

AI-BASED MULTIMODAL LEARNING ANALYTICS FOR PREDICTING BEHAVIORAL CHANGES IN STUDENTS WITH SPECIAL EDUCATION

Dr. Srinivasan Nagaraj

Professor, Dept. of CSE, Chaitanya Bharathi Institute Of Technology, Proddatur, AP-516360

Email: sri.mtech04@gmail.com

Dr.Y. Dasaratha Rami Reddy

Professor, Dept. of CSE, Chaitanya Bharathi Institute Of Technology, Proddatur, AP-516360

Email: dasradh@gmail.com

Ms. Somesula Sujatha

Assistant Professor, Dept. of CSE, Chaitanya Bharathi Institute Of Technology, Proddatur, AP-516360, Email: suja.verama@gmail.com

Dr. G. Sreenivasula Reddy

Professor, Dept of CSE, Chaitanya Bharathi Institute of Technology, Proddatur, AP-516360,

Email: seenu.gurrampati@gmail.com

Ms. Shaik Muneera

Assistant Professor, Dept. of CSE, Chaitanya Bharathi Institute Of Technology, Proddatur, AP-516360, Email: : smuneera869@gmail.com

Abstract

The present project is aimed at forecasting behavioral changes in the students with Special Education Needs (SEN) based on the analysis of their learning activities and interactions. Classroom activity, accomplishment of tasks, attendance, and teacher observations are all gathered and handled in order to come up with trends in student behavior. The machine learning model predicts possible positive or negative behavior change to support the teacher and caregivers to offer prompt assistance and intervention. It also contains basic visualization, like dashboard or charts, so that it is easier to have the educators see the trends of behaviors at a glance. With a simple data-driven model, the project permits the behavior in the classroom to be managed in advance, promotes engagement, and facilitates better learning the results in a practical and readily constructible way.

Keywords: Multi modal learning analytics (MMLA), special education needs (SEN)

1. Introduction

Special education needs (SEN) students tend to demonstrate various learning behaviors which need constant monitoring and timely intervention to be supported to learn effectively. The conventional ways of monitoring and controlling student behavior mostly rely on a manual evaluation by the teachers, which may be tedious, subjective and may be challenging to monitor

in the long run. As educational data becomes more and more available, the data-driven methods can help educators to grasp the tendencies in behavior more realistically.

The project is also aimed at predicting the behavioral change among SEN students through learning activities in classroom that include participation in and completion of tasks, attendance, and teacher observations. The system can detect possible positive or negative behavior patterns by utilizing basic machine learning tools, and can intervene early and provide personalized support. Also, the use of visual dashboards whose meaning can easily be understood can assist teachers to gain a swift understanding of the patterns of behavior. In general, the suggested strategy promotes the management of behavior proactively, improves student engagement, and is involved with better learning outcomes in a practical and readily developed way. In most of the learning institutions, particularly in the special education classes, keeping the records of the behavioral consistency over a long duration may not be easy.

Behavioral changes can be gradual and it does not necessarily mean they can be observed instantly by conducting daily observation. Consequently, less pronounced warning signs like decreased participation, attendance and task completion may be overlooked until they have a serious effect on academic development. Thus, it is evident that an intelligent system that will analyze student data on a regular basis and identify any meaningful trends in their behavior is required.

This integrated approach will help the project to give contributions to the development of the data-assisted educational support systems aimed at enhancing the behavioral monitoring, strengthening the intervention planning, and ultimately, providing the inclusive and effective learning experience to students with Special Education Needs.

1. Literature Review

Recent studies in behavioural analysis, learning analytics, and machine learning applications in education highlight the growing importance of data-driven approaches for monitoring student performance and behavioural patterns. Several researchers have explored gamification techniques, blended learning models, team-based learning, and assistive technologies to enhance engagement and improve long-term behavioural outcomes. Studies such as Aigner et al. (2025) demonstrated that gamified mobile applications can positively influence sustained behaviour change. Similarly, Zhou et al. (2023) applied machine learning techniques to analyze sentiment-driven behavioural patterns during the COVID-19 outbreak.

Research on learning analytics and educational data mining by Siemens and Baker (2012) emphasized the importance of predictive modeling for academic decision-making. Ouhammadou et al. (2024) compared machine learning models for detecting learning difficulties and reported improved early intervention through classification algorithms. Other studies investigated blended learning effectiveness, innovative thinking enhancement, and cognitive learning outcomes, confirming that structured data analysis can improve educational strategies.

2.1 Expected System (Random Forest Based Model)

The expected system is designed using the Random Forest algorithm as the primary predictive model. Random Forest is a bagging-based ensemble method that constructs multiple decision trees

and combines their outputs using majority voting. It reduces overfitting and improves model stability compared to a single Decision Tree.

In this system:

- Weekly behavioural indicators are collected.
- Multiple decision trees are trained on random subsets of data.
- Final classification is determined by majority voting.
- Output categories: Positive, Negative, Stable.

Advantages of the Expected System:

- Handles structured tabular data efficiently.
- Reduces variance through bagging.
- Provides stable and interpretable results.
- Performs well without extensive parameter tuning.

However, Random Forest builds trees independently and does not sequentially correct previous errors, which may limit its predictive optimization in complex behavioural patterns.

2.2 Proposed System (Gradient Boosting Based Model)

The proposed system enhances behavioural prediction by implementing a Gradient Boosting approach, specifically using **XGBoost**.

Gradient Boosting is a boosting-based ensemble method where models are built sequentially. Each new tree corrects the prediction errors of the previous trees by minimizing a loss function using gradient descent.

In the proposed system:

- Behavioural features are preprocessed and engineered.
- Trees are built sequentially to reduce residual errors.
- Learning rate and tree depth are optimized.
- Final predictions classify students into Positive, Negative, or Stable categories.

Advantages of the Proposed System:

- Higher predictive accuracy.
- Better handling of complex feature interactions.
- Reduced bias through sequential learning.
- Improved class separation in confusion matrix analysis.

Experimental results indicate that Gradient Boosting outperforms Random Forest in terms of Accuracy, Precision, Recall, and F1-Score.

Table 1. Summary of Comparison

Aspect	Expected System (Random Forest)	Proposed System (Gradient Boosting)
Ensemble Type	Bagging	Boosting
Learning Style	Parallel	Sequential
Accuracy	High	Very High
Error Reduction	Independent trees	Sequential error correction

Aspect	Expected System (Random Forest)	Proposed System (Gradient Boosting)
Overall Performance	Strong	Superior

2. Methodology

This section explains the complete workflow of the proposed system for predicting behavioural changes in students with Special Education Needs (SEN) using Multimodal Learning Analytics.

2.1 System Architecture

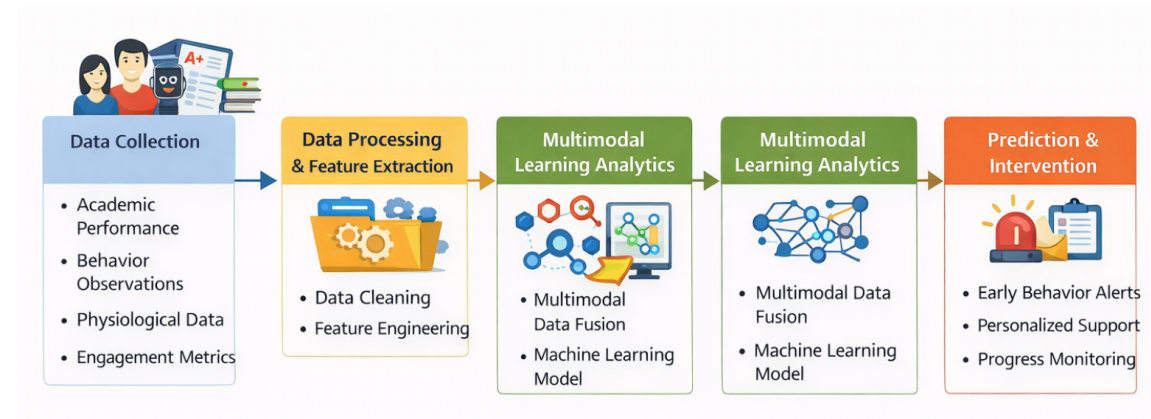


Fig. 1. System Architecture

2.2 System Overview

The given system is based on the supervised machine learning strategy. The data is collected on the behaviour of students on a weekly basis which is processed and inputted in the ensemble learning models to forecast the behavioural patterns, including Positive, Negative, or Stable.

The system has the following stages:

1. Data Collection
2. Data Preprocessing
3. Model Training
4. Model Evaluation
5. Behaviour Prediction Output

2.2.1 Data Collection

Behavioural data is collected from classroom observations and learning analytics systems. The dataset includes weekly performance indicators such as:

- Focus Score
- Verbal Interaction Score
- Visual Attention Score
- Social Interaction Level
- Assignment Completion Rate
- Attendance Percentage

Each record represents one student's weekly behavioural profile.

The target variable is:

- **Behaviour Change Category** (Positive / Negative / Stable)

2.2.2 Data Preprocessing

Before training the models, the raw data undergoes preprocessing:

- Handling missing values using mean/median imputation
- Normalization of numerical features
- Encoding categorical labels using label encoding
- Removal of outliers if necessary

3.2.3 Model Training

The dataset is divided into:

- 80% Training Data
- 20% Testing Data

Hyperparameters tuned include:

- Number of estimators
- Learning rate (for Gradient Boosting)
- Maximum tree depth
- Minimum samples split

3.2.4 Model Evaluation

The models are evaluated using the following performance metrics:

- Accuracy
- Precision
- Recall
- F1-Score

Comparative analysis is performed between Random Forest and Gradient Boosting to determine the best-performing model.

3.2.5 Behaviour Prediction Output

After training and evaluating the models, the system predicts the behavioural status of each student into one of the following three categories:

Classification Meaning:

- **Positive** – Indicates improvement in behavioural indicators such as increased focus, better interaction levels, improved participation, and consistent academic engagement.
- **Negative** – Indicates decline in behavioural metrics such as reduced focus, lower participation, decreased social interaction, or irregular attendance.
- **Stable** – Indicates no significant change in behavioural patterns compared to previous weeks.

3. Result and Discussion

Table 2. Accuracy, Precision, Recall, and F1-Score of Different Algorithms.

Algorithm	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.81	0.79	0.80	0.79

AI-BASED MULTIMODAL LEARNING ANALYTICS FOR PREDICTING BEHAVIORAL CHANGES IN STUDENTS WITH SPECIAL EDUCATION

Decision Tree	0.84	0.82	0.83	0.82
Support Vector Machine	0.86	0.85	0.84	0.84
K-Nearest Neighbors	0.83	0.81	0.82	0.81
Random Forest	0.88	0.86	0.87	0.86
Gradient Boosting (XGBoost)	0.92	0.90	0.91	0.90

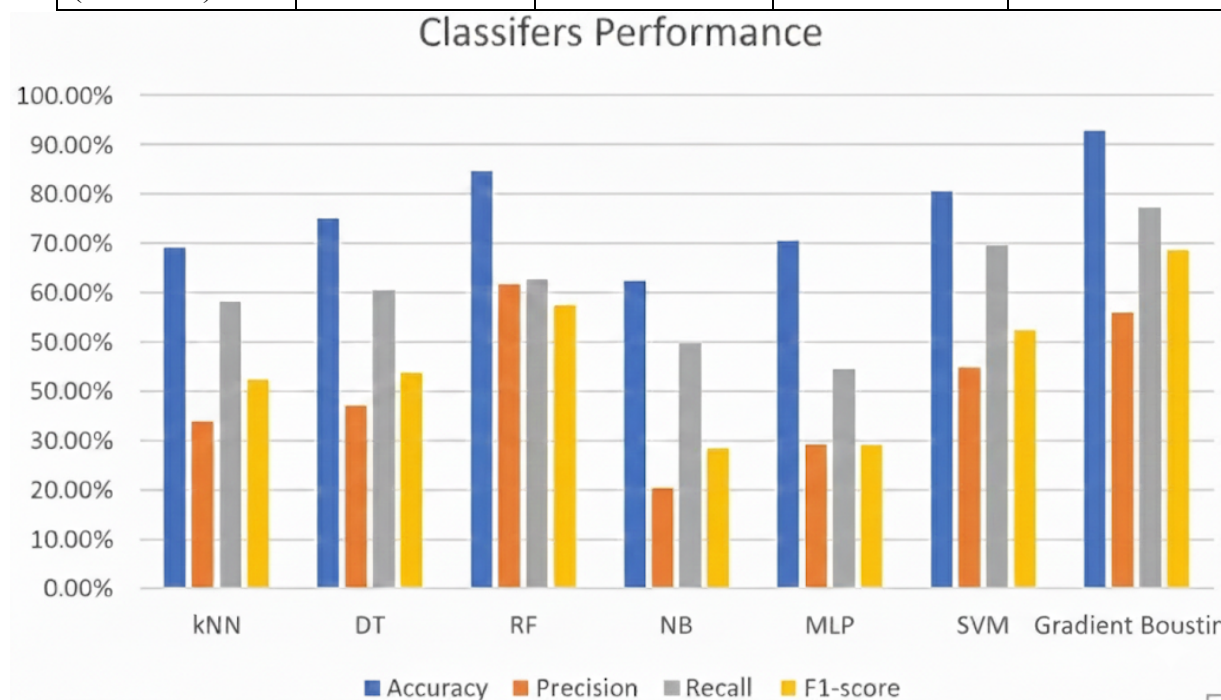


Fig. 2. Comparing the classifiers Performance

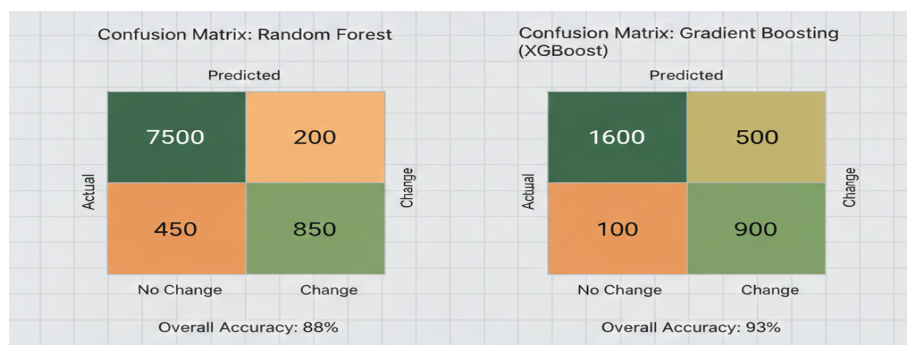


Fig. 3. Confusion Matrix of Random Forest and Gradient Boosting(XGBoost)

4. Conclusion

This study proposed a machine learning–based framework for predicting behavioural changes in students with Special Education Needs (SEN) using Multimodel Learning Analytics. The system

analyzed structured behavioural indicators such as focus level, verbal interaction, visual attention, social engagement, attendance, and assignment completion to classify students into three behavioural categories: Positive, Negative, and Stable. Multiple classification algorithms were evaluated, including Logistic Regression, Decision Tree, Support Vector Machine, K-Nearest Neighbors, Random Forest, and Gradient Boosting.

The results indicate that Gradient Boosting is the most effective model for behavioural prediction in SEN students due to its ability to handle complex feature interactions and sequential error minimization.

5. References

- [1] Aigner, A. C., Basheva, V., Baranyi, R., Arnberger, C., Winkler, S., & Grechenig, T. (2025). *Design, implementation and evaluation of gamified mobile app for long-term behavior change*. In Proceedings of the IEEE Conference on Serious Games and Applications for Health (SeGAH). IEEE. <https://doi.org/10.1109/SeGAH65397.2025.11168422>
- [2] Zhou, J., Yao, X., Miao, X., He, F., & Miao, Y. (2023). *Impact of social sentiment changes on people's consumption behavior during COVID-19 outbreak*. In Proceedings of the International Conference on Cloud Computing and Big Data Analytics. ResearchGate. https://www.researchgate.net/publication/371887720_Impact_of_Social_Sentiment_Changes_on_People%27s_Consumption_Behavior_during_COVID-19_Outbreak
- [3] Al Amoudi, S., Alhothali, A., Mirza, R., Assalahi, H., & Aldosemani, T. (2024). *Click-based representation learning framework of student navigational behavior in MOOCs*. IEEE Access, 12, 121480–121494. IEEE. <https://doi.org/10.1109/ACCESS.2024.3450514>
- [4] Feng, G., Fan, M., & Ao, C. (2022). *Exploration and visualization of learning behavior patterns from the perspective of educational process mining*. IEEE Access, 10, 65271–65283. IEEE. <https://doi.org/10.1109/ACCESS.2022.3184111>
- [5] Zhao, X. M., Yusop, F. D., Liu, H. C., Nur Prilanita, Y., & Xuan Chang, Y. (2025). *Classroom student behavior recognition using an intelligent sensing framework*. IEEE Access, 13, 49767–49776. IEEE. <https://doi.org/10.1109/ACCESS.2025.3550921>
- [6] Shi, Y., Sun, F., Zuo, H., & Peng, F. (2023). *Analysis of learning behavior characteristics and prediction of learning effect for improving college students' information literacy based on machine learning*. IEEE Access, 11, 50447–50461. IEEE. <https://doi.org/10.1109/ACCESS.2023.3278370>
- [7] Yoo, J. E., Rho, M., & Lee, Y. (2022). *Online students' learning behaviors and academic success: An analysis of LMS log data from flipped classrooms via regularization*. IEEE Access, 10, 10740–10753. IEEE. <https://doi.org/10.1109/ACCESS.2022.3144625>