INVERGE JOURNAL OF SOCIAL SCIENCES
https://invergejournals.com/
ISSN (Online): 2959-4359, ISSN (Print): 3007-2018
Volume 4 Issue 1, 2025

DEEP LEARNING-DRIVEN STUDENT PERFORMANCE ANALYSIS: DETECTING ANOMALIES AND PREDICTING ACADEMIC SUCCESS

Muhammad Nadeem Gul [1], Muhammad Arif [2], Sidra Gulzar [3], Ghulshan Naveed [4], Waseem Abbasi [5]

Affiliations:

[1,2,3,4] Department of Computer Science & IT, Superior University, Lahore, 54000, Pakistan

[1] prof.nadeemgul@gmail.com
[2] md.arif@superior.edu.pk
[3] sidragulzar25@gmail.com
[4] gulshanawan718@gmail.com

[5]Department of Computer Science, The University of Lahore, Sargodha Campus, Sargodha 40100, Pakistan
[5] waseemabbasi97@gmail.com

Copyright:
Author/s

License:

Abstract

Accurately predicting student performance and identifying anomalies in academic datasets has become increasingly crucial for enhancing educational outcomes and enabling data-driven interventions in modern learning environments. Traditional statistical methods and conventional machine learning approaches often struggle with the multidimensional nature and increasing scale of contemporary student datasets, which encompass diverse academic, behavioral, and socio-demographic variables. This study explores advanced deep learning techniques; including Autoencoders for unsupervised anomaly detection, Recurrent Neural Networks with Long Short-Term Memory architectures for temporal pattern recognition, and Deep Neural Networks for comprehensive performance prediction to address these challenges. The proposed framework demonstrates significant improvements in detecting subtle performance anomalies that often precede academic difficulties, while simultaneously predicting longitudinal success patterns with greater accuracy than traditional methods. By leveraging the hierarchical feature learning capabilities of deep architectures, our system enables early identification of at risk students through continuous analysis of complex, nonlinear relationships in educational data, allowing institutions to implement timely, personalized interventions. Research studies have empirically validated the effectiveness of these models in educational contexts, showing superior performance in measuring student achievement patterns and predicting learning outcomes. The findings contribute to theoretical advancements in educational analytics but also provide practical insights for curriculum designers and policy makers seeking to optimize instructional strategies. Furthermore, the study establishes significant benchmarks for educational contexts by demonstrating how deep learning can enhance both teaching methodologies and student support systems through data-driven insights. This research makes a substantial contribution to the growing field of Educational Data Mining by proposing a robust deep learning framework that serves as both a predictive tool and a baseline for future studies in student performance analysis, while also addressing critical challenges in model interpretability and implementation scalability within real-world educational settings.

Keywords:

Performance prediction, Deep learning, Autoencoders, Recurrent Neural Networks (RNN), Deep Neural Net- works (DNN), Anomaly and outliers detection.

Introduction

Educational technology improvements have delivered a substantial change to how students get evaluated at school. Deep learning (DL) serves as a fundamental analysis tool for large datasets that contain student demographic along with socio-economic information and behavioural indicators as well as academic characteristics in the big data era. Although detecting irregularities and extreme cases in student achievement stands as a current research problem, Deep Neural Networks (DNNs) among deep learning models show better predictive potential relative to traditional machine learning methods which enhances success rate and risk detection accuracy in student assessment (Abatal et al., 2025). The research analyses deep learning methods to generate improved educational understanding for better outcomes in student performance.

Our study includes different pieces of information from multiple sources including ethnicity and test preparation behaviours alongside gender and family background. Rather than handle general student results, deep learning systems can accurately identify cases where students step far away from academic predictions. Researchers use convolutional neural networks (CNNs) to evaluate university teaching quality and overcome difficulties in offering subjective real-time results (Gao, 2025). The CNN models show accurate results for teacher evaluation by processing multiple student performance sources and outcome ratings and delivering a 92% success rate. Through deep learning, researchers have discovered important capabilities for examining academic risk and providing prompt support. The study applied model selection framework with Reinforcement Learning (RL) for anomaly detection in time series data (Ghanim & Awad, 2025). The Neutrosophic Deep Learning Model (Shitaya et al., 2025) provided accurate predictions about student dropout rates by processing uncertain student data successfully. The combination of deep learning with learning analytics alongside educational data mining helps recognize students at risk prematurely so proper academic support measures can be implemented right on time (Vaidya & Sharma, 2024).

The evaluation of student performance anomalies and outliers generates meaningful feedback, which helps teachers, and officials create better educational approaches to prevent any student from being neglected. The implementation of deep learning for performance detection follows a deep learning approach that supports educational data mining (EDM) through its ability to deliver actionable insights. Institutions across the board utilize deep learning models to forecast student outcomes which bring advantages to both school administration and their teachers alongside parents and their students. The performance examination using deep learning-based methods occurs through analysis of publicly available student data (Kaggle, n.d.).

Problem Statement

The precise identification of student performance patterns together with data anomaly detection stands as an ongoing major challenge in the field. Supervisory methods demonstrate ineffective data processing capabilities when dealing with complex educational datasets at a large scale. This research evaluates deep learning models through Autoencoders alongside Recurrent Neural Networks using Long Short-Term Memory and Deep Neural Networks for detecting anomalies and outcome prediction while dispensing with the need for labelled data. The research utilizes advanced models to achieve better accuracy and operational speed for educational evaluations.

The proposed framework addresses the following research objectives:

1. Deep learning to detect abnormal behaviour in quantitative scores of students
2. Evaluating the effectiveness of Autoencoders, RNN with LSTM, and DNN in anomaly detection
3. Exploring the relationships of core demographic and academic characteristics to the degree of performance variability
4. Providing actionable insights to enhance academic support systems

Contributions

This research will help expand the knowledge base in the field of Educational Data Mining (EDM) as it applies deep learning models in an effort to predict the performance of students and for use in identifying anomalous cases. The deep learning techniques identified include Autoencoder, Recurrent Neural Network

(RNN) using Long Short-Term Memory (LSTM) and Deep Neural Network (DNN). In this analysis, feature extraction methods are used with assistance of data preprocessing techniques; however, evaluation of model performance is based on various parameters.

This research employs deep learning to detect anomalies while making performance predictions, which helps educational institutions and their stakeholders base their decisions on data. Educational assessment methods and teaching practice enhancement, as well as academic intervention strength become possible through the valuable research data. The research operates both as guidance for decision-making and as a standard that future educational data mining research can use for comparisons.

Organization

The subsequent sections organize the content as follows: Section II evaluates existing research about educational data analytics specifically regarding real-time student performance prediction and assessment methods. The study utilizes the dataset explained in Section III while presenting its complete methodology for performance prediction. The analysis of simulated results appears in Section IV with essential findings presented. Section V provides concluding statements while describing prospective research directions that can boost student performance prediction model accuracy and application abilities.

