EAI Endorsed Transactions
on Tourism, Technology and Intelligence
Research Article EAI.EU

The Role of Machine Learning in Smart Education: Taxonomy, Challenges, and Use Cases

Premisha Premananthan[1,*], M. Fahim[2]

[1]Memorial University, Canada
[2]Queen's University Belfast, UK

Abstract

Education is a powerful domain of any country where the changes happened in this domain will reflect all other domains as well. The technical advancement should start with education domain or else there is no strength to that particular advancement. After COVID-19 cause severe upheaval to almost all the industries, in education, the adaptation significantly impacts the development of smart education. Even the developing countries were in the position to adapt the technological advancement through this pandemic. Machine learning plays pivotal role in the technological improvement. The intrusion of smart education fosters an abundance of electronic data and solutions. Machine learning techniques are used to implement models to analyze these larger datasets. In recent years, there have been plenty of studies which address the changes in education and model solutions using various machine learning techniques, such as Supervised, Unsupervised, Semi-supervised, Deep learning and Reinforcement learning techniques.This paper provides an overview, challenges and future directions of research on machine learning techniques applied in education with different levels.

Received on 02 Apr 2024; accepted on 10 Aug 2024; published on 23 Sep 2024

Keywords: Analysis, artificial intelligence, machine learning, smart education, review

doi:10.4108/XX.X.X.XX

1. Introduction

Education has the actual need for digitalization; even with technical improvements in other domains, the actual development only shows when the education sector, the key to powering up other domains, has reached those technical advancements. COVID-19 had a greater impact on each and every sector, and the education industry is no exception. Looking back at the changes in the education industry during the pandemic when the digitalization in the education was quickly implemented, there were positive impacts. Various online platforms were open to students and educators, providing opportunities for seamless learning/teaching experience and an environment to interact. In developing countries, the impacts were even more significant: online learning has provided an access to high-quality education and training, and to some extent, narrowing the gap between developed and developing countries.

In the education sector, the incorporation of technologies can be seen in teaching, learning, administration and in the use of intelligent systems. Smart education [1] includes applications tailored for education, such as smart classrooms, virtual learning environments, virtual reality (VR), augmented reality (AR) and mixed reality (MR) technologies and Internet of Things (IOT). Traditionally depending on the teachers, educational pedagogy now has been tremendously adapted through the use of smart devices. However, while smart education has brought positives into education, there remains inequality towards achieving this smart education to all around the world. There are countries that lack technical facilities and knowledge towards providing smart education to their students. In future there are so many chances to overcome this issues.

Amongst other methods, machine learning is an underlying technology in smart education. As we shall see in the next sections, machine learning techniques

*Corresponding author. Email: ppremanantha@mun.ca

allow the prediction of students' outcome, subject allocation, personalized learning experience, and so on. In this paper, previous studies are analyzed and clearly structured towards the achievements of smart education and its limitations. Educational pedagogy has been tremendously adapted through the use of smart devices whereas in past it was mostly dependent on the teachers, students and parents but now smart devices play major role in here.However these smart education brings so much of positively in the education domain still there are inequality towards achieving this smart education to all around the world. There are countries with lack of technical facilities and knowledge towards providing smart education to their students. In future there are so many chances to overcome this issues. In this paper, the previous studies were analyzed and clearly structured towards the achievements of smart education and it's limitations. In this study, we analyze and draw some key points regarding machine learning techniques, their applications towards education sector improvements, challenges while developing machine learning models to solve the problems in education technologies, and the doors of machine learning applications in smart education.

2. An Introduction to Machine Learning Techniques

Machine learning is one of the major component of artificial intelligence (AI) that aims at the implementation of algorithms that make computers learn from and develop forecasts or solutions based on data. The major goal is to enable machines to learn patterns from data automatically.

2.1. Supervised Machine Learning Techniques

Supervised machine learning used labelled input samples to learn the relationship between input and output. These techniques are mostly used for implementing complex models to get accurate predictions. In [2], the authors predicted the learners' scores based on the log data of a selected e-learning platform. A fresh batch of students from a selected engineering department was chosen. Out of 12 weeks of a semester duration, they collected 9 weeks of log data, except for the exam's week. The learners' behavior data consists of 56 features from the online learning platform. They then used multiple ML algorithms such as linear regression, support vector machines, K-nearest neighbours, multi-layer perceptron, and decision trees to predict the learners' scores. Each prepared dataset was used in the above mentioned five algorithms with k-fold cross validation. Out of these five algorithms, the performance of multilayer perceptron was quite high because it can handle more complex data better than the other models chosen in this study. For future research, the authors suggest including learners' personal data such as age, prior studies, gender, and native language, which will interrupt the prediction score, to get a better prediction of the learners' score.

