Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

SRTTU JECEI

Research paper

MVO-Autism: An effective pre-treatment with High Performance for Improving Diagnosis of Autism Mellitus

***K. Ali Mohsin Alhameedawi*[1, 3]*, R. Asgarnezhad*[1, 2]**

[1]*Department of Computer Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.*

[2]*Department of Computer Engineering, Faculty of electrical and Computer Engineering, Technical and Vocation University (TVU), Tehran, Iran.*

[3]*Department of Computer Engineering, Al-Rafidain University of Baghdad, Baghdad, Iraq.*

Article Info

Article History:
Received 25 May 2021
Reviewed 01 July 2021
Revised 16 September 2021
Accepted 24 September 2021

Keywords:
Data mining
Pre-processing
Machine learning techniques
Autism mellitus

*Corresponding Author's Email Address:
razyehan@gmail.com

Abstract

Background and Objectives: Autism is the most well-known disease tha occurs in any age people. There is an increasing concern in appealing machin learning techniques to diagnose these incurable conditions. But, the poo quality of most datasets contains the production of efficient models for th forecast of autism. The lack of suitable pre-processing methods outline inaccurate and unstable results. For diagnosing the disease, the technique handled to improve the classification performance yielded better results, an other computerized technologies were applied.

Methods: An effective and high performance model was introduced to addres pre-processing problems such as missing values and outliers. Several base classifiers applied on a well-known autism data set in the classification stage Among many alternatives, we remarked that combine replacement with th mean and improvement selection with Random Forest and Decision Tre technologies provide our obtained highest results.

Results: The best-obtained accuracy, precision, recall, and F-Measure values o the MVO-Autism suggested model were the same, and equal 100% outperforms their counterparts.

Conclusion: The obtained results reveal that the suggested model can increas classification performance in terms of evaluation metrics. The results ar evidence that the MVO-Autism model outperforms its counterparts. Th reason is that this model overcomes both problems.

Introduction

Autism diagnosis was a challenging problem in all of the world. It is a pervasive neurodevelopmental disease with signs of lack of social interaction and interchange. Autism Spectrum Disorder (ASD) is a neurodevelopmental disease connected with high health care expenses. Regrettably, long-term diagnostic cause to methods is not cost-effective. The financial influence of this disease has revealed an urgent need to develop easy-to-implement and effective screening methods for autism spectrum datasets. However, these data sets are scarcely getting it hard to achieve comprehensive analyzes to increase the effectiveness, sensibility, etc. The effects of autism include a person's abilities in socialization, produces repeated behaviors, and influences expressionism or verbalism communicating with disorders that range from easy to strict [1]. Autism signs are most noticeable and clear to recognize in kids with two or three years old. In [2], authors show that 1 in 68 teenagers has autism. Then, the current authors worked on a dataset related to screening autism for adults and improving the classification of autism cases. The current authors also reprocessed this data set through the reprocessing method using missing values are replaced using mean and detect outlier with k-nearest neighbor (KNN) method.

The development of studies reveals that children with autism disease have a lowered motive to visit social incentives; also decreased awareness to presentations of people [3]-[5]. These investigations are oriented to promote the social motive theory of autism to reduce time wasted visiting people [4]-[7]. Newly, many studies purpose to improve machine learning systems for the diagnosis of autism. With growing of these studies, the impact of data property for saving vital information matures necessary.

The vital challenges for the autism dataset are of concern: the classification performance in terms of accuracy, to maximize the robustness of disease investigation, a lack of possible dataset, and speedy detection of autism [8]-[10]. Our offered work is an adequate model to address the first challenge.

In this article, we revealed that a proper pre-processing can determines a good prediction performance. For this prediction, the pre-processing techniques can enhance the data property and obtained more reliable results.

Many researchers investigated to predict autism disease newly. But a few studies used extensive preprocessing stages.

