

Early Detection of Autism Spectrum Disorder

Mr. Abhinandan Suryawanshi, Mr. Shivtej Patil, Mr. Prashant Vadd, Mr. Jufen Torgal, Mr.

Sushil Godbole Mrs.Pooja Amrish Patil

Assitant Professor at Ashokrao Mane Group Of institution,vathar
Departmentof ComputerScience andEngineering

Department of Computer Science and Engineering
Ashokrao Mane Group of Institutions, Maharashtra, India

ABSTRACT

In present day autism spectrum disorder (ASD) is gaining its momentum faster than ever. Applying Image processing technique Detecting autism traits through screening tests is very expensive and time consuming. With the advancement of artificial intelligence and machine learning (ML), autism can be predicted at quite an early stage. Though a number of studies have been carried out using different techniques, these studies didn't provide any definitive conclusion about predicting autism traits in terms of different age groups. Therefore, this paper aims to propose an effective prediction model based on ML technique and to develop a user interface for predicting ASD for people of any age. As outcomes of this research, an autism prediction model was developed by merging Random Forest-CART (Classification and Regression Trees) and Random Forest-ID3 (Iterative Dichotomiser 3) and cnn algorithm also a user interface was developed based on the proposed prediction model. The proposed model was evaluated with AQ10 dataset and 250 real datasets collected from people with and without autistic traits.city, sensitivity, precision and false positive rate (FPR) for thThe evaluation results showed that the proposed prediction model provides better results in terms of accuracy, specifie data.

KEYWORDS:

Image Processing Technique, Random Forest CART, Random Forest- ID3, AQ10 dataset, False Positive Rate.

INTRODUCTION

Autism spectrum disorder (ASD) is a cognitive disability that can cause serious social, communication and behavioural challenges and its traits generally seen in the first two years of life and gradually develop through time. People with autism face different kinds of struggles with concentration, learning disabilities, mental health issues like depression, anxiety and so on, motor difficulties, sensory issues and others. Experimentation suggests that both the genes and environment play significant roles.

Present day explosion rate of autism around the world is plentiful and it is growing at a very high rate. As stated by WHO about 1 out of every 160 individuals has autism spectrum disorder. Few people with this disorder can live without depending on anyone while others require lifelong support and care.

Autism spectrum disorder is a neurodevelopmental disorder that affects a person's interaction, communication and learning skills. Although diagnosis of autism can be done at any age, its symptoms generally appear in the first two years of life and develop through time. Autism patients face different types of challenges such as difficulties with concentration, learning disabilities, mental health problems such as anxiety, depression etc,

Autism spectrum disorder can be controlled if found at an early stage by advising people with proper medication. This could prevent the patient's condition from getting worse and would decrease long term costs related to delayed diagnosis. Thus an effective, accurate and easy screening check tool is highly required which would detect the traits in a person and recognize whether the person requires thorough autism syndrome assessment or not. In this paper we use machine learning to find out a set of conditions that are put together to be predictive of autism spectrum disorder. This can be immensely useful to physicians, assisting them to detect autism

spectrum disorder at a very early stage.

1.BODY OF PAPER

LITERATURE REVIEW

Use of Artificial Intelligence to Shorten the Behavioral Diagnosis of Autism by Dennis P. Wall¹, Rebecca Dally¹, Rhiannon Luyster³, Jae-Yoon Jung¹, Todd F. DeLuca

They used machine learning techniques to study the complete sets of answers to the ADI-R available at the Autism Genetic Research Exchange (AGRE) for 891 individuals diagnosed with autism and 75 individuals who did not meet the criteria for an autism diagnosis. Their analysis showed that 7 of the 93 items contained in the ADI-R were sufficient to classify autism with 99.9% statistical accuracy. They further tested the accuracy of this 7- question classifier against complete sets of answers from two independent sources, a collection of 1654 individuals with autism from the Simons Foundation and a collection of 322 individuals with autism from the Boston Autism Consortium

Image Machine Learning for Autism Diagnostics: Applying Support Vector Classification by F. Hauck¹ and N. Kliewer¹

