Journal of Disability Research

2024 | Volume 3 | Pages: 1–15 | e-location ID: e20240003
DOI: 10.57197/JDR-2024-0003

EASDM: Explainable Autism Spectrum Disorder Model Based on Deep Learning

El-Sayed Atlam[1,2,*], Mehedi Masud[3], Mahmoud Rokaya[2,4], Hossam Meshref[3], Ibrahim Gad[2,*] and Abdulqader M. Almars[1]

[1]Department of Computer Science, College of Computer Science and Engineering, Taibah University, Yanbu 966144, Saudi Arabia
[2]Computer Science Department, Faculty of Science, University of Tanta, Tanta, Gharbia, Egypt
[3]Department of Computer Science, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia
[4]Department of Information Technology, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia

Correspondence to:
El-Sayed Atlam*, e-mail: satlam@yahoo.com, satlam@taibahu.edu.sa
Ibrahim Gad*, e-mail: ibrahim.gad@science.tanta.edu.eg

Received: September 5 2023; Revised: November 27 2023; Accepted: October 4 2023; Published Online: February 3 2024

ABSTRACT

A neuro-developmental disorder known as autism spectrum disorder (ASD) affects a significant portion of the global population. Those with ASD frequently struggle to interact and communicate with others and may engage in restricted or repetitive behaviors or interests. The symptoms of autism begin early in childhood and can continue into adulthood. Machine learning and deep learning (DL) models are employed in clinical research for the early identification and diagnosis of ASD. However, the majority of the existing models lack interpretability in their results for ASD diagnosis. The explainable artificial intelligence (XAI) concepts can be used to provide transparent and understandable explanations for models' decisions. In this work, we present an explainable autism spectrum disorder model based on DL for autism disorder detection in toddlers and children. The primary objective of this study is to better understand and interpret the classification process and to discern the significant features that contribute to the prediction of ASD. The proposed model is divided into two distinct components. The first component employs a DL model for autism disorder detection. The second uses an XAI technique known as shapley additive explanations (SHAP) to emphasis key characteristics and explain the model's outcomes. The model showed perfect performance on the training set, with an accuracy of 1 and a receiver operating characteristic score of 1. On the test set, the model achieved an accuracy score of 0.9886, indicating that it performed nearly as well as on the training set. The experimental results demonstrate that the proposed model has the capability to accurately predict and diagnose ASD while also providing explanatory insights into the obtained results. Furthermore, the results indicate that the proposed model performs competitively compared to the state-of-the-art models in terms of accuracy and F1-score. The results highlight the efficacy and potential of the proposed model in accurately predicting ASD in binary classification tasks.

KEYWORDS

explainable AI, autism spectrum disorder, decision tree, DL, logistic regression, ML

INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex condition that affects millions of people worldwide. According to the WHO, ASD affects 1 in 160 children at any given time worldwide (Mahmud et al., 2018). The symptoms of ASD include persistent difficulties in social communication, limited interests, and repetitive behavior (Tuchman et al., 2009; Maenner et al., 2021). In addition, people with ASD can exhibit several other characteristics such as delayed language skills, delayed movement skills, anxiety, stress, excessive worry, unusual mood or emotional reactions, etc. Early symptoms of this condition may appear at 3 years of age and may persist for the remainder of the person's life (Raj and Masood, 2020). Early diagnosis of ASD can be critical and beneficial for a patient's treatment. According to Ramana and Paolucci (Anirudh and Thiagarajan, 2021; Claudio et al., 2023), the diagnosis of ASD early in childhood can facilitate the development of social skills in children. Furthermore, children who receive medical care prior to turning two exhibit higher intelligence quotient (IQs) than those who do not receive it until later in life (Alkahtani et al., 2023; Georgoula et al., 2023).

The use of deep learning (DL) and machine learning technology has proven to be exceptionally effective in supporting the early diagnosis of ASD (Adilakshmi et al., 2023). In order to gain a better understanding of ASD, a variety of studies have been proposed, encompassing diverse approaches such as facial-feature extraction (Guillon et al., 2014; Aldhyani et al., 2022), eye-tracking techniques (Kanhirakadavath and Chandran, 2022), facial expression identification (Akter et al., 2017; Mujeeb Rahman and Subashini, 2022; Alkahtani et al., 2023), biomedical imaging processing (Jiang and Chen, 2008), and voice identification (Schelinski et al., 2016). For example, Yolcu et al. introduced a convolutional neural network (CNN) model designed for automated facial expression recognition and the detection of neurological disorders (Yolcu et al., 2019). Another suggested approach involves leveraging machine learning and DL technologies to develop an eye-tracking system, specifically designed to assist in the early screening of autism in children (Haq et al., 2022; Kanhirakadavath and Chandran, 2022). A novel temporal voice recognition system was proposed by Schelinski, which uses functional magnetic resonance imaging in order to investigate the neural mechanisms involved in voice recognition (Schelinski et al., 2016).

Although current studies have shown a notable capability in identifying ASD, it is noteworthy to mention that these studies often fall short in providing explicit explanations for the observed results. The interpretation of results holds significant value for clinicians as it facilitates their understanding of the decision-making process and enhances their diagnostic capabilities. Explainable artificial intelligence (XAI) is critical in bridging the gap between the complex internal operations of AI models and the critical need for human comprehension. It is accomplished through the development of approaches and strategies targeted at increasing the transparency and comprehension of AI models. By making AI models more explainable, clinicians can understand why AI models generate predictions, promoting meaningful interpretation and building trust (Haq et al., 2020).

This paper introduces a novel approach called the explainable autism spectrum disorder model (EASDM) for the detection of ASD in toddlers and children. The structure of the proposed EASDM model is divided into two distinct components. The first component incorporates a DL model for detecting ASD. The second component utilizes the XAI technique known as shapley additive explanations (SHAP) to highlight crucial characteristics and provide explanations for the model's predictions. The primary objective of this study is to develop a model capable of accurately identifying ASD while simultaneously offering interpretability and transparency in its outcomes. Incorporating explainable techniques into DL frameworks allows clinicians and researchers to gain insight into the underlying mechanisms contributing to the detection results of ASD, as well as into the important features and patterns used by the model to identify it. We used a publicly available dataset to evaluate the proposed model (Thabtah, 2017, 2018; Thabtah et al., 2018). The experimental results demonstrate that the model can identify ASDs and explain their outcomes. The main summarized points of this article are as follows:

- The EASDM model is proposed for the prediction of ASDs at an early stage.
- The SHAP technique is used to visually interpret individual predictions generated by the EASDM model, emphasize important features, and provide explanations for the model's predictions.
- A comparative case study was conducted using publically available datasets, and the results demonstrated the effectiveness of the proposed model.
- Comparison of the proposed model with state-of-the-art models in terms of accuracy and F1-score for comparative performance analysis); the results highlight the efficacy and potential of the proposed model in accurately predicting ASD in binary classification tasks.

The paper is organized as follows: The next section presents and discusses the related work on ASD. In the Explainable Autism Spectrum Disorder Model section, the proposed model for ASD classification is thoroughly explained. The ASD Data Collection section discusses ASD dataset description. The Experimental Evaluation section presents the experimental results and performance metrics. Finally, the Conclusion section provides the conclusion and outlines future directions.

RELATED WORK

In recent years, at an early stage, there has been a lot of focus on the analysis and classification of ASD detection. Machine learning, especially DL technologies, has shown remarkable performance in assisting in the early identification of ASD. A number of studies have been proposed for ASD detection that employ various techniques such as facial-feature extraction (Guillon et al., 2014), eye-tracking techniques (Kanhirakadavath and Chandran, 2022), facial expression identification (Mujeeb Rahman and Subashini, 2022), biomedical imaging processing (Jiang and Chen, 2008), and voice identification (Schelinski et al., 2016).

Thabtah and Peebles (2020) propose a novel machine learning model that utilizes induction rules for autism detection. The technique offers users knowledge bases and rules that provide insights into the model's classification decisions. In other articles, the variable analysis method is applied that identifies a small number of features for robust ASD classification (Howlader et al., 2018; Thabtah, 2018; Hossain et al., 2019). Decision trees (DTs) and logistic regression (LR), two machine learning techniques, are used to illustrate the efficacy of the suggested model. Akter et al. (2021) employed a combination of ML and ensemble techniques, along with feature transformation methods like standardization and normalization, to achieve more accurate autism detection. Their study's dataset was obtained from University of California, Irvine and Kaggle, and the findings showed that LR performed better than other classifiers.

Moreover, Erkan and Thanh (2019) conducted a study on similar datasets and explored the effectiveness of several ML methods such as k-nearest neighbours and random forest (RF) in identifying ASDs. In Bangladesh, Satu et al. (2019) employed tree-based classifiers to analyze and identify the key characteristics of "normal" or "autistic" patients. Duda et al. (2016) applied 6 ML classifiers on 65 items to examine ASD and attention deficit hyperactivity disorder. However, the dataset used in this study is limited and small.

