



Scripta Technica: Journal of Engineering and Applied Technology

Vol 2 No 1 June 2026, Hal. 353-364
ISSN:3110-0775(Print) ISSN: 3109-9696(Electronic)
Open Access: <https://scriptainteletektual.com/scripta-technica>

Implementasi Algoritma Naive Bayes untuk Klasifikasi Tingkat Stres Mahasiswa Berdasarkan Pola Perilaku Belajar

M. Fhatir Fadlillah Lubis^{1*}, Andhica Romanja Sipahutar², Daniel Pertogi Simbolon³, Elisabeth Givani⁴

¹⁻⁴ Universitas Mandiri Bina Prestasi, Indonesia

Email: elizabethgivani@gmail.com¹

Article Info :

Received:

12-06-2026

Revised:

22-06-2026

Accepted:

25-06-2026

Abstract

This study aims to implement the Gaussian Naive Bayes algorithm to classify students' stress levels based on learning behavior patterns as a machine learning-based early detection approach. The study employs an empirical experimental method using the "Student Stress and Study Behavior" dataset, which includes attributes such as study duration, sleep hours, academic workload, physical activity, social media use, and stress levels as the target variable. The research stages included data preprocessing, attribute transformation, dataset splitting using the hold-out validation method, classification model development, and evaluation using accuracy, precision, recall, and F1-score. The results show that the Gaussian Naive Bayes model achieved an accuracy of 91.27%, a precision of 90.84%, a recall of 91.02%, and an F1-score of 90.93%. The findings indicate that sleep duration and study duration are attributes that significantly contribute to the classification of students' stress levels. This study contributes to the development of an efficient, interpretable, and practical decision support system to support the monitoring of students' mental health.

Keywords: Gaussian Naive Bayes, Machine Learning, Stress Classification, Learning Behavior, Students.

Abstrak

Penelitian ini bertujuan untuk mengimplementasikan algoritma Gaussian Naive Bayes dalam mengklasifikasikan tingkat stres mahasiswa berdasarkan pola perilaku belajar sebagai pendekatan deteksi dini berbasis machine learning. Penelitian menggunakan metode eksperimen empiris dengan memanfaatkan dataset Student Stress and Study Behavior yang mencakup atribut durasi belajar, jam tidur, beban akademik, aktivitas fisik, penggunaan media sosial, dan tingkat stres sebagai variabel target. Tahapan penelitian meliputi prapemrosesan data, transformasi atribut, pembagian dataset menggunakan metode hold-out validation, pembangunan model klasifikasi, serta evaluasi menggunakan accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model Gaussian Naive Bayes memperoleh accuracy sebesar 91,27%, precision sebesar 90,84%, recall sebesar 91,02%, dan F1-score sebesar 90,93%. Temuan menunjukkan bahwa jam tidur dan durasi belajar merupakan atribut yang memiliki kontribusi signifikan terhadap klasifikasi tingkat stres mahasiswa. Penelitian ini berkontribusi dalam pengembangan sistem pendukung keputusan yang efisien, interpretatif, dan aplikatif untuk mendukung pemantauan kesehatan mental mahasiswa.

Kata kunci: Gaussian Naive Bayes, Machine Learning, Klasifikasi Stres, Perilaku Belajar, Mahasiswa.



©2022 Authors.. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.
(<https://creativecommons.org/licenses/by-nc/4.0/>)

PENDAHULUAN

Berbagai laporan internasional mengenai kesehatan mental mahasiswa menunjukkan bahwa stres akademik telah berkembang menjadi salah satu tantangan paling serius dalam ekosistem pendidikan tinggi modern, terutama setelah transformasi pembelajaran digital, meningkatnya intensitas kompetisi akademik, serta perubahan pola interaksi sosial yang memengaruhi perilaku belajar mahasiswa secara signifikan. Perubahan tersebut menghasilkan volume data perilaku belajar yang semakin besar, mulai dari durasi belajar, pola tidur, intensitas penggunaan perangkat digital, hingga konsistensi penyelesaian tugas, yang pada saat bersamaan membuka peluang bagi penerapan teknik data mining dan machine learning sebagai pendekatan prediktif dalam mendeteksi kondisi psikologis secara lebih objektif dibandingkan instrumen konvensional berbasis kuesioner. Algoritma Naive Bayes memperoleh perhatian yang luas karena memiliki kompleksitas komputasi yang rendah, kemampuan klasifikasi yang baik pada data multidimensi, serta performa yang relatif stabil meskipun jumlah data pelatihan terbatas.

Pergeseran paradigma dari pendekatan psikometri menuju pendekatan komputasional menunjukkan bahwa klasifikasi tingkat stres tidak lagi dipandang semata sebagai persoalan psikologi klinis, melainkan telah berkembang menjadi persoalan rekayasa data yang memerlukan integrasi antara teknik klasifikasi statistik, pengolahan data perilaku, dan pengambilan keputusan berbasis probabilistik. Perkembangan tersebut menempatkan penelitian mengenai implementasi algoritma klasifikasi sebagai salah satu agenda penting dalam bidang teknik informatika, khususnya pada pengembangan sistem pendukung keputusan yang mampu memberikan deteksi dini terhadap risiko gangguan kesehatan mental mahasiswa melalui analisis pola perilaku belajar secara otomatis (Widyawati et al., 2022; Susanti, 2024; Hasanah & Fatah, 2025).

Sejumlah penelitian terdahulu memperlihatkan bahwa algoritma machine learning telah digunakan dengan pendekatan dan karakteristik data yang beragam untuk memprediksi tingkat stres maupun kesehatan mental mahasiswa, namun hasil-hasil tersebut memperlihatkan kecenderungan metodologis yang berbeda. Penelitian Widyawati et al. (2022) dan Susanti (2024) menunjukkan bahwa Naive Bayes mampu menghasilkan klasifikasi tingkat stres yang baik pada lingkungan pembelajaran hybrid melalui pemanfaatan atribut perilaku akademik, sementara Megantara et al. (2023) memperluas pendekatan tersebut dengan mengintegrasikan domain expert sehingga proses klasifikasi tidak hanya bertumpu pada probabilitas statistik, tetapi juga mempertimbangkan pengetahuan pakar dalam menentukan tingkat stres mahasiswa tingkat akhir. Kajian yang dilakukan Sivi et al. (2025) memperlihatkan bahwa Naive Bayes juga memiliki potensi untuk memprediksi kesehatan mental mahasiswa menjelang penyusunan skripsi, sedangkan Simangunsong et al. (2026) mengembangkan implementasi Naive Bayes untuk mendeteksi burnout pada kelompok usia muda yang memiliki karakteristik psikologis lebih luas dibandingkan stres akademik semata. Pada sisi lain, penelitian Hasanah dan Fatah (2025) serta Megawati et al. (2026) memperlihatkan performa algoritma Support Vector Machine dalam klasifikasi stres dan kesehatan mental, sedangkan Rusdiansyah et al. (2026) memperkenalkan integrasi Natural Language Processing untuk mengeksplorasi informasi tekstual sebagai sumber prediksi kondisi psikologis mahasiswa. Sintesis terhadap berbagai penelitian tersebut menunjukkan bahwa perkembangan penelitian bergerak menuju eksplorasi berbagai algoritma klasifikasi dengan karakteristik data yang semakin kompleks, sehingga isu utama tidak lagi terbatas pada kemampuan menghasilkan prediksi, melainkan pada bagaimana karakteristik data perilaku belajar memengaruhi efektivitas masing-masing algoritma dalam menghasilkan klasifikasi yang akurat.

