Detection of Autism Spectrum Disorder (ASD) from Natural Language Text using BERT and ChatGPT Models

Prasenjit Mukherjee[1], Gokul R. S.[2], Sourav Sadhukhan[3], Manish Godse[4], Baisakhi Chakraborty[5]

Dept. of Technology, Vodafone Intelligent Solutions, Pune, India[1, 2]
Dept. of Computer Science, Manipur International University, Manipur, India[1]
Dept. of Finance, Pune Institute of Business Management, Pune, India[3]
Dept. of IT, BizAmica Software, Pune, India[4]
Dept. of Computer Science and Engg, National Institute of Technology, Durgapur, India[5]

***Abstract*—ASD may be caused by a combination of genetic and environmental factors, including gene mutations and exposure to toxins. People with ASD may also have trouble forming social relationships, have difficulty with communication and language, and struggle with sensory sensitivity. These difficulties can range from mild to severe and can affect a person's ability to interact with the world around them. Autism spectrum disorder (ASD) is a developmental disorder that affects people in different ways. But early detection of ASD in a child is a good option for parents to start corrective therapies and treatment. They can take action to reduce the ASD symptoms in their child. The proposed work is the detection of ASD in a child using a parent's dialog. The most popular Bert model and recent ChatGPT have been utilized to analyze the sentiment of each statement from parents for the detection of symptoms of ASD. The Bert model has been developed by the transformers which are the most popular in the natural language processing field whereas the ChatGPT model is a large language model (LLM). It is based on Reinforcement learning from human feedback (RLHF) that can able to generate the sentiment of the sentence, computer language codes, text paragraphs, etc. The sentiment analysis has been done on parents' dialog using the Bert model and ChatGPT model. The data has been prepared from various Autism groups on social sites and other resources on the internet. The data has been cleaned and prepared to train the Bert model and ChatGPT model. The Bert model is able to detect the sentiment of each sentence from parents. Any positive sentiment detection means parents should be aware of their children. The proposed model has given 83 percent accuracy according to the prepared data.**

Keywords*—BERT model; ChatGPT model; autism; machine learning; generative AI; autism detection

I. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by challenges in verbal communication, restricted interests, and repetitive behaviors as in [1]. The wide variability observed in individuals with ASD makes it difficult to establish universally applicable markers. As a result, there is a growing need for simpler and more accessible measurement techniques that can be routinely implemented for early detection purposes as in [2]. ASD may be caused by a combination of genetic and environmental factors and is characterized by challenges with social skills, communication, and repetitive behaviors. Early diagnosis is important, as early intervention and treatment can improve outcomes and help those with ASD reach their full potential. Depending on the type of ASD, individuals may have difficulty with communication and social interaction, have repetitive behaviors, have difficulty with change, or have sensory sensitivities. Treatment plans are tailored to the individual and may include therapy, medication, and other services as in [3]. This is because autism is a complex disorder, and its symptoms can vary greatly from person to person. In addition, the symptoms of autism may not be obvious at first and can take some time to manifest. The 'spectrum' in ASD refers to the wide range of symptoms and their severity sometimes called Asperger's syndrome as in [4]. Thus, a comprehensive screening process is needed to properly detect autism. This means that the earlier the detection, the sooner the person can receive the necessary treatment and support, which can help to improve the quality of life for those with autism. In addition, early detection can also help to reduce the costs associated with the disorder, as treatments can be more effective when started early. This is because the signs and symptoms of autism can be very subtle in this age group and can be difficult to identify. Furthermore, the behavior of young children is constantly changing, which makes it even harder to accurately diagnose autism in this age group as in [5]. According to the autism report, the worldwide rate of increment of autism over the past decade is six children among 1000 children as in [6].

People with autism have difficulty in social interaction and communication and may have restricted and repetitive patterns of behavior, interests, or activities. Additionally, they may experience sensory sensitivities, have difficulty with transitions between activities, and have difficulty understanding abstract concepts as in [7]. Early detection of autism is important as it can help children to receive the necessary interventions and support to aid in their development. However, the lack of access to resources and the costs associated with screening tests make it difficult for families to get the help they need. While there is no cure for autism, there are effective treatments that focus on helping individuals with autism learn and develop skills and cope with the challenges they face. These treatments may include behavior therapy, speech and language therapy, and occupational therapy. These symptoms can include difficulty with communication, sensory processing, and motor skills, as well as problems with memory, self-regulation, and social interactions. Additionally, these symptoms should interfere with the child's ability to function in home, school, and other social settings. Early diagnosis of ASD is a key to helping children reach their full potential and receiving the necessary treatment and interventions. Early diagnosis also allows for timely referral to other specialists and access to educational and behavioral therapies, which can help improve outcomes for children with ASD as in [8]. AI can analyze patterns and trends in large amounts of data more quickly and accurately than a human can. AI algorithms can be used to study genetic data, environmental factors, and symptom progression over time in order to identify or predict ASD. Higher dimensional data requires more sophisticated algorithms for analysis. Also, data from multiple sources can have different formats and structures, making it even more difficult to analyze. Therefore, artificial intelligence methods are needed to better analyze these types of data, which can be used for the prognosis and diagnosis of ASD as in [7]. Machine learning algorithms have the capability to analyze large amounts of data and identify patterns in the data that can be used to distinguish between individuals with ASD and those without as in [9]. Any machine learning model can be used but improvement of accuracy, precision, and recall will reduce the time complexity of the ASD diagnosis model as in [10]. Classification models are good to use for the detection of ASD of a person after improving the accuracy of the classification model as in [11]. By using these algorithms, physicians can better diagnose individuals with ASD and ensure they are receiving the proper treatment. By speeding up the autism diagnosis process, doctors and caregivers can quickly identify symptoms and provide the appropriate treatments that can help improve the quality of life for those with ASD as in [12].

The proposed research work detects the sentence that contains positive words that are pointing to ASD symptoms. The proposed system contains two machine learning models that are BERT model, and the ChatGPT model. A dataset has been prepared to train the BERT model. The dataset has been prepared from parents' dialogues who are actually parents of autistic children. Their experiences have been utilized to detect ASD symptoms. Parents of autistic children are spent most of their time with their children and their experience is the key to detecting ASD symptoms perfectly. The data has been collected from organizations that are working with autistic children. Some data has been collected from the social networks group. The dataset has been prepared according to the binary classification. A sentence containing positive words regarding autism sentence denotes true (1) and without positive words in a sentence denotes false (0). BERT model follows the most popular transformer architecture to do sentiment analysis where this model is trained with the prepared dataset that has been described in Section III. The ChatGPT model of OpenAI is another popular large language model in recent years. The architecture of this model uses reinforcement learning from human feedback (RLHF) architecture to solve many critical tasks. The example data from the dataset has been sent to this model to understand the data structure and then new data send for predicting the sentiment for the detection of ASD symptoms. Both models are used to detect ASD symptoms. The detail of both models with the dataset has been discussed in Section III. The result of both models has been discussed in Section IV. The limitation and conclusion have been given in Sections V and VI respectively. Finally, this research paper ends with a future work.

