SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I7P112

Volume 11 Issue 7, 117-127, July 2024

Original Article

Ensemble Machine Learning for Classification of Autism Spectrum Disorder in Toddlers and Adults

Yogita Dubey[1], Atharva Patrikar[2], Padmesh Kimmatkar[3], Prajwal Dehankar[4], Punit Fulzele[5]

[1,2,3,4]*Department of Electronics and Telecommunication Engineering, Yeshwantaro Chavan College of Engineering, Maharashtra, India.*
[5]*Department of Research and Development, Datta Meghe Institute of Higher Education and Research, Maharashtra, India.*

[1]*Corresponding Author : yogeetakdubey@yahoo.co.in*

Received: 09 May 2024 Revised: 08 June 2024 Accepted: 08 July 2024 Published: 27 July 2024

Abstract *- Autism Spectrum Disorder (ASD) is a neurological disorder. A Person with ASD always faces challenges in communicating socially and is involved in repetitive behaviors. Accurate and timely diagnosis is essential for efficient assistance and action. This paper presents the ensemble machine learning methodology to classify ASD in both toddlers and adults. Four distinct algorithms, Gradient Boosting (GB), Histogram Boosting (HB), Extreme Gradient Boosting (XGB) and Adaptive Boosting (ADB), are used. Explorative data analysis is performed to show the impact of behavioural features and individual characteristics on ASD in toddlers and adults. Quantitative analysis demonstrates that XGB outperforms with classification accuracy, log loss and F1-score, precision and recall. The findings indicate that the ensemble machine learning methodology has great potential to improve the diagnostic procedures for ASD, possibly resulting in an earlier and more accurate diagnosis of the condition.*

Keywords *- Ensemble machine learning, Boosting algorithms, Autism classification, Parametric analysis, Quantitative assessment, Toddlers and Adults.*

1. Introduction

Autism Spectrum Disorder (ASD) is a complicated neurodevelopmental disorder that impacts behavior, social interaction, and communication. Since ASD is becoming more common, it is critical to create reliable and effective diagnostic instruments. ASD symptoms typically fall into two groups and can range in severity from mild to severe. The first is social contact and communication, which makes it hard to keep up a conversation, read social cues, and build relationships. The second is characterized by limited interests and repetitive actions, which lead to recurrent behaviors, insistence on routines, and extreme interest in particular subjects or pursuits. Although ASD is usually identified in the early stages of childhood, symptoms can last a person's entire life. Conventional diagnostic techniques mostly rely on behavioral evaluations and clinical knowledge, and they are frequently subjective and time-consuming.

Machine Learning (ML) plays an important role in the diagnosis and classification of ASD [1, 2]. Akter et al. proposed ML models for early diagnosis of ASD, emphasizing the importance of timely intervention [3]. Parikh, Li, and He utilized optimized ML models combined with personal characteristic data to improve the diagnosis of ASD. Their approach highlights the potential of personalized data to enhance the accuracy of ML models [4]. Thabtah and Peebles introduced a new ML model based on rule induction for autism detection. This model stands out for its interpretability, making it easier for clinicians to understand and trust the diagnostic process [5]. Thabtah conducted a comprehensive review of ML applications in ASD behavioral research, suggesting future directions to improve the effectiveness and applicability of ML in this field [6, 8]. Stevens et al. utilize unsupervised ML techniques to find and analyze behavioral phenotypes in ASD, emphasizing the role of unsupervised learning in uncovering hidden patterns in behavioral data [7].

Khan et al. propose a neural network approach combined with sequential feed-forward selection for ASD classification, showcasing the advancements in neural network methodologies for this application [9]. Chaidi and Drigas reviewed the application of ML in the context of emotional expression and understanding in individuals with ASD, highlighting the potential of ML to address the social and emotional challenges associated with ASD [10]. Rahman et al. examine various ML methods for feature selection and classification in ASD, providing insights into effective techniques for managing ASD data and improving classification accuracy [11]. Zheng, Deng, and Wang apply logistic regression models to the classification of ASD,

demonstrating the efficacy of these models in processing clinical and behavioral data for diagnostic purposes. They reported an accuracy of 89.7%, showcasing the robustness of simpler models in certain scenarios [12].

ML models, such as support vector machine [13], deep learning [14], logistic regression [15], random forest [16], and neural network [17], have achieved high accuracy rates, ranging from 85% to 95% for the classification of ASD. ML approaches have been successfully applied for ASD classification in both todllers and adults.

For instance, models tailored to pediatric populations demonstrated robust performance, while those designed for adult populations also showed promising results with accuracy of 95% and 93% [18-21]. Parlett-Pelleriti et al. reviewed the use of unsupervised ML in ASD research, highlighting its potential to uncover hidden patterns and phenotypes in ASD data without predefined labels [22].

Simeoli et al. conducted a systematic review of ML for motion analysis in early ASD detection, reflecting the growing integration of movement data in diagnostic models [23]. Aarthi and Kannimuthu provide a comprehensive analysis of various ML algorithms used in ASD research, offering a detailed comparison of their performance and applicability [24]. Jayanthi et al. investigate the use of ML for monitoring the mental health status of individuals with ASD, showcasing the potential of ML to support ongoing mental health assessments and interventions. The survey demonstrated that the accuracy ranged from 85% to 95% depending on the ML algorithm and the dataset used. The review emphasized the importance of feature selection and data preprocessing in achieving high accuracy [25].

Survey shows that the accuracy of ML models for ASD classification is generally high, with many studies reporting accuracies above 90%. However, the choice of algorithm, the quality of the dataset, and the specific features used are critical factors that influence the overall performance of these models. Continued research and development in this area are likely to yield even more accurate and robust models in the future. There is a need to develop ML models that are specifically tailored for early detection in infants and toddlers. There is a growing demand for interpretable and explainable ML models that provide accurate classifications and explain which factors contribute more to ASD diagnosis.

This can help medical professionals and caregivers understand and trust ML based recommendations. Section 2 gives the methodology adapted for ASD classification along with a detailed analysis of features contributing to ASD in toddlers and adults. Section 3 presents the discussion of simulation results. The research's findings are concluded in Section 4.

