

Klasifikasi Status Akademik Mahasiswa Menggunakan Decision Tree

Zaehol Fatah¹, Sinta Bella²

¹ Universitas Ibrahimy, Teknologi Informasi, Situbondo

² Universitas Ibrahimy, Sistem Informasi, Situbondo

Email: wahyunishinbell@gmail.com

Abstrak

Penelitian ini membahas pemodelan klasifikasi status mahasiswa berdasarkan data akademik dengan menerapkan algoritma *Decision Tree* menggunakan RapidMiner. Data yang dianalisis meliputi variabel seperti indeks prestasi kumulatif (IPK), jumlah SKS, lama studi, dan faktor usia. Proses klasifikasi dilakukan melalui tahapan pra-pemrosesan data, pembangunan model, serta evaluasi menggunakan confusion matrix dan metrik akurasi. Hasil analisis menunjukkan bahwa model yang dibentuk mampu mengidentifikasi pola-pola yang memengaruhi kelulusan mahasiswa, dengan tingkat akurasi mencapai 66,67%. Pendekatan ini diharapkan dapat menjadi alat bantu strategis dalam mendukung pengambilan keputusan akademik berbasis data dan memperbaiki sistem pemantauan kelulusan di lingkungan perguruan tinggi.

Kata Kunci: Klasifikasi, Decision tree, RapidMiner, Status Mahasiswa, Data Akademik

Abstarct

This study focuses on modeling the classification of student graduation status using the Decision Tree algorithm implemented through RapidMiner. The academic data includes variables such as Grade Point Average (GPA), total credit hours, study duration, and age. The classification process involves data preprocessing, model building, and evaluation using a confusion matrix and accuracy metrics. The results indicate that the model effectively identifies patterns influencing student graduation outcomes, achieving an accuracy rate of 66.67%. This approach is expected to serve as a strategic tool to support data-driven academic decision-making and to improve student progress monitoring systems in higher education institutions.

Keywords: classification, decision tree, RapidMiner, student status, academic data

PENDAHULUAN

Pendidikan tinggi merupakan jenjang lanjutan setelah pendidikan menengah yang diselenggarakan oleh perguruan tinggi. Salah satu tanggung jawab utama institusi pendidikan tinggi adalah menghasilkan lulusan yang berkualitas dan kompeten. Untuk itu, pemantauan terhadap proses dan hasil belajar mahasiswa sangat penting, salah satunya melalui indikator akademik seperti indeks prestasi kumulatif (IPK) dan lama

studi yang turut menentukan predikat kelulusan

Data akademik mahasiswa, yang mencakup Indeks Prestasi Kumulatif (IPK), Indeks Prestasi Semester (IPS), jalur masuk, dan data demografis, menyimpan informasi berharga yang dapat digunakan untuk memprediksi status akademik mahasiswa. Namun, data tersebut sering kali tidak dimanfaatkan secara optimal. Oleh karena itu, penerapan teknik data mining menjadi penting untuk menggali

informasi tersembunyi dalam data tersebut.

Ketepatan waktu kelulusan menjadi indikator penting dalam menilai efisiensi sistem pembelajaran dan mutu institusi. Mahasiswa yang mampu menyelesaikan studi tepat waktu mencerminkan keberhasilan proses akademik dan berdampak langsung terhadap akreditasi program studi. Oleh karena itu, diperlukan mekanisme prediktif yang mampu mengidentifikasi potensi keterlambatan kelulusan mahasiswa sejak dini.

Seiring dengan kemajuan teknologi informasi, data akademik kini memiliki peran strategis sebagai bahan dasar pengambilan keputusan. Teknik data mining, khususnya klasifikasi, memungkinkan institusi pendidikan mengolah data historis untuk mengungkap pola yang tersembunyi dan menghasilkan informasi prediktif yang bernilai. Salah satu algoritma yang populer dalam klasifikasi adalah Decision Tree, karena memiliki struktur aturan yang mudah dipahami dan divisualisasikan. Penggunaan perangkat lunak RapidMiner memungkinkan proses klasifikasi dilakukan secara sistematis—mulai dari prapemrosesan data, seleksi atribut, hingga evaluasi model klasifikasi.