II. Related Works

Autism Spectrum Disorder (ASD) is a significant healthcare concern among children today, and it is considered one of the primary domains of interest in healthcare research. Artificial intelligence (AI) has become an increasingly popular approach for studying and treating ASD, and numerous studies have explored the use of AI in this field. In this section, we review some of the most important AI-based research on mental health issues, including ASD that have been conducted to date. These acoustic features are indicative of the underlying neurological mechanisms of the disorder. By analyzing these features, researchers can identify patterns that are unique to ASD speech and can identify individuals with autism at an early stage, which can then allow for early intervention and increased success in managing the disorder. These features are used to quantify the differences in speech production between children with ASD and those of a normal population. By analyzing these features, filter features formants (F1 to F5) researchers can identify potential differences in speech production between the two groups and better understand the impact of ASD on speech production. Significant changes in a few acoustic features are observed for ASD-affected speech as compared to normal speech. These changes indicate a difference in the production of speech sounds in individuals with ASD. This can help researchers understand the impact of ASD on speech production and suggest areas for further investigation. This is because the formants and dominant frequencies of the vowel /i/ are the most distinct between ASD and normal children. Additionally, the vowel 'I' is the most common vowel in speech and thus has the most data points to compare between groups. It provides a detailed analysis of how computer algorithms can be used to detect signs of autism in children. The authors provide evidence of how accurate these systems can be and how they can be used to aid in early diagnosis and intervention as in [13]. Early diagnosis of ASD is important in order to provide the necessary interventions and therapies to improve the quality of life for those affected. The development of easily implemented and effective screening methods can help to identify children with autism at an earlier age and provide them with the support they need. By using the Logistic Regression model, the researchers are able to capture the complex patterns of the dataset and accurately predict the outcome of ASD disease. Additionally, the algorithm will allow for rapid classification of the behaviors in the dataset, allowing for faster and more accurate predictions. These challenges include the need to ensure data privacy, to be able to trust the results of the AI system, to have the right infrastructure in place, and to ensure that the system is able to properly interpret the data. Additionally, healthcare organizations must be able to adjust the system as new data becomes available and as the needs of the organization change as in [14]. What is often overlooked, however, is the importance of the quality of the interaction between the child and the caregiver. A supportive and responsive environment with frequent turns and back-and-forth exchanges between the adult and the child is essential for language development. We use observational methods to analyze the language used by caregivers in interactions with their children to see how well it predicts language development. The authors in [15] looked at the child's cognitive, social, and linguistic abilities to see if they are contributing factors to language development. Authors in [15] have used a longitudinal corpus of conversation data to measure how much caregivers repeat their children's words, syntax, and semantics, and whether this repetition is predictive of language development beyond other established predictors. Results: This is because by repeating and mirroring the child's language, the caregiver is providing the child with immediate feedback and reinforcement. This helps the child to better understand their language and helps them to develop their language skills in a more effective manner. Language acquisition is a process that relies on more than just memorization of information. It requires an interactive conversational model where the learner is engaged in meaningful dialogue with a teacher. Our open-source scripts provide a platform for researchers to explore this model in different languages and contexts as in [15]. NLP techniques are used to analyze the text and identify language patterns and emotions. This allows the system to identify users who may be experiencing depression, anxiety, or other mental health issues, and direct them to appropriate resources or suggest personalized interventions to help them cope with their issues. It also provides a comprehensive review of the literature on the various techniques and approaches used for data collection, processing, and analysis, as well as their advantages and limitations. Furthermore, this paper outlines the opportunities and challenges associated with the use of online data writing in mental health support and intervention. By using data from social media and other sources, researchers can detect certain patterns in language and behavior that can be used to identify individuals who may be in need of psychological assistance. Additionally, computational techniques can be used to label and diagnose mental health issues. Finally, through the use of machine learning and natural language processing, interventions can be generated and personalized for individuals in order to help them manage their mental health. This review seeks to provide a comprehensive overview of the existing literature in these fields and to identify gaps in the research and opportunities for future research and development. Additionally, it seeks to provide a common language to facilitate further collaborations between the different disciplines as in [16]. People with autism struggle with communication, interaction, and understanding social cues. They may be overly sensitive to sound, light, and other sensory inputs. They often have difficulty in understanding other people's feelings, and may also have difficulty forming relationships. They may also have unusual behavior patterns, such as repetitive movements or focusing on one topic for long periods of time. By using machine learning algorithms, doctors can identify potential markers of autism in a child at a much earlier stage than before. Early detection of autism can allow parents to seek out treatments and therapies that may help reduce the severity of autism in their children. We will use data from clinical studies, medical records, and other sources to develop a model that can accurately predict the outcome of autism diagnosis in children. We will analyze the data to identify patterns and develop an algorithm that can be used to predict outcomes. We will also evaluate existing machine learning models to determine which ones may be suitable for our proposed work. This method allows us to compare the results of the two groups of patients and look for any correlations between the data and how it may impact the diagnosis and treatment of future patients. Additionally, this approach can help us to identify any potential trends in the data which could further inform our understanding of the condition. By using LR, NB, DT, and KNN algorithms, we can apply different techniques to the data set, such as feature engineering and feature selection. This will help to ensure that the most important factors influencing the diagnosis of autism is being taken into account and that the predictive models are accurate and reliable as in [17]. The ASD QA dataset provides a unique challenge for MRC models, as it requires them to understand the context of the reading passage in order to correctly identify the answer. It is particularly difficult because the answers are not always straightforward and may require some inference or deduction. The dataset was created by selecting a set of questions from the ARC dataset and extracting the corresponding answer segments from passages in Wikipedia. The questions are generated from the passages using a variety of techniques. By adding these questions, we can evaluate the system's ability to identify and classify unanswerable questions. This is important to ensure that the system is not simply guessing, but rather, it is using the contextual information in the reading passage to accurately identify answerable and unanswerable questions. Each answer contains also positional tags (start and end) denoting the numerical positions of the first and last symbols of the answer span in a reading passage. 5% of the questions in ASD QA are unanswerable, which means that corresponding reading passages do not contain any answers to them. This split was chosen so that the model could be trained on the majority of the data, tested and validated on a smaller portion of the data to ensure accuracy, and then tested on an unseen portion of the data to ensure that it is generalizing correctly. Byte pair encoding is a method of compressing words into smaller units of meaning, which reduces the overall vocabulary size and makes it easier for the models to learn. This is beneficial for both types of models, as it reduces the amount of noise in the data, allowing for more accurate predictions as in [18]. One possible risk factor that has been identified is the presence of co-occurring mental health conditions, such as depression, anxiety, and obsessive-compulsive disorder. These conditions can be exacerbated in young people with ASD due to the difficulties they face in forming social relationships, communicating effectively, and coping with sensory overload. EHRs can provide a large amount of data, but they are often not standardized, making it difficult to accurately extract the data needed to create a valid cohort. Developing systems to accurately extract and organize the data would allow researchers to better understand the relationship between ASD and suicide risk. The systematic approach utilizes NLP to identify the suicidal language in the EHR corpus and determine if the language is positive or negative. The performance of the classification tool is then evaluated on a subset of 500 patients. The precision score indicates the proportion of relevant results to the total relevant results retrieved, and the recall score measures the proportion of relevant results retrieved to the total amount of relevant results. The F1 score combines these two scores, providing a single value that is indicative of the system's overall performance. In this case, all of these scores were very high, indicating that the NLP classification tool was highly accurate in recognizing positive suicidality. The application has been validated against existing research and has been found to be effective in identifying potential risk factors for suicidality among individuals with ASD. It has also been found to be useful in predicting suicide risk and providing automated surveillance within clinical settings as in [19]. MRI imaging modalities have the capability to detect subtle brain abnormalities that are associated with ASD, such as changes in the brain's structure, connectivity, and even in its chemistry. This makes it an invaluable tool for diagnosing and monitoring ASD. fMRI uses magnetic fields and radio waves to measure blood flow in the brain and identify any abnormalities or discrepancies in brain activity. sMRI uses high-resolution images to map the structure of the brain and detect any abnormalities in the brain anatomy. These two modalities work together to help clinicians diagnose ASD with greater precision. These systems use AI to analyze brain images, such as MRI and fMRI scans, to assess an individual's brain structure and connectivity. The AI algorithms can detect subtle differences in brain structures which can be used to diagnose ASD more accurately and quickly by specialists. ML algorithms are used to analyze the image data, identify the relevant features, and detect any abnormalities that could be indicative of ASD. DL applications are used to further analyze the data and identify patterns that may be indicative of ASD. This allows for more accurate and reliable diagnoses. Deep Learning (DL) techniques employ large datasets of MRI images and AI algorithms to create models that can detect patterns in the images that are associated with ASD. These models can then be used to automate the diagnosis of ASD and provide more accurate and timely results. We compare the accuracy and training times of ML and DL models to show that DL models can learn faster and achieve higher accuracy. We also discuss the importance of feature selection and data pre-processing in improving the accuracy of the models. Finally, we suggest the potential of combining AI techniques with MRI neuroimaging to detect ASDs as in [20]. Diagnosis of ADHD is typically based on a combination of self-reported symptoms, observations by parents and teachers, and psychological tests. Therefore, it can be difficult to accurately diagnose ADHD since there is no concrete, medical evidence to support a diagnosis. The proposed method was tested on a publicly available ADHD-200 dataset, which showed that 4-D CNN had better performance than traditional methods in terms of accuracy, specificity, and sensitivity. Furthermore, it was also demonstrated that 4-D CNN could effectively detect subtle differences in RS-fMRI data between ADHD and healthy individuals. These models take advantage of the spatiotemporal information of the RS-fMRI images to extract useful information about the brain. For example, the feature pooling model can be used to detect patterns in the data that are consistent across multiple time frames, while the LSTM model can be used to analyze the temporal dynamics of the data. The spatiotemporal convolution model can be used to identify spatial structures in the data. The approach is designed to reduce the computational cost of training deep learning models on fMRI data. By breaking the fMRI frames into shorter pieces with a fixed stride, the data can be processed more quickly and efficiently. The results of the evaluations demonstrate that our 4-D CNN method is more effective at accurately diagnosing ADHD than traditional methods. This method can be used to create a powerful and efficient tool that can be used to accurately diagnose ADHD and provide clinicians with the information they need to make informed decisions as in [21]. A comparative analysis has been done with proposed models and some similar machine learning models that are able to detect mental disorders. The proposed analysis has been described in Table I. This table contains 'Models', 'Description', 'Dataset', 'Accuracy', and 'Remarks' as attributes.

