Predicting autism spectrum disorder through sentiment analysis with attention mechanisms: a deep learning approach

Murali Anand Mareeswaran[1,2], Kanchana Selvarajan[1]
[1]Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India
[2]Department of Computer Science and Applications, Faculty of Science and Humanities, SRM Institute of Science and Technology, Vadapalani, Tamil Nadu, India

Article Info

Article history:

Received Apr 24, 2024
Revised Sep 8, 2024
Accepted Sep 29, 2024

Keywords:

Adam optimizer
Autism spectrum disorder
BiLSTM
N-gram
Sentiment analysis

ABSTRACT

Autism spectrum disorder (ASD) is considered a spectrum disorder. The availability of technology to identify the characteristics of ASD will have major implications for clinicians. In this article, we present a new autism diagnosis method based on attention mechanisms for behavior modeling-based feature embedding along with aspect-based analysis for a better classification of ASD. The hybrid model comprises a convolutional neural network (CNN) architecture that integrates two bidirectional long short-term memory (BiLSTM) blocks, together with additional propagation techniques, for the purpose of classification the origins of Autism Tweet dataset; the proposed work takes Autism Tweet dataset and preprocesses them to employ n-gram to extract features of which the features of the ASD behavior are fed to generate the significant behavior for classification. The model takes into account both behavior-guided features across every aspect of the Class/ASD to provide higher accuracy using Adam optimizer. The experimental values inferred that the n-BiLSTM technique reaches maximum accuracy with 98%.

Corresponding Author:

Kanchana Selvarajan
Department of Computer Science, Faculty of Science and Humanities
SRM Institute of Science and Technology
Kattankulathur, Chengalpattu, Tamil Nadu, 603203, India
Email: kanchans@srmist.edu.in

1. INTRODUCTION

Autism spectrum disorder (ASD) is a prevalent and complex neurodevelopmental disorder that impacts individuals throughout their lives. Characterized by challenges in communication, social interactions, and restricted behaviors, early diagnosis of ASD is crucial for implementing effective interventions. Despite the reliance on neurophysiological signals and behavioral assessments for diagnosis, there are currently no definitive biological markers or straightforward diagnostic procedures available, making the process challenging. General practitioners often screen for ASD symptoms in patients, but definitive diagnosis may require referral to specialists, leading to delays in intervention. Early detection is vital, as timely support can significantly improve the quality of life for individuals with ASD and their families Shahamiri and Thabtah [1].

Recent advancements in machine learning (ML) and deep learning (DL) techniques have shown promise in predicting and detecting various conditions, including ASD, based on physical and physiological indicators Garg *et al.* [2]. However, the analysis of ASD is complicated by overlapping symptoms with other mental health issues, leading to potential false positives. Understanding the features that ML and DL models utilize in predicting ASD can aid healthcare professionals in making informed decisions during initial assessments. Our motivation stems from the understanding that early identification of ASD can reduce symptom severity and enhance treatment outcomes.

Sentiment analysis, a robust natural language processing (NLP) technique, has emerged as a novel approach in healthcare for predicting ASD. Individuals with ASD often display distinct linguistic and emotional patterns in their communication, which can be leveraged to identify potential signs of the disorder. By employing intelligent systems and attention mechanisms, our research aims to develop a sophisticated tool for ASD prediction that harnesses sentiment analysis to analyse extensive textual and auditory data. The integration of attention mechanisms enhances the model's ability to focus on relevant cues within the input data, thereby improving prediction accuracy.

In this article, we explore the fusion of cutting-edge technology and compassionate care. By harnessing the potential of sentiment analysis and attention mechanisms, we aspire to offer a tool that can assist healthcare professionals and caregivers in the early identification of autism, enabling more timely interventions and assistance for people with autism spectrum. This research represents a promising stride towards a brighter future for those affected by ASD, illustrating the profound impact that intelligent systems can have on the field of healthcare and neurodevelopmental disorders. For NLP tasks, such as sentiment analysis, attention based bidirectional long short-term memory (BiLSTM) and Adam optimizer is a prevalent and effective combination. With this configuration, we have constructed a neural network architecture that makes use of the Adam optimizer for effective training, BiLSTM layers for sequence processing, and an attention mechanism to concentrate on a significant part of the input data.

The various existing work proposed by several researchers to detect the ASD are presented in this section. Joudar *et al.* [3] given the systematic information on healthcare sector and physician’s works which included in identification of ASD. According to their study that the most of the articles have discussed the issues, traditional approach to detect and also inferences of ML approach. In addition to this, the authors also proposed a decision-making method to detect and prioritizing the patients accordingly. Mareeswaran and Selvarajan [4] identified the ASD in children by various ML techniques on non-clinical dataset. After handling the missing values the finding suggested that support vector machine (SVM) can identify ASD effectively in its beginning phases and yields 96%. Kashef [5] proposed enhanced convolutional neural network (CNN) to identify the ASD. The proposed method achieved an accuracy of 80% from pattern identified the anterior and posterior brain region’s function which acted as the evidence of ASD.