Meskipun perkembangan penelitian tersebut menunjukkan kemajuan yang cukup signifikan, sejumlah keterbatasan konseptual maupun empiris masih tampak jelas sehingga belum menghasilkan kerangka klasifikasi yang benar-benar komprehensif. Sebagian besar penelitian terdahulu masih berfokus pada konteks mahasiswa tingkat akhir, mahasiswa dalam pembelajaran hybrid, maupun mahasiswa yang sedang menghadapi penyusunan skripsi sehingga generalisasi terhadap perilaku belajar mahasiswa secara lebih luas belum memperoleh perhatian yang memadai (Megantara et al., 2023; Sivi et al., 2025; Megawati et al., 2026). Di sisi metodologis, sebagian penelitian lebih menitikberatkan pada perbandingan performa algoritma tanpa menguraikan hubungan probabilistik antara variabel perilaku belajar dengan tingkat stres sebagai dasar interpretasi hasil klasifikasi, sementara penelitian lain mengembangkan algoritma yang lebih kompleks seperti Support Vector Machine maupun Natural Language Processing tanpa mempertimbangkan kebutuhan implementasi praktis yang menuntut efisiensi komputasi tinggi dan kemudahan interpretasi model (Hasanah & Fatah, 2025; Rusdiansyah et al., 2026). Penelitian mengenai penanganan ketidakseimbangan data melalui pendekatan SMOTE juga menunjukkan bahwa kualitas prediksi sangat dipengaruhi oleh distribusi data, sehingga performa algoritma tidak dapat dinilai hanya berdasarkan tingkat akurasi semata tanpa mempertimbangkan karakteristik dataset yang digunakan (Surojudin et al., 2026). Kondisi tersebut memperlihatkan adanya celah penelitian berupa belum optimalnya integrasi antara pemilihan atribut perilaku belajar, efisiensi algoritma probabilistik, serta interpretabilitas model klasifikasi yang dapat digunakan sebagai dasar pengambilan keputusan pada lingkungan pendidikan tinggi.

Keberadaan celah tersebut memiliki implikasi ilmiah maupun praktis yang cukup besar karena keterlambatan dalam mengidentifikasi tingkat stres mahasiswa berpotensi meningkatkan risiko penurunan prestasi akademik, gangguan kesehatan mental, burnout, bahkan penghentian studi sebelum waktu kelulusan. Dari teknik informatika, kebutuhan terhadap sistem klasifikasi yang mampu melakukan deteksi dini secara cepat, ringan, dan mudah diimplementasikan menjadi semakin mendesak mengingat institusi pendidikan memerlukan solusi yang dapat diterapkan pada berbagai kondisi

infrastruktur teknologi tanpa beban komputasi yang tinggi. Algoritma Naive Bayes menawarkan karakteristik tersebut melalui pendekatan probabilistik yang sederhana, namun efektivitas implementasinya tetap sangat bergantung pada representasi variabel perilaku belajar yang digunakan sebagai atribut klasifikasi. Fakta bahwa berbagai penelitian sebelumnya menghasilkan pendekatan yang berbeda dalam menentukan variabel prediktor menunjukkan bahwa hubungan antara pola belajar mahasiswa dan tingkat stres belum sepenuhnya dipahami secara konsisten, sehingga diperlukan penelitian yang secara khusus mengevaluasi implementasi Naive Bayes dengan memanfaatkan atribut perilaku belajar sebagai dasar pembentukan model klasifikasi yang lebih representatif terhadap kondisi akademik mahasiswa (Widyawati et al., 2022; Susanti, 2024; Simangunsong et al., 2026; Surojudin et al., 2026).

Posisi penelitian ini berada pada irisan antara bidang machine learning, data mining, dan sistem pendukung keputusan untuk kesehatan mental mahasiswa dengan menempatkan pola perilaku belajar sebagai sumber informasi utama dalam proses klasifikasi tingkat stres. Berbeda dari penelitian sebelumnya yang lebih menitikberatkan pada evaluasi performa berbagai algoritma atau pada konteks mahasiswa tertentu, penelitian ini memfokuskan perhatian pada implementasi algoritma Naive Bayes sebagai model probabilistik yang mampu menjelaskan hubungan antara atribut perilaku belajar dengan probabilitas tingkat stres secara lebih sederhana dan mudah diinterpretasikan. Pendekatan tersebut diharapkan dapat memperkuat pemahaman mengenai kontribusi masing-masing variabel perilaku belajar terhadap hasil klasifikasi sekaligus menyediakan dasar empiris bagi pengembangan sistem deteksi dini yang dapat diimplementasikan secara luas pada lingkungan perguruan tinggi. Fokus tersebut juga mempertegas bahwa kontribusi penelitian tidak hanya terletak pada aspek peningkatan akurasi klasifikasi, tetapi juga pada penguatan interpretabilitas model sebagai karakteristik penting dalam penerapan machine learning pada domain pendidikan.

Penelitian ini bertujuan mengimplementasikan algoritma Naive Bayes untuk mengklasifikasikan tingkat stres mahasiswa berdasarkan pola perilaku belajar dengan memanfaatkan atribut-atribut akademik sebagai variabel prediktor dalam proses klasifikasi. Penelitian ini diharapkan menghasilkan model klasifikasi yang efisien, mudah diinterpretasikan, serta mampu mendukung pengembangan sistem deteksi dini tingkat stres mahasiswa. Dari sisi teoretis, penelitian ini memperluas pemahaman mengenai penerapan model probabilistik dalam klasifikasi perilaku akademik pada domain pendidikan tinggi. Dari sisi metodologis, penelitian ini menawarkan kerangka implementasi yang mengintegrasikan tahapan praproses data, pemodelan Naive Bayes, dan evaluasi hasil klasifikasi sebagai pendekatan yang dapat direplikasi untuk pengembangan penelitian maupun aplikasi serupa pada konteks pendidikan lainnya.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen empiris dengan tujuan mengimplementasikan algoritma *Gaussian Naive Bayes* untuk mengklasifikasikan tingkat stres mahasiswa berdasarkan pola perilaku belajar. Arsitektur sistem terdiri atas empat tahapan utama, yaitu akuisisi data, prapemrosesan (preprocessing), pemodelan, dan klasifikasi. Dataset yang digunakan merupakan dataset sekunder *Student Stress and Study Behavior* yang memuat atribut-atribut perilaku belajar mahasiswa, seperti durasi belajar, jam tidur, beban akademik, aktivitas fisik, penggunaan media sosial, serta variabel lain yang berhubungan dengan tingkat stres sebagai kelas target. Tahap prapemrosesan meliputi pemeriksaan nilai hilang, penghapusan data duplikat, transformasi variabel kategorikal menggunakan *Label Encoding*, serta normalisasi apabila diperlukan agar distribusi data lebih seragam. Seluruh proses implementasi dikembangkan menggunakan bahasa pemrograman Python dengan pustaka *Pandas* untuk manipulasi data, *Scikit-learn* untuk pemodelan *Gaussian Naive Bayes*, serta *NumPy* sebagai pendukung komputasi numerik. Setelah data diproses, dataset dipisahkan menjadi data latih dan data uji menggunakan skema *hold-out* dengan komposisi 80% untuk pelatihan model dan 20% untuk pengujian, kemudian model *Gaussian Naive Bayes* dibangun berdasarkan distribusi probabilitas masing-masing atribut terhadap kelas tingkat stres sehingga menghasilkan model klasifikasi yang dapat digunakan untuk memprediksi tingkat stres mahasiswa berdasarkan karakteristik perilaku belajarnya.

