Detection of Autism Spectrum Disorder (ASD) Symptoms using LSTM Model

PRASENJIT MUKHERJEE[1,2], MANISH GODSE[3], BAISAKHI CHAKRABORTY[4]

[1]Department of Technology,
Vodafone Intelligent Solutions,
Pune,
INDIA

[2] Department of Computer Science,
Manipur International University,
Manipur,
INDIA

[3] Department of IT,
Bizamica Software,
Pune,
INDIA

[4] Department of Computer Science and Engineering,
National Institute of Technology,
Durgapur,
INDIA

Abstract: - Autistic children will often exhibit certain behaviors that are unique to them and that are not typical of neurotypical children. Parents will become familiar with these patterns over time and will be able to use this knowledge to answer questions about their child's behavior. Deep learning models are very useful to solve critical problems in the healthcare domain. Detection of ASD at the early age of a child is a challenging task. Recent research reveals that there is an increasing trend of ASD among children. Communication, eye contact, social behavior, and education are very poor for those who suffer from ASD. The proposed research work has been done to detect ASD symptoms in a child. Data has been collected from the various autism groups from social sites and organizations that are working on special children. A Deep learning model like the Long-Short Term Memory (LSTM) model has been used to detect the sentiment of parents' dialog. LSTM is the most popular deep learning model that can able to solve complex natural language problems. The proposed LSTM model has been trained with prepared data and accuracy is 97% according to the prepared data.

Key-Words: - LSTM, Deep Learning, Autism Detection, Machine Learning, ASD dataset, BERT Cosine.

Received: March 22, 2023. Revised: November 19, 2023. Accepted: December 22, 2023. Published: February 20, 2024.

1 Introduction

Autism spectrum disorder (ASD), which is identified as an imbalance in brain functioning, can be described as a neurodevelopmental problem, [1]. Individuals with ASD face difficulties in verbal and non-verbal communication as well as social interactions. Because of these complications, their interpersonal skills and quality of life are greatly hampered. In the recent past, the World Health Organization, in 2019, released a report related to the prevalence of ASD and it has been estimated that 1 in 160 children is affected by this, [2]. Due to the absence of objective interpretative mechanisms for ASD, clinical diagnosis of ASD faces notable challenges, [3]. The clinical diagnosis of ASD used by the presented methodology depends primarily on behavioral assessment, yet the accuracy of the diagnosis is undermined by the considerable heterogeneity of ASD and the diverse range of clinical symptoms present, [4]. Depicted by variations in symptoms, ASD can be categorized into three distinct subtypes. These subtypes comprise ASD, Asperger's disorder (APD), and pervasive developmental disorder not otherwise specified (PDD-NOS).A specific profile of characteristics is represented by each sub-type that allows a more refined understanding of the diverse presentation of ASD within the stretch as in, [5]. Over the last two decades, due to the considerable advancements in neuroscience research, certain biomarkers have been identified which is a step in the right direction to the understanding of neural mechanisms associated with ASD. In this endeavor an important role has been played by functional brain imaging modalities, allowing the characterization of these biomarkers as described in, [6]. Parallel to this trend, the utilization of artificial intelligence (AI) models within the domain of medical diagnosis has a strong exponential, with a particular priority on psychiatric disorders.The prospect of the application of AI technology is crucial in augmenting diagnostic capabilities, including those associated with psychiatric conditions. It is the predictive power of AI through which medical professionals are assured to gain precious insights and can enhance their expertise to diagnose accurately and understand complex psychiatric disorders, [7]. ASD, which is a complex disorder, has many factors that can contribute to the development of it. Moreover, there is no multifunctional approach to managing the symptoms. There are many problems to fulfill the individual's needs who suffer from ASD. A right combination of interventions is needed which are child specific to address the issue to identify the type of ASD and the corresponding treatment. The process, however, is long and complicated. Moreover, there is a lack of awareness about ASD among the masses and little educated people are not free to discuss the issue as they feel threatened by social stigma and discrimination. There is a lack of funding and resources for research related to ASD treatments and therapies. As a result, sufficient caregiving is limited and expensive as in, [8]. Social skills, communication, ability to regulate emotions, and behavior are affected by ASD. The signs and symptoms of ASD can be identified by early screening, and then appropriate interventions are addressed and correspondingly, the relevant treatments are applied. The quality of life for people with ASD can be improved by the early interventions that help them reach their full potential. The ASD diagnosis requires a comprehensive assessment of the child's behavior, communication, and developmental milestones and the process is complex. Additionally, families in low-income communities often lack access to those services because the cost of diagnosis and treatment can be extensively expensive as in, [9]. Early intervention strategies can help early diagnosis, the severity of the symptoms thus can be reduced and the overall functioning of the individual can be improved. The associated healthcare costs can also be reduced by early diagnosis because any potential problems can be identified early which leads to more effective and less expensive treatments. Since interventions are most fruitful when delivered early in life, early detection of ASD can help minimize the long-term economic implications of the condition. Machine learning algorithms enable to identify patterns by analyzing the large amounts of data in a short period, which can help diagnose the disorder accurately. The time and resources needed to address the disorder can be helped by this, and the accuracy of the diagnosis process can be reduced as in, [10]. Computers can analyze data quickly, accurately, and cheaply with the intervention of AI and ML. Detection of signs of autism earlier than ever before and minimizing the cost and time involved with traditional screening tests are helped by the application of AI and ML as in, [11]. Enormous research has evolved in numerous areas of healthcare system development, including the field of security, [12] and the improvement of healthcare quality and services, [13]. The incorporation of artificial intelligence (AI) technology benefited healthcare providers by significantly empowering paramedics to support necessary preliminary treatments ignoring the late arrival of medical specialists, and thus it provides more cost-effective healthcare delivery at an initial stage as in, [14]. Several noteworthy studies have contributed to the progress of medical diagnostics. The authors in, [15], [16] and [17], have made significant contributions in this area, conducting valuable research in medical diagnostic techniques. Their findings have helped advance the field, enabling more accurate and efficient diagnoses, ultimately leading to improved patient care and outcomes.

The proposed research work has stated the detection of ASD symptoms from parents' dialogues. As the first step, the sentiment analysis has been described using parents' dialogues with autistic children. Data has been collected from the organizations of ASD children and social sites. A parent of an autistic child spends maximum time with their autistic child and they are well aware of the ASD symptoms of their baby. The proposed collected data consisted of parents' experiences and thoughts about their ASD children. The dataset has been prepared from the collected data. Each sentence is labeled as true (1) if the sentence contains many words that indicate ASD symptoms. If the sentence does not contain any words that are related to ASD symptoms, then it will be marked as false (0). The first approach is the LSTM model which is the most popular type of RNN model in NLP. The LSTM model has been trained using the proposed dataset for sentiment prediction which is related to ASD symptoms. Section 2 discusses Related works on ASD, Section 3 elaborates on the system architecture for ASD detection, Section 4 focuses on results, and Section 5 points out the limitations. Section 6 deals with conclusion of the proposed work and section 7 winds up the paper by discussing possible future work on ASD detection.

