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

Tri-Attentive BiLSTM Framework for Integrated Student Performance Evaluation, Stress Detection, and Score Prediction

P. Nagendra Kumar[1], K. Chiranjeevi[2], B. Poojitha[2], K. Balakrishna[2]

[1]Professor, [2]Assistant Professor, [1,2]Department of Computer Science and Engineering (AI & ML)
[1,2]Geethanjali Institute of Science and Technology, Nellore-Bombay Highway, S.P.S.R, Andhra Pradesh 524137, India

Abstract

Today's students face a heavy academic workload, long study hours, and continuous pressure to perform well, which affects both their mental and physical health. Increasing stress, fear of failure, and lack of support often push vulnerable students toward severe outcomes, including depression and suicide. Using the provided student dataset containing features such as study hours, sleep duration, attendance, stress level, past scores, and extracurricular activity, we perform preprocessing and Exploratory Data Analysis (EDA) to understand behavior patterns and data quality. The existing system applies classical machine learning models like Random Forest, Support Vector Machine (SVM), Gradient Boosting, Decision Tree, and Linear Regression, but these handle each output separately and fail to learn deeper relationships between student factors. To address this, we propose an integrated Attention-based Bidirectional Long-Short Term Memory (BiLSTM) model combined with Optimal Decision Rule List (ODRL) feature extraction for joint prediction of student performance grade (classification), stress level (classification), and exam score (regression). The Attention-BiLSTM captures important behavioral sequences, while ODRL extracts meaningful features. Results show that the proposed approach outperforms the existing models, providing a more accurate and holistic understanding of student academic and emotional conditions.

Keywords—Student Performance Prediction, Stress Level Classification, Attention-Based BiLSTM, Optimal Decision Rule List (ODRL), and Multi-Task Learning.

1. INTRODUCTION

In India, students pursue education across a wide range of sectors, including schools, engineering colleges, medical institutions, degree programs, and various vocational and professional courses, each carrying its own academic structure, pressure levels, and expectations. Because of these differences, students experience stress in many forms—academic pressure, competitive environments, long study hours, fear of failure, parental expectations, and workload imbalance. Such stress not only affects their emotional well-being but also has a direct impact on their academic performance, causing fluctuations in focus, motivation, learning capacity, and, ultimately, in their exam scores. High stress often leads to reduced performance, while manageable stress levels allow students to maintain consistency and achieve better results. In this context, it becomes important to examine student performance grading, stress levels, and score outcomes together, as this combined understanding helps reveal how stress influences academic progress across various fields of study.

A holistic evaluation approach supports educational institutions in identifying struggling students early, designing effective support systems, reducing academic pressure, and promoting healthier learning environments. This is especially crucial in India's diverse and competitive education system, where students from different backgrounds face unique challenges, and addressing stress-related performance issues can significantly improve overall academic success and well-being.

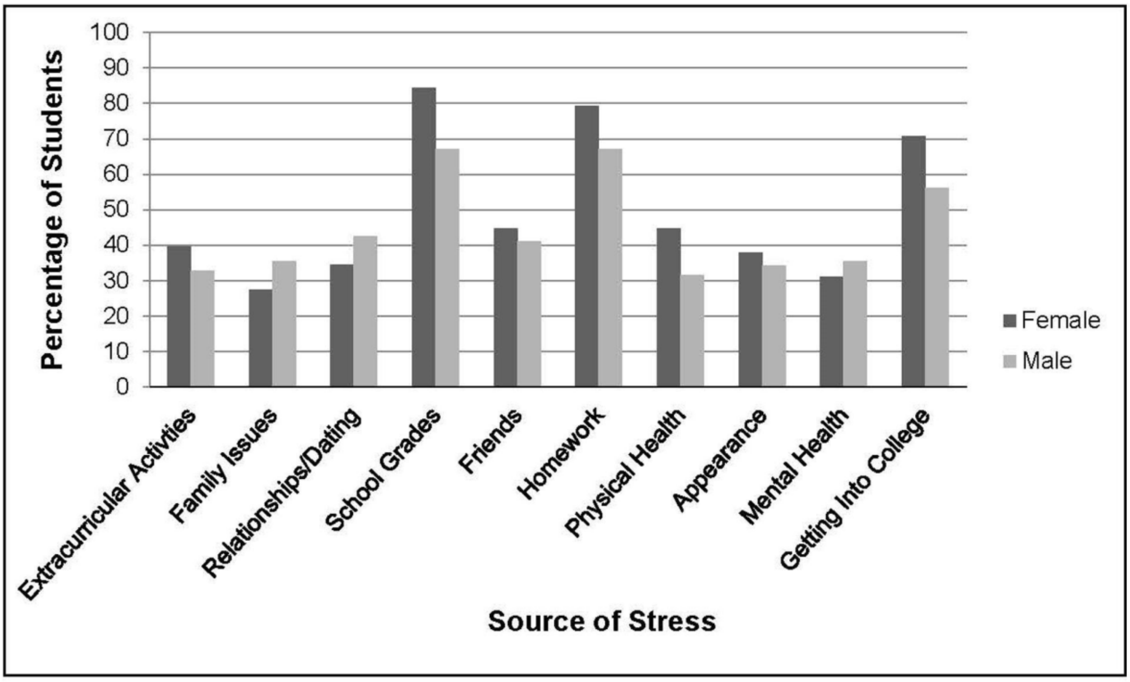


Fig. 1: Stress level Analysis

In Fig 1, the bar graph presents a comparative analysis of stress levels among male and female students across various sources of stress. The results indicate that 40% of females and 33% of males experience stress from extracurricular activities, while 28% of females and 25% of males are impacted by family issues. Stress related to relationships or dating affects 35% of females and 30% of males. Academic pressure is shown to be a major concern, with 84% of females and 67% of males reporting high stress due to school grades. Additionally, stress from friends affects 45% of females and 40% of males, demonstrating a moderate influence on student well-being.

