

Prediction of Student Academic Performance Using Machine Learning Approach

[Prediksi Kinerja Akademik Mahasiswa Menggunakan Pendekatan Machine Learning]

Berliana Aulia Agustin¹⁾, Uce Indahyanti²⁾, Suprianto³⁾, Irwan Alnarus Kautsar⁴⁾

1) Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

2) Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

3) Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

4) Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

Abstract. *One of the most important factors in improving educational standards is predicting students' work habits. The purpose of this research is to use machine learning to evaluate the work habits of university students. Three classification algorithms - Naive Bayes, Decision Tree, and Support Vector Machine (SVM) - were used to analyze students' work-related data, including midterms, final exams, assignments, and presentations. The data came from the activity log of an e-learning master's program in information studies. The findings showed that the Decision Tree algorithm provided the highest accuracy of 98.51% with 60% of the data used for training and 40% for analysis. The SVM algorithm also performed well with 98.00% accuracy, while the Naive Bayes algorithm achieved 95.99% accuracy. This research provides insight into the potential of machine learning in assessing students' work habits academically and can be used to improve the quality of education and develop systems to detect irregularities.*

Keywords - Machine Learning, Academic Performance, Naive Bayes, Decision Tree, Support Vector Machine.

Abstrak. *Salah satu faktor terpenting dalam meningkatkan standar pendidikan adalah memprediksi kebiasaan kerja mahasiswa. Tujuan dari penelitian ini adalah untuk menggunakan machine learning untuk mengevaluasi kebiasaan kerja mahasiswa. Tiga algoritma klasifikasi - Naive Bayes, Decision Tree, dan Support Vector Machine (SVM) - digunakan untuk menganalisis data yang berhubungan dengan pekerjaan dari para mahasiswa, termasuk UTS, UAS, tugas-tugas, dan presentasi. Data berasal dari log aktivitas program magister e-learning di bidang studi informasi. Temuan penelitian menunjukkan bahwa algoritma Decision Tree memberikan akurasi tertinggi sebesar 98,51% dengan 60% data yang digunakan untuk pelatihan dan 40% untuk analisis. Algoritma SVM juga memberikan kinerja yang baik dengan akurasi 97,01%, sedangkan algoritma Naive Bayes mencapai akurasi 95,99%. Penelitian ini memberikan wawasan tentang potensi pembelajaran mesin dalam menilai kebiasaan kerja mahasiswa secara akademis dan dapat digunakan untuk meningkatkan kualitas pendidikan dan mengembangkan sistem untuk mendeteksi penyimpangan.*

Kata Kunci – Machine Learning, Kinerja Akademik, Naive Bayes, Decision Tree, Support Vector Machine.

I. PENDAHULUAN

Salah satu faktor terpenting dalam pembangunan suatu bangsa adalah pendidikan. Tujuan utama dari institusi pendidikan tinggi adalah memberikan pendidikan berkualitas tinggi kepada para mahasiswa. Metode untuk meningkatkan kualitas yang lebih sesuai dengan kebutuhan pendidikan adalah dengan menekankan kinerja mahasiswa dalam penilaian akademik. Dengan melaksanakan hal ini, organisasi kemudian dapat menggunakan hasil penilaian untuk mengembangkan sistem peringatan dini untuk meningkatkan kinerja siswa dalam pekerjaan dan kualitas pendidikan yang tersedia[1].

Mahasiswa merupakan subjek yang sangat penting bagi keberhasilan akademik. Salah satu faktor utama yang mempengaruhi tingkat akademik adalah kemampuan siswa untuk menyelesaikan studi mereka tepat waktu, namun hampir di setiap tahun ajaran baru, ketika jumlah siswa meningkat, terdapat beberapa mahasiswa yang tidak dapat menyelesaikan studi mereka tepat waktu, yang mengakibatkan peningkatan jumlah data mahasiswa[2]. Dengan data yang diambil berupa data mata kuliah algoritma pemrograman dan algoritma struktur data.

