A Machine Learning Way to Classify Autism Spectrum Disorder

https://doi.org/10.3991/ijet.v16i06.19559

Sujatha R, Aarthy SL
Vellore Institute of Technology, Vellore, India

Jyotir Moy Chatterjee
Lord Buddha Education Foundation, Kathmandu, Nepal

A. Alaboudi
Shaqra University, Shaqra, Saudi Arabia

N Z Jhanjhi (✉)
Taylor's University, Selangor, Malaysia
noorzaman.jhanjhi@taylors.edu.my

Abstract—In recent times Autism Spectrum Disorder (ASD) is picking up its force quicker than at any other time. Distinguishing autism characteristics through screening tests is over the top expensive and tedious. Screening of the same is a challenging task, and classification must be conducted with great care. Machine Learning (ML) can perform great in the classification of this problem. Most researchers have utilized the ML strategy to characterize patients and typical controls, among which support vector machines (SVM) are broadly utilized. Even though several studies have been done utilizing various methods, these investigations didn't give any complete decision about anticipating autism qualities regarding distinctive age groups. Accordingly, this paper plans to locate the best technique for ASD classification out of SVM, K-nearest neighbor (KNN), Random Forest (RF), Naïve Bayes (NB), Stochastic gradient descent (SGD), Adaptive boosting (AdaBoost), and CN2 Rule Induction using 4 ASD datasets taken from UCI ML repository. The classification accuracy (CA) we acquired after experimentation is as follows: in the case of the adult dataset SGD gives 99.7%, in the adolescent dataset RF gives 97.2%, in the child dataset SGD gives 99.6%, in the toddler dataset AdaBoost gives 99.8%. Autism spectrum quotients (AQs) varied among several scenarios for toddlers, adults, adolescents, and children that include positive predictive value for the scaling purpose. AQ questions referred to topics about attention to detail, attention switching, communication, imagination, and social skills.

Keywords—ASD, ML, SVM, KNN, RF, NB, SGD, AdaBoost, CN2, AQ, CA.

1 Introduction

The autism spectrum disorder (ASD) screening process differs according to age. Two global classification systems for ASD diagnosis, namely, the Diagnostic Statistical Manual (DSM-5), which is provided by the American Psychiatric Association and considers the condition as a single diagnosis by removing subgroups, and the International Classification of Disease (ICD-11), created by the World Health Organization (WHO). According to the DSM, autism and intellectual disability occur concurrently. By contrast, the ICD provides a detailed guide to distinguish autism prevailing with and without an intellectual disability; it also considers historical data on loss of previous skill in the diagnostic process. The most difficult aspect of diagnosing ASD is that no single pathognomonic feature exists and all symptoms revolve around the modification of an individual's behavioral profile, which varies according to age and severity.

In the prevailing system, classification is carried out using datasets of cases collected from a versatile group. The data depend on the autism diagnostic observation schedule (ADOS) & autism diagnostic interview (ADI), which is conducted in a clinical setting. ADOS sessions are 30–45 minutes long, and the examiner records the provided responses. ADI refers to interviews of suspected autism individuals over 18 with their parents or caregivers in the clinic. The interview is performed in five phases using a questionnaire that probes areas related to communication, social development, play, restricted behavior, and general skills. The individual's responses are evaluated by using scoring algorithms, and 3 main domains, namely, etymological & communiqué, societal relation, constrained & tedious behavior, are assessed. Cumulative scores exceeding the corresponding cut-off values indicate a positive syndrome that must be addressed immediately by proper diagnosis. Determining the most prominent features from a massive dataset is challenging work that must be done by careful analysis. Data processing tasks also present a potential hurdle in managing missing values in attributes. The rest of the process of applying machine learning mostly relies upon the quality of information taken into consideration. Automation based on the diagnostic perspective must be fine-tuned.

ASD is a neurodevelopment disorder that can occur in adults, adolescents, children, and toddlers. Leo Kanner refers to autism as a prototypical condition with a spectrum of presentations and phenotypes that become more subtle in terms of behavioral features when a change in the environment occurs. It is categorized through interactive irregularities in communiqué and shared societal relation, organized thru outlines of monotonous, controlled, & typecast safeties & actions. These issues typically extent during infantile & are probable to increase in the intensity of diverse ages. The heterogeneousness of the exaggerated entities & their hereditary intricacy has helped researchers identify the causes of ASD. Diagnosis of ASD is a lengthy process and varies from individual to individual. Symptoms also change across one's lifespan. ASD can be difficult to detect in young children, and parent raises the concern after the persistent monitoring of the children which delays the process of early diagnosis.

Our contribution to this work is as follows:

- We designed to classify the patient is affected by autism or not based on the various attributes using a machine learning process.
- For the new record, prediction of syndrome and treatment at the earliest stage can be facilitated to prevent worse condition at the earliest.
- The health-related sector requires more accurate and precise estimation at the earliest.

This article is prepared as follows: Related work is presented followed by the proposed method via a workflow. Results and a discussion are then provided, and the conclusions are summarized.

2 Literature Review

ASD denotes a neurodevelopmental issue categorized by confinements in social associations, correspondence, & conduct that become progressively regular [1]. The reasons for ASD is generally connected to hereditary & neural factors; however, they are fundamentally analyzed by utilizing non-hereditary factors identified with conduct, such as social cooperation, play, creative thinking, monotonous practices, and correspondence, amongst others [2]. Prevailing approximations disclose that approximately 1.5% of the populace is on the range, & many persons on the range are believed to remain undetected [3]. Accordingly, the need for quick analyzing amenities conforming by developing alertness of ASD [4]. Wall et al. proposed numerous data mining techniques in an alternating decision-tree algorithm (ADTree) to moderate the count of substances present in the ADOS-Revised test. This work intended to hasten ASD analysis so that members, including family, could utilize the necessary services provided. To accomplish this goal, the authors removed instances of non-ASD cases & then investigated the classification frameworks produced by the ADTree calculation on an imbalanced dataset. The WEKA software was subsequently utilized to evaluate the classification accuracy obtained by using the ADTree method. Subsequently examining the outcomes of the ADTree calculation, the authors found that among the 29 objects included in the ADOS-Revised test, solitary 8 features appear in the classification framework; thus, the group believed that the 29 items could be represented by only these 8 items. There is a necessity to reconsider the features includes within ASD diagnostic tool to satisfy a smaller number of items sets though keeping up the sensitivity & validity of the test [5, 6].