2.2. Unsupervised Machine Learning Techniques

The study in [3] implemented a partly automatic metacognition assessment using unsupervised machine learning techniques. Metacognitive skills are vital in the learning process, specially when it comes to undergraduate students. While most of the metacognitive skills can be assessed through interview or think-aloud procedures, this is not possible with a larger number of students. The data was collected in an Indonesian engineering course by using a metacognitive awareness inventory questionnaire with 52 questions. The model used two algorithms: K-means and expectation maximization clustering algorithms to measure the compactness of the clusters by using silhouette coefficient. The datasets were clustered into two: one consisting of who performed quite well and the other consisting who did not. The pre-processing results using K-means was used for identifying characteristics of the two clustered in terms of eight metacognitive sub-components. The proposed method provided an approach to help determine students who would need further support.

2.3. Deep Learning Techniques

Deep learning is a very popular neural network-based technique that is implemented in an H2O framework to extract patterns and classes from information within data in a layered model. It mimics a human brain where each neuron is modified based on the available information and also predicts the output, acting as a classifier. They excel at learning representations from complex and unstructured data, such as images, audio, and text.

A deep learning technique was introduced in [4] to allocate subjects to faculty members. The efficient subject allocation is required in order to utilize each individual's academics. The system needs to consider their academic loads, administrative works, further research works, experiences, deep knowledge and practices of Bloom taxonomy and qualifications. Previously, data mining techniques had been used to analyze the effectiveness of this subject allocation, but those results did not have the practical implications the educational domain. The authors implemented a model with regression with 20 input parameters and 4 output classes. An artificial neural network (ANN) was used as information processing system for this regression model to allocate the subjects based on the individual qualifications in research and innovation, academic and administrative duties given to them. The proposed technique was proven to be efficient and allow faculty members to handle both academic work

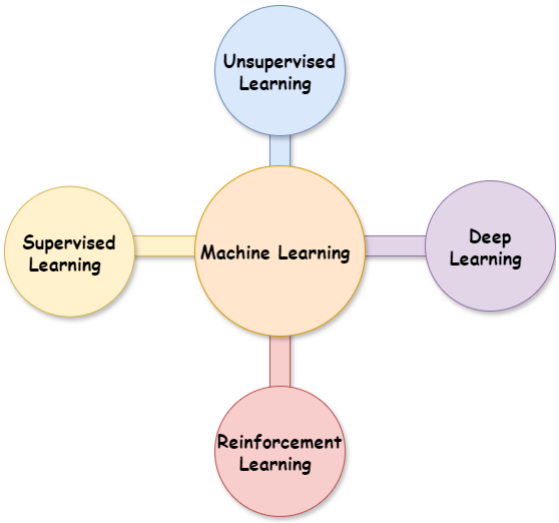

Figure 1. Machine Learning Techniques

load and administrative tasks while ensuring quality of teaching.

2.4. Reinforcement Learning Techniques

Reinforcement learning is a very powerful algorithm where an agent sequentially learns and makes decisions by interacting with the environment without any supervision in order to maximize the rewards. This desirable method is based on exploration, which means the interactions with the environment, and exploitation, which indicates the learning through and observations of how it responds, similar to the trial-and-error method.

A deep reinforcement learning model was implemented in [5] in order to resolve all the major issues in personalized learning environment. This KNIGHT (AI in Education at Hochschule für Technik Stuttgart) framework was used to achieve two major goals. One is to provide satisfied personalised learning environment to students and the other is to utilize the instructor's potential to maximum level. In this framework, they also used multi-modal learning analytics to utilize multiple sources to create a strong platform of student learning. This model is a bundle of education and technology innovations; it also provides various solutions towards students, educators and administrators to improve learning outputs. Also, there were also other studies on personalized learning platforms to learners using reinforcement learning guided frameworks based on Q-Learning techniques [6, 7].

3. Appliances of machine learning techniques towards classification of education

The applications of machine learning techniques spread in all the domains. In education, this techniques used in various ways. However, there are many education levels, according to which the applications of machine learning should be adaptive. In this section, we consider the major classification levels of a child's education according to the age division.

