



[Digital Transformation, Sustainable Innovations and Development in the African Continent]

[Thursday 31st July and Friday 1st August 2025]

[KWUST 2025 Conference Proceedings]

ABSTRACT NO. 47

A novel AI-driven model for student dropout risk analysis with explainable AI insights

Dr. Charles K. Njogu, Mr. Eugene Okoth & Mr. Job Maringa

Kiriri Women's University of Science and Technology, Kenya

ABSTRACT

The increasing number of students dropping out of school due to social, economic, personal (e.g., depression or persistent failure), and health issues is a growing concern for governments, educators, and guardians. Identifying and analyzing the factors contributing to student dropout is crucial. Various machine learning, analytical, and statistical models have been proposed to address this issue. However, the existing models have several limitations in providing a precise and automated system for predicting dropout risk and analyzing the factors behind this. Besides, generating a balanced dataset is also a limitation as 'Dropouts' are less than the 'Non-dropouts'. Moreover, selecting significant features contributing to student dropout and non-dropout is also very important in developing a model. However, this study introduces a comprehensive machine learning (ML) and explainable AI (XAI) based methodology to address these limitations. Firstly, the imbalanced dataset problem was handled using the Upsampling technique by adjusting the minority class 'Dropout'. Then, the feature selection method Recursive Feature Elimination (RFE) is used with Cross-Validation (CV) as the RFE-CV method to select the most significant features. After preprocessing, this study proposed a hybrid model named the Hybrid Logistic Regression and Neural Network (HLRNN) model, which predicts student dropout with 96% accuracy, outperforming other experimented models as well as the parent models Logistic Regression and Artificial Neural Network with 2% and 3% accuracy. Finally, the XAI model The SHapley Additive exPlanations (SHAP), and Local Interpretable Model-agnostic Explanations (LIME) are deployed to analyze the risk factors associated with student dropout. This approach aims to assist institutions and educational stakeholders in formulating policies for student retention, enabling early intervention to reduce dropout rates.

Key Words:

AI-driven model for student dropout risk analysis