Kinerja model dievaluasi melalui proses pengujian terhadap data uji yang tidak terlibat selama tahap pelatihan sehingga kemampuan generalisasi model dapat diukur secara objektif. Validasi dilakukan menggunakan metode *hold-out validation* dengan nilai *random state* yang ditetapkan agar

eksperimen dapat direproduksi pada penelitian selanjutnya. Prediksi yang dihasilkan dibandingkan dengan label aktual untuk membentuk *confusion matrix* sebagai dasar perhitungan berbagai metrik evaluasi. Kinerja model dianalisis menggunakan nilai accuracy, precision, recall, dan F1-score, yang dihitung melalui pustaka *Scikit-learn* sehingga mampu menggambarkan kemampuan model dalam mengidentifikasi setiap kategori tingkat stres secara komprehensif. Laporan klasifikasi (*classification report*) digunakan untuk mengevaluasi performa setiap kelas secara individual, sedangkan analisis hasil difokuskan pada kemampuan algoritma dalam mengenali hubungan probabilistik antara pola perilaku belajar dan tingkat stres mahasiswa. Seluruh parameter eksperimen, perangkat lunak, tahapan prapemrosesan, skema pembagian data, serta prosedur evaluasi didokumentasikan secara rinci untuk menjamin konsistensi implementasi dan memungkinkan reproduktifitas penelitian pada dataset maupun lingkungan komputasi yang berbeda.

HASIL DAN PEMBAHASAN

Implementasi Algoritma Naive Bayes dan Karakteristik Dataset

Penelitian ini diawali dengan proses implementasi algoritma Gaussian Naive Bayes terhadap dataset *Student Stress and Study Behavior* yang telah melalui tahapan prapemrosesan sesuai rancangan metodologi. Dataset memuat sejumlah atribut perilaku belajar yang diasumsikan memiliki hubungan terhadap tingkat stres mahasiswa, meliputi durasi belajar, jam tidur, aktivitas fisik, penggunaan media sosial, beban akademik, dan label tingkat stres sebagai variabel target. Tahapan prapemrosesan berhasil menghilangkan data duplikat, menangani nilai yang tidak valid, serta mentransformasikan atribut kategorikal menjadi bentuk numerik menggunakan *Label Encoding*, sehingga seluruh variabel memenuhi kebutuhan proses komputasi algoritma Gaussian Naive Bayes. Efektivitas tahapan prapemrosesan merupakan faktor penting karena kualitas data masukan sangat memengaruhi probabilitas posterior yang dihitung oleh algoritma klasifikasi berbasis Bayes (Widyawati et al., 2022).

Distribusi data setelah prapemrosesan memperlihatkan bahwa setiap atribut memiliki variasi nilai yang cukup representatif sehingga memungkinkan pembentukan distribusi probabilitas pada masing-masing kelas tingkat stres. Variabel jam tidur dan durasi belajar menunjukkan variasi yang relatif tinggi dibandingkan atribut lain, sedangkan aktivitas fisik memperlihatkan penyebaran data yang lebih homogen. Kondisi tersebut memberikan indikasi bahwa atribut perilaku belajar memiliki karakteristik statistik yang berbeda sehingga kontribusinya terhadap proses klasifikasi tidak dapat diperlakukan secara identik. Karakteristik seperti ini sejalan dengan hasil penelitian Susanti (2024) yang menunjukkan bahwa perbedaan distribusi atribut perilaku akademik memberikan pengaruh terhadap pembentukan probabilitas kelas pada algoritma Naive Bayes.

Proses pembagian dataset menggunakan skema *hold-out validation* menghasilkan proporsi 80% sebagai data pelatihan dan 20% sebagai data pengujian sesuai rancangan eksperimen. Komposisi tersebut memberikan jumlah data pelatihan yang memadai untuk membangun distribusi probabilitas setiap atribut sekaligus menyediakan data pengujian yang independen guna mengevaluasi kemampuan generalisasi model. Pemisahan data dilakukan menggunakan nilai *random state* yang tetap sehingga setiap proses eksperimen dapat direplikasi dengan hasil yang konsisten. Pendekatan validasi semacam ini banyak digunakan dalam penelitian klasifikasi kesehatan mental karena mampu memberikan estimasi performa model secara objektif tanpa meningkatkan kompleksitas komputasi secara berlebihan (Sivi et al., 2025).

Model Gaussian Naive Bayes kemudian dibangun melalui proses estimasi probabilitas prior dan probabilitas kondisional dari setiap atribut terhadap kelas tingkat stres. Setiap atribut diperlakukan sebagai variabel independen sehingga probabilitas posterior dihitung berdasarkan asumsi independensi bersyarat antarfitur sebagaimana prinsip dasar Teorema Bayes. Pendekatan tersebut menghasilkan model klasifikasi yang relatif sederhana namun tetap mampu memproses data berdimensi cukup banyak dengan waktu komputasi yang efisien. Karakteristik tersebut menjadikan Naive Bayes masih relevan digunakan pada sistem klasifikasi kesehatan mental yang membutuhkan kecepatan inferensi tinggi dibandingkan algoritma yang lebih kompleks (Megantara et al., 2023).

Hasil eksplorasi data menunjukkan bahwa beberapa atribut memberikan kecenderungan hubungan yang lebih kuat terhadap kategori tingkat stres dibandingkan atribut lainnya. Jam tidur yang lebih pendek dan durasi belajar yang tidak konsisten lebih sering muncul pada kelompok mahasiswa dengan tingkat stres tinggi, sedangkan pola belajar yang stabil cenderung ditemukan pada kategori stres rendah. Hubungan tersebut menunjukkan bahwa perilaku belajar memiliki nilai diskriminatif yang

cukup baik dalam proses pembentukan model klasifikasi berbasis probabilistik. Temuan tersebut memiliki konsistensi dengan penelitian Simangunsong et al. (2026) yang menjelaskan bahwa perubahan pola aktivitas harian merupakan indikator penting dalam identifikasi kondisi psikologis berbasis machine learning.

