iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60859

Early Performance Prediction of Students in Higher Education

N. Vamshi Vardhan Reddy[1], K. Manohar Reddy[2], P. Arun[3], D. Saidulu[4]
[1, 2, 3]UG Students, [4]Associate Professor
[1, 2, 3]Department of Computer Science and Engineering-(Internet of Things)
[4] Department of Information Technology
[1, 2, 3, 4]Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India

***Abstract:** An essential component of any evaluation of an educational institution is assessing the learning outcomes of its pupils. Student performance is important when addressing issues with the learning process and is one of the main variables used to quantify learning outcomes. The area of study referred to as the idea of using data knowledge to improve educational systems gave rise to educational data mining, or EDM. The creation of techniques for analyzing data gathered from educational settings is known as EDM. This allows for a more thorough and accurate knowledge of pupils as well as an improvement in their academic performance. Assessing the learning outcomes of students is an essential component of assessing any educational establishment. When addressing issues with the learning process, student performance is an important metric utilized to quantify learning outcomes. The area in which the possibility to use data knowledge to improve educational systems gave rise to a field of study known as educational data mining, or EDM. EDM is the process of creating tools to assess data that is gathered from educational settings. Students' academic performance is improved and more accurate and detailed information about them can be learned as a result. The paper's dimensionality reduction method by T-SNE algorithm uses academic achievement tests (AAT), general aptitude tests (GAT), admission scores, first-level courses, and other early-stage criteria for the method of clustering. This makes it possible for the study to look into how these factors and GPAs are related. As for the classification approach, the paper presents experiments on multiple machine learning models that predict student performance in the early stages using a variety of characteristics such course grades and admission exam scores.*
***Keywords:** Educational data mining (EDM), Academic achievement tests (AAT), General aptitude tests (GAT), Admission scores, First-level courses, GPA*

I. INTRODUCTION

In our culture, education is a vital component that holds great importance. Research in many sectors, including education, has been impacted by information and communication technology. For instance, as the recent COVID19 epidemic has shown, several nations have employed a variety of e-Learning environments. One of the most crucial aspects of providing its students with a high-quality education, in the eyes of a higher education institution, is their academic achievement. It can be challenging to identify the important variables influencing a student's performance in the early stages of their schooling. To address the issues with kids' academic performance, many useful strategies have been employed. However, it might not be simple to generalize these techniques to all educational situations. The use of technology to predict student performance has advanced recently, but there are still gaps that need to be filled in order to analyze and improve the accuracy of student performance using new features and data mining techniques[1]. These techniques also present both clustering and classification techniques to identify the impact of early student performance on the GPA.

II. EXISTING SYSTEM

One of the main elements used to evaluate any educational system is the performance of the students in their learning. In order to address concerns with the learning process and determine learning outcomes, student performance is vitally necessary. Educational data mining (EDM) is a topic of study that has grown out of the ability to use data knowledge to enhance educational systems[2]. The development of investigative methods for studying data collected from educational environments (EDM) enables more precise and in-depth student understanding as well as improved educational outcomes. In recent years, there has been a notable surge in the utilization of machine learning (ML) technologies.

A. Existing System Disadvantages
1) Interconnecting appropriate micro services
2) Storage and analytics for such a large volume of data.

III. PROPOSED SYSTEM

Based on a real prototype implementation and performance testing, this study offers a thorough assessment. In our configuration, an edge server serves as both the IoT infrastructure's administrative controller and a means of meeting the latency and privacy requirements of the applications.

By implementing various microservices in an isolated and independent manner, connecting them to create an IoT application, and possibly sharing microservices between concurrently running IoT applications to improve interoperability, we illustrate the usefulness of this architecture[3]. Lastly, we offer a thorough performance analysis that emphasizes CPU and memory usage in addition to application latency.

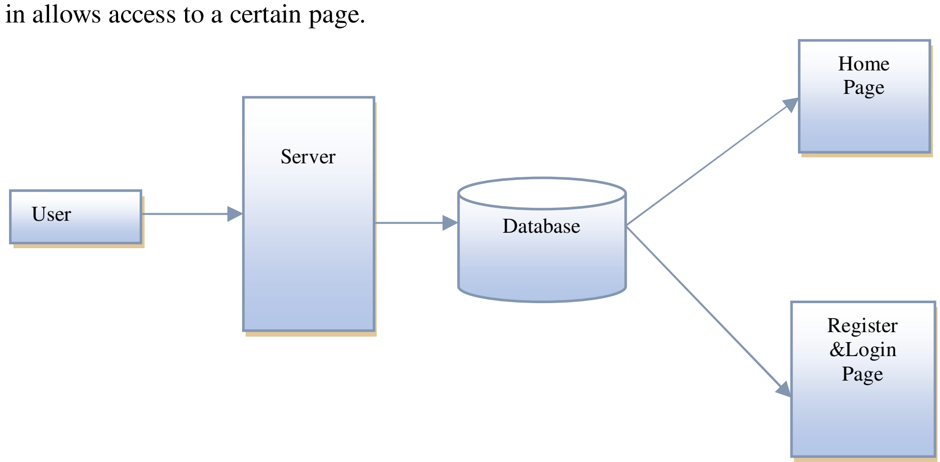
A. Proposed System Advantages
1) Infrastructure and meeting latency and privacy limitations for applications
2) Operating concurrently to improve compatibility.
3) To ensure safe semantic best matching

