iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** X **Month of publication:** October 2023

DOI: https://doi.org/10.22214/ijraset.2023.56313

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

Student Performance Prediction

Prof. Madhuri Thorat[1], Nilesh Sonawane[2], Shubham Bhore[3], Santosh Doiphode[4], Vinayak Sunkewar[5]
B.E. in Computer Engineering, JSPM Narhe Technical Campus, Narhe, Pune

Abstract: ***Assessing student's learning performance is a fundamental aspect of evaluating educational systems, playing a pivotal role in addressing challenges within the learning process and measuring learning outcomes. The emergence of educational data mining (EDM) as a research field has harnessed the power of data and knowledge to enhance education systems. EDM involves the development of techniques to analyze data collected from educational environments, offering a more comprehensive understanding of students and facilitating improvements in educational outcomes.***
The integration of machine learning (ML) technology has witnessed substantial growth, enabling researchers and educators to leverage data mining insights to predict and simulate educational processes, including success rates, dropouts, and more. This research paper explores the analysis of students' performance through data mining methods. It employs classification technique to discern the early-stage impact on GPA. In the classification methodology, various machine learning models are experimented with to predict student performance in the early stages, incorporating diverse features such as course grades and admission test scores. The paper employs different assessment metrics to evaluate model performance. The findings underscore the potential of educational systems to proactively address the risk of student failures during their initial stages of education.
Keywords: EDM, ML

I. INTRODUCTION

The aim of Student Performance Prediction is to provide educators, administrators, and policymakers with valuable insights that can be used to improve the quality of education and support students' individual needs.By identifying students who may be at risk of falling behind or excelling, educational institutions can tailor their strategies to better meet the diverse needs of their students. This field of study is particularly relevant today, as educational institutions face numerous challenges, including increasing class sizes, limited resources, and the need to adapt to rapidly changing learning environments. Student performance prediction can offer solutions to these challenges by enabling proactive interventions, personalized learning plans, and data-informed decision-making. In our study, we have developed a system to predict how well students will do in college. To do this, we're looking at their academic performance. We consider factors like their admission scores,
their grades in initial courses, the results from academic achievement tests, general aptitude tests, and considering data of previously placed students with precious packages. What's innovative about our research is that we're using SVM machine learning algorithm to analyze all of these factors from admission scores and first-level course scores to predict a student's performance early on.
This approach is quite unique and hasn't been tried before. Additionally, we're exploring a new way of deciding when to move a student to a different level of education. We do this by calculating the difference between a student's grade and the grade that comes after or before it.To test how well our methods work, we're using advanced classification models. Our goal is to improve our ability to predict how students will do in college using these new and innovative techniques.

II. OBJECTIVES

1) To provide educators with real-time insights into students' academic performance, enabling them to identify areas of improvement promptly and tailor teaching methods accordingly.
2) To identify students who are at risk of falling behind academically or facing behavioural issues, enabling timely intervention strategies to support their progress and well-being.
3) To offer personalized learning recommendations based on individual students' strengths and weaknesses, ensuring a customized educational approach that maximizes their potential.

III. LITERATURE SURVEY

Prediction of academic performance of students beforehand provides scope to universities to lower their dropout rate and help the students in improving their performance. In this field, research is being done to find out which algorithm is best to use and which features should be considered while predicting the academic performance of students.

©IJRASET: All Rights are Reserved

This kind of research work has been increasing over the years. This paper performs a survey on the techniques used in various research papers for academic performance prediction and also point out the limitations if any, in the methodology used.[1]

This work presented two prediction models for the estimation of student's performance in final examination. The work made use of the popular dataset provided by the University of Minho in Portugal, which relate to the performance in math subject and it consists of 395 data samples. Forecasting the performance of students can be useful in taking early precautions, instant actions, or selecting a student that is fit for a certain task. The need to explore better models to achieve better performance cannot be overemphasized. Most of earlier work on the same dataset used K-Nearest Neighbor algorithm and achieved low results, while Support Vector Machine algorithm was rarely used, which happens to be a very popular and powerful prediction technique. To ensure better comparison, we applied both Support Vector Machine algorithm and K-Nearest Neighbor algorithm on the dataset to predict the student's grade and then compared their accuracy. Empirical studies outcome indicated that Support Vector Machine achieved slightly better results with correlation coefficient of 0.96, while the K-Nearest Neighbor achieved correlation coefficient of 0.95.[2]