TABLE I. COMPARATIVE ANALYSIS BETWEEN PROPOSED MACHINE LEARNING MODELS AND SIMILAR MACHINE LEARNING MODELS BASED ON MENTAL DISORDER DETECTION

Sl.No.	Models	Description	Dataset	Accuracy	Remarks
	Similar Type Machine Learning Models in Mental Disorders				
1	Logistic Regression [14]	The Logistic Regression model has been used in the diagnosis and detection of ASD.	Screening data of a group of toddlers related to autism	100%	The proposed dataset has been prepared with 1054 instances and 18 attributes to train and test the model to detect ASD.
2	Logistic Regression(LR), Naïve Bayes(NB), Decision Tree(DT), K-Nearest Neighbour(KNN) [17]	These models are used to predict ASD using ASD patients and normal patients' data.	The dataset contains data related to normal patients and ASD patients.	99.37%, 96.59%, 100%, and 97.73%	The dataset contains a large number of irrelevant and missing data. Many pre-processing techniques have been applied to clean datasets.
3	4-D CNN [21]	This deep learning model training is based on the fMRI frames dataset to detect ADHD automatically.	This dataset consists of the fMRI frames data.	71.3%	The proposed deep learning model 4d-CNN is able to detect ADHD using resting-state functional magnetic resonance imaging (rs-fMRI).

4	Random Forest, Ridge, SVM [20]	,These models use sMRI and fMRI neuroimaging data to detect ASD.	NDAR and ABIDE datasets have been used that are containing sMRI and fMRI neuroimaging data.	72%, 71%, and 75.3	Software tools are needed to pre-process these kinds of MRI data. The brain extraction tools (BET), FMRIB software libraries (FSL), statistical parametric mapping (SPM), and Free Surfer are used to preprocess MRI data.
5	KNN, RF, SVM [20]	These models use sMRI and fMRI neuroimaging data to detect ASD.	NDAR and ABIDE datasets have been used that are containing sMRI and fMRI neuroimaging data.	81%, 81%, and 78.63%	Software tools are needed to pre-process this kind of MRI data. The brain extraction tools (BET), FMRIB software libraries (FSL), statistical parametric mapping (SPM), and Free Surfer are used to preprocess MRI data.
6	Logistic Regression, Random Forest [20]	These models use sMRI and fMRI neuroimaging data to detect ASD.	UCI dataset has been used which is containing sMRI and fMRI neuroimaging data.	97.54% and 100%	Software tools are needed to pre-process this kind of MRI data. The brain extraction tools (BET), FMRIB software libraries (FSL), statistical parametric mapping (SPM), and Free Surfer are used to preprocess MRI data.
7	Siamese Verification Model and GERSVMC [20]	These models use sMRI and fMRI neuroimaging data to detect ASD.	NDAR and ABIDE datasets have been used that are containing sMRI and fMRI neuroimaging data.	87% and 96.8%	Software tools are needed to pre-process this kind of MRI data. The brain extraction tools (BET), FMRIB software libraries (FSL), statistical parametric mapping (SPM), and Free Surfer are used to preprocess MRI data.
	Proposed Models in Mental Disorder (Autism Spectrum Disorder)				
8	Proposed BERT Model	The BERT model has been proposed to predict positive ASD symptoms from parents' dialogue.	Parents' Dialogues of Autistic Children in text format from SAHAS-Durgapur, India, and Social Sites.	83%	The data has been collected in text form. The parents' dialogues about their autistic children are very useful because they shared their experiences and thoughts about their autistic children. A parent of an autistic child is the best source to understand the ASD symptoms patterns.
9	Proposed OpenAI ChatGPT Model	ChatGPT model has been used to predict positive ASD symptoms from parents' dialogue.	Parents' Dialogues of Autistic Children in text format from SAHAS-Durgapur, India, and Social Sites.	Fine-Tuned and Pre-trained model of OpenAI with High Accuracy	The data has been collected in text form. The parents' dialogues about their autistic children are very useful because they shared their experiences and thoughts about their autistic children. A parent of an autistic child is the best source to understand the ASD symptoms patterns.

III. ARCHITECTURE

Two advanced machine learning models have been used to identify ASD symptoms from parents' dialogues. BERT has been used as the first model to detect the symptoms from the parents' dialogue whereas the ChatGPT model has been used as a second model to detect ASD symptoms from the given dataset. KNN and Random forest are the last two classifiers that are also used to identify ASD symptoms from a given dataset.

A. Dataset

The dataset is prepared by analyzing parents' dialogues in which they described their thoughts and experiences related to their own child with Autism Spectrum Disorder (ASD). These dialogues were obtained from various social networks and organizations where children with special needs receive therapies for communication, speech, and behavior. Table II provides a few examples of these dialogues. Parents' dialogues are a valuable source of data for identifying all possible symptoms of ASD. We used these dialogues to create a dataset for training and testing our proposed machine-learning models.

The dataset was meticulously curated from the content presented in Table II, with a thorough examination of each sentence to ascertain its relevance as a potential symptom of Autism Spectrum Disorder (ASD). It is important to note that the identification of ASD symptoms is not constrained by a fixed set of criteria, allowing for a comprehensive exploration of various indicators. To enhance the dataset and bolster the accuracy of machine learning models, a promising avenue could involve the augmentation of dialogues from parents who have firsthand experience with autistic children. This approach not only facilitates the discovery of additional symptoms but also offers a valuable opportunity to refine machine learning algorithms. Table III provides a glimpse of the dataset, offering illustrative examples of the data's composition.

Table III provides a comprehensive overview of the dataset structure proposed in this research endeavor. The dataset comprises three distinct columns: Serial Number, Comments, and Sentiment. To compile this dataset, textual excerpts were extracted from parents' dialogues, wherein each sentence underwent a rigorous evaluation to determine its association with Autism Spectrum Disorder (ASD) symptoms. Sentences that were indicative of ASD symptoms were categorized with a label of 1 (representing true), while those not exhibiting such symptoms were assigned a label of 0 (indicating false). Referring to Table III, it becomes evident that the Comments column, under serial numbers 1, 3, and 4, highlights sentences bearing true ASD symptoms, while serial numbers 2 and 5 correspond to sentences devoid of ASD symptoms. This meticulously curated ASD symptom-based dataset is now poised for the training of advanced machine learning models such as BERT and ChatGPT, holding promise for enhancing our understanding and analysis of ASD-related linguistic cues.

TABLE II. EXAMPLE OF PARENTS' DIALOGUES

Sl. No.	Parents' Dialogues
1.	My son is 20, has autism, high functioning. Drives, works but still struggles with socializing like most. eye contact can be minimal, talking is minimal (no small talk) He is being bullied now in the work place to the point his ptsd from childhood bullying is affecting him. I try to empower him but he is scared of getting Hit by his coworker
2.	He cries every night and every day about going to nursery. Unfortunately the past few weeks have been awful, he is very upset about going to nursery again and is now asking if he is going the next day at bath time and then when he wakes up in the morning he screams and cries and has to be physically dressed and put in the car.
3.	I am new here I just want to ask about any private speech and language therapist in ENGLAND at Oldham, my 2 years baby is not talking yet, sometime not responding. I am very worried for my son. Please help me for guiding.
4.	My daughter-in-law spoils him let's him have and do whatever he wants. I'm sure it's because they don't want to hear his screams Now she's temporary out of the picture. Her oldest son age 12 opens the fridge and lets him grab whatever because mom let him. I put a stop to it because food is going to waste within a day it makes it hard when the one year old wants milk and there's none because it gets spoiled. Am I wrong in not letting the 5 year old not have his way?
5.	My 9 year old little man was recently diagnosed with Autism and waiting for further appointments for ADHD. Does anyone else's child suffer with bad meltdowns and constantly angry and stressed?

TABLE III. EXAMPLE DATA IN THE PROPOSED DATASET

Sl. No.	Comments	Sentiment
1.	He is playing with toys and spin wheels every time	1
2.	Please help me understand autism because my son is 4 years old now	0
3.	My little girl is 3 years old but she is able to speak.	1
4.	She mumbles but can speak any proper word.	1
5.	I was really surprised when she came home and asked her mom about me.	0

TABLE IV. LIST OF LABELS WITH ASD PROBLEMS

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

Table IV serves as a reference for the association between ASD problems and their respective labels. Each label in Table IV represents a specific ASD problem, with Label 1 indicating "Speech Problem," Label 2 denoting "Sensory" issues, and Label 3 representing "Behavior" problems. Additionally, Table IV includes labels for other ASD problems. Once the sentiment of a sentence related to ASD symptoms is predicted, the proposed system utilizes Table IV to determine the corresponding label associated with the identified problem. In the system flow, when a sentence is classified as positive (1), it becomes the input for the Spacy cosine similarity model. This model compares the positive sentence with the dataset presented in Table V, which contains numerous positive sentences accompanied by their respective labels. Each label in Table V corresponds to an ASD problem as defined in Table IV. As described in the Proposed System Flow section, the cosine similarity model performs a similarity check between the predicted positive sentences and the sentences within the dataset. This process aims to identify the sentence in the dataset that is most similar to the input sentence. The label associated with this highly similar sentence is then selected by the system. By utilizing the information from Table IV and Table V, the proposed system matches positive sentences, which indicate ASD symptoms, with their respective labels. This matching process allows for the identification of specific ASD problems based on the BERT cosine similarity scores calculated by the system.

