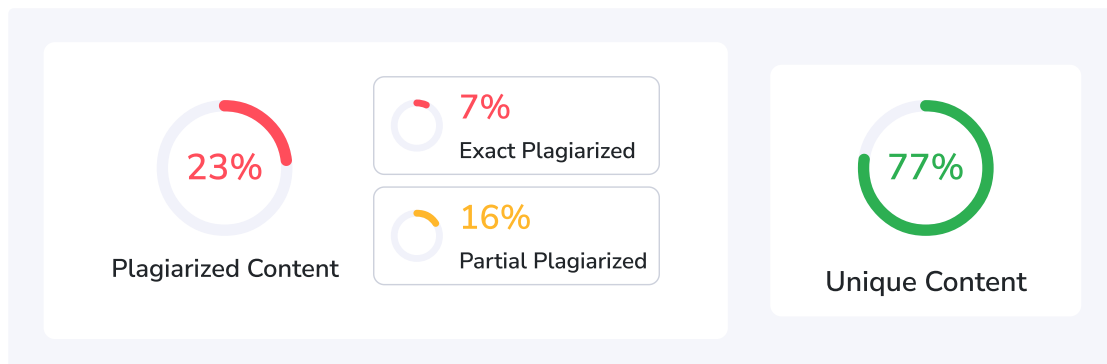


Plagiarism Scan Report By SmallSEOTools

Report Generated on: Jul 28,2025



Total Words: 999

Total Characters: 7961

Plagiarized Sentences: 10.12

Unique Sentences: 33.88 (77%)

Content Checked for Plagiarism

Prediksi Performa Akademik Mahasiswa Berdasarkan Gaya Hidup dan Dukungan Sosial Menggunakan Random Forest Royan Ristu Prayoga 1) 1) Informatika, Universitas Amikom Yogyakarta email : royanristuprayoga@students.amikom.ac.id 1) Jurnal Ilmiah DASI Vol. No. ISSN: 1411-3201

Abstraksi Performa akademik mahasiswa dipengaruhi oleh berbagai faktor, termasuk gaya hidup dan dukungan sosial yang diterima selama proses studi. Penelitian ini bertujuan untuk memprediksi performa akademik mahasiswa dengan memanfaatkan metode data mining, khususnya algoritma Random Forest, yang terbukti efektif dalam menghasilkan prediksi yang akurat. Data yang digunakan merupakan hasil survei terkait kebiasaan hidup mahasiswa dan tingkat dukungan sosial yang diperoleh. Proses penelitian meliputi tahap preprocessing data, pemberian label pada atribut non-numerik, dan validasi model menggunakan teknik cross-validation. Model dievaluasi menggunakan metrik Confusion Matrix, termasuk accuracy, precision, recall, dan F1-score. Hasil eksperimen menunjukkan bahwa algoritma Random Forest mampu mengklasifikasikan performa akademik mahasiswa dengan akurasi yang tinggi, serta memberikan informasi yang berguna bagi pihak kampus dalam pengambilan keputusan strategis untuk meningkatkan kualitas pembelajaran dan dukungan terhadap mahasiswa. Kata Kunci : performa akademik, gaya hidup, dukungan sosial, data mining, Random Forest

Abstract Student academic performance is influenced by various factors, including lifestyle and the social support received during their studies. This research aims to predict student academic performance using data mining methods, specifically the Random Forest algorithm, which has proven effective in generating accurate predictions. The data used is derived from surveys regarding student lifestyle habits and the level of social support obtained. The research process includes data preprocessing, labeling non-numeric attributes, and model validation using cross-validation techniques. The model is evaluated using Confusion Matrix metrics, including accuracy, precision, recall, and F1-score. Experimental results indicate that the Random Forest algorithm can classify student academic performance with high accuracy, providing valuable information for the campus in making strategic decisions to improve the quality of learning and support for students.

