Tech Science Press

Autism Spectrum Disorder Diagnosis Using Ensemble ML and Max Voting Techniques

A. Arunkumar[1,*] and D. Surendran[2]

[1]Computer science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, 641008, India
[2]Computer Science and Engineering, KPR Institute of Engineering and Technology, Coimbatore, 641407, India
*Corresponding Author: A. Arunkumar. Email: arunskcet2021@gmail.com

Abstract: Difficulty in communicating and interacting with other people are mainly due to the neurological disorder called autism spectrum disorder (ASD) diseases. These diseases can affect the nerves at any stage of the human being in childhood, adolescence, and adulthood. ASD is known as a behavioral disease due to the appearances of symptoms over the first two years that continue until adulthood. Most of the studies prove that the early detection of ASD helps improve the behavioral characteristics of patients with ASD. The detection of ASD is a very challenging task among various researchers. Machine learning (ML) algorithms still act very intelligent by learning the complex data and predicting quality results. In this paper, ensemble ML techniques for the early detection of ASD are proposed. In this detection, the dataset is first processed using three ML algorithms such as sequential minimal optimization with support vector machine, Kohonen self-organizing neural network, and random forest algorithm. The prediction results of these ML algorithms (ensemble) further use the bagging concept called max voting to predict the final result. The accuracy, sensitivity, and specificity of the proposed system are calculated using confusion matrix. The proposed ensemble technique performs better than state-of-the art ML algorithms.

Keywords: SVM; autism disorder; Kohonen SONN; max voting; ensemble machine learning technique; random forest; SMO–SVM; bootstrap gradient boosting

1 Introduction

According to the author, autism spectrum disorder (ASD) is the condition when human beings have difficulties in interaction and communication. This miscommunication is due to negative influences in the nervous system of humans. The nervous system tends to affects the eyes, emotional hormones, and health of patients with autism. Symptoms and severity of ASD vary from one person to another. Most commonly identified symptoms are social communication, interactions, and obsessive, cyclic behaviors. In April 2020, 1 out of 54 children was recognized to be affected by ASD. This disorder begins at childhood and continues over adolescence to adulthood. Sometimes patients with ASD live quite independently, but sometimes few need lifelong special care. The symptoms experienced by a person

with ASD can be further reduced by psychosocial evidence-based treatment and parent skill training programs. Beyond all treatment, first and foremost is early detection. This paper focuses on various machine learning (ML) algorithms to detect the early symptoms of ASD.

ASD is usually found in childhood at around 2 to 3 years [1]. Genetic and environmental factors are assumed to combine their influence in autism, which is not an illness but rather a neurological repair that does not support the child to concentrate on thinking, learning, work, and problem solving ability. Patients with autism cannot exhibit gestures, expression, and feelings in communication. Recent survey reports from the World Health Organization imply that autism is the fastest growing disorder worldwide. According to the Autism Center of Excellence, 1 in 68 children had ASD in 2019 [2]. The early diagnosis of this disorder is very important. Various autism categories are given by Fergus in [3] as ailment, namely, autistic, Asperger, and pervasive ailments. The lower disorder is autistic, the medium disorder is Asperger ailment, and the higher disorder is pervasive development ailment. The major challenges of a child with ASD are as follows:

a) No concentration on surrounding events
b) Same words, names, and situations that may be repeated again and again
c) Interaction among the peoples that is not normal
d) No gestures and facial expressions during communication
e) High sensitivity in their feeling when we touch and speak
f) Voice that is rude and sounds high.
g) Showing abnormal body postures to others

ASD may be caused by the gene of parents, another family person who is autistic, complicated deliveries, and missing vaccination in children. This paper analyzes the ensemble ML techniques for predicting ASD with a high accuracy at an early stage. The literature review presents various researchers' detection techniques and the ML algorithms used in this area.

Abnormal human brain development leads to problems such as ASD. A person with ASD has extreme difficulty facing people and socially interacting with them. The person's entire life is affected by ASD. To date, various researchers have identified genetic and environmental factors that cause ASD. If this syndrome is detected in early life, its effect can be reduced, but it cannot be fully cured. The major risk factors analyzed are low weight during birth, first baby has ASD and second baby also has possibilities, old age of parents, and late marriages. Patients with ASD have difficulties such as the following:

a) Giggling laugh and louder cry
b) Not sensing pain in the body
c) Totally missing eye contact
d) Not showing wishes on anything
e) Always liking to stay alone
f) May attach to several inappropriate objects

Patients with ASD will not be interested in any constraints. They repeatedly show consistent conduct and behave as follows:

a) They tend to repeat the same words most of the time.
b) They will feel if their schedule tends to change such as shifting home and leaving friends.
c) They may remember small facts and numbers in their mind.
d) They are less sensitive to noise, lights, and pains.

These symptoms cannot be cured but can be reduced by early-stage detection. The early detection and treatment of ASD will improve the quality of life of patients, but no medical test for detecting autism earlier is available. Judgment can only be made based on behavioral symptoms. ASD in adolescents will be recognized by their teachers and parents in school. The treatment for health in school will examine the symptoms and give suggestions. Then, the student will be sent to a doctor for study and examination of ASD. This process is very difficult because ASD is similar to mental health issues. This problem motivates us to initiate artificial intelligence (AI) technologies for detecting ASD earlier.

