Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 1s (2025)

Performance Comparison of Deep Learning Algorithm for Autism Spectrum Disorder Prediction

[1]Tintu Varghese, [2] [2]Dr K Devasenapathy

*[1]Research Scholar, Karpagam Academy of Higher Education, Coimbatore, tinta.varughese@gmail.com[1],9188086651

*[2]Assistant Professor, Karpagam Academy of Higher Education, Coimbatore,
drdevasenapathy.k@kahedu.edu.in[2],9600965375

Article History:

***Received:** 04-09-2024*

***Revised:** 19-10-2024*

***Accepted:** 02-11-2024*

Abstract:

Autism Spectrum Disorder is a complex neurodevelopmental condition that significantly impacts social interaction and conduct in individuals. Early diagnosis and timely intervention are crucial for effectively managing the condition and enhancing overall outcomes. This study concentrates on creating predictive models through deep learning algorithms to assist in the early identification of ASD. The study employs an extensive dataset that covers various features linked to ASD traits, such as speech and language development, learning disorders, genetic factors, and behavioral characteristics. The investigation encompasses the assessment of three recurrent neural network (RNN) architectures-standard RNN, Long Short Term Memory(LSTM), and Gated Recurrent Unit(GRU) –for their effectiveness in predicting ASD. The dataset is split into 80% for training and 20% for testing.

This study compares RNN, LSTM, and GRU algorithms. The classification results show that the LSTM and GRU models exhibited similar accuracy, with values of 71.03% and 70.78% respectively.

Keywords: Autism Spectrum Disorder, Deep Learning, LSTM, GRU, RNN

1. INTRODUCTION

Autism Spectrum Disorder(ASD) presents a notable challenge within the domain of neurodevelopmental disorders, influencing social interactions, communication skills, and behavioral patterns in individuals. Early identification and intervention are vital in improving outcomes and boosting the well-being of individuals with ASD."Recently, incorporating advanced machine learning techniques, specifically deep learning algorithms has emerged as a promising avenue for predicting modeling in healthcare. Early detection of Autism Spectrum Disorder (ASD) in children is crucial for providing timely support and interventions, significantly enhancing their developmental outcomes and overall quality of life[1]. Early identification also enables timely access to health services, which can help mitigate the challenges faced by individuals with ASD and their families[2]. Various methods for detecting ASD in children include analyzing family medical history and identifying specific biomarkers[3]. This study explores the utilization of deep learning, particularly focusing on RNN, LSTM, and GRU for predicting Autism Spectrum Disorder (ASD).

By using a dataset that encompasses features related to traits, associated with autism spectrum disorder.We conduct data preprocessing and employ RNN, LSTM, and GRU models. Performance metrics, including accuracy, F1 score, and precision, are employed to assess the predictive effectiveness of each algorithm. The selected elements for predicting ASD encompass speech and language development, learning disorders, genetic factors, and social and behavioral characteristics. The study seeks to improve the strength and applicability of the developed models by utilizing a diverse and comprehensive dataset.

This article presents characteristics of RNN, LSTM, and GRU algorithms for Autism spectrum disorder prediction. Part 2 focuses on literature review .

.II. LITERATURE REVIEW

Vaishali R, Sasikala R., et al. [4] proposed an approach for identifying Autism Spectrum Disorder (ASD) by optimizing behavior sets. The study employed an ASD diagnosis dataset with 21 attributes, sourced from the UCI Machine Learning Repository. The authors utilized a bio-inspired binary firefly feature selection framework in their experiments.

The underlying hypothesis suggests that a machine learning model can achieve higher classification accuracy by utilizing a minimal subset of features.M. S. Mythili, A. R. Mohamed Shanavas, et al. [5] conducted an investigation focused on the detection and classification of Autism Spectrum Disorder (ASD) and the assessment of autism severity levels. The primary objective was to identify autism-related issues and evaluate the severity of the condition. To study students' behavioral and social interaction dynamics, the researchers applied Neural Networks, Support Vector Machines (SVM), and Fuzzy techniques, utilizing the WEKA software for data analysis. Through the application of these classification techniques, the research aimed to gain insights into the intricacies of autism and provide a nuanced insight into the spectrum of severity levels connected with the disorder. In paper[6], the main aim is to identify the most significant traits related to ASD and automate the diagnosis process using classification techniques. The study analyzes datasets related to Autism Spectrum Disorder (ASD) spanning different age ranges, including toddlers, children, adolescents, and adults.