Raj and Masood [6] detected the ASD in children by using several ML and DL models on non-clinical dataset from 3 difference age groups. The final result showed up 98.30% accuracy in SVM and CNN after handling missing values. With this, the authors can reach highest accuracy by using CNN for other two dataset also. So, they, concluded that the CNN model has detected the ASD with good accuracy than other models. Bayram *et al.* [7] have presented a method to detect ASD on resonance imaging resting-state functional magnetic resonance imaging (rs-fMRI). Long short term memory (LSTM), CNN and hybrid model were utilized by the author to identify ASD and they achieved 74.74% accuracy and 72.95% sensitivity in recurrent neural network (RNN) based model. Gong *et al.* [8] proposed a method to identify the entities related to molecular mechanism for autism BiLSM and conditional random field (CRF) from a text collection dataset. The result achieved 76.81% of f1*score with the extraction of 9146, 145, 7680, 1058, 981 of proteins, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), cell type and cell line respectively.

Alkahtani *et al.* [9] conducted study based on facial landmark detection. Different ML approaches, including logistic regression (LR), random forest (RF), linear support vector classifier (SVC), decision tree, multi-layer perceptron (MLP) classifier, gradient boosting, and K-nearest neighbors (KNN), are utilized with MobileNetV2 and hybrid visual geometry group (VGG)-19. MobileNetV2 model attained an accuracy of 92%. Rabbi *et al.* [10] detected autism in children using five different algorithms such as MLP, RF, Adaboost, gradient boosting algorithm and CNN. The authors have achieved 92.31% accuracy on CNN. Arumugam *et al.* [11] developed a prediction system based on the pictures using CNN in which dataset was taken from Kaggle. The system reached the accuracy rate of 91% with the loss rate of 0.53 with the splitting ratio of 20:80 for testing and training. Kavitha and Siva [12] proposed a particle swarm optimization (PSO)-CNN model to detect ASD and also analyzed with three other algorithms like SVM, Naive Bayes (NB), and LR. The proposed model achieved 99.1% accuracy with more efficiency.

Deng *et al.* [13] diagnosed fMRI of brain using 3D-CNN. The authors have got 74% accuracy, 76% specificity and 69.9 % sensitivity for publicly available autism dataset. Ali *et al.* [14] proposed a hybrid DL model to detect autism in which the hybrid model consist of CNN with BiLSTM which classify the output from EEG. Kaur and Gupta [15] the system achieved 97.7% accuracy where the distribution ratio was about 70:30 of training and testing respectively. The study highlights how the VGG16 CNN, which outperformed a regular CNN by 68.54%, can increase the accuracy of autism diagnosis. The study emphasizes the usefulness of VGG16 and picture data preprocessing.

Jiang *et al.* [16] presented CNNG model by combining CNN and gated recurrent units (GRU). Spatial and temporal features were extracted by CNN and GRU respectively, and classify them by sigmoid function. The proposed model has got 72.46% accuracy by extracting the feature of spatial and temporal from fMRI. Sherkatghanad *et al.* [17] proposed a CNN model to detect and classify the ASD dataset and also, they compared with three different supervised models like SVM, KNN and RF with pre-processed ABIDE dataset. Finally, the result showed that 70.2% accuracy have achieved with low complex and faster than other models.

Karuppasamy *et al.* [18] identified ASD using brain signals from large dataset with CNN model and achieved 95% accuracy. Aghdam *et al.* [19] diagnosed the imaged based ASD from ABIDE I and ABIDE II dataset with the age group rages from 5 to 10 years. The CNN model with Adam and Adamax optimizer were also employed and concluded that the proposed model of CNN with Adamax optimizer achieved good result. Giarelli *et al.* [20] concurs that it is more common in men than in women, with 81% of the sample's participants being men, and the DSM-IV diagnosis being used.

Abdelwahab *et al.* [21] diagnosed ASD with supported vector machine, RF, LR, NB, KNN, and decision tree were among the ML techniques used to identify the ASD, KNN model achieved 95.65%. Murugaiyan and Uyyala [22] proposed a model for speech emotion detection, speech aspect recognition and aspect based sentimental analysis by using DCNN, BiLSTM and rule-based classifier respectively. The suggested hybrid model obtained the accuracy of 93.28%, 91.45%, 92.12%, and 90.45% for 4 different datasets. Fan *et al.* [23] proposed a DL model called DeepASDPred to identify the RNA risk in autism. Initially the authors have used K-mer to encode the RNA and fuse with corresponding gene expression and also utilized Chi-square test, LR model for feature extraction. The final subset is fed to the CNN model and LSTM for training and classification.

Finally, the 10-fold cross-validation were used and showed that the model has performed well. Sudha and Vijaya [24] proposed LSTM based model to classify gene sequence causing ASD. The experimental result showed that LSTM based RNN model performed better than DNN and LSTM. Amirbay *et al.* [25] eye-tracking data was used to create LSTM+CNN and LSTM+Autoencoder models for autism diagnosis, which showed promising accuracy and clinical use. Belen *et al.* [26] given the systematic review in autism which uses computer vision technique. The authors have also reviewed the publicly available dataset to analyze the work and they concluded that CV approach is better than traditional approaches.

Tao and Shyu [27] proposed a model called SP-ASDNET which combines the CNN and LSTM model to classify ASD in respect to the images. The proposed model achieved 74.22% accuracy on validation dataset. The proposed model achieved the accuracy of 98% with more prediction. Corti *et al.* [28] 691,582 people (188 bots that generated 59,104 tweets) were identified from the 2,458,929 tweets generated in 2020, while 684,032 users (230 bots that generated 50,057 tweets) were identified from the 2,393,236 total tweets from 2019. Only a small portion of the entire dataset consists of the COVID-ASD tweets. The study collected data through online questionnaires, revealing the both teachers and parents rated. The major contribution of this article is as follows:

- A novel DL model is proposed to classify and use the tweet dataset to early identify ASD.
- We proposed a model with a Bi-LSTM neural network that incorporated with Adam optimizer with the benefits of learning features bi-directional which helps to increase the accuracy of the suggested method.
- A text-based representation model used n-gram for feature extraction and pre-processed the dataset. The attention mechanism is used to improve accuracy by giving more importance to one particular node.

