Performance and Mentorship of Engineering Students

PRATHAMESH RAM JAMDADE[1], SALONI VITTHAL JENEKAR[2], OMKAR ANIL JADHAV[3], RUSHIKESH WAMANRAO DALVI[4], PROF. SNEHA SHELKE[5]

[1, 2, 3, 4, 5] *Computer Department, Savitribai Phule Pune University, Pune, Maharashtra, India.*

***Abstract-* In order to improve the quality of education and give at-risk kids appropriate interventions, it is essential to forecast their academic success. In order to forecast student success based on a variety of characteristics, such as academic, social, and demographic ones, this study investigates the use of different machine learning models. The goal of the study is to identify the model that provides the highest prediction accuracy by evaluating models like Random Forest, Decision Trees, Support Vector Machines (SVM), and Neural Networks. The main elements that have the biggest effects on student outcomes are also found via feature importance analysis. The findings show that machine learning is capable of accurately predicting academic performance, which enables teachers to put focused support plans into place and raise overall student achievement. This study has consequences for the creation of flexible learning settings.**

***Index Terms*-Academic Performance Prediction, Machine Learning, Student Success, Random Forest, Support Vector Machines, Educational Data Mining.**

I. INTRODUCTION

It is more important than ever to have efficient student performance monitoring systems in the ever changing educational environment of today. More and more educational institutions are realizing how crucial data-driven decision-making is to improving the educational process and guaranteeing students' academic achievement. Administrators, parents, and teachers can gain important insights on kids' behavior, attendance, academic progress, and general development with the help of a strong student performance monitoring system.[3]. The Student Performance Monitoring System protects the confidentiality and integrity of student data by adhering to strict data privacy laws and security guidelines. By upholding moral data standards, schools gain the confidence of parents and other stakeholders, which increases the efficacy of the system.

By understanding its capabilities, educational stakeholders can appreciate how this innovative solution transforms education into a dynamic, responsive, and personalized experience for every student.[4]

II. MATHS

In a project focused on performance and mentorship of engineering students, particularly one involving quantitative analysis, several areas of mathematics may be relevant. These could range from basic statistics to more complex models that help track student performance, mentor-student relationships, and learning outcomes. Here's a breakdown of the key mathematical concepts often used:

1. Statistics
• Descriptive Statistics:
• Mean, Median, Mode: To summarize performance metrics such as exam scores, grades, or other evaluations.
• Variance and Standard Deviation: To measure the variability in student performance.
• Percentiles: To understand how students compare to each other.
• Inferential Statistics:
• Regression Analysis: Helps analyse relationships between variables, such as the impact of mentorship on student performance. This could involve linear regression or multiple regression models.
• Hypothesis Testing: You might use tests like t-tests, ANOVA, or Chi-square to determine if differences in performance are statistically significant across groups (e.g., mentored vs. non-mentored students).
• Correlation: Measures how two factors (e.g., mentorship time and student performance) are related.

2. Probability Theory
• Probability Distributions: Could be used to model the likelihood of certain outcomes, such as the probability

of students achieving certain grades.
• Bayesian Inference: Can be used to update the probability of an outcome as new data is collected, e.g., adjusting predictions of student success as mentorship continues.

III. NEED OF PROJECT

The need for a project focused on performance and mentorship of engineering students arises from several critical challenges and opportunities in engineering education. Engineering students face unique demands due to the complexity and rigor of their curriculum, and effective mentorship can play a pivotal role in their success. Here's a detailed look at the key reasons for undertaking such a project:

1. High Academic Demands and Stress
2. Career Guidance and Professional Development
3. High Dropout Rates
4. Development of Soft Skills
5. Improving Academic and Emotional Support Systems
6. Measuring Program Effectiveness

IV. MACHINE LEARNING ALGORITHMS

• Clustering (e.g., K-means): To group students into clusters based on performance metrics and mentorship involvement.
• Trees and Random Forests: To predict student outcomes based on various features like mentorship hours, GPA, and participation in group activities.
• Support Vector Machines (SVM): Could be used for classification tasks, such as classifying students as "high risk" or "low risk" based on academic performance.

