International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Advancing Student Learning Assessment: A Novel Hybrid Neural Network Approach Integrating GRU and LSTM Architectures.

Taneja Sanjay Devkishan[*1], Dr. Sanjay Kumar Singh[2], Dr. Ajay Kumar Bharti[3]

Submitted: 28/12/2023 **Revised**: 04/02/2024 **Accepted**: 12/02/2024

Abstract: Timely and accurate assessment of student learning is essential for the effective functioning of educational institutions, guiding program development and instructional strategies. This study introduces an innovative method employing a Hybrid Neural Network that integrates Gated Recurrent Units (GRU) and Long Short-Term Memories (LSTM) architectures. This fusion forms a unique Hybrid Neural Network, capitalizing on the distinctive features of GRU and LSTM to enhance the reliability and predictability of student assessments. The research not only propels the field of student performance evaluation but also unveils a ground-breaking application for the LSTM-GRU Hybrid Neural Network design. Performance metrics, such as Mean Squared Error (MSE) and Loss, were meticulously analyzed. The Hybrid Neural Network demonstrated superior performance, boasting an MSE of **0.236** and a Validation MSE of **0.285**. Furthermore, the Loss was **0.236**, and the Validation Loss was **0.285**. Comparative evaluations against conventional LSTM and GRU models underscored the significant performance enhancements achieved by the Hybrid Neural Network. In evaluating the effectiveness of past approaches, our study unveiled that the proposed Hybrid Neural Network consistently outperformed in terms of MSE, suggesting unparalleled performance compared to existing studies. This research contributes significantly to the evolving landscape of educational assessment methodologies, highlighting the transformative potential of Hybrid Neural Networks for elevating evaluation accuracy and predictive capabilities.
Our findings advocate for the widespread adoption of this innovative approach in educational institutions, paving the way for improved student assessment mechanisms. As institutions strive for enhanced learning outcomes, the proposed Hybrid Neural Network stands as a beacon for advancing the frontier of educational assessment technology.

Keywords: *Deep Learning, LSTM, GRU, Hybrid Neural Network, Student Performance Assessment, Exploratory Data Analysis*

1. Introduction

There is a direct correlation between students' mental health and their academic success. The quality of schooling can largely be gauged by students' performance in the classroom. Academically underperforming students have been shown in related studies to be at significantly higher risk of anxiety, depression, and suicide than their high performing counterparts. The goal of achievement prediction is to identify children who are at high academic risk in advance. This serves to remind administrators, teachers, and students themselves of the need to take timely targeted intervention activities to avoid poor performance, which can include failing courses, quitting out, staying out, and other similar behaviors.

As a result, there has been a significant amount of focus and investigation directed on predicting student accomplishment. Assessment and evaluation of student performance by means of deep learning have recently emerged as a potentially fruitful strategy in the field of education. Various forms of deep learning, including neural networks, are being implemented to predict timely students' academic success based on a variety of criteria. This includes Prior grades, attendance, conduct, and social interactions [1]. This has made it possible to provide early intervention for students who are at danger of falling behind in their studies. In addition, the marking process for specific types of assignments and tests has been automated with the use of deep learning algorithms, which has allowed teachers to give students feedback that is both prompt and reliable. Educators are now able to obtain insights into the learning patterns and behaviors of their students thanks to educational data mining and learning analytics, which are powered by deep learning. This enables educators to make decisions that are data-driven, which improves the learning experience [2]. Deep learning is an important component of Intelligent Tutoring Systems (ITS), which help tailor each student's educational experience to meet their unique needs. Through ongoing analysis of a student's interactions and success Intelligent Tutoring are able to modify its instructional strategy to better meet the requirements of each particular learner. In addition, Natural Language

[1] Research Scholar, Amity Institute of Information Technology,
Amity University Uttar Pradesh Lucknow Campus, Lucknow, India,
Email ID: sdtaneja@lko.amity.edu, ORCID - https://orcid.org/0000-0001-8404-9386
[2] Professor, Amity Institute of Information Technology,
Amity University Uttar Pradesh Lucknow Campus, Lucknow, India,
Email ID: sksingh1@amity.edu, ORCID - https://orcid.org/0000-0002-9400-6715
[3] Professor, Department of Computer Science,
Babu Sunder Singh Institute of Technology and Management, Lucknow, India,
Email ID: ajay_bharti@hotmail.com, ORCID - https://orcid.org/0000-0001-6879-5151
**Corresponding/Main Author Email ID: sdtaneja@lko.amity.edu*

Processing (NLP) approaches that make use of deep learning have been implemented for the purpose of automating the scoring of student essays by evaluating the substance, coherence, and language usage in the essays [3]. Deep learning has also been utilized by Massive Open Online Courses (MOOCs) for the purpose of performance analysis. This has made it possible for platforms to evaluate the efficacy of course content, recognize areas in need of development, and optimize the learning experience on a huge scale. In addition, deep learning has been applied to the analysis of sentiment and emotions, which provides insights into the emotional states of students while they are studying. This has led to the development of emotionally intelligent learning environments that promote students' well-being and engagement. In spite of these breakthroughs, it is vital to examine the consequences of data privacy and ethics when putting deep learning models into practice in educational settings [4]. For the sake of protecting student information and ensuring that data security is not compromised, it is of the utmost importance to ensure responsible data collection and analysis.

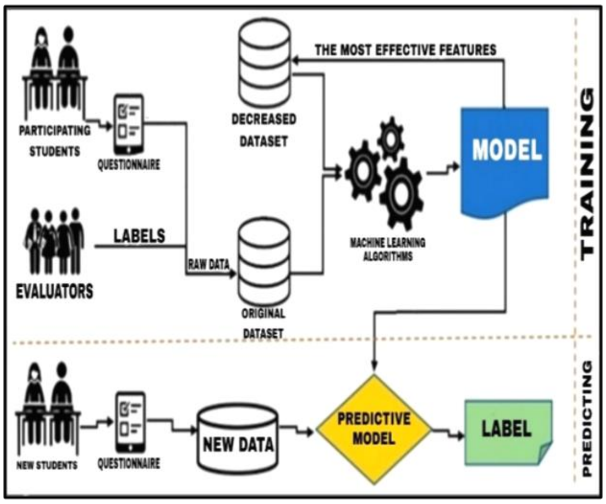


Fig. 1. Student Performance Evaluation

Exploring the most recent developments and breakthroughs in the IEEE domain and other academic sources is still essential for further improving student performance assessment and evaluation utilizing deep learning approaches, since research in this field continues to move at a rapid pace. An effective method for evaluating academic accomplishments is to conduct student performance evaluations using a hybrid neural network. This type of network combines Networks with Gated Recurrent Units (GRUs) and Long Short-Term Memories (LSTMs).