We conducted a comparative experiment to determine the effectiveness of the proposed architecture based on increased accuracy. The article is divided into the following sections. The materials and methods used to forecast ASD are briefly covered in the second part. Section 3 covers the experimental analysis and finding discussion in detail. The result and discussion of proposed work organizational structure is explained in section 4. Section 5 presents the future work and concludes the study.

2. METHOD

Data collection was carried out using a web scraping technique, which collects unstructured data from the web. These unstructured data are transformed as metadata that can be saved and analyzed in the database. The analysis was performed on a significant amount of information generated by collecting data on Twitter in order to describe the social media conversations around ASD and, consequently, understand how various individuals view the subject to examine their behavior. The data collection was done by using hashtags and keywords to find tweets that served as the foundation for the data collection procedure. The basic process for the data collection is shown in Figure 1. In Twitter and other social media, a hashtag (#) is used to represent a specific topic and keywords. After conducting research and evaluating the data we have identified the hashtags and keywords associated with trending topics, on ASD that were generated from

English tweets. These include #autism, #ASD, #Autis, #Actually Autistic, #Autism Acceptance, #Autism Speaks, #Asking Autistics, # Autism parent, #Autchat, #Autism Awareness, #World Autism Awareness Day.

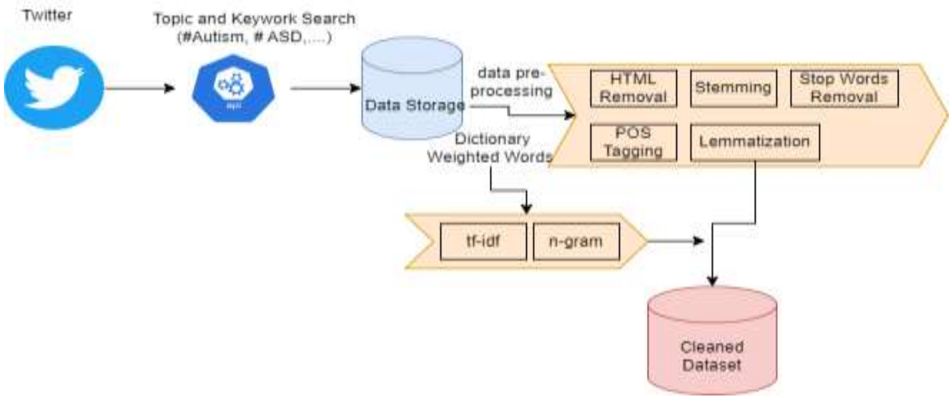

Figure 1. Data collection process of autism

2.1. Dataset description

Pre-processing is an essential step in improving the text's clarity by eliminating noise, irrelevant repetitions, removal of HTML tags, URLs, and so on, which makes deep models operate more efficiently. Contrarily, social media users speak informally and no standardly, and noise needs to be thoroughly processed before being shared on the network. We use a variety of pre-processing techniques, including lemmatization, the elimination of non-unicode and non-English characters, URL replacement, and user handlers, to eliminate all noise and prepare the text for training.

2.2. Model structures

2.2.1. LSTM and Bi-LSTM model

In this section, the models of LSTM and BiLSTM are discussed. LSTM which is an advanced model of RNN that is deeply used in DL. The traditional RNN had the difficulty of learning long-term dependencies in the network whereas LSTM addressed this issue by introducing a memory cell that holds the data for a long time. Three network gates regulate the memory cell: input, output, and forget gates. These gates decide which data to insert, delete, and release from the memory cell. Data added to a memory cell is managed by an input gate. Data erased from a memory cell is managed by a forget gate. Furthermore, the output gate controls the information transmitted by the memory cell. LSTM networks can acquire long-term dependencies due to their capability of selectively retaining or discarding information as it traverses the network.

Unlike LSTM, BiLSTM the input sequence is processed both forward and backward. BiLSTM comprises two LSTM layers: the forward unit will process the input sequence and the other in the backward direction. Typically, the final output is formed at each time step by concatenating the hidden states from both directions. Because of its ability to process data in both directions, the network is well-suited for tasks that require an understanding of both the past and the future. Specifically, this feature allows the network to take into account both the previous and the subsequent context for each time step.

2.2.2. Attention mechanism

Attention mechanism in DL models is used to enhance the ability to process the data by giving importance to selective elements. This process will give prediction accuracy and high computational efficiency. The attention mechanism can simultaneously assign distinct weights according to different features to overcome the information loss caused by extended sequences in LSTM and increase information processing efficiency through differentiated weight assignment. Specifically, it can choose to discard unimportant information and assign greater weights to critical information. Thus, adding the attention mechanism could increase the accuracy of ASD prediction even further.

2.2.3. N-BiLSTM with attention mechanism

This section gives an insight of intelligent system uses a DL architecture that incorporates an attention mechanism. The model performs sentiment analysis on text data by concentrating on identifying particular textual patterns that might be suggestive of sentiments associated with ASD. By using the attention mechanism, the model can highlight significant information in the text by giving different weights to different words or phrases. Initially, the data is pre-processed and used n-gram for feature extraction. The output of the feature extraction is fed to the bidirectional LSTM in which it is capable of capturing the data from past and future words in a sentence. One or more fully connected layers, frequently with activation functions the attention mechanism.