The authors suggest the potential of machine learning, specifically the MLP classifier, in enhancing the accuracy of ASD diagnosis.

Automation of the diagnosis process using these techniques may contribute to more efficient and reliable early detection.

In the paper, the authors [7] introduced a powerful approach for diagnosing Autism Spectrum Disorder (ASD) using deep learning on facial images. The method involves training a convolutional neural network (CNN) with a dataset, incorporating pre-processing and data synthesis. The trained model is then assessed on an independent test set. The novel approach, which simultaneously applies pre-processing and augmentation during training, outperforms recent methods, achieving a remarkable 98.9% accuracy, sensitivity, and specificity, with a 99.9% Area Under the Curve (AUC). Notably, the algorithm integrates explainable AI techniques, enhancing clarity for clinicians and providing a clear understanding of the ASD diagnosis model's decision-making process. This research contributes to the field by delivering a highly accurate and transparent approach to ASD diagnosis from facial images.

Kang's studies utilize a Support Vector Machine (SVM) to identify cases of Autism Spectrum Disorder (ASD) in children, relying on both EEG and eye-tracking data [8]. The classification results from these combined data sources show potential for improvement. The studies suggest employing both EEG and eye-tracking data together improves the precision of detecting ASD compared to relying solely on Electroencephalography or eye-tracking data separately. This observation underscores the importance of leveraging multiple modalities in the diagnostic process. The studies contribute to the field's understanding of the complementary nature of EEG and eye-tracking data in improving the classification of ASD cases in children.

The study [9] explores the use of document classification algorithms to improve the efficiency of measuring the occurrence of Autism Spectrum Disorder (ASD) in children in the United States, a process traditionally managed through labor-intensive procedures by the Centers for Disease Control and Prevention (CDC). While random forest methods have shown promise, they still lag behind human classification accuracy. The research aims to investigate whether newly available document classification algorithms can help reduce this disparity.

III.DATASET DESCRIPTION

Data preprocessing plays a pivotal role in machine learning and data analysis by readying the dataset for analysis or model training. It encompasses various steps to cleanse, transform, and manipulate the data, ensuring its quality and alignment with the selected machine learning algorithms. The dataset used in this study was acquired from Kaggle. The dataset comprises 2,000 instances, encompassing diverse demographic and clinical profiles.

The dataset incorporates the following features.

1. Age –The age of individuals

2.Speech Delay/Language Disorder(Binary)-existence(1) or lack(0) of speech delay or language disorder

3. Genetic Disorder(Binary): existence (1) or lack (0) of known genetic disorders associated with ASD.

4. Depression(Binary): existence (1) or lack (0) of depression.

5. Global Developmental Delay/Intellectual Disability(Binary): existence (1) or lack (0) of global developmental delay or intellectual disability.

6. Social/Behavioural Issues(Binary): existence (1) or lack (0) of social or behavioral issues.

7. Anxiety Disorder(Binary): existence (1) or lack (0) of an anxiety disorder.

8. Sex(Categorical): Gender of the individual (Male, Female)

9. Jaundice(Binary): existence (1) or lack (0) of jaundice during infancy.

10. Family member with ASD(Binary): existence (1) or lack (0) of a family member diagnosed with autism.

Most of the features are binary, representing the existence or lack of specific characteristics.

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 2 (2025)

The dataset underwent a number of preparation processes before the analysis started to make sure it was suitable for machine learning applications.

Table 1 presents a summary of the preprocessing steps applied to the dataset and the corresponding results.

