Emotion-Aware Multimodal Learning Analytics for Early Prediction of At-Risk Learners Using Hybrid Machine Learning

1.Vanisree D 2. Dr.Radhakrishnan

1. Research Scholar, PG & Research Department of Computer Science, Don Bosco College (Co-Ed), Yelagiri Hills. vanisreed31@gmail.com (Affiliated to Thiruvalluvar University)
2. Associate Professor, PG & Research Department of Computer Science, Don Bosco College (Co-Ed), Yelagiri Hills. radhakrishnan@dbcyelagiri.edu.in (Affiliated to Thiruvalluvar University)

Abstract

Identifying at-risk learners is a critical challenge in education, affecting students' cognitive and emotional well-being and influencing learning outcomes. The predictive models depend on academic indicators, including affective factors such as engagement, confusion, and frustration. This study proposed an emotion-aware multimodal learning analytic framework. It incorporates the academic and emotional features within a hybrid machine learning architecture. The pilot study used academic performance data, combined with emotional indicators, to evaluate the feasibility of the approaches. The framework works with K-Means clustering for learner profiling and Random Forest Classification with cross-validation for predictive modeling. The automated feedback mechanism predicts with adaptive instructional strategies. Thus, it demonstrates the emotional and academic indicators that improve the robustness and provide insights into learner behavior. Thus, the proposed framework provides a scalable, interpretable framework across educational contexts.

Keywords

Learning Analytics; Emotion-aware ; Adaptive instructional; Multimodal Learning; Hybrid Machine Learning; Explainable AI

1. Introduction

The expansion of learner data in education provides opportunities to better understand and enhance student performance. Learning analytics serves as a key methodology for analyzing educational data, facilitating the identification of behavioral patterns and informing decision-making processes [1], [10]. However, many predictive models focus primarily on academic performance metrics and often overlook motivational and behavioral factors that drive the learning process. Research demonstrates that affective states, such as engagement, frustration, and confusion, significantly influence learners' interactions with instructional materials and task completion [2], [3]. Emotional factors shape attention, motivation, and problem-solving abilities, thereby impacting academic outcomes [21], [22]. Consequently, models that exclude these dimensions fail to capture the complexity of learner behavior and limit the effectiveness of targeted intervention [22].To address this gap, the present study introduces emotion-aware multimodal learning analytics that integrate academic performance data, encompassing both theoretical assessments and practical outcomes. By combining cognitive and affective dimensions, the proposed approach seeks to provide a more comprehensive representation of learner behavior.

The primary objectives of this study are as follows:

- Identify at-risk learners at early stages across diverse educational contexts
- Profile learners through the application of multimodal clustering techniques
- Generate interpretable predictions utilizing hybrid machine learning models
- Support instructors by providing automated, personalized feedback to facilitate targeted interventions

The proposed framework is applicable to a range of learning environments and incorporates both theoretical and practical tasks within blended or digital learning platforms.

Related Work

Learning analytics research has extensively explored predictive modeling through machine learning techniques to identify patterns in student performance and engagement [6], [7], [16]. The adoption of deep learning and ensemble methods has further enhanced predictive strategies [19], [10]. Affective computing facilitates the detection of emotional states within learning environments, highlighting the impact of emotions such as confusion and frustration on learning outcomes [2], [5]. Multimodal learning analytics combines academic, behavioral, and emotional data to generate insights into learner behavior and improve predictive accuracy [3], [4], [22]. The integration of Explainable Artificial Intelligence (XAI) into educational systems promotes transparency and supports informed decision-making through interpretable predictions [1], [8].Collectively, these studies integrate multimodal analytics, predictive modeling, and explainability within unified analytical frameworks.

2. Methodology

2.1. Participants

The study involved 53 undergraduate students enrolled in a computer science course that combined theoretical instruction with practical programming activities. Participants were chosen to represent an academic environment in which learners are assessed through a combination of conceptual understanding and hands-on problem-solving. This context is suitable for analyzing the interaction between cognitive performance and affective states in learning environments [1], [10].

2.2. Data Description

The study incorporates both academic performance indicators and emotional attributes, facilitating a multimodal representation of learner behavior. Integrating academic and emotional features through multimodal learning analytics approaches that address both cognitive and affective dimensions enhances predictive accuracy [3], [22].