These layers assist in connecting the final prediction to the features that were extracted from the BiLSTM and attention layers. A single neuron with a sigmoid activation function makes up the final output layer. This works well for tasks requiring binary classification, such as determining whether a subject has ASD or not. For binary classification tasks, the loss function can be used to find the difference between the expected and real labels. Finally, the model is evaluated with the performance measure for the proposed model by accuracy, precision, recall and F1*score.

2.3. Adam optimizer

Adam optimizer stands for adaptive moment estimation, an adaptive learning technique that improves the training speed of DL models. Based on the standard gradient descent history, it analyses and customize the learning rate. It is incorporated with two optimization technique called root mean square propagation and momentum. The momentum technique uses exponentially weighted average to accelerate the gradient and it can be written as (1) and (2).

$$v_{t+1} = v_t - \alpha n_t \quad (1)$$

Where,

$$n_t = \beta n_{t-1} + (1-\beta)\left[\frac{\partial O}{\partial v_t}\right] \quad (2)$$

n_t is represented as aggregation of gradients at *t* time, n_{t-1} is a previous *t* time of gradients, v_t is a weight at *t* time, v_{t+1} is *t+1* time weight, β is moving average, ∂O is the loss function, ∂v_t is the weight at time *t*. Root mean square propagation is an additional adaptive learning in Adam optimizer which helps to improve AdaGrad. This algorithm takes exponential moving average instead of taking cumulative sum and represented in (3) and (4).

$$v_{t+1} = v_t - \frac{\alpha_t}{(w_t + \omega)^{1/2}} * \left[\frac{\partial O}{\partial v_t}\right] \quad (3)$$

Where,

$$w_t = \beta w_{t-1} + (1-\beta) * \left[\frac{\partial O}{\partial v_t}\right]^2 \quad (4)$$

v_t is the weight at *t* time, v_{t+1} represents the weight at *t+1*, ∂O is the loss function, α_t is the learning rate of time t. $w_t$is the past sum of square gradient and β is parameter of moving average. Now, the Adam optimizer by taking root mean square propagation and momentum together which is represented in (5).

$$n_t = \beta_1 n_{t-1} + (1-\beta_1)\left[\frac{\partial O}{\partial v_t}\right], w_t = \beta_2 w_{t-1} + (1+\beta_2)\left[\frac{\partial O}{\partial v_t}\right]^2 \quad (5)$$

The using of BiLSTM with an attention mechanism is justified by the need to capture complex linguistic patterns in the text data, which are crucial for accurately predicting ASD. The attention mechanism specifically helps in focusing on the most relevant parts of the input sequence, thereby improving the model's interpretability and performance.

3. EXPERIMENTAL ANALYSIS

This section discussed the experimental study of the proposed system. This article collected data from around 1000 from Twitter with a keyword search by using Twitter API. Once the dataset is completed with the pre-processed stage, the entire dataset is divided in an 80:20 ratio, with 80% allotted for training and 20% for testing.

3.1. Measuring the n-BiLSTM with attention-ASD method performance

The confusion matrix is one type of evaluating the performance of the model which is comprised of true positive, true negative. To identify the efficiency of n-BiLSTM with Attention mechanism, the evaluation metrics are calculated in terms of accuracy, precision. recall and F1*score in (6) to (9). The evaluation equation related to these parameters is represented as:

$$Acc_{asd} = \frac{TP_{asd} + TN_{asd}}{TP_{asd} + FP_{asd} + TN_{asd} + FN_{asd}} \quad (6)$$

$$Pre_{asd} = \frac{TP_{asd}}{TP_{asd} + FP_{asd}} \quad (7)$$

$$Rc_{asd} = \frac{TP_{asd}}{TP_{asd} + FN_{asd}} \quad (8)$$

$$F1 * Score_{asd} = \frac{2Pre_{asd}\, Rc_{asd}}{Pre_{asd} + Rc_{asd}} \quad (9)$$

The confusion matrix is the combination of true class and predicted class. Figure 2 shows the ASD confusion matrix for the proposed algorithm. Figure 3 provides the comparison of jaundice and ASD disease for male and female. The plot compared with almost 1000 data in which 525 is female and 475 is male. While analysing the count plot, the autism has reached higher count than jaundice and also when comparing between female and male, the affected ratio of female is higher than male.

During the network training stages, the inputs to the proposed model consist of a Glove with 200 dimensions and embedding initialization created by combining n-gram weighting with additional parameters. We implemented the attention mechanism after obtaining the feature from the concatenation layer and receiving the output from the feature extraction process, which fed into the BiLSTM layer. The BiLSTM has a kernel regularization rate of 0.001, a dense size of 128, and a batch size of 256. The binary classifier uses a sigmoid function after the dense layer. Lastly, we used the Adam optimizer to determine the learning rate of the model and binary cross-entropy to train it. The proposed algorithm was compared with 3 DL algorithms such as CNN, LSTM and BiLSTM as indicated in Table 1. The CNN model has got an accuracy of 80% whereas the LSTM and BiLSTM outperformed the CNN model with an 82% accuracy and 84% indicating that the BiLSTM has a good accuracy rate in detecting ASD.

