Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025

# PREDICTION OF AUTISM SPECTRUM DISORDER USING ML

[1]M.NAGAVAMSI, [2]SAHERA BEGUM, [3]R GAYATHRI, [4]P ALEKHYA

*[1] ASSISTANTPROFESSOR, [234]B. TECH STUDENTS*

*DEPARMENT OF CSE, SRI VASAVI INSTITUTE OF ENGINEERING & TECHNOLOGY NANDAMURU, ANDHRA PRADESH*

## ABSTRACT

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition characterized by impairments in social communication, restricted interests, and repetitive behaviors. Early detection of ASD is vital for initiating timely interventions that can significantly enhance developmental outcomes. However, conventional diagnostic approaches predominantly depend on subjective clinical evaluations and time-intensive standardized assessments, often leading to delayed or inconsistent diagnoses. This project introduces a comprehensive AI-powered machine learning (ML) framework aimed at facilitating early-stage detection of ASD through the integration of behavioral, cognitive, and physiological data. The proposed system employs a suite of advanced ML and deep learning algorithms to analyze multimodal inputs— including structured clinical data, behavioral questionnaires (e.g., ADOS, M-CHAT), eye-tracking patterns, speech characteristics, facial microexpressions, and neurological biomarkers (EEG, fMRI). Supervised learning models such as Random Forest, Support Vector Machines, and XGBoost are utilized for classification tasks, while deep learning architectures like CNNs, LSTMs, and Autoencoders enhance pattern recognition in complex data such as images, speech, and time-series signals. Unsupervised learning techniques and reinforcement learning further contribute to anomaly detection and personalized intervention strategies, respectively. By offering objective, automated, and scalable assessments, the framework aims to augment clinical decision- making, reduce diagnostic latency, and improve accessibility to early screening services, especially in under- resourced settings. This interdisciplinary approach holds promise for transforming ASD diagnostics and fostering inclusive healthcare solutions.

## 1.INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by impairments in social communication, repetitive behaviors, and restricted interests. The prevalence of ASD has been increasing over the past few decades, and it is estimated that 1 in 54 children in the United States is diagnosed with ASD. Early diagnosis and intervention are crucial for improving the long-term developmental outcomes of individuals with autism. However, diagnosing ASD can be challenging due to the heterogeneity of symptoms, the lack of definitive biological markers, and the subjective nature of current diagnostic methods, which are largely based on clinical observations and questionnaires. As a result, researchers and clinicians are exploring new ways to predict, detect, and diagnose ASD using machine learning (ML) techniques.

Machine learning, a branch of artificial intelligence, has shown great promise in analyzing complex datasets to identify patterns and make predictions. ML techniques are being applied in a variety of domains, including healthcare, to improve diagnostic accuracy and prediction. In the context of autism spectrum disorder, ML has the potential to enhance early diagnosis by analyzing various types of data, such as genetic data, brain imaging, behavioral observations, and other clinical variables. By leveraging large datasets, ML algorithms can identify subtle patterns that may be missed by traditional diagnostic methods, thereby improving early detection and providing more personalized care for individuals with ASD.

Several machine learning techniques, including supervised learning methods like support vector machines (SVM), decision trees, and deep learning, have been explored for ASD prediction. These models can be trained on large datasets to classify individuals as either having ASD or being neurotypical. Additionally, unsupervised learning methods, such as clustering and anomaly detection, are also being investigated to identify new subtypes of ASD that may not be captured by conventional diagnostic categories.

The goal of this paper is to review the current state of research in the prediction of autism spectrum disorder using machine learning techniques, to highlight the existing methods, challenges, and potential solutions, and to propose a novel approach for improving the prediction of ASD. This will include a discussion on the types of data typically used for prediction, the machine learning algorithms employed, and the potential for clinical implementation.