Keywords : Academic performance, Lifestyle, Social support, Data mining, Random Forest

Jurnal Ilmiah DASI Vol. No. ISSN: 1411-3201

1. Pendahuluan Tingkat kelulusan atau graduation rate merupakan salah satu indikator utama dalam mengukur kinerja sebuah lembaga akademik [1]. Di berbagai negara, termasuk Indonesia, angka putus kuliah masih menjadi tantangan serius yang berdampak pada masa depan generasi muda. Studi menunjukkan bahwa lebih dari 15% mahasiswa di negara-negara maju seperti Amerika Serikat, Jerman, dan Inggris mengalami dropout selama masa studi mereka [1]. Di Indonesia sendiri, tingkat mahasiswa yang melanjutkan ke jenjang perguruan tinggi masih rendah, yaitu hanya sekitar 6% dari total lulusan SMA, dengan tingkat putus kuliah yang cukup tinggi di berbagai provinsi, seperti Bengkulu dan Jawa Timur [3]. Penelitian sebelumnya telah menunjukkan bahwa rendahnya performa akademik merupakan salah satu indikator utama dari

tingginya angka putus kuliah [1][2]. Oleh karena itu, penting bagi institusi pendidikan untuk memahami faktor-faktor yang memengaruhi keberhasilan akademik mahasiswa guna merumuskan strategi pencegahan yang efektif. Faktor-faktor tersebut mencakup kondisi sosial ekonomi, gaya hidup, dukungan sosial, hingga variabel personal seperti stres, kebiasaan belajar, dan kesehatan [1][3]. Seiring berkembangnya teknologi, pendekatan berbasis Educational Data Mining (EDM) dan algoritma machine learning seperti Decision Tree, Naïve Bayes, Random Forest, dan K-Nearest Neighbor telah banyak digunakan untuk memprediksi performa akademik mahasiswa [2][3]. Di antara algoritma tersebut, Random Forest memiliki keunggulan dalam hal akurasi dan stabilitas prediksi pada data pendidikan yang kompleks [3]. Berbagai penelitian menunjukkan bahwa Random Forest mampu menghasilkan akurasi hingga 91% dalam memprediksi keberhasilan akademik [3], menjadikannya salah satu algoritma yang efektif untuk digunakan dalam sistem pendukung keputusan pendidikan. Namun demikian, sebagian besar penelitian sebelumnya lebih banyak menitikberatkan pada variabel akademik seperti IPK dan absensi, tanpa memasukkan faktor gaya hidup dan dukungan sosial secara menyeluruh sebagai variabel penting dalam model prediksi [1][3]. Padahal, aspek-aspek tersebut memiliki peran signifikan dalam menentukan keberhasilan studi mahasiswa. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan model prediksi performa akademik mahasiswa dengan mempertimbangkan faktor gaya hidup dan dukungan sosial menggunakan algoritma Random Forest. Dengan pendekatan ini, diharapkan institusi pendidikan dapat memperoleh wawasan yang lebih luas dalam memahami karakteristik mahasiswa dan menyusun kebijakan yang lebih tepat sasaran untuk meningkatkan kualitas pembelajaran dan menekan angka putus kuliah.

2. Metode Penelitian

2.1 Jenis Penelitian

Ini adalah penelitian kuantitatif prediktif yang menggunakan data mining dan machine learning untuk memprediksi performa akademik mahasiswa. Tujuan utamanya adalah membangun model yang bisa memperkirakan performa ini berdasarkan data gaya hidup dan dukungan sosial, dengan memanfaatkan algoritma Random Forest Regressor. Penelitian ini menggunakan pendekatan kuantitatif karena semua tahapannya melibatkan pengumpulan, pemrosesan, dan analisis data numerik (kuantitatif) untuk menjawab pertanyaan penelitian secara statistik dan sistematis. Sifat prediktifnya cocok karena fokus utama penelitian ini adalah mengembangkan model untuk memprediksi performa akademik dari variabel-variabel yang sudah dikumpulkan [2].

