian. SVM mampu menangani hubungan tersebut dengan lebih baik dibanding Naive Bayes yang bekerja berdasarkan asumsi independensi atribut. Menurut Khaldi et al. (2025), SVM memiliki tingkat adaptabilitas yang lebih tinggi pada data praktis dengan pola non-linear. Turgay et al. (2024) juga menegaskan bahwa penggunaan kernel trick dapat meningkatkan akurasi prediksi secara signifikan.

Sementara itu, performa Naive Bayes yang lebih rendah dalam penelitian ini tetap menunjukkan hasil kompetitif. Hal ini sejalan dengan penelitian Akanbi (2023) yang menunjukkan bahwa Naive Bayes masih efektif dalam klasifikasi performa siswa. Kelebihan utama Naive Bayes terletak pada kecepatan komputasi dan kesederhanaan implementasi. Namun, ketika dataset memiliki korelasi antaratribut yang kuat, asumsi independensi dapat mengurangi akurasi model. Temuan ini mendukung penelitian Howay dan Suhirman (2023) serta Lestari et al. (2022) yang menunjukkan bahwa Naive Bayes lebih cocok untuk data sederhana dengan struktur atribut yang tidak terlalu kompleks.

Dari sisi distribusi data, penelitian ini juga menemukan bahwa keseimbangan kelas memengaruhi performa kedua algoritma. Hal ini sejalan dengan penelitian Zhang, Sadaoui, dan Mouhoub (2015) yang menjelaskan bahwa imbalanced dataset dapat menyebabkan bias klasifikasi. Gunawan et al. (2024) membuktikan bahwa teknik balancing seperti NearMiss dapat meningkatkan performa model. Dalam penelitian ini, distribusi data yang relatif seimbang membantu kedua model mencapai performa optimal, meskipun SVM tetap menunjukkan keunggulan yang lebih konsisten.

Secara praktis, hasil penelitian ini memberikan implikasi bahwa institusi pendidikan dapat menggunakan SVM sebagai model utama dalam sistem evaluasi pembelajaran berbasis data. Implementasi model ini dapat mendukung sistem deteksi dini terhadap siswa dengan risiko penurunan performa akademik. Selain itu, hasil ini memperkuat temuan Ramadhani dan Widyaningrum (2022), Nasrulloh et al. (2025), dan Hafidh et al. (2025) bahwa SVM memiliki performa lebih stabil dalam berbagai konteks klasifikasi pendidikan. Dengan demikian, pemilihan algoritma yang tepat menjadi faktor penting dalam meningkatkan kualitas pengambilan keputusan di bidang pendidikan.

KESIMPULAN

Penelitian ini berhasil melakukan perbandingan kinerja algoritma Support Vector Machine (SVM) dan Naive Bayes dalam klasifikasi data evaluasi pembelajaran berdasarkan empat indikator evaluasi, yaitu accuracy, precision, recall, dan F1-score. Hasil pengujian menunjukkan bahwa SVM memiliki performa lebih unggul dibandingkan Naive Bayes pada seluruh parameter yang diuji. SVM memperoleh nilai accuracy sebesar 89,6%, precision 88,9%, recall 90,2%, dan F1-score 89,5%, sedangkan Naive Bayes memperoleh accuracy 84,3%, precision 83,7%, recall 85,1%, dan F1-score 84,4%. Temuan ini menunjukkan bahwa SVM lebih efektif dalam menangani pola data evaluasi pembelajaran yang kompleks dan memiliki keterkaitan antarvariabel.

Keunggulan SVM terletak pada kemampuannya membangun hyperplane optimal dalam ruang multidimensi sehingga mampu menghasilkan klasifikasi yang lebih stabil dan akurat. Sementara itu, Naive Bayes tetap memiliki keunggulan dari sisi efisiensi komputasi dan kemudahan implementasi, namun performanya cenderung menurun ketika diterapkan pada data dengan korelasi atribut yang tinggi. Hal ini mempertegas bahwa karakteristik dataset menjadi faktor utama dalam menentukan pemilihan algoritma klasifikasi.

Secara praktis, hasil penelitian ini dapat dijadikan dasar bagi institusi pendidikan dalam mengembangkan sistem evaluasi pembelajaran berbasis machine learning, terutama untuk mendukung deteksi dini terhadap potensi penurunan performa akademik siswa. Dari sisi teoretis, penelitian ini memperkuat literatur sebelumnya yang menunjukkan konsistensi keunggulan SVM pada klasifikasi data pendidikan multidimensi.

Penelitian ini masih memiliki keterbatasan pada jumlah variabel yang digunakan dan cakupan dataset yang masih terbatas pada satu wilayah. Oleh karena itu, penelitian selanjutnya disarankan menggunakan dataset yang lebih besar, lebih beragam, serta menambahkan algoritma lain seperti Random Forest, Decision Tree, atau Deep Learning. Selain itu, penerapan teknik balancing data yang lebih kompleks juga dapat dieksplorasi untuk meningkatkan akurasi klasifikasi pada data yang tidak seimbang.

