STAR
Research Journal

Available online at www.starresearchjournal.com (Star International Journal)

COMPUTER SCIENCE

Harnessing Machine Learning: A Comprehensive Review of Research on Autism Spectrum Disorder

[1]G. DIVYA
Research Scholar, Department Of Computer Science,
A.V.V.M Sri Pushpam College (Autonomous), Poondi,Thanjavur (Dt), Affliated to Bharathidasan University,Thiruchirappalli, Tamilnadu (Mail Id-gdivya19.mca@gmail.com)

[2]Dr. V. Maniraj
Associate Professor, Research Supervisor, Head of the Department,
Department of Computer Science, A.V.V.M Sri Pushpam College (Autonomous),
Poondi,Thanjavur(Dt), Affliated To Bharathidasan University,Thiruchirappalli, Tamilnadu,
(Mail Id-manirajv61@gmail.com)

Abstract

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition characterized by challenges in social communication, repetitive behaviors, and heightened sensory sensitivities. Timely and accurate diagnosis is crucial for delivering effective interventions and support. Recent developments in machine learning have demonstrated significant potential for improving the diagnostic process for ASD through a range of innovative techniques.This Paper reviews the recent advancements in machine learning techniques applied to the diagnosis and early detection of Autism Spectrum Disorder (ASD). Traditional diagnostic methods often face challenges related to subjectivity and time consumption, leading to delays in intervention. This paper highlights various machine learning algorithms, including Decision Tree Classifiers, Multilayer Perceptrons, and Weighted C4.5, which have demonstrated high accuracy rates in predicting ASD. Through rigorous analysis of diverse datasets, these techniques reveal critical features that facilitate early diagnosis. The findings underscore the importance of tailored machine learning models in enhancing clinical decision-making and improving outcomes for individuals at risk of ASD. Ultimately, this research advocates for the integration of advanced machine-learning methods into routine clinical practice to enable timely and effective interventions.

Keywords: ASD, MLP, neurodevelopment, diagnostic, WCBA, C4.5 algorithm.

1. Introduction

Autism spectrum disorder (ASD) is a complex neurodevelopmental 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

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 neurodevelopmental 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.

Martinez & Patel, (2024) reported on recent advancements in AI models for ASD diagnosis, achieving impressive accuracy rates of up to 98.17%. The authors noted that these models leverage a combination of deep learning techniques and traditional machine learning algorithms, leading to robust performance across different datasets. Their ongoing research into ensemble methods and hybrid models suggests a promising future for AI applications in healthcare, facilitating timely intervention crucial for improving outcomes for individuals with ASD.

Gupta et.al (2024) analyzed that Autism Spectrum Disorder (ASD), commonly referred to as autism, is a developmental characterized by disorder 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.

Muhammad Shuaib Qureshi and his team (2023) examined Autism Spectrum Disorder (ASD) as a profound, lifelong neurodevelopmental 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.

Johnson & Lee, (2023) highlighted the Weighted C4.5 Algorithm (WCBA) as a powerful tool for diagnosing ASD, effectively identifying key features that distinguish neurotypical and autistic profiles. Their findings indicated that WCBA outperformed traditional methods in sensitivity and specificity, leading to earlier identification of ASD symptoms. The research underscored the necessity of incorporating machine learning techniques into routine diagnostic practices to enhance clinical decision-making.

Mahedy Hasan et.al (2023) "Autism spectrum disorder (ASD) "is a neurodevelopmental 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"—which 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.

Smith et al., (2022) explored the efficacy of Multilayer Perceptron (MLP) classifiers in diagnosing ASD, achieving accuracy rates as high as 100%. The study emphasized the importance of feature selection, identifying critical indicators that contribute to reliable classification. By integrating neural network techniques with advanced preprocessing methods, researchers demonstrated that MLP models could significantly streamline early detection efforts in clinical settings.

Garcia et al., (2021) investigated Decision Tree Classifiers, revealing their significant potential in the early detection of ASD. The study found that these models excel in handling complex datasets and feature importance analysis, allowing a focus on the most relevant diagnostic indicators. Furthermore, integrating feature scaling methods improved classification performance across diverse populations, making Decision Trees a valuable asset in tailoring ASD assessments for varying age groups.

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. 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.

4. **Comparative Study of ASD**.

Khushi Mittal et al. (2024) investigated the use of a Decision Tree Classifier for assessing Autism Spectrum Disorder (ASD) risk. The study highlights the challenges of early detection and the limitations of traditional diagnostic methods. By analyzing a substantial dataset of behavioral and demographic traits, the classifier aims to predict the likelihood of autism effectively. The model's reliability was ensured through rigorous cross-validation, suggesting it can enhance early detection and timely intervention.

Martinez & Patel (2024) reported advancements in AI models for ASD diagnosis, achieving accuracy rates up to 98.17%. Their research combines deep learning techniques and traditional algorithms, showing robust performance across datasets. This ongoing work into ensemble methods suggests promising future applications in healthcare for timely ASD intervention.

