SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 7608-7611
Open Access. Online & Print

DOI: 10.5281/zenodo.122126587

ARTIFICIAL INTELLIGENCE DRIVEN MULTIMODAL ANALYTICS FOR HOLISTIC STUDENT SUCCESS PREDICTION

Vinay G[1], Rajeev Ranjan[2]

[1,2] School of Computer Science and Applications, REVA University, Bengaluru, India

Received: 11/12/2024
Accepted: 25/02/2025

Corresponding author: Vinay G (vinay.g@reva.edu.in)

ABSTRACT

This paper describes a multimodal analytics frame- work for predicting student performance using various data types. Existing models mainly utilize educational data, thus failing to consider other factors such as quality of sleep, stress, and time devoted to studies. This framework includes unified ingestion, preprocessing, time series, and quality assurance of various data types. This paper has utilized data collected from 1,200 undergraduate students over a semester, consisting of grade point average, attendance, quality of sleep, perceived stress, time devoted to studies, and learning engagement. After achieving 94.7% data completeness, various experiments were conducted on the data using XGBoost, TabNet, and a stacking ensemble. The best model has recorded an accuracy of 91.2%, an improvement of 10.9% compared to a single academic-based model. The feature importance plot has shown that wellness factors contribute to performance prediction. This framework will be helpful in early identification of at-risk students, considering various factors such as privacy and fairness.

KEYWORDS: Educational data mining, learning analytics, multimodal data fusion, student performance prediction, wellness analytics, machine learning

1. INTRODUCTION

The growth in the number of digital learning environments has resulted in an increase in the quantity of available data for educational analysis. Learning analytics and educational data mining have been used to support prediction, intervention, and decision making in educational institutions, based on available data records [1-3], [7]. The existing models, however, have shown bias towards grade-based factors, including students' past performance and attendance. This has resulted in an inability to account for other factors affecting students' performance. Not only academic history, but wellness and behavior factors also affect the performance of the students. Previous research indicates the relationship between sleep habits, psychological distress, and mental health with academic performance [8-11]. Moreover, student signals related to digital platforms can identify the pattern of self-regulated learning, which is not possible to identify using transcript data alone [2-3]. A predictive model that does not consider these modalities is likely to miss important sources of risk. The second problem is methodology. The student data is not only scattered but also collected at different times and levels of completeness. This makes temporal alignment and quality control difficult [12-14]. Hence, even if data is relevant and available for use, it may not be readily integrated into a coherent and consistent predictive process. This paper fills these gaps by proposing a multimodal data fusion framework for student success prediction. The framework incorporates academic, wellness, and behavioral information, validation and consistency checks before model training, and performance evaluation of multiple machine learning models with interpretability and fairness considerations. The results of this work demonstrate the benefits of multimodal modeling for student success prediction.

2. RELATED WORK

Existing work on educational analytics has shown the potential benefits of prediction modeling for student outcomes, especially when large institutional datasets are available [1–3]. The participation-centric and interpretable approaches also show the potential benefits of student behavior observed during the semester on the prediction of final performance [4]. However, many such studies still focus on more narrow aspects of student interaction with academics or platforms. Recent work in multimodal and cross-source integration has shown that integrating different types of signals can lead to better quality models. A survey on educational data mining emphasizes the importance of better data fusion pipelines [2], [3]. Other related research in learning analytics emphasizes the importance of not having a one-size-fits-all approach by considering individual contexts [6]. There is also related work by Chen *et al.* on how latent patterns in educational technology can be uncovered from complex data sets [5]. Wellness factors are particularly important, although they tend to be underrepresented in predictive models. Both sleep loss and sleep duration have been linked to academic performance [8], [9], and psychological distress, as well as mental health issues, have been demonstrated to impact higher education performance [10], [11]. This suggests that predictive systems should take account of physical and mental well-being, not just past academic performance.

Moreover, issues of governance have been highlighted in the literature, which is essential in managing combinations of self-reported and administrative data, managing missing data, managing inconsistent time stamps, and managing con- sent boundaries [12], [13]. The ethical aspects of learning analytics have also been highlighted, which include issues of privacy, transparency, and intervention ethics [14]. The proposed framework extends the existing literature by considering multimodal integration, validation, and ethics in one predictive process.

3. PROPOSED METHODOLOGY

This section presents the design of the multimodal frame- work and the evaluation process used to assess predictive performance.

3.1. Framework Overview

In the proposed framework, it is suggested to integrate the three data streams, which include academic records, wellness data, and behavioral activity logs. It is proposed to design the workflow with three stages: data collection, multi-modal integration, and predictive modeling. It is suggested to design the workflow in this way to ensure the smooth transition from the heterogeneous raw data to the consolidated representation of the student learning conditions.

The framework is intended to facilitate

continuous monitoring rather than assessment. The alignment of the information collected over time can be used to monitor the evolution of the students' states during the semester, which in turn can be used to support the estimation of risk. Fig. 1 illustrates the diagram of the extracted framework from the source paper.

3.2. Data Modalities

The academic data collected is based on institutional records, which include grade point average, midterm and assignment grades, percentage of attendance, and status of progress in courses. These are a direct measure of academic achievement in the past and current academic record.

The data on wellness are collected using a structured self- report method, which includes sleep duration, sleep quality, perceived stress, physical activity, and quality of social relationships. These are the variables that provide a description of the physical and psychological setting in which learning occurs.

Behavioral data is collected from a learning environment and includes weekly study hours, interaction with a learning management system, library patterns, and assignment submittal. The data collected is a combination of engagement and self-regulated learning behavior.

3.3. Data Preprocessing and Temporal Alignment

The quality of data is ensured through a preprocessing pipeline in which noisy data is removed, missing data is imputed, and numerical data is normalized. All these operations help in reducing measurement bias.

Since the original data is collected at daily, weekly, and peri- odic intervals, the framework incorporates temporal alignment prior to model training. This means the data is transformed into a standard semester-based timeline to ensure the analysis of academic, wellness, and behavior variables together.

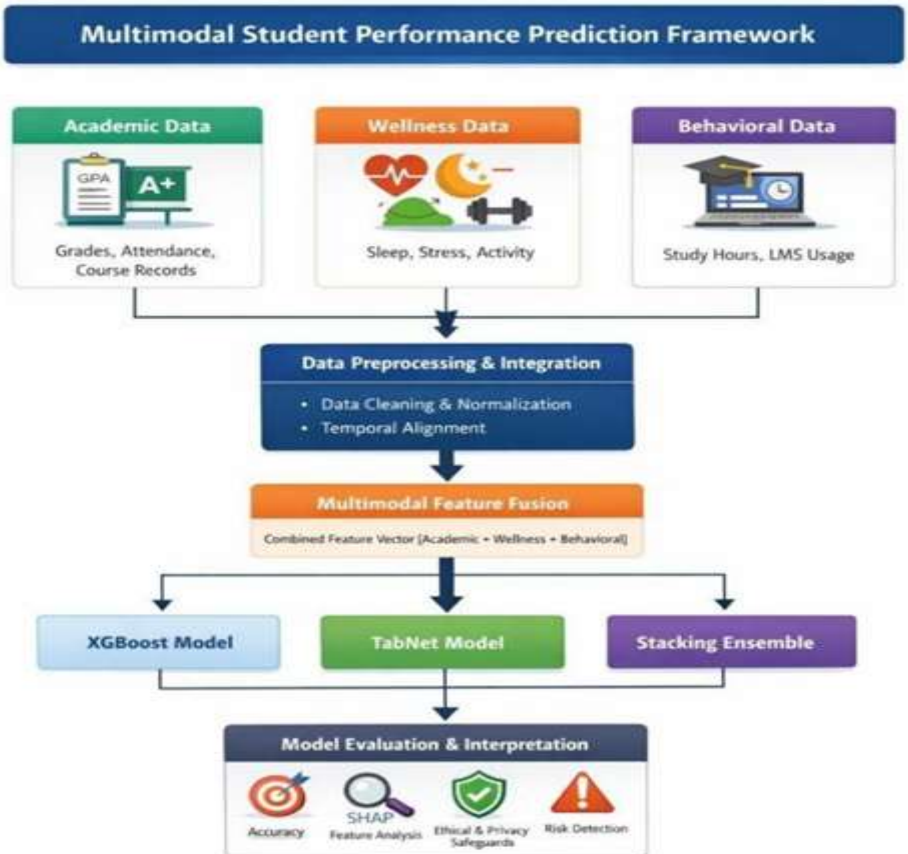

Figure 1: Multimodal student performance prediction framework.

Table 2: Predictive Performance of Evaluated models

Model	Accuracy	F1-score
XGBoost	87.7%	Not reported
TabNet	90.4%	Not reported
Stacking ensemble	91.2%	90.8%

3.4. Multimodal Fusion and Prediction Models

The framework employs early fusion, in which all modality specific features are concatenated into a single feature vector before the commencement of training. This choice enables the model to learn cross-modal dependencies directly rather than combining the results of separate predictions made by the model during execution.

The fused representation is defined in (1).

$$x = [x_{\text{academic}}, x_{\text{wellness}}, x_{\text{behavioral}}] \quad (1)$$

where

x_{academic} represents the academic feature vector,

x_{wellness} represents the wellness feature vector, and

$x_{\text{behavioral}}$ represents the behavioral feature vector.

The prediction models to be compared in this evaluation are XGBoost [16], TabNet [15], and a stacking ensemble model consisting of several strong learners.

Table 2: Accuracy by Feature Configuration

Configuration	Accuracy
Academic only	79.1%
Academic plus behavioral	80.8%
Academic plus wellness	85.4%
Full multimodal XGBoost	87.7%

This comparison enables a comparison of traditional gradient boosting methods and attention-based tabular learning methods in a multimodal context.

3.5. Evaluation Protocol and Safeguards

The evaluation of the model is done by performing stratified 10-fold cross-validation. The evaluation metrics used in the paper are accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve.

Interpretability is ensured via feature importance analysis and Shapley Additive Explanations. These tools can be used to determine which variables contribute the most to the prediction outcome [17]. Ethical measures are also integrated into the framework. Data collection is ensured via informed consent, and the model output is evaluated for privacy and fairness issues, like the previously established guidelines for learning analytics [14], [18].

4. EXPERIMENTAL RESULTS

The multimodal framework has shown impressive predictive performance for all models. The stacking ensemble model has shown the highest performance among all models, achieving 91.2% accuracy and 90.8% F1-score. TabNet has shown 90.4% accuracy, and XGBoost has shown 87.7% accuracy. This implies that the interaction between academic, wellness, and behavioral variables can be successfully captured by both attention-based models and ensemble models. As the results from the ablation test demonstrate, the inclusion of wellness and behavioral information improves the prediction performance beyond the academic features. The performance on the baseline with only the academic features reached 79.1%, while the performance with the addition of the behavioral features reached 80.8%, and the performance with the addition of the wellness features reached 85.4%. The performance with the full multimodal XGBoost approach reached 87.7%, while the stacking ensemble boosted the final performance to 91.2%. This further reinforces the understanding that the prediction process for student success must consider the overall conditions.

The framework also supports early warning use cases. By week 6, the model achieved 84.7% precision and approximately 79% recall for at-risk student detection while maintaining a low false positive rate. This level of performance is suitable for targeted academic support and timely intervention. In feature analysis, it is evident that the prior grade point average is the strongest individual predictor. However, the study hours, sleep duration, and stress level also contribute substantially to the prediction. The wellness indicators contribute to approximately 36.3% of the overall predictive contribution.

5. DISCUSSION

The results show that multimodal data fusion benefits student performance prediction over academic-only baselines. The 10.9 percentage gain over the academic-only model suggests that educational success is a holistic outcome, driven by learning behaviors, well-being, and academic achievement. The study also indicates that the variables of wellness are not peripheral. The information provided by sleep duration, stress, and study behaviors is arguably more useful for intervention than the information provided by the transcripts. This is particularly important for early warning systems, where the aim is not only to classify risk but also to understand the nature of the support that may be required.

Two limitations should be noted: some of the wellness data is self-reported and therefore subject to response biases; and the data is collected from a single institution and therefore may not generalize well to other settings. Future work should include testing across multiple institutions and exploring objective data from sensing systems as well as real-time intervention pipelines.

6. CONCLUSION

In this paper, a multimodal framework for student success prediction is proposed by integrating academic, wellness, and behavioral data into a unified machine learning process. The framework enhances the predictive accuracy of student success predictions and facilitates earlier detection of students who are at risk of failing. The findings indicate that educational analytics systems should move beyond academic-only models and incorporate broader contextual evidence of student well-being and engagement. Future research should focus on real-time multimodal monitoring, broader institutional validation, and individualized intervention strategies.

REFERENCES

[1] G. Siemens and P. Long, "Penetrating the fog: Analytics in learning and education," *EDUCAUSE Review*, vol. 46, no. 5, pp. 30–32, 2011.

[2] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 3, e1355, 2020.

[3] R. S. Baker and P. S. Inventado, "Educational data mining and learning analytics," in *Learning Analytics*. New York, NY, USA: Springer, 2014, pp. 61–75.

[4] W. Xing *et al.*, "Participation-based student final performance prediction model through interpretable genetic programming," *Computers in Human Behavior*, vol. 47, pp. 168–181, 2015.

[5] X. Chen *et al.*, "Detecting latent topics and trends in educational technologies," *Computers and Education*, vol. 151, 103855, 2020.

[6] D. Gasevic *et al.*, "Learning analytics should not promote one size fits all," *The Internet and Higher Education*, vol. 28, pp. 68–84, 2016.

[7] R. Ferguson, "Learning analytics: Drivers, developments and challenges," *International Journal of Technology Enhanced Learning*, vol. 4, no. 5–6, pp. 304–317, 2012.

[8] M. Hysing *et al.*, "Sleep and academic performance in later adoles- cence," *Journal of Sleep Research*, vol. 25, no. 3, pp. 318–324, 2016.

[9] G. Curcio, M. Ferrara, and L. De Gennaro, "Sleep loss, learning capacity and academic performance," *Sleep Medicine Reviews*, vol. 10, no. 5, pp. 323–337, 2006.

[10] H. M. Stallman, "Psychological distress in university students," *Australian Psychologist*, vol. 45, no. 4, pp. 249–257, 2010.

[11] D. Eisenberg, E. Golberstein, and J. B. Hunt, "Mental health and academic success in college," *The B.E. Journal of Economic Analysis and Policy*, vol. 9, no. 1, 2009.

[12] A. Pardo *et al.*, "Combining student self-reported and administrative data," *Computers in Human Behavior*, vol. 72, pp. 618–630, 2017.

[13] S. Knight and S. Buckingham Shum, "Theory and learning analytics," in *Handbook of Learning Analytics*, 2017, pp. 17–22.

[14] P. Prinsloo and S. Slade, "An elephant in the room: Learning analytics and ethics," in *Handbook of Learning Analytics*, 2017, pp. 308–320.

[15] S. O. Arik and T. Pfister, "TabNet: Attentive interpretable tabular learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, pp. 6679–6687.

[16] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[17] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, 2017.

[18] S. Barocas, M. Hardt, and A. Narayanan, *Fairness and Machine Learning*. Cambridge, MA, USA: MIT Press, 2019.