International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Review: An Approach for Secondary School Students Performance using Machine Learning and Data Mining

Palak Patel[1], Tejas Thakkar[2], Mayur Patel[3], Ami Trivedi[4]

Submitted: 05/12/2023 **Revised**: 15/01/2024 **Accepted**: 25/01/2024

Abstract: The measurement of Students' Academic Performance (SAP) stands as a crucial gauge for assessing the standing of students within an academic institution. This metric enables instructors and education administrators to obtain a precise assessment of students across various courses in a specific semester. Additionally, it serves as an invaluable indicator for students to reflect on their strategies, encouraging improvements for enhanced performance in subsequent semesters. Each institution establishes its own set of criteria for evaluating student performance. This variation arises from a deficiency in researching existing prediction techniques, leading to the quest for the most effective methodology in predicting students' academic progress and performance. Another significant factor is the insufficient exploration of relevant factors influencing student achievement in specific courses. To comprehensively comprehend this issue, a thorough literature survey on predicting student performance through the application of data mining techniques is suggested. This paper objective is to enhance student academic performance by employing advanced machine learning techniques. The utilization of these techniques not only aims to elevate student outcomes but also holds the potential to yield benefits for faculty members, students, educators, and the overall institutional management.

Keywords: *Attribute selection; Ensemble learning; Educational data mining; Hybrid model; Learning analytics; Machine learning; Online learning;*

1. Introduction:

Educational data mining (EDM)is a predictive models to identify students who may be at risk of poor performance, allowing for timely intervention and support. Supporting Decision-Making for Instructors and Students assists both instructors and students in making informed decisions [1]. Instructors can use data-driven insights to tailor their teaching methods and provide targeted support to struggling students. Students, on the other hand, can benefit from personalized feedback and recommendations based on their learning patterns and performance. Monitoring and defining the learning progress of students is crucial for adapting instructional strategies [2]. EDM helps in tracking individual and group progress over time at different stages of their learning journey. The selection and optimization of algorithms are essential in EDM. Different machine learning (ML) algorithms may be applied to analyze educational data, and researchers aim to determine the most effective ones for specific tasks [3]. This involves comparing algorithms in terms of accuracy, efficiency, and suitability for the educational context. Student retention is a key focus within EDM research, as it directly impacts the overall performance and reputation of educational institutions. By analyzing various factors such as major selection, academic probation, personal reasons, and financial challenges, EDM can help identify patterns that may lead to student dropout. This knowledge enables institutions to implement targeted interventions and support mechanisms to improve retention rates [4].

[1]*Computer Science, The CVM University, palakben.patel@cvmu.edu.in*
[2]*Computer Science, The CVM University, tejas.thakkar@cvmu.edu.in*
[3]*Computer Science, The CVM University, mayur.patel@cvmu.edu.in*
[4]*Computer Science, The CVM University, ami.trivedi@cvmu.edu.in*

