International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

EARLY PREDICTING OF STUDENTS PERFORMANCE INHIGHER EDUCATION

BANDARI RAVI
Assistant Professor, Guru Nanak Institute of Technology, CSE Department, Hyderabad

ABSTRACT: Evaluating students' learning performance is a fundamental aspect of evaluating any educational institution. When addressing challenges related to the learning process, student performance is critical, and it is one of the key factors used to quantify learning outcomes. The topic of research known as educational data mining (EDM) has grown out of the potential to leverage data knowledge to enhance educational systems. EDM is the development of methods to analyze data collected from educational environments, enabling a more complete and precise understanding ofstudents and the enhancement of their educational results. Evaluating the students' learning results is a crucial part of evaluating any educational institution. One of the key variables used to quantifylearning outcomes is student performance, which is significant when addressing problems with thelearning process. The field of research known as educational data mining, or EDM, was born out of the potential to leverage data knowledge to enhance educational systems. EDM is the process of developing methods for evaluating information obtained from educational environments. This makes it possible to learn more precise and in-depth information about students and enhances theiracademic achievement.

I. INTRODUCTION

Education is an important element and plays a significant role in our society. Information and communication technology has affected many fields of research, specifically in the education field. For example, as we have seen in many countries used various e-Learning environments due to therecent pandemic COVID19. A higher education institution considers the academic performance ofstudents as one of the most important issues regarding presenting quality education to its students. Understanding the significant factors in student's performance at early stage of their education is complex. Various effective tools have been used to overcome the students' performance challengesin academia. However, these tools may not be easy to generalize in all circumstances of education.In the recent years, with the advances of the application of technologies to forecasting students' performance, there are still gaps to be filled in order to analysis and improve the accuracy of students performance using new features and data mining methods and present both clustering and classification techniques to identify the impact of students performance at early stage with on theGPA. The learning process includes a lot of student performance. Identifying students who are more likely to have poor academic success in the future requires making predictions about studentperformance. If the data has been transformed into knowledge, it may be useful and used in predictions. As a result, the information might improve the quality of education and learning and help students in achieving their academic objectives. Data mining techniques are used in the studyarea known as educational data mining (EDM) to analyze information derived from educational backgrounds. EDM implementation also aids in the planning of strategies for raising student performance. As a result, it will improve teaching and learning and the students' experience in theeducational institution. Academic success is important because it is strongly linked to the positiveoutcomes we value. One of the academic success factors is the academic students' performance inthe college or university. The cumulative academic achievement for each student still indicates thesuccess of every college or university. Also, the other factors we can use in analysis and predict the academic students' performance are aptitude test, GPA of secondary school and the name of the school which the student graduated from. We believe that the performance of the students in

the first year in college can be used as a factor to predict the performance of student in the rest ofyears of his/her studies. These factors lead to early remedy for students and take actions to improvestudent per for mince Artificial intelligence techniques have been applied on educational data to reveal the significant reasons behind student performance. The contributions of the paper are as follows:

We propose a framework for predicting students' performance using student's academicperformance and his/her social relationships features.

We use admission scores, his/her first level courses scores and academic achievement test (AAT)and general aptitude test (GAT).

We explore a new way of using admission, his/her first level courses scores, and AAT and GATbyt SNE dimensionality reduction. To the best of our knowledge this attempt is a first of its kind to use features from both admission scores and his/her first level courses scores to early predict student's performance using machine learning.

We also explore a new way of using increasing the threshold of relocating which is to compute the absolute difference between a grade and following grade after or before.

We use a state-of-the-art classification model to evaluate the effectiveness of our proposed idea.We organize the paper as follows: Section II provides the literature related to students performanceprediction techniques used in the field of Education. Section III provides details of the used datasetincludes data characterization and correlations. Section IV provides the research methodology followed by the paper. We evaluate and analysis our proposed method and report findings in Section V. Finally, we conclude our work in Section VI.

