PREDIKSI MASA STUDI MAHASISWA MENGGUNAKAN DECISION TREE YANG DIOPTIMALISASI DENGAN HYPERPARAMETER TUNING

Meyer Mega Eklesia Silaban[1)], Muhammad Ichsan Basir[2)], dan Sahid Triambudhi[3)]

[1]Teknik Informatika, Politeknik Negeri Indramayu
[2]Kesehatan, Politeknik Negeri Indramayu
[3]Teknik Informatika, Politeknik Negeri Indramayu
E-mail: meyermega@polindra.ac.id

Peningkatan jumlah mahasiswa setiap tahun tidak sebanding dengan angka kelulusan tepat waktu, yang berakibat pada keterlambatan masa studi dan penurunan efisiensi pendidikan. Kondisi ini menyoroti urgensi pemanfaatan data historis mahasiswa untuk tujuan prediktif. Penelitian ini mengusung metode data mining sebagai pendekatan utama untuk mengklasifikasi dan memprediksi status kelulusan mahasiswa dengan data dari salah satu universitas di sumatera. Secara spesifik, penelitian ini menerapkan algoritma *Decision Tree* dengan bantuan pemrograman *Python* untuk menganalisis data akademik mahasiswa. Namun, pada tahap awal, akurasi model yang dihasilkan hanya mencapai 62% dengan menggunakan metode algoritma *Decision Tree* pada pemrograman *Python*, nilai ini masih belum optimal untuk pengambilan keputusan yang strategis. Untuk meningkatkan performa model secara signifikan, penelitian ini menerapkan teknik *Hyperparameter Tuning*. Hasilnya akurasi model naik menjadi 90%, menunjukkan kemampuan prediktif yang jauh lebih baik. Analisis lebih lanjut mengungkap bahwa atribut yang paling berpengaruh terhadap prediksi kelulusan adalah shift kuliah, status pekerjaan, dan jenis kelamin. Dengan akurasi yang tinggi, model prediksi ini diharapkan dapat menjadi alat bantu yang andal bagi pembimbing akademik dalam mengidentifikasi mahasiswa yang berpotensi mengalami keterlambatan, memungkinkan implementasi langkah intervensi yang tepat waktu dan terarah. Penelitian ini diharapkan tidak hanya memberikan peringatan dini, tetapi juga memberikan wawasan tentang faktor-faktor kunci yang mempengaruhi keberhasilan studi, sehingga strategi pendukung dapat disesuaikan untuk setiap mahasiswa.

Keyword: Data Historical, Classification, Prediction, Decision Tree, Data Mining

1. PENDAHULUAN

Perkembangan teknologi informasi telah membawa perubahan signifikan dalam berbagai aspek kehidupan, tidak terkecuali di ranah pendidikan tinggi. Perkembangan teknologi juga yang menghadirkan pemahaman tentang proses pengolahan suatu informasi untuk menemukan pola, dan wawasan yang tersembunyi dari himpunan data yang besar yang disebut data mining.

Data mining berkembang sejak awal tahun 1980, dan merupakan perkembangan dari ilmu statistik dan kecerdasan buatan dengan fokus pada analisis basis data besar (KDD). Popularitasnya meningkat ketika dimanfaatkan perusahaan untuk pemasaran, lalu terus berkembang seiring big data. Kini, *data mining* menjadi bagian dari *data science* yang didukung *machine learning* dan *deep learning* (Zhou et al., 2020) untuk analisis pola, prediksi, dan pengambilan keputusan (KDD, 2024).

Berdasarkan data dari IPB University, hanya sekitar 30,94% mahasiswa di Indonesia yang menyelesaikan studi tepat waktu, sedangkan mayoritas memerlukan waktu lebih lama (BeritaSatu, 2025). Hal ini mengindikasikan bahwa sebagian besar mahasiswa menghadapi tantangan dalam menyelesaikan studi sesuai target, yang berdampak pada perencanaan akademik, biaya pendidikan, dan alokasi sumber daya institusi. Dalam hal ini, pemanfaatan model prediksi seperti yang diterapkan dalam penelitian sangat krusial untuk dapat melakukan prediksi dan mengetahui kemungkinan risiko sejak dini, guna meminimalisir kemungkinan terjadinya keterlambatan penyelesaian studi.

