A Systematic Review on Student Performance Analysis Using Machine Learning

(IJCSNT) International Journal of Communication Systems and Network Technologies
ISSN Number: 2053-6283
Volume 14, Issue No. 2, 83–99
©The Author(s) 2025
https://journals.mirlabs.in/index.php/ijcsnt
DOI-10.18486/ijcsnt/14.2.008
IJCSNT

Priyanka Asthana[1], Manish Maheshwari[2]

Abstract

In educational system, the student performance is a core component which pays a significant role in the development of any school/college. The student performance handling is very crucial in learning process and it is one of the significant factors of learning. In the field of research educational data mining (EDM), by using the data knowledge education system can be improved. Advanced machine learning (ML) methods can predict student's performance with key features based on academic, behavioral, and demographic data. Significant works have predicted the student's performance based on the primary and secondary data sets derived from the student's existing data. These works have accurately predicted student's performance but did not provide the metrics as suggestions for improved performance. In this paper, we present the review of literature on the field of early student performance prediction using machine learning and also discuss machine learning techniques with their advantages and disadvantages.

Keywords

EDM, Student Performance, Machine Learning, CNN, Early Prediction

Received: 11 April 2025; **Revised:** 22 July 2025; **Accepted:** 20 August 2025;
Published: 31 August 2025;

[1, 2]Dept of Computer Science & Applications, Makhanlal Chaturvedi University, Bhopal, India.

Corresponding author:
Email: asthana.priyanka.1984@gmail.com

Introduction

Every aspect of our lives, including our schooling, has been impacted by the COVID-19 pandemic. The epidemic has led to a notable advancement in e-learning resources and digital platform learning through remote systems. For instance, as a result of this pandemic, we have observed that several nations have begun using e-learning platforms for instruction. The education system has encountered numerous challenges as a result of technological advancements, particularly in the areas of online content development and teaching and learning. As a result, the system believes that inadequate instruction has a negative impact on students' academic performance. Early on in the e-learning process, it might be challenging to identify the important variables. A variety of efficient tools have been developed to obtain information on e-learning and enhance students' academic performance[1–3]. These days, machine learning (ML) and artificial intelligence (AI)[4,5] are important in a number of fields, including educational data mining (EDM), the commercial sector, and the health sector. In order to assess and examine student academic performance in the field of education, educational data mining employs a variety of data mining techniques, including classification, clustering, time series analysis, support vector machines, and association rule mining. One of the key methods for extracting or revealing hidden patterns that aids the educational system in precisely assessing students' performance is the use of electronic data mining, or EDM.Machine learning makes it easier to detect kids that perform poorly academically, which benefits the growth and development of all educational systems by enabling early prediction of these children. The educational system employs a method known as education data mining (EDM) to obtain the vital information from students using data mining. This method aids in improving students' learning and experiences in the classroom. Several data mining methods have been applied to forecast different educational outcomes, such as performance, accomplishment, attainment, retention, and dropout rate[6]. In the subject of education, data mining techniques are incredibly helpful, especially when it comes to analyzing and forecasting students' academic performance.

Some feature selection and data preprocessing methods were used to improve the performance of the machine learning models. These include data normalization techniques, which simplify the input data and allow the machine learning algorithms to find predictive patterns, data encoding for non-numeric features, and testing datasets for null values. Using the random forest algorithm and the feature importance technique, the most pertinent features are chosen to be trained in the machine learning models. Mean absolute error (MAE) and root mean square error (RMSE) are two performance metrics used to assess the efficacy of the models[1,5]. This research study adds to the growing subject of machine learning in education by proposing a novel ML-based approach for forecasting semester GPA and illustrating the effectiveness of predictive analytics in improving student performance and achievement.

Review of Literature

Systematic assessments of electronic databases draw on a wide range of data sources. The student performance prediction system usage of machine learning to determine research gaps in education.

2.1 Source of data

The following methods have been implemented to gather data for this systematic review from various electronic databases:

- Google Scholar
- IEEE Xplore
- Web address for Elsevier Science Direct
- Web address for the Springer Link
- Web of Science

2.2 Related Work

Kendagor et al. (2023) developed a CNN (convolutional neural network) framework for the student performance prediction in relation to significant factor. This methodology employs the pre-process data which was collected from various universities, which was used to train and test the CNN model.The evaluation of model was done using accuracy, precision and recall parameter, to confirm its toughness andefficacy in predicting dropout student. The outcomes ofthis work can assist as an appreciated resource for futureresearch in this region. [7]

Blundoet al. (2023) anticipated the framework of an innovative process to train dropout students of the universities by selecting the reduced set of features. The idea of this process is to weekly classify those students who are going toward dropout or not with acceptable precision.In the reason of uncertainty, as we receive new data from the university, the decision of classification changed for the next week. This approach provides the awareness to the student for the course timelines and also take care them, and it also offers several benefits such as opportunity to build the real time educational decision support system (EDSS) which supports the decision by providing the sufficient information and other one is to preserve the available resources for the future use and wasting them with students who are classified as the dropout student. The last one is to predict the student behaviour and characteristics which helps in identifying the poor students by employing rule mining approach. [8]