Today, DL (Atlam et al., 2003; Malki et al., 2020a,b, 2021; Farsi et al., 2021) algorithms play a significant role in the classification of ASD. In several studies, it has been shown that DL is more effective in classifying ASD than

ML (Almars et al., 2021; Alwateer et al., 2021; Badawy et al., 2023). Raj and Masood (2020) demonstrated that DL methods, specifically CNNs, outperformed traditional ML methods for ASD detection in adults, children, and adolescents. The suggested model achieved impressive scores of 99.53, 98.30, and 96.88%, respectively. The long short-term memory-recurrent Neural Network model is proposed for automated ASD detection (Carette et al., 2018; Atlam et al., 2022; Noor et al., 2022). The dataset used in their study consists of prerecorded videos of children aged 8 to 10 years. According to the experimental results, the model achieved an impressive average accuracy and a maximum accuracy of 98%. However, the main limitation of the model is the possibility of overfitting due to the small dataset. Heinsfeld et al. (2018) investigated the functional brain patterns that aid in the diagnosis of ASD using the DL approach. The proposed model is evaluated and tested on resting-state functional magnetic resonance imaging data. Deep neural networks (DNNs) have also been proposed in several studies to comprehend the brain states of patients with ASD (Plis et al., 2014; Koyamada et al., 2015; Abraham et al., 2017). However, the aforementioned models achieved low accuracy scores in ASD classification.

Furthermore, in the realm of e-healthcare, recent studies have utilized clinical data in tandem with intelligent systems to identify various disease types. In a notable study, Haq et al. deployed two ensemble learning algorithms, namely Ada Boost and Random Forest, to perform feature selection (Haq et al., 2020). Additionally, the performance of these classifiers was scrutinized in comparison to wrapper-based feature selection algorithms. Subsequently, the classification of diseases was undertaken leveraging the DT methodology. Another investigation focused on the development of a diagnosis method for an early-stage disease using a CNN (Haq et al., 2022). This method improved the CNN model's predictive power by integrating data augmentation (DA) and transfer learning (TL) methods. The study sought to increase the model's accuracy in identifying early illness indications by applying these techniques. The Internet of Medical Things was also applied by Sultan Ahmad et al. to obtain high quality patient data (Ahmad et al., 2022). The study then introduces a hybridized methodology that leverages a combination of Gated Recurrent Unit and CNN as a classification model for disease diagnosis. In 2023, Almars et al. presented an intelligent system that employs the artificial gorilla troops optimizer in conjunction with DL and machine learning for the detection of ASD (Almars et al., 2023). While the experimental results of the aforementioned research demonstrate a promising application for deployment in the e-healthcare landscape, the algorithms lack explainability regarding model decisions.

Previous investigations focused on demonstrating a considerable competence in diagnosing ASD; nonetheless, it is worth noting that these studies frequently fell short in offering precise reasons for the observed results. The interpretability of algorithmic outcomes is a crucial consideration, particularly in healthcare, where transparency in decision-making is essential for building trust and fostering collaboration between medical professionals and advanced computational models. In other words, providing interpretations can assist them in understanding the decision-making process, thereby enhancing their diagnostic abilities XAI is crucial for bridging the gap between AI models' complex internal operations and the critical need for human comprehension. In the literature, only a few studies have been proposed to provide explanations of the model's outcomes (Kaiser et al., 2020; Payrovnaziri et al., 2020; Biswas et al., 2021). However, there is still space for improvement in this area. In this paper, we introduce a new approach called the Explainable Autism Spectrum Disorder model, which utilizes a DL model and explainable techniques for the detection of ASD in children.

EXPLAINABLE AUTISM SPECTRUM DISORDER MODEL

The purpose of this research is to identify ASD and provide an interpretation of model outcomes. To achieve that, the suggested model employs a DL model to detect ASD, and the XAI technique was used to identify the significant characteristics that aid the model in detecting ASD. Figure 1 shows the essential steps of the proposed model. It has three main components: (1) ASD data collection and preparation, (2) AI model building and training, and (3) model interpretation (performance analysis and interpretation).

ASD data collection

The ASD screening dataset was employed in this study for model construction, training, and evaluation. Therefor, two distinct datasets are collected from a public database maintained by Thabtah (2018) and Thabtah et al. (2018) that includes almost 1758 children and toddlers who were seen and reported. We have selected 15 relevant features from the 21 available in the datasets, 14 of these features have been utilized as inputs and one as output.

This toddler ASD screening dataset has significant features that can be used to improve the detection of ASD cases and identify autistic symptoms (Ribeiro and Guestrin, 2016; Alarifi and Young, 2018; Anirudh and Thiagarajan, 2021). In order to accurately differentiate between individuals with ASD and those without, Thabtah conducted a study in the field of behavioral science. In this study, Thabtah identified and recorded 10 distinct qualities (A1-A10) that were derived from an individual's behavior and other related characteristics.

We took into account this particular dataset and chose suitable characteristics because it was predominantly used by various researchers for their ASD investigations and because the most promising and necessary features needed to classify the data were taken into account (Ribeiro and Guestrin, 2016). In order to choose the most promising aspects that contribute to the prediction results, the explainable approach has also been used.

The data in Table 1 present information about individuals who have been assessed for ASD using the Autism Spectrum Quotient (AQ) questionnaire. The AQ questionnaire consists of 50 statements that assess different aspects of social and

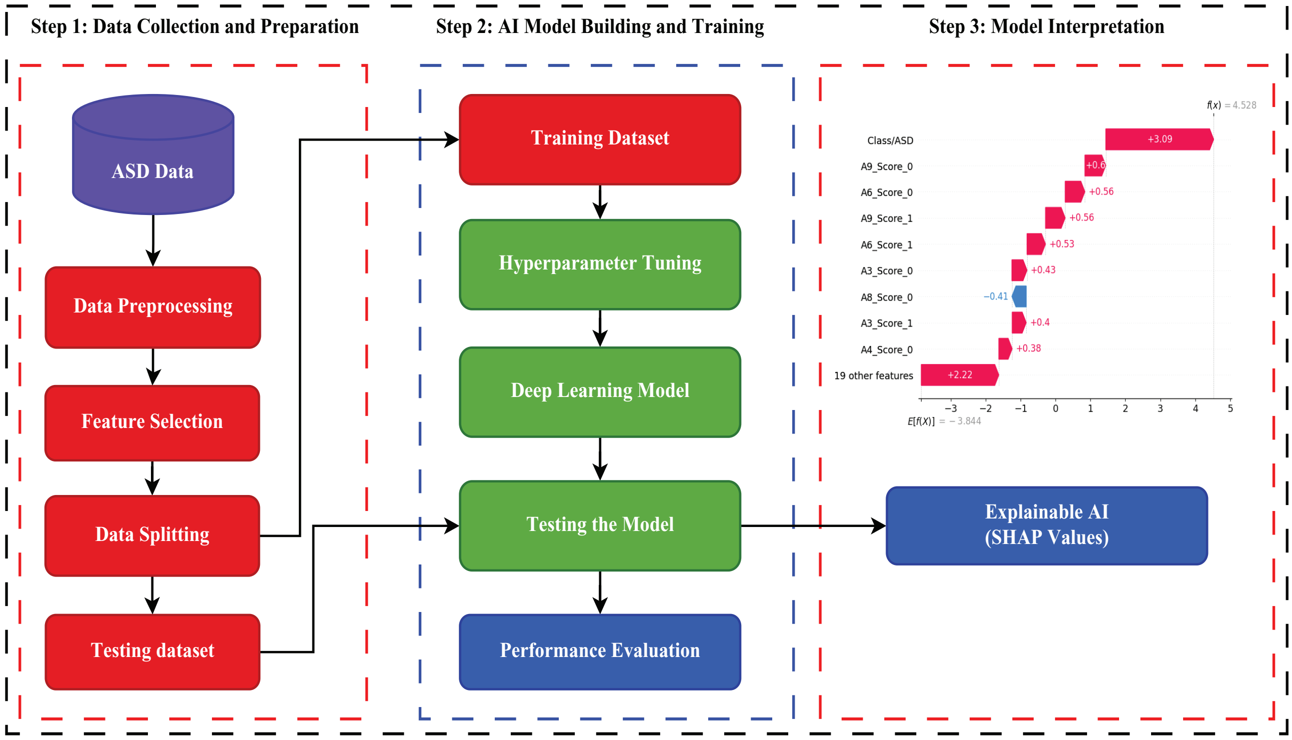

Figure 1: Framework for three sequential stages: (1) data collection and preparation (2) AI model building and training; and (3) model interpretation. Abbreviation: ASD, autism spectrum disorder.

Table 1: ASD data.

A1_score	A2_score	…	Gender	Ethnicity	Jaundice	Autism	Contry_of_res	…	Relation	Class/ASD
0	1	…	F	White-European	No	No	United States	…	Self	No
1	1	…	M	Latino	No	Yes	Brazil	…	Self	No
2	1	…	M	Latino	Yes	Yes	Spain	…	Parent	Yes

Abbreviation: ASD, autism spectrum disorder.

communication skills, as well as restricted and repetitive behaviors and interests.

The data include scores for each of the 10 questions on the AQ questionnaire (columns A1 through A10), as well as information about the individual's gender, ethnicity, whether they were born with jaundice, whether they have been diagnosed with autism, their country of residence, whether they have used an autism-related app before, their age group, their relationship to the person who completed the questionnaire, and whether they were diagnosed with ASD (the final column).

The first row of Table 1 shows the scores of the first individual, who is a female of White-European ethnicity living in the United States. She did not have jaundice at birth, has not been diagnosed with ASD, and has not used an autism-related app before. Her age group is "18 and more" and she completed the questionnaire herself. Her scores on the AQ questionnaire ranged from 0 (for A6 and A10) to 1 (for A2, A3, A4, A5, A7, A8, and A9).

The second row shows the scores of a male of Latino ethnicity living in Brazil. He did not have jaundice at birth, has been diagnosed with ASD, and has used an autism-related app before. His age group and relationship to the person who completed the questionnaire are the same as those of the first individual. His scores on the AQ questionnaire ranged from 0 (for A4, A6, and A10) to 1 (for A1, A2, A3, A5, and A9).