2 Related Works

Autism Spectrum Disorder (ASD) is a notable health issue among children, and thus healthcare researchers have come forward to devise methods of detection of this disorder. Artificial Intelligence (AI) has come up as a significant tool for studying and treating ASD, and various studies have employed AI-based methods to make people aware of this disorder. In addition, other mental health issues have been addressed by AI, and various important works in this field have been incorporated in this section on related research. People with ASD face an obstacle related to social interactions and communication, and more often they have difficulty in understanding and replying to social indications. Sensory sensitivities are also experienced by them which can lead them to respond to certain sounds or textures in an exaggerated manner. This disorder has also other behaviors such as repetitive actions and limited interests. Providing the necessary support and interventions to the children is very important which can help them handle their social and communicational difficulties. ASD detection and treatment face a lack of resources and knowledge. For example, many parents are unaware of the developmental milestones that can specify ASD.~~,~~ Children with ASD are not accurately and easily diagnosed because of lack of enough diagnostic tools. ASD is a complex and multifactorial disorder comprising a variety of environmental and genetic factors. The study is difficult because each case may have a unique set of symptoms. Additionally, many of the present diagnostic tests are high-priced and time-consuming. To identify correlations between certain traits and the presence of ASD and to analyze data from numerous sources, the paper examines the extent to which supervised machine learning algorithms may be applied for ASD detection. To identify patterns by this analysis is the goal of the paper to be used as markers for early diagnosis of ASD. Traditional machine learning algorithms are worthwhile in identifying features that can differentiate between individuals with and without ASD. However, deep learning architectures have the prospect of improving the classification process accurately by leveraging larger datasets and more complex feature sets as in, [18]. Autism appears from a complex interaction of genetic and environmental influences that disturb brain development. Challenges in social interaction, communication, and repetitive behaviors are typified by the condition. It has been investigated that the genesis of this syndrome is because of genetic predispositions, environmental factors, and lifestyle choices. However, the specific cause remains difficult to find, with the current consensus suggesting a compound, miscellaneous nature of ASD. Access to appropriate care is disrupted by the scarcity of skilled professionals and resources for diagnosing and dealing with ASD. Proper detection and classification of ASD is a complex process which means establishing an accurate biomarker for accurate detection of ASD is difficult. Authors, [18], have stated that traditional machine learning models like decision trees and support vector machines are capable of identifying ASD according to symptoms. However deep learning models are more capable of identifying ASD from high-dimensional data accurately. According to accuracy and efficacy, deep learning models are better for the detection of ASD symptoms as in, [18]. Difficulty in understanding and responding to verbal and nonverbal indications, trouble in social interactions, and difficulty in expressing one's thoughts and feelings characterize ASD. Sensory issues create difficulties among individuals with ASD. They are facing difficulties in communication, eye contact, emotions, and behavior. The symptoms of ASD may vary from one to one who is suffering from ASD. In childhood time, symptom identification is difficult but after childhood, the symptoms of ASD can be detected more prominently. Authors, [19], have analyzed the first dataset that is related to ASD and detected symptoms of ASD. The second dataset contains 965 instances with 16 attributes whereas the third dataset contains 1019 instances with 13 attributes. The Second dataset is related to the emotions and the third dataset is related to motor skills as in, [19]. Authors, [19], have used Convolutional Neural Networks (CNNs) for better accuracy and perfection. CNN [19] can handle high-dimensional data with 99.53%, 98.30%, and 96.88% accuracies according to datasets as in, [19]. The work in, [20], helps to recognize an ASD class where data is collected from various sources, including teachers, parents, and medical professionals. It then utilizes algorithms to analyze the data and provide an output on ASD diagnosis. The dataset includes various attributes like age, gender, nationality, and class, collected from the UCI ML repository and Kaggle, the dimension was reduced to reduce the number of variables and standardization was applied. With the minimized dataset, the DNNPC model can then be used to properly classify the ASD as in, [20]. The number of ASD patients is increasing where researchers opine that environmental, genetic, and neurological factors may be contributing to the increment in cases. Additionally, the syndrome of ASD can differ greatly between individuals, making its diagnosis complicated. There is often a lack of scientific affirmation and validity in the tests for ASD~~,~~ which means that they may not always properly reflect the condition of an individual. Furthermore, since ASD is a disorder in the spectrum, it is complicated to evaluate properly someone's exact level of functioning. Automated diagnosis approaches are faster and more precise than traditional methods. This can help families to get the proper treatment and support in the minimum possible time. It reduces financial load and enhances the quality of life for those having ASD. To analyze the data from various modalities, such as audio, video, and text, and to recognize patterns that can help to point out the characteristics connected with ASD, the DANN model was designed as in, [21]. It has succeeded in differentiating various types of ASD, like low-functioning or high-functioning autism. The ABIDE repository has been used to benchmark the model and assess its activities against methods of standard machine-learning models. The outcomes show that the DANN model was capable of properly classify ASD patients with 0.732 accuracy, which was notably higher than the outcomes of other models that were tested. This exhibits the potency of the model in consolidating various scales of brain functional connectives and private characteristic data for ASD categorization. By using more validation methods, the model was capable of exhibiting a high level of activeness on unseen data, pointing out that the model had learned a nonexclusive depiction of the data and was robust to the outliers. This advocates that the model could be useful in healthcare applications as in, [21]. ASD is characterized by the struggle in communication and social reciprocation, as well as confinement and repetitive behavior. People with ASD can have obstacles with everyday activities and tasks and may require appropriate support to gain their full potential abilities. These tests are often labor intensive and require the existence of trained specialist(s) who through several screening sessions decide to detect ASD and its type or intensity. This a time-consuming process. Authors of, [22], have used six private characteristics; age, sex, handedness, and three individual measures of IQ which identify the Personal Characteristic Data (PCD) of an individual to construct a novel predictive model of ASD detection that suggested that PCD can supply valuable insights into the detection of ASD. To enhance the understanding of the biological basis of autism and to advance better diagnostic and treatment methods are the goals of this project. The models were able to detect differences in brain functioning between ASD and non-ASD large datasets were used. The data was divided into various subsets, then the model was trained on one subset and then examined on another subset. This eliminates data overfitting. Therefore, this study results in a mean AUC (SD) of 0.646 (0.005), followed by a k-nearest neighbor with a mean AUC (SD) of 0.641 (0.004) ensuring the efficiency of clinical ASD detection. Such models could allow earlier identification, more effective intrusions, and better personalized treatments for ASD people as in, [22]. This is because ASD is a complicated neurological disorder that can manifest in a variety of symptoms and the syndrome may be subtle and tough to distinguish from other mental health issues. Additionally, diagnosis can be complex by the fact that the disorder is extremely individual-specific. Machine learning technology analyzes large data sets that would otherwise be more complicated to process and recognize patterns that could be used to properly diagnose disorders of mental health. ML could also be used to detect crucial treatments for these disorders and enhance better interventions. ASD-DiagNet utilizes deep learning to recognize patterns in the fMRI data that can be used to distinguish between those with ASD and those without. The proper detection of ASD is directly related to the symptoms that should be identified correctly. Neurological and biological factors are important parts of ASD symptoms. Authors [23] have used the Auto-Encoder technique to extract features from ASD related data. The parameters of the machine learning model, [23] have been optimized by the Single-layer perceptron (SLP). These two techniques have been applied to make the datasets in various shapes and sizes. The machine learning models can be trained with good accuracies and as a result, the machine learning models can predict more accurately as in, [23]. This is because the referred model uses a blend of convolutional neural networks and transfer learning techniques to acquire knowledge from a large number of data samples and recognize patterns for the detection of the disorder. Furthermore, the model is maximized to run in a fraction of the time in comparison to other methods, which makes it much more practical and value-effective as in, [23]. To have better understanding of the differences in hemodynamic fluctuations between ASD and Typically developing (TD) children, a multilayer artificial neural network has been used in [24]. To notice the differences in hemodynamic reply to an auditory oddball activity between the ASD and TD groups, a study was conducted in this work to identify any differences in neural activation patterns between the two groups. To pull out features from the raw data, (comma) CNNs are utilized, while to capture the temporal dependencies between the features GRUs are applied. By combining the two, CGRNN can properly recognize patterns in the data and then classify them successfully. This approach enabled the authors of, [24], to get appropriate needful features from the data and capture the temporal dynamic characteristics of brain actuation. As a large set of data gets trained, the probability of over-fitting is reduced. The use of DL networks results in a more accurate classification of outcomes than traditional methods which depend on a single layer of neurons.– It achieves 85.0% sensitivity, 92.2% accuracy, and 99.4% specificity. The multilayer neural network CGRNN can recognize features that relate to ASD, even in a short period. Supervised learning models can produce better results in certain cases of ASD. The range of accuracies of these supervised learning models may be between 0.78 to 0.86 where the F1 score is between 0.72 and 0.84. The authors, [25], have stated that NBSVM, [25] is the best model according to 10 train-test cycles. The support vector machine can select nonlinear relationships between the input variables and the output labels where NBSVM, [25], will not be considered for recognizing complex patterns. Authors, [25], have stated that a deep learning model may be a good choice for complex pattern recognition. Researchers were capable of classifying brain activation patterns of patients with ASD by using deep learning algorithms, and then those patterns are used to identify ASD patients in large datasets. It is possible to detect distinct patterns in the brain that may be connected to ASD by comparing the brain imaging data of ASD patients with the control patients. Paper, [26], discusses and elaborates on the possibility of deep learning models being able to accurately differentiate ASD from typically developing (TD) controls, based on a comparison of functional MRI (fMRI) brain scans. The model identified certain regions of the brain that contributed most to the differentiation, which is presented in the results mentioned in, [26]. Image is a crucial source in ASD detection where Content-Based Image Retrieval (CBIR) can play a vital role in detecting ASD. Authors of, [27], use CBIR to extract color, shape, texture, and spatial layout from an image for index representation which will be helpful for data preparation. Authors, [28], have implemented their task using K-means clustering techniques that can be applied on the numerical variables. Such kind of unsupervised machine learning models can be implemented in the healthcare domain where data scientists are using data for the improvement of the services in the healthcare domain as in, [29].