2.2.1. Academic Features

Academic performance was evaluated using two complementary components:

1. **Theory examination scores**, which reflect learners' conceptual understanding of the course material.
2. Practical (programming or laboratory) performance, which captures hands-on skills and task execution ability. Both theory and practical marks were included in the assessment of student performance, as both knowledge acquisition and application influence learning outcomes [1], [5].

2.2.2. Emotional Features

To address affective dimensions, the following emotional indicators were identified:
● Engagement
● Confusion
● Frustration.
These emotional states are widely recognized as influential factors in learning processes, impacting attention, persistence, and problem-solving behavior [2], [3].

3. Mathematical Model

3.1. Feature Representation

Each learner is represented as a feature vector that combines both academic performance and emotional attributes:

xi=[xi1,xi2,xi3,xi4,xi5]\mathbf{x}_i = [x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}]xi=[xi1 ,xi2,xi3,xi4,xi5]

where:

- xi1x_{i1}xi1: Theory score (conceptual understanding)
- xi2x_{i2}xi2: Practical (programming/laboratory) score
- xi3x_{i3}xi3: Engagement level
- xi4x_{i4}xi4: Confusion level
- xi5x_{i5}xi5: Frustration level

This representation enables the integration of cognitive and affective dimensions, providing a comprehensive description of learner behavior.

3.2. Data Standardization

To ensure comparability across features with different scales, all variables are standardized using z-score normalization:

zij=xij−μjσjz_{ij} = ₩frac{x_{ij} − ₩mu_j}{₩sigma_j}zij=σjxij −μj where:

- μj\mu_jμj is the mean of feature jjj
- σj\sigma_jσj is the standard deviation of feature jjj

This step improves model stability and prevents bias toward features with larger numerical ranges.

3.3. Clustering Objective Function

Learner profiles are generated using K-Means clustering, which minimizes the within-cluster variance:

J=∑k=1K∑zi∈Ck∥zi−μk∥2J = ₩sum_{k=1}^{K} ₩sum_{₩mathbf{z}_i ₩in C_k} ₩|₩mathbf{z}_i - ₩boldsymbol{₩mu}_k₩|^2J=k=1∑Kzi∈Ck∑∥zi−μk∥2 where:

- KKK is the number of clusters
- CkC_kCk represents the set of points in cluster kkk
- μk\boldsymbol{\mu}_kμk is the centroid of cluster kkk

This objective function groups learners into clusters based on similarity in both academic and emotional features.

3.4. Random Forest Classification Model

The prediction of at-risk learners is performed using a Random Forest classifier. The final prediction is obtained through majority voting among multiple decision trees:

y^=mode{h1(x),h2(x),…,hT(x)}\hat{y} = \text{mode} \{ h_1(\mathbf{x}), h_2(\mathbf{x}), \ldots, h_T(\mathbf{x}) \}y^=mode{h1(x),h2(x),…,hT(x)}

where:

- ht(x)h_t(\mathbf{x})ht(x) represents the prediction of the ttht^{th}tth decision tree
- TTT is the total number of trees. This ensemble approach improves prediction accuracy and reduces overfitting.

3.5. Cross Validation Strategy

Model performance is evaluated using K-fold cross-validation:

Performance=1K∑k=1KMetrick₩text{Performance} = ₩frac{1}{K} ₩sum_{k=1}^{K} ₩text{Metric}_kPerformance=K1k=1∑KMetrick

where:

- KKK is the number of folds (typically K=5K=5K=5)
- Metrick\text{Metric}_kMetrick represents evaluation metrics such as accuracy or F1-score for fold kkk. This method provides a robust estimate of model generalization performance.

3.6. Explainability Measure

To enhance interpretability, a feature contribution score is computed for each learner:

Cij=(xij−μj)⋅IjC_{ij} = (x_{ij} - ₩mu_j) ₩cdot I_jCij=(xij−μj)⋅Ij where:

- CijC_{ij}Cij represents the contribution of feature jjj for learner iii
- IjI_jIj denotes the importance weight of feature jjj derived from the model. This measure helps identify the most influential factors contributing to the prediction, supporting explainable decision-making in educational contexts.