The third row shows the scores for a male of Latino ethnicity living in Spain. He was born with jaundice, has been diagnosed with ASD, and has not used an autism-related app before. His age group and relationship to the person who completed the questionnaire are the same as those of the first two individuals. His scores on the AQ questionnaire ranged from 0 (for A4 and A6) to 1 (for A1, A2, A3, A5, A7, A8, and A9).

The final column indicates whether each individual was diagnosed with ASD based on their scores on the AQ questionnaire. The first two individuals were not diagnosed with ASD, while the third individual was diagnosed with ASD.

The data of Table 2 present the frequency of categorical variables in a dataset. Each row of the table corresponds to a different variable, and the columns of the table provide information about that variable. The first column, labeled "Unique", shows the number of unique values that the variable can take on. For example, the "Ethnicity" variable has 11 unique values, meaning that there are 11 different ethnicities in the dataset. The second column, labeled "Top", shows the most frequently occurring value of the variable. For example, the most frequent ethnicity in the dataset is "White-European", occurring 233 times.

Table 2: Frequency of categorical variables.

	Unique	Top	Freq
Ethnicity	11	White-European	233
Jaundice	2	No	635
Autism	2	No	613
Contry_of_res	67	United States	113
Used_app_before	2	No	692
Age_desc	1	18 and more	704
Relation	5	Self	617

The third column, labeled "Freq", shows the frequency of the most frequently occurring value of the variable. For example, the value "no" occurs most frequently for both the "jaundice" and "autism" variables, occurring 635 and 613 times, respectively.

The data also reveal other interesting information, such as the fact that all individuals in the dataset are 18 years or older (age_desc = 1, 18 and more, 704 occurrences) and that most individuals have not used an autism-related app before (used_app_before = 2, no, 692 occurrences). Finally, the data show that the dataset contains individuals from 67 different countries, with the majority of individuals residing in the United States (contry_of_res = 67, United States, 113 occurrences).

Data preparation

These steps include three main phases: (1) data preprocessing, (2) feature selection/extraction, and (3) data splitting. In the following, we describe each phase in detail

- **Data preprocessing:** This step involves identifying and handling any inconsistencies, errors, or missing values present in the collected data. Techniques such as removing duplicates, imputing missing values, or correcting inconsistencies may be employed to ensure data integrity. Moreover, this step involves transforming the data to ensure compatibility with the chosen AI model. This step includes scaling numerical features and encoding categorical variables. Before using a DL model, it is essential to normalize the data because different attributes have different scales and values. All attribute data were normalized in the range [-1, 1] using the Z normalization approach, which removes the mean and scales the data to unit variance, as represented in Equation 1.

$$\text{Normalized Value} = \frac{X - \text{Mean}}{\text{Standard Deviation}} \tag{1}$$

- **Feature selection/extraction:** This step involves identifying the most relevant features that are crucial for the analysis. Statistical techniques, domain expertise, or feature importance methods may be utilized to select or extract informative features that contribute significantly to the study objectives.
- **Data splitting:** The preprocessed dataset is divided into subsets for training, testing, and validation. Common approaches include random splitting or stratified sampling to ensure representative subsets for each phase. The dataset was split into training and testing sets, with a ratio of approximately 70% training data and 30% testing data.

AI model building and training

In this step, a DL model is employed as the primary approach for classifying ASDs. The proposed model in this study is specifically designed to leverage the power of neural networks and learn intricate patterns and relationships within the ASD-related data. The details of the building and training steps are presented in the following sections.

Building step

In this study, a DNN is employed as the DL model for classifying ASDs. DNNs, also known as feedforward neural networks, are widely recognized for their effectiveness in various tasks, including classification and regression. The DNN architecture consists of multiple layers, which include an initial input layer, one or more intermediate hidden layers, and a final output layer. Every layer is composed of a collection of interconnected artificial neurons or nodes. In a DNN, information flows in a unidirectional manner from the input layer through the hidden layers to the output layer as shown in Figure 2.

The hidden layers of the DNN play a crucial role in learning and capturing complex patterns and representations

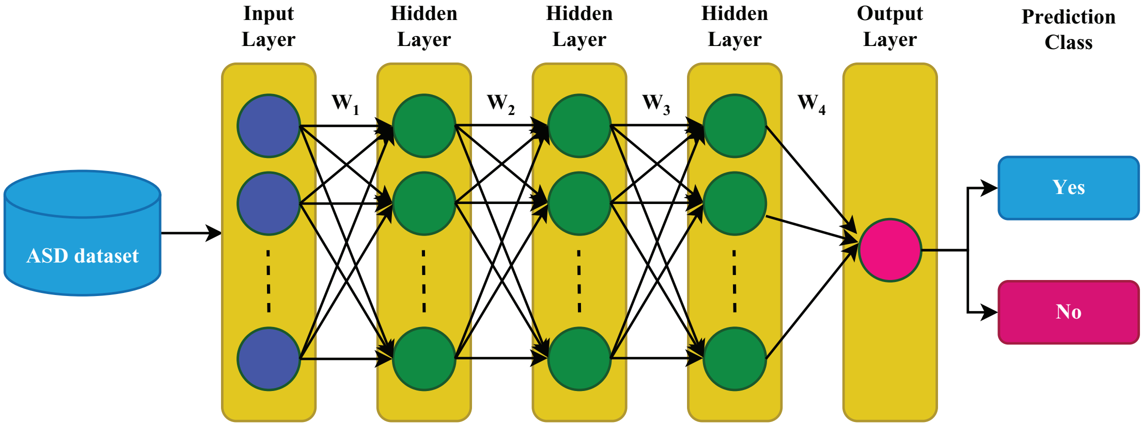

Figure 2: A DNN model for ASD classification. Abbreviations: ASD, autism spectrum disorder; DNN, Deep Neural Network.

within the input data. Every hidden layer neuron receives weighted inputs from the previous layer, processes them using an activation function, and passes the transformed information to the subsequent layer. This process allows the DNN to progressively extract higher-level features and representations as the data propagate through the network.

The implementation of the DNN model is typically facilitated by DL frameworks like TensorFlow or Keras. These frameworks provide a rich set of tools, functions, and libraries for constructing, training, and evaluating DNN models efficiently. The DNN model is a sequential neural network with six layers, including four hidden layers with 10 nodes each and a rectified linear unit (ReLU) activation function, followed by a fifth hidden layer with five nodes and a ReLU activation function, and finally an output layer with a single node and a sigmoid activation function. The ReLU activation function is commonly used in neural networks because of its ability to introduce nonlinearity into the model, which can improve its performance in classifying complex patterns in the data. The sigmoid activation function, on the other hand, is used to produce a probability output for the binary classification task.

The Adam optimizer is used for optimization, which is a popular optimization algorithm used for DL models. In this study, the binary cross-entropy loss function is employed for ASD classification task, which is a commonly used loss function in classification tasks. In addition, the accuracy metric, which is a commonly used statistic for classification tasks, is used to evaluate the model's performance. The "compile()" function in the Keras library is utilized to specify the optimizer, loss function, and evaluation metric for the model.

Training step

To train the proposed model, a diverse and representative dataset of individuals with and without ASD is utilized. This dataset comprises a comprehensive range of relevant features, such as demographic information, behavioral assessments, and neuroimaging data, which are essential for accurate classification.

The model is trained using a well-established optimization algorithm, such as stochastic gradient descent, to reduce the classification error and fine-tune the model's parameters. To train a model on the training data using a custom number of epochs and batch size, the "fit()" function in Keras is used. The model is evaluated on the validation data after each epoch to monitor its performance during training. The number of epochs and batch size are adjusted to improve the model's performance. Evaluation of the DNN's performance involves various metrics such as accuracy, recall, precision, and F1-score. The model's generalization ability and robustness may be assessed using techniques like cross-validation, where the dataset is split into multiple subsets for training and validation.

Model interpretation

The XAI framework is utilized to identify ASD and provide meaningful interpretations of the outcomes generated by the model. The framework leverages advanced AI techniques to analyze and interpret complex data patterns associated with ASD.

During the training phase, the model learns to recognize patterns and relationships within the data that are indicative of ASD. This process involves optimizing the model's parameters in order to reduce the difference between its predictions and the actual labels of ASD. Once the model is trained, the XAI framework focuses on providing interpretability of the model's outcomes. This is achieved through various techniques designed to shed light on the decision-making process of the model and the factors driving its predictions.

Feature importance analysis is a common interpretability technique in XAI. This analysis aims to identify the features that have the most significant impact on the model's predictions. By quantifying the contribution of each feature, researchers and clinicians can gain insights into the factors that are most relevant for identifying ASD. Additionally, the XAI framework may use visualization methods to provide a more intuitive understanding of the model's outcomes. This can include visual representations of the model's internal workings, such as heat maps highlighting regions of importance in neuroimaging data or DTs illustrating the decision rules employed by the model.

In this study, SHAP (shapley additive explanations) force plot is the technique used to visually interpret individual predictions generated by the EASDM model. It is designed to provide a detailed explanation of the factors contributing to a specific prediction for a given instance. The SHAP force plot displays the features that influence the prediction and their corresponding SHAP values, which represent the contribution of each feature to the prediction. The SHAP force plot also includes a reference value, which represents the average prediction for the entire dataset. In the next section, we visualize the interpretation of the model's predictions.

