

Analisis Perbandingan Metode *Decision Tree* Dan *Naive Bayes* Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa

Ni Kadek Juniari¹⁾, Ni Luh Gede Pivin Suwirmayanti²⁾, Gede Herdian Setiawan³⁾

Sistem Informasi¹⁾, Sistem Komputer²⁾³⁾

Institut Teknologi dan Bisnis STIKOM Bali

Denpasar, Indonesia

e-mail: 210030575@stikom-bali.ac.id¹⁾, pivin@stikom-bali.ac.id²⁾, herdian@stikom-bali.ac.id³⁾

Abstrak

Kelulusan tepat waktu mahasiswa merupakan indikator penting dalam menilai kualitas pendidikan tinggi. Penelitian ini ditujukan guna menganalisis dan melakukan perbandingan metode *Decision Tree* dan *Naive Bayes* untuk memprediksi kelulusan tepat waktu mahasiswa Program Studi Sistem Komputer di ITB STIKOM Bali. Data yang dipergunakan yakni data historis mahasiswa angkatan 2018-2020, yang telah melalui proses prapemrosesan untuk meningkatkan kualitas dataset. Metode *Decision Tree* dikenal karena transparansi dalam pengambilan keputusan, sedangkan *Naive Bayes* efektif dalam memanfaatkan distribusi probabilistik. Penelitian ini menunjukkan hasil bahwa metode *Decision Tree* memiliki nilai AUC mencapai 0,919, CA 0,897, F1-Score 0,896, Precision 0,896, Recall 0,897, dan MCC 0,809. Untuk Metode *Naive Bayes* menunjukkan hasil nilai AUC mencapai 0,721, CA 0,554, F1-Score 0,576, Precision 0,617, Recall 0,554, dan MCC 0,264. Evaluasi dalam penelitian ini dilakukan menggunakan *Confusion Matrix*, yang berperan dalam menentukan kualitas metode klasifikasi yang digunakan. Hasil penelitian menghasilkan nilai perbandingan algoritma *Decision Tree* dan *Naive Bayes* dalam meningkatkan efisiensi kelulusan mahasiswa, sehingga dapat membantu institusi pendidikan dalam memilih pendekatan analisis prediktif yang paling sesuai di masa depan

Kata kunci: Kelulusan Tepat Waktu, *Decision Tree*, *Naive Bayes*, Data mining .

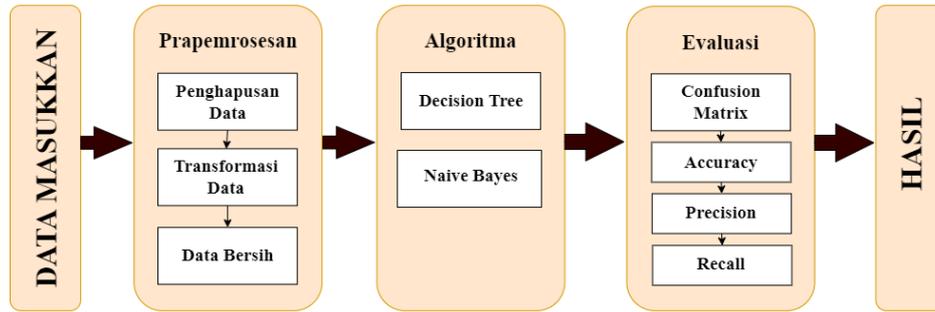
1. Pendahuluan

Kelulusan tepat waktu merupakan salah satu dari sekian indikator fundamental pada penilaian kualitas pendidikan tinggi, sebagaimana diatur dalam PERMENRISTEKDIKTI Nomor 53 Tahun 2023 [1]. Ketepatan waktu kelulusan memiliki dampak signifikan terhadap akreditasi institusi dan efisiensi sistem pendidikan tinggi. Namun, di ITB STIKOM Bali, banyak mahasiswa tidak lulus tepat waktu, menciptakan tantangan dalam pengelolaan data mahasiswa demi memberikan informasi yang bermanfaat untuk pengambilan keputusan. Pada era data yang semakin kompleks, teknik *data mining* menjadi solusi yang relevan untuk mengolah data besar guna menemukan pola dan prediksi tertentu [2].