ML in ASD

Recent advances in AI make researchers use ML to a greater extent in disease prediction. ASD is a highly increasing disorder that needs more scope through technologies for early prediction. In this technology, the pattern is trained with procedures such that an ML algorithm performs based on an observed pattern and returns the results [4]. Learning algorithms have three types: 1. supervised, 2. unsupervised, and 3. semi supervised [5]. Supervised learning is based on classification and regression techniques with induction of input dataset. Unsupervised learning is based on data description. Semi supervised learning is supervised and unsupervised, that is, labeled and unlabeled data. To improve the performance of autism detection, ML algorithms are studied. Various ML algorithms are used to test the accuracy and specificity of the proposed ensemble learning algorithms. Prediction accuracy increases by decreasing ASD prediction time and less feature selection (FS) models.

Section 2 reviews the related literature. Section 3 explains different ML algorithms combined to perform an accurate early prediction of diseases. Section 4 evaluates the performance of the proposed result with existing techniques. Section 5 discusses the conclusion and future work.

2 Related Works

This part explains previous research work on ASD prediction. Among all ML algorithms, the focus is on an algorithm used to predict diseases with a high accuracy.

The support vector machine (SVM) classifier is used in article [6] for autism disorder prediction. The SVM classifier is used to categorize the feature with a high accuracy. It one of the best ML techniques used for classification. The author selects a dataset with 1200 children with autism and 500 children with no autism disorder. Various cross validation stages also used in this work for screening the best. It achieves a prediction accuracy of 86%. The prediction of diabetes and lung cancer uses machine and AI technique in articles [7,8]. In this model, they propose random forest (RF) classifier and regression tree algorithm, and dichotomize in RF to obtain a high accuracy for childhood (4–12 years), adolescence (age 12–18 years), and adulthood (19 years and above) data. The tree-based ML algorithm is proposed in [9] for predicting autism in various age groups of people. Datasets are categorized into different features based on region and age. J48 classifier and decision tree are used in this work. The ML classifiers proposed in Altay et al. [10] are used to distinguish autism from unaffected persons.

Linear discriminant analyses and k-nearest neighbor (KNN) algorithm perform well in classifying the data with 90% accuracy. Instrument for ASD detection is used by Kosmicki et al. [11]. The instrument is designed based on ML algorithms for autism diagnosis. The observation schedule method is used to check various behavior subsets that vary between children with or without ASD. Here, the advantages of eight ML techniques are taken for design with two modules, and 4500 persons are tested. Accuracy is achieved but takes more time to predict. The facial expressions of patients with autism are predicted. The individual face is captured as a dataset, and face recognition with ML-based image classification is developed for prediction.

The mobile application-based image processing technique proposed in Bone et al. [12] is used to predict the autism disorder. A child normally uses various factors in behavioral change. The following tools are used for screening: eye contact, voice recognition, and positive response. The eyes of the child are screened, and their focus on the screen is monitored. A normal person will focus, whereas an abnormal will not focus. The voice of an affected person frequency is lower than that of a normal person. These prediction tools are used in mobile application for easy processing. The interactive feature using mobile advance technology for autism affected person is suggested, and 3D-based animation is postulated for the children affected by ASD and the family for interaction. The author names the animated character Woody. Every animated scene is printed with expression as Woody is happy or unhappy. It helps the child with autism to understand the action clearly without difficulties. The ML classifier SVM [13,14] is further proposed for autism prediction. The interview dataset is used as the trained model. Here, linear kernel SVM and greedy backward elimination method are applied for the trained dataset. Radial function is also used with SVM for predicting with 80% accuracy. In the article [15,16], a different ML is deeply studied, and SVM is the best ML classifier among RF, decision tree, naïve Bayes, and KNN. Rule-based ML [17–20] is used for autism disorder prediction and improves the classification model and prediction rate. The performance of this work encourages our model to use ML techniques. Decision tree classifiers [21,22] with ML are used for autism disorder prediction in Bangladesh. In this paper, sequential minimal optimization with support vector machine (SMO–SVM), neural network, and RF are selected for ensemble learning technique. This combination of ML works with a high prediction rate.

3 Proposed Ensemble Learning with Bagging Prediction Method

3.1 Problem Statement

ASD is a neurological disorder disease or behavioral disease that entails lifelong interaction and communication problems for a person. The ASD problem of a person may start from toddler age and childhood, and continue to adolescence until adulthood. This disability is not curable but can be diagnosed at an early stage. This early detection can help the person improve the treatment. Various ML algorithms have been used to diagnose ASD. It can be identified at the early stage detection of 2 years of age based on symptoms. In the research direction of medical diagnosis of this disease, finding the best methods for early prediction of ASD is still in research.

3.2 Ensemble Learning with Bagging Prediction Model

Ensemble learning is the improvement of ML results combined with several ML algorithmic models results to provide a better prediction than single ML algorithms. The ensemble models can be divided into bagging and boosting.