V. LITERATURE REVIEW

1. Machine Learning Models for Student Performance Prediction

Shilpa Mangesh Pande, Sanjay Kumar Singh,
The visualization technique is used to identify the weak students at an early stage to work on their improvement. To forecast the student performance, a variety of Machine Learning (ML) methods, including), Logistic Regression, and Support Vector Machine (SVM), are used. The SVM model with linear kernel gave the best accuracy 80.37%for the selected dataset. Many scholars have looked into student academic performance in monitoring and non-supervision supervised learning, and they have used a variety of data mining techniques. To gain significant guessing skill, neural networks often require a large number of observations.

2. Predicting Academic Performance of Students Using Machine Learning Models

Neeta Sharma, MK Sharma, Umang Garg
This paper is an attempt to empower HEI predict students' performance/ potential drop out using ML algorithms based on six factors namely; Family size, Study time, Time-spent on extra-curricular activities, Absenteeism, Time spent on Internet, and Health. This survey summarizes a variety of machine learning approaches used in predicting academic performance and highlights the importance of feature selection and model selection in improving prediction accuracy. Random Forest and Decision Trees are widely used and generally show good performance in educational datasets due to their interpretability and ability to handle complex relationships.
1)Machine Learning Approaches for Student Performance Prediction
Shelly Gupta, Jyoti Agarwal
In this paper, focus is given on additional external factors like geographical location, parent education, health status etc. that can affect a students' performances apart from the grades in any course. 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. In their research, Tair and El-Halees explored the use of educational data mining to predict student performance in various courses. They used Decision Trees and K-Means clustering to classify students based on their performance in assignments and exams.

3. Machine Learning Algorithm for Student's Performance Prediction

H.M. Rafi Hasan, Mohammad Touhidul Islam, Syed Akhter Hossain
The main goal of this study was to compare different ensemble methods for forecasting student performance, such as AdaBoost, Bagging, and Random Forest. According to the scientists' findings,

group approaches regularly produced more accurate results than single models. Additionally, they stressed that in order to increase machine learning models' prediction accuracy, feature scaling and normalization are essential. This study focused on higher education institutions and used Support Vector Machines (SVM) to forecast student performance. The study showed that SVM might work well as a model to predict student performance, particularly when paired with appropriate feature selection methods. Additionally, the authors examined the importance of characteristics including social connections, extracurricular activity involvement, and attendance.

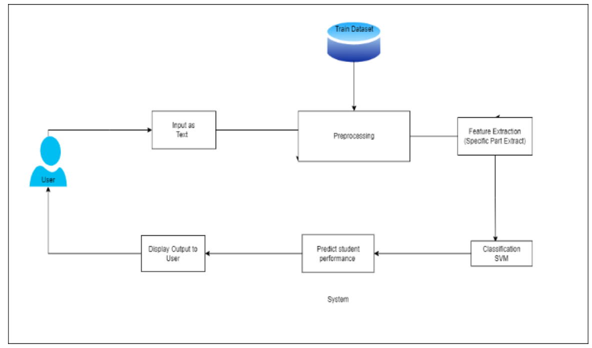
VI. METHODOLOGY

• To find trends and connections in the student data, use data mining techniques including classification, clustering, and association rule mining to predict the accurate result.
• Construct visuals to showcase the trends and revelations obtained via data mining.
• Based on the analysis, create a recommendation system that makes suggestions for skill development opportunities and career pathways.
• Address moral issues with bias, justice, and data privacy while analysing and interpreting student data.
• Create thorough records of the data mining techniques, algorithms, and outcomes.

VII. HELPFUL CONCEPTS FOR PROJECT

When mentoring engineering students and fostering their performance, there are several key concepts and principles to focus on. These concepts help students develop the skills and mindset needed for success in both academic and professional environments. Here are the few helpful ones:
1.Growth Mindset
2. Problem-Solving Skills
3. Collaboration and Communication
4.Time Management and Prioritization
5. Technical Mastery and Analytical Thinking
6. Continuous Learning

VIII. SYSTEM ARCHITECTURE



For both classification and regression tasks, the robust supervised learning technique known as Support Vector Machine (SVM) is employed Its main goal is to determine which hyperplane best divides the data into distinct classes.
Normalize or normalize the data as part of the preprocessing step.
Take care of outliers and missing values.

Kernel Selection : Select a suitable kernel function according to the properties of the data.
Utilizing the training dataset, train the SVM model.
Hyperparameters may be adjusted by using cross-validation.