TABLE V. DATASET FOR COSINE SIMILARITY CHECK

Sl. No.	Positive Sentences	Label
1.	I am unable to demonstrate potty training to him during the daytime when his father is at work.	7
2.	He primarily mumbles and doesn't use coherent words.	1
3.	When I beckon him, he doesn't respond by coming to me.	10
4.	He requires visual cues to comprehend my communication.	6
5.	Her frustration deeply affects me and I'm hoping to hear some success stories.	9

In the subsequent sections, we have detailed each model, including the algorithm employed in its implementation. This dataset served as the foundational training data for these models, and the outcomes of each model have been extensively examined and deliberated upon in the Results and Discussion section.

B. BERT Deep Learning Model

BERT is a deep learning model for natural language processing that helps to understand the meaning of a text or sentence according to the context. The BERT model is trained with the Wikipedia dataset and it can be fine-tuned for better accuracy. BERT stands for Bidirectional Encoder Representations from Transformers which means the entire model is based on transformers which is itself a deep learning model. Each output is well connected with each input in this deep learning model and weights are calculated dynamically according to the input-output connection. BERT has the capability of a Masked Language Model (MLM) where a word from the sentence is hidden at the training time and later this model predicts the hidden word based on the context. BERT is dependent on the self-attention mechanism which is possible for the bidirectional transformers. The encoder and decoder are connected in the form of sequence to sequence model inside the transformer. The mathematical representation is-

$$Attention\,(A, B, C) = softmax\,(AB^T / \sqrt{d_b})\,C$$

A, B, and C are the embedding vector that is transformed by a weight matrix inside the transformer and training of transformers means finding weight matrices. The transformer model becomes a language model when weight matrices are learned. The BERT model contains a set of rules to represent the input text. Input text converts into the embedding parts which is a combination of three embeddings. According to Fig. 1, the token embeddings will be created from the tokenization of the input text. The segment embedding is needed to understand unique embedding from the given two sentences like “my baby is Autistic, she likes to play”. The model can understand better the sentence to distinguish between them. The BERT model uses positional embedding which helps to understand the position of each word in a sentence. Summations of These three embeddings are representing the input inside the BERT model.

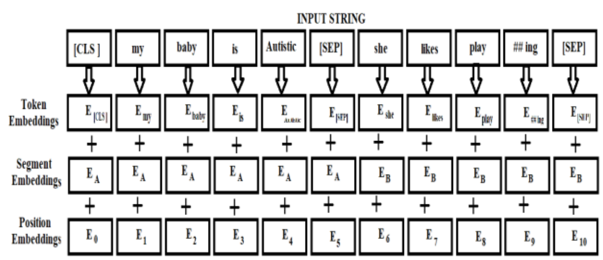

Fig. 1. Input text embedding in BERT model.

Fig, 2 of the Transformer architecture in [22] shows that it has two parts. The first part is Encoder and the second part is Decoder. After the conversion of input text to input embeddings, Encoder starts to handle input embeddings. The Encoder block contains two layers. The first layer is the Multi-Head Attention layer which is connected to the second layer named Feed Forward Neural Network. The encoder block encoded the data and sends it to the Decoder block where the Multi-Head Attention layer and Feed Forward Neural Network are connected but one extra layer is also connected which is Masked Multi-Head Attention. Different masked are handled by this layer. The proposed autism-related dataset has been applied in the BERT model to understand the sentence sentiment for identifying positive and negative symptoms of an autistic child.

The proposed system using the BERT model reads data from the proposed prepared dataset. Some NLP-related tasks have been done that are tokenization and embedding. The embedding and labeled data will be split for creation train and test data for the BERT model. The training data will be used for model training purposes as well as testing data will be used for the model performance purpose. The algorithm of the proposed system using BERT model has been given below:

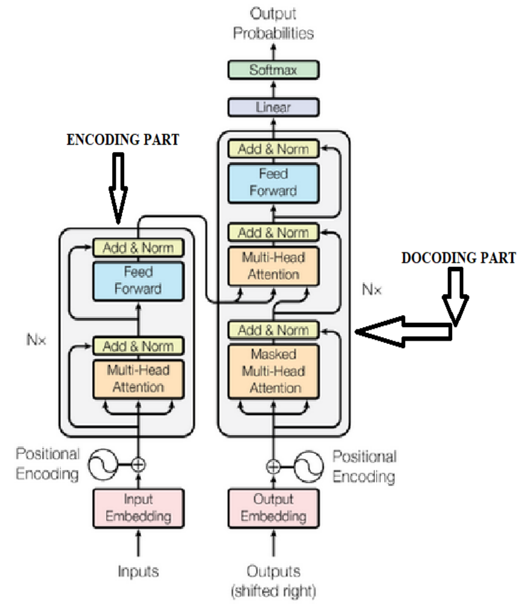

Fig. 2. The architecture of the transformer [22].

Proposed BERT Algorithm:
Pseudo Code:
Step 1: Read data from CSV file.
Step 2: X=data from csv

$x_1 = [a_1, a_2, a_3, a_4, a_5, \ldots\ldots\ldots a_n]$ is a user text column inside the dataset.

$y_1 = [r_1, r_2, r_3, r_4, r_5, \ldots\ldots\ldots r_n]$ *is a label data column inside the dataset*

Step 3: BERT model name selection.

```
// Set BERT model name
MODEL_NAME = 'bert-base-cased'
// Build a BERT based tokenizer
tokenizer = BertTokenizer.from_pretrained(MODEL_NAME)
```

Step 4: Train and test data creation.

```
df_train, df_test = train_test_split(X, test_size=0.2, random_state=RANDOM_SEED)
df_val, df_test = train_test_split(df_test, test_size=0.5, random_state=RANDOM_SEED)
def create_data_loader(df, tokenizer, max_len, batch_size):
  ds = GPReviewDataset(
    reviews=df.Comments.to_numpy(),
    targets=df.Sentiment.to_numpy(),
    tokenizer=tokenizer,
    max_len=max_len
  )
  return DataLoader(
    ds,
    batch_size=batch_size,
    num_workers=0
  )

// Create train, test and val data loaders
BATCH_SIZE = 16
```

train_data_loader = create_data_loader(df_train, tokenizer, MAX_LEN, BATCH_SIZE)
val_data_loader = create_data_loader(df_val, tokenizer, MAX_LEN, BATCH_SIZE)
test_data_loader = create_data_loader(df_test, tokenizer, MAX_LEN, BATCH_SIZE)

Step 5: Create BERT model for sentiment analysis.
bert_model = BertModel.from_pretrained(MODEL_NAME)

```
# Build the Sentiment Classifier class
class SentimentClassifier(nn.Module):

  // Constructor class
  def __init__(self, n_classes):
    super(SentimentClassifier, self).__init__()
    self.bert = BertModel.from_pretrained(MODEL_NAME)
    self.drop = nn.Dropout(p=0.3)
    self.out = nn.Linear(self.bert.config.hidden_size,
n_classes)

  // Forward propagaion class
  def forward(self, input_ids, attention_mask):
    _, pooled_output = self.bert(
      input_ids=input_ids,
      attention_mask=attention_mask,
      return_dict=False

    )
    // dropout layer addition
    output = self.drop(pooled_output)
    return self.out(output)

// Instantiate the model and move to classifier
model = SentimentClassifier(len(class_names))
```

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

C. *ChatGPT (Large Language Model)*

Today, ChatGPT has taken a very strong position to handle natural language processing tasks. ChatGPT model is an advancement of Large Language Model (LLM). ChatGPT model has been developed using reinforcement learning from human feedback (RLHF). RLHF is a reinforcement learning which is a machine learning algorithm where an agent learns everything from the environment using policy. According to Fig. 3, the agent may take action (A_t) according to the policy which affects the environment where the agent is present. According to Fig. 4, the agent may take action (A_t) according to the policy which affects the environment where the agent is present. According to this action taken, a new state (S_t) is generated and it returns a reward (R_t). Rewards are nothing but the feedback signals which indicated the reinforcement learning agent to tune action policy. In the next step, The RL agent modifies the policy to take sequences of actions when it goes through training episodes and as a result, it maximizes the rewards.

