International Journal of Advance and Applied Research
www.ijaar.co.in
ISSN – 2347-7075 Impact Factor – 8.141
Peer Reviewed Bi-Monthly
Vol. 6 No. 22 March - April - 2025

Enhancing Student Performance Prediction and Course Recommendation Using Machine Learning

Kalokhe Anil Sopan[1], Kharade Vaishnavi Santosh[2], Shinde Gauri Krushnath[3], Kumbhar Vijaykumar Sambhajirao[4]

[1]*Research Scholar, Department of Computer Science, Shivaji University, Kolhapur (MH), India.*

[2]*Student, Department of BBA(CA), Vidya Pratishthan's, Arts, Science and Commerce College, Baramati, Pune (MH), India*

[3]*Student, Department of BBA(CA), Vidya Pratishthan's, Arts, Science and Commerce College, Baramati, Pune (MH), India*

[4]*Research Guide, Department of Computer Science, Shivaji University, Kolhapur (MH), India.*

Corresponding Author – Kalokhe Anil Sopan

DOI - 10.5281/zenodo.15296422

Abstract:

Predicting student performance is crucial in education, enabling early interventions and personalized learning. Machine learning (ML) algorithms offer data-driven methods to forecast academic success and identify at-risk students. Features such as demographics, study habits, attendance, and prior academic performance are analysed to build predictive models. The research compares the accuracy of different supervised learning algorithms in forecasting final grades or test scores. The results show that ML can enhance predictions, providing valuable insights for tailored educational strategies and early interventions. Key factors like data preprocessing, and model evaluation are essential for achieving accurate predictions. Additionally, the paper addresses ethical concerns around ML use in education, emphasizing that these tools should complement, not replace, human judgment in decision-making processes.

Keywords: Educational data mining, E-learning, Machine learning, Student academic performance prediction, Support Vector Machine

Introduction:

The ability to predict student performance is a key challenge in the field of education, with significant implications for both educators and learners. Using educational machine learning methods, we could potentially improve the performance and progress of students more efficiently in an efficient manner.

The real goal is to have an overview of the systems of artificial intelligence that were used to predict academic learning. Student's performance in the educational process can literally be defined as something that is obtained from changes in the behavior of students based on their experiences, besides that learning outcomes are also a realization of the potential or capacity possessed by students [1].This research paper explores the application of various machine learning techniques to predict student performance, aiming to offer early intervention strategies and support personalized learning pathways [2]. Students, educator and academic institutions could benefit and also have an impact [3]. With the rapid growth of data and computational power, machine learning (ML) has emerged as a powerful tool to address this gap. By analysing historical

data, including student demographics, academic records, behavioural patterns, and engagement metrics, ML algorithms can predict future academic performance with remarkable accuracy.

Students' performance is a term used for measuring not only students' achievements but also the quality of educational institutions [4]. Traditional methods of assessing student outcomes often rely on periodic evaluations and subjective judgments, which may fail to identify students at risk of underperforming in a timely manner. Analysing this data provides insight into the student learning process and student achievement. Further analysis can identify academic, demographic, and social factors affecting student academic success. Student academic success is measured by assessing student performance across academic subjects [5].

The study also examines the importance of feature selection, data pre-processing, and model evaluation in improving prediction accuracy. Prediction of student's performance can be used as a basis for early intervention on the potential failure of students to achieve learning objectives; and at the same time able to make changes to learning strategies in order to facilitate student diversity [6].

The specific focus of this thesis is education. The aim is to predict student performance. Data about students is used to create a model that can predict whether the student is successful or not, based on other properties [7].Techniques such as decision trees, support vector machines, and neural networks are assessed for their predictive capabilities in determining student outcomes, such as final grades or standardized test scores.

