

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan tinggi memiliki peran penting dalam menciptakan lulusan yang kompeten. Namun, tingkat kelulusan tepat waktu masih menjadi tantangan yang dihadapi oleh banyak institusi pendidikan. Berbagai faktor memengaruhi kelulusan mahasiswa, seperti prestasi akademik, keterlibatan dalam kegiatan kampus, dan kinerja pada mata kuliah tertentu. Oleh karena itu, prediksi kelulusan menjadi penting untuk membantu institusi mengambil langkah-langkah preventif bagi mahasiswa yang mungkin berisiko tidak lulus tepat waktu.

Penelitian ini menunjukkan bahwa meskipun telah ada berbagai studi yang menggunakan algoritma seperti C4.5 dan Naïve Bayes untuk memprediksi kelulusan, masih terdapat variasi dalam akurasi yang dicapai, yang berkisar antara 72% hingga 99,64%. Hal ini menunjukkan bahwa ada kebutuhan untuk penelitian lebih lanjut yang dapat mengidentifikasi faktor-faktor yang lebih spesifik dan relevan yang mempengaruhi kelulusan, serta untuk menguji algoritma lain yang mungkin memberikan hasil yang lebih baik. Selain itu, penggunaan dataset yang lebih besar dan beragam juga dapat membantu dalam meningkatkan akurasi prediksi dan memberikan wawasan yang lebih mendalam tentang faktor-faktor yang mempengaruhi kelulusan mahasiswa. (Wahyudi, 2023)

Metode yang digunakan dalam prediksi kelulusan mahasiswa ini memiliki beberapa tantangan, terutama dalam memilih algoritma yang sesuai. Decision Tree, meskipun sederhana, terkadang kurang akurat jika dibandingkan dengan Random Forest dalam skenario data yang kompleks. Random Forest mampu mengatasi beberapa kelemahan dari Decision Tree dengan mengurangi risiko overfitting melalui agregasi beberapa pohon keputusan yang dibangun secara acak

Mengidentifikasi bahwa prestasi akademik mahasiswa sangat bervariasi, yang dapat disebabkan oleh berbagai faktor yang tidak terdeteksi sebelumnya. Untuk mengatasi masalah ini, diusulkan metode data mining sebagai solusi untuk mengidentifikasi model yang tepat, dengan tujuan membantu dosen dalam memperbaiki setiap permasalahan yang muncul. Penelitian ini menerapkan teknik klasifikasi seperti Decision Tree C4.5 dan Random Forest untuk memprediksi prestasi akademik mahasiswa dengan mengeksplorasi berbagai parameter yang relevan. (Latifah, 2020)

Bahwa terdapat kesulitan dalam memprediksi tingkat kelulusan siswa secara akurat hanya dengan menggunakan nilai tes penerimaan. Penelitian ini menekankan bahwa nilai SMA/MA merupakan prediktor yang lebih baik untuk kelulusan tepat waktu di perguruan tinggi, karena keberhasilan akademik tidak hanya bergantung pada kemampuan kognitif, tetapi juga pada kompetensi pengaturan diri yang lebih baik. Selain itu, faktor-faktor seperti self-efficacy, budaya akademik, dan harapan masa depan juga berperan penting dalam mempengaruhi kelulusan siswa. Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan menggunakan algoritma Support Vector Machine (SVM) dan Random Forest (RF) untuk memprediksi kelulusan siswa Madrasah Aliyah Swasta (MAS), serta melakukan komparasi performa antara kedua algoritma tersebut. (Darmawan et al., 2023)

Penelitian ini diharapkan dapat memberikan wawasan lebih mendalam mengenai efektivitas kedua algoritma tersebut dalam konteks prediksi kelulusan. Dengan membandingkan Decision Tree dan Random Forest, institusi dapat menentukan metode mana yang lebih efektif digunakan untuk memprediksi kelulusan mahasiswa, serta mengidentifikasi faktor-faktor kunci yang paling berpengaruh terhadap kelulusan mahasiswa. Penelitian ini tidak hanya bermanfaat bagi pengembangan sistem akademik, tetapi juga bagi kebijakan pendidikan di masa depan yang lebih berbasis data.