Daozong, et al. (2023) anticipated an innovative approach for predicting the university student performance based on multi-feature fusion and attention mechanism. The anticipated model mainly focuses on the student historical academic records by analysing the relationship between courses and students and also extracted the essential features of university students. Together with, the attention mechanism mainly explore the association between diverse dimensional features. This work, collected a triplet set of associated courses and students' actual historical records through which data analysis is performed to show the correlation between courses and also authenticate the efficacy of different dimensional features. The experimentation of our anticipated and existing machine learning method is performed using the accuracy measuring parameter and found that the anticipated model gives better prediction and precision rate of 72.5%. [9]

Joseph Chi-Ho, et al.(2023) investigate the association between GCDA participation and academic success before and after higher education utilizing comprehensive and representative data (over 10,000 records) spanning three years. Triangulation of the data using DA and machine learning (ML) comes after the formulation and validation of hypotheses. Supervised learning is used to build the predictors of academic achievement and GCDA involvement, and characteristics to enhance predictions are assessed. Using novel techniques based on evolutionary algorithms and stacking in machine learning, we generate predictors. The effects of both the intensity and breadth of participation are also investigated. The results demonstrate that students' academic performance is positively impacted by their involvement in GCDA. Our cutting-edge techniques improve the likelihood of student participation. Our extensive data analysis, machine learning, and hypothesis validation research provides valuable insights into the construction of GCDAs.[10]

Alhazmi et al. (2023)gives a data-mining-based appraisal of student performance. In order to determine how early student performance affects the GPA, the study offers both clustering and classification algorithms. The research uses the dimensionality reduction mechanism for the clustering technique of the T-SNE algorithm to investigate the association between these parameters and GPAs. Academic achievement tests (AAT), general aptitude tests (GAT), and entrance scores from first-level courses are among the early indicators that are also taken into account. The study details trials on different machine learning models that forecast early student success for the categorization technique based on a range of criteria, including course grades and admission exam outcomes. We use multiple evaluation metrics to examine the models' quality. The results suggest that educational institutions can lower the odds of early student failure.[11]

Qin, Kuangyu, et al. (2023)provide a neural network model that combines a number of internal and external components, data-level sample synthesis, multi-classification cost-sensitive learning techniques, and other features to enhance teaching quality and management and provide early warning of student performance in higher education. According to experimental data, the model surpasses conventional machine learning techniques and may be used in teaching scenarios including a combination of online and offline instruction. It also has a higher accuracy than earlier prediction mechanisms that simply took certain academic traits into account.[12]

Farokhi, Soheila, et al. (2023) implies that the information contained in these structural links, as shown by an interaction graph, can improve the task's performance. In order to verify this, we create a special knowledge graph for a sizable MOOC dataset that will be made accessible to the scientific community at large. Moreover, we extract latent structural information embedded in the dataset's entity interactions by using graph embedding techniques. These methods can be applied to a variety of jobs and do not require ground truth labelling. Lastly, we improve the effectiveness of predictive machine learning models in student assignment grade prediction by merging entity specific characteristics, behavioural features, and extracted structural features. Our tests show that the prediction performance of downstream assessment tasks can be greatly enhanced by structural factors.[13]

Nabizadeh, Amir Hossein, et al.(2022)suggest a technique for estimating a student's final course grade based just on their performance information for the current semester. It helps course instructors identify student types early on in the course so they can better support them. It also helps students analyze how much work they need to put into the course to get the required grades. Our approach divides students into a number of groups at first according on the experience points (XP) they accrue over the course of a semester. Next, we create and add virtual students to the smaller clusters in order to estimate cluster size and balance clusters. Lastly, we use a feature selection strategy to exclude student traits that are not important. Then, using three distinct algorithms, we project their ultimate grades. We have used data from a course spanning nine years, with data from 679 students, to evaluate the effectiveness of our strategy with alternative ways. Only four weeks after the course began, our method achieved an average accuracy of 78.02%, outperforming all other methods, according to the data. It demonstrates that we can accurately forecast final marks, which may improve students' learning objectives.[14]

Bujang, SitiDianah Abdul, et al. (2021)offers a comprehensive analysis of machine learning techniques that improve prediction accuracy when predicting students' final grades in courses taken during the first semester. Two modules will be the subject of this paper. We first evaluate the accuracy performance of six widely used machine learning techniques: Decision Tree (J48), Support Vector Machine (SVM), Naïve Bayes (NB), K-Nearest Neighbor (kNN), Logistic Regression (LR), and Random Forest (RF), using a dataset of 1282 real student course scores. Second, in order to reduce the overfitting and misclassification results brought about by imbalanced multi-classification, we created a multiclass prediction model based on the oversampling Synthetic Minority Oversampling Technique (SMOTE) using two features selection approaches. The compiled results show that when the recommended model is combined with RF to produce a significant improvement, the maximum f-measure of 99.5% is reached. This proposed model yields comparable and promising results, and it has the potential to enhance the prediction performance model for imbalanced multiclassification in student grade prediction.[15]

Liu, Dong, et al. (2020)assume that each element affects student performance differently and thus improving prediction accuracy requires combining the features of exercise and student behavior. As a result, by combining the attention mechanism with the knowledge tracing model and making full use of the exercise and student behavior data, our research offers a novel framework for forecasting student success. More specifically, we first automatically capture feature representation using machine learning. Next, a fusion attention technique based on a recurrent neural network architecture is used to predict student performance. Extensive experiments on a real dataset show our method's effectiveness and practicality. In contrast to previous methods, ours achieves up to 98% accuracy.[16]