Beberapa penelitian yang fokus pada prediksi pekerjaan mahasiswa telah dilakukan, baik dalam pendidikan tradisional maupun pembelajaran online[3]. Machine learning digunakan untuk teknik data mining, yang merupakan metodologi yang ketat yang telah menunjukkan hasil dalam domain klasifikasi dan prediksi. Salah satu tugas pembelajaran mesin yang paling umum adalah memprediksi variabel target dalam data yang sebelumnya tidak dapat

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.

Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

diamati: klasifikasi. Tujuan dari klasifikasi adalah untuk memperkirakan variabel target (kelas) dengan mengembangkan model klasifikasi berdasarkan kumpulan data pelatihan. Model ini kemudian diterapkan untuk memperkirakan kelas data uji[1].

Prediksi kinerja mahasiswa dalam mata kuliah inti memiliki efek penting dalam bidang pendidikan untuk meningkatkan hasil akademik terutama untuk keberhasilan dalam sebuah studi mahasiswa, yang berdampak pada masa depan juga pada reputasi dan kualitas pendidikan[4]. Kinerja akademik berperan dalam membantu pengembangan program pembelajaran yang efektif. Misalnya dengan mengetahui pola pikir mahasiswa melalui metode pengenalan metode pengajaran yang spesifik dengan kebutuhan mahasiswa[4].

Penelitian sebelumnya yang dilakukan oleh Ariawan, dkk [5] dalam penelitiannya algoritma k-means clustering digunakan untuk mendapatkan nilai kuliah yang maksimal, mahasiswa harus meningkatkan tingkat akademis mereka dengan model prediksi yang secara akurat menentukan ambang batas kelulusan mereka dengan menggunakan sistem penilaian tertentu.

Penelitian tambahan yang dilakukan oleh Indahyanti, dkk [3] menjelaskan tentang evaluasi kinerja akademik mahasiswa dengan menggunakan beberapa algoritma klasifikasi, termasuk Decision Tree, Random Forest, dan AdaBoost dengan menggunakan teknik root mean square error (RMSE). Tiga tingkat prediksi pekerjaan akademik (tingkat penguasaan) dihasilkan dari proses pemodelan di setiap mata kuliah atau semester: tingkat pemula, tingkat menengah, dan tingkat lanjut. Hasil pemodelan menunjukkan bahwa algoritma RF menghasilkan prediksi yang lebih akurat daripada algoritma Decision Tree dan AdaBoost, dengan rata-rata 75,79% dan RMSE 0,44, masing-masing di bawah ambang batas.

Penelitian yang dilakukan oleh Rohman, dkk [6] menggunakan algoritma Decision Tree untuk menganalisis data dari program magister Universitas Pandanaran untuk menghasilkan model dan aturan.

Oleh karena itu, tujuan dari penelitian ini adalah untuk memperkirakan rata-rata kinerja mahasiswa dengan pendekatan tiga algoritma serupa, seperti Naive Bayes Classifier, Decision Tree, dan Support Vector Machine. Ketiga algoritma tersebut diperlukan untuk mencari akurasi terbaik. Selain itu, diharapkan temuan penelitian ini dapat menjadi panduan untuk penelitian di masa depan tentang hubungan antara tingkat keberhasilan mahasiswa dan tingkat kelulusan, yang mana hal ini adalah suatu yang paling penting.

II. METODE PENELITIAN

A. Alokasi dan Waktu

Penelitian ini dilakukan pada tanggal 17 Mei – 12 Februari 2025 pada prodi Informatika, Universitas Muhammadiyah Sidoarjo

B. Alat dan Bahan

Alat :

1. Komputer atau PC
2. Ms Excel
3. Program Python

Bahan :

1. Data nilai mahasiswa yang meliputi nilai UTS, UAS, Tugas, Presensi, Aslab sebanyak 300 data.
2. Data kategori mahasiswa yang termasuk aslab atau non aslab

C. Metodologi Penelitian

Dalam penelitian ini, peneliti akan menjelaskan metodologi yang digunakan dalam analisis prediksi kinerja akademik mahasiswa menggunakan teknik machine learning. Metodologi ini mencakup berbagai proses sistematis, termasuk pengumpulan data, pra-pemrosesan, pengembangan dan pelatihan model pembelajaran mesin, serta evaluasi dan validasi hasil prediksi.