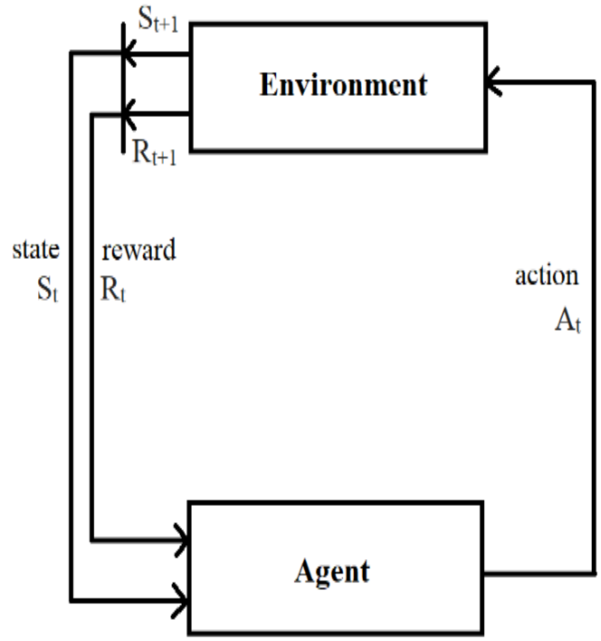

Fig. 3. Reinforcement learning basic diagram.

ChatGPT (GPT stands for Generative Pre-trained Transformer) is a large language model (LLM) developed by OpenAI. ChatGPT is trained on a massive amount of text and code data from the internet. The dataset includes text from books, articles, websites, and code from GitHub repositories. A brief description has been given according to Fig. 4.

1) The training process for ChatGPT involves several stages. First, the model is initialized with a data transformation stage where tokenization and vectorization methods are used to prepare text data for training. These are important steps because they help the ChatGPT model to understand the meaning of the text and generate text that is both coherent and grammatically correct.

2) The tokenization stage involves breaking down text into smaller units, called tokens. Tokens can be individual words, phrases, or even characters. The type of tokenization used for ChatGPT is called WordPiece tokenization. In this tokenization method text breaks down into tokens that are still meaningful, even if they are not individual words. For example, the word "running" would be broken down into the tokens "run" and "ing".

3) After completion of tokenization, it is converted into vectors. Vectors are mathematical objects that represent the meaning of a token. The vectors for each token are learned during the training process. The vectors are then used by the CHATGPT model to generate text, translate languages, and perform other tasks.

4) Once the text data is vectorized then, it is trained using a transformer Model (unsupervised learning) which is a neural network architecture that has been shown to be very effective for natural language processing tasks. It consists of encoder and decoder layers. The encoder layers take in a sequence of text tokens and transform them into a sequence of hidden representations. The decoder layers then use these hidden representations to generate a new sequence of text tokens. to process the new sequence of text tokens distributed learning is a technique used. It involves splitting the model across multiple machines, each of which is responsible for training a subset of the model's parameters. This allows the model to be trained more quickly and efficiently than if it were trained on a single machine.

5) In distributed learning technique, two methods are involved, data parallelism and model parallelism. In data parallelism, the training data is split across the machines, and each machine trains its own copy of the model. In model parallelism, the model's parameters are split across the machines, and each machine trains its own subset of the parameters. ChatGPT used this distributed learning technic to train massive datasets of text and code, which allowed it to achieve the best performance on a variety of natural language processing tasks. It uses an Azure-based supercomputing platform to process the data and models.

6) After completion of model training, it can be fine-tuned for specific natural language processing tasks, such as language translation, text summarization, and question answering. Fine-tuning involves training the model on a smaller amount of task-specific data, which allows it to learn to perform the task more accurately and efficiently.

7) After completion of model training, it can be fine-tuned for specific natural language processing tasks, such as language translation, text summarization, and question answering. Fine-tuning involves training the model on a smaller amount of task-specific data, which allows it to learn to perform the task more accurately and efficiently.

The Reinforcement learning from human feedback helps to enhance the training process of RL agents where humans are also included. An architectural diagram of LLM has been given in Fig. 5. According to Fig. 5, RLHF in language models consists in three phases where the first phase is related to the huge data training because LLM requires a huge amount of data to be trained. The LLM is pre-trained using unsupervised learning and it creates coherent outputs. Some of the output may be aligned or not aligned according to the goal of users. In the second phase, a reward model has been created where another LLM model uses the generated text from the main model to produce quality scores. The second LLM has been modified for scalar value instead of a sequence of text tokens. A dataset of LLM-generated text labeled will be generated for quality. A prompt will be given to the main LLM to generate several outputs. The human evaluator will intervene here to evaluate generated output on the basis of good and bad. In the third phase, the main LLM is an RL agent where it takes several prompts from a generated text and training set. The output of this LLM model is passed to the reward model that provides the score and aligns with the human evaluator. Finally reward model creates output according to the higher score. ChatGPT uses this RLHF framework to handle a large number of natural language queries. The engineers of OpenAI have modified the model for fine-tuning on pre-trained GPT-3.5.

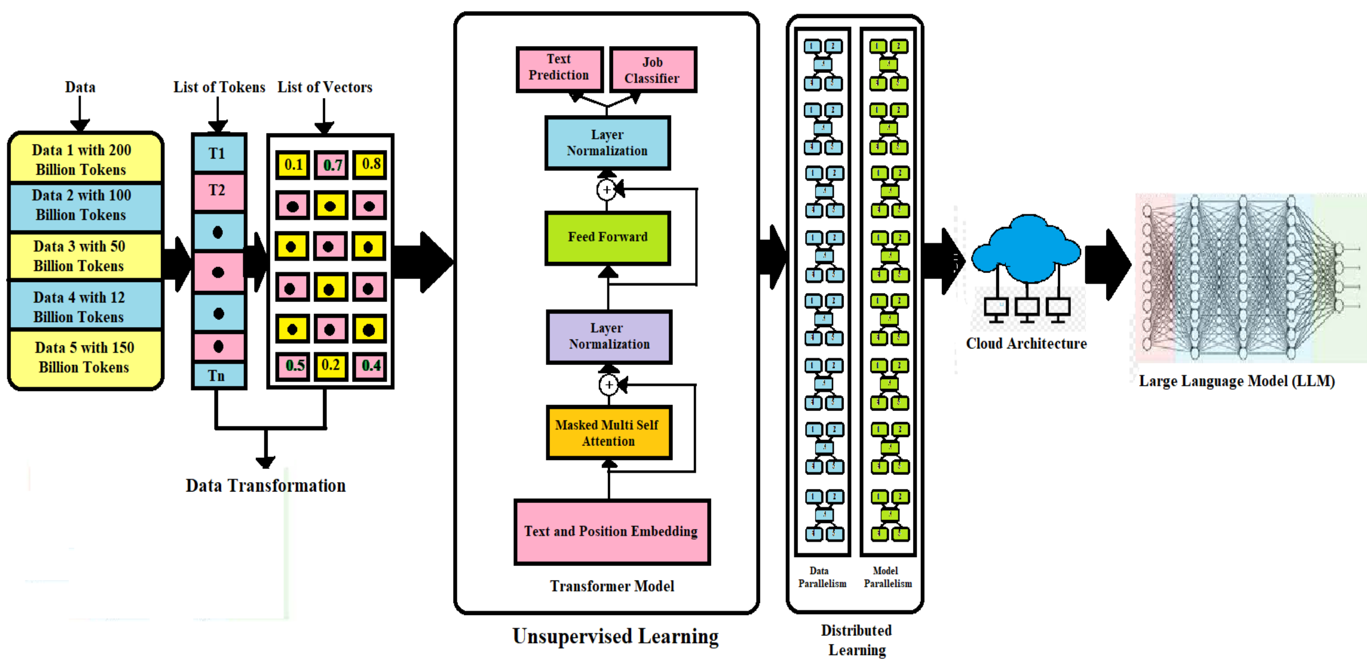

Fig. 4. ChatGPT model architecture.

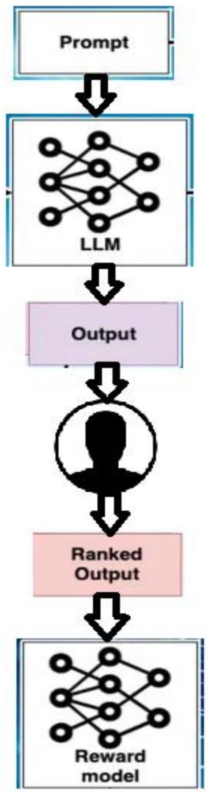

Fig. 5. Large language model diagram.

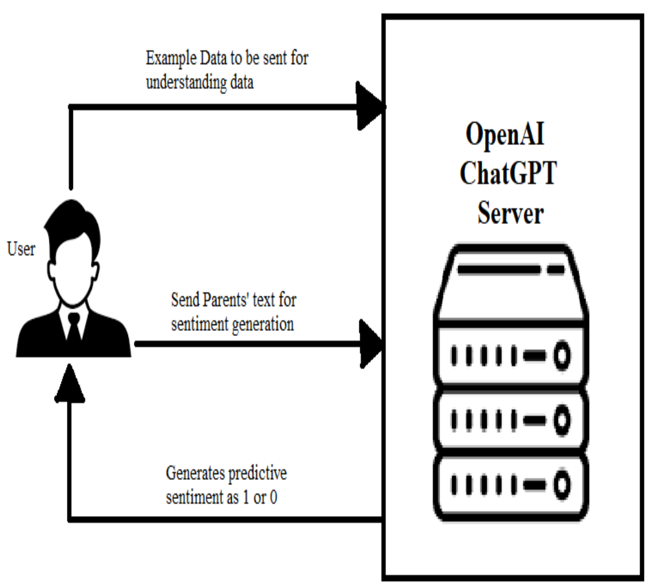

Fig. 6. ChatGPT request-response basic diagram.