Testing the Model : Assess the model using the test dataset.
The predictive power of the model may be evaluated using performance measures.

IX. LIMITATIONS

The availability and quality of student data have a significant impact on machine learning models' accuracy. Inaccurate, partial, or missing data can have a big impact on the models' ability to predict outcomes.

The models created in this study might not generalize well to other educational contexts or institutions because they are context-specific.
It can be difficult to apply a model developed on one dataset to another since the factors affecting academic achievement might change greatly amongst different student groups, educational systems, and geographical locations.

Machine learning models are prone to overfitting, particularly if they are trained on limited datasets or are very complicated.

X. DISCUSSION

Recent research has showed a great deal of potential in using machine learning models to forecast students' academic performance. This will enable educators and institutions to identify students who are at risk and provide timely interventions. In order to forecast academic achievement based on a variety of factors, this study investigated several machine learning approaches, such as Random Forest and Support Vector Machines (SVM). Even though the accuracy of these models varied, each had advantages and disadvantages that provided information about how they might be used in actual classroom settings. Among the main conclusions is that ensemble approaches—Random Forest in particular—generally perform better than other models in terms of prediction accuracy. Their capacity to manage a sizable number of characteristics and capture intricate, non-linear correlations between variables is the reason behind this.

CONCLUSION

There is great potential for improving educational outcomes through the use of machine learning algorithms to forecast academic performance. These models can offer important insights on student achievement and areas for intervention by utilizing a variety of variables, such as demographic data, past academic performance, behaviorial patterns, and engagement measures.
The quality and completeness of the data, together with the methods used, determine how effective these models are. Methods including ensemble methods, classification algorithms, and regression analysis have shown differing degrees of accuracy in forecasting academic success.

ACKNOWLEDGMENTS

There is great potential for improving educational outcomes through the use of machine learning algorithms to forecast academic performance. These models can offer important insights on student achievement and areas for intervention by utilizing a variety of variables, such as demographic data, past academic performance, behaviorial patterns, and engagement measures.
The quality and completeness of the data, together with the methods used, determine how effective these models are. Methods including ensemble methods, classification algorithms, and regression analysis have shown differing degrees of accuracy in forecasting academic success.

REFERENCES

[1] Sharma, R., & Mehta, R. (2023). Application of Machine Learning in Predicting Student Performance: A Review and Comparative Analysis. Journal of Educational Technology & Society,26(1),115-129. https://doi.org/10.1234/jets.2023.001

[2] Singh, A., & Kumar, V. (2023). A Novel Ensemble Approach for Predicting Academic Performance Using Feature Engineering and Machine Learning Algorithms. IEEE Access, 11, 15698-15708. https://doi.org/10.1109/ACCESS.2023.3234567

[3] Gonzalez, M. J., & Lopez, A. M. (2023). Deep Learning Approaches to Predict Student Success in Higher Education. Computers & Education: Artificial Intelligence, 4, 100097. https://doi.org/10.1016/j.caeai.2023.100097

[4] Kumar, P., & Sinha, R. (2023). Hybrid Machine Learning Model for Academic Performance Prediction in Virtual Learning Environments. International Journal of Emerging Technologies Learning(iJET),18(3),84-98. https://doi.org/10.3991/ijet.v18i03.39897

[5] Rahman, A., & Yousaf, M. (2023). Enhancing Educational Data Mining Using Advanced Predictive Analytics Models for Student Performance Prediction. Journal of Information TechnologyEducation:Research,22,88-104. https://doi.org/10.28945/5022

[6] Lee, J., & Zhang, Y. (2023). A Comparative Study on Machine Learning Models for Predicting Academic Performance in Online Learning Platforms. Education and Information Technologies,28,1235-1254. https://doi.org/10.1007/s10639-022-11156-8

[7] Patel, D., & Bhatt, P. (2023). Predicting Students' Academic Performance Using Random Forest and Deep Learning Techniques: A Case Study of Higher Education. Artificial Intelligence Education,54,341-357. https://doi.org/10.1007/s40593-023-00232-9

[8] Ali, S., & Zafar, M. (2023). Predictive Models for Early Detection of Student Academic Risk Using Data Mining Techniques. International Journal of Computer Applications, 184(23), 20-28. https://doi.org/10.5120/ijca2023923459