Table 1 describes a comparative analysis between proposed LSTM models with similar kinds of machine learning models that can diagnose mental disorders. The proposed table contains "Models", "Description", "Dataset", "Accuracy", and "Remarks".

Table 1. Comparative Analysis between the proposed LSTM model and similar machine learning models to detect Mental Disorders

Sl.No.	ML Models	Description	Dataset of Models	Accuracy	Remarks
	Machine Learning Models Related to Mental Disorder (Autism Spectrum Disorder)				
1	Logistic Regression, SVM, Naïve Bayes, KNN, [19]	These Traditional Models can detect ASD from ASD screening data.	The dataset contains screening data for Adults, Children, and Adolescent	96.69%, 98.11%, 96.22%, and 95.75%	The dataset contains 20 attributes and preprocessing techniques are used like removing null values and normalization tasks.
2	ANN, CNN, [19]	These advanced models are used to detect ASD from ASD screening data.	The dataset contains screening data for Adults, Children, and Adolescent	97.64% and 99.53%	The dataset contains 20 attributes and preprocessing techniques are used like removing null values and normalization tasks.
3	Deep Neural Network Prediction and Classification (DNNPC), [20]	This model is a deep learning-based classifier that detects ASD among children.	The UCI repository has been used to collect ASD data.	92%	This model is trained in two phases. First, this model is trained with missing data and then it is trained with complete data in the second phase.
4	multichannel DANN, [21]	This model is integrated with multiple layers of neural networks, attention mechanisms, and feature fusion to detect ASD automatically.	ABIDE dataset has been used to classify the ASD subjects from TDC subjects.	73.2%	These experiments have been done on the dataset. The k-fold cross-validation and leave-one-site-out cross-validation have been designed to complete the experiments.
5	KNN, SVM, Decision Tree, Logistic Regression, Random Forest, and Neural Networks, [22]	These models have been used for ASD diagnosis where ABIDE data repositories have been utilized for dataset preparation.	The ABIDE repository has been used to prepare a dataset using six PCD features like interest—age at testing, sex, handedness, full-scale IQ, verbal IQ, performance IQ	61.8%, 54.7%, 59.1%, 57.2%, and 62%	Some fixed data points have been taken from ABIDE data repository to train all the models.
	Proposed LSTM Model in Mental Disorder (Autism Spectrum Disorder)				
6	Proposed LSTM Model	The LSTM model has been to predict positive ASD symptoms from parents' dialogue. This proposed model has been trained with the proposed dataset which is prepared from the collected parents' dialogues.	Parents' Dialogues of Autistic Children in text format from SAHAS- Durgapur, India, and Social Sites.	97%	The data has been accumulated in textual format. The discourse from parents discussing their experiences and perspectives regarding their children with autism is highly valuable. A parent with an autistic child serves as an optimal resource for grasping the patterns of ASD symptoms.

3 Proposed System Architecture for ASD Detection

3.1 Dataset

The dataset has been curated by collecting dialogues from parents who have shared their experiences and thoughts concerning their autistic children. These dialogues were gathered from various social networks and organizations dedicated to the therapy of special children, specifically focusing on communication, behavior, and speech enhancement. An example of some parent dialogues can be found in Table 2. These parent dialogues serve as critical data, providing a wealth of potential ASD symptoms that can be identified. This information is utilized to create a comprehensive dataset for training and testing the proposed machine learning models.

Table 2. Example of Parents' Dialogues

Sl. No.	Parents' Dialogues
1.	Does anyone have advice on how they bring their children out into busy places my son is 14 months he's being tested for ASD and ADHD but he breaks down when we're outside, especially near traffic he will cry to the point he is sick.
2.	Today I found out my son have Level 3 of autism. I wasn't surprised to hear he is autistic because he shows all the signs . Me and his dad moved to AZ from IL and I really haven't had much luck meeting some ppl who has kid(s) that are autistic and some ppl don't be understanding. Looking for more mommy friends or family
3.	Help! Almost three year old is biting and picking her nails. One of her fingers is bleeding and she won't stop it. Screams at me if I try to distract or anything. She's biting them and shouting and crying that it hurts and then straight back into her mouth. Nothing is working she only started this yesterday. Please help
4.	Hello, I am new in this group and I have a query about speech therapist in Oldham, England. My baby is just 2 years and he is not able to speak properly. Please help me.
5.	My 3 yo is toilet trained (thankfully), however we have noticed when he is in a socially awkward situation he has been have little accidents and today full release of his bladder. Is this common in Autism/Sensory Processing Disorder? Thank you.

