Journal of Computer and Information Systems Ampera
Vol. 3, No. 2, May 2023 e-ISSN: 2775-2496
10.51519/journalcisa.v4i2.393 | Published By APTIKOM SUMSEL

Penerapan Data Mining untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Decision Tree C4.5

Uci Suriani

Program Studi Teknologi Rekayasa Multimedia, Politeknik Darussalam, Palembang, Indonesia
Email: uci.suryani1@gmail.com

Abstrak

Penelitian ini bertujuan untuk memprediksi tingkat kelulusan mahasiswa menggunakan metode klasifikasi dan algoritma decision tree C4.5. Data yang digunakan meliputi data mahasiswa dan Kartu Hasil Studi (KHS) dengan kriteria Indeks Prestasi Semester (IPS) dan Indeks Prestasi Kumulatif (IPK). Data dikelola melalui tahapan *Knowledge Discovery in Database* (KDD) dengan bantuan tool RapidMiner untuk memfasilitasi prediksi tingkat kelulusan mahasiswa. Penerapan data mining dengan metode klasifikasi dan algoritma C4.5 digunakan untuk menganalisis tingkat kelulusan mahasiswa berdasarkan informasi yang dihasilkan. Proses perhitungan data menggunakan algoritma C4.5 menunjukkan bahwa tingkat mahasiswa yang terlambat lulus atau tidak tepat waktu lebih rendah dibandingkan dengan tingkat mahasiswa yang lulus tepat waktu. Penelitian ini melibatkan pengujian melalui metode cross-validation dan evaluasi model *confusion matrix* yang menghasilkan akurasi prediksi sebesar 99.64%. Selain itu, evaluasi menggunakan metrik AUC menunjukkan nilai sebesar 99.5%, menandakan bahwa model memiliki kemampuan hampir sempurna dalam melakukan klasifikasi. Temuan ini menegaskan bahwa model dapat memprediksi tingkat kelulusan mahasiswa dengan akurat.

Kata Kunci: Algoritma C4.5, *Decision tree, Data mining*, Klasifikasi kelulusan mahasiswa, Sistem pengambilan keputusan

1. PENDAHULUAN

Pada era informasi saat ini, sistem pendidikan tinggi dihadapkan pada tuntutan untuk mengoptimalkan proses pengambilan keputusan terkait kelulusan mahasiswa. Menentukan kelulusan mahasiswa tidak hanya didasarkan pada satu faktor tunggal, tetapi melibatkan berbagai atribut seperti nilai ujian, kehadiran, kelas yang diikuti, dan faktor-faktor lainnya.

Dalam bidang *data mining*, algoritma C4.5 telah dikenal sebagai salah satu algoritma pembelajaran mesin yang efektif untuk membangun model keputusan berbasis pohon [1]. Algoritma ini mampu menangani data kategorikal dan numerik, serta memiliki kemampuan untuk melakukan pemilihan atribut yang relevan untuk pengambilan keputusan. Dengan demikian, implementasi algoritma C4.5 pada data mining dapat dijadikan pendekatan yang potensial dalam menentukan kelulusan mahasiswa [2].

Penelitian terdahulu yang dilakukan oleh Hermawanti, dkk. [3], melakukan implementasi Algoritma C4.5 untuk prediksi kelulusan tepat waktu. Teknik klasifikasi yang digunakan yaitu decision tree dengan penerapan algoritma C4.5. Masukkan yang digunakan yaitu berupa atribut dari data mahasiswa meliputi IP (Indeks Prestasi) per semester dari semester 1 sampai 7, dan nilai BTQ di semester 6. Data mahasiswa tersebut merupakan data contoh training yang digunakan dalam penyusunan *decision tree*. Pengujian ini menggunakan *data training* mahasiswa yang sudah lulus pada tahun 2015 sampai 2018. Pengetahuan ini dapat dimanfaatkan oleh pihak Program Studi Teknik Informatika UMMI sebagai langkah preventif untuk menghindari penurunan kelulusan mahasiswa setiap tahunnya.