The CNN model was unable to identify characteristics over time steps, which is required to detect ASD. To improve the sentence's context, the LSTM model could only store data about its prior status; it was unable to retain information about its subsequent status. Whereas, the BiLSTM model enhanced the performance of memory cell and also has the ability to process the data in the forward and backward direction. So, the accuracy improved than CNN and LSTM. The proposed model with hybrid n-BiLSMT-attention ASD model improved the accuracy rate of 90% and 92 of precision, 95 of recall, and 93 of F1*score. The combination network with attention-based enables the model to learn deeply which helps to detect and improve the accuracy. The above result demonstrations that the n-BiLSTM-attention ASD model delivered a good performance rate in terms of accuracy.

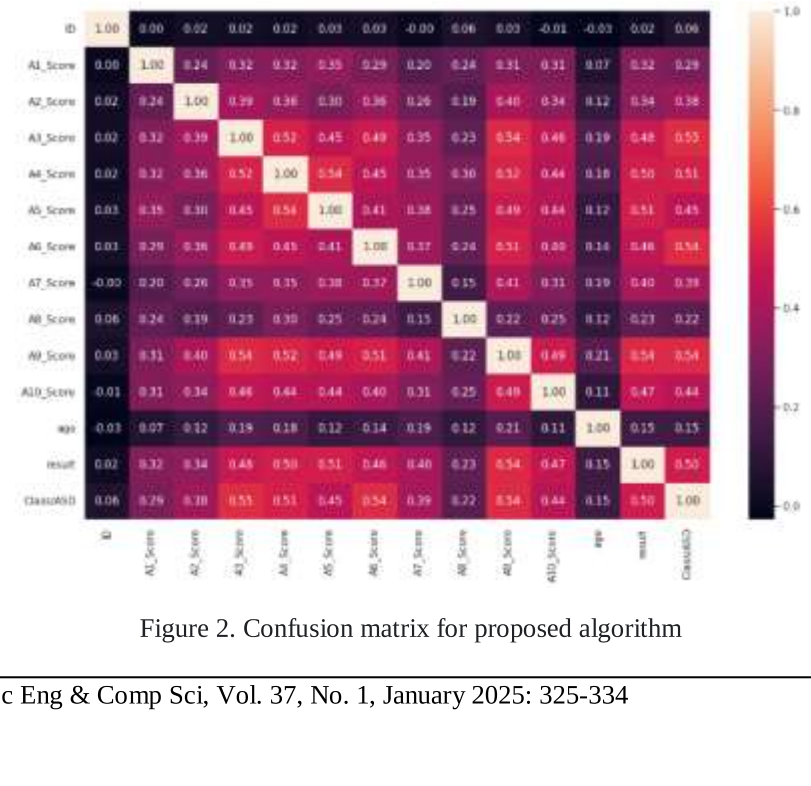

Figure 2. Confusion matrix for proposed algorithm

Table 2 represents that the CNN with Adam optimizer and LSTM with Adam optimizer delivered comparable results of 86% accuracy whereas, the BiLSMT with Adam optimizer resulted in 91% accuracy. To increase the suggested models accuracy rate we employed the Adam optimizer. We optimized with the n-BiLSTM-attention model to improve the accuracy rate and resulted in 98% of accuracy whereas we obtained 90% of accuracy without Adam optimizer.

Figure 4 represents error rate between predicted and actual value of DL model. From the proposed model, the training loss reaches about 0.05% whereas the validation loss has reached 0.06% with 100 epochs. Figure 5 depicts the accuracy rate of training and validation for the proposed model. The training accuracy has reached 99% whereas the validation accuracy has reached 98% for the proposed model with 100 epochs.

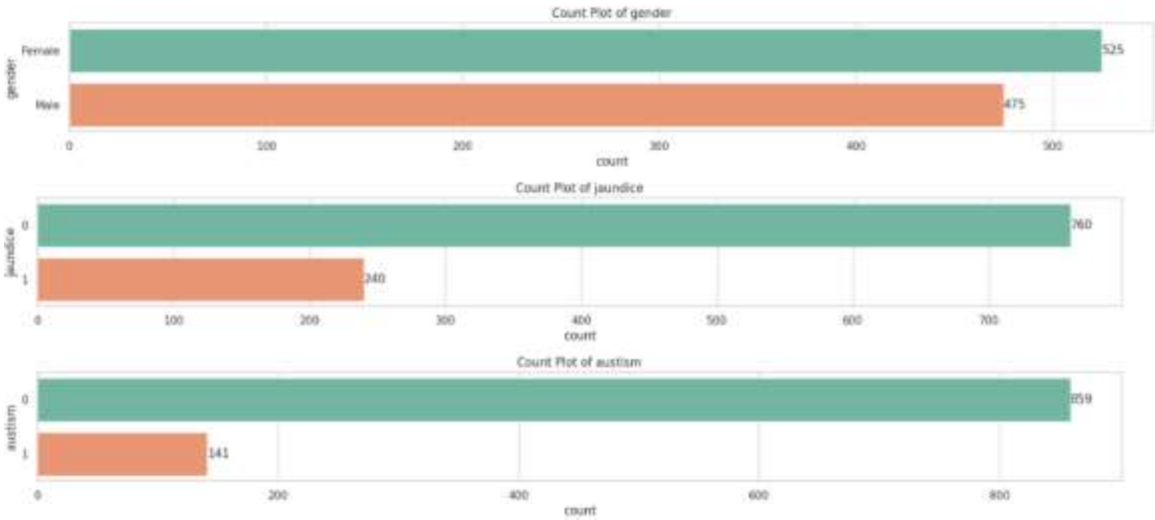

Figure 3. Comparison between jaundice and autism

Table 1. Performance comparison of the suggested system with various algorithms

Models	Accuracy	Precision	Recall	F1 score
CNN	80	77	79	78
LSTM	82	79	80	79
BiLSTM	84	82	84	83
n-BiLstm-attention	90	92	95	93