IV. METHODOLOGIES

We take it for granted that the data owner is reliable and that the data users have permission from them. With the use of current security protocols like SSL and TLS, the owner and users can communicate securely. Our technique is resistant to a more difficult security model for the cloud server than the "semi-honest server" found in existing safe semantic searching schemes[4][5]. In our case, the unscrupulous cloud server tries to get sensitive data and provide false or misleading search results, but it doesn't intentionally remove or alter the documents that are outsourced. Consequently, under such a security model, our secure semantic scheme ought to ensure the verifiability and confidentiality

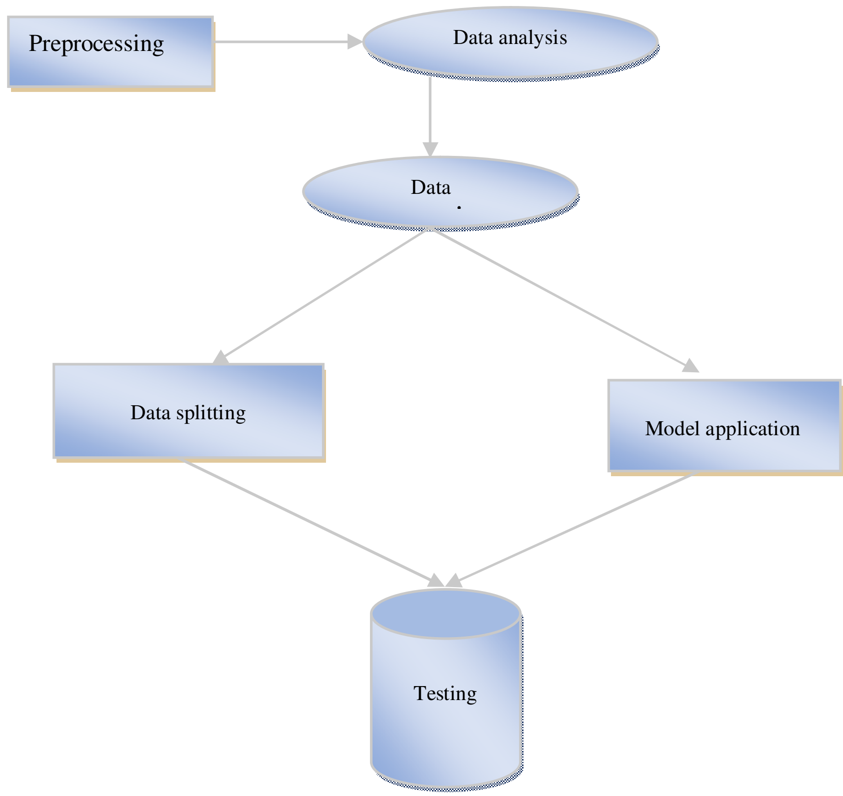
A. User

We create the project's windows in this module. All users can securely log in using these windows. Users can only connect to the server by providing their login and password in order to establish a connection. The user can log in straight to the server if they have previously left; otherwise, they must register their information, including their email address, password, and username. In order to maintain the upload and download rates, the server will create an account for each user[6]. The user ID will be set to name. Typically, logging in allows access to a certain page.