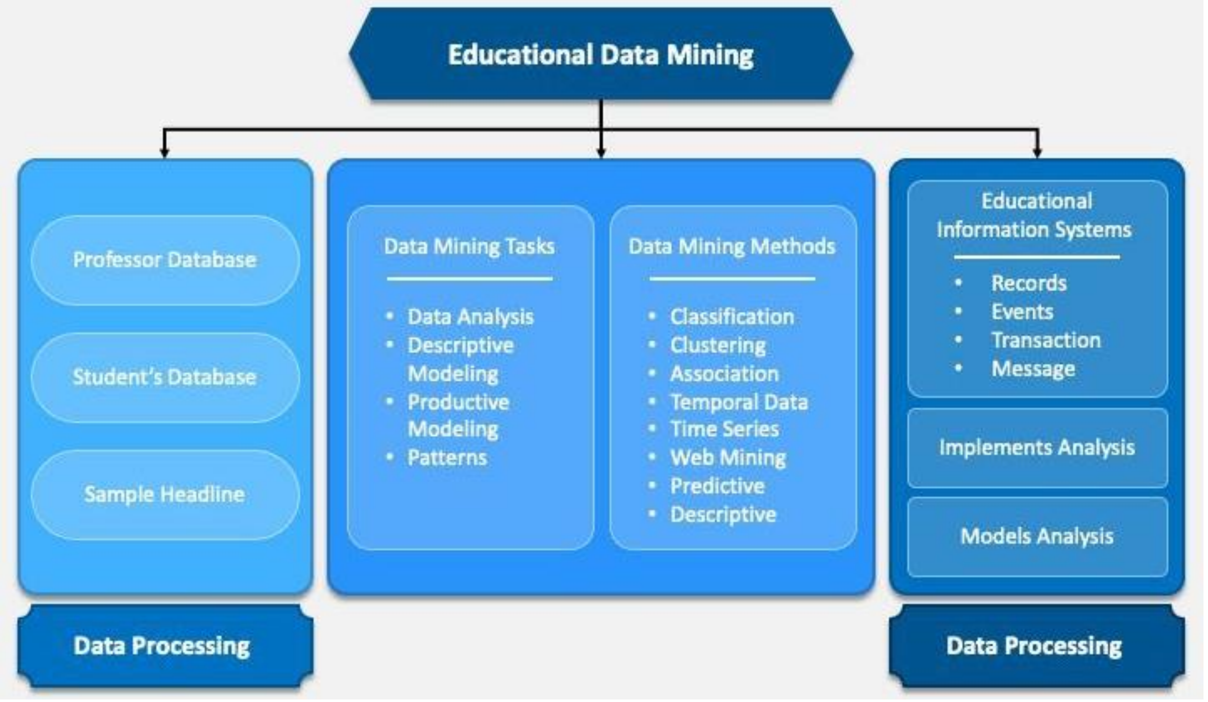

Fig 1 illustrates methods & task of EDM

Educational data mining is defined as a systematic inquiry focused on uncovering potential insights within unique datasets derived from educational contexts. Subsequently, it employs various approaches and methodologies to examine how students engage in diverse learning environments. Figure 1 illustrates specific domains where EDM is extensively applied, including computer-based education, machine learning, computer science, learning analytics, statistics, and pattern recognition.

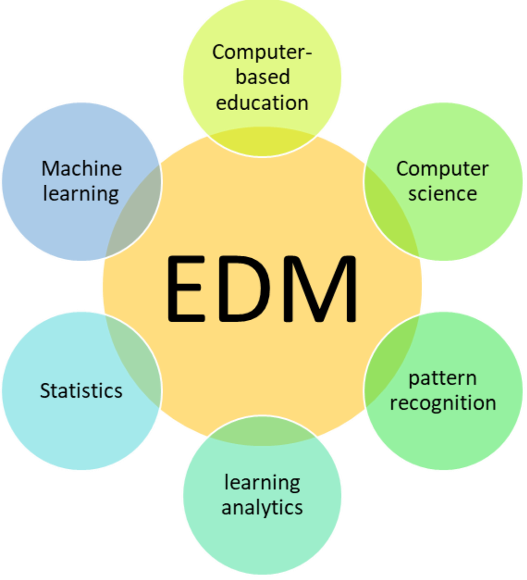

Figure 2 illustrates specific domains of EDM

2. Significance of Guided Pathways:

The potential benefits of guided pathways in improving student success, including clearer academic planning, increased retention rates, and better alignment with career goals.

- **Enhanced Student Success:** Guided pathways provide a structured and clearly defined roadmap for students to follow throughout their academic journey. By offering well-defined program maps and academic plans, students can make informed decisions, reducing the likelihood of taking unnecessary courses and ensuring efficient progress toward degree completion [5].
- **Improved Retention Rates:** The structured nature of guided pathways helps institutions identify potential stumbling blocks early in students' academic journeys. By offering targeted support and interventions, institutions can address challenges that may lead to student attrition, thereby improving overall retention rates [6].
- **Increased Graduation Rates:** With a clear and well-defined path, students are more likely to stay on track and complete their degrees within the expected timeframe. Guided pathways contribute to higher graduation rates by minimizing academic delays, reducing the number of course retakes, and optimizing credit transfer processes [7].
- **Alignment with Career Goals:** Guided pathways often incorporate a focus on career exploration and alignment with students' long-term goals. This helps students connect their academic pursuits with their desired careers, fostering motivation and a sense of purpose throughout their educational journey [8].
- **Efficient Resource Utilization:** Educational institutions can benefit from guided pathways by optimizing resource allocation. With clearer program structures and better-defined course sequences, institutions can streamline course offerings, allocate faculty resources more effectively, and reduce the overall cost of education [9].
- **Facilitation of Transfer Processes:** Guided pathways facilitate smoother transfer processes for students who may transition between institutions or programs. Clear articulation agreements and well-structured pathways enable students to transfer credits seamlessly, minimizing disruptions to their academic progress [10].
- **Enhanced Advising and Support Services:** The implementation of guided pathways often involves strengthened advising and support services. Academic advisors can provide more targeted guidance, helping students navigate their chosen pathways and overcome challenges. This personalized support contributes to students' overall success and satisfaction [11].
- **Alignment with Educational Trends:** Guided pathways align with contemporary educational trends that emphasize student-centered learning and outcomes-based education. This approach acknowledges the diverse needs of students and provides a framework that accommodates various learning styles and paces [12].
- **Data-Driven Decision-Making:** Guided pathways, when coupled with educational data mining and analytics, enable institutions to make data-driven decisions. By analyzing student performance and engagement data within these pathways, institutions can identify areas for improvement, refine academic plans, and enhance overall program effectiveness [13].