Preprocessing Stage	Dataset size	Features	Target Encoding	Categorical Encoding
Original Dataset	2000	10	No	No
After Removing Irrelevant Features	2000	8	No	No
After Imputing Missing Values	2000	8	No	No
After Target Variable Encoding	2000	8	Yes	No
After One-Hot Encoding of Categorical Columns	2000	14	Yes	Yes
Training Set	1600	14	Yes	Yes
Testing Set	400	14	Yes	Yes

Table 1:Dataset overview at various preprocessing steps

In this table,

- “Preprocessing Stage” describes the different stages of preprocessing.
- “Dataset Size” indicates the number of instances in the dataset at each stage.
- “Features” represents the number of features remaining after each pre-processing stage.
- “Target Encoding” specifies whether the target variable has been encoded into numerical format(Yes/No)
- “Categorical Encoding” indicates whether categorical variables have been encoded using one-hot encoding(Yes/No)

A preliminary examination revealed minimal missing values and imputation techniques were applied for handling them. To facilitate the training of machine learning models, the target variable 'Autism' was encoded into a numerical format using label encoding. This transformation replaced 'Yes' with 1 and 'No' with 0. Categorical variables in the feature set were subjected to one-hot encoding. This technique creates binary vectors for each category, ensuring compatibility with the deep learning model. The dataset was partitioned randomly into training (80%) and testing (20%) sets.

IV.MODEL ARCHITECTURE

Deep Learning, a method within machine learning, has garnered considerable attention [10]. It excels in predicting models with high complexity and finds application across various domains. [11].

Three deep learning architectures were implemented and compared: Recurrent Neural Network(RNN), Long Short Term Memory(LSTM), and Gated Recurrent Unit(GRU). Each model comprised initial layers for capturing sequential dependencies, followed by Dense layers to learn complex patterns. Batch normalization and dropout layers were introduced to mitigate overfitting and enhance model performance. The models underwent training using the training data, and their performance was assessed on the validation data. Training parameters, including the number of epochs, batch size, and optimizer configurations, were fine-tuned for each architecture. Accuracy, F1 score, and precision were tracked during training to monitor model convergence and performance.

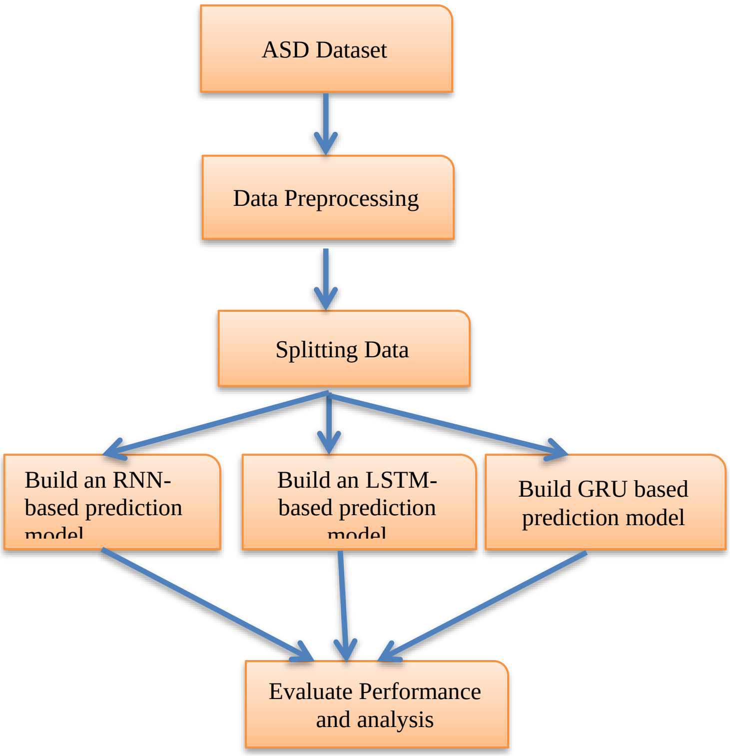

Figure 1:Model Architecture

1. Recurrent Neural Network (RNN):

In simpler terms, recurrent neural networks (RNNs) are structured to interpret sequential data by allowing information to flow not just forward but also backward through the network. This is achieved through loops within the network that connect hidden units. The internal connections facilitate RNNs in efficiently leveraging previous information to anticipate future outcomes, making them adept at recognizing patterns in sequential data. Specifically, RNNs excel at capturing temporal dependencies between data points that might be widely spaced apart in a sequence. This ability to grasp long-range relationships is crucial for tasks like time series forecasting, language translation, and speech recognition.