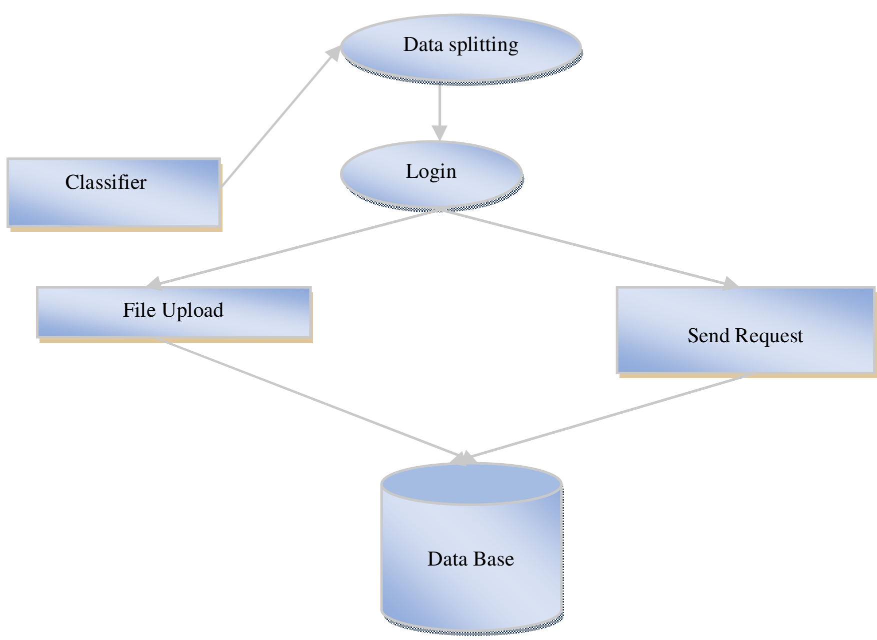
B. Pre processing

The first module where users can sign up and log in is this one. Users have the opportunity to search files by name after logging in. Users of the data can also download a file that displays encrypted data. A trapdoor request can also be sent to the server by a data user. After the server approves the request, the user can obtain the owner's permissions and download the file in plain text.



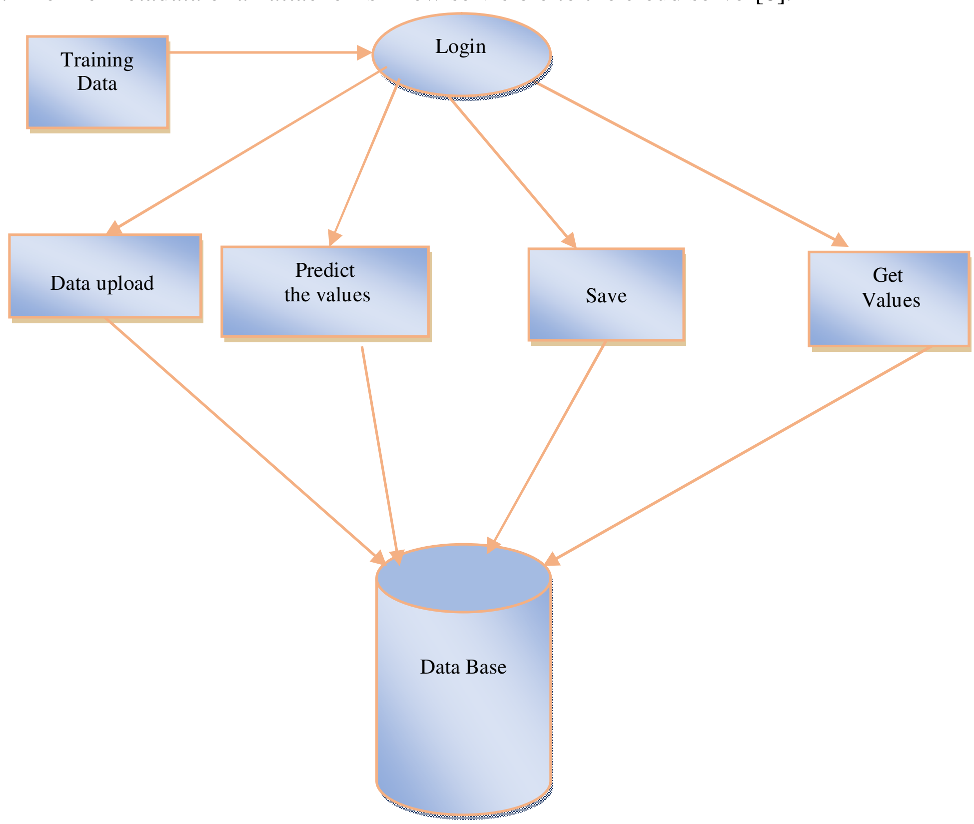
C. Classifier

This project's second module is this one. The data owner has to log in and register for this module. The files will be uploaded into the database by the data owner[7]. Requests may be sent from the data owner to the data user.



D. Training Data

This project's third module is this one. Login to Cloud Server via this module. All data owners' details will be visible after logging in. The cloud server has access to every user's data. Every stored data file is visible to the cloud server. The user can request keys from the cloud server. The file metadata of an attacker is likewise visible to the cloud server[8].



V. REQUIREMENTS

These are the requirements for doing the project. Without using these tools & software's we can't do the project. So we have two requirements to do the project. They are

A. Hardware Requirements

[9][10]The hardware requirements should be a comprehensive and uniform definition of the entire system since they might form the foundation of a contract for the system's implementation. Software engineers use them as the foundation for their system designs. It focuses on the functionality of the system rather than the best way to use it.