Tabel 1. Karakteristik Dataset Setelah Tahap Prapemrosesan

Variabel	Jenis Data	Metode Prapemrosesan	Peran dalam Model
Sleep Hours	Numerik	Normalisasi	Fitur
Study Hours	Numerik	Normalisasi	Fitur
Academic Workload	Kategorikal	Label Encoding	Fitur
Physical Activity	Kategorikal	Label Encoding	Fitur
Social Media Usage	Numerik	Normalisasi	Fitur
Stress Level	Kategorikal	Label Encoding	Target

Karakteristik pada Tabel 1 memperlihatkan bahwa seluruh atribut telah berhasil dikonversi ke dalam format numerik sehingga dapat diproses oleh Gaussian Naive Bayes tanpa memerlukan transformasi tambahan pada tahap pelatihan model. Dominasi atribut perilaku belajar menunjukkan bahwa penelitian ini lebih menekankan hubungan antara aktivitas akademik mahasiswa dengan kondisi psikologis dibandingkan penggunaan indikator biologis maupun data klinis. Struktur dataset tersebut memberikan keuntungan karena variabel yang digunakan relatif mudah diperoleh melalui survei akademik sehingga memungkinkan implementasi sistem klasifikasi pada berbagai lingkungan perguruan tinggi. Pendekatan serupa juga digunakan oleh Arnita et al. (2025) yang menunjukkan bahwa atribut perilaku sehari-hari memiliki kontribusi signifikan dalam membentuk model klasifikasi tingkat stres menggunakan algoritma Naive Bayes.

Nilai probabilitas yang dihasilkan selama proses pelatihan menunjukkan bahwa setiap atribut memiliki kontribusi yang berbeda terhadap pembentukan kelas tingkat stres. Atribut jam tidur dan durasi belajar menghasilkan distribusi probabilitas posterior yang lebih dominan dibandingkan penggunaan media sosial maupun aktivitas fisik sehingga keduanya menjadi indikator utama dalam proses klasifikasi. Perbedaan kontribusi tersebut mengindikasikan bahwa Gaussian Naive Bayes tidak hanya memanfaatkan jumlah atribut, tetapi juga mempertimbangkan kekuatan distribusi probabilistik masing-masing variabel terhadap kelas target. Karakteristik tersebut sejalan dengan penelitian Tsania et al. (2026) yang menjelaskan bahwa atribut perilaku harian memberikan bobot probabilistik lebih tinggi dibandingkan variabel pendukung ketika digunakan untuk memprediksi kondisi kesehatan mental mahasiswa.

Analisis terhadap hasil pelatihan memperlihatkan bahwa model mampu membentuk batas klasifikasi yang stabil meskipun setiap atribut memiliki distribusi yang berbeda. Asumsi independensi bersyarat yang digunakan oleh Gaussian Naive Bayes tidak menyebabkan penurunan performa secara signifikan karena pola hubungan antaratribut masih dapat direpresentasikan melalui distribusi probabilitas masing-masing fitur. Kondisi tersebut memperlihatkan bahwa karakteristik dataset cukup sesuai dengan pendekatan probabilistik sehingga proses inferensi dapat dilakukan secara efisien tanpa memerlukan optimasi model yang kompleks. Temuan tersebut mendukung hasil penelitian Erkamim et al. (2025) yang menunjukkan bahwa Naive Bayes tetap mampu menghasilkan klasifikasi yang kompetitif pada data perilaku meskipun menggunakan asumsi independensi antarvariabel.

Implementasi model juga memperlihatkan efisiensi komputasi yang tinggi selama proses pelatihan maupun prediksi. Waktu pelatihan berlangsung relatif singkat karena estimasi parameter hanya memerlukan perhitungan rerata dan varians setiap atribut pada masing-masing kelas. Efisiensi tersebut memberikan keuntungan bagi pengembangan sistem deteksi dini yang memerlukan respons cepat ketika jumlah data pengguna terus bertambah. Kondisi tersebut memiliki kesesuaian dengan hasil penelitian Jananto (2025) yang menempatkan Naive Bayes sebagai salah satu algoritma data mining dengan kompleksitas komputasi rendah namun tetap mampu mempertahankan kualitas klasifikasi pada berbagai jenis data pendidikan.

Karakteristik model yang sederhana juga memberikan keuntungan dari aspek interpretabilitas karena setiap hasil prediksi dapat ditelusuri berdasarkan probabilitas posterior setiap atribut. Interpretabilitas tersebut menjadi nilai tambah dibandingkan sejumlah algoritma yang menghasilkan akurasi tinggi tetapi sulit menjelaskan dasar pengambilan keputusan model. Lingkungan pendidikan membutuhkan model yang dapat dipahami oleh pengambil kebijakan sehingga hasil klasifikasi dapat diterjemahkan menjadi strategi intervensi akademik secara lebih rasional. Perspektif tersebut memiliki kemiripan dengan penelitian Chanafy dan Sulistiani (2026) yang menekankan pentingnya keseimbangan antara akurasi klasifikasi dan kemudahan interpretasi model dalam sistem prediksi kesehatan mental.

Hasil implementasi memperlihatkan bahwa model telah memenuhi tujuan utama penelitian, yaitu menghasilkan sistem klasifikasi yang mampu memanfaatkan pola perilaku belajar sebagai dasar identifikasi tingkat stres mahasiswa. Kemampuan tersebut menunjukkan bahwa hubungan antara perilaku akademik dan kondisi psikologis dapat direpresentasikan menggunakan pendekatan probabilistik yang sederhana namun informatif. Nilai praktis dari implementasi ini terletak pada kemudahan integrasi model ke dalam sistem informasi akademik tanpa memerlukan sumber daya komputasi yang besar. Karakteristik tersebut memperkuat relevansi penerapan machine learning pada bidang pendidikan sebagaimana ditunjukkan dalam penelitian Afrillia et al. (2025) yang menegaskan bahwa model klasifikasi yang efisien memiliki peluang implementasi lebih luas pada sistem pemantauan kesehatan mental berbasis institusi.

Evaluasi Kinerja Model Klasifikasi dan Analisis Hasil Prediksi

Tahap evaluasi dilakukan setelah model Gaussian Naive Bayes selesai dibangun menggunakan data pelatihan dan selanjutnya diuji pada data pengujian yang tidak terlibat selama proses pembentukan model. Pendekatan tersebut bertujuan untuk mengukur kemampuan generalisasi model ketika dihadapkan pada data baru yang belum pernah dipelajari sebelumnya. Hasil pengujian menunjukkan bahwa model menghasilkan *accuracy* sebesar 91,27%, yang mengindikasikan kemampuan klasifikasi berada pada kategori sangat baik untuk dataset perilaku belajar mahasiswa. Tingkat akurasi tersebut memperlihatkan bahwa pendekatan probabilistik masih memiliki daya prediksi yang kompetitif pada permasalahan klasifikasi tingkat stres sebagaimana dijelaskan oleh Widyawati et al. (2022).

Evaluasi tidak hanya berfokus pada nilai akurasi karena distribusi prediksi pada setiap kelas juga memengaruhi kualitas model secara keseluruhan. Nilai *precision* mencapai 90,84%, sedangkan *recall* sebesar 91,02%, sehingga menunjukkan bahwa model mampu mengenali sebagian besar data aktual pada masing-masing kategori tingkat stres tanpa menghasilkan proporsi kesalahan klasifikasi yang tinggi. Keseimbangan kedua metrik tersebut mengindikasikan bahwa proses pembelajaran model tidak mengalami kecenderungan dominan terhadap salah satu kelas tertentu. Karakteristik evaluasi seperti ini dipandang lebih representatif dibandingkan penggunaan akurasi sebagai satu-satunya indikator performa klasifikasi (Chanafy & Sulistiani, 2026).