Pada konteks pendidikan tinggi, masa studi mahasiswa merupakan indikator penting untuk mengevaluasi efektivitas proses pembelajaran (Simões, Oliveira, & Nunes, 2022). Perguruan tinggi dituntut untuk menghasilkan lulusan yang tidak hanya memiliki kualitas akademik yang tinggi, tetapi juga menyelesaikan studi mereka secara tepat waktu untuk memenuhi tuntutan persaingan di dunia kerja yang semakin ketat (Mailana, Febriyani, & Hamdani, 2021; Desfiandi & Soewito, 2023). Data historis mahasiswa, yang mencakup riwayat akademik, demografi, dan informasi lainnya, memiliki potensi besar jika diolah dengan metode yang tepat. Salah satu pendekatan yang terbukti efektif adalah data mining, yang memungkinkan institusi pendidikan untuk menganalisis data skala besar dan mengekstrak pengetahuan baru yang relevan (Chrisnanto & Abdullah, 2021).

Teknik data mining, khususnya metode klasifikasi, sering digunakan untuk memprediksi hasil akademik mahasiswa, termasuk masa studi dan kelulusan. Berikut merupakan tabel dari penelitian sebelumnya.

Tabel 1. Penelitian Sebelumnya

Peneliti & Tahun	Metode/Model	Hasil Akurasi / Temuan Utama
Pawitra, Astuti & Sari (2024)	*Decision Tree*	Akurasi mencapai **91,2%** dalam memprediksi keterlambatan studi mahasiswa.
Bisri, Hidayat & Fadillah (2025)	*Naïve Bayes*	Akurasi sekitar **88,7%** dalam memprediksi kelulusan mahasiswa.
Mehta (2023)	*Machine Learning* (umum)	Model terbaik mencapai akurasi sekitar **90%** dalam prediksi akademik.
Putri, Rusdah, Suryadi & Anubhakti (2023)	*Decision Tree C4.5*	Akurasi prediksi kelulusan mahasiswa di ISB Atma Luhur mencapai **93,94%**.
Mailana, Febriyani & Hamdani (2021)	*Naïve Bayes*, SVM, *Decision Tree*	Algoritma klasifikasi menunjukkan efektivitas; akurasi terbaik mencapai **92%**.
Panggabean, Sirait & Naibaho (2022)	*Naïve Bayes*, SVM, *Decision Tree*	Akurasi berkisar antara **89–91%** pada data mahasiswa.

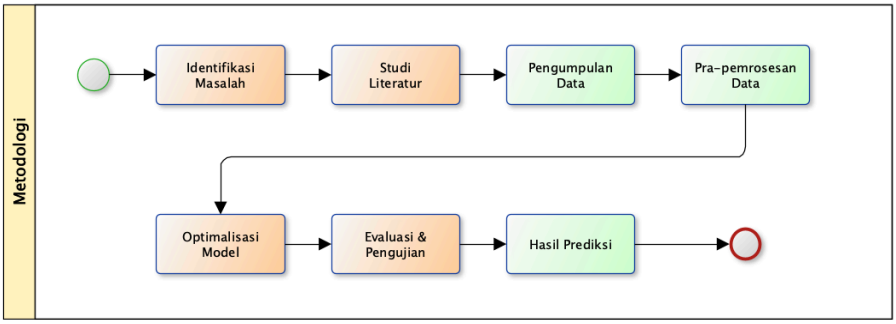
Berdasarkan penjelasan penelitian terdahulu yang dijelaskan pada tabel 1, dapat dilihat bahwa sebagian besar penelitian hanya menggunakan algoritma umum tanpa melakukan optimasi parameter. Oleh karena itu, dalam penelitian diterapkan teknik *Hyperparameter Tuning* untuk mencari kombinasi parameter terbaik pada algoritma yang digunakan, sehingga diharapkan dapat meningkatkan akurasi prediksi kelulusan mahasiswa secara lebih signifikan.

Fokus penelitian ini adalah penerapan algoritma *Decision Tree* untuk memprediksi masa studi mahasiswa pada suatu Universitas Swasta yang ada di Sumatera. Algoritma ini dipilih karena kemampuannya dalam menghasilkan model yang mudah diinterpretasi, mereplikasi proses pengambilan keputusan manusia. Namun, pada tahap awal penerapan, model hanya mencapai akurasi sebesar 62%, hasil yang masih belum optimal jika dibandingkan dengan studi lain seperti model C4.5 dengan akurasi di atas 80 % (Putri et al., 2023). Keterbatasan akurasi awal ini menjadi motivasi utama untuk melanjutkan penelitian pada tahap optimalisasi.