Hence, we applied a combination of missing value replacement and outlier detections. In the next paper, we use ensembles and heuristic methods in this context [11]-[13]. Here, the current authors studied the impact of preprocessing stages through machine learning techniques on autism disease in an iterative manner. In the first stage, missing values are identified and replaced through the mean. Then, we detected outliers through the KNN method. In the second step, we apply six algorithms for the classification stage to improve the classification performance. Our model classified performing supervised learning approaches such as a binary class Support Vector Machines (SVM), Feed-Forward Neural Network (NN), Random Forest (RF), Naïve Bayes (NB,) and Decision Tree (DT) [14].

The Innovations of this research are of concern:

- Proposing a method for addressing the missing values and outlier challenges.
- Offering a high-performance model for classification tasks using base methods on the autism dataset.

The paper is designed as following: After a review of the related works, the proposed model is shown in and estimated by the experiments. Eventually, the paper offers a conclusion.

Related Work

In this section, we review the most number of works for advancing the better performance of autism diagnosis. Most of these works concentrated principally on promoting classification performance through modifying the design of the classifier.

Here, the works that focus on discussions and answer conversations of autism patients are of concern:

In 2020, authors [15] designed an intelligent agent to communicate with children with autism in the face of violent extremism through generating vital characteristics. The obtained results showed two points: 1) having self-report conclusions, and 2) produce significant characteristics to estimate interaction abilities as evidenced through the good enough accuracy of these areas. The authors observed that ICON2 has fit to begin discussions and answer conversation associates in the first work. ICON2 produced 82.93% initiation and 89.33% proper responses. Their result is similar to other operators with conversational skills intended for TD personnel [16]-[18]. The accuracy values of SVM obtained 67.47% and 69.10% for Experiments I and II, respectively. In 2017, author [19] proposed a model to consistently identify the actions of a child with autism disease presenting repeated action through audio/video sensors. As it showed results as shown in the table below, Table 1. This model is applied to recognize hand/arm movements. The author's investigation revealed that the repeated models described signs of tension or emotion. They demonstrated that both the wrist-based accelerometer and the sound sensor provide self-stimulating ways [20].

Table 1: The obtained results for both the wrist-based accelerometer sensor and the sound sensor in 2017

Exercises	S1	S2	S3	S4
The first activities, Flap	96%	–	–	91%
The second activities, Drum	–	95%	–	–
The thirds activities, Punch	–	–	84%	–

Note: S=Self-stimulatory pattern

Here, the next content in related work is about works in ML area are of concern:

In 2015, the authors [21] presented a study to show a relationship between therapy power and time for supervision. They found age and gender to master training results specific to children's ABA systems on the spectrum. The methods used here are incorporated into the clinical.

The preparation also indicates the influence of ML and AI in this area. They randomly divided their data as follows: 65% for training and 30% for testing, and 5% for validation. They used a backpropagation network with Bayesian adjustment [22].

Table 2 summarizes a comparison among our proposed model and its counterparts.

Table 2: The comparison among the existing works and our model on the autism datasets

Works	Method	P	R	A
Jiao et al. 2009	Logistic model tree	-	95	87
Zhang et al. 2020	Intelligent agent, Game	-	-	89.68
Maenner et al. 2016	RF	-	84.5	86.5
Liu et al. 2016	ML	-	93.10	88.51
Altay and Ulas 2018	LDA	86.96	95.24	90.80
Altay and Ulas 2018	KNN	86.96	81	88.51
Bi et al 2018	Random SVM Cluster	-	-	96.15
Anirudh and J. Thiagarajan 2018	Bootstrapping G-CNN	-	-	70.86
Yulita et al. 2019	Regression	79.1	79.6	79.6
Yulita et al. 2019	Random Committee	87.5	87.7	87.7
Kazeminejad and sotero 2019	SVM	-	-	95
Cheng et al. 2019	SVM	-	-	92.70
Kazemine jad and Sotero 2019	SVM	-	-	95
Kong et al 2019	DNN			97.38
Raj and Masood 2019	NB, SVM, LR, KNN, NN, CNN	-	-	99.53
Sherkatghanad et al 2020	CNN	-	-	70.22
MVO-Autism (proposed)	**Base classifiers**	**100**	**100**	**100**