Table 2. Results obtained by using Adam optimizer

Models	Accuracy	Precision	Recall	F1 score
CNN–Adam optimizer	80	77	79	78
LSTM-Adam optimizer	82	79	80	79
BiLSTM-Adam optimizer	84	82	84	83
n-BiLstm–Attention-Adam optimizer	90	92	95	93

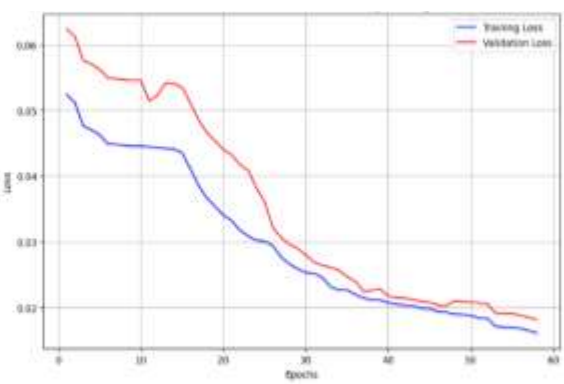

Figure 4. Training and validation loss of n-BiLSTM-attention ASD model

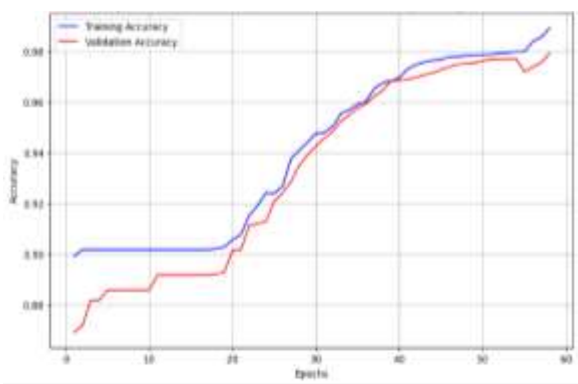

Figure 5. Training and validation accuracy of n-BiLSTM-attention ASD model

4. RESULTS AND DISCUSSION

In this study, we proposed the n-BiLSTM-attention model for the detection and analysis of ASD. The incorporation of the attention mechanism allowed our model to focus on specific elements within the input data, significantly contributing to its impressive accuracy of 98%. We utilized TwitterAPI to scrape ASD-related keywords for both training and testing purposes, ensuring the model's applicability and reliability in the context of ASD prediction. Rigorous preprocessing methods were employed to eliminate biases, noise, and potential confounding variables, enhancing the model's ability to generalize across a wide range of input data. The preprocessing phase included the use of the n-gram method for feature extraction, with the processed output subsequently fed into the n-BiLSTM-attention ASD algorithm.

We also employed the attention mechanism to improve the accuracy by providing more importance to one particular an input sequence. The accuracy was achieved 90% when using attention mechanism later the Adam optimizer was also used to improve the training speed. Additionally, we also observed that our proposed model is more significant and efficient when compared to existing work as tabulated in Table 3. After adding Adam optimizer to our model, we achieved 98% of validation accuracy and 99% of training accuracy with the loss of 0.05, 0.06 for training loss and validation loss respectively. The accuracy measure for few models in existing work is compared with our proposed model as depicted in Figure 6. Apart from the impressive 98% overall accuracy, we evaluated the system's performance using an extensive set of evaluation metrics. The model's strengths and possible areas for improvement in particular ASD prediction disciplines were more fully understood with the computation of F1-score, precision and recall.

Our system's intelligent applications have promising practical applications. Due to sentiment analysis's high accuracy in predicting ASD, people can benefit from early intervention and support. It could have a revolutionary effect on clinical practice, especially in settings with limited resources, allowing for the prompt identification and individualized treatment of people on the autism spectrum.

Table 3. Comparison of our proposed model with existing works

Author	Models	Accuracy
Raj and Masood [6]	CNN	96%
Alkahtani *et al.* [9]	MobileNetV2 and VGG19	92%
Sherkatghanad *et al.* [17]	CNN	70%
Abdelwahab *et al.* [21]	RF	89.23%
Proposed model-n-BiLSTM-attention ASD		**98%**

