iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78840

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

A Comprehensive Review of Machine Learning Techniques for Student Academic Performance Prediction

Dr. P. C. Khanzode[1], Shubham Dande[2], Nishant Sonone[3], Abhinav Surjuse[4], Amit Khadse[5], Shreyash Kuyate[6]
Department of Computer Engineering Sipna College of Engineering and Technology Amravati, India

Abstract: Student academic performance prediction is a crucial topic in Educational Data Mining (EDM) and Learning Analytics, which can be used to help at-risk students by undertaking timely actions. The given paper is a systematic review of machine learning methods used in this field. It analyzes a range of approaches, starting with interpretable models such as Multiple Linear Regression and Decision Trees to ensemble and deep learning high-performance models such as Random Forest and Neural Networks. The review highlights the central role of feature engineering and is discussing predictors of academic and behavioral data, social-economic and psychological conditions. One of the broad implications of this paper is providing a comparative analysis of these methods with an emphasis on the continuing trade-off between predictive accuracy and model inter-pretability. Moreover, the disconnect between theory and real-world, full-stack deployment systems, which are more and more critical when it comes to actual usability, is also critically discussed in this review. Major gaps in the research, such as excessive use of synthetic data, lack of practical testing, and ethics, are determined. Lastly, the paper presents future directions which include the use of Explainable AI (XAI), federated learning in privacy and creation of real-time adaptive feedback systems.

Index Terms: Educational Data Mining, Learning Analytics, Student Performance Prediction, Machine Learning Review, Explainable AI, Full-Stack Deployment.

I. INTRODUCTION

Contemporary educational organizations produce enormous volumes of group information about students. Nonetheless, others still use the approach of reactive assessment, meaning that they do not diagnose struggling children until it is too late to do anything about them [10]. This paradigm constitutes a great inefficiency systemically. Educational Data Mining (EDM) and Learning Analytics (LA) provide a more active way to do so by using computational methods to find useful patterns in educational data that can identify at-risk students early on [1], [5].

The value of such early prediction systems is quite evident in the literature. Research has repeatedly shown that a timely intervention, based on predictive knowledge can lead to a significant increase in student retention rates and academic performance [12]. These systems are based on machine learning (ML) models, which can recognize complicated,

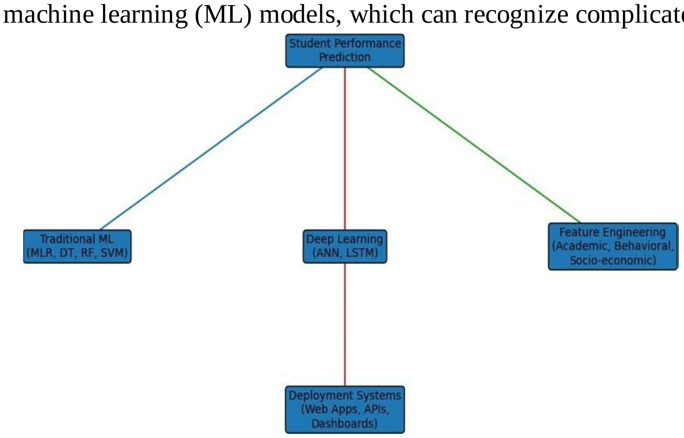


Fig. 1. Taxonomy of Student Academic Performance Prediction Methods Reviewed in This Paper.

non-obvious relationships within multifaceted student data [14].

However, the academic community faces several persistent challenges. First, a wide variety of ML algorithms exist, each presenting a distinct trade-off between predictive accuracy and model interpretability—a critical factor for stakeholder trust in educational settings [16]. Second, student performance is influenced by a complex interplay of academic, behavioral, psychological, and socio-economic factors, making feature selection a non-trivial task [17]. Third, a significant gap per- sists between theoretical ML models and practically deployed systems that educators can use on a daily basis [9].