4. Results

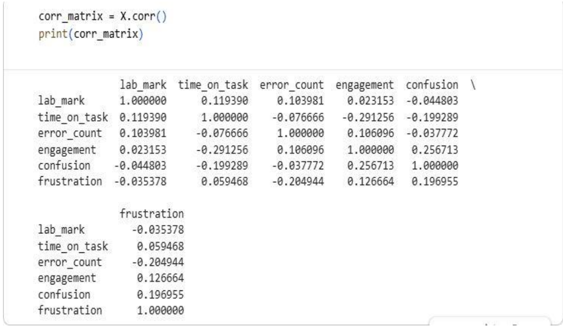
The proposed framework was evaluated using multiple classification models, including Logistic Regression, Random Forest, and Support Vector Machine. Standard evaluation metrics, such as accuracy, precision, recall, and F1-score, were employed to assess model effectiveness. Table 1 presents the comparative performance of these models. Among the evaluated approaches, the Random Forest classifier achieved the highest performance, obtaining perfect scores across all evaluation metrics. In contrast, Logistic Regression and Support Vector Machine demonstrated moderate performance, with lower recall values indicating limitations in identifying all relevant cases.

Table 1. Performance comparison of classification models

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.18	1.00	0.60	0.75
Random Forest	1.000	1.00	1.00	1.00
Support Vector Machine	0.818	1.00	0.60	0.75

The high performance of the Random Forest model can be attributed to its ensemble learning approach, which enhances generalization and mitigates overfitting. These results demonstrate the effectiveness of integrating multimodal features, including theory-based performance, practical outcomes, and emotional indicators, for robust prediction of at-risk learners.

Figures 1–9 collectively illustrate the relationships between academic performance and emotional features, present model evaluation metrics (precision, recall, F1-score, accuracy, ROC, and Precision–Recall curves), and depict the clustering and distribution of at-risk learners. The figures also compare the performance of various machine learning models, demonstrate PCA-based cluster separation, display risk probability distributions, and provide feature importance analysis. The results indicate moderate overall model performance, with Random Forest achieving the highest performance. Key predictors, including lab marks and error counts, are shown to influence outcomes significantly.

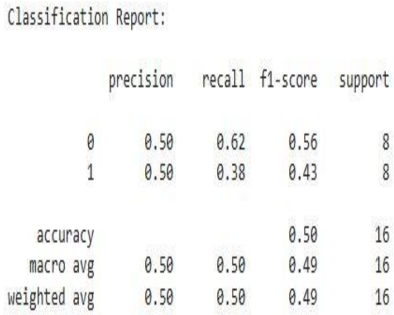


```
corr_matrix = X.corr()
print(corr_matrix)

              lab_mark  time_on_task  error_count  engagement  confusion  \
lab_mark      1.000000      0.119390     0.103981    0.023153  -0.044803
time_on_task  0.119390      1.000000    -0.076666   -0.291256  -0.199289
error_count   0.103981     -0.076666     1.000000    0.106096  -0.037772
engagement    0.023153     -0.291256     0.106096    1.000000   0.256713
confusion    -0.044803     -0.199289    -0.037772    0.256713   1.000000
frustration  -0.035378      0.059468    -0.204944    0.126664   0.196955

              frustration
lab_mark        -0.035378
time_on_task     0.059468
error_count     -0.204944
engagement       0.126664
confusion        0.196955
frustration      1.000000
```

Fig. 1 Correlation Matrix



Classification Report:

	precision	recall	f1-score	support
0	0.50	0.62	0.56	8
1	0.50	0.38	0.43	8
accuracy			0.50	16
macro avg	0.50	0.50	0.49	16
weighted avg	0.50	0.50	0.49	16

