Healthcare Analytics 7 (2025) 100379

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

An exploration of machine learning approaches for early Autism Spectrum Disorder detection

Nawshin Haque, Tania Islam, Md Erfan *

Department Of CSE, University of Barishal, Karnakathi, Barishal, 8200, Bangladesh

ARTICLE INFO

Keywords:
Autism Spectrum Disorder
Logistic Regression
Support Vector Classifier
K-Nearest Neighbour
Decision Tree
Random Forest

ABSTRACT

Autism Spectrum Disorder is a neurodevelopmental condition impacting an individual's repetitive behaviours, social skills, verbal and nonverbal communication abilities, and capacity for acquiring new knowledge. Manifesting typically in early childhood, specifically between 6 months and 5 years, the symptoms of autism exhibit a progressive nature over time. This study explores the application of Logistic Regression, Support Vector Classifier, K-Nearest Neighbour, Decision Tree, and Random Forest for predicting Autism in children and toddlers by leveraging advancements in machine learning. The efficacy of these techniques is evaluated using publicly accessible datasets specific to both age groups. The findings indicate remarkable performance, with the toddler dataset achieving a mean Intersection over Union (mIoU) of 100% for Support Vector Classifier and 99.80% for Logistic Regression. Similarly, the children dataset demonstrates outstanding results, achieving an mIoU of 100% for Support Vector Classifier and 99.96% for Logistic Regression. Furthermore, all algorithms achieved 100% accuracy on the children (age 4–11) dataset collected from real-world sources. Logistic Regression, Random Forest, Support Vector Classifier, and Decision Tree attained 100% accuracy and mIoU with the real-world dataset. These results underscore the potential of machine learning in aiding the early detection of ASD in children and toddlers, offering promising avenues for future research and clinical applications.

1. Introduction

Autism Spectrum Disorder (ASD) is a term used to describe a constellation of early-appearing social communication deficits and repetitive sensory–motor behaviours associated with a strong genetic component as well as other causes [1,2]. Individuals with ASD may exhibit a diverse array of symptoms, and the severity of these symptoms can vary widely, hence the term "spectrum". Common features of ASD include difficulties in understanding social cues, challenges in forming and maintaining relationships, repetitive movements or speech patterns, and a preference for routine and sameness [3,4]. Additionally, sensory sensitivities, such as heightened reactions to sounds, lights, or textures, are also frequently observed [5,6]. These symptoms underscore the importance of early diagnosis and personalized intervention to support individuals with ASD.

In recent years, epidemiological studies have shown a rapid increase in the prevalence of ASD. According to the Centers for Disease Control and Prevention's (CDC) most recent study, one in 36 children has an autistic diagnosis as of 2023. From one in 44 children two years ago, this represents an increase [7]. ASD is thought to affect 1 in 160 children in the Southeast Asian area. ASD affects approximately 2 out of every 1000 children in Bangladesh, according to recent data from Bangabandhu Sheikh Mujib Medical University (BSMMU). Whereas, the prevalence of urban regions is greater than that of rural locations [8].

Autism manifests in diverse ways, with symptoms varying widely among individuals. Recognizing and understanding these manifestations are crucial for effective intervention and support. Common signs include [9–11], challenges in pragmatic language, often evident in early childhood, necessitating professional assessment for communication delays. Poor eye contact, difficulty interpreting facial expressions, gestures, and tone of voice are prevalent nonverbal communication struggles [4]. Some may exhibit repetitive behaviours like hand-flapping or engaging in ritualistic actions, becoming distressed with minor disruptions to routine [3]. Additionally, self-injurious behaviours, such as biting or head-banging, can occur [12]. Treatment approaches encompass medical, sensory, nutritional, and behavioural interventions [5]. Understanding these symptoms aids in recognizing and addressing the complexities of ASD.

Observation emerges as a more effective diagnostic method for identifying behavioural changes in children compared to autism-specific brain imaging. This emphasis on observation is crucial, given that

* Corresponding author.
E-mail address: merfan@bu.ac.bd (M. Erfan).

https://doi.org/10.1016/j.health.2024.100379
Received 31 December 2023; Received in revised form 8 December 2024; Accepted 25 December 2024
Available online 6 January 2025

observable behavioural changes indicative of ASD can be discerned in infants as young as 6 months old [13].

Early diagnosis and intervention have been shown to significantly improve the long-term outcomes and quality of life for individuals with ASD [14]. Early intervention, when began at an early age, results in greater increases in language development, social skills, and adaptive behaviours [15]. It also assists people with ASD in developing effective communication skills, which reduces communication difficulties [16]. Timely behavioural treatments, such as hostility and tantrums, can minimize problematic behaviours and improve overall quality of life [17]. Early intervention promotes the development of social skills, allowing people with ASD to interact positively and form meaningful connections [18]. Early intervention is cost-effective since it reduces the need for intensive interventions and special education services in the long run [19].

However, in the context of Bangladesh, access to specialized schools for autistic children is limited, with notable institutions like ‘Proyash’ [20] concentrated in the capital city, Dhaka. Unfortunately, such facilities are scarce in other cities, leaving children in those areas without adequate treatment and training opportunities, which represents a significant challenge for their development and future prospects. Traditional diagnostic methods relying on clinical assessment, while essential, are subjective and time-consuming, and access disparities to specialized healthcare services may lead to delayed diagnosis and intervention [21].

This research endeavors to achieve the primary objective of identifying autism at an early stage by utilizing screening tools and employing a range of machine learning algorithms. The specific algorithms considered in this study include Logistic Regression (LR), Support Vector Classifier (SVM), K-Nearest Neighbour (KNN), Decision Tree (DT), and Random Forest (RF).

Main contributions of this research are,

- **Model Evaluation**: Detailed evaluation of five classical machine learning algorithms for early ASD detection.
- **High Accuracy**: Achieved approximately 100% accuracy for the ASD Child dataset across all models; perfect accuracy for the toddler dataset with LR and SVM.
- **mIoU Performance**: Impressive mIoU scores, with SVM achieving 100% and LR achieving 99.80% for toddlers; 100% with SVM and 99.96% with LR for children.
- **Real-world Validation**: All classifiers demonstrated 100% accuracy and mIoU on real-world datasets.
- **Promising Algorithm**: Identified SVM as a particularly promising approach for diagnosing ASD.

Structured into distinct chapters, this study delves into a comprehensive review of relevant literature in Chapter 2, followed by an exploration of the methodology employed in Chapter 3. Chapter 4 presents the results obtained and their evaluation, while Chapter 5 offers a detailed discussion, synthesis of key insights, contributions, and outlines avenues for future work in the field.

2. Literature review

2.1. Overview of ASD detection using machine learning

Autism Spectrum Disorder is a neurological condition characterized by impairments in speech and behaviour. The rising prevalence of ASD highlights the critical need for early diagnosis, which can significantly improve outcomes for affected individuals. Machine learning (ML) has emerged as a powerful tool in ASD detection, leveraging health, behavioural, and demographic data to identify patterns indicative of the condition. Classical ML techniques, such as Support Vector Machines, Random Forest (RF), and Logistic Regression, have demonstrated significant potential, while advanced deep learning (DL) methods, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, are increasingly being utilized to tackle more complex datasets.

2.2. Machine learning-based approaches to ASD detection

Classical ML models have been widely applied in ASD research, particularly using structured datasets. Bone et al. [22] utilized SVMs to achieve 89.2% sensitivity and 59% specificity in detecting ASD traits across a wide age range (4–55 years). Allison et al. [23] employed the ‘Red Flags’ tool alongside Autism Spectrum Quotient, narrowing down to Autism Spectrum Quotient-10 (AQ-10) with over 90% accuracy for screening ASD in children and adults. Raj et al. [24] achieved 99.53%, 98.30%, and 96.88% accuracy for adult, child, and adolescent datasets using CNNs trained on AQ-10 responses. SH Hammed et al. [25] employed Gradient Boosting, achieving 87% accuracy, addressing limitations of existing tools by incorporating medical tests and demographic features.

Fadi Thabtah et al. [26] aligned their ASD screening model with DSM-5 criteria, emphasizing the need for updated diagnostic standards. Kazi Shahrukh Omar et al. [27] introduced a hybrid RF-CART and RF-ID3 model, achieving 97.1%, 93.78%, and 92.26% accuracy for adults, adolescents, and children, respectively. Mousumi Bala et al. [28] demonstrated SVM’s superior performance with 97.82% accuracy for toddlers and 99.61% for children using feature selection methods. Jyotismita Talukdar et al. [29] evaluated Naive Bayes, LR, SVM, and RF models, achieving 93.69% accuracy for toddlers and 90% for adolescents, showcasing the effectiveness of ML in non-clinical datasets.

Advanced DL methods also show potential in ASD detection. Jacob et al. [30] employed CNNs and Contextual Graph Recurrent Neural Networks (CGRNNs), extracting relevant information from UCI and real-time datasets. Zhao et al. [31] applied kinematic feature analysis with KNN, achieving 86.37% accuracy. Mostafa et al. [32] used fMRI data with LDA, achieving 77% accuracy after feature selection. Li et al. [33] employed virtual machines and LSTM models, achieving 92.6% accuracy, highlighting the flexibility of DL methods in handling complex datasets.

2.3. Public datasets in ASD research

Publicly available datasets have significantly contributed to ASD research. The Kaggle Toddler dataset, with 1050 records, has been instrumental in training models on early behavioural patterns. Talukdar et al. [29] used this dataset to evaluate ML algorithms, reporting accuracies exceeding 90%. The UCI Child dataset, comprising 732 records, has been utilized by Bala et al. [28] and others to assess classifier performance in structured environments.

Real-world datasets provide valuable insights. Vandewouw et al. [34] utilized datasets from the POND and HBN networks, employing Similarity Network Fusion (SNF) to identify ASD traits in children. Jacob et al. [30] integrated clustering techniques and ML models for data-driven subgroup analysis. Crippa et al. [35] achieved 96.7% accuracy using kinematic analysis of upper-limb movements in children aged 2–4 years.

2.4. Deep learning applications in ASD detection

Deep learning methods excel in handling unstructured and high-dimensional data. Heinsfeld et al. [36] utilized CNNs with ABIDE datasets, achieving 70% mean accuracy. Sherkatghanad et al. [37] extended this work, achieving 70.22% accuracy by improving feature extraction. Wall et al. [38] employed ADTree models, achieving high sensitivity and specificity in ASD screening. Liu et al. [39] and Wang et al. [40] demonstrated strong performance using RF and SVM, emphasizing the importance of robust datasets and feature selection.

2.5. Limitations of existing research

Despite the advancements in ASD detection, several challenges remain. Many studies rely on small, homogeneous datasets, limiting the

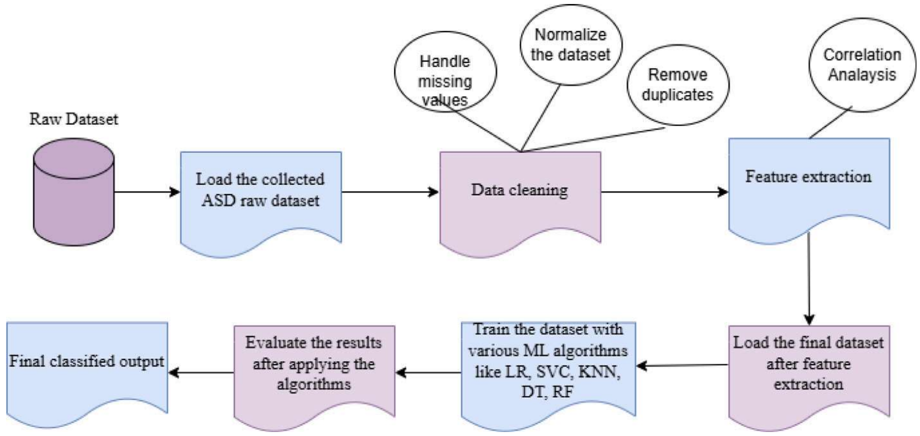


Fig. 1. Workflow for Autism Spectrum Disorder detection via machine learning.

generalizability of findings. Overfitting, lack of feature diversity, and dataset-specific constraints are common issues. Mazumdar et al. [41] highlighted cultural and regional variations, as seen in the ISAA scale in India, emphasizing the need for datasets that reflect local contexts. Bekerom et al. [42] utilized ML techniques, including Naive Bayes, RF, and SVM, to identify ASD traits in children, but challenges in feature diversity persisted.

2.6. Gaps and motivation for the current study

This study addresses the limitations of existing research by integrating three datasets: the Kaggle Toddler dataset, the UCI Child dataset, and a real-world dataset collected in Bangladesh. By combining structured, publicly available datasets with real-world data, the study evaluates ML models across diverse conditions. Furthermore, the research focuses on classical ML techniques, emphasizing simplicity and resource efficiency, while comparing them to state-of-the-art methods to highlight trade-offs in performance, generalizability, and practicality.

3. Methodology

This section describes the proposed system's methodology for detection of ASD from the data sets based on the screening tool questions. The flowchart of the methodology is shown in Fig. 1.

The Fig. 1 delineates the sequential steps of the research workflow. The initial phase involves the collection of raw ASD datasets, followed by preprocessing tasks such as data cleaning and feature extraction. Subsequently, a suite of machine learning algorithms, including Logistic Regression, K Nearest Neighbour, Support Vector Classifier, Decision Tree, and Random Forest, is applied to the preprocessed dataset. The models were implemented and evaluated using Python (version 3.10). The following libraries and frameworks were employed:

Table 1
Libraries and frameworks used in implementation.

Category	Details
Programming Language	Python (version 3.10)
Libraries/Frameworks	
Machine Learning	scikit-learn
Data Preprocessing	pandas, NumPy
Data Visualization	matplotlib, seaborn

The libraries listed in Table 1 play a crucial role in ensuring reproducibility and robustness in implementations, offering standardized tools and frameworks for consistent results across diverse environments Their performances were evaluated using accuracy, precision, recall, and F1-score metrics. Hyperparameter tuning was conducted for all models using grid search. For example, the regularization parameter C for Logistic Regression, the number of neighbours (K) for KNN, and the kernel coefficient (γ) for SVC were optimized during training. The anticipated results from the application of these algorithms constitute the expected outcomes of this research endeavor.

3.1. Dataset used

The datasets are based on AQ-10, a screening technique created by Baron-Cohen et al. [43]. To assist the tool's clinical implementation across diverse settings, Allison et al. [23] developed the Q-CHAT-10 for toddler and AQ 10-child, condensed versions of the original AQ. This study utilizes three datasets: two publicly available datasets (UCI child dataset and Kaggle toddler dataset) and one proprietary real-world dataset collected from autism and mainstream schools in Bangladesh.

1. **UCI Child Dataset:** Dataset of children for this study's purposes was gathered from the UCI Repository [44], which is open to the public. The UCI child dataset contains 292 records of children aged between 4–11 years. It includes 21 attributes such as AQ-10 questionnaire responses, gender, ethnicity, and ASD diagnosis labels. The dataset is balanced, with an approximately equal distribution of ASD-positive and ASD-negative cases, making it suitable for evaluating machine learning models.
2. **Kaggle Toddler Dataset:** The dataset of toddlers is collected from a site named Kaggle [45]. The Kaggle toddler dataset comprises 1050 records of toddlers aged 18–36 months. The dataset includes 18 attributes, such as AQ-10 scores, parental observations, and ASD diagnosis labels. Similar to the UCI dataset, this dataset is also balanced, with roughly equal representation of ASD-positive and ASD-negative cases.
3. **Real World Child Dataset:** The real-world dataset comprising 250 samples was collected anonymously from autism and mainstream schools in Bangladesh as part of a proprietary study. To ensure anonymity, no personally identifiable information (PII) was recorded during data collection. Each participant was assigned a unique identifier, and the data were aggregated at the institutional level before analysis. The study adhered to strict ethical guidelines, with verbal or written informed consent obtained from guardians or caregivers, and the process was overseen by institutional review boards. The dataset includes AQ-10 screening questionnaire responses, demographic details (age, gender), and behavioural observations recorded by trained professionals. To minimize bias, data were collected from multiple autism and mainstream schools representing diverse socioeconomic backgrounds. Approximately 30% of the participants in the dataset were diagnosed with ASD, providing a balanced

mix of ASD-positive and ASD-negative cases. This representative distribution supports robust model training and evaluation while mitigating over-representation of any particular subgroup.

Moreover, ASD is a neurodevelopmental condition characterized by a wide range of symptoms that can manifest differently across individuals and evolve throughout various life stages. The core symptoms associated with ASD are not static, but rather can evolve and transform over time on a broader scale. This synthesized dataset will undergo periodic updates by healthcare professionals, ensuring that it remains current and relevant.

3.2. Data prepossessing

Data pre-processing is a crucial step in transforming raw data into a format that is both meaningful and comprehensible for subsequent analyses. Real-world datasets often exhibit errors and null values, contributing to inconsistency and incompleteness. The significance of well-preprocessed data cannot be overstated, as it forms the basis for obtaining reliable and meaningful outcomes.

To address the challenges posed by incomplete and inconsistent data, a range of data pre-processing techniques is employed. These techniques includes,

- **Handling missing values:** Null values are removed from the dataset, ensuring completeness before further analysis. Removing missing values eliminates potential bias and ensures that analyses are conducted on a complete and reliable dataset. This approach simplifies data processing and minimizes the risk of introducing errors in subsequent analyses. For example, in the children dataset, the initial number of instances was 292, but after removing all the missing values, the dataset was reduced to 242 instances. Notably, there were no missing values in the toddler dataset and the newly collected dataset of children.
- **Data discretization:** Continuous variables are converted into categorical variables. Discretization simplifies data analysis by reducing the number of unique values, making it easier to interpret and analyse patterns within the dataset.
- **One-hot encoding:** Categorical variables are transformed into binary vectors to represent different categories. For example, in all three datasets, the sex column includes two categories: male and female. One-hot encoding is used here to convert these categories into numerical values (e.g., 0 for female and 1 for male), making them suitable for machine learning algorithms that require numerical input.
- **Data reduction:** Dimensionality reduction techniques, such as principal feature selection methods, are employed to decrease the number of variables in the dataset. This step is important for simplifying the dataset, reducing computational complexity, and enhancing the performance of machine learning models. A correlation-based feature selection method was employed to identify and retain the most relevant features for model training. This method evaluates the correlation between features and the target variable, ensuring the inclusion of attributes with the highest predictive power while minimizing redundancy.

3.3. Training and testing model

The datasets (toddler, child, and real-world) were divided into two segments: training and testing. A loop was conducted to systematically adjust the size of the training dataset from 60% to 90% and the testing dataset from 10% to 40%. Preprocessing steps, such as feature scaling and encoding, were performed independently within each iteration of the loop to ensure no information from the test set was used during training. Additionally, K-fold cross-validation was employed by dividing the dataset into five subsets, training and evaluating the models multiple times on different combinations of these subsets. These measures ensured robust evaluation and prevented data leakage.

After splitting the dataset into training and testing sets using train-test split, the validation set is not explicitly separated. However, during K-fold cross-validation, the training set is further partitioned into k subsets (folds), where each fold serves as both training and validation data iteratively. This means that each fold acts as a validation set for the rest of the training data. Therefore, even though the validation set is not explicitly defined after the initial split, it is effectively incorporated into the training process through K-fold cross-validation.

K-fold cross-validation is chosen because it maximizes the use of available data by splitting the training set into k subsets (folds), allowing each fold to act as a validation set while the remaining folds serve as the training set. This method ensures that each data point is used for both training and validation, enhancing the reliability and robustness of the model's performance evaluation. By rotating the validation set across k iterations, K-fold cross-validation reduces the risk of overfitting and provides a more comprehensive assessment of the model's generalization ability compared to a single validation split. This approach allows us to assess the model's performance and generalization ability while ensuring that the model is not overfitting to the training data.

3.3.1. Logistic regression

In binary classification problems, the result or goal variable is categorical and has two classes, commonly denoted by the numbers 0 and 1. In these situations, the statistical and machine learning model of choice is logistic regression. Despite its name, it is utilized more for classification than regression jobs. The likelihood that an input falls into a specific class is modelled by logistic regression [46]. The equation for logistic regression can be expressed as Eq. (1):

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n)}} \tag{1}$$

Where:

$P(Y = 1|X)$ represents the probability of the outcome variable Y being 1 given the predictor variables X.

e is the base of the natural logarithm.

$\beta_0, \beta_1, \beta_2, \dots, \beta_n$ are the coefficients of the logistic regression model.

$X_1, X_2, \dots, X_n$ are the predictor variables.

This equation models the log-odds of the probability of the outcome variable being 1 as a linear combination of the predictor variables, transformed by the logistic function to ensure that the predicted probabilities lie between 0 and 1.

Logistic Regression was employed as a baseline model for binary classification tasks in this study. The implementation used the liblinear solver, which is well-suited for small datasets and supports L1 and L2 regularization. Default regularization parameters were applied to ensure simplicity and reproducibility. Additionally, hyperparameter tuning was performed using grid search to identify optimal configurations. This ensured that the model was appropriately calibrated for the datasets used, balancing interpretability and predictive performance.

3.3.2. K-nearest neighbour (KNN)

The most straightforward among supervised learning techniques is K-Nearest Neighbour (KNN), which handles both classification and regression problems. KNN operates under the assumption that similar data points are in close proximity to each other. The parameter 'K' determines the number of ing points considered for predictions, and its selection is crucial for minimizing errors [47]. Here, it is configured with K = 7, using Euclidean distance. The equation for KNN classification can be described as Eq. (2):

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n)}} \tag{2}$$

Where:

$\hat{y}(x)$ represents the predicted class label for the data point x.

y_i represents the class labels of the k-nearest neighbour of data point x.

$I(\cdot)$ is the indicator function that returns 1 if the condition is true and 0 otherwise.

$\arg\max_{y_i}$ returns the class label that appears most frequently among the k-nearest neighbour.

3.3.3. Support vector classifier (SVC)

A supervised machine learning approach used largely for classification problems is a support vector classifier (SVC), commonly referred to as a support vector machine (SVM) classifier. It is a kind of discriminative classifier that seeks to identify the optimal hyperplane in a high-dimensional space for dividing data points into several groups [48]. The decision function for SVM can be defined as Eq. (3):

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} \tag{3}$$

Where:

$f(x)$ represents the decision function that predicts the class label for the input x.

α_i are the Lagrange multipliers obtained during training.

y_i are the class labels of the training data. $K(x_i, x)$ is the kernel function that computes the similarity between the input x and the training data x_i.

b is the bias term.

The SVC model was tested with different kernel functions, including linear, polynomial, and radial basis function (RBF), to determine the best fit for the datasets. Hyperparameter tuning was performed using grid search to optimize the regularization parameter (C) and the kernel coefficient (γ). The final configuration, which used an RBF kernel with optimized parameters, was selected based on performance across validation datasets.

3.3.4. Decision tree

An approach for supervised machine learning that is utilized for both classification and regression tasks is the Decision Tree. Visually, it resembles an inverted tree with branches and leaves, serving as a predictive modelling tool. Internal nodes in decision trees represent features (attributes or variables), branches represent decisions or rules, and leaves correspond to outcomes or class labels (in the case of classification). Decision trees are intuitive and interpretable, making them widely used in a variety of machine learning tasks. Here, the gini impurity criterion and limited depth are used to prevent overfitting.

The decision function for a decision tree can be represented as a series of if-else conditions based on the feature values. Given a feature vector x, the decision tree traverses the tree from the root node to a leaf node, where a prediction is made based on the majority class or the average value of the target variable in that leaf node [49]. The equation for the decision function of a decision tree can be expressed as Eq. (4):

$$f(x) = \begin{cases} c_1 & \text{if } x \le t_1 \\ c_2 & \text{if } t_1 < x \le t_2 \\ \vdots & \\ c_k & \text{if } x > t_{k-1} \end{cases} \tag{4}$$

Where:

$c_1, c_2, \ldots, c_k$ are the class labels or regression values associated with each leaf node.

$t_1, t_2, \ldots, t_{k-1}$ are the decision thresholds for each decision node.

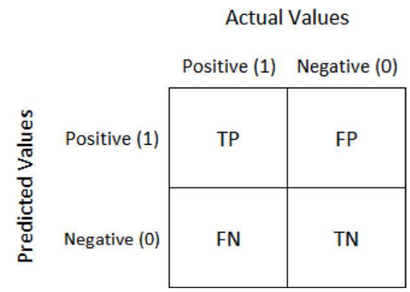


Fig. 2. Confusion matrix [51].

3.3.5. Random forest

For classification and regression tasks, a Random Forest is an ensemble machine learning technique. It is a development of decision trees and intended to address some of its shortcomings while preserving the readability and usability of individual decision trees. The great prediction accuracy and robustness of random forests are well recognized [50]. Random forest is configured here with 100 estimators and Gini impurity as the split criterion. The decision function for a random forest classifier can be expressed as Eq. (5):

$$f(x) = \text{mode}\left(f_1(x), f_2(x), \ldots, f_n(x)\right) \tag{5}$$

Where:

$f(x)$ represents the decision function of the Random Forest classifier.

$f_i(x)$ represents the decision function of the i-th decision tree in the forest.

mode represents the most frequently occurring class label among the predictions of the individual trees.

4. Result and evaluation

The assessment of outcomes will be conducted in terms of precision, recall, f1 score and accuracy. This evaluation will be facilitated by utilizing essential tools such as the confusion matrix, classification report, and ROC Curve. The result is determined by how precisely the model was trained.

- **Precision:** Precision is defined as the ratio of correctly predicted positive observations (True Positives (TP)) to the total predicted positive observations (True Positives + False Positives (FP)). As noted by Powers [52], precision assesses the model's ability to avoid false positives, making it particularly useful in scenarios where the cost of false positives is high. The formula for precision can be written as Eq. (6):
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{6}$$
- **Recall (Sensitivity or True Positive Rate):** Recall is defined as the ratio of correctly predicted positive observations (True Positives (TP)) to the total actual positive observations (True Positives + False Negatives (FN)). As noted by Powers [52], recall is particularly valuable in scenarios where identifying all positive cases is critical, such as in medical diagnoses. The formula for recall can be written as Eq. (7):
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{7}$$
- **F1 Score:** The F1 Score represents the harmonic mean of precision and recall, offering a balanced metric that takes into account both false positives and false negatives equally. As noted by Powers [52], the F1 Score is particularly valuable in scenarios where a balanced consideration of precision and recall is necessary. The

formula for the F1 Score can be written as Eq. (8):

$$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{8}$$

- **Accuracy:** Accuracy is the ratio of correctly predicted observations (True Positives (TP) + True Negatives (TN)) to the total number of observations in the dataset. As noted by Powers [52], accuracy provides an overall measure of correctness but can be misleading in imbalanced datasets where the majority class dominates. The formula for accuracy can be expressed as Eq. (9):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{9}$$

- **Mean Intersection over Union (mIoU):** Mean Intersection over Union (mIoU) is a metric used to evaluate the accuracy of a model's classification across multiple classes. It measures the extent of agreement between the predicted and ground truth class assignments for each class. As discussed by Rahman and Wang [53], IoU is calculated as the ratio of the overlap between the predicted and ground truth class assignments to their union. Mean IoU is then computed by averaging the IoU values across all classes, providing a comprehensive assessment of the model's performance across the entire classification task. The formula for IoU can be expressed as Eq. (10):

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{10}$$

- **Confusion Matrix:** A confusion matrix is used to assess a classification model's performance, particularly in binary classification problems. It provides a detailed overview of the model's predictions and their alignment with the actual data, summarizing the counts of TP, TN, FP, and FN. As explained by Kohavi and Provost [54], the confusion matrix clarifies the discrepancies between the expected and actual class labels, making it an essential tool for evaluating classification performance (see Fig. 2).
- **ROC Curve:** A graphical depiction used to assess the effectiveness of binary classification models is the Receiver Operating Characteristic (ROC) curve. It demonstrates the trade-off between the genuine positive rate (sensitivity) and the false positive rate (1-specificity) for various threshold values [55].

4.1. Performance analysis

For ASD screening data for toddlers and children, experimental results of several machine learning algorithms approaches with all characteristics selection have been demonstrated. In this step, all characteristics are chosen to determine the projected model's accuracy, mean intersection over union (mIoU), precision, recall, and F1 score. K = 7 was utilized in the implementation of KNN. To overcome the overfitting issue, K fold cross validation was done, where K ranged from 2 to 10. The most consistent results were observed with K = 5. Additionally, performance metrics across training, validation, and test sets were monitored to ensure the models' generalizability.

Data leakage risks were carefully mitigated. No features directly revealed the target label, preprocessing steps were conducted post data splitting, and stratified sampling ensured consistent class distributions across datasets. The splitting process ensured no overlap between training and test data, and cross-validation was performed without leakage between folds.

4.1.1. Toddler dataset

A loop was conducted, adjusting the size of the training dataset from 60% to 90% and testing dataset from 10% to 40%. The accuracy graph and the IoU graph for the toddler dataset are as follows:

From the Fig. 3, it is seen that accuracy of Logistic Regression ranges from 99.74% to 100%, K-Nearest Neighbour ranges from 93.59% to 97.10%, Support Vector Classifier accuracy does not change with the test size, it shows 100% accuracy in every test set, Decision Tree ranges from 93.80% to 97.90% and Random Forest ranges from 91.88% to 97%.

From the Fig. 4, it is seen that IoU value of Logistic Regression ranges from 99.63% to 100%, K-Nearest Neighbour ranges from 90.87% to 95.80%, Support Vector Classifier IoU does not change with the test size, it shows 100% IoU value in every test set, Decision Tree Ranges from 91.66% to 96% and Random forest ranges 97.10% from to 99.10%.

The results, as detailed in Table 2, provide a comprehensive evaluation of the performance metrics for the machine learning models.

The assessment of diverse machine learning models on the ASD toddler diagnosis dataset reveals a range of accuracies spanning from 95.80% to 100% on the original dataset. Notably, the K-Nearest Neighbour demonstrates the lowest accuracy, achieving 95.80%. Conversely, Logistic Regression and Support Vector Classifier exhibit the highest prediction accuracy, attaining a perfect score of 100% on the original dataset.

From Table 3, K-Fold Cross Validation, with K set to 5, is utilized to validate the dataset, resulting in mean accuracy values for machine learning models ranging from 91.84% to 100%. Notably, Logistic Regression and Support Vector Classifier achieve the highest accuracy, both attaining a perfect score of 100%.

The confusion matrix and the ROC curve for the toddler dataset using the suggested machine learning models are as follows:

Fig. 5(a), 5(b), 5(c), 5(d), 5(e) show the confusion matrix for Logistic Regression, Support Vector Classifier, Decision Tree, and Random Forest classifier, considering the Children dataset. In the confusion matrix, each row represents the instances in an actual class, while each column represents the instances in a predicted class. The text labels on the left side of the matrix indicate the true label, which is the actual class of the data point, while the text labels along the top of the matrix indicate the predicted label, which is the class that the algorithm predicted the data point belongs to. The numbers in the body of the confusion matrix represent the number of data points in each category. From the confusion matrices, it is evident that the false positive and false negative values for Logistic Regression and Support Vector Classifier are 0. However, K-Nearest Neighbour, Decision Tree, and Random Forest exhibit higher false negative and false positive values compared to these algorithms.

Fig. 6(a), 6(b), 6(c), 6(d), 6(e) shows the ROC curves for Logistic Regression, Support Vector Classifier, Decision Tree, and Random Forest classifier, considering Toddler dataset. In the ROC curve, the x-coordinate indicates False Positive Rate (FPR) and y-coordinate indicates True Positive Rate (TPR).

4.1.2. Children dataset

A loop was conducted, adjusting the size of the training dataset from 60% to 90% and testing dataset from 10% to 40%. The accuracy graph and the IoU graph for the children dataset are as follows:

From the Fig. 7, it is seen that accuracy of K-Nearest Neighbour ranges from 96% to 98%. The accuracy of Logistic Regression, Support Vector Classifier, Decision Tree and Random Forest remain unchanged with the test size, they show 100% accuracy in every test set.

From the Fig. 8, it is seen that IoU value of K-Nearest Neighbour ranges from 93.33% to 97.18%. The IoU value of Logistic Regression, Support Vector Classifier, Decision Tree and Random Forest remain unchanged with the test size, they show 100% IoU value in every test set.

The results, as detailed in Table 4, provide a comprehensive evaluation of the performance metrics for the machine learning models for the children dataset.

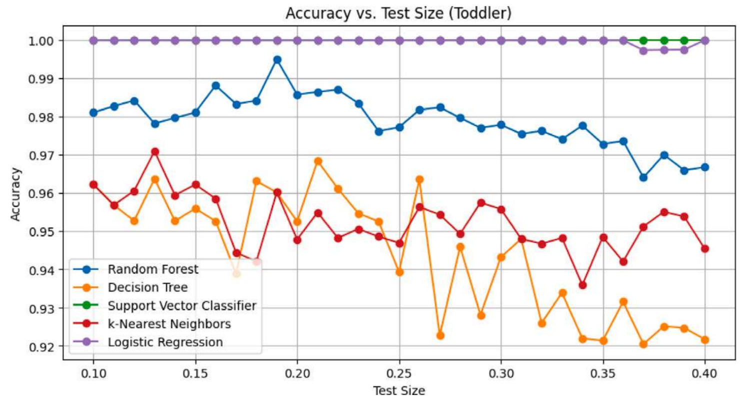


Fig. 3. Accuracy graph for Toddler dataset.

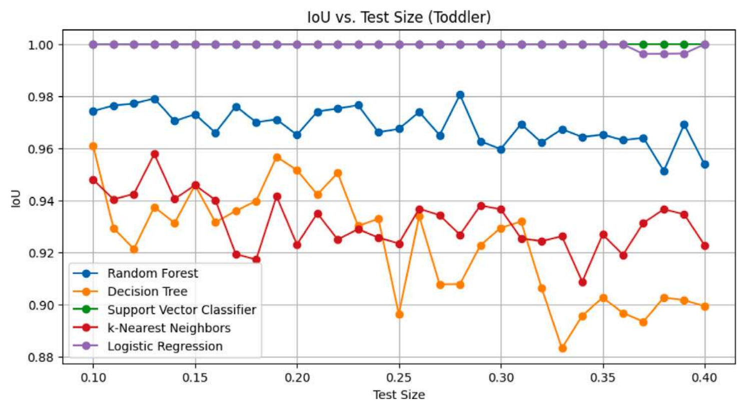


Fig. 4. IoU graph for Toddler dataset.

Table 2
Performance analysis for Toddler dataset.

Methods	Accuracy (%)	mIoU (%)	Precision (%)		Recall (%)		F1-score (%)	
			0	1	0	1	0	1
Logistic Regression	100	99.80	100	100	100	100	100	100
K-Nearest Neighbour	95.80	90.19	87	99	99	93	93	96
Support Vector Classifier	100	100	100	100	100	100	100	100
Decision Tree	96.20	92	91	99	97	95	94	97
Random Forest	99.10	94.79	97	99	97	99	97	99

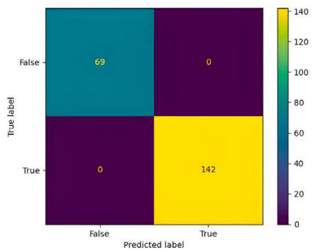
Table 3
K-Fold cross validation score for Toddler dataset.

Methods	Mean accuracy (%)	Standard Deviation of accuracy (%)
Logistic Regression	100	0
K-Nearest Neighbour	94.02	2.24
Support Vector Classifier	100	0
Decision Tree	91.84	2.33
Random Forest	96.20	1.58

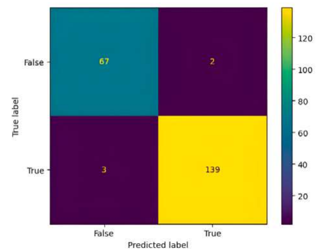
Table 4
Performance analysis for Children dataset.

Methods	Accuracy (%)	mIoU (%)	Precision (%)		Recall (%)		F1-Score (%)	
			0	1	0	1	0	1
Logistic Regression	100	99.96	100	100	100	100	100	100
K-Nearest Neighbour	98	93.17	97	100	100	95	98	98
Support Vector Classifier	100	100	100	100	100	100	100	100
Decision Tree	100	92.43	100	100	100	100	100	100
Random Forest	100	91.92	100	100	100	100	100	100

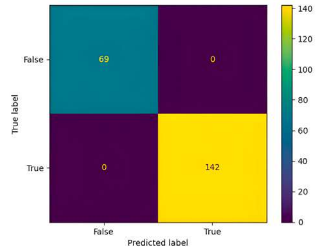
The assessment of different machine learning models on the ASD toddler diagnosis dataset reveals an accuracy range from 97.80% to 100% on the original dataset. The K-NN classifier with K = 7 exhibits the lowest accuracy at 97.80%. Conversely, Logistic Regression,



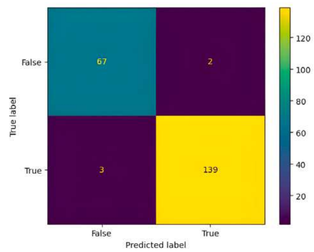
(a) Logistic Regression



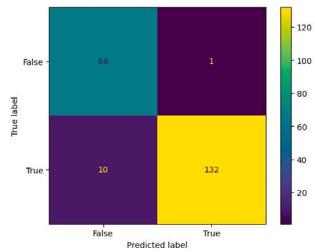
(b) K-Nearest Neighbour



(c) Support Vector Classifier

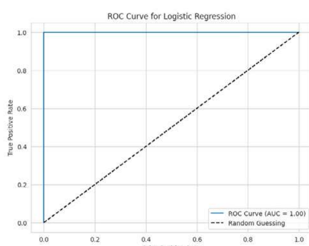


(d) Decision Tree

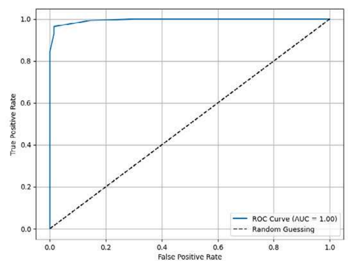


(e) Random Forest

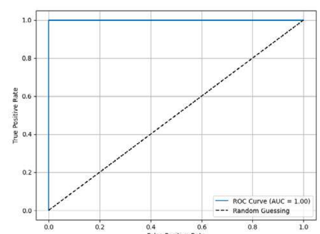
Fig. 5. Confusion matrix for Toddler dataset.



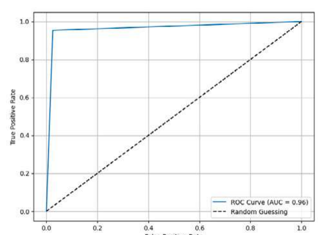
(a) Logistic Regression



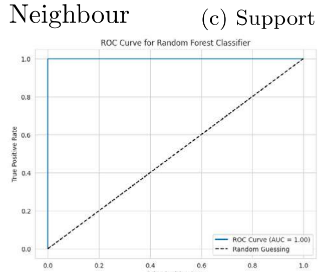
(b) K-Nearest Neighbour



(c) Support Vector Classifier



(d) Decision Tree



(e) Random Forest

Fig. 6. ROC curve for Toddler dataset.

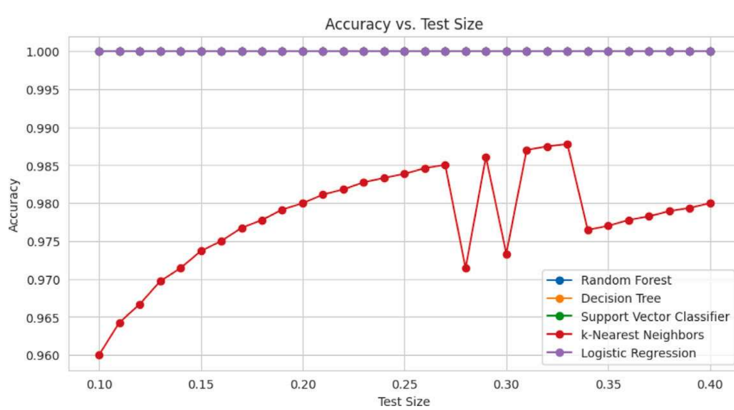


Fig. 7. Accuracy graph for children dataset.

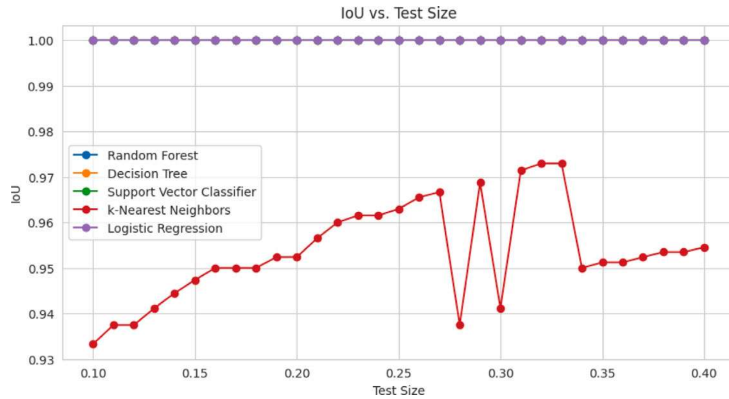
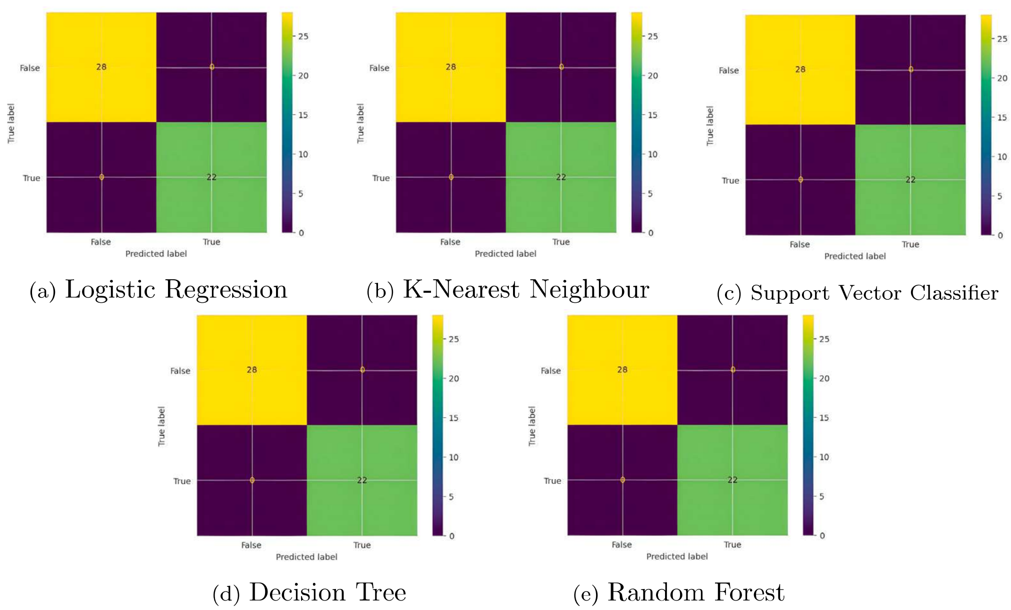


Fig. 8. IoU graph for children dataset.

Table 5
K-Fold cross validation score for Children dataset.

Methods	Mean accuracy (%)	Standard deviation of accuracy (%)
Logistic Regression	100	0
K-Nearest Neighbour	98.38	1.50
Support Vector Classifier	100	0
Decision Tree	100	0
Random Forest	100	0



(a) Logistic Regression (b) K-Nearest Neighbour (c) Support Vector Classifier

(d) Decision Tree (e) Random Forest

Fig. 9. Confusion matrix for children dataset.

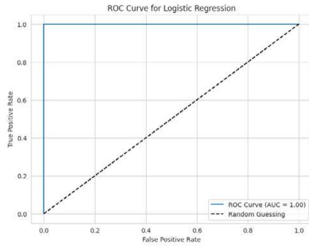
Support Vector Classifier, Decision Tree, and Random Forest models showcase the highest prediction accuracy, all achieving a perfect score of 100% on the original dataset.

From Table 5, it is observed that the dataset is validated using K-Fold Cross Validation with K set to 5, yielding mean accuracy values for machine learning models ranging from 98.38% to 100%. Among the models evaluated, Logistic Regression, Support Vector Classifier, Decision Tree, and Random Forest exhibit the highest accuracy, each achieving a perfect score of 100%.

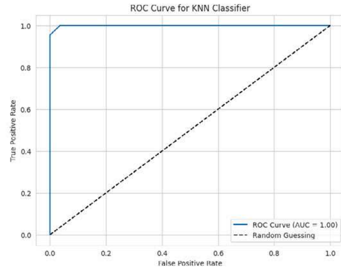
The confusion matrix and the ROC curve for the children dataset using the suggested machine learning models are as follows:

Fig. 9(a), 9(b), 9(c), 9(d), 9(e) shows the confusion matrix for Logistic Regression, Support Vector Classifier, Decision Tree, and Random Forest classifier, considering Children dataset. In the confusion matrix, each row represents the instances in an actual class, while each column represents the instances in a predicted class. The text labels on the left side of the matrix indicate the true label, which is the actual class of the data point, while the text labels along the top of the matrix indicate the predicted label, which is the class that the algorithm predicted the data point belongs to. The numbers in the body of the confusion matrix represent the number of data points in each category. From the confusion metrics, it is seen that the false positive and false negative values for Logistic Regression, Support Vector Classifier, Decision Tree and Random Forest are 0 here. K-Nearest Neighbour's accuracy is less here so it has higher false negative and false positive value than those algorithms.

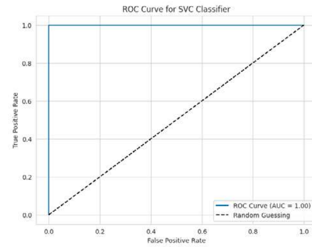
Fig. 10(a), 10(b), 10(c), 10(d), 10(e) shows the ROC curves for Logistic Regression, Support Vector Classifier, Decision Tree, and Random Forest classifier, considering Children dataset. In the ROC curve, the x-coordinate indicates False Positive Rate (FPR) and y-coordinate indicates True Positive Rate (TPR).



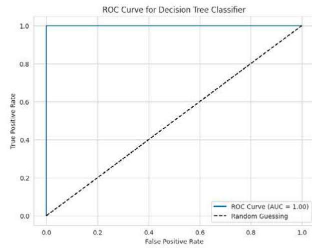
(a) Logistic Regression



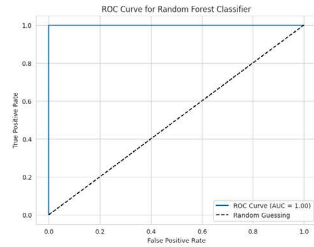
(b) K-Nearest Neighbour



(c) Support Vector Classifier



(d) Decision Tree



(e) Random Forest

Fig. 10. ROC curve for children dataset.

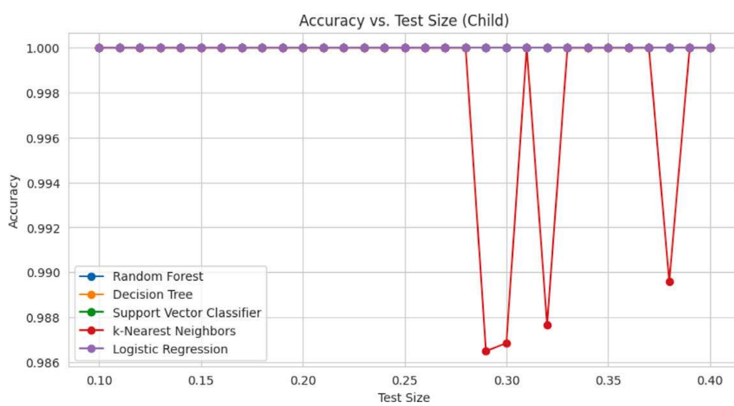


Fig. 11. Accuracy graph for Real World Children dataset.

4.1.3. *Real World Children dataset*

A loop was conducted, adjusting the size of the training dataset from 60% to 90% and testing dataset from 10% to 40%. The accuracy graph and the IoU graph for the toddler dataset are as follows:

From the Fig. 11, it is seen that accuracy of K-Nearest Neighbour ranges from 98.6% to 100%. The accuracy of Logistic Regression, Support Vector Classifier, Decision Tree and Random Forest remain unchanged with the test size, they show 100% accuracy in every test set.

From the Fig. 12, it is seen that IoU value of K-Nearest Neighbour ranges from 93.33% to 100%. The IoU value of Logistic Regression, Support Vector Classifier, Decision Tree and Random Forest remain unchanged with the test size, they show 100% IoU value in every test set.

The results, as detailed in Table 6, provide a comprehensive evaluation of the performance metrics for the machine learning models for the real world children dataset.

The assessment of different machine learning models on the ASD toddler diagnosis dataset reveals an accuracy of 100% Logistic Regression, Support Vector Classifier, K-Nearest Neighbour, Decision Tree, and Random Forest. But the mIoU of KNN is a little bit less, which is 99.60%.

Table 7 reveals that the dataset is validated using K-Fold Cross Validation with K set to 5, producing mean accuracy values for machine learning models in the range of 99% to 100%. Notably, Logistic Regression, Support Vector Classifier, Decision Tree, and Random Forest stand out by achieving a perfect accuracy of 100%.

The confusion matrix and the ROC curve for the real world children dataset using the suggested machine learning models are as follows.

Fig. 13(a), 13(b), 13(c), 13(d), 13(e) show the confusion matrix for Logistic Regression, Support Vector Classifier, Decision Tree, and Random Forest, considering the Real World Children dataset. In the confusion matrix, each row represents the instances in an actual class, while each column represents the instances in a predicted class. The text labels on the left side of the matrix indicate the true label, which is the actual class of the data point, while the text labels along the top of the matrix indicate the predicted label, which is the class that the algorithm predicted the data point belongs to. The numbers in the body of the confusion matrix represent the number of data points in each category. From the confusion matrices, it can be observed that the false positive and false negative values for Logistic Regression, K-Nearest Neighbour, Support Vector Classifier, Decision Tree, and

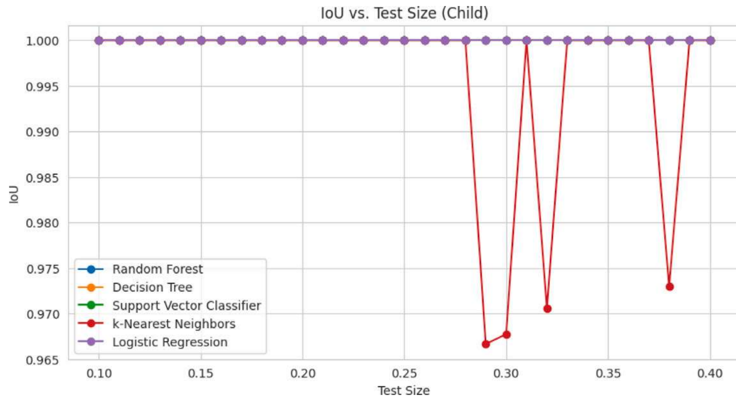


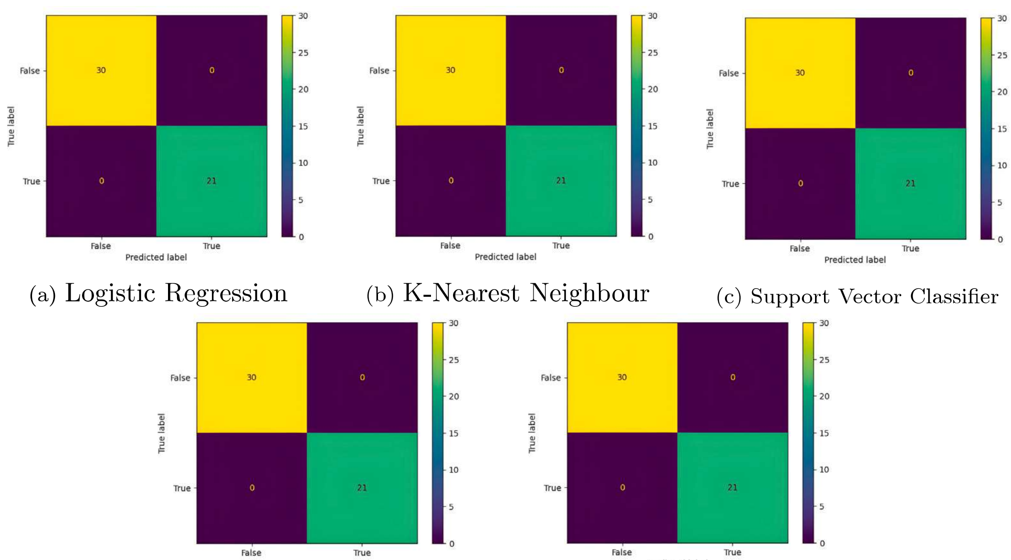
Fig. 12. IoU graph for real world children dataset.

Table 6
Performance analysis for Real World Children dataset.

Methods	Accuracy (%)	mIoU (%)	Precision (%)		Recall (%)		F1-Score (%)	
			0	1	0	1	0	1
Logistic Regression	100	100	100	100	100	100	100	100
K-Nearest Neighbour	100	99.60	100	100	100	100	100	100
Support Vector Classifier	100	100	100	100	100	100	100	100
Decision Tree	100	1oo	100	100	100	100	100	100
Random Forest	100	100	100	100	100	100	100	100

Table 7
K-Fold cross validation score for Real World Children dataset.

Methods	Mean accuracy (%)	Standard deviation of accuracy (%)
Logistic Regression	100	0
K-Nearest Neighbour	99	1.22
Support Vector Classifier	100	0
Decision Tree	100	0
Random Forest	100	0



(a) Logistic Regression (b) K-Nearest Neighbour (c) Support Vector Classifier

(d) Decision Tree (e) Random Forest

Fig. 13. Confusion matrix for real world children dataset.

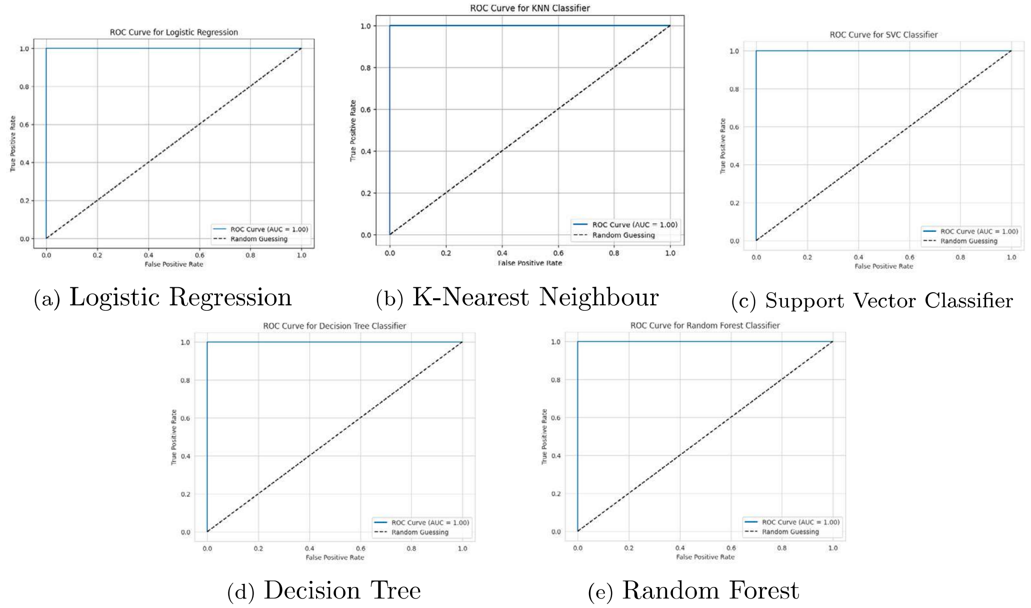


Fig. 14. ROC curve for real world children dataset.

Table 8
Comparison of the proposed study with state-of-the-art techniques.

Study	Methodology	Dataset	Key findings and comparison with current study
Multimodal approaches for early autism spectrum disorder detection using machine learning [56]	Multimodal Deep Learning	Neuroimaging, genetics, and behavioural data	Integration of multiple modalities enhances diagnostic accuracy. Unlike this study, it requires complex data sources and computational resources.
Deep learning models for joint attention detection in children with autism spectrum disorder [57]	Joint Attention–Based Deep Learning	Video recordings of children	Focused on joint attention behaviours for ASD detection. This study uses structured datasets like AQ-10 responses, whereas this approach leverages video data.
A deep learning approach to detecting ASD using resting-state fMRI [58]	Deep Learning with Resting-State fMRI	Resting-state fMRI data	Advanced fMRI biomarkers improve automated ASD diagnosis. This study avoids the high costs and technical requirements of fMRI data.
CNN-based analysis of fMRI data for autism spectrum disorder detection [59]	CNN on fMRI Data	Resting-state fMRI data	High accuracy in early ASD detection among children. Similar to Huda et al. (2024), this approach uses costly and specialized neuroimaging data.
Meta-analysis of deep learning techniques for autism spectrum disorder diagnosis [60]	Deep Learning Meta-Analysis	Multiple datasets	Highlights the effectiveness of DL methods in ASD detection. This study emphasizes interpretability and real-world applicability.

Random Forest are 0.

Fig. 14(a), 14(b), 14(c), 14(d), 14(e) shows the ROC curves for Logistic Regression, Support Vector Classifier, Decision Tree, and Random Forest classifier, considering Children dataset. In the ROC curve, the x-coordinate indicates FPR and y-coordinate indicates TPR.

4.1.4. *Comparative analysis*

The proposed methodology exhibits superior accuracy compared to several research papers studied throughout the research process. A comparison Table 8 is presented below, highlighting the performance of the proposed methodology in contrast to the findings of other research papers.

4.1.5. *Result discussion*

The evaluation of classical machine learning algorithms, including Logistic Regression, Support Vector Classifier, k-Nearest Neighbours, and Decision Tree Classifier, was conducted on three datasets: the UCI child dataset, Kaggle toddler dataset, and a proprietary real-world dataset. The results revealed that classical machine learning models are effective for Autism Spectrum Disorder (ASD) detection. Precision, recall, and F1-score were used to evaluate model performance, ensuring reliability and robustness in results. Notably, the models achieved high overall accuracy and precision–recall balance, with some variation across datasets.

The results demonstrated that Support Vector Classifier achieved a balanced performance across all datasets, while Logistic Regression performed consistently but exhibited a slight sensitivity to data variations. These models effectively leveraged structured datasets like AQ-10 responses and demographic attributes. The performance metrics suggest that the simplicity and consistency of the datasets played a significant role in the models achieving high precision and recall values. However, the generalizability of these models to more diverse populations needs further exploration.

Feature importance analysis highlighted that AQ-10 responses were the most significant predictors of ASD, alongside demographic variables like age and gender. These insights align with clinical observations, offering interpretability and practical relevance in healthcare applications. The proprietary dataset introduced diversity in the training process, which contributed to robust model development despite its smaller size compared to the public datasets.

While the current study primarily focused on evaluating the models on structured datasets, the diversity of the datasets used — including public datasets and a real-world proprietary dataset — allowed for an implicit assessment of model performance under varying conditions. These datasets differ in terms of data distribution, feature patterns, and sample demographics, which provided a measure of the models' ability to generalize across different scenarios. The consistent performance of the models across these datasets, achieving high accuracy and mIoU scores, demonstrates their robustness to naturally occurring variations in data quality and distribution.

When compared to state-of-the-art techniques such as deep learning approaches, classical machine learning models demonstrated competitive performance. While deep learning models using neuroimaging or behavioural data often achieve high accuracy (e.g., 92%–96%), these require significant computational resources and specialized data, making them less accessible in resource-constrained environments. In contrast, the classical models in this study achieved comparable results with minimal computational requirements, emphasizing their practicality for early ASD detection in real-world settings. That is how the robustness of the proposed models was evaluated using controlled experiments to simulate real-world scenarios.

Despite these promising results, the study acknowledges its limitations. The proprietary dataset, while valuable, was relatively small, potentially limiting the generalizability of the models. Additionally, the current study focuses primarily on structured data, limiting its applicability to scenarios involving unstructured data such as neuroimaging or video recordings. These limitations highlight areas for future improvement. Future research will focus on expanding datasets, exploring deep learning approaches for unstructured data, and incorporating techniques to enhance robustness under varied real-world conditions.

5. Conclusion and future work

This study demonstrated the potential of classical machine learning algorithms, such as Logistic Regression and Support Vector Classifier, for the early detection of Autism Spectrum Disorder (ASD) across toddlers, children, and real-world datasets, achieving 100% accuracy and high mIoU scores. The simplicity and homogeneity of the datasets likely contributed to these results, though improvements in generalizability to diverse, real-world scenarios remain necessary. Future work includes testing the models on larger datasets exceeding 10,000 samples, incorporating neuro-image data and image processing techniques for enhanced autism detection, and exploring deep learning methodologies such as video analysis. The study will be expanded to include adolescents and adults to ensure comprehensive applicability across age groups and varied clinical profiles. These efforts aim to address current limitations and ensure broader real-world impact in clinical and low-resource settings.

Ethics approval

The authors confirm that the study did not involve human or animal subjects, thus, formal ethics approval was not required.

Funding

The authors would like to clarify that the research was conducted without external funding. Therefore, there are no funding agencies or grant numbers associated with this study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] C. Lord, M. Elsabbagh, G. Baird, J. Veenstra-Vanderweele, Autism spectrum disorder, Lancet 392 (10146) (2018) 508–520.
[2] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders (DSM-5), American Psychiatric Association Publishing, 2013.
[3] U. Frith, Autism: Explaining the Enigma, Blackwell Publishing, 2003.
[4] S. Baron-Cohen, Mindblindness: An Essay on Autism and Theory of Mind, MIT Press, 1997.
[5] G. Dawson, K.M. Zanolli, Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder, Dev. Psychopathol. 24 (3) (2012) 759–770.
[6] S.J. Rogers, G. Dawson, Early start Denver model: A comprehensive early intervention approach for toddlers with autism, J. Cogn. Behav. Psychother. 10 (2010) 163–174.
[7] Autism Parenting Magazine, Autism statistics, 2023, Autism Parenting Magazine. Retrieved from https://www.autismparentingmagazine.com/autism-statistics/.
[8] M.H.K. Khan, Autism spectrum disorder and its care in Bangladesh, 2023, The Daily Star. Retrieved from https://www.thedailystar.net/health/news/autism-spectrum-disorder-and-its-care-bangladesh-1859095.
[9] Autism.org, Symptoms of autism, 2023, Retrieved from https://autism.org/symptoms-of-autism/.
[10] Centers for Disease Control and Prevention, Autism spectrum disorder (ASD), 2023, Retrieved from https://www.cdc.gov/autism.
[11] World Health Organization, Autism spectrum disorders, 2023, Retrieved from https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders.
[12] J.L. Matson, K.R. Smith, Behavioral treatment of self-injury and related problems in children with autism spectrum disorders, Res. Autism Spectr. Disord. 2 (2) (2008) 322–336.
[13] P. Jain, M.K. Gupta, R. Kumar, A.K. Singh, Autism spectrum disorder detection using machine learning techniques, Int. J. Eng. Technol. (IJET) 9 (8) (2017) 1585–1592.
[14] G. Dawson, S. Rogers, J. Munson, M. Smith, J. Winter, J. Greenson, A. Donaldson, J. Varley, Randomized, controlled trial of an intervention for toddlers with autism: the early start denver model, Pediatrics 125 (1) (2010) e17–e23.
[15] S.J. Rogers, L.A. Vismara, Evidence-based comprehensive treatments for early autism, J. Clin. Child Adolesc. Psychol. 37 (1) (2008) 8–38.
[16] R. Paul, Interventions to improve communication in autism, Child Adolesc. Psychiatr. Clin. North Am. 17 (4) (2008) 835–856.
[17] T. Smith, Outcome of early intervention for children with autism, Clin. Psychol. Sci. Pract. 6 (1) (1999) 33.
[18] C. Kasari, S. Freeman, T. Paparella, Joint attention and symbolic play in young children with autism: A randomized controlled intervention study, J. Child Psychol. Psychiatry 47 (6) (2006) 611–620.
[19] J.W. Jacobson, J.A. Mulick, G. Green, Cost–benefit estimates for early intensive behavioral intervention for young children with autism—general model and single state case, Behav. Interv. Theory Pract. Resid. Commun. Based Clin. Prog. 13 (4) (1998) 201–226.
[20] S. Correspondent, Army's Proyash to reach out further, 2016, The Daily Star. Retrieved from https://www.thedailystar.net/city/armys-proyash-reach-out-further-1328926.
[21] K.E. Zuckerman, O.J. Lindly, B.K. Sinche, Parental concerns, provider response, and timeliness of autism spectrum disorder diagnosis, J. Pediatr. 166 (6) (2015) 1431–1439.
[22] D. Bone, S.L. Bishop, M.P. Black, M.S. Goodwin, C. Lord, S.S. Narayanan, Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion, J. Child Psychol. Psychiatry 57 (8) (2016) 927–937.
[23] C. Allison, B. Auyeung, S. Baron-Cohen, Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls, J. Am. Acad. Child Adolesc. Psychiatry 51 (2) (2012) 202–212.
[24] S. Raj, S. Masood, Analysis and detection of autism spectrum disorder using machine learning techniques, Procedia Comput. Sci. 167 (2020) 994–1004.
[25] S.H. Hammed, A. Albahri, Unlocking the potential of autism detection: Integrating traditional feature selection and machine learning techniques, Appl. Data Sci. Anal. 2023 (2023) 42–58.
[26] F. Thabtah, Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment, in: Proceedings of the 1st International Conference on Medical and Health Informatics 2017, 2017, pp. 1–6.
[27] K.S. Omar, P. Mondal, N.S. Khan, M.R.K. Rizvi, M.N. Islam, A machine learning approach to predict autism spectrum disorder, in: 2019 International Conference on Electrical, Computer and Communication Engineering, ECCE, IEEE, 2019, pp. 1–6.

[28] M. Bala, M.H. Ali, M.S. Satu, K.F. Hasan, M.A. Moni, Efficient machine learning models for early stage detection of autism spectrum disorder, Algorithms 15 (5) (2022) 166.

[29] J. Talukdar, D.K. Gogoi, T.P. Singh, A comparative assessment of most widely used machine learning classifiers for analysing and classifying autism spectrum disorder in toddlers and adolescents, Heal. Anal. 3 (2023) 100178.

[30] S.G. Jacob, M.M.B.A. Sulaiman, B. Bennet, Algorithmic approaches to classify autism spectrum disorders: a research perspective, Procedia Comput. Sci. 201 (2022) 470–477.

[31] Z. Zhao, X. Zhang, W. Li, X. Hu, X. Qu, X. Cao, Y. Liu, J. Lu, Applying machine learning to identify autism with restricted kinematic features, IEEE Access 7 (2019) 157614–157622.

[32] S. Mostafa, L. Tang, F.-X. Wu, Diagnosis of autism spectrum disorder based on eigenvalues of brain networks, Ieee Access 7 (2019) 128474–128486.

[33] J. Li, Y. Zhong, J. Han, G. Ouyang, X. Li, H. Liu, Classifying ASD children with LSTM based on raw videos, Neurocomputing 390 (2020) 226–238.

[34] M.M. Vandewouw, J. Brian, J. Crosbie, R.J. Schachar, A. Iaboni, S. Georgiades, R. Nicolson, E. Kelley, M. Ayub, J. Jones, et al., Identifying replicable subgroups in neurodevelopmental conditions using resting-state functional magnetic resonance imaging data, J. AMA Netw. Open 6 (3) (2023) e232066.

[35] A. Crippa, C. Salvatore, P. Perego, S. Forti, M. Nobile, M. Molteni, I. Castiglioni, Use of machine learning to identify children with autism and their motor abnormalities, J. Autism Dev. Disord. 45 (2015) 2146–2156.

[36] A.S. Heinsfeld, A.R. Franco, R.C. Craddock, A. Buchweitz, F. Meneguzzi, Identification of autism spectrum disorder using deep learning and the ABIDE dataset, NeuroImage Clin. 17 (2018) 16–23.

[37] A. Fredo, A. Jahedi, M. Reiter, R.-A. Müller, Diagnostic classification of autism using resting-state fMRI data and conditional random forest, Age 12 (2.76) (2018) 6–41.

[38] D.P. Wall, J. Kosmicki, T. Deluca, E. Harstad, V.A. Fusaro, Use of machine learning to shorten observation-based screening and diagnosis of autism, Transl. Psychiatry 2 (4) (2012) e100.

[39] W. Liu, M. Liu, D. Yang, M. Wang, T. Tao, Automatic diagnosis of autism based on functional magnetic resonance imaging and elastic net, in: 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference, ITOEC, IEEE, 2020, pp. 104–108.

[40] Y. Wang, J. Wang, F.-X. Wu, R. Hayrat, J. Liu, AIMAFE: Autism spectrum disorder identification with multi-atlas deep feature representation and ensemble learning, J. Neurosci. Methods 343 (2020) 108840.

[41] A. Mazumdar, B. Chatterjee, M. Banerjee, S. Shanker, Machine learning based autism screening tool—a modified approach, Multimedia Tools Appl. (2024) 1–18.

[42] B. van den Bekerom, Using machine learning for detection of autism spectrum disorder, in: Proc. 20th Student Conf. IT, 2017, pp. 1–7.

[43] S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin, E. Clubley, The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians, J. Autism Dev. Disord. 31 (2001) 5–17.

[44] F. Thabtah, Autistic spectrum disorder screening data for children, 2017, http://dx.doi.org/10.24432/C5659W, UCI Machine Learning Repository.

[45] F.F. Thabtah, Autism screening data for toddlers, 2019, URL https://www.kaggle.com/datasets/fabdelja/autism-screening-for-toddlers.

[46] D.W. Hosmer Jr., S. Lemeshow, R.X. Sturdivant, Applied Logistic Regression, vol. 398, John Wiley & Sons, 2013.

[47] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inform. Theory 13 (1) (1967) 21–27.

[48] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297.

[49] J.R. Quinlan, Induction of decision trees, Mach. Learn. 1 (1) (1986) 81–106.

[50] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.

[51] Sarang Narkhede(Towards DataScience), Understanding confusion matrix, 2023, URL https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62.

[52] D.M. Powers, Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation, J. Mach. Learn. Technol. 2 (1) (2020) 37–63.

[53] M.A. Rahman, Y. Wang, Optimizing intersection-over-union in deep neural networks for image segmentation, Int. Symp. Vis. Comput. (2016) 234–244.

[54] R. Kohavi, F. Provost, Glossary of terms, Mach. Learn. 30 (2–3) (1998) 271–274.

[55] T. Fawcett, An introduction to ROC analysis, Pattern Recognit. Lett. 27 (8) (2006) 861–874.

[56] J. Dcouto, R. Pradeepkandhasamy, Multimodal approaches for early autism spectrum disorder detection using machine learning, Sensors 23 (5) (2023) 1228, http://dx.doi.org/10.3390/s23051228.

[57] J. Ko, S. Lee, Y. Kim, Deep learning models for joint attention detection in children with autism spectrum disorder, J. AMA Netw. Open 6 (5) (2023) e230123, http://dx.doi.org/10.1001/jamanetworkopen.2023.0123.

[58] M.A. Huda, A. Rahman, M.T. Chowdhury, A deep learning approach to detecting ASD using resting-state fMRI, Cogn. Neurodyn. 18 (1) (2024) 45–58, http://dx.doi.org/10.1007/s11571-024-10176-z.

[59] Q. Feng, J. Xu, CNN-based analysis of fMRI data for autism spectrum disorder detection, Children 10 (10) (2023) 1654, http://dx.doi.org/10.3390/children10101654.

[60] Z. Ding, H. Wang, J. Liu, Meta-analysis of deep learning techniques for autism spectrum disorder diagnosis, BMC Psychiatry 24 (1) (2024) 216, http://dx.doi.org/10.1186/s12888-024-06116-0.