- PROCESSOR : DUAL CORE 2 DUOS.
- RAM : 2GB DD RAM
- HARD DISK : 250 GB

B. Software Requirements

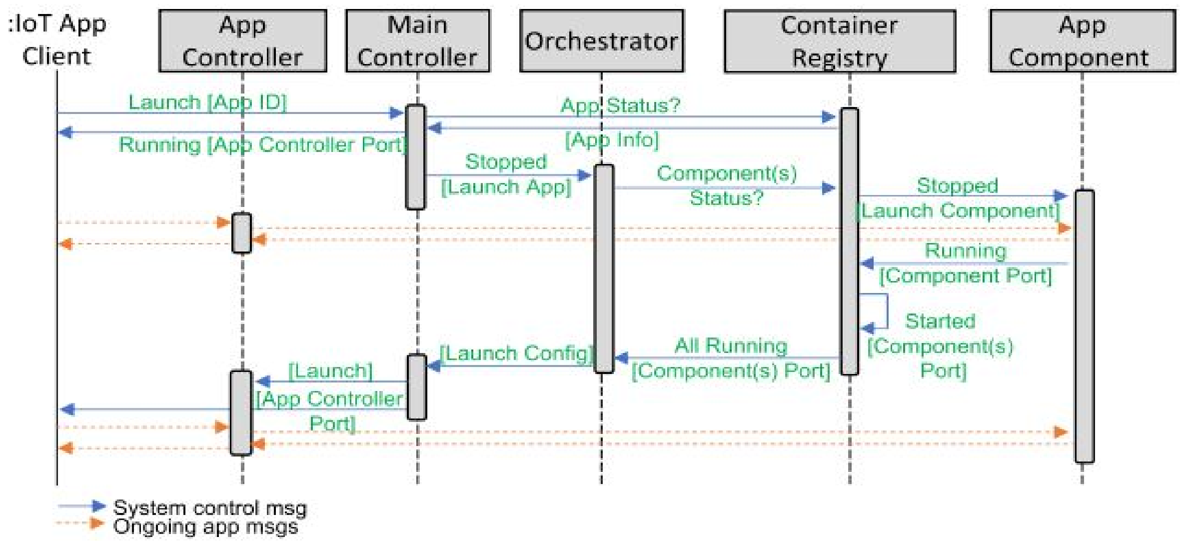
The system specification is found in the software requirements paper. It ought to have a description and a list of prerequisites. Rather than focusing on how the system should operate, it is a list of what it should perform. The foundation for developing the software requirements specification is provided by the software requirements[11]. It is helpful for cost estimation, organizing team activities, carrying out tasks, managing teams, and monitoring the teams' advancement during the development process.

- FRONT END : J2EE (JSP, SERVLET)
- BACK END : MY SQL 5.5
- OPERATING SYSTEM : WINDOWS 7
- IDE : ECLIPSE

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

VI. SYSTEM ARCHITECTURE

The data owner for this project must register all details before logging in. A document may be uploaded by the data owner. The data user may get a request from the data owner. A data user can utilize an uploaded document to search for a query. There is a download option for the file, which displays the encryption format. The cloud server receives requests from data users as well. A cloud server may have a login. The key will be accepted by it[13]. The entire data set is visible to the cloud server as well. All user data is visible to the cloud server as well. The whole stored data set is visible to the cloud server. A key request from the user may be approved by the cloud server. After receiving the request, the data owner might provide the user with a secret key. [14]After then, the user can download a file. In the event that the user enters the wrong keys, a permanent block is imposed. There is an attack on the file.



VII. LITERATURE SURVEY

By extending the Cloud Computing paradigm to the network's edge, fog computing makes it possible for new types of services and applications. Low latency and location awareness, broad geographic dispersion, mobility, a high number of nodes, e) the predominance of wireless access, f) the strong presence of streaming and real-time applications, and g) heterogeneity are features that define the fog. In this study, we suggest that the aforementioned features make the Fog the right platform for several important Internet of Things (IoT) applications and services, such as Smart Cities, Smart Grids, Connected Vehicles, and Wireless Sensors and Actuators Networks (WSANs) in general[15]. Inspired by the strong need to maximize the use of non-general purpose devices in order to achieve computational goals at a lower cost, we put forth a novel scheduling model for microservices across heterogeneous cloud-edge environments. Our model describes an architecture with heterogeneous machines that can process distinct microservices using a specific mathematical formulation. Since each novel model requires an early solution risk analysis, we modified the Clouds simulation framework to make it appropriate for an experiment involving those kinds of systems.

In this study, we describe two applications of our suggested scheduling system in practice. We further give some experimental results based on the created simulation tool for an unbiased understanding of the first case[16]. Our interpretation of the experimental data reveals that, when utilizing a micro service-oriented approach, certain very basic scheduling algorithms may perform better than others in specific scenarios that are commonly found in cloud-edge environments. The Internet of Things, or IoT, has garnered a lot of interest lately and is expected to have a significant impact on several sectors in the near future because of its many prospective applications. However, because of their limited power and resource availability, the ongoing and quick expansion of IoT devices also presents new difficulties. Seamless connectivity for mobile IoT is one of the issues[17]. Second, Internet of Things devices have the potential to stream massive amounts of data, offering a way to effectively lower data transmission service costs. Lastly, there are issues with service deployment and management at mobile IoT Edge Gateways. In this situation, a containerized virtualization solution might be extremely helpful in facilitating the effective deployment and management of microservices to offer seamless communication. Using Docker container-based microservices architecture, this article presents a lightweight container-based virtualization technique for the Internet of Things that enables the efficient deployment of applications within a virtualized environment.