Gambar 1. Diagram Tahap Penelitian

1. Pengumpulan Data

Pengumpulan data sekunder dilakukan dengan menggunakan data kelulusan mahasiswa program studi informatika pada mata kuliah Algoritma Pemrograman dan Algoritma Struktur Data yang diperoleh dari log aktivitas elearning umsida. Variable yang digunakan antara lain nilai uts, nilai uas, nilai tugas, dan presensi. Peneliti memilih beberapa variable diatas karena peneliti mendapatkan data dari log aktivitas machine learning[4].

2. Identifikasi Data

Data yang digunakan berupa data kinerja mahasiswa berupa log aktivitas elearning untuk mata kuliah Algoritma Pemrograman dan Algoritma Struktur Data. Data yang diperoleh kemudian diperiksa untuk memastikan apakah masih terdapat kekurangan serta apakah data tersebut telah sesuai dengan permasalahan.

3. Persiapan Data

Data pada penelitian ini memiliki beberapa atribut. Atribut yang relevan dengan kinerja akademik mahasiswa yaitu nilai uts, uas, tugas, dan presensi yang akan diteliti. Dataset penelitian sebanyak 300.

4. Pemodelan Klasifikasi

Pemilihan model klasifikasi yang digunakan pada penelitian ini sangat penting untuk mendapatkan hasil yang akurat. Beberapa model klasifikasi yang digunakan seperti :

1. Naïve Bayes
2. Decision Tree
3. Support Vector Machine (SVM)

5. Evaluasi dan Validasi Model

Analisis data menggunakan pemrograman Python dengan menggunakan teknik confusion matrix dan F1 Score untuk memberika akurasi hasil kinerja model yang telah dievaluasi berdasarkan parameter-parameter tersebut[7].

6. Hasil dan Pembasan

Menganalisis hasil pemodelan klasifikasi dengan pendekatan mencari akurasi terbaik dari masing masing algoritma yang digunakan , Menyusun laporan hasil penelitian yang mencakup metode, eksperimen, dan hasil.

D. Draft Output

1. Pelabelan data menjadi tiga klasifikasi yaitu:
 - a. Lulus
 - b. Tidak lulus
 - c. Remidi
2. Akurasi pemodelan dalam presentase

III. HASIL DAN PEMBAHASAN

Data Selection

Sumber data awal yang digunakan dalam penelitian ini adalah data nilai mahasiswa pada mata kuliah algoritma pemrograman dan algoritma struktur data fakultas sains dan teknologi Universitas Muhammadiyah Sidoarjo tahun 2021-2022. Data nilai mahasiswa terdiri dari presensi, keaktifan, tugas, uts, uas, na, aslab, dan ukm yang didapat dari log aktivitas machine learning[8].

Preprocessing Data

Berikut data mentah nilai mahasiswa pada log aktivitas machine learning Universitas Muhammadiyah Sidoarjo. Dataset berisikan atribut data mentah yang di dapat dari log aktivitas machine learning. Data nilai mahasiswa yang diperoleh seperti Presensi, Keaktifan, Tugas, UTS, UAS, NA, Aslab, dan UKM. Untuk lebih jelas seperti tabel di bawah ini.