Early indications of student's progress can help academics to increase their learning strategies and focus on different educational practices to make the learning experience successful. ML application can help academics to predict the weaknesses in learning processes and as a result they can actively engage such students in better learning experience. We applied logistic regression, linear discriminant analysis, K-nearest neighbors, classification and regression trees, Gaussian Naive Bayes and support vector machines on historical data of student grades and developed a model to predict the grades of students. Our experiments show Linear discrimination analysis as the most effective approach to correctly predict the student's performance outcome in final exams. Out of total 54 records, 49 were predicted by model, giving 90.74% of accuracy.[3]

This paper is about how the application of machine Learning have huge impact in teaching and learning for further improvement in learning environment in higher education. Due to the interest of students in online and digital courses increased rapidly websites such as Course Era, Udemy etc became very influential. We implement the new applications of machine learning in teaching and learning considering the students background, students past academic score and considering other attributes. As the sizes of classes are large, it would be difficult to assist each individual student in each open learning course, this can increase the bar of the dropout rate at the end of the course. In this paper we are implementing linear regression which is a machine learning algorithm to predict the student's performance in academics.[4]

This paper presents methods to improve the prediction of student academic performance using feature selection by removing misclassified instances and Synthetic Minority Over-Sampling Technique. It compares the performance of seven students' academic performance prediction models, namely Naive Bayes, Sequential Minimum Optimization, Artificial Neural Network, k-Nearest Neighbor, REPTree, Partial decision trees, and Random Forest. The data were collected from 9,458 students at the Rajabhat Maha Sarakham University, Thailand during 2015 - 2018. The model performances were evaluated with precision, recall, and F-measure. The experimental results indicated that the Random Forest approach significantly improves the performance of students' academic performance prediction models with precision up to 41.70%, recall up to 41.40% and F-measure up to 41.60%, respectively.[5]

Today, predictive analytics applications became an urgent desire in higher educational institutions. Predictive analytics used advanced analytics that encompasses machine learning implementation to derive high-quality performance and meaningful information for all education levels. Mostly know that student grade is one of the key performance indicators that can help educators monitor their academic performance. During the past decade, researchers have proposed many variants of machine learning techniques in education domains. However, there are severe challenges in handling imbalanced datasets for enhancing the performance of predicting student grades. Therefore, this paper presents a comprehensive analysis of machine learning techniques to predict the final student grades in the first semester courses by improving the performance of predictive accuracy. Two modules will be highlighted in this paper. First, we compare the accuracy performance of six well-known machine learning techniques namely Decision Tree (J48), Support Vector Machine (SVM), Naïve Bayes (NB), K-Nearest Neighbor (kNN), Logistic Regression (LR) and Random Forest (RF) using 1282 real student's course grade dataset. Second, we proposed a multiclass prediction model to reduce the overfitting and misclassification results caused by imbalanced multi-classification based on oversampling Synthetic Minority Oversampling Technique (SMOTE) with two features selection methods. The obtained results show that the proposed model integrates with RF give significant improvement with the highest f-measure of 99.5%. This proposed model indicates the comparable and promising results that can enhance the prediction performance model for imbalanced multi-classification for student grade prediction.[6]

Student performance prediction is very important to understand a student progress rate. It is said that 'Prevention is better than the cure'. In this Research, we are trying to find out student's current status and predict his/her future results.

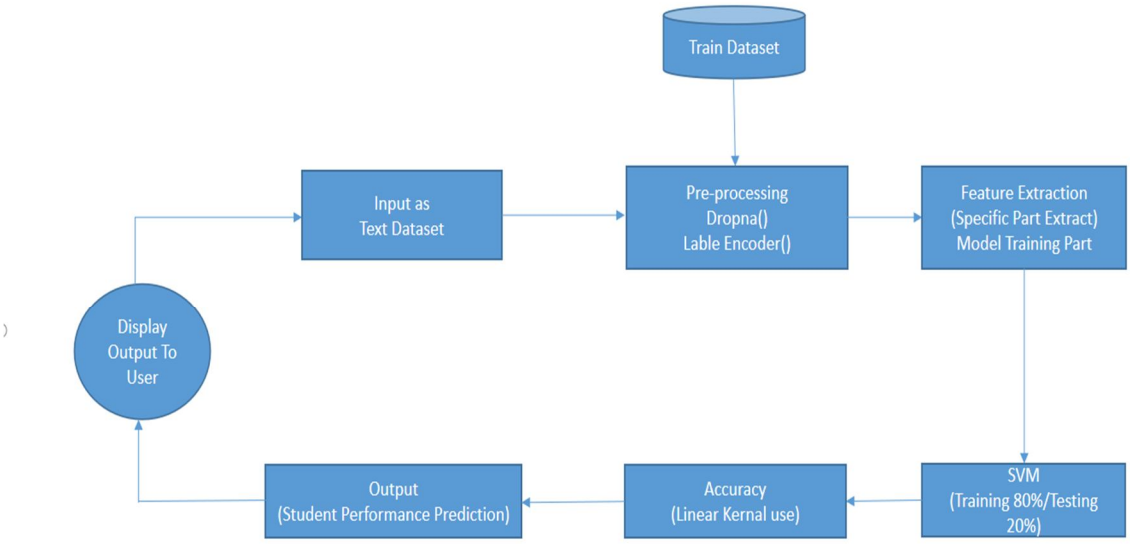
©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

