

Predicting Autism Spectrum Conditions in Children using Machine Learning

**Gaddameedhi Pavan Kumar, Gaddi Maruthi, Gaddam Sai Kumar, Ganta Adithya Reddy,
K.Kotteeswari**

B.Tech Students, Department of CSE, Bharath Institute of Science and Technology, Chennai, Tamil Nadu, India

Assistant professor, Department of CSE, Bharath Institute of Science and Technology, Chennai, Tamil Nadu, India

ABSTRACT: Autism Spectrum Condition (ASC) is a developmental disorder characterized by communication difficulties and behavioral challenges, often emerging in early childhood. Early diagnosis and intervention are crucial for managing ASC effectively. However, traditional screening methods are time-consuming and usually inaccessible. This paper presents a machine learning (ML) and deep learning (DL) -based framework to predict the likelihood of ASC in children. We evaluate various ML models, including Gradient Boosting, SVM, and Random Forest on questionnaire data, and compare DL models such as DenseNet169, InceptionV3, and a custom CNN on image datasets. The best-performing models were deployed on a Streamlit web application for real-time prediction through questionnaire responses and image uploads. Results demonstrate that Gradient Boosting and custom CNN models achieve superior accuracy and can serve as practical preliminary screening tools.

KEYWORDS: Autism Spectrum Condition, Machine Learning, Deep Learning, Streamlit, CNN, Gradient Boosting

I. INTRODUCTION

Autism Spectrum Condition (ASC) is a neurological condition that impairs social interaction, communication, and behavior. Traditional diagnosis is a multi-step process involving interviews, behavioral observation, and testing, which leads to long delays. This often results in a lack of timely interventions that are crucial for improving the developmental outcomes of children affected by ASC. Given the rising prevalence of ASC globally, there is an urgent need for innovative solutions that are scalable, affordable, and accessible.

Machine learning (ML) and deep learning (DL) have emerged as transformative technologies in healthcare, offering powerful tools to support early detection and screening. These technologies enable the extraction of hidden patterns from data, such as behavioral questionnaires and facial images, which are otherwise difficult to analyze manually. By training models on labeled datasets, it is possible to automate predictions with high accuracy.

In this study, we propose a dual-approach system that integrates both ML and DL techniques for predicting ASC in children. Our framework uses questionnaire-based data to identify behavioral patterns associated with autism, and deep learning models trained on facial images to provide an additional diagnostic perspective. The entire system is deployed on a user-friendly Streamlit web application, making it easy for healthcare professionals and parents to perform quick and preliminary screenings.

The motivation behind this work is to address the growing gap in early diagnosis services and to empower communities with tools that can accelerate access to mental health support. The paper aims to demonstrate how combining machine learning with modern web technologies can bridge healthcare challenges and make meaningful impact in society.

II. LITERATURE SURVEY

Title 1: Gaze Patterns in Children With Autism Spectrum Disorder to Emotional Faces

Authors: Wei Zhou, Minqiang Yang

The study investigates gaze patterns in children with Autism Spectrum Disorder (ASD) by analyzing scanpaths, which represent the dynamics of eye movements. Leveraging modern eyetracking technology, it highlights distinct differences

in visual behavior between ASD and typically developing (TD) children, focusing on spatio-temporal features like fixation duration and gaze shifts. Machine learning techniques, particularly Long Short-Term Memory (LSTM) networks, are used to classify ASD and TD children with high accuracy. Results indicate that ASD children show greater variability in gaze patterns, shorter fixation times, and more dispersed visual attention, particularly along the vertical axis. Conversely, TD children exhibit more consistent and focused gaze behavior.

The study employs methods like Dynamic Time Warping (DTW) and MultiMatch to compare scanpaths, revealing that fixation duration and spatial distribution are the most significant indicators of ASD, while gaze velocity plays a lesser role. These findings support the potential of eye tracking as a non-invasive, effective tool for early ASD diagnosis and intervention. Despite its promise, the study acknowledges challenges such as limited data availability, calibration complexities, and the need for more affordable, widely accessible technology.