The dataset has been compiled from the text provided in Table 2. Every sentence has been examined to determine whether it indicates a symptom of ASD. As ASD does not exhibit a fixed set of symptoms for identification, increasing the number of dialogues from parents who have children with autism could lead to a broader range of symptom identification. Additionally, it offers the advantage of providing a more robust training dataset for machine learning models, potentially enhancing their accuracy. A selection of data from the suggested dataset is presented in Table 3.

Table 3. Example data in the proposed dataset

Sl. No.	Comments	Sentiment
1.	At 3 he had 0 words and now he consistently says about 50-75 words.	1
2.	So my son turned 18 in January & at the end of the school yr.	0
3.	I don't think I can handle him because he is very aggressive.	1
4.	when I call him not much eye contact and also he's not talking.	1
5.	I have a doctor appointment coming up next week.	0

The structure of the dataset used in the proposed study is outlined in Table 3, where the first column is the Serial Number, the second column contains Comments, and the third column shows Sentiment. The dataset was developed using text from the dialogues of parents. Each sentence from these parent dialogues was considered and analyzed to determine if it signifies a symptom of ASD. If it was a symptom, it was labeled as 1 (true), and if not, it was labeled as 0 (false). As per Table 3, the comments with serial numbers 1,3, and 4 represent true ASD symptoms, while those with serial numbers 2 and 5 are not indicative of ASD symptoms. This ASD symptom-focused dataset has now been prepared for LSTM training.

Table 4. List of Labels with ASD Problems

Sl. No.	Label	ASD Problems
1.	1	Problem of Speech
2.	2	Problem of Sensory
3.	3	Problem of Behaviour
4.	4	Problem of Special Education
5.	5	Problem of Social Interaction
6.	6	Problem of Eye Contact
7.	7	Problem of Cognitive Behaviour
8.	8	Problem of Hyper Active
9.	9	Problem of Child Psychological
10.	10	Problem of Attention

Table 4 illustrates the relationship between various ASD problems and their corresponding labels. Specifically, Label 1 is associated with the "Speech Problem," while Label 2 and Label 3 correspond to the "Sensory" and "Behavior" problems, respectively. Additional ASD problems are also represented by labels listed in Table 4. After the prediction of sentiment, which is related to ASD symptoms, the proposed system will use Table 4 to identify the problem according to the labels which are associated. The cosine similarity model will use each positive sentence to compare with each sentence inside the dataset that has been given in Table 5. Table 5 contains multiple positive sentences with their respective labels. Each label in Table 5 signifies an ASD problem based on the associations outlined in Table 4. The cosine similarity model, as explained in the Proposed System Flow section, performs a similarity check between the predicted positive sentences and the dataset sentences to identify the sentence with the highest similarity. The system selects the label associated with this highly similar sentence.

Table 5. Dataset for Cosine Similarity check

Sl. No.	Positive Sentences	Label
1.	Hello, I need some advice on potty training for my 4 years old girl. She is non responsive.	7
2.	My 7y old just decided she hates school and does not want to go.	4
3.	My baby is a girl and she is just 3 years old. But I am not hearing any sound from her. Maybe she is nonverbal. Please help me.	1
4.	he's on the move always and always into something.	8
5.	He can concentrate on TV but cannot even have a little attention.	10

The LSTM model has been described with the proposed algorithm in the next sections where this dataset has been utilized to train this model.

3.2 The LSTM Model

The Long Short-Term Memory (LSTM) is an advanced model of Recurrent Neural Network (RNN) which is a deep learning model. The output of the previous step is an input of the current step in RNN where RNN is not able to store data for prediction in a long-term basis. The prediction result is more accurate on current data on RNN. This is the main disadvantage of the simple RNN model. This problem has been solved by the LSTM model which is itself a RNN type model. The LSTM can store data on a long-term basis. LSTM has been widely used in classification problems due to its feedback connectivity. The LSTM can handle not only single data points, but it can also handle complete data streams. According to Figure 1, a neural network and multiple memory blocks are the main structure of the LSTM model. Four units are there in the LSTM model where the cell is the first unit, the input Gate and output gate are as second and third units, and the last one is the forget gate. The flow of information process inside a cell is managed by the three gates. The main work of this cell is to remember values in long time intervals. The cell inside the LSTM model stores information and it acts as a memory where other gates manipulate memory. The input gate takes responsibility to use input value for changing the memory and here sigmoid function is used to allow either 0 or 1. The tanh function is used to assign weights to the data for determining their importance according to the score of -1 to 1. The mathematical equation of the input gate in the n LSTM model has been given below:

$$Input_t = Sigmoid\ (W_{input}.[h_t - 1, X_t] + b_{input})$$
$$T_c = tanh(W_T.[h_t -1, X_t] + b_T)$$

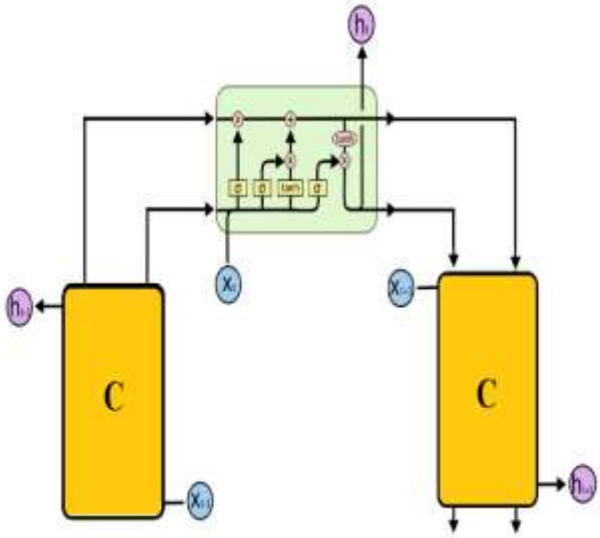

Fig. 1: LSTM Architectural Diagram

The forget gate is used to remove information from the cell using the sigmoid function. For each number in cell state T_c -1, This gate looks into the preceding state (h_t-1) where the input is X_t and generates a number between 0 and 1. The blocks input and cell are used to identify the output where the sigmoid function is used to allow 0 or 1 and tanh function is used for the determination of values (0 and 1). After this, tanh function is used to assign weight to the provided values on a scale of -1 to 1 and multiply it with a sigmoid value. The mathematical equation of this gate has been given below:

$$output_t = sigmoid\ (W_{output}\ [h_t -1, X_t] + b_{output})$$
$$h_t = output_t * tanh\ (T_c)$$