Nilai F1-score sebesar 90,93% memperlihatkan bahwa harmonisasi antara *precision* dan *recall* berada pada tingkat yang stabil selama proses pengujian. Kondisi tersebut menunjukkan bahwa model mampu mempertahankan keseimbangan antara kemampuan mengidentifikasi data positif dan meminimalkan kesalahan prediksi pada kelas lain. Stabilitas tersebut menjadi salah satu indikator bahwa distribusi probabilitas yang dibangun selama proses pelatihan mampu merepresentasikan karakteristik data perilaku belajar secara memadai. Temuan ini memiliki kesesuaian dengan penelitian Arnita et al. (2025) yang melaporkan bahwa Gaussian Naive Bayes menghasilkan performa optimal ketika atribut yang digunakan memiliki hubungan langsung dengan karakteristik psikologis responden.

Proses evaluasi juga menunjukkan bahwa model tidak mengalami indikasi *overfitting* yang berarti selama eksperimen berlangsung. Perbedaan performa antara data pelatihan dan data pengujian berada pada rentang yang relatif kecil sehingga kemampuan generalisasi model tetap dapat dipertahankan. Kondisi tersebut menunjukkan bahwa jumlah data pelatihan mampu membentuk distribusi probabilitas yang cukup representatif tanpa menyebabkan model terlalu menyesuaikan diri terhadap karakteristik data latih. Fenomena tersebut sejalan dengan hasil penelitian Jananto (2025) yang menjelaskan bahwa algoritma Naive Bayes memiliki risiko *overfitting* yang relatif rendah dibandingkan beberapa algoritma klasifikasi berbasis pembelajaran kompleks.

Interpretasi terhadap hasil evaluasi memperlihatkan bahwa performa model tidak hanya dipengaruhi oleh algoritma yang digunakan, tetapi juga oleh kualitas proses prapemrosesan data.

Penghapusan data yang tidak konsisten serta transformasi atribut kategorikal berhasil mengurangi potensi kesalahan selama proses estimasi probabilitas posterior. Hubungan tersebut menunjukkan bahwa peningkatan kualitas data memberikan dampak langsung terhadap stabilitas hasil klasifikasi yang dihasilkan model. Temuan tersebut memperkuat pandangan Megantara et al. (2023) bahwa efektivitas Naive Bayes sangat dipengaruhi oleh representasi atribut sebelum proses pembelajaran dilakukan.

Tabel 2. Hasil Evaluasi Kinerja Model Gaussian Naive Bayes

Metrik Evaluasi Nilai (%) Interpretasi		
Accuracy	91,27	Sangat baik
Precision	90,84	Prediksi positif akurat
Recall	91,02	Identifikasi kelas efektif
F1-Score	90,93	Keseimbangan precision dan recall
Error Rate	8,73	Tingkat kesalahan rendah

Hasil pada Tabel 2 memperlihatkan bahwa seluruh metrik evaluasi berada di atas 90%, sedangkan *error rate* hanya mencapai 8,73%, sehingga menunjukkan bahwa model memiliki tingkat kesalahan klasifikasi yang relatif rendah. Keseimbangan nilai accuracy, precision, recall, dan F1-score mengindikasikan bahwa performa model tidak hanya tinggi secara numerik, tetapi juga konsisten pada setiap aspek evaluasi klasifikasi. Pola tersebut memperlihatkan bahwa pendekatan Gaussian Naive Bayes mampu memanfaatkan hubungan probabilistik antaratribut perilaku belajar secara efektif untuk menghasilkan keputusan klasifikasi yang stabil. Hasil tersebut memiliki kemiripan dengan penelitian Hasanah dan Fatah (2025), meskipun penelitian tersebut menggunakan Support Vector Machine yang memiliki kompleksitas komputasi lebih tinggi dibandingkan Naive Bayes.

Analisis terhadap *confusion matrix* menunjukkan bahwa sebagian besar data uji berhasil ditempatkan pada kelas yang sesuai dengan label aktualnya. Kesalahan klasifikasi masih ditemukan pada sejumlah sampel yang memiliki karakteristik perilaku belajar berada di antara kategori stres sedang dan tinggi sehingga probabilitas posterior kedua kelas menjadi relatif berdekatan. Kondisi tersebut menggambarkan bahwa batas antar kelas pada data psikologis cenderung bersifat gradual dan tidak selalu dapat dipisahkan secara tegas menggunakan satu atribut dominan. Temuan tersebut memiliki kesesuaian dengan penelitian Surojudin et al. (2026) yang menjelaskan bahwa distribusi data kesehatan mental sering menghasilkan area klasifikasi yang saling tumpang tindih sehingga memengaruhi tingkat kesalahan prediksi.

Hasil evaluasi memperlihatkan bahwa atribut jam tidur memberikan kontribusi probabilistik paling besar terhadap proses pengambilan keputusan model. Durasi belajar dan beban akademik menempati urutan berikutnya sebagai variabel yang memiliki daya diskriminasi tinggi dalam membedakan kategori tingkat stres mahasiswa. Aktivitas fisik dan penggunaan media sosial tetap memberikan pengaruh terhadap hasil klasifikasi, meskipun bobot probabilitasnya relatif lebih rendah dibandingkan atribut utama. Pola tersebut memperkuat hasil penelitian Tsania et al. (2026) yang menyatakan bahwa perubahan pola tidur merupakan indikator yang paling konsisten dalam berbagai model prediksi kesehatan mental berbasis machine learning.

Performa model juga menunjukkan bahwa pendekatan Gaussian Naive Bayes masih mampu bersaing dengan algoritma klasifikasi lain yang memiliki kompleksitas lebih tinggi. Support Vector Machine sering menghasilkan akurasi yang sedikit lebih tinggi pada data berdimensi kompleks, tetapi memerlukan proses optimasi parameter yang lebih intensif dan waktu komputasi yang lebih panjang. Naive Bayes menawarkan keseimbangan antara efisiensi komputasi, kemudahan implementasi, dan interpretabilitas hasil sehingga lebih sesuai untuk pengembangan sistem deteksi dini pada lingkungan akademik yang membutuhkan respons cepat. Perspektif tersebut sejalan dengan hasil penelitian Kurniawan et al. (2026) serta Megawati et al. (2026) yang menempatkan Support Vector Machine sebagai model berakurasi tinggi, namun dengan kebutuhan komputasi yang lebih besar dibandingkan pendekatan probabilistik.

Perbandingan terhadap algoritma klasifikasi lainnya juga memperlihatkan bahwa karakteristik dataset memiliki pengaruh yang lebih besar dibandingkan pemilihan algoritma secara semata. Penelitian mengenai Random Forest, K-Nearest Neighbor, Decision Tree, maupun pendekatan berbasis Natural Language Processing menunjukkan bahwa peningkatan performa model sering kali dipengaruhi oleh kualitas atribut, distribusi data, serta strategi prapemrosesan yang digunakan sebelum pelatihan dilakukan. Kondisi tersebut memperlihatkan bahwa keberhasilan sistem klasifikasi tidak hanya ditentukan oleh kompleksitas algoritma, melainkan oleh kesesuaian antara karakteristik data dan mekanisme pembelajaran model. Argumentasi tersebut didukung oleh Widjaya et al. (2025), Adhelia et al. (2026), Utomo et al. (2026), Rusdiansyah et al. (2026), dan Afrillia et al. (2025) yang menunjukkan bahwa kualitas representasi data merupakan determinan utama dalam peningkatan performa klasifikasi.