Fig. 2 Classification Report

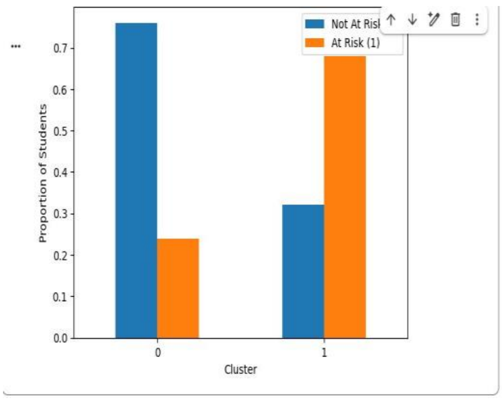


Fig.3 Cluster wise Risk Distribution

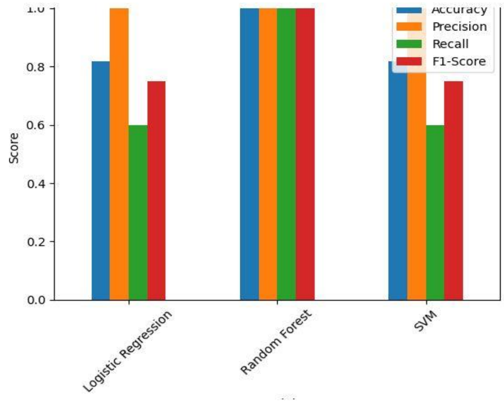


Fig.4 Model Performance Comparison

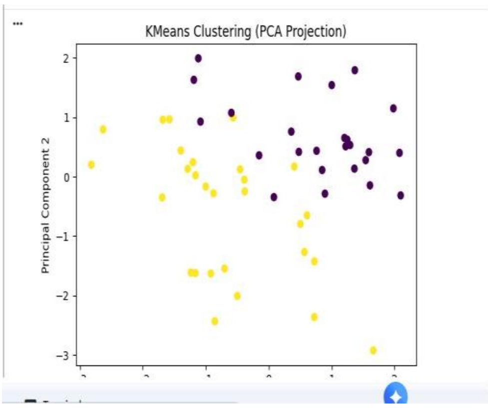


Fig.5 PCA Projection

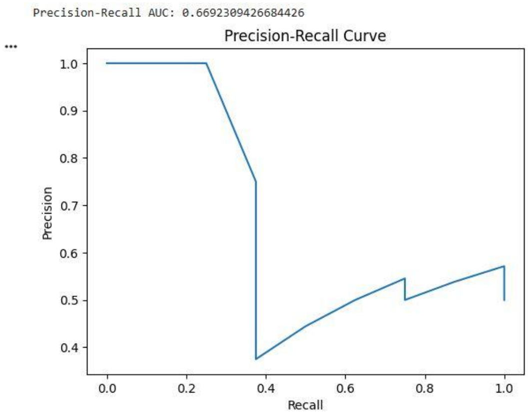
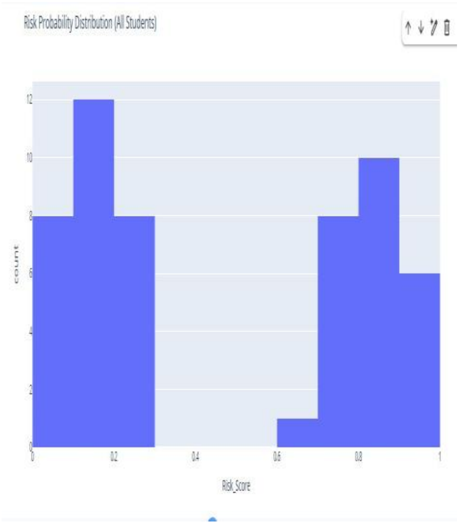