Meskipun teknologi ini menawarkan berbagai keunggulan, tantangan utama yang dihadapi adalah bagaimana memprediksi dan meningkatkan **kinerja akademik mahasiswa** secara akurat. Ketersediaan data dalam jumlah besar dari aktivitas pembelajaran digital membuka peluang untuk diterapkannya teknik **data mining**, khususnya algoritma klasifikasi seperti **Decision Tree**, guna membangun model prediktif yang dapat mengidentifikasi mahasiswa yang berisiko mengalami kegagalan akademik sejak dini.

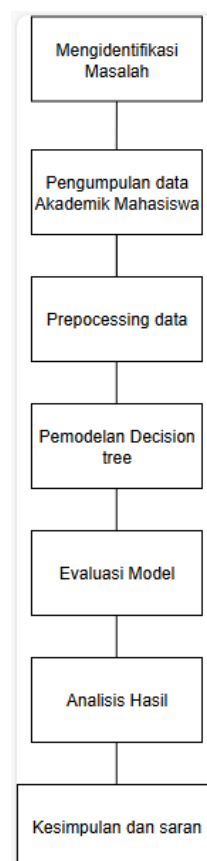
Penelitian ini bertujuan untuk membangun model klasifikasi status kelulusan mahasiswa berdasarkan data akademik seperti nilai UTS, UAS, dan nilai akhir. Model ini tidak hanya memprediksi kelulusan, tetapi juga

mengidentifikasi variabel-variabel yang memiliki pengaruh signifikan dalam proses klasifikasi. Hasil dari penelitian ini diharapkan dapat mendukung formulasi kebijakan akademik dan penerapan intervensi dini terhadap mahasiswa yang berisiko tidak menyelesaikan studi tepat waktu.

METODE

Desain penelitian ini dimulaidari beberapa tahapan yang akan di jalankan dalam penelitian ini. Dalam melakukan proses klasifikasi data prediksi kelulusan tepat waktu mahasiswa menggunakan pendekatan penelitian decision tree. Dan aplikasi yang di gunakan adalah Rapidminer sebagai alat untuk mengolah data.

Tahapan penelitian ini dapat dilihat pada gambar 1 berikut :



Gambar 1. Tahapan Penelitian

- a. Identifikasi Masalah Langkah awal ini bertujuan untuk merumuskan secara jelas isu yang ingin dipecahkan. Dalam konteks ini,

fokusnya adalah mencari tahu faktor-faktor apa saja yang berpengaruh terhadap ketepatan waktu kelulusan mahasiswa.

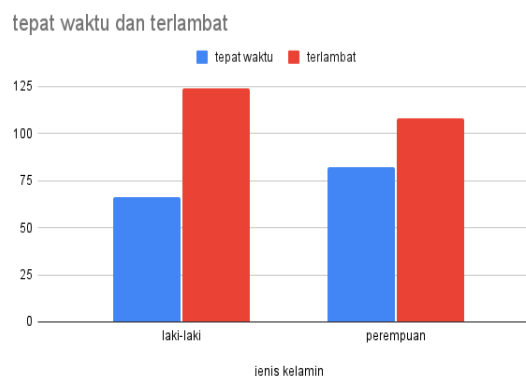
- b. Pengumpulan Data Akademik Mahasiswa Pada tahap ini, data yang relevan dikumpulkan dari sistem akademik, seperti nilai IPS tiap semester, IPK, umur, dan status mahasiswa. Data ini akan menjadi dasar dalam membangun model analisis prediktif.
- c. Pemrosesan Data (Data Preprocessing) Data mentah seringkali tidak langsung siap pakai. Oleh karena itu, dilakukan pembersihan data, penanganan nilai kosong, normalisasi, dan konversi data kategorikal agar layak untuk dianalisis oleh algoritma machine learning.
- d. Pemodel Decision Tree Pada tahap ini, algoritma Decision Tree diterapkan untuk membentuk model klasifikasi. Model ini akan menghasilkan struktur pohon keputusan yang dapat digunakan untuk memprediksi kelulusan mahasiswa berdasarkan atribut-atribut input.
- e. Evaluasi Model Setelah model dibuat, perlu dilakukan evaluasi performa menggunakan metrik seperti akurasi, precision, recall, dan confusion matrix. Tujuannya adalah untuk memastikan bahwa model memiliki tingkat keandalan yang memadai.
- f. Analisis Hasil Hasil dari model akan ditelaah lebih lanjut untuk memahami pola-pola penting yang ditemukan. Misalnya, atribut mana yang paling dominan dalam memengaruhi kelulusan mahasiswa.
- g. Simpulan dan Rekomendasi Tahap akhir berupa penarikan kesimpulan dari keseluruhan proses dan pemberian saran yang dapat digunakan oleh pihak kampus dalam merancang kebijakan pendidikan

berbasis data.