Tabel 3.1 Atribut Data Mentah

Presensi	Keaktifan	Tugas	UTS	UAS	NA	Aslab	UKM	Lulus
75	75	70	75	75	73.8	Tidak	Tidak	Remidi
75	75	81	70	75	75.3	Tidak	Tidak	Lulus
...	Lulus
100	85	79	80	75	81.6	Tidak	Tidak	Lulus
100	85	85	83	83	85.9	Tidak	Tidak	Lulus

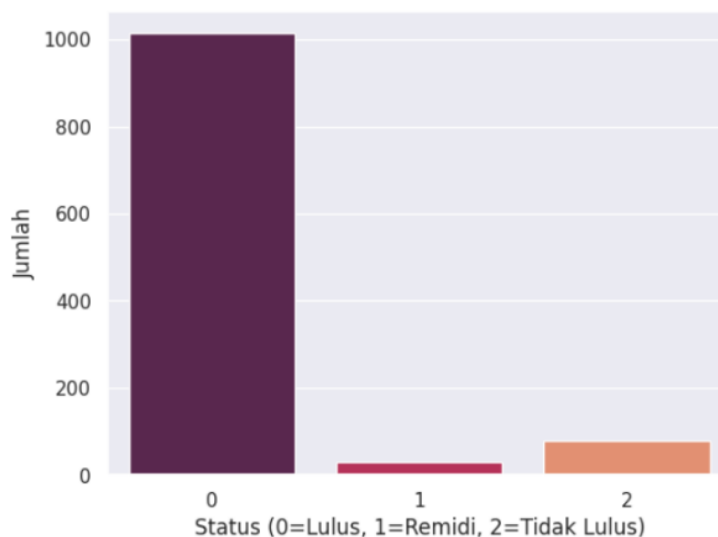
Transformasi

Selanjutnya data tersebut dijadikan data numerik, data yang akan diolah seperti terlihat pada Tabel 3.1

Tabel 3.2 Data Numerik

Presensi	Keaktifan	Tugas	UTS	UAS	NA	Aslab	UKM	Lulus
75	75	70	75	75	73.8	0	0	1
75	75	81	70	75	75.3	0	0	0
...	0
100	85	79	80	75	81.6	0	0	0
100	85	85	83	83	85.9	0	0	0

Berikut merupakan grafik distribusi data berdasarkan atribut mata kuliah algoritma pemrograman dasar dan algoritma struktur data di fakultas sains dan teknologi Universitas Muhammadiyah Sidoarjo dan jenis kelulusan pada tiap mahasiswa.



Gambar 2. Grafik Persebaran Data

Proses Klasifikasi Naïve Bayes

Rumus umum Teorema Bayes menjadi dasar Naive Bayes

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

Keterangan :

X = Sampel data yang memiliki *class* (label) yang tidak diketahui.

C = Hipotesis bahwa **X** adalah data *class* (label).

P(C) = Probabilitas hipotesis **C**.

P(X) = Peluang dari data sampel yang diamati (probabilitas **C**).

P(X|C) = Probabilitas berdasarkan kondisi pada hipotesis.

Proses klasifikasi Naïve Bayes, yang pertama adalah pemanggilan library, kedua import dataset, ketiga memisahkan fitur dan label, keempat membagi data training dan testing, kelima membangun model naïve bayes, keenam latih model pada data pengujian, ketujuh melakukan prediksi dan menghitung akurasi terhadap evaluasi model dengan hasil ini sudah melakukan pengujian meliputi nilai accuracy, nilai precision, nilai recall dan nilai f1-score[9][10].

Tabel 3.3 Hasil Model Naïve Bayes Akurasi Terbaik

	Precision	Recall	F1-Score	Score
0	1.00	0.95	0.98	400
1	0.75	1.00	0.86	12
2	0.73	1.00	0.84	37
Accuracy	0.83	0.98	0.96	449
Macro Avg	0.97	0.96	0.89	449
Weighted Avg			0.84	449
Akurasi NB : 95.99%				