According to Fig. 6, the pre-trained ChatGPT model accepts example data to understand the structure of the data. The dataset is containing structured data. ChatGPT needs example structured data to understand the pattern for response generation but there is a limitation of sending example data to the ChatGPT. ChatGPT reads data as a token and within this token limitation example data has to be sent. Now, ChatGPT is ready to generate responses to new requests. The algorithm of the proposed sentiment analysis using the ChatGPT model has been described below.

Proposed ChatGPT algorithm:
Pseudo code:
Step 1: Initialize API key in a variable
```
import openai as ai
ai.api_key ='API_Key'
```
Step 2: Initialize the data as text in a variable
```
text1='''Comments
Sentiment
How does speech therapy help with a nonverbal to speak
        1
because all they do there is play with toys with him every time
        0
I'm confused guys help my son is 3years old now
0
he does is mumbles only no proper words
1
but he goes to speech therapy every month
1
Does it help
0
```
Step 3: Define the ChatGPT model inside method:
```
def generate_gpt3_response(user_text, print_output=False):
    completions = ai.Completion.create(model='text-davinci-003',        #davinci:ft-personal-2023-02-02-06-04-19',
    temperature=0.5,
    prompt=user_text,
    max_tokens=100,
      n=1,
    stop=[" Human:", " AI:"],
  )

  // Displaying the output can be helpful if things go wrong
  if print_output:
    print(completions)
  // Return the first choice's text
  return completions.choices[0].text
```

Step 4: Call the method to identify the sentiment of the sentence.
```
text2="My baby is not responding and his eye contact is very low."
text3=text1+" "+text4+" "+text2
#result=generate_gpt3_response()
result=generate_gpt3_response(text3)
print(result)
```

D. Proposed System Architecture

According to Fig. 7, a general architecture has been given where each model is associated with the BERT cosine similarity model. The working flow of BERT and ChatGPT has been given below.

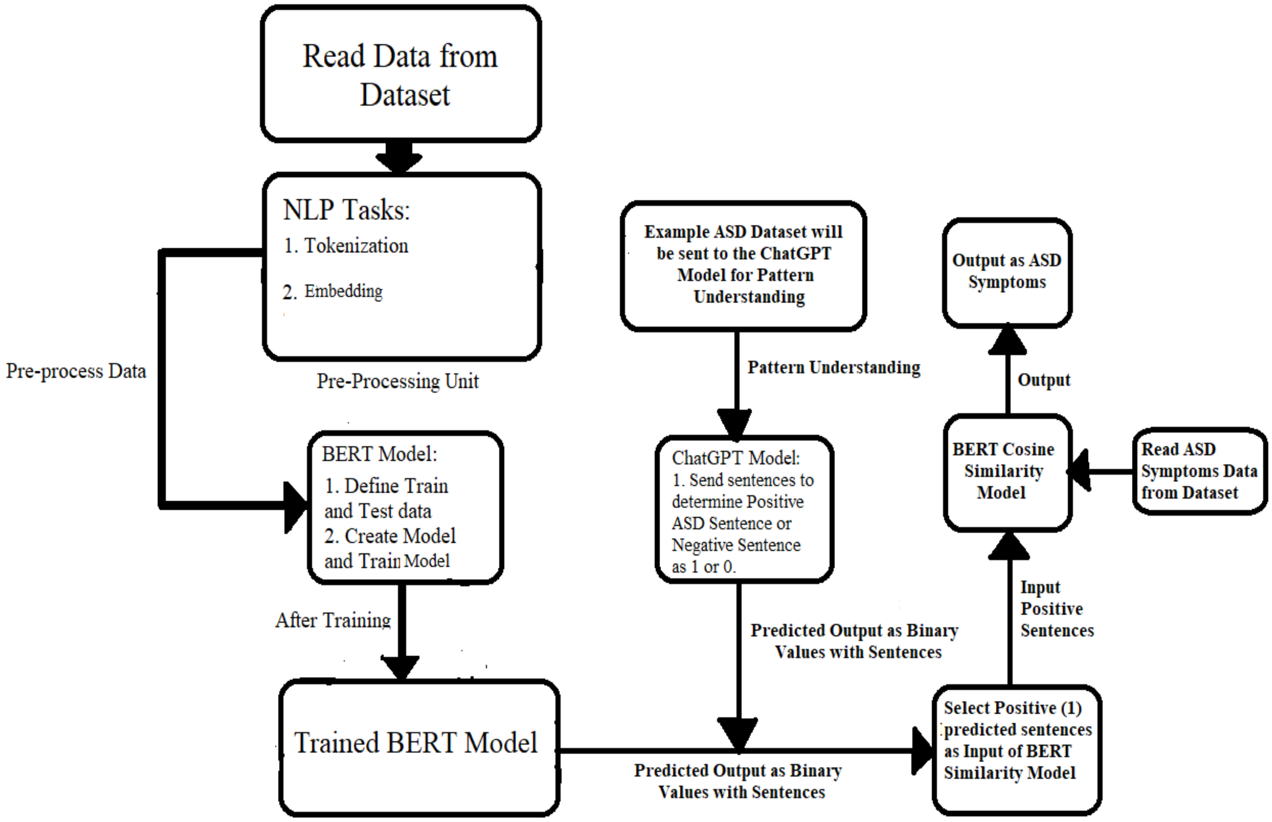

Fig. 7. General flow of proposed system architecture.

Below is the algorithm summarizing the steps:

1) Select positive sentences from the input text.
2) Apply the BERT Cosine Similarity Model.
3) Calculate the cosine similarity between the input sentence and each positive sentence from the ASD symptoms dataset.
4) Select the sentence with the highest cosine similarity score.
5) Retrieve the label associated with the selected sentence from Table IV, indicating the corresponding ASD problem.

By utilizing the BERT Cosine Similarity Model and leveraging the labels from Table IV, the proposed system can effectively identify ASD problems based on the similarity between input sentences and the ASD symptoms dataset.

The Algorithm in Python pseudo-code:

```
from sentence_transformers import SentenceTransformer, util
import pandas as pd
import pandasql as ps

// Intialize dataset in Dataframe

df = pd.read_csv(r"ASD_Symptoms.csv",encoding='Latin-1')
// List Array declare to store 'Comments', 'Sentiment value' and 'Cosine Score value

comments=[]
sentiment=[]
cosine_value=[]
// Function for bracket remove from Cosine value in Python
def listToStringWithoutBrackets(list1):
   return str(list1).replace('[','').replace(']','')
// BERT Cosine calculation function
def BERT_Cosine(strs):
  for ind in df.index:
   #print(df['Comments'][ind],df['Sentiment'][ind])

   sentences = [df['Comments'][ind],strs]

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

   #Compute embedding for both lists
   embedding_1= model.encode(sentences[0], convert_to_tensor=True)
   embedding_2 = model.encode(sentences[1], convert_to_tensor=True)

   comments.append(df['Comments'][ind])
   sentiment.append(df['Sentiment'][ind])
   score=util.pytorch_cos_sim(embedding_1,embedding_2)

cosine_value.append(listToStringWithoutBrackets(score.tolist()))

  dfc=pd.DataFrame(
   {'Comments': comments,
```

```
 'Sentiment': sentiment,
 'Cosine_Scores': cosine_value
})

// Dataframe to csv conversion with Cosine Score of each sentence for extraction of Max Cosine value with corresponding label value.
 dfc.to_csv('ASD_Cosine_Data.csv")
dfc['Cosine_Scores']=dfc['Cosine_Scores'].astype('float64')
  i = dfc['Cosine_Scores'].idxmax()
   return dfc['Sentiment'][i]

strs=pd.read_csv("ASD_New_Data.csv")
for st in strs['Comments']:
      result=BERT_Cosine(st)
      print(st,"=",result)
```

IV. RESULTS AND DISCUSSION

A. Result and Discussion of BERT Model

1) Result: The BERT model has been evaluated using some popular metrics like F1, Precision, Recall, Support, etc. The proposed research uses the BERT model as the binary classifier that will detect the sentences as positive or negative regarding ASD symptoms detection. The report of this binary classification has been given in Table III. But, before understanding the classification report, the confusion matrix of this proposed BERT model has been elaborated for a better clear view of this model.