EXPERIMENTAL EVALUATION

This section aims to determine the effectiveness of EASDM in detecting autism disorder in toddlers and children, utilizing the dataset that was collected. The Python programming language was used to carry out the actual implementation of the proposed framework. The model's performance test was conducted on the Google Colab cloud computing platform. This platform is equipped with a central processing unit operating at a frequency of 2.6 GHz and a memory capacity of 32 GB.

To construct the classification model, the preprocessed data were split into training and testing sets in an 80:20 ratio. Specifically, 80% of the data were assigned for model training, while the remaining 20% was used to evaluate the model's performance. The present study employs a variety of hyperparametric classifiers, such as random trees, LR, and a DL model. The classifiers were utilized to construct the classification model using the training data, and subsequently, the model's performance was evaluated using the test data. The suggested model's efficiency was evaluated using a variety of criteria, including accuracy, recall, precision, and F1-score.

The following subsections provide a comprehensive exposition of the evaluation metrics employed in the study. The ASD Data Collection section explains the dataset used in this study. The Performance Evaluation section provides a comprehensive analysis of the evaluation metrics utilized. The experimental findings are then presented in the sections LR Performance, XGB Classifier, and DL model and XAI, which demonstrate the results in both simulated and tabular forms.

Performance evaluation

The classification metrics should be used to evaluate the proposed model's performance. While classification accuracy is a commonly used metric, it may not be the most appropriate one when dealing with imbalanced datasets where one class has a much larger representation than the others. Therefore, several other performance metrics have been developed, including precision, recall (also known as sensitivity), F1-score, and the receiver operating characteristic (ROC) curve. The ROC curve is typically plotted on a graph with the sensitivity (y-axis) and specificity (x-axis) of each parameter. The best classifier should have a curve that passes through the graph's top-left corner, showing high sensitivity and specificity. These performance metrics, which are commonly used in the field of machine learning, give an in-depth assessment of the classification model's performance. This confusion matrix shows the number of predictions generated by a classifier that are true positives (TPs), true negatives (TN), false negatives (FN), and false positives (FP). The diagonal cells represent accurately classified observations, while the off-diagonal cells represent misclassified observations. Finally, the calculation of accuracy, precision, and recall metrics involves the utilization of mathematical equations (2, 4, 5) that rely on the values of TP, FP, FN, and TN (Gad and Hosahalli, 2020).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \tag{2}$$

$$\text{F1-score} = 2\,\frac{\text{precision} \star \text{recall}}{\text{precision} + \text{recall}} \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

Machine learning and XAI

This section examines the performance of machine learning models and XAI approaches in detail. Specifically, the development and evaluation of an LR model is discussed in the LR Performance section, while an XGBoost (XGB) model is elaborated on in the XGB Classifier section. Both models are analyzed thoroughly using quantitative metrics and visualizations to provide insights into their predictive capabilities on the given dataset. The explanations produced by XAI techniques for each model are also evaluated to determine their effectiveness in representing the models' reasoning and decision-making processes.

LR performance

Table 3 represents the evaluation of the efficiency of the LR model on both the training and validation sets. The table has two rows for each metric: one for the training set and one for the validation set. The two columns for each set show the accuracy and ROC score of the model. In this case, the model performs perfectly on both the training and validation datasets, as demonstrated by accuracy and ROC scores of 1.

Table 4 represents the classification report of the LR model and provides information about the model's performance. The table's first two rows provide the precision, recall, and F1-score metrics for classes 0 and 1, as well as the support, which indicates the number of occurrences in each class. In this particular scenario, the model demonstrates perfect precision, recall, and F1-score for both classes, demonstrating its accurate classification of all cases.

According to the table, the fourth row represents the overall accuracy of the model, representing the proportion of instances that can be classified correctly. In this case, the model has an accuracy of 1, indicating that it correctly classified all instances. In the table's final rows, the precision, recall, and F1-score metrics are shown in macro and weighted averages. The macro average is a statistical measure that computes the average performance across all classes, whereas the weighted average incorporates the level of support for each individual class. In this particular instance, it can be observed that both the macro and weighted average performance metrics have a value of 1. This means that the model performed perfectly across both classes.

In a binary classification problem, the confusion matrix shows the true and false predictions of the two classes (ASD and nonASD) in the rows and columns of Figure 3. The rows of the matrix display the actual class labels, while the columns display the predicted class labels. The main diagonal elements of the matrix indicate successfully

Table 3: Logistic regression model performance.

	Training set		Validation set	
Metric	Accuracy	ROC	Accuracy	ROC
Value	1.0	1.0	1.0	1.0

Abbreviation: ROC, receiver operating characteristic.

Table 4: Classification report of the logistic regression model.

	Precision	Recall	F1-score	Support
0	1.00	1.00	1.00	127
1	1.00	1.00	1.00	49
Accuracy	1.00			176
Macro average	1.00	1.00	1.00	176
Weighted average	1.00	1.00	1.00	176

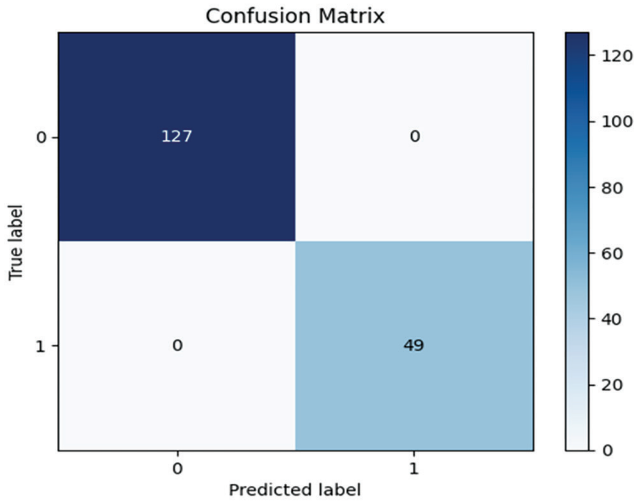

Figure 3: The confusion matrix for the logistic regression model.

categorized cases, while the off-diagonal members reflect erroneously classified examples. In this case, the model correctly predicted all instances of both classes, with 127 TPs and 49 TNs. There were no FPs or FNs in the predictions, indicating that the model performed perfectly on the given dataset.

A summary plot is a graphical representation of the feature importance scores generated by a machine learning model. It is frequently used to discover the most important elements in deciding the outcome variable and to comprehend the relative importance of each item on the model's predictions. The summary plot typically displays the feature importance scores in the descending order, with the most significant features at the top of the plot. Additionally, the plot may highlight features that have a significant impact on the outcome variable for specific individuals, even if they are not the most significant features on the average. The summary plot is a useful tool for interpreting the results of a machine learning model and identifying areas where further investigation may be necessary.

The summary plot of the LR model is presented in Figure 4. The plot indicates that, on average, A5_Score_0 is the most significant feature in determining the outcome variable. However, it is also observed that other features, such as A9_Score_1, may have a greater impact on the outcome variable for specific individuals.

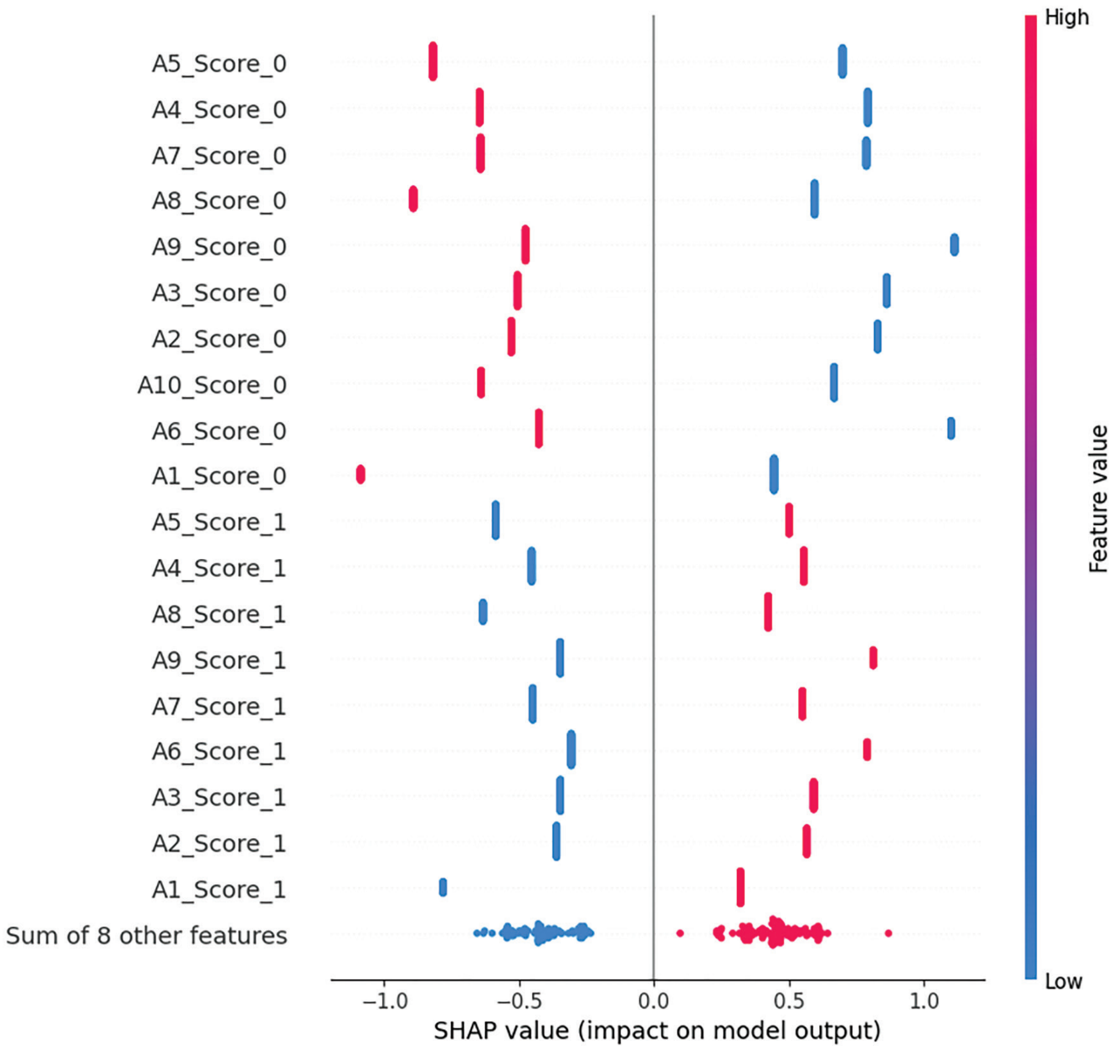

Figure 4: The summary plot for the logistic regression model.

