A COMPREHENSIVE STUDY FOR IDENTIFICATION OF FAST AND SLOW LEARNERS USING MACHINE LEARNING APPROACH

Shrawani Rayalkar[1], Chirag Nere[2], Dr. Hrushikesh Kulkarni[2*]

[1]*School of Data science,*[2] *School of Mechatronics Engineering, Symbiosis Skills and Professional University, Pune*

---***---

Abstract - *This research paper presents a detailed investigation into the identification of fast and slow learners in an academic setting through the implementation of a machine learning-based ensemble model. The proposed ensemble model utilizes a training dataset to generate rules, subsequently applied to a testing dataset for predicting academic performance.*

The goal is to have a better understanding of how students learn and identify the settings in which they learn to improve educational outcomes and to gain insights into and explain educational phenomena. The biggest challenge is to improve the quality of the educational processes so as to enhance student's performance. Thus, it is crucial to set new strategies and plans for a better management of the current processes. It is concerned with developing methods for exploring the unique types of data that come from educational environments. The paper uses trains various machine learning model using real time data to find the prediction and accuracy of the models and hence finds out which works accurately and hence helping in academic actions can be taken for each student accordingly

Experimental results, conducted using a dataset from a computer science department, demonstrate the model's efficacy with an accuracy of 90.83%. The ensemble model not only provides valuable insights for instructors to scrutinize and enhance student performance but also aids learners in self-assessment and corrective actions. The study concludes with potential extensions, suggesting the development of a recommender system for course selection based on performance predictions and incorporating indirect factors affecting academic performance. The integration of predictive analytics in the evolving landscape of educational technology is also discussed, highlighting its potential in assisting students through personalized recommendations based on their predicted performance.

Key Words: *Slow learners, fast learners, Education performance, machine learning*

1. INTRODUCTION

Educational data has become an important resource in this modern time, contributing much to the welfare and growth of the society. Educational system is becoming more competitive because of the number of schools and institutions which are growing rapidly. The educational schools are focusing more on improving various aspects and one important factor among them is quality learning.

A fast learner is someone who gains the skills of being a strategic thinker and a good listener and applies them to learning quickly. A fast learner not only learns things at small interval of time but also gains enough knowledge than a normal person.A slow learner work on tasks more slowly, have poor memory and difficulties understanding concepts and subject taught to them.

Prediction of student's performance is challenging task as it depends on many factors such as grades, class performance, demographic data and emotional features. It is important for the teachers to forecast the future performance of a student based on his past performances, identifying weak students at an early stage so that additional material and special attention can be facilitated to avoid the risk of failure [1].

Besides this, various statutory and regulatory bodies such as National Assessment and Accreditation Council (NAAC) and the National Accreditation Board (NBA) which are accrediting the higher education institutions also stress the identification of the learning levels of the students and accordingly steer the teaching–learning process for them. If the weak learners are identified at the start of the semester, the respective subject teachers can plan their academic activity for a better understanding of the subject matter by such students, and their results improve as they go to higher semesters. To improve the students' performance by tailored teaching–learning activities for the slow and fast learners is the key outcome of this research [3].

2. Materials and Methods

The methodology for this research involved the identification of fast and slow learners in a college environment to facilitate the adaptation of academic teaching methods for personalized learning. The study utilized a dataset comprising records of approximately 1000 students, encompassing their performance marks in three subjects and the corresponding number of study hours per day. Five distinct machine learning models, namely Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), were employed to classify students into fast, slow, and average learners based on their academic achievements and study habits. The models underwent training and evaluation processes, with performance metrics such as accuracy, precision, recall, and F1 score assessed to determine their effectiveness. The results revealed that SVM exhibited the highest accuracy, while Logistic Regression showed the lowest. The findings from this comprehensive methodology aim to inform educational institutions about potential modifications in teaching approaches to cater to the diverse learning paces of students, thereby contributing to a more effective and inclusive educational environment.

A literature review was presented using a broad array of data about students and courses collected by institutions and learning analytics to improve student's success and retention [5]. Academic analytics measure, collect, decipher, report, effectively share data, and identify student strengths and weaknesses [6].

3. Results and discussion

The research evaluated five machine learning models, each with its unique characteristics and applicability. The Support Vector Machine (SVM) emerged as the most accurate model, achieving an accuracy of 97%. The Logistic Regression model demonstrated an accuracy of 86%, while the Decision Tree, Random Forest, and KNN models showed accuracies of 90%, 93%, and 87%, respectively as mentioned in Table 1. The study provides a detailed breakdown of precision, recall, and F1-score for each model, offering insights into their performance across different learner categories.

Logistic Regression can be employed to classify students into fast and slow learners based on their academic performance, study time, and attendance. It is effective when the relationship between input features and the output is approximately linear. It models the probability of an instance belonging to a particular class, making it suitable for binary classification tasks. Logistic Regression estimates the coefficients for each input feature, applying a logistic function to the linear combination of these coefficients. This maps the input features to a probability between 0 and 1, and a threshold is applied to make the final classification.