Related Work

In the field of educational data mining, identifying students at risk of low academic success is crucial for improving learning performances and enabling targeted interventions (Nassif et al., 2021). While many educational departments have adopted predictive models, challenges persist in ensuring accuracy and stability due to limited technological investment in advanced analytical tools. Deep learning approaches, such as Autoencoders, Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM), and Deep Neural Networks (DNN), provide promising solutions for detecting anomalies in student academic performance (Bulusu et al., 2020). However, their implementation remains challenging due to the complexity of deep learning models and the need for substantial computational resources. Additionally, factors such as student motivation and external influences further complicate academic predictions, highlighting the need for robust and adaptable deep learning frameworks (Alam & Mohanty, 2022).

Different standardized tests most often give instructors a narrow picture of their pupils' abilities since students might not perform optimally due to various factors. That is why student performance outcomes cannot be determined solely by test results in a particular subject, as students may excel in some areas while struggling in others (Kamalov et al., 2021). Deep learning-based anomaly detection helps identify cases requiring special attention, enabling tailored strategies for individual students. Using Autoencoders, RNN-LSTM, and DNN, institutions can monitor student behaviour throughout the study period to identify potential struggling students (Wang et al., 2021). Remote learning difficulties can be addressed early through deep learning diagnostic tools, allowing teachers and parents to intervene promptly (Vaidya & Sharma, 2024).

To provide a comprehensive overview of the literature, a Systematic Literature Review (SLR) was conducted to identify deep learning models for anomaly detection in education (Nassif et al., 2021). The review analysed 290 articles published between 2000-2020, identifying 43 applications, 29 models, and 22 datasets. It also discusses improving generalization by recognizing out-of-distribution (OOD) data and adversarial inputs that could lead to incorrect predictions using Autoencoders and RNN-LSTM architectures. Related research (Bulusu et al., 2020) categorizes various deep learning approaches based on their assumptions and methods. The detection of anomalies associated with academic dishonesty has also been explored through deep learning in academic integrity monitoring (Kamalov et al., 2021). Another study (López-García et al., 2023) uses continuous assessment data with RNN models and anomaly detection approaches to identify low scores, promoting fairness in remote learning environments.

Deep learning models show significant potential in modelling student academic performance. Research by Hussain et al. (2021) demonstrates that deep learning outperforms traditional regression models in assessing student performance, achieving a lower mean absolute error (1.61 vs. 1.97) and reduced loss (4.7 vs. 6.7) compared to linear regression models. These findings highlight the potential of deep learning techniques for identifying complex patterns in student performance, enabling early interventions and optimized support systems.

Table 1

Related Work

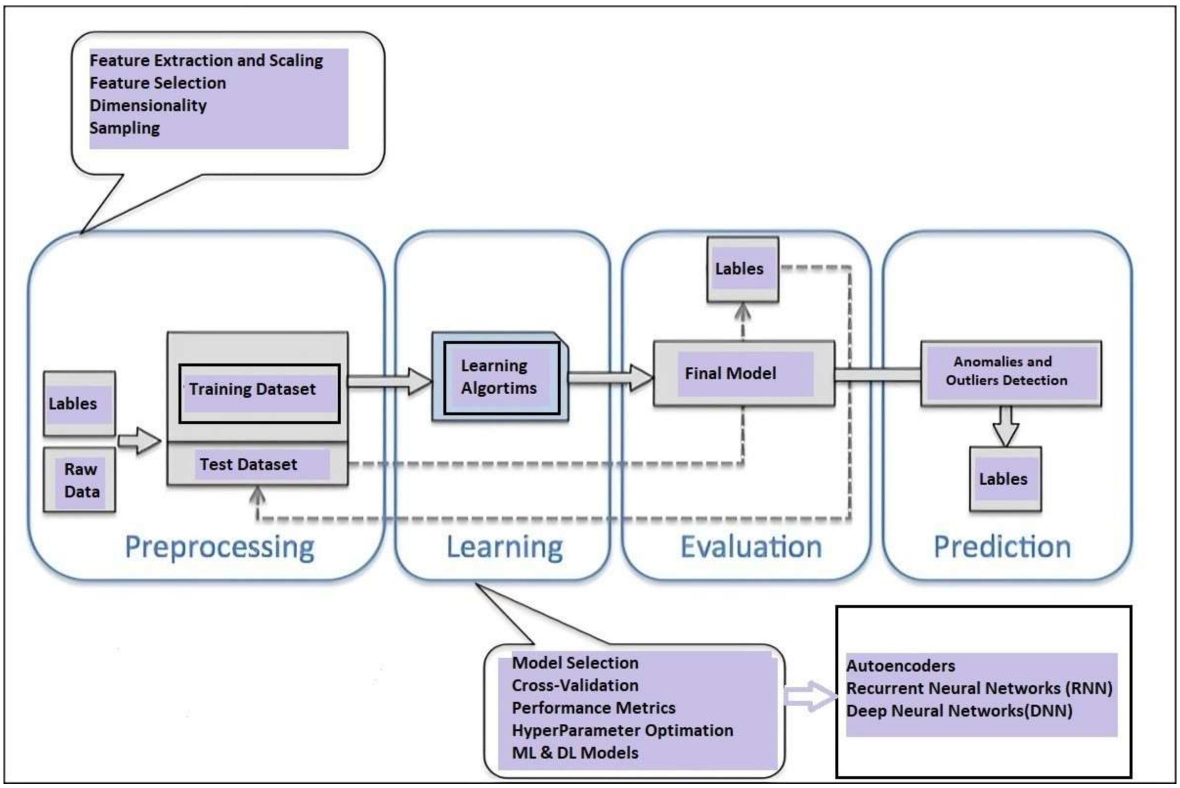
Ref	Study Focus	Machine Learning Techniques Used	Key Findings	Accuracy/Results	Dataset
Nassif et al. (2021)	Systematic review of ML models for anomaly detection	KMeans, Isolation Forest, DBSCAN	Identified 43 applications, 29 ML models, 22 datasets	Highlights unsupervised anomaly detection techniques	Articles published (2000-2020)
Kamalov et al. (2021)	Detecting cheating in online exams	RNN with anomaly detection	Effective identification of abnormal scores	Preserved academic integrity in remote exams	Continuous assessment data
López-García et al. (2023)	Predicting academic failure	XGBoost	Identified at-risk students in Colombia	Improved failure predictions	Industrial University of Santander student data
Alruwais & Zakariah (2023)	Predicting classroom engagement	CATBoost	Achieved high accuracy and recall for engagement prediction	Accuracy: 94.64%, Recall: 100%	Virtual Learning Environment (VLE) data
Pek et al. (2022)	Predicting at-risk students using ensemble methods	Naïve Bayes, Random Forest, SVM (stacking)	Achieved 94.8% accuracy with demographic + academic data, 98.4% with only academic data	High performance in predicting at-risk students	Demographic and academic datasets
Hussain et al. (2021)	Deep learning for academic performance	Deep Learning, Regression models	DL outperformed traditional methods for performance prediction	MAE: 1.61, Loss: 4.7	Academic performance data
Wang et al. (2021)	Anomaly detection using student-teacher frameworks	Pre-trained teacher-student network	Detected anomalies efficiently with multi-scale feature matching	Surpassed state-of-the-art performance on MVTec dataset	MVTec dataset
Al-Fairouz & Al-Hagery (2020)	Educational Data Mining (EDM) for hidden patterns	Classification, Regression, Association Rules	Identified issues in student performance and enhanced academic advising	Improved decision-making	Data from College of Business and Economics (CBE)
Alam & Mohanty (2022)	Guide for educators to predict student performance	Data mining techniques (various)	Made data mining tools accessible to educators	Enhanced student success prediction accuracy	Educational datasets
Huang et al. (2020)	Risk identification for at-risk students	Classification methods, Spearman correlation	Improved accuracy with high feature significance	Enhanced risk identification accuracy	Datasets from 3 universities (Taiwan, Japan)

