International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Autism Spectrum Disorder Pre-Diagnosis

Jeyalakshmi.S[1], Adithya.D[2], Dhilipkumar.B[3], Subhashini.V[4]

[1]Associate Professor, Department of Artificial Intelligence and Data Science, SRM Valliammai Engineering College, Kattankulathur.
[2,3,4]Students, Department of Artificial Intelligence and Data Science, SRM Valliammai Engineering College, Kattankulathur.S

ABSTRACT
Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that affects communication, behavior, and social interactions. Early and accurate detection of ASD is crucial for timely intervention and support. This project presents a machine learning-based approach to detect ASD using behavioral and physiological data. The dataset used in this study includes key attributes related to ASD symptoms, and various preprocessing techniques were applied to enhance data quality. We implemented multiple machine learning models, to analyze patterns indicative of ASD. Feature selection techniques were employed to identify the most significant predictors of ASD, improving model performance. The results demonstrate that our optimized model achieves high accuracy in distinguishing individuals with ASD from neurotypical individuals. Comparative analysis with existing methods highlights the efficiency and robustness of our approach. This project contributes to the field of medical diagnostics by providing an automated, data-driven solution for ASD detection. The findings emphasize the potential of machine learning in improving diagnostic processes, reducing manual effort, and aiding healthcare professionals in early identification and intervention. Future work may involve expanding the dataset, integrating multimodal data sources such as genetic and imaging data, and enhancing model interpretability to ensure greater clinical relevance. Our research highlights the potential of artificial intelligence in the medical field, particularly in assisting early ASD diagnosis. By leveraging data-driven approaches, we contribute to the growing efforts in utilizing technology for better healthcare outcomes.

KEYWORDS: Autism Spectrum Disorder (ASD), Machine Learning, Early Detection, Classification Models, Feature Selection, Medical Diagnostics.

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that affects communication, social interaction, and behavior. It manifests in early childhood and continues throughout life, with varying degrees of severity among individuals. Traditional ASD diagnosis relies on clinical assessments and behavioral evaluations conducted by specialists using tools such as the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R). However, these methods are often time-consuming, subjective, and dependent on expert availability, leading to delays in early intervention. With advancements in artificial intelligence, machine learning (ML) has emerged as a promising solution for automating and improving the accuracy of ASD detection.

Machine learning models analyze large datasets containing behavioral, physiological, genetic, and neuroimaging data to identify patterns associated with ASD symptoms. The models enhance early screening by identifying key features such as facial expressions, eye-gaze patterns, speech characteristics, and behavioral responses. Feature selection and preprocessing techniques further improve model efficiency, ensuring higher accuracy in distinguishing ASD from neurotypical individuals.

The application of ML in ASD detection offers significant advantages, including early diagnosis, reduced manual effort, cost-effective screening, and scalability. By automating the diagnostic process, ML minimizes human error and increases accessibility to ASD screening in remote and underdeveloped regions. By leveraging machine learning, this research contributes to the growing efforts in utilizing AI for medical diagnostics, offering a data-driven approach to ASD detection that can aid healthcare professionals in early intervention and treatment planning.

2. RELATED WORK

[1] A study by Hasan et al. (2023) proposed a machine learning framework for the early detection of ASD, utilizing ensemble models such as Random Forest (RF), AdaBoost, and k-Nearest Neighbors (k-NN). The research highlighted the effectiveness of feature selection techniques in improving classification accuracy and reducing computational complexity.

[2] Ahmed et al. (2022) introduced an eye-tracking-based approach for ASD detection using Convolutional Neural Networks (CNNs) and Random Forest classifiers. Their model analyzed gaze patterns and eye movement behaviors, demonstrating high accuracy in distinguishing individuals with ASD from neurotypical controls. However, challenges included dataset limitations and variability in eye-tracking data.

[3] Sukumaran and Govardhanan (2021) explored a voice-based ASD prediction model using Mel-Frequency Cepstral Coefficients (MFCCs) and Gaussian Mixture Models (GMMs). Their research focused on detecting emotional and speech abnormalities in children with ASD. The findings indicated that vocal tone and pitch variations could serve as reliable indicators of ASD, though dataset diversity remained a concern.

[4] Hossain et al. (2021) investigated the impact of feature selection methods such as Recursive Feature Elimination (RFE) and Correlation-Based Feature Selection (CFS) in ASD classification. The study demonstrated that selecting key behavioral attributes improved model accuracy and interpretability while reducing overfitting risks.