Decision Trees are applicable to classify students based on their academic performance, study time, and attendance. It can capture non-linear relationships between features and the target variable. They partition the feature space based on the most informative features, creating a tree structure that facilitates classification. Decision Trees make decisions by recursively splitting the data based on the feature that maximally separates the classes. The decision at each node is determined by feature values, and the final classification is made at the leaf nodes.

Random Forest can enhance the classification accuracy by combining multiple Decision Trees for categorizing fast and slow learners. It mitigates overfitting by aggregating predictions from various Decision Trees. It introduces randomness in the tree-building process, leading to diverse trees and increased robustness. Random Forest builds multiple Decision Trees on different subsets of the data and averages their predictions. This ensemble approach reduces overfitting and improves generalization.

SVM can be employed for binary classification of fast and slow learners based on their academic performance and study time. SVM is effective in high-dimensional spaces and can handle non-linear relationships through the kernel trick. It identifies the hyperplane that best separates the classes. SVM seeks the hyperplane with the maximum margin between classes. It transforms the data into a higher-dimensional space if needed, allowing for better separation. The support vectors are instances closest to the decision boundary.

KNN can classify students into fast and slow learners based on the similarity of their academic performance, study time, and attendance. KNN is suitable for capturing complex relationships and works well with small to moderately-sized datasets. It classifies instances based on the majority class of their k-nearest neighbors. KNN calculates distances between instances in the feature space and assigns a class based on the majority class among the k-nearest neighbors. It's a lazy learner as it does not build an explicit model during training.

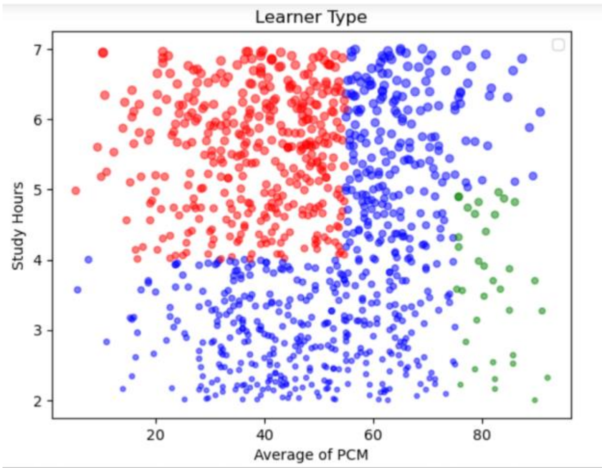

Fig -1: Plot showing the average of 3 subjects against the study hours where RED is slow learners, BLUE is average learners and GREEN is fast learners.

Table 1: Accuracy of models and their classification report for each category

Sr No.	Model Name	Accuracy	Category	Precision	Recall	F1 score
1	Logistic regression	0.86	Average	0.84	0.94	0.89
			Fast	1	0.3	0.46
			Slow	0.89	0.81	0.85
2	Decision Tree	0.9	Average	0.33	0.5	0.4
			Fast	1	0.94	0.97
			Slow	0.92	0.92	0.92
3	Random Forest	0.93	Average	0.5	1	0.67
			Fast	1	0.94	0.97
			Slow	1	0.92	0.96
4	Support Vector Machine	0.97	Average	0.67	1	0.8
			Fast	1	0.94	0.97
			Slow	1	1	1
5	KNN	0.87	Average	0.25	0.5	0.33
			Fast	0.94	1	0.97
			Slow	1	0.75	0.86

The scatter plot in Figure 1. visually illustrates the distribution of 1000 students based on study hours and marks, highlighting predominant average and slow learners. The classification is depicted using various colors: red represents slow learners who spend more time studying but achieve lower marks, green represents fast learners who study for shorter durations but attain higher marks, and blue represents average learners falling in between.

Upon examination, it becomes evident that the majority of students in the dataset are either average or slow learners. The x-axis represents the study hours, ranging from 2 to 7 hours, while the y-axis depicts the marks achieved, spanning from 0 to 100. This visualization allows for a clear understanding of the distribution of students across different learner categories, highlighting the diverse study habits and academic outcomes within the dataset. Notably, a substantial portion of students falls into the average learner category, indicating a balance between study hours and academic performance.

On the other hand, a significant number of slow learners, characterized by extended study hours and comparatively lower marks, are also present. Additionally, the plot showcases a distinct group of fast learners who achieve high marks despite investing fewer hours in studying. Machine learning models, including SVM with 97% accuracy, demonstrate efficacy in learner classification. The outcomes can inform educational strategies, aiding early identification of underperforming students.

Implementing these models in the education system enables personalized interventions. For instance, targeted support mechanisms for slow learners, advanced challenges for fast learners, and tailored approaches for average learners could optimize academic outcomes. Integrating predictive analytics fosters a data-driven educational environment, guiding educators to proactively address student needs and enhance overall learning outcomes.

Despite providing valuable insights, the study has limitations, including a relatively small and potentially unrepresentative sample, a focus on limited features, and the analysis primarily considers marks in subjects and study hours as features. Other relevant factors influencing academic performance, such as extracurricular activities, socio-economic background, or mental health aspects, are not included. The exclusion of these variables could impact the depth of the model's understanding.

The analysis primarily considers marks in subjects and study hours as features. Other relevant factors influencing academic performance, such as extracurricular activities, socio-economic background, or mental health aspects, are not included. The exclusion of these variables could impact the depth of the model's understanding. The representativeness of this sample in comparison to a larger student population or different academic disciplines could be a limitation. Variability in academic settings and student demographics may influence the generalizability of the results.

4. Conclusion

It is paramount to recognize and accommodate the diverse learning profiles exhibited by students. By tailoring educational strategies to meet the needs of both advanced and slow learners, educational institutions can establish an environment that not only challenges the intellectually curious but also provides support for those encountering educational obstacles. The proposed interventions outlined in this comprehensive plan aim to address individual needs and offer the essential resources required for academic success.

Our prior discussion on the application of Support Vector Machines (SVM) in the education system aligns seamlessly with the call for data-driven approaches to identify fast and slow learners. Integrating machine learning models, as previously explored, can enhance these initiatives by providing valuable insights derived from data patterns, empowering educators to make well-informed decisions.

In conclusion, a holistic approach that combines targeted educational interventions with data-driven methodologies is vital for fostering a more inclusive, responsive, and effective educational system.

This approach not only supports academic excellence but also addresses the distinctive challenges faced by students, ensuring a comprehensive and enriching educational experience. Conclusions drawn from this work include the identification of principles and generalizations from the results, recognition of exceptions or challenges to these principles, exploration of theoretical and practical implications, and the formulation of recommendations based on the findings. This multifaceted approach contributes to the creation of a well-rounded educational system that caters to the diverse needs of students, fostering an environment conducive to both academic achievement and personal growth.

Initiatives for Advanced Learners(SRM)

The strategies for advanced learners focus on providing opportunities for intellectual growth and challenging their capabilities. Offering separate assignments with challenging problems, encouraging participation in advanced projects, and involving research scholars in seminars contribute to a stimulating academic atmosphere. Additionally, initiatives such as career guidance, participation in conferences, and exposure to advanced legal topics through debates aim to broaden their perspectives. Encouraging research projects and paper publications, along with activities in clubs and committees, helps nurture a spirit of innovation and leadership. The provision of individual mentorship, specialized projects, and participation in technical events further supports their academic journey [2].

Initiatives for Slow Learners

For slow learners, the initiatives are geared towards providing tailored support to address specific weaknesses. Special classes, revision of courses, and group study systems aim to reinforce foundational knowledge. Remedial classes and individual counseling by subject teachers offer personalized assistance, while forming study groups facilitates peer-to-peer learning. Special assignments and varied instructional techniques help in developing a better understanding of difficult topics. Communication with parents and systematic performance updates ensure a collaborative effort in the student's educational journey. Initiatives like NPTEL sessions, workshops, and compensatory teaching provide additional resources to bridge gaps in understanding. The emphasis on self-learning materials and multiple assessments helps in gradual improvement, while leadership activities and regular counseling sessions aim to boost confidence and foster positive interactions with parents [2].

3. Acknowledgment

We would like to express my sincere gratitude to all those who contributed to the completion of this research paper. We are also thankful to the participants who provided valuable data for the analysis. This work would not have been possible without the encouragement and assistance from colleagues and friends. Your insights and feedback have been instrumental in shaping the outcome of this research. We are deeply appreciative of the collaborative efforts that made this project a reality.

REFERENCES

1. Bharadi, V. A., Prasad, K. K., & Mulye, Y. G. (2023). Classification of Slow and Fast Learners Using Deep Learning Model. In Computational Intelligence: Select Proceedings of InCITe 2022 (pp. 461-472). Singapore: Springer Nature Singapore.

2. Bharadi, V. A., Prasad, K. K., & Mulye, Y. G. (2020). Using Deep Learning Techniques for the Classification of Slow and Fast Learners. Department of Information Technology, 135.

3. Bin Mat U, Buniyamin N, Arsad P. M., Kassim R., (2013). "An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention" IEEE 5th Conference on Engineering Education (ICEED), 126-130, DOI: 10.1109/ICEED.2013.6908316

4. Baepler, Paul Murdoch, Cynthia James, (2010) "Academic Analytics and Data Mining in Higher Education," International Journal for the Scholarship of Teaching and Learning: 4(2):17.

5. Vital, T. P., Sangeeta, K., & Kumar, K. K. (2021). Student classification based on cognitive abilities and predicting learning performances using machine learning models. International Journal of Computing and Digital Systems, 10(1), 63-75.

6. ”The institution assesses the learning levels of the students and organies special Programmes for advanced learners and slow learners” weblink-https://www.gdcthannamandi.com/naacfile/Special%20Programmes%20for%20advanced%20lerners.pdf