This paper provides a systematic review of the literature on ML-based student performance prediction. The primary objectives are:

1) To catalog and compare the ML techniques commonly applied in student performance prediction.
2) To examine the role and importance of feature engineer- ing.
3) To analyze the growing body of work on full-stack, deployable prediction systems.
4) To identify persistent research gaps and propose produc- tive future research directions.

II. LITERATURE REVIEW

The literature on student performance prediction can be divided into the major ML methods applied, the features obtained, and the systems provided. Fig. 1 displays a taxonomy of these areas.

A. Traditional Machine Learning Approaches

Classical ML techniques have been central to EDM studies, typically due to their performance-inter-pretability trade-off [13].

1) Linear Regression: Multiple Linear Regression (MLR) is among the oldest methods because it is the most interpretable. Direct examination can be performed by teachers on model coef- ficients to interpret the weighted contribution of each predic- tor [11]. Experiments by Kaur and Sharma [3] suggest that when the relationship between the variables is somewhat linear, then MLR can be competitive with more intricate mod- els. To its detriment, however, is its basic limitation in the underlying linearity assumption, which might not be able to represent the non-linear influence of variables such as severe sleep deprivation or high-stress levels [17].
2) Decision Trees: Decision Tree algorithms are inter-pretable (hierarchical) sets of if-then rules. According to early EDM surveys by Romero and Ventura [1] and Kotsiantis [2], many other applications of such algorithms as C4.5 and CART to predict course grades and risk of dropping out were reported. Although good at mixed data types and exposing non- linearities, single Decision Trees are infamously overfitting and unstable, with minor differences in training data causing radically different models [13].
3) Random Forest: A random forest is an ensemble algorithm, which overcomes the flaws of a single Decision Tree by combining the results of a set of trees. It is a common fact in the literature that Random Forest is more predictive accurate than linear models and single trees [12], [20]. It is characterized by its ability to model complex interactions without a lot of feature engineering. The cost of this performance, however, is interpretability which produces a black box that can hardly be trusted or examined by educators [6].
4) Support Vector Machines: Support Vector Machines (SVMs) have been used on classification and regression problems in EDM [2]. Maximum-margin provides good theoretical guarantees against overfitting. Researchers, however, say that SVMs do not scale to large data sets and do not directly produce probabilistic results, making their use in larger applications in practical educational systems a challenge [10].

B. Deep Learning Approaches

As large learning datasets have increased, researchers have been looking at deep learning, where highly complex pattern models are possible.

1) Artificial Neural Networks: Artificial Neural Networks (ANNs) and Multi-layer Perceptrons (MLPs) are capable of modeling complex, non-linear relations that simple mod-els do not represent [18]. It has been claimed to improve accuracy over conventional models, particularly on large and rich datasets [1], [19]. But the main disadvantage is that they have. Absoluteness of lack of interpretability. This is a major obstacle in the context of education where transparency and fairness are paramount [6].
2) Recurrent and Time-Series Approaches: In situations that have longitu- dinal data (e.g. weekly quiz scores, LMS activity logs), Recurrent Neural Networks (RNNs) and Long Short- Term memory (LSTM) networks have been considered. Such models are able to learn time-based dependencies, e.g. detecting negative patterns of engagement, which then result in bad behavior [9], [14]. But they cannot be widely used because they require a strong time-series data collection machinery, which is not present in most institutions.

C. Predictor Selection and Feature Engineering.

The selection of the predictors is no less important than the selection of the algorithm.

1) Academic and Behavioral Characteristics: The body of literature excessively depends on academic and behavioral information. Past scores, hours of study, attendance, and practice engagement are characteristics that are repeatedly rated as strong predictors [3], [11], [12]. These characteristics are not particularly difficult to gather and are intuitive to teachers.
2) Socio-Economic and Psychological Features: One of the commonly accepted disadvantages of the literature is the tendency to exclude socio-economic and psychological aspects [1], [17]. Other factors such as family income, parental education, student motivation, and mental health are strong determinants of academic success. The fact that they are not provided in models means that the risks of attributing systems-level disadvantages to personal failure are high and also restricts the quality of predictive understandings.