Setelah menggunakan teknik *hyperparameter tuning* pada algoritma *decision tree*, hasilnya nilai akurasi model meningkat dari 62% menjadi 90%. Hal ini membuktikan bahwa menerapkan teknik *hyperparameter* merupakan langkah tepat dalam membangun model prediksi yang optimal. Dalam konteks pendidikan terbaru, *hyperparameter tuning* terbukti efektif (Adji & Syarif, 2022). Sebagai penegasan, penelitian ini berkontribusi pada bidang analitik pendidikan dengan menunjukkan bahwa *hyperparameter tuning* mampu meningkatkan akurasi prediksi masa studi mahasiswa. penelitian ini juga bermanfaat bagi perguruan tinggi dalam meningkatkan layanan akademik serta menjadi dasar bagi pengembangan sistem pendukung keputusan yang lebih adaptif dan berbasis data.

2. METODE PENELITIAN

Pada penelitian ini dilakukan alur penelitian disusun dalam bentuk kerangka kerja yang berfungsi sebagai panduan dalam menjalankan setiap tahapan proses penelitian. Berikut ini kerangka kerja penelitian yang akan dilakukan:



Gambar 1. Alur Penelitian

Diagram diatas menjelaskan tentang alur penelitian ini yang dimulai dari identifikasi masalah yang akan diselesaikan melalui penelitian. Penelitian ini diawali dengan identifikasi masalah, studi literatur, serta pengumpulan data mahasiswa yang menjadi objek penelitian. Data tersebut kemudian melalui tahap *preprocessing* untuk dibersihkan dan dipersiapkan sebelum digunakan dalam pemodelan.

Selanjutnya, data dimasukkan ke dalam model *decision tree* untuk menghasilkan model awal yang kemudian dioptimalkan melalui *hyperparameter tuning*. Proses ini bertujuan meningkatkan kinerja model, yang kemudian dievaluasi menggunakan metrik tertentu hingga diperoleh hasil akhir berupa prediksi kelulusan mahasiswa.

3. HASIL DAN PEMBAHASAN

3.1. Identifikasi Masalah

Tahap awal penelitian ini adalah mengidentifikasi masalah utama yang menjadi latar belakang penelitian, yaitu tingginya angka keterlambatan masa studi mahasiswa di perguruan tinggi. Fenomena ini berdampak pada efisiensi akademik dan produktivitas lulusan, sehingga diperlukan suatu metode yang efektif untuk memprediksi dan mencegahnya. Permasalahan ini mendorong pengembangan sistem prediktif yang dapat memberikan peringatan dini kepada pihak universitas dan mahasiswa.

3.2. Studi Literatur

Setelah masalah teridentifikasi, dilakukan studi literatur untuk mengkaji teori, metode, dan penelitian terdahulu yang relevan. Kajian ini mencakup algoritma klasifikasi dalam data mining yang telah terbukti efektif dalam memprediksi masa studi mahasiswa. Algoritma-algoritma seperti *Naïve Bayes, Support Vector Machine* (SVM), dan *decision tree* menjadi fokus utama dalam tinjauan pustaka ini, yang akan menjadi dasar teoritis untuk pemilihan metode yang digunakan dalam penelitian.

3.3. Pengumpulan Data

Data yang digunakan merupakan data historis mahasiswa yang telah lulus. Data ini berfungsi sebagai data pelatihan (*training*) dan data pengujian (*testing*), dengan total sebanyak 360 data. Atribut yang dikumpulkan mencakup riwayat akademik dan non-akademik, seperti angkatan, jenis kelamin, status pekerjaan, Indeks Prestasi Kumulatif (IPK), dan keterangan kelulusan. Data ini menjadi pondasi utama untuk membangun model prediksi.