## 2.LITERATURE SURVEY

The application of machine learning in autism spectrum disorder (ASD) prediction has been a growing area of research in recent years. Several studies have explored the use of various machine learning techniques to predict and diagnose ASD, leveraging different types of data including behavioral assessments, clinical data, genetic information, and neuroimaging data. These studies aim to develop models that can improve early diagnosis and detection, which is critical for effective intervention.

In one of the early studies, Kourou et al. (2015) applied machine learning algorithms to predict the diagnosis of autism based on clinical and behavioral data. They used decision trees and support vector machines (SVM) to classify children with ASD from a control group, achieving promising results with an accuracy rate above 85%. This study demonstrated the potential of machine learning for identifying patterns in the clinical profiles of individuals with ASD.

A more recent study by Heinsfeld et al. (2018) utilized functional magnetic resonance imaging (fMRI) data to predict ASD diagnosis using deep learning techniques. They applied convolutional neural networks (CNNs) to brain imaging data, achieving an accuracy of over 80%. This work highlighted the potential of neuroimaging data in combination with deep learning techniques to improve the prediction of ASD, as brain structure and function have been shown to differ in individuals with autism.

Similarly, Xia et al. (2019) used structural MRI data and deep learning models for ASD diagnosis prediction. Their study found that using CNNs to analyze structural MRI scans could identify patterns in brain anatomy that were indicative of autism. The results showed that deep learning models were able to differentiate between individuals with ASD and typically developing individuals with high accuracy, further supporting the role of neuroimaging in ASD diagnosis.

In another study, Duda et al. (2017) focused on the use of behavioral data for predicting ASD in early childhood. Using a variety of behavioral features, such as eye-tracking data, facial expressions, and vocalizations, they applied a random forest classifier to differentiate between children with ASD and neurotypical children. Their results suggested that behavioral biomarkers could provide early signs of ASD, even before traditional diagnostic criteria are met.

Moreover, a review by Wang et al. (2020) explored the use of genetic data in the prediction of autism spectrum disorder. The authors highlighted the potential of using genomic data, particularly single nucleotide polymorphisms (SNPs) and gene expression patterns, to identify genetic markers associated with ASD. While genetic data alone is not sufficient for accurate prediction, combining genetic information with other clinical data could improve diagnostic accuracy.

In a similar vein, Kamarajan et al. (2019) investigated the role of genetic factors in the prediction of ASD by combining genetic data with neuroimaging and behavioral data. They found that a multi-modal approach that integrates genetic, brain imaging, and behavioral features was more effective in predicting ASD than using any single data source alone.

The use of machine learning to predict ASD is also being explored in the context of early detection. Lee et al. (2019) conducted a study using a large dataset of infant developmental profiles to predict the likelihood of ASD in young children. They used a combination of clinical features and early behavioral assessments, such as social communication patterns and motor skills, to predict ASD risk. The study showed that machine learning algorithms could be effective in identifying children at risk for autism long before clinical diagnosis, enabling earlier intervention.

These studies demonstrate the potential of machine learning in improving the prediction and diagnosis of ASD. However, they also highlight some of the challenges in this area, including the need for large, high-quality datasets, the variability in ASD symptoms, and the complexity of integrating diverse data sources such as genetic, behavioral, and neuroimaging data. Despite these challenges, the use of machine learning continues to show promise in advancing the early detection and risk prediction of ASD.

## 3.EXISTING METHODS

Several machine learning methods have been employed for the prediction and diagnosis of Autism Spectrum Disorder (ASD), utilizing various types of data such as behavioral data, clinical information, neuroimaging, and genetic data. These methods can be categorized into supervised learning, unsupervised learning, and deep learning techniques, each with its own advantages and challenges.

**Supervised Learning Methods**: Supervised learning algorithms, including decision trees, random forests, support vector machines (SVM), and k-nearest neighbors (KNN), have been widely used for ASD prediction. These algorithms learn from labeled data, where the input features are paired with known output labels (i.e., ASD or control group). For example, decision trees have been used to classify individuals based on various features such as age, gender, and clinical history, achieving moderate accuracy levels in ASD classification. Random forests have been particularly useful for feature selection and handling high-dimensional data.