Prasenjit Mukherjee,
Manish Godse, Baisakhi Chakraborty

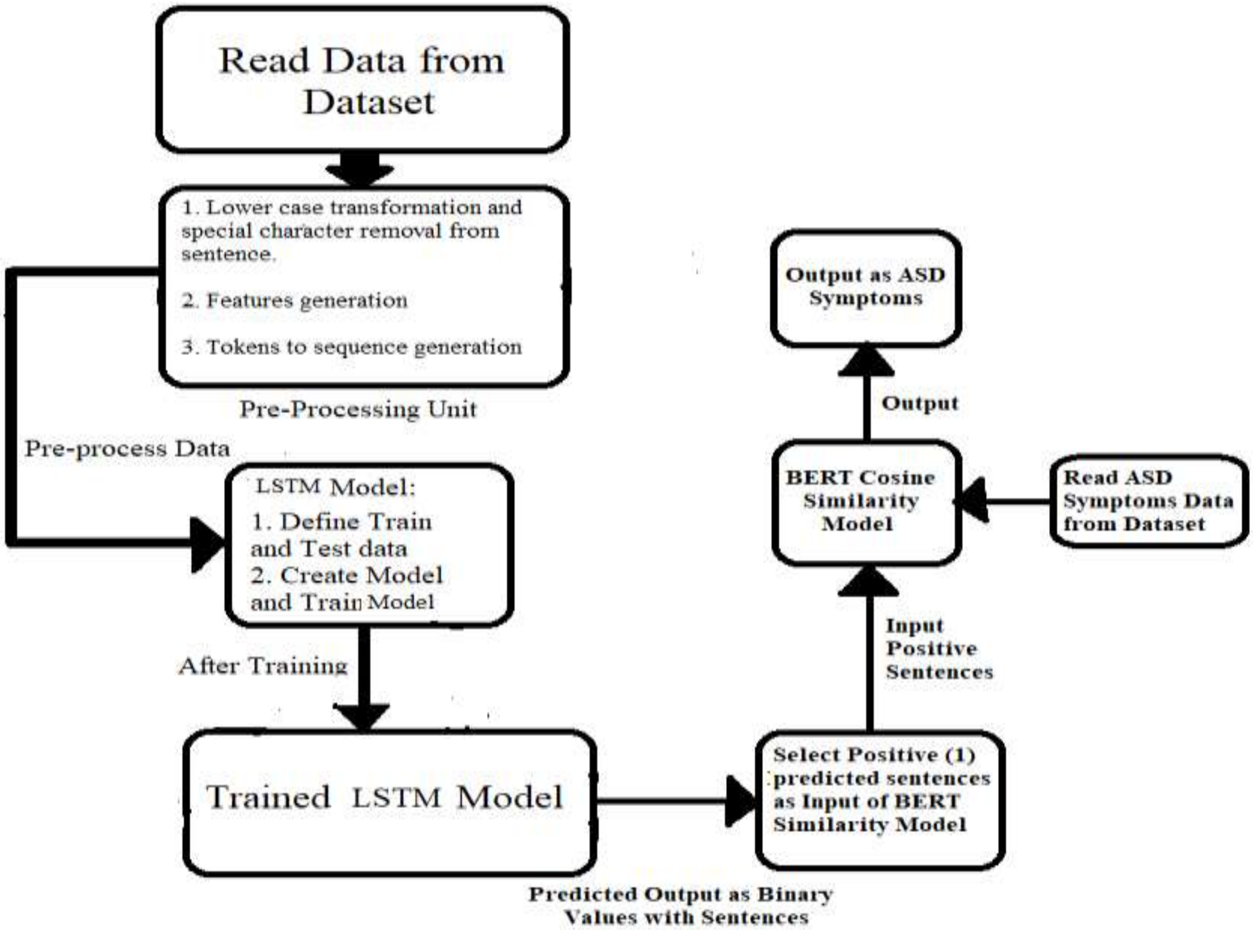

Fig. 2: Architectural Diagram of Proposed System

3.3 The Proposed ASD Detection System

3.3.1 Sentiment Analysis using LSTM Model

According to Figure 2, the proposed system will read the dataset. Then each word from the sentences of the dataset will be transformed into lowercase words if a word contains any uppercase alphabet. Now special characters have been removed from each sentence. After completion of this step, the features have been generated using the tokenizer in a sequence as input with labeled output data to the proposed LSTM model. Before inputting prepared data into the model, the prepared data is split into the training data and testing data. The LSTM model will be created, and training data will be sent to train the model. After completion of this step, the LSTM model can understand the pattern for the sentiment analysis of Parents' dialogues. In the final step, the trained model will be tested using the testing data. After completing the testing phase, the proposed LSTM model is ready to predict the sentiment as 0 or 1 on new data. Any parent can send their experience of their child as a text into the proposed system as input, the positive or negative sentences will be detected according to the ASD symptoms.

Pseudo Code:

Step 1: Read the Dataset in variable data.
Step 2: Pre-processing task on data
// data[‘Comments’] stores sentences and data[‘Sentiment’] stores binary value 0 and 1.
data = data[['Comments','Sentiment']]
// At first covert all the sentences in lower case and remove special characters from the sentences.
data['Comments'] = data['Comments'].apply(lambda x: x.lower())
data['Comments'] = data['Comments'].apply((lambda x: re.sub('[^a-zA-z0-9\s]','',x)))
Step 3: Feature Generation from text.
// Maximum feature generation has been initialize
max_fatures = 2000
// Tokenize each sentences from the data[‘Comments’] column
tokeniz = Tokenizer(num_words=max_fatures, split=' ')
tokeniz.fit_on_texts(data['Comments'].values)

// Sequence generation for training and testing data
X =
tokeniz.texts_to_sequences(data['Comments'].values
)
X = pad_sequences(X)
Step 4: LSTM Deep Learning Model creation
embedim = 128
lstmout = 196
modl = Sequential()
modl.add(Embedding(max_fatures,
embedim,input_length = X.shape[1]))
modl.add(SpatialDropout1D(0.4))
modl.add(LSTM(lstm_out, dropout=0.2,
recurrent_dropout=0.2))
modl.add(Dense(2,activation='softmax'))
modl.compile(loss = 'categorical_crossentropy',
optimizer='adam',metrics = ['accuracy'])
Step 5: Training and Testing Data Creation
batch_size = 32
modl.fit(X_train, Y_train, epochs = 7,
batchsize=batch_size, verbose = 2)
valsize = 50
X_valid = X_test[-valsize:]
Y_valid = Y_test[-valsize:]
X_test = X_test[:-valsize]
Y_test = Y_test[:-valsize]
Step 6: Model Evaluation
scores,accuracy = modl.evaluate(X_test, Y_test,
verbose = 2, batch_size = batchsize)
print("Scores= %.2f" % (scores))
print("Accuracy=: %.2f" % (accuracy))
Step 7: Prediction using proposed LSTM model on
validation data

```
Pos_Cnt, Neg_Cnt, Pos_Correct, Neg_Correct = 0,
0, 0, 0
for x in range(len(X_validate)):
   result =
modl.predict(X_validate[x].reshape(1,X_test.shape[
1]),batch_size=1,verbose = 2)[0]
   rst1.append(np.argmax(Y_validate[x]))
   rst2.append(np.argmax(result))

   if np.argmax(result) == np.argmax(Y_valid[x]):
      if np.argmax(Y_valid[x]) == 0:
         Neg_Correct = Neg_Correct +1
      else:
         Pos_Correct = Pos_Correct + 1

   if np.argmax(Y_valid[x]) == 0:
      Neg_Cnt = Neg_Cnt +1
   else:
      Pos_Cnt = Pos_Cnt +1
```

