

Chapter 1

Machine Learning for Predictive Learning Analytics

Agnieszka Szmelter-Jarosz

Faculty of Economics, University of Gdańsk, Poland

ABSTRACT

This research examines the role of machine learning in predictive learning analytics and its applications in educational settings. The main focus has been on how to prepare data, select and train machine learning models, evaluate and validate them, and finally, effectively integrate these models into educational systems. The findings show that combining data-driven analytics with machine learning algorithms can help improve the quality of learning, personalize education, and enhance educational interventions. In addition to technical requirements, this approach requires attention to ethical, social, and educational policy considerations in order to serve as an effective tool in education.

1. INTRODUCTION

In recent decades, the digital transformation in education has led to the accumulation of a vast amount of data related to teaching and learning processes. Data that previously remained in the form of traditional teacher-student interactions or limited educational activities is now recorded and stored in a structured manner with the spread of virtual learning environments, learning management systems (LMS), massive online courses (MOOCs), and personalized learning platforms (Sghir et al., 2023). This change has not only provided new opportunities for more accurate monitoring and analysis of the learning process, but also made it possible

DOI: 10.4018/979-8-3373-5117-9.ch001

to predict learner behaviors and educational outcomes. In this regard, a new field called Learning Analytics has emerged that analyzes learning data with the aim of extracting patterns, better understanding learning processes, and improving educational decision-making (Nozari, Szmelter-Jarosz, & Ghahremani-Nahr, 2022).

Among the various branches of learning analytics, predictive learning analytics has gained a special place due to its ability to predict academic performance, identify educational risks, and optimize interventions. Predictive analytics, unlike descriptive and diagnostic analytics, goes beyond observing the past and attempts to predict future events based on existing data (Mubarak et al., 2021). Applying this type of analytics in education can play an effective role in preventing academic failure, dropout, or reduced learner engagement, and at the same time, it enables the design of personalized learning paths (Guryanova et al., 2020).

Meanwhile, machine learning, as one of the most important sub-branches of artificial intelligence, plays a key role in implementing predictive analytics. Machine learning's ability to discover hidden patterns in data, identify complex relationships between learning variables, and generate models that can generalize to new data has made it a powerful tool for learning analytics (Nozari, Ghahremani-Nahr, & Szmelter-Jarosz, 2024). Machine learning algorithms, whether classifiers to predict learning outcomes or regressions to predict continuous values, have been widely used in various fields of education (Samanpour et al., 2017).

One of the most important advantages of using machine learning models in learning analytics is their ability to be applied in adaptive learning environments and personalized learning systems. Such models can continuously analyze data on learner interactions, performance changes, or learning behaviors to adjust the learning path in real time to suit the individual needs of each user (Nozari, Tavakkoli-Moghaddam, & Dolgui, 2024). This feature is especially important in open learning environments, such as massively online courses or self-paced e-learning, where teachers or course designers cannot directly interact with each learner (Alam & Mohanty, 2022).

However, implementing machine learning-based predictive learning analytics is not without its challenges. One of the main obstacles is the quality and structure of educational data. Data obtained from educational systems often have characteristics such as heterogeneity, class imbalance, incomplete data, or noise, which can affect the accuracy and reliability of predictive models (Nozari, Ghahremani-Nahr, & Szmelter-Jarosz, 2023). On the other hand, choosing the right algorithm, tuning model parameters, and interpreting the outputs also require technical knowledge and a deep understanding of the nature of educational data (Hlosta et al., 2022).

Another important challenge is the ethical and social considerations associated with educational data analytics. Learning analytics, especially in its predictive form, often deals with sensitive student and user data, which can raise serious concerns about privacy, information security, and even the social consequences of inaccu-

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/machine-learning-for-predictive-learning-analytics/393416

Related Content

An Intelligent Approach for Virtual Chemistry Laboratory

Shikha Mehta, Monika Bajaj and Hema Banati (2016). *Handbook of Research on Advanced Hybrid Intelligent Techniques and Applications* (pp. 483-517).

www.irma-international.org/chapter/an-intelligent-approach-for-virtual-chemistry-laboratory/140465

Exploring the Advances of Artificial Intelligence in Cryptocurrencies and Darknet Forensics

Ramy El-Kady (2027). *Encyclopedia of Modern Artificial Intelligence* (pp. 1-35).

www.irma-international.org/chapter/exploring-the-advances-of-artificial-intelligence-in-cryptocurrencies-and-darknet-forensics/406039

Machine Learning-Based Demand Forecasting in Supply Chains

Real Carbonneau, Rustam Vahidov and Kevin Laframboise (2007). *International Journal of Intelligent Information Technologies* (pp. 40-57).

www.irma-international.org/article/machine-learning-based-demand-forecasting/2426

Taxonomy on Ambient Computing: A Research Methodology Perspective

Diganta Sengupta (2020). *International Journal of Ambient Computing and Intelligence* (pp. 1-33).

www.irma-international.org/article/taxonomy-on-ambient-computing/243445

Factors Influencing Patient Adoption of the IoT for E-Health Management Systems (e-HMS) Using the UTAUT Model: A High Order SEM-ANN Approach

Manish Dadhich, Kamal Kant Hiran, Shalendra Singh Rao and Renu Sharma (2022). *International Journal of Ambient Computing and Intelligence* (pp. 1-18).

www.irma-international.org/article/factors-influencing-patient-adoption-of-the-iot-for-e-health-management-systems-e-hms-using-the-utaut-model/300798