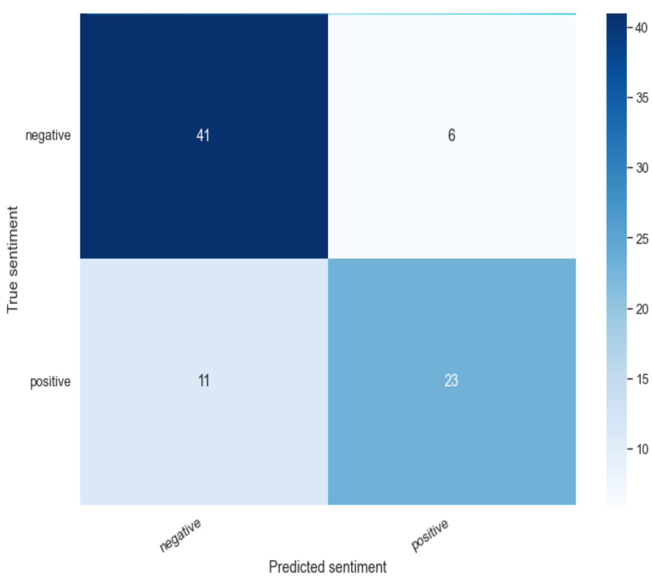

Fig. 8. Confusion matrix of proposed BERT model.

2) Discussion: According to Fig. 8. The proposed confusion matrix has two axes. The Y-axis is a true sentiment and X-axis is a predicted sentiment. True sentiment has both values positive and negative whereas predicted sentiment has the same positive and negative values. The number of sentences is 11 which are true positive sentiments but predicted as negative. In the next step, the number of true positives is 29 and the model has predicted it as true. According to Fig. 7, true negative sentences are 41, and the proposed model has predicted these sentences as negatives whereas six true negative sentences have been predicted by this proposed model as positives. The number of correct predictions according to the true sentiment is greater than wrong predictions. The F1 score, Precision, Recall, and Support have been given in Fig. 8.

The F1 score is an important metric for the evaluation of machine-learning models. The F1 score will be calculated by the combination of Precision and the Recall value. The equation of the F1 score calculation has been given here.

$$F1= (2 * (Precision * Recall)) / (Precision + Recall)$$

The Precision can be calculated using this formula where the number of True Positives (TP) is divided by the Total Number of True Positives (TP) and False Positives.

$$Precision= (TP/ (TP+FP))$$

The Recall can be formulated by the number of True Positives (TP) divided by the total number of True Positives (TP) and False Negatives (FN).

$$Recall= (TP/ (TP+FN))$$

The macro average computes the arithmetic mean according to the class. It is a straightforward average calculation method where the weighted average method refers to the proportion of each class support that is related to the sum of all support values. The count of correct occurrences of the class in the proposed dataset is called support. The imbalance support indicates the low scores of the report then the classification model can be remodified again.

	precision	recall	f1-score	support
negative	0.84	0.87	0.85	47
positive	0.81	0.76	0.79	34
accuracy			0.83	81
macro avg	0.82	0.82	0.82	81
weighted avg	0.83	0.83	0.83	81

Review text: My 3 years baby is noneverbal and less eye contact
Sentiment : positive

Fig. 9. Classification report of proposed BERT model.

Fig. 9 shows the classification report of the proposed BERT model where precision and recall values of negative and positive classes are 0.84, 0.87, 081, and 0.76 with support values 47 and 34. The F1 score of the negative and positive classes are 0.85 and 0.79. The accuracy of the proposed model is 0.83 (83%) according to the F1 score with 81 support values. The precision and recall values of macro and weighted averages are 0.82, 0.82, 0.83, and 0.83 where f1 scores are 0.82 and 0.83 with support values 81 and 81. An example sentence has been sent to the proposed BERT model for sentiment prediction.

Review text: “My 3 years baby is nonverbal and less eye contact”

This sentence directly indicates the positive sentence of ASD symptoms because ASD children are nonverbal and have less eye contact. The predicted result can be seen in Fig. 9 as “Sentiment: positive”.

B. Result and Discussion of ChatGPT Model

1) Result: A pre-trained ChatGPT model has been used where the base model is ‘text-davinci-003’ which is a very powerful base language model in ChatGPT. Author Junjie Ye and et.al. described ‘text-davinci-300’ language model performance on various datasets. The classification report on ‘‘text-davinci-300’ language model has been elaborated in detail. The GPT series models like GPT-3, CodeX, InstructGPT, and ChatGPT have been considered to evaluate the performance on nine natural language understanding (NLU) tasks using 21 datasets as in [23]. The base models have been used to train their datasts. OpenAI has given the opportunity to train their base modeel on a particular dataset where each dataset has to be designed according to their dataset standard. However, the concern is that the base model is not fine-tuned as a ChatGPT pre-trained model. An API key is needed to use base ChatGPT base models and fine-tuned pre-trained base models. The fine-tuned pre-trained model has been considered in this research to utilize the advantages of fine-tuning and pre-trained. The main advantage is the accuracy of the prediction of the sentences regarding ASD symptoms.

The following code will show the initialization part of the fine-tuned pre-trained “text-davinci-003” base model of ChatGPT.

```
def generate_gpt3_response(user_text, print_output=False):
    completions = ai.Completion.create(model='text-davinci-003',    #davinci:ft-personal-2023-02-02-06-04-19',
    temperature=0.5,
    prompt=user_text,
    max_tokens=100,
        n=1,
    stop=[" Human:", " AI:"],
  )
```

2) Discussion: The output from the proposed ChatGPT model according to the new parent’s dialogues is good. An example sentence with sentiment value has to be sent to the ChatGPT model then it can understand the pattern and according to the pattern, it will start prediction according to the defined RLHF algorithm.

The output can be seen from the given Fig. 10 on sentences of parents’ dialogues. The sentences are-

a) “*My* baby is not responding and his eye contact is very low."

b) “*from* last few days my son is obsessed with the placement of various things"

The first sentence is 1 means positive according to ASD symptoms whereas the second sentence is negative and ChatGPT returns 0.

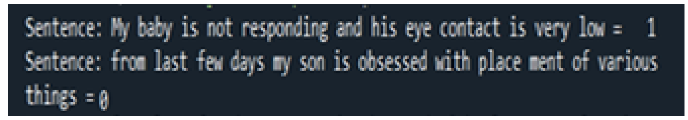

Fig. 10. Output of the proposed ChatGPT model.

C. Result and Discussion of BERT Cosine Model

1) Result: The proposed cosine similarity model is responsible for identifying ASD symptoms from positive ASD sentences that have been predicted by the machine learning algorithms. The model's output is depicted in Fig. 11, where the sentence " She does not have any eye contact when I call" is labeled with 6.

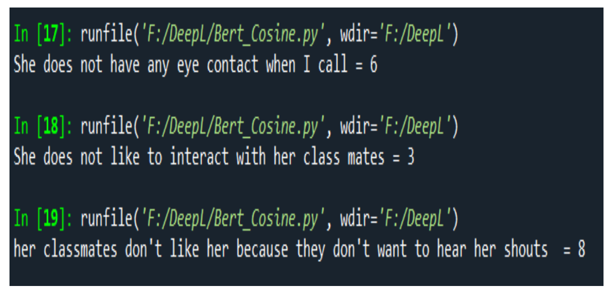

Fig. 11. Example output result of BERT cosine similarity.

2) Discussion: Similarly, the sentences " She does not like to interact with her classmates" and " her classmates don't like her because they don't want to hear her shouts" are labeled with 3 and 8, respectively. Referring to Table IV, we can determine that label 6 corresponds to Eye Contact problems, label 3 indicates Behaviour problems, and label 8 represents hyperactive problems. These labels provide insight into the specific ASD problems associated with the detected symptoms. Upon identifying the ASD problems, appropriate therapies can be initiated based on the specific problem identified. Tailoring the interventions according to the detected problems is crucial in providing targeted support to individuals with ASD. These targeted therapies have the potential to significantly reduce the symptoms associated with ASD and improve overall well-being. The ability of the proposed system to accurately identify ASD problems through the analysis of positive sentences enables a more focused and tailored approach to therapeutic interventions. This personalized approach holds great promise in positively impacting individuals with ASD and enhancing their quality of life. By leveraging the system's ability to identify ASD problems accurately, individuals with ASD can receive more effective and personalized support, leading to improved outcomes and overall well-being.

V. LIMITATION

The proposed system leverages both transformer-based and LLM-based machine learning models to enhance its performance. Quantum machine learning models can be employed to improve accuracy by applying them to a large dataset. Additionally, collecting more data, especially accurate ASD-related parent dialogs, can significantly enhance the training and testing scores of the models. According to the language model, it has been observed that they are not very efficient in calculations like aggregation-type natural language response generation. As an example, ChatGPT is not able to do the right calculation of the sum of multiple float values at a time. This is the main limitation of language models. It is vital to ensure that all components of the system are functioning properly for the cosine similarity part to work effectively. Any issues or malfunctions in the system's components can impact the performance of the cosine similarity model, as it relies on accurate processing and selection of positive sentences.

