ISSN No:-2456-2165

Optimizing E-Learning Platforms using Machine Learning Algorithms

Peter N.Mulei[1], Salesio Kiura[2]
[1,2]School of Computing and Information Technologies,
Technical University of Kenya, P.O Box 52428-00200–Nairobi-KENYA

Abstract:- The proliferation of e-learning platforms and blended learning environments has spurred a great deal of study on how to improve educational processes. The problem with the e-learning platforms, give the content as whole without considering the level of cognitivity of learners.One key factor is being able to forecast student performance with accuracy. Early in the learning process, it is useful to detect low-performing pupils based on a high forecast accuracy of their performance. But in order to accomplish these goals, a lot of student data needs to be examined and forecast using a variety of machine-learning models. Machine learning algorithms have shown to be a useful tool for focusing performances at different learning levels when used to forecast learners' actions based on their performance and background. For the purpose of enhancing learning outcomes, early student performance prediction is helpful. utilizing clever and flexible components to provide students with a personalized learning environment.Differentiating prediction levels by different machine-learning models may be the result of variations in socioeconomic conditions. Specialized scope classifiers are then merged into an ensemble to robustly forecast student achievement on learning objectives independently of the student's specific learning settings. Personalized Learning Environmentsimprove the educational process by offering specialized services that are based on the preferences of the learners.

Keywords: *Machine Learning Algorithms, E-learning, Ensemble, Personalized, and Adaptive.*

I. INTRODUCTION

Machine learning is becoming widely used as a result of the quick advancement of technology. Learners' logs are recorded in electronic learning environments, and a multitude of learning attributes are noted. The e-learning logs are analyzed using machine learning techniques to forecast how well the students will learn[6] Through the process of forecasting students' progress, instructors can step in early to prevent students from failing and increase their chances of achieving success. The three primary algorithms in machine learning are classification, clustering, and regression[4]

E-learning can be facilitated by the Learning Management System (LMS). Algorithms for machine learning have been utilized to help provide educational materials to students. Additionally, it enables instructors to evaluate and comprehend how pupils behave and perform as they progress through the learning process. In e-learning, personalization highlights how every learner is different and has a different learning style. Numerous approaches and structures exist for customized education. Owing to the proliferation of networks and technical advancements, e-learning systems that support instructors and students in achieving individualized learning are desperately needed. One of the most well-liked individualized learning strategies is learning styles, which take learners' learning preferences into account [14]

II. RELATED WORK

A. Introduction.

Predicting students' academic success is turning into a crucial function in the computer-supported intelligent education system. But in order to forecast students' academic achievement, traditional machine learning-based techniques can only take use of the sparse discriminative characteristics of students' behaviour in unbalanced academic datasets[5]But the majority of studies on AI prediction models that have been conducted have mostly focused on the development and optimization of AI models, i.e., using multiple algorithms to develop models with higher prediction accuracy; this is in contrast to applying AI models to provide real-time and continuous feedback and improve the quality of student learning[8].

A number of research groups have emerged as a result of the greater usage of educational data, including learning analytics, which aims to predict learners' behaviour and find ways to inspire them while they learn. Gathering, compiling, reviewing, and analysing student login records is one aspect of learning analytics. This process allows for the inference of the learning environment and optimizes the performance of both lecturers and students(Viberg et al., 2018; Capuano & Toti 2019). With the use of technology-enhanced learning platforms and an abundance of available educational data, it is possible to mine student learning behaviour, correct problems, optimize the learning environment, and facilitate data-driven decision-making. Virtual learning environments are a valuable addition to the learning analytics paradigm as they efficiently furnish statistics for examining, reporting, and reflecting on students' learning experiences and how those experiences impact their individual performances. Instructors can have real-time access to student information by using machine-learning algorithms, which enables them to step in early in the course. Predictive models are frequently constructed using machine learning algorithms and student data. With these machine approaches, categorical and numerical predictor variables can be handled.

B. Machine Learning Algorithms in E-learning Platforms

Based on the kind of learning signal or feedback that a learning system has access to, machine learning systems can be broadly divided into four categories: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. In order to provide meaningful information, machine learning may extract and analyze data, search for patterns and correlations from huge data sets, and more. Several methods employed in the development of e-learning platforms include Neural networks, which consist of an interconnected network of nodes that collaborate to provide solutions, Utilizing Support Vector Machine (SVM) to classify data in order to produce the optimal solution space In order to identify a method for identifying the most appropriate route that leads to problem solving, decision trees generate images of decisions [13]

C. Algorithms in E-learning Systems.

A learning management system e-learner's likelihood of finishing the registered course must be predicted by a machine. The machine uses its prior knowledge and experiences to complete the assignment. Learner behaviour data collected during the course, including interactions, responses to activities, and outcomes. The creation of algorithms that can take in input data, analyze it statistically, and forecast an output while updating those outputs as new data becomes available is the fundamental idea of machine learning[2].

D. Supervised Learning

Accepting labeled examples such as an input where the desired outcome is provided—is the first step in the creation of supervised learning. It offers a dataset with labels in addition to features. This method is frequently applied in situations where past data indicates the likelihood of future events. Supervised learning falls into two categories: regression and classification. Regression uses a continuous label, whereas classification uses a discrete label.

E. Unsupervised learning.

Unsupervised learning makes use of data without prior examples with the aim of using data exploration to identify similarities between the items. Actually, it's a method of identifying labels directly from the data. The prediction model is created by the unsupervised learning algorithm during the training phase, and it attempts to fit its parameters to the best summarized regularities it can find in the data.

F. Reinforcement Learning.

The algorithm of reinforcement learning finds out via trial and error which action yields the highest reward. Finding the best policy is the objective. In this kind of learning, the learner, the surroundings, and the activities are the three key elements.

G. Semi-supervised learning

Both labeled and unlabeled data are used in semi-supervised learning, and it is ideal for the task to already have a "prediction problem." The model then gains the ability to arrange data and predict as well. Classification, regression, and prediction techniques are applied to this kind of learning (e.g. facial recognition on videos or photographs)

H. Current machine Algorithms used in e-learning platforms.

- *Decision Tree*

Utilizing a predictive machine-learning model, it determines the value of a new sample by utilizing different attribute values from the accessible data. A decision tree's internal nodes represent the various qualities, its branches indicate potential values for these attributes in the observed samples, and its terminal nodes indicate the dependent variable's ultimate value (classification)[3]For learners' datasets to be subjected to predictive analysis throughout the learning process, data mining is required. Decisions regarding the learners' future steps will be made using the information or knowledge that data mining techniques can produce, which is the prediction of learner performance and attainment[3]

- *Naïve Bayes algorithm.*

It is a supervised learning approach for classification issues that is based on the Bayes theorem. Its primary application is in text categorization, where a high-dimensional training dataset is used. One of the most straightforward and efficient classification methods is the Naïve Bayes classifier, which aids in the rapid development of machine learning models with rapid prediction capabilities. It may be used to generate data from any institution utilizing data mining techniques for real-time predictions and the prediction of student performance and achievement, which will help learners make decisions[3]

- *Logistic Regression*

This categorization may be used to forecast distinct results. Its advantages include efficiency and a lower tendency to overfit. It is one of the supervised learning algorithms that is helpful in statistical techniques and to predict the outcome of dependent variables. It is also the simplest algorithm that doesn't require a lot of processing power, straightforward to build, and quite quick. The logistic regression algorithm has a very low overfitting rate and is highly efficient. This algorithm is accurate and basic[12] and it is simple to update to take into account fresh data. The fact that it cannot solve nonlinear issues is one of its drawbacks. This algorithm may cause the training dataset to become overfit when it uses a high dimensional dataset[12]

- *Random Forest (RF)*

It is an ensemble approach with numerous trees that is based on decision trees. Tree votes are the classifications that every tree in the forest offers. The vote tree with the highest vote is selected as the final forecast by the forest fusion. Because so many decision trees are utilized in the process, random forest is an effective and extremely accurate strategy**(**Lassaad K. et al., 2022)Out-of-bag data is the remaining data that does not have a decision tree and is used to evaluate the decision tree's performance. Using feature ranking, which takes into account a feature's significance across the entire dataset, it adopts the feature