ASD prediction-based ML requires cautious examination, particularly when managing diagnostic strategies employing techniques in the clinical setting. Limiting the ADOS-Revised test to eight items may result in misleading results because exercises must be directed by the clinician on an experiment before the grouping [6,8]. Duda et al. [7] conducted a realistic investigation associating numerous smart methods to differentiate amongst ASD & attention deficit hyperactivity disorder (ADHD). Six methods were differentiated on a dataset with 65 items obtained from the Simons Simplex Collection version 15.41. Information was gathered by utilizing a parent-directed survey symptomatic strategy called the Social Responsiveness Scale. A preprocessing stage was conducted by the author to (1) dispose of occurrences that had at least four missing qualities, (2) balance data collection by using the under-sampling procedure, and (3) diminish information dimensionality by using feature selection strategies. Chu et al. [9] explored several approaches to separate ADHD and obstructive sleep apnea (OSA) by using the data of 217 kids who had been diagnosed as having ADHD, OSA, or a combination of ADHD & OSA as per the Diagnostic & Statistical Manual of Mental Disorders (fourth edition; DSM IV) standards. Information was gathered by utilizing a diverse diagnostic tool, and three ML techniques were used to infer classifiers that could help clinicians & doctors improve diagnostic criteria. Detailed outcomes demonstrated that 17 highlights show significant distinctions amongst 3 groups of pervasive developmental disorders (PDDs), especially in the Child Behavior Checklist (CBCL). Moreover, compared with the neural network and CHAID algorithm, the decision tree generated classifiers faster.

Wolfers et al. [10] researched issues identified with PDDs, having trivial sample extents, exterior legitimacy, & ML methodic difficulties, lacking focus on ASD. Lopez Marcano [11] inspected the appropriateness of various methods, for example, neural system & RF, to minimize the period required for ASD diagnosis. Maenner et al. [12] examined the RF method on a dataset obtained from the Georgia Autism & Developmental Disabilities Monitoring Network using expressions & words acquired in youngsters' formative assessments. The dataset comprised 5,396 assessments for 1,162 offspring, 601 of whom were on the range. The RF classifiers were assessed on an autonomous test informational collection containing 9,811 assessments of 1,450 youngsters. The outcomes revealed that RF achieves approximately 89% predictive ability & 84% sensitivity. Thabtah dissected limitations related to testing reads that embraced ML for ASD classification [13-15]. The goal of [36] is through a literature survey:

a) To outline the emotional turn of events and instruction of people on the range
b) To present the discoveries of examinations
c) To present and raise key worries about the emotional intelligence of kids range of autism
d) Bring up issues about the advancement of instructive techniques pointed toward improving the emotional improvement of people in the autism range and subsequently the advancement of social sentiments their maternal aptitudes

[37] gives a brief & agent depiction of the job that artificial intelligence plays these days at the evaluation of autism. [38] features the point-by-point research occurred between 2010 - 2020, while looking at the effect of robots on medically introverted kids through their connection, utilization of craftsmanship, programming, etc.

Table 1. Comparative analysis of various existing works with their advantages and limitations

Sl. No.	Year	Advantages	Limitations
1	2020 [31]	Behavior inflexibility (BI) for youngsters with formative inabilities was built by utilizing a multi-step procedure with the help of a parent	Studies need to analyze the focalized and unique legitimacy of the Behavioral Inflexibility Scale (BIS) using a multi-method approach.
2	2020 [32]	An unsupervised online learning model was built for ASD grouping.	Models must be prepared by using the dataset, rather than simply employing a pre-trained model.
3	2020 [33]	Utilizing the Stockholm Youth Cohort, authors analyzed anxiety syndrome amongst mentally imbalanced adults (n = 4,049) with and without scholarly inability against a population control (n = 217,645).	More investigation is necessary to govern the reasons for anxiety among individuals through ASD. Future research is expected to improve the understanding of the phenomenologist of anxiety syndromes and enhance methods to estimate and treat anxiety.
4	2019 [34]	Gaussian mixed models and hierarchical clustering were applied to distinguish among social phenotypes of ASD and assess treatment reactions over scholarly phenotypes.	A limitation of the present investigation is the absence of information from institutionalized appraisals.
5	2018 [35]	Ongoing investigations on mental imbalance were examined. This work not only articulated previously mentioned issues but also suggested ways to improve AI use in ASD in terms of conceptualization, execution, and information.	No implementation work was shown.

3 Proposed Methodology

3.1 Workflow

The data available in the UCI repository were obtained for our work and collected with the help of a mobile application (hereinafter referred to as an app) developed to perform four ASD screening methods, namely, the autism-spectrum quotient (AQ) of Adult, Adolescent, Child, and Toddler. Dataset available in the UCI repository includes clean data without missing values. Since the dataset has approximately 21 features and 1 class labeled autism and non-autism. The features are age, sex, jaundice during birth, ASD for any family member, questions A1–A10, and ASD score from the application was used to classify the work as autism or non-autism. Principal components were retrieved by using a principal component analysis (PCA)-algorithm then applied to a minimized dataset. We considered five eigenvectors from the given data and co-occurrence matrices and then fed the system to different classifier algorithms with a cross-fold value of 10. The system was classified by using the different algorithms, and the best algorithm for early diagnosis of ASD was identified by using precision, recall, F1 score, and accuracy values. Figure 1 illustrates the proposed workflow.