D. System Implementation and Practical.

An increasing amount of evidence acknowledges that the value of a model can only be achieved through successful implementation. A number of studies have been done to present the integration of machine learning models with web based systems, with the emphasis on usability and real time prediction capabilities [7], [9].

Lightweight Python frameworks, such as FastAPI and Flask, are often mentioned as having been used to create API backends, with user-facing dashboards implemented in JavaScript libraries, like React [4]. This full stack solution seeks to close the disconnect between raw ML research and real-world classroom use, allowing capabilities such as batch processing of whole cohorts, and transparent analytics to stakeholders [5].

III. DISCUSSION

A critical review of the literature shows that there are some overarch- ing themes and tensions that characterize the present-day student performance prediction research.

A. The Accuracy Interpretability Trade-Off.

The overwhelming theme is the inevitable trade-off between predictive quality and model explanability. On the one hand, ensemble and deep learning models such as Random Forest and ANNs are always the most accurate models in a comparative analysis [12], [18]. Their modeling complex capabilities. Their predictive power arises because of non-linear interactions. Conversely, they have a black box character, which is a major obstacle to implementation in an educational setting. Explanations are needed by stakehold- ers such as teachers, administrators, and students to trust and act on predictions, which Linear Regression and Deci- sion Trees directly provide in highly interpretable forms, despite typically being less accurate than more advanced methods [3], [16]. This trade-off requires institutions to decide which prediction is most accurate

B. The Persistence of Linear Models.

Nevertheless, in spite of the emergence of more complicated algorithms, MLR is still very topical and is actively applied in real life [3], [11]. Three major strengths make it appealing to this day. To start with, its unrecognized interpretability gives teachers an opportunity to use model coefficients to generate actionable information (e.g., an additional hour of learning per day is related to a 5- point score increase). Second, it is very fast and can be deployed in real time, at very low latency in web systems without special hardware. Third, it is simple and thus provides a strong and effective foundation upon which the more complicated models may be modeled.

C. Deep Learning Learning Obstacles.

Deep learning has transformed other disciplines, but its use as a predictive tool in education has been slow and limited. The literature attributes a number of reasons to this. The first challenge is the overwhelming degree of interpretability, which is most of the time intolerable in stakes of learning decision making [6]. Secondly, deep learning algorithms have high data requirements and need large-scale data to train successfully, which cannot be achieved by various institutions [19]. Lastly, the sheer difficulty of training and implementing these models becomes a major technical obstacle to many learning institutions, who might not have dedicated teams of data scientists.

D. Practical and Ethical Issues.

Moving out of a theoretical model and into the practical tool is not an easy task. Most of the high-performing models described in the academic literature are actually trained on clean and synthetic datasets [8], not on the noisy and incomplete institutional data of the real world. This causes a gap in real-world validation. Moreover, the introduction of such systems also poses great ethical issues. Historical information can be learnt and reproduced through models that are trained with historical data [21]. Problems with data privacy, security, and student permission within laws such as GDPR and FERPA also need to be considered, which much of the technical literature generally ignores.

IV. COMPARATIVE ANALYSIS

Through literature review, Table I provides a comparative overview of popular machine learning methods by performance, interpretability, and applicability, grounded on their deployment and performance capabilities. Based on Table I, it is obvious that there is a trade-off. Ensemble and deep learning methods always perform better than their counterparts. Greater precision, the high interpretability and deployability of linear models guarantee their further use in most real-world education.