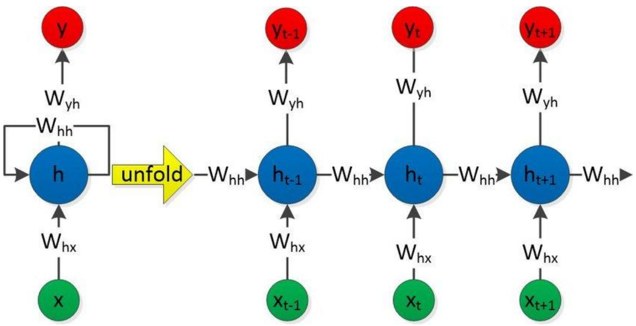

Figure 2:RNN architecture

Figure 1 illustrates the RNN architecture where given an input series $x = \{x1, x2, \dots . xT\}$, the RNN iteratively computes the hidden state sequence $h = \{h1, h2, \dots, hT\}$

and the output sequence $y = \{y1, y2, \dots, yT\}$

using the below equations.

$ht = f(ht = f(Yhxxt + Yhhht - 1 + ch)$---(1)

yt=g(Yyhht+cy)-----------*(2)*

Here Yhx, Yhh, and Yyh represent

the input-hidden weight matrix, hidden-hidden weight matrix, and hidden-output weight matrix. The ch, cy denote the biases of the hidden layer and the output layer, respectively.

Further, the activation functions for the output layer and the hidden layer are g(.) and f(.) respectively.

.The Recurrent Neural Network utilizes hidden state hT at timestep t to retain information from previous time steps, capturing all relevant information contained within each preceding time step.

2. Long Short-Term Memory(LSTM):

The Long Short-Term Memory architecture was selected to tackle the issue of vanishing gradients and accurately model long-term relationships within the dataset. The model architecture featured a single LSTM layer with 64 units and utilized a ReLU activation function. To ensure regularization, Batch Normalization, and dropout layers were incorporated. For the binary classification output, a Dense layer with a sigmoid activation function was employed. An LSTM cell usually comprises several crucial components:

1. **Cell State (Ct):** Within the LSTM cell, the cell state functions as the internal memory. It can carry information across extensive sequences, which makes it adept at capturing long-term dependencies.

2. **Hidden State (ht):** At a given time step, it represents the output of the LSTM cell. It holds the information that the cell considers relevant for the current prediction or processing step.

4. **Input Gate (i):** It controls the influx of information into the cell state, determining which values from the input and the previous hidden state must be modified and integrated into the cell state.

5. **Forget Gate (f):** It determines the values based on the previous cell state, excluding the input. It controls the information that should be removed from the cell state.

6. **Output Gate (o):** It dictates which part of the memory state is to be revealed as the hidden state for a present time interval.

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 2 (2025)

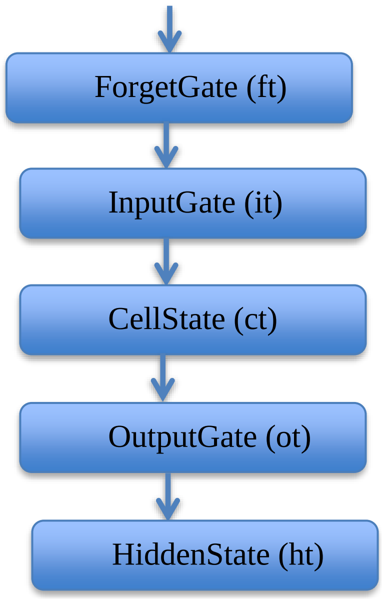

Figure 3:LSTM Architecture

In this diagram:

- The arrows represent the flow of information.
- Each gate (input gate, forget gate, output gate) is like a small neural network layer that takes inputs, applies weights, and produces outputs.
- The forget gate determines which aspects of the prior cell state to retain or discard. This gate evaluates the importance of the previous information and decides whether to keep it in the cell state or remove it.
- The input gate assesses the relevance of incoming data from the current input and the previous hidden state, deciding which parts should be incorporated into the cell state. It updates the cell state by incorporating this new information based on its significance.
- The output gate identifies the portion of the cell state that is to be revealed as the hidden state. It controls what data is output from the cell state to be used for the current time step's processing or prediction.