sangat penting untuk kualitas pendidikan, banyak perguruan tinggi yang belum memiliki sistem yang efektif untuk memprediksi dan mencegah keterlambatan kelulusan. Penelitian ini mengidentifikasi bahwa faktor-faktor seperti lama bimbingan, penghasilan orang tua, dan status pekerjaan mahasiswa berkontribusi terhadap ketepatan waktu kelulusan, namun belum ada eksplorasi yang mendalam mengenai atribut-atribut ini dalam konteks prediksi. Selain itu, meskipun penelitian ini menggunakan beberapa algoritma data mining, masih terdapat kebutuhan untuk mengembangkan metode yang lebih komprehensif dan akurat dalam memprediksi kelulusan tepat waktu mahasiswa, sehingga dapat membantu perguruan tinggi dalam mengantisipasi dan meningkatkan tingkat kelulusan. (Satrio Junaidi et al., 2024)

Penggunaan algoritma Decision Tree dan Random Forest dalam memprediksi kelulusan mahasiswa memberikan pendekatan yang komprehensif. Decision Tree menawarkan interpretabilitas yang tinggi, sementara Random Forest memberikan akurasi dan stabilitas yang lebih baik. Penelitian ini menunjukkan bahwa pemilihan model yang tepat sangat penting dalam mengembangkan strategi intervensi yang efektif untuk meningkatkan tingkat kelulusan mahasiswa. (Budiyantara & A, 2018)

1.2 Identifikasi Masalah

1. Bagaimana tingkat akurasi algoritma Decision Tree dan Random Forest dalam memprediksi kelulusan mahasiswa berdasarkan data akademik
2. Faktor-faktor apa saja dalam data akademik mahasiswa yang paling mempengaruhi prediksi kelulusan menurut algoritma Decision Tree dan Random Forest
3. Seberapa efektif algoritma Decision Tree dan Random Forest dalam mengidentifikasi mahasiswa yang berpotensi tidak lulus tepat waktu

4. Bagaimana perbandingan performa kedua algoritma tersebut dalam hal akurasi, presisi, dan recall dalam memprediksi kelulusan mahasiswa
5. Algoritma manakah yang lebih efisien dan akurat digunakan untuk prediksi kelulusan mahasiswa berdasarkan data akademik

1.3 Tujuan Penelitian

1. Menganalisis dan membandingkan akurasi algoritma Decision Tree dan Random Forest dalam memprediksi kelulusan mahasiswa berdasarkan data akademik.
2. Mengidentifikasi faktor-faktor dalam data akademik mahasiswa yang paling berpengaruh terhadap prediksi kelulusan, seperti Nama Mahasiswa, SKS, IPK, dan NPM.
3. Menentukan algoritma yang lebih efisien dan akurat untuk digunakan dalam prediksi kelulusan mahasiswa, dengan membandingkan kinerja Decision Tree dan Random Forest berdasarkan metrik evaluasi seperti accuracy, precision, dan recall.

1.4 Manfaat penelitian

1. Bagi Institusi Pendidikan Penelitian ini dapat membantu institusi pendidikan untuk memahami lebih baik faktor-faktor akademik yang memengaruhi kelulusan mahasiswa. Dengan model prediktif yang dihasilkan, institusi dapat membuat program bimbingan yang lebih terfokus dan strategi intervensi bagi mahasiswa yang berisiko tidak lulus tepat waktu.
2. Bagi Peneliti dan Akademisi Hasil penelitian ini memberikan kontribusi tambahan pada bidang Educational Data Mining, khususnya dalam perbandingan performa algoritma decision tree dan random forest dalam konteks pendidikan.

1.5 Batasan Masalah

1. Penelitian ini berfokus pada pengembangan metode yang lebih komprehensif dan akurat dalam memprediksi kelulusan tepat waktu mahasiswa.

2. Faktor yang berkontribusi terhadap kecepatan waktu kelulusan seperti IPK dan SKS.

3. Penelitian ini membandingkan algoritma Decision Tree dan Random Forest dalam hal akurasi, presisi, dan recall.