Metode *Decision Tree* dan *Naive Bayes* telah banyak digunakan dalam penelitian terkait prediksi kelulusan mahasiswa. *Decision Tree* unggul dalam transparansi pengambilan keputusan, sedangkan *Naive Bayes* lebih efektif dalam memanfaatkan distribusi probabilistik. Beberapa penelitian sebelumnya menunjukkan keberhasilan metode ini, seperti penelitian [3] yang menggunakan *Decision Tree* C4.5 dengan tingkat akurasi 90% dan penelitian [4] yang mencapai akurasi 99,67% dengan metode yang sama. Sementara itu, penelitian oleh [5], [6] menunjukkan bahwa *Naive Bayes* mampu memprediksi kelulusan dengan akurasi masing-masing 85% dan 90%, dengan keunggulan pada *precision* dan *recall*. Penelitian yang dilakukan [7] juga melaporkan akurasi di atas 90% menggunakan *Naive Bayes* dengan alat *data mining* seperti Weka dan RapidMiner.

Berdasarkan potensi besar dari kedua metode ini dalam analisis prediktif, penelitian ini ditujukan guna melaksanakan perbandingan kinerja *Decision Tree* dan *Naive Bayes* untuk memperkirakan kelulusan tepat waktu mahasiswa Program Studi Sistem Komputer ITB STIKOM Bali menggunakan data historis dari angkatan 2018–2020. Hasil penelitian ini diharap mampu berkontribusi menjadi referensi strategis bagi institusi dalam meningkatkan efisiensi waktu kelulusan mahasiswa dan pemilihan algoritma yang paling sesuai. Dalam proses memprediksi kelulusan mahasiswa, penelitian ini akan menggunakan aplikasi *Orange*. Selain itu, peneliti akan menguji hasil kinerja *Decision Tree* dan *Naive Bayes* dengan *Confusion Matrix* dilihat juga nilai *accuracy*, *precision* dan *recall*.

2. Metode Penelitian



Gambar 1. Alur Penelitian

2.1 Dataset

Dalam penelitian ini *Dataset* yang digunakan adalah data historis mahasiswa Program Studi Sistem Komputer Angkatan 2018-2020. Tabel 1 merupakan deskripsi dari *Dataset* tersebut.

Tabel 1. Deskripsi *Dataset*

<i>Dataset</i>	Jumlah Fitur	Jumlah Data	Tipe Data
Mahasiswa Program Studi Sistem Komputer Tahun Angkatan 2018-2020	12	368	Numerikal, Kategorikal

2.2 Prapemrosesan *Dataset*

Sebelum *Dataset* diolah menggunakan algoritma, diperlukan untuk melakukan prapemrosesan. Penelitian [8] menunjukkan bahwa teknik ini digunakan untuk mendapatkan *dataset* berkualitas tinggi. Prapemrosesan data dilaksanakan melalui proses eliminasi data yang tidak selaras atau melakukan perubahan data ke dalam wujud yang lebih mudah diolah sistem [9]. Prapemrosesan data yakni tahap fundamental pada *data mining* karena sebagian besar data mentah yang terdapat dalam basis data bersifat belum diproses, tidak lengkap, dan *noisy* [10]. Dalam proses *data mining*, basis data perlu menjalani prapemrosesan, yang melibatkan pembersihan data (*data cleaning*), penghapusan data dan transformasi data. Pembersihan data yakni proses membersihkan dan menyiapkan kumpulan data agar dipergunakan atau dianalisis pada model pengajaran mesin. Tujuan utama dari pembersihan data yakni guna membereskan masalah seperti nilai yang tidak ada, *outliers*, atau inkonsistensi format [11]. Adapun mengubah tipe data, memperkaya, menghapus duplikat atau nol dari data, atau mengagregasi data adalah beberapa contoh transformasi data [12].