Deep learning techniques have significantly improved feature matching and accuracy in educational anomaly detection. An innovative approach (Wang et al., 2021) uses a student-teacher network structure where a teacher model transfers knowledge to help a student model learn normal data distributions. This architecture enables effective detection of anomalies at different scales, outperforming previous methods. Furthermore, deep learning models are increasingly applied in educational data mining to analyse historical student records and identify patterns that improve learning outcomes (Al-Fairouz & Al-Hagery, 2020). These models help pinpoint academic strengths and weaknesses, facilitating targeted interventions. Additional research (Alam & Mohanty, 2022) emphasizes deep learning's role in developing achievement metrics, selecting optimal features, and enhancing prediction models to ensure student success.

Deep learning models have proven effective in student risk identification through behavioural and academic performance analysis. As Panchenko (2018) emphasizes, early prediction of student performance enables instructors to provide timely support to struggling students while offering additional motivation to high performers. Analysis of 6,597 students' behavioural data revealed that academic success strongly correlates with diligence, organization, and sleep patterns (Acharya & Sinha, 2014). These findings demonstrate how advanced analytics can extract valuable educational insights from large datasets, supporting academic performance prediction and intervention strategies.

Figure 1

Proposed Model of Students for making Predictions: Detecting Anomalies and Outliers



Recent Advances in Deep Learning for Educational Outcomes

Recent research demonstrates the growing application of deep learning in forecasting student outcomes and evaluating learning behaviours. Issah et al. (2023) provide detailed analysis of artificial neural networks (ANNs) in modelling relationships between learner characteristics and academic performance. Compared to conventional machine learning methods, deep learning models excel at capturing complex, non-linear relationships within educational data, significantly improving prediction accuracy. Further research by Pallathadka et al. (2023) develops a comprehensive deep learning framework for classifying and forecasting both teacher and student performance. These advancements highlight deep learning's transformative potential in educational data mining for performance prediction and student support systems.

Student performance prediction has emerged as a critical focus in Educational Data Mining (EDM), particularly as institutions increasingly adopt data-driven approaches for academic improvement (Alam & Mohanty, 2022). Deep learning architectures including Autoencoders, Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) layers, and Deep Neural Networks (DNN) have shown remarkable potential in analysing multidimensional student data. These models effectively combine demographic and academic features to identify learning patterns, predict performance trajectories, and flag at-risk students (Hussain et al., 2021). While deep learning outperforms traditional methods in capturing complex educational data relationships, challenges persist in optimal feature selection, data pre-processing pipelines, and handling class imbalances - areas requiring further research to improve model robustness and interpretability (Nassif et al., 2021).

Methodological Approach

This study implements three deep learning architectures for anomaly detection in student performance data: Autoencoders (5% anomaly detection rate), RNN-LSTM (4.7%), and DNN (4.9%). These models analyze behavioural and academic patterns to identify performance risks, enabling early intervention strategies (Vaidya & Sharma, 2024). The integration of continuous assessment data with deep learning tools improves prediction accuracy and provides comprehensive performance insights, facilitating the development of adaptive learning systems and data-driven academic interventions.

Methodology

The research methodology employs deep learning techniques; Autoencoders, RNN, and DNN, in a comprehensive framework for student performance prediction and anomaly detection (see Figure 1). The conceptual model specifically targets early identification of at-risk students through advanced anomaly detection algorithms. The methodological pipeline includes:

1. Data pre-processing and feature engineering
2. Experimental design and model configuration
3. Training and validation procedures
4. Performance evaluation metrics

This structured approach enhances predictive accuracy while providing actionable insights for educational institutions to implement targeted student support mechanisms (Gao, 2025). The findings contribute to both theoretical advancements in educational data mining and practical applications for learning optimization.

Dataset Collection

The dataset used for this research obtained from Kaggle (n.d.) is clean and well-suited for making student performance prediction analysis. It was heavily analyzed and cleaned using the necessary elements in Python on Google Colab before being fitted into the chosen model. The dataset includes 8 features with 1000 samples, where 5 features are categorical while 3 are numerical; these are ideal features as that range helps design a model for deep learning effectively (Alam & Mohanty, 2022).

- **Gender**: Denotes the gender of the student (male or female)
- **Race**: Categorizes students into five distinct groups: A, B, C, D, or E
- **Parent Education Level**: Specifies the highest level of education attained by the student's parents, including categories such as school level, high school, college, associate degree, bachelor's degree, and master's degree
- **Lunch Consumption**: Indicates whether students receive a Standard or Reduced lunch before classes or exams
- **Exam Preparation**: Denotes whether the student has prepared for the exam (marked as 'yes' or 'no')
- **Grades**: The final three columns represent the marks obtained by students in mathematics, reading, and writing assessments, each scored out of a maximum of 100

INVERGE JOURNAL OF SOCIAL SCIENCES
https://invergejournals.com/
ISSN (Online): 2959-4359, ISSN (Print): 3007-2018
Volume 4 Issue 1, 2025

The dataset is free from duplicate entries, which enhances the quality of analysis (Hussain et al., 2021). Data types for each attribute were determined, and checks for null values were performed to ensure data quality. The absence of missing values in the variables means the dataset contains complete cases, making it ready for further analysis.

Visualization and Plotting

This section presents a graphical analysis of the student dataset to examine mean distributions and derive meaningful insights. Visualizations, including Histograms, Kernel Density Estimation (KDE) plots, and combined Histogram-KDE plots, provide a structured overview of the data (Pek et al., 2022). To facilitate gender-specific analysis, separate datasets for male and female students were created, revealing distinct performance patterns, as illustrated in Figure 2. These visualizations help identify trends and variations in student outcomes, supporting data-driven decision-making. Gender distribution shown as a histogram indicates there are 518 males (51.8%) and 482 females (48.2%) in the dataset.