In this proposed work, the early diagnosis of ASD is predicted using the ASD dataset such as Toddlers, Children, Adolescents, and Adults. The input dataset is divided into training and testing data. The proposed prediction is divided into four phases: preprocessing, FS, classification using ML algorithms, and ensemble learning for final prediction. The input raw dataset is preprocessed to remove the missing values, and the features are selected using bootstrapped gradient descent FS algorithm. The dataset with selected features are then classified using ensemble learning with three ML algorithms such as SMO–SVM, Kohonen self-organizing neural network (SONN), and RF. The prediction results of these ML algorithms are then subjected to ensemble learning using the bagging concept called max voting to predict the final result. The proposed methodology is depicted in Fig. 1, in which different structures of ML algorithms are constructed from the input data. The derived classifiers are integrated to create a global classifier that consists of an ensemble of outputs.
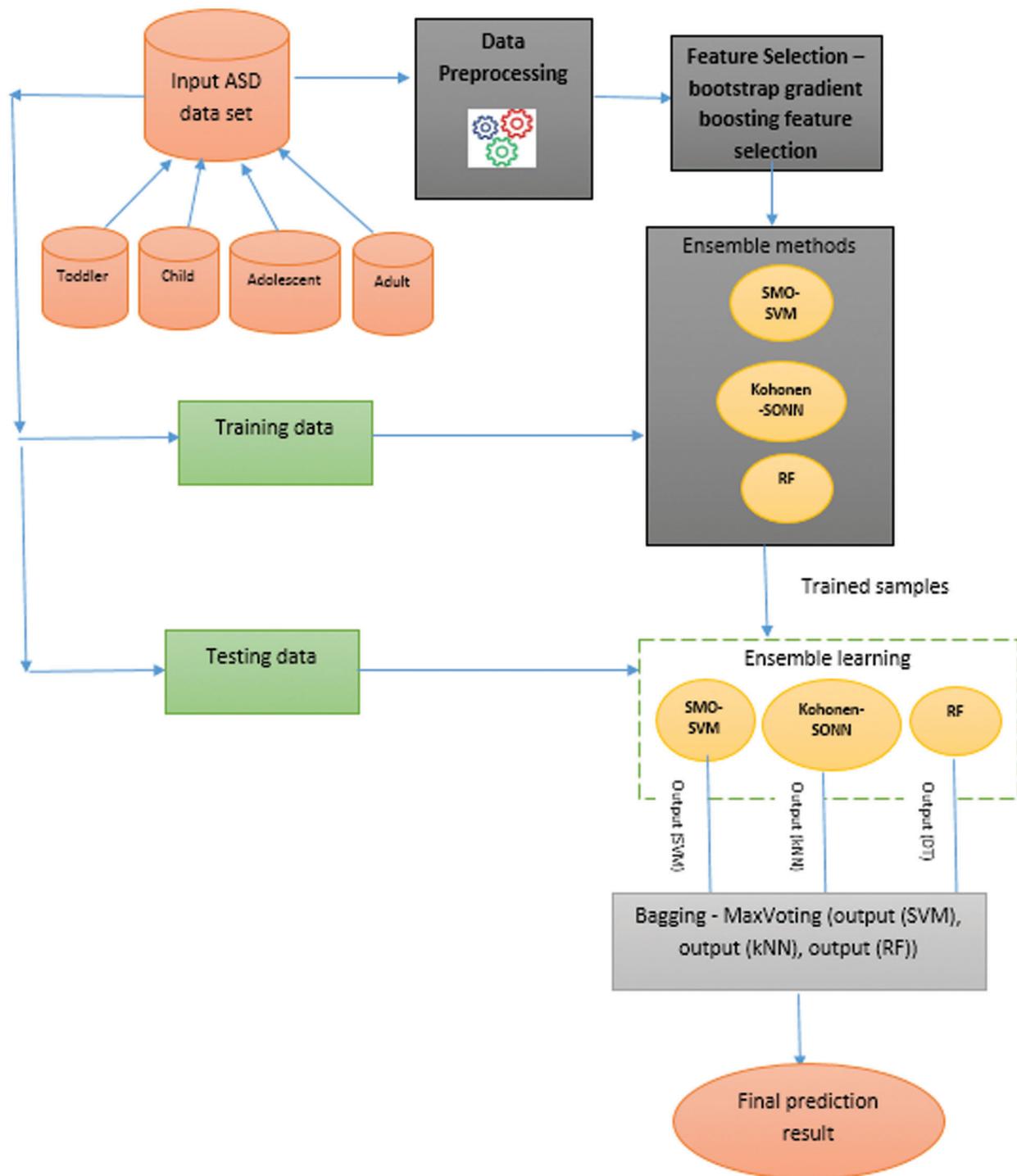
Figure 1: Architectural overview of proposed ASD prediction system

3.3 Datasets

The datasets are collected from the UCI repository [23]. The AQ-10 dataset contains four datasets: Toddlers, Children, Adolescents, and Adults. The details of these datasets are presented in Tab. 1. These

datasets have 20 common attributes that are used for prediction and one class attribute. The attributes are described in Tab. 2.

Table 1: ASD dataset list

S. No.	Dataset name	Attribute type	No. of attributes	No. of instances
1.	Toddlers	Categorical, continuous, binary	21	1054
2	Children	Categorical, continuous, binary	21	292
3	Adolescents	Categorical, continuous, binary	21	104
4	Adults	Categorical, continuous, binary	21	704

Table 2: Attribute description

Attribute ID	Attribute name	Type	Description
1	Age	Number	Toddlers (month), children, adolescents, and adults (year)
2	Gender	String	Male or female
3	Ethnicity	String	List of common ethnicities in text format
4	Born with jaundice	Boolean	Whether the case was born with jaundice
5	Family member with PDD	Boolean	Whether any immediate family member has PDD
6	Who is completing the test	String	Parent, self, caregiver, medical staff, and clinician
7	Country of residence	String	List of countries in text format
8	Used the screening application before	Boolean	Whether the user has used a screening application
9	Screening Method Type	Integer	Type of screening methods selected based on age category
10	Response of Q1	Binary	Does your child look at you when you call his/ her name?
11	Response of Q2	Binary	How easy is it for you to get eye contact with your child?
12	Response of Q3	Binary	Does your child point to indicate that s/he wants something?
13	Response of Q4	Binary	Does your child point to share interest with you?

3.4 Preprocessing

Data preprocessing is important to transform the raw data into a meaningful, understandable format. The collected dataset from is preprocessed to remove the features with missing values that are not needed for further processing. The removed records are not needed for further processes such as FS and

classification. This removal of unwanted features and missing values records improves classification accuracy [24]. Good preprocessed data lead to good result. Various algorithms such as missing values, data discretization, and dimensionality reduction are used to handle inconsistent data. The missing values are handled using a missing value imputation technique [25] such as continuous, categorical attributes are input using mode, and binary attributes' missing values are handled with a random number.