Tabel 2. *Dataset* Sebelum Prapemrosesan

NIM	Tahun Masuk	IPS 1	IPK 1	IPS 2	IPK 2	IPS 3	IPK 3	IPS 4	IPK 4	IPS 5	IPK 5	IPS 6	IPK 6	IPS 7	IPK 7	IPS 8	IPK 8	IPS 9	IPK 9	IPS 10	IPK 10	IPS 11	IPK 11	IPS 12	IPK 12	IPK	Tahun Lulus		
18000006	2018	3.06	3.06	3.33	3.19	2.89	3.08	3.42	3.89	3	3.19	2.77	3.12	2.67	3.08	3.67	3.1										3.1	2024/2022	
18000007	2018	3.78	3.78	3.61	3.69	3.39	3.57	3.42	3.93	3.45	3.52	0.66	3.21	0	3.2	3.11	3.32	0	3.32	1	3.32	1	3.32	3.26	3.34	3.34	3.34	2024/2026	
18000008	2018	3.58	3.58	3.58	3.56	3.57	3.56	3.42	3.52	3.46	3.58	3.17	3.5	3.29	3.51	0	3.51	3.33	3.55								3.55	2023/2023	
18000009	2018	3.11	3.11	3.5	3.33	3.05	3.24	3.1	3.2	3.65	3.33	3.08	3.28	3.5	3.31	4	3.38	3	3.37								3.37	2022/2023	
18000011	2018	3.58	3.58	3.58	3.57	3.24	3.45	3.25	3.4	3.58	3.46	3.42	3.45	3.53	3.56	4	3.57										3.57	2022/2022	
18000012	2018	6.97	6.97	3.36	3.43	3.16	3.34	3	3.24	3.8	3.36	3.19	3.38	2.89	3.32	2	3.31	2	3.31	4	3.35	3.5	3.36				3.36	2022/2022	
18000014	2018	3.42	3.42	3.42	3.42	2.88	3.26	3.58	3.46	3.43	3.48	3.68	3.4	3.45	3.5	3.75	3.51										3.51	2022/2022	
18000016	2018	3.33	3.33	3.11	3.22	3.05	3.16	3.21	3.89	3.31	3.2	2.79	3.21	3.44	3.27	2.38	3.26	3	3.26								3.26	2022/2023	
18000016	2018	6.97	6.97	3.47	3.49	3.21	3.39	2.79	3.21	3.46	3.27	3.17	3.28	3.33	3.29	3.25	3.29										3.29	2022/2022	
18000017	2018	3.97	3.97	3.79	3.88	3.65	3.77	3.96	3.82	3.86	3.84	3.64	3.9	3.63	3.81	4	3.82	4	3.82								3.82	2022/2023	
20000100	2020																												
20000107	2020	3.22	3.22	3.75	3.49	3.7	3.57	3.54	3.56	4	3.66	3.58	3.65	3.83	3.69												3.69	2023/2024	
20000107	2020	3.39	3.39	3.53	3.46	3.74	3.57	3.91	3.66	4	3.74	3.82	3.81	3.71	3.9	4	3.81										3.81	2023/2024	
20000103	2020	3.89	3.89	3.78	3.83	3.87	3.85	3.79	3.83	3.83	3.82	3.88	3.83	3.71	3.82	4	3.82										3.82	2023/2024	
20000126	2020	3.39	3.39	3.5	3.44	3.47	3.45	3.85	3.56	3.17	3.49	3.42	3.48	3.5	3.51	4	3.53										3.53	2024/2025	
20000127	2020	3.44	3.44	3.17	3.31	3.62	3.4	3.88	3.56	3.25	3.49	3.33	3.46	3.38	3.52	4	3.54										3.54	2024/2025	
20000129	2020	3.72	3.72	3.94	3.83	3.83	3.83	3.89	3.85	3.77	3.83	3.9	3.84	3.69	3.83	3.5	3.82										3.82	2023/2024	
20000130	2020	3.22	3.22	3.44	3.33	3.52	3.41	3.46	3.42	3.63	3.46	3.54	3.47	3.33	3.48	3	3.45										3.45	2024/2025	
20000132	2020	4	4	3.94	3.97	3.9	3.95	3.96	3.95	4	3.96	3.91	3.95	4	3.95	3.5	3.94										3.94	2023/2024	
20000137	2020	3.61	3.61	3.63	3.65	3.45	3.58	3.91	3.68	3.55	3.65	3.91	3.7	3.21	3.84	4	3.68										3.68	2024/2025	
20000142	2020	3.28	3.28	3.58	3.42	3.78	3.56	3.91	3.65	3.94	3.71	3.82	3.75	4	3.75	3.71	3.76										3.76	2024/2025	
20000144	2020	3.61	3.61	3.44	3.53	3.69	3.69	3.83	3.66	3.58	3.64	3.82	3.67	3.8	3.69	3.5	3.69										3.69	2023/2024	
20000147	2020	3.67	3.67	3.61	3.64	4	3.76	3.83	3.78	4	3.83	3.79	3.82	3.8	3.82	4	3.83										3.83	2023/2024	
20000148	2020	3.86	3.86	3.28	3.17	3.63	3.33	3.5	3.38	3.95	3.36	3.82	3.44	3.67	3.48	3	3.47										3.47	2023/2024	
20000149	2020	3.94	3.94	3.94	3.94	3.91	3.89	3.79	3.86	3.75	3.83	3.75	3.81	4	3.82	3	3.8										3.8	2023/2024	
20000157	2020	3.29	3.28	3.36	3.32	3.61	3.43	3.38	3.42	4	3.57	3.54	3.57	3.5	3.56												3.56	2023/2024	
20000159	2020	3.33	3.33	3.09	3.21	3.55	3.33	3.72	3.42	2.82	3.28	3.35	3.41	3.5	3.35	3.75	3.49										3.49	2016/2016	