Figure 2

Gender Visualization

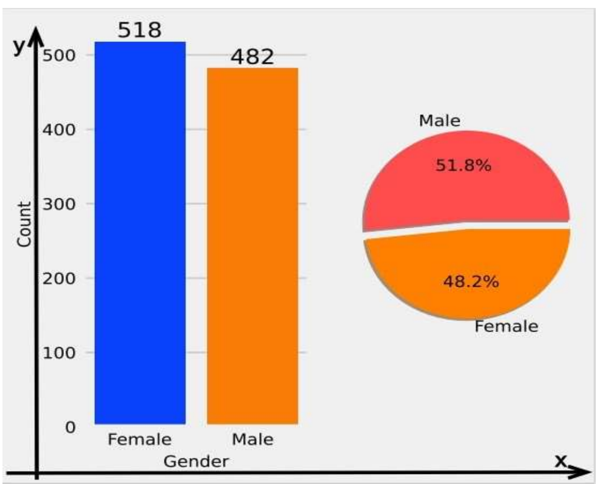


Figure 3

Ethnicity Visualizations

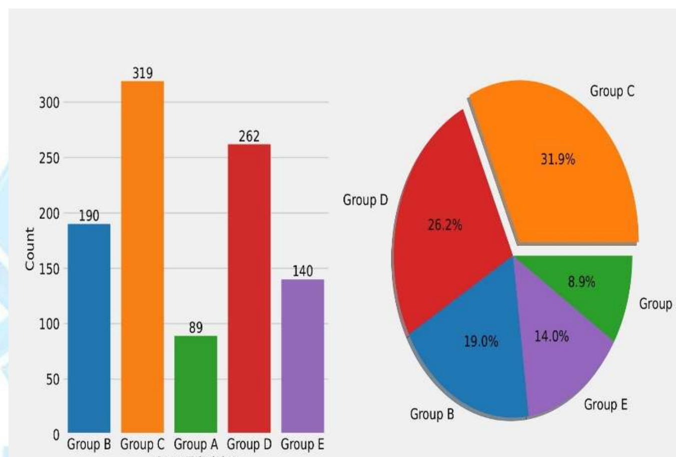
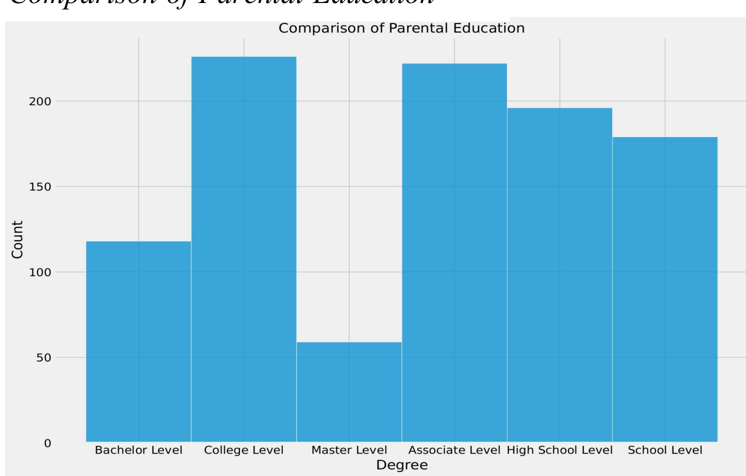


Figure 4

Comparison of Parental Education



Total 518 males and 482 females we have total 1,000 individuals in the dataset. From this plot, we can see that 51.8% of dataset is males and 48.2% is females. Figure 3 illustrates the distribution of the ”race” attribute, showing that groups C and D are the most prevalent, while group A has the lowest

INVERGE JOURNAL OF SOCIAL SCIENCES
https://invergejournals.com/
ISSN (Online): 2959-4359, ISSN (Print): 3007-2018
Volume 4 Issue 1, 2025

representation. The figure also provides a proportional breakdown of each racial or ethnic category, offering a clearer perspective on demographic composition. Subsequently, the distribution of parental education levels was graphically illustrated to analyze the data across various categories, including School Level, High School Level, College Level, Associate Level, Bachelor Level, and Master Level. This visualization, presented in Figure 4, provides an insightful depiction of the parental education distribution within the dataset.

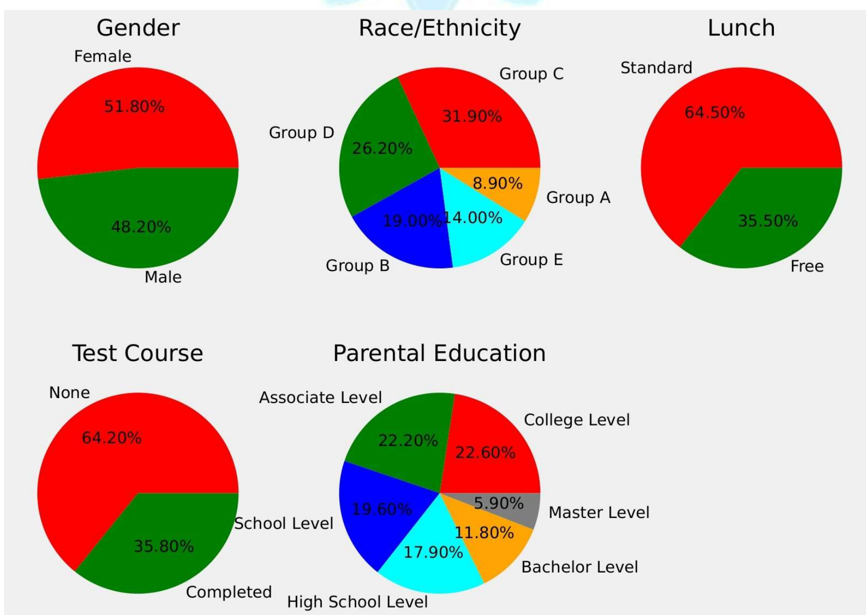
The comparative analysis plot indicates that the largest proportion of students' parents fall into the "College Level" category, followed by the "Associate Level" and then the "High School Level" category. Figure 4 illustrates that students whose parents have attained a "Master Level" or "Bachelor Level" education demonstrate significantly better academic performance compared to those from other educational backgrounds. Figure 3 illustrates the distribution of the "race" attribute, showing that groups C and D are the most prevalent, while group A has the lowest representation. The figure also provides a proportional breakdown of each racial or ethnic category, offering a clearer perspective on demographic composition.

Pie Plot and Multivariate Analysis

Figure 5 employs a pie chart to effectively encapsulate the selected dataset, providing a clear and comprehensive overview of all categorical features. This visualization aids in informed decision-making regarding student attributes and their associated factors, highlighting the intricacies and significance of the dataset. The plot also reveals a linear relationship between scores, indicating that performance increases consistently across all subjects. Notably, student performance closely correlates with factors such as lunch consumption, race/ethnicity, and parental education level, while gender and test preparation show weaker correlations with academic outcomes. Based on these findings, researchers can make informed decisions. Female students demonstrate higher pass rates and top scores, while the analysis indicates that test preparation courses may have a limited impact on overall student achievement. Therefore, an emphasis on graduation levels and course completion may yield greater benefits for students.