Note: P=Precision, R=Recall, and A=Accuracy

In 2018, the authors [23] studied children aged four to 11 years diagnosed with the ASD classification method. The Linear Discriminant Analysis (LDA) in conjunction with the KNN algorithm is used in the classification stage. The authors reveal that the LDA provided a more favorable result than the KNN in terms of accuracy value. They achieved a value of F-measure 90.91% for the LDA and 89.13% for the KNN, respectively. Also, for LDA, a value of sensitivity 95.24%, a value of specificity as 86.67%, a value of precision of 86.96% were obtained, respectively. The value of the F-measure was 90.91%. For LDA and KNN, the accuracy values were 90.80% and 88.51%, as well. While the sensitivity value was 97.62%, the specificity value was 80%, the precision value was 82%, and the F-Measure value was 89.13%. The authors [23] studied several tools for distinguishing autism. In [24], the authors applied the classification method. Wenbo Liu et al. suggested a framework for the recognition of autism through these methods [25]. In [26], the Authors were studied by applying four ML methods like SVM, multilayer perceptron, and logistic trees. Also, SVM was applied for ASD investigation [27]. The obtained sensitivity and specificity values were 88% and 86%, respectively. The authors [28] suggested an Application that could be categorized based on ML techniques. The Regression and Stochastic Committee were applied as classifiers. The result revealed that the performance of the random committees was better than that regression. The performance of classification was higher than 85% in terms of evaluation metrics. Moreover, it performed the reconfiguration process to overwhelm the problems of class imbalance. It seems that the proposed model helped improve the performance of both classifiers. With solving the imbalance problem, the accuracy value was improved from 64.6% to 79.6% for regression and from 66.3% to 87.7% for the random committee. With sampling, the random committee obtained an improvement of about 8.1% for the performance of classification through regression. Also, the authors in [29] proposed a classification system using graph-based approach to diagnosis the ASD problem. They showed that graph neural networks can improve the performance of their system. They proposed a bootstrapping version of graph CNN (G-CNN). The highest accuracy of the system was 70.86%. Due to the less classification accuracy of single SVM, the authors in 2018, applied multiple SVMs to classify ASD problem. Their results showed the highest accuracy of 96.15% using random SVM cluster [30].

In 2019, the authors applied a deep NN (DNN) classifier to handle ASD diagnosis using classification task. They extracted connectivity features which ranked using f-score measurement. Their obtained results showed that the highest accuracy was 97.38% [31]. The authors investigate the effect of ML techniques such as NB, SVM, LR, KNN, NN, and CNN to predict Autism diagnosis using three ASD datasets. In a pre-processing, the authors removed all null values or normalized them to handle the missing values problem for the used datasets. They achieved the highest accuracies equal to 99.53, 98.30, and 96.88% for the three used datasets [32]. Cheng et al. in 2019 proposed a new diagnosis system for the ASD problem. They used new features which these generate through SVM and the accuracy was 92.7% [33]. The authors applied ASD dataset to investigate the impacts of graph theoretical metrics through SVM. The best accuracy, sensitivity, and specificity were 95, 97, and 95%, respectively [34].

In 2020, the authors concentrated the automated detection of ASD through CNN. The suggested system classified the used dataset and an accuracy of 70.22% obtained [35].

As you have seen in the related work, missing value and outliers were effective in the data quality. Hence, we applied some methods to solve these problems. In this article, we used some methods for classification such as a combination of NN, KNN, DT, NB, SVM, and RF, which have proven successful and yielded the best results. After reviewing the related work, we affirmed that the preprocessing stage conjunction with a fit classifier is more reliable and points to higher results.

Data Collection

Here, we explain two used datasets related to autism disease. The autism datasets of the UCI ML repository were used.

The first applied dataset consists of 104 instances and 21 features. All Attributes contain Nominal and Numerical values. This dataset is learned from UCI Machine Learning Database Repository.