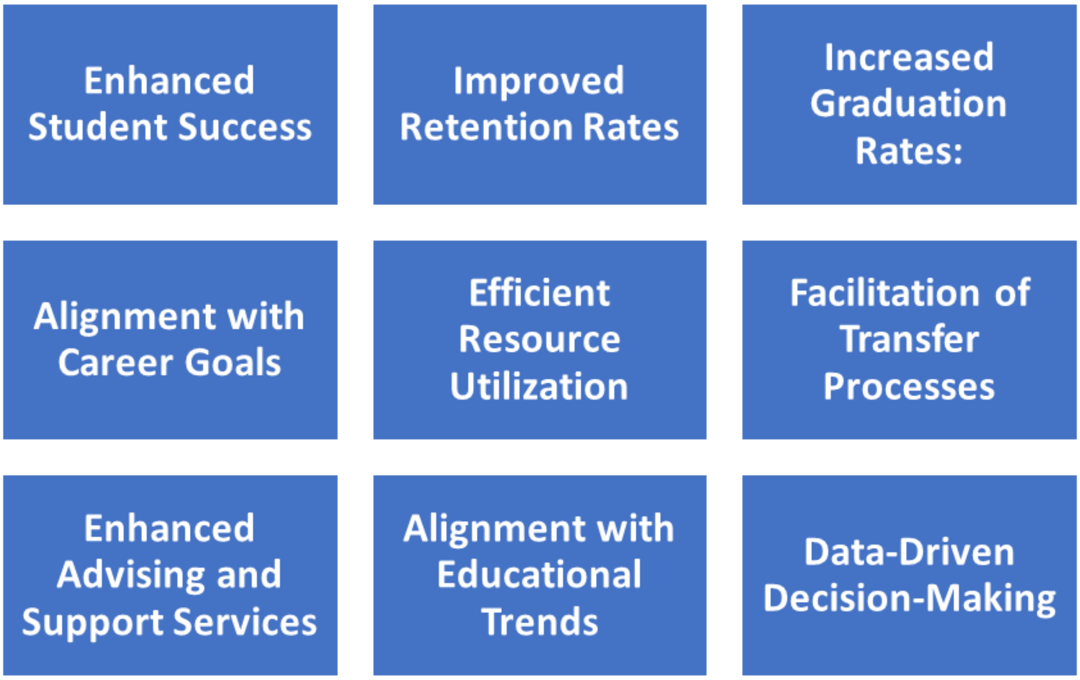

Figure 3 illustrates significance of guided pathways

3. Factors Influencing Student Performance:

Student performance is influenced by a myriad of factors, encompassing a combination of individual, social, economic, and educational elements. Understanding these factors can provide valuable insights into improving educational outcomes. Here are some key factors that influence student performance [33][34].

Individual Characteristics:

- Intellectual Ability: Students' inherent cognitive abilities and aptitude for learning play a crucial role in their academic performance.
- Motivation: The level of interest in the subject can significantly influence student's engagement and effort.
- Learning Style: Individuals have different learning preferences.

Family Background:

- Parental Involvement: Supportive and involved parents often contribute to better student outcomes. This includes providing a conducive home environment and participating in their child's education.
- Socioeconomic Status: Economic factors, such as income and access to resources, can influence the quality of a student's educational experience [35][36].

School Environment:

- Quality of Teaching: The effectiveness of teachers and their ability to communicate, inspire, and adapt teaching methods to diverse student needs is critical.
- Class Size: Smaller class sizes often allow for more individualized attention and better interaction between students and teachers.
- School Facilities: Adequate infrastructure, resources, and a safe and supportive learning environment contribute to student success.

Peer Relationships:

- Social Dynamics: Positive peer relationships and a supportive social environment can foster a sense of belonging and motivation to excel academically.
- Peer Pressure: Negative peer pressure or a disruptive social environment can have adverse effects on academic performance [37].

Health and Well-being:

- Physical Health: A student's overall health, including nutrition, exercise, and sleep, can impact their cognitive abilities and concentration in class.
- Mental Health: Emotional well-being, stress levels, and mental health issues can affect a student's ability to focus and succeed academically [38].

Cultural and Community Factors:

- Cultural Background: Cultural influences can shape a student's attitude toward education and learning.
- Community Support: The level of community support for education, including access to extracurricular activities and community programs, can impact student performance.

Educational Policies and Practices:

- Curriculum Design: The relevance and effectiveness of the curriculum can influence how well students engage with the material.
- Assessment Methods: Fair and effective assessment methods can accurately reflect a student's understanding and promote learning [39][40].

Others:

- **Previous Academic Performance:** The student's academic performance in earlier grades can be indicative of their preparedness for the 12th grade.
- **Study Habits and Time Management:** Effective study habits and time management skills contribute significantly to academic success.
- **Parental Involvement and Support:** Supportive and involved parents can positively impact a student's motivation and success.
- **Teacher Quality and Classroom Environment:** The quality of teaching and the overall classroom environment can influence a student's engagement and understanding of the subject matter.
- **Peer Influence:** The friends a student associates with can impact their study habits, motivation, and overall academic performance.
- **Access to Resources:** Availability of resources like textbooks, online materials, and a conducive study environment can affect performance.
- **Extracurricular Activities:** Balancing extracurricular activities with academics can impact overall well-being and academic performance.
- **Health and Well-being:** Physical and mental health play crucial roles in a student's ability to focus and perform well academically.

Factors influencing stream selection in graduation:

- **Interest and Passion:** Students often choose a stream based on their interests and passion for a particular subject or field.
- **Career Aspirations:** Future career goals and aspirations can guide students toward specific streams that align with their intended professions.
- **Academic Strengths and Weaknesses:** Personal academic strengths and weaknesses may influence the choice of a stream that aligns with a student's abilities.
- **Counseling and Guidance:** Effective guidance from teachers, career counselors, and parents can help students make informed decisions.
- **Market Trends and Job Opportunities:** Awareness of current market trends and job opportunities can impact stream selection, especially for career-oriented fields.

- **Parental Influence:** Parental expectations and influence can sometimes play a role in stream selection, either positively or negatively.
- **Financial Considerations:** The financial implications of pursuing a particular stream or course may also influence decision-making.
- **Peer Influence:** Peer discussions and societal perceptions may play a role in shaping a student's choices.