[5] Bala et al. (2022) emphasized the importance of Explainable AI (XAI) techniques in ASD diagnosis, ensuring transparency and trustworthiness in ML models. The research proposed a hybrid approach combining deep learning with traditional ML techniques, making AI-driven ASD detection more interpretable for healthcare professionals.

[6] A study by Mahedy et al. (2023) evaluated ensemble learning methods, comparing Support Vector Machines (SVM), Decision Trees (DT), and Logistic Regression for ASD detection. Their findings suggested that ensemble models outperformed individual classifiers by enhancing prediction robustness and reducing false positives.

[7] Crippa et al. (2015) developed a deep learning-based ASD classification model that analyzed motor behavior and facial expressions. The study utilized Convolutional Neural Networks (CNNs) to process movement data, achieving high accuracy in distinguishing ASD individuals from neurotypical ones. However, the model required large datasets and computational resources for effective training.

[8] Pietrucci et al. (2022) explored genetic and neuroimaging data as predictive markers for ASD diagnosis. The study applied machine learning algorithms such as Random Forest (RF) and Support Vector Machines (SVM) to analyze genetic variations and brain imaging patterns. This multimodal approach improved diagnostic precision but faced challenges related to data availability and ethical concerns in genetic testing.

[9] Dawson et al. (2012) proposed a behavioral analysis model that examined social interaction patterns in children with ASD. The study leveraged video-based machine learning techniques to track eye contact, facial expressions, and response delays during interactions. Results demonstrated that AI-driven behavioral tracking could serve as an early screening tool, though real-world applications required further validation.
[10] Subah et al. (2021) developed a mobile-based autism screening tool that utilized machine learning models to analyze questionnaire responses and behavioral patterns. The study implemented Natural Language Processing (NLP) and Decision Trees to evaluate parent-reported symptoms and predict ASD likelihood. While the approach increased accessibility to ASD screening, challenges included response bias and limited real-time behavioral assessments.

3. PROPOSED SYSTEM

The proposed system aims to enhance the early detection of Autism Spectrum Disorder (ASD) by integrating a machine learning model with a web-based user interface. The system collects user responses to the AQ-10 questionnaire, along with demographic and medical history factors such as age, gender, ethnicity, jaundice history, and relation to the individual being tested. The frontend consists of a simple and interactive form that transmits data to the backend via a Flask API. The backend processes the input, applies necessary data preprocessing techniques such as feature encoding and normalization, and feeds it into a trained Random Forest Classifier for ASD prediction. This model is chosen for its high accuracy, robustness, and ability to handle non-linear relationships between features. The system outputs an ASD likelihood score, which is displayed on a results page with recommendations for further clinical evaluation if necessary. By leveraging both behavioral and personal factors, the proposed approach provides a more comprehensive assessment compared to AQ-10 alone, ensuring better reliability and aiding in early diagnosis.