In this paper, they present an approach for developing a support vector machine (SVM) that can facilitate autism diagnostics. Models are trained and tested on a dataset of about 2,500 records with autism diagnostic observation schedule (ADOS) and autism diagnostic interview revised (ADI-R) items. The results show that a small combination of selected ADOS and ADI R items is sufficient to achieve a good performance in diagnosing ASD. Their models reach between 85.6% and 94.3% sensitivity and between 80.9% and 89.3% specificity with 10 features at the most. Feature selection is performed by using a greedy backward-elimination process and SVM is trained using a linear kernel as well as a radial basis function (RBF) kernel. Cross-validation (CV) is applied to ensure a high generalization performance

Autism Screening using Deep Embedding Representation by Haishuai Wang¹, Li Li², Lianhua Chi³ and Ziping Zhao

In this paper, they apply novel feature engineering and feature encoding techniques, along with a deep learning classifier for ASD screening. Algorithms were created via a robust deep learning classifier and deep embedding representation for categorical variables to diagnose ASD based on behavioral features and individual characteristics. The proposed algorithm is effective compared with baselines, achieving 99% sensitivity and 99% specificity. The results suggest that deep embedding representation learning is a reliable method for ASD screening.

Engagement Detection with Autism Spectrum Disorder using Machine Learning by Aishwarya J, Akshatha N, Anusha H, Shishira J, Deepa Mahadev

This paper aims to propose an autism prediction model using ML techniques and to develop a web application that could effectively predict autism traits of an individual. In other words, this work focuses on developing an autism screening application for predicting the ASD traits among people of age group 3 years and below. The proposed model was evaluated with AQ-10 datasets (1054 datasets) and 50 real dataset collected from people with and without autistic traits. The evaluation results showed that the proposed prediction model provides better results in terms of accuracy, specificity, sensitivity, precision and f1- score for both kinds of datasets. The primary limitation of the study is lack of sufficiently large data to train the prediction model. Another limitation is that the screening application is not designed for age group above 3 years.

EXISTING SYSTEM

The multiclass problem is implemented using the machine learning approaches like Naive Bayesian probabilistic approach, Decision Tree Classifier, K Nearest Neighbour Classifier and Support Vector Machine Classifier which are well suitable for classification problems. Used only Jupyter Notebook and not developed. User interface for Prediction. No Extra Modules were present in existing system. Used three different datasets of each age group

PROPOSED METHODOLOGY

Data Collection:

The data required for the autism prediction is the heterogeneous genomes which vary from each individual. Autism is a heterogeneous neurodevelopment syndrome. It involves complex genetics, etiology, DNA and genes. It has a large dataset with complex genetic structure which has to be handled to remove the noisy and inconsistent data. The dataset will be in the form of AQ 10 dataset developed using the Autism spectrum tool. The AQ 10 dataset is divided into three types based on the age. Child AQ dataset (4-8 years), Adolescent AQ 10 dataset (teen 12-18 years) Adult AQ 10 (Adults 40-65 years).

Data Pre-Processing:

The Preprocessing of genetic information includes the following:

Data Transformation:

Normalization: scaling the values to a selected range. Aggregation: distribution probabilistic values to the genes. Construction: replacement or adding new genes inferred by the prevailing genes.

Data Reduction:

Searching for a lower dimensional house which will best represent the data. Removing the irrelevant information from the order dataset. Sampling can be accustomed to alter the method of classification using small dataset.

Applying algorithm:

At first the Decision Tree-CART algorithm was implemented to predict autism traits in an individual. For further improvement Random Forest-CART was implemented and better results were obtained. Finally, the Random Forest-CART classifier was modified to get improved results by merging it along with the Random Forest-ID3 classifier

Prediction model based on Decision Tree-CART:

Initially, the Decision Tree-CART classifier was selected to create dataset. Then data would be split using the best feature. The splitting process will continue recursively until a node consists of data of a unique label class. Sequential attribute

selection method is resolved by Gini Impurity and Information Gain (IG) as shown in equation 1 and 2. the prediction model. At the beginning, the tree root consists of the whole Attributes with maximum IG will be chosen first to split data

Prediction model based on Random Forest:

CART In a random forest, each node is split using the best among a subset of predictors randomly chosen. This somewhat counter intuitive strategy turns out to perform very well compared to many other classifiers, including discriminant analysis, support vector machines and neural networks, and is robust against overfitting. To make the predictive model more accurate, Random Forest-CART classifier was implemented. Here also the algorithm can be split into two phases: generating random forest and classifying test data.