The hybrid network is able to identify deep temporal relationships and patterns in student data, such as study habits and test scores, by capitalizing on the strengths of both LSTM and GRU. The memory retention capabilities of the LSTM and the gating mechanisms of the GRU work together to make it possible for the model to successfully capture both short-term and long-term dynamics. This allows for more accurate forecasts of student performance as well as insights into the factors that influence the educational results of students. This novel technique improves the quality of the assessment process and gives educators greater insights that may be adapted into support and intervention measures for individual students [5].

The novel approach of using a Hybrid Neural Network architecture, notably integrating Components such as Gated Recurrent Units and Long Short-Term Memories, for student performance assessment improves accuracy and predictability in educational evaluations. This unique evaluation method capitalizes on LSTM and GRU's strengths to capture subtle temporal patterns and dependencies in student data. This unique methodology not only predicts student results more accurately but also proposes a new paradigm for using neural network architectures in educational evaluation. This unique technique promises to improve student performance assessment by delivering a fresh perspective and a more robust predictive model.

2. Literature Review

Niu 2022 et al. the use of a C# script and prefix system might enhance practical fake action proofs by constructing a virtual scene and importing Unity3D motors. Later, finish the cognitive aid module by employing an authorization technique to address the problem of arranging permissions in scenes involving several entities. Our system has the ability to anticipate text, video, and some flexible functions, and it may grant individualized rights to different pupils. Our solution is modular, consisting of separate Spring Cloud components. Maximize the system's design by relying even more on Redis. The system's clear virtual auditoriums can be easily used in government-supervised chemistry classrooms. In addition to improving learning efficiency and expanding access, it may successfully address authority challenges in real-world contexts. In this vein, share a suite of clever classroom behavior systems that leverage deep learning and other cutting-edge forms of education technology. It's a great addition to any classroom's perception ministry! Using deep vision techniques including face perception and expression analysis, the system can perform status monitoring and classroom discussion services at peak efficiency.

Extensive testing findings have proven our technique to be competitive in performance [6]. Kamal 2022 et al. provides a strategy for identifying and predicting student performance based on metaheuristics and machine learning. Initially, features are selected using a relief algorithm. Machine learning classifiers such as Back

Propagation Neural Network (BPNN), Random Forest (RF), and Naïve Bayes (NB) are used to make sense from student performance data. Classification and predicting students' academic success are two areas where BPNN excels [7]. Tan 2022 et al. researches how the deep learning mobile terminal's facial picture evaluation algorithm can be used for managing student check-in. For the purpose of streamlining the student check-in process, a deep learning network is employed to recognize faces in images. Researchers develop a cascaded convolution network-based face key point identification method and a deep learning network-based face detection algorithm that relies on candidate regions to identify faces. To improve the ineffectiveness of face identification and detection, present an extraction approach for face binary characteristics and conduct tests to evaluate its efficacy. Simulation results demonstrate that enhanced deep learning network-based face identification can reduce retrieval time and boost face picture classification precision [8]. Alsariera 2022 et al. in this work, used pre-existing Machine Learning (ML) methods and essential criteria for forecasting student success. Through a thorough search of multiple online databases, find papers published among 2015 and 2021 that are relevant. The researchers looked at data from 39 studies. Conclusions Decision trees, artificial neural networks, support vector machines, kernel neural networks, linear regressions, or Naive Bayes classifiers were the most common ML models used. Our results also demonstrated that Artificial Neural Network (ANN) is superior to alternative models in terms of efficiency and precision. The majority of the input (e.g., predictive features) used to foretell a student's success was based on academic, demographic, internal evaluation, and family/personal attributes. Our research shows that this field is attracting increasing attention and a wide variety of machine learning techniques. In addition, there is mounting evidence that ML can be used to target and improve specific aspects of academic performance [9]. Zhang 2022 et al. traditional methods for evaluating ideological or political theory courses make it hard to resolve the contradiction among the content for these courses or the ideological knowledge for the educated, and this problem has become increasingly prominent as schools try to gauge students' ideological awareness within test scores. Using deep learning technology to create a new mechanism to evaluate ideological or political education in higher education can improve the quality of both formative and summative assessments [10].

Table 1. Literature Summary

Author / Year	***Method***	***Performance Matrices***	***References***
Waldeyer / 2022	Diagonally Weighted Least Square	**RMSE= 0.3, TLI=0.96, CFI=0.97**	[11]
Mai / 2022	Louvain Method	**Accuracy =74%, F1-Score =73.5%.**	[12]
Alshabandar / 2020	Random Forest (RF)	**RMSE= 8.1**	[13]
Pereira / 2019	Evolutionary Algorithm	**Accuracy =75.55%**	[14]
Rimadana / 2019	Time Structure Questionnaire (TSQ)	**Accuracy =84%**	[15]
Abu Zohair / 2019	Multiple Machine Learning Algorithms	**Accuracy =76.3%**	[16]

3. Proposed Methodology

Discuss the proposed technique in this section, which includes information about the dataset, data pre-processing, and class balancing as a vital step for avoiding bias and maximizing the model's ability to learn from a variety of situations. Label encoding turns category variables into numerical representations, which makes it easier to understand a model and to execute computations on it. Exploratory Data Analysis (EDA), which is used for data visualization and statistical analysis, and the hybrid neural network, that was presented for modelling is going to be employed for the prediction and evaluation of student performance depicts in figure 2. The novelty of employing a Hybrid Neural Network architecture, specifically integrating Components such as Gated Recurrent Units and Long Short-Term Memories, for student performance assessment lies in its pioneering approach to predictability in educational evaluations. By synergistically harnessing the strengths of both LSTM and GRU, this innovative methodology transcends traditional assessment techniques, capitalizing on their respective abilities to capture intricate temporal patterns and dependencies within student data. This amalgamation presents a unique framework that not only ensures a more reliable prediction of student outcomes but also introduces a novel paradigm for leveraging neural network architectures in the realm of educational evaluation. The emergent novelty of this approach promises to revolutionize student performance assessment by offering a fresh perspective and a more robust predictive model.

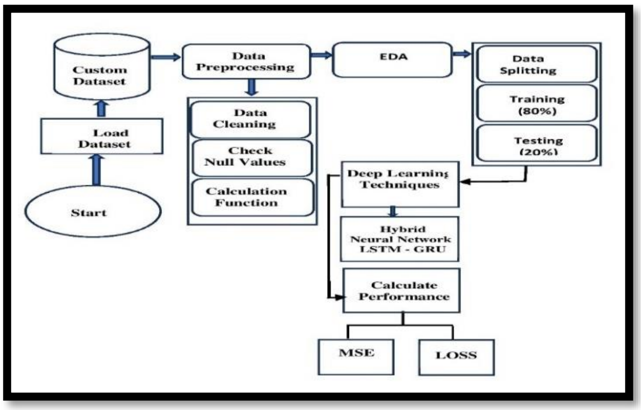


Fig. 2. Proposed Flowchart

a) Data Collection

There are 21 different categories here, such as "Program of Study," "Aggregate% Marks in Class-X," "Aggregate% Marks in Class-XII," and so on. In addition to "CGPA", "SGPA", "Highest SGPA" and "Student Behavior," other relevant variables are also provided. Student mark sheets and student conduct were used to compile this data, which includes a wide range of performance indicators. In the application process, you were asked questions like "Self-Study per day (In Hours)", "Time spent in extracurricular activities (In Hours)", "Do you regularly access virtual learning platforms", "Do you smoke", "Do you drink", "Are you exposed to social media", and "Do you have regular access to virtual learning platforms". A large family size, both parents working, annual family income in lacs, and current health status are all necessities. The data contains these fields, which can be used to evaluate students' progress.