3. Gated Recurrent Unit (GRU):

The Gated Recurrent Unit architecture was designed as a variant of Long Short-Term Memory, offering similar capabilities with a reduced computational cost. The model included a single GRU layer with 64 units and employed a ReLU activation function. Batch Normalization and dropout layers were used to ensure regularization. For the binary classification output, a Dense layer with a sigmoid activation function was utilized as the final layer.

GRU Components:

1. **Update Gate (z_t):** The degree to which the merging of the previous hidden state with the candidate's new hidden state is determined.

2. **Reset Gate (r_t)**: The Reset gate influences the candidate's new hidden state by controlling how much of the past information to forget.

This process helps in determining which previous information is retained and which is discarded, shaping the new hidden state accordingly.

3. **Candidate Hidden State (h'_t):** The new hidden state candidate creates an updated representation by integrating information from the previous hidden state and present input.

4. **Hidden State (h_t):** The update gate regulates the extent to which the candidate state influences the current state, effectively blending past information with new inputs to form the updated hidden state. The final hidden state is formed by blending the current hidden state with the candidate's new hidden state, regulated by this gate.

Key performance indicators, including accuracy, F1 score, and precision were used to evaluate the model, to comprehensively assess classification effectiveness. Classification reports were generated to provide detailed insights into the model's ability to correctly classify individuals with or without Autism Spectrum Disorder.

V. RESULT AND DISCUSSION

While the current study compares the performance of various deep learning algorithms for predicting Autism Spectrum Disorder, it is essential to acknowledge the existing research gaps that provide avenues for future exploration. Firstly, the majority of existing literature primarily focuses on the accuracy of predictive models, neglecting the interpretability and explainability of these models, which are crucial aspects in the context of medical diagnostics. Addressing this gap would contribute to building trust in the practical application of deep learning algorithms for ASD prediction, ensuring that the generated insights are not only accurate but also transparent and understandable for clinicians and stakeholders.

Furthermore, while the study identifies the most effective deep learning algorithm for ASD prediction based on the utilized dataset, there is an opportunity to explore the impact of different feature sets and data representations on model performance. Investigating the influence of varied input features, such as different types of behavioral data or additional biomarkers, could provide valuable insights into optimizing model inputs for enhanced accuracy and clinical relevance. Addressing these research gaps would both strengthen the current findings and foster the creation of more robust and adaptable deep-learning models for ASD prediction, fostering their integration into clinical practice and improving the overall understanding of the disorder.

In this study, the models were evaluated using key performance metrics including accuracy, precision, and F1 score. The results obtained for each architecture are summarized below:

	RNN	LSTM	GRU
Accuracy	70.78%	71.03%	70.78%
Precision	63.41%	63.87%	63.41%
F1 Score	77.00%	77.50%	77.61%

Table 2:Summary Results

The LSTM model achieved the highest accuracy among the three architectures, closely followed by GRU and RNN. The marginal differences suggest that all models demonstrated comparable overall predictive performance.

Precision assesses a model's effectiveness in accurately identifying positive cases. The proportion of true positive predictions to the total number of predicted positive outcomes is represented by it. The model's precision in predicting a favorable outcome is evaluated by measuring the accuracy of positive predictions.

The LSTM and GRU models exhibited slightly higher precision compared to the RNN model. A thorough assessment of a model's performance is provided by the F1 score, which takes into account both precision and recall. All three models demonstrated high F1 scores, indicating an effective balance between precision and recall.

To offer a visual representation of the comparative results, the following chart illustrates the performance metrics achieved by the RNN, LSTM, and GRU models. This chart aims to provide an overview and facilitate a more intuitive understanding of the comparative analysis.