4. Integration of Machine Learning in Educational Data Mining:

The vast volume of digital content across various domains has spurred research and the advancement of diverse disciplines aimed at simplifying the search, organization, and analysis of this content. Among these disciplines, data mining and machine learning have arisen to automate the analysis of information [14]. These fields involve the identification of patterns and relationships within raw data, playing a crucial role in addressing intricate challenges.

Data Mining Techniques:

Classification: Involves categorizing data into predefined classes or groups based on certain features. It's commonly used for predicting outcomes or assigning items to predefined categories.

Clustering: Focuses on grouping similar data points together based on inherent similarities, without predefined classes. It helps discover natural patterns within the data.

Association Rule Mining: Identifies relationships and associations among variables in large datasets. It's commonly used in market basket analysis to discover patterns of co-occurrence [15].

Classification: Predicting student performance, identifying customer preferences, or diagnosing medical conditions.

- Bayes Classifier: Utilizes Bayes' theorem to predict the probability of a data point belonging to a particular class.
- Function Classifier: Employs mathematical functions to model the relationship between input features and output classes.
- Lazy Classifier: Defers the process of model building until predictions are needed, making it computationally efficient.
- Meta Classifier: Combines the outputs of multiple base classifiers to improve overall performance.
- Rules Classifier: Generates a set of rules based on the data, providing interpretable insights.
- Tree Classifier: Constructs a tree-like structure to represent decisions based on input features.
- PMML (Predictive Model Markup Language) Classifier: Uses a standardized language for representing predictive models, allowing for model interchange between different applications.

Clustering: Grouping similar documents, segmenting customer bases, or identifying anomalies in network traffic.

- Non-hierarchical:
  - **K-Means Clustering:** K-Means is a non-hierarchical clustering algorithm that partitions data points into a predetermined number (K) of clusters. It minimizes the within-cluster variance by iteratively updating cluster centroids and assigning data points to the nearest centroid [16].
  - **Expectation Maximization (EM):** EM is a probabilistic model-based clustering algorithm that iteratively estimates the parameters of a statistical model. It has applications in situations where data points may belong to different clusters with certain probabilities [17].
  - **Farthest First Clustering:** Farthest First is a non-hierarchical clustering algorithm that selects the initial cluster centers by choosing the data points that are farthest apart. It then assigns each remaining point to the nearest cluster center.
- Hierarchical:
  - **Hierarchical Agglomerative Clustering:** Hierarchical Agglomerative Clustering is a hierarchical method that starts with individual data points and gradually merges them into larger clusters.
  - **Partitioning Around Medoids (PAM):** PAM is a non-hierarchical clustering algorithm that identifies representative data points (medoids) for each cluster.

Association Rule Mining: Analyzing shopping basket data, identifying frequent itemsets, and suggesting product recommendations.

- **Apriori Algorithm:** Apriori is a classical algorithm used for discovering frequent itemsets and generating association rules.
- **FP-Growth (Frequent Pattern Growth):** FP-Growth is an alternative algorithm designed to improve efficiency compared to Apriori. It builds a compact data structure called an FP-tree, which allows for faster mining of frequent itemsets without explicitly generating candidate sets.

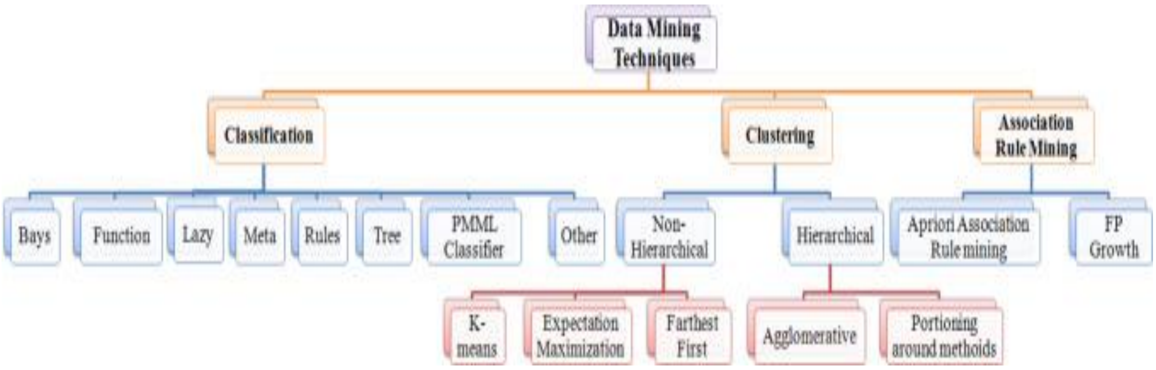

Fig 4 illustrates Data Mining Techniques

Machine Learning

Supervised Learning:

Supervised learning constitutes a ML approach is trained on a labeled training data, signifying that input features are associated with corresponding output labels. The primary objective is for the model to acquire a mapping from inputs to outputs, enabling it to make predictions on unseen data [18].

Semi-Supervised Learning:

Semi-supervised learning represents an approach that amalgamates both labeled and unlabeled data for training purposes. The model encounters a dataset where only a subset of examples are labeled, facilitating learning from both labeled and unlabeled instances [19].

Unsupervised Learning:

Unsupervised learning entails training a model on data devoid of explicit labels. The primary objective is for the algorithm to uncover patterns, relationships, or structures within the data [20].

