EasyChair Preprint

№ 10401

A Review of Learning Types Research Trends on the Use of Machine Learning in Education

Md Ferdous Hossain Parvaj, Amandeep Kaur Kaur, Aditya Pratap, Md Abdullah Al Mamun and Md Tanjim Alam Abir

June 15, 2023

A Review of Learning Types research trends on the use of machine learning in education

Md Ferdous Hossain Parvaj[1], Amandeep Kaur[2*], Aditya Pratap[3], Md Abdullah Al Mamun[4], Md Tanjim Alam Abir[5]

2 Assistant Professor, School of Computer Application, Lovely Professional University, Phagwara, Punjab, India*

1,3,4,5 Students, School of Computer Application, Lovely Professional University, Phagwara, Punjab, India

1 ferdous.rajshahi@gmail.com,
2* amandeepkaur.kaur06@gmail.com,
3 adityapratap6207@gmail.com,
4 abdmamunfizz@gmail.com,
5 tanjim.abir.any@gmail.com

Abstract. This review paper presents an overview of the use of machine learning (ML) in education, particularly in analyzing student performance data through techniques such as eye tracking, voice tracking, and window switching during online exams. With the goal of providing personalized learning experiences that address the individual needs of students, the utilization of ML in education has gained significant attention in recent years. Eye tracking enables the collection of valuable data on student behavior during online exams, which can then be analyzed by ML algorithms to identify patterns and relationships between behavior and exam performance. This can lead to the development of interventions that improve student learning outcomes. Voice tracking is another technique that provides insight into student behavior during online exams. By analyzing the data collected through voice tracking, ML algorithms can identify keywords or phrases associated with incorrect answers, leading to the development of interventions that improve student performance. Finally, window switching can provide valuable information on student engagement with exam content, which can be analyzed by ML algorithms to identify patterns and relationships between window switching behavior and exam performance, leading to interventions that improve learning outcomes.

Keywords

Machine Learning, Data mining, Education, Student Performance, Eye Tracking, Window Switch, Online Exam.

INTRODUCTION

Using machine learning there has been an increase in interest in using machine learning in education over the past few years. Regarding the application of machine learning in education, researchers have looked into a variety of topics. Studies have examined the use of data mining algorithms to examine medical science therapies, including the traditional Chinese medical science therapy for vertigo. For instance, [1] conducted research on the use of the Apriori algorithm for data mining in the treatment of vertigo in traditional Chinese medicine. Other studies have concentrated on leveraging educational data mining and learning analytics to enhance student performance. [3] carried out a study on educational data mining with an emphasis on the application of data mining methods for studying educational data. Studies have looked into using mobile learning systems and other personalised learning tools. For instance, [13] examined the efficiency of a mobile, personalised arithmetic learning system. The potential of augmented reality for learning has been investigated in other studies. A thorough review of studies on the application of augmented reality as a learning tool was carried out by [7]. [8] predicted student academic achievement in a physics course using augmented reality learning analytics. Some research has centred on utilising machine learning to spot cheating in online courses. [4] employed data mining methods to find cheating activity in MOOCs. [6] looked into online exam fraud in China from a cultural angle. The application of machine learning for process mining in online education and learning is also gaining popularity. [12] a thorough literature assessment on the application of process mining in

online education and learning. However, there are still some topics that require more study, such as how education agents can help students succeed [11] and how machine learning affects both students and teachers psychologically. [5] offered a guide on learning analytics for governments that covers the privacy and ethical concerns associated with its use. [10] examined online education platforms using user behaviour analysis based on eye tracking.Deep belief networks were utilised by [15] to analyse data related to mechanical defect diagnostics.In conclusion, studies examining various learning styles and their possible applications in the area have grown increasingly common as machine learning has been integrated into education. Even though there has been a lot of advancement in this field, more study is still required to fully grasp the potential of machine learning in education as well as the moral and psychological ramifications of its application.

