

Prediksi Status Akademik Mahasiswa Menggunakan Algoritma Decision Tree Berdasarkan Data Aktivitas dan Nilai Akademik

Nanda Eirlangga¹, Dimas Ajirul Anam¹, Neddy Satria Tama¹, Wardianto¹, Annisa Nur Azizah²

¹Program Studi Informatika, Universitas Nurul Huda

²Program Studi Matematika, Universitas Nurul Huda

E-mail: eirlanggananda@mail.com

Abstrak

Perkembangan teknologi informasi menghasilkan data akademik mahasiswa dalam jumlah besar yang dapat dimanfaatkan untuk mendukung proses pengambilan keputusan di perguruan tinggi. Salah satu pemanfaatannya adalah melakukan prediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik menggunakan teknik data mining. Penelitian ini bertujuan untuk menerapkan algoritma Decision Tree dalam memprediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik. Dataset yang digunakan terdiri dari 500 data mahasiswa dengan atribut jenis kelamin, umur, status menikah, kehadiran, partisipasi diskusi, nilai tugas, aktivitas e-learning, IPK, dan status akademik. Tahapan penelitian meliputi pengumpulan data, preprocessing data, pembagian data menjadi data training dan data testing, proses klasifikasi menggunakan algoritma Decision Tree, serta evaluasi model menggunakan confusion matrix dan classification report. Hasil penelitian menunjukkan bahwa model Decision Tree mampu melakukan prediksi status akademik mahasiswa dengan tingkat akurasi sebesar 64%. Hasil evaluasi juga menunjukkan bahwa atribut kehadiran, nilai tugas, aktivitas e-learning, dan IPK memiliki pengaruh terhadap hasil klasifikasi status akademik mahasiswa. Dengan demikian, algoritma Decision Tree dapat digunakan sebagai salah satu metode untuk membantu proses evaluasi akademik mahasiswa di perguruan tinggi.

Kata kunci : Data Mining, *Decision Tree*, Prediksi, Status Akademik, Mahasiswa

Abstract

The development of information technology has generated a large amount of academic data that can be utilized to support decision-making processes in higher education institutions. One of its applications is predicting students' academic status based on activity data and academic performance using data mining techniques. This study aims to implement the Decision Tree algorithm to predict students' academic status based on activity data and academic achievement. The dataset used consists of 500 student records with attributes including gender, age, marital status, attendance, discussion participation, assignment scores, e-learning activities, Grade Point Average (GPA), and academic status. The research stages include data collection, data preprocessing, dataset splitting into training and testing data, classification using the Decision Tree algorithm, and model evaluation using a confusion matrix and classification report. The results show that the Decision Tree model is capable of predicting students' academic status with an accuracy rate of 64%. The evaluation results also indicate that attendance, assignment scores, e-learning activities, and GPA significantly influence the classification of students' academic status. Therefore, the Decision Tree algorithm can be used as an alternative method to support the academic evaluation process in higher education institutions.

Keywords : Data Mining, *Decision Tree*, Prediction, Academic Status, Students

PENDAHULUAN

Perkembangan teknologi informasi pada era digital telah menghasilkan peningkatan jumlah data yang sangat besar di berbagai bidang, termasuk bidang pendidikan. Perguruan tinggi sebagai institusi pendidikan menghasilkan berbagai jenis data akademik mahasiswa setiap semester, seperti data kehadiran, nilai tugas, partisipasi diskusi, aktivitas pembelajaran daring (*e-learning*), serta

indeks prestasi kumulatif (IPK). Data-data tersebut umumnya hanya digunakan sebagai arsip akademik dan administrasi, padahal jika diolah dengan baik dapat menghasilkan informasi yang bermanfaat dalam mendukung proses pengambilan keputusan akademik [1], [2].

Salah satu permasalahan yang sering dihadapi oleh perguruan tinggi adalah kesulitan dalam melakukan pemantauan dan evaluasi status akademik mahasiswa secara cepat dan tepat. Jumlah mahasiswa yang terus meningkat setiap tahun menyebabkan proses evaluasi akademik menjadi semakin kompleks. Dalam banyak kasus, mahasiswa yang mengalami penurunan prestasi akademik baru diketahui setelah memperoleh hasil belajar yang kurang memuaskan. Kondisi tersebut dapat berdampak pada keterlambatan masa studi, rendahnya tingkat kelulusan, hingga meningkatnya risiko putus kuliah. Oleh karena itu, diperlukan suatu metode yang mampu membantu pihak perguruan tinggi dalam mengidentifikasi kondisi akademik mahasiswa sejak dini sehingga dapat dilakukan tindakan pencegahan dan pembinaan secara lebih efektif [3], [4].

Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah data mining. Data mining merupakan proses penggalian informasi, pola, hubungan, dan pengetahuan baru dari sekumpulan data berukuran besar menggunakan teknik statistik, kecerdasan buatan, dan machine learning [1], [10].

Penerapan data mining memungkinkan organisasi memperoleh informasi yang sebelumnya tidak terlihat sehingga dapat digunakan sebagai dasar dalam pengambilan keputusan. Dalam bidang pendidikan, penerapan data mining dikenal sebagai *Educational Data Mining* (EDM), yaitu pemanfaatan teknik analisis data untuk memahami karakteristik, perilaku, dan performa mahasiswa berdasarkan data akademik yang tersedia [4], [15].

Berbagai penelitian menunjukkan bahwa data akademik mahasiswa dapat dimanfaatkan untuk melakukan prediksi performa dan status akademik mahasiswa. Atribut seperti kehadiran, nilai tugas, aktivitas pembelajaran, partisipasi diskusi, dan IPK memiliki hubungan yang erat dengan keberhasilan akademik mahasiswa. Oleh karena itu, analisis terhadap data akademik menjadi salah satu langkah penting dalam mendukung proses evaluasi dan peningkatan kualitas pendidikan di perguruan tinggi [7], [8], [12].

Metode klasifikasi yang banyak digunakan dalam penelitian data mining adalah algoritma *Decision Tree*. Algoritma ini bekerja dengan membentuk struktur pohon keputusan berdasarkan atribut yang paling berpengaruh terhadap hasil klasifikasi. Setiap cabang pada pohon keputusan merepresentasikan aturan tertentu yang digunakan dalam proses pengambilan keputusan. Keunggulan utama algoritma *Decision Tree* adalah kemampuannya menghasilkan model yang mudah dipahami, mudah diinterpretasikan, serta mampu memberikan hasil klasifikasi yang cukup baik pada berbagai jenis dataset [5], [6].

Selain itu, algoritma ini banyak digunakan dalam penelitian pendidikan karena mampu menjelaskan hubungan antar atribut yang memengaruhi hasil akademik mahasiswa [11].

Beberapa penelitian sebelumnya telah menerapkan algoritma *Decision Tree* untuk memprediksi tingkat kelulusan, prestasi belajar, dan status akademik mahasiswa. Hasil penelitian menunjukkan bahwa metode tersebut mampu membantu proses analisis data akademik dan memberikan informasi yang berguna bagi institusi pendidikan dalam melakukan evaluasi akademik [7], [14].

Dengan memanfaatkan model klasifikasi yang dihasilkan, perguruan tinggi dapat mengidentifikasi mahasiswa yang berpotensi mengalami permasalahan akademik sehingga dapat dilakukan pendampingan lebih awal. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menerapkan algoritma *Decision Tree* dalam memprediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik. Dataset yang digunakan terdiri dari 500 data mahasiswa dengan atribut jenis kelamin, umur, status menikah, kehadiran, partisipasi diskusi, nilai tugas, aktivitas e-learning, IPK, dan status akademik mahasiswa. Penelitian dilakukan melalui

tahapan pengumpulan data, preprocessing data, pembentukan model klasifikasi menggunakan algoritma *Decision Tree*, serta evaluasi hasil prediksi menggunakan data pengujian [9], [13].

Roadmap penelitian dimulai dari identifikasi permasalahan, pengumpulan dataset mahasiswa, preprocessing data, pembagian data menjadi data training dan data testing, proses klasifikasi menggunakan algoritma *Decision Tree*, hingga evaluasi hasil model menggunakan akurasi, *classification report*, dan *confusion matrix*. Outcome yang diharapkan dari penelitian ini adalah diperolehnya model prediksi yang mampu membantu proses evaluasi akademik mahasiswa secara lebih efektif. Adapun output penelitian berupa artikel ilmiah dan model klasifikasi yang dapat digunakan sebagai referensi dalam pengembangan sistem pendukung keputusan akademik di lingkungan perguruan tinggi.