3.3.2 The BERT Cosine Similarity Model

The proposed system adopts a two-step approach for ASD symptom identification. Initially, it filters out negative sentences from the input text, focusing solely on positive sentences. These positive sentences are then utilized as input for the BERT Cosine Similarity Model. In the second step, the BERT Cosine Similarity Model processes each positive sentence from the ASD symptoms dataset (Table 5) and computes the cosine similarity score between the input sentence and each sentence in the dataset. By comparing the input sentence with the dataset sentences, the model identifies the sentence with the highest cosine similarity score. The system subsequently selects the label associated with this highly similar sentence. According to Table 4, which establishes the correspondence between ASD problems and labels, this label indicates a specific ASD problem. The cosine similarity model independently applies this process to each input sentence, allowing the system to identify ASD problems based on the highest similarity scores and their corresponding labels. The algorithm summarized below outlines the steps involved:

1. A positive sentence will be selected from the input text.
2. In the next step, the BERT cosine similarity model will be applied.
3. The cosine similarity will be calculated between the input sentence and each positive sentence from the dataset of ASD symptoms.
4. In this step, identify the sentence according to the highest cosine similarity score. After identifying the positive sentence, retrieve the corresponding label from Table 4 which is associated with the ASD problem.

By leveraging the BERT Cosine Similarity Model and utilizing the labels from Table 4, the proposed system effectively identifies ASD problems by analyzing the similarity between input sentences and the ASD symptoms dataset.
The algorithm in Python-pseudo code has been given below.

The Algorithm in Python pseudo-code:
from sentence_transformers import
SentenceTransformer, util
import pandas as pd
import pandasql as ps

// Dataframe df to be initialized by the dataset
df =
pd.read_csv(r"ASD_Symptoms.csv",encoding='Lati
n-1')

```
// List Array declare to store ‘Comments’,
‘Sentiment value’, and ‘Cosine Score value

Comments1=[]
Sentiment1=[]
cosine_value1=[]
// Function for bracket remove from Cosine value in
Python
def StringWithoutBrackets(lists):
    return str(lists).replace('[','').replace(']','')
// Calculation with BERT Cosine function
def BERTCosine(strs1):
  for indx in df.index:
    #print(df['Comments'][indx],
df['Sentiment'][indx])

    sentence = [df['Comments'][indx], strs1]

    modl = SentenceTransformer('sentence-
transformers/all-MiniLM-L6-v2')

    # Embeddings to be computed for both lists
    embeddings1= modl.encode(sentence[0],
convert_to_tensor=True)
    embeddings2 = model.encode(sentence[1],
convert_to_tensor=True)

    comments1.append(df['Comments'][indx])
    sentiment1.append(df['Sentiment'][indx])
    scored=util.pytorch_cos_sim(embeddings1,
embeddings2)

cosine_value.append(StringWithoutBrackets(scored.
tolist()))

  dfd=pd.DataFrame(
    {'Comment': comments,
     'Sentiment': sentiment,
     'CosineScores': cosine_value
    })

//Dataframe to CSV which contains Cosine Scores
of sentences with corresponding label values.

 dfd.to_csv('ASD_Cosine_Data1.csv”)


dfd['CosineScores']=dfc['CosineScores'].astype('flo
at64')
  i = dfd['CosineScores'].idxmax()
    return dfd['Sentiment'][i]
strs1=pd.read_csv("ASD_Cosine_Data1.csv")
for sts in strs1['Comment']:
        rslt=BERT_Cosine(sts)
        print(sts,"=",rslt)
```

The result of this proposed algorithm has been discussed in the Result and Discussion section.

4 Results and Discussion

4.1 Result and Discussion of Proposed LSTM Model

The proposed LSTM model can handle sentiment analysis on the new data. Parents who are unaware of the ASD symptoms of their child will get the benefit of detecting ASD symptoms as early as possible without going through a long process of ASD detection. The evaluation results of this proposed LSTM model have been discussed here one by one.

	precision	recall	f1-score	support
0	0.96	0.97	0.96	1738
1	0.97	0.97	0.97	2001
accuracy			0.97	3739
macro avg	0.97	0.97	0.97	3739
weighted avg	0.97	0.97	0.97	3739

Fig. 3: Classification Report of Proposed LSTM Model

Many matrices are there to evaluate the performance of a machine learning model where F1 score is one of them. The precision and recall values are used to calculate the score of F1. F1 score indicates a positive prediction ability of the machine learning model. The equation is given below:

F1 = (2 x (Precision x Recall)) / (Precision + Recall)
Precision can be calculated by the true positive (TP) and false positive (FP) whereas Recall can be calculated by the true positive (TP) and false negative. The equations are given below:

Precision = TP ÷ (TP + FP)
Recall = TP ÷ (TP + FN)

The average F1 score for multiple classes can be calculated by two methods. Macro-Averaging method and weighted method are used to calculate average F1 scores. On the other hand, weighted averaging takes the relative proportions of each class in the dataset, where the weight of each class is governed by its support, which is the number of true occurrences of the class in the dataset. Scores for those classes are indeed lower if the support values of these certain classes are very imbalanced. In such cases, it may be necessary to adjust the model or the dataset to address the imbalance and improve the overall performance.

The classification report of this model has been given in Figure 3 where the precision and recall value of class 0 and 1 is 0.96, 0.97, 097, and 0.97. The F1 score is 0.96 and 0.97 of class 0 and 1 with support values 1738 and 2001. The accuracy of the proposed LSTM model according to the F1 score is 0.97 with a support value of 3739 which indicates a strong predictive model. The precision and recall values of the macro average and weighted average are 0.97, 0.97, 0.97, and 0.97. The F1 score of the macro average and weighted average are 0.97 and 0.97 with support values 3739 and 3739.

Fig. 4: Confusion Matrix of Proposed LSTM Model

According to Figure 4, The Y-axis has the Actual 0s and Actual 1s whereas the X-axis has predicted 0s and predicted 1s. 50 sentences have been used to represent this confusion matrix using the proposed LSTM model. 23 sentences have been predicted by the proposed LSTM as 0 where these are actual 0s. No sentences have been predicted as 0 but they are 1. The proposed LSTM model has predicted 26 sentences as 1 which is 1 and 1 sentence has been predicted as 1 which is 0. According to the confusion matrix, it is a clear observation that the proposed LSTM can detect sentences as ASD symptoms from the parent's dialogue.