Further highlights that academic workload also plays a significant role, as 78% of females and 66% of males feel stressed due to homework. Concerns related to physical health affect 45% of females and 30% of males, whereas appearance-related stress impacts 38% of females and 34% of males. Mental health concerns are reported by 30% of females and 35% of males, revealing a nearly balanced distribution. Lastly, stress associated with getting into college is notably high, with 70% of females and 55% of males experiencing pressure related to future academic admissions.

2. LITERATURE SURVEY

Balcioğlu, et al. [1] The Ensemble Model (EM) consistently outperformed other models, achieving the highest overall F1 measure, precision, recall, and accuracy. These results highlight the potential of data-driven techniques in informing educational stakeholders' decision-making processes, enabling targeted interventions, and facilitating personalized learning strategies tailored to students' needs. By identifying at-risk students early in the academic year, institutions can provide additional support to improve academic outcomes and retention rates.

Alnasyan, et al. [2] utilized KANFormer model is developed, a novel deep learning (DL) architecture that integrates multi-head self attention (MHSA) mechanisms with Kolmogorov–Arnold networks (KAN) to effectively captures complex patterns within large-scale student datasets, enabling more accurate predictions of student performance. KANFormer processes student demographic, academic, and engagement features from the open university learning analytics dataset (OULAD) while applying a feature elimination process to enhance dataset quality, reduce redundancy, and mitigate multicollinearity, ensuring optimal predictive performance.

Kesgin , et al. [3] proposed A comprehensive fairness analysis is conducted, considering sensitive attributes such as gender, school type, and socioeconomic factors, including parental education (Medu and Fedu), cohabitation status (Pstatus), and family size (famsize). Using the AIF360 library, we compute the demographic parity difference (DP) and Equalized Odds Difference (EO) to assess model biases across diverse subgroups. Our results demonstrate that XGBoost achieves high predictive performance (accuracy: 0.789; F1 score: 0.803) while maintaining low bias for socioeconomic attributes, offering a balanced approach to fairness and performance.
Ab Rahman, et al. [4] developed a systematic literature review on machine learning technology for predicting student performance, analysing 51 relevant articles from Scopus and Science Direct databases between 2019 and 2023 using the PRISMA method. The findings reveal that the primary motivation for employing machine learning in educational institutions is to improve predictive accuracy, identify early interventions, and optimise decision-making processes. Supervised machine learning approaches such as Decision Trees, Linear Models, and Neural Networks are commonly used.
Alalawi, et al. [5] utilized Learning Analytics Intervention (LAI) studies have emerged as an approach that aims to address this gap. In LAI studies, student risk levels are predicted and disseminated to relevant stakeholders (academics, administrators and students) using learning analytics (LA) tools for targeted interventions. The interventions themselves are mainly left under the discretion of the academics and/or administrators, who are aware of the learning context and have the authority to make decisions, with LA tools facilitating this process. LAI studies have shown success in improving outcomes (e.g. improve pass rates, retention, grades), but their uptake has been slow.
Saqr, et al. [6] proposed advanced predictive methods, namely machine learning, where the goal is to predict continuous variables like grades. The chapter uses advanced and popular AI/machine learning algorithms like Random Forest, K-Nearest Neighbor, Linear Regression, Neural Networks, and Support Vector Machines. The chapter provides a practical guide to building and evaluating predictive models with R using two approaches: one is the classic approach for predictive modeling with R, and the other more modern approach using the tidymodels suite.
Ahmed, et al. [7] proposed challenges in performance analysis, quality education delivery, and student evaluation through machine learning (ML) models. Ten regression models including K-Nearest Neighbors Regressor, Linear Regression, CatBoost, XGBoost, AdaBoost, and ensemble voting regression (VR) algorithm based on the top five heterogeneous regressors as base models are employed to predict academic outcomes. Two datasets with distinct feature sets and sizes were used to evaluate the generalizability of the models. The first dataset comprises 10,000 samples and six features focused on study behaviors, prior performance, and extracurricular activities.
Doma, et al. [8] developed a highly efficient and accurate hybrid machine learning model to classify student stress levels, by taking into account the psychological, physiological, environmental, academic and social stress factors. Student stress level classification will enable early intervention, tailored support, evaluation of stress management programs, resource allocation and improvement of academic success. This also provides us with a deeper understanding of student stressors and how they are related. To build our hybrid model, we first implement five base models namely, Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR) and Extreme Gradient Boosting (XGBoost).
Asto Buditjahjanto, et al. [9] developed training and testing of SVM to be able to classify student learning outcomes correctly. The study results indicated that 36 stress variables could be grouped into 7 stress factors using factor analysis, namely conflict, learning readiness, participation, workload, fairness, focusing, and suitability. Four levels were used to categorize the learning results of students, such as very good, good, fair, and poor. The accuracy and precision of the SVM classification values