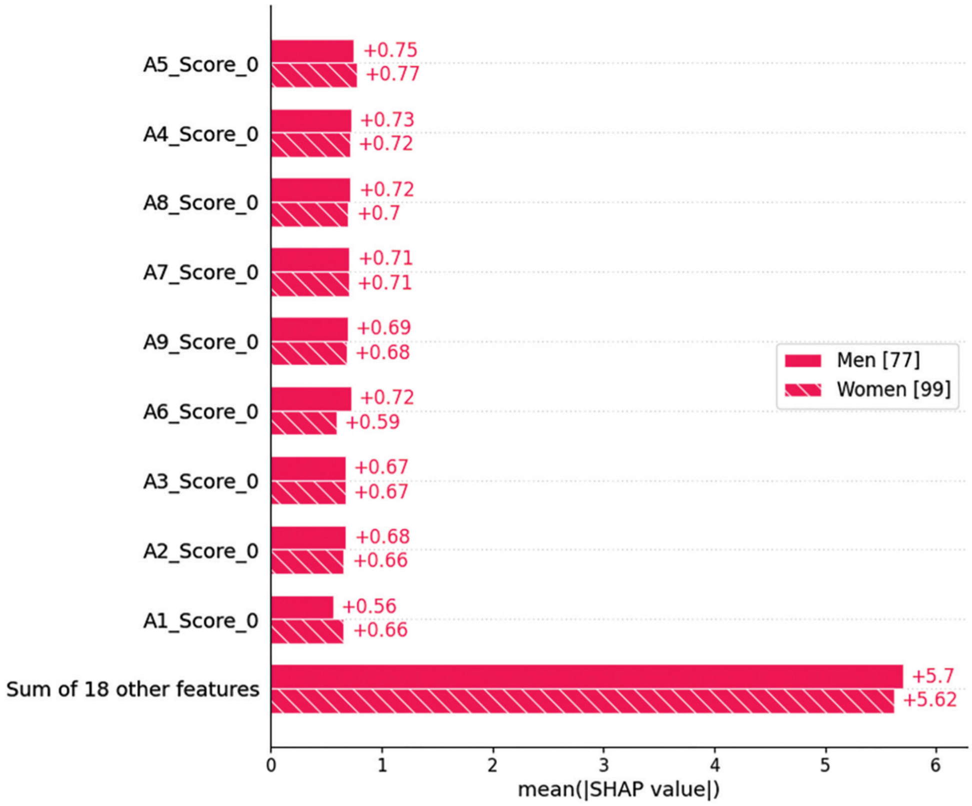

Figure 5: The Cohort bar plot for the logistic regression model.

The bar plot in Figure 5 displays the cohort analysis results for the LR model. The plot is generated by utilizing a dictionary of explanation objects. This dictionary is then used to create a multiple-bar plot, where each bar type represents one of the cohorts represented by the explanation objects. In the following analysis, we utilize this methodology to generate a comprehensive overview of feature importance, distinguishing between male and female individuals. The number within the brackets represents the frequency of occurrences in each cohort. Additionally, the feature importance values are displayed in red on the right side of the feature names for each cohort.

Figure 6 shows the summary plot for the LR model. The bar plot displays the SHAP values linked to a local feature's importance. The default setting for the bar plot limits the display to a maximum of 10 bars. In this plot, each bar represents the SHAP value associated with a specific feature. The feature importance values are displayed in red on the right side of the feature names. Finally, it is clear that the feature A5_Score_0 holds the highest level of significance in determining the outcome variable on average.

XGB classifier

The force plot of the XGB model is shown in Figure 7. It is clear that, in this plot, higher values indicate a greater likelihood of negative outcomes. As a result, the features represented by the red bars in the plot contribute positively to the likelihood of a positive outcome (i.e. have a value of 1), whereas the features represented by the negatively valued bars have a negative impact on the outcome variable, lowering the likelihood of a positive result.

The summary plot of the XGB model is presented in Figure 8. The plot indicates that, on average, A5_Score_0 is the most significant feature in determining the outcome variable. However, it is also observed that other features, such as contry_of_res, may have a greater impact on the outcome variable for specific individuals.

A SHAP-based heat map is a graphical representation of the feature effects generated by a machine learning model, particularly in the context of XGB models for the prediction of ASD. It is commonly utilized to illustrate the relationship between the features and the outcome variable and to identify patterns in the data. In particular, the SHAP-based heat map depicts the shapley additive explanations (SHAP) values of each feature, which denote the contributions of that feature to the prediction of ASD by the XGB model. The SHAP values are derived by taking into account all possible feature subsets and their projected ASD probability, which are then utilized to quantify the feature significance values.

The SHAP-based heat map typically displays the feature effects in a color-coded format, where warmer colors indicate a positive effect and cooler colors indicate a negative effect as shown in Figure 9. Furthermore, the heat map may highlight features that have a significant effect on the outcome variable for specific individuals or under specific conditions, as well as interactions between features that may affect the prediction of ASD.

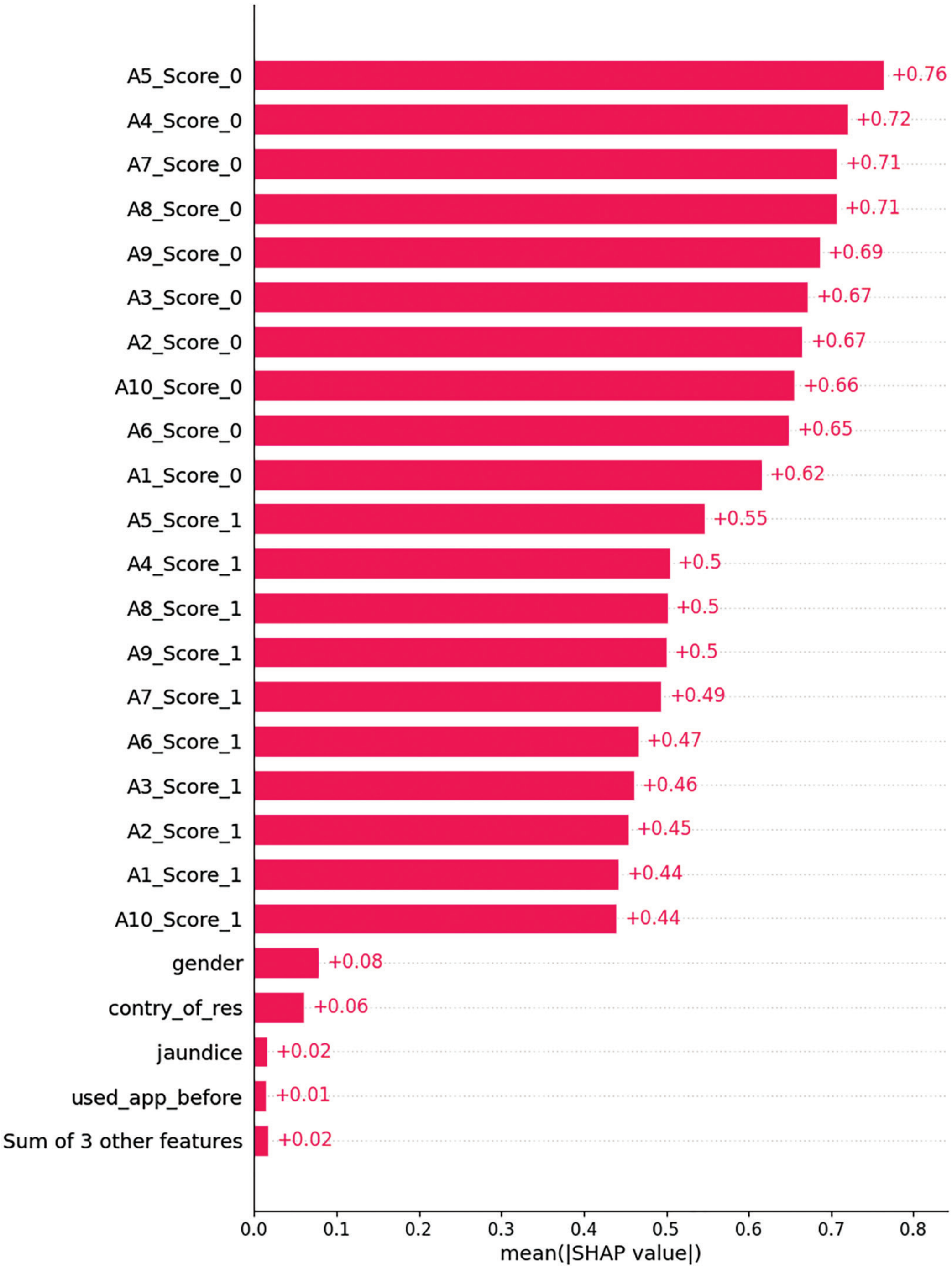

Figure 6: The summary plot for the logistic regression model.