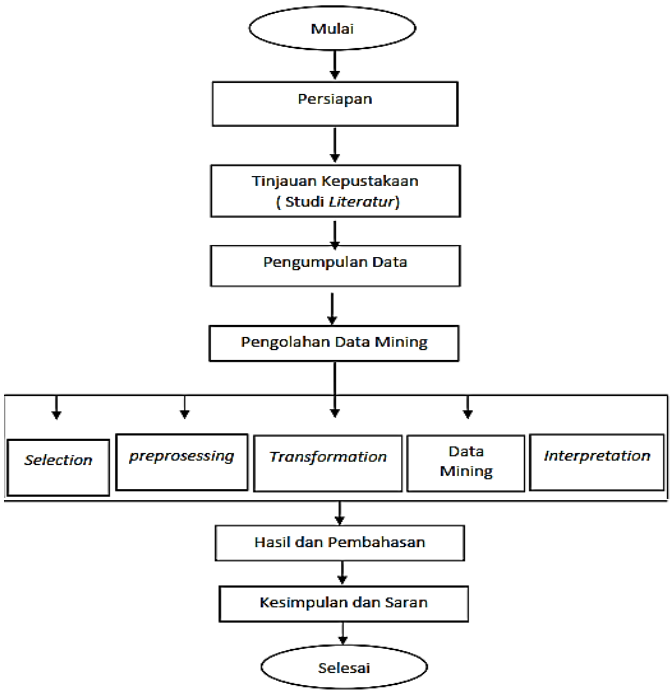
Penelitian selanjutnya dilakukan oleh Rahman dkk. [4] melakukan prediksi kelulusan mahasiswa menggunakan algoritma C4.5. Jumlah kelulusan mahasiswa sangat mempengaruhi penilaian akreditasi bagi suatu perguruan tinggi. Banyak faktor yang sangat berpengaruh terhadap ketepatan waktu kelulusan bagi mahasiswa seperti, jenis kelamin, umur, status pernikahan, IPK, dan status pekerjaan. Kelima faktor inilah yang akan dijadikan variabel input dalam menentukan klasifikasi kelulusan mahasiswa. Variabel-variabel tersebut akan diolah dengan algoritma KNN (*K-Nearest Neighbor*), *Naive Baye*s, C4.5. Data mahasiswa diseleksi sebanyak 300 mahasiswa. Data preprocessing menggunakan data mahasiswa yang terdiri dari data pribadi (jenis kelamin, umur, status pernikahan, dan status pekerjaan) dan data akademik (IPK). Berdasarkan data kelulusan, 203 siswa lulus tepat waktu dan 97 siswa lulus terlambat. Setelah dilakukan transformasi, keseluruhan data dapat digunakan karena tidak ada nilai yang kosong. Data yang diubah adalah umur (muda : 19 - 24, tua : 25 - 50) dan IPK (besar : 3 - 4, kecil : 1 - 2.9). Hasil *confusion matrix* menunjukkan bahwa *Naive Bayes* mempunyai akurasi 100% dan AUC 100% lebih tinggi dibandingkan dengan C4.5 dan KNN. Sehingga

algoritma *Naive Bayes* mempunyai kinerja yang lebih baik dibandingkan dengan KNN dan C4.5.

Maka penelitian ini bertujuan untuk mengimplementasikan algoritma C4.5 pada *data mining* guna menentukan kelulusan mahasiswa. Dalam implementasi ini, akan digunakan dataset yang terdiri dari atribut-atribut relevan yang berkaitan dengan prestasi akademik mahasiswa. Dengan memanfaatkan kemampuan algoritma C4.5 dalam memodelkan data dan mengambil keputusan, diharapkan dapat dikembangkan suatu model yang dapat memprediksi kelulusan mahasiswa secara akurat.