were 0.89 and 0.70, respectively. The SVM classification results showed better results in comparison to comparable methods of classification like Decision Trees, Random Forests, K-Nearest Neighbor (K-NN), Linear Discriminant Analysis (LDA), and Gaussian Naive Bayes (GNB).
Gedam, et al. [10] proposed model was trained and tested on a dataset of 200 participants, which involves applying four different stressors. Nine ML algorithms were investigated for both multivariate and univariate features. The physiological data was collected using a novel device developed using an Arduino microcontroller and low-cost sensors such as ECG, GSR, and ST sensors. The findings reveal that the suggested model detects mental stress with an accuracy of 96.17%, with the XGBoost method outperforming other algorithms in multivariate analysis. Univariate feature analysis found that XGBoost regularly demonstrated good accuracy, showing its dependability for detecting mental stress.
Deena, et al. [11] proposed models that can precisely classify and predict stress levels using a well selected dataset intended to capture elements contributing to student stress. Given the global increase in student stress levels, it is imperative that we deal with this issue to avert serious consequences. This work proposes a machine learning algorithm that measures students' stress levels in classroom environments by leveraging recent advances in computer science, especially in the field of healthcare. The examination examines personal attributes and significant stressors, encompassing academic demands, psychological states, and social engagements.
Ali Hassan, et al. [12] proposed machine learning with an empathetic understanding of the human experience, our approach strives to pave the way for a more holistic and personalized educational ecosystem. The insights derived from this study aspire to guide educators, administrators, and policy makers in crafting nurturing environments that empower students to excel both academically and emotionally. Embarking on a journey to unravel the myriad factors influencing students' stress, we consider academic pressures, social dynamics, and personal experiences. Harnessing machine learning algorithms such as decision trees, neural networks, we navigate expansive student populations.
Biswas, et al. [13] utilized a Fine-tuned Squirrel Search-driven Light Gradient Boosting Machine (FSS-LGBM) model, which integrates advanced optimization techniques with Light GBM for enhanced predictive performance. The FSS-LGBM model validated superior performance associated with convolutional machine learning models, achieving higher accuracy (98%), recall (99%), precision (97%), and F1-score (97%) in student stress prediction. The advanced machine learning model developed in this study, incorporating FSS-LGBM, offers a promising approach for predicting student stress.
Suraj Arya, et al. [14] proposes various machine learning and deep learning models, including support vector machine (SVM), Random Forest, Gradient Boosting, AdaBoost, CatBoost, LightGBM, ExtraTree, XGBoost, logistic regression, K-nearest neighbor (KNN), Naive Bayes, decision tree, multi-layer perceptron (MLP), and artificial neural network (ANN). The Naive Bayes model achieved an accuracy of 90%, while SVM had the lowest test accuracy at 85.45%.
Arias, et al. [15] developed PRISMA framework to search through various databases from 2019 to 2024, such as ACM, IEEE Explore, and others, sifting through 249 articles. We chose 38 articles to look at closely, while we excluded 179 and identified 32 duplicates. Our review found that supervised models, like Random Forest (RF), K-Nearest Neighbor (KNN), and support vector machine (SVM), are the most common and effective, detecting stress with about 95% accuracy. Physiological datasets, such as WESAD, and real-time data collection are the most common sources for measuring stress.

3. PROPOSED SYSTEM

The proposed system develops a trivariate intelligent prediction framework that simultaneously performs student performance grading, stress level classification, and score estimation by leveraging both traditional machine learning and deep learning models. As shown as fig 2 The system begins with

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

structured academic, behavioral, and demographic data and applies systematic preprocessing and exploratory analysis to understand learning patterns. Multiple CART-based baseline models are implemented to establish comparative performance, followed by an advanced Attention Recurrent Trees CART architecture that integrates Attention-based BiLSTM feature extraction with decision tree learning. To further enhance prediction accuracy and ranking stability, a Student CART model combining MLP and CatBoost is proposed. The final system is deployed using Flask to provide real-time classification and regression outputs, enabling effective academic performance assessment and decision support.

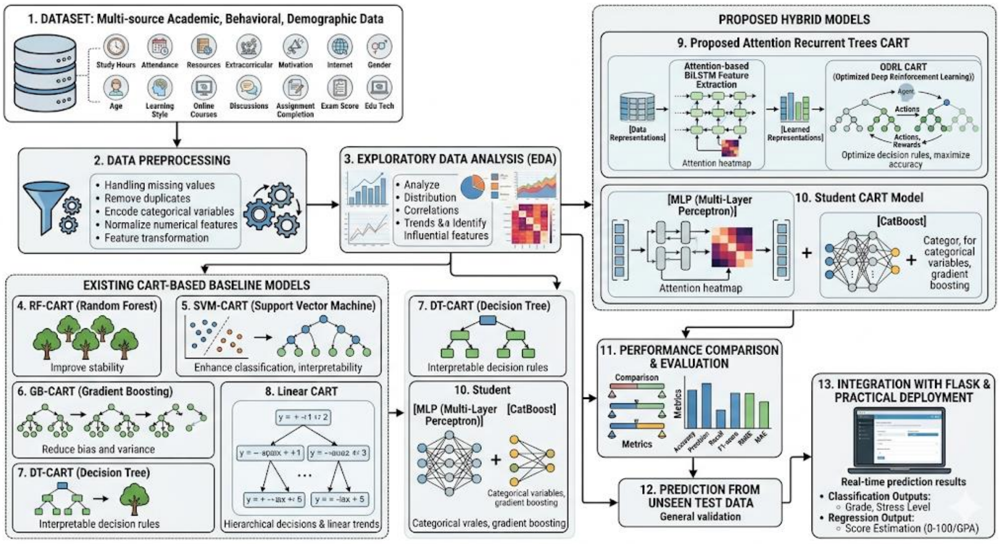


Fig. 2: Proposed system architecture.

Step 1: Dataset: The dataset consists of academic, behavioral, technological, and demographic attributes including Study Hours, Attendance, Resources, Extracurricular activities, Motivation, Internet access, Gender, Age, Learning Style, Online Courses, Discussions, Assignment Completion, Exam Score, and Edu Tech usage. These features are used to predict Stress Level, Final Grade, and continuous score values. The dataset supports both classification and regression tasks, making it suitable for trivariate learning objectives.

Step 2: Data Preprocessing: In this step, raw data is cleaned by handling missing values, removing duplicates, and correcting inconsistent entries. Categorical variables such as Gender, LearningStyle, and StressLevel are encoded using suitable encoding techniques, while numerical features are normalized or standardized. Feature selection and transformation are performed to improve model efficiency and reduce noise.

Step 3: Exploratory Data Analysis (EDA): EDA is conducted to analyse data distribution, correlations, and trends among academic and behavioural variables. Statistical summaries and visual analysis are used to identify influential features affecting student performance and stress levels. This step helps in understanding data imbalance, feature relationships, and potential outliers.

Step 4: Existing Random Forest-CART (RF-CART): The RF-CART model uses an ensemble of decision trees to capture non-linear relationships within the dataset. It improves prediction stability by aggregating multiple CART trees trained on random feature subsets. This model serves as a strong baseline for both classification and regression tasks.

Step 5: Existing Support Vector Machine-CART (SVM-CART): SVM-CART integrates margin-based learning with CART-based decision boundaries. It enhances classification performance by separating complex data patterns while maintaining interpretability through tree-based decisions.