b) Pre-processing

By enhancing the readiness of data for analysis, pre-processing greatly contributes to the quality and effectiveness of machine learning models. As such, class balancing is a crucial process for minimizing bias and maximizing the model's ability to learn from a wide range of situations. To make a model easier to understand and to perform computations on, label encoding converts category variables into numerical representations. The model can then prioritize features in accordance with their relative relevance, as determined by the weights assigned to each column. The model can be tuned to the importance of various qualities through the use of weights, leading to more precise predictions. The evaluation process can be simplified by combining metrics like MSE Loss into a single computation performed by a general-purpose function. This calculation function can then be used iteratively on individual rows, permitting in-depth examination of the role that each data point played in the model's final performance. This fine-grained assessment reveals patterns, strengths, and shortcomings that add to an overall comprehension of the model's behavior. Pre-processing takes raw data and puts it into a refined and informative form that machine learning models may use to make accurate predictions and judgments.

c) Exploratory Data Analysis (EDA)

EDA, or "Exploratory Data Analysis," is when a student performance record is analyzed in a methodical manner in order to expose the dataset's intrinsic patterns, trends, or anomalies. EDA is able to find critical insights about variables such as study time, prior test scores, attendance, and other aspects through the use of data visualization tools and statistical metrics. Because of this approach, we are able to recognize possible correlations within the data, as well as distribution features and possible outliers. EDA lays the groundwork for informed decision-making by directing the creation of predictive models such hybrid neural networks, such as LSTM-GRU architectures, to reliably assess and estimate student performance based on significant data-driven observations. In other words, EDA provides the basis for informed decision-making.

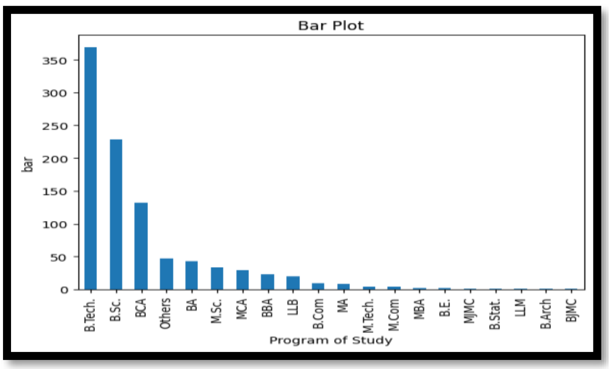


Fig. 3. Bar Plot of Program of Study

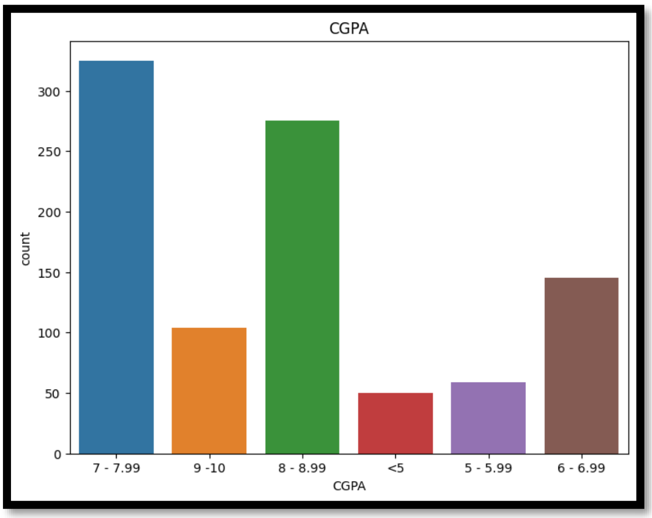


Fig. 4. Bar Plot of Program of CGPA

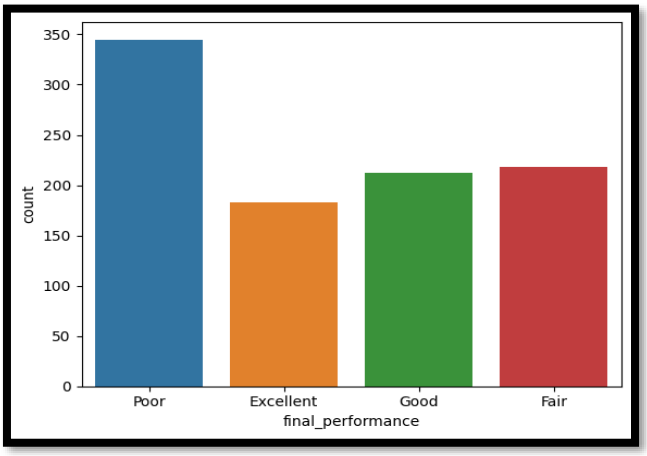
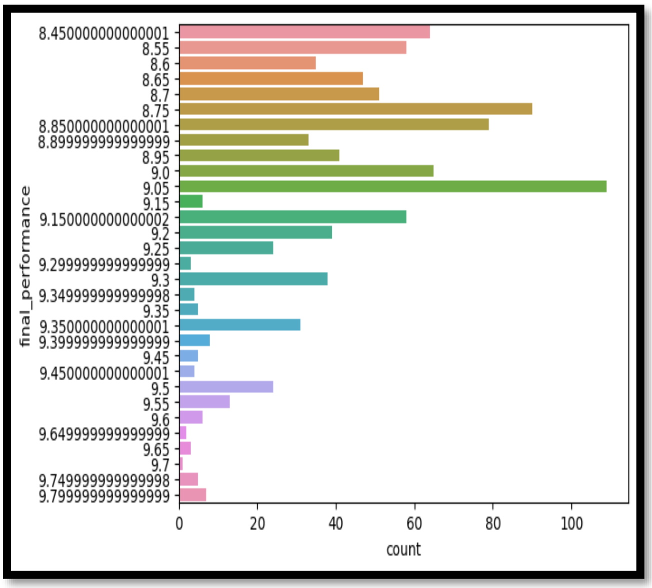