Pada Tabel 2 merupakan data histori mahasiswa Program Studi Sistem Komputer Angkatan 2018-2020 sebanyak 376 mahasiswa. Data tersebut yakni data mentah yang belum dilaksanakan prapemrosesan.

2.3 Decision Tree

Pohon keputusan atau *Decision Tree* dipakai dalam melaksanakan pengklasifikasian sampel data dengan kelas yang tidak diketahui ke dalam kelas yang tersedia. Pembuatan pohon keputusan terdiri dari proses perubahan bentuk data (tabel) dalam metode pohon, melakukan perubahan metode pohon menjadi aturan, lalu selanjutnya membuat aturan lebih sederhana [13].

$$Gain(S, A) = Entropy(S) - \sum_i^n 1 \frac{|S_i|}{|S|} \times Entropy(S_i) \tag{1}$$

Rumus $Gain(S, A)$ menghitung pengurangan *entropy* setelah *dataset* S_i dibagi berdasarkan atribut A . Nilai gain diperoleh dengan mengurangi *entropy* awal S dengan rata-rata tertimbang dari *entropy* setiap subset S_i , di mana bobotnya adalah proporsi ukuran subset $|S_i|$ terhadap ukuran total *dataset* $|S|$. Rumus *entropy* tersebut yakni:

$$Entropy(S) = \sum_{i=1}^n -P_i \times \log_2 P_i \tag{2}$$

Rumus $Entropy(S)$ menghitung tingkat ketidakpastian atau keacakan dalam *dataset* S . Nilainya diperoleh dengan menjumlahkan hasil kali antara probabilitas P_i dari setiap kelas i dengan logaritma basis 2 dari P_i , dikalikan dengan -1 untuk memastikan nilainya positif.

2.4 Naïve Bayes

Algoritma ini memakai metode probabilitas serta statistik yang diusulkan oleh ilmuwan Inggris bernama *Thomas Bayes*, yakni melaksanakan prediksi kemungkinan masa depan atas dasar pengalaman masa lalu [14]. *Naïve Bayes* yakni salah satu algoritma pembelajaran induktif yang paling efisien serta efektif dalam *machine learning* maupun *data mining* [15]. *Naïve Bayes* menggunakan perhitungan probabilitas dan statistik *Teorema Bayes*.

$$P(H|X) = \frac{P(X|H)}{P(X)} P(H) \tag{3}$$

Rumus $P(H|X)$ yakni *Teorema Bayes*, yang menghitung probabilitas hipotesis H diberikan bukti X . Nilainya diperoleh dengan mengalikan probabilitas kondisi $P(H|X)$ dengan prior $P(H)$ lalu membaginya dengan probabilitas keseluruhan $P(X)$, yang memastikan hasilnya terukur dengan benar terhadap semua kemungkinan hipotesis.