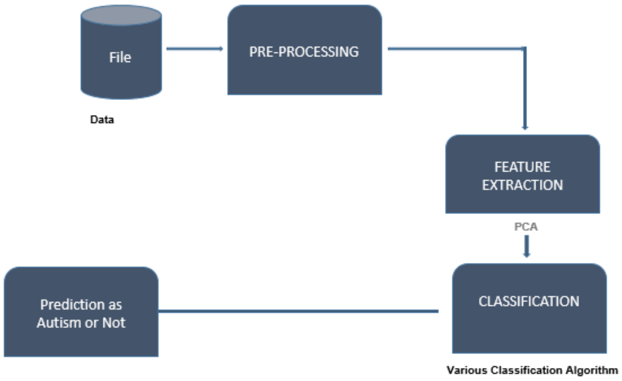

Fig. 1. Workflow

3.2 Data collection & description

Four classes include data adolescent, autism data adult, autism child process, and toddler. The dataset included the following attributes: age, sex, jaundice during birth, ASD for any family member, residence, previous app use, screening, language, and classes. The screening test was conducted among age groups of 4–11 years, 12–16 years, and 17 years and older. Upon completion of the test by the user (questions A1–A10), a screen appeared so that the user can review and modify his/her responses. Before the data gets saved, it allows the users to verify the filled data as part of quality assurance. The value "0" or "1" is recorded based on the response given by the participants. The attributes and their data types are illustrated in Table 1.

Table 2. Attribute Description

Attribute	Format	Description
Age	Number	Toddler in months, child, adolescent, and adult in years
Sex	String	Male or female
Ethnicity	String	Common ethnicities in text format
Jaundice	Boolean (Yes or No)	Born with jaundice?
Family_ ASD	Boolean (Yes or No)	Does an immediate family member have an ASD?
The test is taken from	String	Parent, self, caregiver, medical staff, clinician, etc.
Residence	String	List of countries in text format
Previous app use	Boolean (Yes or No)	Was a screening test taken?
Screening method type (A1–A10)	Binary (0,1)	Question method type
Score	Integer	Values generated based on conditions
Screening type	Integer	Age of the individual
Language	String	Regional language
Class	Boolean (Yes or No)	Class description

3.3 Attribute extraction

The attributes were extracted by using PCA. Certain rules are associated with attribute extraction, as discussed below. The main idea of PCA is to diminish the dimensionality of the dataset variables available in the given input data. Principal components are created in the process of PCA using orthogonal transformation by transferring the set of possible correlated components or variables into a set of linearly uncorrelated variables. In the flow of work, we used five principal components (PC1–PC5) derived from the set of data inputs after preprocessing. These vectors have been used as feature extraction variables for the rule-based algorithm described in our previous work [16].

The steps involved in PCA begins with normalization of the data, identification of covariance matrix, computation of the eigenvalues, and vectors followed by choosing the principal component than creating the attribute vector.

4 Classification Algorithms

According to the workflow for ASD diagnosis and prediction, the dataset is first framed, after which feature selection is conducted. The severity of autism is calculated by applying machine learning classification algorithms. After a review of their characteristics, the following supervised classification algorithms are applied.

4.1 SVM

The goal of SVM is to compute a hyperplane in the N-dimensional field. To separate as two classes of data points, the number of hyperplanes selected. We intend to get an aircraft with the greatest margin, i.e., the optimal distance among two-class data points. The trustworthiness of future data points is improved by optimizing the margin gap. Hyperplanes are conclusion limits that help classify data points. The data points on each side of the hyperplane can be allocated to separate groups. The hyperplane dimension also relies on the number of characteristics. If the input number is 2, for example, the hyperplane is only one line. A hyperplane is a two-dimensional plane if the number of features to be entered is 3. The number of features approaching 3 is difficult to imagine. Vectors supporting the hyperplane are similar data points and influence the hyperplane position and orientation. Support vectors are used to optimize the margin of the classifier. The elimination of support vectors would change the hyperplane’s location. These concepts were used to build our SVM [17-18].

4.2 K-NN

KNN is a data mining algorithm utilized for classification purposes. The steps involved in the method are as follows:

1) Obtain the unclassified data
2) Evaluate the distance from new data to all other already categorized (Euclidian, Manhattan, Minkowski, or weighted) data
3) Calculate k value
4) Review the list of classes at the minimum distance, counting the number of every appearing class
5) Selection of the class that occurs most often as the right one
6) Classify actual data with the class obtained in (5)

The distance between two points can be easily calculated using several formulas [19-20].

The formula for the Euclidean distance is as follows and a, b stands for points:

$$D = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2} \qquad (1)$$

4.3 RF

A decision tree is an abstract typical form that can be used as a building block for an RF. This paradigm is interpretable because the classifications are familiar before a decision reached (in an ideal world), how the issue affecting the data is built are the technical details of a decision tree. The decision tree in the Classification and Regression Trees (CART) algorithm is constructed by evaluating the questions (called node splits) contributing to the largest reduction in Gini impurities when responded. This indicates that the decision tree tries to create nodes involving a high ratio of datasets (data points) from a single class by locating values in the attributes that split the data cleanly into classes [21-22].

$$I_G(n) = 1 - \sum_{i=1}^{f}(p_i)^2 \qquad (2)$$

4.4 NB

A classifier is a model used to distinguish between objects based on certain characteristics. The NB classifier is a deterministic prediction system model. The cluster is focused on the principle of Bayes. Finding the likelihood of A occurring as B is happening is conducted by using NB [23]. Here B is the proof, and A is the assumption. The predictions/features here are believed to be independent, i.e., one function has no impact on the other. The various kinds of NB Classifiers namely Multinomial, Bernoulli, and Gaussian.

4.5 AdaBoost (AB)

AB is a sub-algorithm used for machine learning established by Yoav Freund & Robert Schapire, who received the Nobel Prize for their research in 2003. It can be used to enhance performance in combination with several other learning algorithms. The performance of other optimization algorithms is incorporated into a weighted sum representing the boosted classifier's overall results. While AdaBoost is prone to loud outliers and data, it is less vulnerable than most other learning algorithms to overfitting issues in several situations. AdaBoost is frequently known as the satisfactory out-of-the-field classifier. However, the pattern is introduced at every level of the AB set of rules [24-25].