Step 6: Existing Gradient Boosting-CART (GB-CART): GB-CART builds decision trees sequentially, where each tree corrects the errors of the previous one. This model effectively reduces bias and variance, improving prediction accuracy for student grades and score estimation.

Step 7: Existing Decision Tree-CART (DT-CART): DT-CART constructs a single decision tree using recursive binary splitting. It provides an interpretable baseline model that highlights key decision rules influencing student performance and stress classification.

Step 8: Existing Linear CART: Linear CART combines linear regression models at the leaf nodes of a CART structure. This approach captures both linear trends and hierarchical decision rules, improving regression accuracy for score estimation.

Step 9: Proposed Attention Recurrent Trees CART (Combination of Attention BiLSTM Feature Extraction with ODRL CART): This step combines Attention-based BiLSTM feature extraction with ODRL CART to form an Attention Recurrent Trees CART model. The BiLSTM captures bidirectional dependencies and uses attention to highlight critical features influencing performance, stress, and scores. These learned representations are fed into ODRL CART to optimize decision rules and improve classification and regression accuracy

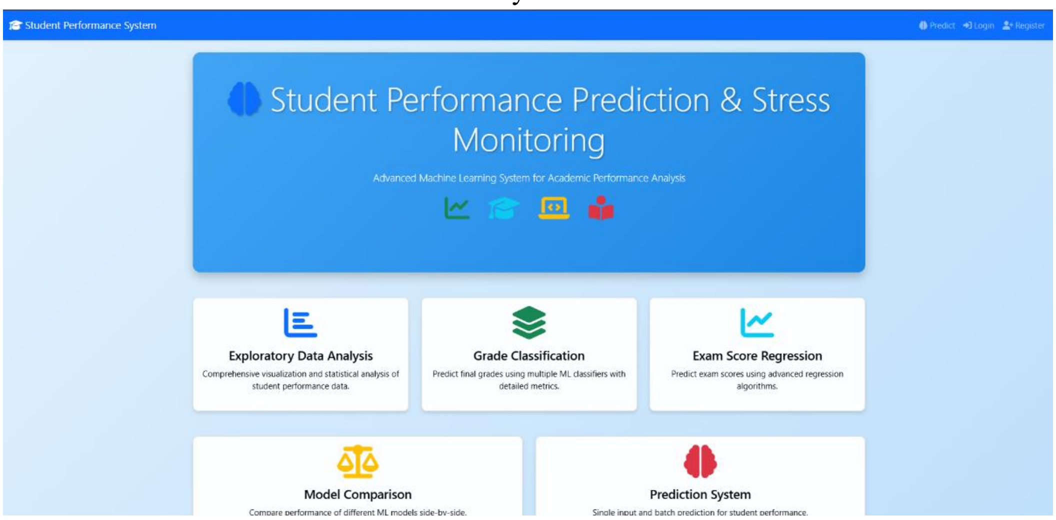
Step 10: Performance Comparison: All implemented models are evaluated and compared using metrics such as accuracy, precision, recall, F1-score, RMSE, and MAE. Comparative analysis demonstrates the effectiveness of the proposed model over traditional CART-based approaches.

Step 11: Prediction From Test Data: The best-performing model is applied to unseen test data to generate predictions for student grades, stress levels, and estimated scores. This step validates the generalization capability of the proposed framework.

Step 12: Integration with Flask: The finalized model is integrated into a Flask-based web application. The system provides real-time classification outputs such as performance grade and stress level categories, along with regression-based score estimation in the range of 0–100 or GPA format, enabling practical deployment and user interaction.

4. RESULTS ANALYSIS

Fig. 3 illustrates the home page interface of the student performance and stress monitoring system, which integrates multiple analytical modules into a single platform. It depicts how functionalities such as EDA, classification, regression, and model comparison are organized for user interaction. The figure highlights that the system supports both classification outputs such as FinalGrade (0–3) and StressLevel (0–2), along with regression output ExamScore (0–100). It represents the ability of the system to handle multiple prediction tasks simultaneously. The interface reflects how users can initiate analysis workflows and access different modules efficiently.