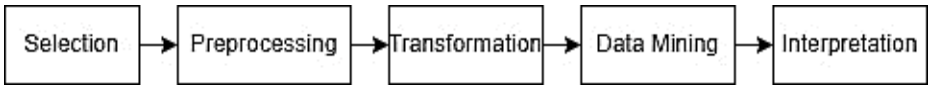
2. METODOLOGI

Data yang digunakan dalam penelitian ini adalah data mahasiswa aktif dari tahun 2015 sampai 2017 dan data Kartu Hasil Studi (KHS) yang diperoleh melalui Unit PUSTIPD UIN Raden Fatah. Proses eksperimen dilakukan melalui beberapa tahapan, seperti yang diilustrasikan pada Gambar 1 di bawah ini:



Gambar 1. Tahapan Penelitian

Penelitian ini mengimplementasikan pengolahan *data mining* dengan mengikuti tahapan yang terstruktur dalam *Knowledge Discovery in Database* (KDD) [5]. Tujuannya adalah untuk menghasilkan informasi yang relevan sesuai dengan urutan proses yang telah ditetapkan. Tahapan-tahapan yang digunakan dalam penelitian ini dapat dilihat dengan jelas pada Gambar 2 sebagai berikut.



Gambar 2. Tahapan *Knowledge Discovery in Database* (KDD)

2.1. *Data Selection*

Tahap awal ini merupakan langkah penting untuk memastikan data yang digunakan dalam proses KDD adalah data yang tepat dan berkualitas. Sumber data yang diambil dalam penelitian ini berasal dari tahun-tahun mahasiswa 2015 hingga 2017 yang tercatat pada Fakultas Tarbiyah dan Keguruan serta Fakultas Sains dan Teknologi UIN Raden Fatah Palembang. Pertama-tama, dilakukan penentuan kriteria atau parameter yang akan digunakan untuk memilih data yang relevan sesuai dengan tujuan KDD dan jenis masalah yang ingin diselesaikan. Sebagai contoh, jika tujuan KDD adalah untuk menentukan kelulusan mahasiswa, maka kriteria seleksi mungkin mencakup nilai-nilai akademik, jumlah kredit yang diambil, atau informasi pribadi mahasiswa.

Data dari berbagai sumber dikumpulkan sebagai langkah awal dalam proses seleksi dan kemudian disaring sesuai dengan kriteria seleksi yang telah ditetapkan. Pada tahap ini, dilakukan identifikasi dan penghapusan data duplikat atau data yang redundan. Menghilangkan data duplikat menjadi langkah penting untuk menjaga integritas dan konsistensi data yang digunakan dalam analisis.

2.2. *Data Preprocessing*

Pada tahap *Data Preprocessing* (Pra-pemrosesan Data), langkah selanjutnya setelah pemilihan data pada tahapan sebelumnya (*Data Selection*) adalah mempersiapkan data agar siap untuk diolah dan dianalisis secara lebih mendalam [6]. Tahap ini sangat krusial karena

kualitas data yang baik dan terstruktur akan memberikan dampak positif pada hasil analisis dan pengambilan keputusan. Beberapa hal yang dilakukan pada tahap *data preprocessing* pada penelitian ini yaitu pembersihan data, pengurangan data, dan normalisasi data.

2.3. *Transformation*

Data Transformation merupakan tahapan yang penting untuk mempersiapkan data agar lebih cocok dengan metode analisis berikutnya, seperti algoritma *Machine Learning* [7]. Proses transformasi data ini bertujuan untuk meningkatkan kualitas data dan memberikan wawasan lebih dalam mengenai faktor-faktor yang mempengaruhi tingkat kelulusan mahasiswa.

Tujuan utama dari *data transformation* adalah mengubah format dan representasi data agar lebih sesuai dan relevan untuk analisis lanjutan. Beberapa teknik yang digunakan pada tahapan ini antara lain transformasi atribut, transformasi skala, pembentukan fitur baru, dan diskritisasi data [8]. Dalam transformasi atribut, nilai atau representasi dari atribut tertentu diubah agar sesuai dengan kebutuhan analisis. Contohnya, atribut "Usia" dalam bentuk tahun dapat diubah menjadi kelompok usia untuk memudahkan analisis pengaruh usia terhadap kelulusan.