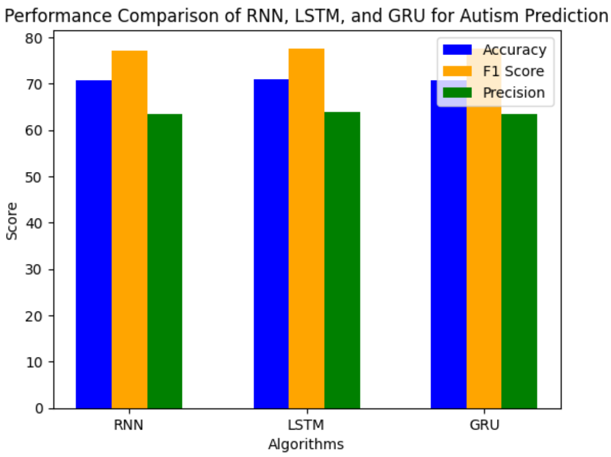

Figure 4:Performane comparison of RNN, LSTM, and GRU

V. CONCLUSION

In this study, we conducted a comprehensive comparison of three recurrent neural network architectures, LSTM and GRU-for the task of autism prediction. Across the performance metrics, the LSTM and GRU models exhibited similar accuracy, with values of 71.03% and 70.78%, respectively. The F1 scores were close, with LSTM leading at 77.50% and GRU following closely at 77.61%. Precision scores were also comparable, with LSTM slightly outperforming the other models at 63.87%.

Our research extends to the body of knowledge by providing insight into whether deep learning techniques are appropriate for predicting autism.

Future research could explore hybrid models, hyperparameter tuning, and the inclusion of additional features to further enhance predictive performance. Despite the results, this study has limitations, including the reliance on a specific dataset and potential biases inherent in the data. our comparative analysis offers valuable insights into the performance of RNN, LSTM, and GRU models for autism prediction. The findings contribute to the knowledge enhancement in the application of deep learning in healthcare and lay for future advancements in this field.

REFERENCES

[1] Metcalfe, D., McKenzie, K., McCarty, K., & Murray, G. (2020). Screening tools for autism spectrum disorder, used with people with an intellectual disability: A systematic review. *Research in Autism Spectrum Disorders*, 74.

[2] Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *International Journal of Research and Practice*, *12*(5), 537-556.

[3] Bridgemohan, C., Cochran, D. M., & Howe, Y. J. (2019). Investigating potential biomarkers in autism spectrum disorder. *Frontiers in Integrative Neuroscience*, *13*(31).

[4] Vaishali, R., & Sasikala, R. (2018). A machine learning-based approach to classify Autism with optimum behavior sets. *International Journal of Engineering & Technology*, *7*(4), 18.

[5] Mythili, M. S., & Mohamed Shanavas, A. R. (2014). A study on Autism spectrum disorders using classification techniques. *International Journal of Soft Computing and Engineering (IJSCE)*, *4*, 88-91.

[6] Hossain, M. D., Kabir, M. A., Anwar, A., & Islam, M. Z. (2021). Detecting autism spectrum disorder using machine learning techniques. *Health Information Science and Systems, 9*(1). https://doi.org/10.1007/s13755-021-00145-9

[7] Alam, M. S., Rashid, M. M., Faizabadi, A. R., Mohd Zaki, H. F., Alam, T. E., Ali, M. S., Gupta, K. D., & Ahsan, M. M. (2023). Efficient deep learning-based data-centric approach for autism spectrum disorder diagnosis from facial images using explainable AI. *Technologies, 11*, 115. https://doi.org/10.3390/technologies11050115

[8] Kang, J., Han, X., Song, J., Niu, Z., & Li, X. (2020). The identification of children with autism spectrum disorder by SVM approach on EEG and eye-tracking data. *Computers in Biology and Medicine*, *120*, 1-5.

[9] Lee, S. H., Maenner, M. J., & Heilig, C. M. (2019). A comparison of machine learning algorithms for the surveillance of autism spectrum disorder. *PLoS ONE, 14*(9),e0222907. https://doi.org/10.1371/journal.pone.0222907

[10] Medsker, L., & Jain, L. (2001). Recurrent neural networks. *Design and Applications*.

[11] Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural networks. In *ICML (3), vol. 28* (pp. 1310–1318).

[12] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, *18*(7), 1527-1554.

[13] Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., & Dehmer, M. (2020). An introductory review of deep learning for prediction models with big data. *Frontiers in Artificial Intelligence, 3*, 4.