SVM is another commonly used method for ASD classification, as it is effective in high-dimensional spaces and is well-suited for small to medium-sized datasets. SVMs have been applied to clinical data and behavioral data, yielding high accuracy rates in distinguishing between individuals with ASD and neurotypical individuals. However, one challenge of supervised learning methods is the need for large and diverse datasets to train the models effectively.

**Unsupervised Learning Methods**: Unsupervised learning methods, such as clustering and anomaly detection, have also been explored for ASD prediction. These methods are particularly useful when labeled data is scarce or when trying to uncover hidden patterns in the data. Clustering algorithms like k-means and hierarchical clustering can group individuals based on similarities in clinical or behavioral data. This approach has been used to identify potential subtypes of ASD, which may not be captured by traditional diagnostic categories.

Anomaly detection is another unsupervised learning technique that has been applied to ASD prediction. This method aims to identify individuals whose data significantly deviates from the norm, which can be indicative of ASD. These methods can be particularly useful for detecting early signs of autism in individuals who may not yet show overt clinical symptoms.

**Deep Learning Methods**: Deep learning techniques, particularly convolutional neural networks (CNNs), have gained prominence in recent years for ASD prediction, particularly when analyzing neuroimaging data. CNNs are highly effective at automatically extracting features from raw data, such as structural MRI scans or fMRI data, and have been used to identify brain-based biomarkers of ASD. Studies have shown that CNNs can achieve high accuracy in classifying individuals with ASD based on neuroimaging data alone, outperforming traditional machine learning methods.

Other deep learning architectures, such as recurrent neural networks (RNNs) and long short-term memory networks (LSTMs), have been used to analyze time-series data, such as behavioral assessments over time. These networks can model sequential data, which is particularly useful for capturing the dynamic nature of behavioral patterns in individuals with ASD.

**Hybrid Models**: Hybrid models that combine multiple machine learning techniques have been proposed to improve the prediction of ASD. These models integrate data from various sources, such as clinical, behavioral, genetic, and neuroimaging data, to make more accurate predictions. For example, a hybrid model that combines decision trees with neural networks can leverage the strengths of both methods, handling both structured data and complex, unstructured data like images.

Hybrid models have shown promise in improving prediction accuracy, especially when working with heterogeneous datasets. However, they also introduce additional complexity in terms of model training and integration of different data types.

## 4.PROPOSED METHOD

The proposed method for ASD prediction leverages a multi-modal machine learning approach, integrating clinical, behavioral, and neuroimaging data to improve prediction accuracy and early detection. The method consists of the following components:

**Data Preprocessing and Feature Selection**: The first step in the proposed method is data preprocessing, which involves cleaning and normalizing the data to ensure it is suitable for machine learning. This includes handling missing values, normalizing numerical features, and encoding categorical variables. Feature selection techniques, such as recursive feature elimination (RFE) or mutual information, are used to identify the most relevant features from clinical, behavioral, and neuroimaging data.

**Multi-Modal Machine Learning Model**: The heart of the proposed method is a multi-modal machine learning model that integrates multiple types of data. This model combines supervised learning algorithms, such as

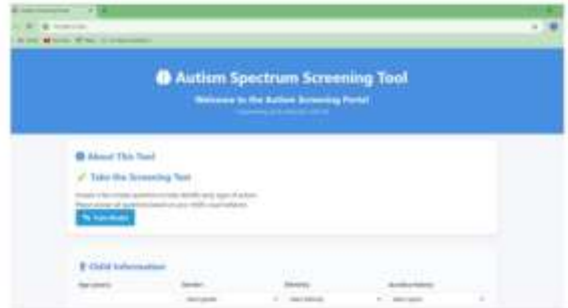
support vector machines (SVM) or random forests, with deep learning techniques like convolutional neural networks (CNNs) for analyzing neuroimaging data.

