IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

Student Academic Monitoring System

Prof. Vanashri.N. Sawant[1], Aniruddh Salunkhe[2], Maitreyee Patil[3], Vedant Yadav[4], NishitVetal[5]
Professor, Department of Computer Engineering[1]
Students, Department of Computer Engineering[2-5]
NBN Sinhgad Technical Institute Campus, Pune, India

Abstract: *The increasing complexity of educational data demands advanced analytical methods beyond traditional evaluation metrics. This study investigates the application of Educational Data Mining (EDM) to analyze and predict students' academic performance effectively. It combines clustering methods, particularly an enhanced K-means algorithm, and deep learning techniques like Convolutional Neural Networks (CNN) to provide a comprehensive evaluation framework. The proposed approach focuses on improving the accuracy of performance prediction by determining optimal clustering numbers and using labeled data for deep learning. Results demonstrate significant improvements in identifying at-risk students and enhancing educational decision-making.*

Keywords: Academic records, Attendance tracking, HTML (Hyper Text Markup Language) ,PHP CSS, JQuery, Bootstrap, Node.js, Course registration. component, formatting, style, styling, CNN, EDM

I. INTRODUCTION

Educational Data Mining (EDM) has emerged as a critical field, focusing on extracting meaningful patterns from vast datasets to understand student behavior and performance. Traditional evaluation methods, relying solely on absolute scores, often fail to capture the nuances of student learning. These methods may overlook factors like course difficulty or varying grading standards. As a result, there is a growing need for a more robust, data-driven approach to evaluate and predict academic outcomes.

This study aims to address this gap by integrating clustering algorithms with deep learning techniques to analyze student performance. The research specifically explores:

1. Objective determination of clustering numbers to group students accurately based on performance.
2. Development of a predictive model using CNN to forecast future academic outcomes and provide early warnings for at-risk students.

II. LITERATURE SURVEY

1. Descriptive Models of EDM:

Students can be grouped based on the performance by the help of K-means clustering and identified patterns for high as well as low achievers among them. However, the optimal choice of number of clusters has to be specified. The improved techniques try to fine-tune clustering for higher accuracy and handling noisy data.

2. Predictive Models of EDM:

Predictions regarding academic outcomes are also done by using machine learning models, especially CNN`s. These deep learning methods are very good at handling large volumes of data, though they are often criticized to be "black boxes." Current efforts in explainable AI seem to improve model transparency.

3. Hybrid Approaches in EDM:

Hybrid models combine clustering on student group with CNN`s for predictive performance, thus absorbing unsupervised as well as supervised learning. These combinations have demonstrated better accuracy rates in the prediction of at-risk students. This integration makes predictions even more reliable and overarching.

4. Challenges in Current Research:

Data quality is mostly low, deep models have poor interpretability, and generalizability is limited. Predictive ability might be compromised due to inconsistent or missing data as complex models are not transparent. All the aforementioned must be overcome by EDM if these tools are going to be used more frequently in educational contexts.

5. Emerging Trends and Future Directions:

Current trends in Edm include multi-source data and NLP usage in student behavior and feedback analysis. Future studies involve adaptive learning systems for personal education on the respective basis of students' target values. This shift aims to enrich the outcome of student performance through real-time analysis improvement.

III. METHODOLOGY

1. Preprocessing the Data:

Cleaning: Incomplete records were dropped, missing data were handled by mean imputation, and duplicates were eliminated for quality assurance.

Integration: Combining course information based on relevance to reduce redundancy and averaging scores for semester-wise courses.

Transformation: Scores changed to standard for analytical purposes, data types transformed to make them compatible with algorithms, Reduction: Filtering by non-essential features like credit hours while giving importance to relevant academic scores and student identifiers.

2. Clustering Analysis:

An improved version of K-means is applied for cluster-based grouping of students based on their performances. A novel statistical measure, Rk2, for the choice of the best number of clusters (k) has been proposed. This statistical measure balances intra-cluster compactness with inter-cluster separation and provides more objective and accurate clustering results than earlier methods.

3. Discriminant Analysis:

Use Bayesian Discriminant Analysis, by which one may get the posterior probabilities of every data point as such evidence that these clusters created by the K-means algorithm are reliable enough to be used in predictions.

4. Deep Learning with CNN:

To train the CNN model, we utilize the labeled data of the clustering analysis. For our model architecture, we include several convolutional layers that extract features, pool layers that reduce dimensions, and finally, the full connection layers with final classification. Use of Activation Functions: In order to increase the models capability of dealing with nonlinearity, we use ReLU. The output of the last layers uses Softmax for obtaining probabilities in class prediction.

5. Evaluation Metrics:

Accuracy, precision, recall, and F1-score are used to evaluate the performance of the model. Cross-validation is applied using random hold-out with 80% as training and 20% as testing or shuffle 5-fold cross-validation to assure the generalizability of the model and avoid overfitting.

IV. CHALLENGES

1. Issues of Descriptive Models:

The issue with most descriptive models such as k-means clustering is that they are vulnerable to the 'best' number of clusters, which may lead to mistakes in identification. the models are highly sensitive to initial data points from where they work. Such models with noisy and incomplete data may continue to be challenging in the quality of the patterns identified. such models may not tend to capture deeper contextual factors that influence the performance of the students.

2. Limitations of Predictive Models:

predictive models, in particular deep learning techniques such as CNN`s, require extremely large data bases of high quality to train appropriately. other than these suffer education data from missing or inconsistent values. these models also create problems related to transparency in decisions made in case these models are kept opaque. deep learning models can thus prove less appealing to educators who would favor more interpretable instruments.

3. Issues With Hybrid Methods:

Although hybrid models that amalgamate clustering and deep learning have better performance, they inherit the flaws of both methods. the number of clusters to use is still unknown, and clustering errors could spill over into the predictive stage. generally, the combination between unsupervised and supervised learning calls for proper calibration to further maintain accuracy and computational efficiency. computationally, they require much more resources than most educational institutions can afford.

4. Issues With Data Quality and Interpretability:

The educational data often contains missing values, duplicate entries, and inconsistent entries. thus, poor quality of data can increase the chances of inaccuracy not only in descriptive models but also in predictive models. another aspect is that the interpretation of models like CNN`s and ensemble methods have really become complicated and are very hard to understand for educators, so they might not be able to understand the actual decision-making process which is behind a prediction. lack of transparency may also pose limitations to the adoption of these advanced data mining tools in classrooms.

5. Issues with Emerging Trends and Future Directions:

all of this would call for state-of-the-art data processing capabilities; the integration of multi-modal data sources and the application of nlp in educational analysis is quite demanding. for example, combining data types, such as clickstream data and text feedback, presents challenges of standardization and cleaning the dataset for analysis. adaptive learning systems that demand tremendous computational power and sophisticated algorithms require real-time data processing capabilities. privacy and security of student data is also becoming a great concern while using such advanced analytics.

V. FUTURE SCOPE

The Student Academic Monitoring System has significant scope for future progress. It can have newer types of data sources, psychological tests or even extracurricular activities that can be added to monitor the student more holistically. Adding predictive analytics can make it possible to provide a real-time insight into the student progress by incorporating algorithms such as Support Vector Machines or Naïve Bayes. Also, with an integration of web-based or mobile interface, students, parents, and teachers will be able to view performance on any device

This system may also be developed for working environments, which will help organizations assess employee performance and spots where they need improvement.

VI. CONCLUSION

The Student Academic Monitoring System has the ability to identify at-risk students, which aids in helping that individual progress into success in academics. The mechanism assists the students and faculty by monitoring the key performance indicators from time to time. However, it's not limited to the monitoring of academic performances. Instead, it looks forward to creating a collaborative environment by permitting teachers and parents to offer interventions. Although the system has disadvantages, such as a dependence on curricula, which may require teaching to be inflexible in most cases, it can easily adapt and scale to almost any educational and organizational requirement, hence leading to a more supportive and proactive approach toward student development.

REFERENCES

[1]. Huirong Yang, Wenjie Zhang. Data mining in college student education management information system. International Journal of Embedded Systems, 2022, 15(3):279-287.

[2]. Can Yi. The information association model of college students' archives based on a cloud computing management platform. International Journal of Information Technology and Management, 2022, 21(1):30-45.

[3]. Bansah Abednego Kofi, Darko Agyei Douglas. Perceived convenience, usefulness, effectiveness and user acceptance of information technology: evaluating students’ experiences of a Learning Management System. Technology, Pedagogy and Education, 2022, 31(4):431-449.

[4]. Md. Tofail Ahmed, Md. Humaun Kabir and Sujit Roy. Web Based Student Registration and Exam Form Fill-up Management System for Educational Institute. International Journal of Information Engineering and Electronic Business(IJIEEB), 2022, 14(2):47-62.

[5]. Lv Y. Student Behavior Analysis System Based on Kmeans Algorithm Under the Background of Smart Campus. In2022 International Conference on Computer Network, Electronic and Automation (ICCNEA) 2022 Sep 23 (pp. 174-177). IEEE.

[6]. Gupta S, Agarwal J. Machine Learning Approaches for Student Performance Prediction. In2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) 2022 Oct 13 (pp. 1-6). IEEE.

[7]. Sudais, M., Safwan, M., Khalid, M. A., & Ahmed, S. (2022). Students’ Academic Performance Prediction Model Using Machine Learning. Research Square; 2022. DOI: 10.21203/rs.3.rs-1296035/v1.