The second applied dataset is a new dataset related to autism disease. It included 20 features to be utilized for additional analysis especially in determining substantial autistic traits and developing the classification of ASD cases. In this dataset, ten behavioral features plus ten characteristics were recorded.

Two explanations included in the autism datasets are as follows:

1. Many limited datasets on autism-linked are open and generative. Therefore, we are working on a different dataset associated with autism for adults that has been downloaded from the UCI Machine Learning Database Repository, which contains 20 features, particularly in identifying influencing autism and improving the classification performance of autism.

2. In the mentioned dataset, ten behavioral characteristics (AQ-10-Adolescent) in conjunction with ten individuals' features studies.

The MVO-Autism Model

The main weakness of the autism data set is the absence of quality. Besides, the data set yields a high volume of missing values and outliers. An extra problem is the need for a good classifier. To succeed in these circumstances, the proposed model, MVO-Autism, is of concern:

In the first stage of our model, we investigate assigning the best values to the missing values using a reprocessing method to replace the missing values using mean and detect outliers with KNN methods. The second solution defines the data reduction task. Subgroup selection techniques are used for the attribute optimization algorithm classification. The proposed model applied an SVM classifier for predicting autism. Also, we applied other methods such as RF, NN, NB, and DT. Experiments exhibit that the DT and the RF are two of the most reliable classifiers for diagnosing autism on the Autism dataset. It is why we choose them in this research. As for the rest of the methods used in this paper, NN and SVM, and NB methods also showed good enough results but less good than the two mentioned algorithms. Fig. 1 shows the stages of the proposed model.

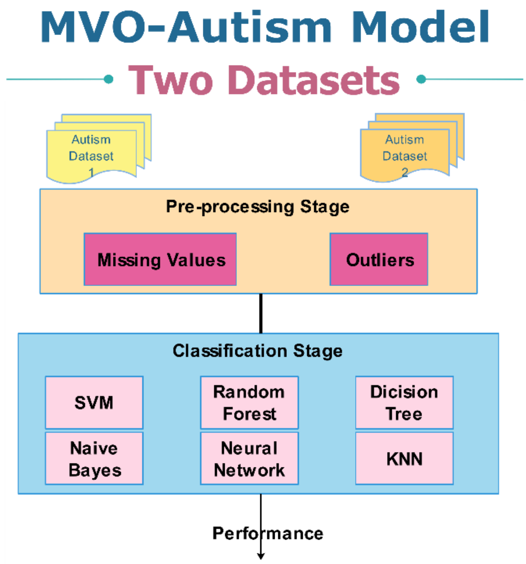


Fig. 1: The proposed MVO-Autism model.

Here, we describe the applied techniques in the proposed model in detail. Our model examines the impact of two techniques, i.e., mean and KNN, for missing values and outliers on classification performance. Moreover, our model applied six classifiers which are explained in the following. After cleaning, the algorithms were applied to decrease the repetition of features and obtained the related characteristics in the used dataset. After preprocessing and data modification methods, an SVM classifier was used to classify cases. Descriptions of the used techniques and our model in a staged manner are of concern:

A. Pre-processing Stage

Here, we used methods to Solve the Missing Data Problem by Replace the Missing value with the mean and detect the outlier through the KNN method. In the applied dataset, the "0" value is assigned to missing values. Two principle tasks in the preprocessing are of concern:

- Replace Missing Values through Mean

Here, missing values are substituted using the mean. Hence, all missing attribute values will be replaced with the mean of all values for the particular characteristic. In some of the cases, this value can be approximately nearby to the actual value, and thus we will treat the data loss with this command [36].

Note that other methods such as imputation, KNN, Average, and PCA were used, but the result was not improved.

- Detect Outlier through KNN

Here, our model detects the outliers by specifying the KNN. The *k* is assigned to the number of neighbors, and this value is denoted by a number. Also, *k* equals five or *k* equals ten, as appropriate for our work. KNN technique substitutes outliers with a similar value from the closest tuple through Euclidean distance [36].