METODE PENELITIAN

Penelitian ini menggunakan metode data mining dengan pendekatan klasifikasi menggunakan algoritma *Decision Tree* untuk memprediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik. Metode klasifikasi dipilih karena mampu mengelompokkan data ke dalam kategori tertentu berdasarkan atribut yang dimiliki. Algoritma *Decision Tree* merupakan salah satu metode klasifikasi yang banyak digunakan karena mampu menghasilkan model yang mudah dipahami, mudah diinterpretasikan, dan memiliki performa yang baik dalam proses pengambilan keputusan [3], [6].

Dalam penelitian ini, kategori yang diprediksi adalah status akademik mahasiswa, yaitu "Lulus" dan "Tidak Lulus". Data yang digunakan dalam penelitian ini merupakan dataset mahasiswa yang terdiri dari 500 data dengan 10 atribut, yaitu nama, jenis kelamin, umur, status menikah, kehadiran (%), partisipasi diskusi, nilai tugas, aktivitas e-learning, IPK, dan status akademik. Atribut status akademik digunakan sebagai variabel target (*target variable*), sedangkan atribut lainnya digunakan sebagai variabel prediktor dalam proses klasifikasi. Penelitian dilakukan menggunakan perangkat keras berupa laptop dengan sistem operasi Windows 10. Adapun perangkat lunak yang digunakan adalah Python yang dijalankan melalui Google Colab. Beberapa pustaka (*library*) yang digunakan dalam penelitian ini meliputi Pandas untuk pengolahan data, NumPy untuk komputasi numerik, Matplotlib untuk visualisasi data, serta Scikit-learn untuk implementasi algoritma *Decision Tree* dan evaluasi model [2], [11].

Tahapan penelitian dimulai dengan proses pengumpulan dataset mahasiswa yang digunakan sebagai sumber data penelitian. Dataset kemudian dianalisis untuk mengetahui struktur data, jumlah atribut, tipe data, serta kelengkapan data. Hasil pemeriksaan menunjukkan bahwa dataset terdiri dari 500 data mahasiswa dan seluruh atribut memiliki nilai lengkap sehingga tidak ditemukan data yang hilang (*missing value*). Tahap berikutnya adalah preprocessing data. Preprocessing merupakan tahapan penting dalam data mining yang bertujuan untuk meningkatkan kualitas data sebelum dilakukan proses pemodelan dan analisis lebih lanjut [1], [10].

Pada tahap ini dilakukan transformasi data kategorikal menjadi data numerik menggunakan teknik encoding agar dapat diproses oleh algoritma *Decision Tree*. Atribut jenis kelamin, status menikah, dan status akademik dikonversi ke dalam bentuk numerik sehingga seluruh data dapat digunakan dalam proses klasifikasi. Setelah preprocessing selesai, data dibagi menjadi dua kelompok menggunakan metode *train-test split*. Sebanyak 80% data digunakan sebagai data training dan 20% digunakan sebagai data testing. Dari total 500 data, diperoleh 400 data training dan 100 data testing. Data training digunakan untuk membangun model klasifikasi, sedangkan data testing digunakan untuk mengukur kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya. Pembagian data training dan testing merupakan prosedur yang umum digunakan dalam machine learning untuk menghindari *overfitting* serta mengukur performa model secara objektif [2], [5].

Pada tahap pemodelan, algoritma *Decision Tree* diterapkan untuk membentuk model klasifikasi berdasarkan atribut yang terdapat pada dataset. Algoritma ini bekerja dengan membangun struktur pohon keputusan yang terdiri dari node dan cabang yang merepresentasikan aturan

keputusan. Model yang terbentuk kemudian digunakan untuk melakukan prediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik [3], [6].

Tahap terakhir adalah evaluasi model. Evaluasi dilakukan menggunakan beberapa metrik, yaitu accuracy, classification report, dan confusion matrix. Nilai accuracy digunakan untuk mengukur tingkat ketepatan model dalam melakukan klasifikasi. Classification report digunakan untuk melihat nilai precision, recall, dan f1-score pada masing-masing kelas. Sedangkan confusion matrix digunakan untuk mengetahui jumlah prediksi yang benar dan salah pada setiap kategori status akademik [2], [11].

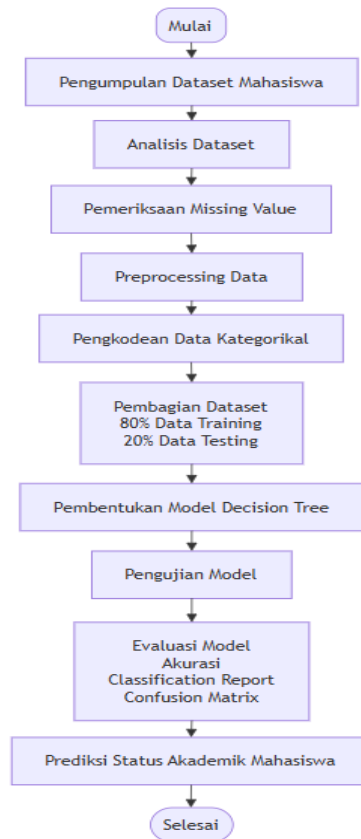
Alur penelitian secara keseluruhan dimulai dari pengumpulan dataset, pemeriksaan kualitas data, preprocessing data, pembagian data training dan testing, pembentukan model menggunakan algoritma Decision Tree, evaluasi model, serta prediksi status akademik mahasiswa berdasarkan data yang diberikan. Tahapan penelitian tersebut dilakukan untuk menghasilkan model klasifikasi yang dapat digunakan sebagai pendukung proses evaluasi akademik mahasiswa di perguruan tinggi.

Tabel 1. Variabel Penelitian

No	Variabel	Jenis
1	Jenis Kelamin	Input
2	Umur	Input
3	Status Menikah	Input
4	Kehadiran (%)	Input
5	Partisipasi Diskusi	Input
6	Nilai Tugas	Input
7	Aktivitas E-Learning	Input
8	IPK	Input
9	Status Akademik	Target

Berdasarkan Tabel 1, penelitian ini menggunakan delapan variabel input dan satu variabel target yaitu status akademik mahasiswa. Variabel input digunakan sebagai atribut prediktor dalam proses klasifikasi, sedangkan variabel target digunakan sebagai kelas yang akan diprediksi oleh algoritma Decision Tree.

Tahap selanjutnya adalah preprocessing data. Preprocessing dilakukan untuk memastikan data yang digunakan memiliki kualitas yang baik sebelum memasuki proses klasifikasi. Berdasarkan hasil pemeriksaan dataset, seluruh atribut memiliki nilai lengkap sehingga tidak ditemukan data yang hilang (*missing value*). Dengan demikian, proses pembersihan data tidak diperlukan karena dataset telah memenuhi syarat untuk digunakan pada tahap berikutnya. Setelah proses pemeriksaan data selesai, dilakukan transformasi data kategorikal ke dalam bentuk numerik menggunakan teknik encoding. Proses ini dilakukan karena algoritma *Decision Tree* hanya dapat mengolah data dalam bentuk numerik. Atribut jenis kelamin, status menikah, dan status akademik dikonversi menjadi nilai numerik sehingga seluruh atribut dapat diproses oleh model klasifikasi. Tahap berikutnya adalah pembagian dataset menjadi data training dan data testing menggunakan metode *train-test split*. Pembagian data dilakukan dengan perbandingan 80:20, di mana 80% data digunakan sebagai data training dan 20% digunakan sebagai data testing. Dari total 500 data mahasiswa, diperoleh 400 data training dan 100 data testing. Data training digunakan untuk membangun model *Decision Tree*, sedangkan data testing digunakan untuk mengevaluasi kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya.

Gambar 1. Alur Penelitian Prediksi Status Akademik Mahasiswa Menggunakan Algoritma Decision Tree

Berdasarkan Gambar 1, penelitian diawali dengan proses pengumpulan dataset mahasiswa yang terdiri dari 500 data. Selanjutnya dilakukan analisis dataset dan pemeriksaan missing value untuk memastikan kualitas data yang digunakan dalam penelitian. Setelah itu dilakukan preprocessing data melalui proses pengkodean data kategorikal menjadi data numerik agar dapat diproses oleh algoritma Decision Tree.

Tahap berikutnya adalah pembagian dataset menjadi data training sebesar 80% dan data testing sebesar 20%. Data training digunakan untuk membentuk model klasifikasi menggunakan algoritma Decision Tree, sedangkan data testing digunakan untuk menguji kemampuan model dalam melakukan prediksi. Model yang telah dibangun kemudian dievaluasi menggunakan nilai akurasi, classification report, dan confusion matrix. Tahap terakhir adalah melakukan prediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik yang dimiliki.

HASIL DAN PEMBAHASAN

Pada penelitian ini digunakan dataset mahasiswa yang terdiri dari 500 data dengan atribut jenis kelamin, umur, status menikah, kehadiran, partisipasi diskusi, nilai tugas, aktivitas e-learning, IPK, dan status akademik. Dataset tersebut telah melalui tahap preprocessing dan siap digunakan dalam proses klasifikasi menggunakan algoritma Decision Tree. Hasil penelitian meliputi pembagian dataset, evaluasi model, serta analisis hasil klasifikasi yang diperoleh.

Tabel 2. Hasil Pemeriksaan Missing Value

Atribut	Missing Value
Nama	0
Jenis Kelamin	0
Umur	0

Status Menikah	0
Kehadiran (%)	0
Partisipasi Diskusi	0
Nilai Tugas	0
Aktivitas E-Learning	0
IPK	0
Status Akademik	0

Berdasarkan Tabel 2, seluruh atribut pada dataset memiliki nilai missing value sebesar 0. Hasil ini menunjukkan bahwa dataset yang digunakan memiliki kualitas data yang baik dan tidak memerlukan proses penanganan data hilang sebelum dilakukan tahap klasifikasi.

Tabel 3. Hasil Transformasi Data (Encoding)

Atribut	Kategori	Nilai
Jenis Kelamin	Laki-laki	0
Jenis Kelamin	Perempuan	1
Status Menikah	Belum Menikah	0
Status Menikah	Menikah	1
Status Akademik	Tidak Lulus	0
Status Akademik	Lulus	1

Berdasarkan Tabel 3, atribut kategorikal diubah ke dalam bentuk numerik menggunakan teknik encoding. Transformasi ini diperlukan agar seluruh atribut dapat diproses oleh algoritma Decision Tree yang bekerja menggunakan data numerik.

Tabel 4. Pembagian Dataset

Jenis Data	Jumlah
Data Training	400
Data Testing	100
Total Data	500

Berdasarkan Tabel 4, dataset yang digunakan dalam penelitian ini berjumlah 500 data mahasiswa. Dataset kemudian dibagi menjadi dua bagian menggunakan metode train-test split dengan perbandingan 80:20. Sebanyak 400 data digunakan sebagai data training untuk membangun model Decision Tree, sedangkan 100 data digunakan sebagai data testing untuk mengukur kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya.

Tabel 5. Hasil Evaluasi Model

Metode	Akurasi
Decision Tree	64%

Berdasarkan Tabel 3, algoritma Decision Tree menghasilkan tingkat akurasi sebesar 64% dalam memprediksi status akademik mahasiswa. Nilai akurasi tersebut menunjukkan bahwa model mampu mengklasifikasikan data mahasiswa dengan tingkat ketepatan yang cukup baik berdasarkan atribut yang digunakan dalam penelitian. Hasil ini menunjukkan bahwa algoritma Decision Tree dapat digunakan untuk membantu proses prediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik.

Tabel 6. Clasficion Report

Kelas	Precision	Recall	F1-Score
Tidak Lulus	0.62	0.57	0.59
Lulus	0.66	0.70	0.68

Berdasarkan Tabel 6, kelas Lulus memiliki nilai recall tertinggi sebesar 0,70 yang menunjukkan bahwa model mampu mengenali sebagian besar mahasiswa dengan status lulus. Sementara itu, nilai precision pada kelas Lulus sebesar 0,66 menunjukkan bahwa sebagian besar prediksi lulus yang dihasilkan model sesuai dengan kondisi sebenarnya.

Tabel 7. Confusion Matrix

	Prediksi Tidak Lulus	Prediksi Lulus
Aktual Tidak Lulus	26	20
Aktual Lulus	16	38

Berdasarkan Tabel 7, model berhasil mengklasifikasikan 26 data mahasiswa dengan status tidak lulus dan 38 data mahasiswa dengan status lulus secara benar. Namun masih terdapat 20 data mahasiswa yang sebenarnya tidak lulus tetapi diprediksi lulus serta 16 data mahasiswa yang sebenarnya lulus tetapi diprediksi tidak lulus. Hasil tersebut menunjukkan bahwa model masih mengalami beberapa kesalahan klasifikasi sehingga tingkat akurasi yang diperoleh belum mencapai nilai yang optimal.

Berdasarkan hasil evaluasi tersebut, model memperoleh tingkat akurasi sebesar 64%. Nilai akurasi ini menunjukkan bahwa model telah mampu mengenali pola hubungan antara atribut kehadiran, partisipasi diskusi, nilai tugas, aktivitas e-learning, dan IPK terhadap status akademik mahasiswa. Meskipun demikian, tingkat akurasi yang diperoleh masih dapat ditingkatkan melalui penambahan jumlah data, proses seleksi fitur, maupun penggunaan metode klasifikasi lainnya yang memiliki kemampuan prediksi lebih tinggi. Hasil pengujian juga menunjukkan bahwa atribut aktivitas akademik memiliki pengaruh terhadap hasil klasifikasi status akademik mahasiswa. Mahasiswa yang memiliki tingkat kehadiran tinggi, nilai tugas yang baik, aktivitas e-learning yang aktif, serta IPK yang tinggi cenderung memiliki status akademik yang lebih baik dibandingkan mahasiswa dengan nilai atribut yang rendah. Oleh karena itu, atribut-atribut tersebut dapat digunakan sebagai indikator dalam proses evaluasi akademik mahasiswa di perguruan tinggi.

Hasil penelitian menunjukkan bahwa algoritma Decision Tree mampu digunakan untuk memprediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik. Tingkat akurasi yang diperoleh sebesar 64% menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengidentifikasi pola hubungan antara atribut input dan status akademik mahasiswa. Hasil tersebut membuktikan bahwa teknik data mining dapat dimanfaatkan sebagai alat bantu dalam proses pengambilan keputusan akademik. Atribut kehadiran, partisipasi diskusi, nilai tugas, aktivitas e-learning, dan IPK merupakan faktor yang berpengaruh terhadap hasil klasifikasi. Mahasiswa yang memiliki tingkat kehadiran tinggi, aktif mengikuti pembelajaran, serta memperoleh nilai akademik yang baik cenderung diklasifikasikan ke dalam status akademik yang lebih baik. Temuan ini menunjukkan bahwa aktivitas belajar mahasiswa memiliki hubungan yang erat dengan keberhasilan akademik yang dicapai. Algoritma Decision Tree memiliki beberapa keunggulan, yaitu mudah dipahami, mudah diinterpretasikan, dan mampu menghasilkan aturan keputusan yang jelas. Struktur pohon keputusan yang dihasilkan dapat membantu pihak perguruan tinggi dalam memahami faktor-faktor yang memengaruhi status akademik mahasiswa. Selain itu, proses implementasi algoritma relatif sederhana sehingga cocok digunakan dalam penelitian klasifikasi data akademik. Meskipun demikian, hasil penelitian menunjukkan bahwa model masih memiliki keterbatasan yang ditunjukkan oleh tingkat akurasi sebesar 64%. Nilai tersebut mengindikasikan bahwa masih terdapat data yang

belum dapat diklasifikasikan secara tepat oleh model. Hal ini dapat disebabkan oleh variasi karakteristik data yang cukup tinggi serta adanya kemiripan atribut antara mahasiswa yang memiliki status akademik berbeda.

Untuk meningkatkan performa model pada penelitian selanjutnya, dapat dilakukan penambahan jumlah dataset, penerapan teknik seleksi fitur, maupun perbandingan dengan algoritma klasifikasi lainnya seperti Random Forest, Naïve Bayes, dan Support Vector Machine. Dengan demikian, diharapkan model prediksi yang dihasilkan dapat memberikan tingkat akurasi yang lebih tinggi dan menghasilkan prediksi yang lebih baik dalam mendukung proses evaluasi akademik mahasiswa.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Decision Tree dapat digunakan untuk memprediksi status akademik mahasiswa berdasarkan data aktivitas dan nilai akademik. Penelitian ini menggunakan dataset sebanyak 500 data mahasiswa yang terdiri dari beberapa atribut, yaitu jenis kelamin, umur, status menikah, kehadiran, partisipasi diskusi, nilai tugas, aktivitas e-learning, IPK, dan status akademik. Melalui tahapan preprocessing, pembagian data training dan testing, serta proses klasifikasi menggunakan algoritma Decision Tree, diperoleh model yang mampu melakukan prediksi status akademik mahasiswa dengan tingkat akurasi sebesar 64%. Hasil penelitian menunjukkan bahwa atribut kehadiran, partisipasi diskusi, nilai tugas, aktivitas e-learning, dan IPK memiliki pengaruh terhadap hasil klasifikasi status akademik mahasiswa. Mahasiswa yang memiliki tingkat kehadiran tinggi, aktif dalam proses pembelajaran, serta memperoleh capaian akademik yang baik cenderung memiliki status akademik yang lebih baik. Dengan demikian, penerapan teknik data mining menggunakan algoritma Decision Tree dapat membantu perguruan tinggi dalam melakukan evaluasi dan pemantauan perkembangan akademik mahasiswa secara lebih efektif. Untuk penelitian selanjutnya, disarankan melakukan penambahan jumlah dataset, penerapan teknik seleksi fitur, serta membandingkan algoritma Decision Tree dengan metode klasifikasi lainnya seperti Random Forest, Naïve Bayes, dan Support Vector Machine agar diperoleh model prediksi dengan tingkat akurasi yang lebih tinggi. Selain itu, penelitian selanjutnya dapat menggunakan data akademik yang lebih beragam sehingga hasil prediksi yang diperoleh menjadi lebih optimal dan representatif.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Allah SWT atas segala rahmat dan karunia-Nya sehingga penelitian ini dapat diselesaikan dengan baik. Penulis juga mengucapkan terima kasih kepada Program Studi Informatika Universitas Nurul Huda yang telah memberikan dukungan dan fasilitas selama proses penelitian berlangsung.

Ucapan terima kasih juga disampaikan kepada dosen pengampu mata kuliah Data Mining atas bimbingan, arahan, dan masukan yang diberikan selama pelaksanaan penelitian. Selain itu, penulis mengucapkan terima kasih kepada seluruh pihak yang telah membantu dalam penyediaan data, proses penelitian, serta penyusunan artikel ini sehingga penelitian dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2012.
- [2] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. Cambridge, MA, USA: Morgan Kaufmann, 2017.
- [3] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.
- [4] F. Gorunescu, *Data Mining: Concepts, Models and Techniques*. Cham, Switzerland: Springer, 2017.
- [5] S. Raschka and V. Mirjalili, *Python Machine Learning*, 3rd ed. Birmingham, UK: Packt Publishing, 2019.

- [6] C. C. Aggarwal, *Machine Learning for Text*. Cham, Switzerland: Springer, 2018.
- [7] A. Romero and S. Ventura, "Educational Data Mining and Learning Analytics: An Updated Survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 3, pp. 1–21, 2020.
- [8] M. F. Ridho and S. Haryati, "Implementasi Data Mining untuk Prediksi Status Akademik Mahasiswa Menggunakan Algoritma C4.5," *Jurnal Informatika*, vol. 10, no. 1, pp. 55–62, 2020.
- [9] A. B. Kurniawan and D. A. Prasetyo, "Penerapan Algoritma Decision Tree untuk Prediksi Kelulusan Mahasiswa," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 2, pp. 101–108, 2021.
- [10] S. Wahyuni and M. Fikri, "Prediksi Prestasi Akademik Mahasiswa Menggunakan Metode Decision Tree," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 8, no. 1, pp. 35–42, 2022.
- [11] M. T. Rahman, M. S. Islam, and S. Hasan, "Student Academic Performance Prediction Using Machine Learning Algorithms," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 45–53, 2021.
- [12] P. Cortez and A. Silva, "Using Data Mining to Predict Secondary School Student Performance," *Expert Systems with Applications*, vol. 39, no. 17, pp. 12957–12966, 2020.
- [13] S. B. Kotsiantis, "Machine Learning Applications in Education: A Survey," *International Journal of Computer Science Issues*, vol. 18, no. 2, pp. 15–24, 2021.
- [14] N. Aljohani and H. Cristea, "Predicting Student Performance Using Classification Techniques," *Education and Information Technologies*, vol. 27, no. 4, pp. 5123–5141, 2022.
- [15] R. M. A. Rahman and M. N. Islam, "Educational Data Mining for Student Performance Analysis: A Systematic Review," *IEEE Access*, vol. 12, pp. 22541–22560, 2024.