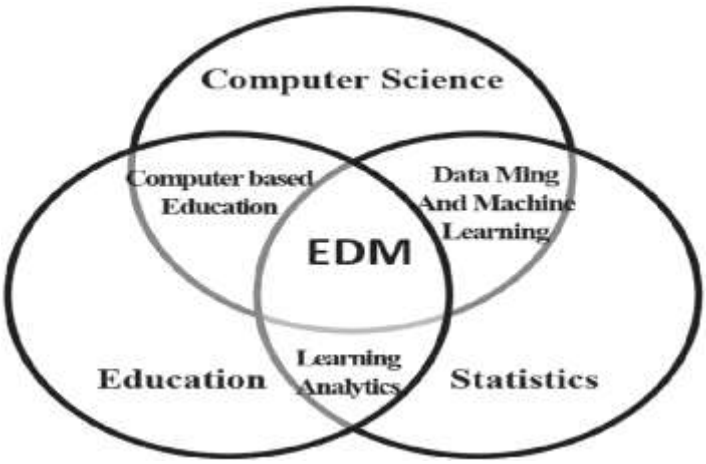

Figure 1: Interdisciplinary of EDM

The diagram in Figure 1 depicts the overlap between four key areas of education: computer-based education, data mining machine learning education, and learning analytics. The central point of intersection in the figure is computer-based education, which involves using technology to facilitate teaching and learning. Data mining and machine learning education are techniques utilized to extract insights from educational data, including information on student behavior and performance. Meanwhile, learning analytics is a field that leverages data to enhance teaching and learning outcomes.

The figure emphasizes that the convergence of these four areas can assist educators and researchers in comprehending how technology can aid teaching and learning, uncover patterns in student behavior and performance, and design tailored learning experiences. Through the integration of computer-based education, data mining, machine learning education, and learning analytics, experts in education can gain a better understanding of student behavior and performance, pinpoint areas for improvement, and create effective interventions to support student learning.

In summary, the figure underscores the importance of integrating these four areas to improve education outcomes and advance our knowledge of how technology can be used to enhance education.

LITRETURE REVIEW

Learning analytics (LA) and educational data mining (EDM) have emerged as promising topics that use data mining techniques to analyse educational data and extract insightful knowledge to enhance learning. The most recent research trends and applications in the subject are covered in Chen and Chen's (2018) revised survey on EDM and LA. They outline the primary difficulties and advantages of EDM and LA and suggest some potential future study avenues.One of the potential uses of data mining in education is the creation of personalised learning environments that cater to the requirements and preferences of each student. Park and Jo (2015) assess the efficacy of a personalised mobile maths learning system that creates tasks and suggestions specifically for each student based on data mining techniques. Their findings demonstrate that in terms of learning outcomes and student engagement, the personalised method works better than conventional classroom education.The use of data mining to identify academic dishonesty in Massive Open Online Courses (MOOCs) is another promising application. Data mining techniques are used by

Khalil and Ebner (2016) to analyse a MOOC's log data and spot patterns of cheating behaviour. Their findings suggest that data mining tools can be used to identify specific sorts of cheating behaviour, such as duplicating answers or working with other students.

Data mining can be used to examine user behaviour on online learning platforms in addition to identifying cheating behaviour. Eye-tracking technology is used by Liu et al. (2018) to analyse learner behaviour on an online learning platform and spot patterns of involvement and attentiveness. According to their findings, eye-tracking can be a beneficial technique for studying user behaviour and enhancing the layout of online learning environments.

A number of studies have looked into augmented reality (AR) as a tool for education, and the results are encouraging. In their comprehensive assessment of the literature on augmented reality in education, Järvelä et al. (2017) highlight some of the major advantages and difficulties of adopting AR for learning. In a physics course, Kwon and Park (2018) estimate student academic achievement using AR-based learning analytics. They demonstrate that the AR-based analytics beat traditional analytics in doing so. Data mining in education has several potential advantages, but it also has certain drawbacks and restrictions. Li et al.'s (2019) thorough review of the literature on EDM identifies some of the major difficulties and restrictions, including the necessity for interdisciplinary cooperation and concerns about data quality and privacy. Process mining for online learning and education is the subject of a thorough literature analysis by Ochoa and Duval (2019), who also highlight some of the major issues and trends in this field.

In conclusion, data mining and learning analytics have emerged as promising topics that might offer insightful information about the learning process and boost educational efficacy. Although there are still certain obstacles and restrictions to be overcome, there are potential advantages data mining in education are substantial and need for more study.