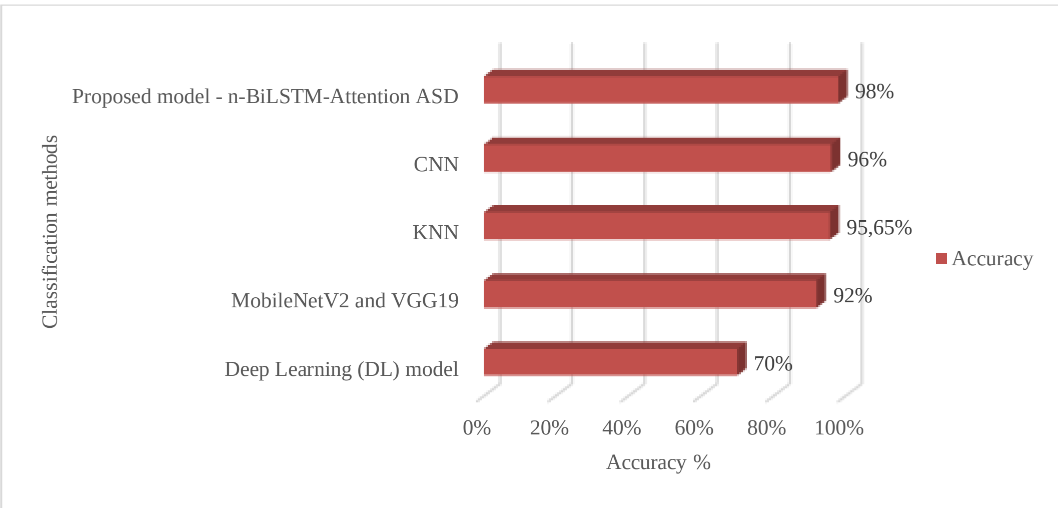

Figure 6. Comparison graph of proposed model with existing model

5. CONCLUSION

Autism disorder is a complicated neurodevelopment disorder that often shows symptoms at an early stage. Sentiment analysis is a viable method for looking for possible ASD indicators in text-based data because people with ASD frequently have unique communication styles. This paper reports on an experimental analysis of an intelligent system that uses textual data to predict ASD by using sentiment analysis and an attention mechanism. The experimental analysis of intelligent systems using sentimental analysis in the detection of ASD using the n-Bi-LSTM-attention mechanism demonstrated more effectively and efficiently. The integration of sentimental analysis and attention-based methods provides a promising approach to detecting an early detection of ASD which helps to treat it at an earlier stage. The study showcases the potential of sentiment analysis and DL for early ASD detection, enabling timely interventions. It highlights the importance of integrating NLP and DL in healthcare to tackle complex medical challenges. These methods could revolutionize the diagnosis and management of neurodevelopmental disorders, improving the lives of those affected by ASD. The proposed system achieved 98% of validation accuracy and 99% of training accuracy. In future, we would like work to improve the system's predictive capability by exploring the multi-model dataset of ASD related videos and images. Clinical trials and real-world applications are also essential to confirm the system's usefulness in early ASD detection.

REFERENCES

[1] S. R. Shahamiri and F. Thabtah, "Autism AI: a new autism screening system based on artificial intelligence," *Cognitive Computation*, vol. 12, no. 4, pp. 766–777, Jul. 2020, doi: 10.1007/s12559-020-09743-3.

[2] A. Garg *et al.*, "Autism spectrum disorder prediction by an explainable deep learning approach," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1459–1471, 2022, doi: 10.32604/cmc.2022.022170.

[3] S. S. Joudar, A. S. Albahri, and R. A. Hamid, "Triage and priority-based healthcare diagnosis using artificial intelligence for autism spectrum disorder and gene contribution: A systematic review," *Computers in Biology and Medicine*, vol. 146, p. 105553, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105553.

[4] M. A. Mareeswaran and K. Selvarajan, "A computational intelligent analysis of autism spectrum disorder using machine learning techniques," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 1, pp. 807–816, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp807-816.

[5] R. Kashef, "ECNN: Enhanced convolutional neural network for efficient diagnosis of autism spectrum disorder," *Cognitive Systems Research*, vol. 71, pp. 41–49, Jan. 2022, doi: 10.1016/j.cogsys.2021.10.002.

[6] S. Raj and S. Masood, "Analysis and detection of autism spectrum disorder using machine learning techniques," *Procedia Computer Science*, vol. 167, pp. 994–1004, 2020, doi: 10.1016/j.procs.2020.03.399.

[7] M. A. Bayram, İ. Özer, and F. Temurtaş, "Deep learning methods for autism spectrum disorder diagnosis based on fMRI images," *Sakarya University Journal of Computer and Information Sciences*, vol. 4, no. 1, pp. 142–155, Apr. 2021, doi: 10.35377/saucis.04.01.879735.

[8] L. Gong, X. Zhang, T. Chen, and L. Zhang, "Recognition of disease genetic information from unstructured text data based on BiLSTM-CRF for molecular mechanisms," *Security and Communication Networks*, vol. 2021, pp. 1–8, Feb. 2021, doi: 10.1155/2021/6635027.

[9] H. Alkahtani, T. H. H. Aldhyani, and M. Y. Alzahrani, "Deep learning algorithms to identify autism spectrum disorder in children-based facial landmarks," *Applied Sciences (Switzerland)*, vol. 13, no. 8, p. 4855, Apr. 2023, doi: 10.3390/app13084855.

[10] M. F. Rabbi, S. M. M. Hasan, A. I. Champa, and M. A. Zaman, "A convolutional neural network model for early-stage detection of autism spectrum disorder," in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, Feb. 2021, pp. 110–114, doi: 10.1109/ICICT4SD50815.2021.9397020.

[11] S. R. Arumugam, S. G. Karuppasamy, S. Gowr, O. Manoj, and K. Kalaivani, "A deep convolutional neural network based detection system for autism spectrum disorder in facial images," in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Nov. 2021, pp. 1255–1259, doi: 10.1109/I-SMAC52330.2021.9641046.

[12] V. Kavitha and R. Siva, "Classification of toddler, child, adolescent and adult for autism spectrum disorder using machine learning algorithm," in *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Mar. 2023, pp. 2444–2449, doi: 10.1109/ICACCS57279.2023.10112932.

[13] J. Deng, M. R. Hasan, M. Mahmud, M. M. Hasan, K. A. Ahmed, and M. Z. Hossain, "Diagnosing autism spectrum disorder using ensemble 3D-CNN: a preliminary study," in *2022 IEEE International Conference on Image Processing (ICIP)*, Oct. 2022, pp. 3480–3484, doi: 10.1109/ICIP46576.2022.9897628.

