

ABSTRAK

Penerapan *Learning Analytics* dan *Educational Data Mining* dalam lingkungan *Electronic Learning System* memungkinkan pendidik untuk memahami karakteristik mahasiswa guna mendukung personalisasi pembelajaran. Penelitian ini bertujuan untuk memprediksi gaya belajar mahasiswa berdasarkan *Felder-Silverman Learning Style Model*, khususnya pada dimensi *processing* dan *input*, dengan memanfaatkan data log aktivitas mahasiswa di Universitas Amikom Purwokerto. Berbeda dengan penelitian sebelumnya yang umumnya menggunakan klasifikasi label tunggal, penelitian ini menerapkan pendekatan *multi-label classification* untuk menangkap kombinasi gaya belajar secara simultan.

Tahapan pengembangan model diawali dengan penanganan *missing value* menggunakan metode imputasi median. Selanjutnya, tantangan *class imbalance* yang ekstrem pada *dataset* diatasi melalui pendekatan *Random Oversampling*. Untuk pemodelan klasifikasi *multi-label*, algoritma *Random Forest* diterapkan dan dikomparasikan kinerjanya dengan empat pendekatan lain, yaitu *XGBoost*, *Self-Training*, *Support Vector Machine*, dan *RBF Network*. Validasi model kemudian dilakukan menggunakan teknik *Nested Cross-Validation* guna menjamin objektivitas hasil.

Hasil eksperimen menunjukkan bahwa algoritma *Random Forest* merupakan model terbaik dengan performa paling konsisten. Model ini mencapai skor *F1-Macro* sebesar (0,7063), mengungguli *XGBoost* (0,7052) dan *Self-Training* (0,6922). Selain itu, *Random Forest* menunjukkan stabilitas tertinggi dengan standar deviasi terendah ($\pm 0,0794$) setelah melalui proses *hyperparameter tuning*. Penelitian ini menyimpulkan bahwa data perilaku dari log aktivitas ELS dapat diandalkan untuk memprediksi gaya belajar mahasiswa secara akurat, yang bermanfaat sebagai dasar rekomendasi strategi pembelajaran adaptif.

Kata kunci: *FSLSM*, *Multi-Label Classification*, *Random Forest*, *Random Oversampling*, *Electronic Learning System*.

ABSTRACT

The implementation of Learning Analytics and Educational Data Mining within an Electronic Learning System (ELS) environment enables educators to understand student characteristics to support personalized learning. This study aims to predict student learning styles based on the Felder-Silverman Learning Style Model, specifically focusing on the processing and input dimensions, by utilizing student activity log data at Universitas Amikom Purwokerto. Unlike previous studies that predominantly use single-label classification, this study applies a multi-label classification approach to capture learning style combinations simultaneously.

The model development phase begins with handling missing values using the median imputation method. Subsequently, the challenge of extreme class imbalance in the dataset is addressed through a Random Oversampling approach. For multi-label classification modeling, the Random Forest algorithm is applied and its performance is compared against four other approaches: XGBoost, Self-Training, Support Vector Machine, and RBF Network. Model validation is then conducted using the Nested Cross-Validation technique to ensure objective and unbiased results.

Experimental results indicate that the Random Forest algorithm is the best model with the most consistent performance. This model achieved an F1-Macro score of 0.7063, outperforming XGBoost (0.7052) and Self-Training (0.6922). Furthermore, Random Forest demonstrated the highest stability with the lowest standard deviation (± 0.0794) after undergoing hyperparameter tuning. This study concludes that behavioral data from ELS activity logs is reliable for accurately predicting student learning styles, serving as a beneficial basis for adaptive learning strategy recommendations.

Keywords: FSLSM, Multi-Label Classification, Random Forest, Random Oversampling, Electronic Learning System.