B. Classification Stage

In the classification stage, the actual packages are analyzed, and the headers are extracted. With header values, the data structure created during the preprocessing stages crossed to find the best matching base. Since the classification stage works in the data path, the classification speed is important. We applied six algorithms to improve the classification performance, including DT, NB, RF, NN, SVM, and KNN classifiers. Here, we explain the classifiers in detail.

- SVM

The SVM is one of the popular base classifiers in ML techniques. In SVM classification, the separation purpose is formulated as a linear collection of kernels correlated with support vectors [37]-[27]. As you see, Fig. 2 presents what was indicated.

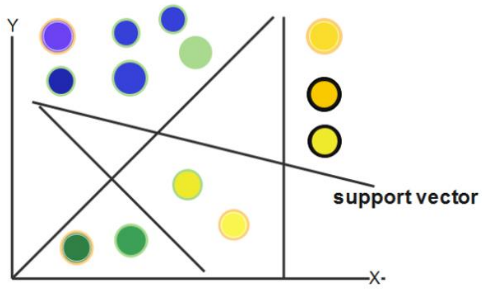


Fig. 2: The steps of the SVM working

- RF

RFs are considered to be the best algorithm applied to this data. The obtained accuracy of it was 100%. Regression and other functions that work by producing multiple decision trees at training and outputting a layered class or average/average prediction for individual trees [38]-[39]. Random decision forests are true of the habit of DTs of over-allocation of their training set. Fig. 3 shows a schematic diagram of a random set of decisions.

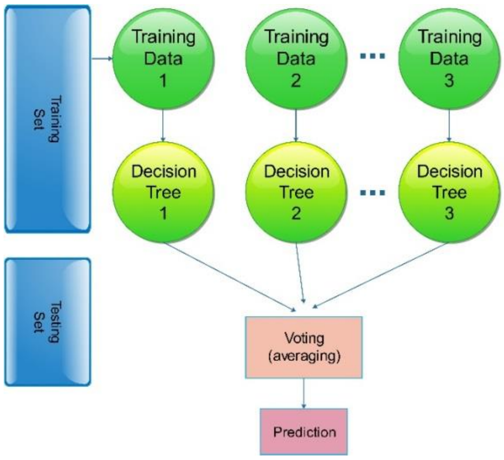


Fig. 3: The training and testing phases in conjunction with voting in the RF classifier.

- NN

NN. is a method that achieved significant results in a diversity of areas associated with pattern recognition, etc. The limitation of parameters in Artificial Neural Networks (ANNs) is the most advantageous regard for Convolutional Neural Networks (CNNs). This performance advised authors to move closer to resolve complicated tasks, which were not otherwise attainable with classic ANNs. A NN is a network consisting of neurons. The synthetic or contract data was used to improve the classification performance, and it showed satisfactory results [36].

- DT

DT showed very high results and was classified as one of the best techniques used in this paper, with accuracy reaching 100%. DTs are simple yet effective classification algorithms. One of its main advantages is that it provides human-readable classification rules. DTs have several drawbacks, one of which is the need to sort all numerical features to determine where to split a node. This becomes expensive in terms of runtime and memory size, especially when DT is trained on big data [36].

- NB

Due to simplicity in both the training and classification phases, the NB classifier has been widely used [10]. Several researchers have proven that it is effective enough to classify text in many areas [40]. The NB models recognize each characteristic to provide decision-making on an equal footing independently. This theorem presents a method of computing the following possibility, $P(c/x)$, from $P(c)$, $P(x)$, and $P(x/c)$. NB implies that the impact of predictor (x) on an addressed class (c) is autonomous of other predictors. This theory is named class qualified autonomous, (1).

$$P(c\,|\,x)=\frac{P(x\,|\,c)\,P(c)}{P(x)} \tag{1}$$

$$P(x)=P(x1\,|\,c)\times P(x2\,|\,c)\times\ldots\times P(xn\,|\,c)\times P(c)$$

where P(c|x) is the following possibility of the target, P(c) is the previous possibility of the target, P(x|c) is the reasonableness which is the possibility of predictor assigned class, and P(x) is the previous possibility of predictor.