Halde et al. (2016) find social skill-based determinants of academic achievement utilizing knowledge discovery in databases (KDD) and use Random Forest classification methods to clean the data. A sociodemographic questionnaire and the Social Skills Scale (SSS) were administered in person to 93 general psychology students at Indoamerica University in order to gather data. The study's findings suggest that social skills, including the capacity to refuse contacts and voice disagreement, as well as gender, are related to indicators of academic success. These results point to the need for support program design, academic counseling, and social skill development in order to enhance academic achievement and future professional success. To sum up, the study offers a fresh viewpoint for those working in student welfare departments to raise academic achievement levels.[17]

Sahlaoui, Hayat, et al. (2021) proposes a method that, by improving performance and providing an explanation for why a student's performance is achieving a particular score, improves the prior student performance forecast. Machine learning methods were used to test and propose a prediction model. Our models perform better than other models created using the same dataset. The suggested model achieves an accuracy of over 98% by combining frameworks of data level and algorithm techniques, which is a 20.3% improvement over earlier work models. As a balancing technique for upsampling data, we use the default strategy of the synthetic minority oversampling method (SMOTE) to oversample all classes to the number of samples in the majority class. Furthermore, we utilize group methods. We use a simple grid search method provided by scikit-foundation to get the optimal model parameters. The robustness of the novel model is proven through hyperparameter tuning and ten-fold cross-validation. Additionally, a novel visual and intuitive approach is utilized to help determine the components that have the greatest influence on the score, supporting the interpretation and understanding of the entire model and providing a visual representation of feature attributions for the machine learning model at the observation level. As a result, SHAP values are an effective tool that ought to be included in the framework for predicting student performance. By acquiring the explanation and prediction generated by the experiment, teachers can identify students who are at risk early on and offer appropriate encouragement in a fortunate way.[18]

Sirait, Jaime Christ Bonar, et al. (2023)use a comprehensive strategy to address the skewed distribution of data by applying the Support Vector Machine – Synthetic Minority Oversampling Technique (SVM-SMOTE). We presented a novel method called Deep Feature Synthesis to further improve prediction skills. This sophisticated method produces 36 new features. Our study's conclusion highlights the random forest algorithm's superiority, particularly when combined with the well-thought-out use of Random Search to create a harmonious blend of Deep Feature Synthesis and SVM-SMOTE. Especially, our model obtains a remarkable 69% accuracy rate, along with 47% precision, 52% recall, and 47% F1-score.[19]

Dahroug et al. (2023) forecast the student's performance, a number of machine learning classifiers were used, including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Support Vector Machine (SVM), and Naïve Bayes (NB). In addition, a hybrid model was employed to lower the dimensionality of the data and enhance the classifiers' overall performance. To assess the suggested hybrid model, an actual dataset from the UCI machine learning library is used. According to the published results, CNN-RNN performs better than alternative techniques with an accuracy of 89.1%.[20]

Mohamed Mohsen Elsaid, et al. (2023)provides an approach for predicting student performance (SPP) that forecasts students' academic successes based on a range of factors, including demographic data, academic background, and behavioral tendencies. It does this by utilizing machine learning techniques. To guarantee a thorough and precise analysis, the suggested SPP approach makes use of a real-world dataset that has been gathered from several educational institutions. The first step of the suggested technique is data preparation, which involves organizing and cleaning the data in preparation for analysis. This procedure includes things like scaling the data, addressing missing numbers, and changing variables as needed. The most crucial features for the SPP model were chosen using the RFE selection technique. After testing several machine learning classifiers, the top performance was determined to be the linear regression technique. The evaluation's findings demonstrated that, with a mean absolute error (MAE) of 0.23 and a root mean square error (RMSE) of 0.29, the linear regression approach can be utilized to forecast student performance.[21]

Pallavi, et al. (2023) suggests the "Learning Coefficients" assessed by automated adaptive assessment based on trajectory analysis. Learning coefficients also give pupils quantitative measures to help them concentrate more on their academics and perform better in the future. Positive Pearson's coefficient correlation is used to determine the learning coefficients' dependence on other key features before choosing them as the primary characteristics for predicting students' success. The study also compares the effectiveness of regression-based machine learning models on the same dataset, including decision trees, random forests, support vector regression, linear regression, and artificial neural networks. Comparing linear regression to other models, the results indicate that it achieved the best accuracy of 97%.[22]

Raja, et al. (2023) Graph Convolutional Networks (GCN) and graph structures are used in Predicting Student Performance using Integrated Similarity Modeling (PRISM) to enhance the prediction of student performance. PRISM uses data on academic performance and personal attributes to create a graph in which each student is represented as a node and the edges indicate the relationships between them. Using this network structure as a guide, GCN is used to forecast pupils' academic success. An actual educational dataset is used to assess the suggested pipeline, and it is contrasted with conventional machine-learning techniques. The suggested GCN-based model performs better in predicting student performance, as shown by the experimental findings.[23]

Majid et al. (2023) analyze student behaviour patterns using data mining approaches to forecast academic performance.The results of this study indicate that a variety of parameters, including resource (page) views, activity gaps, grades from the previous semester, grades from prerequisite courses, and assessments of first-term examinations, significantly correlate with student success. This study can be used by educators and teachers to identify pupils who require additional support so they can provide it.[24]

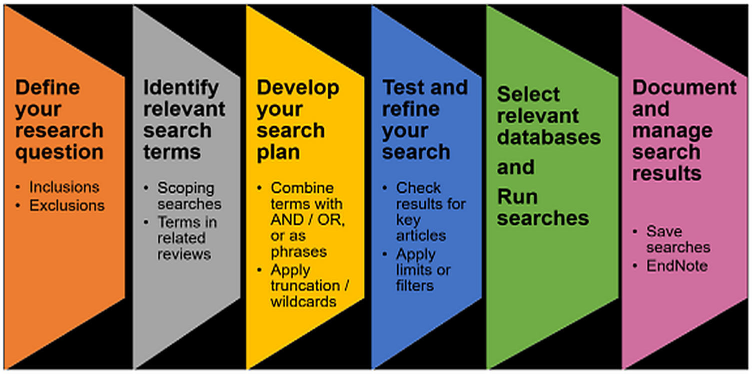

Figure 1. Flowchart for systematic literature review[31]

Table 1. Criteria for drafting systematic literature review questions[31]

Criteria	**Descriptions**
Population	Pre-tertiary and higher institutions of learning
Intercession	Approaches, algorithms and procedures employed in forecasts
Consequence	Pre-eminent performance indicators, top features or factors, as well as actual prediction strategies or methods
Perspective	Educational formations and Maximum performance accomplishment of Students

Through a thorough and comprehensive analysis of published literature spanning the years 2010 to 2023, we have added to the body of knowledge on the application of machine learning techniques to forecast student performance in the context of early detection and management (EDM). Prior research studies have focused on DM techniques generally, even if ML methods have demonstrated potential in forecasting student performance. Our research is distinctive in that we concentrate on machine learning techniques and algorithms, along with their application in forecasting academic achievement. Furthermore, we determined the characteristics and variables influencing the prediction models as well as the feature selection techniques and algorithms used in the literature. Lastly, we have examined the corrective measures that were implemented for those kids who were anticipated to be at-risk in the literature.

We are able to determine the new fields and ongoing uses of machine learning algorithms in predicting academic performance, at-risk kids, and retention through our review of the literature. This evaluation

Table 2. Early prediction of student performance using static and dynamic data

Approach	Algorithms	Attributes
Dynamic	NB, BN, DT, CR, ADTree, J48, RF, SMOTE, BM, SVM, NB, BN, and	Data on student achievement, reading assignments, and quizzes
Static	Item Response Theory (IRT), Adaboosting Tree, CART, ICRM2, SVM, NB, DT, ID3, DL, and kNN	Data on enrollment and demographics
Both	ICRM2 with ARTMAP, RF, CART, NB, KNN, DT, LR, SVM, DL, ID3, and	Details on pre-college admission and transcripts

reveals topics that need further investigation in addition to highlighting the literature's primary areas of concentration. Furthermore, the paper offers an examination of the most well-known machine learning methods, platforms, algorithms, and tools for forecasting student success.

Machine Learning

Within the field of artificial intelligence, machine learning (ML) advances the notion that machines can self-learn how to tackle a given problem if they are provided with appropriate data[1]. Through the use of sophisticated mathematical and statistical techniques, machine learning (ML) equips robots to carry out autonomously intellectual tasks that have historically been completed by humans. In the hopes of offloading various jobs related to communication network design and operation to machines, the concept of automating complicated operations has sparked a lot of attention in the networking industry.

3.1 Classification of Machine Learning

Supervised Learning

Algorithms for supervised machine learning are ones that require outside help. The train and test datasets are separated from the input dataset. There is an output variable in the train dataset that needs to be classed or forecasted. Every algorithm picks up some sort of pattern from the training dataset and uses it to make predictions or classify data from the test dataset[25]. Fig. 1 shows the supervised machine learning algorithms' workflow.