selection approach. The RF has extremely high accuracy and robustness. (Benny et al., 2020)In comparison to other algorithms, such as Logistic Regression (LR), which achieved accuracy of 87.99%, precision of 89.19%, recall of 93.90%, and f1-score of 0.9148, Qingxiang & Jiesen 2021 report that the iterative RF algorithm performed exceptionally well in predicting students' performance, with an accuracy of 88.55%, precision of 88.21%, recall of 95.86%, and f1-score of 0.9187. The general regression neural network (GRNN) model had the worst performance. GRNN obtained 84.35% accuracy, 86.40% precision, 90.75% recall, and 0.8852 f1-score. The RF algorithm performs better in classification when compared to other algorithms. Large-scale data processing, large-scale variable parameter support, and intuitive feature importance evaluation are all capabilities of this system. Due to its bagging effect, the RF algorithm has demonstrated through numerous algorithm practices and competitions that it has superior robustness and stability, as well as good classification performance[11]

➢ *Support Vector Machine*

It's a machine learning method that was created mostly with binary classification in mind. Supervised learning techniques, or SVMs, help analyze datasets and categorize them into classes[10] The SVM method works well in large dimensions and is incredibly memory efficient. Among the benefits are that it performs best in large dimensions and functions effectively when there is a distinction between the classes in the data. It uses the kernel to handle nonlinear data and is reasonably memory efficient. Some drawbacks include the inability to function properly in the presence of noise in the dataset, unsuitability for larger datasets, and lengthy training times [10]

I. Personalization.

To personalize something is to create something that is tailored to a person's needs. The e-learning platforms ought to adjust to various user attributes, including abilities and preferences. In order to increase a learner's pleasure and success in the course, the functionality allows the system to specifically address the needs and characteristics of each individual learner, including their degrees of experience, prior knowledge, cognitive capacities, skills, interests, preferences, and learning styles. The goal of the personalized learning strategy is to give each student an efficient, effective, and customized learning route so they may all take part in the process[7].

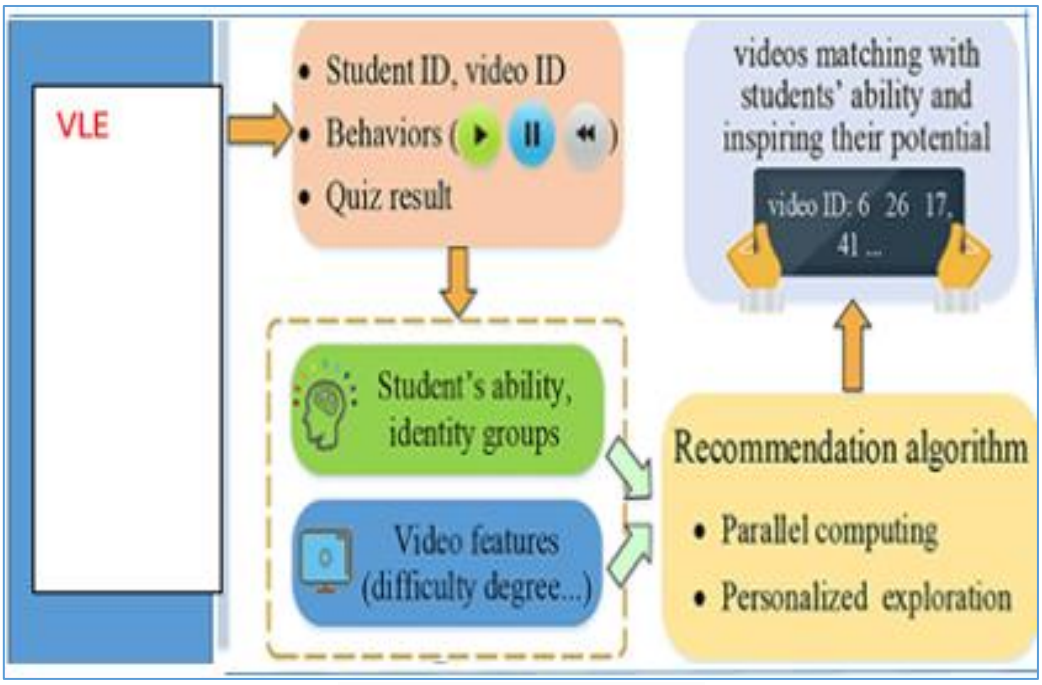

Fig. 1: Personalisation processing.

➢ *Learning Paths vs Learning Styles*

The learning route targets a set of objectives by the time it is finished, but it strives to accomplish one goal at a time. The information gathered by the learning management system about users can be used to define the learner's profile, understand their behaviour, and pinpoint their requirements and challenges. Observing the activities that students take on the system is one method to support them; these actions might lead to pathways called Learning Paths (LP). Activities include things like tasks, discussions, forums, and quizzes. A learning path is intended to give students a direct path through a course of study. It simplifies course material into digestible portions so that students can quickly assimilate knowledge and move through each course without difficulty[5]

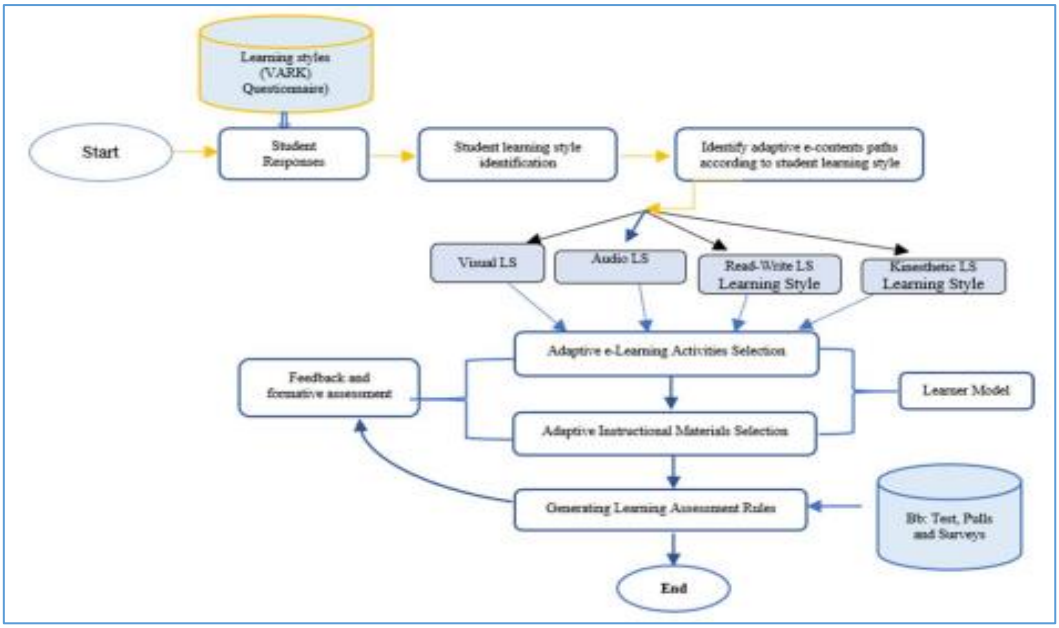

Fig. 2: Path through Learning Style [1]

III. METHODOLOGY

Various Machine Learning techniques are used to study student progress and engagement in different virtual learning environment activities (VLE). These techniques are suitable for both domain and categorical educational attributes. The log files in VLE are a typical source of information regarding user-system interaction, these records include every login and activity, and the model was constructed using the information from thelogs. Weintegrated the model into the Moodle and monitored the learners and validate the model using the cross validation with five folds. The clicking position and behavior are key in the learner's monitoring activities [8] Pre-processing methods, the dataset, and specifics of the machine learning model were used. Overall information was taken into account, including demographic information, clickstream data for each student every quarter, and student engagement on the VLE portal [15]

IV. RESEARCH DESIGN

Figure 3.Give the explicit process which was carried out during the model development.

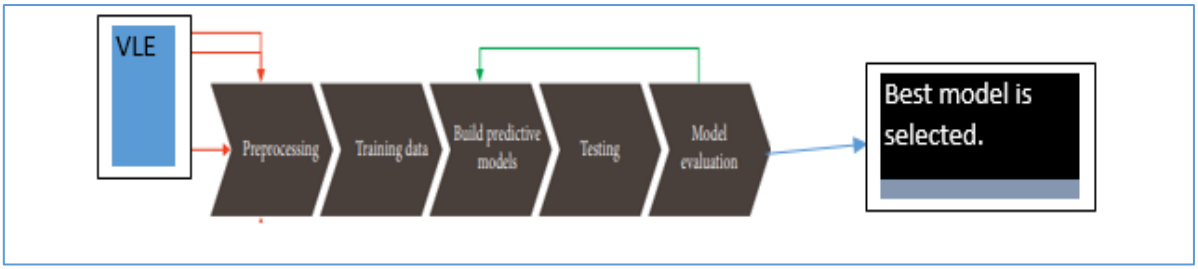

Fig. 3: Algorithm modeling Process

V. RESULTS AND CONCLUSION

E-learning platforms are being utilized more and more to deliver effective learning services, particularly since the World Health Organization declared the COVID-19 pandemic to be worldwide. Institutions of higher learning have embraced e-learning. We categorized machine learning algorithms used in e-learning and helped with personalization and adaptation, which helped create a positive learning environment for the learners. In this paper, we presented a thorough review of the efforts made to apply information and communication technologies to improve e-learning services. This paper also explains how to offer more flexible and customized course delivery and personalized learning by leveraging data collected in big data-based e-learning systems. Additionally, demonstrated a learning style pattern that, when the optimal algorithm is applied to the learning platform, promotes the optimal learning path in terms of cost and time. Student motivation and outcomes are significantly impacted by the path analysis model, individual needs, and the perceived value of e-learning. Furthermore, the perceived value of e-learning outcomes is significantly impacted, albeit indirectly, by student motivation[14]. [15] asserts that early intervention methods for at-risk students should be provided in order to identify them. According to the findings, the suggested model can classify data with an accuracy of 84%–93%.The results demonstrate that deep artificial neural networks perform better than the standard logistic regression and support vector machine models. The support vector machine achieves 79.95%–89.14% accuracy, whereas logistic regression achieves 79.82%–85.60% [14]

REFERENCES

[1]. Hassan A&El-Sabagh, (2021) Adaptive e-learning environment based on learning styles and its impact on the development of students' engagement.
[2]. Fan Ouyang et al., (2023) Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course.
[3]. Agung &Wahyu, (2021).Comparison J48 and Naïve Bayes Methods in Educational Analysis. Comparison J48 and Naïve Bayes Methods in Educational Analysis - IOPscience
[4]. Huang, A. Y. Q., Lu, Huang, (2020). Predicting students' academic performance by using educational big data and learning analytics: Evaluation of classification methods and learning logs. Interact. Learn.
[5]. David et al., (2021), A New Learning Path Model for E-Learning Systems. Available from: https://www.researchgate.net/publication/350559518_A_New_Learning_Path_Model_for_E-Learning_Systems [accessed Oct 28 2022].
[6]. Lu, O. H., Huang, A. Y., Huang, J. C. (2018). Applying learning analytics for the early prediction of students' academic performance in blended learning. J. Educ. Technol. Soc. 21, 220–232.
[7]. Hussein & Al-Chalabi, (2021), Adaptive e-learning environment based on learning styles and its impact on the development of students' engagement.
[8]. Jiang, G., Feng, X., Liu, W., and Xingjun, L. (2020). Clicking position and user posting behavior in online review systems: a data-driven agent-based modeling approach. Inform. Sci. 512, 161–174. doi: 10.1016/j.ins.2019.09.053
[9]. Lassaad K (2022), Using Ensemble Learning Algorithms to Predict Student Failure and Enabling Customized Educational Paths.
[10]. Maryam et al (2020). Major depressive disorder assessment via enhanced k-nearest neighbor method and signals. Physical and Engineering Sciences in Medicine, 43(3):1007–1018, 2020.
[11]. Qingxiang & Jiesen (2021) Application of Random Forest Algorithm in Physical Education.
[12]. Chao Pan, Cheng Shi, Honglang Mu, Jie Li, and Xinbo Gao. Eeg-based emotion recognition using logistic regression with gaussian kernel and laplacian prior and investigation of critical frequency bands. Applied Sciences, 10(5), 2020.
[13]. Pallavi Asthana & Bramah Hazela (2020) Applications of Machine Learning in Improving Learning Environment
[14]. Fethia Yahiaoui et al., (2023), The Impact of e-Learning Systems on Motivating Students and Enhancing Their Outcomes during COVID-19: A Mixed-Method Approach.
[15]. https://www.journalijar.com/upload-articles/