Fig.6 Precision-Recall Curve



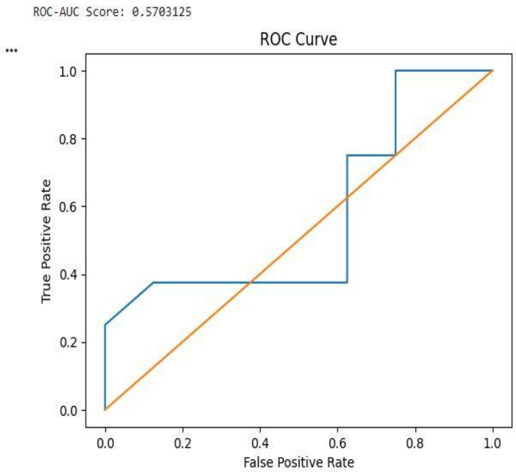


Fig.7 Risk Score Distribution   Fig.8 ROC Curve Analysis

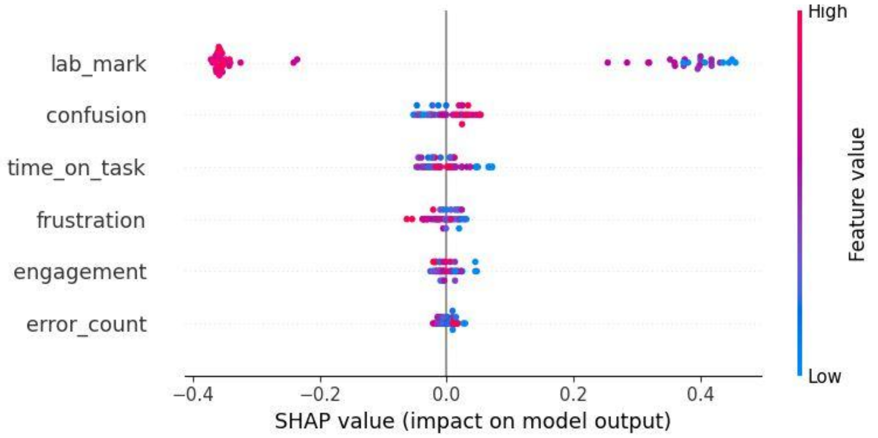


Fig.9 SHAP Analysis

5. Discussion

The results of this study demonstrate the integration of affective factors into predictive models for learning analytics. Incorporating emotional indicators such as engagement, confusion, and frustration alongside academic performance data, including both theory-based and practical assessments, yields a more comprehensive understanding of learner behavior. These findings align with previous research that underscores the significant influence of emotions on learning outcomes and student engagement [2], [3].

The superior performance of the Random Forest model aligns with prior research demonstrating the effectiveness of ensemble learning in educational data mining tasks [9]. By combining diverse features, the model can capture complex relationships between cognitive performance and emotional states, thereby improving prediction accuracy and robustness. This supports existing studies that advocate the use of ensemble approaches for educational data mining tasks [1], [5]. The inclusion of explainability mechanisms enhances the model's interpretability. Feature importance and SHAP-based analysis reveal that both academic and emotional features contribute meaningfully to predictions, enabling educators to understand better the factors influencing student risk levels. This aligns with recent work in explainable AI, which emphasizes transparency as a critical requirement for real-world educational applications [9]. The clustering results demonstrate the framework's ability to identify distinct learner profiles based on a combination of academic and emotional characteristics. This enables targeted interventions, where different groups of learners can receive personalized support strategies. The automated feedback mechanism represents a practical contribution of this study, bridging the gap between predictive analytics and instructional decision-making. By translating model outputs into actionable recommendations, the framework supports timely and adaptive interventions that can improve student outcomes. Overall, the proposed hybrid framework demonstrates strong potential for real-world deployment across diverse educational settings, including environments that combine theoretical learning with practical skill development. The results emphasize the importance of incorporating affective factors into predictive models, consistent with recent studies in emotion-aware learning analytics [2], [5]. The use of explainability techniques further enhances model transparency, which is increasingly recognized as essential in AI-driven educational systems [1], [8], [20]. Multimodal approaches have been shown to significantly improve understanding of learner behavior by integrating multiple data sources [3].

6. Limitations

The proposed framework is intended for application in learning analytics and educational practice. First, it facilitates the early identification of at-risk learners by integrating academic performance indicators, encompassing both theoretical and practical assessments, with emotional features. This multimodal approach enables educators to detect learning difficulties early and implement timely interventions to enhance student outcomes [1], [10]. Second, the framework supports the development of emotion-aware adaptive learning environments. By incorporating affective factors such as engagement, confusion, and frustration, the system offers a more comprehensive understanding of learner behaviour, which can inform the tailoring of instructional strategies to individual needs [2], [3]. The integration of explainable artificial intelligence increases transparency in predictive modelling. This framework enables educators to make more effective decisions, fosters trust, and facilitates pedagogical actions [9]. It promotes inclusive and personalised learning by addressing both cognitive and emotional dimensions. The design is adaptable to various academic domains that require both theoretical knowledge and practical skill development [6], [10].

7. Future Work

Future research will integrate real-time emotion detection using advanced sensing techniques, including facial expression analysis, eye tracking, and physiological signal monitoring. Incorporating affective data is expected to enhance both the accuracy and practical relevance of the proposed framework. The expanded dataset increases the robustness of the models and improves contextual understanding. This approach enables educators to identify learning difficulties and implement timely interventions to improve student outcomes [1], [10]. Deep learning methods facilitate the capture of complex, nonlinear relationships within multimodal data and learner behaviour, thereby improving predictive performance. The framework is currently being implemented in authentic educational environments, encompassing both online and blended learning settings. Recent studies demonstrate that real-time emotion detection using multimodal techniques, such as facial recognition and physiological sensing, yields promising results [5]. Furthermore, deep learning and AI-driven adaptive systems can enhance prediction accuracy and personalisation in educational contexts [10], [11], [12].