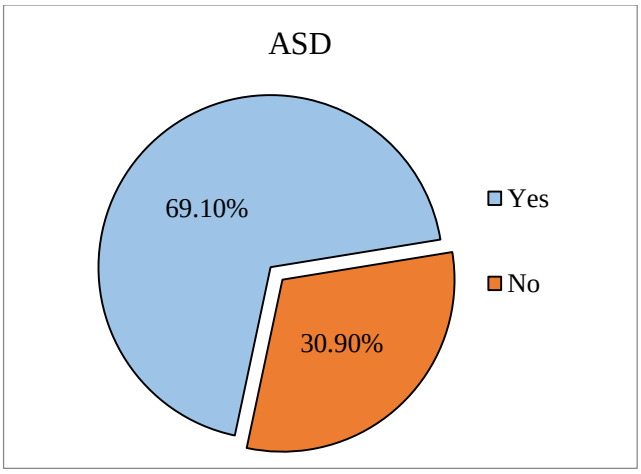


Fig. 1 Percentage of ASD and Non-ASD cases in Toddlers dataset

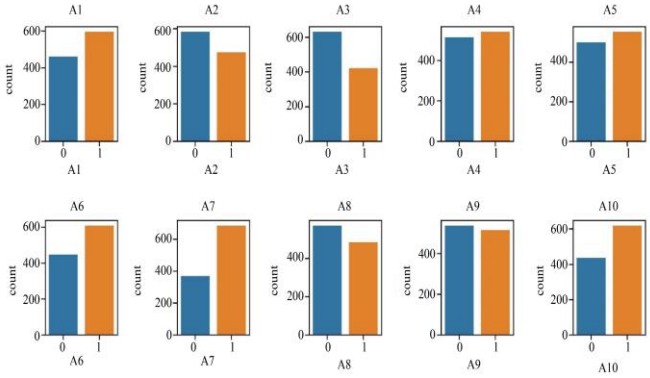


Fig. 2 ASD and Non-ASD scores

2. Materials and Methods

The methodology employed for the prediction of autism disorder using ensemble boosting algorithms is discussed in this section, along with a detailed analysis of parameters contributing to ASD in toddlers and adults.

2.1. Dataset Description

Toddler's datasets contain ten behavioural features. For the reference study, the dataset was collected from Kaggle [26]. In total, there were 1054 instances in the dataset and 17 attributes, including class variables. From Figure 1, an estimation can be made that around 69.1% of toddlers are affected by ASD.

2.1.1. Attributes

There are 10 Questions within Q-Chat-10 named as A1 to A10. The possible answers to these questions are "Always, Usually, Sometimes, Rarely and Never", which are recorded as "1" or "0". For A1-A9, the responses like Sometimes / Rarely / Never is recorded as "1". For A10, the responses Always / Usually / Sometimes are recorded as "1". A score of more than 3 in Q-chat 10 for any individual has a possibility of ASD. The distribution of ASD and Non-ASD scores for A1-A10 chat is shown in Figure 2.

2.1.2. Ethnicity

Ethnicity significantly impacts ASD cases, influencing prevalence, diagnosis, and access to treatment. Minority children often face delays in diagnosis due to cultural differences, language barriers, and healthcare biases. Access to specialized ASD services is limited for these groups, compounded by socioeconomic factors and lack of insurance. Cultural perceptions and stigma can deter families from seeking help, leading to underreporting. Figure 3 shows that Native Indian and Pacifica ethnicity are more prone to ASD.