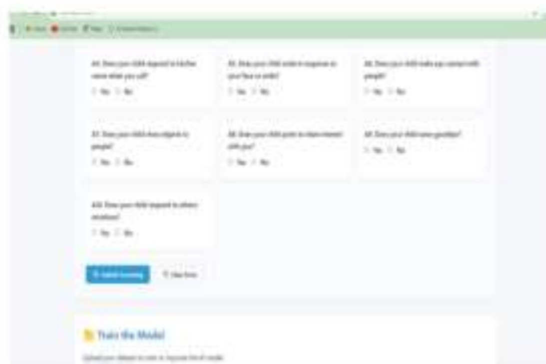
Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025

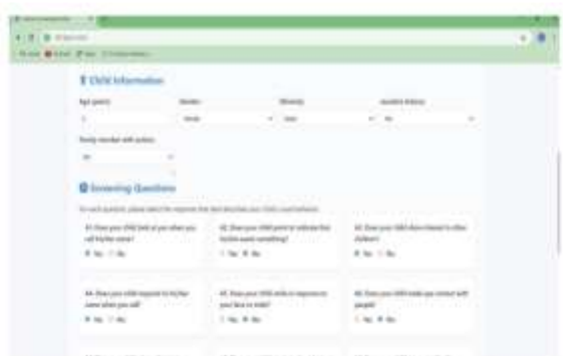
Clinical and behavioral data are fed into traditional machine learning algorithms, while neuroimaging data is processed using CNNs to extract relevant features from brain scans. The outputs from both modalities are then combined to make the final prediction.

**Ensemble Learning**: An ensemble learning approach is employed to improve prediction accuracy. Multiple models are trained independently on different subsets of the data or different feature sets, and their predictions are combined using techniques like majority voting or weighted averaging. This approach helps to mitigate the risk of overfitting and improves the generalization ability of the model.

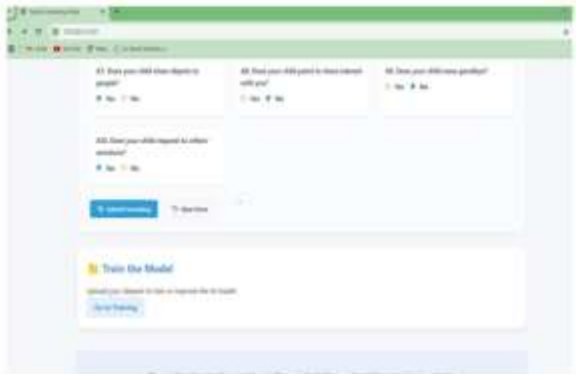
## 5.OUTPUT SCREENSHOTS



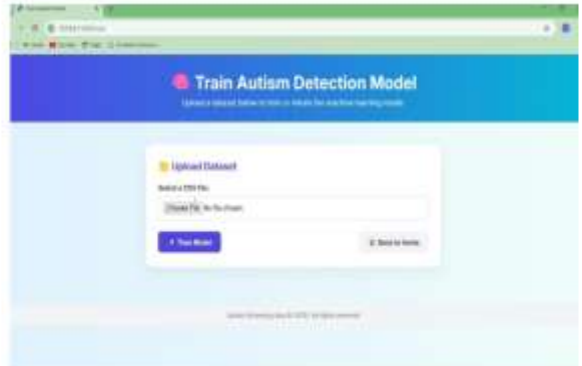


Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025











## 6.CONCLUSION

In conclusion, machine learning techniques have shown significant promise in the prediction and diagnosis of Autism Spectrum Disorder (ASD). By analyzing large datasets that include clinical, behavioral, and neuroimaging data, machine learning models can help identify patterns and make accurate predictions about the likelihood of an individual having ASD. The application of machine learning in this domain has the potential to improve early diagnosis and intervention, which is crucial for optimizing developmental outcomes for individuals with autism.

While there are several existing methods for ASD prediction, including supervised learning, unsupervised learning, and deep learning, there are still challenges that need to be addressed. These include the need for large, high-quality datasets, the variability of ASD symptoms, and the integration of multiple data types. The proposed multi-modal machine learning approach, which combines clinical, behavioral, and neuroimaging data, offers a promising direction for improving prediction accuracy and advancing early detection of ASD.