V. RESEARCH GAPS

We have identified the following gaps in the literature that any future research should fill:

1) Overfitting to Synthetic Datasets: many studies test their models on clean, synthetic datasets [8]. It results in artificially boosted performance measurements, and obscures the difficulty of operating in the real world and making noisy data work.
2) Absence of Longitudinal and Prospective validation: Majority of the studies conduct retrospective validation on a fixed data set. Longitudinal studies that de- ploy a model and assess its prospective performance over time in a reality institutional setting are not available.
3) The Interpretability-Actionability Gap: Sometimes interpretable models, although interpretable, do not provide a clear picture of how educa- tors can put these findings into practice in a form that would be effective. A gap exists between an explanation of a model and an action of a teacher that requires research.
4) Missing Feature Sets: As mentioned, important fac- tors such as student mental health, motivation as well as socio- economic background are not consistently included, constraining predictive ability and completeness of the models [17].
5) Ethical and Fairness Audits: There is no standardized model of how prediction models are audited to predict fairness and bias. The means of identifying and addressing racial, gender and socio-economic biases are urgently required [21].

VI. FUTURE DIRECTIONS

Using the identified gaps we suggest the following directions on how to conduct future research:

1) Explainable AI (XAI) Introspection: The most likely direction is the systematized version of XAI frameworks such as SHAP [6] and LIME. Such tools may be used to give local, per-student descriptions of the black box models, which may address the accuracy-interpretability trade- off.
2) Privacy-Preserving Predic- tion Federated Learning): Federated learning is a direction to take in the future to ensure privacy and allow collaboration across institutions without sharing raw student data. This method trains a world model using decentralized data and maintains student privacy [22].
3) Real-Time Adaptive Feedback Systems: Future systems: Future systems no longer need to predict, but intervene. ML mod- els have the power to drive real time adaptive learning systems that deliver students a learning path, customized feedback and resources in response to their projected learning and performance.
4) Hybrid Modeling: Scientists ought to study hybrid models that integrate strengths of various methods, including a deep learning model to engineer.

TABLE I
COMPARATIVE ANALYSIS OF ML APPROACHES FOR STUDENT PERFORMANCE PREDICTION

Model Family	Example Algo- rithm	Typical Accuracy	Interpretability	Pros	Cons
Linear Models	Multiple Linear	Moderate to High	High	Highly interpretabl	Assumes linear relationships;

	Regression			e coefficients; computationally efficient; easy to deploy [3], [11].	may underfit complex data.
Tree-Based Models	Decision Tree (C4.5)	Moderate	High	Intuitive rule-based structure; handles non-linear and cate- gorical data well [1], [13].	Prone to overfitting; can be un- stable.
Ensemble Models	Random Forest, XGBoost	High to Very High	Low	High accuracy; robust to over- fitting; handles complex inter- actions natively [12], [20].	”Black box” nature; computa- tionally more intensive.
Deep Learning	ANN, LSTM	Very High	Very Low	Can model extremely com- plex patterns; state-of-the-art for large datasets [18], [19].	Requires huge datasets; lowest interpretability; high computa- tional cost for training.
Kernel-Based	Support Vecto r Machine	Good	Low	Strong theoretical foundation; effective in high-dimensional spaces [2].	Poor scalability with dataset size; less intuitive than other models.

features that are then fed into an interpretable linear model.

5) Causal Inference: The field should move from cor- relation to causation. Causal inference methods could help determine which interventions are most effective for specific student subgroups, providing more actionable insights than purely predictive models.

VII. CONCLUSION

This paper has provided a comprehensive review of the machine learning techniques used for student academic perfor- mance prediction. The analysis reveals a dynamic field charac- terized by a fundamental tension between model accuracy and interpretability. While complex ensemble and deep learning models offer superior predictive power, the transparency and practicality of traditional methods like Multiple Linear Regres- sion ensure their continued relevance, especially in deployed systems. A significant gap persists between models developed in a research context and those validated in real-world educational settings. Future work must prioritize real-world validation, address the ethical dimensions of algorithmic prediction, and integrate Explainable AI to build systems that are not only accurate but also trustworthy and actionable for educators. The ultimate goal is to augment human judgment, providing data- driven insights that enable more effective and equitable support for all students.