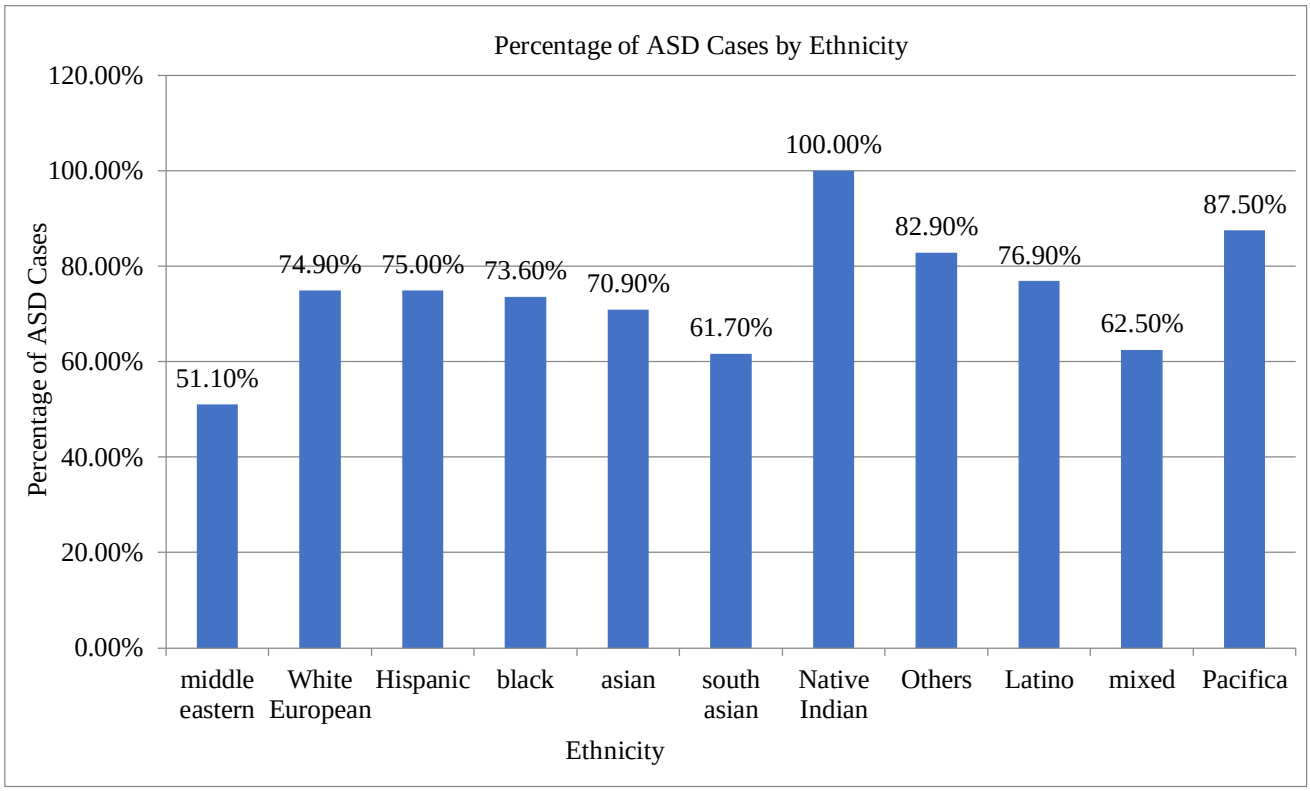


Fig. 3 Percentage of ASD cases by ethnicity

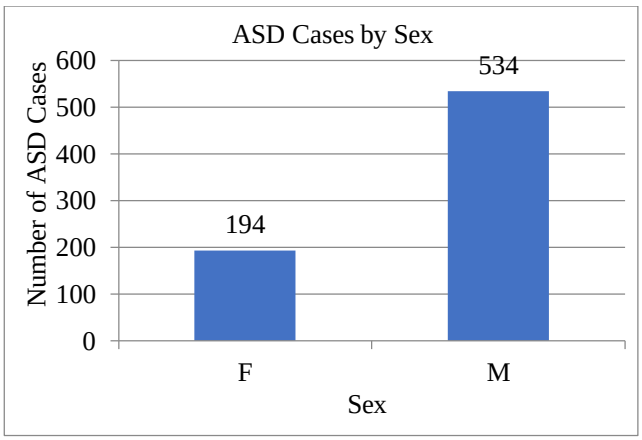


Fig. 4 Number of ASD cases by gender

2.1.3. Sex

Sex plays a significant role in ASD cases, affecting prevalence, diagnosis, and manifestation of symptoms. Boys are diagnosed with ASD more frequently than girls. This discrepancy is thought to be influenced by both biological and social factors. Biologically, there may be genetic and hormonal differences that contribute to the higher prevalence of ASD in boys. Certain genetic mutations and the influence of sex hormones like testosterone might play a role in this disparity. However, the exact mechanisms remain an active area of study. Figure 4 shows that males are more prone to ASD in comparison to females.

2.1.4. Jaundice

Neonatal jaundice, characterized by high bilirubin levels in newborns, has been investigated for its potential link to ASD.

Some studies suggest that severe or prolonged jaundice might increase the risk of developing ASD due to the neurotoxic effects of elevated bilirubin, which can damage brain tissues involved in social behaviour, communication, and repetitive behaviours. The newborn period is critical for brain development, and conditions affecting the brain during this time could influence neurodevelopmental outcomes. Figure 5 shows that toddlers having jaundice are more prone to ASD in comparison to normal toddlers.

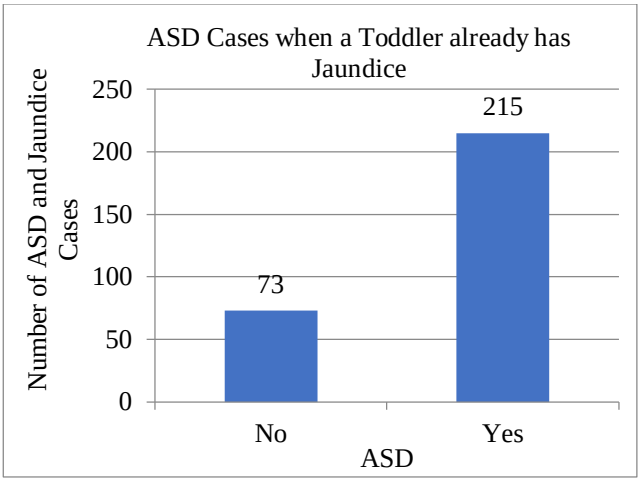


Fig. 5 Impact of jaundice on ASD

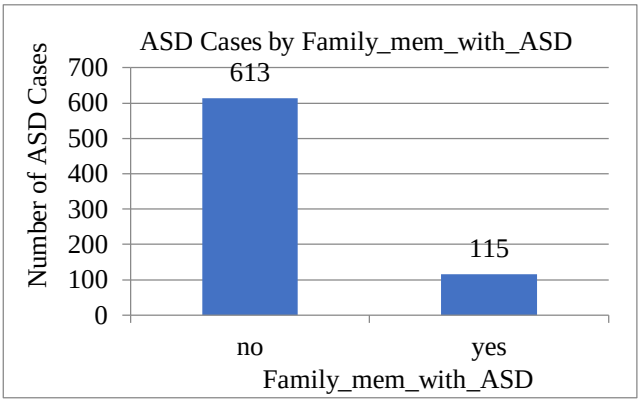


Fig. 6 ASD cases by family members

2.1.5. Family Member

ASD is influenced by genetics, with studies showing a strong hereditary component. Specific genetic mutations and family patterns indicate that genetics play a significant role. However, ASD also involves a complex interaction with environmental factors, meaning it is not solely a genetic disease but rather a condition influenced by both genetic and environmental elements. Figure 6 shows that most of the children with ASD around the world do not have their family members with ASD, which means ASD is not a genetic disease.

2.1.6. Age

The impact and management of ASD in toddlers vary with age. Figure 7 shows that toddlers aged 36 months have the most ASD cases around the world.

Early signs, noticeable from 12 to 18 months, include lack of eye contact and delayed speech, and early diagnosis allows for crucial early interventions. From 18 to 36 months, symptoms like repetitive behaviours and difficulty with social interactions become more apparent, necessitating intensive therapies.

2.1.7. Q-Chat-10-Score

Q-CHAT-10 is a screening tool for the assessment of the risk of ASD. It consists of 10 questions that assess communication skills and behaviours associated with ASD. Their Qchat-10 score influences the impact of ASD in toddlers, a screening tool used to assess risk for ASD.

A higher score indicates a higher chance of ASD, prompting earlier diagnosis and intervention. Toddlers with elevated scores may exhibit more pronounced symptoms, such as limited social interaction and communication challenges. Figure 8 shows that the toddlers with a Qchat-10-Score value greater than 3 have ASD.