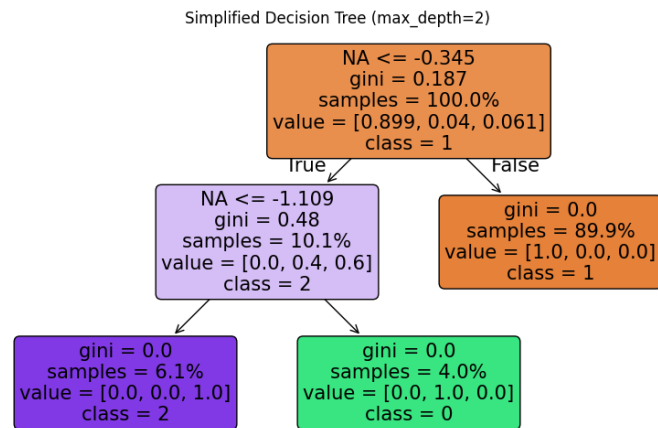
Pemodelan Klasifikasi

- **Naïve Bayes**

Proses pemodelan klasifikasi menggunakan metode Naïve Bayes menunjukkan hasil akurasi yang bervariasi berdasarkan pembagian data antara data training dan data testing. Pada pembagian 30% data testing dan 70% data training, akurasi yang diperoleh adalah 95.85%. Sementara itu, pada pembagian 40% data testing dan 60% data training, akurasi meningkat signifikan hingga mencapai 95.99%[11].

Proses Klasifikasi Decision Tree

Berikut merupakan pohon keputusan (Decision Tree) yang dibuat menggunakan algoritma Decision Tree Classifier dengan kedalaman maksimum (`max_depth=2`)[12]. Berikut adalah penjelasan elemen-elemen utama dalam pohon keputusan tersebut:

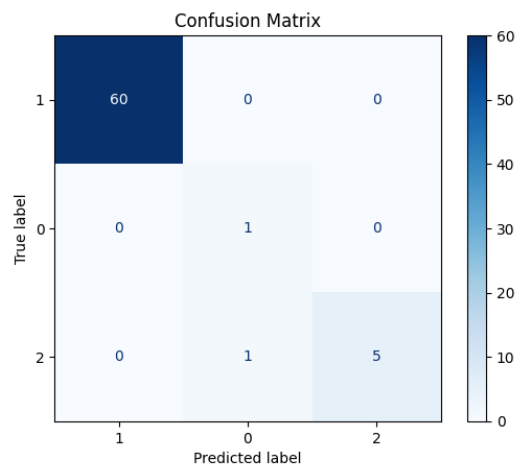


Gambar 3. Pohon Keputusan

Gambar 3 diatas mengenai keputusan terkait penelitian dengan prediksi kinerja akademik mahasiswa di Universitas Muhammadiyah Sidoarjo. Diagram pohon keputusan tersebut diperoleh dari hasil proses cleaning dan data transformasi[13].

- NA adalah fitur yang memiliki pengaruh besar terhadap keputusan kelulusan.
- Jika $NA \leq -0.345$, maka ada kemungkinan lebih besar bahwa mahasiswa **tidak lulus** (kelas 2).
- Jika $NA > -0.345$, maka hampir semua mahasiswa **lulus** (kelas 1).
- **Semakin kecil nilai NA**, semakin besar kemungkinan mahasiswa tidak lulus.

Perhitungan Tingkat Akurasi



Gambar 4. Confusion Matrix

Perhitungan Accuracy

$$\text{Accuracy} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Data}} = \frac{60+1+5}{60+0+0+0+1+0+0+1+5} = \frac{66}{67} = 0.9851$$

Perhitungan Precision

$$\text{Precision} = \frac{\text{True Positive (TP)}}{\text{True Positiv (TP)+False Positive (FP)}}$$

- Precision Kelas 1

$$P_1 = \frac{60}{60+0+0} = \frac{60}{60} = 1.00$$

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.

Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

- Precision Kelas 0

$$P_0 = \frac{1}{1+0+1} = \frac{1}{2} = 0.50$$

- Precision Kelas 2

$$P_2 = \frac{5}{5+0+0} = \frac{5}{5} = 1.00$$

Perhitungan Recall

$$\text{Recall} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Negative (FN)}}$$

- Recall Kelas 1

$$R_1 = \frac{60}{60+0+0} = \frac{60}{60} = 1.00$$

- Recall Kelas 0

$$R_0 = \frac{1}{1+0} = \frac{1}{1} = 1.00$$

- Recall Kelas 2

$$R_2 = \frac{5}{5+1} = \frac{5}{6} = 0.8333$$

Perhitungan F1-Score

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- F1 Score Kelas 1

$$F1_1 = 2 \times \frac{1.00 \times 1.00}{1.00 + 1.00} = 1.00$$

- F1 Score Kelas 0

$$F1_0 = 2 \times \frac{0.50 \times 1.00}{0.50 + 1.00} = 2 \times \frac{0.50}{1.50} = \frac{1.00}{1.50} = 0.67$$

- F1 Score Kelas 2

$$F1_2 = 2 \times \frac{1.00 \times 0.8333}{1.00 + 0.8333} = 2 \times \frac{0.8333}{1.8333} = \frac{1.666}{1.8333} = 0.909$$

Dari hasil yang Anda berikan, tampaknya Anda mengukur akurasi model Decision Tree dengan berbagai pembagian data training dan testing. Berikut adalah analisisnya:

Analisis Hasil:

- 70% training, 30% testing → 98.00% akurasi
Akurasi menurun sedikit, kemungkinan karena lebih banyak data digunakan untuk pengujian.
- 60% training, 40% testing → 98.51% akurasi
Akurasi sedikit meningkat dibandingkan 30% testing, tetapi masih lebih rendah dari 80%-20%.

Tabel 3.4 Hasil Model Decision Tree Akurasi Terbaik

	Precision	Recall	F1-Score	Score
0	1.00	1.00	1.00	60
1	0.50	1.00	0.67	1
2	1.00	0.83	0.91	6
Accuracy	0.83	0.94	0.99	67
Macro Avg	0.99	0.99	0.86	67
Weighted Avg			0.99	67
Akurasi DT : 98.51%				

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.

Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

Proses Klasifikasi Support Vector Machine

Pada algoritma **Support Vector Machine (SVM)**, akurasi model dipengaruhi oleh proporsi data training dan testing yang digunakan. Dengan **40% training dan 60% testing**, model mencapai akurasi **97.01%**, menunjukkan bahwa SVM mampu mempelajari pola dengan baik dari sebagian besar data. Saat proporsi data testing ditingkatkan menjadi **30%** dan training dikurangi menjadi **70%**, akurasi sedikit menurun menjadi **96.00%**, kemungkinan karena lebih banyak data yang digunakan untuk evaluasi sehingga model memiliki lebih sedikit data untuk belajar. Secara keseluruhan, hasil ini menunjukkan bahwa **SVM cukup stabil dalam menangani berbagai proporsi data training dan testing**, dengan perbedaan akurasi yang relatif kecil[14].

Tabel 3.5 Hasil Model Support Vector Machine Akurasi Terbaik

	Precision	Recall	F1-Score	Score
0	0.97	1.00	0.98	60
1	0.00	0.00	0.00	1
2	1.00	0.83	0.91	6
Accuracy	0.66	0.61	0.97	67
Macro Avg	0.96	0.97	0.63	67
Weighted Avg			0.96	67
Akurasi SVM : 98.00%				

Perbandingan Akurasi terbaik

Tabel yang ditampilkan berisi perbandingan tiga algoritma klasifikasi berdasarkan waktu proses, ketelitian (akurasi), dan pembagian data terbaik antara data pelatihan dan pengujian. Berikut penjelasan masing-masing aspek:

1. Algoritma: Menyebutkan tiga metode machine learning yang dibandingkan, yaitu:
 - Naïve Bayes
 - Decision Tree (terdapat kesalahan penulisan: "Decission Tree" seharusnya "Decision Tree")
 - Support Vector Machine (SVM)
2. Waktu Proses: Mengukur berapa lama setiap algoritma membutuhkan waktu untuk melakukan komputasi.
 - Naïve Bayes: 18.47 ms (terlama)
 - Decision Tree: 9.58 ms
 - SVM: 0.05 ms (tercepat)
3. Ketelitian (Akurasi): Persentase keakuratan algoritma dalam membuat prediksi.
 - Naïve Bayes: 95.99%
 - Decision Tree: 98.51% (tertinggi)
 - SVM: 97.01%
4. Pembagian Data Terbaik: Rasio pembagian data antara pelatihan dan pengujian, yang dalam semua kasus adalah 40:60.