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

Fig. 3: Home Page Interface for Student Performance & Stress Monitoring

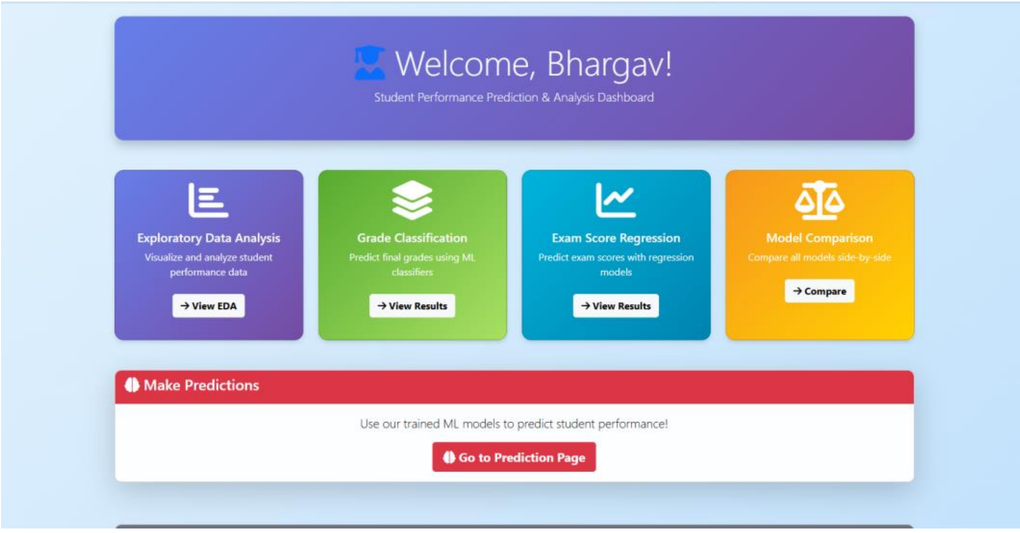


Fig. 4: Student Dashboard Interface

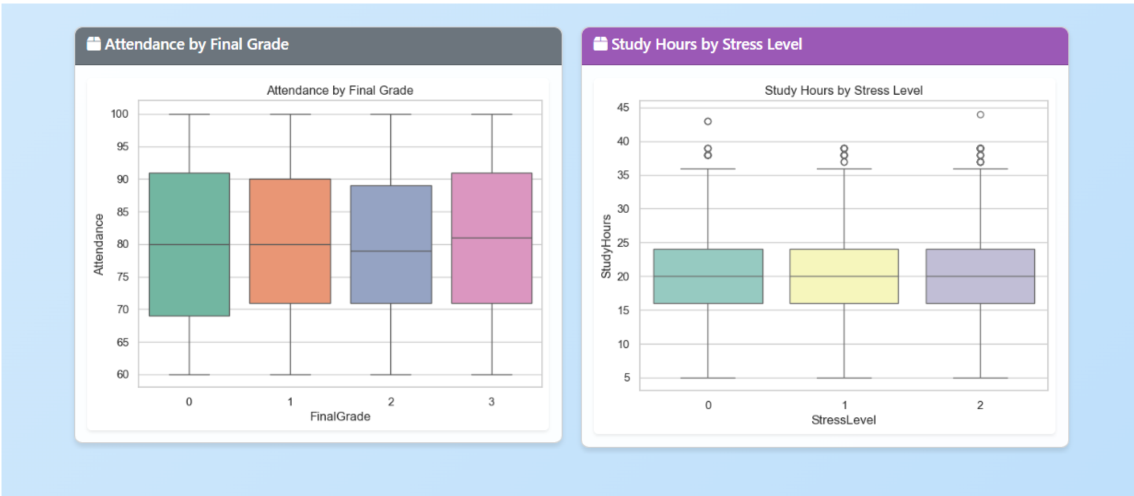
Fig 4 depicts the student dashboard interface, which provides access to core analytical modules of the system. It illustrates how users can navigate through EDA, classification, regression, and model comparison functionalities. The figure highlights that classification outputs include FinalGrade (0–3) and StressLevel (0–2), while regression predicts ExamScore values typically ranging from 40 to 100. It represents the integration of multiple models such as RF, SVM, GB, DT, and BARM within a single interface. The dashboard enables real-time interaction with prediction modules and result visualization. It reflects the structured workflow for executing different analytical operations.

Fig. 5 (a) illustrates the EDA results focusing on distributions, relationships, and correlations among features. It depicts the distribution of FinalGrade across values 0 to 3 and StressLevel across values 0 to 2, showing variations in student outcomes. The figure highlights the distribution of ExamScore values, which typically range between 40 and 100. It represents correlation patterns among features such as StudyHours, Attendance, and AssignmentCompletion. The analysis shows how different variables contribute to academic performance and stress prediction. It supports feature selection and model optimization processes.

Fig. 5 (b) depicts the relationship between attendance, study hours, and student outcomes through statistical visualization. It illustrates how Attendance values (ranging from 60 to 100) vary across FinalGrade categories (0–3). The figure highlights the distribution of StudyHours (approximately 5 to 45 hours) across StressLevel categories (0–2). It represents how behavioral factors influence both performance and stress levels. The analysis identifies trends and variability within different groups of students. These insights help in understanding the impact of study behavior on outcomes.

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

(a)



(b)

Fig. 5: EDA for (a) Distributions, Relationships, and Correlations (b) Attendance and Study Behavior Analysis

Confusion Matrix (CM)

	Pred 0	Pred 1	Pred 2	Pred 3
True 0	1398	22	18	7
True 1	15	1218	13	12
True 2	12	16	1431	15
True 3	16	18	9	1220

Classification Report (CR)

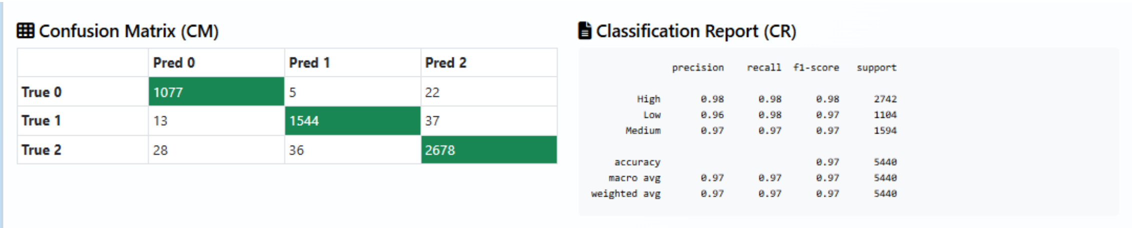
	precision	recall	f1-score	support
Average	0.97	0.97	0.97	1474
Excellent	0.97	0.97	0.97	1445
Good	0.96	0.97	0.96	1258
Poor	0.97	0.97	0.97	1263
accuracy			0.97	5440
macro avg	0.97	0.97	0.97	5440
weighted avg	0.97	0.97	0.97	5440

Fig. 6: Confusion Matrices and Classification Report obtained for Final Grade Classification -BARM Models

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

Fig. 6 illustrates the confusion matrix and classification report for the proposed BARM model. It depicts that the model achieves the highest accuracy of approximately 0.97, indicating superior classification performance. The confusion matrix shows very high diagonal values such as 1398, 1218, 1431, and 1220 with minimal misclassification. The classification report confirms that precision, recall, and F1-scores are consistently around 0.97 across all classes (0–3). This indicates excellent prediction capability and balanced performance. The model effectively distinguishes between all categories with high reliability.

Fig. 7 illustrates the confusion matrix and classification report for the proposed BARM model. It depicts that the model achieves the highest accuracy of approximately 0.97, indicating superior performance. The confusion matrix shows very high diagonal values such as 1077, 1544, and 2678 with minimal misclassification. The classification report confirms that precision, recall, and F1-scores are consistently around 0.97–0.98 across all classes (0–2). This indicates highly accurate and balanced predictions. The model effectively distinguishes between all stress levels with minimal error.



Confusion Matrix (CM)

	Pred 0	Pred 1	Pred 2
True 0	1077	5	22
True 1	13	1544	37
True 2	28	36	2678

Classification Report (CR)

	precision	recall	f1-score	support
High	0.98	0.98	0.98	2742
Low	0.96	0.98	0.97	1104
Medium	0.97	0.97	0.97	1594
accuracy			0.97	5440
macro avg	0.97	0.97	0.97	5440
weighted avg	0.97	0.97	0.97	5440

Fig. 7: Confusion Matrices and Classification Report obtained for Stress Level Classification - BARM Models

Fig. 8 illustrates the scatter plot for the proposed BARM model in ExamScore regression. It depicts a strong alignment of predicted values with the diagonal reference line across the full range (40–100). The distribution of points closely follows the ideal prediction line, indicating high correlation. Minimal deviation from the reference line suggests accurate predictions. The model effectively captures variations in input data. The spread of points reflects balanced and consistent performance.

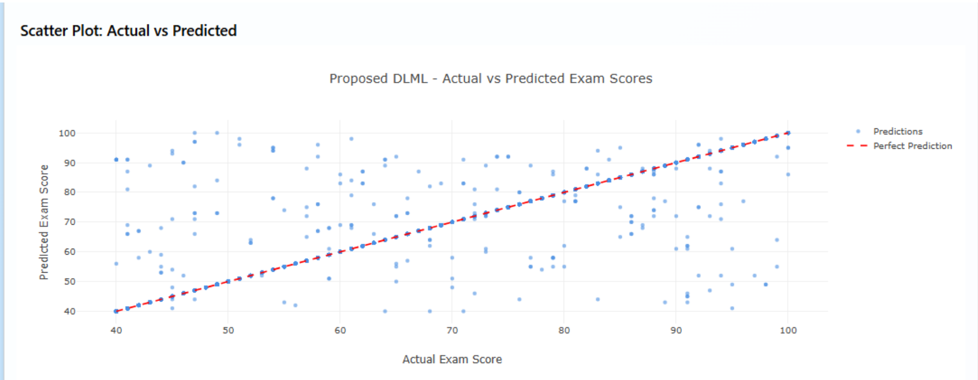


Fig. 8: Scatter Plots obtained for Exam Score Regression - BARM Models

Fig. 9 illustrates the comparative analysis of classification and regression models using performance metrics. It depicts that the proposed BARM model achieves the highest accuracy and F1-score (approximately 0.97–0.98) in both FinalGrade and StressLevel classification tasks. The figure shows that baseline models such as RF perform moderately well, while SVM, GBC, DTC, and LRC show comparatively lower performance. In regression analysis, the BARM model achieves the highest R^2 score (around 0.90) and the lowest RMSE (approximately 5), indicating superior predictive accuracy. Other models exhibit lower R^2 values and higher RMSE, reflecting weaker performance. The comparison highlights the effectiveness of the hybrid approach.

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

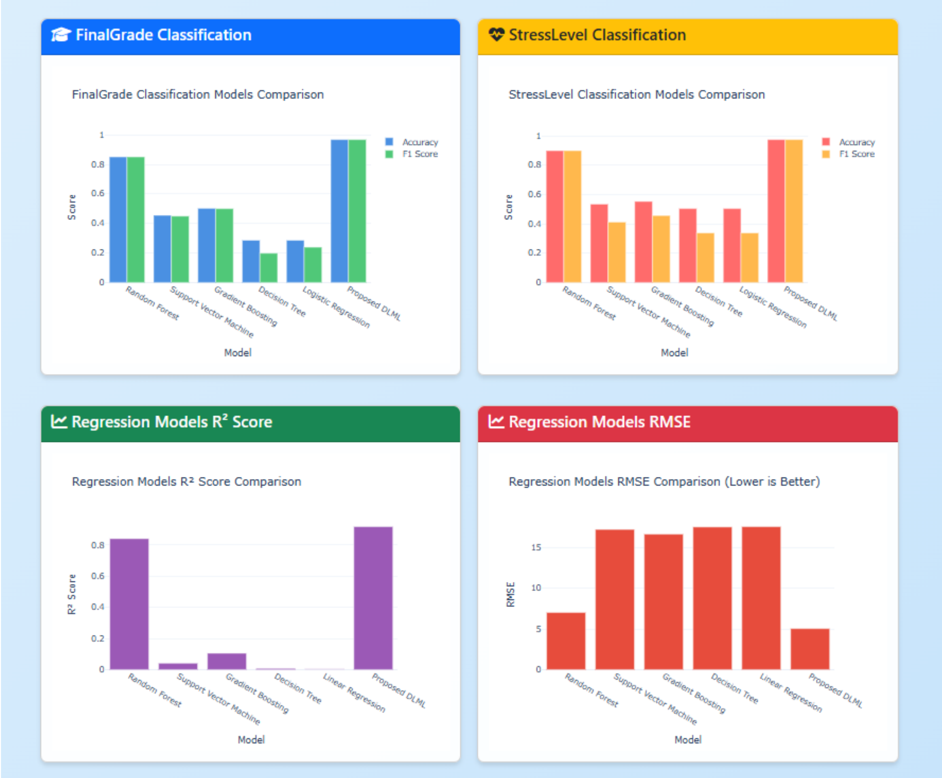
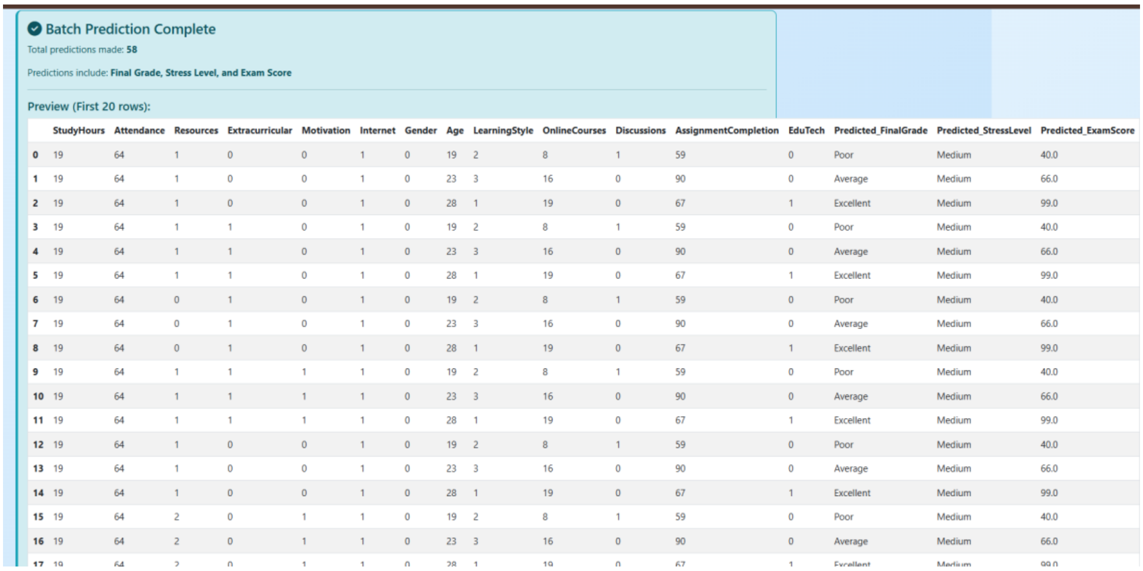


