



## LEVERAGING MACHINE LEARNING FOR A COMPREHENSIVE ANALYSIS OF AUTISM SPECTRUM DISORDER: A SYNTHESIS OF SCHOLARLY INSIGHTS

G. Divya\* & Dr. V. Maniraj\*\*

\* Research Scholar, Department of Computer Science, A.V.V.M Sri Pushpam College (Autonomous), Poondi, Thanjavur, Affiliated to Bharathidasan University, Thiruchirappalli, Tamil Nadu

\*\* Associate Professor & Research Supervisor, Head of the Department, Department of Computer Science, A.V.V.M Sri Pushpam College (Autonomous), Poondi, Thanjavur, Affiliated to Bharathidasan University, Thiruchirappalli, Tamil Nadu

**Cite This Article:** G. Divya & Dr. V. Maniraj, "Leveraging Machine Learning for a Comprehensive Analysis of Autism Spectrum Disorder: A Synthesis of Scholarly Insights", International Journal of Computational Research and Development, Volume 9, Issue 2, July - December, Page Number 91-94, 2024.

**Copy Right:** © DV Publication, 2024 (All Rights Reserved). This is an Open Access Article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited.

### Abstract:

Autism Spectrum Disorder (ASD) is a complex developmental condition that is defined by difficulties in social interaction, repetitive behaviors, and increased sensitivity to sensory stimuli. Early and accurate diagnosis is crucial for ensuring effective intervention and support. Recent advancements in machine learning have demonstrated promising potential in improving ASD diagnosis using a variety of techniques. This paper reviews recent research on how machine learning is applied to ASD diagnosis, emphasizing the effectiveness of different classification models, feature selection approaches, and their impact on diagnostic accuracy. Key methods include Multilayer Perceptron (MLP) classifiers, which have achieved up to 100% accuracy, along with other advanced AI models that have shown diagnostic accuracies of 98.17%. Additional techniques, such as the Weighted C4.5 Algorithm (WCBA) and Decision Tree Classifiers, have proven successful in identifying critical features and enhancing early detection. Combining feature scaling with various machine learning algorithms has further improved classification performance across diverse age groups. These developments underscore the potential of machine learning to enhance ASD diagnosis, leading to more precise and earlier identification, which is essential for effective intervention and management.

**Key Words:** ASD, MLP, Neurodevelopment, Diagnostic, WCBA, C4.5 Algorithm.

### 1. Introduction:

Autism spectrum disorder (ASD) is a complex neuro developmental condition that influences how individuals communicate, socialize, and perceive sensory input. The disorder varies widely in its symptoms and severity, making early diagnosis and intervention crucial for enhancing development and overall well-being. ASD presents a broad spectrum of experiences and challenges. It is typically diagnosed in early childhood and persists throughout life, influencing various aspects of daily functioning.

### 2. Key Features:

- **Difficulties in Social Communication:** People with ASD frequently struggle with understanding and engaging in social interactions. They may struggle with nonverbal communication, such as eye contact, body language, and facial expressions, making it challenging to form and maintain relationships.
- **Repetitive Patterns of Behavior and Narrowed Interests:** Common behaviors include repetitive movements (e.g., hand-flapping, rocking) and a strong focus on specific interests or activities. Individuals might insist on routines and become distressed by changes in their environment or schedule.
- **Sensory Perception Variations:** People with ASD often exhibit either heightened or diminished responses to sensory input. They may be particularly sensitive to lights, sounds, textures, or tastes, or they might engage in sensory-seeking behaviors..
- **Diagnosis and Early Intervention:** ASD is typically diagnosed through a combination of developmental history and behavioral assessments, often beginning around age 2 to 4, though signs can be observed earlier. Prompt recognition and action are key to improving developmental results. Support strategies may involve behavioral techniques, speech and occupational therapy, and individualized educational assistance, all adapted to meet the person's specific needs.

**Causes and Risk Factors:** The underlying reasons for ASD are not completely known, but research indicates that a blend of genetic and environmental elements contributes to its development. Evidence shows that genetic tendencies are a major factor, though environmental influences may also contribute. There is no single known cause, and multiple factors likely interact to influence the development of ASD.

**Impact and Support:** The impact of ASD varies widely among individuals, from those with high-functioning autism who may have minimal support needs to those requiring substantial assistance with daily living activities. Support strategies are tailored to the individual's needs and may involve educational accommodations, therapeutic interventions, and community support.