Fig. 5. Bar Plot of Final Performance



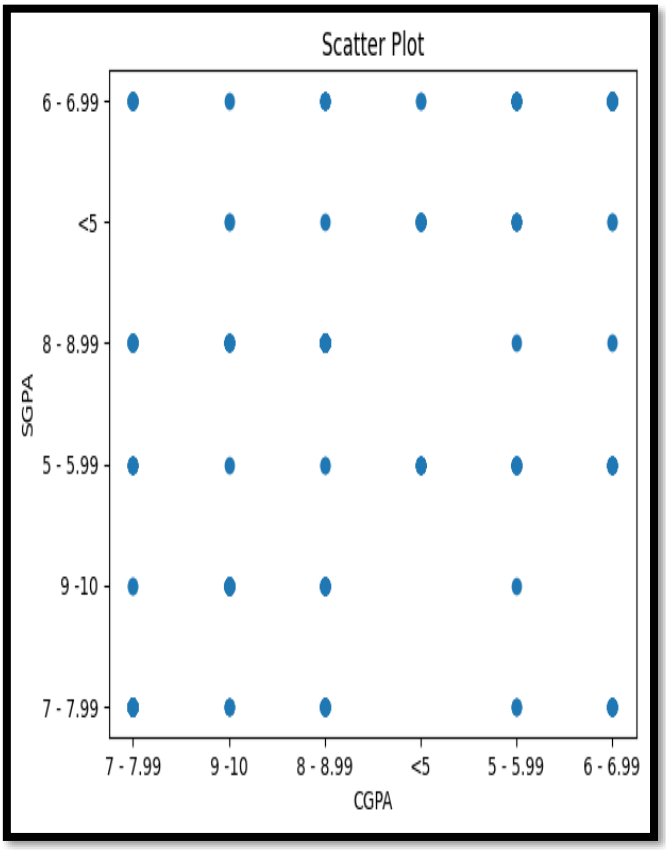


Fig. 7. Scatter Plot of CGPA

Figure-3 through Figure-8 present the visual representations generated through Exploratory Data Analysis (EDA). These include a Bar plot illustrating the distribution of programs of study, Scatter plots showcasing the relationship with Cumulative Grade Point Average (CGPA), Violin plots depicting the distribution of smoke-related information, Boxplots illustrating the distribution of the number of siblings, Count plots visualizing the distribution of final performance categories, Bar plots displaying the distribution of final performance, and a Correlation Matrix representing the interrelationships among the analyzed variables.

d) Deep Learning & Modelling

Some popular types of deep learning models are the Long Short-Term Memory, Gated Recurrent Unit and advanced neural network designs that are designed to handle sequential and time-series data. They have demonstrated remarkable efficacy in a variety of tasks, including speech recognition, natural language processing, and sequential data analysis, amongst others. Both LSTM and GRU incorporate memory cells and gating mechanisms into their architecture, which enables them to recognize long-term dependencies in sequences. This allows them to circumvent the vanishing gradient problem. LSTMs provide an extended memory capability by retaining information pertinent to long sequences by utilizing cell states to accomplish this. They are made up of input gates, forget gates, and output gates, and their purpose is to control the flow of information into and out of memory cells. This gives memory cells the ability to efficiently learn and recall temporal patterns. GRUs, a subtype of LSTMs, make use of a reduced number of gating mechanisms, which helps to simplify the design while keeping the same level of performance.

Fig. 6. Count Plot of Final Performance

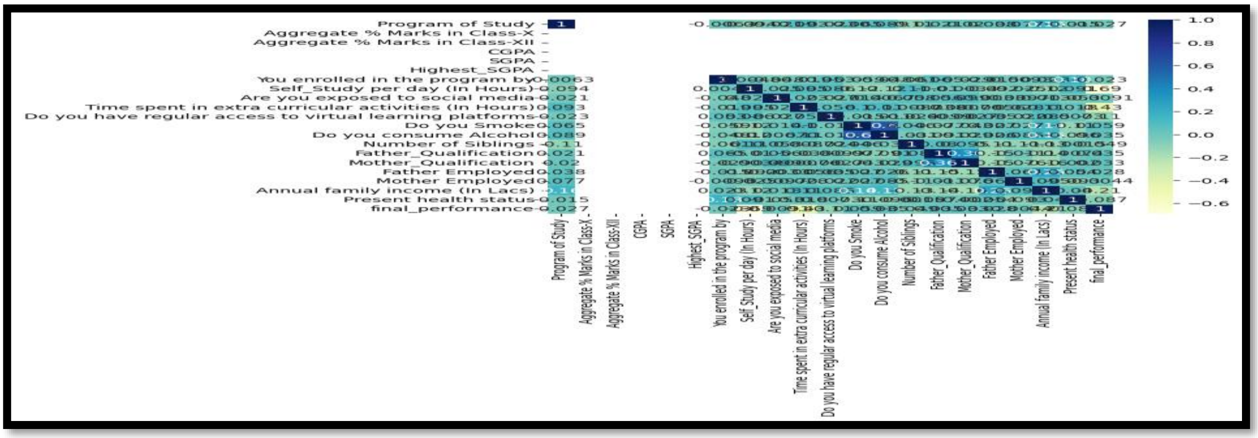


Fig. 8. Correlation Matrix

Both LSTM and GRU models are excellent in modelling sequences, which makes them excellent options for tasks such as time-series forecasting, sentiment analysis, and evaluating student achievement. They are equipped to handle complicated patterns in sequential data as a result of their ability to record and remember long-range dependencies, which in turn enables them to make more accurate predictions and get useful insights from the data. LSTM Equations given are below:

LSTM Equations-

$$\boldsymbol{i_t = \sigma(x_t U^I + h_{t-1} W^i)} \quad (1)$$

$$\boldsymbol{f_t = \sigma(x_t U^f + h_{t-1} W^f)} \quad (2)$$

$$\boldsymbol{o_t = \sigma(x_t U^o + h_{t-1} W^o)} \quad (3)$$

$$\boldsymbol{\widetilde{C}_t = tanh\,(x_t U^g + h_{t-1} W^g)} \quad (4)$$

$$\boldsymbol{C_t = \sigma(f_t * C_{t-1} + i_t * \widetilde{C}_t} \quad (5)$$

$$\boldsymbol{h_t = tanh\, tanh\,(C_t) \; * o_t} \quad (6)$$

GRU Equations-

$$\boldsymbol{z_t = \sigma(W_z \cdot [h_{t-1}, x_t])} \quad (7)$$

$$\boldsymbol{r_t = \sigma(W_r \cdot [h_{t-1}, x_t])} \quad (8)$$

$$\boldsymbol{\widetilde{h}_t = tanh\,(W.[r_t * h_{t-1}, x_t])} \quad (9)$$

$$\boldsymbol{h_t = (1 - z_t) * h_{t-1} + z_t * \widetilde{h}_t} \quad (10)$$

Figure 10 depicts the training log of LSTM model across 100 epochs, with tracked loss and mean squared error (MSE) metrics. The graph illustrates a consistent decrease in both training and validation losses, signifying progressive performance enhancement. The concluding epochs reveal a converged model, achieving a low training loss of 1.0730 and a matching validation loss of 1.0997. Figure 16 depicts training log illustrates the progression of GRU model across 100 epochs. Both training and validation losses, along with mean squared error (MSE) metrics, are monitored. The model demonstrates a consistent decline in losses, indicative of improved performance. The final epochs exhibit a well-converged model with minimal training loss (0.4179) and corresponding validation loss (0.5424).