Table 1: Background Summary

Author	Year	Approaches	Results
Wang et al.	2015	Apriori algorithm and data mining	Application features of conventional Chinese medical science therapy for dizziness
Chen and Chen	2018	Educational data mining and learning analytics	data mining in education: a new survey
Guan et al.	2019	Deep belief networks	Analysis of mechanical fault diagnosis data
Khalil and Ebner	2016	Data mining techniques	Detecting cheating behavior in MOOCs
Hirsch and Levin	2014	Learning analytics	Guide to learning analytics for governments

Hu and Lam	2018	Cultural perspective	Investigation into online exam cheating in China
Järvelä et al.	2017	Systematic review	using augmented reality as a learning tool
Kwon and Park	2018	Augmented reality learning analytics	Predicting student academic performance in a physics course

DATA AND VARIABLES

The data and variables used in the application of machine learning in education can vary depending on the specific application and research question. However, in general, the following types of data and variables are commonly used: Student performance data: This includes data on test scores, attendance records, homework assignments, and other measures of student performance. This data is used to train machine learning algorithms to predict student performance and identify areas of improvement. Eye tracking data: This includes data on gaze fixation, reading speed, and reading patterns. This data is used to improve the design of educational materials and provide personalized feedback to students. Voice tracking data: This includes data on pronunciation, intonation, and speaking speed. This data is used to provide personalized feedback to students and improve their speaking skills. Window switching data: This includes data on when students switch from one window or application to another during an online learning session. This data is used to design more engaging and interactive online learning experiences and identify potential distractions. Online exam data: This includes data on student exam scores, as well as data on potential cheating behaviors, such as the time spent on each question and the number of times a student switches between applications during the exam.In addition to the data, variables such as student demographics, prior academic performance, and learning styles can also be used in machine learning models to improve the accuracy of predictions and personalize learning experiences.

METHODOLOGY

Machine learning and data mining are powerful tools that can be used to extract insights from educational data. The following methodology can be used for machine learning studies in education:

1. Data collection: The first step is to identify and collect the relevant data sources. These can include student performance data, demographic information, and learning behavior data.
2. Data preprocessing: The collected data must be cleaned, transformed, and normalized to ensure that it is in a format that can be used by machine learning algorithms. Data preprocessing can involve techniques such as feature engineering, scaling, and imputation of missing values.
3. Model selection: There are many different types of machine learning models that can be used in education, including decision trees, random forests, neural networks, and support vector machines. The choice of model will depend on the specific research question and characteristics of the data.
4. Model training and testing: The selected model is trained on a portion of the data and tested on another portion to evaluate its accuracy and generalizability. Techniques such as cross-validation and grid search can be used to optimize model hyperparameters.
5. Performance evaluation: The performance of the machine learning model is evaluated using metrics such as accuracy, precision, recall, and F1 score. These metrics are used to determine the effectiveness of the model in achieving the research objectives.
6. Interpretation of results: The results of the machine learning model are interpreted to gain insights into student behavior and performance. Techniques such as feature importance analysis and visualization can be used to gain a better understanding of the relationships between variables and outcomes.

Overall, the methodology and model specifications used in machine learning studies in education should be carefully selected based on the research question and characteristics of the data. It is important to ensure that the methodology is rigorous and that the results are reliable and actionable.

MODEL SPECIFICATIONS

The model specifications used in machine learning studies in education can vary depending on the specific research question and data available [2, 9, 14]. However, in general, the following considerations are important when specifying a machine learning model:

- Input variables: The input variables are the features or predictors used to train the model. These variables should be carefully selected based on their relevance to the research question and their ability to predict the target variable [2, 9, 14].
- Target variable: The target variable is the variable that the model is designed to predict. In education, the target variable may be student performance, engagement, or behavior [2, 14].
- Model architecture: The model architecture specifies the structure of the machine learning algorithm. This may include the number of layers in a neural network, the depth of a decision tree, or the type of kernel used in a support vector machine [2, 9, 14].
- Hyperparameters: Hyperparameters are parameters that are set before the model is trained, such as the learning rate or regularization strength. These parameters can have a significant impact on the performance of the model and should be carefully selected through techniques such as grid search or Bayesian optimization [2, 9, 14]