REFERENCES

[1] C. Romero and S. Ventura, “Educational data mining: A survey from 1995 to 2005,” Expert Systems with Applications, vol. 33, no. 1, pp. 135–146, 2007.

[2] Kotsiantis, “Educational data mining: A review of the state of the art,” University of Peloponnese, 2007.

[3] T. M. Kaur and S. Sharma, “Student performance prediction using ma- chine learning techniques,” Int. J. of Recent Technology and Engineering, vol. 8, no. 4, pp. 2277–2281, 2019.

[4] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[5] G. Siemens and R. S. Baker, “Learning analytics and educational data mining: Towards communication and collaboration,” in Proc. 2nd Int. Conf. on Learning Analytics and Knowledge, 2012, pp. 252–254.

[6] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 4765–4774.

[7] FastAPI Framework Documentation, 2024. [Online]. Available: https://fastapi.tiangolo.com/

[8] N. (nikhil7280), “Student Performance (Multiple Linear Regression),” Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/ nikhil7280/student-performance-multiple-linear-regression

[9] M. Hussain, W. Zhu, W. Zhang, and S. M. R. Abidi, "Student engage- ment predictions in an e-learning system and their impact on student performance," Computational Intelligence and Neuroscience, vol. 2018, 2018.

[10] A. Namoun and A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," Applied Sciences, vol. 11, no. 1, p. 237, 2021.

[11] O. Oyerinde and C. Chia, "Predicting students' academic performances— A learning analytics approach using multiple linear regression," Int.J. of Computer Applications, vol. 157, no. 4, pp. 37–44, 2017.

[12] A. Rastrollo-Guerrero, J. A. Go´mez-Pulido, and A. Dura´n-Dom´ınguez, "Analyzing and predicting students' performance by means of machine learning: A review," Applied Sciences, vol. 10, no. 3, p. 1042, 2020.

[13] D. Kabakchieva, "Predicting student performance by using data mining methods for classification," Cybernetics and Information Technologies, vol. 13, no. 1, pp. 61–72, 2013.

[14] S. Aulck, N. Velagapudi, J. Blumenstock, and J. West, "Predicting student dropout in higher education," arXiv preprint arXiv:1606.06364, 2016

[15] C. Romero, S. Ventura, and E. Garc´ıa, "Data mining in course manage- ment systems: Moodle case study and tutorial," Computers & Education, vol. 51, no. 1, pp. 368–384, 2008.

[16] V. Belle and I. Papantonis, "Principles and practice of explainable artificial intelligence," Frontiers in Big Data, vol. 4, p. 688969, 2021

[17] M. T. H. Alyahyan and D. Du¨s¸tego¨r, "Predicting academic success in higher education: literature review and best practices," International Journal of Educational Technology in Higher Education, vol. 17, no. 1, p. 3, 2020.

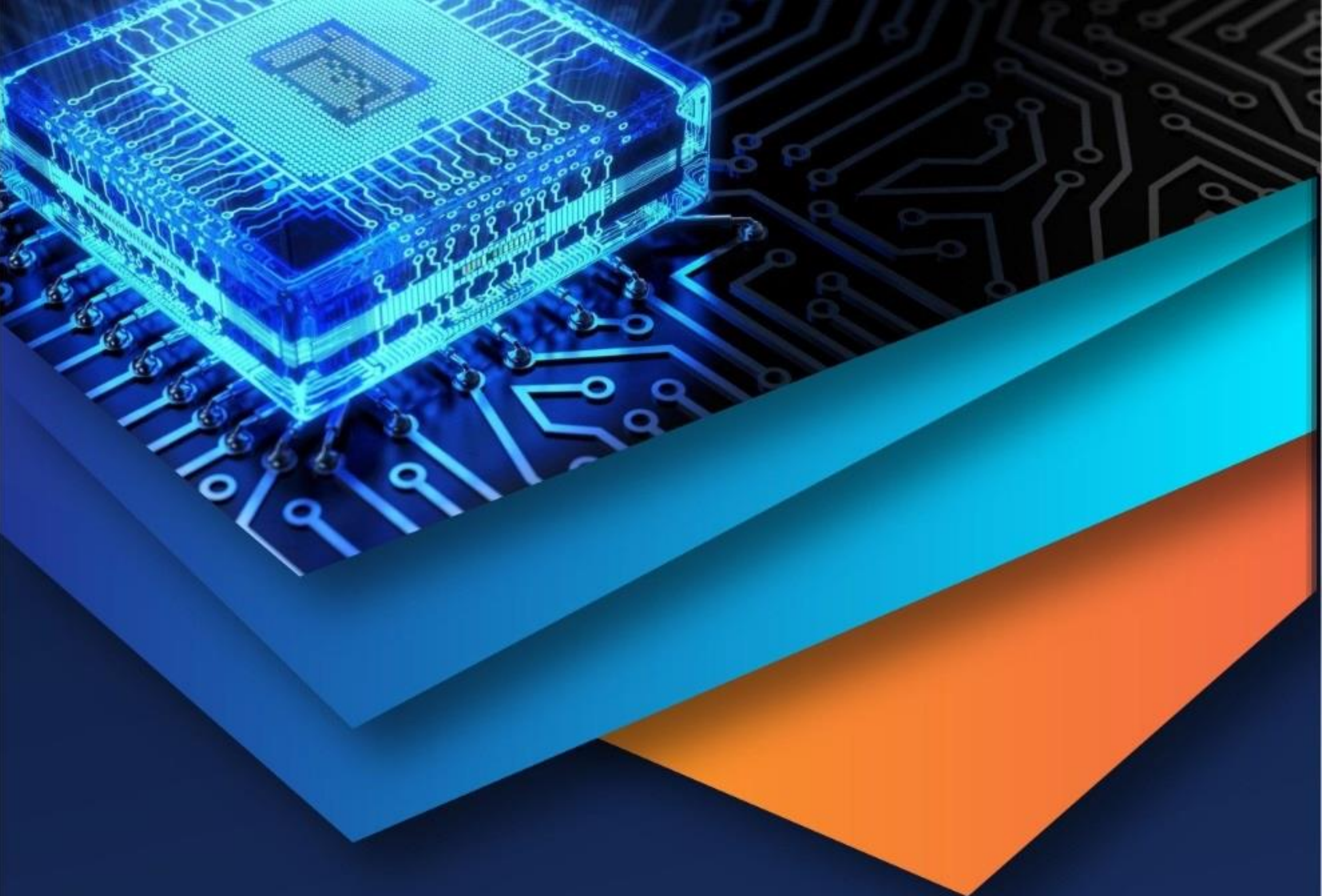
[18] I. E. Livieris, K. Drakopoulou, and P. Pintelas, "A CNN-LSTM model for dropout prediction in e-learning," in 2020 11th International Con- ference on Information, Intelligence, Systems and Applications (IISA), 2020, pp. 1-6.

[19] A. J. Al-Radaideh, E. M. Al-Shawakfa, and M. I. Al-Najjar, "Mining student data using decision trees," in 2006 International Arab Confer- ence on Information Technology, 2006.

[20] B. M. F. Al-Shargabi, F. H. Al-Hadhrami, and A. A. Al-Dhaqm, "A systematic review of machine learning techniques for students' academic performance prediction," IEEE Access, vol. 11, pp. 69022-69040, 2023.

[21] R. S. Baker and K. H. Yacef, "The state of educational data mining in 2009: A review and future visions," Journal of Educational Data Mining, vol. 1, no. 1, pp. 3-17, 2009.

[22] S. Li, Y. Liu, and Q. Yang, "Federated learning for privacy-preserving collaborative learning in education," in Proceedings of the 2021 AAAI Spring Symposium on AI for Education, 2021.



10.22214/IJRASET

INDEX COPERNICUS
45.98

ISRA JIF
IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY