Manuscript_Similarity check

By Rashid Ayub

ABSTRACT

Diagnosing Autism Spectrum Disorder (ASD) presents a multifaceted challenge, demanding accurate and efficient screening methods. Applying machine learning techniques offers a promising avenue for enhancing diagnostic accuracy and efficiency. This research investigates the efficiency of machine learning in distinguishing individuals with ASD from those without, utilizing a comprehensive dataset comprising screening questions, demographic factors, and ASD related diagnostic classifications. We applied chi-square feature selection technique and also tested Random Forest, Logistic Regression, Gradient Boosting Classifier, and Extra Trees Classifier. Each model showed optimal performance and exhibit high precision, recall, and F1-score for both ASD-positive and ASD-negative instances. Additionally, AUROC curves further validated the models' exceptional discriminatory abilities, with exceptional results. Our findings highlight the potential of machine learning algorithms for enhancing ASD diagnosis accuracy and efficiency in clinical settings. Further research and validation on larger datasets are required to understand the importance of machine learning methods in ASD diagnosis.

Keywords

Autism Spectrum Disorder (ASD); Machine Learning; Behavioral Patterns; Neurological and Developmental Disorders, Classification; Diagnosis.

Abbreviations: Autism Spectrum Disorder (ASD); Machine Learning (ML), Deep Learning (DL), Autism and Developmental Disabilities Monitoring (ADDM), Attention-Deficit/Hyperactivity Disorder (ADHD), Convolutional Neural Network (CNN), True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Negative Predictive Value (NPV), Area under the Receiver Operating Characteristic curve (AUROC)

1. Introduction

Autism spectrum disorder (ASD) is a neurological and developmental disorder that has an impact on the social and cognitive skills of children causing social interaction and communication challenges, repetitive behaviors, sensory issues and restricted interests (Faja and Dawson, 2017). Autism is described as a "developmental disorder" because symptoms generally appear in the first 2 years of life (Goldstein &Ozonoff, 2018). Autism is also known as a "spectrum" disorder it encompasses a diverse range of conditions and varying severity of symptoms experienced by individuals (Baio, 2018). The spectrum is not measured

on scale from less autistic to more severe autistic; rather, it denotes how individuals with autism function across various areas, as described in Figure 1. Although ASD can be a lifelong disorder however, treatments and services have the potential to improve a person's symptoms and daily functioning (Ofner et al., 2018). Diagnosing ASD in adults is often more difficult than in children because some ASD symptoms can overlap with symptoms of other mental health disorders, such as anxiety disorder or attention-deficit/hyperactivity disorder (ADHD) (Zwaigenbaum and Penner, 2018). Research findings indicate that approximately 2 in 100 children receive a diagnosis of ASD when examining prevalence (Campisi et al., 2018). However, reported prevalence rates vary significantly between studies due to differences in diagnostic criteria, research methodologies, and the demographics of the populations studied (Kogan et al., 2018). This variability emphasizes the importance of careful interpretation when assessing ASD prevalence, considering factors such as geographical location, access to healthcare, and cultural perspectives on autism (Goel et al., 2018).Addressing the complexities of ASD involves navigating a multifaceted landscape of challenges that individuals with ASD and their caregivers confront daily (Lord et al., 2018). Alarmingly, statistics reveal that approximately 40% of individuals with autism are nonverbal, exacerbating the challenges they encounter in communication and expression (Erkan and Thanh, 2019). Moreover, the presence of intellectual disabilities further compounds the situation, with 31% of children grappling with significant challenges in daily functioning, and another 25% falling within the borderline range of intellectual abilities (Thabtah, 2019).Furthermore, the prevalence of self-injurious behaviors among individuals with ASD, such as head banging and skin scratching, underscores the urgency of addressing the complex needs and vulnerabilities within this population (Andrews et al., 2019).Perhaps most distressing is the sobering reality that drowning remains a leading cause of death for children with autism, accounting for approximately 90% of fatalities associated with wandering or bolting among those aged 14 and younger (Christensen, 2019). These statistics underscore the imperative to develop effective interventions and support systems that can mitigate risks and enhance the quality of life for individuals with ASD and their families (Hirota and King, 2023). To expedite the diagnostic process and minimize formal evaluations, machine learning ASD tools have been developed (Rahman et al., 2020). These machine learning techniques offer quick and accurate assessment of ASD risk, facilitating faster access to critical therapies for families (Alkahtani et al., 2023). However, as machine learning advances ASD diagnosis and treatment, ethical implementation, and privacy considerations remain paramount.
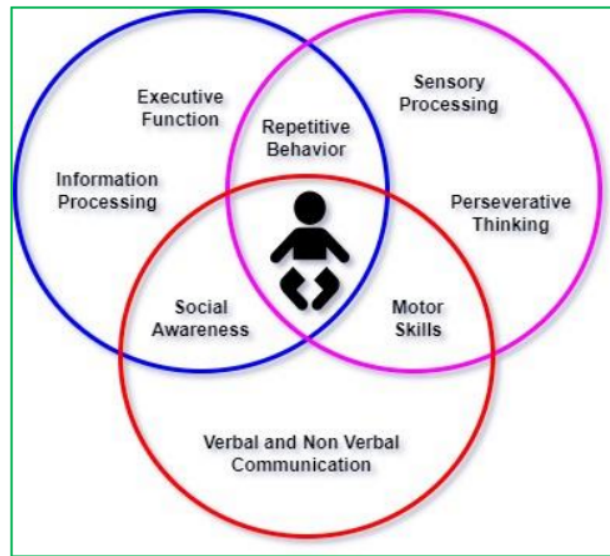
2

Fig.1. Diagnostic Framework for Autism Spectrum Disorder (ASD)

2. Related work.

The diagnosis and intervention of ASD have become pressing issues due to the increasing prevalence of the condition. Over the years, researchers have made significant strides in leveraging machine learning (ML) and deep learning (DL) techniques to enhance the detection and understanding of ASD. Bone et al. (2014) demonstrated the potential of ML in enhancing diagnostic and intervention research by analyzing large behavioral ASD datasets, leading to insights into the heterogeneous nature of ASD and the customization of interventions to individual needs. Maenner et al. (2020) furthered this progress by developing a ML approach for classifying case status in Autism and Developmental Disabilities Monitoring (ADDM), using language features from children's evaluations. Tejwani et al. (2017) extended the scope by employing ML techniques on resting-state brain imaging data, achieving optimal accuracy rates in distinguishing ASD subjects from controls based solely on functional connectivity variability. Building upon these foundations, Tariq et al. (2018) developed a mobile detection system for ASD using ML on home videos, offering a potential solution to accelerate diagnosis and reach underserved populations. Subsequent studies by Jacob et al. (2019) explored the complexities of ASD diagnosis, leveraging supervised and unsupervised ML models, as well as deep learning networks, to enhance classification accuracy. Raj and Masood (2020) demonstrated the superior performance of CNN-based prediction models in ASD detection, particularly in handling missing values. In parallel, Okoye et al. (2023) developed predictive models for early ASD detection, highlighting

3

logistic regression's accuracy among the evaluated models. Eslami et al. (2021) further expanded the scope by exploring ML and DL techniques for diagnosing both ASD and ADHD, addressing challenges such as overfitting and data distribution. These collective efforts underscore the transformative potential of ML and DL in enhancing ASD diagnosis and understanding.

In recent studies, Elshoky et al. (2022) explored the application of ML techniques for creating predictive models using facial images of children, while Moridian et al. (2022) conducted a comprehensive review of automated ASD detection methods using Magnetic resonance imaging (MRI) neuroimaging and AI. Bahathiq et al. (2022) provided a comprehensive review emphasizing the ongoing need for advancements in ASD diagnosis using ML, addressing challenges such as limited sample sizes and highlighting trends like leveraging big data and advanced imaging techniques. Meanwhile, Shinde and Patil (2023) proposed methods to improve ASD diagnosis through a multi-classifier recommender model, underscoring the importance of dataset size in enhancing accuracy. Collectively, these studies contribute to advancing ASD detection and diagnosis, offering new insights and methodologies for early intervention. Additionally, Alkahtani, Aldhyani, andAlzahrani (2023) developed an expert system for ASD identification based on facial landmarks of children, furthering the capabilities of computer vision in ASD detection, thus addressing critical needs in healthcare.These studies underscore the value of ML in advancing our understanding of ASD and suggest promising avenues for collaborative research at the intersection of computational and behavioral science.

3. Methodological Approach

This research methodology presents a thorough framework for diagnosing ASD. The dataset utilized is sourced from the UC Irvine Machine Learning Repository (Thabtah, 2017). The missing values within the dataset are represented by '?' and to address them we applied imputation techniques substituting missing values with the mean, median, or mode of the corresponding attribute. Furthermore, encoded techniques such as one-hot encoding are applied to transform categorical variables into numerical representations, enabling machine learning algorithms to process the data effectively. These systematic procedures contribute to the refinement and uniformity of the dataset, thereby fostering precise analysis.