VI. CONCLUSION

The proposed system is designed to process natural language text extracted from parents' dialogues to detect ASD symptoms. It utilizes sentiment analysis techniques to determine whether a sentence expresses positive or negative sentiments regarding ASD symptoms. To accomplish this task, the system employs BERT and ChatGPT models that have been trained on the provided dataset. After performing sentiment analysis, the system selects only the positive sentences for further analysis using the cosine similarity model. The system utilizes an ASD symptoms dataset, where each sentence is labeled with a value corresponding to a specific ASD symptom. By calculating the cosine similarity between the input sentence and each sentence in the ASD symptoms dataset, the system identifies the label value of the sentence with the highest similarity score. This label value indicates the specific ASD problem associated with the input sentence. One significant advantage of the proposed system is that it relies solely on text-based analysis. By leveraging text-based analysis and prioritizing affordability and accessibility, the proposed system has the potential to facilitate ASD detection and support in underserved regions, thus bridging the gap in ASD diagnosis and intervention.

FUTURE WORK

To further enhance the output and accuracy of the proposed the provided dataset can be a valuable approach. This model can leverage the dataset to make predictions and improve the system's performance. By incorporating these models into the system and training them using the proposed dataset for ASD detection, the system can benefit from their advanced techniques and potentially achieve superior results. However, it is important to consider that transformer-based and LLM-based models like BERT and ChatGPT may not perform optimally when dealing with aggregation-type sentences. Therefore, it is crucial to carefully assess the sentences in the dataset and choose appropriate models accordingly. In summary, training the Quantum machine learning models using the proposed dataset for ASD detection represents a promising future development for the system.

ACKNOWLEDGMENT

The authors extend their appreciation to the Manipur International University, Imphal, India for supporting this Post-Doctoral (D.Sc.) research work on Autism.

REFERENCES

[1] C. Lord, M. Elsabbagh, G. Baird, J. Veenstra-Vanderweele, “Autism spectrum disorder", Lancet, vol. 392, pp. 508-520, 2018.

[2] Azian Azamimi Abdullah, Saroja Rijal, Satya Ranjan Dash, "Evaluation on Machine Learning Algorithms for Classification of Autism Spectrum Disorder (ASD)", International Conference on Biomedical Engineering, pp.1-7, 2019.

[3] Maitha Rashid Alteneiji, Layla Mohammed Alqaydi, Muhammad Usman Tariq, "Autism Spectrum Disorder Diagnosis using Optimal Machine Learning Methods", International Journal of Advanced Computer Science and Applications, vol. 11(9), pp.252-260, 2020.

[4] D. Wall, J. Kosmicki, T. Deluca, E. Harstad, and V. Fusaro, “Use of machine learning to shorten observation-based screening and diagnosis of autism", Translational psychiatry, vol. 2(4), pp. e100, 2012.

[5] Kruthi C H, Tejashwini H N , Poojitha G S, Shreelakshmi H S, Shobha Chandra K, "Detection of Autism Spectrum Disorder Using Machine Learning", International Journal of Scientific Research & Engineering Trends, vol. 7(4), pp. 2267-2271, 2020.

[6] F. Tabtah, "Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fullfillment", 1st International Conference on Medical and Health Informatics, pp.1-6, 2017.

[7] D. Aarthi, M. Udhayamoorthi, G. Lavanya, “Autism Spectrum Disorder Analysis Using Artificial Intelligence: A Survey", International Journal of Advanced Research in Engineering and Technology, vol. 11(10), pp. 235-240, 2020.

[8] T. Amalraj Victoire, A. Ramalingam, A. Naresh, K.M. Nasimudeen, M. S. Jaya Kumar, “An Efficient Approach to Detect Autism in Child Using Machine Learning and Deep Learning”, Journal of Theoretical and Applied Information Technology, vol. 99(20), pp. 4759-4769, 2021.

[9] R. Gandhi, towardsdatascience. Retrieved june 7, 2018, fromtowardsdatascience.com:https://towardsdatascience.com/supportvector-machine-introduction-to-machine-learning-algorithms-934a444fca47, 2018.

[10] F. Tabtah, "A mobile app for ASD screening", www.asdtests.com [accessed December 20th, 2017.

[11] F. Tabtah, "Machine Learning in Autistic Spectrum Disorder Behavioral Research: A Review: To Appear in Informatics for Health and Social Care Journal. December, 2017.

[12] American Psychiatric Association, (2013, 10 10). American Psychiatric Association, Retrieved from American Psychiatric Association: https://www.psychiatry.org/, 2013.

[13] Abhijit Mohanta, Vinay Kumar Mittal, "Autism Speech Analysis Using Acoustic Features", 16th Intl. Conference on Natural Language Processing, pp. 85-94, 2019.

[14] Sherif Kamel, Rehab Al-harbi, “Newly Proposed Technique for Autism Spectrum Disorder Based Machine Learning", International Journal of Computer Science & Information Technology (IJCSIT), vol 13(2), pp. 1-15, 2021.

[15] Riccardo Fusaroli, Ethan Weed, Deborah Fein, Letitia Naigles, "Caregiver linguistic alignment to autistic and typically developing children: a natural language processing approach illuminates the interactive components of language development", Cognition, pp. 1-66, 2023.

[16] Rafael A Calvo, David Nicholas Milne, M. Sazzad Hussain, Helen Christensen, "Natural language processing in mental health applications using non-clinical texts", Natural Language Engineering, vol. 23(5), pp.649-685, 2017.

[17] Amrutha S M, K R Sumana, "Autism Spectrum Disorder Detection Using Machine Learning Techniques", International Research Journal of Engineering and Technology (IRJET), vol. 8(8), pp. 1252-1254, 2021.

[18] Firsanova Victoria Igorevna, "The Description of the Autism Spectrum Disorder Question Answering Dataset", 27th International Conference on

Computational Linguistics and Intelligent Technologies Dialogue, pp. 1-8, 2021.

[19] Johnny Downs, Sumithra Velupillai et al., "Detection of Suicidality in Adolescents with Autism Spectrum Disorders: Developing a Natural Language Processing Approach for Use in Electronic Health Records", AMIA-2017, pp. 641-649, 2017.

[20] Parisa Moridian, Navid Ghassemi et al., "Automatic Autism Spectrum Disorder Detection Using Artificial Intelligence Methods with MRI Neuroimaging: A Review", Front Mol Neurosci, pp. 1-51, 2022.

[21] Zhenyu Mao,Yi Su et al., "Spatio-temporal deep learning method for ADHD fMRI classification", Information Sciences, pp.1-11, 2019.

[22] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin, "Attention Is All You Need", 31stConferenceonNeuralInformationProcessingSystems(NIPS2017), pp. 1-15, 2017.

[23] Junjie Ye, Xuanting Chen, Nuo Xu et al., "A Comprehensive Capability Analysis of GPT-3 and GPT-3.5 Series Models", ArXiv, pp. 1-47, 2023

AUTHORS' PROFILE

Prasenjit Mukherjee has 14 years of experience in academics and industry. He completed his Ph.D. in Computer Science and Engineering in the area of Natural Language Processing from the National Institute of Technology (NIT), Durgapur, India under the Visvesvaraya PhD Scheme from 2015 to 2020. Presently, He is working as a Data Scientist at Vodafone Intelligent Solutions, Pune, Maharashtra, India, and doing his Post Doctoral (D.Sc.) in Computer Science from Manipur International University, Imphal, Manipur, India.

Gokul R S, a senior data scientist with over six years of experience in leading and executing data-driven projects for various domains. He is currently working as a deputy manager in Vodafone Intelligent Solutions, Pune in the area of Artificial Intelligence and Analytics. He received a B.Tech degree from Visvesvarava Technological University, Karnataka. His research areas of interest include AI automation, machine learning, natural language processing, computer vision, and big data, prompt engineering, Quantum technologies, etc.

Sourav Sadhukhan has above 5 years of experience in Law and Management. He completed his Graduation in LLB from Calcutta University, Kolkata, India, and Post Graduate Diploma in Management from Pune Institute of Business Management, Pune, India. Presently he is a student of Executive Post Graduation in Data Science and Analytics from the Indian Institute of Management, Amritsar, India.

Dr. Manish Godse has 27 years of experience in academics and industry. He holds Ph.D. from Indian Institute of Technology, Bombay (IITB). He is currently working as an IT Consultant in the Bizamica Software, Pune in the area of Artificial Intelligence and Analytics. His research areas of interest include automation, machine learning, natural language processing and business analytics. He has multiple research papers indexed at IEEE, ELSEVIER, etc.

Dr. Baisakhi Chakraborty received the PhD. degree in 2011 from National Institute of Technology, Durgapur, India in Computer Science and Engineering. Her research interest includes knowledge systems, knowledge engineering and management, database systems, data mining, natural language processing, and software engineering. She has several research scholars under her guidance. She has more than 60 international publications. She has a decade of industrial and 22 years of academic experience.