**Awareness and Advocacy:** Increasing awareness and understanding of ASD are critical for improving the quality of life for individuals with the condition. Advocacy efforts focus on enhancing public knowledge, reducing stigma, and ensuring access to resources and services. As societal attitudes evolve, there is a growing recognition of the diverse abilities and contributions of individuals with ASD.

In summary, Autism Spectrum Disorder is a multifaceted condition that affects various aspects of life. Ongoing research, improved diagnostic practices, and increased public awareness are essential for supporting individuals with ASD and promoting their inclusion and well-being in society.

### 3. A Synthesis of ASD:

Hossain et.al (2021) explored that “Autism Spectrum Disorder (ASD)”, a condition affecting brain development, frequently presents with sensory sensitivities such as unusual responses to auditory, olfactory, or tactile stimuli. While genetic factors are a major contributor to its development, early identification and intervention can significantly improve outcomes. Recently, machine learning techniques have been developed as an additional tool to traditional diagnostic methods, which can be both lengthy and expensive.

This research aims to highlight the most critical features (Hossain, 2021) and streamline the diagnostic process using various classification methods to enhance precision. It reviews ASD datasets from different age ranges-toddlers, children, adolescents, and adults-evaluating advanced classification and feature selection methods to identify the most effective classifier and feature set for each demographic. The results show that the multilayer perceptron (MLP) classifier surpasses other techniques, achieving perfect accuracy with a minimal number of attributes across all datasets. Additionally, we found that the 'relief F' feature selection technique is the most effective for ranking the significant attributes in all four datasets.

Alwidian et.al (2020) Autism Spectrum Disorder (ASD) is a mental health condition that impacts cognitive, linguistic, communicative, and social functions. Recently, data mining techniques have been employed to uncover critical features of ASD and the connections between them. This article examines the use of Association Classification (AC) (Alwidian, 2020) to predict autism diagnosis. To do this, seven prominent algorithms were chosen to evaluate how well the AC technique identifies feature relationships that could facilitate early autism diagnosis, especially in children.

The performance of these algorithms was measured using standard evaluation metrics such as Precision, Accuracy, F-Measure, and Recall. A comparative performance analysis showed that the WCBA algorithm generally performed the best, achieving 97% accuracy in most cases, although many of the algorithms demonstrated high accuracy in this context.

Gupta et.al (2024) analyzed that Autism Spectrum Disorder (ASD), commonly referred to as autism, is a developmental disorder characterized by persistent challenges in speech, communication, repetitive behaviors, nonverbal communication, and focus. Despite the absence of a complete cure, early diagnosis is critical for managing the condition. In this study, we propose a methodology aimed at enhancing the early detection of ASD. Utilizing a dataset comprising 1,986 patients and 27 features sourced from Kaggle, and developed an artificial intelligence model to predict ASD.

The model's performance was evaluated against several established classifiers (Gupta, 2024), including “K-Nearest Neighbors (KNN)”, “Random Forest (RF)”, “Decision Tree (DT)”, “Gradient Boosting Classifier (GBC)”, “Support Vector Machine (SVM)”, and “Naïve Bayes (NB)”. Our machine learning model achieved an accuracy rate of 98.17%, surpassing the accuracy rates of other classifiers, which ranged from 68% to 93%.

These results demonstrate that our proposed model offers a highly effective approach to ASD prediction, achieving the highest accuracy among the classifiers evaluated. The study's findings confirm the superiority of our model in accurately predicting ASD, making it a valuable tool for early diagnosis.

Mahedy Hasan et.al (2023) “autism spectrum disorder (ASD) “is a neuro developmental disorder that significantly impacts daily life. While complete eradication is challenging, early intervention can mitigate the severity of the condition. This paper proposes an effective framework for evaluating various “Machine Learning (ML)” techniques to enhance early ASD detection. The framework incorporates (S. M. Mahedy Hasan, 2023) Four techniques for standardizing features-“Quantile Transformer, Power Transformer, Normalizer, and Max Abs Scaler”-were implemented across four standard ASD datasets, each corresponding to different age groups: Toddlers, Adolescents, Children, and Adults.. Eight ML algorithms, including Ada Boost, Random Forest, and Support Vector Machine, were used for classification.

The study identified the best-performing ML algorithms and FS techniques for each dataset by comparing classification outcomes using metrics such as accuracy, ROC curve, and F1-score. Notably, Ada Boost achieved the highest accuracy of 99.25% for Toddlers and 97.95% for Children, while Linear Discriminant Analysis excelled with 97.12% accuracy for Adolescents and 99.03% for Adults. Additionally, feature importance was ranked using four Feature Selection Techniques, offering valuable insights for healthcare practitioners in ASD screening. The proposed framework demonstrated superior results compared to existing methods, underscoring the importance of fine-tuning ML models for effective ASD prediction across different age groups.

Khushi Mittal et.al (2024) study explores the application of machine learning techniques, specifically a Decision Tree Classifier, for assessing the risk of autism spectrum disorder (ASD). Autism, (Khushi Mittal, 2024) a complex neuro developmental disorder, presents unique challenges in early detection and intervention. Traditional diagnostic methods often rely on time-intensive and subjective assessments, leading to delayed diagnoses. In this research, a Decision Tree Classifier is utilized to analyze a substantial dataset of behavioral and demographic traits associated with ASD. The algorithm is trained on a diverse dataset that includes individuals diagnosed with ASD and those without, allowing it to identify patterns and correlations within the data.

The primary goal of this research is to evaluate the classifier's effectiveness in accurately predicting the likelihood of autism based on input features. The reliability of the model is ensured through rigorous cross-validation and statistical evaluation methods. The proposed machine learning approach has the potential to enhance early detection processes, facilitating timely intervention and support for individuals at risk of ASD.

Muhammad Shuaib Qureshi and his team (2023) examined Autism Spectrum Disorder (ASD) as a profound, lifelong neuro developmental condition that impacts social communication, cognitive abilities, behaviors, and activities. The symptoms of autism generally become more recognizable and easier to diagnose in children aged two to three years. Contemporary research on ASD often involves prediction models using established machine learning techniques, including Support Vector Machines, Random Forests, Multilayer Perceptrons, Naive Bayes, Convolutional Neural Networks, and Deep Neural Networks.

These models are commonly assessed using metrics such as accuracy, precision, and recall. This study aims to evaluate and compare these machine learning techniques in predicting ASD, focusing on various factors such as application types, simulation approaches, comparison methods, and input data. The goal is to create a unified framework for researchers in the field of autism prediction. The Random Forest method demonstrated the highest performance, with an accuracy rate of 89.23%,

surpassing other traditional machine learning techniques. Additionally, the study provides detailed diagrams of the workflows of the evaluated models, offering insights into their core functions and structures.

**4. Comparative Study of ASD:**

Recent studies on machine learning techniques for diagnosing “autism spectrum disorder (ASD) “highlight significant advancements and varying approaches. Hossain et al. (2021) demonstrated that the Multilayer Perceptron (MLP) classifier achieved 100% accuracy across different age groups by effectively utilizing the 'relief F' feature selection method. Alwidian et al. (2020) found that the Weighted C4.5 Algorithm (WCBA) excelled in predicting ASD with 97% accuracy by leveraging association classification to identify key feature correlations. Gupta et al. (2024) proposed an AI model that outperformed traditional classifiers, reaching 98.17% accuracy in ASD prediction, showcasing its potential for early detection. Mahedy Hasan et al. (2023) highlighted the effectiveness of Ada Boost and Linear Discriminant Analysis in specific age groups, emphasizing the importance of feature scaling methods for improved classification. Khushi Mittal et al. (2024) demonstrated that Decision Tree Classifiers can enhance early detection by analyzing behavioral and demographic traits. Muhammad Shuaib Qureshi et al. (2023) provided a comparative analysis of traditional machine learning algorithms, with Random Forest achieving the highest accuracy at 89.23%. Overall, these studies underscore the potential of machine learning in improving ASD diagnosis, each offering unique insights into classifier performance and feature importance.

**5. Summarized Analysis:**

Here's a comparative summary of the machine learning techniques for diagnosing “autism spectrum disorder (ASD) “based on the studies by the six authors, presented in tabular format:

Author	Year	Focus	Datasets/ Subjects	Techniques Evaluated	Key Findings	Best Performing Method	Accuracy
Hossain et al.	2021	Classification and feature selection	Toddlers, Children, Adolescents, Adults	Multilayer Perceptron (MLP), 'Relief F' Feature Selection	MLP achieved 100% accuracy across all datasets	MLP, 'Relief F' Feature Selection	100% across all datasets
Alwidian et al.	2020	Association Classification	Not specified	Weighted C4.5 Algorithm (WCBA), seven other algorithms	WCBA performed best with 97% accuracy; good accuracy across algorithms	WCBA (Weighted C4.5 Algorithm)	0.97
Gupta et al.	2024	Classification Model Development	1,986 patients, 27 features	KNN, Random Forest (RF), Decision Tree (DT), GBC, SVM, NB	AI model achieved 98.17% accuracy, surpassing other classifiers	AI Model	0.9817
Mahedy Hasan et al.	2023	Feature Scaling and ML Techniques	Toddlers, Children, Adolescents, Adults	Ada Boost, Random Forest, SVM, Linear Discriminant Analysis (LDA), FS Methods	Ada Boost, LDA	Ada Boost and LDA	Ada Boost: 99.25% (Toddlers), LDA: 99.03% (Adults)
Khushi Mittal et al.	2024	Decision Tree Classification	Behavioral and demographic traits	Decision Tree Classifier	Enhanced early detection and pattern recognition	Decision Tree Classifier	Not specified
Muhammad Shuaib Qureshi et al.	2023	Comparison of ML Algorithms	Not specified	SVM, Random Forest, MLP, Naive Bayes, CNN, Deep Neural Networks	Random Forest performed best with 89.23% accuracy	Random Forest	0.8923

Figure 1: The comparative Statement of various Researchers

- Best Performing Methods: The MLP classifier by Hossain et al. and the AI model by Gupta et al. achieved the highest accuracy overall, with 100% and 98.17%, respectively.
- Top Algorithms: The WCBA algorithm by Alwidian et al. and Ada Boost by Mahedy Hasan et al. performed exceptionally well in their respective evaluations, achieving 97% and 99.25% accuracy.
- Feature Importance: Studies highlighted the effectiveness of various feature selection techniques, with 'Relief F' and FS methods showing significant improvements in classification performance.
- Comparative Accuracy: Random Forest, while not the highest in accuracy compared to MLP or AI models, was still notably effective, achieving 89.23% accuracy according to Muhammad Shuaib Qureshi et al.

Figure 1 outlines the different approaches implemented in the research, results, and best-performing techniques across different studies, illustrating the advancements and effectiveness of machine learning in ASD diagnosis.

**6. Conclusion:**

Recent advancements in machine learning for “autism spectrum disorder (ASD) “diagnosis have demonstrated significant progress and varied approaches. Machine learning models, such as Multilayer Perceptron (MLP) and advanced AI systems, have achieved exceptional accuracy rates, with some reaching 100% and 98.17%, respectively, underscoring their effectiveness in ASD diagnosis. Techniques like the Weighted C4.5 Algorithm (WCBA) and Decision Tree Classifiers have also proven effective, with accuracy rates of 97% and notable improvements in early detection through behavioral and demographic

analysis. Furthermore, the exploration of feature scaling methods and various algorithms, including Ada Boost and Linear Discriminant Analysis (LDA), has highlighted their importance in enhancing classification performance for different age groups. Overall, these studies collectively illustrate the robust capabilities of machine learning in improving ASD diagnosis and emphasize the need for continued refinement and innovation in diagnostic methodologies.

#### **7. References:**

1. Abdelwahab, M. M.-K. (2024). Autism spectrum disorder prediction in children using machine learning. *Journal of Disability Research* .
2. Alwidian, J. A. (2020). Predicting autism spectrum disorder using machine learning technique. *International Journal of Recent Technology and Engineering* , 4139-4143.
3. Gupta, A. I. (2024). Autism spectrum disorder detection using machine learning. In *AIP Conference Proceedings*,.
4. Hossain, M. D. (2021). Detecting autism spectrum disorder using machine learning techniques: An experimental analysis on toddler, child, adolescent and adult datasets. *Health Information Science and Systems*, 1-13.
5. Khushi Mittal, K. S. (2024). Applying Machine Learning for Autism Risk Evaluation Using a Decision Tree Classification Technique. *Journées Francophones d'Ingénierie des Connaissances*.
6. Muhammad Shuaib Qureshi, M. B. (2023). Prediction and Analysis of Autism Spectrum Disorder Using Machine Learning Techniques. *Journal of Healthc Eng*.
7. Praveen, P. M. (2024). Identification of “autism spectrum disorder (ASD) “using Feature-based Machine Learning Classification Model. In *2024 2nd International Conference on Sustainable Computing and Smart Systems* .
8. Rajagopalan, S. S. (2024). Machine Learning Prediction of Autism Spectrum Disorder From a Minimal Set of Medical and Background Information. *JAMA Network*.
9. S. M. Mahedy Hasan, M. P. (2023). A Machine Learning Framework for Early-Stage Detection of Autism Spectrum Disorders. *IEEE* .