

Comparison of Machine Learning Algorithms for Autism Spectrum Disorder Classification

Erina S. Dewi¹, Elly M. Imah^{1*}

¹Department of Mathematics, Universitas Negeri Surabaya, Indonesia

*Corresponding author: ellymatul@unesa.ac.id

ABSTRACT

Autism Spectrum Disorder is one of the fastest-growing neurodevelopmental disorders in the world. These neurodevelopmental disorders often attack children, affecting social development and behavior. Effective early detection of ASD is needed to reduce the risk of ASD in children. This study classified the ASD using KNN, SVM, Random Forest, Backpropagation, and Deep Learning. The Children ASD dataset consists of 292 subjects, 49 ASD subjects and 243 normal subjects. This adolescents ASD dataset consisting of 104 subjects with 63 ASD subjects and 41 normal subjects. The classification process is done by applying some parameters to each algorithm. The algorithms performance in classifying ASD dataset was compared. Based on the specificity and sensitivity value of the Random Forest algorithm with full features is the best algorithm compared to other algorithms in classifying ASD in children and adolescents.

Keywords: Autism Spectrum Disorder, KNN, SVM, Random Forest, Deep Learning.

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that can affect the development of social and communication behavior, stereotypical behavior, and interest in a matter [1]. Etiopathogenesis of ASD is still unknown with certainty, but several studies have argued that ASD is generally influenced by several biological factors such as genetic defects, inflammation of the brain, and abnormal conditions during pregnancy [2]. Besides that, genetic and environmental interactions such as autoimmunity can also be a factor causing ASD [3].

There are around 1.47% of the world's child population experiencing ASD [4]. Based on the National Health Interview Survey in 2014, ASD cases experienced a significant increase with an estimated prevalence of 1:45 and an increase in ASD cases was most prevalent in children [5]. The prevalence of ASD cases in children in the United States reaches 1:59, the prevalence is almost one-third of ASD cases that occur in children in the world [4]. According to the World Health Organization (WHO), an estimated 1 in 160 children in the world has ASD [6]. The prevalence of autism according to ASA (Autism Society of America) data in 2000 is 60 per 10.000 births, with a total of 1:250 populations [7].

Early detection of ASD in children needs to be done to ensure that children with ASD get appropriate support from the surrounding environment [8]. Besides, early detection of ASD can also speed up health services, to reduce the burden of ASD sufferers and their environment [9]. To get ASD detection results reliable can be done in the age range of 18 to 24 months [10]. Some methods that can be used in the detection of ASD in children can be through consideration of family medical history or with biomarkers in children [11].

Several studies have been conducted as an effort to detect early cases of ASD, especially in children. Screening ASD has been done by Charman, by comparing Social Communication Questionnaire (SCQ), Social Responsiveness Scale (SRS), and Children's Communication Checklist (CCC) methods. The SCQ method has a sensitivity of 0.86 and a specification of 0.76. While the SRS method has a sensitivity of 0.78 and a specification of 0.67. And the CCC method has specifications of 0.93 and specifications 0.46. Three screening methods can identify ASD in school-age children [12].

Kang's studies identifying cases of ASD in children using Support Vector Machine (SVM) based on EEG and eye-tracking data [13]. The classification results using EEG and eye-tracking data can be improved then

only used EEG or eye-tracking. Research on the classification of epilepsy and ASD using feature selection DWT and cross-correlation as well as methods ANN, KNN, SVM, and LDA classification based on EEG was conducted by Ibrahim [14]. The combination of feature selection DWT and Shannon Entropy method with the KNN classification method is the best in EEG classification of three classes, normal, epilepsy, and autism.

Other studies conducted by Hadoush in identifying the severity of ASD in children using empirical mode decomposition and second-order difference plot [15]. Data features of different IMFs, SODP plot patterns, elliptical areas, and CTM values indicate children with low and severe ASD levels. Based on the ANN classification model, the EMD value measurement results can be used as an automatic tool that can distinguish the severity of ASD children.

Some studies only screen ASD manually. Some studies have classified using machine learning but not many use the Random Forest and Deep Learning methods. Besides, there is still no ASD classification based on DSM-5 results comparing the performance of several methods of machine learning. So that in this study using the KNN, SVM, Random Forest, Deep Learning, and Backpropagation methods and compare the performance of these methods in the classification of ASD in children based on the results of DSM-5.

