International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

BEYOND ACADEMIC METRICS: PREDICTING STUDENT PERFORMANCE USING ENGAGEMENT AND CONTEXTUAL INDICATORS

Author: Bibhu Ranjan Sahoo
Co-Author: Dr S Rangaswamy
University: Bharatiya Engineering Science and Technology Innovation University

Abstract
This study investigates the role of engagement and contextual indicators in predicting student academic performance using machine learning techniques. Traditional academic metrics such as attendance and grades often fail to capture behavioral and psychological aspects of learning. Using a dataset of 177 student records, this study evaluates Logistic Regression, Decision Tree, and Random Forest models to analyze the influence of engagement-related factors on academic outcomes. Experimental results show that psychological engagement is the most influential predictor of performance. Logistic Regression and Random Forest achieved the highest prediction accuracy of 83%. The findings highlight the importance of incorporating engagement indicators into modern educational analytics systems.

1. Introduction and Literature

Educational institutions increasingly rely on data-driven approaches to understand and improve student performance. Traditional evaluation methods typically focus on academic indicators such as attendance and assessment results. However, recent research suggests that behavioral engagement and learning motivation play critical roles in academic success [1], [2].

Machine learning techniques have been widely applied in educational research to analyze student behavior and predict academic outcomes. Studies have demonstrated the effectiveness of predictive models in identifying patterns of student engagement and motivation [3], [4]. These models allow educators to analyze large datasets and uncover relationships between psychological engagement and academic achievement.

Several researchers have explored engagement detection and behavioral analysis in classroom environments using machine learning algorithms [2], [5]. Other studies have investigated how supportive learning environments and students' beliefs influence learning motivation and adoption of emerging technologies such as artificial intelligence [6].

Furthermore, recent work examining the impact of artificial intelligence usage on students' autonomous learning behavior suggests that engagement with technology can significantly influence learning outcomes [7]. Despite these advances, many predictive studies still focus primarily on academic metrics rather than behavioral indicators.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Table 1: Literature Review Table

Ref	Authors	Dataset / Subject	Methods / Algorithms	Key Findings / Results
[1]	Dalal, P.; Beniwal, G.; Sharma, V.; Garg, P.; Ahmed, K.	Student motivation and engagement data collected from learning platforms and academic records	Logistic Regression, Decision Tree, Random Forest, SVM	Machine learning models can effectively predict student engagement and motivation using behavioral and academic indicators.
[2]	Kaur, A.; Mustafa, A.; Mehta, L.; Dhall, A.	Classroom video data capturing student facial expressions and behavior	Computer Vision, Deep Learning, Engagement Detection Models	Automatic engagement detection can identify attentive and disengaged students using facial and behavioral cues.
[3]	Fredricks, J.; Blumenfeld, P.; Paris, A.	Educational research on behavioral, emotional, and cognitive engagement	Conceptual framework and empirical analysis	Student engagement is a multidimensional construct strongly linked to academic achievement and learning outcomes.
[4]	Wang, F.; King, R. B.; Chai, C. S.; Zhou, Y.	University students learning Artificial Intelligence	Structural Equation Modeling (SEM), Survey-based	Supportive learning environments and expectancy-value beliefs significantly influence students' willingness
		courses	analysis	to learn AI.
[5]	Huang, X.; Hew, K.	University student online learning engagement data	Machine Learning Analysis of learning activity logs	Machine learning methods can identify patterns of student engagement in online learning environments.
[6]	Baker, R. S.; Inventado, P. S.	Educational data from learning management systems and tutoring systems	Educational Data Mining techniques	Educational data mining enables prediction of student behavior, performance, and engagement through large-scale educational datasets.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

[7]	Romero, C.; Ventura, S.	Educational datasets from online learning systems	Data Mining and Machine Learning approaches	Educational data mining techniques help discover patterns in student learning and support personalized education.
[8]	Siemens, G.	Learning analytics systems and educational data platforms	Learning Analytics frameworks	Learning analytics provides tools for measuring and improving learning outcomes using educational data.
[9]	Ferguson, R.	Educational technology environments and digital learning platforms	Learning Analytics methodologies	Learning analytics supports data-driven decision-making in education to improve student success and engagement.
[10]	Kotsiantis, S.	Educational datasets used in academic performance prediction	Machine Learning techniques including classification algorithms	Machine learning methods are effective in predicting student performance and identifying at-risk students.
[11]	Tempelaar, D.; Rienties, B.; Giesbers, B.	Higher education student engagement datasets	Statistical analysis and learning analytics models	Student engagement positively correlates with academic performance in higher education.
[12]	Henrie, C.; Halverson, L.; Graham, C.	Technology-mediated learning environments	Engagement measurement frameworks and learning analytics	Measuring engagement in online learning environments is essential for understanding student learning behavior.