Selanjutnya, transformasi skala digunakan untuk menyesuaikan skala data yang berbeda pada atribut yang berbeda pula. Misalnya, atribut "IPK" dengan skala 0 hingga 4 dan atribut "Jumlah Mata Kuliah" dengan skala lebih besar. Melalui normalisasi, kedua atribut ini dapat diubah menjadi rentang yang seragam, membantu mencegah dominasi atribut dengan skala besar dalam analisis. Dengan *data transformation* yang tepat, data yang diolah akan memberikan hasil analisis yang lebih akurat dan bermanfaat dalam pengambilan keputusan mengenai kelulusan mahasiswa.

2.4. *Data Mining*

Setelah melakukan proses *preprocessing* dan transformasi data sesuai dengan teknik data mining klasifikasi dan algoritma C4.5, langkah berikutnya adalah melakukan proses analisis data mining. Karena volume data yang digunakan cukup besar, proses analisis data mining dilakukan secara manual untuk mempermudah pemahaman tentang penggunaan

teknik klasifikasi dengan algoritma C4.5. Data yang digunakan dalam proses analisis ini adalah *data training*, yang telah melalui proses data transformation sehingga siap untuk dilakukan proses *data mining*.

Proses pembentukan *data training* didasarkan pada data yang telah ada, dan harus melewati beberapa tahapan. Pertama, data harus diintegrasikan untuk menggabungkan data dari berbagai sumber menjadi satu kesatuan yang terstruktur. Setelah melalui proses data integrasi, data kemudian harus disaring (seleksi) untuk menentukan atribut-atribut mana yang dapat mempengaruhi kelulusan mahasiswa, yang disebut sebagai data target. Data target berisi atribut-atribut yang relevan dan mendukung proses *data mining* untuk memprediksi tingkat kelulusan mahasiswa. Dengan melakukan langkah-langkah ini, *data training* dapat terbentuk dengan baik dan siap digunakan dalam proses analisis data mining dengan algoritma C4.5 (Tabel 1).

Tabel 1. Contoh *Data Training*

Nim	Ips 1	Ips 2	Ips 3	Ips 4	Ips 5	Ips 6	Ips 7	Ips 8	IPK	Status Kelulusan
10260008	2.95	2.83	3.00	3.25	3.09	3.18	3.16	3.20	3.11	Terlambat
10260009	3.58	2.64	3.64	2.83	3.12	3.27	3.18	3.14	3.07	Terlambat
10260018	3.17	3.45	2.64	3.67	3.45	3.36	3.27	3.14	3.28	Tepat
10260029	3.08	2.68	3.42	3.25	3.36	3.45	3.40	3.42	3.38	Tepat
10260030	3.25	3.50	3.27	2.92	3.45	3.27	3.55	3.36	3.15	Terlambat
10260032	3.00	3.59	2.82	3.12	3.45	3.33	3.36	3.30	3.00	Terlambat
10260037	3.67	2.82	3.64	3.20	3.55	3.47	3.45	3.29	3.20	Terlambat
10260043	3.33	3.14	3.36	2.75	3.55	3.36	3.27	3.55	2.95	Terlambat

Selanjutnya, data training yang terdiri dari setiap atribut yang telah disebutkan sebelumnya akan digunakan untuk membentuk sebuah pohon keputusan menggunakan algoritma C4.5. Tahap-tahap dalam algoritma C4.5 untuk membangun *decision tree* adalah sebagai berikut [9]:

1. Menyiapkan *data training*. Data diperoleh dari koleksi sebelumnya dan dikategorikan ke dalam kelas yang sesuai.
2. Menentukan akar dari pohon keputusan. Akar dipilih dengan menghitung *gain* dari setiap atribut, dan atribut yang memiliki gain tertinggi akan menjadi akar pertama dalam pohon keputusan. Sebelum menghitung *gain*, penting untuk terlebih dahulu menghitung nilai entropy. Berikut adalah persamaan (1) yang digunakan untuk perhitungan nilai *entropy*.

$$Entropy(S) = -\sum_{i=1}^{n} pi * log2\ pi \quad (1)$$

Dimana *S* merupakan himpunan kasus, *A* merupakan fitur, *n* merupakan partisi *S*, dan *Pi* merupakan jumlah proporsi dari *Si* terhadap *S*.