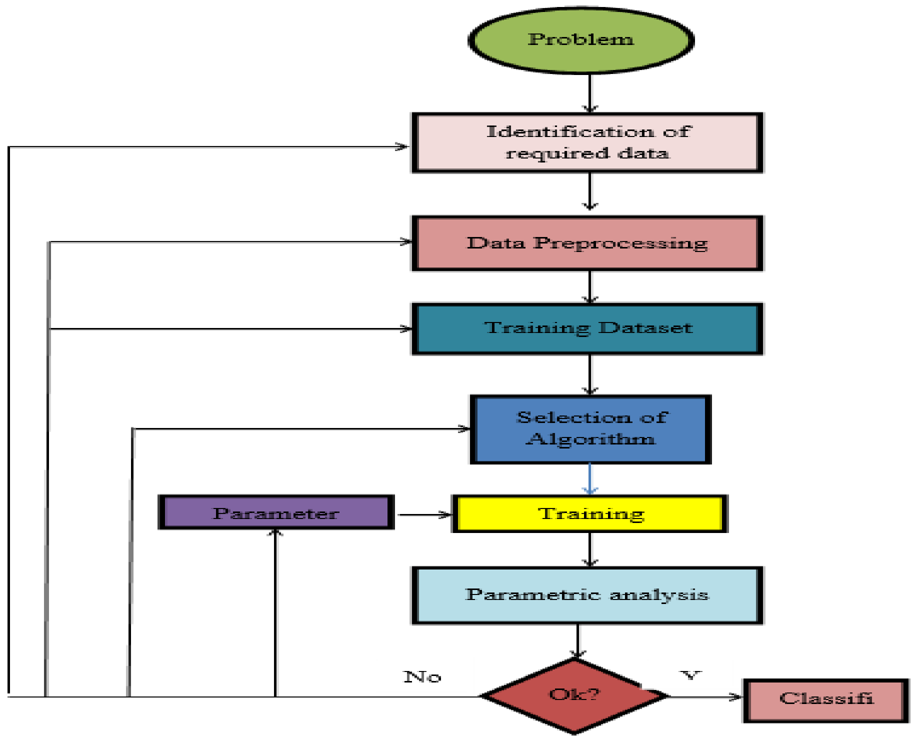

Figure 2. Workflow of supervised machine learning

Benefits of Supervised Learning:

- Explicit Feedback: Labeled data, which offers explicit feedback on the model's predictions, is a prerequisite for supervised learning. The training and enhancement of the model benefits greatly from this input.
- Predicted Accuracy: When trained on representative, high-quality data, supervised learning models can attain a high level of predicted accuracy. They work well for tasks like regression and classification.
- Generalization: Suitable for real-world applications, well-trained supervised models are able to apply their expertise to generate precise predictions on novel, unseen data points.
- Interpretability: Models produced by some supervised learning algorithms, such as decision trees and linear regression, enable users to comprehend the connections between input characteristics and predictions.
- Broad Range of Applications: Supervised learning has applications in computer vision, natural language processing, healthcare, finance, and other fields.

- Availability of Tools and Libraries: There are numerous tools, libraries (e.g., sci-kit-learn, TensorFlow, PyTorch), and resources available for implementing and experimenting with supervised learning algorithms.

Drawbacks of Supervised Learning:

- Requirement for Data Labeling: Large datasets can be costly and time-consuming to get labeled data, which is necessary for supervised learning.
- Restricted to Labeled Data: The model's capacity to handle unfamiliar or unexpected scenarios is limited because it can only predict outcomes based on data that is comparable to what it was trained on.
- Bias and Noise in Labels: If labeled data has errors or biases, the model may pick up on these biases and continue to use them, which could result in unfair or incorrect predictions.
- Overfitting: When a model learns training data too well, it runs the danger of collecting noise rather than underlying patterns. To lessen this, regularization techniques are frequently needed.
- Feature Engineering: One of the most important steps in creating efficient supervised learning models is choosing and designing pertinent features. Inadequate feature selection may result in less than ideal performance.
- Scalability: Using big datasets to train intricate models can be computationally costly and time-consuming, requiring a significant amount of processing power.
- Limited to Labeled Data Distribution: Supervised models may not function properly when presented with data from a different distribution because they are limited by the labeled data distribution.
- Privacy Issues: The usage of labeled data in some applications may give rise to privacy issues since it may reveal personally identifiable information.
- Unbalanced Data: trained models may find it difficult to correctly forecast minority classes when faced with imbalanced datasets (such as those pertaining to uncommon disease identification).
- Concept Drift: Over time, there could be a shift in the way input features relate to the target variable. Retraining supervised models frequently may be necessary to adjust to these modifications.

Unsupervised Learning

Few features are learned from the data by the unsupervised learning algorithms. It recognizes the class of the data when it is introduced by using the previously learned features. Its primary applications are in feature reduction and clustering.

Benefits of Unsupervised Learning:

- Finding Hidden Patterns: Data may have hidden structures and patterns that are not immediately visible upon manual inspection. These can be found by unsupervised learning techniques. This may result in insightful discoveries and a better comprehension of the underlying data.
- Data exploration: An effective technique for exploratory data analysis is unsupervised learning. It helps you find patterns, outliers, and possible areas of interest by enabling you to view and summarize large, complex datasets.
- Dimensionality Reduction: High-dimensional data can be made less dimensional while retaining valuable information by using techniques like Principal Component Analysis (PCA) and t-SNE. This may result in modeling and visualization that is more effective.

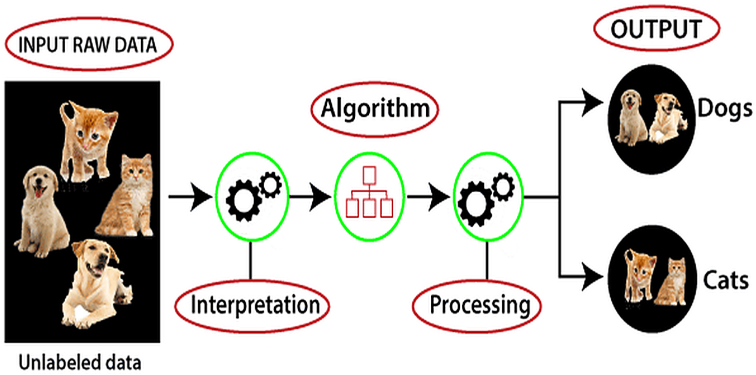

Figure 3. Unsupervised learning

- Clustering: The act of assembling comparable data points into groups can be accomplished through unsupervised learning. Activities such as anomaly detection, picture segmentation, and customer segmentation can benefit from clustering.
- Feature Engineering: In later supervised learning tasks, unsupervised learning can produce new features or data representations. Word embeddings acquired using unsupervised methods, for instance, can enhance models for natural language processing.
- Anomaly Detection: Unsupervised learning is useful for applications like fraud detection and quality control because it can identify anomalies or strange patterns in data.
- Lower Labeling Costs: Labeling data for supervised learning can occasionally be costly or time-consuming. By producing features or insights that improve the performance of supervised models, unsupervised learning can contribute to the reduction of the demand for labeled data.

Drawbacks of Unsupervised Learning:

- Absence of Ground Truth: Not having labeled data or a ground truth to assess the caliber of outputs is one of the primary drawbacks of unsupervised learning. This makes evaluating the precision of pattern recognition or clustering difficult.
- Subjectivity: Depending on the algorithm, hyperparameters, and preprocessing techniques selected, the outcomes of unsupervised learning can be arbitrary. There is no one "right" answer; rather, different strategies may provide different results.
- Interpretability: It can be challenging to understand certain unsupervised learning models, particularly deep learning models. It can be difficult to comprehend why a model recognized certain patterns or reached a specific judgment.
- Computationally Intensive: When working with huge datasets or high-dimensional data, certain unsupervised learning algorithms can be computationally intensive. This may call for a large amount of processing power.

- The “curse of dimensionality” states that unsupervised learning may be difficult with high-dimensional data since many algorithms’ distance measurements lose their efficacy in high-dimensional domains. In order to solve this problem, dimensionality reduction techniques are frequently required.
- Difficulty in Evaluation: Since there might not be a definite ground truth to compare against, evaluating the quality of clustering or pattern finding might be difficult. Unsupervised learning evaluation measures are frequently application-specific.

Semi-Supervised Learning

The power of both supervised and unsupervised learning can be combined with semi-supervised learning methods. When unlabeled data is already available and obtaining labeled data is a laborious procedure, it can be beneficial in some machine learning and data mining domains[26]. Semi-supervised learning falls into a lot of different types[27]. Below is a discussion of a few of them:

1) Models that generate: One of the earliest semi-supervised learning techniques, generative models assume a structure such as $p(x,y) = p(y)p(x|y)$, where $p(x|y)$ represents a mixed distribution, such as in Gaussian mixture models. The mixed components can be identified in the unlabeled data. For each component, one labelled example is sufficient to verify the distribution of the mixture.

2) Self-Training: A classifier is trained using some labelled data in self-training. Then, unlabeled data is given into the classifier. In the training set, the predicted labels and the unlabeled points are combined. The process is then carried out once more. The term “self-training” comes from the fact that the classifier is self-learning.

3) Transductive SVM: An expansion of SVM is the transductive support vector machine, or TSVM. Both labelled and unlabeled data are taken into account in TSVM. In order to maximize the margin between the labelled and unlabeled data, it is utilized to label the unlabeled data. It is NP-hard to find a precise answer using TSVM.

Benefits of Semi-supervised learning techniques

- Enhanced accuracy: By utilizing both labelled and unlabeled data, the semi-supervised learning approach can enhance accuracy.
- Lower labelling costs: Labelling is a time-consuming and expensive procedure. However, by utilizing semi-supervised learning, we may lower the labelling costs by combining supervised and unsupervised learning strategies.
- Using both labelled and unlabeled data: The semi-supervised learning approach makes use of both labelled and unlabeled data, making it an effective use of data.

Drawbacks of Semi-supervised learning techniques

- Not totally dependable: Because semi-supervised learning relies on predictions, its dependability may be questionable. There’s also a chance for mistakes.
- Choosing the right unlabeled data is crucial because the wrong choice could impair the model’s performance.
- Accurate labelling: It’s critical to accurately classify datasets that have been labelled. To do that, knowledgeable data engineers are needed.

Reinforcement Learning

With reinforcement learning, decisions are made based on the best course of action to pursue in order to get a better outcome. The learner is unaware of the proper path of action prior to being presented with a circumstance. Future actions and circumstances may be affected by the learner's activity. Trial-and-error search and delayed outcome are the only two requirements for reinforcement learning [28]. Fig. 3 shows the general model [29] for reinforcement learning. The Reinforcement Learning Model [29].The agent in the figure receives the input function (I), the current state (s), the state transition (r), and the input (i) from the environment. The agent creates a behaviour B and executes an action that results in an outcome based on these inputs.