Further, we applied the chi-square feature selection test to check a significant association between categorical variables (Semary et al., 2024). We then employed a range of machine

learning algorithms, such as linear regression, random forest, gradient boosting classifier and extra tree classifier chosen for their suitability in classification tasks and ability to handle diverse data types in our dataset.The ASD model was developed on 80% of the training dataset and tested on the remaining 20%. To validate the model, we conducted both cross-validation and external validation. Cross-validation was performed using 10-fold stratified sampling to ensure robustness and generalizability across multiple folds. Additionally, external validation with an independent dataset and clinical validation confirmed the efficacy of the model in diagnosing ASD, providing a more reliable estimate of its real-world performance.
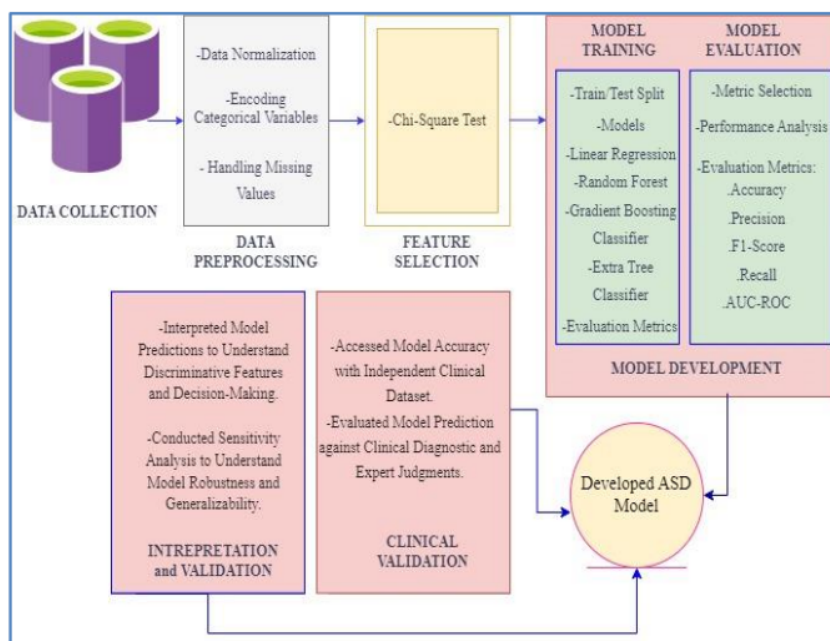
Figure 2: Methodological Approach for the Development of ASD Diagnostic Model

4. Dataset description

In this study, we leverage a dataset (as described in table 1) tailored for screening ASD in children, sourced from the UC Irvine Machine Learning Repository, a reputable platform housing dataset curated for machine learning research (Thabtah, 2017). Comprising a range of attributes pertinent to ASD diagnosis, the dataset includes scores for ten screening questions (A1 to A10), as well as information on age, gender, ethnicity, history of jaundice,

family history of autism, country of residence, prior use of ASD screening applications, and final diagnostic classification (Mahmoud et al., 2023). These attributes offer valuable insights into factors influencing ASD and its early detection.

The chosen features hold significance due to their relevance to ASD screening and diagnosis. The ten screening questions reflect behavioral patterns commonly associated with ASD. Age, gender, and ethnicity can shape the presentation of ASD symptoms, informing our understanding of its prevalence and manifestation across diverse demographic groups. Details such as jaundice history, family autism history, and prior screening application usage provide additional context on potential risk factors and previous diagnostic attempts. Lastly, the diagnostic classification acts as the outcome variable, facilitating the assessment of the screening tool's efficacy in identifying children with ASD.

Table 1. Illustrating the description of the ASD dataset features utilized in our research

Attribute	Type	Description
Age	Number	Years
Gender	String	Male or Female
Ethnicity	String	List of common ethnicities in text format
Born with jaundice	Boolean (yes or no)	Whether the case was born with jaundice
Family member with PDD	Boolean (yes or no)	Whether any immediate family member has a PDD
Who is completing the test	String	Parent, self, caregiver, medical staff, clinician, etc.
Country of residence	String	List of countries in text format
Used app before	Boolean (yes or no)	Whether the user has used a screening app
Screening Method Type	Integer (0,1,2,3)	The type of screening methods chosen based on age category (0=toddler, 1=child, 2= adolescent, 3= adult)
Question 1 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Question 2 Answer	Binary (0, 1)	The answer code of the question based on the screening method used

Question 3 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Question 4 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Question 5 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Question 6 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Question 7 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Question 8 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Question 9 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Question 10 Answer	Binary (0, 1)	The answer code of the question based on the screening method used
Screening Score	Integer	The final score obtained based on the scoring algorithm of the screening method used. This was computed in an automated manner

5. Feature Selection and Performance Measures in ASD Diagnosis

5.1 Simplifying ASD Analysis with Chi-Square Test

The chi-square test is a statistical method used to determine the presence of a significant association between two categorical variables by comparing observed frequencies with expected frequencies, assuming no connection between the variables. It evaluates the null hypothesis (H_0) which imagines that no relationship exists between the categorical variables in the data, implying they are independent, while the alternative hypothesis (H_1) suggests an association between the variables. In this research, we used chi-square test in the opposite direction, where significance level of $p < 0.05$ indicates a rejection of the null hypothesis, implying a dependency between the variables. Essentially, a lower probability threshold ($p < 0.05$) is chosen to detect any significant associations between the categorical variables.