II. LITERATURE REVIEW

F. Bonomi, R. Milito, J. Zhu, and S. Addepalli discussed that Fog computing extends the Cloud Computing paradigm to the edge of the network, thus enablinga new breed of applications and services. Defining characteristics of the Fog are: a) Low latency and location awareness; b) Wide-spread geographical distribution; c) Mobility; d) Very large number of nodes, e) Predominant role of wireless access, f) Strong presence of streaming and realtime applications, g) Heterogeneity. In this paper we argue that the above characteristics make theFog the appropriate platform for a number of critical Internet of Things (IoT) services and applications, namely, Connected Vehicle, Smart Grid, Smart Cities, and, in general, Wireless Sensors and Actuators Networks (WSANs).

I.-D. Filip, F. Pop, C. Serbanescu, and C. Choi, explained that Motivated by the high-interest in increasing the utilization of no general purpose devices in reaching computational objectives with a reduced cost, we propose a new model for scheduling microservices over heterogeneous cloud-edge environments. Our model uses a particular mathematical formulation for describing an architecture that includes heterogeneous machines thatcan handle different microservices. Since any new model asks for an early risk-analysis of the solution, we improved the Clouds simulation framework to be suitable for an experiment that includes that kind of systems. In this paper, we discuss two examples of real-life utilizations of ourproposed scheduling architecture. For an objective appreciation of the first example, we also include some experimental results based on the developed simulation tool. As a result of our interpretation of the experimental results we find out that some very simple scheduling algorithmsmay outperform some others in given situations that are frequently present in cloud-edge environments when we are using a micro service-oriented approach.

A. S. Gaur, J. Budakoti, and C.-H. Lung, explained that Internet of Things (IoT) has drawn a great deal of attention and is envisioned in varioussectors in the near future due to its promising benefits. However, the constant and rapid growth inIoT devices also brings new challenges due to constrained power and resources associated to them.One of the challenges is to provide seamless connectivity in mobile IoT. Secondly, IoT devices may stream enormous amount of data; hence, providing a solution that can effectively reduce service cost of data transfer. Finally, there are challenges in management and deployment of services running at mobile IoT Edge Gateway. In this context, containerized virtualization solutioncould play a key role in support of efficient management and deployment of microservices to provide seamless connectivity. This paper proposes a lightweight container-based virtualization technology for IoT, which employs Docker container-based microservices architecture for effectively deploying applications in a virtualized ecosystem. We evaluated the performance of the proposed solution on real IoT testbed using Raspberry Pi 3 as a mobile IoT Edge Gateway fornetwork handover decision making among various alternatives, such as Wi-Fi, Radio, and Satellite. The results demonstrated better performance compared with the native environment, i.e.,the one without introduction of a virtualization layer. The results also showed that the Docker container produces negligible resource overhead and can be used on resource constrained mobile IoT Edge Gateway devices like Raspberry Pi 3 for efficiently managing IoT application and services.

A. Samanta and J. Tang, In recent years, the rapid development of mobile edge computing (MEC) provides an efficient execution platform at the edge for Internet-of-Things (IoT) applications. Nevertheless, the MEC also provides optimal resources to different micro services, however, underlying network conditions and infrastructures inherently affect the execution process in MEC. Therefore, in the presence of varying network conditions, it is necessary to optimally execute the available task of end users while maximizing the energy efficiency in edge platform and we also need to provide fair Quality-of-Service (QoS). On the other hand, it is necessary to schedule the micro services dynamically to minimize the total network delay and network price. Thus, in this article, unlike most of the existing works, we propose a dynamic micro service scheduling scheme for MEC. Wedesign the micro service scheduling framework mathematically and also discuss the computationalcomplexity of the scheduling algorithm. Extensive simulation results show that the micro servicescheduling framework significantly improves the performance metrics in terms of total network delay, average price, satisfaction level, energy consumption rate (ECR), failure rate, and networkthroughput over other existing baselines.