Reinforcement Learning:

Within machine learning, an agent may learn to make decisions by interacting with its surroundings. This process is known as reinforcement learning. Based on its behaviors, the agent receives feedback in the form of incentives or punishments [21].

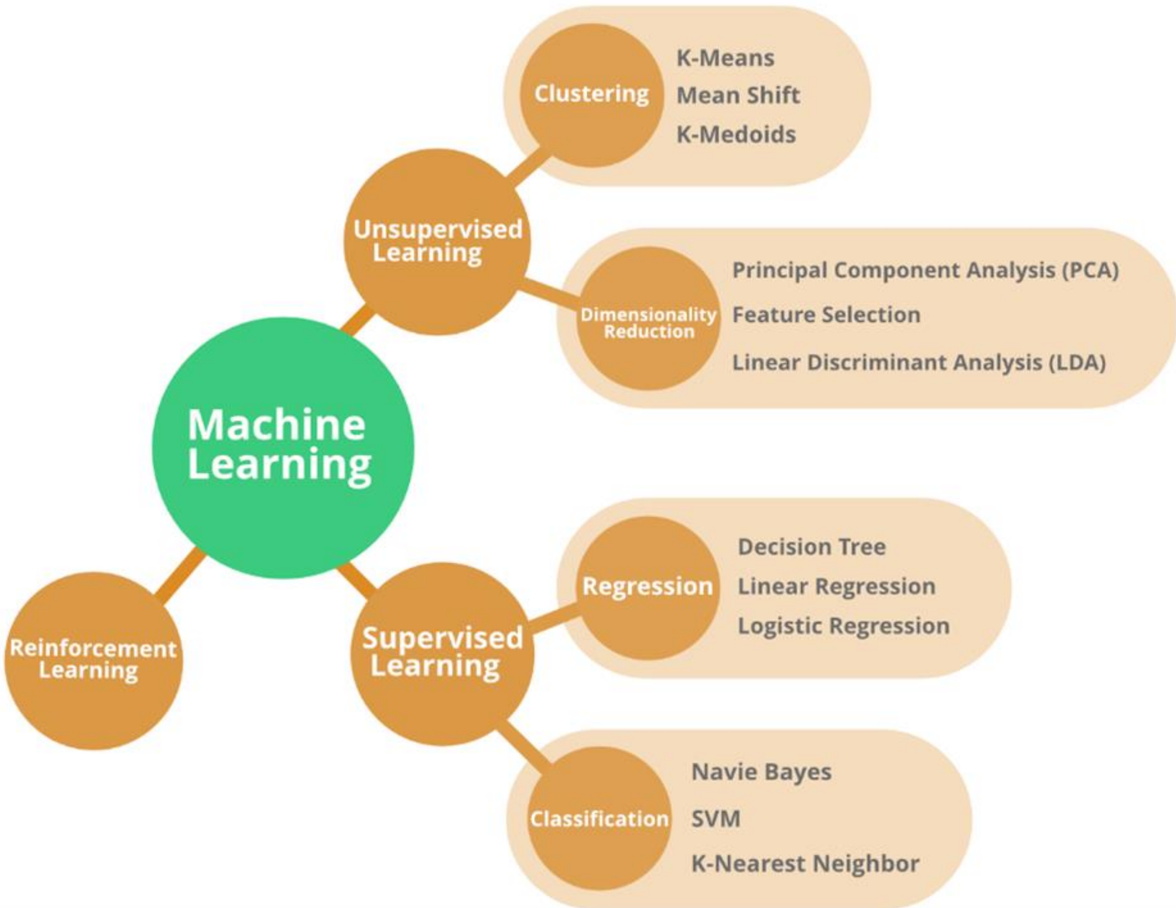

Fig 5 Machine learning techniques

5. Related Work Analysis:

In [22], authors have focuses on forecasting student performance through a singular form of educational data. Random Forest is employed, and the evaluation suggests that Decision Tree and Artificial Neural Network models are promising options for student performance prediction tasks.

In[23], authors have Address the challenge of accurate and efficient academic performance prediction, this work utilizes C4.5 Decision Tree, Multilayer Perceptron, Naïve Bayes, and Random Forest. Naïve Bayes outperforms MLP and CDT across specific ratios.

In [24], authors have Identify factors influencing academic performance, this study employs Synthetic Minority Oversampling Technique, Naïve Bayes, and Random Forest. Naïve Bayes exhibits exceptional performance with a prediction accuracy of 99.34%, closely followed by Random Forest with 98.7%.

In [25], authors have Analyz students' performance across diverse tests, the study uses Long Short Term Memory network, Random Forest, and Gradient Boosting. Despite the superiority of Deep Learning technologies, there's a noteworthy distinction in predictive capabilities.

In [26], authors have Focus on early detection of students at risk, this study utilizes Decision Tree. The findings emphasize the significance of consistently monitoring high-risk students for poor performance.

In[27], authors have Determin key variables impacting academic performance, this work uses ensemble techniques. Logitboost in conjunction with Random Forest demonstrates exceptional performance with an accuracy rate of 99.8%.

In [28], authors have Address the challenge of predicting academic performance, the study employs Synthetic Minority Oversampling Technique. The findings suggest that addressing class imbalance positively impacts predictive capabilities.