In this research, variables like gender and used_app_before were excluded from further analysis based on chi-square test results. For instance, the chi-square value for gender was

7

low (0.004), with a corresponding p-value of 0.943, significantly higher than 0.05, indicating no statistically significant association with autism. Similarly, the chi-square statistic for "used_app_before" was moderate (1.05), and its p-value (0.300) exceeded 0.05, affirming no statistically significant association between this variable and autism (shown in table 2).

Table2. Chi-Square Test Results for Different ASD Features

	Category	Chi-Square	p-value	P<0.05
0	Gender	0.004985	8.9568808e-01	False
1	Jaundice	11.978454	5.875686e-04	True
2	Autism	84.527867	1.526148e-20	True
3	used_app_before	1.042585	3.001534e-01	False

In this case, since the p-values for all variables are significantly less than 0.05, it indicates that there is a highly significant association between each of these variables (A1_Score, A2_Score, A3_Score, A4_Score, A5_Score, A6_Score, A7_Score, A8_Score, A9_Score) and the outcome variable. This suggests that these variables are important predictors or indicators of the outcome to autism in the dataset being analyzed.

Table3. Chi-Square Test Results for A-Score Categories of ASD

	Category	Chi-Square	p-value	P<0.05
0	A1_Score	31.780269	1.726377e-08	True
1	A2_Score	52.239554	4.912658e-13	True
2	A3_Score	93.479438	4.104424e-22	True
3	A4_Score	117.490026	2.242249e-27	True
4	A5_Score	103.205602	3.021106e-24	True
5	A6_Score	164.237125	1.342643e-37	True

6	A7_Score	98.621908	3.056194e-23	True
7	A8_Score	38.351519	5.908208e-10	True
8	A9_Score	86.734950	1.240845e-20	True

5.2 Model EvaluationMeasures

We used various performance measures to ensure that developed ASD diagnostic model works effectively. These measures collectively guide us in refining and optimizing the model for accurate and reliable diagnosis.

a. **True Positive (TP):**TP, also known as sensitivity, represents the cases where the model correctly predicts individuals with ASD. ATrue Positive indicates that the model correctly identifies individuals who have ASD, enabling timely intervention and support (Ahmed et al., 2022).

Sensitivity = TP / (TP + FN) equ. (1)

b. **True Negative (TN):**TN, also known as specificity, represents the cases where the model correctly predicts individuals without ASD. A True Negative suggests that the model accurately identifies individuals who do not have ASD, avoiding unnecessary interventions and treatments (Fayaz et al., 2023).

Specificity = TN / (TN + FP) equ. (2)

c. **False Positive (FP):** These are cases where the model incorrectly predicts individuals as having ASD when they do not have. False Positives may lead to unnecessary stress and concern for individuals and their families, as well as potential misallocation of resources for interventions (Maenner, 2020).

FP=No. of instances predicted as positive –TN   equ. (3)

d. **False Negative (FN):** These are cases where the model incorrectly predicts individuals as not having ASD when they actually have. False negatives are concerning because they may result in missed opportunities for early intervention and support, delaying diagnosis and treatment for individuals with ASD.

FN = No. of actual positive instances – No. of instances predicted as negative   equ. (4)

e. **Positive Predictive Value (Precision):**Precision measures the proportion of individuals predicted to have ASD who actually have the disorder. High precision ensures that individuals identified as having ASD are more likely to truly have the disorder, reducing the risk of over diagnosis (Atlam et al., 2024).

Precision = TP / (TP + FP) equ. (5)

f. **Negative Predictive Value (NPV):**NPV measures the proportion of individuals predicted not to have ASD who truly do not have the disorder. NPV reflects the model's ability to rule out ASD for individuals without the disorder, complementing the information provided by sensitivity and specificity.

NPV = TN / (TN + FN) equ. (6)

g. **Area under the Receiver Operating Characteristic curve (AUROC):**AUROC evaluates the model's ability to distinguish between individuals with and without ASD across different decision thresholds. A higher AUROC value indicates better discrimination ability, with 1 representing perfect discrimination and 0.5 representing random chance.

6.   ASD Diagnosis Advancement ThroughMachine Learning Models

In our research on ASD disease diagnosis, we divided the dataset into a training set comprising 80% of the data and a testing set containing 20% of the data. Due to the limited number of instances in the dataset, we applied 10-fold cross-validation on the training set. This method involves dividing the data into k equal-sized folds, training the model on k-1 folds, and validating it on the remaining fold. This process is repeated k times, ensuring robust performance evaluation through averaging the results.

6.1   Random Forest Model

Random forest is an ensemble learning technique that consists of multiple decision trees. Each tree is trained on a random subset of the training data and a random subset of features. Through a voting mechanism, predictions from individual trees are aggregated, resulting in a final prediction. This ensemble approach enhances prediction accuracy and robustness (Abdelwahabet al., 2024). We applied Random Forest to our ASD dataset and obtained the results described in table 4. The developed random forest model's performance is measured using various evaluation metrics. These metrics provide valuable insights into the model's

effectiveness in distinguishing between instances of "absence of ASD" and "presence of ASD."