3.1. Early childhood education

Early childhood includes children under the age of five. Education at this level mostly focuses on sensory development, basic motor skills, language development, basic social skills and early literacy and numeracy skills in order to prepare for the formal schooling. Basic motor skills are very important to build the sports skills in future. In early education, children are mostly encourage to follow these motor skills. But there are some lack of proper training and followup of motor skills. In order to automatically track, evaluate and present the skills in early childhood, a research work implemented a toolkit called Early childhood actions monitoring and analytics system. This automation system used to integrate with machine learning techniques, computer vision with Test of gross motor development- 2nd edition toolkit to evaluate the activities including walking, running, jumping and sitting up from a given video input.In this paper authors used BlazePose model to track the 3-D poses and to validate the system involved physical training expert collaboration. Also this paper suggested two future works such as identifying motor skills which children are not explicitly monitored and

recommendation of the deviations of the corrective actions by highlighting the difference from reference action [8].

Most young adults loves to study through fun segments like images, videos , audios and games rather than texts. The authors in [9] developed a mobile application using deep neural network to automatically identify the fruit's images name as audio and text format. In that research implement,using mobile app camera children can take any fruit's picture and processed it will output the exact name not only textual format but also audio format as well with 93.29% accuracy.To develop the system, Convolution Neural Network was primarily used with Firebase database.This application helps the young adult to encourage self-learning with fun element.

3.2. Primary Education

Primary education consists of the children from the age of six to twelve years in the aim of providing the basic education in subjects like reading, writing, mathematics, science, and social studies. Emphasis on foundational skills and personal development.

The COVID played major role in the online learning system improvement. The authors proposed a personalised assistive e-learning web system with the help of machine learning in order to fill the gap of primary students schooling experience in the pandemic situation. The system majorly focused on three main activities: drawing activities, mathematical activities and speaking activities. These proposed web system worked as a personal teacher to each students and special attention was developed in each students learning progress. In each activity they had personalised progress report with grades and suggested activities to improve the lower grades. All three activities used Convolutional Neural Network for training. All these activities used separate datasets for mathematical activities used datasets from Modified National Institute of Standards' dataset as training dataset and processed the sample with OpenCV toolkit, for kinder art activities used Google quickdraw dataset and speaking activities used spectral representation of the auditory (Librosa).The authors implemented a personalised recommendation system to the primary students with the aim of developing cognitive skills with customised learning activities based on each students knowledge level.This will help the students to improve self learning with effective suggestion without manual interaction and all the outputs of their performance are

3.3. Secondary education

Secondary education is mostly focused on carreer development skills and more specialized subjects, and development of critical thinking skills to the age of 12 to 18 years old.

A research team worked with the secondary education students who are members of intergenerational co-design team also who dont have prior programming skills in order to demonstrate the way of children's understanding using teachable machines. The authors used adult co designers to engage with children via existing teachable interface provided by Google abbreviated as Gteach. The selected children used their own designs like origami shapes and experience the machine's classifier's performance. The sessions were video recorded for the collection of the demonstration with set of training with different children.The results extend and reaffirm evidence from prior work and reveal new understandings regarding children's interactions with machine teaching [10].

3.4. Higher education

Higher education system needs more careful and rapidly adaptive curriculum changes towards the competitive industry environment.The research [11] conducted based on the curriculum improvements of higher education level in order to learn basic machine learning and data analysis models and tools with the aim of awareness of cyber threats analysis. Nowadays the most of the higher studies students faced the cyberthreats as they are the one who used the platform mostly.The research method implemented with the approval of University of Texas A & M IRB. The first set of execution of these new curriculum took part from teachers with the help of their data collection. The cyberdata anaytics combines data science and cybersecurity in order to identify and mitigate various cyberthreats. Waikato Environment for Knowledge Analysis tool was used to analyze the data to get the final output to integrate the machine learning into cybersecurity education systems.

4. Role of machine learning in education

4.1. Student's Learning Environment Improvements

Improving traditional classroom interaction between student and teacher is the key aspect when it comes to a student's learning environment's digitalization. This digitalization should ensure fair interaction between each student and the teacher. To achieve that, an omni-directional interaction model was implemented by the authors. The major aim of this model was to promote diversified interaction between teachers and students. In this model, there is no direct teacher-student interaction, but rather teacher-technology-student interaction, which will lead to neutral communication in the classroom. Traditional person-to-person interaction relies one-to-many communications, and it cannot

ensure the fairness of teacher-student communication. But through this omni-directional model, there will be human-machine interaction with inter-subjective communication. This will be a major platform to the student's active participation and omni-directional interaction in the learning environment [12].