DAFTAR PUSTAKA

- Akanbi, O. B. (2023). Application of Naive Bayes to students' performance classification. *Asian Journal of Probability and Statistics*, 25(1), 35–47. <https://doi.org/10.9734/ajpas/2023/v25i1536>
- Azhari, M., Situmorang, Z., & Rosnelly, R. (2021). Perbandingan akurasi, recall, dan presisi klasifikasi pada algoritma C4.5, Random Forest, SVM dan Naive Bayes. *Jurnal Media Informatika Budidarma*, 5(2), 640–651. <https://doi.org/10.30865/MIB.V5I2.2937>
- Chrisinta, D., & Simarmata, J. E. (2023). Comparative study of Support Vector Machine and Naive Bayes for sentiment analysis on lecturer performance. *Journal of Research in Mathematics Trends and Technology*, 5(1), 1–7. <https://doi.org/10.32734/jormtt.v5i1.15864>
- Eligo, W. M., Leng, C., Kurika, A. E., & Basu, A. (2022). Comparing supervised machine learning algorithms on classification efficiency of multiclass classifications problem. *International Journal of Emerging Trends in Engineering Research*, 10(6), 346–360. <https://doi.org/10.30534/ijeter/2022/081062022>
- Gunawan, W., Devianto, Y., & Sari, A. P. (2024). Imbalanced data NearMiss for comparison of SVM and Naive Bayes algorithms. *Computer Engineering and Applications Journal*, 13(3), 34–43. <https://doi.org/10.18495/comengapp.v13i03.485>
- Hafidh, M. H. F., Rahmadden, Pratama, A., Adrianto, S., & Cahyo, M. R. D. (2025). Komparasi algoritma C4.5, SVM, dan Naive Bayes untuk klasifikasi gaya belajar siswa SMK berdasarkan lingkungan pembelajaran. *Jurnal INSTEK (Informatika Sains dan Teknologi)*, 10(2), 350–361. <https://doi.org/10.24252/instek.v10i2.59910>
- Herijanto, C. K., & Wahyuningsih, Y. (2024). Perbandingan klasifikasi label tunggal untuk soal ujian fisika menggunakan Naïve Bayes dan K-Fold Cross Validation. *Jurnal Teknologi Terpadu*, 10(1), 40–45. <https://doi.org/10.54914/jtt.v10i1.1210>
- Howay, S., & Suhirman, S. (2023). Comparison of SVM, NBC, and KNN classification methods in determining students' majors at SMK N02 Manokwari. *Journal of Computer Science and Technology Studies*, 5(1), 15–23. <https://doi.org/10.32996/jcsts.2023.5.1.3>
- Hidayatunnisa'i, Kusrini, & Kusnawi. (2023). Perbandingan kinerja metode Naïve Bayes dan Support Machine (SVM) dalam analisis kualitas butir soal. *Jurnal Fasikom*, 13(2), 173–180. <https://doi.org/10.37859/jf.v13i02.5087>
- Ilmawan, L. B., & Mude, M. A. (2020). Perbandingan metode klasifikasi Support Vector Machine dan Naïve Bayes untuk analisis sentimen pada ulasan tekstual di Google Play Store. *ILKOM Jurnal Ilmiah*, 12(2), 154–161. <https://doi.org/10.33096/ILKOM.V12I2.597.154-161>
- Khaira, U., Aryani, R., & Hardian, R. W. (2023). Komparasi algoritma Naïve Bayes dan Support Vector Machine (SVM) pada analisis sentimen kebijakan Kemdikbudristek mengenai kuota internet selama Covid-19. *Jurnal Processor*, 18(2), 183–191. <https://doi.org/10.33998/processor.2023.18.2.897>
- Khalidi, M. I., Erraissi, A., Hain, M., & Banane, M. (2025). In-depth comparison of supervised classification models: Performance and adaptability to practical requirements. *International Journal of Advanced Computer Science and Applications*, 16(8), 624–635. <https://doi.org/10.14569/IJACSA.2025.0160862>
- Lestari, U., Romadhani, T., Suraya, S., & Fatkhayah, E. (2022). Sentiment analysis for extracting student opinion data on higher education services using the Naive Bayes classifier and Support Vector Machine methods: Case study AKPRIND Institute of Science and Technology Yogyakarta. *Jurnal TAM (Technology Acceptance Model)*, 13(1), 51–56. <https://doi.org/10.56327/jurnaltam.v13i1.1220>
- Nasrulloh, A., Yusuf, M., Mas'ud, I., Toifur, T., Ikhwanudin, A., & Syamhalim, A. (2025). Comparison of Naive Bayes, Decision Trees and SVM algorithms for sentiment classification of JMO applications. *CCIT Journal*, 18(2), 212–225. <https://doi.org/10.33050/ccit.v18i2.3510>

- Pridiptama, R. P., Wasono, W., & Amijaya, F. D. T. (2024). Perbandingan algoritma Support Vector Machine dan Naïve Bayes pada klasifikasi penyakit tekanan darah tinggi: Studi kasus Klinik Polresta Samarinda. *BASIS: Jurnal Ilmiah Matematika*, 3(1), 1–16. <https://doi.org/10.30872/basis.v3i1.1264>
- Ramadhani, L. K., & Widyaningrum, B. N. (2022). Perbandingan metode klasifikasi Naïve Bayes dan Support Vector Machine pada predikat kelulusan mahasiswa. *Ledger: Journal Informatic and Information Technology*, 1(3), 150–158. <https://doi.org/10.20895/ledger.v1i3.877>
- Simarmata, J. E., Weber, G.-W., & Chrisinta, D. (2024). Performance evaluation of classification methods on big data: Decision Trees, Naive Bayes, K-Nearest Neighbors, and Support Vector Machines. *Jurnal Matematika, Statistika dan Komputasi*, 20(3), 623–638. <https://doi.org/10.20956/j.v20i3.32970>
- Turgay, S., Han, M., Erdoğan, S., Kara, E. S., & Yilmaz, R. (2024). Evaluating the predictive modeling performance of Kernel Trick SVM, Market Basket Analysis and Naive Bayes in terms of efficiency. *WSEAS Transactions on Computers*, 23, 56–66. <https://doi.org/10.37394/23205.2024.23.6>
- Zhang, S., Sadaoui, S., & Mouhoub, M. (2015). An empirical analysis of imbalanced data classification. *Computer and Information Science*, 8(1), 151-162. <https://doi.org/10.5539/CIS.V8N1P151>