IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 1, March 2024, pp. 807~816
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i1.pp807-816

A computational intelligent analysis of autism spectrum disorder using machine learning techniques

Murali Anand Mareeswaran[1,2], Kanchana Selvarajan[1]
[1]Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science and Technology, Chengalpattu, India
[2]Department of Computer Science and Applications, Faculty of Science and Humanities, SRM Institute of Science and Technology, Chennai, India

Article Info

Article history:

Received Jun 2, 2023
Revised Oct 1, 2023
Accepted Oct 21, 2023

Keywords:

Artificial intelligence
Autism spectrum disorder
Decision tree
Random forest
Support vector machine

ABSTRACT

Children between the ages of 12 and 24 months who have autism spectrum disorder (ASD) experience abnormalities in the brain that result in undesirable symptoms. Children with ASD struggle to comprehend what others are trying to say and or feel, and they experience extreme anxiety in social situations. Additionally, they have a hard time making friends and even living independently. The defective genes, which control the brain and govern how brain cells communicate with one another, are the primary cause of ASD because they alter brain function. Our primary goal is to assist therapists and parents of children with ASD in using current technologies, such as human intelligence and artificial intelligence, to treat ASD and assist those youngsters in obtaining better social interaction and societal integration. For the purpose of doing an early analysis of ASD, the data is divided into the following three categories: age, gender, and jaundice symptoms. The performance of machine learning algorithms can be influenced by a variety of factors, such as the size of the dataset and quality of the dataset, the choice of features, and the tuning of hyper-parameters. In this work, the support vector machine (SVM) yields 96% as the highest classification accuracy.

Corresponding Author:

Kanchana Selvarajan
Department of Computer Science, Faculty of Science and Humanities
SRM Institute of Science and Technology
Kattankulathur, Chennai, Tamil Nadu, 603203, India
Email: kanchans@srmist.edu.in

1. INTRODUCTION

In recent years, autism spectrum disorder (ASD) has increased drastically and happens in children as early as within 12 months to before they reach three; Boys are more likely to have it than girls, with a ratio of 4 to 1. Due to increasing technologies it can be detected during pregnancy using a prenatal ultrasound which can identify early signs of autism thereby reducing the risk of having children born with ASD. ASD can happen in pregnancy due to an advanced parental gap during conception, maternal obesity, diabetes or immune system disorders. It is more common among children born prematurely and with very low birth weight. The mental disorders (DSM-IV) is a set of official criteria established by the American Psychiatric Association. The primary aim of this manual is to create a system for categorizing disorders and establishing clear diagnostic criteria for the disorders that are included in it (American Psychiatric Association, 1994) [1]. If the children have any retardness or abnormalities within 3 years and develop at least any one of the following three areas are included: (a) social interaction, (b) social communication through language, and (c) Imaginative or symbolic play, then there is fair chance for them to have ASD. DSM-IV for ASD also has three standard behaviors such as (a) Non-verbal behaviors, which include the absence of eye contact, facial expressions, body postures, and gestures in social interactions, (b) Verbal behaviors, which include speech impairments, delayed or underdeveloped spoken language, the repetitive use of language and stereotyped language, and play that is immature for their age, (c) Physical behaviors such as repetitive manners (hand or finger twisting or fluttering) and preoccupation in mind.

Genetic and environmental factors mainly contribute to ASD [2], [3] because each individual has a unique set of strengths and challenges. As mentioned before, since there is no permanent cure early diagnosis helps to reduce the symptoms but it takes several years for the process. Researchers in a study [4] states that early diagnosis helps to improve learning, communications and social skills and ability to think thereby improving brain development. In order to discover ASD in children the criteria's mentioned previously play a major role and clinicians identify them better by interacting with them, observing their body language and changes in their behaviour. It can be identified correctly by a family doctor or paediatrician because the identification and diagnosis of ASD is complicated [5]. No single medical test can correctly identify it; only experienced and well-trained doctors have a clear knowledge in this highly specialized area. The predicted performance utilized semi-supervised learning (SSL) method on the dataset to demonstrate significant advancements to increase the classifier result in comparison to other methods [6]. Machine learning (ML) plays an important role in analyzing the dataset since the data consists of surveys for various age groups ranging from children to adults [7]. It helps to predict the likelihood of a person having autism, thereby enabling healthcare persons to prioritize their resources.

The first three years are crucial to identify ASD, in a study found that diagnosis in three to six years is possible and one in 150 children suffer from ASD. As per the study, males are impacted three to four times more frequently than females [8]. Additionally, some cases allow for diagnosis as early as the second year of a child's life, and such early diagnosis is based on their behavior, social communication abilities [9], [10] and overall functioning, irrespective of their spoken language abilities. The screening process begins from 18 months of age and repeats again from 24 months to 36 months of age. Giarelli *et al.* [11] also agrees the same fact that it is predominant in male than in females with a percentage of 81% males contributing for the sample and the DSM-IV diagnosis were followed. The data was obtained from schools, clinical health agencies and the patients who came to the clinicians for diagnosis. Girls without any impairment may be identified at a later age than boys [12]. To achieve ASD classification accuracy, the dataset includes 16 attributes 703 patients and non-patients. The experimentations were in the simulation environment and analyzed using waikato environment for knowledge analysis (WEKA) platform. Cross-validation techniques were employed to derive the prediction of the ASD status [13].

ML models are used to detect ASD at an early stage, and the dataset is available in various stages here for toddler dataset support vector machine (SVM) is used, for children Adaboost algorithm is used, Glmboost for adolescence and adaboost for adults is used [14]-[17]. ML methods give good predictions for ASD and can be used in the early stages for identification and diagnosis. In order to investigate the underlying structural and strategic foundations of ASD, researchers have employed deep learning techniques [18], utilizing 14 distinct models including convolutional neural network (CNN) and recurrent neural network (RNN) to analyze 1,000 magnetic resonance imaging (MRI) scan images [19]. This approach enables the identification of specific brain structures that are indicative of a complex psychiatric condition, while also streamlining and optimizing the diagnostic process for clinicians in terms of time and resources. Retrospective data from 359 patients with ASD were examined, and training and testing were conducted using them [20]. A genetic algorithm (GA) and SVM classification algorithms to analyze ribonucleic acid (RNA)-seq data of Anopheles Gambia mosquitoes is employed [21].

In current diagnosis, the autism diagnostic observation schedule (ADOS) is used by clinicians and task-based functional magnetic resonance imaging (fMRI) is used and the dataset is splitted into three with a total of 157 patients with 92 mild symptoms, 32 moderate and 33 severe symptoms [22]. The system described here effectively extracts accurate discriminant features by analysing brain functionality [23], enabling the identification of ASD cases with high precision. Machine learning (ML) techniques have been utilized to yield favourable results, and future research efforts may involve the application of deep learning and multi-model fusion methods to further improve accuracy. The class/ASD dataset [24], [25] has been employed, with subgroups based on age and verbal intelligence quotient (VIQ) used for analysis via the RF model. Although the obtained classification accuracy is relatively low, it can serve as a useful reference for clinical diagnosis and early detection of ASD [26]. Explored the classification of images using pre-trained CNN models, namely VGG16, VGG19, InceptionV3, and ResNet101 [27]. Explored the performance of the extreme learning machine (ELM) classifier and long short-term memory (LSTM) classifier. The wavelet kernel with db8 achieved the highest accuracy of 83.33% in the ELM classifier, outperforming other kernels.

ML plays an important role in analysing the dataset since the data consists of surveys for various age groups ranging from children to adults. Our primary goal is to assist therapists and parents of children with ASD in using current technologies, such as human intelligence and artificial intelligence, to treat ASD and assist those youngsters in obtaining better social interaction and societal integration. For the purpose of doing an early analysis of ASD, the data is divided into three categories: age, gender, and jaundice symptoms. The distribution of ASD was investigated with respect to gender. The analysis revealed a higher prevalence of ASD in males compared to females. Our classification results demonstrated that the SVM method outperformed other classification techniques, achieving the highest accuracy of 96% for ASD diagnosis. These findings suggest that SVM can effectively identify ASD in its early stages. The rest of the paper is organized as follows section 2 gives a brief detail of the materials and methods employed to predict ASD, section 3 discusses about the results and evaluation followed by the conclusion in section 4.

2. METHOD

The class/ASD dataset contains more positive samples and negligible or no negative samples thereby affecting the efficiency of the classifier. The efficiency depends both on positive and negative samples so random sampling is performed to balance the unbalanced data. After balancing the data the feature selection methods are applied to select the valuable features. The training set is then kept and processed by four different machine learning classifiers, which includes decision tree, random forest, support vector machine (SVM), and logistic regression, after principal component analysis (PCA) is done to further reduce the data by removing the noise and other irrelevant data. The working process of ASD is described in Figure 1.

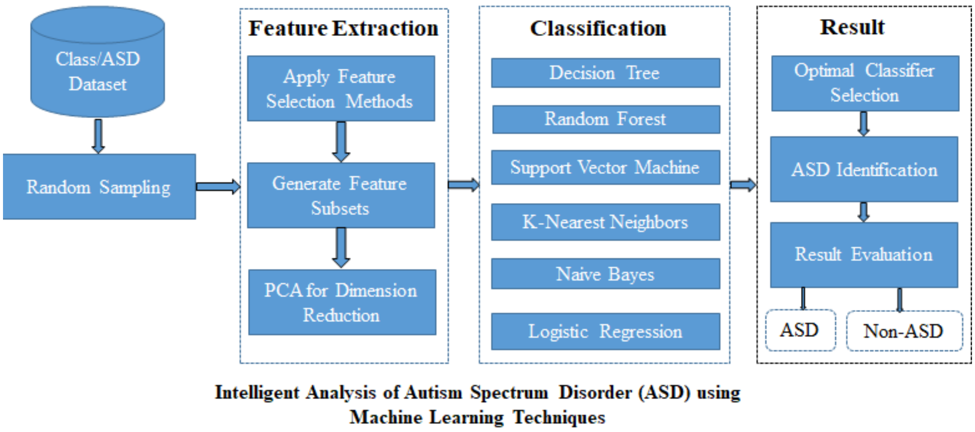

Figure 1. Working process of ASD

2.1. Dataset description

ASD happens in children at an early age, so diagnosis at an earlier age is necessary to improve the condition of children. Due to the complexity and heterogeneity involved in the diagnosis it is important to select the salient features leading to favourable outcomes. Large-scale samples are necessary in order to obtain favourable outcomes but collecting all the data from a single laboratory cannot obtain sufficient efficiency. The Class/ASD in response to the problem has collected brain image data containing functional and structural data from laboratories around the world to advance our knowledge on the neural bases of autism.

2.2. Work flow description

The initial stage in the workflow is to analyze the Class/ASD dataset for any missing values, repetition of data and negative values on the data and the main disadvantage that happens in the earlier step of processing is unbalancing of data, when the negative data is more compared to positive data the classifier can bias towards the side where the data is available in abundance, so balancing them becomes an essential and tedious task for the results to be accurate and favourable. Random sampling is performed to balance the unbalanced data by oversampling the under samples were balancing the data from each section of the data has higher probability to be selected in order to oversample, here, the chances of having sampling error is lower compared to non-probability sampling methods. The next phase is the feature extraction process where feature selection techniques are applied first to yield best results because not every feature contributes equally towards the efficiency i.e. in order to identify a disease in our case it is ASD we analyze all the features and select the most relevant features that contribute more towards it. The feature selection techniques generate three feature subsets such as age, gender and jaundice which influences ASD more compared to other features in the dataset. PCA is performed on these three feature subsets finally to reduce any noise present in them before performing the five machine learning classification algorithms. The three feature subsets that make up the categorization strategy are age, gender, and jaundice. Different classification techniques, including decision trees, random forests, support vector machine (SVM), K-nearest neighbors, Naive Bayes, and logistic regression, are used to classify these feature subsets.

The three subsets are processed further to predict the classification accuracy for ASD. In the gender subset, there are two categories namely male and female, here which of the two gender contributes more towards ASD is analyzed. The age category consists of a wide range of categories ranging from new born to senior citizens where the data is processed to find which age groups are more vulnerable towards ASD. The third feature includes jaundice which is also a major contributor towards ASD, here again the gender category comes into play. The best among them is identified using the classification accuracy where SVM outperforms the classification accuracy compared to other ML algorithms for all the three subsets. ASD identification is done using a classifier and the results are evaluated using SVM.

The evaluation set for further processing is decided and the training samples are obtained from random sampling, the SVM is utilized to classify the subsets effectively with an accuracy of 93% which is more compared to other machine learning classification algorithms. In ML analysis the data are analyzed, compared and ASD identification is done where the result consists of data evaluated using both testing and training sets. SVM helps to classify the data more accurately which is evident through the accuracy obtained and other classification algorithms show lower accuracy which proves the SVM is superior to other methods. The results are useful to family doctors or paediatricians which helps them to identify children at an earlier stage and diagnosis can be started at an earlier stage and the data obtained through this study can be useful in medical fields as well as to other doctors in this field and can even be stored in publicly available databases.

2.3. Classification methods

2.3.1. Logistic regression

The classification algorithm uses supervised learning techniques to predict newer class data based on training. Here, the classifier model is trained from a set of given observations in order to predict newer outcomes, so four different types of classification techniques have been used on age, gender and jaundice subsets to yield the best predictive outcome. Logistic regression is a statistical analysis method to predict a binary outcome, based on previous outcomes by comparing the relationship between one or more existing independent variables and is represented as,

$$f\big(E(x)\big) = \alpha + \beta\, y1 + \gamma\, y2 \tag{1}$$

where $f()$ represents thelink function, E(x) is the expectation of target variable and $\alpha + \beta y1 + \gamma y2$ is the linear predictor (α, β, γ to be predicted) [16]. The link function role is to 'link' the expectation of x to linear predictor.

2.3.2. Decision tree

The decision tree has a hierarchical tree structure with a root node, branches, internal nodes, and leaf nodes [25]. It also uses both classification and regression methods. Entropy is defined as the amount of information needed to accurately represent the data.

$$Entropy = -\sum_{i=1}^{n} pr_i * \log(pr_i) \tag{2}$$

$$Gini\ index = 1 - \sum_{i=1}^{n} pr_i^2 \tag{3}$$

pr_i is the probability that an arbitrary tuple belongs to class Ci

2.3.3. Random forest

The random forest is a meta-estimator that puts different decision tree classifiers into different sub-samples of the dataset and uses mean to improve the predictive accuracy and control over fitting. Random forest algorithms can also be used for both classification and regression problems because they are used by data scientists to predict stocks, e-commerce, medicine has many other applications in numerous fields as well. The importance for each feature in a random forest algorithm is calculated as (4):

$$fi_i = \frac{\sum j{:}node\ j\ splits\ on\ feature\ i\ ni_j}{\sum k{\in}all\ nodes\ ni_k} \tag{4}$$

where fi_i $represents$ the importance of feature i and ni_j $represents$ the importance of node j.

2.3.4. Support vector classifier

Support vector classifier or SVC is used for both classification as well as regression problems. SVC yields favourable outcomes for all the three subsets such as age, gender and jaundice. SVC generates the optimum decision boundary or line that can divide n-dimensional space into classes, allowing us to quickly classify data that comes in the future. A hyperplane is the name of this optimal decision boundary [17]. Mathematically, the hyperplane equation for classification, or dividing the points, is written as,

$$H: W^T(x) + b = 0 \tag{5}$$

Here, b is the Intercept and bias term of the hyperplane equation. The distance of a hyperplane equation $W^T\phi(x) + b = 0$ from a given point vector $\phi(x_0)$ as,

$$\mathrm{d_H}\big(\phi(\mathrm{x_0})\big) = \frac{|\mathrm{w^T}(\phi(\mathrm{x_0})+\mathrm{b})|}{\|\mathrm{w}\|_2} \tag{6}$$

$\|\mathrm{w}\|_2$ is the Euclidean norm for the length w and is given by,

$$\|w\|_2 := \sqrt{w_1^2 + w_2^2 + w_3^2 \dots \dots . w_n^2} \tag{7}$$

In order to maximize the minimum distance is given by (8) then,

$$w^* = arg_w \max\, [min_n \mathrm{d_H}\big(\phi(\mathrm{x_0})\big)] \tag{8}$$

If the point is substituted from the positive group in the hyperplane, then equation for value greater than 0 is,

$$\mathrm{w^T}(\phi(\mathrm{x_0}) + \mathrm{b}) > 0 \tag{9}$$

If the point is substituted from the negative group in the hyperplane, then equation for value less than 0 is,

$$\mathrm{w^T}(\phi(\mathrm{x_0}) + \mathrm{b}) < 0 \tag{10}$$

$$y_n[\mathrm{w^T}(\phi(\mathrm{x_0}) + \mathrm{b})] = \begin{cases} \geq 0\ if\ correct \\ < 0\ if\ incorrect \end{cases} \tag{11}$$

$$w^* = arg_w\ max\left[min_n \frac{|\mathrm{w^T}(\phi(\mathrm{x_0})+\mathrm{b})|}{\|\mathrm{w}\|_2}\right] = arg_w\ max\left[min_n \frac{y_n|\mathrm{w^T}(\phi(\mathrm{x_0})+\mathrm{b})|}{\|\mathrm{w}\|_2}\right] \tag{12}$$

$$w^* = arg_w\ max \frac{1}{\|w\|_2}[min_n y_n|\mathrm{w^T}(\phi(\mathrm{x_0}) + \mathrm{b})|] \tag{13}$$

$$w \to cw, b \to cb \tag{14}$$

$$(cw)^t\Phi(\mathrm{x_n}) + (cb) = c\big(\mathrm{w^T}(\phi(\mathrm{x_n}) + \mathrm{b})\big) \tag{15}$$

$$w^* = arg_w\ max \frac{1}{\|w\|_2}, \text{ s.t } min_n y_n[\mathrm{w^T}(\phi(\mathrm{x_n}) + \mathrm{b})] = 1 \tag{16}$$

3. RESULTS AND DISCUSSION

The class label for ASD has been displayed with respect to gender has more males affected with ASD compared to females. It represents the entire dataset taken into consideration for validation which consists of class labels 0 and 1. The class 1 in the graph represents those people who are affected by ASD i.e 161 and the class 0 represents the data for people who are not affected by ASD i.e. 639 here the negative data is more compared to positive class leading to biasing issues so it needs to be balanced in order to yield favorable and correct outcome. It displays the information of which gender people are most affected by ASD, the male count is more compared to female count consisting of 424 negative data and 106 positive data for males and females with a count of 215 for negative data and 55 for positive class. After balancing the unbalanced data and applying the data balancing techniques the results have been displayed in Figure 2, where the age has been considered from new born to 80 years of age for individuals, the graph shows more positive class in early stage from new born to until 40 years of age and then gradually decreases with increase in age.

The third feature subset of jaundice has been displayed in Figure 3 including the gender factor; it displays the class label for ASD with people suffering from jaundice. The graph also displays which gender people are affected more by ASD who have jaundice as well. Figure 4(a) represents the entire dataset taken into consideration for validation which consists of class labels yes and no. The class no in the graph represents those people who are not affected by ASD and the class yes represents the data for people who are affected by ASD. Figure 4(b) displays the information of which gender people are most affected by ASD, the male count is more compared to female count. The class labels for ASD with respect to jaundice are portrayed in Figure 5 where the positive data or positive classes are more compared to negative data. The training accuracy and the training loss for the training set are displayed in Figure 6 for ASD. The confusion matrix for the classification and consisting of scores from A1 score to A10 scores including age factors, result and Class/ASD and the values in the matrix ranges from 0 to 1 and higher values close to 1 are more significant and they in turn represents the data that are correctly classified. Figure 7 yields the accuracy for training and testing tests where the training set has more accuracy compared to testing test, in the testing phase the class labels will be removed and validated for accuracy in order to test the classifier accuracy.

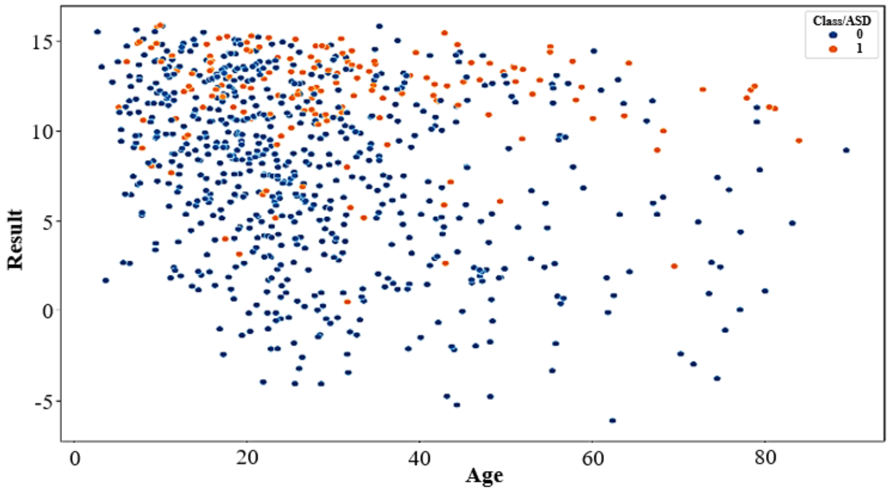

Figure 2. Classification of autism spectrum disorder with age

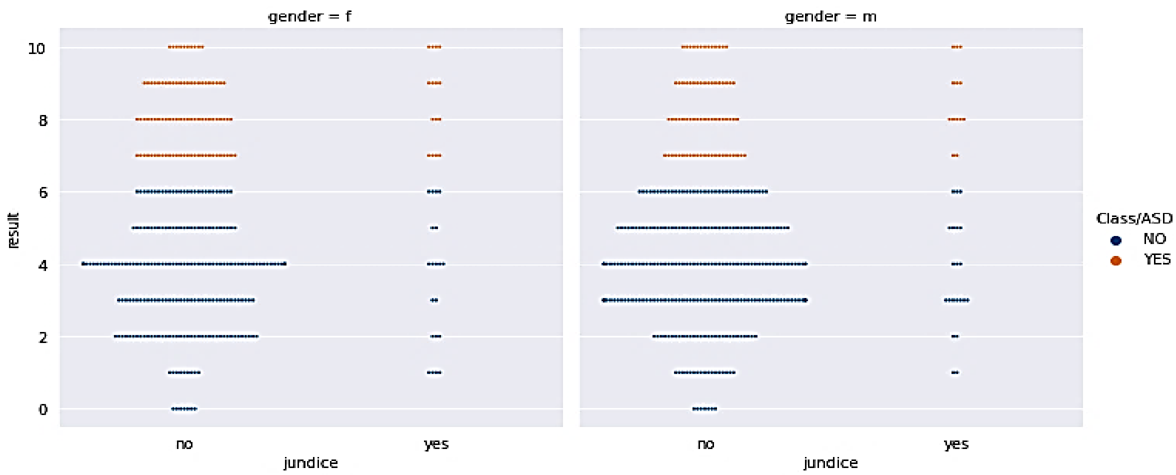

Figure 3. Class label for ASD with jaundice in male and female

(a)

(b)

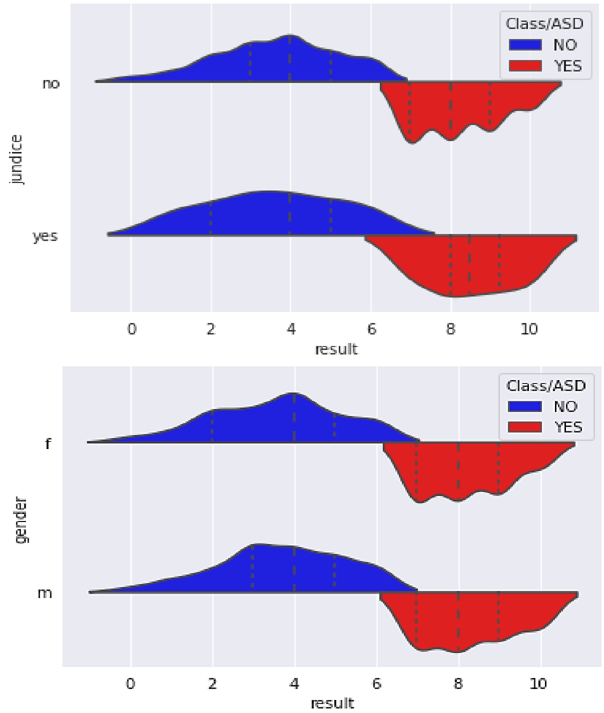

Figure 4. Class label for ASD with (a) jaundice and (b) jaundice gender

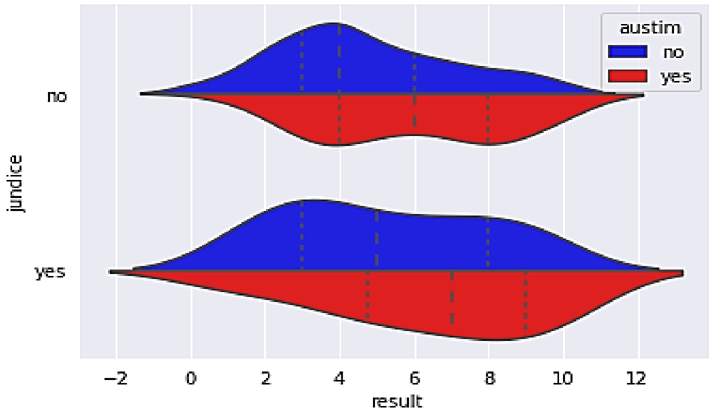

Figure 5. Class label for ASD with jaundice

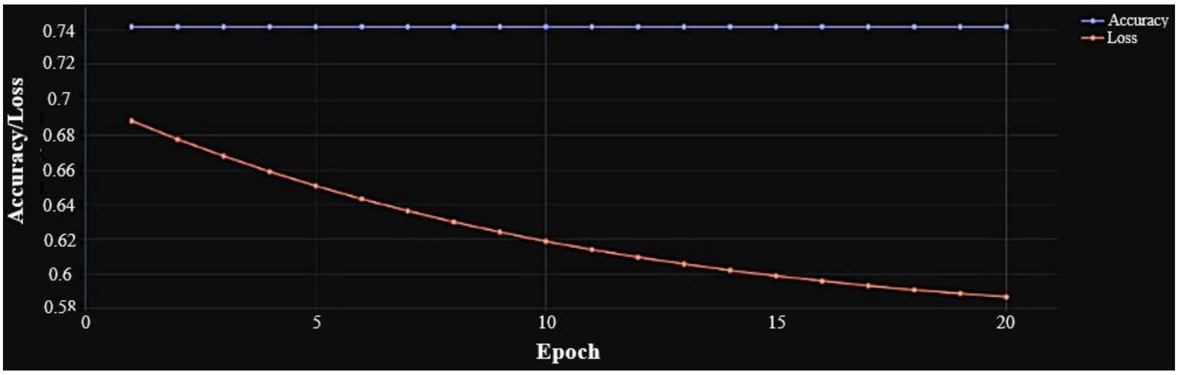

Figure 6. Training accuracy vs training loss

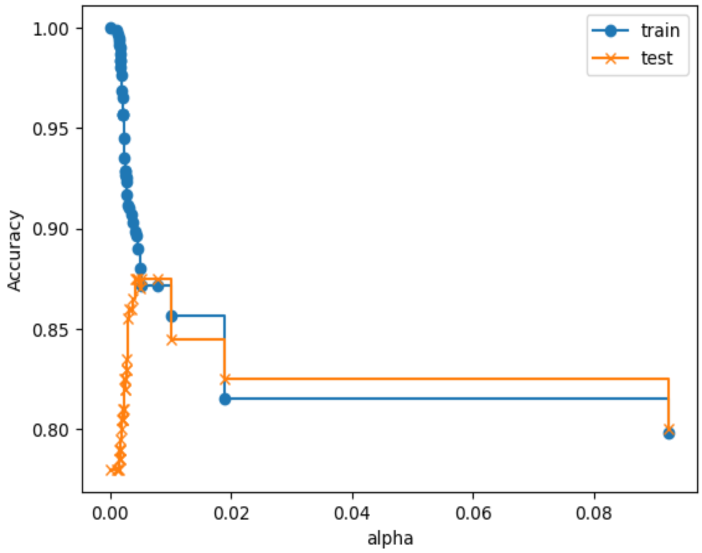

Figure 7. Accuracy vs alpha for training and testing sets

Finally, the accuracy measure in comparison with the four classification methods along with artificial neural network (ANN)-neural network mentioned previously has been displayed in Figure 8. The SVM method has the highest classification accuracy of 0.96 compared with other classification methods for ASD. The ASD labels have been correctly identified in SVM with a higher accuracy compared with other classification techniques and are able to predict ASD in early stages.

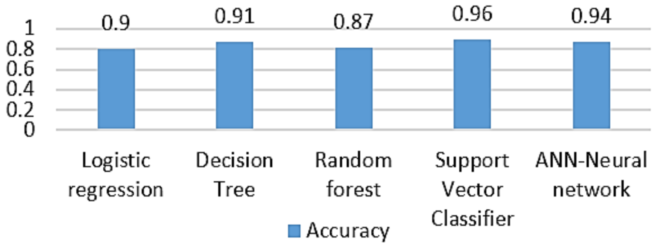

Figure 8. Accuracy comparison with machine learning algorithms

It is observed that the LSTM classifier achieved an accuracy of 83.78% in detecting capable dyslexic children. SVM proved to be the optimum classifier for distinguishing normal, poor dyslexic, and capable dyslexic children, achieving accuracy rates of 91.67% and 93.55% in detecting poor and capable dyslexic children, respectively [27]. Our classification results demonstrated that the SVM method outperformed other classification techniques, achieving the highest accuracy of 96% for ASD diagnosis. These results highlight the superiority of our proposed models over existing approaches. In our paper, we used a small data set to conduct our experiment. In the future, it is suggested to use a large-size data set to compare our classification models. The CNN technique can interpret brain biomarkers in ASD patients using fMRI [26].

4. CONCLUSION

In today's world, the occurrence of ASD has increased compared to previous years and early diagnosis is very effective for the individuals to lead a successful life in society. There is no complete cure for ASD so early diagnosis helps to improve the condition of individuals by continuous monitoring, changing the methods of training according to the severity of ASD and taking care of social and environmental changes. ML based prediction methods are becoming more welcomed since the probability of diagnosis is high compared to other methods used previously. The class/ASD dataset further helps to identify ASD accurately by considering the three main criteria. The gender, age and jaundice are the three most important characteristics that contribute more towards ASD. The gender category includes more males prone to getting ASD compared to females while in age category new-borns and children below five years of age are more vulnerable towards ASD. The jaundice feature along with gender as the sub feature helps to identify that child in young age along with jaundice and age factor contributes more towards ASD. The extracted features are evaluated using SVM based classification with accuracy of 96% and the severity is labelled for all data. SVM helps to classify the data more accurately which is evident through the accuracy obtained and other classification algorithms show lower accuracy which proves the SVM is superior to other methods. The results are useful to family doctors or paediatrician which helps them to identify children at an earlier stage and diagnosis can be started at an earlier stage and the data obtained through this study can be useful in medical fields as well as to other doctors in this field and can be stored in publicly available databases for references. Our proposed method yields best probability results and in the future deep learning techniques can be used to improve the accuracy further.

REFERENCES

[1] P. E. Tanguay, J. Robertson, and A. Derrick, "A dimensional classification of autism spectrum disorder by social communication domains," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 37, no. 3, pp. 271–277, 1998, doi: 10.1097/00004583-199803000-00011.

[2] G. B. Schaefer and N. J. Mendelsohn, "Genetics evaluation for the etiologic diagnosis of autism spectrum disorders," *Genetics in Medicine*, vol. 10, no. 1. pp. 4–12, Jan. 2008, doi: 10.1097/GIM.0b013e31815efdd7.

[3] J. D. Murdoch and M. W. State, "Recent developments in the genetics of autism spectrum disorders," *Current Opinion in Genetics and Development*, vol. 23, no. 3. pp. 310–315, Jun. 2013, doi: 10.1016/j.gde.2013.02.003.

[4] J. L. Matson, J. Wilkins, and M. González, "Early identification and diagnosis in autism spectrum disorders in young children and infants: How early is too early?," *Research in Autism Spectrum Disorders*, vol. 2, no. 1, pp. 75–84, Jan. 2008, doi: 10.1016/j.rasd.2007.03.002.

[5] P. A. Filipek *et al.*, *The screening and diagnosis of autistic spectrum disorders*. London: Routledge, 2001.

[6] S. Karim, N. Akter, and M. J. A. Patwary, "Predicting Autism Spectrum Disorder (ASD) meltdown using Fuzzy Semi-Supervised Learning with NNRW," in *2022 International Conference on Innovations in Science, Engineering and Technology, ICISET 2022*, 2022, pp. 367–372, doi: 10.1109/ICISET54810.2022.9775860.

[7] R. F. Eagle, R. G. Romanczyk, and M. F. Lenzenweger, "Classification of children with autism spectrum disorders: A finite mixture modeling approach to heterogeneity," *Research in Autism Spectrum Disorders*, vol. 4, no. 4, pp. 772–781, Oct. 2010, doi: 10.1016/j.rasd.2010.02.001.

[8] R. J. Landa, "Diagnosis of autism spectrum disorders in the first 3 years of life," *Nature Clinical Practice Neurology*, vol. 4, no. 3. pp. 138–147, Mar. 2008, doi: 10.1038/ncpneuro0731.

[9] G. Russell, C. Steer, and J. Golding, "Social and demographic factors that influence the diagnosis of autistic spectrum disorders," *Social Psychiatry and Psychiatric Epidemiology*, vol. 46, no. 12, pp. 1283–1293, Dec. 2011, doi: 10.1007/s00127-010-0294-z.

[10] M. Valicenti-Mcdermott, K. Hottinger, R. Seijo, and L. Shulman, "Age at diagnosis of autism spectrum disorders," *Journal of Pediatrics*, vol. 161, no. 3, pp. 554–556, Sep. 2012, doi: 10.1016/j.jpeds.2012.05.012.

[11] E. Giarelli *et al.*, "Sex differences in the evaluation and diagnosis of autism spectrum disorders among children," *Disability and Health Journal*, vol. 3, no. 2, pp. 107–116, Apr. 2010, doi: 10.1016/j.dhjo.2009.07.001.

[12] J. Shuster, A. Perry, J. Bebko, and M. E. Toplak, "Review of factor analytic studies examining symptoms of autism spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 44, no. 1, pp. 90–110, Jan. 2014, doi: 10.1007/s10803-013-1854-3.

[13] N. A. Mashudi, N. Ahmad, and N. M. Noor, "Classification of adult autistic spectrum disorder using machine learning approach," *IAES International Journal of Artificial Intelligence (IJAI)*, vol. 10, no. 3, pp. 743–751, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp743-751.

[14] T. Akter *et al.*, "Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders," *IEEE Access*, vol. 7, pp. 166509–166527, 2019, doi: 10.1109/ACCESS.2019.2952609.

[15] F. Ke, S. Choi, Y. H. Kang, K.-A. Cheon, and S. W. Lee, "Exploring the Structural and Strategic Bases of Autism Spectrum Disorders With Deep Learning," *IEEE Access*, vol. 8, pp. 153341–153352, 2020, doi: 10.1109/ACCESS.2020.3016734.

[16] M. F. Taraque *et al.*, "Early Stage Prediction of Autism Spectrum Disorder: Analyzing Different Hyperparameter Tuned Machine Learning Classifier," in *2022 4th International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*, Dec. 2022, pp. 1–4, doi: 10.1109/ICECTE57896.2022.10114518.

[17] S. Liang, A. Q. M. Sabri, F. Alnajjar, and C. K. Loo, "Autism Spectrum Self-Stimulatory Behaviors Classification Using Explainable Temporal Coherency Deep Features and SVM Classifier," *IEEE Access*, vol. 9, pp. 34264–34275, 2021, doi: 10.1109/ACCESS.2021.3061455.

[18] N. A. Ali, A. R. Syafeeza, A. S. Jaafar, and M. K. M. F. Alif, "Autism spectrum disorder classification on electroencephalogram signal using deep learning algorithm," *IAES International Journal of Artificial Intelligence (IJAI)*, vol. 9, no. 1, pp. 91–99, 2020, doi: 10.11591/ijai.v9.i1.pp91-99.

[19] A. S. Mohanty, P. Parida, and K. C. Patra, "ASD classification for children using deep neural network," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 461–466, Nov. 2021, doi: 10.1016/j.gltp.2021.08.042.

[20] J. Maharjan *et al.*, "Machine learning determination of applied behavioral analysis treatment plan type," *Brain Informatics*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40708-023-00186-8.

[21] O. A. Marion, O. A. Micheal, and O. Olugbara, "A genetic algorithm for prediction of rna-seq malaria vector gene expression data classification using svm kernels," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 2, pp. 1071–1079, 2021, doi: 10.11591/eei.v10i2.2769.

[22] O. A. Marion, O. A. Micheal, and O. Olugbara, "A genetic algorithm for prediction of rna-seq malaria vector gene expression data classification using svm kernels," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 2, pp. 1071–1079, 2021, doi: 10.11591/eei.v10i2.2769.

[23] E. Delbruck, M. Yang, A. Yassine, and E. D. Grossman, "Functional connectivity in ASD: Atypical pathways in brain networks supporting action observation and joint attention," *Brain Research*, vol. 1706, pp. 157–165, Mar. 2019, doi: 10.1016/j.brainres.2018.10.029.

[24] F. F. Thabtah, "Autistic Spectrum Disorder Screening Data for Children," *UCI Machine Learning Repository*, 2017.

[25] K. Kangra and J. Singh, "Comparative analysis of predictive machine learning algorithms for diabetes mellitus," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 3, pp. 1728–1737, Jun. 2023, doi: 10.11591/eei.v12i3.4412.

[26] S. A. Akbar, K. H. Ghazali, H. Hasan, Z. Mohamed, W. S. Aji, and A. Yudhana, "Rapid bacterial colony classification using deep learning," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 1, pp. 352–361, Apr. 2022, doi: 10.11591/ijeecs.v26.i1.pp352-361.

[27] A. Z. A. Zainuddin, W. Mansor, K. Y. Lee, and Z. Mahmoodin, "Machine learning and deep learning performance in classifying dyslexic children's electroencephalogram during writing," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 6, pp. 6614–6624, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6614-6624.

BIOGRAPHIES OF AUTHORS

Murali Anand Mareeswaran completed his M.Sc. in Loyola College, Chennai, India. He is currently working as an Assistant Professor in the Department of Computer Science and Applications, SRM Institute of Science and Technology, Vadapalani, Chennai, Tamil Nadu, 60026, India. At present, He is pursuing his Ph.D. in the Department of Computer Science under the guidance of Dr. S. Kanchana, Assistant Professor, Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, 603203, India. He can be contacted at email: muraliam@srmist.edu.in.

Kanchana Selvarajan obtained her Ph.D. degree by Bharathiyar University, Coimbatore. She has 15 years of teaching experience. She is currently working as an Assistant Professor in the Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, 603203, India. She has published more the 20 research papers in national and international journals. Her area of research is machine learning, data mining, block chain and IoT. Currently, she is guiding four Ph.D. scholars. She can be contacted at email: kanchans@srmist.edu.in.