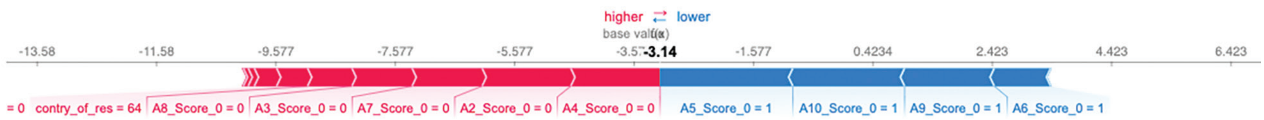

Figure 7: The force plot of the XGB model. XGB, XGBoost.

DL model and XAI

Table 5 displays the accuracy and ROC scores of the DNN model for different sets. The sets include the training set, the test set, and the validation set. Accuracy measures how many predictions were correct, while the ROC score evaluates the ability of the model to distinguish between "Yes" and "No" samples. The model showed perfect performance on the training set, with an accuracy of 1 and an ROC score of 1. On the test set, the model achieved an accuracy score of

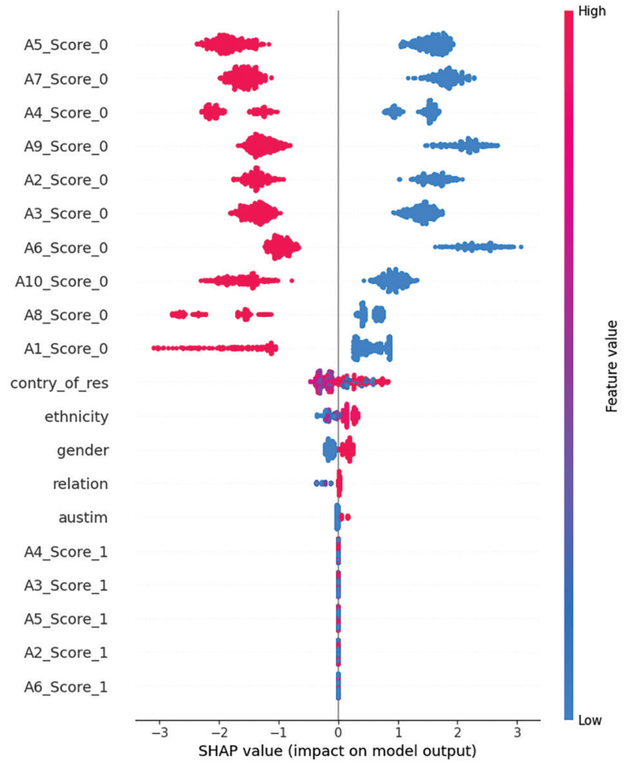

Figure 8: The summary plot the XGB model. XGB, XGBoost.

0.9886, indicating that it performed nearly as well as on the training set. The accuracy score of the model on the validation set was 0.9829, which was slightly lower than the accuracy score on the test set. The validation set's ROC score is 0.976, indicating that the model's ability to differentiate between "Yes" and "No" samples is slightly lower than on the training and testing sets.

The confusion matrix gives a concise representation of the accuracy of predictions for each class (ASD and nonASD) in a binary classification task. Figure 10 assigns a distinct row and column to each of the four possible results: TP, TN, FP, and FN. The figure displays accurately labeled examples through its diagonal elements, while misclassified examples are represented by the nondiagonal elements.

Table 5: Accuracy and ROC scores for different sets.

Set	Accuracy	ROC
Train	1.0000	1.0000
Test	0.9886	0.9786
Validation	0.9829	0.9760

Abbreviation: ROC, receiver operating characteristic.

In the present case, the model demonstrated outstanding efficiency by achieving 126 TPs and 48 FNs. Consequently, it accurately predicted all instances within each respective group. The presence of only one FP and one FN in the predictions suggests a high level of accuracy in the given dataset.

The force plot illustrates the position of the "output value" in relation to the "base value." Additionally, it is possible to observe the features that exert a positive effect on the forecast, denoted by the red color, as well as those that have a negative impact, represented by the blue color, as well as the magnitude of the impact. The force plot of the DL model is presented in Figure 11. The plot indicates that, on average, A8_Score_1 is the most significant feature in determining the outcome variable. However, it is also observed that other features, such as A1_Score_0, may have a greater impact on the outcome variable for specific individuals.

Comparative performance analysis

Table 6 summarizes the predictive performance of several machine learning models for binary ASD prediction. The three columns in the table represent the training set accuracy, training set ROC, and validation set accuracy for each model.

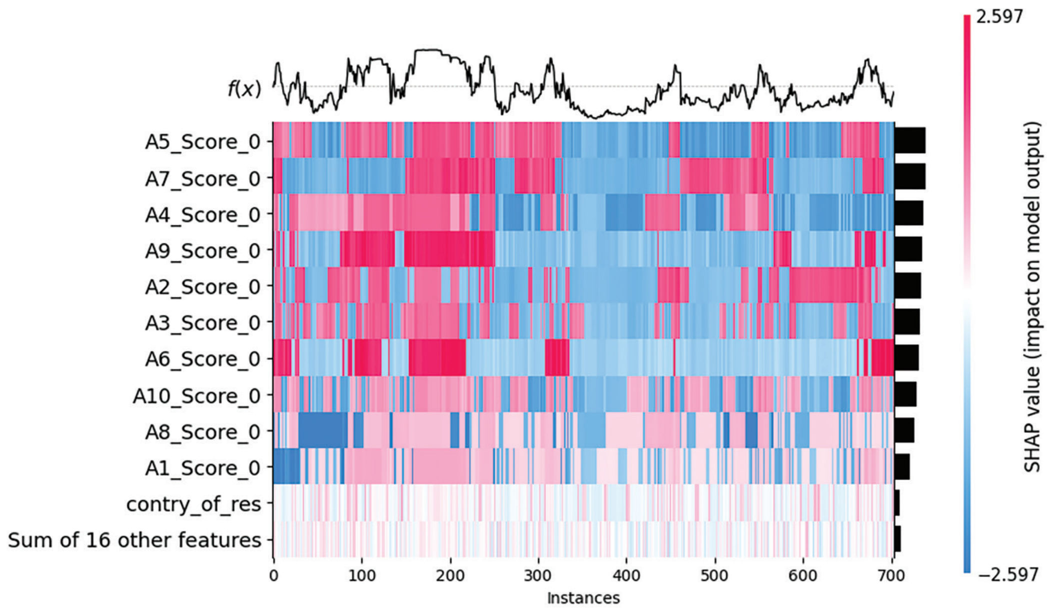

Figure 9: Heat map of the XGB model. XGB, XGBoost.

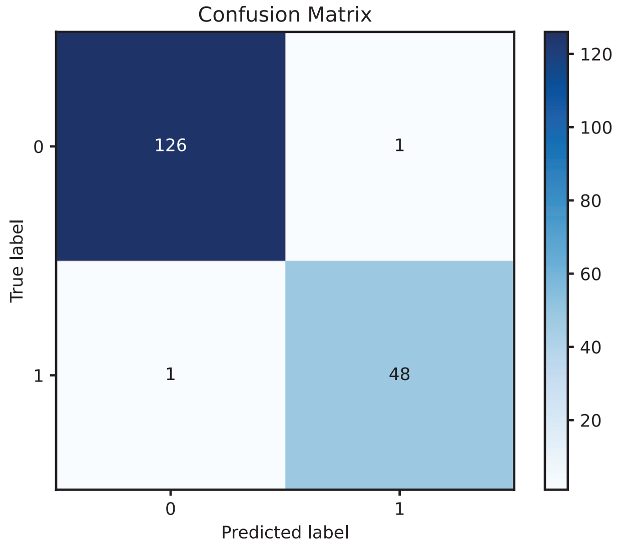

Figure 10: The confusion matrix of the deep learning model.

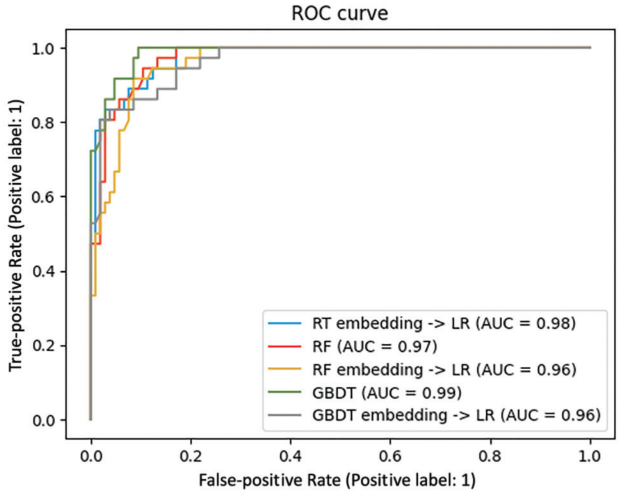

Figure 12: The ROC curves for binary ASD prediction models. Abbreviations: ASD, autism spectrum disorder; LR, logistic regression; ROC, receiver operating characteristic.