2.2 Alur Penelitian

Penelitian ini diawali dengan pengumpulan data sekunder dari dataset `enhanced_student_habits_performance_dataset.csv`, yang memuat informasi gaya hidup, kebiasaan belajar, dan faktor sosial mahasiswa. Data kemudian diproses melalui penghapusan atribut tidak relevan, penanganan nilai kosong dan duplikat, serta transformasi data kategorikal menggunakan Label Encoding. Fitur numerik distandarisasi untuk meningkatkan kinerja model. Setelah dilakukan pemisahan fitur (X) dan target (y), data dibagi menjadi data latih dan data uji dengan rasio 80:20. Model yang digunakan adalah Random Forest Regressor dengan parameter utama: 200 pohon keputusan (`n_estimators`) dan kedalaman maksimum 10. Model dilatih pada data latih dan dievaluasi menggunakan metrik MAE, MSE, RMSE, dan R^2 [2].

2.3 Pengambilan Dataset

Data yang dimanfaatkan dalam penelitian ini berasal dari Kaggle, sebuah platform sumber data publik, di bawah judul "Enhanced Student Habits Performance Dataset." Dataset ini dirancang khusus untuk memfasilitasi analisis dan prediksi performa akademik mahasiswa, dengan mempertimbangkan beragam faktor seperti kebiasaan belajar, gaya hidup, dan dukungan sosial. Dataset tersebut dapat diakses secara bebas melalui tautan: <https://www.kaggle.com/datasets/aljarah/xAPI-Edu-Data>. Fitur-fitur tersebut pada dataset ini meliputi IPK sebelumnya (`previous GPA`), persentase kehadiran, rata-rata jam belajar per hari, tingkat stres, waktu penggunaan perangkat elektronik (`screen time`), dan skor kemampuan manajemen waktu (`time management score`). Variabel-variabel ini berperan sebagai prediktor (variabel independen) untuk memodelkan nilai ujian (`exam_score`), yang merupakan representasi dari performa akademik mahasiswa sebagai variabel target. Setiap entri atau baris dalam dataset ini merepresentasikan satu individu mahasiswa, lengkap dengan kombinasi nilai dari seluruh fitur yang disebutkan. Dataset telah melalui serangkaian pra-proses, termasuk penghapusan data

Plagiarized Sources

<https://journals.ums.ac.id/indigenous/article/view/1610>

Model dievaluasi menggunakan metrik Confusion Matrix, termasuk accuracy, precision, recall, dan F1-score. [↗](#)

<https://jurnal.unt.ac.id/index.php/it/article/download/9099/4575>

The model is evaluated using Confusion Matrix metrics, including accuracy, precision, recall, and F1-score. [↗](#)

<https://encord.com/blog/classification-metrics-accuracy-precision-recall>

Keywords : Academic performance, Lifestyle, Social support, Data mining, Random Forest Jurnal Ilmiah DASI Vol. [↗](#)

<http://jidsaa.com/index.php/welcome/article/view/229>

Berbagai penelitian menunjukkan bahwa Random Forest mampu menghasilkan akurasi hingga 91% dalam memprediksi keberhasilan akademik [3], menjadikannya salah satu algoritma yang efektif untuk digunakan dalam sistem pendukung keputusan pendidikan. [↗](#)

<https://rcf-indonesia.org/home/index.php/jsit/article/view/464>

Metode Penelitian 2.1 Jenis Penelitian Ini adalah penelitian kuantitatif prediktif yang menggunakan data mining dan machine learning untuk memprediksi performa akademik mahasiswa. [↗](#)

<https://rcf-indonesia.org/home/index.php/jsit/article/view/464>

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan model prediksi performa akademik mahasiswa dengan mempertimbangkan faktor gaya hidup dan dukungan sosial menggunakan algoritma Random Forest. [↗](#)

<https://journal.umkendari.ac.id/decode/article/view/1103>

Setelah dilakukan pemisahan fitur (X) dan target (y), data dibagi menjadi data latih dan data uji dengan rasio 80:20. [↗](#)

<http://etheses.uin-malang.ac.id/66007/2/200605110154.pdf>

Dataset tersebut dapat diakses secara bebas melalui tautan:

<https://www.kaggle.com/datasets/aljarah/xAPI-Edu-Data>. [↗](#)

<https://www.kaggle.com/datasets/aljarah/xAPI-Edu-Data>

Dataset telah melalui serangkaian praproses, termasuk penghapusan data [↗](#)

<https://ejurnal.umbima.ac.id/index.php/scientific/article/download/399/212>