Interpretasi hasil evaluasi mengindikasikan bahwa implementasi Gaussian Naive Bayes telah memenuhi tujuan penelitian sebagai model klasifikasi yang efisien, stabil, dan memiliki kemampuan generalisasi yang baik terhadap data perilaku belajar mahasiswa. Nilai evaluasi yang konsisten pada berbagai metrik memperlihatkan bahwa model mampu menghasilkan prediksi yang dapat dijadikan dasar pengembangan sistem pendukung keputusan untuk deteksi dini tingkat stres di lingkungan perguruan tinggi. Karakteristik tersebut menunjukkan bahwa pendekatan probabilistik tetap memiliki relevansi tinggi dalam penelitian machine learning terapan, khususnya ketika sistem memerlukan transparansi proses pengambilan keputusan dan efisiensi komputasi. Kesimpulan ini memiliki keselarasan dengan hasil penelitian Sarita et al. (2024), Hasanah dan Fatah (2025), serta Simangunsong et al. (2026) yang menegaskan bahwa model klasifikasi berbasis machine learning dapat menjadi instrumen efektif dalam mendukung identifikasi kondisi kesehatan mental pada populasi mahasiswa.

Interpretasi Ilmiah, Perbandingan dengan Penelitian Terdahulu, dan Implikasi Implementasi Sistem

Implementasi Gaussian Naive Bayes pada penelitian ini memperlihatkan bahwa hubungan antara pola perilaku belajar dan tingkat stres mahasiswa dapat direpresentasikan secara efektif melalui pendekatan probabilistik. Model mampu mengidentifikasi kecenderungan perubahan tingkat stres berdasarkan kombinasi beberapa atribut perilaku tanpa memerlukan proses optimasi parameter yang kompleks. Karakteristik tersebut menunjukkan bahwa distribusi probabilitas setiap atribut telah mampu menangkap pola penting yang membedakan kategori stres rendah, sedang, dan tinggi secara konsisten. Interpretasi ini mendukung pandangan Widyawati et al. (2022) bahwa algoritma Naive Bayes memiliki kemampuan yang baik dalam memodelkan hubungan antara perilaku akademik mahasiswa dengan kondisi psikologis menggunakan pendekatan probabilitas bersyarat.

Hasil penelitian memperlihatkan bahwa atribut jam tidur dan durasi belajar memiliki kontribusi yang lebih dominan dibandingkan atribut pendukung lainnya dalam membentuk keputusan klasifikasi. Dominasi kedua variabel tersebut mengindikasikan bahwa perubahan pola belajar tidak hanya memengaruhi performa akademik, tetapi juga menjadi indikator dini terhadap munculnya tekanan psikologis. Hubungan tersebut menunjukkan adanya keterkaitan antara pengelolaan waktu belajar dan kemampuan mahasiswa mempertahankan keseimbangan aktivitas sehari-hari selama proses perkuliahan berlangsung. Temuan tersebut memiliki kesesuaian dengan penelitian Susanti (2024) serta Sivi et al. (2025) yang menunjukkan bahwa perubahan pola belajar merupakan salah satu prediktor utama dalam klasifikasi kesehatan mental mahasiswa.

Perspektif probabilistik yang digunakan oleh Gaussian Naive Bayes memberikan keunggulan pada aspek interpretabilitas hasil dibandingkan sejumlah algoritma klasifikasi yang lebih kompleks. Setiap keputusan klasifikasi dapat dijelaskan berdasarkan probabilitas posterior dari masing-masing atribut sehingga proses pengambilan keputusan model menjadi lebih transparan bagi pengguna sistem. Transparansi tersebut memiliki nilai penting ketika hasil prediksi digunakan sebagai dasar penyusunan kebijakan akademik maupun layanan konseling di lingkungan perguruan tinggi. Argumentasi tersebut sejalan dengan Megantara et al. (2023) yang menekankan bahwa interpretasi model merupakan aspek penting dalam pengembangan sistem diagnosis berbasis machine learning pada bidang kesehatan mental.

Kinerja model yang diperoleh juga menunjukkan bahwa pendekatan sederhana tidak selalu menghasilkan performa yang lebih rendah dibandingkan metode klasifikasi yang memiliki struktur komputasi lebih kompleks. Beberapa penelitian memang melaporkan bahwa algoritma seperti Support

Vector Machine mampu memberikan peningkatan akurasi pada karakteristik data tertentu, namun peningkatan tersebut sering diikuti oleh kebutuhan optimasi parameter yang lebih rumit dan waktu pelatihan yang lebih panjang. Efisiensi komputasi yang dimiliki Gaussian Naive Bayes memberikan keuntungan praktis bagi institusi pendidikan yang membutuhkan sistem klasifikasi dengan sumber daya perangkat keras yang terbatas. Pola tersebut konsisten dengan hasil penelitian Hasanah dan Fatah (2025), Megawati et al. (2026), serta Kurniawan et al. (2026) yang memperlihatkan bahwa pemilihan algoritma perlu mempertimbangkan keseimbangan antara performa prediksi dan kompleksitas implementasi.

Analisis terhadap hasil penelitian menunjukkan bahwa kualitas representasi data memiliki pengaruh yang lebih besar terhadap keberhasilan model dibandingkan kompleksitas algoritma itu sendiri. Variabel yang mampu merepresentasikan kondisi akademik mahasiswa secara akurat menghasilkan distribusi probabilitas yang lebih stabil sehingga meningkatkan kualitas keputusan klasifikasi. Fenomena tersebut memperlihatkan bahwa tahapan prapemrosesan merupakan fondasi utama sebelum proses pelatihan model dilakukan pada berbagai pendekatan machine learning. Kesimpulan tersebut memperkuat hasil penelitian Chanafy dan Sulistiani (2026), Widjaya et al. (2025), serta Jananto (2025) yang menjelaskan bahwa kualitas data menjadi determinan utama dalam peningkatan performa model klasifikasi berbasis data mining.

Tabel 3. Perbandingan Karakteristik Gaussian Naive Bayes dengan Algoritma Klasifikasi Lain

Algoritma	Kompleksitas Komputasi	Interpretabilitas	Kebutuhan Optimasi	Kesesuaian Deteksi Dini	untuk
Gaussian Naive Bayes	Rendah	Tinggi	Rendah	Sangat Tinggi	
Support Vector Machine	Tinggi	Sedang	Tinggi	Tinggi	
Random Forest	Sedang	Sedang	Sedang	Tinggi	
K-Nearest Neighbor	Sedang	Tinggi	Rendah	Sedang	
Decision Tree	Rendah	Tinggi	Rendah	Tinggi	

Karakteristik pada Tabel 3 memperlihatkan bahwa Gaussian Naive Bayes memiliki keseimbangan yang baik antara efisiensi komputasi dan kemudahan interpretasi dibandingkan beberapa algoritma klasifikasi lainnya. Keunggulan tersebut menjadikan algoritma ini lebih sesuai untuk sistem pemantauan kesehatan mental yang memerlukan proses inferensi cepat serta hasil prediksi yang mudah dipahami oleh tenaga pendidik maupun konselor. Perbedaan karakteristik antaralgoritma menunjukkan bahwa tidak terdapat satu metode yang selalu unggul pada seluruh kondisi data, sehingga pemilihan model harus mempertimbangkan tujuan implementasi dan karakteristik dataset yang digunakan. Perspektif tersebut didukung oleh penelitian Utomo et al. (2026), Erkamim et al. (2025), Adhelia et al. (2026), Arnita et al. (2025), serta Rusdiansyah et al. (2026) yang memperlihatkan bahwa efektivitas algoritma klasifikasi sangat dipengaruhi oleh konteks penerapan dan struktur data yang dianalisis.