2. METHOD

2.1. Machine Learning Algorithms

2.1.1. K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is the simplest algorithm in the artificial learning algorithm [16]. To classify existing samples, the closest sample is chosen based on a certain distance. The new sample class is determined through the closest neighbor k class which often appears. The selection of k is done a priori with cross-validation or heuristics. In binary classification, k chosen is an odd number to avoid the similarity of the number of classes of the nearest closest neighbor [17]. To calculate the closest distance can be done with the distance Euclidean, City- block, and Chebychev.

$$D(x, p) = \begin{cases} \sqrt{(x-p)^2}, & \text{Euclidean} \\ |x-p|, & \text{City - block} \\ \text{Max}(|x-p|), & \text{Chebychev} \end{cases} \quad (1)$$

2.1.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) is one of the algorithms in machine learning that is often used in classifications in various fields [18]. SVM looks for an optimal hyperplane as a separator of two classes in the

process classification. An optimal hyperplane is obtained from optimal margins. Margin value is the distance between classes. Support vectors are the points in each class that are the closest to the hyperplane.

Suppose there are two classes in the dataset namely $z_i \in \mathbb{R}^n, n > 1, i = 1, 2, 3, \dots$ and $y_i = \{+1, -1\}$ by hyperplane $h(z) = \langle w, z \rangle + b$, the decision rule is defined as :

$$f(z) = \begin{cases} +1, h(z) \geq 1 \\ -1, h(z) < 1 \end{cases} \quad (2)$$

Then find the minimum value of:

$$\frac{1}{2} \|w\|^2 = \frac{1}{2} (w_1^2 + w_2^2) \quad (3)$$

to find the optimal hyperplane with the following conditions:

$$\langle w_1, z_1 \rangle + \langle w_2, z_1 \rangle + b \geq 1 \quad (4)$$

Next, to determine the data class, the decision function is calculated $\text{sign}(f(x))$

$$f(z_i) = \sum_{i=1}^m a_i y_i K(z_i, z) + y \quad (5)$$

where m is a lot of support vectors, a_i is the weight of each data, and $K(z_i, z)$ is a kernel function [19].

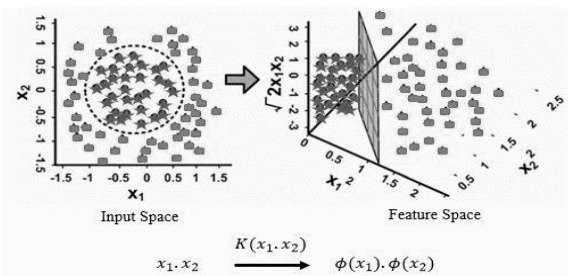


Figure 1 Non-linear SVM.

2.1.3. Random Forest

Random Forest is a combination of classification and regression trained on a dataset that has the same size as the training set and called bootstraps. Bootstraps were obtained from a random sample selection on the training set [20]. In the classification process, each tree represents one class, and the forest predicts which class has the votes most. The training algorithm in random forest uses aggregating bootstraps or bagging techniques. Suppose training set $X = x_1, \dots, x_n$ and response $Y = y_1, \dots, y_n$. Bagging repeated predictions sample select a random sample by replacing the training set and x' is obtained by the average prediction of all regression tree at x' is:

$$\hat{f} = \frac{1}{B} \sum_{b=1}^B f_b(x') \quad (6)$$

or with take the majority of classes on trees.

2.1.4. Deep Learning

Deep Learning is one method of machine learning that has a lot of attention [21]. Deep learning can predict models with high complexity and can be applied to various applications [22]. In this study, the instance iterator used is the default instance iterator with a simple ARFF file without spatial interpretation. The network configuration used is the neural network configuration and the iterator listener settings used are the epoch listener where this listener is used to evaluate the model when training for every N epochs. In this study, the CNN model used is CustomNet.

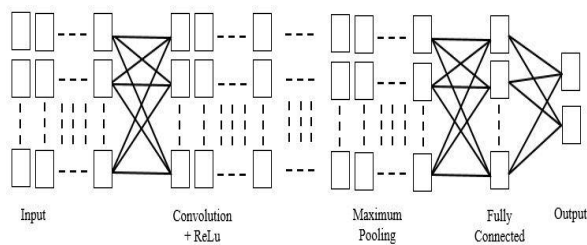


Figure 2 CNN architecture.