Blockchain is another technology [13] that is widely used in the education sector. For various reasons, we can use blockchain technologies in education to create smart contracts where, in a virtual environment, educators and students can agree to an agreement based on how the contents were used to evaluate a student's assignments and grades as well. With the blockchain integration, more traditional methods were replaced with effective digitalization such as digital certificates, customized learning paths, which can minimize early dropouts, and attendance monitoring, where the pattern of each student's participation can be actively managed. Also, blockchain integrated machine learning techniques are used in the education sector to leverage the system. Maintaining secure documentation like degree certificates was tough in the traditional way, but nowadays there is secure credential verification and transparent record maintenance in order to digitalize effective documentation.

4.2. Student's Performance Prediction

The identification of factors affecting high-school students' each semesters was very important because each semester the students' behaviors and performance have vast changes. In order to identify the features, a machine learning-based performance evaluation model was developed [14]. To implement this model, random forest, support vector machines, logistic regression and artificial neural network (ANN) techniques were used. Moreover, the importance of features was calculated using the Boruta algorithm. The ANN accuracy was reduced because of the smaller number of data points; hence, the dataset was originally collected from surveys. SVM's accuracy was the highest compared to the other two algorithms. It was revealed from the study that the student's attention is the major factor which affected their semester performance.

To measure the student's learning performances, the authors in [15] used ensemble deep learning techniques. They implemented the models on the five given dataset with two simulations. On the first set of simulations, they used traditional machine learning algorithms: Gaussian naïve Bayes, support vector machine, decision tree, multi-layer perceptron, random forest, linear discriminant analysis, and quadratic discriminant analysis. Out of those traditional machine learning algorithms, multi-layer perceptron outperformed. In the second simulation, the convolutional neural network (CNN) was used, which was proven to be more accurate than the previous traditional algorithms. As a conclusion, the CNN algorithms outperformed other traditional machine learning algorithms, and when it comes to augmented datasets, more accuracy can be achieved.

4.3. Virtual Reality (VR) and Augmented Reality (AR) in Education

Augmented reality (AR) allows users to interact with digital elements such as smartphones, tablets and AR glasses such as Microsoft HoloLens and Google Glass in order to overlay digital information. Virtual reality (VR), on the other hand, creates a full immersion in a virtual world using controllers and motion tracking that isolates the users from the real world. Another one is mixed reality (MR), which compresses the real and virtual world in order to produce a new environment which will enable more interaction between them. In the education domain, the use of these AR/VR/MR is pivotal, especially in subjects which have lot of practical interactions such as physics, mathematics, and medical science. In this regard, the study in [16] implemented an AR/VR teacher for physics and mathematics subjects to assist the student in a more interactive way. Using AR glasses, they created a virtual environment which helps students do their homework, assist in their practical sessions and review subject materials without taking a long time. Unity 3D and Android studio software development platforms were used to develop this entire environment with the help of Oculus VR headsets. The students can record the sessions as well. This environment acts as a library to the students, but rather than having data only, it visualizes those data into objects in front of the students. Using this system, student's attention was increased by 60%.

The review paper in [17] analyzed past research papers using AR, VR and MR technologies in education and identified future research gaps in this domain. For this knowledge mapping, 2092 research papers were collected from the Web of Science, from which indicates nine hot-spots and trends in these simulation technologies were specified. Most of the reviewed papers resolved issues from children's education, medical studies and engineering studies using VR/AR and MR technologies. Also, the learning through games and visualization led to the students' knowledge acquisition at a high level.

4.4. Natural Language Processing for Language Learning

Natural language processing enables the machines to deal with human languages via text or voice. Mostly in the education domain several studies were done through text-based natural language processing.

After the COVID 19 the education system faced major changes like online education system. This online education system was not appropriate in developing countries because they have the issues availability of tools, stable internet connection and knowledge of usage. To analyse the basic issues faced by the students there was a study conducted on collecting survey questions via online and offline mode at Bangaladesh. On this study the authors first labelled manually the collected 50005 Bangala and Romanized bangala texts by multi labelling using sentiment analysis. To classify the students' opinions about the online learning they used multiple classifiers such as Support Vector Machine (SVM), Multinomial Naïve Bayes, Random Forest and Logistic Regression machine learning algorithms. The final evaluation was showed the highest performance was given by SVM classifier. The result showcased most problem identified was financial problem followed by unstable internet problem.

5. Conclusion

In this review study, there are lot of technical advancements and impacts of various issues towards education domain were analyzed. Specially the Machine learning techniques used in this domain was quite different from other domains. Each levels of education system used their own advancements in order to achieve their goals.Specially the customised goals for each levels with technical improvements were vital to the students' performance increment. Most of the smart devices introduced in this COVID - 19 to most of the countries. Also COVID brings so many hidden issues of traditional system in the public stage to address the solutions. In this review paper we address the adaptations and future gaps to find solutions in the education domain.