We assessed the effectiveness of the suggested method on an actual Internet of Things testbed by utilizing a Raspberry Pi 3 as a portable IoT Edge Gateway to make decisions about network handover between Wi-Fi, radio, and satellite, among other options. When compared to the native environment—that is, the one without the addition of a virtualization layer—the results showed improved performance. The outcomes also demonstrated that the Docker container uses very little resources and can be effectively utilized to manage IoT applications and services on mobile IoT Edge Gateway devices with limited resources, such as the Raspberry Pi 3[18]. The swift advancement of mobile edge computing (MEC) in recent times offers an effective edge execution platform for Internet-of-things (IoT) applications. Even so, the MEC also offers various micro services the best resources available; yet, the MEC's execution process is intrinsically impacted by the infrastructures and underlying network conditions. Therefore, it is essential to maximize energy economy in the edge platform, ensure fair Quality-of-Service (QoS), and execute end users' available tasks as efficiently as possible in the presence of fluctuating network conditions. To reduce the overall network latency and cost, however, it is imperative to dynamically schedule the microservices. Several IoT devices are deployed to interact with the physical environment for detecting and actuating purposes, and the edge gateway serves as the core of this local network. Through microservices built on several separate server modules, the gateway offers management functions[19].

VIII. EXISTING ALGORITHM

A. Academic Achievement Tests (AAT)

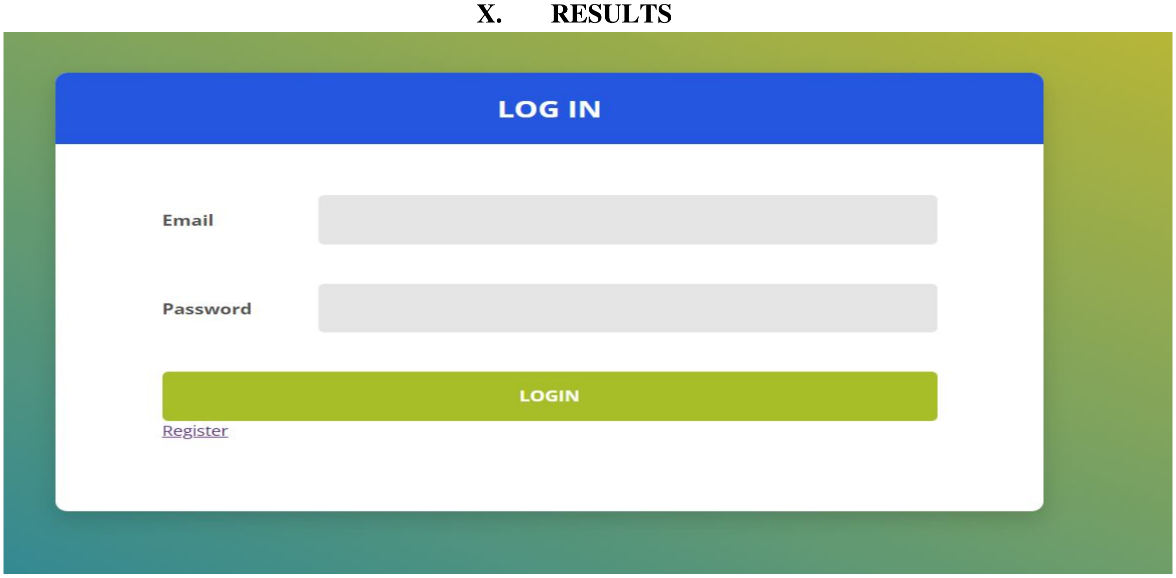
Tests of general aptitude and academic accomplishment (AAT) to investigate the connection between these variables and GPAs. The study reports on trials on various machine learning models that predict early student performa[20]nce based on a variety of characteristics, such as course grades and admission exam results, for the classification technique. To analyze the models' quality, we employ several evaluation measures[21]. The findings imply that early student failure risks can be reduced by educational institutions. Because academic achievement is closely associated with the results we value, it is vital. The performance of academic students in the college or university is one of the variables that determines their academic achievement.