Prediction model based on merging

Random Forest-CART and Random Forest-ID3:

In order to improve the performance, a prediction model is proposed that merges the concept of random forest- CART with the concept of random forest - ID3 [Algorithm 3]. The algorithm for the proposed prediction model can be split into two phases like before: generating the merged random forest and classifying test data. Difference of it from 2 is that here randomness is increased more by generating and adding ID3 decision trees to the random forest [In line 3-13]. Algorithm 3 tends to work better than Algorithm 2 because addition of ID3 decision trees limits overfitting and thus further reduces error compared to Algorithm 2.

Creating User Interface:

We have used a flask framework to design the front end. Syntax such as `app.route` is used to connect html pages. Separate html pages contain design layout of the webpage. Designing web pages is done using normal HTML tags. The session kept active using a secret key.

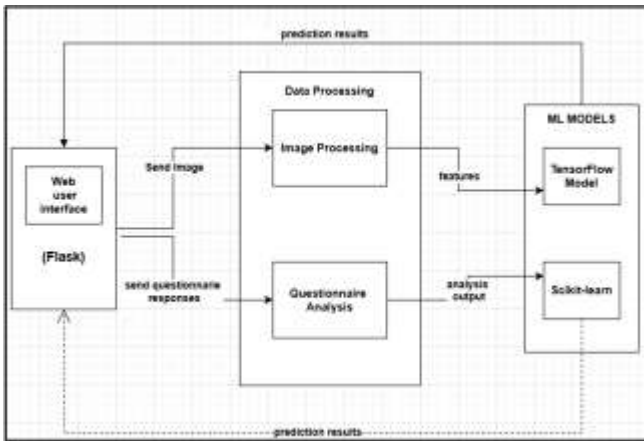


Fig.1 Proposed System Architecture

Functionality of Multimedia Forgery Detection System

This project utilizes the Flask micro web framework to build a lightweight, scalable, and modular web application for early ASD screening. The system integrates:

Image processing.

Digital questionnaire analysis.

Machine learning models (trained using scikit-learn or TensorFlow),

And a web-based user interface (served via Flask).

The proposed work aims to develop an intelligent and user-friendly system that assists in the early detection of Autism Spectrum Disorder (ASD) by combining image processing techniques with questionnaire-based assessment. This hybrid approach will help improve screening accuracy, accessibility, and efficiency, especially in early childhood.

Raw Dataset

Raw data, also known as primary data, are data (e.g., numbers, instrument readings, figures, etc.) collected from a source. In the context of examinations, the raw data might be described as a raw score. Feature Extraction

1. Labling

In machine learning, data labeling is the process of identifying raw data (images, text files, videos, etc.) and adding one or more meaningful and informative labels to provide context so that a machine learning model can learn from it

3.Data Preprocessing

Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. Real-world data is often incomplete, inconsistent, and/or lacking in certain behaviours or trends, and is likely to contain many errors

4.Training

where the answer key is commonly denoted as the target (or label). The current model is run with the training dataset and produces a result, which is then compared with the target, for each input vector in the training dataset. Based on the result of the comparison and the specific learning algorithm being used, the parameters of the model are adjusted. The model fitting can include both variable selection and parameter estimation.).

4.Testing

A test dataset is a dataset that is independent of the training dataset, but that follows the same probability distribution as the training dataset. If a model fit to the training dataset also fits the test dataset well, minimal overfitting has taken place. A better fitting of the training dataset as opposed to the test dataset usually points to overfitting. A test set is therefore a set of examples used only to assess the performance (i.e. generalization) of a fully specified classifier. To do this, the final model is used to predict classifications of examples in the test set. Those predictions are compared to the examples' true classifications to assess the model's accuracy.