After the outcome, teachers can give him/her proper advice to avoid the poor result and also can groom the student. By finding out the dependencies for final examinations. Which courses he/she should take in the upcoming semester (roles of adviser/teacher). Every year a lot of students lag behind because of lack of proper advice and monitoring. A teacher can't monitor each and every single student at once. If a system can help a Teacher about the students like which student needs which kind of help. Then it will be much helpful for both teachers and student. The aim is helping the student to avoid his/her predicted poor result using Artificial Intelligence. If a student could know what will be his/her result in the future and notify him/her what to do to avoid his/her the bad results by predicting the final examinations mark. This research would be helpful for the students and teachers with The highest accuracy of 94.88%.[7]

IV. IMPLEMENTATION DETAILS OF MODULES

1) Gather relevant data including academic records, attendance, behavioral data, and any other pertinent information.
2) Identify the most relevant features (variables) that influence student performance and engineer new features if needed.
3) Choose appropriate machine learning algorithms (e.g., regression, classification, clustering) and train models using historical data.
4) Integrate machine learning models with the existing SPMS infrastructure for real-time data analysis.
5) Implement algorithms for real-time data analysis and prediction, enabling instant feedback for educators and administrators.
6) Continuously monitor model performance and evaluate its accuracy, precision, recall, and other relevant metrics.

Architecture Diagram:



V. FUTURE SCOPE

The research presented in this paper provides a compelling vision of the future of education, driven by data-driven insights and enhanced student performance prediction. This vision encompasses a wide range of exciting opportunities for future research and development in the field of educational technology.

One of the key areas for future exploration is the integration of advanced data sources. As technology continues to evolve, educational institutions can tap into a wealth of data, including information from online learning platforms, biometric data to assess student engagement, and sentiment analysis of student feedback. The utilization of this broader set of data inputs will undoubtedly lead to more refined and accurate predictive models.

Furthermore, the future holds great promise for the development of enhanced predictive models. By leveraging cutting-edge technologies such as deep learning and natural language processing, these models can provide even more precise and nuanced predictions of student performance. This will enable educators and institutions to tailor interventions and support at an unprecedented level of granularity.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VI. CONCLUSION

In conclusion, the project presented in this research paper addresses a critical need in the field of education: the ability to predict and improve student performance in college.

By leveraging the power of educational data mining (EDM) and machine learning techniques, this research offers a unique and innovative approach to assessing and enhancing the learning experience. The specific objectives outlined for this project contribute to the broader goal of improving educational outcomes and ensuring that students receive the support they need to succeed.

By predicting student performance and identifying at-risk students early in their educational journey, this research enables educational institutions to take proactive measures to prevent academic failures. The creation of personalized learning plans, informed by data-driven insights, ensures that each student's strengths and weaknesses are taken into account, leading to a more tailored and effective educational experience.