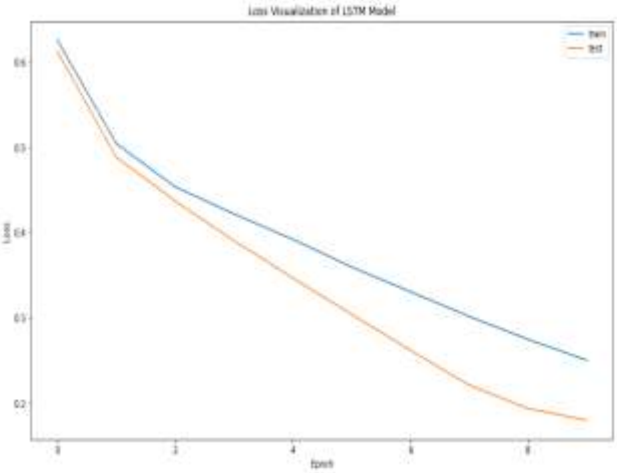

Fig. 5: Loss Visualization of Proposed LSTM Model

According to Figure 5, two loss learning curves have been shown which are blue lines and yellow lines. The Y-axis is a Loss against each epoch as the X-axis. 10 epochs have been considered here where it has been seen that each line goes downwards after epochs. The blue line refers to the loss during the training using training data whereas the yellow line refers to the loss during testing using test data. Both lines indicate that losses (error) are reduced after a certain time interval means completing epochs. Reduction of loss or error means the model is ready with good accuracy. The accuracy curves are plotted in Figure 6. According to Figure 6, The Y-axis refers to the accuracy whereas X-axis refers to the epoch. The blue line is an accuracy line of the proposed LSTM model during the training using the training data. The yellow line is an accuracy line of the proposed LSTM model during testing using test data. Both lines are increasing towards 1.0.

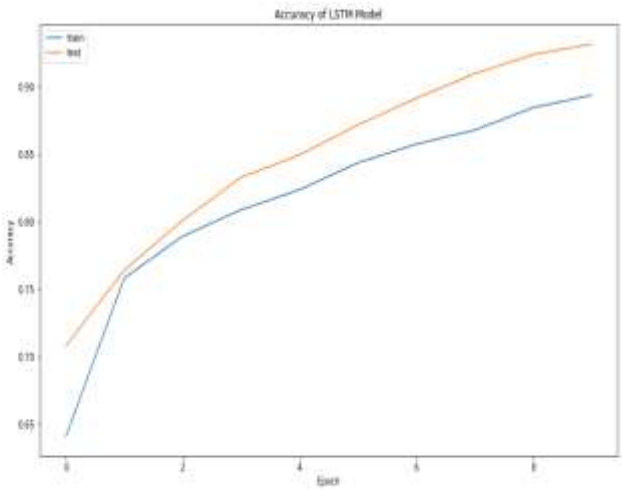

Fig. 6: Accuracy Visualization of the Proposed LSTM Model

4.2 Result and Discussion of BERT Cosine-Similarity Model

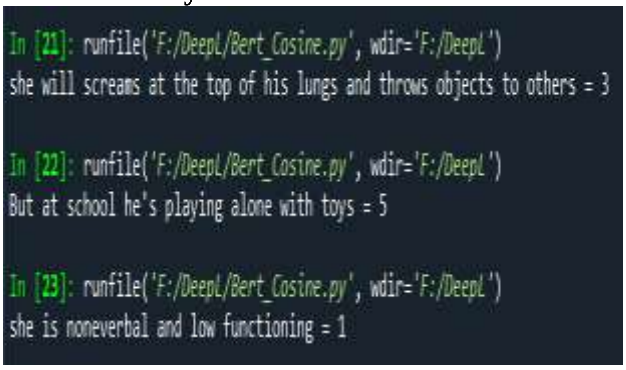

Fig. 7: Example Output of BERT Cosine Similarity Model

The proposed system employs a cosine similarity model to identify ASD symptoms from positive ASD sentences, predicted by machine learning algorithms. The system's output, as illustrated in Figure 7, assigns labels to sentences such as "*she will screams at the top of his lungs and throws objects at others*" (labeled 3), *"But at school, he's playing alone with toys"* (labeled 5), and *"she is nonverbal and low functioning"* (labeled 1). Referring to Table 4, we can associate label 3 with Behaviour problems, label 5 with Social Interaction problems, and label 1 with Speech problems. These labels provide valuable insights into the specific ASD problems related to the identified symptoms. Proper therapies can be initiated based on the specific problem identified if the ASD problems are detected. Tailored interventions play a crucial role in delivering targeted support to people with ASD. Accurate identification of ASD problems by the proposed system through the analysis of positive sentences allows for a focused and personalized procedure for therapeutic interventions. This personalized procedure promises a positive impact on ASD patients and enhances their quality of life.

5 Limitation

The overall system performance degrades with large datasets as the LSTM model may on work optimally for large datasets. LSTM models have shown restrictions in working with complicated computations, like aggregation-type natural language response generation. For example, models like LSTM may fail to perfectly calculate the sum of multiple float values concurrently.

6 Conclusion

To detect ASD symptoms, the proposed system is designed to analyze natural language text extracted from parent dialogues. The sentiment (positive or negative) expressed in sentences related to ASD symptoms is determined by sentiment analysis techniques. The proposed system uses the LSTM Model on the supplied dataset. Only the positive sentences are chosen for additional analysis using the BERT cosine similarity model after performing sentiment analysis. An ASD symptoms dataset is leveraged by the system, where each sentence is tagged with a value corresponding to a particular ASD symptom. Through the computation of cosine similarity between the input sentence and the sentences associated with the ASD symptoms dataset, the system decides the label that demonstrated the highest score signifying that the input sentence has similarity to the sentence associated with a specific ASD problem. The system aims to bridge the gap in ASD diagnosis and intervention by leveraging the ability of text-based analysis, supplying valuable insights and assistance to needy individuals and communities. Moreover Large Language Model (LLM) can be launched to uncover the symptoms from the parents' conversations and such kind of LLM-based systems are perfect for the generation of outcomes.

7 Future Work

Training the BERT and ChatGPT models using the supplied dataset can be an important step for the further development of the output and accuracy of the proposed system. These models have proven to be successful in several natural language processing activities and can leverage the dataset to make predictions, potentially enhancing the system's performance. BERT, which is a deep learning model, excels in classification activities and can come up with more accurate ASD identification. On the other side, ChatGPT, which is a large language model, gives appropriate prediction abilities and can supply important insights. Some hybrid processes can be applied to handle aggregation-type reply generation by using these models. In summary, future work based on training the BERT and ChatGPT models using the supplied dataset for ASD identification holds assurance for developing the system further. However, to achieve optimal performance, a proper study needs to be done to consider the dataset size and select the proper models for our future work.

Acknowledgement:
The authors extend their appreciation to the Manipur International University, Imphal, India for supporting this post-doctoral research work on Autism.