3. Hitung nilai *gain* dengan menggunakan persamaan (2) sebagai berikut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{Si}{S} Entropy(Si) \quad (2)$$

Dimana S = Himpunan Kasus, A = Atribut, n = Jumlah partisi atribut A, |Si| = Jumlah kasus pada partisi ke-i, dan |S| = Jumlah kasus dalam S.

4. Ulangi pada langkah kedua sampai semua record terpartisi dengan sempurna.
5. Proses partisi pohon keputusan akan berhenti jika terpenuhi tiga kondisi berikut: pertama, semua rekaman pada simpul N memiliki kelas yang sama; kedua, tidak ada atribut dalam rekaman yang akan dipartisi lebih lanjut; dan ketiga, tidak ada rekaman di dalam cabang yang kosong. Pada titik ini, pembentukan pohon keputusan selesai karena telah mencapai kondisi akhir yang sesuai.

2.5. *Interpretation*

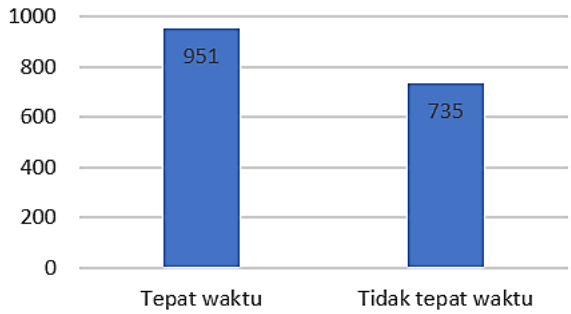
Tahap ini merupakan bagian dari proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan sesuai dengan fakta atau hipotesis yang sudah ada sebelumnya [5]. Pola informasi yang dihasilkan dari proses data mining perlu disajikan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Dalam metode decision tree, pola atau informasi tersebut berupa aturan (*rules*) yang diperoleh dari *decision tree* yang telah dibangun.

Tahap interpretasi ini penting karena hasil analisis dan model prediksi yang telah dibangun pada tahapan sebelumnya (misalnya, dengan menggunakan algoritma *machine learning*) perlu diartikan agar hasilnya dapat dipahami dan diterapkan dalam konteks pengambilan keputusan yang lebih luas. Interpretasi yang tepat memungkinkan pihak yang berkepentingan untuk mengambil langkah-langkah yang sesuai berdasarkan pola dan informasi yang ditemukan dari *data mining*.

3. HASIL DAN PEMBAHASAN

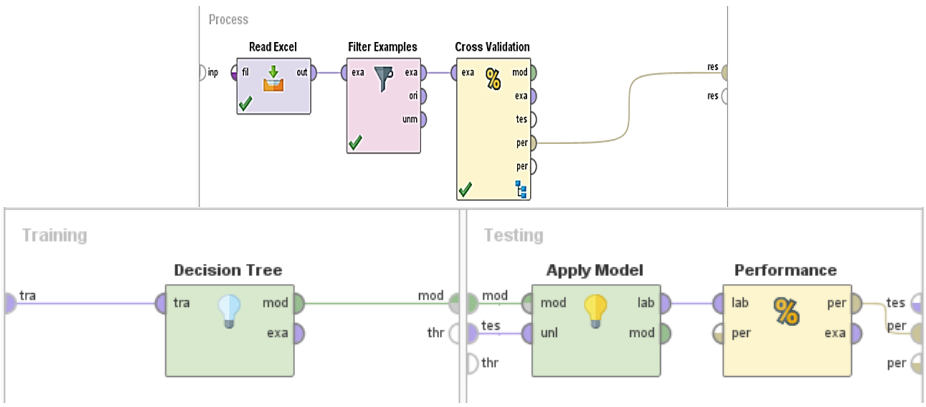
Eksperimen pada penelitian ini memanfaatkan *software* RapidMiner untuk analisis data dan algoritma C4.5 untuk pemodelan. Proses data