4. ARCHITECTURE DIAGRAM

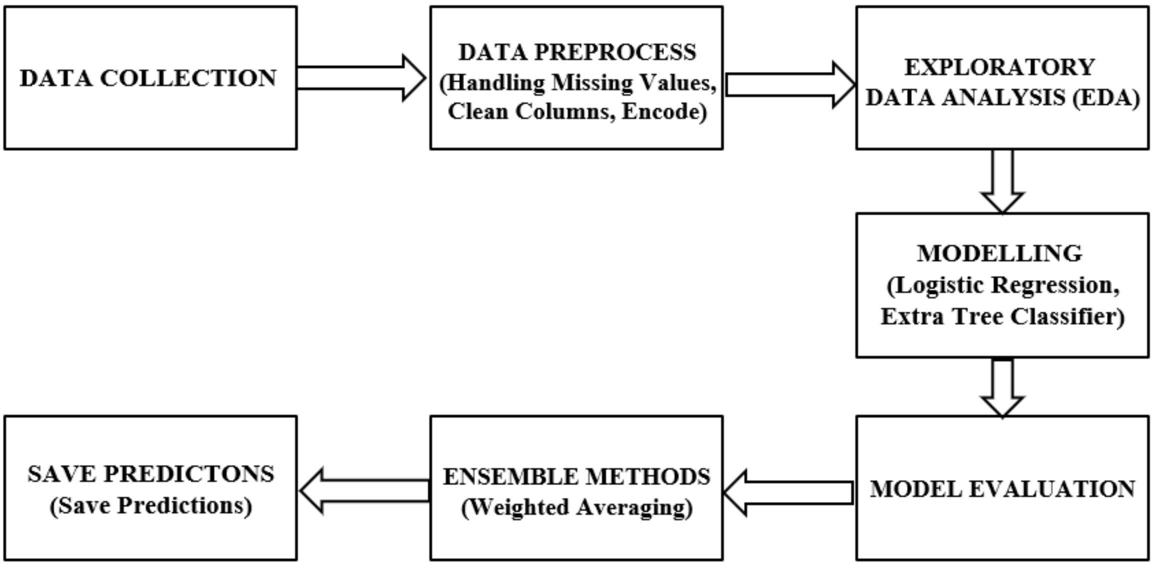

Fig 4.1 System Architecture

5. MODULE DESCRITION

5.1 LOAD DATA

This module is responsible for loading the dataset from a CSV file. The dataset is read into a Pandas Data-

Frame, and unnecessary columns such as 'age_desc', 'used_app_before', and 'austim' are removed. Additionally, missing values are handled by dropping rows with null values. This step ensures that the dataset is clean before moving to further processing.

5.2 DATA PREPROCESSING

In this step, categorical variables (text-based data) are converted into numerical values using Label Encoding, making them suitable for machine learning models. The target column (which indicates whether a person has Autism Spectrum Disorder) is identified and separated from the feature variables. Then, StandardScaler is applied to normalize numerical features so that all values are on a similar scale. Finally, the dataset is split into training (80%) and testing (20%) sets to train and evaluate the machine learning model effectively.

5.3 FEATURE SCALING

Feature selection helps identify the most relevant features from the dataset that contribute to autism prediction. A Random Forest Classifier is trained on the dataset, and the feature importance scores are extracted. These scores help rank the features based on their significance in making predictions. By selecting only the most important features, we can improve the model's efficiency and accuracy while reducing computational complexity.

5.4 MODEL SELECTION AND VALIDATION

Different machine learning models, such as Logistic Regression, Random Forest Classifier, and Support Vector Classifier, are evaluated to identify the one that best fits the dataset. Each model is tested based on its performance metrics, and comparisons are made to determine the most accurate and reliable option. The model that demonstrates the highest accuracy and efficiency is then selected and utilized for making predictions in the autism detection system.

5.5 UI CREATION AND INTEGRATION

The user interface is designed according to the planned structure, ensuring a seamless experience. Flask is utilized to run the UI as a web application, enabling efficient deployment and interaction. Additionally, Flask serves as the bridge between the frontend and backend, facilitating smooth integration and data exchange. With these implementations, the project is successfully completed, achieving its intended functionality.

6. RESULT AND DISCUSSION

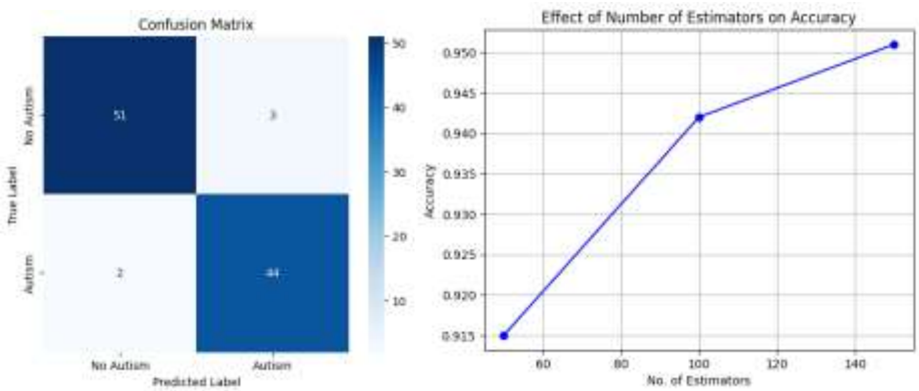

Fig 6.1 Performance Analysis

7. EXPECTED OUTPUT

Fig 7.1 Home Page

Fig 7.1 About Autism

Fig 7.3 Prediction Page 1

Fig 7.4 Prediction Page 2

Fig 7.5 Prediction Page 3

Fig 7.6 Result Page

8. CONCLUSION

The use of machine learning for Autism Spectrum Disorder (ASD) detection offers a promising approach to early diagnosis and intervention. By analyzing behavioral and physiological data, this study demonstrates how AI-driven models can identify ASD patterns with high accuracy. The application of preprocessing techniques, feature selection, and classification models ensures improved performance, making the detection process more efficient and reliable. This automated approach reduces reliance on traditional manual screening methods, which are often time-consuming and subject to human errors.