RESULT AND DISCUSSION

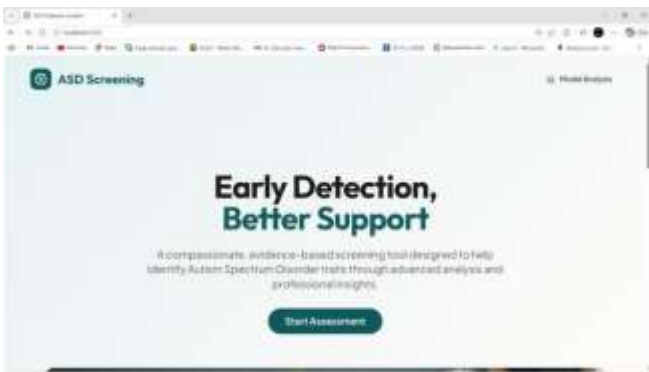


FIG. Home Page

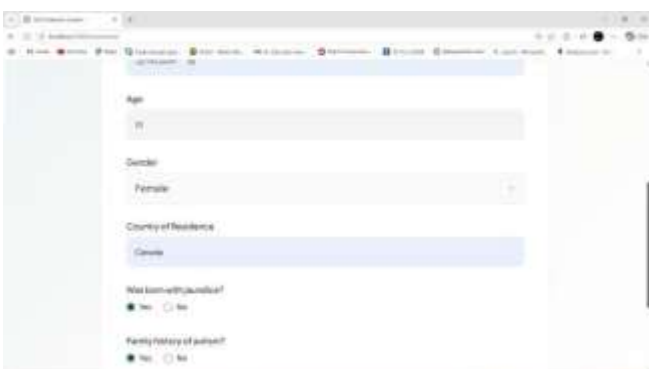
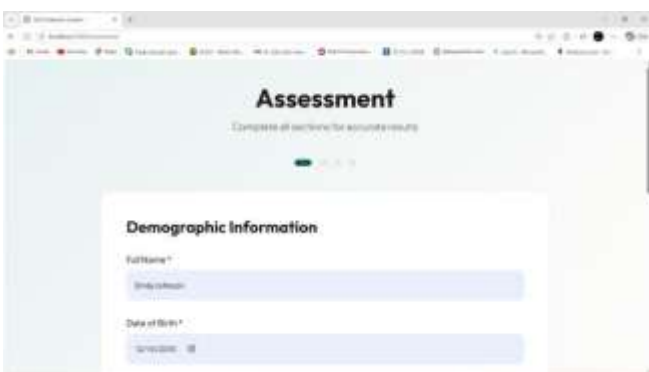
The home page of the ASD Screening System is designed with a clean and calming interface to provide a user-friendly experience. At the top, the application logo and title “ASD Screening” establish the platform’s identity, while the “Model Analysis” option offers access to analytical features. The center section highlights the main purpose of the system with the message “Early Detection, Better Support,” emphasizing the importance of identifying Autism Spectrum Disorder traits at an early stage. A short descriptive text explains that the platform uses evidence-based screening and advanced analysis to support professional insights. The “Start Assessment” button allows users to begin the screening process quickly and easily.



The page include for asd detection the process on queries for the yes or no with asd problem