W. Jin, R. Xu, T. You, Y.-G. Hong, and D. Kim discussed that Edge computing is an emerging computing paradigm that distributes the computational capabilityto the edge of networks for enabling the computation near to the environment where the sensors and actuators are deployed. Therefore, from the network edge, heterogeneous solutions can be provided to the Internet based on sufficient computing ability. Nevertheless, computing and networking resources are constrained for devices in the network edge. Providing secure services from edge computing is a challenge based on constrained resources. In this paper, we propose a secure edge computing to provide management of device, data, user and additional services basedon deploying independent microservices providers with a security gateway on an edge gateway. The edge gateway is the hub of a local network where multiple IoT devices are deployed to interactwith the physical environment for sensing and actuating. The gateway provides the management functionalities through microservices based on multiple independent server modules. Each gateway-centric local network has a standalone management service based on the gateway. For providing secure edge computing services through the edge gateway, a security gateway is deployed on the proposed edge gateway to provide Representational State Transfer Application Programming Interfaces to expose the security services to the Internet instead of microservices from management modules.

J. Islam, E. Harjula, T. Kumar, P. Karhula, and M. Ylianttila. Explained that The following topics are dealt with: 3G mobile communication; 5G mobile communication; telecommunication traffic; Internet of Things; quality of service; telecommunication network reliability; virtualization; telecommunication security; Long Term Evolution; cellular radio.

III. METHODOLOGY

The study of this paper in, the most recent findings on service placement computation and offloading in fog were presented. Specifically, this study suggests a novel category of optimization techniquesfor tackling service placement issues in IoT applications running on fog nodes. The methods andoptimization objectives play a significant role in this categorization. However, reinforcement learning methods were not investigated as a solution to the offloading issues in fog computing inthis study. In addition, the authors of reviewed resource management strategies in fog computingenvironments considering heterogeneity, resource limitations, and unknown future fog computing traffic. In addition, they presented significant management issues including resource scheduling, resource allocation, task offloading, and resource provisioning.

EXISTING SYSTEM:

Students learning performance is one of the core components for assessing any educational systems. Student’s performance is very crucial in tackling issues of learning process and one of the important matters to measure learning outcomes. The ability to use data knowledge to improveeducation systems has led to the development of the field of research known as educational data mining (EDM). EDM is the creation of techniques to investigate data gathered from educational settings, allowing for a more thorough and accurate understanding of students and the improvement of educational outcomes for them. The use of machine learning (ML) technology has increased significantly in recent years. Researchers and teachers can use the measurements ofsuccess, failure, dropout, and more provided by the discipline of data mining in education to predictand simulate education processes. Therefore, this work presents an analysis of student’s performance using data mining methods.

Existing System Disadvantages:

- Interconnecting appropriate micro services
- Storage and analytics for such a large volume of data.

PROPOSED SYSTEM

This paper provides a detailed evaluation based on an actual prototpye implementation and performance measurement. In our setup, an edge server fulfills dual roles of being an administrative controller of the IoT infrastructure and satisfying application's latency and privacyconstraints. We demonstrate the utility of this architecture by isolated and independent implementation of different microservices, constructing an IoT application by interconnecting these microservices, and potential sharing of microservices between different IoT applications running simultaneously to enhance interoperability. Finally, we provide an extensive performanceevaluation focusing on application latency as well as CPU and memory consumption.

PROPOSED SYSTEM ADVANTAGE

- Infrastructure and satisfying application's latency and privacy constraints
- Running simultaneously to enhance interoperability.
- For secure semantic optimal matching.

We assume that the data owner is trusted, and the data users are authorized by the data owner. The communication channels between the owner and users are secure on existing security protocols such as SSL, TLS. With regard to the cloud server, our scheme resists a more challenging securitymodel which is beyond the "semi-honest server" used in other secure semantic searching schemes.In our model, the dishonest cloud server attempts to return wrong/forged search results and learn sensitive information, but would not maliciously delete or tamper with the outsourced documents. Therefore, our secure semantic scheme should guarantee the verifiability, and confidentiality undersuch a security model

MODULES NAME:

1. User
2. Pre processing
3. Classifier
4. Training Data

1. User

In this module we design the windows for the project. These windows are used for secure login for all users. To connect with server user must give their username and password then only they can able to connect the server. If the user already exits directly can login into the server else user must register their details such as username, password and Email id, into the server. Server will create the account for the entire user to maintain upload and download rate. Name will be set as user id. Logging in is usually used to enter a specific page.

2. Pre processing

This is the first module Data User can register and Login. After login Data User have an option of searching the files as a file name. Data user can also have a download file it will show an encrypteddata. Data user can also send a trapdoor request to the server. Server can accept the request and then data user can takes permissions from the owner then the file it will downloaded in plain text.

3. Classifier

This is the Second module of this project. In this module Data Owner should register and Login. Data Owner will Uploads the files into the database. Data owner can also send request to the datauser.

4. Training Data

This is the third module of this project. In this module Cloud Server can login. After login it willsee all data owners' information. Cloud server can see all users' information. Cloud server can see an all stored data files. Cloud server can give keys request to the user. Cloud server can also see an attacker information of file.

IV. IMPLEMENTATION

1. Data Collection and Preparation:

Identify data sources:

- Admissions data (high school GPA, standardized test scores)
- Student information system (course enrollment, grades)
- Learning management system (online activity, attendance)
- Demographic data (optional, with student consent)

Data cleaning and pre-processing: Ensure data quality by addressing missing values, inconsistencies, and formatting issues.

Feature engineering: Create new features based on existing data (e.g., time spent on assignments, course difficulty).

2. Model Building and Training:

Choose a machine learning algorithm:

- Classification algorithms (e.g., Logistic Regression, Decision Trees) are common for predicting student success (pass/fail, high/low GPA).
- Regression algorithms (e.g., Linear Regression) can predict specific grade values.

Split data into training and testing sets: The training set is used to build the model, and the testing set evaluates its performance on unseen data.

Train the model: The algorithm learns patterns from the training data to predict student performance.

3. Model Evaluation and Deployment:

Evaluate model performance: Metrics like accuracy, precision, and recall assess how well the model predicts student success.

Refine the model: Based on evaluation results, adjust the model parameters or try different algorithms for improvement.

Deployment: Integrate the model into an existing system or create a user interface for educators or advisors to access predictions.

4. Intervention and Support:

Identify at-risk students: Use the model's predictions to flag students with a high likelihood of struggling.

Develop targeted interventions:

- Offer academic support services (e.g., tutoring, study skills workshops).
- Connect students with mentors or advisors.
- Provide early alerts to allow students to adjust study habits.

Monitor student progress: Track the effectiveness of interventions and adjust them as needed.

Additional Considerations:

Security and Privacy: Ensure student data is anonymized or handled according to data privacy regulations.

Transparency and Explainability: Educators and students should understand how predictions are made and the limitations of the model.

Ethical Use: Avoid using predictions for discriminatory purposes or labeling students.

Tools and Resources:

Open-source software libraries like scikit-learn (Python) offer tools for data analysis and machine learning.

Educational data mining platforms can be helpful for managing and analyzing student data.

V. EXPERIMENTAL RESULTS

This project is implements like web application using COREJAVA and the Server process is maintained using the SOCKET & SERVERSOCKET and the Design part is played by CascadingStyle Sheet.

SNAPSHOTSHOME

PAGE:

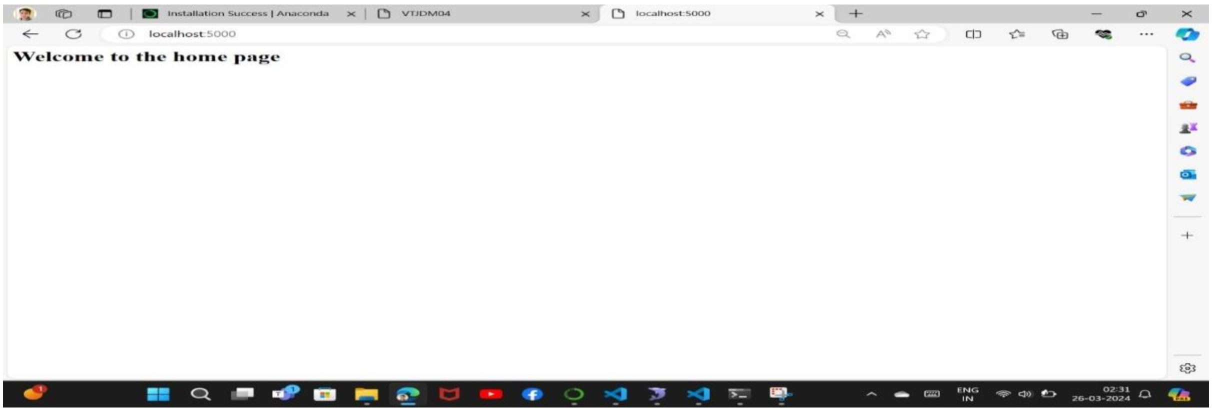

Fig: Home Page

The home page of our project on predicting student performance in higher education is designedto be welcoming and informative. At the top, you'll find our logo and navigation menu to help you explore. The main section features a catchy title explaining our goal: helping students succeedby predicting their performance early on. Below that, we explain what the project is about and why it matters in simple terms. We highlight key features like data analysis methods and prediction

accuracy. Further down, you'll see charts and graphs illustrating trends in student performance, making complex information easy to understand. Overall, our home page aims to engage visitorsand show them the value of our project in a clear and accessible way.

RESULT PAGE:

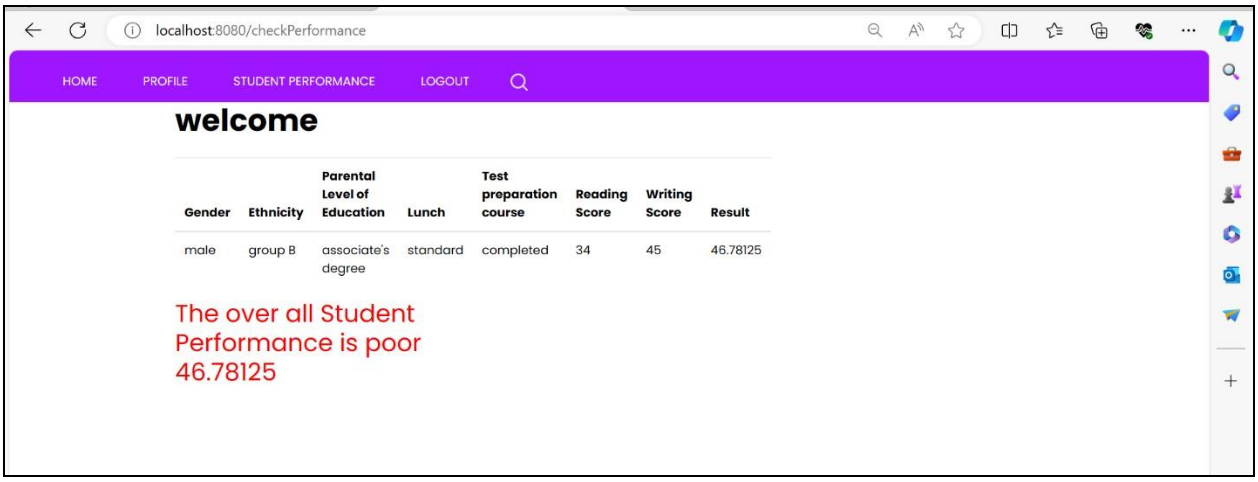

Fig: Result Page

The result page of our project for predicting student performance in higher education offers straightforward insights. At the top, you'll see our logo and navigation menu to help you get around. The main section gives a summary of the student's expected performance, including howaccurate the prediction is. Below that, you'll find details about the student, like their name and program, alongside their predicted grades or performance levels for each subject. We also providehelpful tips or suggestions based on the predictions. Overall, our result page aims to make it easyfor students to understand their expected performance and how they can improve.

VI.CONCLUSION

The student performance is a vital issue. It is difficult to deal with this issue. This paper presentedan analysis of the results data mining research to develop models of students' performance prediction. Our paper showed the use of machine learning algorithms to be better understand efficiency of the algorithms with data dimensionality reduction by T-SNE. It uses four factors such as admission scores and first level courses, academic achievement test (AAT) and general aptitudetest (GAT).In the future, we would like to use deep learning architectures to construct the prediction and improve performance. It can be combined non-academic features with academic features.