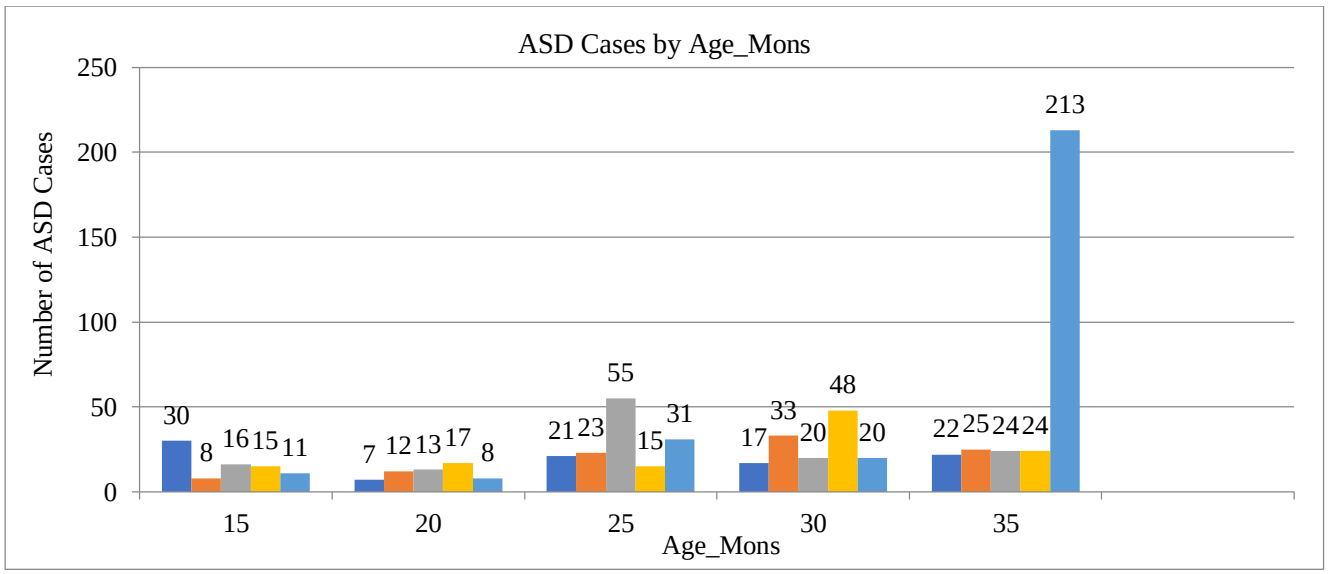


Fig. 7 ASD cases in Toddlers as per age in months

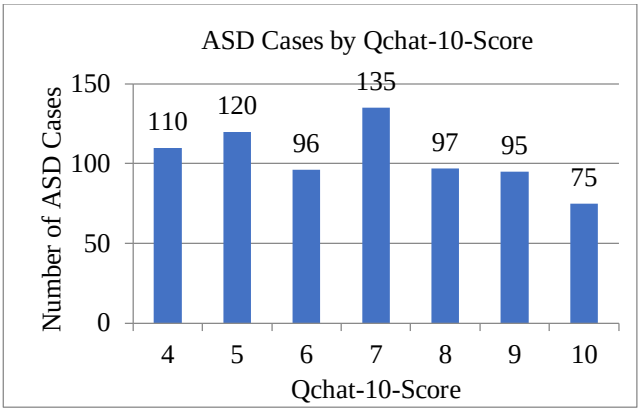


Fig. 8 ASD cases by Qchat-10 Score

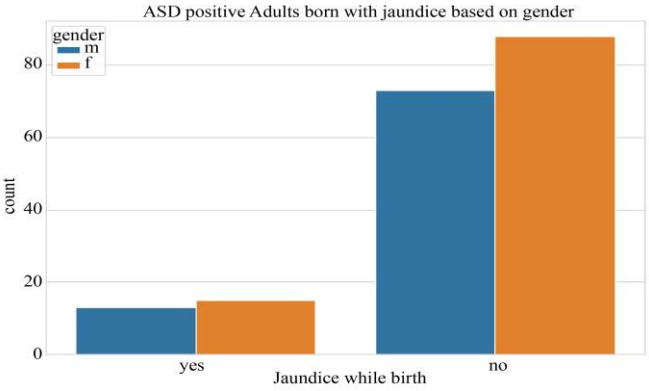


Fig. 9 ASD cases for adults born with jaundice based on gender

The adult dataset also has ten behavioural features (AQ-10-Adult) and ten individual characteristics. For the reference study, the dataset was collected from Kaggle [27]. In total, there were 704 instances in the dataset, and 20 features were the same as toddlers. In the adult dataset, 29.55% of the population has ASD, and 70.44% are non-ASD.

2.1.8. Jaundice

The jaundice data is analyzed while birth based on gender in the dataset. Studies have shown that almost 6-7 times more (in adults) of non-jaundice were born with ASD positive. Infants born jaundiced have a tenuous connection to ASD. Additionally, it has been shown that boys are more likely than girls to have ASD (by about 4-5 times). However, we observe a smaller ratio in adults. It is depicted in Figure 9.

2.1.9. Age

The age distribution of adults with ASD typically shows a higher prevalence in younger age groups, reflecting increased diagnosis rates in recent years due to greater awareness and improved diagnostic practices. Some individuals may not be diagnosed until adulthood, leading to variability in diagnosis ages. The majority of individuals with ASD are between the ages of 20 and 30. Adults have a decreasing number as they get older. As illustrated in Figure 10, adults with autism develop coping mechanisms to facilitate a better aging process.

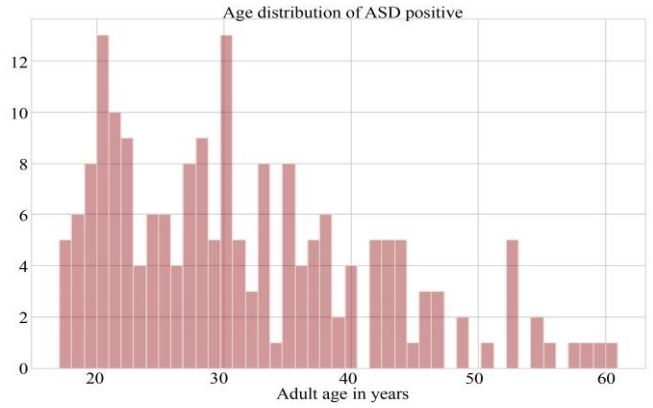


Fig. 10 ASD cases in adults as per age

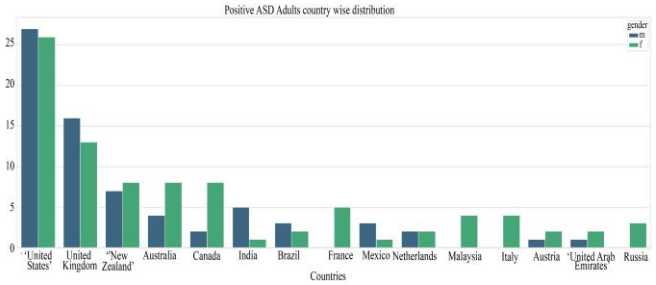


Fig. 11 Gender specific country wise ASD cases in adults

2.1.10. Country of Residence

The impact of ASD on adults can vary significantly depending on their country of residence due to differences in healthcare systems, availability of support services, cultural attitudes, and economic resources. In countries with robust healthcare and social support systems, adults with ASD may have better access to diagnostic services, leading to improved quality of life. Whereas, in countries with limited resources for mental health, individuals with ASD might face greater challenges in accessing necessary services, receiving proper diagnosis, and integrating into society. The developed nations that are most impacted are, in fact, the US, Canada, Australia, and the UK. On the other hand, we can clearly distinguish the female population from the male population, as depicted in Figure 11.