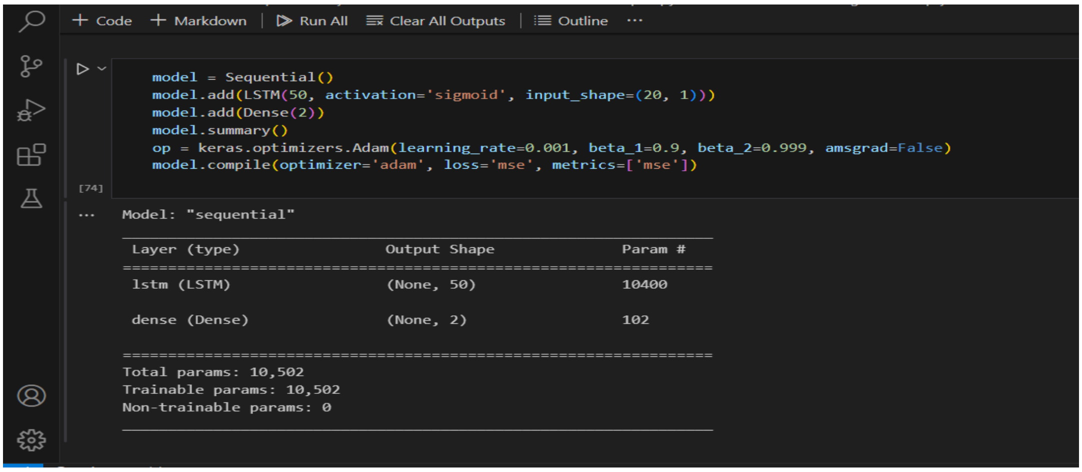


Fig. 9. Python Script for creating the LSTM Model with two LSTM layers and Summary of LSTM Model created with two LSTM layers

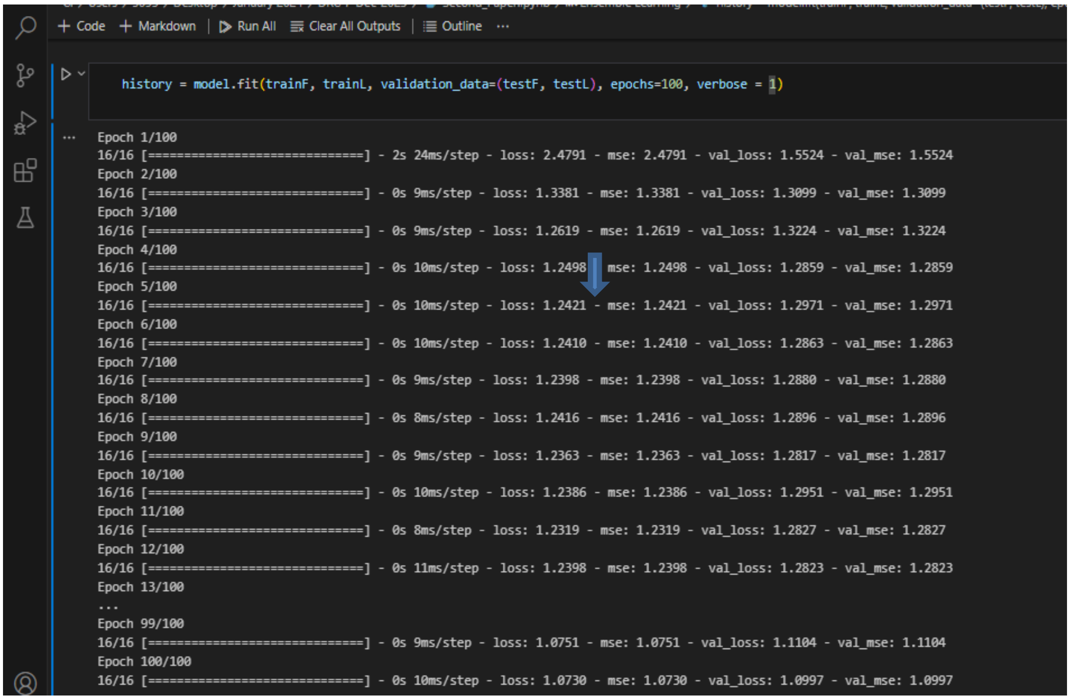


Fig. 10. Training log for a LSTM model showing over 100 epochs

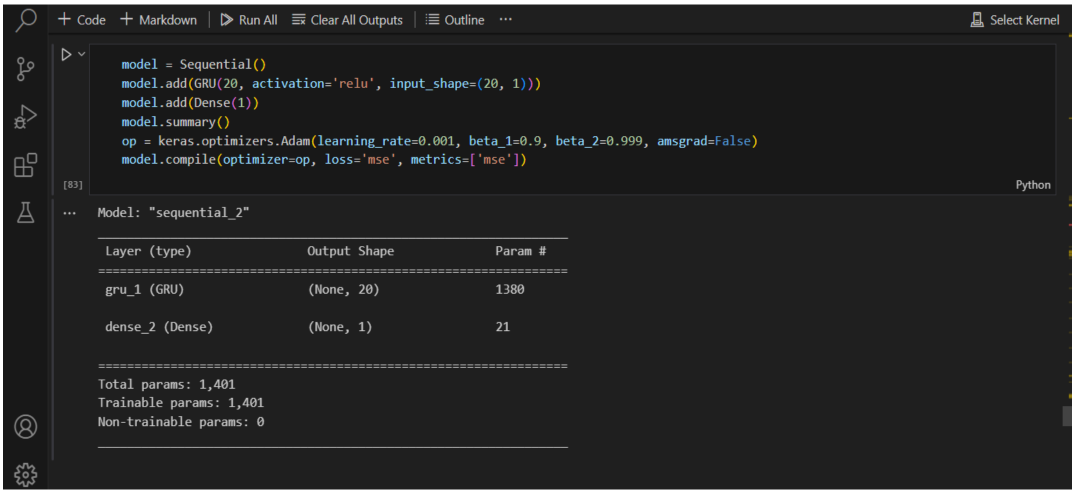


Fig. 11. Python Script for creating the GRU Model with one GRU layers and Summary of GRU Model created with one GRU layers