Figure 5

Multi Variant Analysis and Pie Plot



Dataset Pre-Processing

Before applying machine learning classifiers and regression models to the dataset, it is essential to eliminate less influential and irrelevant features. This strategy aims to enhance model performance and achieve outcomes that are more favorable.

Dataset Feature Engineering

Raw datasets, when applied directly to classifiers and machine learning algorithms, often yield sub-optimal results. As a result, data preprocessing is essential for ensuring the effectiveness of algorithmic applications. During the feature-engineering phase, the dataset is transformed to optimize and enhance the accuracy of the selected classification features. As part of this transformation, a new column labeled ”Total” was introduced, which aggregates the Mathematics, Reading, and Writing scores to provide a holistic measure of overall student performance. The updated dataset, now incorporating the ”Total” column, is presented in Table 2.

Additionally, details about the data types for each attribute, with categorical attributes represented as strings and numerical attributes assigned as integers. It is important to note that not all values in the dataset are whole numbers. Therefore, a necessary step involves transforming these values to suit the proposed model. This transformation includes converting categorical or non-numeric features into a binary format, where ”1” denotes ”true” and ”0” represents ”false.” This conversion allows for the creation of a new dataset for representing the different options for each feature.

Table 2

Sample Student Attributes Dataset

Gender	Race	Parental Level of Education	Lunch	Preparations	Computer Science Score	Reading Score	Writing Score	Total
Male	Group B	Bachelor Level	Standard	None	74	73	76	223
Male	Group C	College Level	Standard	Completed	67	88	87	242
Female	Group B	Master Level	Standard	None	91	96	94	281
Male	Group A	Associate Level	Free/ Reduced	None	49	59	60	168
Female	Group C	College Level	Standard	None	77	79	76	232

Visualization of Dataset after Feature Engineering

The data is transformed and feature engineering is performed to re-visualize the dataset, with the goal of extracting accurate results, facilitating predictions, and assessing student performance. Figure 6 presents a visualization depicting the relationship between gender and the results obtained by the students. The graph in Figure 6 clearly shows that female students tend to achieve higher and more favorable grades compared to their male peers. This observation highlights a notable gender difference in academic performance within the dataset.

Figure 7 visualizes the relationship between race/ethnicity and students’ corresponding grades. This graphical representation aims to highlight recognizable patterns or differences in student performance across different racial or ethnic groups. Figure 7 illustrates that students from race/ethnicity group C tend to achieve higher scores than other groups, indicating a notable performance trend. Additionally, after data transformation and feature engineering, the correlation between parental education levels and student test scores is analyzed, providing insights into how parental background influences academic achievement.

INVERGE JOURNAL OF SOCIAL SCIENCES
https://invergejournals.com/
ISSN (Online): 2959-4359, ISSN (Print): 3007-2018
Volume 4 Issue 1, 2025

Figure 6
Relationship of Race and Total Score

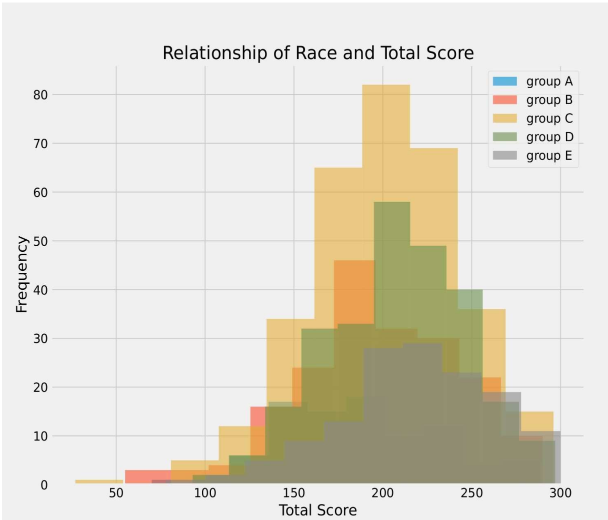


Figure 7
Relationship of Total Score and Gender

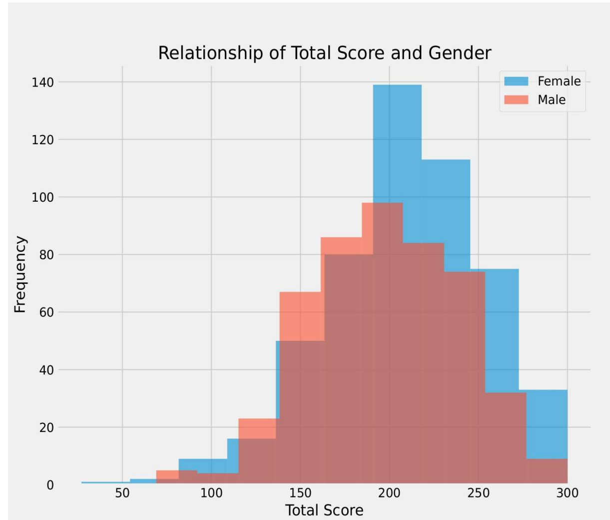


Figure 8 visualizes the correlation between students' grades and their parents' education level. The plot suggests that students with parents who have attained college or graduate-level education, as well as those with high school-educated parents, tend to perform better. While this indicates a potential link between parental education and student achievement, further analysis is needed to establish a definitive relationship.

Figure 8
Relationship of Parents and Overall Score

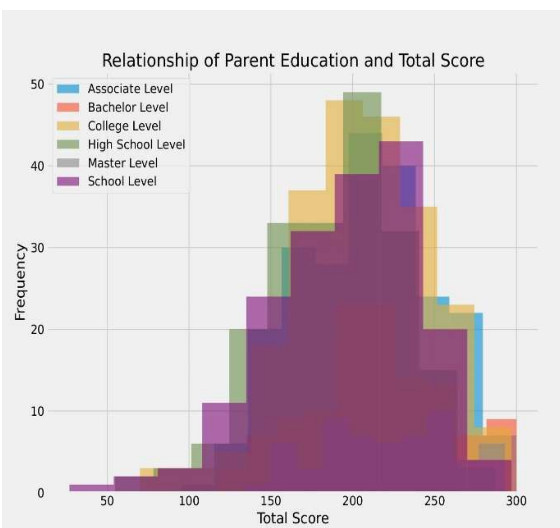


Figure 9
Lunch Plan and Total Score

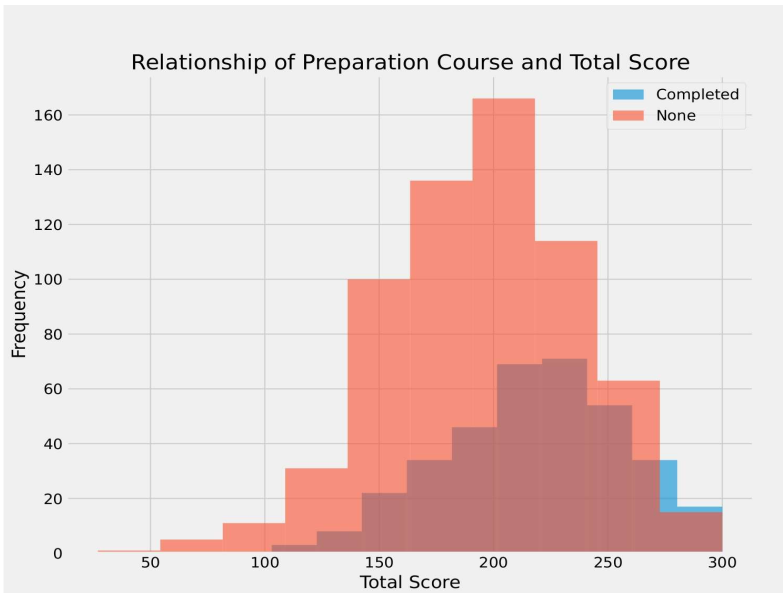
Figure 9 presents a histogram depicting the relationship between student performance and lunch preferences. The distribution of total scores across different lunch plans helps assess whether lunch type influences academic achievement.