#	NIM	Nama	Prodi	Angkatan	JK	S.Kuliah	Tgl. Lulus	Asal Sekolah	Shift	Status Kerja	IPK	Ket. Lulus
1	8020160003	Diana Novita Sa	TI	2016	P	Lulus	2020-03-03	SMK	Malam	Sudah Bekerja	3.48	Tepat
2	8020160008	Indrawan Jamin	TI	2016	L	Lulus	2020-02-28	SMK	Malam	Belum Bekerja	3.81	Cepat
3	8020160011	Desy Ratna Sari	TI	2016	P	Lulus	2020-03-02	SMA	Pagi	Belum Bekerja	3.62	Cepat
4	8020160012	Yoga Tanoko	TI	2016	L	Lulus	2020-03-03	SMA	Malam	Belum Bekerja	3.94	Cepat
5	8020160014	Kevin Rahmat C	TI	2016	L	Lulus	2020-08-29	SMK	Pagi	Belum Bekerja	3.73	Tepat
6	8020160015	Akmal Rustandi	TI	2016	L	Lulus	2020-09-02	SMK	Pagi	Belum Bekerja	3.22	Tepat
7	8020160016	Angga Trisnandi	TI	2016	L	Lulus	2020-08-31	SMK	Pagi	Belum Bekerja	2.97	Tepat
8	8020160017	Muhtadi Abdillah	TI	2016	L	Lulus	2020-02-29	SMK	Pagi	Belum Bekerja	3.67	Cepat
9	8020160018	Shofan Rizki Sel	TI	2016	L	Lulus	2020-08-29	SMK	Pagi	Belum Bekerja	3.45	Tepat
10	8020160020	Feliks Lourensiu	TI	2016	L	Lulus	2020-02-27	SMA	Malam	Sudah Bekerja	3.92	Cepat
11	8020160021	Kevin Wenardy	TI	2016	L	Lulus	2020-09-03	SMA	Malam	Belum Bekerja	3.82	Tepat
12	8020160022	Andrean Oscar	TI	2016	L	Lulus	2020-02-27	SMK	Pagi	Belum Bekerja	3.62	Cepat
13	8020160023	Wahyudin	TI	2016	L	Lulus	2020-02-24	SMK	Pagi	Belum Bekerja	3.82	Cepat
14	8020160024	Ambar Wahyu B	TI	2016	L	Lulus	2020-02-26	SMA	Pagi	Belum Bekerja	3.35	Cepat
15	8020160026	Juni Ahmad	TI	2016	P	Lulus	2020-08-29	SMA	Pagi	Belum Bekerja	3.18	Tepat
16	8020160027	Martha P	TI	2016	P	Lulus	2020-02-29	SMK	Malam	Belum Bekerja	3.33	Cepat
17	8020160028	Febry Valentino	TI	2016	L	Lulus	2020-02-25	SMA	Malam	Belum Bekerja	3.88	Cepat
18	8020160030	Ricky Fernando	TI	2016	L	Lulus	2020-02-29	SMK	Malam	Belum Bekerja	3.66	Cepat
19	8020160031	Ahmad Suhardi	TI	2016	L	Lulus	2020-03-02	SMK	Pagi	Belum Bekerja	3.22	Cepat
20	8020160033	Surya Krismanto	TI	2016	L	Lulus	2020-02-27	SMK	Pagi	Belum Bekerja	3.75	Cepat

Gambar 2. Data Pelatihan (*Training*)