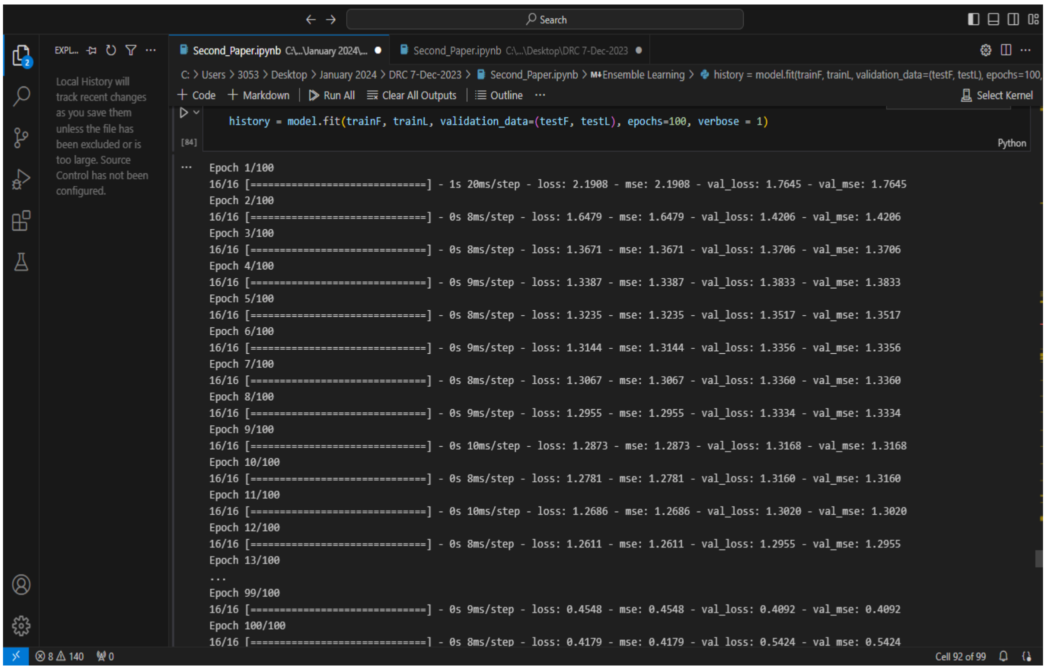


Fig. 12. Training log illustrates the progression of a GRU model across 100 epochs

- **Hybrid Neural Network**

In order to improve sequence modelling and prediction, a hybrid neural network that blends Long-term storage and recurrent units that are gated architectures can integrate the strengths of both models. This novel strategy takes advantage of LSTM's capability of capturing long-term dependencies as well as GRU's efficiency in terms of computer processing. The hybrid network optimally learns complicated patterns within sequential data, such as student performance records, as a result of the fusion of various architectures, which enables accurate and nuanced predictions while simultaneously decreasing the complexity of the computational process. This synergy between LSTM and GRU enables the model to provide substantial insights into student outcomes and assists educators in making informed decisions to promote personalized learning paths. LSTM and GRU are both recurrent neural networks.

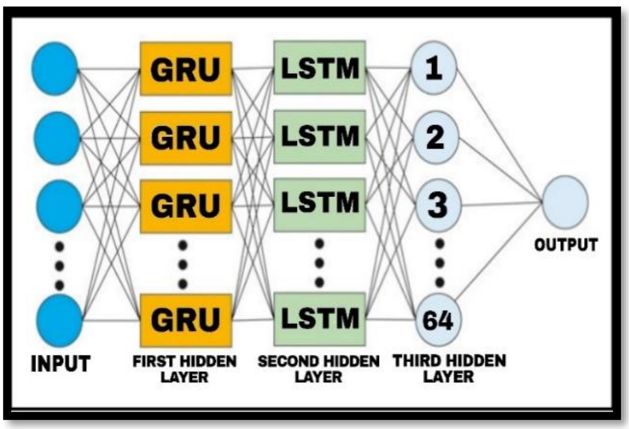


Fig. 13. Hybrid Neural Network Architecture

4. Result & Discussion

A wide range of performance assessment criteria will be utilized in order to make an overall assessment of the performance of a Hybrid Neural network. The performance graph, which is available on this project, is used to compute these metric values, and the graph may be accessed here. Calculations will be done to determine Measures of performance like Mean Squared Error and Loss will be applied to determine the model's efficiency.

a) MSE (Mean Square Error)

The MSE is currently the most widely used metric that is used to evaluate the quality of photographs. Because it serves as an all-encompassing standard, it is to one's advantage for the number to be as close as possible to 0. To communicate the approximate monetary value of this squared error loss, we can make use of a risk function that is denoted by the notation MSE. The mean squared error (MSE) is frequently found to be non-zero, as opposed to zero, either because of random factors or because the estimator chose to ignore information that could have led to a more accurate estimate.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} \blacksquare \, (Y_i - \hat{Y}_i)^2 \qquad (11)$$

b) Loss

A loss is incurred due to an inaccurate projection. When considered in these other situations, loss is a measurement of the prediction error associated with the model for a particular case. If a model's prediction is incorrect, then losses will be bigger; otherwise, there will be no losses. Models need to be trained in order to discover a collection of biases or weights that have a low average loss.

$$Loss = -\frac{1}{m}\sum_{i=1}^{m} \quad Yi.lo\,g(Yi) \qquad (12)$$

Figure 15 depicts the training log details the progress of a Hybrid Model (LSTM+GRU) model across 100 epochs. The notation "16/16 [========]" suggests batches processed during training. The model exhibits a consistent decrease in both training and validation losses, highlighting its learning efficacy. In the final epochs, the well-converged model achieves a low training loss of 0.2363 and a corresponding validation loss of 0.2855.

Table-3 details the Mean Squared Error (MSE) values for each model and offers this information as part of the Performance Evaluation of the previously developed techniques.

Notably, the MSE was found to be the lowest within the context of the proposed work, which indicates that it performed better than the studies that were listed earlier in the paragraph.

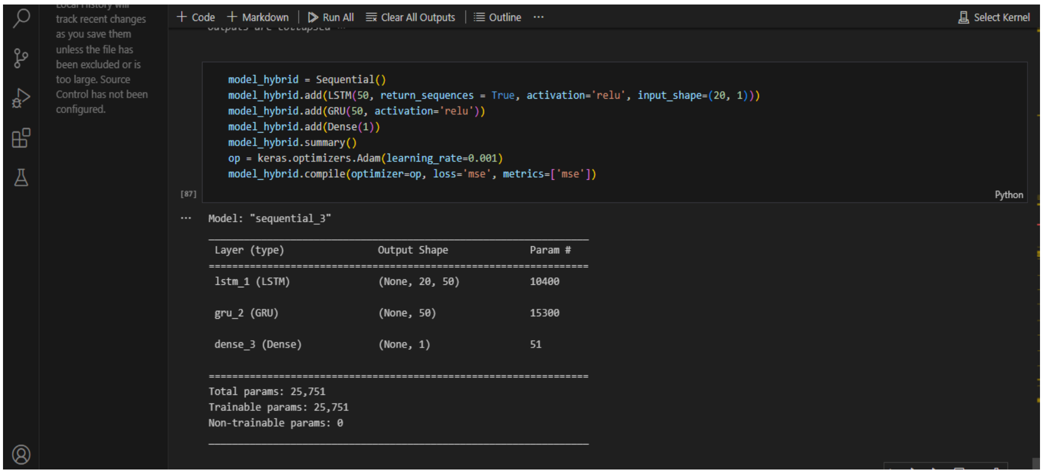


Fig. 14. Python Script for creating the Hybrid Model (LSTM + GRU) with one LSTM and one GRU layer and The Summary of Hybrid Model (LSTM+GRU) created with one LSTM and one GRU layer