2.5 Evaluasi

Pelaksanaan uji dilaksanakan dengan menghitung akurasi algoritma *Decision Tree* serta *Naïve Bayes* dengan *Confusion Matrix*. *Confusion Matrix* yakni tabel yang mengklasifikasikan banyaknya data uji yang benar serta banyaknya data uji yang salah. Keakuratan hasil diukur dengan nilai *recall*, *precision*, *accuracy* [16]. Sesudah akurasi dari setiap algoritma diperoleh, selanjutnya dilakukan perbandingan hasil akurasi algoritma *Decision Tree* dengan *Naïve Bayes* hingga diperoleh algoritma yang berakurasi paling baik untuk memperkirakan kelulusan mahasiswa.

3. Hasil dan Pembahasan

Data yang dipergunakan dalam proses *training* dan *testing* sudah melalui proses prapemrosesan data hingga data sudah siap melalui pengolahan. Adapun prapemrosesan data yang dilakukan yakni penghapusan serta transformasi data. Pada penghapusan *dataset*, data yang memiliki nilai 0 serta atribut Tahun Masuk dan Tahun Lulus dihapus. Kemudian pada transformasi *dataset*, peneliti memberikan label Status Kelulusan serta mengisi data yang kosong menggunakan nilai rata-rata. Berikut ialah *Dataset* dalam Tabel 3 yang hendak melalui pengolahan sejumlah 368 data mahasiswa.

Tabel 3. *Dataset* Prediksi Kelulusan Mahasiswa

No	IPS1	IPS2	IPS3	IPS4	IPS5	IPS6	IPS7	IPS8	IPS9	IPS10	IPS11	Status Kelulusan
1	3.06	3.33	2.89	3.42	3	2.77	2.67	3.67	3.10125	3.10125	3.10125	Tepat
2	3.11	3.5	3.05	3.1	3.65	3.08	3.5	4	3	3.33222	3.33222	Terlambat
3	3.58	3.56	3.24	3.25	3.58	3.42	3.53	4	3.52	3.52	3.52	Tepat
4	6.97	3.36	3.16	3	3.8	3.19	2.88	2	2	4	3.5	Terlambat
5	3.42	3.42	2.98	3.58	3.43	3.08	3.45	3.75	3.38875	3.38875	3.38875	Tepat
....
366	3.94	3.94	3.81	3.79	3.75	3.75	4	3	3.7475	3.7475	3.7475	Tepat
367	3.28	3.36	3.61	3.38	4	3.54	3.5	3.52428571	3.52429	3.52429	3.52429	Cepat
368	3.33	3.08	3.55	3.72	2.82	3.35	3.5	3.75	3.3875	3.3875	3.3875	Tepat

Berdasarkan Tabel 3 dapat dilihat atribut status kelulusan mejadi label target dengan kategori Cepat bagi masa studi 3,5 tahun, Tepat untuk masa studi sama dengan 4 tahun serta Terlambat untuk masa studi lebih dari 4 tahun. Berikutnya nilai akurasi dari hasil pengujian bisa disaksikan atas dasar hasil akurasi selayaknya ditampilkan dalam Gambar 2 serta Gambar 3.

3.1 Perbandingan Hasil Akurasi *Decision Tree* dan *Naïve Bayes*

Berikut ini merupakan hasil akurasi dari metode *Decision Tree* dan *Naïve Bayes*. Pada Gambar 2 menunjukkan bahwa akurasi metode *Decision Tree* digunakan untuk analisis data, menghasilkan model yang sangat akurat dengan nilai *AUC* mencapai 0,919, *CA* 0,897, *F1-Score* 0,896, *Precision* 0,896, *Recall* 0,897, dan *MCC* 0,809. Model ini juga menunjukkan kinerja yang baik untuk memprediksi data positif dengan tingkat akurasi yang tinggi.

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.919	0.897	0.896	0.896	0.897	0.809

Gambar 2. Hasil Akurasi Metode *Decision Tree*

Model	AUC	CA	F1	Prec	Recall	MCC
Naive Bayes	0.721	0.554	0.576	0.617	0.554	0.264

Gambar 3. Hasil Akurasi Metode *Naïve Bayes*

Adapun Gambar 3, menunjukkan bahwa metode *Naïve Bayes* memiliki kinerja optimal yang rendah, dengan menunjukkan nilai *AUC* mencapai 0,721, *CA* 0,554, *F1-Score* 0,576, *Precision* 0,617, *Recall* 0,554, dan *MCC* 0,264.