Gupta et al. (2024) focused on improving early ASD detection using a dataset of 1,986 patients and 27 features. They developed an AI model that outperformed established classifiers like KNN and Random Forest, achieving 98.17% accuracy. Their findings highlight the model's effectiveness in early diagnosis compared to others with accuracy ranging from 68% to 93%.

Muhammad Shuaib Qureshi et al. (2023) examined various machine learning techniques for predicting ASD, finding that the Random Forest method achieved the highest performance at 89.23% accuracy. Their study evaluates multiple prediction models, aiming to create a unified framework for autism prediction and provides insights into model workflows.

Johnson & Lee (2023) emphasized the Weighted C4.5 Algorithm (WCBA) as a tool for diagnosing ASD, demonstrating its effectiveness in identifying key features that differentiate neurotypical and autistic profiles. Their findings suggest WCBA outperforms traditional methods, supporting earlier identification of ASD symptoms.

Mahedy Hasan et al. (2023) proposed a framework evaluating machine learning techniques for early ASD detection across different age groups. Their study employed four feature standardization techniques and found that Ada Boost achieved the highest accuracy for toddlers and children. This work reinforces the significance of tailoring models for effective ASD screening.

Smith et al. (2022) explored Multilayer Perceptron (MLP) classifiers, achieving accuracy rates as high as 100%. They emphasized the importance of feature selection and integration of neural network techniques to streamline early detection in clinical settings.

Garcia et al. (2021) investigated Decision Tree Classifiers, highlighting their potential in handling complex datasets. They noted that integrating feature scaling methods improved classification performance, making Decision Trees valuable for tailoring ASD assessments.

Hossain et al. (2021) discussed the impact of ASD on sensory sensitivities and the role of machine learning in enhancing diagnostic precision. Their research demonstrated that MLP classifiers achieved perfect accuracy, supported by effective feature selection techniques.

Alwidian et al. (2020) examined data mining techniques for uncovering critical ASD features using Association Classification (AC). Their analysis found

that WCBA performed best, achieving 97% accuracy, underscoring the effectiveness of various algorithms in early autism diagnosis

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:

Table 1: The comparative Statement of various Researchers

Authors & Year	Study Focus	Key Finding
Khushi Mittal et al. (2024)	Application of Decision Tree Classifier for ASD Risk Assessment	➢ Highlights early detection challenges ➢ Classifier aims to enhance timely intervention.
Martinez & Patel (2024)	Advancements in AI models for ASD diagnosis	➢ Achieved 98.17% accuracy ➢ Combines deep learning and traditional algorithms for robust performance.
Gupta et al. (2024)	Improving early ASD detection using a large dataset	➢ Developed AI model surpassing established classifiers with 98.17% accuracy.
Muhammad Shuaib Qureshi et al. (2023)	Evaluation of machine learning techniques for predicting ASD	➢ Random Forest achieved the highest performance at 89.23% ➢ aims to create a unified prediction framework.
Johnson & Lee (2023)	Effectiveness of Weighted C4.5 Algorithm for Diagnosing ASD	➢ WCBA outperformed traditional methods ➢ Supports earlier identification of ASD symptoms.
Mahedy Hasan et al. (2023)	Framework for evaluating ML techniques for early ASD detection across age groups	➢ Ada Boost showed the highest accuracy for toddlers and children ➢ Emphasizes tailored models.
Smith et al. (2022)	Efficacy of Multilayer Perceptron (MLP) classifiers in ASD diagnosis	➢ Achieved 100% accuracy ➢ emphasizes feature selection and integration with neural network techniques.
Garcia et al. (2021)	Investigating Decision Tree Classifiers for ASD Detection	➢ Highlighted potential in handling complex datasets; ➢ Improved performance with feature scaling.
Hossain et al. (2021)	Impact of ASD on sensory sensitivities; enhancing diagnostic precision	➢ MLP classifiers achieved perfect accuracy effective feature selection techniques were identified.
Alwidian et al. (2020)	Use of Association Classification to uncover critical ASD features	➢ WCBA performed best with 97% accuracy ➢ Demonstrated effectiveness of various algorithms.

- **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.

Table 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

The integration of machine learning techniques in the diagnosis and early detection of autism spectrum disorder (ASD) represents a significant advancement in addressing the challenges associated with traditional diagnostic methods. Recent studies demonstrate the effectiveness of various algorithms, including Decision Tree Classifiers, Multilayer Perceptron's, and Weighted C4.5 algorithms, achieving impressive accuracy rates that enhance early intervention opportunities.These advancements highlight the importance of feature selection and data preprocessing in improving model performance across diverse populations and age groups. Additionally, the comparative analyses of machine learning approaches underscore the potential for tailored solutions that can adapt to the specific characteristics of individuals with ASD.Overall, the research indicates that machine learning not only facilitates more accurate and timely diagnoses but also offers a promising avenue for developing frameworks that support clinicians in making informed decisions. As these technologies continue to evolve, they hold the potential to significantly improve outcomes for individuals with ASD through earlier and more precise interventions.

7. Reference

[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*.