Figure 9 illustrates the relationship between students' performance in math, reading, and writing based on lunch type. The histogram reveals that students who receive a standard lunch tend to achieve higher total scores compared to those receiving free or reduced-price lunch. This visualization helps assess the potential influence of nutritional access on academic performance, offering insights into correlations between lunch type and student outcomes.

In the final stage of visualization, after feature engineering, Figure 10 depicts the relationship between students' preparation status and their scores. This analysis highlights the impact of course preparation on academic performance, offering valuable insights into the role of prior readiness in student success. Figure 10 illustrates the correlation between students' course preparation levels and their scores. Notably, a significant proportion of students fall into the "none" category, indicating a lack of course preparation. This trend suggests a potential impact on student performance. Despite this, the analysis considers all five-preparation categories comprehensively, ensuring a holistic evaluation of their effects on academic outcomes.

Figure 10

Test Preparation and Total Score



Data preprocessing and cleaning are some of the critical steps that need to be carried out in data preprocessing for later use in the deep learning model. The idea of the train set and the test set are followed to ensure that an accurate result is achieved when testing the performance of the model. This helps the models generalize well and avoid the rigidity of memorizing data by taking a portion of the data to learn from it. Appropriate data conditioning means better achievements of the goals of deep learning since the indicators of student performance are determined more accurately.

Results and Discussion

This section applies multiple anomaly detection algorithms to the student performance dataset, focusing on deep learning techniques: Autoencoders, RNN (LSTM), and DNN. The effectiveness of each algorithm is evaluated based on the percentage of anomalies detected, anomaly characteristics, computational efficiency, and robustness. These comparisons highlight the strengths of deep learning in identifying inconsistencies in student performance data.

The findings highlight the varying effectiveness of deep learning techniques in anomaly detection. While traditional methods like KMeans provide general clustering, RNN (LSTM) and DNN excel at capturing intricate patterns of deviation. These results emphasize the value of deep learning in identifying outliers in student performance, enhancing the potential of educational data mining for targeted interventions.

Anomalies and Outliers Detection using Autoencoders

The autoencoders uses student performance data analysis through feature reconstruction and deviation detection based on the generated reconstruction error. A model contains two parts, which encompass an encoder to transfer information to reduced dimensions then a decoder to regenerate the data. The detection of

INVERGE JOURNAL OF SOCIAL SCIENCES
https://invergejournals.com/
ISSN (Online): 2959-4359, ISSN (Print): 3007-2018
Volume 4 Issue 1, 2025

anomalies takes place when reconstruction errors exceed the 95th percentile metric. A combination of Adam optimizer and mean squared error loss function serves as an optimization method to achieve effective detection of mathematical and reading and writing score outliers.

Figure 11
3D Scatter Plot of Student Performance with Anomalies Detected Using Autoencoders

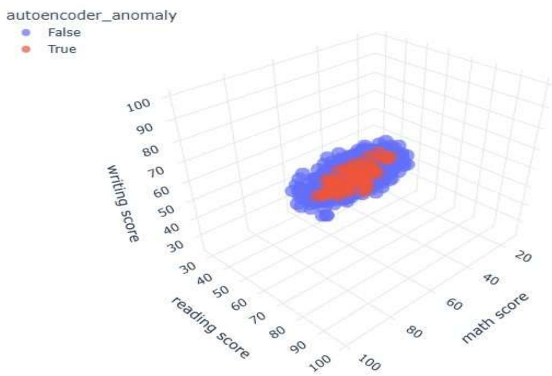
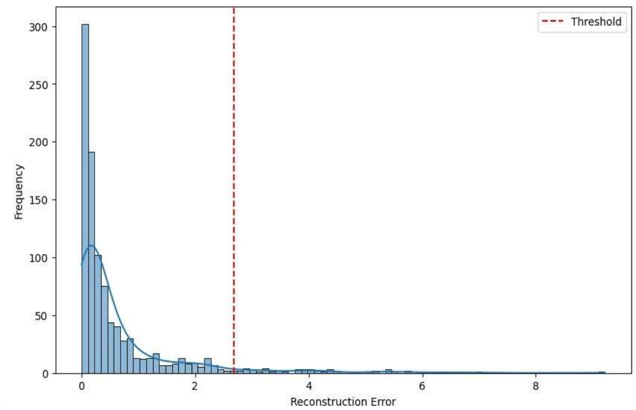


Figure 12
Distribution of Reconstruction Errors with Threshold for Anomaly Detection



Through the Autoencoder model the 5% anomalies in the data set were determined to be 50 observations. We display the discovered anomalies on a 3D scatter plot where math reading and writing scores mark each axis. Students who perform differently from the typical pattern are displayed in red for easy recognition of their peculiar results. A histogram of reconstruction errors helps us see normal data values clearly differ from points that stand out. The red line in the threshold divides typical values of data from problem points. The Autoencoder effectively finds rare performance outliers to help users find unique patterns and check for possible data problems.

Anomalies and Outliers Detection Using Recurrent Neural Networks (RNN) with LSTM

The RNN with LSTM layers was employed for anomaly detection in student performance data, analyzing math, reading, and writing scores. Before applying PCA, the data was standardized using Standard Scaler. The dataset was then reformatted to fit the LSTM structure, treating each sample as a one-step sequence. The model consisted of an LSTM layer with 50 units and a Dense output layer, using Mean Squared Error (MSE) as the loss function. Training was conducted for 50 epochs with a batch size of 32, and 20% of the data was reserved for validation. Anomalies were identified based on reconstruction error, with a threshold set at the 95th percentile. The LSTM model detected 47 anomalies, accounting for 4.7.