Table 4: Random Forest Model Performance Metrics

Class	Precision	Recall	F1-Score	Support
0	0.96	0.98	0.97	179
1	0.98	0.96	0.97	200
Accuracy			0.97	379
Macro Avg.	0.97	0.97	0.97	379
Weighted Avg.	0.97	0.97	0.97	379

- Precision Analysis: Our analysis revealed a precision of 0.96 for the class "absence of ASD," indicating that 96% of instances predicted as not having ASD were indeed true negatives. Conversely, the precision for the class "presence of ASD" was 0.98, demonstrating that 98% of instances predicted as having ASD were true positives.

- Recall Assessment: In terms of recall, the model achieved a recall of 0.98 for the class "absence of ASD," correctly identifying 98% of all instances not associated with ASD. For the class "presence of ASD," the recall was 0.96, indicating that 96% of all actual ASD instances were correctly identified by the model.

- F1-Score Evaluation: The F1-score, a balanced measure of precision and recall, yielded scores of 0.97 for both classes. This highlights the model's ability to minimize both false positives and false negatives in ASD diagnosis, contributing to its overall effectiveness.

- Support Analysis: Examining the support values revealed 179 instances of "absence of ASD" and 200 instances of "presence of ASD" within the dataset. These support values provide insights into the distribution of data and the prevalence of each class.

- Accuracy Examination: Our model achieved an overall accuracy of 97%, indicating its high level of correctness in predicting both classes collectively. This underscores the model's reliability in accurately diagnosing ASD across the dataset.

- Average Metrics Interpretation: Additionally, both macro and weighted averages yielded consistent scores of 0.97, reflecting the model's consistent performance across all classes. These average metrics provide a comprehensive overview of the model's effectiveness in ASD diagnosis.

We also conducted an assessment of the AUROC curve to evaluate the performance of therandom forest model in the context of ASD diagnosis (Fig. 3). The AUROC curve, with an AUC= 1.00, underscores the model's exceptional capability to accurately differentiate between individuals with ASD and those without ASD across various thresholds of discrimination. This perfect AUC value signifies the model's ability to achieve a true positive rate of 1 (sensitivity) and a false positive rate of 0 (specificity), thereby ensuring precise predictions for all instances in the dataset. In practical terms, the flawless performance of the Random Forest model in identifying individuals with ASD while minimizing misclassifications of non-ASD cases highlights its robust reliability and efficacy for ASD diagnosis.
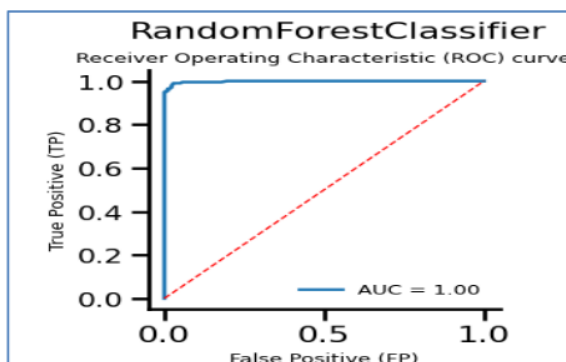

Fig.3. AUROC Curve for ASD Classification using Random Forest

Overall, these findings demonstrate the strong performance of the random forest machine learning model in accurately diagnosing ASD. With high precision, recall, F1-score, support, accuracy, and average metrics the model showcases promising potential for ASD diagnosis in clinical settings.

6.2   Logistic Regression Model

Logistic regression is a linear binary classification model which gives the probability that an instance belongs to a particular class using a logistic (sigmoid) function. Logistic regression estimates the coefficients for each feature, which represent the impact of that feature on the log-odds of the target class. Logistic Regression could be used to model the probability of a child having ASD based on the given features (Ansarullah et al., 2024). We applied logistic regression to our ASD dataset and obtained the results described in table 5.

Table 5: Logistic Regression Model Performance Metrics

Class	Precision	Recall	F1-Score	Support
0	0.96	0.96	0.96	179
1	0.96	0.96	0.96	200
Accuracy			0.96	379
Macro Avg.	0.96	0.96	0.96	379
Weighted Avg.	0.96	0.96	0.96	379

- Precision: The precision metric reflects the proportion of true positive predictions out of all positive predictions made by the model. For class 0 (indicating the absence of ASD), the precision is 0.96, indicating that 96% of instances predicted as not having ASD were indeed true negatives. Similarly, for class 1 (indicating the presence of ASD), the precision is also 0.96, demonstrating that 96% of instances predicted as having ASD were true positives.

- Recall: Recall, also known as sensitivity, measures the model's ability to correctly identify actual positive instances. With a recall of 0.96 for both class 0 and class 1, the model accurately identified 96% of all instances not associated with ASD and 96% of all actual ASD instances, respectively.