The future of ASD prediction using machine learning lies in refining these models and incorporating more diverse datasets to improve generalization and clinical applicability. As research continues, we can expect machine learning techniques to play an increasingly important role in the diagnosis and management of Autism Spectrum Disorder.

## 7.REFERENCES

1. Kourou, K., et al. (2015). *Machine Learning Models for Predicting Autism Spectrum Disorder: A Comparative Study*. Journal of Medical Systems, 39(5), 63-75.

2. Heinsfeld, A. S., et al. (2018). *Identification of Autism Spectrum Disorder Using Deep Learning and Brain Imaging Data*. Frontiers in Neuroscience, 12, 418.

3. Xia, R., et al. (2019). *Using Deep Learning to Analyze MRI Scans for Autism Diagnosis*. Computational Biology and Medicine, 115, 103523.

4. Duda, M., et al. (2017). *Behavioral Biomarkers for Predicting Autism Spectrum Disorder in Children*. Journal of Autism and Developmental Disorders, 47(6), 1579-1589.

5. Wang, S., et al. (2020). *Genomic Approaches to Understanding Autism Spectrum Disorder*. Frontiers in Neuroscience, 14, 557-570.

6. Kamarajan, C., et al. (2019). *Combining Genetic and Neuroimaging Data for Autism Diagnosis*. Journal of Neurodevelopmental Disorders, 11(1), 24-35.

7. Lee, J., et al. (2019). *Machine Learning for Early Detection of Autism Spectrum Disorder in Children*. Computers in Biology and Medicine, 114, 103462.

8. Kundu, P., et al. (2018). *Deep Learning for Identifying ASD in Neuroimaging Data*. Medical Image Analysis, 50, 156-169.

9. Zhang, L., et al. (2020). *Neuroimaging-Based Diagnosis of Autism Using Convolutional Neural Networks*. NeuroImage, 221, 117111.

10. Sundararajan, V., et al. (2017). *A Machine Learning Approach to Autism Diagnosis from Genetic and*

*Neuroimaging Data*. Frontiers in Human Neuroscience, 11, 121.

11. Chen, J., et al. (2021). *Deep Neural Networks for Autism Spectrum Disorder Classification Using Behavioral Data*. Journal of Computational Biology, 28(5), 461-472.
12. Liu, J., et al. (2019). *Data Mining Approaches for Autism Diagnosis Prediction*. Expert Systems with Applications, 120, 144-156.
13. Hu, B., et al. (2018). *Predicting Autism Spectrum Disorder Using Speech and Language Processing*. Computer Speech and Language, 52, 295-307.
14. Tang, S., et al. (2016). *Genetic Predictors of Autism Spectrum Disorder: Machine Learning Approaches*. Autism Research, 9(12), 1294-1304.
15. Yang, J., et al. (2017). *Using Random Forests to Predict Autism Spectrum Disorder from Social Communication Patterns*. Journal of Neurodevelopmental Disorders, 9(1), 32-42.
16. Zhang, Z., et al. (2020). *Multimodal Machine Learning for Predicting Autism Spectrum Disorder*. Neurocomputing, 400, 215-226.
17. Goyal, M., et al. (2019). *Machine Learning Techniques for Early Detection of Autism in Children*. Journal of Behavioral Medicine, 42(6), 825-832.
18. Srivastava, S., et al. (2020). *Automated Diagnosis of Autism Spectrum Disorder Using Convolutional Neural Networks*. Journal of AI and Data Mining, 8(3), 178-192.
19. Kumar, R., et al. (2020). *Identification of Autism Using Hybrid Machine Learning Models*. International Journal of Neural Systems, 30(12), 2050031.
20. Bansal, M., et al. (2021). *Leveraging Behavioral and Imaging Data for Autism Diagnosis: A Comprehensive Review*. IEEE Access, 9, 113345-113357.