A variety of DM methods are employed to predict student performance, identify slow learners, and dropouts (Hardman et al., 2013; Kaur et al., 2015). Early prediction is a new phenomenon that includes assessment methods to support students by proposing appropriate corrective strategies and policies in this field [8]. With rapid advancements in artificial intelligence and computer science, numerous advanced AI technologies have found broad application within engineering systems. Con currently, machine learning has attracted considerable attention in Educational Data Mining (EDM) and has been utilized extensively for predicting students' academic performance [9]. The objective of this study is to analyse the role of machine learning in predicting student performance and assess the accuracy and reliability of different models. By doing so, we aim to contribute to the growing body of knowledge in educational data mining and artificial intelligence, ultimately supporting efforts to improve student success rates and learning experiences.

Literature Review:

Lidia Sandra, Ford Lumbangaol, Tokuro Matsuo [1], stated that the findings from this systematic literature are expected to contribute to understanding for educational institutions to be able to use the most appropriate machine learning algorithm based on their needs and datasets to ensure the achievement of the ultimate learning goal, namely the learning success of students. Further research on the diversity of datasets used to predict student learning achievement and the development of machine learning algorithms and models that are able to facilitate the diversity of datasets to predict student learning achievement will be very useful for further development.

J. Dhilipan1, N. Vijayalakshmi, S. Suriya, Rockeye Christopher [2], focuses on the student academic growth analysis using machine learning techniques. For analysis Binomial logical regression, Decision tree, Entropy and KNN classifier are used. This process can help the instructor to decide easily about performance of the students and

schedule better method for improving their academics. In future additional features are added to our dataset to acquire better accuracy.

Jovana Jović, Emilija Kisić, Miroslava Ras Popovic Milić, Dragan Domazet, Kavitha Chandra [3], observed that student academic performance is a key quality indicator for universities. In this study, we evaluated the accuracy of six machine learning algorithms to predict whether students will pass or fail the final exam. The algorithms tested were NB, LDA, LR, DT, KNN, and SVM, and were assessed based on accuracy and precision. Future work will explore more machine learning algorithms and include additional features to improve the model for predicting student performance and supporting the learning process.

Research Methodology:

To predict student performance, start by clearly defining the type of performance you're predicting, such as final exam scores, overall grades, likelihood of passing, or identifying at-risk students. Next, gather relevant data, which can include demographic information, attendance, participation, homework scores, previous academic history, and study habits. Once the data is collected, perform data pre-processing by cleaning the data, handling missing values, and detecting outliers.

To analyse student performance data, we utilize fundamental statistical techniques such as mean and correlation analysis. The mean (average) helps summarize key student behaviors, such as the typical number of study hours per week or the average assignment completion rate. This provides a central value for comparison. Additionally, correlation analysis is used to identify relationships between variables, such as the impact of attendance on final grades. Pearson's correlation coefficient measures the strength of these relationships, helping to determine whether factors like study time have a significant influence on academic success. Furthermore, to evaluate the predictive accuracy of machine learning models, metrics such as accuracy and precision are applied. These statistical approaches ensure a data-driven evaluation, improving the reliability of student performance predictions.

Result and Discussion:

Result:

The results of this research paper indicate that machine learning (ML) algorithms, such as decision trees, support vector machines, and neural networks, can effectively predict student performance using historical data. The study finds that factors like demographics, study habits, attendance, and prior academic performance play a significant role in forecasting final grades or test scores. A comparative analysis of different supervised learning techniques demonstrates that ML models enhance prediction accuracy, offering valuable insights for personalized learning and early interventions. Additionally, the study highlights the importance of feature selection, data preprocessing, and model evaluation in improving predictive performance. While ML proves to be a powerful tool in education, ethical concerns related to data privacy and fairness must be considered to ensure responsible implementation**.**

Discussion:

Question 1: How do you rate your understanding of course materials? With the answer options Excellent – I grasp concepts easily, Good – I understand most concepts with some effort, Average – I.

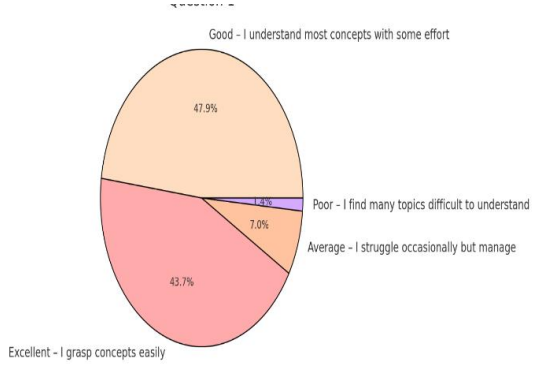

Figure1: The result of understanding of course materials

The pie chart represents the distribution of responses to a question about self-assessed understanding of concepts. The different segments indicate various levels of comprehension: Excellent (43.4%): A significant portion of respondents find it easy to grasp concepts. Good (46.9%): The largest segment, indicating that most people understand concepts with some effort. Average (7.6%): A smaller portion struggles occasionally but manages. Poor (0.4%): A very small number of people find many topics difficult to understand. Very Poor (0.4%): An equally small percentage requires significant help to understand concepts.

Overall, the majority of respondents rate themselves as either "Good" or "Excellent," suggesting that most people feel confident in their understanding, with only a minor percentage struggling significantly.

Question 2: How often do your complete assignments on time? Always – I submit all assignments before the deadline, often – I meet deadlines most of the time, sometimes – I struggle but manage to submit, rarely – I frequently submit late, never – I often miss deadlines.

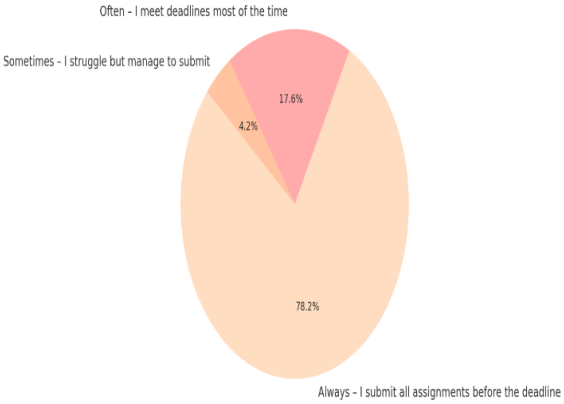

Figure2: The result of do your complete assignments on time

The pie chart shows that most respondents (78.2%) consistently submit assignments on time. A smaller group (17.6%) usually meets deadlines but may occasionally miss them. Only a few (4.2%) struggles but still manage to submit. Overall, the majority have strong time management skills, while a small percentage find deadlines challenging.

Question 3: How many hours do you study outside of class per week? More than 20 hours, 15 – 20 hours,10 – 14 hours,5 – 9 hours, Less than 5 hours.

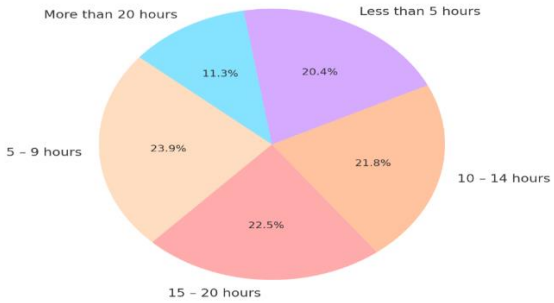

Figure3: The result of how many hours do you study

The pie chart represents the distribution of time spent on a particular activity per week. The largest portion (23.9%) spends 5–9 hours, followed closely by those spending 15–20 hours (22.5%) and 10–14 hours (21.8%). A smaller group (20.4%) spends less than 5 hours, while the smallest percentage (11.3%) spends more than 20 hours. This suggests that most

people dedicate a moderate amount of time, with fewer individuals at the extreme ends of the spectrum.

Question 4: How actively do you participate in class discussions? Very actively – I contribute regularly, occasionally – I participate when confident, rarely – I speak only when required, never – I prefer to listen rather than speak.