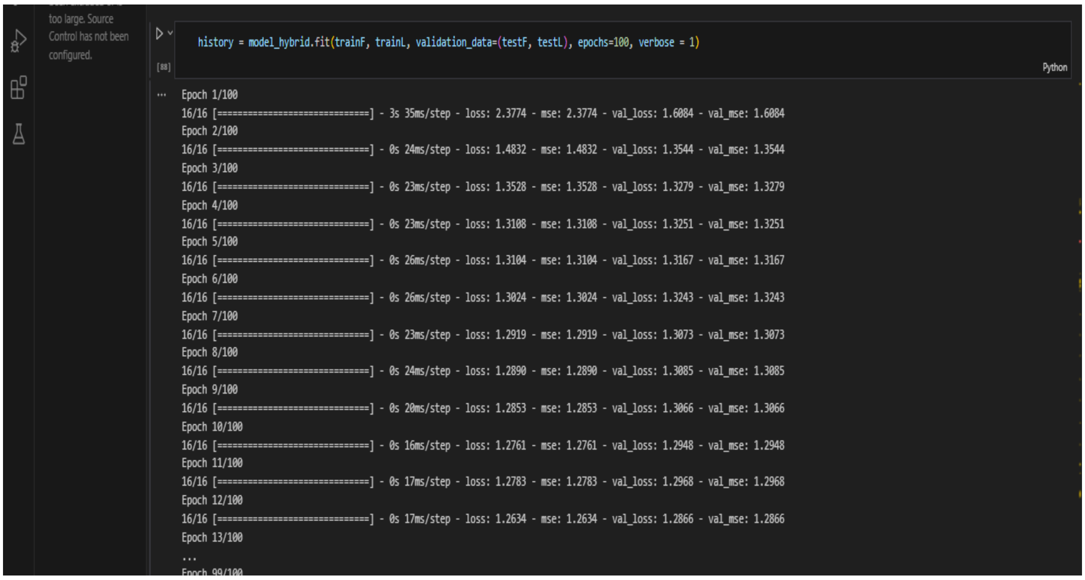


Fig. 15. Training log details the progress of a **Hybrid (LSTM+GRU) model** across 100 epochs

Table 2. Hyper Parameter Details

Model	***Hybrid LSTM-GRU***
Activation Function	**ReLU**
Input shape	**20,1**
Optimizer	**Adam**
Metrics	**MSE, Loss**

Epochs	**100**
Total Parameters	**25,751**

Table 3. Performance Evaluation of Deep Learning and Hybrid Neural Network

Model	*MSE*	*Val MSE*	*Loss*	*Val Loss*
LSTM	**1.07**	**1.09**	**1.07**	**1.09**
GRU	**0 .417**	**0.542**	**0.417**	**0.542**
Hybrid LSTM-GRU	**0.236**	**0.285**	**0.236**	**0.285**

Table 4. Performance Evaluation of Existing Work

Author / Year	*Models*	*Parameters*	*References*
Adil/ 2023	**LRR**	**MSE=1.81**	[17]
Dessain/ 2022	**ML Models**	**MSE=0.7**	[18]
Li/ 2021	**DNN**	**MSE=0.78**	[19]

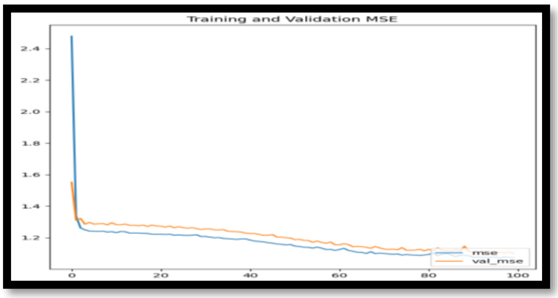


Fig. 16(a). MSE and Loss of LSTM

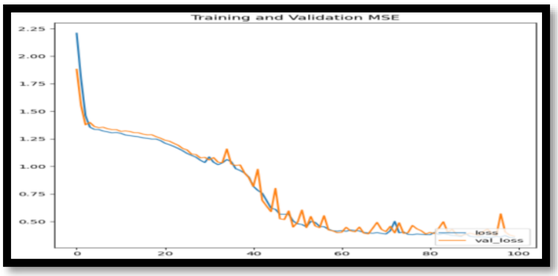


Fig. 16(b). MSE and Loss of LSTM

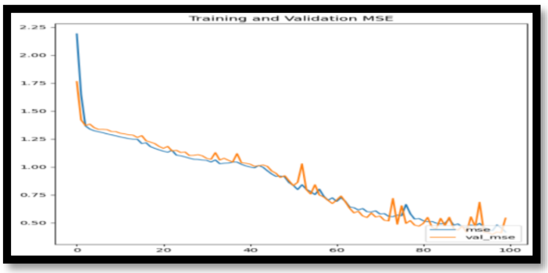


Fig. 17(a). MSE and Loss of GRU

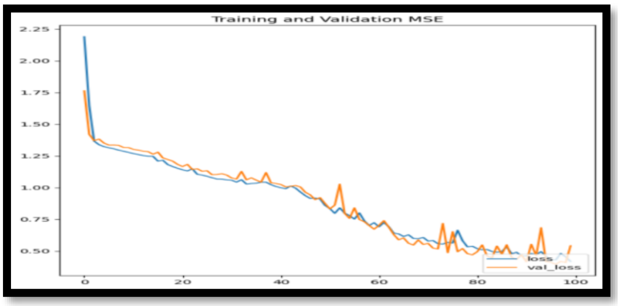


Fig. 17(b). MSE and Loss of GRU

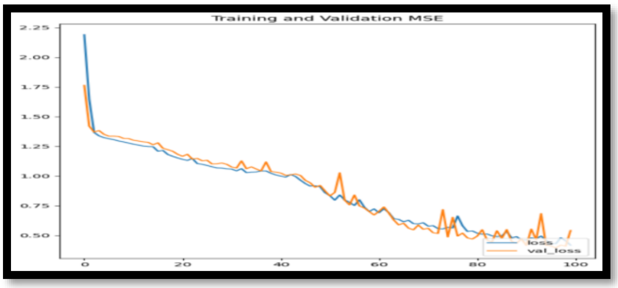


Fig. 18(a). MSE and Loss of Hybrid Neural Network

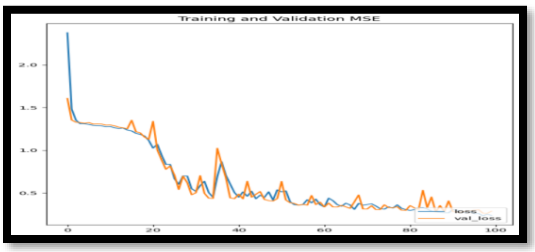


Fig. 18(b). MSE and Loss of Hybrid Neural

Network

Figure 16(a & b), 17(a & b), and 18(a & b) depict the Mean Squared Error (MSE) and Loss graphs associated with The Gated Recurrent Unit (GRU), Long-Short Term Memory (LSTM) and a Hybrid Neural Networks (HNN) "LSTM-GRU Model". It is evident from these figures that the Hybrid Neural Network model attains superior performance compared to its LSTM and GRU counterparts.