Tabel 3.6 Perbandingan Tiga Algoritma

Algoritma	Waktu Proses	Ketelitian	Pembagian Data Terbaik
Naive Bayes	18.47 ms	95.99%	40 : 60
Decission Tree	9.58 ms	98.51%	40 : 60
Support Vector Machine	0.05 ms	97.01%	40 : 60

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dalam memprediksi kinerja akademik mahasiswa dengan menggunakan machine learning, beberapa temuan utama dapat diidentifikasi. Sebagai permulaan, algoritma yang digunakan, yaitu Naïve Bayes, Decision Trees, dan Support Vector Machines (SVM), menunjukkan kinerja yang baik dalam memprediksi kinerja akademik mahasiswa dengan tingkat akurasi yang berbeda-beda. Decision Trees memiliki tingkat akurasi tertinggi, yaitu 98,51%, diikuti oleh SVM (97,01%) dan Naïve Bayes (95,99%)[15]. Waktu proses yang dibutuhkan oleh masing-masing algoritma berbeda secara signifikan, dengan SVM menjadi yang tercepat yaitu 0.05 ms, Decision Tree 9.58 ms, dan Naïve Bayes menjadi yang paling lambat yaitu 18.47 ms. Ketiga, rasio data yang ideal antara pelatihan dan pengujian adalah 40:60 untuk akurasi algoritmik. Secara keseluruhan, penelitian ini menunjukkan bahwa machine learning dapat menjadi alat yang efektif untuk mengevaluasi kebiasaan kerja mahasiswa secara akademis, dengan Decision Trees sebagai pilihan yang paling dapat diandalkan, meskipun SVM menawarkan keunggulan dalam hal kecepatan pemrosesan. Diharapkan hasil penelitian ini dapat memberikan kontribusi pada pengembangan sistem peringatan dini dan peningkatan kualitas pendidikan di perguruan tinggi.

UCAPAN TERIMA KASIH

Puji syukur kami panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya sehingga penelitian dan penyusunan laporan ini dapat diselesaikan dengan baik. Penulis mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan bantuan selama proses penelitian ini berlangsung.

Ucapan terima kasih yang tulus disampaikan kepada Dosen Pembimbing, yang telah memberikan bimbingan, arahan, serta masukan yang berharga dalam penyusunan penelitian ini. Terima kasih juga kepada seluruh dosen Program Studi Informatika Universitas Muhammadiyah Sidoarjo yang telah memberikan ilmu dan wawasan yang sangat bermanfaat selama masa perkuliahan.

Penulis juga ingin menyampaikan apresiasi kepada keluarga dan teman-teman atas doa, dukungan, dan motivasi yang diberikan selama penelitian ini dilakukan. Tanpa dukungan mereka, penelitian ini tidak akan dapat terselesaikan dengan baik. Semoga penelitian ini dapat memberikan manfaat bagi pengembangan teknologi informasi. Penulis menyadari bahwa masih terdapat banyak kekurangan dalam penelitian ini, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk penyempurnaan di masa mendatang.