[14] N. A. Ali, S. A. Radzi, A. S. Jaafar, and N. K. Nor, "ConVnet BiLSTM for ASD classification on EEG brain signal," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 18, no. 11, pp. 77–94, Aug. 2022, doi: 10.3991/ijoe.v18i11.30415.

[15] N. Kaur and G. Gupta, "Refurbished and improvised model using convolution network for autism disorder detection in facial images," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 29, no. 2, pp. 883–889, Feb. 2023, doi: 10.11591/ijeecs.v29.i2.pp883-889.

[16] W. Jiang *et al.*, "CNNG: a convolutional neural networks with gated recurrent units for autism spectrum disorder classification," *Frontiers in Aging Neuroscience*, vol. 14, p. 948704, Jul. 2022, doi: 10.3389/fnagi.2022.948704.

[17] Z. Sherkatghanad *et al.*, "Automated detection of autism spectrum disorder using a convolutional neural network," *Frontiers in Neuroscience*, vol. 13, p. 482737, Jan. 2020, doi: 10.3389/fnins.2019.01325.

[18] S. G. Karuppasamy, D. Muralitharan, S. Gowr, S. R. Arumugam, E. A. Devi, and K. Maharajan, "Prediction of autism spectrum disorder using convolution neural network," in *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, Apr. 2022, pp. 1096–1100, doi: 10.1109/ICOEI53556.2022.9776657.

[19] M. A. Aghdam, A. Sharifi, and M. M. Pedram, "Diagnosis of autism spectrum disorders in young children based on resting-state functional magnetic resonance imaging data using convolutional neural networks," *Journal of Digital Imaging*, vol. 32, no. 6, pp. 899–918, Dec. 2019, doi: 10.1007/s10278-019-00196-1.

[20] E. Giarelli *et al.*, "Sex differences in the evaluation and diagnosis of autism spectrum disorders among children," *Disability and Health Journal*, vol. 3, no. 2, pp. 107–116, Apr. 2010, doi: 10.1016/j.dhjo.2009.07.001.

[21] M. M. Abdelwahab, K. A. Al-Karawi, E. M. Hasanin, and H. E. Semary, "Autism spectrum disorder prediction in children using machine learning," *Journal of Disability Research*, vol. 3, no. 1, pp. 1–9, Jan. 2024, doi: 10.57197/JDR-2023-0064.

[22] S. Murugaiyan and S. R. Uyyala, "Aspect-based sentiment analysis of customer speech data using deep convolutional neural network and BiLSTM," *Cognitive Computation*, vol. 15, no. 3, pp. 914–931, May 2023, doi: 10.1007/s12559-023-10127-6.

[23] Y. Fan, H. Xiong, and G. Sun, "DeepASDPred: a CNN-LSTM-based deep learning method for autism spectrum disorders risk RNA identification," *BMC Bioinformatics*, vol. 24, no. 1, p. 261, Jun. 2023, doi: 10.1186/s12859-023-05378-x.
[24] V. P. Sudha and M. S. Vijaya, "Recurrrent neural network based model for autism spectrum disorder prediction using codon encoding," *Journal of The Institution of Engineers (India): Series B*, vol. 103, no. 2, pp. 599–605, Apr. 2022, doi: 10.1007/s40031-021-00669-4.
[25] A. Amirbay *et al.*, "Development of an algorithm for identifying the autism spectrum based on features using deep learning methods," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 5, pp. 5513–5523, Oct. 2024, doi: 10.11591/ijece.v14i5.pp5513-5523.
[26] R. A. J. de Belen, T. Bednarz, A. Sowmya, and D. Del Favero, "Computer vision in autism spectrum disorder research: a systematic review of published studies from 2009 to 2019," *Translational Psychiatry*, vol. 10, no. 1, p. 333, Sep. 2020, doi: 10.1038/s41398-020-01015-w.
[27] Y. Tao and M.-L. Shyu, "SP-ASDNet: CNN-LSTM based ASD classification model using observer ScanPaths," in *2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Jul. 2019, pp. 641–646, doi: 10.1109/ICMEW.2019.00124.
[28] L. Corti, M. Zanetti, G. Tricella, and M. Bonati, "Social media analysis of Twitter tweets related to ASD in 2019–2020, with particular attention to COVID-19: topic modelling and sentiment analysis," *Journal of Big Data*, vol. 9, no. 1, p. 113, Nov. 2022, doi: 10.1186/s40537-022-00666-4.

BIOGRAPHIES OF AUTHORS

Murali Anand Mareeswaran completed his M.Sc. in Loyola College, Chennai, India. He is currently working as an Assistant Professor in the Department of Computer Science and Applications, SRM Institute of Science and Technology, Vadapalani, Chennai, Tamil Nadu, 600026, India. At present, he is pursuing his Ph.D. in the Department of Computer Science under the guidance of Dr. S. Kanchana, Assistant Professor, Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, 603203, India. He can be contacted at email: muraliam@srmist.edu.in.

Dr. Kanchana Selvarajan obtained her Ph.D. degree by Bharathiyar University, Coimbatore. She has 15 years of teaching experience. She is currently working as an Assistant Professor in the Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, 603203, India. She has published more the 20 research papers in national and international journals. Her area of research is machine learning, data mining, block chain, and IoT. Currently, she is guiding four Ph.D. Scholars. She can be contacted at email: kanchans@srmist.edu.in.