4.6 SGD

Stochastic refers to a random probability-related scheme or method. Thus, in SGD, several samples, rather than the whole set of data for each iteration, are randomly chosen. In the descending gradient, the term batch indicates the maximum set of data from a sample used to measure the gradient of iteration. The goal of the SGD is to find a better way of traveling the error surface so that minimum error value is achieved quickly without resorting to brute force search, therefore it is very costly to perform computationally. SGD solves this problem because, in SGD, only one sample is used for iteration. The batch is spontaneously mixed and chosen to carry out the computation process [26].

$$for\ i\ in\ range\ (m){:}\ \theta_j - \alpha\ (\hat{y}^i - y^i)x_j^i \tag{3}$$

4.7 CN2 rule induction

CN2 rule induction is a classification algorithm that works on rules based on a condition followed by a prediction class [27] on different datasets.

1. Adult

```
If PC1<=1.034 AND PC1>=1.034 AND PC1<=1.067 AND
PC1>=1.067 AND PC1<=1.2963 AND PC1>=1.2972 AND
PC1<=1.434 AND PC1>=1.434
ELSE IF PC1<=1.080 AND PC2>=0.260 AND PC2>=3.464
THEN
CLASS=NO
```

2. Adolescent

```
IF PC1<=-0.507 AND PC4>=-0.736 AND PC1<=-0.330 AND
PC3>=- 0.782 AND PC2>=2.276 AND PC4>=1.297
THEN
CLASS=YES
ELSE IF PC1>=0.507 AND PC1>=0.507 AND PC4>=2.589
THEN
CLASS=NO
```

3. Child

```
IF PC1<=-0.3214 AND PC2>=0.4409 AND PC4>=0.8436 AND
PC2>=0.7726 AND PC4>=0.8436
THEN
CLASS=YES
ELSE IF PC1>=0.1277 AND PC1<=-0.8413 AND PC1>=-0.8413
AND PC1<=0.563 AND PC1>=-0.5635
THEN
CLASS=NO
```

4. Toddler

```
IF PC1<=1.5616 AND PC5>=-1.0834 AND PC3>=2.563 AND
PC4>=-0.864 AND PC3>=1.959 AND PC5>=2.781
THEN
CLASS=YES
ELSE IF PC1>=1.7457 AND PC1<=1.5118 AND PC1>=1.5118 AND
PC1<=1.6124 AND PC1>=1.6124
THEN
CLASS=N
```

5 Result and Discussion

5.1 Performance metrics

The performance of the entire system architecture was calculated based on F1 scores, precision, recall, and accuracy [28-30].

5.2 F1 Score

F1 score is the weighted standard of precision and recall. The score evaluates false positives and negatives. While it is not as straightforward as exactness, the F1 score is normally more helpful than precision. Precision works best if false positives and negatives have a comparative expense. If the expense of false positives and negatives are altogether different, precision and recall may be more informative. In the adult dataset, the F1 score for AB is 0.993, which is higher than the F1 scores obtained from the other methods. In the adolescent dataset, the F1 score of RF is 0.972, which is higher than the F1 scores obtained from the other methods. In the child dataset, the F1 score of SGD is 0.996, which is higher than the F1 scores obtained from the other methods. In the toddler dataset, the precision rate of SGD is 0.996, which is higher than the F1 scores obtained from the other methods.

$$F1\,Score = 2 * \frac{recall * precision}{recall + precision} \tag{4}$$

5.3 Precision

Precision is the ratio of effectively anticipated optimistic perceptions compared with all-out anticipated positive perceptions. In the adult dataset, the precision rate of SGD is 0.997, which is higher than the precision scores obtained from the other methods. In the adolescent dataset, the precision rate of RF is 0.972, which is higher than the precision scores obtained from the other methods. In the child dataset, the precision rate of SGD is 0.996, which is higher than the precision scores obtained from the other methods. In the toddler dataset, the precision rate of SGD is 0.996, which is higher than the precision scores obtained from the other methods.

$$Precision = \frac{tp}{tp+fp} \tag{5}$$

5.4 Recall

Recall refers to the ratio of effectively anticipated optimistic perceptions versus all perceptions in the real class. In the adult dataset, the recall of SGD is 0.997, which is higher than the recall scores obtained from the other methods. In the adolescent dataset, the recall of RF is 0.972, which is higher than the recall scores obtained from the other methods. In the child dataset, the recall of SGD is 0.996, which is higher than the recall scores obtained from the other methods. In the toddler dataset, the recall of AdaBoost is 0.997, which is higher than the recall scores obtained from the other methods.

$$Recall = \frac{tp}{tp+fn} \tag{6}$$

Deploying the autism dataset on various machine learning algorithms provides insights into the type of algorithm yielding optimal results. Figures 2–5 provide a comparative analysis of these methods.

5.5 Classifier accuracy

In the adult dataset, the highest accuracy (99.7%) was obtained from SGD. In the adolescent dataset, the highest accuracy (97.2%) was obtained from RF. In the child and toddler datasets, the highest accuracies were obtained from SGD and RF.

Figure 2 describes the performance values obtained for SVM, KNN, RF, NB, AB, SGD, and CN2 rule inducer. The F1 score, precision, and recall of each algorithm were obtained for the child ASD dataset, and the RF algorithm yielded the highest value of 0.98.

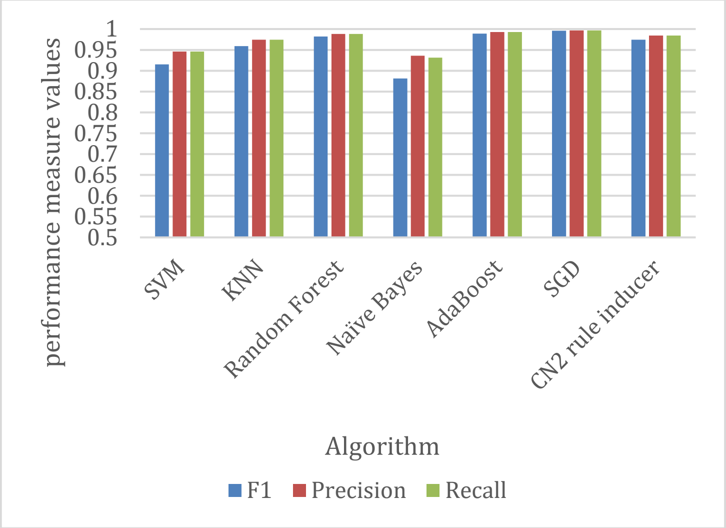

Fig. 2. Performance evaluation on the adult ASD dataset

Figure 3 describes the performance values obtained for SVM, KNN, RF, NB, AB, SGD, and CN2 rule inducer. The F1 score, precision, and recall of each algorithm were obtained for the adolescent ASD dataset, and the RF algorithm yielded the highest value of 0.97.

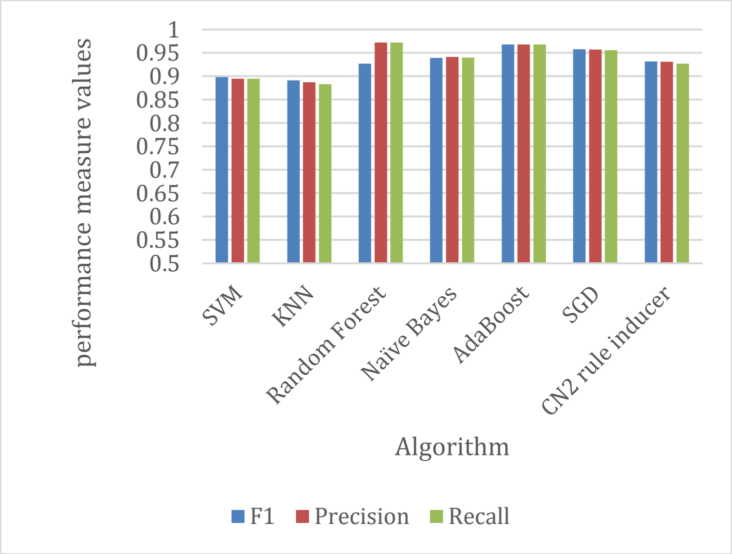

Fig. 3. Performance evaluation on the adolescent ASD dataset

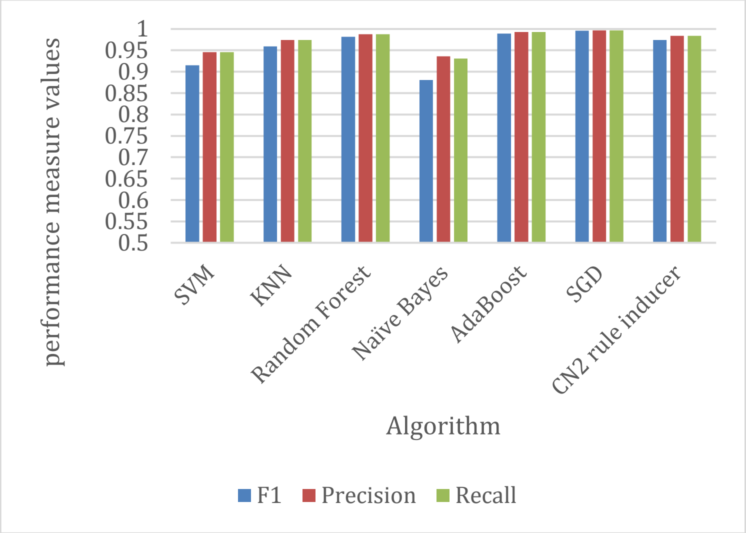

Fig. 4. Performance evaluation on the toddler ASD dataset

Figure 4 describes the performance values obtained for SVM, KNN, RF, NB, AdaBoost, SGD, and CN2 rule inducer. The F1 score, precision, and recall of each algorithm were obtained for the toddler ASD dataset, and the RF algorithm yielded the highest value of 0.99.

Figure 5 describes the performance values obtained for SVM, KNN, RF, NB, AB, SGD, and CN2 rule inducer. The F1 score, precision, and recall of each algorithm were obtained for the child ASD dataset, and the SGD algorithm yielded the highest value of 0.99.

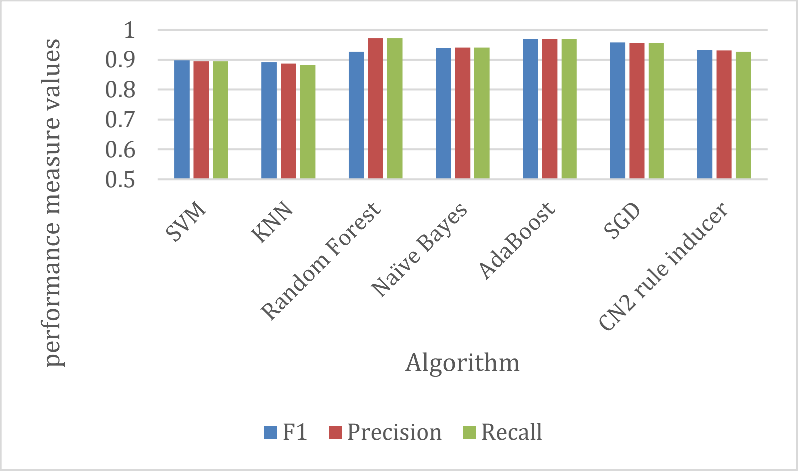

Fig. 5. Performance evaluation on the child ASD dataset

5.6 Accuracy value

Table 3. Comparison between various classification methods

Accuracy	Adult Data	Adolescent Data	Child Data	Toddler Data
SVM	94.6	89.5	94.1	91.7
KNN	97.4	88.3	95.9	97.8
RF	98.8	97.2	98.1	99.7
NB	93.1	94	95.9	95.8
AB	99.3	96.8	97.9	99.8
SGD	99.7	95.6	99.6	99.7
CN2 Rule	98.4	92.7	97.5	99.3

Figure 6 describes the performance values obtained for SVM, KNN, RF, NB, AB, SGD, and CN2 rule inducer and it shows the cumulative accuracy chart. The best algorithm for adult dataset is SGD, which yields an accuracy of 99.3%. The best algorithm for adolescent dataset is RF, which has an accuracy of 97.2%. The best algorithm for child dataset is SGD, which yields an accuracy of 99.6%. Finally, the best algorithms for the toddler dataset are RF and SGD, both of which yield 99.7% accuracy.