- KNN

The KNN classifier exhibited high performance on large data sizes for example, such as near infinity, its error rate is almost as large as Bayes improvement under very moderate conditions. The KNN rating performance can be affected by some issues like k-value choice, distance metrics selection, etc. Recently, several technologies have been developed to succeed the problems. A KNN method can consolidate a confidence portion scale into the traditional KNN [41]-[43]. See KNN Classification in Fig. 4.

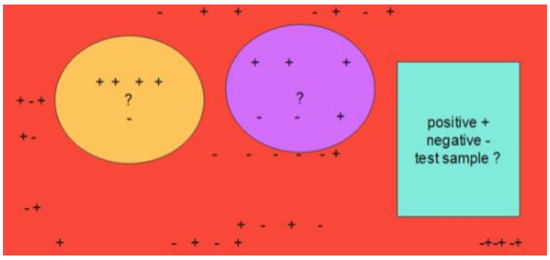


Fig. 4: A classification model through the KNN with k = 5.

Results and Discussion

For implementation, Rapid Miner 9th Software was employed to handle our experiments on the autism dataset. For building the proposed model, SVM, RF, KNN, NN, BN, and DT classifiers were applied. The experiments propose estimating the influence of our pre-processing. To evaluate the performance of classification, four important methods are used, namely accuracy, precision, recall, and measure. These formulas, (2), (3), (4), and (5) are of concern:

$$Accuracy=(TP+TN)\,/\,(P+N) \tag{2}$$

$$Precision\ (Pre)=(TP)\,/\,(TP+FP) \tag{3}$$

$$Recall\ (Rec)=TP\,/\,P \tag{4}$$

$$F-Measure=\frac{2\times Pre\times Rec}{Pre+Rec} \tag{5}$$

Note that TP=True positive, TN=True negative, FP=False positive, and FN=false negative.

Accuracy is a measurement of classification ends to the original values without regarding the class labels. Precision is a measurement of accuracy given that the subject class has been prophesied [40]. The recall is measured as the number of true positives arranged through the both number of true positives and false negatives. The F-Measure is a wide measurement that gives a method to join accuracy and recall into an individual scale to arrests both parts [36]. Table 3 and Table 4 present the comparative results in terms of evaluation metrics.

Table 3: The obtained results for solving missing values through mean and outlier through KNN in terms of precision and recall on dataset 1 and 2

Classifier	Positive Precision	Negative Precision	Positive Recall	Negative Recall
		Dataset 1		
SVM	86.36	**100**	**100**	75
KNN	95	**100**	**100**	91.67
RF	**100**	**100**	**100**	**100**
NB	**100**	92.31	94.74	**100**
NN	94.12	78.57	84.21	91.67
DT	**100**	**100**	**100**	**100**
		Dataset 2		
SVM	77.78	**100**	**100**	73.33
KNN	91.11	97.62	97.62	91.11
RF	**100**	**100**	**100**	**100**
NB	90.70	93.18	92.86	91.11
NN	86.96	95.12	95.24	86.67
DT	**100**	**100**	**100**	**100**

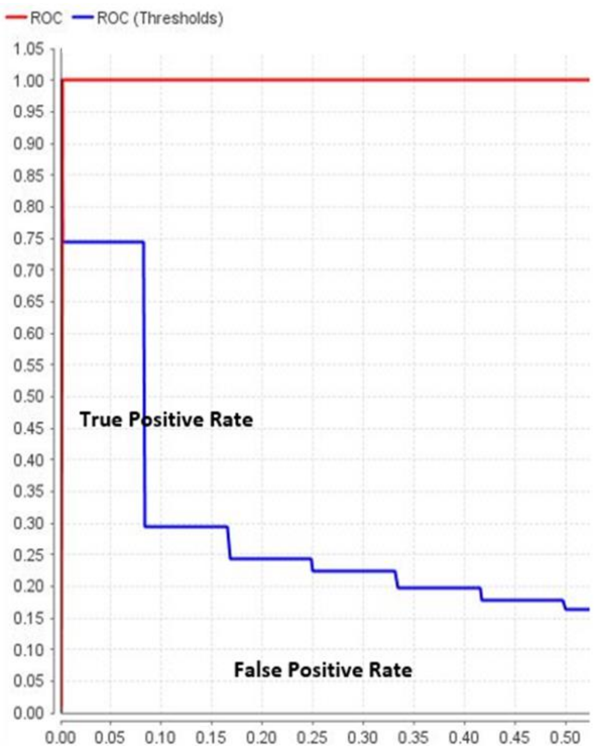
Table 4: The obtained results for solving missing values through mean and outlier through KNN in terms of accuracy and F-Measure on dataset 1 and 2

Classifier	Accuracy	F-Measure
	Dataset 1	
Support Vector Machine	90.32	90.25
K-nearest neighbor	95	**100**
Random Forest	**100**	**100**
Naïve Bayes	96.77	96.75
Neural Network	87.10	87.13
Decision Tree	**100**	**100**
	Dataset 2	
Support Vector Machine	86.21	**100**
K-nearest neighbor	94.25	97.62
Random Forest	**100**	**100**
Naïve Bayes	91.95	93.02
Neural Network	90.80	95.18
Decision Tree	**100**	**100**

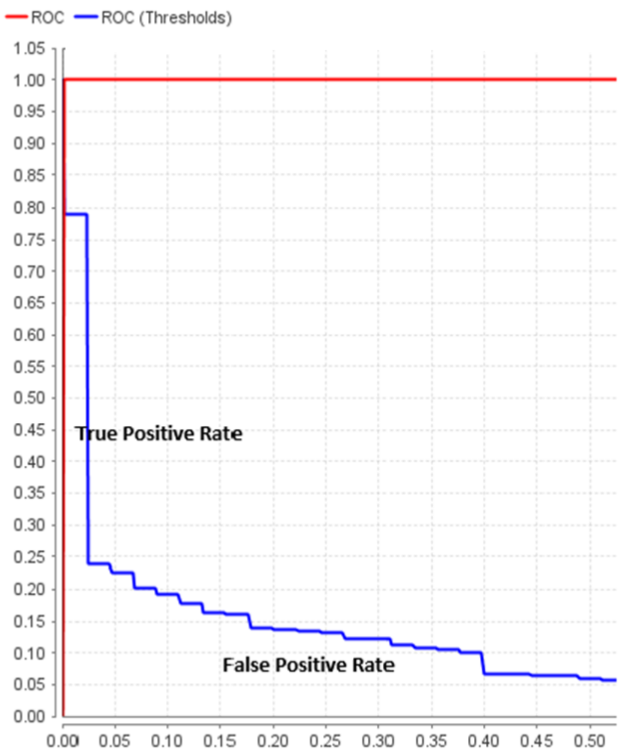
Table 3 presents the obtained results for solving missing values through mean, an outlier through KNN in terms of precision and recall. The obtained results revealed that the DT and RF are the best. When the pre-treatment stage is not performed, the performance of the classification is more moderate than that of cases which applied namely, DT, NN, and NP, RF, and SVM. As you can see, Table 3 exhibits the achieved results for resolving missing values and outliers in terms of precision and recall. Table 4 displays the obtained results for solving missing values and outliers in terms of accuracy and F-Measure. The achieved results confirmed that the DT algorithm and RF algorithm give the best results. The best results of the accuracy for the DT algorithm reached 100% and precision of 100%. Also, the RF algorithm gets accuracy 100% and precision of 100% on the autism dataset. These two algorithms in the two tables are considered the best results in improving classification performance and have shown successful results. Therefore, we used two methods to increase the results significantly. Fig. 5, Fig. 6, Fig. 7, Fig. 8, Fig. 9, and Fig. 10 show the obtained ROC curves for all classifiers.