Figure 13
3D Visualization of Student Performance with Anomalies Detected by RNN

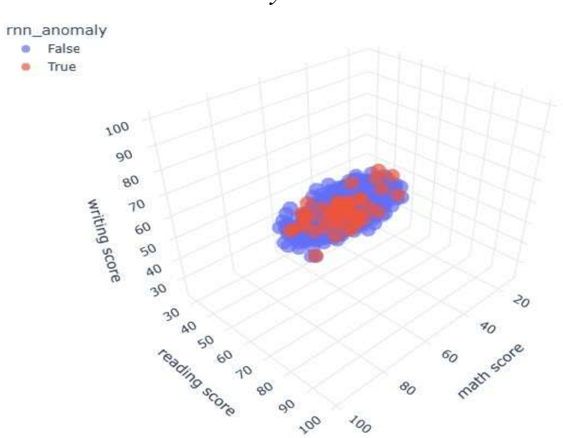
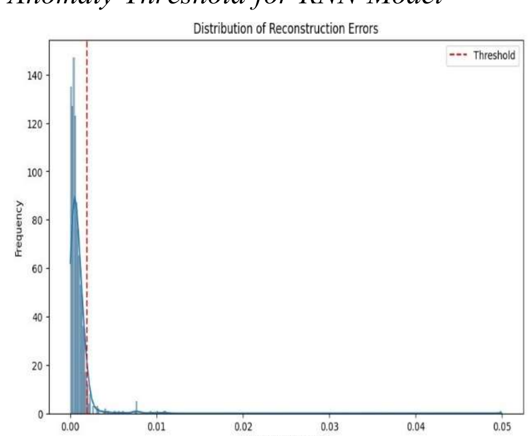


Figure 14
Distribution of Reconstruction Errors with Anomaly Threshold for RNN Model



INVERGE JOURNAL OF SOCIAL SCIENCES
https://invergejournals.com/
ISSN (Online): 2959-4359, ISSN (Print): 3007-2018
Volume 4 Issue 1, 2025

The detected anomalies are visually represented in the 3D scatter plot (Figure 13), where they deviate from the normal distribution of student performance scores. Additionally, the reconstruction error distribution (Figure 14) confirms that most data points exhibit low errors, while anomalies show significantly higher error values. This approach highlights the effectiveness of LSTM-based RNNs in identifying atypical student performance, making it a valuable tool for outlier analysis and detecting potential data quality issues.

Anomalies and Outliers Detection Using Deep Neural Networks

Deep Neural Network used for detecting anomalies in the student performance data set containing math, reading and writing scores. First, data preprocessing is done, and the features are normalized using Standard Scaler because neural networks require features of the same range. DNN had an input layer of 128 neurons, two hidden layers of 64 and 32 neurons and the output layer the activation function being sigmoid. The data is learnt for 50 epochs using mean squared error and the adam optimizer was used as the optimization algorithm.

The reconstruction errors (Figure 16) computed for each data point and what exceeded the 95th percentile of the distribution is considered an outlier. Those pieces of data that had reconstruction errors higher than this value where labeled as an anomaly.To compare the differences, a 3D scatter plot [Figure 15] is created to present the data, and the anomalies were highlighted by coloration for convenience. Threshold at the 95th percentile of the reconstruction error distribution. Data points with reconstruction errors above this threshold are classified as anomalies. A 3D scatter plot (Figure 15) used to visualize the results, with anomalies color-coded for easy identification. It can be identified that the model is able to pin point 49 of the cases, on an overall of 1000 as the database size, which is just 4.9%. This approach also showed how the DNN could easily find outliers, indicative of some forms of student performance that may require closer inspection.

Comparison of Anomaly Detection Algorithms on Student Exam Data

The comparison of the three anomaly detection algorithms Autoencoders, RNN with LSTM, and DNN it shows how that they have potential to improve educational practices and policy making. These models can drive detection of anomalous performance in student data so as to inform instructors of students who may need special monitoring or support. The insights gained from these outlier behaviors could indicate academic issues, personal hardships, or even data errors. Out of the 1000 anomalies known, the Autoencoder was only able to find 50 (5%) and the RNN with (LSTM) managed to find 47 (4.7%). Yet, DNN found anomalies (49, or 4.9%) since its tendency to detect anomalies when there are differing data densities.

Figure 15
3D Visualization of Student Performance with Anomalies Detected by DNN

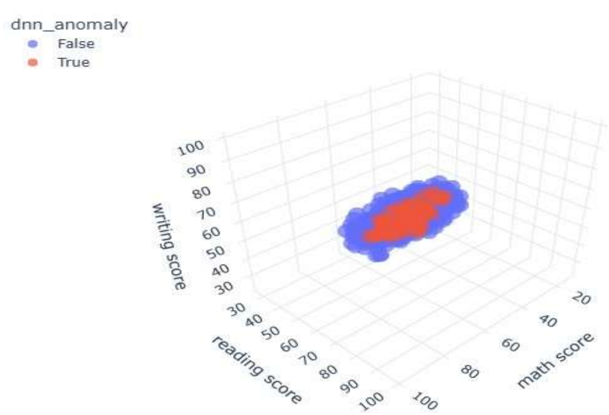
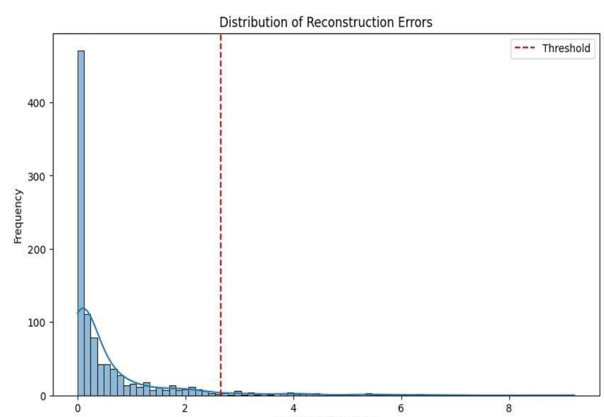


Figure 16
Distribution of Reconstruction Errors with Anomaly Threshold for DNN Model



The findings presented in Table 3 show that machine learning and deep learning techniques can effectively detect anomalies, which can directly benefit educational institutions, policy makers, and parents.

Table 3

Comparison of Anomaly Detection Performance across Models

Model	Strengths	Limitations	Anomalies Detected	Performance Note
Autoencoders	Effective for high-dimensional data; flexible	Requires large datasets; sensitive to architecture	5%	Excellent for complex data patterns
RNN (LSTM)	Suited for sequential data; detects non-linear anomalies well	Requires large datasets and computational resources	4.7%	Effective for sequential anomaly detection
Deep Neural Networks (DNN)	Flexible; effective for complex patterns	Requires large datasets and architecture tuning	4.9%	Good for general anomaly detection

Ability to detect anomalous performance provides educational institutions with a way to direct inter- ventions and develop tailored educational plan for students who are potentially lacking in class or demon- strating unusual learning patterns. Deep learning models such as Autoencoders, RNN with LSTM, DNN can leverage the strengths to detect subtle and non-linear relationship in performance data as compared to traditional methods like KMeans, Isolation Forest. By understanding this, learning professionals can devise more targeted, data informed support of students, with the aim of improving academic outcomes. This research study can help policy-makers to reach evidence-based decisions within the sphere of educational policy. General findings might reveal possible shortcomings attributed to a specific gender, race, ethnic group, or learners with learning disabilities, course delivery or assessment gaps, or data validity and reliability issues. This can be used to frame some potential implications of these findings, namely in relations to areas that students could be potentially struggling, thus using additional targeted funding or program changes.