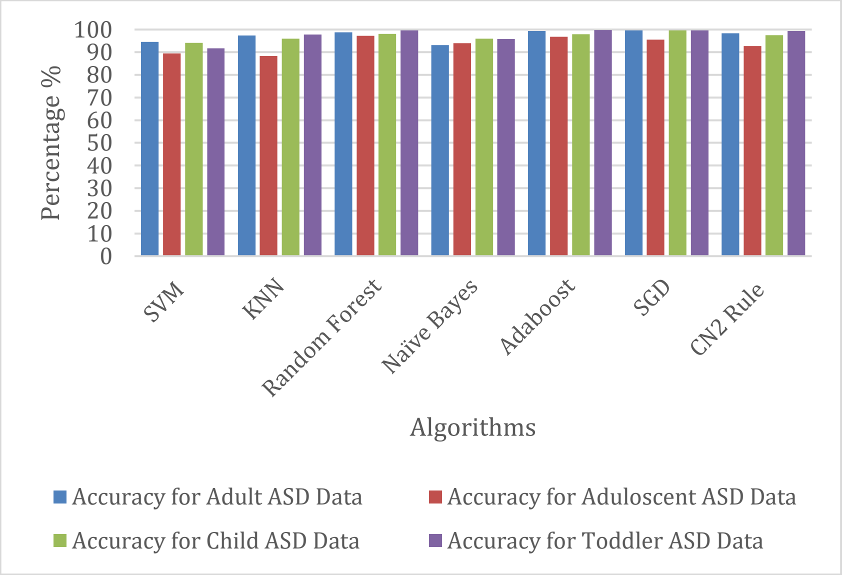

Fig. 6. Comparison of the various algorithm over Autism Spectrum Disorder

6 Conclusion

Awareness of ASD has rapidly increased, and several methods to diagnose and treat the condition as early as possible have been developed. Many researchers worldwide have developed screening and diagnosis methods to detect ASD and assist in its medical diagnosis. In particular, the development of machine learning algorithms provides great support for the medical field. A stakeholder of these projects are patients, the caretakers who can provide the best insight about the patients, medical practitioners, psychologists, behavioral science, and neuroscience.

In this work, we used several classification algorithms to make the best prediction. Various potential supervised classification algorithms were applied over the dataset and trained the model. The performance was evaluated based on accuracy, precision, recall, and F1 score.

The work has shown a predominant result, and the system can be further trained with deep learning procedures to improve the early detection of ASD.

7 References

[1] Ruzich, Emily, Carrie Allison, Paula Smith, Peter Watson, Bonnie Auyeung, Howard Ring, and Simon Baron-Cohen. "Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females." Molecular autism 6, no. 1 (2015): 2. https://doi.org/10.1186/2040-2392-6-2

[2] Ramaswami, Gokul, and Daniel H. Geschwind. "Genetics of autism spectrum disorder." In Handbook of clinical neurology, vol. 147, pp. 321-329. Elsevier, 2018. https://doi.org/10.1016/b978-0-444-63233-3.00021-x

[3] Brugha, Traolach S., Sally McManus, John Bankart, Fiona Scott, Susan Purdon, Jane Smith, Paul Bebbington, Rachel Jenkins, and Howard Meltzer. "Epidemiology of autism spectrum disorders in adults in the community in England." Archives of general psychiatry 68, no. 5 (2011): 459-465. https://doi.org/10.1001/archgenpsychiatry.2011.38

[4] Russell, Ailsa J., Clodagh M. Murphy, Ellie Wilson, Nicola Gillan, Cordelia Brown, Dene M. Robertson, Michael C. Craig et al. "The mental health of individuals referred for assessment of autism spectrum disorder in adulthood: a clinic report." Autism 20, no. 5 (2016): 623-627. https://doi.org/10.1177/1362361315604271

[5] Levy, Sebastien, Marlena Duda, Nick Haber, and Dennis P. Wall. "Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism." Molecular autism 8, no. 1 (2017): 65. https://doi.org/10.1186/s13229-017-0180-6

[6] Kavakiotis, Ioannis, Olga Tsave, Athanasios Salifoglou, Nicos Maglaveras, Ioannis Vlahavas, and Ioanna Chouvarda. "Machine learning and data mining methods in diabetes research." Computational and structural biotechnology journal 15 (2017): 104-116. https://doi.org/10.1016/j.csbj.2016.12.005

[7] Duda, M., R. Ma, N. Haber, and D. P. Wall. "Use of machine learning for behavioral distinction of autism and ADHD." Translational psychiatry 6, no. 2 (2016): e732. https://doi.org/10.1038/tp.2015.221

[8] Khabbaz, Amir H., Ali A. Pouyan, Mansoor Fateh, and Vahid Abolghasemi. "An adaptive RL Based fuzzy game for autistic children." In 2017 Artificial Intelligence and Signal Processing Conference (AISP), pp. 47-52. IEEE, 2017. https://doi.org/10.1109/aisp.2017.8324105

[9] Chu, Kuo-Chung, Hsin-Jou Huang, and Yu-Shu Huang. "Machine learning approach for distinction of ADHD and OSA." In 2016 IEEE/ACM international conference on advances

in social networks analysis and mining (ASONAM), pp. 1044-1049. IEEE, 2016. https://doi.org/10.1109/asonam.2016.7752370

[10] Wolfers, Thomas, Jan K. Buitelaar, Christian F. Beckmann, Barbara Franke, and Andre F. Marquand. "From estimating activation locality to predicting disorder: a review of pattern recognition for neuroimaging-based psychiatric diagnostics." Neuroscience & Biobehavioral Reviews 57 (2015): 328-349. https://doi.org/10.1016/j.neubiorev.2015.08.001

[11] Lopez Marcano, Juan L. "Classification of ADHD and non-ADHD using AR models and machine learning algorithms." PhD diss., Virginia Tech, 2016.