HASIL DAN PEMBAHASAN

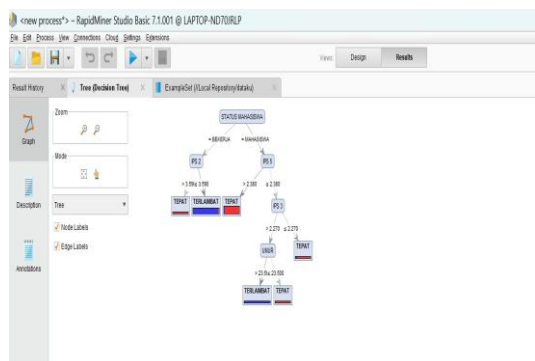
Jumlah data kelulusan yang telah di kumpulkan berjumlah 380 record data yang telah di kumpulkan oleh penulis, Dataset bersifat privat, dikumpulkan secara langsung oleh penulis. Tidak berasal dari lembaga atau website publik. untuk data mahasiswa yang lulus tepat waktu berjumlah 148 mahasiswa, sedangkan yang tidak lulus tepat waktu (terlambat) berjumlah 232 mahasiswa, berikut gambar grafik perbandingan data kelulusan mahasiswa yang tepat waktu dan terlambat. Mahasiswi perempuan memiliki jumlah kelulusan tepat waktu yang lebih tinggi dibandingkan mahasiswa laki-laki (82 vs 66), Namun, jumlah keterlambatan pada mahasiswa laki-laki juga lebih tinggi (124 vs 108).

Secara keseluruhan, proporsi kelulusan tepat waktu: laki-laki 34.7% dan perempuan 43.2%



Gambar 2. Grafik Jumlah Data Set

Dalam penelitian ini, decision tree digunakan untuk memprediksi tingkat kelulusan mahasiswa tepat waktu atau terlambat berdasarkan nilai IPS dan IPK, metode yang digunakan untuk simulasi adalah decision tree yang di proses menggunakan aplikasi rapid minner, berikut adalah hasil pohon keputusan yang menampilkan kriteria atau variabel untuk menentukan hasil :



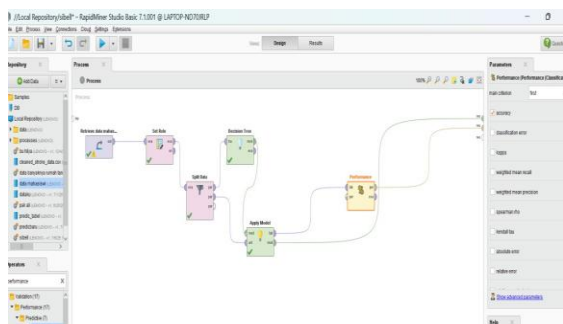
Gambar 3. Hasil Pohon Keputusan Decision tree

Hasil dari pemodelan menggunakan algoritma Decision tree di atas menampilkan proses prediksi kelulusan mahasiswa berdasarkan IPS dan IPK. Berikut adalah tabel dari pohon keputusan:

Tabel ini menunjukkan aturan sederhana yang dihasilkan oleh model Decision Tree berdasarkan nilai UAS mahasiswa. Jika nilainya kurang dari 70.500, maka prediksinya **tidak lulus**; jika nilainya 70.500 atau lebih, maka **lulus**.

Hasil

Model yang dianalisis dengan bantuan RapidMiner tidak hanya membentuk pola Decision Tree, tetapi juga memberikan informasi mengenai tingkat ketepatan prediksi berdasarkan data yang digunakan. Melalui Confusion Matrix, dapat diidentifikasi jumlah prediksi terkait kemajuan hafalan para santri. Berikut adalah visualisasinya dalam bentuk Confusion Matrix.