For parents, anomaly detection can offer signs of academic distress that a child may be pulling in school. Through the use of performance comparison, parents will be able to take prompt actions such as getting a child to seek tutor's help or seek counseling.

The findings of all the presented anomaly detection algorithms highlight the potential of applying the state of the art machine learning approaches in education. Even though deep learning models demand more computational power, they bring ample value in identifying intricate patterns in students' data that can serve as the basis for intervention. While it might be more efficient for traditional methods, some behaviors that deep learning models can easily find in the transactions might not be easily captured by traditional methods. Hence, using these tools can bring about a convergence of efforts of educational institutions, policy makers, and parents in order to have a better responsive education system.

Conclusion

In conclusion, this study highlights the effectiveness of deep learning techniques, Autoencoders, RNN with LSTM, and DNN, in detecting anomalies in student performance data. These models successfully identified outliers, providing valuable insights into academic challenges and data inconsistencies. The findings emphasize the potential of deep learning in enhancing educational analysis, enabling data-driven interventions to support student success.

The comparison of these algorithms demonstrate that although the basic approaches as the deep learning models like Autoencoders, RNN with LSTM or DNN provide higher level of detail in regarding the anomalies and consider nonlinearity in data. All the algorithms discussed in this paper including Autoencoder, RNN with LSTM, and DNN flagged more anomaly cases that are sensitive to differences in data densities.

These anomaly detection techniques offer a better understanding of undesired behavior patterns that should benefit educational institutions, policy makers, and parents. Using such tools, educational stakeholders are able to notice early signs of students in need of help, provide corrective measures that respond to the needs of students and cover for whatever that may be hindering the students from performing optimally. In addition, the policy makers should be able to design the education policy from the findings credited to research to support the policy makers to be in a position to allocate the resource in the right areas. To the parents, finding own performance anomalies provide the first chance to come to the assistance of children and help them to overcome academic difficulties.

Finally, this research highlights the relevance of using higher levels of ML and DL technologies in the systemized education field. Not only do these techniques improve the comprehension of student achievements but also create the basis for relevant and efficient approaches. The future work is to extend the current models, make the new parametric features, and include labeled data to enhance the detection rate that can be used effectively in organizations and educational facilities. This means that the future development of the machine learning models in the context of education will allow creating more supportive environment for students to perform in school, college or university and respond individually to all students.

References

Abatal, A., Korchi, A., Mzili, M., Mzili, T., Khalouki, H., & Billah, M. E. (2025). A comprehensive evaluation of machine learning techniques for forecasting student academic success. *Journal of Electronics, Electromedical Engineering, and Medical Informatics, 7*(1), 1-2.

Acharya, A., & Sinha, D. (2014). Early prediction of students performance using machine learning techniques. *International Journal of Computer Applications, 107*(1), 37-43.

Al-Fairouz, E. I., & Al-Hagery, M. A. (2020). Students performance: From detection of failures and anomaly cases to the solutions-based mining algorithms. *International Journal of Engineering Research and Technology, 13*(10), 2895-2905.

Alam, A., & Mohanty, A. (2022). Predicting students' performance employing educational data mining techniques, machine learning, and learning analytics. In *International Conference on Communication, Networks and Computing* (pp. 166-177). Springer.

Alruwais, N., & Zakariah, M. (2023). Student-engagement detection in classroom using machine learning algorithm. *Electronics, 12*(3), 731.

Bulusu, S., Kailkhura, B., Li, B., Varshney, P. K., & Song, D. (2020). Anomalous example detection in deep learning: A survey. *IEEE Access, 8*, 132330-132347.

Gao, Y. (2025). Deep learning-based strategies for evaluating and enhancing university teaching quality. *Computers and Education: Artificial Intelligence, 7*, 100362.

Ghanim, J., & Awad, M. (2025). An unsupervised anomaly detection in electricity consumption using reinforcement learning and time series forest-based framework. *Journal of Artificial Intelligence and Soft Computing Research, 15*(1), 5-24.

Huang, A. Y., Lu, O. H., Huang, J. C., Yin, C. J., & Yang, S. J. (2020). Predicting students' academic performance by using educational big data and learning analytics: Evaluation of classification methods and learning logs. *Interactive Learning Environments, 28*(2), 206-230.

Hussain, S., & Khan, M. Q. (2023). Student-performulator: Predicting students' academic performance at secondary and intermediate level using machine learning. *Annals of Data Science, 10*(3), 637-655.

Hussain, S., Gaftandzhieva, S., Maniruzzaman, M., Doneva, R., & Muhsin, Z. F. (2021). Regression analysis of student academic performance using deep learning. *Education and Information Technologies, 26*(1), 783-798.

Issah, I., Appiah, O., Appiahene, P., & Inusah, F. (2023). A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. *Decision*

Analytics Journal, 5, 100204.

Kaggle. (n.d.). *Students Performance in Exams*. Retrieved from https://www.kaggle.com/datasets/spscientist/students-performance-in-exams/data

Kamalov, F., Sulieman, H., & Santandreu Calonge, D. (2021). Machine learning based approach to exam cheating detection. *PLOS ONE, 16*(8), e0254340.

López-García, A., Blasco-Blasco, O., Liern-García, M., & Parada-Rico, S. E. (2023). Early detection of students' failure using machine learning techniques. *Operations Research Perspectives, 11*, 100292.

Nassif, A. B., Talib, M. A., Nasir, Q., & Dakalbab, F. M. (2021). Machine learning for anomaly detection: A systematic review. *IEEE Access, 9*, 78658-78700.

Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., & Phasinam, K. (2023). Classification and prediction of student performance data using various machine learning algorithms. *Materials Today: Proceedings, 80*, 3782-3785.

Pek, R. Z., Özyer, S. T., Elhage, T., Özyer, T., & Alhajj, R. (2022). The role of machine learning in identifying students at-risk and minimizing failure. *IEEE Access, 11*, 1224-1243.

Riestra-González, M., del Puerto Paule-Ruíz, M., & Ortin, F. (2021). Massive LMS log data analysis for the early prediction of course-agnostic student performance. *Computers & Education, 163*, 104108.

Shitaya, A. M., Wahed, M. E., Ismail, A., Shams, M. Y., & Salama, A. A. (2025). Predicting student behavior using a neutrosophic deep learning model. *Neutrosophic Sets and Systems, 76*, 288-310.

Vaidya, A., & Sharma, S. (2024). Anomaly detection in the course evaluation process: A learning analytics–based approach. *Interactive Technology and Smart Education, 21*(1), 168-187.

Wang, G., Han, S., Ding, E., & Huang, D. (2021). Student-teacher feature pyramid matching for anomaly detection. *arXiv preprint arXiv:2103.04257*.