ADAPTIVE STUDENT PROFILING USING AI AND ML: STRATEGIES FOR UNIVERSITY ADMISSIONS AND RETENTION

Ihor Bovkun, Oksana Ivashchenko, Ihor Hamaiun
National Technical University «Kharkiv Polytechnic Institute», Kharkiv

The growing complexity of student recruitment and retention in higher education necessitates innovative, data-driven approaches to segmentation and profiling. This study presents a comprehensive review of artificial intelligence (AI) and machine learning (ML) applications in student segmentation, drawing from successful methodologies used across industries such as marketing, finance, and telecommunications. The main objective is to adapt these models to the specific needs of higher education institutions to enhance admissions, student support, and academic retention strategies.

A systematic literature review (SLR) was conducted using databases including Scopus, Web of Science, and Google Scholar. Over 700 publications were initially identified, with a final sample of 40 selected through a Weighted Sum Model (WSM) evaluation based on relevance, methodological rigor, novelty, impact, and interdisciplinary adaptability. Key findings indicate that clustering techniques such as K-Means, DBSCAN, and hierarchical clustering are the most commonly used ML methods for student segmentation. Predictive analytics models–particularly Random Forest, Gradient Boosting, and Support Vector Machines–are widely used to forecast student success, dropout risk, and future employability. Hybrid models, such as K-Means combined with Decision Trees or PCA, were found to improve accuracy and interpretability.

The study also highlights the importance of integrating unstructured behavioural data, such as social media activity, online learning engagement, and chatbot interactions. However, the use of natural language processing (NLP) techniques like Word2Vec and sentiment analysis remains underutilized. Ethical considerations concerning student data usage and compliance with GDPR and FERPA are emphasized as critical to the responsible implementation of AI-driven systems.

This review proposes a three-stage roadmap for future research: (1) data collection and feature engineering from structured and unstructured sources; (2) real-time behavioural tracking and NLP integration; and (3) selection and evaluation of AI/ML models using precision metrics such as F1 score and RMSE. These steps aim to support the development of adaptive, personalized, and ethically compliant student profiling systems.

References:
1. Das S., Nayak J. Customer Segmentation via Data Mining Techniques: State-of-the-Art Review // *In: Computational Intelligence in Data Mining. Smart Innovation, Systems and Technologies, vol 281.* Springer, Singapore, 2022. DOI: https://doi.org/10.1007/978-981-16-9447-9_38.
2. Yunus M., Mutalib S., Abdul Hamid N., Rahman S., Malik A. Predictive analytics of university student intake using supervised methods // *IAES International Journal of Artificial Intelligence.* – 2019. – Vol. 8(4). – P. 367–374. DOI: 10.11591/ijai.v8.i4.pp367-374..