Scope
Volume 16 Number 03 September 2026

Machine Learning for the Differential Diagnosis of ASD and ADHD in Children: A Classification Approach

Kazi Md. Shahiduzzaman[1]; Bijoy Kumar Karmaker[1]; Mamash Kanti Saha[2]; Urmi Saha[2]

[1] Associate Professor; Department of EEE, Jatiya Kabi Kazi Nazrul Islam University (JKKNIU), Mymensingh, Bangladesh

[2] Director; Child Development Institute (CDI) and Welfare School, Mymensingh, Bangladesh

Corresponding Author: **Bijoy Kumar Karmaker**

Abstract

Purpose – This study aims to develop and evaluate a machine learning-based framework capable of distinguishing between Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD) in children using structured, question-based behavioural data, addressing the diagnostic overlap that often delays accurate clinical differentiation. **Design/methodology/approach** – Data were collected from 50 children through a structured questionnaire covering motor, functional and behavioural indicators. After preprocessing (feature scaling and stratified 80:20 train–test splitting), eight supervised machine learning algorithms — Support Vector Machine, Random Forest, K-Nearest Neighbour, Naïve Bayes, Logistic Regression, Decision Tree, Linear Discriminant Analysis and Artificial Neural Network — were trained and evaluated using accuracy, precision, recall, F1-score and false alarm rate. **Findings** – Random Forest, Logistic Regression and Linear Discriminant Analysis achieved 100 per cent accuracy, precision, recall and F1-score, with a 0 per cent false alarm rate, indicating strong discriminative power even with simpler, more interpretable models. The Decision Tree algorithm performed least effectively (80 per cent accuracy, 50 per cent recall). **Research limitations/implications** – The small sample size (N = 50) raises concerns about overfitting and generalisability; validation on larger, more demographically diverse, multi-centre datasets is required. **Practical implications** – The proposed framework offers a scalable, objective and cost-effective screening tool that could support clinicians, particularly in resource-limited settings, in achieving earlier and more accurate differential diagnosis of ASD and ADHD. **Originality/value** – This study contributes a focused, question-based diagnostic approach using gold-standard clinical instruments rather than large heterogeneous datasets, and compares eight classifiers to identify the most clinically reliable models for ASD–ADHD differentiation.

Keywords – Machine learning, Autism Spectrum Disorder, Attention-Deficit Hyperactivity Disorder, Questionnaire data, Early detection, Artificial intelligence in healthcare, Classification, Random Forest, Logistic regression, Paediatric diagnosis.

1. Introduction

Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD) are two of the most common neurodevelopmental diseases. They are all different in behaviour and developmental stages with some common features. These diseases are highly prevalent and can have a high clinical overlap, making them important diagnostic challenges in child and adolescent mental health treatment, and often making it difficult to make accurate distinctions between them (Casseus and M., 2022; Wu, 2024). ASD is mostly linked to problems with social interaction, communication, and repetitive behaviors, while ADHD is characterized by hyperactivity, impulsivity, and inattention (Federico et al., 2023). These conditions have significant and long-term developmental impacts, so they need to be identified early and accurately for timely and effective interventions to be delivered for long-term quality gains. The difficulty of differentiating between different diseases sometimes delays a correct intervention, highlighting the need for more advanced and objective methods with a great degree of diagnostic differentiation (Fernández-Delgado et al., 2024). In this context, machine learning represents a pioneering methodological approach that enables the in-depth analysis of clinical data captured digitally and significantly enhances the accuracy and effectiveness of the diagnosis of neurodevelopmental disorders (Eslami et al., 2019). The use of machine learning makes it possible to build predictive models with high diagnostic accuracy, which is directly useful in the clinic for finding and treating instances of ASD and ADHD early (Grazioli et al., 2024; Shahiduzzaman et al., 2024). One of the significant progresses of recent research is that the use of complicated machine learning and deep learning systems consistently outperform traditional assessment methods in the classification of ASD and ADHD (Jaiswal et al., 2016). Data-driven machine learning approaches are able to identify subtle phenotypic differences, especially their ability to extract diagnostically important features from standardized assessment tools, which would support reliable differentiation between ASD and non-ASD populations (Wolff et al., 2022). This capability represents a paradigm change from subjective, clinician-dependent clinical diagnostics towards objective, reproducible, and scalable diagnostics that minimise the chances of bias (Wolff et al., 2022).

Such methodological innovation would be especially important because of the existence of substantial clinical symptom overlap, with children with ADHD also showing abnormalities in social interaction as seen in children with ASD, and inattention being a central feature of ADHD also often appearing in presentations of ASD. This overlap causes difficulty in diagnosis and a large number of misdiagnoses and delayed diagnosis (Ardulov et al., 2021). In response, the creation and use of advanced machine learning algorithms is a new and important contribution that provides a data-driven basis for improving the accuracy, dependability, and scalability of differential diagnosis in pediatric neurodevelopmental disorders (Fernández-Delgado et al., 2024; Wu, 2024). The existing diagnostic framework heavily relies on subjective assessments such as

clinical observations and self-report questionnaires, which often result in inconsistencies, delays in diagnosis, and potential misdiagnosis, given the considerable overlap in symptoms (Aasim et al., 2024; Jugunta et al., 2023). This is an example of why it is so important to have more objective, more reliable diagnostic tools.

Machine learning, a component of artificial intelligence, offers itself as a possible remedy as it is very efficient at recognising complicated patterns in information and separating between ailments with comparable clinical manifestations. Small high dimensional datasets found in healthcare are especially well suited for Support Vector Machines and other algorithms (Chou et al., 2017). This method could greatly improve the accuracy of differential diagnosis for neurodevelopmental diseases by using computer models to analyze complex clinical data (Wolff et al., 2022). These advanced techniques can analyze a wide range of data, including neuroimaging, genetic markers, and behavioral observations, to identify subtle biomarkers and unique patterns linked to specific disorders, thereby improving the diagnostic capabilities of healthcare providers (Shahiduzzaman et al., 2024; Zolghadr et al., 2024; Ashraf et al., 2024). Moreover, data-driven methods can help to clarify the phenotypic differences and similarities between ASD and ADHD, which is important for developing personalized treatment strategies (Wu, 2024; Zhang et al., 2024). Correctly identifying clinical symptoms as particular to ASD or ADHD is essential for accurate diagnosis and to guarantee suitable healthcare and social assistance (Wolff et al., 2022).