In [29], authors have Comprehensive examination of past academic performance, demographics, and psychological attributes is undertaken using Decision Tree and Naïve Bayes. Notably, past academic performance in Mathematics serves as a robust predictor for future performance in MCE English, highlighting the interconnected nature of students' achievements.

In [30], authors have challenge involves integrating entrepreneurial training into education using Random Forest, C5.0, CART, and Artificial Neural Network. The study suggests a modified ensemble machine learning model with significant potential for supporting educators in online entrepreneurship education.

In [31], authors have study focuses on identifying at-risk students and strategically analyzing courses using an adaptive neuro-fuzzy inference system (ANFIS). The ANFIS outperforms commonly used techniques, indicating its effectiveness in capturing complex relationships within the data.

Table I: Comparative analysis previous work done

References	**Problem Statement**	**Technique**	**Findings**
[22]	A singular form of educational data for forecasting student performance binary and multi-classification .	Random Forest (RF)	The evaluation of Decision Tree and Artificial Neural Network (ANN) models on the chosen datasets suggests that they are also promising options for addressing student performance prediction tasks.
[23]	The utilization of classification algorithms in academic performance prediction poses a significant challenge as it involves developing accurate and efficient models to predict student outcomes	C4.5 Decision tree (CDT), Multilayer Perceptron (MLP), Naïve Bayes (NB) and RF	The Naïve Bayes algorithm demonstrated notably superior performance compared to the MLP (Multi-Layer Perceptron) and CDT (Custom Decision Tree) algorithms across specific ratios.
[24]	Identifying the factors influencing the academic performance of students at the onset is crucial for educational institutions to accomplish their educational objectives.	Synthetic Minority Oversampling Technique; NB; RF	The Naïve Bayes (NB) model exhibited exceptional performance, achieving a prediction accuracy of 99.34%. It was closely followed by the Random Forest (RF) model, which attained an accuracy of 98.7%.
[25]	Conducting an analysis of students' performance across diverse	Long Short Term	Despite the evident superiority of existing performance prediction systems utilizing Deep Learning (DL)

	academic tests is imperative for shaping future skill development.	Memory; RF; and Gradient Boosting	technologies, such as ANN and Recurrent Neural Network , over ML -based systems, there remains a noteworthy distinction in their predictive capabilities.
[26]	The inability to detect and provide support to students facing challenges in their academic journey leads to uncertainty regarding their likelihood of success. Hence, there is an imperative need to formulate an early detection framework that can identify students at risk of poor performance.	Decision Tree	The outcomes of this study will serve as valuable insights for educational policymakers and practitioners. The findings highlight the significance of consistently monitoring students identified as being at a high risk of poor performance.
[27]	Determine the key variables that could potentially impact academic performance. These may include socio-economic background, parental involvement, study habits, extracurricular activities, teacher quality, and student motivation.	Ensemble techniques	The evaluation results indicate that the Logitboost model, in conjunction with Random Forest, has demonstrated exceptional performance, achieving an impressive accuracy rate of 99.8%.
[28]	The challenge at hand revolves around predicting the academic performance of students, a topic of paramount importance due to its pivotal role in shaping educational outcomes. Accurate predictions in this domain are instrumental in the development of practical systems that serve multifaceted purposes, including the advancement of academic performance and the prevention of student dropouts.	Synthetic Minority Oversampling TEchinque (SMOTE)	The findings of the study reveal that the SMOTE surpasses the performance of the Info-Gain filter when combined with SMOTE, leading to notable enhancements in the overall performance of the algorithms under consideration. This suggests that addressing class imbalance through the strategic introduction of synthetic samples positively impacts the predictive capabilities of the models.
[29]	The problem at hand involves a comprehensive examination of three key factors—students' past academic performance, demographics, and psychological attributes—with the objective of discerning their individual and collective influence on prediction outcomes. This investigation is critical for understanding the intricate interplay of these factors in shaping predictive models. By scrutinizing the impact of past academic performance, demographics, and psychological attributes	Decision Tree (DT); Naïve Bayes (NB)	In analyzing top predictors identified by four distinct classifier types, this study has uncovered a significant finding: the past academic performance of students in Mathematics serves as a robust predictor for their subsequent performance in MCE (Malaysian Certificate of Education) English. Conversely, the study also reveals that students' past performance in English is a strong predictor for their future performance in MCE Mathematics. This interrelation between the academic performances in both subjects underscores the interconnected nature of students' achievements.
[30]	The challenge at hand pertains to the effective integration of entrepreneurial training into the educational landscape, with a specific focus on leveraging technology as a means to enhance	RF, C5.0, CART and ANN	this study indicate that the suggested modified ensemble machine learning model holds significant potential in supporting educators and administrators in the realm of online entrepreneurship education. Specifically, the model demonstrates efficacy in identifying students who may benefit from additional support. This capability

	students' capacity to develop viable and lucrative solutions for emerging challenges. While the importance of entrepreneurial skills is acknowledged, there is a need to explore technology in delivering education.		allows for the customization of instructional strategies and the development of targeted interventions, ultimately contributing to the enhancement of students' overall learning experience.
[31]	The challenge at hand involves the identification of students who are at risk within an educational context, coupled with a strategic analysis of courses to pinpoint areas requiring enhancements in content, delivery mode.	ANFIS	The study reveal that the ANFIS approach outperforms commonly used techniques, specifically demonstrating superior predictive accuracy compared to multilinear regression. This suggests that ANFIS, with its adaptive and fuzzy logic-based modeling, provides a more effective framework for capturing and representing complex relationships within the data.