2.1.5. Backpropagation

Backpropagation is an algorithm that adds delta rules to Multi Layers Perceptron (MLP) as a backward phase [23]. In this study, Backpropagation is abbreviated as BP. Backpropagation is a systematic method for multilayer Artificial Neural Network (ANN) training because backpropagation has three layers in its training process, namely the input layer, hidden layer, and output layer. The backpropagation algorithm reduces the error rate by adjusting the weight based on differences in output and desired targets.

In this study, experiments were carried out using WEKA version 3.8.4. Waikato Environment for Knowledge Analysis (WEKA) is a computer program developed to identify information from raw data. WEKA develops machine learning and data mining by applying algorithms to data [24]. K-folds cross-validation is dividing the sample into k samples of the same size, each subsample is taken as validation data to test classification and repeated k times [25]. The inputted data is then cross-validated with 10-fold cross-validation. Data classification is done by four methods, namely K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest, Deep Learning, and Backpropagation. Some parameters are changed in each classification method to determine the effect of parameters on the results of data classification.

2.2. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)

Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition is a diagnostic and taxonomic

tool that is updated in 2013 and published by the American Psychiatric Association (APA). DSM is used as the main alloy for psychiatric diagnoses in the United States [1]. In DSM-5, there are concise and explicit criteria to facilitate an objective assessment of symptom presentations.

3. RESULTS AND DISCUSSION

In this study, there are two datasets, they are child and adolescent data. The dataset that is used in this study comes from the UCI Machine Learning. UCI Machine Learning provides 507 datasets that can be used by the machine learning community in developing research [26]. ASD and normal data were obtained based on observations and weighting of ten features of children and adolescent behavior with ten individual characteristics that were effective in detecting the ASD cases in children and adolescents from controls in behavior science. There are 21 attributes in the dataset. Child data are taken from 292 subjects consisted of 49 ASD subjects and 243 normal subjects. The age range of subjects is 4 to 11 years, with a total of 208 male subjects and 84 female subjects. Adolescents' data are taken from 104 subjects consisted of 63 ASD subjects and 41 normal subjects. The age range of subjects is 12 to 16 years, with a total of 50 male subjects and 54 female subjects [26][27].

The machine-learning algorithm that has been used for ASD classification are KNN, SVM, Random Forest, Deep Learning, and Backpropagation. All algorithm has been run with various parameters. The evaluation measure that is used in this classification is accuracy, precision, recall, ROC, specificity, and sensitivity. Accuracy in classification is the percentage of accuracy of records data that are classified correctly after testing the results of classification [28]. The accuracy value is formulated as (7). Besides accuracy, in classifying data need to consider the level of precision, recall, ROC, specificity, and sensitivity in the classification process. Precision is the level of accuracy between the desired information and the results of predictions made by the system formulated as (8).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (8)$$

The recall is the level of success of a system in classifying data correctly which is formulated as (9). Receiver Operating Characteristic (ROC) is a measure of the performance of a classification problem in determining the threshold of a model. The ROC curve is formed based on the True Positive Rate and the False Positive Rate. The higher the True Positive Rate and the smaller the False Positive Rate, resulting in a better threshold. Specificity is the level of accuracy in

predicting data in a negative class, formulated as (10). Sensitivity is the level of accuracy in predicting data in a positive class, formulated as (11)

$$\text{Recall} = \frac{TP}{TP+FN} \quad (9)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (10)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (11)$$

Based on Figure 3, the highest accuracy is achieved in the ASD classification process in children using the Random Forest algorithm with full features which is 1. Then followed by Backpropagation and SVM algorithm with C=5 and $\epsilon=10^{-9}$ at 0.9965. By using the Deep Learning algorithm with E=15 the accuracy is 0.9589 and the KNN algorithm with K=5 is 0.8973. The higher the accuracy value generated by the system in the classification process, the more precise the system is in predicting all data.

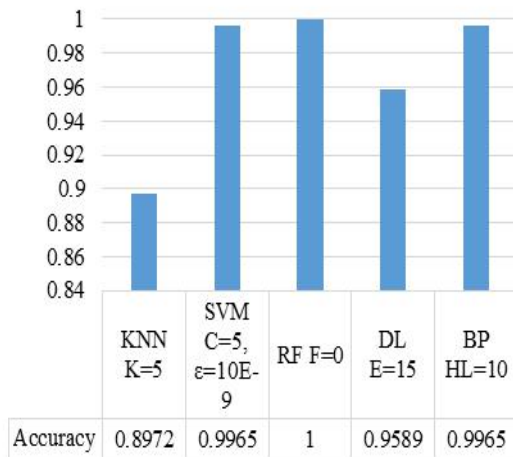


Figure 3 The highest accuracy of each algorithm in the classification process in children.

Based on Figure 4, the highest accuracy is achieved in the ASD classification process in adolescents using the Random Forest algorithm with full features which is 1. Then followed by the KNN algorithm with K=1 is 0.9038. By using Backpropagation and SVM algorithm with C=5 and $\epsilon=10^{-6}$ at 0.8942. And the Deep Learning algorithm with E=15 the accuracy is 0.8846. The higher the accuracy value generated by the system in the classification process, the more precise the system is in predicting all data.

In addition to using accuracy in data classification, a value is needed kappa statistical to strengthen the confidence of the truth of the classification results. Kappa Statistic is a statistic that has a range of values from 0 to 1 and is used to test interrater reliability. The higher the kappa statistics, the better the algorithm used [29]. In this study, the Random Forest algorithm has the highest statistical kappa value, which is 1.

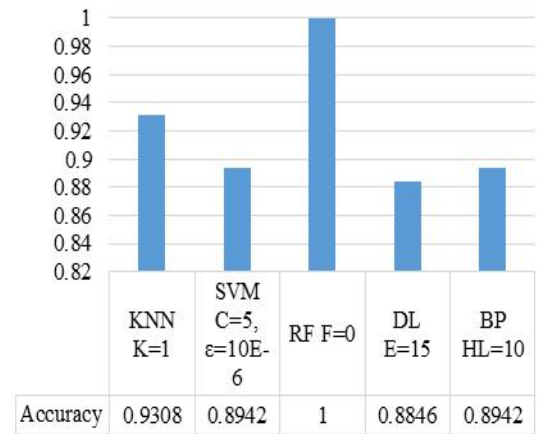


Figure 4 The highest accuracy of each algorithm in the classification process in adolescents

The study was conducted by applying various parameters in each algorithm. The best results classification from each algorithm is seen in Table 1. Based on Table 1, the Random Forest algorithm with full features has the highest precision value with a value of 1, this means that the algorithm Random Forest with full features has a high level of accuracy between the desired information and the predicted results provided by the system, then followed by other algorithms. Random Forest algorithm with full features also has the highest recall value which is 1 compared to other algorithms which means that the Random Forest algorithm with full features also has a very good level of accuracy in classifying data correctly. The threshold generated by the Random Forest algorithm with full features is also the best, this is indicated by the highest ROC value found in the Random Forest algorithm with full features so that it produces the best classification result compared to the Backpropagation, K, SVM, Deep Learning and KNN algorithms.

In predicting data on negative classes (ASD), the Random Forest algorithm with full features, SVM with C=5 and $\epsilon=10^{-9}$ and Backpropagation with 10 hidden layers has a very good level of accuracy with a value of specificity 1. But in predicting data in a positive class (normal), the Random Forest algorithm with full features has the highest level of accuracy with a value of sensitivity 1 followed by the Backpropagation, SVM algorithm, Deep Learning, and KNN. The greater the specificity and sensitivity of the algorithm, the level of accuracy in predicting data is also getting better.

In the ASD classification process in children and adolescents, each algorithm with various parameters is calculated training time and testing time. Training time is the time needed to build a model in the classification process. While testing time is the time needed to test data in the classification process. The best training time and testing time for each algorithm are shown in Figure 5 and Figure 6.

Table 1. The best precision, recall roc, specificity, and sensitivity from each algorithms

Method	Dataset	Class	Precision	Recall	ROC	Specificity	Sensitivity
KNN K=5	Child	No	0.976	0.821	0.973	0.9787	0.8211
		Yes	0.836	0.979	0.973		
		Avg	0.909	0.897	0.973		
SVM C=5 $\epsilon=10^{-9}$		No	1	0.993	0.997	1	0.9933
		Yes	0.993	1	0.993		
		Avg	0.997	0.997	0.995		
RF F=0		No	1	1	1	1	1
		Yes	1	1	1		
		Avg	1	1	1		
DL E=15		No	0.979	0.94	0.988	0.9787	0.9403
		Yes	0.939	0.979	0.947		
		Avg	0.96	0.959	0.968		
BP HL=10	No	1	0.993	1	1	0.9933	
	Yes	0.993	1	1			
	Avg	0.997	0.997	1			
KNN K=1	Adolescent	No	0.970	0.780	0.906	0.9841	0.7804
		Yes	0.873	0.984	0.906		
		Avg	0.911	0.904	0.906		
SVM C=5 $\epsilon=10^{-6}$		No	0.841	0.902	0.896	0.8888	0.9024
		Yes	0.933	0.889	0.896		
		Avg	0.897	0.894	0.896		
RF F=0		No	1	1	1	1	1
		Yes	1	1	1		
		Avg	1	1	1		
DL E=15		No	0.892	0.805	0.952	0.9365	0.8048
		Yes	0.881	0.937	0.952		
		Avg	0.885	0.885	0.952		
BP HL=10	No	0.875	0.854	0.981	0.9206	0.8535	
	Yes	0.906	0.921	0.981			
	Avg	0.894	0.894	0.981			

Based on Figure 5, the time needed for KNN, SVM, Random Forest, Deep Learning, and Backpropagation algorithms for training ASD data is quite short. The training time data on the Deep Learning algorithm is greater than other algorithms. This is because the deep

learning algorithm is a complex model with a large number of parameters that require more time in execution. While the lowest training time is on the KNN algorithm because KNN is one of the simplest methods in data classification. KNN algorithm is faster for

training because it does not require a learning process, there is no process of making a model. The existing dataset is only stored. When there is incoming test data, the class will look for the closest similarity to the test data.

Based on Figure 6, the time needed for the Deep Learning algorithm is greater than other algorithms. While the lowest testing time is on the Random Forest and SVM algorithm. The ASD Screening dataset for children has been used in research publications by Fadi Thabtah [27]. Their research publications explained that the classification of ASD by using machine learning but did not review about machine learning and its accuracy in the classification process. Thabtah only discusses the dataset and DSM-5 [27]. This study wants to complete Thabtah studies with the contribution to find the best machine learning algorithm for ASD classification.

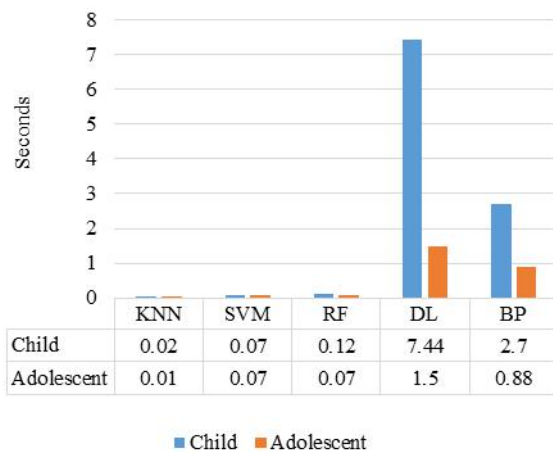


Figure 5 Training time in classification process.

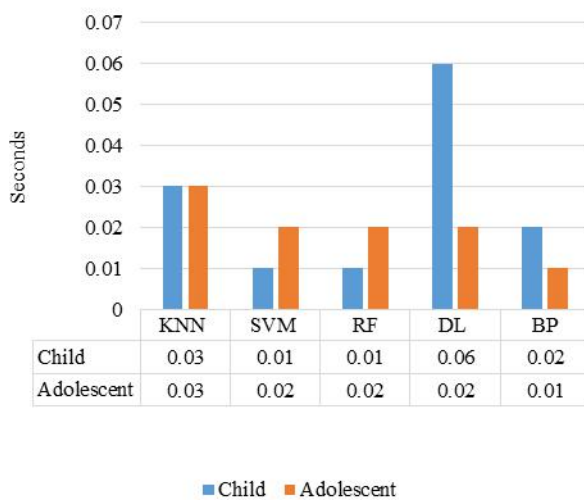


Figure 6 Testing time in classification process.

The previous study with the title a comparison of machine learning algorithms for surveillance of autism spectrum disorder compared the performance of the algorithms across 10 random train-test splits of the data, using classification accuracy, F₁ Score, and the number of positive calls to evaluate their potential use for surveillance [30]. This study wants to compare the performance of the algorithms in machine learning by their accuracy, kappa statistic, training, and testing time in the classification of ASD.

4. CONCLUSION

A previous study with the title ASD data classification in children and adolescents by machine learning but did not review machine learning along with accuracy, only reviewed the dataset and DSM-5. This study compares machine learning in ASD classification with a focus on the performance of the algorithms, with some evaluation measurements are the value of accuracy, kappa statistic, classification time, and other aspects. The best algorithm in ASD classification in children and adolescents is obtained in this study.

In this experiment, based on the accuracy, specification, and sensitivity values for ASD classification, the Random Forest algorithm has the highest performance than the other algorithms. The value of kappa statistics on the Random Forest algorithm is 1, this means that the results of the ASD classification using the Random Forest algorithm have a high level of true confidence. Thus, the Random Forest algorithm is the best in the classification ASD compared by other algorithms. KNN algorithm shows good training and testing time because this algorithm does not need time to build the model or learning phase, and this dataset is very simple so simple algorithm able to handle it. Deep Learning algorithm is no better than the Random Forest algorithm because the dataset used is simple namely 21 features. Deep Learning is intended for datasets with a very large number of features so that if used for data with simple features, it causes high computational costs. So, for this problem, simple algorithms are better performing.

ACKNOWLEDGMENTS

This work is supported by Grant of National Competitive on Applied Research with agreement No. B//11606/UN38.9/LK.04.00/2020 funding by the Ministry of Research, Technology, and Higher Education.

REFERENCES

[1] A. P. Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*-5th ed. Washington, D.C.: American Psychiatric Publishing, 2013.

- [2] H. V. Ratajczak, "Theoretical aspects of autism: causes-a review," *Journal of Immunotoxicol*, vol. 8, no. 1, pp. 68-79, 2011.
- [3] T. Zou, J. Liu, X. Zhang, H. Tang, Y. Song and X. Kong, "Autoantibody and autism spectrum disorder: a systematic review," *Research in Autism Spectrum Disorders*, vol. 75, pp. 1-12, 2020.
- [4] J. Baio, L. Wiggins and D., and L. Christensen, "Prevalence of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, United States, 2014," *MMWR Surveill Summ*, vol. 67, no. 6, pp. 1-23, 2018.
- [5] B. Zablotzky, L. I. Black, M. J. Maenner, L. A. Schieve and S. J. Blumberg, "Estimated prevalence of autism and other developmental disabilities following questionnaire changes in the 2014 national health interview survey," *Natl Health Stat Report*, vol. 87, pp. 1-20, 2015.
- [6] W. H. O. (WHO), "Autism spectrum disorders," *World Health Organization*, 7 November 2019. [Online]. Available: <https://www.who.int>. [Accessed 12 July 2020].
- [7] Kemenpppa, "Hari peduli autisme sedunia: kenali gejalanya, pahami keadaannya," 2 April 2018. [Online]. Available: <https://www.kemenpppa.go.id>. [Accessed: July. 24, 2020].
- [8] D. Metcalfe, K. McKenziea, K. McCartya, and G. Murrayb, "Screening tools for autism spectrum disorder, used with people with an intellectual disability: a systematic review," *Research in Autism Spectrum Disorders*, vol. 74, 2020.
- [9] D. L. Robins, "Screening for autism spectrum disorders in primary care settings," *International Journal of Research and Practice*, vol. 12, no. 5, pp. 537-556, 2008.
- [10] W. Guthrie, L. B Swineford, C. Nottke and A. M Wetherby, "Early diagnosis of autism spectrum disorder: stability and change in clinical diagnosis and symptom presentation," *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, vol. 54, no. 5, pp. 582-590, 2013.
- [11] C. Bridgemohan, D. M. Cochran and Y. J. Howe, "Investigating potential biomarkers in autism spectrum disorder," *Frontiers Integrative Neuroscience*, vol. 13, no. 31, 2019.
- [12] T. Charman, G. Baird, E. Simonoff, T. Loucas, S. Chandler, D. Meldrum and A. Pickles, "Efficacy of three screening instruments in the identification of autistic spectrum disorders," *British Journal of Psychiatry*, vol. 191, pp. 554-559, 2007.
- [13] J. Kang, X. Han, J. Song, Z. Niu and X. Li, "The identification of children with autism spectrum disorder by SVM approach on EEG and eye-tracking data," *Computers in Biology and Medicine*, vol. 120, pp. 1-5, 2020.
- [14] S. Ibrahim, R. Djemal and A. Alsuwailem, "Electroencephalography (EEG) Signal Processing for epilepsy and autism spectrum disorder diagnosis," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 1, pp. 16-26, 2018.
- [15] H. Hadoush, M. Alafeef and E. Abdulhay, "Automated identification for autism severity level: EEG analysis using empirical mode decomposition and second order difference plot," *Behavioural Brain Research*, vol. 362, pp. 240-248, 2019.
- [16] J. Diz, G. Marreiros and A. Freitas, "Applying data mining techniques to improve breast cancer diagnosis," *Journal of Medical Systems*, vol. 40, no. 9, p. 203, 2016.
- [17] W. Cherif, "Optimization of K-NN algorithm by clustering and reliability coefficients: application to breast-cancer diagnosis," *Procedia Computer Science*, vol. 127, pp. 293-299, 2018.
- [18] L. Demidova, E. Nikulchev and Y. Sokolova, "Big data classification using the SVM classifiers with the modified particle swarm optimization and the SVM ensembles," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 5, pp. 294-312, 2016.
- [19] L. Demidova, I. Klyueva, Y. Sokolova, N. Stepanov and N. Tyart, "Intellectual approaches to improvement of the classification decisions quality on the base of the SVM classifier," *Procedia Computer Science*, vol. 103, pp. 222-230, 2017.
- [20] S. Alessia, C. Antonio and Q. Aldo, "Random forest algorithm for the classification of neuroimaging data in alzheimer's disease: a systematic review," *Frontiers in Aging Neuroscience*, vol. 9, p. 329, 2017.
- [21] G. E Hinton, S. Osindero and Y.W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527-1554, 2006.
- [22] F. Emmert-Streib, Z. Yang, H. Feng, S. Tripathi and M. Dehmer, "An introductory review of deep learning for prediction models with big data," *Frontiers in Artificial Intelligence*, vol. 3, p. 4, 2020.
- [23] E. M. Imah and A. Widodo, "A comparative study of machine learning algorithms for epileptic seizure classification on EEG signals," in *International*

Conference on Advanced Computer Science and Information Systems (ICACSIS), Bali, 2017.

- [24] S. Laxmi, A. Kiranmai and A. Jaya, "Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy," *Protection and Control of Modern Power Systems*, vol. 3, no. 29, pp. 1-12, 2018.
- [25] S. Kotagiri Raju and M. R. Murty, "Support vector machine with K-fold cross validation model for software fault prediction," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 21B, pp. 321-334, 2018.
- [26] D. Dua and C. Graff, "UCI machine learning repository," *University of California, Irvine, School of Information and Computer Sciences*, 2019. [Online]. Available: <http://archive.ics.uci.edu/ml>. [Accessed: July. 08, 2020].
- [27] F. Thabtah, "Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment," in *Proceedings of the 1st International Conference on Medical and Health Informatics*, Taichung, Taiwan, 2017.
- [28] J. Han, M. Kamber and J. Pei, *Data Mining Concepts and Techniques*, 2nd Edition, San Fransisco: Elsevier, 2006.
- [29] P. Mittal and N. S. Gill, "A comparative analysis of classification techniques on medical data sets," *International Journal of Research in Engineering and Technology*, vol. 3, no. 6, pp. 454-460, 2014.
- [30] L. Scott H, M. Matthew J and H. Charles M, "A comparison of machine learning algorithms for the surveillance of autism spectrum disorder," *PLoS ONE*, vol. 14, no. 9, pp. 1-11, 2019.