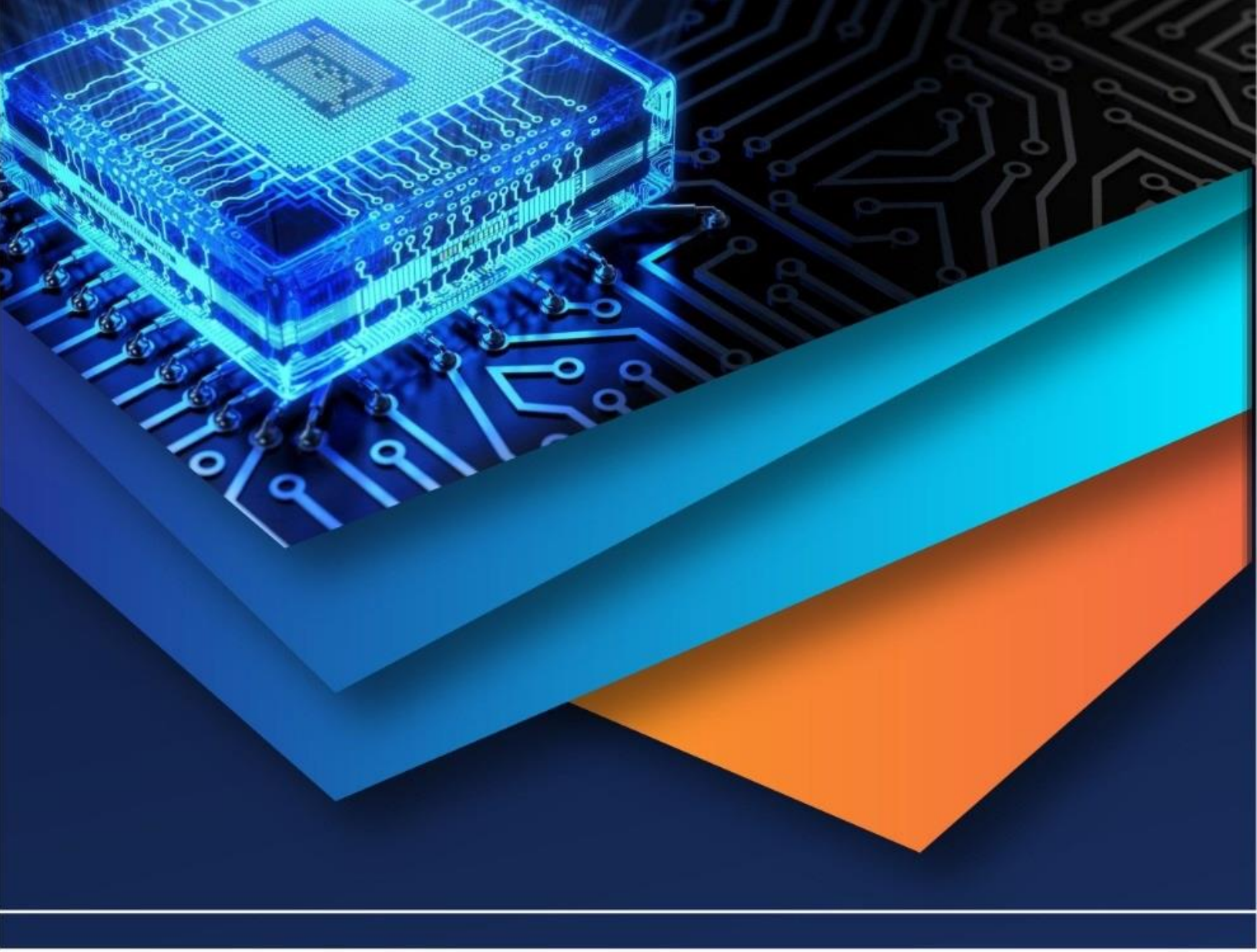
Furthermore, this research project emphasizes adaptability in the face of evolving learning environments, such as those influenced by events like the COVID-19 pandemic. The development of predictive models that can thrive in changing conditions is vital for ensuring the continued success of educational institutions and their students.

The project's innovative use of the Support Vector Machine (SVM) machine learning algorithm and grade difference calculation sets it apart, offering a fresh perspective on performance prediction. These techniques have the potential to enhance the accuracy and effectiveness of predictions, ultimately benefiting both students and educators.

REFERENCES

[1] M. Yagcı, ‘‘Educational data mining: Prediction of students’ academic performance using machine learning algorithms,’’ Smart Learn. Environ., vol. 9, no. 1, pp. 1–19, Dec. 2022.
[2] 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.
[3] 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.
[4] M. Maphosa, W. Doorsamy, and B. Paul, ‘‘A review of recommender systems for choosing elective courses,’’ Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 9, pp. 287–295, 2020.
[5] University of Johannesburg. (2021). Faculty of Engineering & the Built Environment. Accessed: Apr. 6, 2022. [Online]. Available: https://www.uj.ac.za/faculties/engineering-the-built-environment/
[6] W. Doorsamy and K. Padayachee, ‘‘Conceptualising the knower for a new engineering technology curriculum,’’ J. Eng., Design Technol., vol. 17, no. 4, pp. 808–818, Aug. 2019.
[7] G. B. Brahim, ‘‘Predicting student performance from online engagement activities using novel statistical features,’’ Arabian J. Sci. Eng., vol. 47, no. 8, pp. 10225–10243, Aug. 2022.
[8] M. M. Eid. (Oct. 28, 2022). MouseDynamicsDatasetForRCL. [Online]. Available: https://github.com/ErrorLogic1211/MouseDynamics DatasetForRCL
[9] TinkerCad Online Circuit Simulator. Accessed: Jan. 21, 2022. [Online]. Available: https://www.tinkercad.com/

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)