Table 1 summarizes prior studies that explore student engagement, learning analytics, and machine learning approaches in educational research. The literature highlights the growing use of data-driven techniques to analyze learning behavior and predict academic performance. However, many existing studies focus primarily on interaction data or traditional academic indicators. This study extends previous work by incorporating engagement and contextual factors to improve prediction accuracy.

This study aims to address this gap by investigating whether engagement-related indicators can improve the prediction of student academic performance compared to traditional metrics.

2. Experiment / Study

The dataset used in this study contains 177 student records and includes the following

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

variables: attendance percentage, psychological engagement score, language barrier score, and family support score. The target variable represents student performance level (0 = low performance, 1 = high performance).

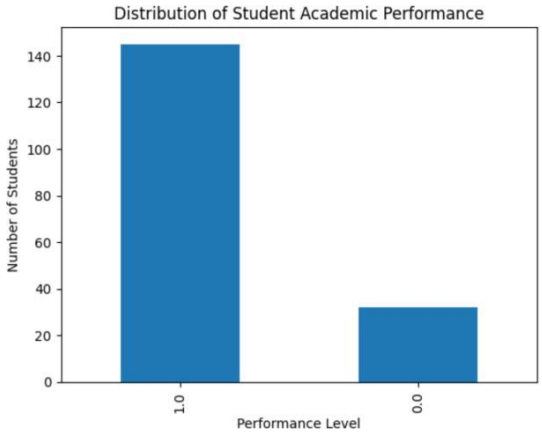

Figure 1: Distribution of student performance levels in the dataset.

Exploratory data analysis was conducted to identify relationships between engagement indicators and academic performance. Correlation analysis revealed that psychological engagement had the strongest relationship with performance, followed by attendance and family support.

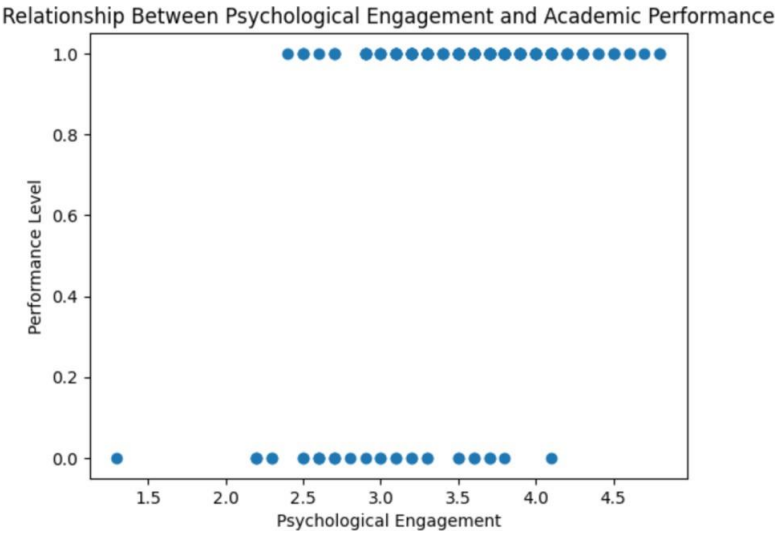

Figure 2: Relationship between psychological engagement and academic performance.

Three machine learning models were implemented: Logistic Regression, Decision Tree, and Random Forest. The dataset was divided using an 80–20 train–test split, resulting in 141 training samples and 36 testing samples.

3. Result and Discussion

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

The experimental results demonstrate that machine learning models can effectively predict student academic performance using engagement indicators. Logistic Regression and Random Forest achieved the highest accuracy of 83%, while the Decision Tree model achieved an accuracy of 78%.

Table 2: Machine Learning Model Performance

Model	Accuracy
Logistic Regression	83%
Decision Tree	78%
Random Forest	83%

The results presented in **Table 2** indicate that Logistic Regression and Random Forest provide the highest prediction accuracy.