Fig. 9: Model Comparative Analysis

Fig. 10 depicts the batch prediction interface of the system, which processes multiple input records simultaneously. It illustrates the structured tabular output containing predicted values for FinalGrade (0–3), StressLevel (0–2), and ExamScore (40–100). The figure shows that predictions are generated for multiple samples, enabling large-scale analysis. The interface supports efficient handling of batch datasets. It highlights the integration of prediction results with input features. The system ensures consistent output formatting for all records.



Batch Prediction Complete

Total predictions made: **58**

Predictions include: **Final Grade, Stress Level, and Exam Score**

Preview (First 20 rows):

	StudyHours	Attendance	Resources	Extracurricular	Motivation	Internet	Gender	Age	LearningStyle	OnlineCourses	Discussions	AssignmentCompletion	EduTech	Predicted_FinalGrade	Predicted_StressLevel	Predicted_ExamScore
0	19	64	1	0	0	1	0	19	2	8	1	59	0	Poor	Medium	40.0
1	19	64	1	0	0	1	0	23	3	16	0	90	0	Average	Medium	66.0
2	19	64	1	0	0	1	0	28	1	19	0	67	1	Excellent	Medium	99.0
3	19	64	1	1	0	1	0	19	2	8	1	59	0	Poor	Medium	40.0
4	19	64	1	1	0	1	0	23	3	16	0	90	0	Average	Medium	66.0
5	19	64	1	1	0	1	0	28	1	19	0	67	1	Excellent	Medium	99.0
6	19	64	0	1	0	1	0	19	2	8	1	59	0	Poor	Medium	40.0
7	19	64	0	1	0	1	0	23	3	16	0	90	0	Average	Medium	66.0
8	19	64	0	1	0	1	0	28	1	19	0	67	1	Excellent	Medium	99.0
9	19	64	1	1	1	1	0	19	2	8	1	59	0	Poor	Medium	40.0
10	19	64	1	1	1	1	0	23	3	16	0	90	0	Average	Medium	66.0
11	19	64	1	1	1	1	0	28	1	19	0	67	1	Excellent	Medium	99.0
12	19	64	1	0	0	1	0	19	2	8	1	59	0	Poor	Medium	40.0
13	19	64	1	0	0	1	0	23	3	16	0	90	0	Average	Medium	66.0
14	19	64	1	0	0	1	0	28	1	19	0	67	1	Excellent	Medium	99.0
15	19	64	2	0	1	1	0	19	2	8	1	59	0	Poor	Medium	40.0
16	19	64	2	0	1	1	0	23	3	16	0	90	0	Average	Medium	66.0
17	19	64	2	0	1	1	0	28	1	19	0	67	1	Excellent	Medium	99.0

Fig. 10: Batch Prediction Interface

Table 1 presents the performance comparison of various classification models for predicting final grades. The Proposed BARM model significantly outperforms all baseline models, achieving the highest accuracy, precision, recall, and F1 score (0.9682). Among traditional methods, RF shows strong

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

performance with balanced metrics around 0.85. In contrast, SVM and GB demonstrate moderate results with accuracies below 0.55. DT and LR perform poorly, indicating limited capability in capturing complex patterns. The consistent metrics of BARM suggest robust and reliable classification performance.

Table 1: FinalGrade – Model Performance

Model	Accuracy	Precision	Recall	F1 Score
RF	0.8509	0.8514	0.8509	0.8509
SVM	0.4553	0.4605	0.4553	0.4498
GB	0.5026	0.5069	0.5026	0.5003
DT	0.2864	0.1534	0.2864	0.1984
LinR	0.2860	0.2804	0.2860	0.2394
Proposed BARM	0.9682	0.9682	0.9682	0.9682

Table 2 compares different models for stress level classification. The Proposed BARM model again achieves the best performance, with all evaluation metrics reaching 0.9741. Random Forest is the second-best model, delivering high accuracy and balanced precision and recall values close to 0.90. SVM, GB show moderate effectiveness but with noticeable drops in recall and F1 score. DT and LR yield relatively weak results, indicating poor generalization. The higher precision of SVM suggests some class bias despite lower overall accuracy.

Table 2: StressLevel – Model Performance

Model	Accuracy	Precision	Recall	F1 Score
RF	0.8985	0.8988	0.8985	0.8986
SVM	0.5344	0.6422	0.5344	0.4119
GB	0.5526	0.6036	0.5526	0.4565
DT	0.5040	0.2541	0.5040	0.3378
LinR	0.5040	0.2541	0.5040	0.3378
Proposed BARM	0.9741	0.9741	0.9741	0.9741

Table 3 presents regression model performance for predicting exam scores using error metrics and R^2 score. The Proposed BARM model achieves the lowest MAE (0.8270), MSE (25.5347), and RMSE (5.0532), along with the highest R^2 score (0.9179), indicating excellent prediction accuracy. RF performs well among baseline models with relatively low error values and a strong R^2 of 0.8409. Other models, including SVM and GB, show higher errors and significantly lower R^2 scores. DT and LinR on exhibit the weakest performance, with minimal explanatory power. The results demonstrate that BARM effectively captures underlying patterns in exam score prediction.