IX. PROPOSED ALGORITHM

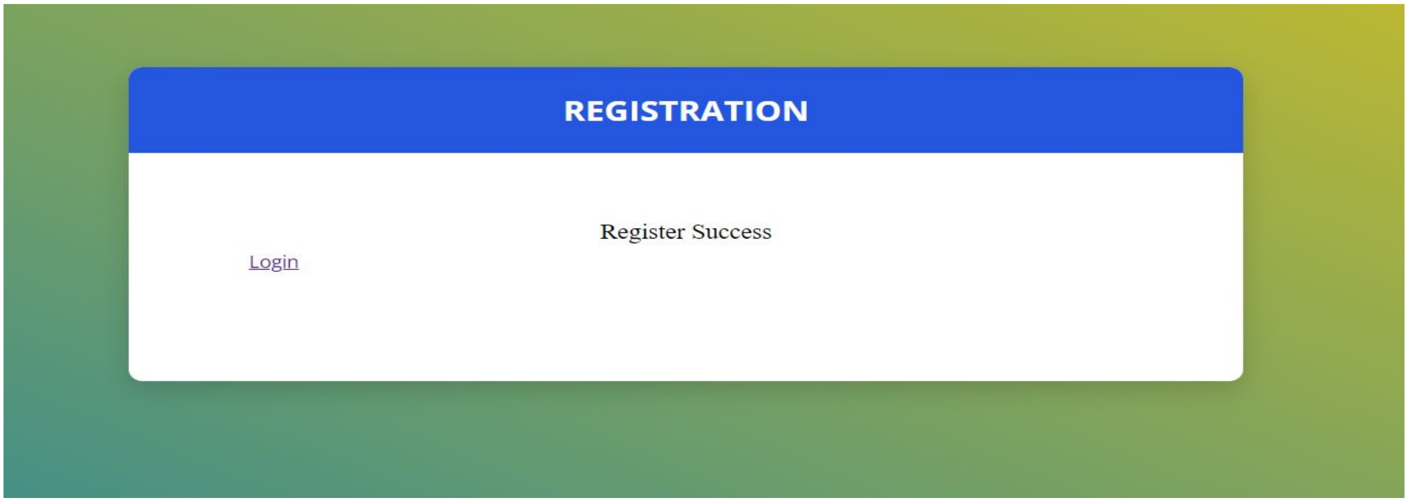
A. Educational Data Mining (EDM)

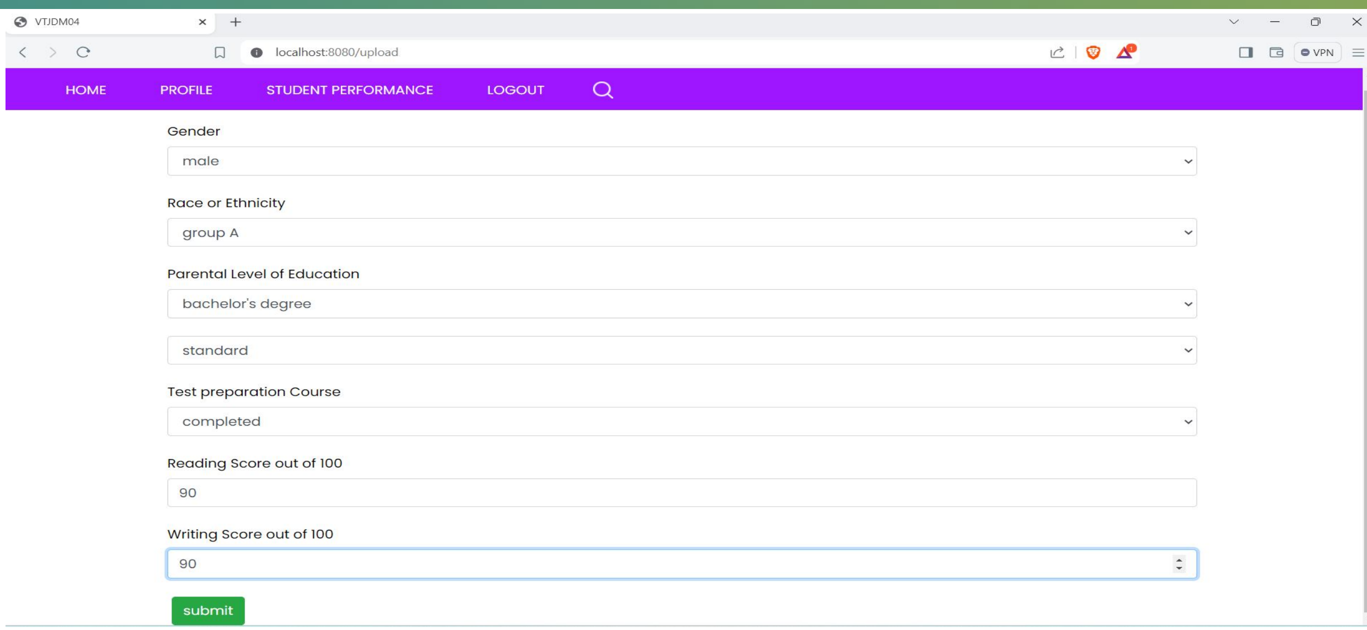
Student performance is an integral part of the learning process. Predicting student performance is necessary in order to identify pupils who are more likely to experience low academic accomplishment in the future. [22]Predictions may benefit from the data if it has been converted into knowledge. Consequently, the data may enhance the standard of instruction and learning and assist students in reaching their learning goals. In the field of research known as educational data mining (EDM), information collected from educational backgrounds is analyzed using data mining techniques[23]. Implementing EDM helps with strategy development as well as improving student performance. It will thereby enhance instruction and learning as well as the educational institution's student experience.