This work presents the utilization of machine learning to distinguish between several neurodevelopmental profiles: persons diagnosed only with ASD, exclusively with ADHD, those with comorbid ASD and ADHD, and neurotypical controls. This separation is achieved with the use of specific, diagnostically relevant questions from established clinical instruments (the Autism Diagnostic Observation Schedule and the Autism Diagnostic Interview-Revised) rather than relying on larger, less specific datasets. This focused method makes sure that it is clinically relevant and makes it easier to integrate into diagnostic workflows. The main goal is to create a strong classification model based on machine learning that can accurately tell the difference between these subtle diagnostic categories utilizing a structured, question-based dataset from gold-standard clinical examinations. The research is driven by the need to address persistent diagnostic challenges, enhance the availability and objectivity of diagnostic support, particularly in resource-limited settings with restricted expertise and traditional diagnostic methods often being time-consuming and subjective. This work proposes a data-driven, scalable, objective methodology which makes use of the analytical capability of machine learning. It represents a significant advance in bridging the gap between traditional, subjective systems of diagnosis and advanced mathematical approaches and calls for earlier and more precise identification and is conducive to the implementation of individualized and effective intervention strategies that can improve long-term developmental outcomes. The next parts talk about the background, methods, results, and debate of this machine learning framework.

2. Literature review

Traditionally, the diagnosis of Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder has been based on long, subjective behavioral observations and clinical assessment, which may lack diagnostic accuracy and reliability. Machine learning is a powerful solution to these diagnostic problems, because it enables identification of complex patterns in data (Mikoláš et al., 2022; Zhao et al., 2021). Recent studies have revealed the possibility of accurate diagnosis of ASD using a significantly reduced number of diagnostic features by employing ML models. This enhances the diagnostic process and aides doctors in making better decisions (Liu et al., 2024; Schulte-Rüther et al., 2021; Wolff et al., 2022). For instance, a data-driven ML method was able to distinguish the ASD and non-ASD groups using a subset of five ASD features from the Autism Diagnostic Observation Schedule and the Autism Diagnostic Interview-Revised instruments, which validated the use of streamlined sets of features for ASD diagnosis (Wolff et al., 2022). In addition, integrating multiple data modalities, such as EEG and questionnaire results, can offer machine learning models input from a broader range of features, which can improve the accuracy of diagnosis and procedures related to these complex neurodevelopmental disorders (Rajagopalan and Tammimies, 2024). The application of machine learning (ML), especially with digitally sourced parent-reported socio-anamnestic data, can produce accurate predictive models for the classification of ADHD and ASD, significantly impacting clinical practice by enabling early psychopathological diagnosis (Grazioli et al., 2024).

These studies may be used to group research by the type of data used. For example, one group of studies looks at how to combine several types of data, such as neuroimaging, EEG, and behavioral assessments, to make diagnoses more accurate. Combination of multiple data streams in a single diagnostic framework has shown significant promise (Ardulov et al., 2021; Bertacchini et al., 2023; Mukherjee et al., 2023; Song et al., 2022). Another group of studies shows how important digitally collected socio-anamnestic data is for early diagnosis, especially insights from parents (Grazioli et al., 2024; Alsharif et al., 2024). In addition, some machine learning algorithms such as decision tree, logistic regression, support vector machine and random forest have been successfully employed to classify and predict ADHD in children, and are effective in different clinical applications.

There are opportunities for further learning in each of these clusters and needs for further research. For example, although there has been progress in the use of multimodal data for diagnosis, the region could greatly benefit from providing standardized protocols and larger data sets to validate results with various populations. Similarly, socio-anamnestic data have been used with success, but further studies are required for the standardization and optimization of the questionnaires to achieve more accurate and reliable parent-reported data. Using a large amount of clinical data along with contemporary computer technology, ML can overcome the limitations of traditional clinical research. This paves the way for novel approaches to disease

diagnosis (Mao et al., 2025). We contribute to this emerging field by proposing a holistic approach for the analysis of missing data in diagnostics, which relies on machine learning for the analysis of these important gaps to support more accurate and objective assessment of neurodevelopmental disorders.

3. Methodology

The methodological framework of this study involves an extensive examination of machine learning techniques for the precise separation of ASD and ADHD, emphasizing the application of validated diagnostic tools and objective data. The machine learning model for identifying ASD and ADHD was created utilizing an organized method that emphasized preparing the data and thoroughly testing the model. This included an initial stage of data preparation, feature engineering, and careful selection of the best classification algorithms to improve diagnostic accuracy and clinical usefulness.

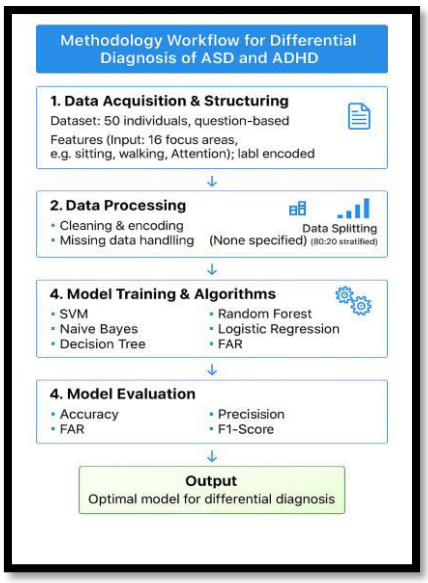


Figure 1. Sample questionnaire instrument used for data collection

3.1 Dataset

The data used for this study was structured and based on questions, involving 50 participants as a pilot sample that demonstrates the difficulties of collecting data on a small scale. The record for each participant contained information on various behavioral and functional aspects that were coded for each of the many different daily activities and phenotypic features, including the name, gender, and diagnostic category of the individual. Data collection took place within a clinical setting where everyone was exposed to a thorough professional assessment and the same diagnostic procedures. Even with these methodological controls, it's crucial to say out loud that the sample is not a good representation of the larger group of children with neurodevelopmental problems in terms of demographics or clinical characteristics. This can make it difficult to transfer the results to other populations. The dependent variable (Type of disorder)