Penerapan Gaussian Naive Bayes dalam penelitian ini memperlihatkan bahwa pendekatan probabilistik masih memiliki relevansi yang kuat untuk mendukung sistem deteksi dini tingkat stres mahasiswa di lingkungan pendidikan tinggi. Model tidak hanya menghasilkan prediksi yang akurat, tetapi juga menyediakan mekanisme inferensi yang mudah dipahami karena setiap keputusan didasarkan pada peluang kemunculan atribut terhadap masing-masing kelas. Karakteristik tersebut memberikan keuntungan bagi institusi pendidikan yang membutuhkan sistem klasifikasi yang dapat dijelaskan secara rasional kepada pengelola akademik maupun tenaga konseling. Perspektif ini memiliki kesesuaian dengan penelitian Simangunsong et al. (2026) yang menjelaskan bahwa interpretabilitas model menjadi aspek penting dalam implementasi sistem prediksi kesehatan mental berbasis machine learning.

Analisis lebih lanjut menunjukkan bahwa implementasi model dapat dikembangkan menjadi bagian dari sistem peringatan dini (*early warning system*) yang terintegrasi dengan sistem informasi

akademik. Data mengenai pola belajar mahasiswa dapat diproses secara berkala sehingga perubahan perilaku yang mengarah pada peningkatan tingkat stres dapat diidentifikasi lebih cepat dibandingkan evaluasi konvensional berbasis observasi manual. Pendekatan tersebut memungkinkan institusi memberikan intervensi akademik maupun layanan konseling sebelum kondisi psikologis mahasiswa berkembang menjadi gangguan yang lebih serius. Temuan tersebut memperkuat hasil penelitian Tsania et al. (2026) serta Sarita et al. (2024) yang menunjukkan bahwa deteksi dini berbasis data perilaku memberikan peluang lebih besar dalam mendukung pencegahan gangguan kesehatan mental pada mahasiswa.

Dari perspektif rekayasa perangkat lunak, model yang dibangun memiliki karakteristik modular sehingga mudah diintegrasikan ke dalam aplikasi berbasis web maupun sistem akademik yang telah digunakan perguruan tinggi. Tahapan prapemrosesan, pelatihan model, dan proses inferensi dapat dikembangkan sebagai layanan yang bekerja secara otomatis ketika data perilaku mahasiswa diperbarui. Struktur implementasi tersebut mendukung pengembangan sistem yang bersifat adaptif terhadap penambahan data baru tanpa memerlukan perubahan besar pada arsitektur aplikasi. Pendekatan serupa telah diperlihatkan oleh Rusdiansyah et al. (2026) melalui integrasi model klasifikasi pada sistem berbasis *Natural Language Processing*, meskipun penelitian ini memilih pendekatan probabilistik yang lebih sederhana dan efisien.

Meskipun hasil penelitian menunjukkan performa yang baik, terdapat sejumlah keterbatasan yang perlu menjadi perhatian dalam pengembangan penelitian berikutnya. Penggunaan satu jenis dataset menyebabkan karakteristik distribusi data masih terbatas pada populasi tertentu sehingga kemampuan generalisasi model terhadap lingkungan akademik yang berbeda memerlukan pengujian lanjutan. Asumsi independensi antaratribut yang digunakan oleh Gaussian Naive Bayes juga berpotensi mengabaikan hubungan kompleks antarvariabel perilaku belajar yang dalam praktiknya dapat saling memengaruhi. Kondisi tersebut sejalan dengan temuan Surojudin et al. (2026), Widjaya et al. (2025), dan Chanafy dan Sulistiani (2026) yang menunjukkan bahwa karakteristik dataset dan distribusi kelas memiliki pengaruh signifikan terhadap performa berbagai algoritma machine learning.

Hasil penelitian membuka peluang bagi pengembangan model klasifikasi yang lebih komprehensif melalui integrasi atribut perilaku belajar dengan data akademik, data linguistik, maupun indikator psikologis lainnya agar kemampuan prediksi dapat terus ditingkatkan. Pendekatan hibrida yang mengombinasikan Gaussian Naive Bayes dengan teknik *feature selection*, penanganan ketidakseimbangan data, atau metode *ensemble learning* berpotensi menghasilkan model yang lebih robust terhadap variasi karakteristik data. Pengembangan tersebut juga dapat diarahkan pada implementasi sistem pemantauan kesehatan mental secara real-time sehingga proses identifikasi risiko berlangsung lebih adaptif terhadap perubahan perilaku mahasiswa. Arah pengembangan tersebut selaras dengan rekomendasi penelitian Afrillia et al. (2025), Adhelia et al. (2026), Utomo et al. (2026), Erkamim et al. (2025), dan Jananto (2025) yang menekankan pentingnya integrasi berbagai pendekatan machine learning untuk meningkatkan akurasi sekaligus mempertahankan interpretabilitas model dalam aplikasi pendidikan.

KESIMPULAN

Penelitian ini membuktikan bahwa implementasi algoritma Gaussian Naive Bayes mampu menghasilkan model klasifikasi tingkat stres mahasiswa berdasarkan pola perilaku belajar dengan tingkat performa yang tinggi dan stabil. Proses implementasi yang melibatkan tahapan prapemrosesan data, pembentukan model probabilistik, serta evaluasi kinerja menunjukkan bahwa karakteristik atribut perilaku belajar memiliki pengaruh penting terhadap kemampuan model dalam membedakan kategori tingkat stres mahasiswa. Hasil pengujian memperlihatkan bahwa model memperoleh nilai accuracy sebesar 91,27%, precision sebesar 90,84%, recall sebesar 91,02%, dan F1-score sebesar 90,93%, yang menunjukkan bahwa algoritma mampu melakukan prediksi secara efektif dengan tingkat kesalahan klasifikasi yang rendah. Analisis lebih lanjut menunjukkan bahwa variabel jam tidur dan durasi belajar menjadi faktor dengan kontribusi paling dominan dalam pembentukan keputusan klasifikasi, sehingga pola keseimbangan aktivitas akademik dan kebiasaan harian dapat menjadi indikator penting dalam identifikasi dini risiko stres mahasiswa. Selain memiliki kemampuan prediktif yang baik, Gaussian Naive Bayes juga memberikan keunggulan berupa efisiensi komputasi dan interpretabilitas model yang mendukung penerapan pada sistem pendukung keputusan di lingkungan pendidikan tinggi. Meskipun demikian, pengembangan lebih lanjut diperlukan melalui penggunaan dataset yang lebih beragam,

optimasi fitur, serta integrasi metode machine learning lanjutan untuk meningkatkan kemampuan adaptasi dan generalisasi model dalam mendeteksi kondisi psikologis mahasiswa secara lebih komprehensif.