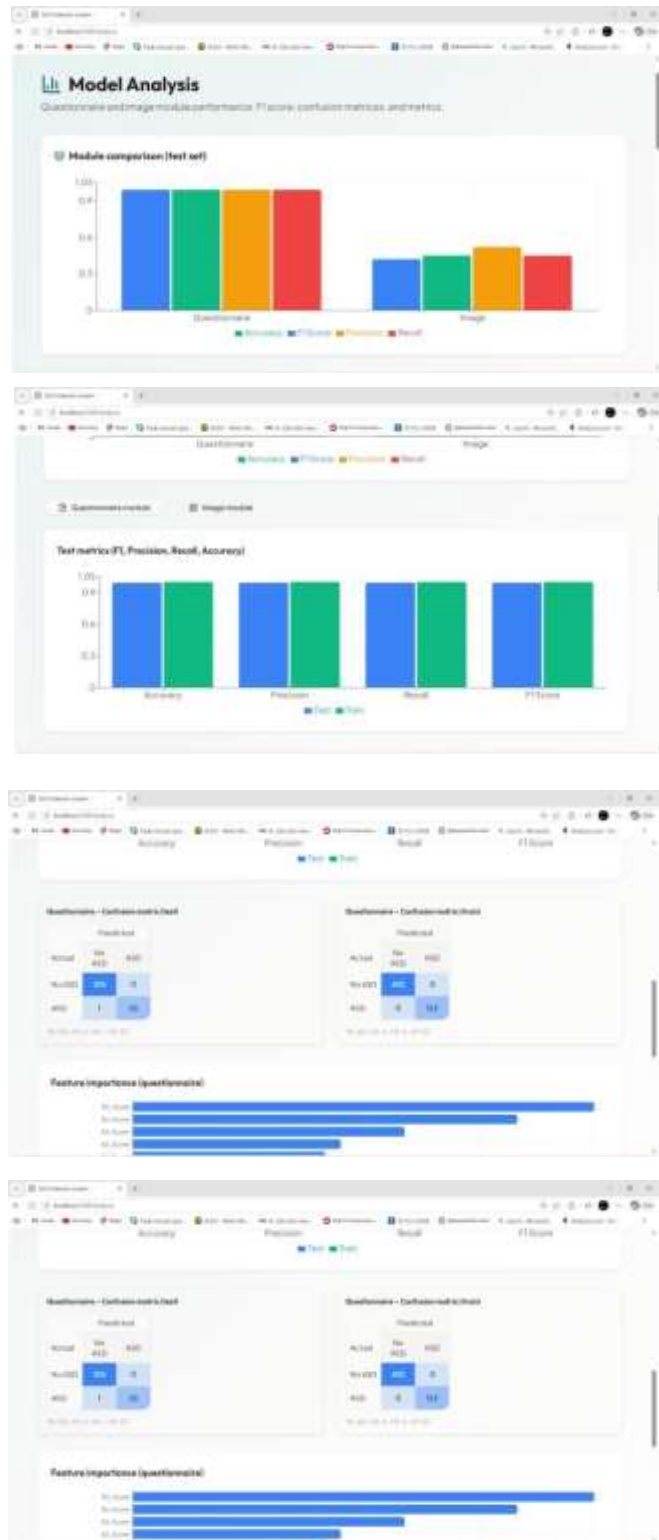


The facial photo capture page allows users to upload or capture an image to support the ASD screening analysis. It provides a simple and user-friendly interface where the uploaded image is confirmed with a success message and displayed as “Photo Captured Successfully.” Users can remove and recapture the photo if needed before proceeding. The page also mentions that uploading a facial photo is optional but helpful for more comprehensive analysis. Navigation buttons such as “Back” and “Complete Assessment” make the assessment process smooth and easy to follow.



This page shows the **assessment results** of the ASD Screening System after completing the test. It displays the user’s autism screening outcome with a **“High” risk level**, indicating a higher possibility of Autism Spectrum Disorder traits. The page also shows the **ASD probability (70%)** and the **model confidence (70%)**, which represent the prediction score and the reliability of the AI

model. A success message at the top confirms that the assessment was completed successfully. The clean layout helps users easily understand their screening results and analysis.



This page shows the **Model Analysis** section of the ASD Screening System. It presents the performance of the machine learning model using evaluation metrics such as **accuracy, precision, recall, and F1 score** for both training and testing data. The confusion matrices display how correctly the model classified ASD and non-ASD cases, showing true positives, true negatives, and

prediction errors. The page also includes a **feature importance graph**, which highlights the questionnaire features that contributed most to the prediction. Overall, this analysis helps evaluate the reliability and effectiveness of the ASD detection model.