5. Conclusion

The assessment of student performance holds paramount significance in facilitating the optimal functioning of educational institutions, serving as a guiding beacon for curriculum development and pedagogical strategies. The reason for this study was to bring a new method to this field by combining the advantages of architectures of Neural Networks like the Gated Recurrent Unit (GRU), and the Long Short-Term Memory (LSTM). The objective is to innovate and enhance the reliability and predictability of grading and other forms of student assessment in forthcoming educational contexts. The outcomes of this research not only usher in an advancement in the precision of student performance evaluation but also unveil an unexplored avenue for harnessing the potential of the LSTM-GRU Hybrid Neural Network paradigm. Performance analysis is conducted using the metrics of Mean Squared Error (MSE) and Loss. **Remarkably, the Hybrid Neural Network emerges as the front runner, boasting the lowest MSE and Loss values.** Specifically, the MSE for the Hybrid Neural Network registers at **0.236**, with a corresponding Validation MSE of **0.285**. Simultaneously, the Loss stands at **0.236**, with the Validation Loss mirroring the same value at **0.285**.Comparatively, the performance of standalone LSTM and GRU models pales in significance when juxtaposed with the prowess of the Hybrid Neural Network, which garners demonstrably superior results. Furthermore, a retrospective evaluation of previously proposed Methodologies is undertaken, revealing a pivotal insight – the proposed work exhibits the highest MSE among these antecedent studies, indicative of its surpassing efficacy in contrast to its predecessors.

Acknowledgements

We heartily thank **Professor Umesh K. Chopra**, **Director, Amity Institute of Information Technology, Amity University Uttar Pradesh Lucknow Campus** for his constant motivation, valuable discussions and guidance during the entire research work, providing a conducive and congenial learning environment to study and research.

Conflicts of interest

The authors declare **no** conflicts of interest.

References

[1] M. Bouhlal, K. Aarika, R. AitAbdelouahid, S. Elfilali, and E. Benlahmar, "*Emotions recognition as innovative tool for improving students' performance and learning approaches*", Procedia Comput. Sci., vol. 175, pp. 597–602, 2020, doi: 10.1016/j.procs.2020.07.086.

[2] V. Matzavela and E. Alepis, "Decision tree learning through a Predictive Model for Student Academic Performance in Intelligent M-Learning environments", Comput. Educ. Artif. Intell., vol. 2, p. 100035, 2021, doi: 10.1016/j.caeai.2021.100035.

[3] P. Bhardwaj, P. K. Gupta, H. Panwar, M. K. Siddiqui, R. Morales-Menendez, and A. Bhaik, "Application of Deep Learning on Student Engagement in e-learning environments", Comput. Electr. Eng., vol. 93, no. April, p. 107277, 2021, doi: 10.1016/j.compeleceng.2021.107277.

[4] S. Oeda and D. Shimizu, "Verification of usefulness of student modeling with real educational data using convex factorization machines", Procedia Comput. Sci., vol. 192, pp. 804–811, 2021, doi: 10.1016/j.procs.2021.08.083.

[5] R. Ahuja and A. Banga, "Mental stress detection in University students using machine learning algorithms", Procedia Comput. Sci., vol. 152, pp. 349–353, 2019, doi: 10.1016/j.procs.2019.05.007.

[6] X. Niu, "Deep-Learning-Guided Student Intelligent Classroom Management System", Appl. Bionics Biomech., vol. 2022, 2022, doi: 10.1155/2022/1961631.

[7] Metaheuristics Method for Classification and Prediction of Student Performance Using Machine Learning Predictors" , vol. 2022, pp. 1–5, 2022.

[8] Y. Tan, "Application Research on Face Image Evaluation Algorithm of Deep Learning Mobile Terminal for Student Check-In Management", Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/3961910.

[9] Y. A. Alsariera, Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, and N. Ali, "Assessment and Evaluation of Different Machine Learning Algorithms for Predicting Student Performance" ,vol. 2022, 2022.

[10] Y. Zhang, T. Liu, and B. Ru, "Effect Evaluation and Student Behavior Design Method of Moral Education in Colleges and Universities under the Environment of Deep Learning", Sci. Program.,

vol. 2022, 2022, doi: 10.1155/2022/5779130.

[11] J. Waldeyer et al., “A moderated mediation analysis of conscientiousness, time management strategies, effort regulation strategies, and University students’ performance” , Learn. Individ. Differ, vol. 100, no. July, p. 102228, 2022, doi: 10.1016/j.lindif.2022.102228.

[12] T. T. Mai, M. Bezbradica, and M. Crane, “Learning behaviours data in programming education: Community analysis and outcome prediction with cleaned data”, Future. Gener. Comput. Syst., vol. 127, pp. 42–55, 2022, doi: 10.1016/j.future.2021.08.026.

[13] R. Alshabandar, A. Hussain, R. Keight, and W. Khan, “Students Performance Prediction in Online Courses Using Machine Learning Algorithms”, Proceedings of Joint Conference of Neural Networks, vol. 02, no. 11, pp. 74–79, 2020, doi : 10.1109/IJCNN48605.2020.9207196.

[14] F. D. Pereira, E. H. T. Oliveira, D. Fernandes, and A. Cristea, “Early performance prediction for CS1 course students using a combination of machine learning and an evolutionary algorithm”, Proc. - IEEE 19th Int. Conf. Adv. Learn. Technol. ICALT 2019, vol. 2161–377X, pp. 183–184, 2019, doi: 10.1109/ICALT.2019.00066.

[15] M. R. Rimadana, S. S. Kusumawardani, P. I. Santosa, and M. S. F. Erwianda, “Predicting Student Academic Performance using Machine Learning and Time Management Skill Data”, 2019 2nd Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2019, pp. 511–515, 2019, doi: 10.1109/ISRITI48646.2019.9034585.

[16] L. M. Abu Zohair, “Prediction of Student’s performance by modelling small dataset size”, Int. J. Educ. Technol. High. Educ., vol. 16, no. 1, 2019, doi: 10.1186/s41239-019-0160-3.

[17] K. Adil and A. Ahmed, “Machine Learning and Deep Learning based Students’ Grades Prediction” pp. 1–18, 2023.

[18] J. Dessain, “Machine learning models predicting returns: Why most popular performance metrics are misleading and proposal for an efficient metric”, vol. 199. 2022. doi: 10.1016/j.eswa.2022.116970.

[19] Li and T. Liu, “Performance Prediction for Higher Education Students Using Deep Learning”, Complexity, vol. 2021, 2021, doi: 10.1155/2021/9958203.