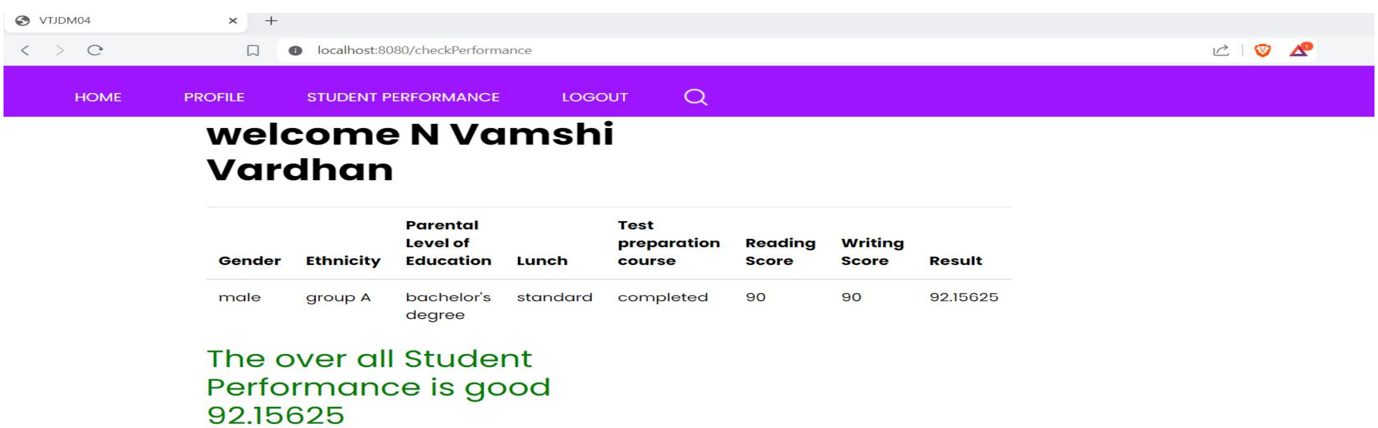
X. RESULTS



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com





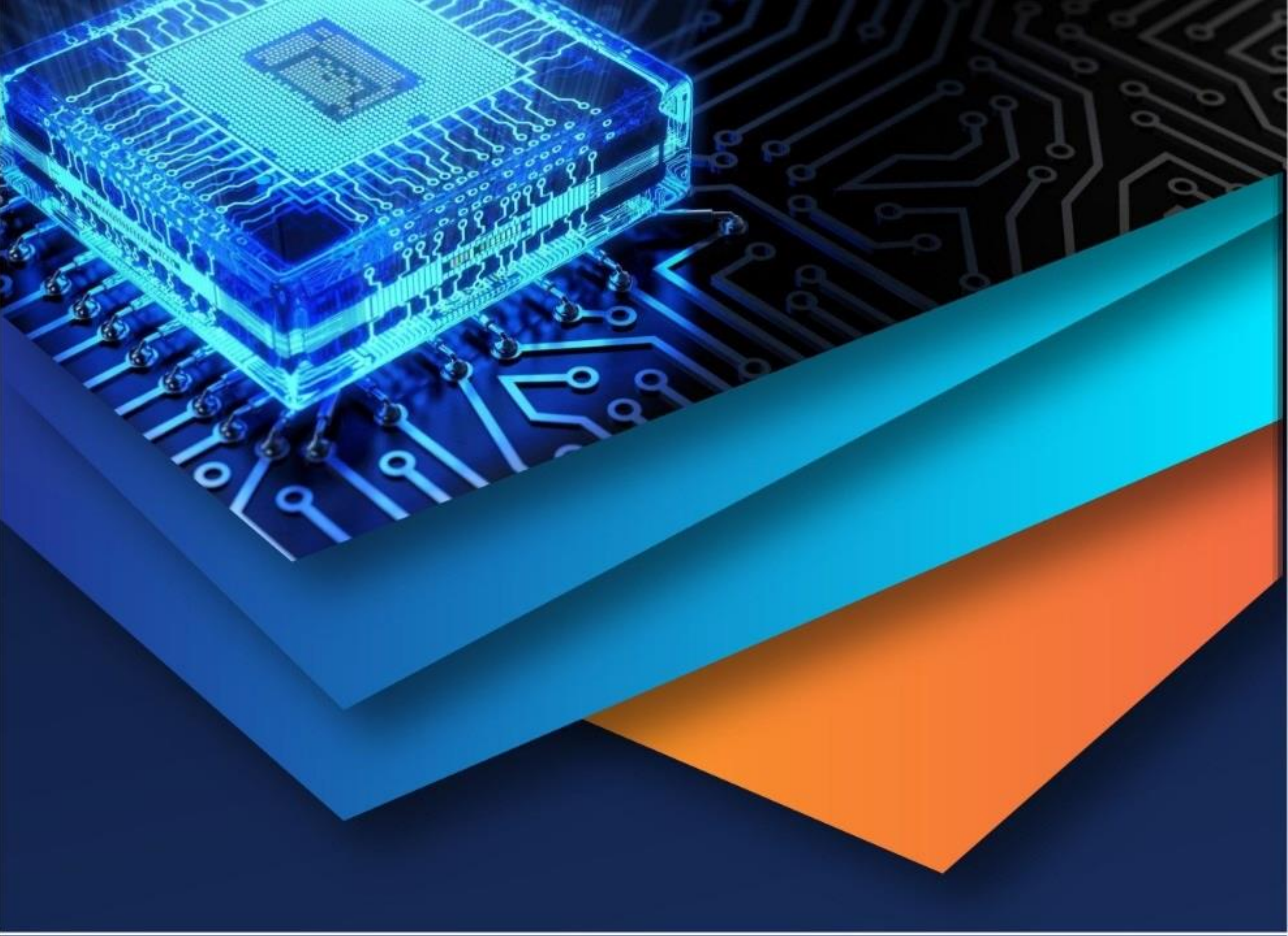


XI. CONCLUSION

The student performance is a vital issue. It is difficult to deal with this issue. This paper presented an 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[24]. It uses four factors such as admission scores and first level courses, academic achievement test (AAT) and general aptitude test (GAT).[25]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

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 measures in Student performance prediction problems,'' 2022, arXiv:2208.10625.
[6] X. Liu and L. Niu, ''A student performance predication approach based on multi-agent system and 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.
[10] X. Xu, J. Wang, H. Peng, and R. Wu, ''Prediction of academic perfor mance associated with internet usage behaviors using machine learning algorithms,'' J. Comput. Hum. Behav., vol. 98, pp. 166–173, Sep. 2019.
[11] T. T. Dien, S. H. Luu, N. Thanh-Hai, and N. Thai-Nghe, ''Deep learning with data transformation and factor analysis for student performance pre diction,'' Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 8, pp. 1–11, 2020.
[12] D. T. Ha, P. T. T. Loan, C. N. Giap, and N. T. L. Huong, ''An empirical study for student academic performance prediction using machine learning techniques,'' Int. J. Comput. Sci. Inf. Secur. (IJCSIS), vol. 18, no. 3, pp. 1–8, 2020.
[13] J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido, and A. Durán-Domínguez, ''Analyzing and predicting students' performance by means of machine learning: A review,'' Appl. Sci., vol. 10, no. 3, p. 1042, 2020.
[14] M. Agaoglu, ''Predicting instructor performance using data mining tech niques in higher education,'' IEEE Access, vol. 4, pp. 2379–2387, 2016.
[15] A. Gonzalez-Nucamendi, J. Noguez, L. Neri, V. Robledo-Rella, R. M. G. García-Castelán, and D. Escobar-Castillejos, ''The prediction of academic performance using engineering student's profiles,'' Comput. Electr. Eng., vol. 93, Jul. 2021, Art. no. 107288.