By leveraging predictive analytics and machine learning models, institutions can analyze various factors such as past academic performance, socio-economic background, and engagement metricsto forecast student outcomes. This proactive approach allows educators to provide personalized support, tailored resources, and targeted interventions to students who may be at risk of academicunderachievement. Moreover, early prediction enables institutions to allocate resources efficientlyand effectively, ultimately enhancing student retention and graduation rates. Embracing predictiveanalytics in higher education empowers institutions to foster a culture of student success and continuous improvement, ensuring that every student has the opportunity to thrive academically and achieve their full potential.

FUTURE ENHANCEMENT

In the future, we would like to use deep learning architectures to construct the prediction and improve performance. It can be combined non-academic features with academic features.

Enhancing the features of early prediction of students' performance in higher education involves several key strategies. Firstly, by incorporating additional data sources such as extracurricular activities and personal interests, institutions can gain a more comprehensive understanding of eachstudent's profile, leading to more accurate predictions. Secondly, real-time monitoring and feedback systems enable educators to intervene promptly when signs of academic struggle arise, preventing issues from escalating and enhancing students' learning experiences.

Integrating predictive models with existing learning management systems streamlines data analysis and facilitates seamless communication between predictive analytics systems and educators. Moreover, personalized learning pathways tailored to each student's strengths and weaknesses can be developed using predictive analytics, maximizing student engagement and success. Collaboration among institutions and researchers fosters knowledge sharing and advancesthe field of predictive analytics in higher education. Finally, prioritizing ethical considerations andtransparency ensures that student data is handled responsibly and maintains trust in predictive analytics initiatives. These feature enhancements collectively strengthen institutions' capacity to predict and support student success, ultimately improving outcomes for all students.

REFERENCES

[1] F. Giannakas, C. Troussas, I. Voyiatzis, and C. Sgouropoulou, ‘‘A deep learning classification framework for early prediction of team-based academic performance,’’ Appl. Soft Comput., vol. 106, Jul. 2021, Art. no. 107355.

[2] H. Hamsa, S. Indiradevi, and J. J. Kizhakkethottam, ‘‘Student academic performance prediction model using decision tree and fuzzy genetic algo rithm,’’ Proc. Technol., vol. 25, pp. 326–332, Jan. 2016.

[3] B. K. Francis and S. S. Babu, ‘‘Predicting academic performance of students using a hybrid data mining approach,’’ J. Med. Syst., vol. 43, no. 6, pp. 1–15, Jun. 2019.

[4] M. Yağcı, ‘‘Educational data mining: Prediction of students’ academic performance using machine learning algorithms,’’ Smart Learn. Environ., vol. 9, no. 1, pp. 1–19, Dec. 2022.

[5] T. Le Quy, T. H. Nguyen, G. Friege, and E. Ntoutsi, ‘‘Evaluation of group fairness measuresin Student performance prediction problems,’’ 2022, arXiv:2208.10625.

[6] X. Liu and L. Niu, ‘‘A student performance predication approach based on multi-agent systemand deep learning,’’ in Proc. IEEE Int. Conf. Eng., Technol. Educ. (TALE), Dec. 2021, pp. 681– 688.

[7] C.RomeroandS.Ventura, ‘‘Educational data mining: A survey from 1995 to 2005,’’ Expert Syst. Appl., vol. 33, pp. 135–146, Jul. 2007.

[8] R. Conijn, C. Snijders, A. Kleingeld, and U. Matzat, ‘‘Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS,’’ IEEE Trans. Learn. Technol., vol. 10, no. 1, pp. 17–29, Jan./Mar. 2017.

[9] S.Helal,J. Li, L. Liu, E.Ebrahimie, S.Dawson,D.J.Murray,andQ.Long, ‘‘Predicting academic performance by considering student heterogeneity,’’ Knowl.-Based Syst., vol. 161, pp. 134–146, Dec. 2018.