The LR and DL models achieved perfect training accuracy and ROC scores of 1.0. However, the DL model exhibited some overfitting with a validation accuracy of 0.983. Among the ensemble models, the random trees embedding model obtained the best training set metrics, with an accuracy of 0.982 and an ROC of 0.989. The gradient boosting classifier and random forest classifier exhibited the lowest training performance but achieved a competitive validation accuracy of 0.922. Overall, the results demonstrate strong predictive capabilities for multiple modeling techniques on this binary classification dataset, with differences emerging in generalization performance.

Figure 12 shows the ROC curves for all the classifiers on the same plot, making it easy to compare their performance. The x-axis displays the FP rate, while the y-axis displays the TP rate. The nearer the ROC curve is to the upper left corner of the figure, the better the model's performance. It is clear that the gradient boosting (GBDT) classifier has the highest ROC value among the models listed. The GBDT classifier has a test set ROC of 0.99 and a training set ROC of 0.92, compared to the other models that have training set ROC values ranging from 0.92 to 0.989. However, it is important to note that the ROC value is only one measure for evaluating the effectiveness of a model and should be examined with other metrics like as accuracy, precision, and recall.

The suggested model's accuracy and F1-score are compared with those of the most advanced models as shown in Table 7. The models evaluated include LR, DL, support vector machine (SVM) (Biswas et al., 2021), and DL (Garg et al., 2022).

LR achieved perfect accuracy and an F1-score of 100%, indicating its ability to accurately classify individuals with and without ASD. DL exhibited an accuracy of 98.86% and an F1-score of 99.2%, demonstrating its strong performance in ASD prediction. SVM (Biswas et al., 2021) showed comparable accuracy and an F1-score of 98.27%,

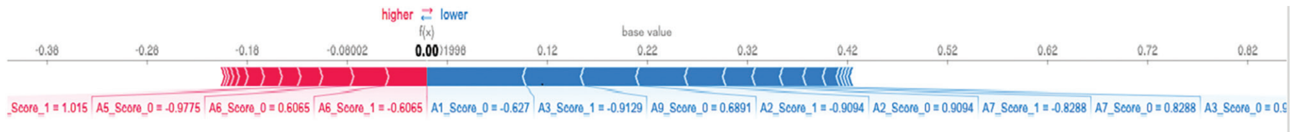

Figure 11: The force plot for deep learning model.

Table 6: Model performance for binary ASD prediction.

Model	Train set accuracy	Train set ROC	Test set accuracy
LR	1.0	1.0	1.0
Deep learning	1.0	1.0	0.988
Random trees embedding	0.982	0.989	0.922
RF embedding	0.973	0.960	0.894
GBDT embedding	0.965	0.944	0.922
Gradient boosting classifier	0.938	0.921	0.922
Random forest classifier	0.929	0.914	0.922

Abbreviations: ASD, autism spectrum disorder; LR, logistic regression; ROC, receiver operating characteristic.

Table 7: Comparison of the proposed model with the state-of-art models.

Model	Accuracy	F1
LR	100	100
Deep learning	98.86	99.2
SVM (Biswas et al., 2021)	98.27	98.27
Deep learning (Garg et al., 2022)	98	98
Federated learning (Farooq et al., 2023)	98	98

Abbreviation: LR, logistic regression.

further supporting its effectiveness in this domain. Another DL model (Garg et al., 2022) achieved an accuracy of 98% and an F1-score of 98%.

These results indicate that the proposed model performs competitively compared to the state-of-the-art models in terms of accuracy and F1-score. Finally, the results highlight the efficacy and potential of the proposed model in accurately predicting ASD in binary classification tasks.

CONCLUSION

For the early prediction of ASD, the suggested methodology combines two newly applied methodologies, namely DL and XAI. This paper's main goal is to recommend the features that will most likely increase the accuracy of ASD prediction in children. The findings show that machine learning models may accurately identify whether or not a person has ASD when they are correctly optimized. The LR and DL models achieved perfect training accuracy and ROC scores of 1.0. However, the DL model exhibited some overfitting with a validation accuracy of 0.9829. The XAI indicates that A8_Score_1 is the most significant feature in determining the outcome variable. However, it is also observed that other features, such as A1_Score_0, may have a greater impact on the outcome variable for specific individuals. It has been found that ML models' explainability can be increased without compromising the model's effectiveness. Finally, future research could focus on applying the same models to varied and large datasets in order to increase the scope of the study. Several steps can be taken by researchers to improve the performance of the models, such as increasing the size and diversity of the training data by collecting more samples, employing DA methods to generate newer training instances, or employing TL techniques to use pretrained models from comparable domains or tasks. Furthermore, other ensemble models can be employed to combine the predictions of multiple models with different architectures, hyperparameters, or initializations, using techniques such as averaging or stacking to reduce model biases and errors and enhance the model's generalization capabilities.

AUTHOR CONTRIBUTIONS

El-Sayed Atlam, Mehedi Masud, and Abdulqader M. Almars conceptualized the study; El-Sayed Atlam, Ibrahim Gad, Mahmoud Rokaya, and Hossam Meshref did the methodology; Ibrahim Gad and Abdulqader M. Almars worked on the software; Abdulqader M. Almars and El-Sayed Atlam performed formal analysis; Mahmoud Rokaya and Mehedi Masud performed the investigation; Abdulqader M. Almars and Ibrahim Gad were responsible for data curation; El-Sayed Atlam and Abdulqader M. Almars drafted the original manuscript; El-Sayed Atlam and Hossam Meshref reviewed and edited the manuscript; Ibrahim Gad performed visualization; and El-Sayed Atlam performed supervision.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest in association with the present study.

ACKNOWLEDGMENTS

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group no KSRG-2023-371, Funder id: http://dx.doi.org/10.13039/501100019345.

REFERENCES

Abraham A., Milham M.P., Di Martino A., Craddock R.C., Samaras D., Thirion B., et al. (2017). Deriving reproducible biomarkers from multi-site resting-state data: an autism-based example. *NeuroImage*, 147, 736-745.

Adilakshmi J., Vinoda Reddy G., Nidumolu K.D., Cosme Pecho R.D. and Pasha M.J. (2023). A medical diagnosis system based on explainable artificial intelligence: autism spectrum disorder diagnosis. *Int. J. Intell. Syst. Appl. Eng.*, 11(65), 385-402.

Ahmad S., Khan S., AlAjmi M.F., Dutta A.K., Dang L.M., Joshi G.P., et al. (2022). Deep learning enabled disease diagnosis for secure internet of medical things. *Comput. Mater. Contin.*, 73(1), 965-979.

Akter T., Satu M.S., Barua L., Sathi F.F. and Ali M.H. (2017). Statistical analysis of the activation area of fusiform gyrus of human brain to explore autism. *Int. J. Comput. Sci. Inf. Secur.*, 15, 331-337.

Akter A., Khan M.I., Ali M.H., Satu M.S. and Uddin M.J. (2021). Improved machine learning based classification model for early autism detection. In: *2021 2nd Int. Conf. on Robotics, Electrical and Signal Processing Techniques (ICREST)*, Dhaka, Bangladesh, 5-7 January 2021; pp. 742-747.

Alarifi H.S. and Young G.S. (2018). Using multiple machine learning algorithms to predict autism in children. In: *Proc. of Int. Conf. Artificial Intelligence*, New York, NY, United States, July 30 – August 2, 2018; pp. 464-467.

Aldhyani T.H., Verma A., Al-Adhaileh M.H. and Koundal D. (2022). Multi-class skin lesion classification using a lightweight dynamic kernel deep-learning-based convolutional neural network. *Diagnostics*, 12(9), 2048.

Alkahtani H., Aldhyani T.H. and Alzahrani M.Y. (2023). Deep learning algorithms to identify autism spectrum disorder in children-based facial landmarks. *Appl. Sci.*, 13(8), 4855.

Almars A.M., Alwateer M., Qaraad M., Amjad S., Fathi H., Kelany A.K., et al. (2021). Brain cancer prediction based on novel interpretable ensemble gene selection algorithm and classifier. *Diagnostics*, 11(10), 1936.

Almars A.M., Badawy M. and Elhosseini M.A. (2023). ASD^2-TL GTO: Autism spectrum disorders detection via transfer learning with gorilla troops optimizer framework. *Heliyon*, 9(11), e21530.

Alwateer M., Almars A.M., Areed K.N., Elhosseini M.A., Haikal A.Y. and Badawy M. (2021). Ambient healthcare approach with hybrid whale optimization algorithm and naive bayes classifier. *Sensors*, 21(13), 4579.

Anirudh R. and Thiagarajan J.J. (2021). Machine learning methods for autism spectrum disorder classification. In: *Neural Engineering Techniques for Autism Spectrum Disorder*, Elsevier, pp. 151-163. 10.1016/b978-0-12-822822-7.00008-9.

Atlam E.S., Fuketa M., Morita K. and Aoe J.i. (2003). Document similarity measurement using field association term. *Inform. Process. Manage. J.*, 39(6), 809-824.

Atlam E., Ewis A., Abd El-Raouf M., Ghoneim O. and Gad I. (2022). A new approach in identifying the psychological impact of COVID-19 on university student's academic performance. *Alex. Eng. J.*, 61(7), 5223-5233.

Badawy M., Almars A.M., Balaha H.M., Shehata M., Qaraad M. and Elhosseini M. (2023). A two stage renal disease classification based on transfer learning with hyperparameters optimization. *Front. Med.*, 10, 1106717.