3.5 FS Using Bootstrap Gradient Boosting Approach

Various FS algorithms such as correlation, gain ratio, and information gain are available for selecting the relevant features from the whole feature set. Our proposed work focuses on the use of ensemble learning in ASD prediction and uses ensemble bagging concept called bootstrap gradient boosting for selecting the relevant features for further classification. The main motivation of this technique is that gradient boosting performs well because they belong to ensemble ML class. In the gradient technique, ensembles are formed from the decision tree by adding one tree at time for ensemble purpose. It generally works like Ada boosting, but bootstrap aggregation is used to improve performance. Here, samples are randomly selected, and ensemble members are fitted. Bootstrap samples are independent of their process. These samples are eventually distributed with a low correlation between the input data samples. Gradient boosting uses a gradient descent algorithm to minimize the loss while adding input data models into the ensemble.

Algorithm: 1 (FS-Bootstrapped Gradient Boosting)

Input: Preprocessed dataset

Output: Selected features

Step 1: The subsets of features are selected randomly from the preprocessed data.

Repetitions of features are eliminated, and identical subsets of features are generated.

Step 2: For i=1 to X (all training data samples)

Step 3: Gradient boosting improves each iteration t by creating a new subset that adds an estimator

called h to create a better subset model. h is represented as

$$F_{(t+1)}(X) = F_t(X) + h(X) = Y \text{ (target)} \tag{1}$$

$$h(X) = y - F_t(X) \tag{2}$$

Step 4: Weights are calculated after each iteration using Eq. (3):

$$\Delta w_i = -\eta \frac{\delta C}{\delta w_i}, \tag{3}$$

where cost function $C(W) = \frac{1}{2} \sum_i \left(\text{target}^i - \text{output}^i \right)^2$ and
η is the learning rate that is updated over time as $\eta(t + 1) = \frac{\eta(t)}{(1 + t \times d)}$,
where d is the decreasing constant.

Step 5: Weight is updated as

$$w_i := w_i + \Delta w_i \tag{4}$$

Step 6: End for

Step 7: Features are ranked in descending order according to weight.

The features at the top have the highest weight and are selected as the relevant features. From the dataset taken for consideration, the attributes such as who completed the screening, age, gender, used the application before, country of residence, and screening score are considered not needed, and these features do not provide any usage for our analysis. These six attributes are removed, and the 14 remaining attributes are selected as relevant features for further processing using the proposed FS approach to improve prediction accuracy.

3.6 Ensemble Learning

The best ML algorithms are combined for ensemble learning. Three ML approaches, namely, SMO–SVM, Kohonen SONN, and RF, are used for classifying the features with a high accuracy. These three algorithms perform better individually and combine to provide further better classification in ASD prediction. Features are classified in the ensemble, and these results are processed using bagging concept called max voting to predict the final result.

3.6.1 SMO–SVM

SVM is proven to be the best classification algorithm for all prediction problems. In this proposed work, sequential minimal optimization is used to train the SVM to improve the accuracy of classification result. The linear classification of SVM is represented in Eq. (5):

$$f(x) = w^T(x) + b \tag{5}$$

Our problem is a binary classification, where the output is predicted as y=1 if f(x)>=0 and y=−1 if f(x)<0. The linear function is also improved with the kernel represented as Eq. (6):

$$f(x) = \sum_{i=0}^{n} \alpha_i y^i \times k(x^i, x) + b \tag{6}$$

The kernel function

$$k(x^i, x) = \exp\left(-\frac{||x_i - x||^2}{2\sigma^2}\right), \tag{7}$$

where $||x_i - x||^2$ is the squared Euclidean distance, and σ is a free parameter. With this standard SVM, SOM can improve the SVM for binary classification using Eq. (8):

$$f'(x) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y^i y^j \alpha_i \alpha_j k(x^I, x), \tag{8}$$

where $0 \leq \alpha_i \leq C$, $i = 1, 2 \ldots n$, and $\sum_{i=1}^{n} \alpha_i y^i = 0$.

The optimal solution is based on the condition as represented in Eq. (9):

$$\alpha_i = 0 \Rightarrow y^i(w^T x^i + b) \geq 1$$

$$\alpha_i = C \Rightarrow y^i(w^T x^i + b) \leq 1$$

$$0 \leq \alpha_i \leq C \Rightarrow y^i(w^T x^i + b) = 1 \tag{9}$$

The classification of SOM trained SVM is shown in Fig. 2 based on the optimal solutions. The input data space is divided using the hyper plane, and the decision is based on the SOM based optimal solution conditions.
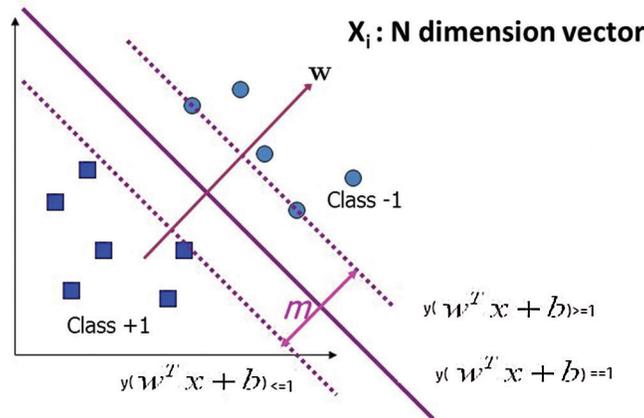
Figure 2: SOM–SVM classification