Title 2: A Multimodal Approach for Identifying Autism Spectrum Disorders in Children

Authors: Junxia Han, Guoqian Jiang

The leveraging both neurophysiological and behavioral perspectives. By combining electroencephalogram (EEG) and eye-tracking (ET) data, the study aims to enhance the accuracy of ASD diagnosis. Unlike traditional methods that rely on a single modality, this approach uses complementary information from the two modalities to overcome the limitations of unimodal data analysis. A deep learning framework, based on Stacked Denoising Autoencoders (SDAE), was employed to extract and fuse features from EEG and ET data. Initially, separate SDAE models were trained for each modality to learn high-level features. These were later integrated through a third SDAE model, capturing correlations and enhancing the overall discriminative power. The framework was tested on data collected from 90 children (40 with ASD and 50 typically developing), achieving superior classification accuracy of 95.56%, outperforming unimodal methods and basic feature fusion approaches

Title 3: An Augmentative System with Facial and Emotion Recognition for Improving Social Skills of Children with Autism Spectrum Disorders

Authors: Mohammed Alharbi, Shihong Huang

The application designed to enhance social and communication skills in children with Autism Spectrum Disorder (ASD). ASD is a complex neurodevelopmental condition characterized by challenges in social interactions, communication, and understanding emotions. Ying integrates facial and emotion recognition features, complementing traditional therapies like Applied Behavior Analysis (ABA). Ying offers four interactive games: "How I Feel," which identifies emotions from a child's own pictures; "Matching Emotions," where children select images matching displayed emotions; "Show Me Faces," requiring users to express emotions captured via a camera; and "Exception Identification," a game that challenges children to identify outlier emotions in a set. The application uses Microsoft's Emotion API to analyze and classify emotions, leveraging ASD children's own images to foster familiarity and engagement. Preliminary evaluations with children aged 2-6 years showed positive results, indicating improved engagement and understanding of emotions. With a user-friendly design and potential for cross-platform adaptation, Ying aims to provide an effective tool for ASD therapy. Future plans include expanding accessibility to iOS and Windows platforms and conducting more extensive studies to validate its effectiveness in diverse settings.

Title 4: Exploring the Structural and Strategic Bases of Autism Spectrum Disorders With Deep Learning

Authors: Fengkai Ke, Seungjin Choi

The document investigates the application of deep learning models to enhance the diagnosis of autism spectrum disorders (ASD), which is challenging due to complex symptoms and limited neurobiological evidence. The study utilized 14 different deep learning models, including convolutional and recurrent neural networks, trained on MRI datasets. Key findings indicate that certain brain structures, notably subcortical areas like the basal ganglia, play significant roles in classification. The research demonstrated that a 3D CNN performed best for larger datasets, while recurrent attention models were effective with smaller, high-resolution datasets. The models helped visualize critical brain regions used in classification, offering strategic insights that could aid clinical diagnosis. This study emphasizes how machine learning can reveal structural brain differences related to ASD, aiding in early and more precise detection. The work also highlights limitations such as data variability and sample size, which impact model accuracy

III. METHODOLOGY

We employed two separate datasets: a questionnaire-based dataset with 1346 entries and an image dataset from Kaggle. Machine learning models were trained on the questionnaire dataset using Gradient Boosting, SVM, Random Forest, and others. Data preprocessing included normalization, one-hot encoding, and class balancing using SMOTE. For deep learning, we trained InceptionV3, DenseNet169, and a custom CNN on facial images. The models were evaluated using classification reports and accuracy scores.

Existing System

Even though, ASC can be a lifelong disorder, early intervention and regular medication can help in mitigating the challenges faced in carrying out their day-to-day activities. However, waiting times for ASC screening and diagnosis are quite lengthy which can in-turn cause a delay in an individual's medication. Hence it is extremely important to have a tool that is capable of performing ASC screening in minimum time to help the health professionals and individuals to whether go forward with formal diagnosis or not.