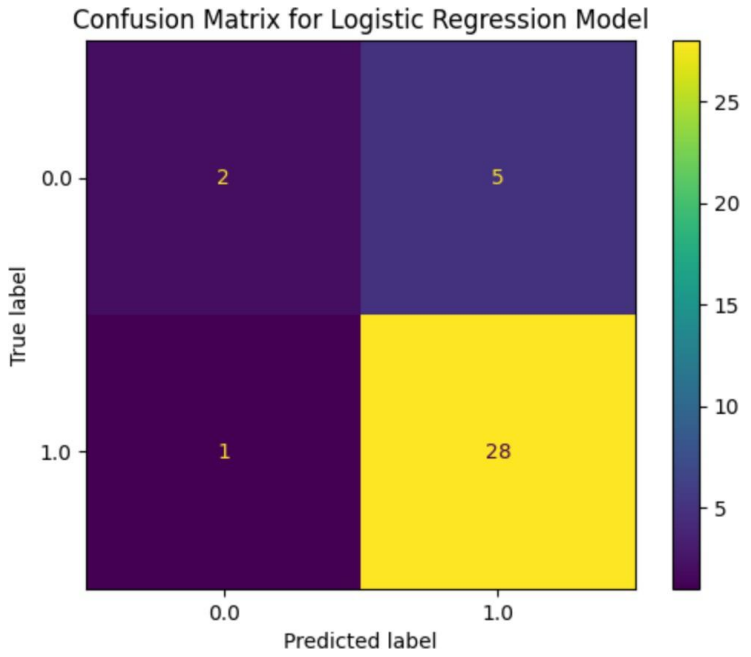

Figure 3: Confusion matrix showing prediction outcomes for the Logistic Regression model.

Feature importance analysis derived from the Random Forest model indicates that psychological engagement is the most significant predictor of academic performance, followed by attendance, language barrier score, and family support score.

Feature importance analysis from the Random Forest model indicates that psychological engagement is the most significant predictor.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Table 3: Feature Importance Scores

Feature	Importance Score
Psychological Engagement	0.47
Attendance	0.24
Language Barrier	0.15
Family Support	0.13

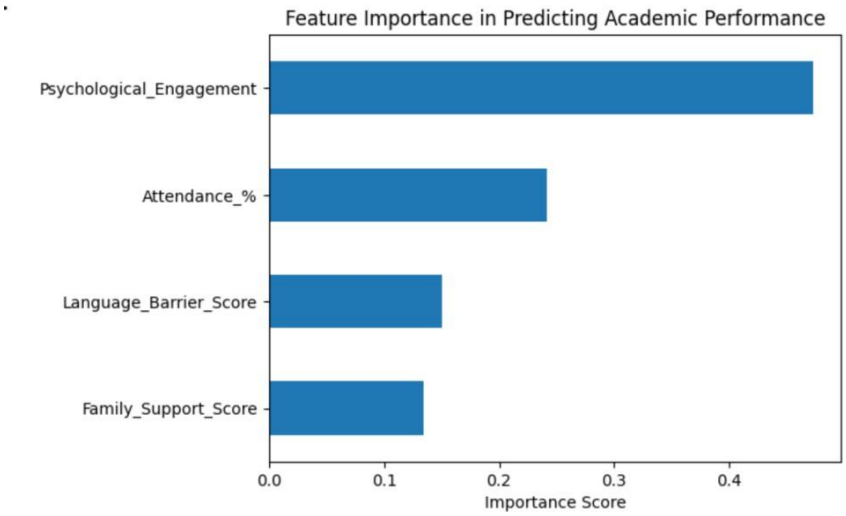

Figure 4: Relative importance of features in predicting academic performance.

As shown in **Table 3** and **Figure 4**, psychological engagement is the most influential predictor of student performance.
These findings reinforce previous research suggesting that engagement-based indicators provide deeper insights into student learning outcomes than traditional academic metrics alone [8], [9].

4. Conclusion

This study examined the predictive value of engagement and contextual indicators in student academic performance using machine learning techniques. Results demonstrate that psychological engagement is the strongest predictor of academic success, while Logistic Regression and Random Forest achieved the highest predictive accuracy of 83%. The findings highlight the importance of incorporating behavioral and motivational indicators into educational analytics systems. Future research may extend this work by incorporating larger datasets and additional learning analytics indicators.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

5. References

[1] Predicting Student Motivation and Engagement through Machine Learning Models.
[2] Student Engagement Detection in Classroom Using Machine Learning Algorithms.
[3] Modeling Students' Behavioral Engagement through Different In-Class Behavior Styles.
[4] University Students' Intentions to Learn Artificial Intelligence: The Roles of Supportive Environments and Expectancy–Value Beliefs.
[5] The Impact of AI Usage on University Students' Willingness for Autonomous Learning.
[6] Baker, R. & Inventado, P. Educational Data Mining and Learning Analytics.

[7] Romero, C., & Ventura, S. Educational Data Mining: A Review of the State of the Art.
[8] Siemens, G. Learning Analytics: The Emergence of a Discipline.

[9] Ferguson, R. Learning Analytics: Drivers, Developments, and Challenges.

[10] Kotsiantis, S. Use of Machine Learning Techniques for Educational Data Mining.

[11] Tempelaar, D., et al. Student Engagement and Academic Performance in Higher Education.
[12] Henrie, C., et al. Measuring Student Engagement in Technology-Mediated Learning.

6. Acknowledgement

The authors would like to thank the academic mentors, institutional support teams, and colleagues who provided guidance and feedback during the development of this study.