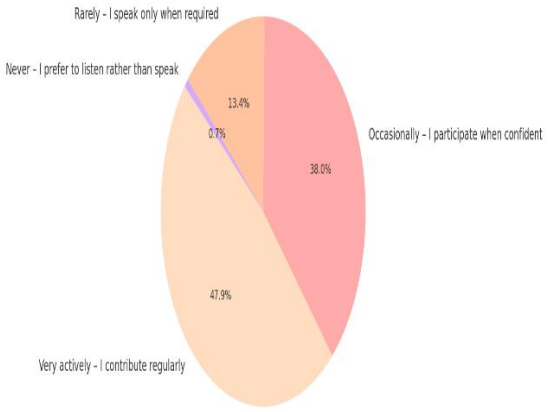

Figure 4: The result of actively do you participate in class discussions

The pie chart represents participation levels in discussions. The largest group (47.9%) contributes very actively, followed by those who participate occasionally when confident (38.0%). A smaller portion (13.4%) speaks only when required, while a minimal percentage (0.7%) prefers to listen rather than speak. This suggests that most individuals engage in discussions, with only a few being passive participants.

Question 5: What study methods do you use most frequently? Self-study, Group study, Online resources (YouTube, Coursera, etc.), Private tutoring, University workshops.

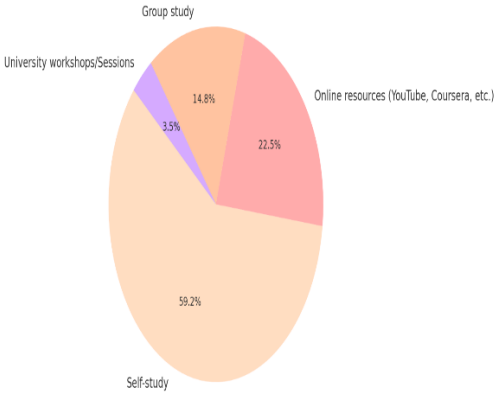

Figure 5: The result of What study methods do you use most frequently

The pie chart illustrates preferred study methods. The majority (59.2%) rely on self-study, followed by online resources like YouTube and Coursera (22.5%). Group study accounts for 14.8%, while only a small portion (3.5%) attends university workshops or sessions. This indicates that most learners prefer independent study, with online resources being a popular supplementary method.

Question 6: Do you use academic resources such as the library, online journals, or workshops? Yes, regularly – I use them often ,Occasionally – Only when needed , Rarely – I prefer other study methods , Never – I don't use academic resources.

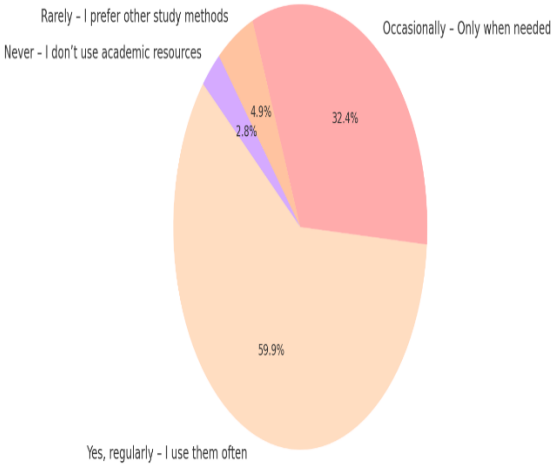

Figure 6: The result of academic resources such as the library, online journals, or workshops

The pie chart illustrates the usage of academic resources. A majority (59.9%) use them regularly, while 32.4% access them only when needed. A smaller portion (4.9%) rarely uses them, preferring other study methods, and 2.8% never use academic resources. This indicates that most learners rely on academic resources, though some use them selectively or not at all.

Question 7: How effectively do you manage your time between coursework and other responsibilities? Very effectively – I have a structured schedule, Moderately well – I manage most tasks on time, Somewhat poorly – I struggle with balancing task, Very poorly – I often miss deadlines and feel overwhelmed.