2.1.11. Ethnicity

The impact of ASD on adults varies with ethnicity due to differences in cultural perceptions, healthcare access, diagnostic practices, and socioeconomic factors. Cultural attitudes towards ASD can affect whether individuals seek or receive a diagnosis and support, with some ethnic groups facing stigma or misconceptions. Access to healthcare also plays a significant role, as minority groups may encounter barriers such as language differences, lack of culturally competent care, and financial constraints, leading to disparities in diagnosis and treatment. ASD cases for white and European country-wise distribution are plotted in Figure 12. The US, UK, Australia, NZ, and Canada have the highest number of positive ASD.

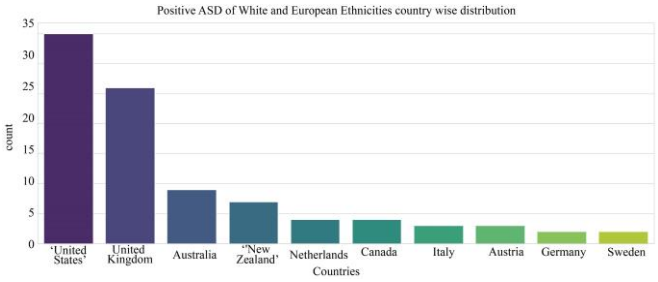


Fig. 12 Count of ASD cases for white and European ethnicities in adults

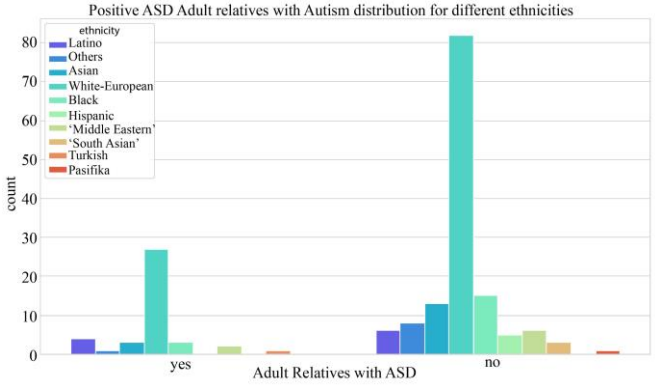


Fig. 13 ASD cases with respect to ethnicity

It is observed from Figure 13 that adults of White and European descent are most likely to have an ASD, followed by Black and Asians. But then also, it is difficult to conclude any genetic relevance for ASD positive.

2.2. Data Preprocessing

ASD data is first pre-processed for further processing. Out of 16 features in the dataset, sex, ethnicity, jaundice, and family with ASD features are categorical, which are converted to numerical using label encoding. There were no null cases in the dataset. Hence, all the samples of the dataset are used for experimentation. Data splitting is done for training (80%) and testing (20) for the application of ML algorithms. The standard scaling is also used to normalize all the features on the same scale.

2.3. Machine Learning (ML) Algorithms

Ensemble ML algorithms are techniques that combine multiple models to improve the performance, robustness, and generalization of predictions. This involves bagging, boosting, stacking, voting and blending. In this paper, four popular boosting ML algorithms, Gradient Boosting (GB), Histogram Boosting (HB), Extreme Gradient Boosting (XGB) and Adaptive Boosting (ADB), are used. These are described below.

2.3.1. Gradient Boosting (GB)

GB is an ensemble ML technique used for classification and prediction that sequentially builds models. The advantage of GB is that it provides the highest accuracy using an optimized loss function on different types of data with insight into feature importance [29]. Given an input-output pair (x, y), the initial prediction for GB using loss function L is

$$F_0(X) = arg\,\min_c \sum_{i=1}^{n} L(y_i, c) \tag{1}$$

The next step is computing the residuals for M number of iterations using

$$r_i^m = -\left[\frac{\partial L(y_i, F(X_i))}{\partial F(X_i)}\right]_{F(X)=F_{m-1}(X)} \quad \text{for } m = 1toM \tag{2}$$

The weak learner $h_m(X)$ is fitted to the residuals r_i^m and the model is updated using,

$$F_m(X) = F_{m-1}(X) + \nu h_m(X) \tag{3}$$

Where, ν is the learning rate.

2.3.2. Extreme Gradient Boosting (XGB)

XGB is a sophisticated gradient boosting approach optimised for speed and efficiency. The advantage of XGB is that it provides regularization to prevent overfitting and handles missing data. The objective function in XGB includes both the loss function and the regularization term given by

$$\phi(\theta) = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{4}$$

Where, n is total training samples, K is the number of trees, θ represents all the parameters of the model (including the weights and structure of the trees) and $\hat{y}_i$ is the predicted output at i^{th} instance. Ω is the regularization term to prevent overfitting [30].

2.3.3. Adaptive Boosting (ADB)

AdaBoost assigns weights to each training instance. Initially, all weights are equal. In each iteration, it trains a weak learner and adjusts the weights of misclassified instances, increasing the influence of hard-to-classify points. Initially, all training instances are assigned equal weights. If there are n instances, each instance is initially assigned a weight $w_i = \frac{1}{n}$. The weak learner h_t is trained on the weighted training data, and its performance is evaluated using a weighted error rate.

$$e_t = \sum_{i=1}^{n} w_i . I\left(y_i \neq h_t(X_i)\right) \tag{5}$$

Where, I is the binary function that returns 1 for incorrect prediction and 0 otherwise. The weak learner's weight α_t is calculated using

$$\alpha_t = \frac{1}{2} ln\left(\frac{1 - e_t}{e_t}\right) \tag{6}$$

and weights of the training instances are updated using

$$w_{i+1} = w_i . exp\left(\alpha_t . I\left(y_i \neq h_t(X_i)\right)\right) \quad (7)$$

These weights are normalized by dividing the sum of all weights [31].

2.3.4. Histogram Boosting (HB)

HB is an efficient implementation of GB that uses histogram-based algorithms. Instead of using the raw feature values, it first bins the feature values into discrete bins (histograms). This binning process reduces computational complexity and memory usage. The key steps in histogram boosting are Data Binning, Gradient Calculation, Split Finding and Tree Building. Given a feature x, we define K bins as

$B_k = \{x | b_{k-1} \leq x \leq b_k\}, k = 1,2,\ldots\ldots\ldots K$, where b_k are the bin edges.

For each bin, B_k compute the sum of gradients G_k and Hessians H_k (second-order gradients) as $G_k = \sum_{i \in B_k} g_i$, $H_k = \sum_{i \in B_k} h_i$,

Where, g_i and h_i are the gradient and Hessian for instance i, respectively. For a potential split at the bin s , the gain is calculated as

$$Gain = \frac{G_{left}^2}{H_{left} + \lambda} + \frac{G_{right}^2}{H_{right} + \lambda} + \frac{\left(G_{left} + G_{right}\right)^2}{H_{left} + H_{right} + \lambda} - \gamma \quad (8)$$

Where, G_{left}, H_{left}, G_{right} and H_{right} is the sum of gradients and Hessians, respectively, for the left and right child nodes. $\lambda$is a regularization parameter. γ is a complexity parameter controlling the minimum gain for a split.

Based on the best splits, the tree is constructed. Each internal node represents a split based on histogram bins, and each leaf node contains a value, which is the average of the target values of the instances falling into that leaf [32].

3. Results and Discussion

This research investigated the efficacy of ensemble-driven machine learning approaches for predicting ASD in toddlers and adults. The performance of four prominent algorithms, Extreme Gradient Boosting, Gradient Boosting, AdaBoost and Histogram-based Gradient Boosting, are compared.

Two separate datasets were employed, one for toddlers and another for adults. Each dataset was evaluated using various performance metrics. The confusion matrix is generated for each dataset using these ML algorithms. Table 1 shows the structure of the confusion matrix for the evaluation of ASD.

Table 1. Confusion matrix for ASD

		Actual Output	
		Patient with ASD	**Patient with No ASD**
Predicted Output	Patient with ASD	Correctly predicted ASD (TP)	Incorrectly predicted ASD when there is no ASD (FP)
	Patient with No ASD	Incorrectly predicted no ASD when it is ASD (FN)	Correctly predicted no ASD (TN)

Based on TP, TN, FP and FN, the following quantitative metrics are defined for ASD classification.

- Accuracy: Proportion of correctly classified instances (both ASD and non-ASD) $Accuracy = \frac{TP+TN}{N}$.
- Precision: Proportion of predicted ASD cases that were truly ASD Precision $= \frac{TP}{TP+FP}$.
- Recall: Proportion of actual ASD cases that were correctly identified Precision $= \frac{TP}{TP+FN}$.
- F1-score: Balances precision and recall by taking their harmonic mean $F_1 - Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+Recall}}$.

Log Loss: Measures the model's capacity to differentiate between ASD and non-ASD cases. Lower values indicate superior performance.

$$LogLoss = -\frac{1}{N}\left(\sum_{i=1}^{N}(y_i \, log(p_i) + (1 - y_i) \, log(1 - p_i))\right),$$

Here y_i is the actual class labels for ASD for i^{th} sample and p_i the predicted probabilities of class with ASD for i^{th} sample, N is the total number of samples in the dataset.

The accuracy of ASD classification obtained using these four ML algorithms is given in Table 2. It is observed that all algorithms achieved high training accuracy (100% for most) on both datasets. This suggests that the models learned the training data well. Testing accuracy is generally lower than training accuracy, but it remains high (>95%) for all algorithms in both datasets. This indicates reasonable generalizability of the models to unseen data. XGB achieved the highest testing accuracy (97.63%) for the toddler dataset. There is a slight difference in performance between some algorithms on the toddler dataset compared to the adult dataset. For instance, HB performs better on the adult dataset (97.54%) compared to the toddler dataset (96.20%).

Overall, Table 2 highlights the effectiveness of these ML algorithms for ASD prediction in both toddlers and adults, with XGB demonstrating the strongest performance on the toddler dataset based on testing accuracy.

Table 2. Accuracy of ASD classification using GB, XGB, ADB and HB ML algorithms

	Toddlers Dataset		Adults Dataset	
Algorithm	**Training**	**Testing**	**Training**	**Testing**
Gradient Boosting	100%	97.63%	100%	97.54%
Extreme Gradient Boosting	100%	97.63%	100%	95.90%
Adaptive Boosting	100%	100%	100%	100%
Histogram Boosting	100%	96.20%	100%	97.54%

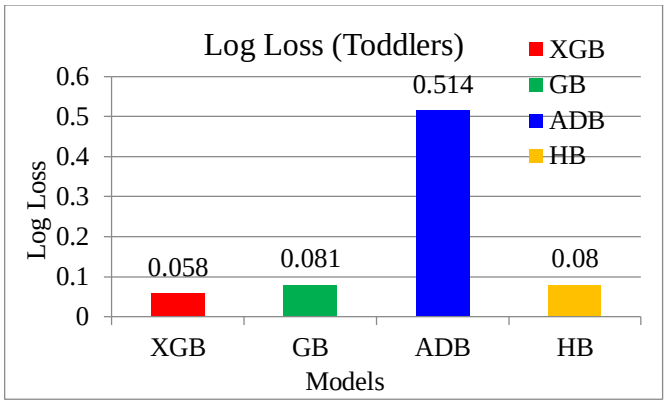
Table 3. Comparative analysis of various quantitative metrics for ASD classification on Toddlers dataset

Sr. No	Algorithm	Precision	Recall	F1-Score
1.	Gradient Boosting	97.64%	97.63%	97.61%
2.	Extreme Gradient Boosting	97.62%	97.63%	97.62%
3.	Adaptive Boosting	100%	100%	100%
4.	Histogram Boosting	96.20%	96.20%	96.20%

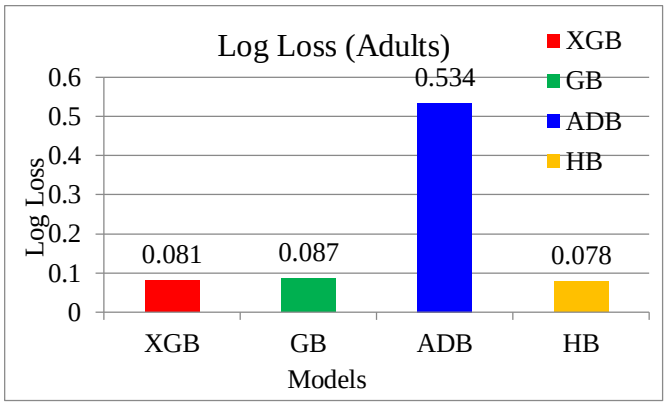
Table 4. Comparative analysis of various quantitative metrics for ASD classification on Adults dataset

Sr. No	Algorithm	Precision	Recall	F1-Score
1.	Gradient Boosting	97.54%	97.54%	97.53%
2.	Extreme Gradient Boosting	95.89%	95.90%	95.89%
3.	Adaptive Boosting	100%	100%	100%
4.	Histogram Boosting	97.54%	97.54%	97.53%

Tables 3 and 4 show a comparative analysis of various quantitative metrics for ASD classification using four ML algorithms. XGB again emerges as a strong performer, achieving high and balanced scores (around 97%) for F1-score, precision and recall in the toddler dataset. The performance of the algorithms can vary across the toddler and adult datasets. For example, GB and HB exhibit very similar metrics in the toddler dataset. However, Histogram-based Gradient Boosting achieves slightly better precision (identifying more true ASD cases out of the predicted ASD cases) and recall (capturing a higher proportion of all true ASD cases) in the adult dataset. ADB, despite achieving 100% accuracy (potentially due to overfitting), has much lower values in the other metrics compared to other algorithms, suggesting it might not be effectively distinguishing between classes.



(a)



(b)

Fig. 14 Comparison of Log Loss obtained using XGB, GB, ADB and HB algorithm on (a) Toddlers dataset, and (b) Adults dataset.

A comparison of Log Loss obtained using XGB, GB, ADB and HB algorithms for ASD classification is shown in Figure 14. As depicted in Figure, XGB outperformed other algorithms with the lowest log loss in both datasets (0.0810 for adults and 0.0578 for toddlers). This indicates its superior ability to distinguish between classes in both age groups. GB has a slightly higher log loss compared to XGB in both datasets (0.0871 for adults and 0.0811 for toddlers). ADB achieved significantly higher log loss compared to all other algorithms in both datasets (0.5344 for adults and 0.5142 for toddlers). This suggests a much poorer ability to differentiate between classes, potentially due to overfitting. HB achieved the second-lowest log loss (0.0780) in the adult dataset, performing slightly better than GB. However, its performance in the toddler dataset (0.0803) was closer to GB. This highlights a potential dependence of this algorithm's effectiveness on the specific dataset characteristics.

A comparison of AUC values obtained using XGB, GB, ADB and HB algorithms for ASD classification is shown in Figure 15. As expected, ADB AUC curves in both toddler and adult datasets likely stick very close to the x-axis (FPR). This indicates very poor discrimination between ASD and non-ASD cases. A near-flat AUC curve for ADB suggests it might be performing no better than random guessing when

classifying ASD in both datasets. The AUC curves for the other algorithms (XGB, GB and HB) likely show more variation compared to AdaBoost.

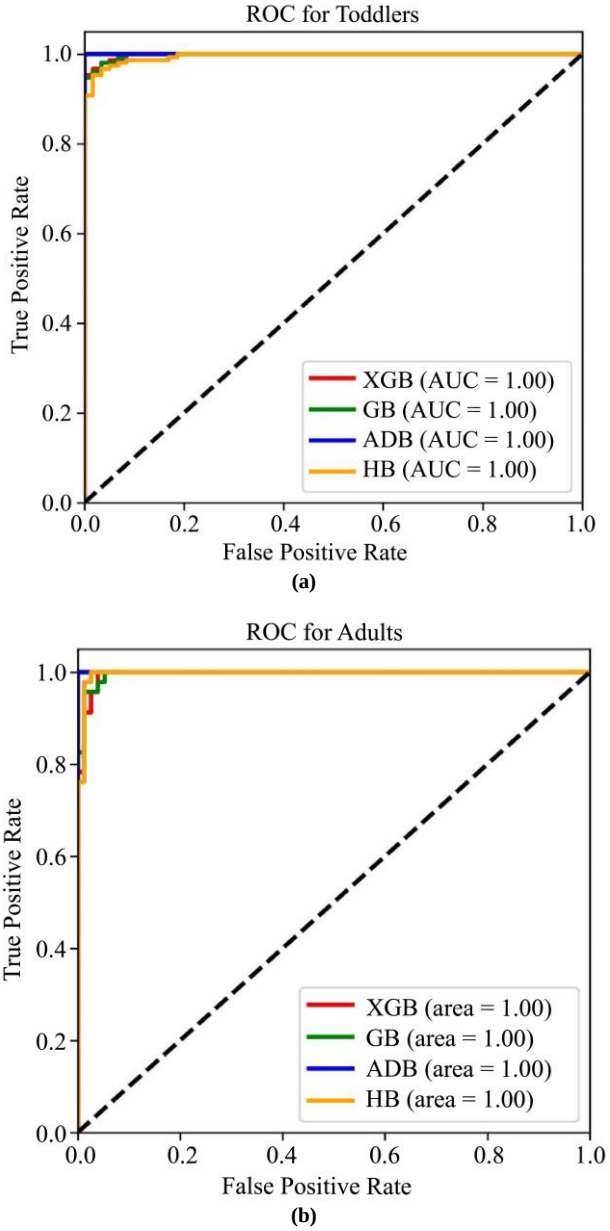


Fig. 15 Comparison of AUC values obtained using XGB, GB, ADB and HB algorithm on (a) Toddlers dataset, and (b) Adults dataset.

The AUC curves for the toddler dataset exhibit more steps or changes along the horizontal axis (FPR). This indicates that the models are struggling to differentiate between some true positives and false positives, leading to more fluctuations in the AUC calculation.

The AUC curves in the adult dataset show more significant changes along the vertical axis (TPR) indicating the models are more effective at correctly classifying true positives (identifying actual ASD cases) in the adult dataset compared to the toddler dataset. The reasons behind these observations could be due to inherent challenges in diagnosing ASD in toddlers compared to adults. Early signs of ASD might be more varied in toddlers, making it more difficult for the models to distinguish them from non-ASD cases.

4. Conclusion

In this paper, ensemble machine learning methods using Gradient Boosting (GB), Histogram Boosting (HB), Extreme Gradient Boosting (XGB) and Adaptive Boosting (ADB) algorithms are used for the classification of ASD in both toddlers and adults. XGB turned out to be the most effective algorithm, showing superior performance with F1-score, log loss, accuracy, precision, and recall.

The exceptional performance of XGB, especially its decreased log loss in the adult and toddler datasets, highlights its potent ability to distinguish between cases with and without ASD. The AUC curves indicated possible differences in model performance across the toddler and adult datasets and graphically confirmed ADB's limitations.

Results show that identifying ASD in toddlers may be more difficult than in adults, based on the variations in the shapes of AUC curves. Based on the quantitative assessment using all metrics, results demonstrate the stability and promise of XGB as a trustworthy technique for classifying ASD in a range of age groups.

Authors Contribution

YD and PK conceived and designed the research. AP performed parametric analysis. PD and AP conducted experiments. PK contributed to the simulation. YD and AM wrote the manuscript. PF validated the results. All authors read and approved the manuscript.

References

[1] Fadi Thabtah, Neda Abdelhamid, and David Peebles, "A Machine Learning Autism Classification Based on Logistic Regression Analysis," *Health Information Science and Systems*, vol. 7, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[2] Ugur Erkan, and Dang N.H. Thanh, "Autism Spectrum Disorder Detection with Machine Learning Methods," *Current Psychiatry Research and Reviews Formerly: Current Psychiatry Reviews*, vol. 15, no. 4, pp. 297-308, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[3] Tania Akter et al., "Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders," *IEEE Access*, vol. 7, pp. 166509-166527, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[4] Milan N. Parikh, Hailong Li, and Lili He, "Enhancing Diagnosis of Autism with Optimized Machine Learning Models and Personal Characteristic Data," *Frontiers in Computational Neuroscience*, vol. 13, pp. 1-5, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] Fadi Thabtah, and David Peebles, "A New Machine Learning Model Based on Induction of Rules for Autism Detection," *Health Informatics Journal*, vol. 26, no. 1, pp. 264-286, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Fadi Thabtah, "Machine Learning in Autistic Spectrum Disorder Behavioral Research: A Review and Ways Forward," *Informatics for Health and Social Care*, vol. 44, no. 3, pp. 278-297, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[7] Elizabeth Stevens et al., "Identification and Analysis of Behavioral Phenotypes in Autism Spectrum Disorder via Unsupervised Machine Learning," *International Journal of Medical Informatics*, vol. 129, pp. 29-36, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] Fadi Thabtah, "An Accessible and Efficient Autism Screening Method for Behavioural Data and Predictive Analyses," *Health Informatics Journal*, vol. 25, no. 4, pp. 1739-1755, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Naseer Ahmed Khan et al., "A Three-Stage Teacher, Student Neural Networks and Sequential Feed Forward Selection-Based Feature Selection Approach for the Classification of Autism Spectrum Disorder," *Brain Sciences*, vol. 10, no. 10, pp. 1-22, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Irene Chaidi, and Athanasios Drigas, "Autism, Expression, and Understanding of Emotions: Literature Review," *International Journal of Online and Biomedical Engineering*, vol. 16, no. 2, pp. 94-111, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Mokhlesur Rahman et al., "A Review of Machine Learning Methods of Feature Selection and Classification for Autism Spectrum Disorder," *Brain Sciences*, vol. 10, no. 12, pp. 1-23, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Yuanrui Zheng, Tingyan Deng, and Yaozheng Wang, "Autism Classification Based on Logistic Regression Model," *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering*, Nanchang, China, pp. 579-582, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] R. Sujatha et al., "A Machine Learning Way to Classify Autism Spectrum Disorder," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 6, pp. 182-200, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Kaushik Vakadkar, Diya Purkayastha, and Deepa Krishnan, "Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques," *SN Computer Science*, vol. 2, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Nurul Amirah Mashudi, Norulhusna Ahmad, and Norliza Mohd Noor, "Classification of Adult Autistic Spectrum Disorder Using Machine Learning Approach," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 3, pp. 743-751, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Md Delowar Hossain et al., "Detecting Autism Spectrum Disorder Using Machine Learning Techniques," *Health Information Science and Systems*, vol. 9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Muhammad Kashif Hanif et al., "Employing Machine Learning-Based Predictive Analytical Approaches to Classify Autism Spectrum Disorder Types," *Complexity*, vol. 2022, no. 1, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Mousumi Bala et al., "Efficient Machine Learning Models for Early Stage Detection of Autism Spectrum Disorder," *Algorithms*, vol. 15, no. 5, pp. 1-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Mengyi Liao, Hengyao Duan, and Guangshuai Wang, "Application of Machine Learning Techniques to Detect the Children with Autism Spectrum Disorder," *Journal of Healthcare Engineering*, vol. 2022, no. 1, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Muhammad Shoaib Farooq et al., "Detection of Autism Spectrum Disorder (ASD) in Children and Adults Using Machine Learning," *Scientific Reports*, vol. 13, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Dhuha Dheyaa Khudhur, and Saja Dheyaa Khudhur, "The Classification of Autism Spectrum Disorder by Machine Learning Methods on Multiple Datasets for Four Age Groups," *Measurement: Sensors*, vol. 27, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Chelsea M. Parlett-Pelleriti et al., "Applications of Unsupervised Machine Learning in Autism Spectrum Disorder Research: A Review," *Review Journal of Autism and Developmental Disorders*, vol. 10, pp. 406-421, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Roberta Simeoli et al., "Using Machine Learning for Motion Analysis to Early Detect Autism Spectrum Disorder: A Systematic Review," *Review Journal of Autism and Developmental Disorders*, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] D. Aarthi, and S. Kannimuthu, "A Comprehensive Analysis of Autism Spectrum Disorder Using Machine Learning Algorithms: Survey," *Proceedings of Power Engineering and Intelligent Systems*, vol. 1, pp. 241-253, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] S. Jayanthi et al., "Mental Health Status Monitoring for People with Autism Spectrum Disorder Using Machine Learning," *International Journal of Information Technology*, vol. 16, pp. 43-51, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Fadi, Autism Screening Data for Toddlers, Kaggle, 2018. [Google Scholar] [Publisher Link]

[27] Faizunnabi, Autism Screening, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/faizunnabi/autism-screening

[28] Fadi Thabtah, "Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment," *Proceedings of the 1st International Conference on Medical and Health Informatics*, Taichung City, Taiwan, pp. 1-6, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[29] Jerome H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189-1232, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[30] Tianqi Chen, and Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, USA, pp. 785-794, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[31] Yoav Freund, and Robert E Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119-139, 1997. [CrossRef] [Google Scholar] [Publisher Link]

[32] Guolin Ke et al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 3149-3157, 2017. [Google Scholar] [Publisher Link]