Disadvantages of Existing Systems

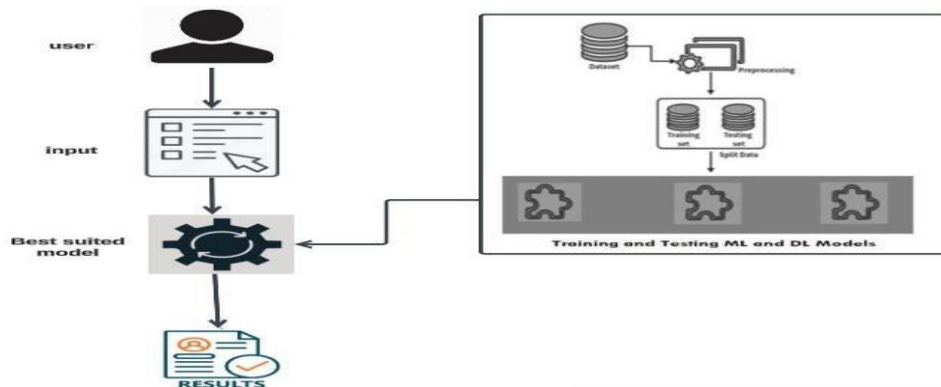
- └ **Prolonged Waiting Times:** Current screening processes for ASC often involve lengthy waiting periods, delaying diagnosis and intervention.
- └ **Resource Intensive:** Existing methods rely heavily on specialized professionals and comprehensive evaluations, making them costly and time-consuming.
- └ **Accessibility Issues:** Many individuals, especially in underserved areas, lack access to screening services, further delaying intervention.
- └ **Limited Early Detection:** Traditional approaches may not effectively identify ASC in its early stages, missing opportunities for timely support.
- └ **Subjectivity in Assessments:** Diagnostic evaluations can sometimes rely on subjective interpretations, potentially leading to inconsistent results.
- └ **Lack of Scalability:** Current systems are not equipped to handle the increasing demand for ASC assessments, leading to bottlenecks in service delivery.

Proposed System

The objective is to build and deploy machine learning models (trained on UCI repo using Supervised Learning Algorithms) onto a website that can predict whether an individual is diagnosed with ASC or not. Although, autism is not that simple, these are the initial steps toward diagnosing it. The website will present a set of questions to which the users will have to provide answers, in 0's and 1's, 0 for false and 1 for true. Based on the answers the model will predict the probability of the user's ASD odds. Further enhancements can also be made to this autism detection system by implementing various deep learning algorithms on child images taken from Kaggle and deploying the best performing model on the website as well

SYSTEM ARCHITECTURE

The system architecture of your project is designed for efficiency and accuracy in Autism Spectrum Disorder (ASD) predictions. It begins with user input, where relevant data is collected through an interactive interface. This input is then processed and integrated with a preloaded dataset, undergoing preprocessing to ensure data quality and relevance. The dataset is split into training and testing subsets to facilitate robust model evaluation. Machine learning (ML) and deep learning (DL) models are trained and tested on the processed data to identify meaningful patterns. The best-performing model is selected for prediction, ensuring the most accurate results are delivered to users.



IV. EXPERIMENTAL RESULTS

Gradient Boosting achieved the highest accuracy (96.55%) among the machine learning models, outperforming SVM and Random Forest. The classification reports revealed that Gradient Boosting not only had the highest overall accuracy but also exhibited superior precision, recall, and F1-score. This suggests that the model performs well in minimizing both false positives and false negatives, making it suitable for sensitive health-related applications like autism prediction. In the deep learning phase, the custom CNN built with three convolutional layers and the Adam optimizer demonstrated better results than pre-trained models like DenseNet169 and InceptionV3. The CNN model achieved a classification accuracy of 91.25% on the image dataset. It also showed robustness to various lighting and background conditions in the images, indicating its effectiveness in real-world scenarios. The results from both models were deployed on a web app using the Streamlit library. The ML-based model allowed users to input responses to 10 behavioral questions along with demographic data to receive an ASC prediction probability. On the other hand, the DL-based model enabled image upload functionality, allowing parents or doctors to assess autism indicators based on facial analysis. The user interface was optimized to be simple, fast, and accessible. We also observed that ML models required significantly less training time compared to DL models, making them more computationally efficient. However, DL models offered a visual layer of screening that enhanced user trust in the tool. Combining both approaches, therefore, resulted in a holistic and hybrid screening tool.