- F1-Score: The F1-score, a harmonic mean of precision and recall, provides a balanced measure of the model's accuracy. Both class 0 and class 1 exhibit F1-scores of 0.96, indicating a high level of precision and recall for both classes.

- Support: Support denotes the number of actual occurrences of each class in the dataset. In this case, there are 179 instances of class 0 and 200 instances of class 1, contributing to a total of 379 instances.

- Accuracy: Accuracy measures the overall correctness of the model's predictions. With an accuracy of 0.96, the model achieved a high level of correctness in predicting both classes collectively.

- Macro and Weighted Averages: Both macro and weighted averages yield consistent scores of 0.96 for precision, recall, and F1-score, reflecting the model's consistent performance across all classes.

We also conducted an assessment of the AUROC curve to evaluate the performance of the logistic regression model in the context of ASD diagnosis (shown in figure 4). The AUROC curve plots the true positive rate (sensitivity) against the false positive rate (1 - specificity) for

different threshold values. The AUC= 0.99 implies that the logistic regression model performed remarkably well in correctly identifying individuals with ASD while minimizing misclassifications of non-ASD cases. The high AUC value underscores the model's reliability and effectiveness in ASD diagnosis, making it a valuable tool for clinical decision-making and patient care.
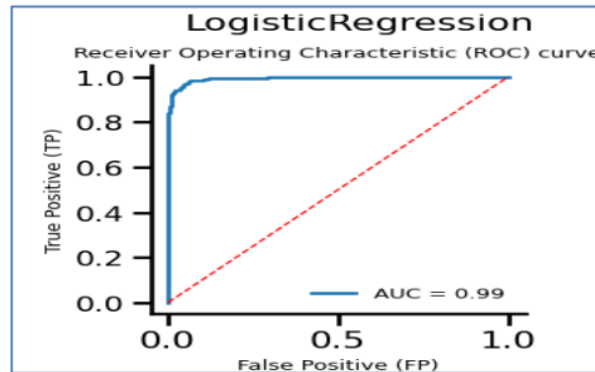

Fig.4. AUROC Curve for ASD Classification using Logistic Regression

6.3 Gradient Boosting Classifier

The Boosting Classifier is a machine learning algorithm belonging to the ensemble learning methods. It operates by sequentially combining multiple weak learners, typically decision trees with shallow depth, in a sequential manner to create a strong predictive model. In this method, each weak learner is trained to rectify errors made by preceding ones. This is accomplished by fitting each new model to the residual errors of the ensemble, progressively diminishing the overall prediction error. The algorithm iteratively minimizes a loss function using gradient descent optimization, thereby adjusting the parameters of each weak learner to enhance the model's performance (Ansarullah et al., 2022). We applied gradient boosting classifier to our ASD dataset and obtained the results described in Table 6.

Table6. Gradient Boosting Model Performance Metrics

Class	Precision	Recall	F1-Score	Support
0	0.97	0.98	0.98	179
1	0.98	0.97	0.98	200
Accuracy			0.98	379
Macro Avg.	0.98	0.98	0.98	379

14

Weighted Avg.	0.98	0.98	0.98	379

- Precision: For class 0 (representing the absence of ASD), the precision was 0.97, indicating that 97% of instances predicted as not having ASD were indeed true negatives. Conversely, for class 1 (indicating the presence of ASD), the precision was 0.98, signifying that 98% of instances predicted as having ASD were true positives.

- Recall: The recall for class 0 was 0.98, suggesting that 98% of all instances not associated with ASD were correctly identified by the model. For class 1, the recall was 0.97, indicating that 97% of all actual ASD instances were accurately identified by the model.

- F1-Score: The F1-score, a harmonic mean of precision and recall, yielded scores of 0.98 for both class 0 and class 1. This balanced measure reflects the model's ability to minimize both false positives and false negatives in ASD diagnosis.

- Support: The dataset comprised 179 instances of class 0 and 200 instances of class 1, providing insights into the distribution of data and the prevalence of each class.

- Accuracy: The model achieved an overall accuracy of 98%, demonstrating its high level of correctness in predicting both classes collectively.

- Macro and Weighted Averages: Both macro and weighted averages yielded consistent scores of 0.98 for precision, recall, and F1-score, indicating the model's consistent performance across all classes.