[16] A. Alshanqiti and A. Namoun, ''Predicting student performance and its influential factors using hybrid regression and multi-label classification,'' IEEE Access, vol. 8, pp. 203827–203844, 2020.
[17] H. A. Mengash, ''Using data mining techniques to predict student perfor manceto support decision making in university admission systems,'' IEEE Access, vol. 8, pp. 55462–55470, 2020.
[18] M. Raihan, M. T. Islam, P. Ghosh, J. H. Angon, M. M. Hassan, and F. Farzana, ''A machine learning approach to identify the correlation and association among the students' educational behavior,'' in Proc. Int. Conf. Comput. Advancements, Jan. 2020, pp. 1–6.
[19] L.S.Maurya,M.S.Hussain,andS.Singh,''Developingclassifiersthrough machine learning algorithms for student placement prediction based on academic performance,'' Appl. Artif. Intell., vol. 35, no. 6, pp. 403–420, May 2021.
[20] N. Aslam, I. Khan, L. Alamri, and R. Almuslim, ''An improved early student's academic performance prediction using deep learning,'' Int. J. Emerg. Technol. Learn. (iJET), vol. 16, no. 12, pp. 108–122, 2021.
[21] D. M. Ahmed, A. M. Abdulazeez, D. Q. Zeebaree, and F. Y. H. Ahmed, ''Predicting university's students performance based on machine learn ing techniques,'' in Proc. IEEE Int. Conf. Autom. Control Intell. Syst. (I2CACIS), Jun. 2021, pp. 276–281.
[22] E. D. L. Evangelista and B. D. Sy, ''An approach for improved students' performance prediction using homogeneous and heterogeneous ensemble methods,'' Int. J. Electr. Comput. Eng. (IJECE), vol. 12, no. 5, p. 5226, Oct. 2022.
[23] L. Van der Maaten and G. Hinton, ''Visualizing data using t-SNE,'' J. Mach. Learn. Res., vol. 9, no. 11, pp. 2579–2605, 2008.
[24] T. Chen and C. Guestrin, ''XGBoost: A scalable tree boosting system,'' in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Aug. 2016, pp. 785–794.
[25] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, ''LIB LINEAR: A library for large linear classification,'' J. Mach. Learn. Res., vol. 9, pp. 1871–1874, Jun. 2008.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY