ABSTRAK

Motivasi belajar dan hasil prestasi belajar siswa merupakan faktor penting yang memengaruhi keberhasilan pendidikan, khususnya pada tingkat Sekolah Menengah Pertama (SMP). Penelitian ini membandingkan performa dua algoritma machine learning, yaitu Random Forest dan Logistic Regression, dalam memprediksi tingkat motivasi dan hasil belajar siswa berdasarkan data kebiasaan belajar, kondisi mental, kehadiran, jam tidur, dukungan keluarga, dan nilai akademik. Proses studi meliputi pra-pemrosesan data , normalisasi , pemisahan data menjadi data pelatihan dan pengujian , pelatihan model , dan evaluasi menggunakan akurasi , sensitivitas , spesifisitas , dan AUC. Berdasarkan temuan studi, Random Forest berkinerja lebih baik dengan akurasi 0,91, sensitivitas 0,91, spesifisitas 0,94, dan AUC 0,94. Sementara itu, Logistic Regression memperoleh akurasi 0.84, sensitivitas 0.84, spesifisitas 0.90, dan AUC 0.95. Temuan ini menegaskan bahwa Random Forest lebih unggul dalam kemampuan prediksi secara keseluruhan, sedangkan Logistic Regression tetap relevan karena interpretabilitasnya. Studi ini bertujuan untuk membantu pengembangan sistem pendukung keputusan berbasis data dalam pendidikan untuk membantu sekolah dalam mengidentifikasi siswa yang memerlukan intervensi dini.

Kata Kunci: Random Forest, Logistic Regression, Motivasi Belajar, Hasil Belajar, Prediksi

ABSTRACT

Student learning motivation and learning outcomes are important factors that influence educational success, especially at the junior high school level. This study compares the performance of two machine learning algorithms, namely Random Forest and Logistic Regression, in predicting student motivation and learning outcomes based on data on learning habits, mental condition, attendance, sleep hours, family support, and academic grades. The study process included data pre-processing, normalization, separation of data into training and testing data, model training, and evaluation using accuracy, sensitivity, specificity, and AUC. Based on the study findings, Random Forest performed better with an accuracy of 0.91, sensitivity of 0.91, specificity of 0.94, and AUC of 0.94. Meanwhile, Logistic Regression obtained an accuracy of 0.84, sensitivity of 0.84, specificity of 0.90, and AUC of 0.95. These findings confirm that Random Forest is superior in its overall predictive ability, while Logistic Regression remains relevant due to its interpretability. This study aims to assist in the development of data-driven decision support systems in education to help schools identify students who require early intervention.

Keywords*: Random Forest, Logistic Regression, Learning Motivation, Academic Performance, Prediction*