9. FUTURE SCOPE

The future scope of this project involves enhancing the accuracy and reliability of Autism Spectrum Disorder (ASD) detection through advanced machine learning techniques. Incorporating multimodal data sources such as genetic information, neuroimaging, and speech analysis can provide a more comprehensive assessment of ASD traits. Additionally, expanding the dataset with diverse demographic information will improve the model's generalizability, making it more suitable for real-world clinical applications. Deep learning architectures like transformers and convolutional neural networks (CNNs) can further enhance feature extraction and classification, leading to more precise and efficient ASD detection.

Another key direction is integrating the model into real-time diagnostic systems through mobile applications or cloud-based platforms, enabling early ASD screening in remote areas. Explainable AI (XAI) techniques can be employed to improve model interpretability, ensuring transparency in decision-making for healthcare professionals. Ethical considerations, such as data privacy and bias mitigation, should also be prioritized to ensure responsible AI deployment in medical diagnostics. Future collaborations with healthcare institutions can refine the model for clinical use, making it a valuable tool for early intervention and better treatment planning for individuals with ASD.

10. REFERENCES

1. Ahmed, E. M. Senan, T. H. Rassem, M. A. H. Ali, H. S. A. Shatnawi, S. M. Alwazer, and M. Alshahrani, ‘‘Eye tracking-based diagnosis andearly detection of autism spectrum disorder using machine learning and deep learning techniques,’’ Electronics, vol. 11, no. 4, p. 530, Feb. 2022
2. Crippa, C. Salvatore, P. Perego, S. Forti, M. Nobile, M. Molteni, and I. Castiglioni, ‘‘Use of machine learning to identify children with autism and their motor abnormalities,’’ J. Autism Develop. Disorders, vol. 45, no. 7, pp. 2146–2156, 2015
3. D. Pietrucci, A. Teofani, M. Milanesi, B. Fosso, L. Putignani, F. Messina, G. Pesole, A. Desideri, and G. Chillemi, ‘‘Machine learning data analysis highlights the role of parasutterella and alloprevotella in autism spectrum disorders,’’ Biomedicines, vol. 10, no. 8, p. 2028, Aug. 2022.
4. F. Z. Subah, K. Deb, P. K. Dhar, and T. Koshiba, ‘‘A deep learning approach to predict autism spectrum disorder using multisite restingstate fMRI,’’ Appl. Sci., vol. 11, no. 8, p. 3636, Apr. 2021.
5. H. Chahkandi Nejad, O. Khayat, and J. Razjouyan, ‘‘Software development of an intelligent spirography test system for neurological disorder detection and quantification,’’ J. Intell. Fuzzy Syst., vol. 28, no. 5, pp. 2149–2157, Jun. 2015.
6. J. Amudha and H. Nandakumar, ‘‘A fuzzy based eye gaze point estimation approach to study the task behavior in autism spectrum disorder,’’ J. Intell. Fuzzy Syst., vol. 35, no. 2, pp. 1459–1469, Aug. 2018.
7. M. Bala, M. H. Ali, M. S. Satu, K. F. Hasan, and M. A. Moni, ‘‘Efficient machine learning models for early stage detection of autism spectrum disorder,’’ Algorithms, vol. 15, no. 5, p. 166, May 2022.
8. M. D. Hossain, M. A. Kabir, A. Anwar, and M. Z. Islam, ‘‘Detecting autism spectrum disorder using machine learning techniques,’’ Health Inf. Sci. Syst., vol. 9, no. 1, pp. 1–13, Dec. 2021.
9. P. Sukumaran and K. Govardhanan, ‘‘Towards voice based prediction and analysis of emotions in ASD children,’’ J. Intell. Fuzzy Syst., vol. 41, no. 5, pp. 5317–5326, 2021.
10. S. M. Mahedy Hasan, M. H. R. Khandoker, M. A. Hossain, and M. Z. Uddin, "Machine Learning Framework for Early-Stage Detection of Autism Spectrum Disorders," IEEE Access, vol. 9, pp. 132494-132505, 2021.