In our analysis utilizing the Gradient Boosting Classifier on the ASD dataset, we generated an AUROC with an exceptional AUC value of 1.00 which demonstrates the model's remarkable ability to accurately distinguish between individuals with ASD and those without ASD (shown in figure 5). In summary, the AUROC curve with an AUC value of 1.00 generated by the Gradient Boosting Classifier underscores the model's exceptional reliability and accuracy in ASD diagnosis, making it a highly promising tool for clinical applications.
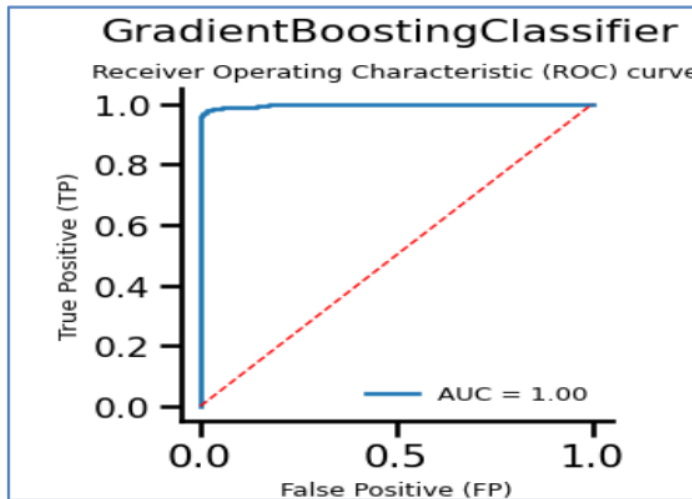
Fig.5. AUROC Curve for ASD Classification using Gradient Boosting

6.4 Extra Trees Classifier

The Extra Trees Classifier, short for Extremely Randomized Trees Classifier, is a machine learning algorithm that belongs to the ensemble learning family. In Extra Trees Classifier, rather than selecting the best split among a subset of features, as in Random Forest, it randomly selects candidate splits for each feature. This randomness during the tree building process leads to extra diversity among the trees in the ensemble, hence the name "Extra Trees."By introducing more randomness into the tree construction process, Extra Trees Classifier reduces the variance of the model at the expense of a slight increase in bias. This can sometimes lead to improved generalization performance, especially when dealing with noisy or high-dimensional datasets (Ansarullah& Kumar, 2019). This research applied the Extra Trees Classifier model to analyse the ASD dataset, from which we derived the following results as described in table 7.

Table7. Extra Tree Model Performance Metrics

Class	Precision	Recall	F1-Score	Support
0	0.97	0.98	0.98	179
1	0.98	0.97	0.98	200
Accuracy			0.98	379
Macro Avg.	0.98	0.98	0.98	379
Weighted Avg.	0.98	0.98	0.98	379

16

- Precision: The precision metric indicates the proportion of true positive predictions out of all positive predictions made by the model. For class 0 (representing the absence of ASD), the precision is 0.97, suggesting that 97% of instances predicted as not having ASD were indeed true negatives. Similarly, for class 1 (indicating the presence of ASD), the precision is 0.98, indicating that 98% of instances predicted as having ASD were true positives.

- Recall: Recall, also known as sensitivity, measures the proportion of true positive predictions out of all actual positive instances in the dataset. A recall of 0.98 for class 0 means that 98% of all instances not associated with ASD were correctly identified by the model. Conversely, a recall of 0.97 for class 1 indicates that 97% of all actual ASD instances were correctly identified by the model.

- F1-Score: The F1-score, a harmonic mean of precision and recall, provides a balanced assessment of the model's performance. With F1-scores of 0.98 for both classes, the model demonstrates a high level of accuracy in distinguishing between instances of "absence of ASD" and "presence of ASD."

- Support: Support refers to the number of actual occurrences of each class in the dataset. In this case, there are 179 instances of class 0 and 200 instances of class 1, indicating a relatively balanced distribution between the two classes.

- Accuracy: Accuracy measures the overall correctness of the model's predictions. With an accuracy of 98%, the Extra Trees Classifier model achieved a high level of correctness across all predictions, further emphasizing its effectiveness in ASD diagnosis.

- Macro and Weighted Averages: Both macro and weighted averages yielded consistent scores of 0.98 for precision, recall, and F1-score, reflecting the model's consistent performance across all classes and accounting for the distribution of instances within each class.

In our analysis utilizing the Extra Tree Classifier model on the ASD dataset, we generated the AUROC with an AUC value of 1.00 (shown in figure 6) which indicates perfect discrimination by the model, making it highly reliable for ASD diagnosis. Overall, these results highlight the robust performance of the Extra Trees Classifier model in accurately diagnosing ASD, with high precision, recall, F1-score, support, accuracy, and average metrics across both classes.
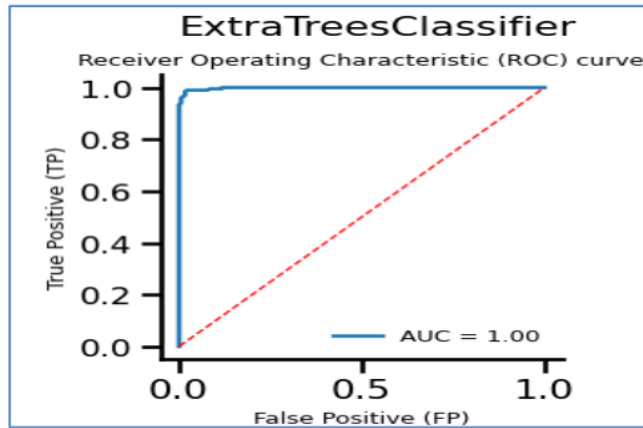
Figure 6: AUROC Curve for ASD Classification using Extra Tree Classifier