3.2 Perbandingan Hasil *Confusion Matrix Decision Tree* dan *Naïve Bayes*

		Predicted			Σ
		Cepat	Tepat	Terlambat	
Actual	Cepat	12	6	2	20
	Tepat	4	181	9	194
	Terlambat	2	15	137	154
Σ		18	202	148	368

Gambar 4. Hasil *Confusion Matrix Decision Tree*

		Predicted			Σ
		Cepat	Tepat	Terlambat	
Actual	Cepat	12	7	1	20
	Tepat	38	109	47	194
	Terlambat	19	52	83	154
Σ		69	168	131	368

Gambar 5. Hasil *Confusion Matrix Naïve Bayes*

Gambar 4 serta Gambar 5 di atas menunjukkan hasil *Confusion Matrix* dari model *Decision Tree* dan *Naive Bayes* dapat dibaca dengan membandingkan kemampuan kedua model dalam memprediksi tiga kelas: "Cepat," "Tepat," dan "Terlambat." Pada model *Decision Tree*, performa terbaik terlihat pada kelas "Tepat," dengan 181 prediksi benar dari total 194 *instance* aktual, sementara *Naive Bayes* hanya mampu memprediksi 109 *instance* dengan benar pada kelas yang sama. Untuk kelas "Cepat," *Decision Tree* memprediksi 12 *instance* dengan benar dari total 20, lebih akurat dibandingkan *Naive Bayes* yang hanya memprediksi 12 dari total 20, namun dengan kesalahan lebih sedikit 6 ke "Tepat" dan 2 ke "Terlambat". Pada kelas "Terlambat," *Decision Tree* memiliki 137 prediksi benar dari total 154, sedangkan *Naive Bayes* hanya memiliki 83 prediksi benar, menunjukkan keunggulan *Decision Tree* dalam mengenali kelas ini. Secara keseluruhan, *Decision Tree* menunjukkan akurasi yang lebih tinggi dengan distribusi kesalahan yang lebih kecil dibandingkan *Naive Bayes*, menjadikannya lebih andal untuk memprediksi data pada ketiga kelas.

4. Kesimpulan

Penelitian ini menerapkan metode *Decision Tree* yang dibandingkan dengan metode *Naive Bayes* dalam memprediksi kelulusan tepat waktu mahasiswa Program Studi Sistem Komputer di ITB STIKOM Bali. Hasil analisis menunjukkan bahwa *Decision Tree* memperoleh nilai *AUC* sebesar 0,919, *CA* 0,897, *F1-Score* 0,896, *Precision* 0,896, *Recall* 0,897, dan *MCC* 0,809. Untuk metode *Naive Bayes* memperoleh nilai *AUC* sebesar 0,721, *CA* 0,554, *F1-Score* 0,576, *Precision* 0,617, *Recall* 0,554, dan *MCC* 0,264. Evaluasi menggunakan *Confusion Matrix* menunjukkan hasil *Decision Tree* mampu mengenali kategori "Tepat" dengan nilai 181 dan "Terlambat" dengan nilai 137. Untuk metode *Naive Bayes* mencapai nilai 109 dan 83 prediksi benar pada kategori yang sama. Berdasarkan hasil analisis ini menunjukkan perbandingan tingkat performa metode *Decision Tree* dan metode *Naive Bayes*. Dimana dari hasil evaluasi terlihat metode *Decision Tree* memiliki nilai lebih tinggi dari metode *Naive Bayes*. Hasil analisis ini dapat mendukung pengambilan keputusan terkait kelulusan mahasiswa di institusi pendidikan tinggi