3.6.2 Kohonen SONN

The Kohonen network is also called SONN, which is a computational method for analyzing high-dimensional data classification. The main objective of Kohonen SONN is to map the arbitrary dimension of input data into a discrete map comprised of neurons. This map is trained to organize data. While training the map, the location of the neuron is not changed where the weights differ based on the value. In self-organization, in the first phase of selection, each neuron has a small weight and input data. In the second phase, the neuron closest to the point is considered the winning neuron, and the neurons near the winning neuron also move toward the point. Euclidean distance is used to find the distance between the neuron and the point. The neuron with the least distance is the winning neuron. This process is repeated for all iterations, and the points are clustered. In this work, Kohonen network is used to determine with a high accuracy whether patients have ASD. The data item with n dimensional Euclidean vectors is represented in Eq. (10):

$$x(t) = |Ed_1(t), \ Ed_2(t), \ \ldots Ed_n(t)|, \tag{10}$$

where t is the index of the data item in the sequence. The ith model is declared as

$$m_i(t) = |\mu_{i1}(t), \mu_{i2}(t), \ldots \mu_{in}(t)| \tag{11}$$

The new value of $m_i(t + 1)$ is calculated by using old $m_i(t)$ and data item x(t) as

$$m_i(t + 1) = m_i(t) + \sigma(t)sk_{ci}(t)x(t) - m_i(t), \tag{12}$$

where $\sigma(t)$ is the scalar vector that defines the size of correction; i is the index of the model in the process; c is the index of the model that has less distance from x(t); sk_{ci} is the smoothing kernel, whose value is one when i=c, and this value decreases while the distance between i and c increases. The new model is computed as

$$m_i = \frac{\sum_j n_j sk_{ji} \bar{x}_j}{\sum_j n_j sk_{ji}}, \tag{13}$$

where n_j is the number of input data items mapped into node j that runs over the neighbor node i. The Kohonen SONN structure and classification are represented in Figs. 3a and 3b, respectively.
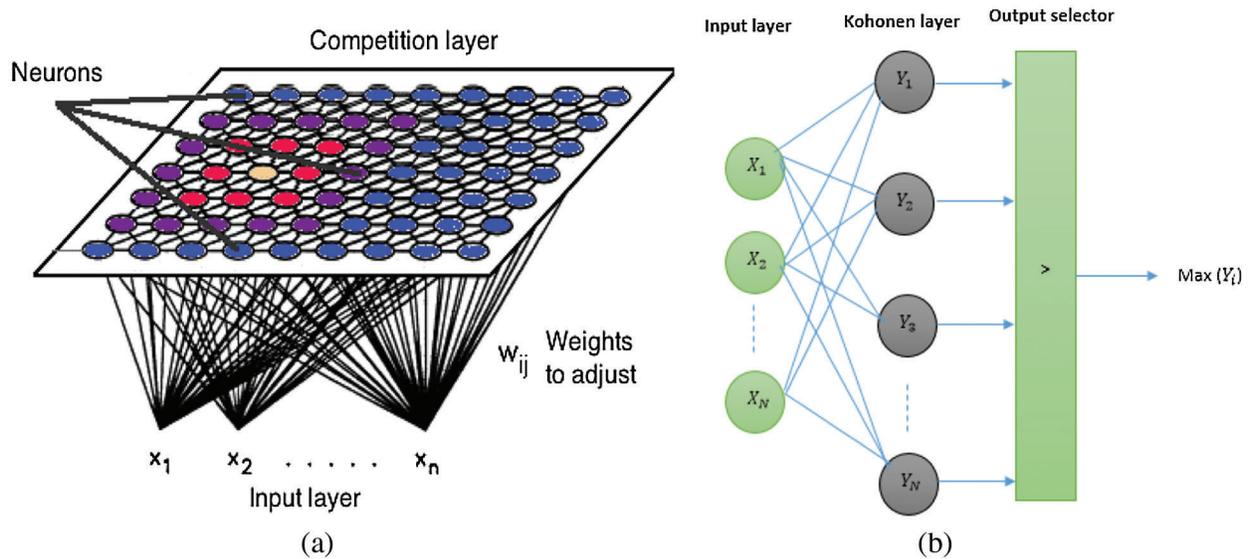
Figure 3: (a) Kohonen SONN (b) Kohonen classification

3.6.3 RF

RF is a classifier based on a decision tree that classifies the data samples as many sub trees. Each tree provides the classification result. Each RF is a collection of decision tree in the form of

$$h(x, \theta_k) \text{ for } k = 1, 2 \ldots K, \tag{14}$$

where θ_k is an independent, identically distributed random feature sample. Each K tree provides the prediction class of data sample X. The subset of a decision tree is constructed based on

$$X = \{(x_I, I)\} \, (|X| = n, x_i \in R^m, \, y_i \in R \tag{15}$$

The probability of the prediction of each subset is represented as

$$p(c|M) = p_1 + p_2 + \ldots + p_n \sum_{i=1}^{n} (p_i(c|M)), \tag{16}$$

where c is the prediction class, M is the feature, p1…pn is the probability of each feature with class value, and n is the number of subsets. If the number of trees is reached, the error is calculated as

$$e \leq \frac{c(1 - s^2)}{s^2}, \tag{17}$$

where c is the correlation between the trees, and s is the metric of strength of the tree.

Algorithm 2: Proposed ASD prediction

Input: Dataset

Output: ASD prediction result (ASD – no ASD)

Step 1: Preprocess: The input dataset is preprocessed using Section 3.2.

Step 2: FS: The preprocessed data are then given as input to FS called

---

Algorithm 2: (continued)

---

bootstrapped gradient boosting presented in Section 3.3 using Algorithm 1.

Step 3: ML: The data with selected features are then given as input to classification

algorithms such as SOM–SVM, Kohonen SONN, and RF

Step 4: Ensemble Learning: The prediction results of each algorithm are calculated separately. The

maximum result is the final result of the proposed ASD prediction.