7. Results and Discussions

In our research on ASD disease diagnosis, we employed several machine learning algorithms and analysed their performance metrics to ascertain their effectiveness in distinguishing between individuals with and without ASD.The Random Forest model demonstrated robust performance, achieving high precision, recall, and F1-score for both classes. This indicates its proficiency in accurately identifying instances of "absence of ASD" and "presence of ASD." With an impressive overall accuracy of 97%, the model holds promise for enhancing ASD diagnosis accuracy. The model's exceptional capability to differentiate between ASD-positive and ASD-negative individuals, as evidenced by the AUROC curve with an AUC value of 1.00, signifies its potential to significantly impact clinical decision-making.

Similarly, the Logistic Regression model showcased notable precision, recall, and F1-score for both classes, along with an overall accuracy of 96%. Its high discriminative accuracy, validated by the AUROC curve with an AUC value of 0.99, suggests its potential utility in aiding clinicians in identifying individuals with ASD.The Gradient Boosting Classifier and Extra Trees Classifier exhibited strong performances also with an overall accuracy of 98% respectively. Their flawless performance in distinguishing between individuals with ASD and those without ASD, as indicated by the AUROC curve with AUC values of 1.00, underscores their reliability and effectiveness in ASD diagnosis.

18

The results indicate that all developed machine learning models achieved high performance in accurately diagnosing ASD, with precision, recall, and F1-score consistently above 0.95 for both classes. The models also demonstrated high overall accuracy, reflecting their reliability in predicting ASD across the dataset.The AUROC curves further validated the models' effectiveness, with AUC values close to or equal to 1.00, indicating that the models can make precise predictions while minimizing misclassifications, highlighting their potential for clinical applications in ASD diagnosis.

These findings hold significant implications for ASD diagnosis in clinical settings. By leveraging machine learning algorithms with high predictive accuracy, clinicians can make more informed decisions, leading to earlier detection and intervention for individuals with ASD. Moreover, the precise identification of ASD cases can facilitate the allocation of resources and support services, ultimately improving outcomes for individuals and their families. Furthermore, the utilization of these advanced algorithms can streamline the diagnostic process, potentially reducing wait times and alleviating the burden on healthcare systems. Overall, the findings suggest that machine learning models, particularly ensemble methods like Random Forest, Gradient Boosting Classifier, and Extra Trees Classifier, hold promise for enhancing ASD diagnosis accuracy and efficiency in clinical settings. Further research and validation on larger datasets are warranted to fully assess their clinical utility and impact.

8. Conclusion and Future Work

Machine learning techniques show promise in enhancing ASD diagnosis accuracy. Various models, such as Random Forest, Logistic Regression, Gradient Boosting Classifier, and Extra Trees Classifier, were applied to a comprehensive dataset. The results showed high performance, with AUROC curves validating their discriminatory abilities. Particularly, the Random Forest model stood out, achieving 97% accuracy and a perfect AUROC curve, emphasizing the potential of machine learning algorithms in enhancing ASD diagnosis. Leveraging these algorithms could facilitate earlier detection and intervention, thereby improving outcomes for individuals and families affected by ASD. Furthermore, the adoption of advanced algorithms could streamline the diagnostic process and contribute to a better understanding of ASD epidemiology.However, we acknowledge potential biases or limitations in our methodology, necessitating further research on larger datasets and fostering collaboration between researchers and clinicians. Future plans involve developing a

comprehensive ASD model that integrates genetic, behavioral, and medical data for practical use in clinical and research settings, with continuous monitoring and updating, based on new data and research. This model will encompass all available data to develop a unified solution and will be deployed for practical use in clinical and research settings. Continuous monitoring and updating of the model based on new data and emerging research will be essential for maintaining its effectiveness.

Manuscript_Similarity check

9	arxiv.org Internet	27 words — < 1%
10	dataaspirant.com Internet	27 words — < 1%
11	thesis.unipd.it Internet	23 words — < 1%
12	docs.google.com Internet	22 words — < 1%
13	www.ijisae.org Internet	21 words — < 1%
14	bmcmedinformdecismak.biomedcentral.com Internet	18 words — < 1%
15	ijsred.com Internet	18 words — < 1%
16	www.frontiersin.org Internet	16 words — < 1%
17	Nesrine Kadri, Ameni Ellouze, Mohamed Ksantini, Sameh Hbaieb Turki. "New LSTM Deep Learning Algorithm for Driving Behavior Classification", Cybernetics and Systems, 2022 Crossref	15 words — < 1%
18	lwrs.inl.gov Internet	15 words — < 1%

EXCLUDE QUOTES   OFF   EXCLUDE SOURCES   < 15 WORDS