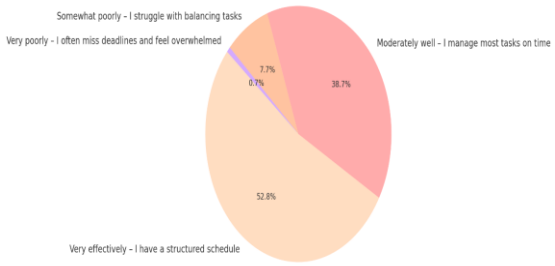

Figure 7: The result of manage your time between coursework and other responsibilities

The pie chart represents task management efficiency. The majority (52.8%) manage tasks very effectively with a structured schedule, while 38.7% handle most tasks on time. A smaller group (7.7%) struggles with balancing tasks, and only 0.7% frequently miss deadlines and feel overwhelmed. This suggests that most individuals have good time management skills, with only a few facing significant challenges.

Question 8: Are you involved in extracurricular activities, internships, or research projects? Yes, multiple activities, Yes, but only one, No, but I plan to, No, not interested.

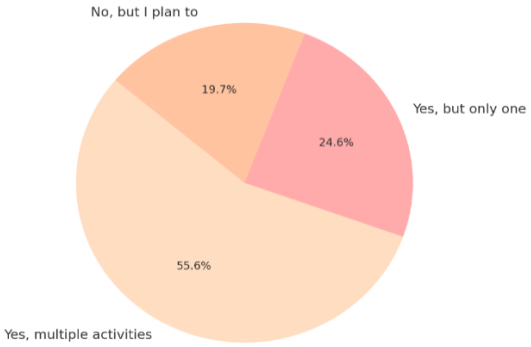

Figure8: The result of extracurricular activities, internships, or research projects

The pie chart illustrates participation in extracurricular activities. The majority (55.6%) engage in multiple activities, while 24.6% participate in only one. A smaller portion (19.7%) do not currently participate but plan to. This suggests that most individuals are actively involved in extracurriculars, with some intending to join in the future.

Question 9: What are the biggest challenges affecting your academic performance? Time management issues, Lack of motivation, financial difficulties, Personal or family issues, Difficulty understanding course material.

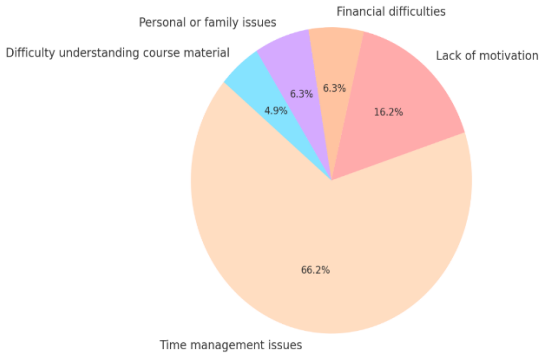

Figure9: The result of biggest challenges affecting your academic performance

The pie chart highlights the main challenges students face. The majority (66.2%) struggle with time management, while 16.2% experience a lack of motivation. Smaller percentages face

financial difficulties (6.3%), personal or family issues (6.3%), and difficulty understanding course material (4.9%). This suggests that effective time management is the biggest obstacle to student success.

Question 10: Do you feel supported by faculty and peers in your academic journey? Yes, very much – I receive a lot of support, somewhat – I get support but not always, not really – I struggle to find support, No, not at all – I feel isolated in my studies.

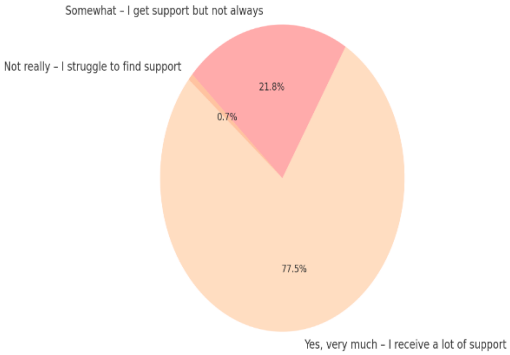

Figure 10: The result of supported by faculty and peers in your academic journey

The pie chart illustrates the level of support students receive. A significant majority (77.5%) report receiving a lot of support, while 21.8% receive some support but not always. Only a small fraction (0.7%) struggle to find support. This indicates that most students feel well-supported, though some still face occasional difficulties.