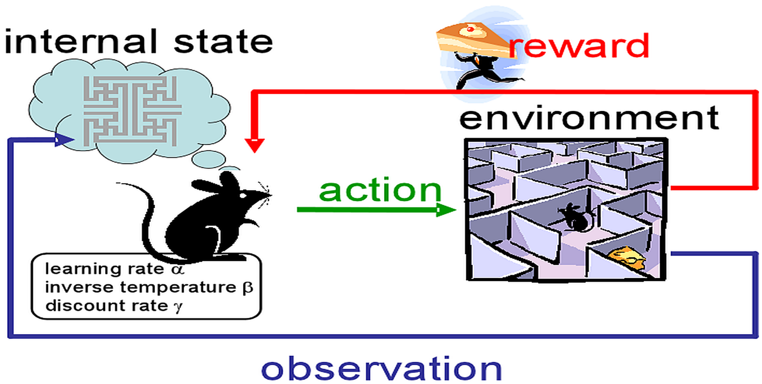

Figure 4. Reinforcement Learning

Benefits of Reinforcement Learning

- Highly complicated issues that traditional methods are unable to tackle can be resolved with reinforcement learning.
- Long-term outcomes are quite hard to produce, hence this strategy is chosen.
- There are many similarities between this learning paradigm and human learning. Thus, it is nearly attained perfection.
- The mistakes made during training can be fixed by the model.
- There is very little likelihood of the same issue recurring after the model has fixed it.
- It can produce the ideal model to address a specific issue.
- Reinforcement learning techniques can be used by robots to teach them to walk.
- Without a training dataset, it will inevitably pick up knowledge from its experience.
- In many tasks, reinforcement learning models can perform better than people. In March 2016, DeepMind's AlphaGo program—a reinforcement learning model—bested world go champion Lee Sedol.
- The goal of reinforcement learning is to maximize a model's performance by helping it to exhibit its optimum behavior in a particular setting.
- It can be helpful in situations where engaging with the environment is the only method to gather knowledge about it.

- Algorithms for reinforcement learning keep the balance between exploitation and exploration. The act of exploring involves attempting new things to determine whether they are superior to others that have already been attempted. The practice of attempting what has historically worked best is known as exploitation. This balance is not accomplished by other learning techniques.

Drawbacks of Reinforcement Learning

- Although there are numerous ways in which reinforcement learning is flawed as a framework, it is precisely this flaw that lends it use.
- An excess of states resulting from reinforcement learning may lessen the outcomes.
- Using reinforcement learning to solve easy issues is not recommended.
- A large amount of data and processing power are required for reinforcement learning. It is obsessed with data. That is why it works really well in video games because one can play the game again and again and again, so getting lots of data seems feasible.
- Reinforcement learning operates under the false assumption that the world is Markovian. The Markovian model depicts a series of hypothetical occurrences where the likelihood of each event is only dependent on the condition reached in the preceding event.
- For genuine physical systems, reinforcement learning is severely constrained by the curse of dimensionality.
- The curse of real-world samples is another drawback. The Wikipedia definition of the curse of dimensionality is “a variety of phenomena that arise when analyzing and organizing data in high-dimensional spaces that do not occur in low-dimensional settings such as the three-dimensional physical space of everyday experience.” Think about the scenario of robots learning, for instance. Robot hardware typically costs a lot of money, needs constant maintenance, and is prone to wear and tear. The expense of fixing a robot system is high.
- Rather than relying solely on reinforcement learning, we can combine it with other methods to address a number of its challenges. Reinforcement learning combined with deep learning is one such combo.

Multitask Learning

The basic objective of multitask learning is to improve the performance of other students. When multitask learning algorithms are used on a task, the process by which the problem was solved or the specific conclusion reached is remembered. The algorithm then applies these processes to solve new tasks or problems of a similar nature. An alternative name for this process of one algorithm assisting another is an inductive transfer mechanism. Students can learn concurrently—rather than individually—and considerably more quickly if they share their experiences with one another[30].

Ensemble Learning

Ensemble learning is a specific kind of learning that occurs when multiple individual learners are united to form one learner. The individual learner could be a neural network, decision tree, Naïve Bayes, etc. Since the 1990s, group learning has gained popularity. It has been noted that a group of learners almost always performs a task more effectively than a single learner[31].

Conclusion

The approach for early predictionof student performance (SPP), which uses machine learning methods to forecast students' academic achievement, has been presented in this study. In this paper, we give an outline of the literature on machine learning algorithms for predicting student performance. After looking through a number of literatures, we discovered that some of the current methods—like decision trees, linear regression, SVM, and Naïve Bayes classifiers, among others—are not more effective in predicting student performance. Thus, in our upcoming work, we will employ the ensemble approach of machine learning to forecast student performance early on.

References

[1] Hamsa, Hashmia, Simi Indiradevi, and Jubilant J. Kizhakkethottam. "Student academic performance prediction model using decision tree and fuzzy genetic algorithm." Procedia technology 25 (2016): 326-332..

[2] Francis, Bindhia K., and SuvanamSasidharBabu. "Predicting academic performance of students using a hybrid data mining approach." Journal of medical systems 43 (2019): 1-15.

[3] Yağcı, Mustafa. "Educational data mining: prediction of students' academic performance using machine learning algorithms." Smart Learning Environments 9.1 (2022): 11.

[4] Zawacki-Richter, Olaf, et al. "Systematic review of research on artificial intelligence applications in higher education–where are the educators?." International Journal of Educational Technology in Higher Education 16.1 (2019): 1-27.

[5] Sekeroglu, Boran, KamilDimililer, and Kubra Tuncal. "Student performance prediction and classification using machine learning algorithms." Proceedings of the 2019 8th International Conference on Educational and Information Technology. 2019.

