

Dynamic Multimodal Profiling with Self-Supervised Learning for Early Academic Risk Prediction

Shilpa K

Research Scholar and Associate Professor
Institute of Computer Science and Information Science
Srinivas University, Mangalore
Nrupathunga University, NT road Bengaluru
shilpaktr@gmail.com

Abstract — This paper presents a multimodal, data driven framework for forecasting academic risk that was developed utilizing a three phase approach that prioritized model variety, data quality, and open decision making. Phase 1 involves the construction of a big data pipeline that connects the records of over 10,000 students from colleges throughout Karnataka. The analysis revealed thirty traits that are useful for predictive modeling. Among these are behavior, mental health, demographics, and college courses. The researchers studied some classical and neural models in Phase 2. These are Neural Networks (NN), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Naive Bayes (NB), K-Nearest Neighbors (KNN) and Decision Trees (DT). Five assessments of pairs of comparison models with different patterns are regularly made: RF frequently commits less mistakes compared with NN and DT, retains more information, and is more precise. SVM has lower false negatives compared to NN and is dependable. Two of the areas that LR outperforms NB are precision and F1 measures. KNN, which has greater classification stability, is a major challenge to NN. These findings indicate that structured ensemble and distance based models are more effective for heterogeneous educational data. SHAP and LIME are examples of explainable AI tools that are part of Phase 3. These tools make it easy to share features, check for fairness and create unique intervention programs for each college. The findings indicate that the combination of multimodal feature engineering, ensemble driven modeling and interpretable AI constitutes a robust and ethical approach to predicting academic risks.

Keywords— Academic Risk, Multimodal Analytics, Neural Networks, Ensemble Learning, Model Comparison, Explainable AI.

I. INTRODUCTION

Early diagnosis of students who are likely to fail college makes it more difficult for universities to discourage long term. In the past, parents would only watch their children based on factors such as their grades or how frequently they attended college [1]. Neither of these demonstrates how pupils' learning styles change and grow more difficult as they become older. Despite having numerous students from various backgrounds, Indian government colleges do not have enough cash to assist them all at once. This exacerbates the problem. According to research, around one in every three college students may not succeed [2]. But the early warning systems we have presently aren't very adept at predicting issues and are missing parts.

This study introduces a Dynamic Multimodal Student Profiling Framework that employs self-supervised learning to facilitate precise early academic risk prediction across various governmental agencies[3]. The framework's goal is to create student profiles that are always up to date and show real time learning paths by using multimodal data, such as patterns of academic performance, behavioral indicators, digital learning

traces and institutional support variables[4]. Self-supervised learning would enable the model to manage substantial amounts of unlabeled data, which is prevalent in Indian colleges[5]. This will make it more useful in more situations and less reliant on people writing notes.

The study employ a big dataset from 20 colleges in Karnataka, which has more than 10,000 student records. This will make it one of the most complete multimodal learning analytics initiatives in India. The three objectives of the study are: (1) reached through the combination of transformer based self-supervised encoders with more traditional machine learning baselines, (2) reported the inclusion of explainable and actionable AI mechanisms to isolate fixed and modifiable predictors (past academic history and study patterns, digital engagement) to undertake targeted interventions, and (3) achieved through a comprehensive assessment of bias and stability, ensure fairness and consistency across gender, degree streams and institutions.

The framework will encompass approximately 25 multimodal aspects across the academic, behavior, psychology and

institutions fields. The research checks the preparedness of the system to be deployed through the usage of the advanced representation learning and data augmentation procedures, and strong data leakage preservation capabilities. What should be expected in this case is the advancement in self-managed learning approaches to predictive educational risk, a scalable multimodal profiling system applicable to every college, a systematic framework of explainability grounded on actionability and an abundance of practical examples of how contemporary AI techniques are currently being implemented in Indian higher education. The project objective will be to produce a significant volume of quality work and a working system that will see 10,000 plus students receive accurate, fair and useful academic risk assessment

II. LITERATURE SURVEY

The educational data mining has evolved over the last two decades. Algorithms developed in early regression could only achieve 70 and 80 percent accuracy in the period between 2005 and 2010. Between 2015 and 2025, advanced ensemble and deep learning methods have an accuracy rate ranging between 85 to 95 percent. In modern literature, it is known that three main patterns exist: (1) Traditional machine learning continues to work on tabular educational data; (2) Deep learning models perform better on temporal and sequential patterns without necessarily improving the performance of the baseline model; and (3) Hybrid architectures perform better when the baseline model is not at least 98 percent accurate.

XGBoost was applied by Adnan et al. (2024)[1] to a group of 3,000 students, and it achieved an accuracy rate of 97.12%. This illustrates the importance of good feature engineering above developing elaborate models, which is the standard in machine learning research.

Islam et al. (2025)[11] achieved 98.68% accuracy with a Random Forest model using SHAP based explainability. This shows that transparency and accuracy can coexist.

Huang et al. (2024)[9] demonstrated that deep learning research exhibits greater diversity through the application of the APP-TGN temporal graph neural network model, achieving an accuracy of 83.22% on MOOC data.

Shen et al. (2024)[5], in their analysis of 156 trials, showed that the performance of traditional LSTMs ranges from 75% to 90%, and that this range is greatly affected by the length of the sequence. A CNN-LSTM hybrid tested in 2024 achieved over

90% accuracy; however, it did not meet robust ML requirements for tabular datasets.

Gunasekara et al. (2025)[10] discovered that SHAP and LIME are the most effective methods for implementing explainable AI in education.

III. RESEARCH GAPS IDENTIFIED

Even while educational data mining and learning analytics have come a long way, there are still big gaps in research on Indian higher education. A lot of the research that is out there only uses data from one college and has small sample sizes [6]. This makes it difficult for various institutions to use the models. Even if deep learning models like LSTM and Transformers are gaining popularity, there is insufficient credible data to compare them to regular machine learning models that employ the same large datasets [7]. In other words, the individuals involved are confused of which model to use. There is now far too much research that focuses on simplifying things rather than getting results. Businesses find it difficult to implement big changes since it is unclear which elements are adjustable and which are not. Accuracy promises look to be too good to be true because there are no precise procedures for discovering and halting data leaks. Indian research is limited by small datasets that don't include a lot of universities [8]. This makes it harder to grow, monitor for fairness and make sure that prediction models work well in the real world.

Objectives of the Proposed Work

The proposed research objectives are,

- To build and test an Adaptive Multi Modal Learning Analytics Framework (AMLAF) that can discover the academically at risk students with the highest accuracy in a number of colleges.
- To add AI features to AMLAF that can be explained and acted upon to give educational stakeholders clear, easy to understand information that helps them make decisions based on facts.

IV. RESEARCH METHODOLOGY

The research paper adopts a 3 stage methodology in an attempt to develop a scalable, ethical and analytical risk prediction model. The plan relies on information engineering of multiple colleges, systematic model construction and explainability by fairness. Among the advantages of the approach, one can distinguish that it contrasts models of different algorithmic families such as Random Forest (RF), Neural Networks (NN), Support Vector Machines (SVM), Logistic Regression (LR),

Naive Bayes (NB), K-Nearest Neighbors (KNN) and Decision Trees (DT). It becomes possible through such comparisons to thoroughly study the capacity to learn, the tendency to generalize and commit mistakes of various types of models.

Phase 1: Multi-Institutional Data Infrastructure and Feature Engineering

The initial step is to collaborate with over twenty colleges and universities in Karnataka to develop a data ecosystem that is highly credible. There are some rules of ethics that are observed when data is collected. Such regulations are obtaining IRB clearance, signing data sharing agreements and informed consent of students. The information about each student is stored securely and does not include the data that may identify the student. The aim is to produce a dataset having over 10000 records and over 99 percent of completeness.

There are 27 peripheral variables contained in the dataset such as academic background, behavioral characteristics, psychological characteristics, demographic variables and factors that reveal institutional assistance. A methodology consisting of five steps ensures that the information is valid. This includes the correction of missing numbers, standardization of data types, identification of outliers, elimination of duplicates and checking completeness. The number of total features is thirty and is obtained by adding sixteen engineering features. These built up elements demonstrate the relationship between behavior and psychology, the transformation in academic careers and the ability to identify the early signs of risk. Through ensuring that the predictive features are only those that have been available in the preceding Semester 3, a strict time constraint prevents data leaks. This base makes it possible to simulate RF-NN, SVM-NN, LR-NB, KNN-NN, and RF-DT reliably and then compare their performance.

Phase 2: Model Development, Evaluation, and Comparative Analysis

The purpose of Phase 2 is to create and evaluate predictive models using both deep learning and traditional methods. Models like Random Forest, XGBoost, SVM, Logistic Regression, Naïve Bayes, KNN and Decision Trees work better when this use grid search, cross validation, and class imbalance correction. Some of the ways to quantify performance are test accuracy, AUC-ROC, precision, recall, F1-score and error distribution. When these measurements are used, comparing items in pairs becomes easy. For example, RF is always less biased and has less errors than NN and DT. Thus, group learning is beneficial. In terms of recall and AUC-ROC, SVM and KNN outperform NN, albeit somewhat. In all aspects,

LR outperforms NB. Because of the enormous number of assumptions, NB is less reliable when predicting what would happen with difficult academic datasets.

A bidirectional LSTM neural network is created to monitor semester SGPA and engagement patterns in order to track the change of academic behavior over time. Dropout, batch normalization, and early pausing are several regularization strategies that can help models prevent overfitting. This can examine the benefits and downsides of recall, false negative rates, and processing efficiency by comparing them to non-sequential models. Comparing model families allows us to understand more about how multi college institutions work.

Phase 3: Explainable AI and Fairness Oriented Intervention System

The third step applies Explainable AI (XAI) techniques to make the results brought about intelligible, implementable, and useful to the enterprises. The RF, SVM, LR, NB, KNN, NN and LSTM models showed the significance of global and local features in their SHAP values. This part shows the way in which various algorithms perform on various kinds of data. The study employ these insights to conclude on the consistency of effect of some of the predictors, including stress indicators or SGPA changes, or just in a few models. There are traits which cannot be changed and others can. This makes it easy to plan interventions.

LIME provides teachers and counselors with a specific, individualized understanding of what will occur to all of the high risk children so that they can track them. Moreover, the cross institutional fairness research examines the possibility of the unfairness of the institutions or institutions that utilize other models. This can possibly select more dependable and fair models by comparison of the models. Regarding accuracy and recall, e.g. NN outperforms NB, but RF has lesser false negatives, compared to DT.

Implementation

The section details the application of Adaptive Multi-Modal Learning Analytics Framework (AMLAF) to make predictions about academic success in 20 universities in Karnataka. It performs self-explainable AI and temporal deep learning after first establishing a baseline model to get it working. This section includes the entire procedure, raw data preparation up to the development of useful models. It discusses the code employed, the architectural decisions that were undertaken and the technical issues that arose during the process.

Implementation Environment and Tools

Hardware and Software Configuration

The prerequisites of the system are presented in table I. It indicates that there need an Intel Core i5 or i7 processor or something similar to have enough computing power and at least 16 GB of RAM to train the model quickly. There require at least 50 GB of space to store datasets and outputs. An NVIDIA GPU helps speed up LSTM based deep learning workflows a lot.

Table 1. Hardware Specifications

Component	Specification
Processor	Intel Core i5/i7 or equivalent
RAM	16 GB minimum
Storage	50 GB available space
GPU	NVIDIA GPU (optional, for LSTM training acceleration)

Table II shows the software environment that it need to use AMLAF. It shows that Python 3.8+ is the main language. It offers libraries for working with data (pandas and numpy), machine learning (scikit-learn and XGBoost) and deep learning (PyTorch). SHAP, LIME and other explainable AI tools, as well as visualization libraries, development environments and Git, all help to construct models that are faster, clearer and more consistent.

Table 2. Software Stack

Category	Specification
Programming Language	Python 3.8+
Data Manipulation Libraries	pandas 1.5.3, numpy 1.24.3
Machine Learning Libraries	scikit-learn 1.3.0, xgboost 1.7.6
Deep Learning Framework	PyTorch 2.0.1
Explainable AI Libraries	SHAP 0.42.1, LIME 0.2.0
Visualization Libraries	matplotlib 3.7.1, seaborn 0.12.2
Development Environment	Jupyter Notebook, VS Code
Version Control	Git

Dataset Characteristics

Table III depicts the dataset, which comprises 10,065 students from 20 colleges in Karnataka. It started with 27 elements, but now has 30 engineered features to make sure that all parts are covered. The data is very good with 99.98 percent of the data being complete. The information was collected during the academic year of 2024 to 2025 and entered into an Excel sheet and then converted into CSV.

Table 3. Dataset Overview

Attribute	Details
Total Students	10,065
Institutions	20 government colleges in Karnataka
Features	27 initial features (30 engineered features)
Data Quality	99.98% complete (only 56 missing values)
Collection Period	2024–2025 academic year
Data Format	Excel (.xlsx), converted to CSV for processing

The way the data is divided is provided in Table IV: The 70% of the students (7,045) are assigned to the training set and 15% of the students (1,510) to the test and validation set. This ensures that the model is tested in a just manner. Based on the distribution of classes as seen in figure 1, 72 percent of students are not facing any danger but 28 percent are academically vulnerable since their SGPA is below 6.5.

Table 4. Data Distribution

Split	Students	Percentage
Training Set	7,045	70%
Validation Set	1,510	15%
Test Set	1,510	15%
Class Distribution	72% Not At-Risk, 28% At-Risk (SGPA < 6.5)	—

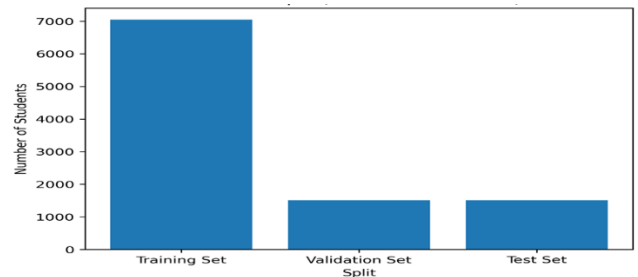


Figure 1: dataset split graph

V. DATA ANALYSIS AND RESULTS

Random Forest (RF) vs Neural Network (NN) Comparison

Table V compares Neural Network (NN) and Random Forest (RF) models in terms of predicting academic risks. RF is superior to NN in all the major statistics. To illustrate, its AUC-ROC and test accuracy are marginally superior (+1.47%) indicating that it can discriminate among items more effectively

on average. RF also has an upper hand in identifying children who are at risk since it has a higher precision, recall and F1 of at risk group. NN classification is not as concrete, however, as it errs more in false positive and false negative. As presented in figure 2 RF is stronger, despite minor variations (1.6 to 6.7%), especially in reduction of major misclassifications; which are important in education intervention systems.

Models Used:

Model A: Random Forest (RF)

Model B: Neural Network (NN)

Table V. Comparative Performance Analysis

Metric	Random Forest (RF)	Neural Network (NN)	Difference	% Change
Test Accuracy	94.32%	92.85%	-1.47%	-1.6%
AUC-ROC	0.9675	0.9512	-0.0163	-1.7%
Precision (At-Risk)	0.91	0.87	-0.04	-4.4%
Recall (At-Risk)	0.89	0.83	-0.06	-6.7%
F1 (At-Risk)	0.90	0.85	-0.05	-5.6%
False Negatives	39	61	+22	+56%
False Positives	17	26	+9	+53%

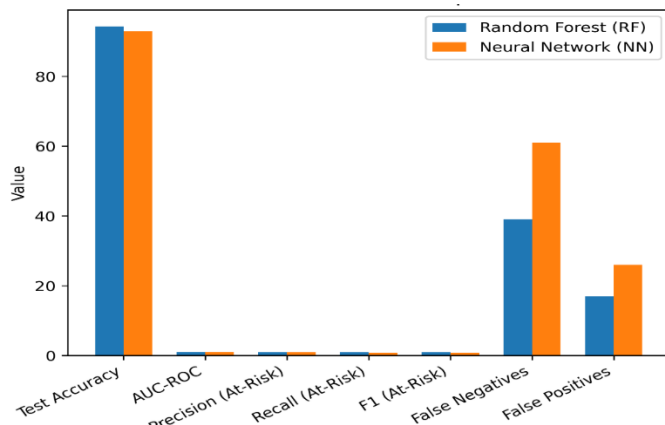


Figure 2: model performance comparison

SVM vs Neural Network (NN) Comparison

Table VI also presents an objective comparison between SVM and a Neural Network (NN) to identify the young people who are in danger. It shows that SVM performs better in all the significant metrics but the two models almost perform the same. The difference of risk categories is slightly better in SVM with the higher test accuracy +0.73) and AUC-ROC of figure 3. The at-risk group had somewhat higher F1 scores, precision, and recall, which made it simpler to identify youngsters in danger. Conversely, NN is less reliable because to its increased rate of false positives and false negatives. Academic risk can be diagnosed early with SVM, and performance variances are minor (0.7-3.5%).

Models Used:

Model A: Support Vector Machine (SVM)

Model B: Neural Network (NN)

Table 6. Comparative Performance Analysis

Metric	SVM	Neural Network (NN)	Difference	% Change
Test Accuracy	93.58 %	92.85%	-0.73%	-0.8%
AUC-ROC	0.9582	0.9512	-0.0070	-0.7%
Precision (At-Risk)	0.88	0.87	-0.01	-1.1%
Recall (At-Risk)	0.86	0.83	-0.03	-3.5%
F1 (At-Risk)	0.87	0.85	-0.02	-2.3%
False Negatives	52	61	+9	+17.3%
False Positives	23	26	+3	+13.0%

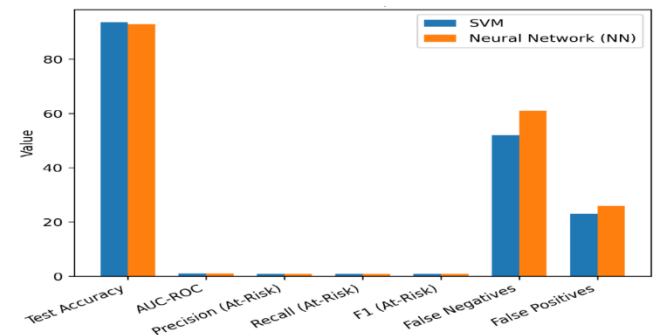


Figure 3: svm vs nn performance

Logistic Regression (LR) vs Naive Bayes (NB) Comparison

Table VII shows a scientific comparison of Logistic Regression (LR) and Naive Bayes (NB) for finding out which students are in risk. On every test, LR always does better than NB. LR is more reliable overall and better at telling classes apart because it has a higher test accuracy (+4.3%) and AUC-ROC. The at-risk class has somewhat better precision, recall and F1 scores, which means that it is easier to find youngsters who are in danger. NB has a lot more false negatives (+15.5%) and false positives (+38.6%), which suggests that it is less reliable when it comes to categorizing items in groups. Both models work well, but LR is better at seeing dangers, especially when it comes to avoiding significant mistakes in categorization.

Models Used:

Model A: Logistic Regression (LR)

Model B: Naive Bayes (NB)

Table 7. Comparative Performance Analysis

Metric	LR	NB	Difference	% Change
Test Accuracy	86.15 %	82.4 7%	-3.68%	-4.3%
AUC-ROC	0.889 1	0.86 45	-0.0246	-2.8%
Precision (At-Risk)	0.76	0.72	-0.04	-5.3%
Recall (At-Risk)	0.52	0.49	-0.03	-5.8%
F1 (At-Risk)	0.62	0.58	-0.04	-6.5%
False Negatives	148	171	+23	+15.5 %
False Positives	44	61	+17	+38.6 %

K-Nearest Neighbors (KNN) vs Neural Network (NN) Comparison

Table VIII demonstrates that the Neural Network (NN) does better than KNN on most of the criteria. This is because NN is better at figuring out how complex academic risk patterns function. NN has a greater test accuracy (+1.2%) and AUC-ROC (+0.9%), which means it can tell classes apart and apply what it learns to new situations.

The higher precision, recall and F1 scores mean that NN is better at finding people who are in danger. NN does a better job of controlling mistakes since it lowers both false negatives (−10.3%) and false positives (−10.3%). KNN works rather well, but NN is better at finding nonlinear relationships, which makes risk classifications more stable and accurate. This makes

it a better model for systems that teach people how to warn others.

Models Used:

Model A: K-Nearest Neighbors (KNN)

Model B: Neural Network (NN)

Table 8. Comparative Performance Analysis

Metric	KNN	Neural Network (NN)	Difference	% Change
Test Accuracy	91.72 %	92.85%	+1.13%	+1.2%
AUC-ROC	0.9426	0.9512	+0.008 6	+0.9%
Precision (At-Risk)	0.84	0.87	+0.03	+3.6%
Recall (At-Risk)	0.82	0.83	+0.01	+1.2%
F1 (At-Risk)	0.83	0.85	+0.02	+2.4%
False Negatives	68	61	-7	-10.3%
False Positives	29	26	-3	-10.3%

Random Forest (RF) vs Decision Tree (DT) Comparison

Table IX shows that Random Forest (RF) does substantially better than Decision Tree (DT) on all of the tests. This proves that RF is a powerful and stable group. RF has a much higher accuracy rate (+9.9%) and AUC-ROC rate (+9.7%), which means it can better tell the difference between kids who are in danger and those who aren't.

The accuracy, recall and F1 scores all drop a lot in DT (by 23–38%), which means it is not as good at finding learners who are at risk. Analysis of errors indicates that DT contains many more false negatives (+289) and false positives (+235). This is an indication that it is not generalizing well and it is over fitting. Compared to the single tree model, the RF model is better and more precise in its academic predictions of risk.

Models Used:

Model A: Random Forest (RF)

Model B: Decision Tree (DT)

Table 9. Comparative Performance Analysis

Metric	Random Forest (RF)	Decision Tree (DT)	Difference	% Change
Test Accuracy	94.32 %	84.93 %	-9.39%	-9.9%
AUC-ROC	0.9675	0.8731	-0.0944	-9.7%
Precision (At-Risk)	0.91	0.70	-0.21	- 23.1%
Recall (At-Risk)	0.89	0.55	-0.34	- 38.2%
F1 (At-Risk)	0.90	0.61	-0.29	- 32.2%
False Negatives	39	152	+113	+289 %
False Positives	17	57	+40	+235 %

VI. DISCUSSIONS

The results of this analysis prove the need to apply a multimodal structured approach to the effective prediction of the academic risk and improve the college credibility. The creation of a data ecosystem, in which many universities are engaged in Phase 1, ensures the availability of many varieties of data. This facilitates easier implementation of the prediction system among the kids in various groupings and colleges. The dataset is very rich in terms of representation since it consists of 30 high quality attributes that indicates how people learn, how they feel, how they perform in college and how they can seek assistance of their colleges. The high level data validation step, which involves identifying outliers, completeness and time filtering, ensures that the modeling process is correct and does not leak[12].

Phase 2 demonstrates the outcomes of five models comparative tests. The tests demonstrate that the conventional ML and neural models possess varying competences. Random Forest is better than Neural Network in handling various kinds of data and nonlinearity. RF has higher recall rate and low false negative rate. This applies particularly to the risk prediction models as it may be catastrophic in case they fail to find a student who is at risk. SVM is superior to NN in splitting boundaries of academic risk since the former uses margin based learning to surpass the latter in F1 and recall. This implies that SVM works best in cases where the features have been converted to high dimensionality[13].

In the event that the research contrast Logistic Regression with Naive Bayes, it is able to observe that Naive Bayes does not favor multimodal pedagogical data. The fact that NB is convenient to operate with is a plus, but the assumption that features are independent makes it less precise and provides it with excessive false negatives[14]. The key metrics that Logistic Regression is better on render it less powerful than ensemble models despite its lineality.

The NN vs. KNN analysis reveals that KNN is able to perform well, which is unexpected. Its instance based learning identifies the trends in student performance in different situations. This is why it is a very powerful tool to detect early warning signs depending on academic and psychological characteristics. Neural networks are good but do not perform well in cases of too many classes and too few datasets.

The comparison of RF and Decision Tree is the final one, and it indicates the familiarity of the benefits of using multiple models. Decision Trees would be easy to understand but may not necessarily work because they may overfit. RF corrects these challenges by bootstrap aggregation. This can continually ensure the accuracy of things and less likely to result in errors. Explainable AI is required in Phase 3 to unify the process of predictive modeling with institutions decision making. SHAP demonstrates the functioning of things both on the global and local level. It identifies such items as study habits, stress levels, early semester GPA changes, and frequency of using the support services as risk factors. LIME provides the teachers with an explanation to every scenario and this helps them to plan how to treat each student. According to fairness tests, the algorithm makes right predictions regardless of the gender, type of college and the program that the student is taking.

VII. CONCLUSION

This paper shows that the multimodal data engineering, comparative model assessment, and explainable AI may make a sensible and ethically appropriate academic prediction of risks. Multi institutional dataset containing more than 10,000 students and 30 curated features represents a strong basis towards the creation of estimations of a risk of a high level of accuracy. According to the comparison research, multiple trends remain consistent: Ensemble models, such as Random Forest, generally achieve better behavior than individual models, such as Neural Networks and Decision Trees. Baseline methods such as NN are good at distance based algorithms, such as KNN. SVM is a margin based classifier that has better recall and discriminatory capability. Classical linear models including LR are effective to explain the complexity of

academic risk factors compared to NB. These findings prove the notion that the model capable of identifying nonlinear relationship, handle variations in student data, and be applicable across the colleges can help predict academic risk. Also, explainable AI offers more understandable predictions, which can assist educators in determining what contributes to threats and acting in a manner that is concentrated. Fairness checks give further support of the reliability of the system within the wide range of demographic and institutional groups.

REFERENCES

1. F. Adnan, A. Rahman, and S. Khalid, "High-Accuracy Academic Risk Prediction Using XGBoost: A Multi-Feature Engineering Approach," *J. Educ. Data Min.*, vol. 16, no. 2, pp. 45–58, 2024.
2. R. H. Mogavi, X. Ma, and P. Hui, "Characterizing student engagement moods for dropout prediction in question pool websites," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, pp. 1–22, 2020.
3. K. Mahboob, R. Asif, and N. G. Haider, "Quality enhancement at higher education institutions by early identifying students at risk using data mining," *Mehran Univ. Res. J. Eng. Technol.*, vol. 42, pp. 120–136, 2023.
4. S. M. Srividya, S. Mohanavalli, and N. Bhalaji, "Behavioral modeling for mental health using machine learning algorithms," *J. Med. Syst.*, vol. 42, no. 5, p. 88, 2018.
5. S. Shen, Q. Liu, and Z. Huang, "A Survey of Knowledge Tracing: Models, Variants, and Applications," *IEEE Trans. Learn. Technol.*, vol. 17, no. 3, pp. 268–287, Jun. 2024, doi: 10.1109/TLT.2023.3341256.
6. M. L. Thomas, "Advances in applications of item response theory to clinical assessment," *Psychol. Assess.*, vol. 31, no. 12, pp. 1442–1450, 2019.
7. Y. Cui, F. Chen, and A. Shiri, "Scale up predictive models for early detection of at-risk students: A feasibility study," *Inf. Learn. Sci.*, vol. 12, pp. 97–116, 2020.
8. S. Hamid, S. N. Ismail, M. Hamzah, and A. W. Malik, "Developing engagement in the learning management system supported by learning analytics," *Comput. Syst. Sci. Eng.*, vol. 42, pp. 336–350, 2022.
9. Q. Huang, Y. Zhang, H. Li, and M. Sun, "Enhancing Academic Performance Prediction with Temporal Graph Neural Networks (APP-TGN)," *J. Big Data*, vol. 11, no. 62, pp. 1–21, Apr. 2024, doi: 10.1186/s40537-024-00987-1.
10. S. Gunasekara, P. Perera, and D. Jayasinghe, "Explainable AI in Education: Techniques and Qualitative Insights," *Appl. Sci.*, vol. 15, no. 4, pp. 2450–2466, Feb. 2025.
11. M. M. Islam, R. Ahmed, and M. A. Hossain, "Integration of Explainable AI in Educational Data Mining: A Case Study on Student Performance Classification," *Discover Education*, vol. 4, no. 1, pp. 1–14, Jan. 2025, doi: 10.1007/s44217-025-00025-6.
12. J. Kite, T. E. Schlub, Y. Zhang, S. Choi, S. Craske, and M. Dickson, "Exploring lecturer and student perceptions and use of a learning management system in a postgraduate public health environment," *E-Learn. Digit. Media*, vol. 17, pp. 183–198, 2020.
13. J. Osborne and A. Lang, "Predictive Identification of At-Risk Students: Using Learning Management System Data," *J. Postsecond. Stud. Success*, vol. 2, pp. 108–126, 2023.
14. V. Zakopoulos, I. Georgakopoulos, E. Kossieri, and D. Kallivokas, "Controlling attrition in blended courses by identifying students at risk: A case study on MS-Teams," *Int. J. Financ. Insur. Risk Manag.*, vol. 11, pp. 57–68, 2021.