Daftar Pustaka

- [1] dikti.kemdikbud.go.id, "Permendikbudristek Nomor 53 Tahun 2023, Perguruan Tinggi Fleksibel Kembangkan Standar Kompetensi Lulusan," dikti.kemdikbud.go.id. Accessed: Jul. 22, 2024. [Online]. Available: <https://dikti.kemdikbud.go.id/kabar-dikti/kabar/permendikbudristek-nomor-53-tahun-2023-perguruan-tinggi-fleksibel-kembangkan-standar-kompetensi-lulusan/>
- [2] S. Pujiono, R. Astuti, and F. Muhamad Basysyar, "Implementasi *Data mining* Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 615–620, 2024, doi: 10.36040/jati.v8i1.8360.
- [3] M. N. Yatimah and D. Komalasari, "Implementasi *Data mining* untuk Prediksi Kelulusan Tepat Waktu Mahasiswa STIMIK ESQ Menggunakan *Decision Tree* C4. 5," in *SNIA (Seminar Nasional Informatika dan Aplikasinya)*, 2021, pp. D21-25.
- [4] F. Rahmadayanti and I. Anggraini, "Penerapan Metode *Decision Tree* Dalam Menentukan Kelulusan Mahasiswa," *Build. Informatics, Technol. Sci.*, vol. 3, no. 3, pp. 441–445, 2021, doi: 10.47065/bits.v3i3.1154.
- [5] N. M. A. Mahar, Vihl Atina, and Nugroho Arif Sudibyo, "Pemodelan Prediksi Kelulusan Mahasiswa Dengan Metode *Naive Bayes* Di Uniba," *J. Manaj. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 148–158, 2023, doi: 10.36595/misi.v6i2.875.
- [6] Y. Apridiansyah, N. D. M. Veronika, and E. D. Putra, "Prediksi Kelulusan Mahasiswa Fakultas Teknik Informatika Universitas Muhammadiyah Bengkulu Menggunakan Metode *Naive Bayes*," *JSAI (Journal Sci. Appl. Informatics)*, vol. 4, no. 2, pp. 236–247, 2021, doi: 10.36085/jsai.v4i2.1701.
- [7] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, "Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode *Data mining Naive Bayes* : Systematic Review," *Fakt. Exacta*, vol. 13, no. 1, p. 35, 2020, doi: 10.30998/faktorexacta.v13i1.5548.
- [8] M. Zainuddin, "Perbandingan 4 Algoritma Berbasis Particle Swarm Optimization (pso) Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa," *J. Ilm. Teknol. Inf. Asia*, vol. 13, no. 1, pp. 1–12, 2019.
- [9] D. A. Nasution, H. H. Khotimah, and N. Chamidah, "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, p. 78, 2019, doi: 10.24114/cess.v4i1.11458.
- [10] D. T. Larose and C. D. Larose, *Discovering knowledge in data: an introduction to data mining*. Wiley., 2020.
- [11] M. F. Febrianto *et al.*, "Prediksi Situasi Lalu Lintas Menggunakan Machine Learning Dengan Algoritma K-Nearest Neighbors Classifier," *Inform. J. Ilmu Komput.*, vol. 20, no. 1, pp. 28–34, 2024, doi: 10.52958/iftk.v20i1.7042.
- [12] H. T. B. Suandi As and L. Zahrotun, "PENERAPAN *DATA MINING* DALAM MENGELOMPOKKAN DATA RIWAYAT AKADEMIK SEBELUM KULIAH DAN DATA KELULUSAN MAHASISWA MENGGUNAKAN METODE AGGLOMERATIVE HIERARCHICAL CLUSTERING (Implementation Of *Data mining* In Grouping Academic History Data Before Students And Stud)," *J. Teknol. Informasi, Komput. dan Apl.*, vol. 3, no. 1, pp. 62–71, 2021, [Online]. Available: <http://jtika.if.unram.ac.id/index.php/JTIKA/>
- [13] S. Mashlahah, *Prediksi Kelulusan Mahasiswa Menggunakan Metode Decision Tree Dengan Penerapan Algoritma C4.5*. 2013.

- [14] N. L. G. P. Suwirmayanti, "Penerapan Metode *Naive Bayes* Untuk Menganalisa Kerusakan Automatic Transmission Pada Mobil," *Jurnal Sist. Dan Inform.*, vol. 10, no. 1, pp. 41–51, 2015.
- [15] S. Syarli and A. A. Muin, "Metode *Naive Bayes* Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)," *J. Ilmu Komput.*, vol. 2, no. 1, pp. 22–26, 2018, [Online]. Available: <https://fikom-unasman.ac.id/ejournal/index.php/jikom/article/view/26/17>
- [16] D. Normawati and S. A. Prayogi, "Implementasi *Naive Bayes* Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform.*, vol. 5, no. 2, pp. 697–711, 2021.
-