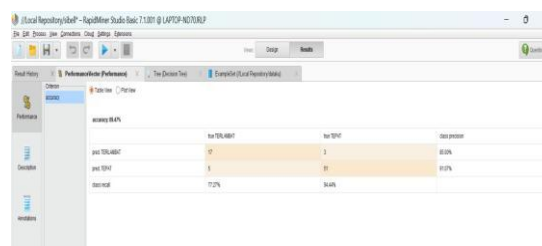


Gambar 4. View Proses Confusion Matrix

Proses klasifikasi dilakukan menggunakan RapidMiner Studio, dimulai dari pengambilan data (Read Excel), penentuan atribut target (Set Role), pembagian data menjadi data latih dan data uji (Split Data), pelatihan model dengan algoritma Decision Tree, penerapan model terhadap data uji (Apply Model), hingga evaluasi performa model (Performance). Evaluasi dilakukan menggunakan metrik-metrik klasifikasi seperti akurasi, precision, recall, dan F-measure.

Gambar 4 menampilkan hasil evaluasi model klasifikasi di RapidMiner. Model memiliki **akurasi 66,67%**, artinya 2 dari 3 prediksi benar:

- a. **Precision** tertinggi untuk kelas “no” (80%), menunjukkan prediksi “tidak” lebih akurat.
- b. **Recall** lebih tinggi untuk kelas “yes” (75%), model lebih sensitif dalam mengenali data berlabel “ya”



Gambar5. Hasil Confusion Matrix

Pembahasan

RapidMiner merupakan perangkat lunak berbasis GUI yang mendukung proses data mining secara visual dan terstruktur. Qitratul Ameliatus et al. (2024) memanfaatkan RapidMiner untuk membangun model klasifikasi kelulusan mahasiswa dan memperoleh akurasi hingga 99%. Keunggulan RapidMiner terletak pada kemudahan integrasi tahapan analisis, mulai dari impor data hingga evaluasi model.

Penelitian ini mengkaji penggunaan algoritma **Decision Tree** sebagai metode klasifikasi untuk memprediksi status keaktifan mahasiswa, khususnya dalam mengidentifikasi mahasiswa yang berpotensi menjadi **non aktif**. Data yang digunakan meliputi berbagai atribut

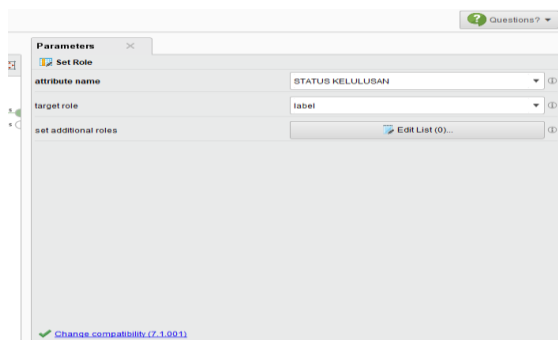
akademik dan demografis, seperti IPK, jumlah SKS yang diambil per semester, lama studi, frekuensi absen, dan data demografi lainnya. Algoritma Decision Tree dipilih karena kemampuannya dalam menghasilkan model yang mudah dipahami dan divisualisasikan melalui aturan keputusan yang jelas.

Berikut adalah langkah-langkah yang dilakukan untuk mendapatkan hasil menggunakan *decision tree* dengan menggunakan *Rapid miner* :

- Import data .CSV melalui repository, setelah data berhasil di import maka masukkan data ke dalam proses view rapid miner.
- Carilah set role di bagian operator lalu masukkan ke dalam proses view rapid miner, dan tentukan target role yang berada dalam parameters dengan memilih atribut yang akan ditetapkan menjadi label.
- Lalu carilah lagi pada bagian operators yaitu decision tree dan masukan juga ke dalam proses view setelah itu sambungkan, data ke set role lalu ke decision tree dan ke res bagian kanan atas bagian dalam view proses
- Jalankan proses



Gambar 6. View Proses



Gambar 7. Pilih Parameter

Pada tahap awal preprocessing data, dilakukan penetapan peran atribut menggunakan operator *Set Role*. Atribut ‘STATUS KELULUSAN’ ditetapkan sebagai label, yang berfungsi sebagai variabel target dalam pemodelan klasifikasi. Penetapan label ini penting agar algoritma machine learning dapat mempelajari pola dari data latih untuk kemudian digunakan dalam proses prediksi.