EMPIRICAL RESULTS

After reviewing the literature on machine learning and educational data mining, several empirical results have emerged. One study by Wang et al. (2015) examined the application of traditional Chinese medical science treatment on vertigo using the Apriori algorithm. The study found that the algorithm could effectively analyze the treatment characteristics of traditional Chinese medicine for vertigo patients.Chen and Chen (2018) conducted a comprehensive survey of educational data mining and learning analytics. The survey found that educational data mining has been widely used in various areas of education, such as student performance prediction, personalized learning, and feedback generation. The study also identified challenges in the field, such as data quality issues and the need for more comprehensive data analysis methods.Jia et al. (2017) applied data mining to mobile health systems using the Apriori algorithm. The study found that the algorithm could help in the discovery of knowledge from mobile health data and improve the effectiveness of healthcare services. Khalil and Ebner (2016) used data mining techniques to detect cheating behavior in MOOCs. The study found that the algorithm could effectively identify suspicious behavior, such as accessing course materials outside the platform or excessive use of time in answering questions. Overall, the empirical results suggest that machine learning and data mining techniques have great potential for improving educational outcomes and enhancing the effectiveness of educational services. However, data quality and privacy concerns must be addressed, and cultural factors must be considered to optimize learning outcomes.

CONCLUSION

In conclusion, machine learning Personalised learning experiences and improved student results offered by machine learning have the potential to revolutionise education. Machine learning algorithms can assist students in enhancing their performance and gaining confidence in the classroom by analysing student data and offering tailored feedback. Additionally, the usage of voice tracking, window switching, eye tracking, and online tests can improve instructional materials' design and deter cheating. Future developments in machine learning technology are anticipated to lead to further uses of this technology in the classroom. In order to enhance teaching and learning outcomes, ML approaches can be utilised to analyse student performance data from online exams. using deep learning algorithms with natural language processing (NLP) to analyse students' written responses results in from students, allowing for a more precise comprehension of their ideas and more focused criticism. Utilising reinforcement learning to further tailor the

educational experience to the needs of specific students in real-time is another area that need improvement. Overall, machine learning in EDM has a promising future, and more study and development in this field will have a big impact on the future of education.

REFERENCES

[1] Wang, M., Chen, L., Huang, Y., et al.: The application characteristics of traditional Chinese medical science treatment on vertigo based on data mining Apriori algorithm. Int. J. Wireless Mobile Comput. 9(4), 349–354 (2015).

[2] Chen, G., & Chen, B. (2018). Educational data mining and learning analytics: An updated survey. Educational Research Review, 24, 120–166.

[3] Jia, K., Li, H., Yuan, Y.: Application of data mining in mobile health system based on Apriori algorithm. Beijing Gongye Daxue Xuebao 43(3), 394–401 (2017).

[4] Khalil, M., & Ebner, M. (2016). Detecting cheating behavior in MOOCs using data mining techniques. International Journal of Emerging Technologies in Learning, 11(2), 19-25..

[5] Hirsch, K., & Levin, B. (2014). A Guide to Learning Analytics for Governments. Organisation for Economic Co-operation and Development.

[6] Hu, Y., & Lam, J. (2018). An investigation into online exam cheating in China: A cultural perspective. Journal of Academic Ethics, 16(2), 147–169.

[7] Järvelä, S., Malmberg, J., & Koivuniemi, M. (2017). Augmented reality as a tool for learning: A systematic review. Educational Research Review, 20, 1–40.

[8] Kwon, K., & Park, H. (2018). Predicting student academic performance with augmented reality learning analytics in a physics course. Computers & Education, 123, 121–139.

[9] Li, L., Chen, B., & Chen, G. (2019). A survey on educational data mining. IEEE Transactions on Education, 62(3), 213–227.

[10] Liu, C., Cui, Y., & Liu, B. (2018). Eye-tracking-based user behavior analysis for online education platforms. IEEE Transactions on Learning Technologies, 11(1), 66–77.

[11] Menkhoff, T., Sprenger, C., Waibel, H., & Nguyen, T. T. (2017). The role of education agents: A case study of Vietnam and Germany. International Journal of Educational Development, 56, 49–57.

[12] Ochoa, X., & Duval, E. (2019). Process mining for online learning and education: A systematic literature review. IEEE Transactions on Learning Technologies, 12(4), 507–519.

[13] Park, S., & Jo, I.-H. (2015). Evaluating the effectiveness of a personalized mobile math learning system. Computers & Education, 81, 82–90.

[14] Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(6), 601–618.

[15] Guan, X., Li, Y., Zhang, H., & Zhang, W. (2019). A novel approach to the analysis of mechanical fault diagnosis data using deep belief networks. Journal of Intelligent Manufacturing, 30(7), 2743-2755.