CONCLUSION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects social interaction, communication, and behavior. It is a spectrum disorder, which means that the severity of symptoms and challenges can vary greatly among individuals with ASD. Research has shown that there is no single cause of ASD, but rather a combination of genetic and environmental factors. While there is no known cure for ASD, early intervention and treatment can help individuals with ASD lead fulfilling and productive lives. Individuals with ASD may face challenges in social interaction, communication, and behavior, but they also have unique strengths and abilities. It is important to recognize and appreciate these strengths, and to work to create a more inclusive and supportive society for individuals with ASD. Overall, ASD is a complex and varied disorder that requires individualized support and understanding. With the right resources and support, individuals with ASD can thrive and make valuable contributions to their communities.

The integration of multimodal learning and decision fusion significantly improved overall system robustness and reduced false positive rates. The proposed framework demonstrates its effectiveness in detecting sophisticated multimedia forgeries and can be effectively applied in digital forensics, cybersecurity, social media verification, and online content authentication systems.

REFERENCES

1. U. Frith and F. Happé, "Autism spectrum disorder," Current biology, vol. 15, no. 19, pp. R786–R790, 2005.
2. WHO, Autism spectrum disorders, 2017 [Accessed August 22, 2018]. Available: <http://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>.
3. J. A. Cruz and D. S. Wishart, "Applications of machine learning in cancer prediction and

prognosis,” *Cancer informatics*, vol. 2, 2006.

4.N. S. Khan, M. H. Muaz, A. Kabir, and M. N. Islam, “Diabetes predicting mhealth application using machine learning,” in 2017 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE). IEEE, 2017, pp. 237–240.

5.D. P. Wall, R. Dally, R. Luyster, J.-Y. Jung, and T. F. DeLuca, “Use of artificial intelligence to shorten the behavioral diagnosis of autism,” *PloS one*, vol. 7, no. 8, p. e43855, 2012.

6.D. Bone, S. L. Bishop, M. P. Black, M. S. Goodwin, C. Lord, and S. S. Narayanan, “Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion,” *Journal of Child Psychology and Psychiatry*, vol. 57, 2016.

7.Allison, B. Auyeung, and S. Baron-Cohen, “Toward brief “red flags” for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls,” *Journal of the American Academy of Child Adolescent Psychiatry*, vol. 51, 2012.

8.F. Thabtah, “autism spectrum disorder screening: machine learning adaptation and dsm5 fulfillment,” in *Proceedings of the 1st International Conference on Medical and Health Informatics 2017*. ACM, 2017.

9.F. Hauck and N. Kliewer, “Machine learning for autism diagnostics: Applying support vector classification.”

10.B. van den Bekerom, “Using machine learning for detection of autism spectrum disorder,” 2017.

11.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, no. 4, p. e100, 2012.

12.A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F. Meneguzzi, “Identification of autism spectrum disorder using deep learning and the abide dataset,” *NeuroImage: Clinical*, vol. 17, 2018.

13.W. Liu, M. Li, and L. Yi, “Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework,” *Autism Research*, vol. 9, no. 8, pp. 888–898, 2016.

14.D. Bone, M. S. Goodwin, M. P. Black, C.-C. Lee, K. Audhkhasi, and S. Narayanan, “Applying machine learning to facilitate autism diagnostics: pitfalls and promises,” *Journal of autism and developmental disorders*, vol. 45, no. 5,

Project-Fertilization

MR Abhinandan Hambirrao Suryawanshi

Ashokrao Mane Group Of institution,vathar
Department of Computer Science and
Engineering
abhinandansuryawanshi6293@gmail.com

MR. Shivtej Suhas Patil

Ashokrao Mane Group Of institution,vathar
Department of Computer Science and
Engineering
shivtejpatil8412@gmail.com

MR. Prashant Pandurang Vadd

Ashokrao Mane Group Of institution,vathar
Department of Computer Science and
Engineering
prashantvadd4670@gmail.com

MR. Jufen Farooq Torgal

Ashokrao Mane Group Of institution,vathar
Department of Computer Science
and Engineering
Jufentorgal2004@gmail.com

MR. Sushil Pravin Godbole

Ashokrao Mane Group Of institution,vathar
Department of Computer Science and
Engineering
sushilgodbole67@gmail.com

Mrs.Pooja Amrish Patil

Assitant Professor at Ashokrao Mane Group
Of institution,vathar Department of
Computer Science and Engineering
papatil@amgoi.edu.in