References:
[1] A.M. Pagnozzi, E. Conti, S. Calderoni, J. Fripp, S.E. Rose, A systematic review of structural MRI biomarkers in autism spectrum disorder: A machine learning perspective, *Int. J. Dev. Neurosci.* 2018, pp. 68–82.
[2] Autism Spectrum Disorders. World Health Organization, [Online]. https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders (Accessed Date: February 7, 2021).
[3] S.N. Hansen, D.E. Schendel, E.T. Parner, Explaining the increase in the prevalence of autism spectrum disorders: The proportion attributable to changes in reporting practices, *JAMA Pediatr*., 2015, pp. 56-62.
[4] A.N. Witwer, L. Lecavalier, Examining the validity of autism spectrum disorder subtypes, *J. Autism Dev. Disord.*, 2008, pp. 1611-1624.
[5] M.O. Mazurek, F. Lu, H. Symecko, E. Butter, N.M. Bing, R.J. Hundley, M. Poulsen, S.M. Kanne, E.A. Macklin, B.L. Handen, A prospective study of the concordance of DSM-IV and DSM-5 diagnostic criteria for autism spectrum disorder, *J. Autism Dev.Disord*., Vol. 47, 2017, pp.2783–2794.
[6] E. Conti, J. Mitra, S. Calderoni, K. Pannek, K. Shen, A. Pagnozzi, S. Rose, S. Mazzoti, D. Scelfo, M. Tosetti, F. Muratori, G. Cioni, A. Guzzetta, Network over-connectivity differentiates autism spectrum disorder from other developmental disorders in toddlers: A diffusion MRI study, *Hum. Brain Mapp*., vol. 38, 2017, pp. 2333–2344.
[7] F. Thabtah, Machine learning in autistic spectrum disorder behavioral research: A review and ways forward, *Inform. Health Soc. Care*, vol. 44, 2019, pp. 278–297.
[8] G. Sandhu, A. Kilburg, A. Martin, C. Pande, H. F. Witschel, E. Laurenzi, E. Billing, A Learning Tracker using Digital Biomarkers for Autistic Preschoolers– Practice Track, *EPiC Series in Computing*, 2022, vol. 84, pp. 219-230.
[9] N V Ganapathi Raju, Karanam Madhavi, G Sravan Kumar, G Vijendar Reddy, Kunaparaju Latha, K Lakshmi Sushma, Prognostication of Autism Spectrum Disorder (ASD) using Supervised Machine Learning Models, *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 8(4), 2019, pp. 1028-1032.
[10] S.R. Surya, Kalpana G., Autism Spectrum Disorder Using KNN Algorithm, *European Journal of Molecular & Clinical Medicine*, vol. 07(09), 2020, pp. 1628-1637.
[11] G.K. Suhas, N. Naveen, M. Nagabanu, Edwin R. Mario, R. Nithish Kumar, Premature Identification of Autism Spectrum Disorder using Machine Learning Techniques, *Advanced Innovations in Computer Programming Languages*, vol. 3(3), 2021, pp. 1-10.
[12] S. Purwanti, B. Nugraha, and M. Alaydrus, Enhancing security on E-health private data using SHA-512, *International Conference on Broadband Communication, Wireless Sensors and Powering, BCWSP*, vol. 2018-January, 2018, pp. 1–4.
[13] H. H. Purba, F. Debora, C. Jaqin, H. Adiyatna, Service quality analysis: An empirical study of customer satisfaction in healthcare, *Jurnal Teknologi dan Manajemen*, vol. 19, no. 1, 2021, pp. 33–38.
[14] Yaya Sudarya Triana, Mohd Azam Osman, Adji Pratomo, Muhammad Fermi Pasha, Deris Stiawan, Rahmat Budiarto, Neural network models selection scheme for health mobile app development, *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12(3), pp. 2023, 1191-1203.
[15] S. Kaur, J. Singla, L. Nkenyereye, S. Jha, D. Prashar, G. P. Joshi, S. El-Sappagh, M. S. Islam, S. M. Riazul Islam, Medical diagnostic systems using artificial intelligence (AI) algorithms: Principles and perspectives, *IEEE Access*, vol. 8, 2020, pp. 228049–228069.
[16] L. Jiang, Z. Wu, X. Xu, Y. Zhan, X. Jin, L. Wang, Y. Qiu, Opportunities and challenges of artificial intelligence in the medical field: current application, emerging problems, and problem-solving strategies, *Journal of International Medical Research*, vol. 49, no. 3, 2021.
[17] W. Walter, C. Haferlach, N. Nadarajah, I. Schmidts, C. Kühn, W. Kern, T. Haferlach, How artificial intelligence might disrupt diagnostics in hematology in the near future, *Oncogene*, vol. 40, no. 25, 2021, pp. 4271–4280.
[18] Shomona Gracia Jacoba, Majdi Mohammed Bait Ali Sulaimanb, Bensujin Bennetc, Algorithmic Approaches to Classify Autism

Spectrum Disorders: A Research Perspective, *The 5th International Conference on Emerging Data and Industry 4.0*, 2022, pp. 470-477.

[19] Suman Raja, Sarfaraz Masoodb, Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques, *Conference on Computational Intelligence and Data Science (ICCIDS)*, 2020, pp. 994-1004.

[20] Ashima Sindhu Mohanty, Priyadarsan Parida, Krishna Chandra Patra, ASD classification for children using deep neural network, *Global Transitions Proceedings*, 2021, pp.461-466.

[21] Ke Niu, Jiayang Guo, Yijie Pan, Xin Gao, Xueping Peng, Ning Li, Hailong Li, Multichannel Deep Attention Neural Networks for the Classification of Autism Spectrum Disorder Using Neuroimaging and Personal Characteristic Data, *Complexity,* vol. 2020, pp. 1-9.

[22] Milan N. Parikh, Hailong Li, Lili He, Enhancing Diagnosis of Autism With Optimized Machine Learning Models and Personal Characteristic Data, *Frontiers in Computational Neuroscience*, vol. 13, article 9, 2019, pp. 1-5.

[23] Taban Eslami, Vahid Mirjalili, Alvis Fong, Angela R. Laird, Fahad Saeed, ASD-DiagNet: A Hybrid Learning Approach for Detection of Autism Spectrum Disorder Using fMRI Data, *Frontiers in Neuroinformatics*, vol. 13, article 70, 2019, pp. 1-11.

[24] Lingyu Xu1, Xiulin Geng, Xiaoyu He, Jun Li, Jie Yu, Prediction in Autism by Deep Learning Short-Time Spontaneous Hemodynamic Fluctuations, *Frontiers in Neuroscience*, vol. 13, article 1120, 2019, pp. 1-12.

[25] Scott H. Lee, Matthew J. Maenner, Charles M. Heilig, A comparison of machine learning algorithms for the surveillance of autism spectrum disorder, *PLOS ONE*, 2019, pp. 1-11,.

[26] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, F. Meneguzzia, Identification of Autism Spectrum Disorder using Deep Learning and the ABIDE Dataset, NeuroImage: *Clinical*, 2017, pp. 1-29.

[27] Vishal Padole, Image Classification by Using Multiclass Support Vector Machines, *WSEAS Transactions on Computer Research*, 2019, pp. 1-8.

[28] Suboh Alkhushayni, Taeyoung Choi, Du'a Alzaleq, Data Analysis Using Representation Theory and Clustering Algorithms, *WSEAS Transactions on Computers*, vol.19, 2020, pp. 310-320, https://doi.org/10.37394/23205.2020.19.38.

[29] Meera Sharma, Sonok Mahapatra, Adeethyia Shankar, Xiaodi Wang, Predicting the Utilization of Mental Health Treatment with Various Machine Learning Algorithms, *WSEAS Transactions on Computers,* vol..19, 2020, pp. 281-295, https://doi.org/10.37394/23205.2020.19.34.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed to the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.