[12] Maenner, Matthew J., Marshalyn Yeargin-Allsopp, Kim Van Naarden Braun, Deborah L. Christensen, and Laura A. Schieve. "Development of a machine learning algorithm for the surveillance of autism spectrum disorder." PloS one 11, no. 12 (2016): e0168224. https://doi.org/10.1371/journal.pone.0168224

[13] Thabtah, Fadi, Firuz Kamalov, and Khairan Rajab. "A new computational intelligence approach to detect autistic features for autism screening." International journal of medical informatics 117 (2018): 112-124. https://doi.org/10.1016/j.ijmedinf.2018.06.009

[14] Thabtah, Fadi. "An accessible and efficient autism screening method for behavioural data and predictive analyses." Health informatics journal 25, no. 4 (2019): 1739-1755. https://doi.org/10.1177/1460458218796636

[15] Thabtah, Fadi, and David Peebles. "A new machine learning model based on induction of rules for autism detection." Health informatics journal (2019): 1460458218824711. https://doi.org/10.1177/1460458218824711

[16] Song, Fengxi, Zhongwei Guo, and Dayong Mei. "Feature selection using principal component analysis." In 2010 international conference on system science, engineering design and manufacturing informatization, vol. 1, pp. 27-30. IEEE, 2010. https://doi.org/10.1109/icsem.2010.14

[17] Gholami, Raoof, and Nikoo Fakhari. "Support vector machine: principles, parameters, and applications." In Handbook of Neural Computation, pp. 515-535. Academic Press, 2017. https://doi.org/10.1016/b978-0-12-811318-9.00027-2

[18] Huang, Shujun, Nianguang Cai, Pedro Penzuti Pacheco, Shavira Narrandes, Yang Wang, and Wayne Xu. "Applications of support vector machine (SVM) learning in cancer genomics." Cancer Genomics-Proteomics 15, no. 1 (2018): 41-51. https://doi.org/10.21873/cgp.20063

[19] Dey, Lopamudra, Sanjay Chakraborty, Anuraag Biswas, Beepa Bose, and Sweta Tiwari. "Sentiment analysis of review datasets using naive bayes and k-nn classifier." arXiv preprint arXiv:1610.09982 (2016). https://doi.org/10.5815/ijieeb.2016.04.07

[20] Gök, Murat. "An ensemble of k-nearest neighbours algorithm for detection of Parkinson's disease." International Journal of Systems Science 46, no. 6 (2015): 1108-1112. https://doi.org/10.1080/00207721.2013.809613

[21] Chen, Wei, Xiaoshen Xie, Jiale Wang, Biswajeet Pradhan, Haoyuan Hong, Dieu Tien Bui, Zhao Duan, and Jianquan Ma. "A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility." Catena 151 (2017): 147-160. https://doi.org/10.1016/j.catena.2016.11.032

[22] Sun, Guanglu, Shaobo Li, Yanzhen Cao, and Fei Lang. "Cervical cancer diagnosis based on random forest." International Journal of Performability Engineering 13, no. 4 (2017): 446-457.

[23] Gao, Chong-zhi, Qiong Cheng, Pei He, Willy Susilo, and Jin Li. "Privacy-preserving Naive Bayes classifiers secure against the substitution-then-comparison attack." Information Sciences 444 (2018): 72-88. https://doi.org/10.1016/j.ins.2018.02.058

[24] Nayak, Deepak Ranjan, Ratnakar Dash, and Banshidhar Majhi. "Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests." Neurocomputing 177 (2016): 188-197. https://doi.org/10.1016/j.neucom.2015.11.034

[25] Wyner, Abraham J., Matthew Olson, Justin Bleich, and David Mease. "Explaining the success of adaboost and random forests as interpolating classifiers." The Journal of Machine Learning Research 18, no. 1 (2017): 1558-1590.

[26] Zou, Difan, Yuan Cao, Dongruo Zhou, and Quanquan Gu. "Stochastic gradient descent optimizes over-parameterized deep relu networks." arXiv preprint arXiv:1811.08888 (2018).

[27] Ge, Zhiqiang, Zhihuan Song, Steven X. Ding, and Biao Huang. "Data mining and analytics in the process industry: The role of machine learning." Ieee Access 5 (2017): 20590-20616. https://doi.org/10.1109/access.2017.2756872

[28] Ye, Yanfang, Tao Li, Donald Adjeroh, and S. Sitharama Iyengar. "A survey on malware detection using data mining techniques." ACM Computing Surveys (CSUR) 50, no. 3 (2017): 1-40. https://doi.org/10.1145/3073559

[29] Rao, K. Sreenivasa, N. Swapna, and P. Praveen Kumar. "Educational data mining for student placement prediction using machine learning algorithms." Int. J. Eng. Technol. Sci 7, no. 1.2 (2018): 43-46. https://doi.org/10.14419/ijet.v7i1.2.8988

[30] Chauhan, Ritu, and Harleen Kaur. "Predictive analytics and data mining: a framework for optimizing decisions with R tool." In Business Intelligence: Concepts, Methodologies, Tools, and Applications, pp. 359-374. IGI Global, 2016. https://doi.org/10.4018/978-1-4666-9562-7.ch019

[31] Lecavalier, Luc, James Bodfish, Clare Harrop, Allison Whitten, Desiree Jones, Jill Pritchett, Richard Faldowski, and Brian Boyd. "Development of the Behavioral Inflexibility Scale for Children with Autism Spectrum Disorder and Other Developmental Disabilities." Autism Research (2020). https://doi.org/10.1002/aur.2257

[32] Liang, Shuaibing, Chu Kiong Loo, and Aznul Qalid Md Sabri. "Autism Spectrum Disorder Classification in Videos: A Hybrid of Temporal Coherency Deep Networks and Self-organizing Dual Memory Approach." In Information Science and Applications, pp. 421-430. Springer, Singapore, 2020. https://doi.org/10.1007/978-981-15-1465-4_42

[33] Nimmo-Smith, Victoria, Hein Heuvelman, Christina Dalman, Michael Lundberg, Selma Idring, Peter Carpenter, Cecilia Magnusson, and Dheeraj Rai. "Anxiety Disorders in Adults with Autism Spectrum Disorder: A Population-Based Study." Journal of Autism and Developmental Disorders 50, no. 1 (2020): 308-318. https://doi.org/10.1007/s10803-019-04234-3

[34] Stevens, Elizabeth, Dennis R. Dixon, Marlena N. Novack, Doreen Granpeesheh, Tristram Smith, and Erik Linstead. "Identification and analysis of behavioral phenotypes in autism spectrum disorder via unsupervised machine learning." International journal of medical informatics 129 (2019): 29-36. https://doi.org/10.1016/j.ijmedinf.2019.05.006