SIMPULAN (PENUTUP)

Berdasarkan hasil penelitian ini, metode decision tree yang diimplementasikan melalui RapidMiner terbukti efektif dalam mengklasifikasikan status mahasiswa berdasarkan data akademik. Model yang dihasilkan mampu mengidentifikasi pola-pola signifikan dari variabel input yang berkontribusi terhadap status kelulusan, melalui proses perhitungan menggunakan metode klasifikasi dan algoritma decision tree, dengan beberapa jumlah atribut, diperoleh hasil bahwa akurasi klasifikasi kelulusan mencapai 66,67%.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih sebesar-besarnya kepada pihak universitas yang telah menyediakan data yang dibutuhkan dalam penelitian ini. Ucapan terima kasih juga ditujukan kepada dosen pembimbing atas arahan dan masukan yang sangat berharga selama proses penyusunan jurnal. Tak lupa, apresiasi diberikan kepada seluruh pihak yang telah memberikan dukungan moral dan teknis yang memungkinkan penelitian ini terselesaikan dengan baik.

DAFTAR PUSTAKA

- Y. E. Yuspita, R. Okra, and M. Rezeki, “Penerapan Algoritma Klasifikasi Untuk Prediksi Tingkat Kelulusan Mahasiswa Menggunakan Rappidminer,” *Djtechno J. Teknol. Inf.*, vol. 6, no. 1, pp. 376– 388, 2025, doi: 10.46576/djtechno.v6i1.6169.

- R. V. A. Suyanto, Eduard Rusdianto, and Ernawati, "Penerapan Algoritma Decision Tree C4.5 dan Metode AdaBoost Untuk Prediksi Kelulusan Mahasiswa," *J. Inform. Atma Jogja*, vol. 5, no. 1, pp. 75–86, 2024, doi: 10.24002/jiaj.v5i1.8646.
- H. Situmorang, "Sistem Informasi Pengelolaan Data Alumni Berbasis Web (Studi Pada Fakultas Sain, Teknologi Dan Informasi) Universitas Sari Mutiara Indonesia," *J. Mahajana Inf.*, vol. 4, no. 1, pp. 28–41, 2019, doi: 10.51544/jurnalmi.v4i1.724.
- U. Al Faruq, M. Ainun Naja Fauzi, I. Fatayasya, E. Daniati, A. Ristyawan, and N. PGRI Kediri, "Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi) 131 Prediksi Data Kelulusan Mahasiswa Dengan Metode Decision Tree menggunakan Rapidminer," *Agustus*, vol. 7, pp. 2549–7952, 2023.
- M. Ridwan, H. Suyono, and M. Sarosa, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Eeccis*, vol. 7, no. 1, pp. 59–64, 2013, doi: 10.1038/hdy.2009.180.
- A. Asro, J. Chaidir, C. Cahairuddin, and J. Friadi, "Evaluasi Kinerja Algoritma Klasifikasi dalam Studi Kasus Prediksi Kelulusan di Universitas XYZ," *Zo. Tek. J. Ilm.*, vol. 19, no. 1, pp. 15–22, 2025, doi: 10.37776/zt.v19i1.1674.
- G. A. Mundok, Amiruddin, and Zulfrianto, "Klasifikasi Waktu Kelulusan Mahasiswa Menggunakan Metode Decision Tree," *J. Ilm. Ilmu Komput.*, vol. 3, no. 1, pp. 31–36, 2024.
- Al-Tameemi, G., Xue, J., Hadi Ali, I., & Ajit, S. (2024). A Hybrid Machine Learning Approach for Predicting Student Performance Using Multi-class Educational Datasets. *Procedia Computer Science*, 238, 888–895.
- Matzavela, V., & Alepis, E. (2021). Decision tree learning through a Predictive Model for Student Academic Performance in Intelligent M-Learning environments. *Computers and Education: Artificial Intelligence*, 2, 100035.
- Adejo, O. W., & Connolly, T. (2018). Predicting student academic performance using multi-model heterogeneous ensemble approach. *Journal of Applied Research in Higher Education*, 10(1), 61–75.
- Ahmed, M. (2024). Student Performance Prediction Using Machine Learning Algorithms. *Applied Computational Intelligence and Soft Computing*.
- Injadat, M. N., Moubayed, A., Bou Nassif, A., & Shami, A. (2020). Systematic Ensemble Model Selection Approach for Educational Data Mining. *arXiv*.
- Yekun, E. A., & Teklay, A. (2019). Student Performance Prediction with Optimum Multilabel Ensemble Model. *arXiv*.
- Wang, Y., Ding, A., Guan, K., Wu, S., & Du, Y. (2021). Graph-based Ensemble Machine Learning for Student Performance Prediction. *arXiv*.