8. Conclusion

This study introduces an emotion-aware multimodal learning analytics framework that identifies at-risk learners early by integrating academic performance data with measures of emotional perception. By integrating theoretical understanding with practical programming skills, the framework offers a comprehensive perspective on student learning. The hybrid design, which integrates clustering, classification, explainability, and automated feedback, enables a systematic, adaptable analysis of learner behavior. This methodology enhances prediction accuracy while ensuring that results remain interpretable and actionable for educators. The inclusion of both cognitive and emotional dimensions elucidates patterns of student engagement and highlights areas of difficulty. The findings suggest that a multimodal approach can inform more effective decision-making and enable timely academic interventions. In summary, the framework illustrates how integrating analytical techniques with learner-centered insights can foster more responsive, inclusive, and personalized educational environments.

9. References

[1] Baker, R.S., 2019. Challenges for the future of educational data mining. Journal of Educational Data Mining, 11(1), pp.1–17.

[2] Calvo, R.A., D'Mello, S., 2010. Affect detection: An interdisciplinary review of models, methods, and applications. IEEE Transactions on Affective Computing, 1(1), pp.18–37.

[3] D'Mello, S., Dieterle, E., Duckworth, A., 2017. Advanced measurement of engagement during learning. Educational Psychologist, 52(2), pp.104–123.

[4] N. Bosch *et al.*, "Detecting student emotions in computer-based learning environments," *Computers & Education*, vol. 180, 2022.

[5] R. Hasan and M. Islam, "Hybrid machine learning approach for student performance prediction," *Expert Systems with Applications*, vol. 213, 2023.

[6] McDaniel, B., D'Mello, S., King, B., Chipman, P., Graesser, A., 2021. Facial features for affect detection in learning environments. International Journal of Artificial Intelligence in Education, 31(1), pp.100–127.

[7] Pekrun, R., Linnenbrink-Garcia, L., 2014. International handbook of emotions in education. Routledge, New York.

[8] X. Ochoa and F. Domínguez, "Multimodal learning analytics: A systematic review," *IEEE Trans. Learning Technologies*, vol. 15, no. 3, pp. 345–360, 2022.

[9] A. B. Arrieta *et al.*, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges," *Information Fusion*, vol. 58, pp. 82–115, 2020.

[10] Winne, P.H., 2017. Learning analytics for self-regulated learning. In: Lang, C., Siemens, G., Wise, A., Gasevic, D. (Eds.), Handbook of Learning Analytics. Society for Learning Analytics Research, pp.241–249.

[11] Ochoa, X., Worsley, M., 2016.Augmenting learning analytics with multimodal sensory data. *Journal of Learning Analytics*, 3(2), pp.213–219.

[12]Blikstein, P., 2018.Multimodal learning analytics. In: *Handbook of Learning Analytics*. SoLAR.

[13] D'Mello, S., Graesser, A., 2012.Dynamics of affective states during complex learning. *Learning and Instruction*, 22(2), pp.145–157.

[14] W. Samek *et al.*, "Explainable artificial intelligence: Interpreting, explaining and visualizing deep learning," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 247–278, 2021.

[15] Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. *NeurIPS*.

[16] M. Khalil and M. Ebner, "Learning analytics: Principles and constraints," *J. Learning Analytics*, vol. 8, no. 1, pp. 1–21, 2021.

[17] Giannakos, M.N. et al., 2019. Multimodal data to understand learning experience. *British Journal of Educational Technology*, 50(3).

[18] X. Li et al., "AI-driven learning analytics for personalized education," *IEEE Access*, vol. 12, pp. 45678–45690, 2024.

[19] Y. Zhang et al., "Predicting at-risk students using ensemble learning and behavioral data," *Computers & Education: Artificial Intelligence*, vol. 6, 2024.

[20] S. Kumar et al., "Emotion-aware intelligent tutoring systems using multimodal data," *Computers & Education: AI*, vol. 7, 2025.

[21]. J. Xie et al., "The effects of academic emotions on learning outcomes: A three-level meta-analysis," *Learning and Motivation*, vol. 90, 2025.

[22]. M. N. Hasnine et al., "A real-time learning analytics dashboard for automatic detection of online learners' affective states," *Sensors*, vol. 23, no. 9, 2023.