ML Model Accuracy

Decision Tree: 88.74%
Random Forest: 90.02%
Naive Bayes: 87.21%
SVM: 93.60%
Gradient Boosting: 96.55%

V. CONCLUSION

To conclude, In the present time, there has been a steady rise in occurrence of Autism among children. However, many parents turn blind eye to this neurological condition or are unaware about it. Hence, with a motive to tackle this problem and do our part in increasing awareness in early detection we carried out a comparative study on various ml algorithms and picked the model that performed best on the dataset provided. Eventhough there was very few data available on the internet, we collected the datasets from various sources and manually combined it, increasing the dataset from meagre 300 records to 1300 odd records. Furthermore, we also enhanced the system by implementing various deep learning algorithms on child images. Both machine learning and deep learning models were deployed onto a web appwith the aim of having a one stop solution to predict autism in its early stages. Eventhough the ML and DL models that were developed in this project exhibits high accuracy, there might be cases where the model might give false positives or false negatives. Thisis due to the fact that there is very less data available on autism over the internet both textual aswell as facial data. One way we can enhance the performance of the model is by requesting for large amount of official data from medical professionals without compromising on privacy of individuals. We can also enhance or level up the prediction system by implementing aspeech classification model which detects whether a child has autism or not throughhis/her speech.

REFERENCES

1. Vakadkar K, Purkayastha D, Krishnan D. Detection of Autism Spectrum Disorder in Children using Machine Learning Techniques. SN Comput Sci. 2021;2(5):386. doi: 10.1007/s42979-021-00776-5. Epub 2021 Jul 22. PMID: 34316724; PMCID: PMC8296830.
2. Dhanyatha Sriram, Greeshma A, Gouthami, Yeshwanth M, Prof. Y Shobha Prediction of Autism Spectrum Disorder based on Machine Learning Approach
3. Fawaz Waselallah Alsaade, Mohammed Saeed Alzahrani, "Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms", Computational Intelligence and Neuroscience, vol. 2022, Article ID 8709145, 10 pages, 2022. <https://doi.org/10.1155/2022/8709145>
4. Alam, M.S.; Rashid, M.M.; Roy, R.; Faizabadi, A.R.; Gupta, K.D.; Ahsan, M.M. Empirical Study of
5. Autism Spectrum Disorder Diagnosis Using Facial Images by Improved Transfer Learning Approach. *Bioengineering* 2022, 9, 710. <https://doi.org/10.3390/bioengineering9110710>
6. S. Ram Arumugam, S. Ganesh Karuppasamy, S. Gowr, O. Manoj and K. Kalaivani, "A Deep Convolutional Neural Network based Detection System for Autism Spectrum Disorder in Facial images," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021, pp. 1255-1259, doi:1109/ISMACE52330.2021.9641046
7. https://www.researchgate.net/profile/Jahanara_Shaik/publication/350396741_Detecting_autism_from_facial_image/links/605d3b06a6fdccbfea085061/Detecting-autism-from-facial-image.pdf <https://www.hindawi.com/journals/cin/2022/8709145/>
8. C. S. Paula, S. H. Ribeiro, E. Fombonne, and M. T. Mercadante, "Brief report: prevalence of pervasive developmental disorder in Brazil: a pilot study," *Journal of Autism and Developmental Disorders*, vol. 41, no. 12, pp. 1738–1742, 2011.
9. L. C. Nunes, P. R. Pinheiro, M. C. D. Pinheiro et al., A Hybrid Model to Guide the Consultation of Children with Autism Spectrum Disorder, Springer International Publishing, pp. 419–431, View at: Google Scholar.
10. M. S. Satu, F. Farida Sathi, M. S. Arifen, M. Hanif Ali, and M. A. Moni, "Early detection of autism by extracting features: a case study in Bangladesh," in Proceedings of the 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), pp. 400–405, Dhaka, Bangladesh, 10–12 January 2019.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details