Biswas M., Kaiser M.S., Mahmud M., Mamun S.A., Hossain M.S. and Rahman M.A. (2021). An XAI based autism detection: the context behind the detection. In: *Brain Informatics*, Springer International Publishing, pp. 448-459. 10.1007/978-3-030-86993-9_40.

Carette R., Cilia F., Dequen G., Bosche J., Guerin J.-L. and Vandromme L. (2018). Automatic autism spectrum disorder detection thanks to eye-tracking and neural network-based approach. In: *Internet of Things (IoT) Technologies for HealthCare*, Springer International Publishing, pp. 75-81. 10.1007/978-3-319-76213-5_11.

Claudio P., Federica G., Riccardo S., Flavio V.A. and Diciotti S. (2023). Early prediction of autism spectrum disorders through interaction analysis in home videos and explainable artificial intelligence. *Comput. Hum. Behav.*, 148(3), 86-79.

Duda M., Ma R., Haber N. and Wall D.P. (2016). Use of machine learning for behavioral distinction of autism and ADHD. *Transl. Psychiatry*, 6(2), e732-e732. 10.1038/tp.2015.221.

Erkan U. and Thanh D.N.H. (2019). Autism spectrum disorder detection with machine learning methods. *Curr. Psychiatry Res. Rev. Former. Curr. Psychiatry Rev.*, 15, 297-308.

Farooq M.S., Tehseen R., Sabir M. and Atal Z. (2023). Detection of autism spectrum disorder (ASD) in children and adults using machine learning. *Sci. Rep.*, 13(1), 9605. 10.1038/s41598-023-35910-1.

Farsi M., Hosahalli D., Manjunatha B., Gad I., Atlam E.-S., Ahmed A., et al. (2021). Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data. *Alex. Eng. J.*, 60(1), 1299-1316.

Gad I. and Hosahalli D. (2020). A comparative study of prediction and classification models on NCDC weather data. *Int. J. Comput. Appl.*, 44, 414-425.

Garg A., Parashar A., Barman D., Jain S., Singhal D., Masud M., et al. (2022). Autism spectrum disorder prediction by an explainable deep learning approach. *Comput. Mater. Contin.*, 71(1), 1459-1471. 10.32604/cmc.2022.022170.

Georgoula C., Ferrin M., Pietraszczyk-Kedziora B., Hervas A., Marret S., Oliveira G., et al. (2023). A phase III study of bumetanide oral liquid formulation for the treatment of children and adolescents aged between 7 and 17 years with autism spectrum disorder (SIGN 1 trial): participant baseline characteristics. *Child Psychiatry Hum. Dev.*, 54, 1360-1372.

Guillon Q., Hadjikhani N., Baduel S. and Rogé B. (2014). Visual social attention in autism spectrum disorder: insights from eye tracking studies. *Neurosci. Biobehav. Rev.*, 42, 279-297.

Haq A.U., Li J.P., Khan J., Memon M.H., Nazir S., Ahmad S., et al. (2020). Intelligent machine learning approach for effective recognition of diabetes in E-healthcare using clinical data. *Sensors*, 20(9), 2649. 10.3390/s20092649.

Haq A.U., Li J.P., Khan I., Agbley B.L.Y., Ahmad S., Uddin M.I., et al. (2022). DEBCM: deep learning-based enhanced breast invasive ductal carcinoma classification model in IoMT healthcare systems. *IEEE J. Biomed. Health Inform.*, 1-12. 10.1109/jbhi.2022.3228577.

Heinsfeld A.S., Franco A.R., Craddock R.C., Buchweitz A. and Meneguzzi F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *Neuroimage Clin.*, 17, 16-23.

Hossain M.A., Saiful Islam S.M., Quinn J.M.W., Huq F. and Moni M.A. (2019). Machine learning and bioinformatics models to identify gene expression patterns of ovarian cancer associated with disease progression and mortality. *J. Biomed. Inform.*, 100, 103313.

Howlader K., Satu M., Barua A. and Moni M. (2018). Mining significant features of diabetes mellitus applying decision trees: a case study in Bangladesh. *BioRxiv*, 481-499. 10.1109/JBHI.2020.3001216.

Jiang X. and Chen Y.-F. (2008). Facial image processing. In: *Applied Pattern Recognition*, (Bunke H., Kandel A. and Last M. eds.) Studies in Computational Intelligence, Vol 91, Springer, Berlin, Heidelberg, pp. 29-48.

Kaiser M.S., Mamun S.A., Mahmud M. and Tania M.H. (2020). Healthcare robots to combat COVID-19. In: *COVID-19: Prediction, Decision-Making, and its Impacts*, Springer, Singapore, pp. 83-97. 10.1007/978-981-15-9682-7_10.

Kanhirakadavath M.R. and Chandran M.S.M. (2022). Investigation of eye-tracking scan path as a biomarker for autism screening using machine learning algorithms. *Diagnostics*, 12(2), 518.

Koyamada S., Shikauchi Y., Nakae K., Koyama M. and Ishii S. (2015). Deep learning of fMRI big data: a novel approach to subject-transfer decoding. *arXiv preprint arXiv:1502.00093*.

Maenner M.J., Shaw K.A., Bakian A.V., Bilder D.A., Durkin M.S., Esler A., et al. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveill Summ.*, 70(11), 1-14.

Mahmud M., Kaiser M.S., Rahman M.M., Rahman M.A., Shabut A., Al-Mamun S., et al. (2018). A brain-inspired trust management model to assure security in a cloud based IoT framework for neuroscience applications. *Cogn. Comput.*, 10(5), 864-873. 10.1007/s12559-018-9543-3.

Malki Z., Atlam E.-S., Ewis A., Dagnew G., Reda A., Elmarhomy G., et al. (2020a). ARIMA models for predicting the end of COVID-19 pandemic and the risk of a second rebound. *J. Neural Comput. Appl.*, 33(7), 2929-2948.

Malki Z., Atlam E.-S., Hassanien A.E., Dagnew G., Elhosseini M.A. and Gad I. (2020b). Association between weather data and COVID-19 pandemic predicting mortality rate: machine learning approaches. *Chaos Solitons Fractals*, 138, 110137.

Malki Z., Atlam E.-S., Ewis A., Dagnew G., Ghoneim O.A., Mohamed A.A., et al. (2021). The covid-19 pandemic: prediction study based on machine learning model. *J. Environ. Sci. Pollut. Res.*, 28(30), 40496-40506.

Mujeeb Rahman K. and Subashini M.M. (2022). Identification of autism in children using static facial features and deep neural networks. *Brain Sci.*, 12(1), 94.

Noor T., Almars A., Alwateer M., Almaliki M., Gad I. and Atlam E.S. (2022). Sarima: a seasonal autoregressive integrated moving average model for crime analysis in Saudi Arabia. *Electronics*, 11(23), 3986-3998.

Payrovnaziri S.N., Chen Z., Rengifo-Moreno P., Miller T., Bian J., Chen J.H., et al. (2020). Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. *J. Am. Med. Inform. Assoc.*, 27(7), 1173-1185. 10.1093/jamia/ocaa053.

Plis S.M., Hjelm D.R., Salakhutdinov R., Allen E.A., Bockholt H.J., Long J.D., et al. (2014). Deep learning for neuroimaging: a validation study. *Front. Neurosci.*, 8, 229.

Raj S. and Masood S. (2020). Analysis and detection of autism spectrum disorder using machine learning techniques. *Proc. Comput. Sci.*, 167, 994-1004. 10.1016/j.procs.2020.03.399.

Ribeiro S.S.M.T. and Guestrin C. (2016). Why should I trust you? Explaining the predictions of any classifier. In: *KDD'16: Proc. of the 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, San Francisco, California, USA, 13-17 August 2016; pp. 1135-1144.

Satu M.S., Sathi F.F., Arifen M.S., Ali M.H. and Moni M.A. (2019). Early detection of autism by extracting features: a case study in Bangladesh. In: *Proc. ICREST*, pp. 400-405. https://arxiv.org/abs/2004.13932.

Schelinski S., Borowiak K. and von Kriegstein K. (2016). Temporal voice areas exist in autism spectrum disorder but are dysfunctional for voice identity recognition. *Soc. Cogn. Affect. Neurosci.*, 11(11), 1812-1822.

Thabtah F. (2017). Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment. In: *ICMHI'17: Proceedings of the 1st International Conference on Medical and Health Informatics*, Taichung City, Taiwan, 20-22 May 2017; pp. 1-6.

Thabtah F. (2018). Machine learning in autistic spectrum disorder behavioral research: a review and ways forward. *Inform. Health Soc. Care.*, 44(3), 278-297. 10.1080/17538157.2017.1399132.

Thabtah F. and Peebles D. (2020). A new machine learning model based on induction of rules for autism detection. *Health Inform. J.*, 26(1), 264-286.

Thabtah F., Kamalov F. and Rajab K. (2018). A new computational intelligence approach to detect autistic features for autism screening. *Int. J. Med. Inform.*, 117, 112-124. 10.1016/j.ijmedinf.2018.06.009.

Tuchman R., Moshé S.L. and Rapin I. (2009). Convulsing toward the pathophysiology of autism. *Brain Dev.*, 31(2), 95-103. 10.1016/j.braindev.2008.09.009.

Yolcu G., Oztel I., Kazan S., Oz C., Palaniappan K., Lever T.E., et al. (2019). Facial expression recognition for monitoring neurological disorders based on convolutional neural network. *Multimed. Tools Appl.*, 78, 31581-31603.