6. Conclusion:

In conclusion, the assessment of Students' Academic Performance (SAP) serves as a pivotal tool in understanding the academic standing of students within an educational institution. This metric not only provides instructors and administrators with a precise evaluation of students across diverse courses but also empowers students to reflect on their strategies, fostering opportunities for improvement in subsequent semesters. The variations in criteria used by different institutions to evaluate student performance highlight the need for a more comprehensive exploration of existing prediction techniques. The quest for the most effective methodology in predicting students' academic progress underscores the importance of research in this domain. Additionally, a significant gap exists in the exploration of factors influencing student achievement in specific courses, further emphasizing the need for a thorough investigation. This paper addresses these gaps by proposing the application of advanced machine learning techniques to enhance student academic performance. By leveraging data mining methods, the objective is not only to elevate student outcomes but also to provide tangible benefits for faculty members, students, educators, and overall institutional management. The integration of these advanced techniques holds the potential to contribute to a more informed and proactive approach to academic support, creating a positive impact on the entire educational ecosystem. Through this endeavor, the aim is to pave the way for improved educational practices, ultimately benefiting both students and institutions alike.

References:

[1] Sciarrone, F. (2018, April). Machine learning and learning analytics: Integrating data with learning. In 2018 17th International Conference on Information Technology Based Higher Education and Training (ITHET) (pp. 1-5). IEEE.

[2] Urbina Nájera, A. B., & Calleja Mora, J. D. L. (2017). Brief review of educational applications using data mining and machine learning. Revista electrónica de investigación educativa, 19(4), 84-96.

[3] Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wiley interdisciplinary reviews: Data mining and knowledge discovery, 10(3), e1355.

[4] Dol, S. M., & Jawandhiya, P. M. (2023). Classification Technique and its Combination with Clustering and Association Rule Mining in Educational Data Mining—A survey. Engineering Applications of Artificial Intelligence, 122, 106071.

[5] Kausar, S., Huahu, X., Hussain, I., Wenhao, Z., & Zahid, M. (2018). Integration of data mining clustering approach in the personalized E-learning system. IEEE access, 6, 72724-72734.

[6] Du, X., Yang, J., Hung, J. L., & Shelton, B. (2020). Educational data mining: a systematic review of research and emerging trends. Information Discovery and Delivery, 48(4), 225-236.

[7] Dabhade, P., Agarwal, R., Alameen, K. P., Fathima, A. T., Sridharan, R., & Gopakumar, G. (2021). Educational data mining for predicting students' academic performance using machine learning algorithms. Materials Today: Proceedings, 47, 5260-5267.

[8] Wibawa, B., Siregar, J. S., Asrorie, D. A., & Syakdiyah, H. (2021, April). Learning analytic and educational data mining for learning science and technology. In AIP conference proceedings (Vol. 2331, No. 1). AIP Publishing.

[9] Okewu, E., Adewole, P., Misra, S., Maskeliunas, R., & Damasevicius, R. (2021). Artificial neural networks for educational data mining in higher education: A systematic literature review. Applied Artificial Intelligence, 35(13), 983-1021.

[10] Nanavaty, S., & Khuteta, A. (2024). A Deep Learning Dive into Online Learning: Predicting Student Success with Interaction-Based Neural Networks. International Journal of Intelligent Systems and Applications in Engineering, 12(1), 102-107.

[11] Dayal, G., Verma, P., & Sehgal, S. (2024). A Comprehensive Review on the Integration of Artificial Intelligence in the Field of Education. Leveraging AI and Emotional Intelligence in Contemporary Business Organizations, 331-349.

[12] Li, L., Yu, H., & Kunc, M. (2024). The impact of forum content on data science open innovation performance: A system dynamics-based causal machine learning approach. Technological Forecasting and Social Change, 198, 122936.

[13] Bendaouia, A., Qassimi, S., Boussetta, A., Benzakour, I., Benhayoun, A., Amar, O., ... & Hasidi, O. (2024). Hybrid features extraction for the online mineral grades determination in the flotation froth using Deep Learning. Engineering Applications of Artificial Intelligence, 129, 107680.

[14] Li, D., Nie, J. H., Wang, H., & Ren, W. X. (2024). Loading condition monitoring of high-strength bolt connections based on physics-guided deep learning of acoustic emission data. Mechanical Systems and Signal Processing, 206, 110908.

[15] Qiu, Q., Tian, M., Tao, L., Xie, Z., & Ma, K. (2024). Semantic information extraction and search of mineral exploration data using text mining and deep learning methods. Ore Geology Reviews, 105863.

[16] Batool, S., Rashid, J., Nisar, M. W., Kim, J., Kwon, H. Y., & Hussain, A. (2023). Educational data mining to predict students' academic performance: A survey study. Education and Information Technologies, 28(1), 905-971.

[17] Batool, S., Rashid, J., Nisar, M. W., Kim, J., Kwon, H. Y., & Hussain, A. (2023). Educational data mining to predict students' academic performance: A survey study. Education and Information Technologies, 28(1), 905-971.

[18] Abideen, Z. U., Mazhar, T., Razzaq, A., Haq, I., Ullah, I., Alasmary, H., & Mohamed, H. G. (2023). Analysis of Enrollment Criteria in Secondary Schools Using Machine Learning and Data Mining Approach. Electronics, 12(3), 694.

[19] Hussain, S., & Khan, M. Q. (2023). Student-performulator: Predicting students' academic performance at secondary and intermediate level using machine learning. Annals of data science, 10(3), 637-655.

[20] Aulakh, K., Roul, R. K., & Kaushal, M. (2023). E-learning enhancement through Educational Data Mining with Covid-19 outbreak period in backdrop: A review. International Journal of Educational Development, 102814.

[21] Baig, M. A., Shaikh, S. A., Khatri, K. K., Shaikh, M. A., Khan, M. Z., & Rauf, M. A. (2023). Prediction of Students Performance Level Using Integrated Approach of ML Algorithms. International Journal of Emerging Technologies in Learning, 18(1).

[22] Chen, Y., & Zhai, L. (2023). A comparative study on student performance prediction using machine learning. Education and Information Technologies, 1-19.

[23] Adane, M. D., Deku, J. K., & Asare, E. K. (2023). Performance Analysis of Machine Learning Algorithms in Prediction of Student Academic Performance. Journal of Advances in Mathematics and Computer Science, 38(5), 74-86.

[24] Alghamdi, A. S., & Rahman, A. (2023). Data mining approach to predict success of secondary school students: A Saudi Arabian case study. Education Sciences, 13(3), 293.

[25] Kukkar, A., Mohana, R., Sharma, A., & Nayyar, A. (2023). Prediction of student academic performance based on their emotional wellbeing and interaction on various e-learning platforms. Education and Information Technologies, 1-30.

[26] Masabo, E., Nzabanita, J., Ngaruye, I., Ruranga, C., Nizeyimana, J. P., Uwonkunda, J., & Ndanguza, D. (2023). Early detection of students at risk of poor performance in Rwanda higher education using machine learning techniques. International Journal of Information Technology, 1-10.

[27] Jalota, C. (2023). An effectual model for early prediction of academic performance using ensemble classification. Journal of Language and Linguistics in Society (JLLS) ISSN 2815-0961, 3(02), 19-33.

[28] Alija, S., Beqiri, E., Gaafar, A. S., & Hamoud, A. K. (2023). Predicting Students Performance Using Supervised Machine Learning Based on Imbalanced Dataset and Wrapper Feature Selection. Informatica, 47(1).

[29] Roslan, M. H. B., & Chen, C. J. (2023). Predicting students' performance in English and Mathematics using data mining techniques. Education and Information Technologies, 28(2), 1427-1453.

[30] Malik, A., Onyema, E. M., Dalal, S., Lilhore, U. K., Anand, D., Sharma, A., & Simaiya, S. (2023). Forecasting students' adaptability in online entrepreneurship education using modified ensemble machine learning model. Array, 19, 100303.

[31] Mehdi, R., & Nachouki, M. (2023). A neuro-fuzzy model for predicting and analyzing student graduation performance in computing programs. Education and Information Technologies, 28(3), 2455-2484.

[32] Gilholm, P., Gibbons, K., Brüningk, S., Klatt, J., Vaithianathan, R., Long, D., ... & Schlapbach, L. J. (2023). Machine learning to predict poor school performance in paediatric survivors of intensive care: a population-based cohort study. Intensive Care Medicine, 1-11.

[33] Narayan, Vipul, et al. "7 Extracting business methodology: using artificial intelligence-based method." Semantic Intelligent Computing and Applications 16 (2023): 123

[34] Narayan, Vipul, et al. "A Comprehensive Review of Various Approach for Medical Image Segmentation and Disease Prediction." Wireless Personal Communications 132.3 (2023): 1819-1848.

[35] Mall, Pawan Kumar, et al. "Rank Based Two Stage Semi-Supervised Deep Learning Model for X-Ray Images Classification: AN APPROACH TOWARD TAGGING UNLABELED MEDICAL DATASET." Journal of Scientific & Industrial Research (JSIR) 82.08 (2023): 818-830.

[36] Narayan, Vipul, et al. "Severity of Lumpy Disease detection based on Deep Learning Technique." 2023 International Conference on Disruptive Technologies (ICDT). IEEE, 2023.

[37] Saxena, Aditya, et al. "Comparative Analysis Of AI Regression And Classification Models For Predicting House Damages İn Nepal: Proposed Architectures And Techniques." Journal of Pharmaceutical Negative Results (2022): 6203-6215.

[38] Kumar, Vaibhav, et al. "A Machine Learning Approach For Predicting Onset And Progression"“Towards Early Detection Of Chronic Diseases “." Journal of Pharmaceutical Negative Results (2022): 6195-6202.

[39] Chaturvedi, Pooja, Ajai Kumar Daniel, and Vipul Narayan. "Coverage Prediction for Target Coverage in WSN Using Machine Learning Approaches." (2021).

[40] Chaturvedi, Pooja, A. K. Daniel, and Vipul Narayan. "A Novel Heuristic for Maximizing Lifetime of Target Coverage in Wireless Sensor Networks." Advanced Wireless Communication and Sensor Networks. Chapman and Hall/CRC 227-242.