In addition to predicting student performance, integrating a course recommendation system can enhance personalized learning pathways. By leveraging machine learning algorithms, such a system can analyze factors like past academic performance, study habits, and engagement levels to suggest suitable courses or additional learning resources. Techniques like collaborative filtering and content-based filtering can recommend courses that align with a student's strengths and learning preferences. For instance, if a student demonstrates strong analytical skills in mathematics but struggles in programming, the system can suggest beginner-friendly coding courses to bridge the gap. Furthermore, classification models such as Support Vector Machines (SVM) or Decision Trees can help predict which courses a student is likely to excel in based on historical data. Integrating this approach with student performance prediction can provide a data-driven, adaptive learning environment, improving academic success rates and student satisfaction.

Conclusion:

The study demonstrates that machine learning techniques can effectively predict student performance, providing valuable insights for early interventions and personalized learning. By analysing key factors such as study habits, demographics, attendance, and prior academic performance, ML models offer a data-driven approach to identifying at-risk students. The findings emphasize the importance of proper data pre-processing, feature selection, and model evaluation to enhance accuracy. While ML presents significant advantages in forecasting student outcomes, ethical considerations, such as data privacy and the complementary role of human judgment, must be addressed. The research highlights the potential of ML in education, encouraging institutions to integrate these technologies responsibly to improve academic success.

References:

1. Lidia Sandra, Ford Lumbangaol, Tokuro Matsuo, "Machine Learning Algorithm to Predict Student's Performance: A Systematic Literature Review," *TEM Journal*. Vol-10, Issue 4, 1919-1927, ISSN 2217-309, DOI: 10.18421/TEM104-56, November 2021.
2. J.Dhilipan, Vijayalakshmi, S.Suriya, Arockiya Christopher, "Prediction of

Students Performance using Machine learning," doi:10.1088/1757-899X/1055/1/012122.

3. Jovana Jovic, Emilija Kisic, Miroslava Raspopovic Milic, Dragan Domazet, Kavitha Chandra, "Prediction of student academic performance using machine learning algorithms," *Proceedings 13th International Conference on eLearning (eLearning-2022),* September 29–30, 2022.
4. Esmael Ahmed, "Student Performance Prediction Using Machine Learning Algorithms," *Hindawi Applied Computational Intelligence and Soft Computing Volume 2024*, https://doi.org/10.1155/2024/4067721.
5. Opeyemi Ojajuni, Foluso Ayeni, Olagunju Akodu, Femi Ekanoye, Samson Adewole,TimothyAyo, Sanjay Misra, Victor Mbarika6Predicting, "Student Academic Performance Using Machine Learning," 481–491, 2021, https://doi.org/10.1007/978-3-030-87013-3_36.
6. Charalampos Dervenis, Vasileios Kyriatzis, Spyros Stoufis, Panos Fitsilis, "Predicting Students' Performance Using Machine Learning Algorithms," *Association for Computing Machinery*, September 16–18, 2022, https://doi.org/10.1145/3564982.3564990.
7. Bashru Aliyu Sani, Badamasi Haruna, "Machine Learning Algorithms To Predict Student's Academic Performance," *Bakolori Journal of General Studies, Vol. 12 No. 2*, 3656-3671.
8. Mustafa Yağcı,"Educational data mining: prediction of students' academic performance using machine learning algorithms,"1-19, https://doi.org/10.1186/s40561-022-00192-z
9. Zheng Luo, Jiahao Mai and Zhanbo Zhu, Caihong Feng, Deyao Kong, Jingyu Liu, Yunhong Ding, Bo Qi, "A Method for Prediction and Analysis of Student Performance That Combines Multi-Dimensional Features of Time and Space," *Mathematics 2024*, 12, 3597. https://doi.org/10.3390/math12223597.