DAFTAR PUSTAKA

- Adhelia, M., Wiwenar, G. R., & Wafa, H. S. (2026). Klasifikasi Tingkat Depresi Mahasiswa Menggunakan Algoritma Decision Tree C4. 5. *JUSIFOR: Jurnal Sistem Informasi dan Informatika*, 5(1), 219-226. <https://doi.org/10.70609/jusifor.v5i1.8831>.
- Afrillia, Y., Fhonna, R. P., & Rahma, M. (2025). Klasifikasi Kesehatan Mental Remaja Tingkat SMA di Kota Lhokseumawe Menggunakan Algoritma Random Forest. *Jurnal Pengabdian kepada Masyarakat Nusantara*, 6(3), 4019-4024. <https://doi.org/10.55338/jpkmn.v6i3.6377>.
- Arnita, A., Simamora, E., Indra, Z., & Manullang, S. (2025). Implementasi Algoritma Naïve Bayes Dalam Melakukan Klasifikasi Tingkat Stres Siswa SMA. *Jati (Jurnal Mahasiswa Teknik Informatika)*, 9(4), 6655-6660. <https://doi.org/10.36040/jati.v9i4.14173>.
- Chanafy, M., & Sulistiani, H. (2026). Analisis Perbandingan Kinerja Algoritma Machine Learning Untuk Classifikasi Kesehatan Mental Mahasiswa. *Smatika Jurnal*, 16(01), 19-29. <https://doi.org/10.32664/smatika.v16i01.2080>.
- Erkamim, M., Nurhayati, N., Heriyani, N., & Riyanto, U. Klasifikasi Kepribadian Introvert dan Extrovert Menggunakan Random Forest, Naïve Bayes, dan K-Nearest Neighbor. *Jurnal Ilmiah Fifo*, 17(2), 202-211. <https://doi.org/10.22441/fifo.2025.v17i2.009>.
- Hasanah, U., & Fatah, Z. (2025). Prediksi Tingkat Stress dan Kesehatan Mental Mahasiswa Menggunakan Algoritma SVM. *Jurnal Mahasiswa Teknik Informatika*, 4(2), 200-207. <https://doi.org/10.35473/jamastika.v4i2.4542>.
- Jananto, A. (2025). Studi Komparatif Algoritma Klasifikasi Data Mining pada Prediksi Prestasi Siswa Berbasis Data Sosiodemografis. *Jurnal Teknik Informatika UNIKA Santo Thomas*, 231-242. <https://doi.org/10.54367/jtiust.v10i2.5840>.
- Kurniawan, W., Putra, O. V., & Utami, A. (2026). Implementasi Algoritma Support Vectore Machine Untuk Klasifikasi Kesehatan Mental Mahasiswa Tingkat Akhir pada Proses Penyusunan Skripsi Berdasarkan Filsafat AL-WUJ? D dalam Islam. *Jurnal Ilmiah IT CIDA*, 11(2), 69-83. <https://doi.org/10.55635/jic.v11i2.295>.
- Megantara, Y. D., Latipah, L., & Falani, A. Z. (2023). Pemanfaatan Metode Naïve Bayes Untuk Diagnosis Tingkat Stres Pada Mahasiswa Tingkat Akhir Berbasis Domain Expert. *Spirit*, 15(1). <http://dx.doi.org/10.53567/spirit.v15i1.276>.
- Megawati, M., Widyaningsih, M., Hadi, A., & Safithry, E. A. (2026). Sistem Klasifikasi Tingkat Stres Mahasiswa Tingkat Akhir Berbasis Web Menggunakan Support Vector Machine. *Jurnal Sains dan Informatika*, 12(1), 21-30. <https://doi.org/10.34128/jsi.v12i1.1877>.
- Rusdiansyah, I., Pangestu, R., Azalia, D., Zhafran, M. F., Saputra, F., & Amsury, F. (2026). Integrasi Model Klasifikasi Tingkat Stress Mahasiswa Berbasis Natural Language Processing. *RIGGS: Journal of Artificial Intelligence and Digital Business*, 4(4), 7823-7831. <https://doi.org/10.31004/riggs.v4i4.4388>.
- Sarita, J., Zainiyah, I. L. L., Maylani, I., & Setyawati, P. (2024, October). Pengaruh Kesehatan Mental Terhadap Stress Akademik Pada Mahasiswa Di Universitas X Dengan Menggunakan Metode Algoritma Decision Tree J48. In *Prosiding Seminar Nasional Sains dan Teknologi Terapan* (No. 1).
- Simangunsong, A., Simanjourang, R. M., Hasugian, P. S., Amalia, F., & Khairunnisa, P. (2026). Analisis Dan Deteksi Burnout Pada Kalangan Anak Muda Dengan Algoritma Naïve Bayes. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 5(1), 158-166. <https://doi.org/10.53513/jursi.v5i1.12473>.
- Sivi, N. A., Dwiatmoko, F., Utami, D., & Hafiz, A. (2025). Penerapan Metode Naive Bayes untuk Memprediksi Tingkat Kesehatan Mental Mahasiswa UNU Lampung Menjelang Skripsi. *Jurnal informasi dan Komputer Vol*, 13(1). <https://doi.org/10.35959/jik.v13i01.687>.
- Surojudin, N., Butsianto, S., & Firmansyah, A. (2026). Perbandingan Kinerja Naïve Bayes dengan dan Tanpa SMOTE untuk Klasifikasi Gangguan Kecemasan Mahasiswa pada Data Tidak Seimbang. *Bulletin of Computer Science Research*, 6(2), 804-812. <https://doi.org/10.47065/bulletincsr.v6i2.1021>.

- Susanti, L. (2024). Klasifikasi Tingkat Stress pada Mahasiswa Teknik Informatika dalam Melakukan Perkuliahan Metode Hybrid Menggunakan Algoritma Naive Bayes. *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 8(3), 243-248. <http://dx.doi.org/10.30998/string.v8i3.17096>.
- Tsania, K. N. B., Wardani, R., & Hanifah, A. I. (2026). Prediksi Pengaruh HP Terhadap Kesehatan Mental Tingkat Kecemasan dan Insomnia Mahasiswa Universitas Islam Lamongan Menggunakan Metode Naive Bayes. *Jurnal Ilmiah Sistem Informasi*, 5(1), 156-165. <https://doi.org/10.51903/syn92p96>.
- Utomo, W. C., Riono, S. H., Triono, W., & Mukharrom, M. S. (2026). Prediksi Tingkat Stress Siswa Berbasis Machine Learning Menggunakan Algoritma K-Nearest Neighbors (KNN). *Nexus: Journal of Information & Computational Science*, 1(1), 31-39.
- Widjaya, F. A., Ngete, M. S., & Tumanggor, D. S. (2025). Perbandingan Algoritma Random Forest dengan K-Nearest Neighbor pada Klasifikasi Tingkat Stress Pelajar. *Jatissi*, 12(4). <https://doi.org/10.35957/jatissi.v12i4.13790>.
- Widyawati, N., Khasanah, M., Rasywir, E., & Feranika, A. (2022). Prediksi Tingkat Stress Pada Mahasiswa Universitas Dinamika Bangsa Jambi Dalam Melakukan Perkuliahan Metode Hybrid Menggunakan Algoritma Naive Bayes. *Jurnal Manajemen Teknologi Dan Sistem Informasi (JMS)*, 2(1), 99-109. <https://doi.org/10.33998/jms.2022.2.1.44>.