

Klasifikasi Kelulusan Berdasarkan Data Aktivitas Menggunakan Decision Tree C4.5

M. Alfa Najmi Fadhil¹, Dr. Kusuma Ayu Laksitowening S.T., M.T.², Gede Agung Ary Wisudiawan S.Kom, M.T.³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹alfanajmi@student.telkomuniversity.ac.id, ²ayu@telkomuniversity.ac.id,

³degunk@telkomuniversity.ac.id.

Abstrak

Penelitian ini memfokuskan pada masalah prediksi kelulusan mahasiswa berdasarkan data aktivitas dataset OULAD (Open University Learning Analytics Dataset). Tantangan utama adalah bagaimana memanfaatkan data aktivitas yang sangat besar dan beragam untuk secara efektif memprediksi kelulusan mahasiswa, sehingga institusi pendidikan dapat memberikan intervensi yang tepat waktu kepada mahasiswa yang berisiko gagal. Dengan memahami pola aktivitas yang berkaitan dengan kelulusan, pihak universitas dapat memberikan intervensi tepat waktu kepada mahasiswa yang berisiko tidak lulus. Penggunaan data tersebut untuk pengambilan keputusan dapat memberikan wawasan yang lebih mendalam mengenai kinerja mahasiswa. Tanpa pemanfaatan data ini, institusi kehilangan kesempatan untuk mendukung mahasiswa dengan cara yang lebih proaktif. Masalah ini penting diangkat karena prediksi kelulusan yang akurat dapat membantu institusi pendidikan dalam mengidentifikasi mahasiswa yang membutuhkan bantuan lebih dini, sehingga meningkatkan tingkat kelulusan secara keseluruhan. Dengan peningkatan popularitas pendidikan jarak jauh, kemampuan untuk memonitor dan mendukung mahasiswa secara efektif menjadi semakin kritis. Dengan memahami pola aktivitas yang terkait dengan kelulusan, institusi dapat merancang strategi yang lebih baik untuk mendukung kesuksesan akademis mahasiswa.

Untuk menyelesaikan masalah ini, digunakan metode Decision Tree C4.5 karena metode ini terkenal akan kemampuannya dalam menangani dataset dengan atribut yang kompleks dan beragam. Metode C4.5 dipilih karena dapat menghasilkan model yang mudah dipahami dan diinterpretasikan, sehingga memudahkan pihak universitas dalam mengambil keputusan. Algoritma ini juga memiliki kemampuan untuk menangani data yang hilang dan atribut kontinu, yang sering ditemukan dalam data aktivitas mahasiswa. Implementasi dari metode ini akan melibatkan pembersihan data, pembagian data menjadi data latih dan data uji, serta pembangunan model untuk memprediksi kelulusan. Dengan menggunakan Decision Tree C4.5, diharapkan dapat diperoleh model prediksi kelulusan yang akurat dan dapat diandalkan. Hasil dari penelitian ini adalah akurasi sebesar 56.81% dengan 15.382 mahasiswa lulus, 7.168 mahasiswa withdrawn dan 6.678 mahasiswa tidak lulus.

Kata kunci : Decision Tree C4.5, Data mining, Klasifikasi Graduasi

Abstract

This study focuses on the problem of predicting student graduation based on activity data from the OULAD (Open University Learning Analytics Dataset) dataset. The main challenge is how to utilize the vast and diverse activity data to effectively predict student graduation, so that educational institutions can provide timely interventions to students at risk of failure. By understanding activity patterns related to graduation, universities can provide timely interventions to students at risk of failure. Using this data for decision-making can provide deeper insights into student performance. Without utilizing this data, institutions miss the opportunity to support students in a more proactive way. This issue is important to raise because accurate graduation predictions can help educational institutions identify students who need assistance earlier, thereby increasing overall graduation rates. With the increasing popularity of distance education, the ability to effectively monitor and support students is becoming increasingly critical. By understanding activity patterns related to graduation, institutions can design better strategies to support students' academic success.

To solve this problem, the Decision Tree C4.5 method is used because this method is known for its ability to handle datasets with complex and diverse attributes. The C4.5 method was chosen because it can produce a model that is easy to understand and interpret, making it easier for universities to make decisions. This algorithm also has the ability to handle missing data and continuous attributes, which are often found in student activity data. The implementation of this method will involve data cleaning, data division into training data and test data, and building a model to predict graduation. By using Decision Tree C4.5, it is expected that an accurate and reliable graduation prediction model can be obtained. The results of this study are an accuracy of 56.81% with 15,382 students graduating, 7,168 students withdrawn and 6,678 students failing.

Keywords: Decision Tree C4.5, Data mining, Graduation Classification