was pre-coded into numbers. This helped the machine learning algorithms to work with, as they were able to directly identify categorical relationships. But the small sample size introduces significant statistical constraints, including reduced power, overfitting potential and potential training bias. Thus, the results must be interpreted with care; these results are preliminary and further validation tests with larger, more demographically varied data sets are needed to ensure reliability and external validity.
Essential motor milestones, functional skills and behavioral characteristics crucially relevant to the diagnostic distinction of ASD and ADHD were rigorously selected from the questionnaire data and will thereupon be called 'Focus Areas'. Such areas include: having difficulty sitting, standing, walking or using hands independently; being unable to communicate independently; having difficulties sleeping; difficulty attending for long periods of time; needing help to take a bath; dressing appropriately; eating less; carrying objects appropriately; feeding oneself appropriately; dressing up easily; holding onto an item properly; having difficulty with change; and having impulsive behavior. This rich set of features is crucial to capture the nuanced symptomatic manifestation and trajectory of development associated with neurodevelopmental disorders, and enables machine learning models to detect subtle yet clinically meaningful differences between ASD and ADHD.
We included additional diagnostic information for individuals with ASD, such as the presence of comorbid ADHD, a sub-population of individuals with ASD (Fernández-Delgado et al., 2024). This information was thoroughly abstracted from DSM-5 criteria or DSM-IV-TR criteria, and specific diagnostic instruments. This stringent approach ensures that the data is clinically valid and reliable for diagnosis, laying a solid foundation for subsequent machine learning analysis. People with comorbid ADHD should also be included as it is important for the models to learn and distinguish the subtle symptoms that often overlap. This helps the models to be more ready for diagnostic problems in practice. Moreover, sociodemographic factors such as gender and the presence of comorbid diseases were carefully assessed. This was done in an attempt to aid in the understanding of the diagnostic situation as a whole, and to give careful consideration to any factors which might have influenced the results. Dealing with these aspects is crucial in order to create classification models that are not only accurate but strong and can be applied to a variety of patient populations. This illustrates the importance of clinical and demographic considerations in the diagnosis of neurodevelopment.

Table 1. Processed dataset representation for ASD and ADHD (Sex: M = Male, F = Female; 1 = Yes, 0 = No)

SL	Sex	Disorder type	Sits indep.	Stands indep.	Walks indep.	Uses hands	Communicates	Sleeps more	Attends long	Needs bath help	Wears clothes	Eats less	Carried properly	Feeds self	Dressed easily	Grips properly	Struggles w/ change	Impulsive behaviour
1	M	1	1	1	1	0	0	0	0	1	1	0	1	1	1	1	0	1
2	M	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	1	0
3	M	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0
4	M	0	1	0	0	1	0	0	1	1	1	0	1	1	1	1	1	0
5	M	1	1	1	1	1	0	0	0	1	1	0	0	0	0	0	0	1
6	M	0	1	1	1	1	1	0	1	0	0	1	0	0	0	0	1	0
7	M	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0
8	M	1	1	1	1	1	0	0	0	1	1	1	1	1	0	1	0	1
9	M	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
10	M	1	1	1	1	1	0	0	0	1	0	1	0	0	0	0	0	1
11	M	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0
12	M	0	1	1	1	1	0	0	1	0	0	0	0	1	0	1	1	0
13	M	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	0
14	M	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	0	0
15	M	1	1	1	1	1	0	0	0	1	1	0	1	1	1	1	0	1
16	M	0	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	0

SL	Sex	Disorder type	Sits indep.	Stands indep.	Walks indep.	Uses hands	Communi-cates	Sleeps more	Attends long	Needs bath help	Wears clothes	Eats less	Carried properly	Feeds self	Dressed easily	Grips properly	Struggles w/ change	Impulsive behaviour
17	M	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
18	M	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	1	0
19	M	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	0
20	M	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	0
21	M	0	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1	0
22	M	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	0	1
23	M	1	1	1	1	1	0	0	0	1	1	0	1	1	1	1	0	1
24	M	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	1	0
25	M	0	1	1	1	1	0	1	1	0	0	0	1	1	0	0	1	0
26	M	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1
27	M	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1

SL	Sex	Disorder type	Sits indep.	Stands indep.	Walks indep.	Uses hands	Communicates	Sleeps more	Attends long	Needs bath help	Wears clothes	Eats less	Carried properly	Feeds self	Dressed easily	Grips properly	Struggles w/ change	Impulsive behaviour
28	M	0	1	1	1	1	0	1	1	0	0	0	1	1	0	0	1	0
29	M	0	1	1	1	1	0	1	1	1	0	1	1	1	0	0	1	0
30	M	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1
31	M	0	1	1	1	1	0	1	1	1	0	1	1	1	0	0	1	0
32	M	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1
33	M	0	1	1	1	1	0	1	1	1	0	1	1	1	0	0	1	0
34	F	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	1
35	F	0	1	1	1	1	0	0	1	1	1	0	1	0	0	0	1	0
36	F	0	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1
37	F	0	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0

SL	Sex	Disorder type	Sits indep.	Stands indep.	Walks indep.	Uses hands	Communi-cates	Sleeps more	Attends long	Needs bath help	Wears clothes	Eats less	Carried properly	Feeds self	Dressed easily	Grips properly	Struggles w/ change	Impulsive behaviour
38	F	0	1	1	1	0	0	0	0	1	1	0	1	0	0	0	1	0
39	F	1	1	1	1	1	0	0	0	1	0	1	1	1	1	1	0	1
40	F	0	1	1	0	1	1	0	1	1	1	0	0	1	1	1	1	0
41	F	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1
42	F	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	1
43	F	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1
44	F	0	1	1	0	1	1	0	1	1	1	0	0	1	1	1	1	0
45	F	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0
46	F	1	1	1	1	1	0	0	0	1	1	0	1	1	1	1	0	1
47	F	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0

SL	Sex	Disorder type	Sits indep.	Stands indep.	Walks indep.	Uses hands	Communi-cates	Sleeps more	Attends long	Needs bath help	Wears clothes	Eats less	Carried properly	Feeds self	Dressed easily	Grips properly	Struggles w/ change	Impulsive behaviour
48	F	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0
49	F	1	1	1	1	1	0	0	0	1	1	0	1	1	1	1	0	1
50	F	0	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0

3.2 Data preprocessing

Pre-processing processes were crucial in ensuring the quality of the data, the effectiveness of the model, and its suitability for machine learning algorithms. Cleaning and encoding: columns that didn't add any useful information were carefully eliminated to get rid of unnecessary data, cut down on noise, and stop possible overfitting. This improved the efficiency and understanding of the model. The label for supervised learning is correctly separated from the objective variable, "Type of disorder."

- Handling missing data: a careful inspection of data did not result in any missing data and no imputation techniques were required.
- Feature selection: Standard Scaler was used for feature scaling to make sure that all features contribute equally to the model's learning process and to keep features with broader numerical ranges from having too much of an effect on the results. This method changed the characteristics so that their average was 0 and their standard deviation was 1. This standardisation is particularly useful in distance-based classifiers such as Support Vector Machines, where the scale of the features has a direct impact on the distances computed and thus the performance of the model.
- Data splitting: the standardized data was divided in a systematic manner into training and testing sets such that 80% of data was assigned to the training set and the remaining 20% to the testing set. Stratified splitting was used to ensure that the original class proportions of ASD and ADHD remained the same in both subsets. This approach is crucial to ensure that the model is trained and tested on representative samples of the entire population. This prevents any bias towards the majority class (if there is any imbalance) and facilitates easier generalization to new data.

3.3 Machine learning algorithms

A diverse set of supervised learning algorithms was employed to get a comprehensive evaluation of how well machine learning can differentiate between ASD and ADHD. These algorithms were selected as they are effective in classification tasks and are able to detect complex patterns in sets of data that are not uniform. This is of great relevance when developing diagnostic models that are reliable. The algorithms selected are known to have unique pattern recognition capabilities, enabling a comparison analysis to determine their performance in predicting in this specific clinical context.

The aim of this study was to investigate Support Vector Machine, which is very effective in handling high dimensional data and optimal separating hyperplane. Very helpful for distinguishing between similar diagnostic categories for relatively nuanced responses to questions. In addition to SVM, the following methods were chosen for their complementary properties and widely used in biomedical classification:

- Random Forest
- K-Nearest Neighbor
- Naïve Bayes

- Logistic Regression
- Linear Discriminant Analysis
- Decision Tree
- Artificial Neural Network

3.4 Performance evaluation metrics

To provide a thorough and clinically relevant analysis of the effectiveness of each model at differentiating between ASD and ADHD, a wide range of standard classification metrics were carefully selected for their ability to provide detailed insights into diagnostic accuracy and reliability, and were carefully evaluated on the held-out test set:

Accuracy: measures the percentage of the instances correctly predicted.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

Precision: how many of the persons we said were positive are really positive.

$$Precision = TP / (TP + FP)$$

Recall: measures the proportion of actual positives that were correctly identified.

$$Recall = TP / (TP + FN)$$

F1-Score: harmonic mean of Precision and Recall, useful for imbalanced data sets.

$$F1\text{-}Score = 2 \times (Precision \times Recall) / (Precision + Recall)$$

False Alarm Rate (FAR): the percentage of cases that are negative but predicted to be positive.

$$FAR = FP / (FP + TN)$$

3.5 Experimental environment

The experiment was conducted on a common personal computer with the widely-used programming language Python within a Jupyter Notebook environment, which allows the process to be iterative and easily replicable. The powerful and efficient libraries Pandas and NumPy were used to change and prepare the data. The model was implemented using the Scikit-learn machine learning library, which is a popular and flexible machine learning tool, and the metrics were carefully evaluated. This ensured the methods were valid and that they could be applied in a variety of scenarios. This uniform computing structure enabled algorithms to be executed uniformly and results to be compared reliably among the models. This rendered the findings more valid.

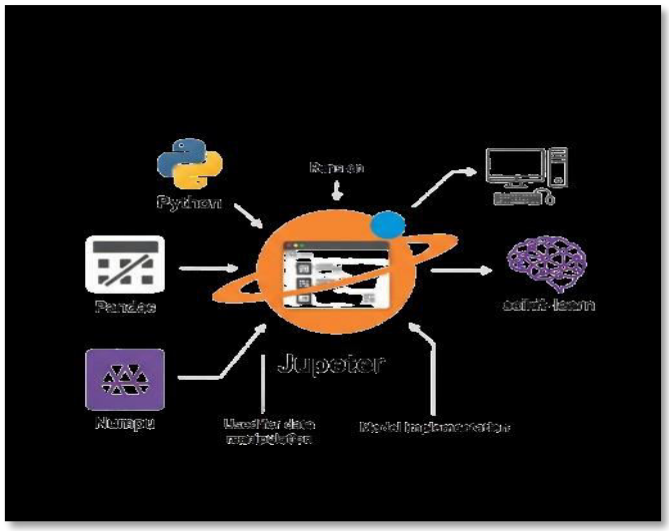


Figure 2. Experimental environment for machine learning model development

The experimental environment, as depicted in Figure 2, was meticulously structured to ensure both methodological rigor and reproducibility.

4. Results

This section provides a comprehensive analysis of the results obtained from the robust testing of the different machine learning models, with a special emphasis on the findings from the performance of models in the chosen classification metrics and their ability to differentiate between ASD and ADHD. The accuracy, precision, recall, F1-Score and false alarm rate (FAR) for each algorithm are compared in order to gain a complete insight into the strengths and limitations of the algorithms in this diagnostic context.

4.1 Model efficiency, robustness, and generalizability considerations

The model was tested and assessed on test data using eight different machine learning algorithms. Table 2 provides a summary of the results.

Table 2. Performance metrics of machine learning algorithms for ASD and ADHD classification

Metric	SVM	RF	KNN	NB	LR	DT	LDA	ANN
Accuracy (%)	93.33	100	93.33	93.33	100	80	100	86.67
Precision (%)	85.71	100	85.71	100	100	100	100	83.33
Recall (%)	100	100	100	83.33	100	50	100	83.33
F1-Score (%)	92.30	100	92.30	90.90	100	66.67	100	83.33
FAR (%)	11.11	0	11.11	0	0	0	0	11.11

In addition to these individual metric scores, the comparative analysis of the algorithms from Table 2 provides insights into the efficiency and robustness of the models,

especially in light of the characteristics of the dataset. A notable finding is that the relatively less complex Logistic Regression (LR) and Linear Discriminant Analysis (LDA) models performed as well as the more complex ensemble model Random Forest (RF) with respect to all metrics, giving 100% accuracy. This is important because LR and LDA are more interpretable and less computation-intensive than RF or Artificial Neural Networks (ANN). They were able to accomplish a 100% differentiation, which indicates that the features extracted from the dataset of questions have a strong discriminative power and hence the classification task is very separable in the feature space. These simpler but equally effective models would be beneficial in some cases where computational resources are limited or when model interpretability is a primary concern for clinical use.

The results show that while certain models were able to achieve 100% accuracy, precision, recall, and F1-Score on the test set, it is important to exercise caution in interpreting the generalizability of these models. Since the data set is small (N = 50), there is a risk of overfitting, meaning that the models fit the data too well and may not perform as well on diverse data that they were not trained on. Stratified splitting was used to make sure to have representative subsets, but the perfect scores might be due to the dataset's particular properties and not necessarily a general robustness. Future studies with larger and more heterogeneous datasets will be important to fully support the generalizability of these high-performing models and their consistent accuracy in more general patient populations. However, the present study shows that with small sample sizes, machine learning is a powerful initial method to discover highly effective diagnostic predictors of ASD and ADHD.

4.2 Comparative performance analysis

The results of testing eight different machine learning algorithms showed considerable differences in their effectiveness for the differential diagnosis of ASD and ADHD and could be clustered in different performance levels. Three algorithms, namely Random Forest (RF), Logistic Regression (LR) and Linear Discriminant Analysis (LDA), formed the top tier and showed exceptional performance with 100% on all major classification metrics: accuracy, precision, recall and F1-Score (Table 2). This consistently strong performance on the test set is a clear indication of their strong ability to discriminate between ASD and ADHD using the questionnaire information available. The results are summarised in Table 3.

Table 3. Comparative analysis focus and key insights

Analysis focus	Derived result / insight	Supported by
Model tiers	Clearly defined three tiers of models: top tier (100% in all core metrics: RF, LR, LDA), mid tier (high accuracy, balanced scores: SVM, KNN, NB, ANN), and lower tier (lowest accuracy: DT).	Data from all metrics.

Analysis focus	Derived result / insight	Supported by
Outperformance	Specifically identifies which models perfectly solve the classification on the test set (RF, LR, LDA — all 100% in accuracy, precision, recall, F1-score). This is a strong finding.	Accuracy, precision, recall, F1-score of 100%.
Generalization score	Uses the F1-score as the single best indicator of a model's stability, as it is the harmonic mean of precision and recall. The highest F1-score models (RF, LR, LDA: 100%) show the best balance of these two measures.	F1-score column.

After that, models such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB) and Artificial Neural Network (ANN) also gave good performance with an accuracy range of 86.67% to 93.33% (Table 2). These models were very effective, but with some differences in their F1-Scores and other measures, suggesting a slightly less well-rounded performance than the best models. Only the Decision Tree (DT) algorithm sat in the lower tier, achieving the lowest accuracy of 80% with significantly lower F1-Score and Recall, indicating that it is the least effective algorithm for the diagnostic task at hand amongst the algorithms evaluated. The F1-Score, being the harmonic mean of Precision and Recall, also emphasized the overall stability and balance in the predictive power of RF, LR and LDA (all 100% F1-Score), which continued to occupy the leading positions in this comparative analysis.

4.3 Error-specific insights (false alarms vs. misses)

A careful analysis of the error measures, such as Recall and False Alarm Rate (FAR), can give important clues about the clinical consequences of the performance of each model. In a diagnostic situation, minimising both the false positive (Type I error) and the false negative (Type II error) is crucial (Table 4).

Table 4. Error analysis and diagnostic reliability of ML models

Error metric	Derived result / insight	Supported by
False alarm rate (FAR)	Identifies the models that are most clinically reliable by having a FAR of 0%. These models (RF, LR, DT, LDA) never incorrectly classified a non-disordered case as positive, indicating high specificity.	FAR column of 0.
False negatives (misses)	By looking at recall, false negatives can be inferred. Models with 100% recall (SVM, RF, KNN, LR, LDA) never miss an actual positive case (zero false negatives), which is the most critical factor in diagnostic screening.	Recall column of 100%.

Error metric	Derived result / insight	Supported by
The "weakest" model	The Decision Tree (DT) has the lowest recall (50%) and lowest F1-score (66.67%), meaning it missed 50% of the actual positive cases. This makes it the least reliable model for screening, despite having a 0% false alarm rate.	Recall (50%) and F1-score (66.67%).

The specificity of several models was very good (0% false alarm rate). These are Random Forest, Naïve Bayes, Logistic Regression, Decision Tree and Linear Discriminant Analysis (Table 2). A 0% FAR is a very desirable characteristic of a diagnostic tool, because it means that those who do not have a neurodevelopmental disorder have a low chance of being classified as such by the tool, which would otherwise lead to unnecessary clinical follow-up and increased anxiety for the patient. On the other hand, SVM, KNN and ANN had a FAR of 11.11%, pointing towards a small percentage of false positive classifications.

At the same time, and perhaps even more important in diagnostic screening, is Recall — the model's ability to correctly label true positives (that is, children who actually have ASD or ADHD) and thereby prevent dangerous false negatives. SVM, Random Forest, K-Nearest Neighbor, Logistic Regression and LDA all obtained an excellent result on Recall with a score of 100% (Table 2). This means that these models were able to correctly identify all of the children with a neurodevelopmental disorder in the test set, without missing any children who actually have one. Such a characteristic is extremely useful for early detection and intervention efforts.

The Decision Tree algorithm, with a 0% FAR, had a much lower Recall of 50%. This means that it was also least likely to detect the actual cases, and more prone to false negatives than other models, and is therefore not the most reliable when used for a screening application where it is important that identification of actual cases is not missed. Overall, the models that achieved a 0% FAR and 100% Recall (Random Forest, Logistic Regression and Linear Discriminant Analysis) appear to be the most clinically robust in balancing the prevention of misdiagnosis and comprehensive detection.

5. Discussion

High performance of some machine learning models in discriminating between ASD and ADHD, especially those with 0% FAR and 100% Recall, make them attractive for use in clinical diagnostic procedures. The main goal of this study was to build and test a machine learning model to enable early identification and differentiation of children with Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD) using a question-based dataset. The results provide strong evidence of the potential of machine learning for improving diagnostic accuracy, which could be a viable way to overcome the limitations in conventional diagnostic procedures.

5.1 Comparative performance and model efficacy

The results of this comparative analysis of eight supervised learning algorithms were very interesting. Interestingly, three classifiers, namely Random Forest (RF), Logistic Regression (LR), and Linear Discriminant Analysis (LDA), consistently obtained remarkable performance with 100% accuracy, precision, recall and F1-Score on all the main classification metrics (Table 2). This excellent classification on the held-out test set is an indicator of their strong capability to distinguish between ASD and ADHD using the behavioural and functional characteristics recorded in the questionnaire. It is interesting to see that the simpler models (LR and LDA), which offer more clarity about the classification, achieved the same 100% accuracy as the more intricate ensemble method (RF). This indicates that the items carefully extracted from the questionnaire have good discriminative power that can even be separated by less complex algorithms between the two neurodevelopmental disorders. This finding resonates with the pursuit of transparency and simplicity in models when data attributes allow it, particularly in clinical contexts where decision-making clarity is appreciated. The comparatively poor performance of the Decision Tree (DT) algorithm (80% accuracy, 50% Recall) indicates that it is not an appropriate algorithm to capture the nuanced patterns required for robust differentiation of this diagnostic task.

5.2 Error-specific insights and clinical utility

A thorough analysis of the error metrics reveals the potential clinical applicability of the best-performing models. A number of algorithms (Random Forest, Logistic Regression, Linear Discriminant Analysis) had a FAR of 0%, indicating that there were no false positives on the test set. This is an important property of any diagnostic instrument, as it reduces the chances of overdiagnosis of typically developing children, which can lead to unwarranted parental anxiety and clinical interventions.

Additionally, these three models (along with SVM and KNN) showed 100% Recall. This implies the ability to accurately detect 100% of those children in the test set who were truly diagnosed with ASD or ADHD, and that no children who were truly diagnosed with either condition were missed. The emphasis in early detection and intervention for neurodevelopmental disorders is on avoiding false negatives, as missed diagnoses can result in delayed or insufficient support and services and may ultimately impact the developmental trajectory. RF, LR, and LDA achieving 0% FAR and 100% Recall makes them particularly strong candidates for a preliminary screening tool or differential diagnostic aid useful for early clinical decision-making, where both specificity and sensitivity are important. Compared to traditional subjective diagnostic techniques, the significant improvement of these results is clearly shown, consistent with previous studies where high variability and delays were observed.

5.3 Addressing methodological challenges and generalizability

The results reached are very promising, but the necessity to reflect on these results within the methodological framework and limitations of the study is paramount. The

use of a small dataset (50 individuals) raises the question of generalizability of the perfect scores achieved for RF, LR and LDA. While high accuracy on a small dataset can sometimes indicate overfitting — where the model learns noise or patterns specific to the training data rather than generalizable patterns — stratified splitting was used to reduce bias in data partitioning. Further validation of these models with larger, more diverse, and multi-center datasets encompassing broader demographic, cultural, and clinical presentations would be definitive.

The 'Focus Areas' derived from the question-based dataset were able to significantly differentiate between ASD and ADHD, supporting the literature on the value of behavioral and functional characteristics in the diagnosis of neurodevelopmental disorders. However, the present model is based on questionnaire data only. One of the major opportunities for future ML applications in neurodevelopmental disorders would be to incorporate multi-modal data (e.g., eye-tracking, EEG, fMRI) to improve diagnostic accuracy and provide a more comprehensive picture of these complex disorders.

5.4 Implications for AI in healthcare

This study makes a meaningful contribution to the field of AI in healthcare, especially with regard to pediatric neurodevelopment. This work has demonstrated the effectiveness of machine learning in the diagnosis of ASD and ADHD, and provided a potentially cost-effective, scalable, and objective method of diagnosis. In regions with limited resources, where access to specialized diagnostic expertise is limited, such a tool may be especially impactful, filling a critical gap in healthcare access and allowing for earlier, more equitable interventions. The framework created here paves the way for further innovation and provides a foundation for developing user-friendly clinical decision-support systems that can enable caregivers and clinicians to make data-driven decisions. The transition to computational diagnostics is an important step toward the transformation of neurodevelopmental healthcare, allowing for timely, accurate and accessible diagnosis, and improving developmental outcomes for children globally.

6. Conclusion

This study was successful in demonstrating that machine learning models can accurately differentiate between ASD and ADHD using questionnaire data. This study explored and compared the application of different machine learning algorithms to improve the differential diagnosis between children with Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder (ADHD) using structured questionnaire data. These two neurodevelopmental diseases are extremely complex to diagnose, with a great deal of overlap in symptoms, and most diagnosis is done through subjective means, which further highlights the importance of developing more objective and precise diagnostic tools.

The major findings of this inquiry are summarised below:

- Multiple machine learning models, including Random Forest, Logistic Regression, and Linear Discriminant Analysis (LDA), achieved exceptional performance, all with 100% accuracy, precision, recall, and F1-Score on the test set. This perfect score demonstrates their high performance for discriminating between ASD and ADHD.
- The high-performing models were also optimal for clinical utility, achieving both 100% Recall and a 0% False Alarm Rate — a double benefit in clinical use, reducing false-positive readings (which lead to misdiagnoses and unnecessary interventions) while avoiding false-negative readings (which would leave actual cases undetected).
- The study confirmed that features derived from structured questionnaires, after proper preprocessing and using advanced machine learning tools, can be highly discriminative for differentiation of neurodevelopmental disorders and can provide an objective alternative to subjective clinical observations.

As the results show, machine learning has great promise in tackling this diagnostic challenge. When studying eight different algorithms, a clear separation in performance was observed:

- Random Forest, Logistic Regression and Linear Discriminant Analysis models performed at the top of the line, with 100% accuracy on all primary classification metrics. The high performance of the simpler models (LR and LDA), matching the more complex Random Forest, points to good discriminative power of the selected features and to model interpretability and computational efficiency, which makes them attractive for clinical use.
- Support Vector Machine, K-Nearest Neighbor, Naïve Bayes and Artificial Neural Network also performed well, with accuracy between 86.67% and 93.33%, though with slightly less balanced performance than the top models.
- The Decision Tree algorithm was found to be the least effective, with an accuracy of 80% and a significantly poorer F1-Score and Recall, making it less suited for this specific diagnostic task.
- Error-specific analysis showed that RF, NB, LR, DT, and LDA achieved a 0% False Alarm Rate, indicating minimal misclassification of neurotypical individuals, while SVM, RF, KNN, LR, and LDA recorded 100% Recall, indicating that all real cases of neurodevelopmental disorder were correctly identified — a crucial factor in diagnostic screening.
- The high performance of multiple models on this small-scale dataset gives a strong initial indication that ML is a powerful tool for identifying good diagnostic predictors, but results must be treated with caution given the risk of overfitting. Studies with larger and more varied datasets are needed to further establish the generalizability and robustness of these findings.

The implications of these findings are significant for the field of paediatric neurodevelopmental healthcare. This machine learning technology could prove to be a powerful, accessible, and potentially cost-effective diagnostic tool to mitigate diagnostic

delays, bridge the gap between available and specialist expertise, and provide earlier, targeted intervention. This is especially important in low-resource areas, where it could help caregivers and clinicians alike improve developmental outcomes for children worldwide.

Acknowledgements

The authors would like to express their heartfelt thanks to all the persons and institutions that helped them achieve the successful completion of this research.

The Child Development Institute and Mymensingh Welfare School have been very helpful and cooperative in conducting the data collection process. They accepted this study very readily, which was a key factor in establishing the basic dataset.

This project received generous support and funding from the Research and Extension Center of Jatiya Kabi Kazi Nazrul Islam University for the Fiscal Year 2024–25 allocated for this purpose. Such financial support was essential to allowing for the research and analysis activities outlined in this paper.

Declaration about conflict of 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.

References:

1. Aasim, S., Butt, M. A., Rana, S. and Banal, R. (2024). Bridging the diagnostic gap in neuropsychiatry: the role of neuro acoustic therapies in accurate diagnosis. Research Square.
2. Alsharif, N., Al-Adhaileh, M. H., Alsubari, S. N. and Al-Yaari, M. (2024). ADHD diagnosis using text features and predictive machine learning and deep learning algorithms. Deleted Journal, 3(7).
3. Alves, C. L., Martinelli, T., Sallum, L. F., Rodrigues, F. A., Toutain, T. G. L. de O., Porto, J. A. M., Thielemann, C., Aguiar, P. de C. and Moeckel, M. (2024). Multiclass classification of autism spectrum disorder, attention deficit hyperactivity disorder, and typically developed individuals using fMRI functional connectivity analysis. PLoS ONE, 19(10).
4. Ardulov, V., Martínez, V. R., Somandepalli, K., Zheng, S., Salzman, E., Lord, C., Bishop, S. and Narayanan, S. (2021). Robust diagnostic classification via Q-learning. Scientific Reports, 11(1).
5. Bertacchini, F., Demarco, F., Scuro, C., Pantano, P. and Bilotta, E. (2023). A social robot connected with ChatGPT to improve cognitive functioning in ASD subjects. Frontiers in Psychology, 14.
6. Federico, A., Zgodic, A., Flory, K., Hantman, R. M., Eberth, J. M., McLain, A. C. and Bradshaw, J. (2023). Predictors of autism spectrum disorder and ADHD: results from the National Survey of Children's Health. Disability and Health Journal, 17(1), 101512.

7. Fernández-Delgado, M., Cruz, S., Cernadas, E., Alateyat, H., Tubío-Fungueiriño, M., Sampaio, A., Carracedo, Á. and Prieto, M. F. (2024). Population-based detection of children ASD/ADHD comorbidity from atypical sensory processing. Applied Intelligence, 54(20), 9906.
8. Grazioli, S., Crippa, A., Buo, N., Ceccarelli, S. B., Molteni, M., Nobile, M., Salandi, A., Trabattoni, S., Caselli, G. and Colombo, P. (2024). Use of machine learning models to differentiate neurodevelopment conditions through digitally collected data: cross-sectional questionnaire study. JMIR Formative Research, 8.
9. Grazioli, S., Crippa, A., Rosi, E., Candelieri, A., Ceccarelli, S. B., Mauri, M., Manzoni, M., Mauri, V., Trabattoni, S., Molteni, M., Colombo, P. and Nobile, M. (2023). Exploring telediagnostic procedures in child neuropsychiatry: addressing ADHD diagnosis and autism symptoms through supervised machine learning. European Child and Adolescent Psychiatry, 33(1), 139.
10. Jaiswal, S., Valstar, M., Gillott, A. and Daley, D. (2017). Automatic detection of ADHD and ASD from expressive behaviour in RGBD data. IEEE International Conference on Automatic Face and Gesture Recognition, 762.
11. Jugunta, S. B., El-Ebiary, Y. A. B., Saravanan, K., Prasad, K. S. R., Koteswari, S., Rachapudi, V. and Rengarajan, M. (2023). Unleashing the potential of artificial bee colony optimized RNN-Bi-LSTM for autism spectrum disorder diagnosis. International Journal of Advanced Computer Science and Applications, 14(11).
12. Liu, X., Han, T., Jiang, Z., Hu, L., Li, W. and Song, C. (2024). Machine learning-based prediction of intellectual disability in children with autism spectrum disorder: using behavioral observation techniques. Research Square.
13. López, C. Q., Vera, V. D. G. and Quintero, M. P. (2025). Diagnosis of ADHD in children with EEG and machine learning: systematic review and meta-analysis. Clínica y Salud, 36(2), 109.
14. Mao, Y. B., Qi, X., He, L., Wang, S., Wang, Z. and Wang, F. (2025). Advanced machine learning techniques reveal multidimensional EEG abnormalities in children with ADHD: a framework for automatic diagnosis. Frontiers in Psychiatry, 16.
15. Mikoláš, P., Vahid, A., Bernardoni, F., Süß, M., Martini, J., Beste, C. and Bluschke, A. (2022). Training a machine learning classifier to identify ADHD based on real-world clinical data from medical records. Scientific Reports, 12(1).
16. Mukherjee, P., Gokul, R., Sadhukhan, S., Godse, M. and Chakraborty, B. (2023). Detection of autism spectrum disorder (ASD) from natural language text using BERT and ChatGPT models. International Journal of Advanced Computer Science and Applications, 14(10).
17. Qin, L., Wang, H., Ning, W., Cui, M. and Wang, Q. (2024). New advances in the diagnosis and treatment of autism spectrum disorders. European Journal of Medical Research, 29(1).

18. Rahman, K. K. M. and Subashini, M. M. (2022). Investigation of eye-tracking scan path as a biomarker for autism screening using machine learning algorithms. Diagnostics, 12(2), 518.
19. Rajagopalan, S. S. and Tammimies, K. (2024). Predicting neurodevelopmental disorders using machine learning models and electronic health records – status of the field. Journal of Neurodevelopmental Disorders, 16(1).
20. Saakyan, W., Norden, M., Eversmann, L., Kirsch, S., Lin, M., Guendelman, S., Dziobek, I. and Drimalla, H. (2025). Improving autism detection with multimodal behavioral analysis. Lecture Notes in Computer Science, 252.
21. Schulte-Rüther, M., Kulvičius, T., Stroth, S., Roessner, V., Marschik, P. B., Kamp-Becker, I. and Poustka, L. (2021). Using machine learning to improve diagnostic assessment of ASD in the light of specific differential diagnosis. medRxiv.
22. Shahiduzzaman, K. Md., Islam, S. and Mohua, S. S. (2024). covMAG: patient allocation management tool for infectious disease based on data mining with machine learning algorithms for COVID-19. Journal of Nazrul University, 10(1), 163.
23. Solek, P., Nurfitri, E., Sahril, I., Prasetya, T., Rizqiamuti, A. F., Burhan, B., Rachmawati, I., Gamayani, U., Rusmil, K., Chandra, L. A., Afriandi, I. and Gunawan, K. (2025). The role of artificial intelligence for early diagnostic tools of autism spectrum disorder: a systematic review. Turkish Archives of Pediatrics, 60(2), 126.
24. Song, C., Jiang, Z., Liu, D. and Wu, L. (2022). Application and research progress of machine learning in the diagnosis and treatment of neurodevelopmental disorders in children. Frontiers in Psychiatry, 13.
25. Sridhar, U. (2025). AI in healthcare: revolutionizing medical diagnosis. World Journal of Advanced Research and Reviews, 26(1), 3229.
26. Wolff, N., Kohls, G., Mack, J. T., Vahid, A., Elster, E. M., Stroth, S., Poustka, L., Kuepper, C., Roepke, S., Kamp-Becker, I. and Roessner, V. (2022). A data driven machine learning approach to differentiate between autism spectrum disorder and attention-deficit/hyperactivity disorder based on the best-practice diagnostic instruments for autism. Scientific Reports, 12(1).
27. Wu, Y. (2024). Systematic review of the effectiveness of machine learning in discriminating ADHD from ASD. Journal of Education Humanities and Social Sciences, 45, 601.
28. Zaheer, A. and Akhtar, A. (2025). Artificial intelligence as a support to diagnose ADHD: an insight of unorthodox approaches: a scoping review. Child Neuropsychology, 1.
29. Zhang, J., Zhang, Z., Sun, H., Ma, Y., Yang, J., Chen, K., Yu, X., Qin, T., Zhao, T., Zhang, J., Chu, C. and Wang, J. (2024). Personalized functional network mapping for autism spectrum disorder and attention-deficit/hyperactivity disorder. Translational Psychiatry, 14(1).

30. Zhao, C., Xu, Y., Li, R., Li, H. and Zhang, M. (2025). Artificial intelligence in ADHD assessment: a comprehensive review of research progress from early screening to precise differential diagnosis. Frontiers in Artificial Intelligence, 8.
31. Zhao, Z., Tang, H., Zhang, X., Qu, X., Hu, X. and Lu, J. (2021). Classification of children with autism and typical development using eye-tracking data from face-to-face conversations: machine learning model development and performance evaluation. Journal of Medical Internet Research, 23(8).
32. Zolghadr, Z., Batouli, S. A. H., Majd, H. A., Shafaghi, L. and Mehrabi, Y. (2024). fMRI-based multi-class DMDC model efficiently decodes the overlaps between ASD and ADHD. Basic and Clinical Neuroscience Journal, 15(3), 367.