REFERENSI

- [1] R.- Annisa and A.- Sasongko, "Prediksi Nilai Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes," *JST (Jurnal Sains dan Teknol.*, vol. 9, no. 1, pp. 1–10, 2020, doi: 10.23887/jstundiksha.v9i1.19488.
- [2] A. Putri *et al.*, "Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 1, pp. 20–26, 2023, doi: 10.57152/malcom.v3i1.610.
- [3] U. Indahyanti, N. L. Azizah, and H. Setiawan, "Pendekatan Ensemble Learning Untuk Meningkatkan Akurasi Prediksi Kinerja Akademik Mahasiswa," *J. Sains dan Inform.*, vol. 8, no. 2, pp. 160–169, 2022, doi: 10.34128/jsi.v8i2.459.
- [4] M. Windarti and A. Suradi, "Perbandingan Kinerja 6 Algoritme Klasifikasi Data Mining untuk Prediksi Masa Studi Mahasiswa," *Telematika*, vol. 12, no. 1, p. 14, 2019, doi: 10.35671/telematika.v12i1.778.
- [5] M. P. A. Ariawan, I. B. A. Peling, and G. B. Subiksa, "Prediksi Nilai Akhir Matakuliah Mahasiswa Menggunakan Metode K-Means Clustering (Studi Kasus : Matakuliah Pemrograman Dasar)," *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 2, pp. 122–131, 2023, doi: 10.25077/teknosi.v9i2.2023.122-131.
- [6] A. Rohman and A. Rufiyanto, "Implementasi Data Mining Dengan Algoritma Decision Tree C4 . 5 Untuk Prediksi Kelulusan Mahasiswa Di Universitas Pandaran," *Proceeding SINTAK 2019*, pp. 134–139, 2019.
- [7] E. Haryatmi and S. Pramita Hervianti, "Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 386–392, 2021, doi: 10.29207/resti.v5i2.3007.
- [8] Satrio Junaidi, R. Valicia Anggela, and D. Kariman, "Klasifikasi Metode Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa dengan Algoritma Naïve Bayes, Random Forest, Support Vector Machine (SVM) dan Artificial Neural Network (ANN)," *J. Appl. Comput. Sci. Technol.*, vol. 5, no. 1, pp. 109–119, 2024, doi: 10.52158/jacost.v5i1.489.
- [9] T. Algorithm, "Penerapan Machine Learning Untuk Memprediksi Kelulusan Mahaiswa Menggunakan Algoritma Decision Tree Application of Machine Learning to Predict Student Graduation Using the Decision," pp. 1–6, 2024.
- [10] Z. Saputra, D. Sartika, and M. Haviz Irfani, "Prediksi Calon Mahasiswa Penerima KIP Pada Universitas Indo Global Mandiri menggunakan Algoritma Decision Tree," *J. Rekayasa Tek. Inform. dan Inf.*, vol. 43, no. 3, pp. 231–240, 2024, [Online]. Available: <https://docs.python.org/3.13/tutorial/index.html>.
- [11] N. Nasution, K. Djahara, and A. Zamsuri, "Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes (Studi Kasus: Fasilkom Unilak)," *J. Teknol. Inf. dan Komun. Digit. Zo.*, vol. 6, no. 2, pp. 1–11, 2015.
- [12] A. S. B. Asmoro, W. S. G. Irianto, and U. Pujiyanto, "Perbandingan Kinerja Hasil Seleksi Fitur pada Prediksi Kinerja Akademik Siswa Berbasis Pohon Keputusan," *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 2, p. 84, 2018, doi: 10.26418/jp.v4i2.29294.
- [13] R. Rachmatika and A. Bisri, "Perbandingan Model Klasifikasi untuk Evaluasi Kinerja Akademik Mahasiswa," *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 3, p. 417, 2020, doi: 10.26418/jp.v6i3.43097.
- [14] A. Nurhidayat, A. Asmunin, and D. F. Suyatno, "Prediksi Kinerja Akademik Mahasiswa Menggunakan Machine Learning dengan Sequential Minimal Optimization untuk Pengelola Program Studi," *J. Inf. Eng. Educ. Technol.*, vol. 5, no. 2, pp. 84–91, 2021, doi: 10.26740/jieet.v5n2.p84-91.

- [15] M. S. Mustafa, M. R. Ramadhan, and A. P. Thenata, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Creat. Inf. Technol. J.*, vol. 4, no. 2, p. 151, 2018, doi: 10.24076/citec.2017v4i2.106.

Conflict of Interest Statement:

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.