[35] Thabtah, Fadi. "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward." Informatics for Health and Social Care 44, no. 3 (2019): 278-297. https://doi.org/10.1080/17538157.2017.1399132

[36] Chaidi, Irene, and Athanasios Drigas. "Autism, Expression, and Understanding of Emotions: Literature Review." International Journal of Online and Biomedical Engineering (iJOE) 16, no. 02 (2020): 94-111. https://doi.org/10.3991/ijoe.v16i02.11991

[37] Anagnostopoulou, Panagiota, Vasiliki Alexandropoulou, Georgia Lorentzou, Andriana Lykothanasi, Polyxeni Ntaountaki, and Athanasios Drigas. "Artificial Intelligence in Autism Assessment." International Journal of Emerging Technologies in Learning (iJET) 15, no. 06 (2020): 95-107. https://doi.org/10.3991/ijet.v15i06.11231

[38] Mitsea, Eleni, Niki Lytra, Antigoni Akrivopoulou, and Athanasios Drigas. "Metacognition, Mindfulness and Robots for Autism Inclusion." https://doi.org/10.3991/ijes.v8i2.14213

8 Authors

Dr. R. Sujatha completed the Ph.D. degree at Vellore Institute of Technology, in 2017 in the area of data mining. She received her M.E. degree in computer science from Anna University in 2009 with university ninth rank and done Master of Financial Management from Pondicherry University in 2005. She received her B.E. degree in computer science from Madras University, in 2001. Has 15 years of teaching experience and has been serving as an associate professor in the School of Information Technology and Engineering at Vellore Institute of Technology, Vellore. Organized and attended several workshops and faculty development programs. She actively involves her in the growth of the institute by contributing to various committees at both academic and administrative levels. She gives technical talks in colleges for the symposium and various sessions. She acts as an advisory, editorial member, and technical committee member in conferences conducted in other educational institutions and in-house too. She has published a book titled software project management for college students. Also has published research articles and papers in reputed journals. She used to guide projects for undergraduate and postgraduate students. Currently guides doctoral students. Interested to explore different places and visit the same to know about the culture and people of various areas. She is interested in learning upcoming things and gets herself acquainted with the student's level. Her areas of research interest include Data Mining, Machine Learning, Software Engineering, Soft Computing, Big Data, Deep Learning, and Blockchain. r.sujatha@vit.ac.in

Dr. S. L Aarthy completed the Ph.D. degree at Vellore Institute of Technology, in 2018 in the area of medical image processing. She received her M.E. degree in computer science from Anna University in 2010. She received her B.E. degree in computer science from Anna University, in 2007. Has 10 years of teaching experience and has been Assistant Professor (Senior) in the School of Information Technology and Engineering at Vellore Institute of Technology, Vellore. Her research area includes Image processing, soft computing, and data mining. She has published a good number of journal papers in her research field. She is a life member of CSI and IEEE. She is also part of various school activity committees. aarthy.sl@vit.ac.in

Jyotir Moy Chatterjee is currently working as an Assistant Professor in the Department of Information Technology at Lord Buddha Education Foundation (Asia Pacific University of Innovation & Technology), Kathmandu, Nepal. Earlier he worked as an Assistant Professor in the Department of Computer Science Engineering at G. D. Rungta College of Engineering & Technology (Chhattisgarh Swami Vivekananda Technical University), Bhilai, India. He is serving as the Young Ambassador of Scientific Research Group of Egypt (SRGE) for 2020-2021. He has been selected as Top 1% of reviewers in Computer Science on Publons global reviewer database 2019 powered by Web of Science Group. He has received his M. Tech in Computer Science & Engineering from Kalinga Institute of Industrial Technology (KIIT), Bhubaneswar, Odisha in 2016, and B. Tech from Dr. MGR Educational & Research Institute, Maduravoyal, Chennai in 2013. His research interests include the internet of things, machine learning & blockchain technology. jyotirchatterjee@gmail.com

Dr Abdulellah A. Alaboudi, has completed his PhD in Computer Sciences from University of Staffordshire, UK. Currently he is working at Department of Computer Science, Shaqra University, Saudi Arabia as Assistant Professor. Postgraduate certification is on his credit from Staffordshire university, UK. He has vast experience as business process reengineer. An ample number of peer reviewed are in his credit. His research areas include, IoT, Cybersecurity, Software Engineering, Wireless Networks, and Machine learning. alaboudi@su.edu.sa

Dr Noor Zaman Jhanjhi is currently working as Associate Professor (School of Computer Science and Engineering, SCE) with Taylor's University Malaysia. He has great international exposure in academia, research, administration, and academic quality accreditation. He worked with ILMA University, and King Faisal University (KFU) for a decade. He has 20 years of teaching & administrative experience. He has an intensive background of academic quality accreditation in higher education besides scientific research activities, he had worked a decade for academic accreditation and earned ABET accreditation twice for three programs at CCSIT, King Faisal University, Saudi Arabia. He also worked for National Commission for Academic Accreditation and Assessment (NCAAA), Education Evaluation Commission Higher Education Sector (EECHES) formerly NCAAA Saudi Arabia, for institutional level accreditation. He also worked for the National Computing Education Accreditation Council (NCEAC). Dr Noor Zaman has awarded as top reviewer 1% globally by WoS/ISI (Publons) recently for the year 2019. He has edited/authored more than 13 research books with international reputed publishers, earned several research grants, and a great number of indexed research articles on his credit. He has supervised several postgraduate students, including master's and PhD. Dr Noor Zaman Jhanjhi is an Associate Editor of IEEE ACCESS, moderator of IEEE TechRxiv, Keynote speaker for several IEEE international conferences globally, External examiner/evaluator for PhD and masters for several universities, Guest editor of several reputed journals, member of the editorial board of several research journals, and active TPC member of reputed conferences around the globe. noorzaman.jhanjhi@taylors.edu.my

Article submitted 2020-10-31. Resubmitted 2020-11-27. Final acceptance 2020-11-28. Final version published as submitted by the authors.