

ABSTRACT

Theodorus Pratama (01679230005)

A PREDICTION APPROACH FOR UNIVERSITY STUDENTS' COMPETENCY LEVELS USING RANDOM FOREST AND PYTHAGOREAN TREE

(xii + 49 pages; 8 figures; 9 tables; 2 attachments)

This study aims to predict university students' competency levels based on their final project scores and supporting course grades using a machine learning-based approach. The method applied involves the Random Forest algorithm to build a predictive model, along with Pythagorean Tree visualization techniques to enhance the interpretability of the prediction results. The dataset used consists of final project scores and supporting course grades of Universitas Pelita Harapan students, processed through preprocessing steps such as data cleaning, normalization, and feature transformation.

The Random Forest model was chosen for its ability to handle complex and multivariable data, producing accurate predictions by leveraging the aggregation of multiple decision trees. To improve the understanding of the analysis results, Pythagorean Tree visualization was implemented, which illustrates the contribution of each feature in the model's decision-making process.

The research findings indicate that the Random Forest algorithm successfully produced predictions with reasonably good accuracy, as measured by an R-squared (R^2) value of 0.346. Additionally, the model demonstrated a Mean Square Error (MSE) of 626,355.437, a Root Mean Square Error (RMSE) of 791.426, a Mean Absolute Error (MAE) of 569.958, and a Mean Absolute Percentage Error (MAPE) of 24.5%. Moreover, the Pythagorean Tree visualization provided an intuitive representation of how attributes such as average grades and categories of final projects and supporting courses influence students' competency levels. This visualization enables academic stakeholders to identify significant patterns within the data, supporting more strategic decision-making in designing educational policies.

Keywords : Student Competence, Random Forest, Pythagorean Tree, Machine Learning, Academic Prediction, Data Visualization, Educational Analysis.

References : 40 (1997 – 2024)

ABSTRAK

Theodorus Pratama (01679230005)

PENDEKATAN PREDIKSI BARU UNTUK TINGKAT KOMPETENSI MAHASISWA UNIVERSITAS MENGGUNAKAN RANDOM FOREST DAN PYTHAGOREAN TREE

(xii + 49 halaman; 8 gambar; 9 tabel; 2 lampiran)

Penelitian ini bertujuan untuk memprediksi tingkat kompetensi mahasiswa universitas berdasarkan nilai tugas akhir dan mata kuliah pendukung mereka dengan menggunakan pendekatan berbasis *machine learning*. Metode yang diterapkan adalah algoritma *Random Forest* untuk membangun model prediksi, serta visualisasi dengan teknik *Pythagorean tree* guna meningkatkan interpretabilitas hasil prediksi. Dataset yang digunakan mencakup nilai tugas akhir dan mata kuliah pendukung mahasiswa Universitas Pelita Harapan, yang diolah melalui tahapan pra-pemrosesan seperti pembersihan data, normalisasi, dan transformasi fitur.

Model *Random forest* digunakan karena kemampuannya dalam menangani data yang kompleks dan multivariabel, menghasilkan prediksi yang akurat dengan memanfaatkan penggabungan banyak pohon keputusan. Untuk meningkatkan pemahaman terhadap hasil analisis, visualisasi *Pythagorean tree* diterapkan, yang memperlihatkan kontribusi setiap fitur dalam pengambilan keputusan model.

Hasil penelitian menunjukkan bahwa algoritma *Random forest* berhasil menghasilkan prediksi dengan akurasi yang cukup baik, yang diukur melalui nilai R-squared (R^2) sebesar 0.346. Selain itu, model ini memiliki *Mean Square Error* (MSE) sebesar 626,355.437, *Root Mean Square Error* (RMSE) sebesar 791.426, *Mean Absolute Error* (MAE) sebesar 569.958, dan *Mean Absolute Percentage Error* (MAPE) sebesar 24.5%. Selain itu, visualisasi *Pythagorean tree* memberikan gambaran intuitif mengenai bagaimana atribut-atribut seperti nilai rata-rata dan kategori tugas akhir dan mata kuliah pendukung berpengaruh terhadap tingkat kompetensi mahasiswa. Visualisasi ini memungkinkan pihak akademik untuk mengenali pola signifikan dalam data, mendukung pengambilan keputusan yang lebih strategis dalam merancang kebijakan pendidikan.

Kata Kunci : Kompetensi Mahasiswa, *Random forest*, *Pythagorean Tree*, Pembelajaran Mesin, Prediksi Akademik, Visualisasi Data, Analisis Pendidikan.

Referensi : 40 (1997 – 2024)