References

[1] Afzal Badshah et al. "Smart Education in Smart Cities: Challenges And Solution". In: *2023 IEEE International Smart Cities Conference (ISC2)*. 2023, pp. 01–08. DOI: 10.1109/ISC257844.2023.10293615.

[2] Worawat Lawanont and Anantaya Timtong. "Smart Education Using Machine Learning for Outcome Prediction in Engineering Course". In: Institute of Electrical and Electronics Engineers Inc., 2022, pp. 63–68. ISBN: 9781665400145. DOI: 10.1109/KST53302.2022.9729078.

[3] I Hidayah et al. "Work in progress: Application of unsupervised learning method toward student's metacognition assessment". In: *2016 IEEE Global Engineering Education Conference (EDUCON)*. 2016, pp. 1175–1178. DOI: 10.1109/EDUCON.2016.7474704.

[4] Nitin Kumar Saxena et al. "Knowledge-Based Recommendation for Subject Allocation Using Artificial Neural Network in Higher Education". In: *IEEE Transactions on Education* 66.5 (2023), pp. 500–508. DOI: 10.1109/TE.2023.3296315.

[5] Muddsair Sharif and Dieter Uckelmann. "Multi-Modal LA in Personalized Education Using Deep Reinforcement Learning Based Approach". In: *IEEE Access* 12 (2024), pp. 54049–54065. DOI: 10.1109/ACCESS.2024.3388474.

[6] Jingheng Wang et al. "Learning path design on knowledge graph by using reinforcement learning". In: *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2023, pp. 3480–3485. DOI: 10.1109/BIBM58861.2023.10386061.

[7] Chelsea William Fernandes et al. "Unleashing the Potential of Reinforcement Learning for Enhanced Personalized Education". In: *2023 IEEE Frontiers in Education Conference (FIE)*. 2023, pp. 1–5. DOI: 10.1109/FIE58773.2023.10342902.

[8] Isuru Supasan Naotunna Andarage et al. "Early Childhood Action Monitoring and Analytics System (ECAMS)". In: *2023 IEEE 28th Pacific Rim International Symposium on Dependable Computing (PRDC)*. 2023, pp. 346–351. DOI: 10.1109/PRDC59308.2023.00056.

[9] Ameya Dikshit et al. "Image-to-Text and Speech-based Learning Aid for Children". In: *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*. 2023, pp. 1–5. DOI: 10.1109/I2CT57861.2023.10126501.

[10] Utkarsh Dwivedi et al. "Exploring Machine Teaching with Children". In: *2021 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 2021, pp. 1–11. DOI: 10.1109/VL/HCC51201.2021.9576171.

[11] Kossi Bissadu and Gahangir Hossain. "Designing a High School Course on Machine Learning for Cyberthreat Analytics". In: *2024 IEEE 14th Annual Computing and Communication Workshop and Conference (CCWC)*. 2024, pp. 0214–0219. DOI: 10.1109/CCWC60891.2024.10427775.

[12] Qi Wu. "Networking College's Smart Education System Based on Omnidirectional Interaction Model". In: *2021 2nd International Conference on Information Science and Education (ICISE-IE)*. 2021, pp. 141–144. DOI: 10.1109/ICISE-IE53922.2021.00039.

[13] T. Swathi et al. "Exploring the Synergy of Machine Learning and Blockchain in The Education Sector". In: *2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*. 2023, pp. 1–3. DOI: 10.1109/SMARTGENCON60755.2023.10442389.

[14] Mostafa Zafari et al. "A practical model for the evaluation of high school student performance based on machine learning". In: *Applied Sciences (Switzerland)* 11 (23 Dec. 2021). ISSN: 20763417. DOI: 10.3390/app112311534.

[15] Worawat Lawanont and Anantaya Timtong. "Smart Education Using Machine Learning for Outcome Prediction in Engineering Course". In: *2022 14th International Conference on Knowledge and Smart Technology (KST)*. 2022, pp. 63–68. DOI: 10.1109/KST53302.2022.9729078.

[16] Lidia-Cristina Bazavan et al. "Virtual Reality and Augmented Reality in Education". In: *2021 30th Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEEIE)*. 2021, pp. 1–4. DOI: 10.1109/EAEEIE50507.2021.9531005.

[17] Hai Zhang et al. "Hotspots and Trends of Virtual Reality, Augmented Reality and Mixed Reality in Education Field". In: *2020 6th International Conference of the Immersive Learning Research Network (iLRN)*. 2020, pp. 215–219. DOI: 10.23919/iLRN47897.2020.9155170.