diolah dengan metode KDD, dan dari banyaknya data mahasiswa yang tersedia, hanya 1686 *record* data mahasiswa yang dapat digunakan dalam penelitian ini. Dari jumlah tersebut, terdapat 951 mahasiswa yang lulus tepat waktu dan 735 mahasiswa yang lulus tidak tepat waktu atau terlambat (lihat pada Gambar 3). Hasil dari implementasi algoritma C4.5 sesuai dengan pemodelan yang dihasilkan oleh *software* RapidMiner.



Gambar 3. Data tingkat kelulusan mahasiswa

Pemodelan data telah diproses oleh RapidMiner untuk mengidentifikasi pola kelulusan dan akurasi data menggunakan model *performance vector*. Dengan menghitung jumlah data yang terklasifikasi dengan benar, kita dapat mengetahui akurasi dan mendapatkan pola kelulusan berupa pohon keputusan dan rules dari pola yang terbentuk.



Gambar 4. Proses klasifikasi model dengan algoritma *Decision Tree* C4.5

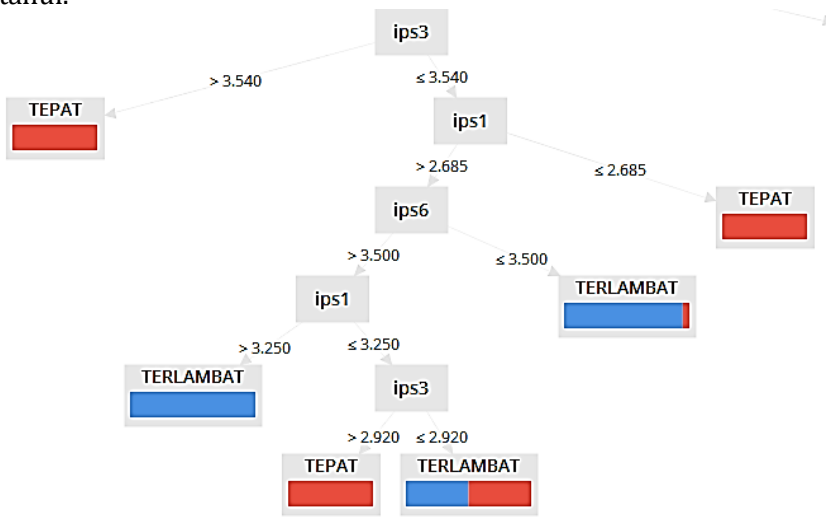
Penggunaan parameter juga mempengaruhi akurasi dan model yang dihasilkan oleh algoritma C4.5. Penelitian ini menggunakan metode *Cross*

Validation untuk menilai akurasi algoritma. Pada proses *Cross Validation*, terdapat subproses yang melibatkan operator Decision Tree sebagai algoritma yang digunakan dalam proses data mining Classification. Selain itu, terdapat operator *Apply Model* dan *Operator Performance* yang merupakan bagian dari pengolahan data mining untuk menghasilkan *Decision Tree* (lihat pada Gambar 4).

3.1 Hasil Percobaan dan Pengujian Model

Dari hasil pemodelan menggunakan algoritma *Decision Tree* (pada Gambar 5), terlihat bahwa terdapat *node* 19 dengan 31 mahasiswa yang status kelulusannya belum diketahui, apakah lulus tepat waktu atau terlambat. Di sisi lain, pada node 22 terdapat 3 mahasiswa yang lulus tepat waktu, dan pada *node* 21 terdapat 28 mahasiswa yang status kelulusannya masih belum diketahui.