Table 3: Exam Score – Model Performance

Model	MAE	MSE	RMSE	R^2 Score
RF	1.7103	49.5114	7.0364	0.8409
SVM	14.5936	298.3599	17.2731	0.0412
GB	14.3744	278.4714	16.6875	0.1051
DT	15.1999	308.8406	17.5739	0.0075
LinR	15.2205	310.1464	17.6110	0.0033
Proposed BARM	0.8270	25.5347	5.0532	0.9179

5. CONCLUSION

This work successfully designed, implemented, and evaluated a Trivariate Attention BiLSTM–ODRL CART framework for jointly predicting student performance grade, stress level, and exam score within a unified educational decision-support system. Unlike traditional machine learning approaches that handle these outcomes independently, the proposed model learns shared representations from academic, behavioral, and demographic attributes such as study hours, attendance, motivation, past scores, and extracurricular involvement. Experimental results show that existing CART-based models such as Random Forest, SVM, Gradient Boosting, and Decision Tree achieved classification accuracies in the range of 85%–90%, with Random Forest reaching approximately 90% accuracy and average F1-scores above 0.87. However, these models showed limitations in capturing temporal dependencies and stress-related behavioral patterns.

The proposed Attention Recurrent Trees CART model, which integrates Attention-based BiLSTM feature extraction with ODRL, demonstrated superior performance by effectively learning bidirectional temporal patterns and emphasizing influential features through attention mechanisms. The optimized decision rule lists improved classification reliability for performance grading and stress detection while maintaining interpretability. In regression tasks, the model achieved lower MAE and RMSE values compared to Linear CART and tree-only models, resulting in more accurate exam score estimation. Overall, the system provides a holistic, accurate, and interpretable solution for student monitoring, enabling early identification of academic risk and stress conditions, and supporting timely educational interventions through real-time Flask-based deployment.

REFERENCES

[1]. Balcioğlu, Yavuz Selim, and Melike Artar. "Predicting academic performance of students with machine learning." Information Development 41, no. 3 (2025): 896-915.

[2]. Alnasyan, Bayan, Mohammed Basheri, Madini Alassafi, and Khozama Alnasyan."Kanformer: an attention-enhanced deep learning model for predicting student performance in virtual learning environments." Social Network Analysis and Mining 15, no. 1 (2025): 25.

[3]. Kesgin, Kadir, Salih Kiraz, Selahattin Kosunalp, and Bozhana Stoycheva. "Beyond performance: explaining and ensuring fairness in student academic performance prediction with machine learning." Applied Sciences 15, no. 15 (2025): 8409.

[4]. Ab Rahman, Noor Fadzilah, Shir Li Wang, Theam Foo Ng, and Amr S. Ghoneim. "Artificial intelligence in education: A systematic review of machine learning for predicting student performance." Journal of Advanced Research in Applied Sciences and Engineering Technology 54, no. 1 (2025): 198-221.

[5]. Alalawi, Khalid, Rukshan Athauda, and Raymond Chiong. "An extended learning analytics framework integrating machine learning and pedagogical approaches for student performance prediction and intervention." International Journal of Artificial Intelligence in Education 35, no. 3 (2025): 1239-1287.

[6]. Saqr, Mohammed, Kamila Misiejuk, Santtu Tikka, and Sonsoles López-Pernas. "Artificial intelligence: using machine learning to predict students' performance." In Advanced Learning Analytics Methods: AI, Precision and Complexity, pp. 41-78. Cham: Springer Nature Switzerland, 2025.

[7]. Ahmed, Wesam, Mudasir Ahmad Wani, Pawel Plawiak, Souham Meshoul, Amena Mahmoud, and Mohamed Hammad. "Machine learning-based academic performance prediction with

explainability for enhanced decision-making in educational institutions." Scientific Reports 15, no. 1 (2025): 26879.

[8]. Doma, Victor, Ali Abd Almisreb, Emine Yaman, Salue Amanzholova, and Nurlaila Ismail. "Classification of Student Stress Levels Using a Hybrid Machine Learning Model." In AI-Driven: Social Media Analytics and Cybersecurity, pp. 77-104. Cham: Springer Nature Switzerland, 2025.

[9]. Asto Buditjahjanto, I., Naim Rochmawati, and Parama Diptya Widayaka. "Stress Level Classification Towards Student Learning Outcomes Using Support Vector Machine." TEM Journal 14, no. 2 (2025).

[10]. Gedam, Shruti, Sandip Dutta, and Ritesh Jha. "Analyzing mental stress in Indian students through advanced machine learning and wearable technologies." Scientific Reports 15, no. 1 (2025): 20610.

[11]. Deena, G., A. Sandhya, and K. Raja. "Machine Learning-Based Classification And Prediction Of Student Stress Levels: A Comparative Study Of Algorithms." Journal of Theoretical and Applied Information Technology 102, no. 19 (2024).

[12]. Ali, Hassan, M. Hamza Amin, Attique Ur Rehman, Sabeen Javaid, Tahir Muhammad Ali, and Azka Mir. "Predictive Modeling of Students' Stress Levels Using Machine Learning Algorithm." In 2024 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), pp. 417-423. IEEE, 2024.

[13]. Biswas, Anusuya. "Modelling an innovative machine learning model for student stress forecasting." Global Perspectives in Management 2, no. 2 (2024): 22-30.

[14]. Suraj Arya, Anju, and Nor Azuana Ramli. "Predicting the stress level of students using supervised machine learning and artificial neural network (ANN)." Indian Journal of Engineering 21 (2024): e9ije1684.

[15]. Arias, Enrique Macías, Jorge Parraga-Alava, and David Zambrano Montenegro. "Stress detection among higher education students: A comprehensive systematic review of machine learning approaches." In 2024 Tenth International Conference on eDemocracy & eGovernment (ICEDEG), pp. 1-8. IEEE, 2024.