[6] E. Alyahyan and D. Düştegör, "Predicting academic success in higher education: Literature review and best practices," Int. J. Educ. Technol. Higher Educ., vol. 17, no. 1, pp. 1_21, Dec. 2020.

[7] Kendagor, Valentine Jemutai, Geoffrey Chemwa, and Wilson Cheruyoit. "A Convolutional Neural Network Framework for Predicting Student Drop Out in Universities." 2023 IEEE AFRICON.IEEE, 2023.

[8] Blundo, Carlo, Vincenzo Loia, and Francesco Orciuoli. "A time-aware approach for MOOC dropout prediction based on Rule Induction and Sequential Three-Way Decisions." IEEE Access (2023).

[9] Sun, Daozong, et al. "A University Student Performance Prediction Model and Experiment Based on Multi-Feature Fusion and Attention Mechanism." IEEE Access (2023).

[10] So, Joseph Chi-Ho, et al. "Analytic study for predictor development on student participation in generic competence development activities based on academic performance." IEEE Transactions on Learning Technologies (2023).

[11] Alhazmi, Essa, and Abdullah Sheneamer. "Early Predicting of Students Performance in Higher Education." IEEE Access 11 (2023): 27579-27589.

[12] Qin, Kuangyu, et al. "Early Warning of Student Performance with Integration of Subjective And Objective Elements." IEEE Access (2023).

[13] Farokhi, Soheila, et al. "Enhancing the Performance of Automated Grade Prediction in MOOC using Graph Representation Learning." 2023 IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA).IEEE, 2023.

[14] Nabizadeh, Amir Hossein, et al. "Early Prediction of Students' Final Grades in a Gamified Course." IEEE Transactions on Learning Technologies 15.3 (2022): 311-325.

[15] Bujang, SitiDianah Abdul, et al. "Multiclass prediction model for student grade prediction using machine learning." IEEE Access 9 (2021): 95608-95621.

[16] Liu, Dong, et al. "Multiple features fusion attention mechanism enhanced deep knowledge tracing for student performance prediction." IEEE Access 8 (2020): 194894-194903.

[17] Halde, Radhika R., Arti Deshpande, and Anjali Mahajan. "Psychology assisted prediction of academic performance using machine learning." 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT).IEEE, 2016.

[18] Sahlaoui, Hayat, et al. "Predicting and interpreting student performance using ensemble models and shapley additive explanations." IEEE Access 9 (2021): 152688-152703.

[19] Sirait, Jaime Christ Bonar, et al. "Predicting Students Performance Using Data Mining Approach (Case Study: IT Del)." 2023 IEEE International Conference on Data and Software Engineering (ICoDSE).IEEE, 2023.

[20] Ali, Donia Ashraf, Mohamed Aborizka, and Ahmed Dahroug. "Prediction of Student Performance by Using Machine Learning Techniques." 2023 4th International Conference on Artificial Intelligence, Robotics and Control (AIRC).IEEE, 2023.

[21] Khoudier, Mohamed Mohsen Elsaid, et al. "Prediction of student performance using machine learning techniques." 2023 5th Novel Intelligent and Leading Emerging Sciences Conference (NILES).IEEE, 2023.

[22] Asthana, Pallavi, et al. "Prediction of Student's Performance With Learning Coefficients Using Regression Based Machine Learning Models." IEEE Access (2023).

[23] Lakshmi, P. Raja, et al. "PRISM: Predicting Student Performance using Integrated Similarity Modeling with Graph Convolutional Networks." 2023 12th International Conference on Advanced Computing (ICoAC).IEEE, 2023.

[24] Khan, Majid, et al. "Utilizing Machine Learning Models to Predict Student Performance from LMS Activity Logs." IEEE Access (2023).

[25] X. Zhu, A. B. Goldberg, “Introduction to Semi – Supervised Learning”, Synthesis Lectures on Artificial Intelligence and Machine Learning, 2009, Vol. 3, No. 1, Pages 1-130

[26] X. Zhu, “Semi-Supervised Learning Literature Survey”, Computer Sciences, University of Wisconsin-Madison, No. 1530, 2005

[27] R. S. Sutton, “Introduction: The Challenge of Reinforcement Learning”, Machine Learning, 8, Page 225-227, Kluwer Academic Publishers, Boston, 1992

[28] L. P. Kaelbing, M. L. Littman, A. W. Moore, “Reinforcement Learning: A Survey”, Journal of Artificial Intelligence Research, 4, Page 237-285, 1996

[29] R. Caruana, “Multitask Learning”, Machine Learning, 28, 41-75, Kluwer Academic Publishers, 1997

[30] D. Opitz, R. Maclin, “Popular Ensemble Methods: An Empirical Study”, Journal of Artificial Intelligence Research, 11, Pages 169-198, 1999.

[31] K. RATHI, V. SHARMA, S. GUPTA, A. BAGWARI and G. S. TOMAR, “Home Appliances using loT and Machine Learning: The Smart Home,” *2022 14th International Conference on Computational Intelligence and Communication Networks (CICN)*, Al-Khobar, Saudi Arabia, 2022, pp. 329-332, doi: 10.1109/CICN56167.2022.10008294.

[32] Issah, Iddrisu, et al. “A systematic review of the literature on machine learning application of determining the attributes influencing academic performance.” Decision Analytics Journal (2023): 100204.