Selanjutnya, pada node 24 terdapat 2 mahasiswa yang lulus tepat waktu, dan pada *node* 23 terdapat 26 mahasiswa yang status kelulusannya belum diketahui. Kemudian, pada node 26 terdapat 18 mahasiswa yang lulus terlambat dan 1 mahasiswa yang lulus tepat waktu, serta pada *node* 25 terdapat 7 mahasiswa yang status kelulusannya juga masih belum diketahui.



Gambar 5. Hasil pemodelan algoritma Decision Tree C 4.5

3.2 Evaluasi Model

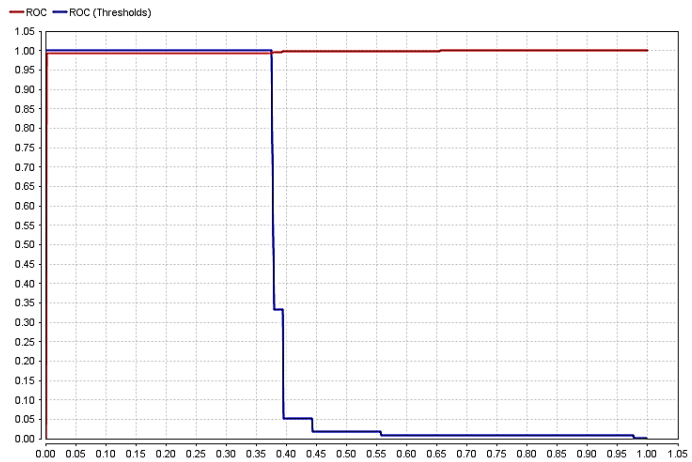
Hasil pemodelan yang telah diproses menggunakan RapidMiner tidak hanya menghasilkan pola pemodelan, tetapi juga memberikan informasi tentang keakuratan data yang digunakan. Melalui *confusion matrix*, didapatkan nilai prediksi tepat waktu sebanyak 945 mahasiswa dan nilai prediksi tidak tepat waktu sebanyak 735 mahasiswa. Dengan demikian, diperoleh nilai akurasi prediksi sebesar 99.64% (lihat pada Gambar 6).

accuracy: 99.64%

	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	735	6	99.19%
pred. TEPAT	0	945	100.00%
class recall	100.00%	99.37%	

Gambar 6. Hasil kinerja *accuracy, precision*, dan *recall* menggunakan Algoritma C 4.5

Disamping itu, dalam penelitian ini digunakan metrik evaluasi AUC (*Area Under the Curve*) untuk mengukur kualitas model klasifikasi, terutama dalam konteks klasifikasi biner. AUC mengukur sejauh mana model dapat membedakan antara dua kelas target, yaitu kelas positif dan kelas negatif. Berdasarkan Gambar 7, hasil evaluasi metrik menggunakan AUC menunjukkan nilai sebesar 99.5%.



Gambar 7. Metrik evaluasi model menggunakan AUC

Nilai AUC berada dalam rentang 0 hingga 1, di mana nilai 1 menunjukkan bahwa model sempurna dalam membedakan kelas positif dan negatif, sementara nilai 0.5 menunjukkan bahwa model tidak lebih baik daripada tebakan acak.

Dalam konteks evaluasi klasifikasi biner, nilai AUC mendekati 1 (100%) menunjukkan bahwa model hampir sempurna dalam memprediksi kelas positif dan negatif. Hal ini berarti model memiliki tingkat akurasi yang sangat tinggi dan memiliki kemampuan yang sangat baik dalam memisahkan antara sampel dari kedua kelas target.

4. KESIMPULAN

Berdasarkan hasil perhitungan *data mining* menggunakan algoritma C4.5, dapat disimpulkan bahwa kelas status kelulusan "*terlambat*" atau lulus tidak tepat waktu lebih kecil daripada kelas status kelulusan "*tepat*" atau lulus tepat waktu. Melalui proses perhitungan menggunakan metode klasifikasi dan algoritma C4.5 dengan atribut yang telah dijelaskan sebelumnya, diperoleh hasil bahwa akurasi klasifikasi kelulusan mencapai 99.64%.