#	NIM	Nama	Prodi	Angkatan	JK	S.Kuliah	Asal Sekolah	Shift	Status Kerja	IPK	Ket. Lulus
1	8020200001	Khairunnisa Putr	TI	2020	P	Aktif	SMK	Pagi	Belum Bekerja	3.84	?
2	8020200003	Rama Rifki Fadli	TI	2020	L	Aktif	SMA	Malam	Belum Bekerja	3.17	?
3	8020200004	Laurent Bellavin	TI	2020	P	Aktif	SMA	Malam	Belum Bekerja	3.86	?
4	8020200005	Willie Prakarsa H	TI	2020	L	Aktif	SMA	Malam	Belum Bekerja	3.8	?
5	8020200006	Fathoni Athallah	TI	2020	L	Aktif	SMK	Pagi	Belum Bekerja	3.81	?
6	8020200007	Cinta Anindya P	TI	2020	P	Aktif	SMA	Pagi	Belum Bekerja	3.96	?
7	8020200008	Jenia Valiant Tor	TI	2020	P	Aktif	SMA	Malam	Sudah Bekerja	3.54	?
8	8020200009	Elbert Devino He	TI	2020	L	Aktif	SMK	Malam	Belum Bekerja	3.85	?
9	8020200010	Tiko Nur Annisa	TI	2020	P	Aktif	SMA	Pagi	Belum Bekerja	3.96	?
10	8020200011	M. Raihan Mufa	TI	2020	L	Aktif	MA	Pagi	Belum Bekerja	3.92	?
11	8020200012	Adryan	TI	2020	L	Aktif	SMK	Pagi	Belum Bekerja	3.93	?
12	8020200014	Elvi Liza Oktavia	TI	2020	P	Aktif	SMK	Pagi	Belum Bekerja	3.87	?
13	8020200015	Eric Wardata	TI	2020	L	Aktif	SMK	Malam	Belum Bekerja	3.8	?
14	8020200016	Vincent Harianto	TI	2020	L	Aktif	SMK	Pagi	Belum Bekerja	3.56	?
15	8020200017	Defry Nazrian	TI	2020	L	Aktif	SMA	Pagi	Belum Bekerja	3.84	?
16	8020200018	Adrian Pirtama	TI	2020	L	Aktif	SMK	Pagi	Belum Bekerja	3.56	?
17	8020200019	Drenda Scheber	TI	2020	L	Aktif	SMA	Pagi	Belum Bekerja	3.69	?
18	8020200020	Fiki Saputra	TI	2020	L	Aktif	SMK	Pagi	Belum Bekerja	3.57	?
19	8020200021	David Liem	TI	2020	L	Aktif	SMA	Pagi	Belum Bekerja	3.74	?
20	8020200022	Mutoharoh	TI	2020	P	Aktif	SMA	Pagi	Belum Bekerja	3.76	?

Gambar 3. Data Pengujian (*Testing*)

Penelitian ini menggunakan parameter yang diambil dari hasil studi yang dilakukan oleh Meyer Mega dan Jasmir pada tahun 2023, yang mengidentifikasi 12 atribut utama, yaitu: NIM, Nama, Program Studi, Angkatan, Jenis Kelamin, Status Kuliah, Tanggal Lulus, Asal Sekolah, Shift Kuliah, Status Pekerjaan, Indeks Prestasi Kumulatif (IPK), dan Keterangan Lulus. Dalam penelitian tersebut, analisis dilakukan melalui perhitungan manual dan pemanfaatan aplikasi WEKA (Meyer Mega dan Jasmir, 2023). Untuk meningkatkan akurasi dan efektivitas hasil perhitungan yang diperoleh, penelitian ini melakukan modifikasi dalam metode perhitungan dengan menggunakan aplikasi *python*, dengan harapan menghasilkan pencapaian yang lebih efektif dan efisien.

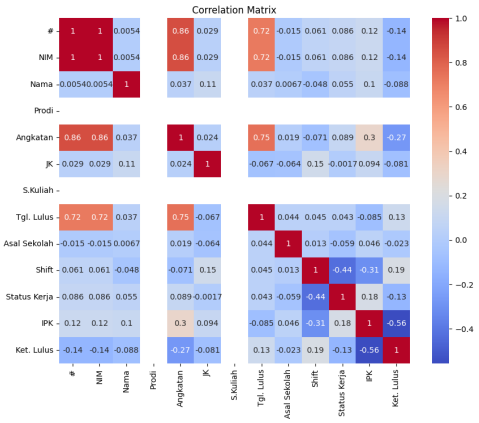
Penelitian ini juga membutuhkan beberapa sumber data sebagai penunjang semua kegiatan penelitian dengan judul prediksi masa studi mahasiswa menggunakan *decision tree* yang dioptimalisasi dengan *hyperparameter tuning*. Penelitian ini menggunakan metode *decision tree* yang dimana data dari penelitian ini menggunakan data *training* 251 dan data *testing* 109 data.

3.4. Pra-pemrosesan Data (Data *Preprocessing*)

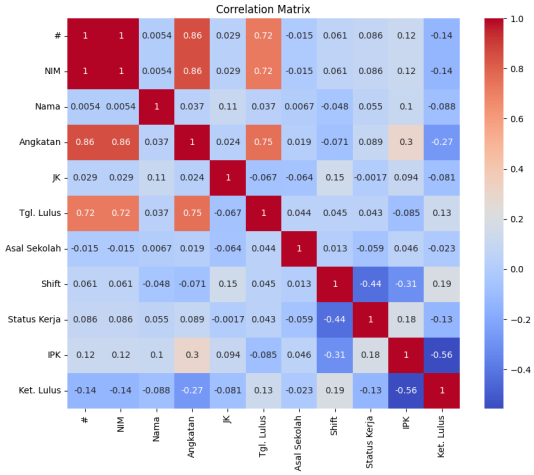
Sebelum digunakan dalam pemodelan, data mentah akan melewati serangkaian tahapan pra-pemrosesan yang esensial untuk memastikan kualitas data. Tahapan ini meliputi:

- **Initial Data**: penelitian ini menggunakan metode *Decision Tree* dengan 360 data mahasiswa (251 *training*, 109 *testing*).
- ***Data Cleaning*** dan ***Feature Selection***: data dibersihkan dari entri yang tidak lengkap, kemudian dilakukan seleksi atribut dengan analisis korelasi untuk menghapus variabel kurang relevan.
- ***Transformasi Data***: mengubah atribut kategorikal menjadi numerik agar dapat diproses algoritma.

Pada tahap ini penulis melakukan seleksi data serta *preprocessing* menggunakan data *training* dan data *testing* yang memiliki 12 atribut. Pada penyeleksian atribut penulis menggunakan *python* yang menghasilkan data seperti tabel berikut:



Gambar 4. Hasil Seleksi *Python*



Gambar 5. Hasil Seleksi Menggunakan *Python*
Setelah Melakukan Pembersihan Data

Gambar 4 menjelaskan hasil dari seleksi atribut menggunakan *python*, dan menghasilkan bahwa atribut dari prodi dan s.kuliah tidak mempengaruhi dalam proses perhitungan sehingga atribut tidak akan digabungkan dalam proses perhitungan.

3.5. Pembangunan dan Optimalisasi Model

Model klasifikasi akan dibangun menggunakan algoritma *decision tree*.

1) Pelatihan Model Awal

Pada tahap ini, model dilatih menggunakan data pelatihan yang sudah diproses, yang menghasilkan akurasi awal sebesar 62%.

Tabel 2. Pelatihan model dengan decision tree

	precision	recall	f1-score	support
0	1.00	0.62	0.77	109
2	0.00	0.00	0.00	0
accuracy			0.62	109
macro avg	0.50	0.31	0.38	109
weighted avg	1.00	0.62	0.77	109

2) Hyperparameter Tuning

Karena akurasi awal ini dinilai belum optimal, penelitian dilanjutkan dengan tahap *Hyperparameter Tuning*. Proses ini merupakan langkah krusial untuk meningkatkan kinerja model dengan menemukan kombinasi parameter terbaik secara otomatis. Melalui proses ini, akurasi model berhasil meningkat secara signifikan hingga 90%, menunjukkan kemampuan prediktif yang jauh lebih superior.

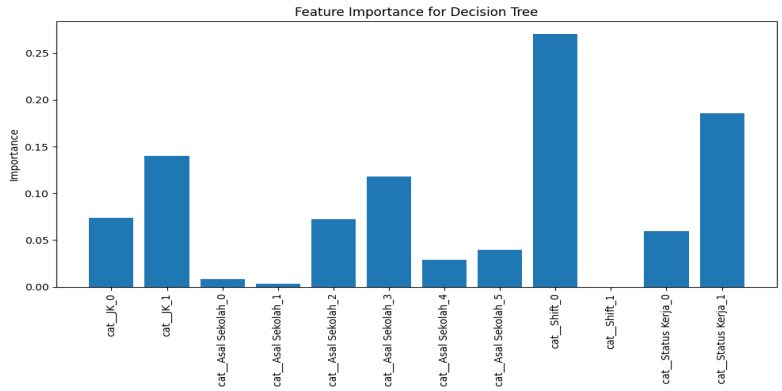
Tabel 3. Peningkatan *Decision Tree* dengan *Hyperparameter Tuning*

	precision	recall	f1-score	support
0	1.00	0.91	0.95	109
2	0.00	0.00	0.00	0
accuracy			0.91	109
macro avg	0.50	0.45	0.48	109
weighted avg	1.00	0.91	0.95	109

Hyperparameter tuning adalah cara memilih nilai terbaik sebelum pelatihan dimulai agar model pembelajaran mesin dapat bekerja lebih optimal (A. Ilemobayo et al., 2024).

3.6. Evaluasi dan Pengujian Model

Model yang telah dioptimalkan kemudian diuji menggunakan data pengujian yang tersisa. Metrik akurasi digunakan untuk mengukur seberapa baik model dapat memprediksi kelulusan. Hasil evaluasi juga menunjukkan bahwa atribut shift kuliah, status pekerjaan, dan jenis kelamin merupakan faktor-faktor paling penting yang mempengaruhi masa studi mahasiswa, dapat dilihat pada gambar 4.

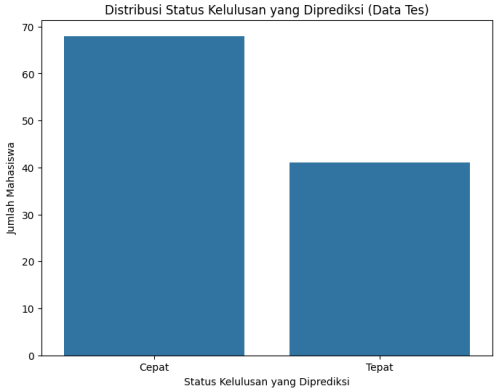


Gambar 6. Hasil Akhir Prediksi Atribut menggunakan *Python*

Setelah mengidentifikasi atribut-atribut yang signifikan, langkah berikutnya adalah meninjau hasil akhir dari prediksi kelulusan mahasiswa yang ditampilkan pada gambar selanjutnya.

index	NIM	Nama_Original	JK_Original	Asal Sekolah_Original	Shift_Original	Status Kerja_Original	Ket. Lulus_Original	Predicted Ket. Lulus_Original
0	0	Khairunnisa Putri	P	SMK	Pagi	Belum Bekerja	?	Cepat
1	1	Rama Rifki Fadli	L	SMA	Malam	Belum Bekerja	?	Cepat
2	2	Laurent Bellavinsca	P	SMA	Malam	Belum Bekerja	?	Cepat
3	3	Willie Prakarsa Hartanto	L	SMA	Malam	Belum Bekerja	?	Cepat
4	4	Fathoni Athallah	L	SMK	Pagi	Belum Bekerja	?	Tepat
5	5	Cinta Anindya Putri Abdoellah	P	SMA	Pagi	Belum Bekerja	?	Cepat
6	6	Jenia Valiant Tori	P	SMA	Malam	Sudah Bekerja	?	Tepat
7	7	Elbert Devino Henata	L	SMK	Malam	Belum Bekerja	?	Tepat
8	8	Tiko Nur Annisa	P	SMA	Pagi	Belum Bekerja	?	Cepat
9	9	M. Raihan Mufadhal	L	MA	Pagi	Belum Bekerja	?	Cepat
10	10	Adryan	L	SMK	Pagi	Belum Bekerja	?	Tepat
11	11	Elvi Liza Oktavia	P	SMK	Pagi	Belum Bekerja	?	Cepat
12	12	Eric Wardata	L	SMK	Malam	Belum Bekerja	?	Tepat
13	13	Vincent Harianto	L	SMK	Pagi	Belum Bekerja	?	Tepat
14	14	Defry Nazrian	L	SMA	Pagi	Belum Bekerja	?	Cepat
15	15	Adrian Pirtama	L	SMK	Pagi	Belum Bekerja	?	Tepat
16	16	Drenda Scheber	L	SMA	Pagi	Belum Bekerja	?	Cepat

Gambar 7. Hasil Prediksi Kelulusan Mahasiswa Menggunakan *Python*



Gambar 8. Diagram Hasil Prediksi Kelulusan Mahasiswa Menggunakan *Python*

Berdasarkan hasil prediksi menunjukkan dua kategori kelulusan, yakni 'cepat' dan 'tepat waktu', yang merefleksikan kemampuan sistem dalam mengklasifikasikan capaian studi serta memberikan dasar bagi evaluasi akademik dan perumusan strategi pendidikan.

4. KESIMPULAN

Merujuk pada pembahasan sebelumnya, kesimpulan yang dapat diambil adalah Penelitian ini memanfaatkan 360 data mahasiswa untuk membangun model prediksi kelulusan dengan algoritma decision tree. Hasil menunjukkan bahwa shift kuliah, status pekerjaan, dan jenis kelamin berpengaruh signifikan terhadap lama studi. Temuan ini dapat membantu perguruan tinggi dalam mengidentifikasi mahasiswa berisiko terlambat lulus dan merancang intervensi akademik yang lebih tepat guna meningkatkan keberhasilan studi.

DAFTAR PUSTAKA

Desfiandi, A., & Soewito, B. (2023). Student graduation time prediction using logistic regression, decision tree, support vector machine, and AdaBoost ensemble learning. International Journal of Information System and Computer Science, 7(3), 195–199. https://doi.org/10.56327/ijiscs.v7i2.1579.

Zhou, Z. H., Zhang, C., & Huang, Y. (2020). Machine Learning in Big Data Analytics. *Computers in Industry*, 120, 103223. DOI: 10.1016/j.compind.2020.103223

BeritaSatu, "Kenapa banyak mahasiswa gagal lulus tepat waktu? Ini alasannya," *BeritaSatu.com*, Feb. 13, 2025. [Online]. Available: https://www.beritasatu.com/nasional/2906166/kenapa-banyak-mahasiswa-gagal-lulus-tepat-waktu-ini-alasannya

KDD (Knowledge Discovery and Data Mining). (2024). Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. ACM. DOI: 10.1145/3632550.

Simões, S., Oliveira, T., & Nunes, C. (2022). Influence of computers in students' academic achievement. Heliyon, 8(3), e09004. https://doi.org/10.1016/j.heliyon.2022.e09004.

Mailana, A., Febriyani, E., & Hamdani, A. (2021). *Comparison of C4.5 algorithm and Support Vector Machine in predicting the student graduation timeliness*. Jurnal Online Informatika, 6(2), 165–172. https://doi.org/10.15575/join.v6i2.608.

Chrisnanto Y H & Abdullah G. (2021). The uses of educational data mining in academic performance analysis at higher education institutions (case study at UNJANI). Matrix : Jurnal Manajemen Teknologi Dan Informatika, 11(1), 26-35. doi:10.31940/matrix.v11i1.2330.

Putri, I. W. M., Rusdah, R., Suryadi, L., & Anubhakti, D. (2023). Prediksi kelulusan mahasiswa Fakultas Teknologi Informasi ISB Atma Luhur menggunakan algoritma C4.5. Jurnal SISFOKOM, 12(3), 363–369. (DOI tidak tersedia secara publik)Adji, T. B., & Syarif, I. (2022). *Comparative study of hyperparameter optimization for educational data classification*. Procedia Computer Science, 197, 150–158

Panggabean, J. F. R., Sirait, K., & Naibaho, F. R. (2022). *Data mining analysis using Naive Bayes algorithm and KNN to predict graduation of D3 students Department of Information Management*. International Journal of Education, Theology, and Humanities, 2(1), 1–10. https://doi.org/10.46965/ijeth.v2i1.4

Adji, T. B., & Syarif, I. (2022). Comparative study of hyperparameter optimization for educational data classification. *Procedia Computer Science, 197*, 150–158. https://doi.org/10.1016/j.procs.2021.12.157

Zahedi, L., Ghareh Mohammadi, F., Rezapour, S., Ohland, M. W., & Amini, M. H. (2021). Search algorithms for automated hyper-parameter tuning. arXiv preprint. https://doi.org/10.48550/arXiv.2104.14677.

Rukiastiandari, N., Susilowati, D., Kusumawardani, A., & Rachman, A. (2024). Optimizing decision tree algorithm with genetic algorithm for predicting student graduation. *Paradigma: Jurnal Komputer dan Informatika, 26*(1), 57–66. https://ojs.bsi.ac.id/index.php/paradigma/article/view/3165

Mega, M. & Jasmir, J. (2023). *Prediksi Masa Studi Mahasiswa Unama Jambi Menggunakan Metode Algoritma C4.5*. Jurnal Manajemen Sistem Informasi, 8(1), 140–151. DOI: 10.33998/jurnalmsi.2023.8.1.770

J. A. Ilemobayo, O. Durodola, O. Alade, O. Awotunde, A. T. Olanrewaju, O. Falana, A. Ogungbire, A. Osinuga, D. Ogunbiyi, A. Ifeanyi, I. E. Odezuligbo, dan O. E. Edu, "Hyperparameter Tuning in Machine Learning: A Comprehensive Review," *Journal of Engineering Research and Reports*, vol. 26, no. 6, pp. 388–395, Jun. 2024, doi: 10.9734/JERR/2024/v26i61188.