Max(output (SMO–SVM), output(Kohonen SONN), output (RF))

Step 5: Output: Return the prediction result.

The proposed ensemble-based ASD prediction system accurately predicts ASD in patients using the proposed series of processes such as preprocessing, FS, and classification. The preprocessing of the input data fills the missing values and removes the redundant records. FS determines the relevant features and classification using the best classifiers that accurately categorize the data. Hence, the proposed ASD prediction system is a complete system that involves all the process steps for a complete prediction system. Each step of the proposed process improves prediction accuracy compared with other existing approaches.

4 Results and Discussions

The proposed ensemble-based ASD prediction system is evaluated with the dataset mentioned in Section 3.1 in terms of accuracy, sensitivity, and specificity by using confusion matrix and classification report.

4.1 Evaluation Metrics

To determine the performance, the classification model with the target and performance measurement is an important process. Metrics are used to evaluate the efficiency and effectiveness of the proposed classification model using the test dataset. In this work, the performance of the proposed model is evaluated by metrics such as accuracy, sensitivity, and specificity using equations with confusion matrix elements, as shown in Tab. 3.

$$\text{Sensitivity (SN)} = \frac{TN}{(TN + FN)} \tag{18}$$

$$\text{Specificity (SP)} = \frac{TN}{(TN + TP)} \tag{19}$$

$$\text{Accuracy (AC)} = \frac{TP + TN}{(TN + TP + FP + FN)} \tag{20}$$

Table 3: Confusion matrix elements

Actual ASD values	Predicted ASD values	
	ASD	No ASD
ASD	True positive (TP)	False positive (FP)
No ASD	False Negative (FN)	True Negative (TN)

The experimental results of the proposed system use 14 attributes from the FS method and evaluate the accuracy of the classification using the three algorithms with max voting. The three algorithms are evaluated separately to assess the performance of the proposed ensemble model. These algorithms are implemented with the four datasets such as Toddlers, Children, Adolescents, and Adults. To understand the importance of FS algorithm in classification, Tab. 4 represents the accuracy of the classifier algorithms with and without FS.

Table 4: Accuracy of classifiers with and without FS

Classifiers	Total number of features	Accuracy before FS	Reduced features	Accuracy after FS
SVM	21	91.2	15	93.42
SMO–SVM	21	93	15	94.6
Kohonen	21	89	15	91
RF	21	92	15	94.37
Proposed ASD prediction system	21	94.78	15	99.54

The evaluated result of this Tab. 4 is illustrated in Fig. 4. The graph shows an improvement in the accuracy of all classifiers before and after FS. Among the classifiers, our proposed ASD prediction system obtains a high level accuracy of 99.54% after applying the proposed FS algorithm. Using only the relevant features for classification improves classification accuracy. The overall performance measures of the classification algorithms individually and proposed are shown in Tab. 5.
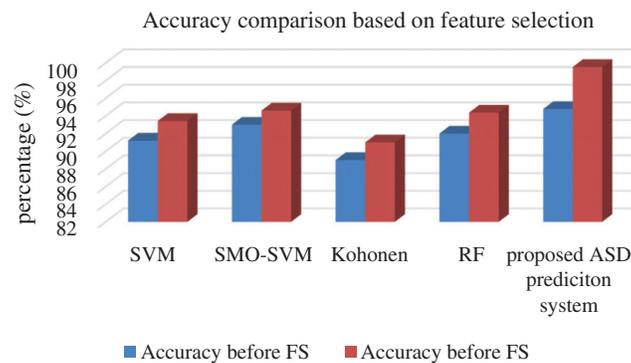

Figure 4: Performance of feature selection algorithm

Table 5: Overall performance of proposed *vs.* existing classifiers for four datasets

Classifiers	Toddlers			Children			Adolescents			Adults		
	SN	SP	Acc	SN	SP	Acc	SN	SP	Acc	SN	SP	Acc
SVM	96	95.75	96.79	92.3	93.23	94.23	90.32	92.32	95.3	92.43	89.34	94.38
SMO–SVM	88.8	92.3	96.22	89.23	93.23	95.98	91.43	94.23	96.3	92.88	95.38	97.37
Kohonen	87.23	93.42	95.32	86.34	92.23	94.26	88.34	93.54	95.3	89.9	94.08	96.89
RF	91.23	90.23	93.42	92.34	91.78	94.27	93	92.4	96.98	94.87	94.03	96.98
Proposed ASD Prediction system	99.33	98.23	99.64	99.12	98.65	99.12	99.01	99.34	99.78	98.98	99.45	99.8

The performance evaluation of various classifiers on four datasets obtains an accuracy in the range of (86%–99.8%). For the evaluation of the Toddlers dataset, our proposed approach obtains 99.64% accuracy that is higher than other algorithms. The next best algorithm is SVM, and the least percentage of accuracy is obtained by the RF approach. The results are illustrated in Fig. 5.
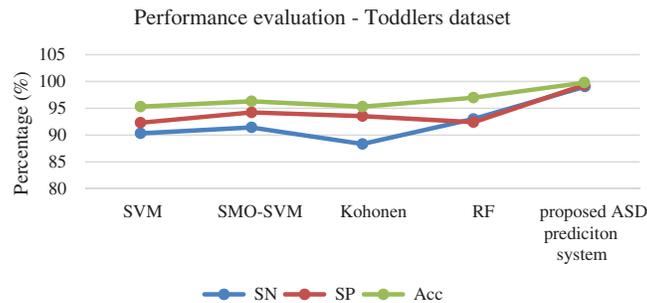

Figure 5: Performance evaluation of classifiers for toddlers dataset

For the evaluation of the Children dataset, our proposed classification approach obtains 99.12% accuracy that is higher than other algorithms. The next best method is SMO–SVM, and the least accuracy is obtained by SVM. The results are illustrated in Fig. 6.
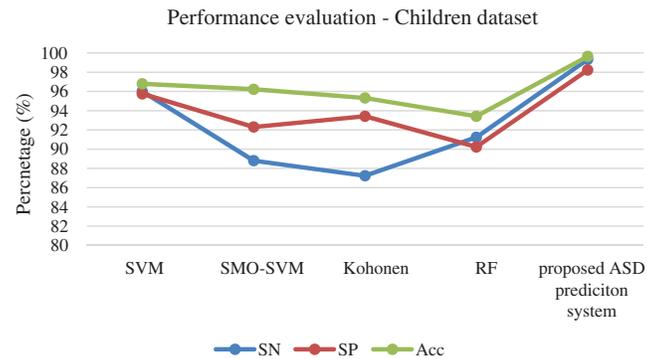

Figure 6: Performance evaluation of classifiers for children dataset

For the evaluation of the Adolescents dataset, our proposed classification approach obtains 99.78% accuracy that is higher than other algorithms. The next best method is RF, and the least accuracy is obtained by SMO and Kohonen. The results are illustrated in Fig. 7.

For the evaluation of the Adults dataset, our proposed classification approach obtains 99.8% accuracy that is higher than other algorithms. The next best method is SMO–SVM, and the least accuracy is obtained by SVM. The results are illustrated in Fig. 8.

The overall performance of our proposed ASD prediction system using ensemble are 99.64% accuracy for the Toddlers dataset, 99.12% accuracy for the Children dataset, 99.78% of accuracy for the Adolescents dataset, and 99.8% accuracy for the Adults dataset. For all four datasets, the proposed model obtains an effective accuracy, and our proposed approach obtains a high accuracy on the Adults dataset.
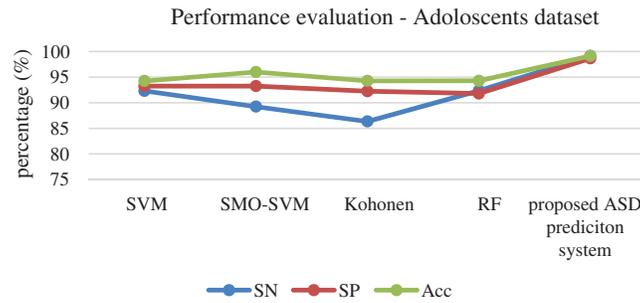
Performance evaluation - Adoloscents dataset

Figure 7: Performance evaluation of classifiers for Adolescents dataset

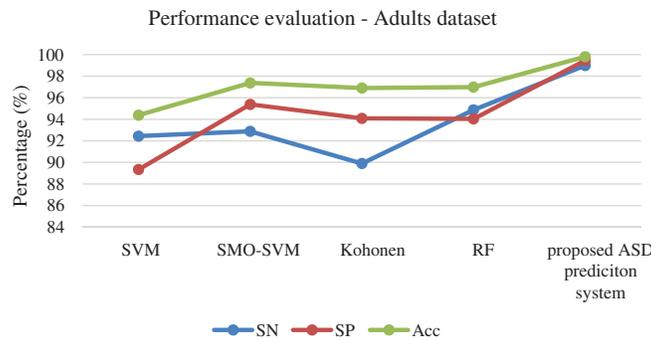Performance evaluation - Adults dataset

Figure 8: Performance evaluation of classifiers for adults dataset

4.2 Classification Errors

During the classification, different types of errors can be observed. Those errors can be expressed as deviation, mean absolute error (MAE), and root mean square error (RMSE), and represent differences between the predicted and observed data. The best classification technique has the lease amount of error. In this proposed work, MAE, RMSE, and relative absolute error (RAE) are calculated using the following equations [26]:

MAE: It is the average of the test samples that is the difference between prediction and actual observation.

$$MAE = \frac{1}{n} \sum_{j=1}^{n} \left| Y_j - \bar{Y}_j \right| \tag{21}$$

RMSE: It is the square root of the average squared differences of prediction and actual observation.

$$RMSE = \sqrt{\frac{\sum_{j=1}^{n} \left(Y_j - \bar{Y}_j \right)^2}{n}} \tag{22}$$

RAE: It is the mean ratio produced by a trivial or base model using the equation

$$RAE = \frac{\sqrt{\sum_{i=2}^{n} \left(P_i - A_i \right)^2}}{\sqrt{\sum_{i=1}^{n} A_i^2}} \tag{23}$$

Classification errors of the proposed model include MAE of 0.04, RMSE of 0.02, and RAE of 0.12, which is the minimum among all the existing classifiers. For the Adolescents dataset, the classification errors of the proposed model include MAE of 0.05, RMSE of 0.02, and RAE of 0.12, which is the minimum among all the existing classifiers. For the Adults dataset, the classification errors of the proposed model include MAE of 0.04, RMSE of 0.01, and RAE of 0.11, which is the minimum among all the existing classifiers. Hence, our proposed ensemble-based ASD prediction system is the best in terms of high accuracy and low error, which obtains a high accuracy on all four datasets and the minimum error rate for all the four datasets.

5 Conclusions

Autism disorder is considered a substantial problem that is difficult to predict and prevent. Having a child with this serious disorder has become challenging for the family. The early diagnosis of ASD is very important for small life. Today, research highly focuses on improving the early prediction and accuracy of autism disorder. This proposed technique achieves 99% of accuracy with an error of 0.02, which is the highest among all other existing ML techniques. The success of ML enables combining the best ML algorithms as ensemble learning and performing a faster prediction in this article. The output of the ensemble learning model is processed with bagging concept. Max voting in bagging concept predicts accurately compared with other prediction algorithms. The future scope of the paper can be improved by collecting all data of patients such as brain MRI, face recognition, body posture, and patient response model to improve prediction accuracy.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] U. Frith and F. Happé, "Autism spectrum disorder," *Current Biology*, vol. 15, no. 19, pp. 786–R790, 2005.

[2] CSR Mandate, "Unlocking unlimited potentials for autistic children," 2017. [Online]. Available: http://www. csrmandate.org/autism-centre-forexcellence-unlocking-unlimited-potentials-for-autisticchildren.

[3] P. Fergus, B. Abdulaimma, C. Carter and S. Round, "Interactive mobile technology for children with autism spectrum condition," in *Proc. CCNC*, Las Vegas, USA, 2014.

[4] W. Liu, M. Li and L. Yi, "Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework," *International Society for Autism Research*, vol. 9, no. 8, pp. 888–898, 2016.

[5] K. H. Kayleigh, N. Marlena, L. Nicholas, C. P. Pelleriti, R. Anden *et al.,* "Applications of supervised machine learning in autism spectrum disorder research: A review," *Review Journal of Autism and Developmental Disorders*, vol. 6, pp. 128–146, 2019.

[6] D. Bone, S. Bishop, M. P. Black, M. S. Goodwin, C. Lord *et al.,* "Use of machine learning to improve autism screening and diagnostic instruments: Effectiveness, efficiency, and multi-instrument fusion," *Journal of Child Psychology and Psychiatry*, vol. 57, no. 8, pp. 927–937, 2016.

[7] F. Thabtah, "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward," *Informatics for Health and Social Care*, vol. 44, no. 3, pp. 278–297, 2019.

[8] K. S. Omar, P. Mondal, N. S. Khan, R. K. Rizvi and M. N. Islam, "A machine learning approach to predict autism spectrum disorder," in *Proc. IECCE*, Bangladesh, IECCE, Cox's Bazar, pp. 1–6, 2019.

[9] M. S. Satu, F. Farida Sathi, M. S. Arifen, M. Hanif Ali and M. A. Moni, "Early detection of autism by extracting features: A case study in Bangladesh," in *Proc. ICREST*, Dhaka, Bangladesh, pp. 400–405, 2019.

[10] O. Altay and M. Ulas, "Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and k-nearest neighbor in children," in *Proc. ISDFS*, Antalya, Turkey, pp. 1–4, 2018.

[11] J. A. Kosmicki, V. Sochat, M. Duda and D. P. Wall, "Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning," *Translational Psychiatry*, vol. 5, no. 2, pp. e514, 2015.

[12] D. Bone, S. L. Bishop, M. P. Black, M. S. Goodwin, C. Lord *et al.,* "Use of machine learning to improve autism screening and diagnostic instruments effectiveness efficiency and multi-instrument fusion," *Journal of Child Psychology and Psychiatry*, vol. 57, no. 8, pp. 927–937, 2016.

[13] F. Hauck and N. Kliewer, "Machine learning for autism diagnostics applying support vector classification," in *Proc*. HIMS, Berlin, Germany, 2017.

[14] A. Sharma and P. Tanwar Deep, "Analysis of autism spectrum disorder detection techniques," in *Proc. ICIEM*, London, UK, 2020.

[15] F. Thabtah and D. Peebles, "A new machine learning model basedon induction of rules for autism detection," *Health Informatics Journal*, vol. 26, no. 1, pp. 264–286, 2020.

[16] T. Akter, S. Satu, I. Khan, M. Hanif Ali and S. Uddin, "Machine learning based models for early stage detection of autism spectrum disorders," *IEEE Access*, vol. 7, pp. 166509–166527, 2019.

[17] R. Vaishali and R. Sasikala, "A machine learning based approach to classify autism with optimum behaviour sets," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 994–1004, 2018.

[18] F. Thabtah, "Autism spectrum disorder screening: Machine learning adaptation and DSM-5 fulfillment," in *Proc. ICMHI*, Kastamonu, Turkey, pp. 1–6, 2017.

[19] M. S. Mythili and A. R. Mohamed Shanavas, "A study on autism spectrum disorders using classification techniques," *International Journal of Soft Computing and Engineering*, vol. 4, pp. 88–91, 2014.

[20] J. A. Kosmicki, V. Sochat, M. Duda and D. P. Wall, "Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning," *Translational Psychiatry*, vol. 5, no. 2, pp. e514, 2015.

[21] B. Li, A. Sharma, J. Meng, S. Purushwalkam and E. Gowen, "Applying machine learning to identify autistic adults using imitation: An exploratory study," *PLOS One*, vol. 12, no. 8, pp. e0182652, 2017.

[22] F. F. Thabtah, "Autistic spectrum disorder screening data for adult," 2017. [Online]. Available: https://archive.ics.uci.edu/ml/machine-learningdatabases/00426.

[23] C. Allison, B. Auyeung and S. Baron Cohen, "Toward brief aIJred flagsâ for autism screening the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 51, no. 2, pp. 202–212, 2012.

[24] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 5, pp. 5–32. 2001.

[25] S. Raj and S. Masood, "Analysis and detection of autism spectrum disorder using machine learning techniques," *Procedia Computer Science*, vol. 167, no. 12, pp. 994–1004, 2020.

[26] M. D. Hossain, M. A. Kabir, A. Anwar and M. Z. Islam, "Detecting autism spectrum disorder using machine learning techniques: An experimental analysis on toddler, child, adolescent and adult datasets," *Health Information Science and Systems*, vol. 9, no. 1, pp. 1–17, 2021.