Hasil akurasi sebesar 99.64% ini menunjukkan bahwa model yang dibangun mampu melakukan prediksi dengan sangat tinggi. Namun, perlu diperhatikan bahwa akurasi tinggi juga dapat disebabkan oleh kompleksitas data yang rendah, yang mengakibatkan model mampu memprediksi dengan cukup akurat. Oleh karena itu, penting untuk melakukan evaluasi lebih lanjut terhadap model dan menguji kehandalan prediksi pada data yang berbeda untuk memastikan keandalan model.

Kemampuan algoritma C4.5 dalam membentuk *decision tree* dan mengidentifikasi atribut yang relevan dalam proses *data mining* menjadi faktor penentu tingginya akurasi prediksi. Dengan demikian, hasil penelitian ini memberikan wawasan dan informasi yang berharga terkait prediksi tingkat kelulusan mahasiswa menggunakan metode klasifikasi dan algoritma C4.5. Namun, untuk aplikasi yang lebih luas dan keputusan pengambilan yang lebih kompleks, perlu dilakukan penelitian dan evaluasi lebih mendalam dengan melibatkan lebih banyak data dan menguji model pada berbagai skenario untuk memastikan keandalan dan generalisasi dari model yang dikembangkan.

REFERENSI

[1] A. Muzakir, H. Syaputra, and F. Panjaitan, "A Comparative Analysis of Classification Algorithms for Cyberbullying Crime Detection: An Experimental Study of Twitter Social Media in Indonesia," *Sci. J. Informatics; Vol 9, No 2 Novemb. 2022DO - 10.15294/sji.v9i2.35149* , Oct. 2022, [Online]. Available: https://journal.unnes.ac.id/nju/index.php/sji/article/view/35149

[2] Y. Mardi, "Data Mining: Klasifikasi Menggunakan Algoritma C4. 5," *J. Edik Inform. Penelit. Bid. Komput. Sains dan Pendidik. Inform.*, vol. 2, no. 2, pp. 213–219, 2017.

[3] S. N. Hermawanti, A. Asriyanik, and A. A. Sunarto, "Implementasi Algoritma C4. 5 Untuk Prediksi Kelulusan Tepat Waktu:(Studi Kasus: Program Studi Teknik Informatika)," *SANTIKA is a Sci. J. Sci. Technol.*, vol. 9, no. 1, pp. 853–864, 2019.

[4] A. F. A. Rahman, "PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN ALGORITMA C4. 5 (STUDI KASUS DI UNIVERSITAS PERADABAN): Array," *Indones. J. Informatics Res.*, vol. 1, no. 2, pp. 70–77, 2020.

[5] C. K. Nwagu, O. C. Omankwu, and H. Inyiama, "Knowledge Discovery in Databases (KDD): an overview," *Int J Comput Sci Inf Secur*, vol. 15, no. 12, pp. 13–16, 2017.

[6] A. Muzakir and R. A. Wulandari, "Model Data Mining sebagai Prediksi Penyakit Hipertensi Kehamilan dengan Teknik Decision Tree," 2016, [Online]. Available: https://journal.unnes.ac.id/nju/index.php/sji/article/view/4610

[7] R. Feldman and I. Dagan, "Knowledge Discovery in Textual Databases (KDT).," in *KDD*, 1995, vol. 95, pp. 112–117.

[8] F. A. Rahman, M. I. Desa, A. Wibowo, and N. A. Haris, "Knowledge discovery database (KDD)-data mining application in transportation," *Proceeding Electr. Eng. Comput. Sci. Informatics*, vol. 1, no. 1, pp. 116–119, 2014.

[9] L. Y. L. Gaol, M. Safii, and D. Suhendro, "Prediksi Kelulusan Mahasiswa Stikom Tunas Bangsa Prodi Sistem Informasi Dengan Menggunakan Algoritma C4. 5," *Brahmana J. Penerapan Kecerdasan Buatan*, vol. 2, no. 2, pp. 97–106, 2021.