Fig. 11 shows a comparison between our model and the mentioned works in this context. The results are evidence that the MVO-Autism model outperforms its counterparts. The reason is that this model overcomes both problems, that is, missing values and outliers. Finally, experiments showed that the utility of pre-treatment methods points to a significant improvement in the predictive model of autism diagnosis. Among all existing methods, our model showed that a compound of replacement with average is a method to replace the lost value with improvement. For performing classification, using techniques such as DT, RF, NB, NN, SVM, and KNN gives high results in terms of predictive accuracy. This finding suggests that the aforementioned group could improve autism diagnosis. However, there is no guarantee that the obtained results are also suitable for different datasets. Hence, different experiments need to be performed on other types of the dataset in this context. The decision tree and random forest algorithm showed the best results in this paper.

According to Figs. 5-10, we found that the best AUC of our model belongs to tree methods like DT and RF.

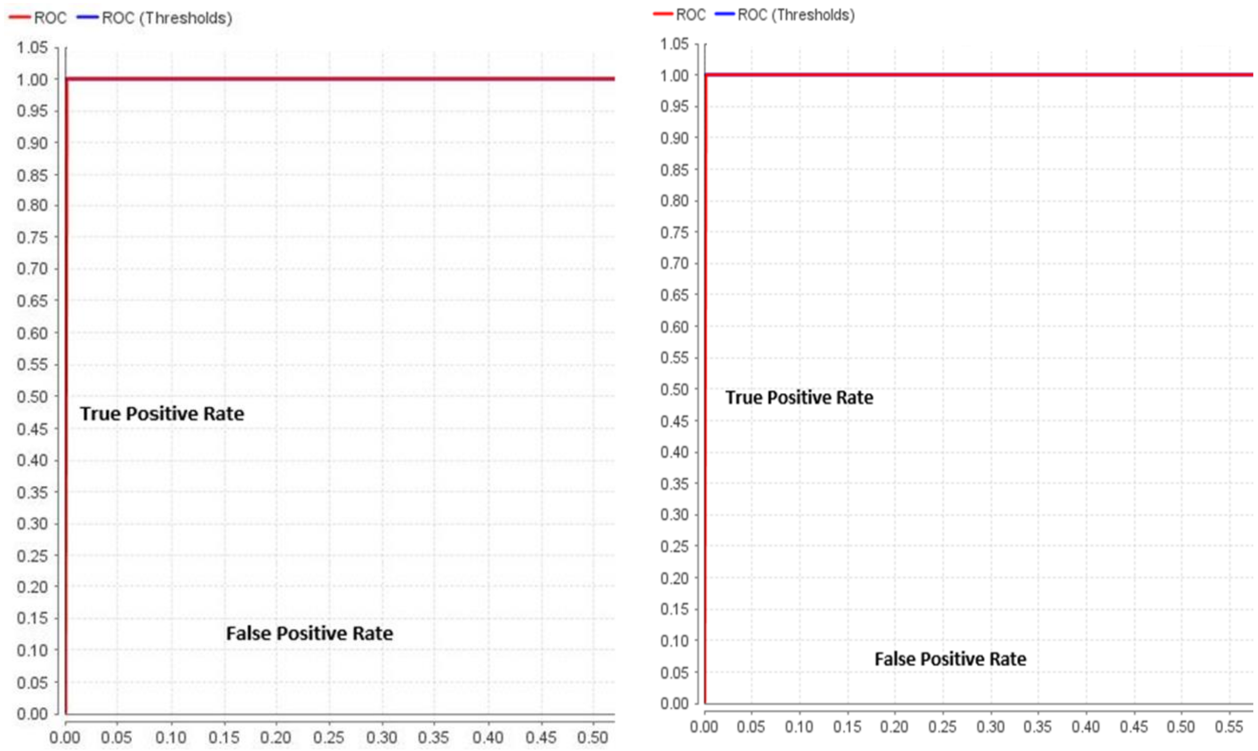


on dataset-1



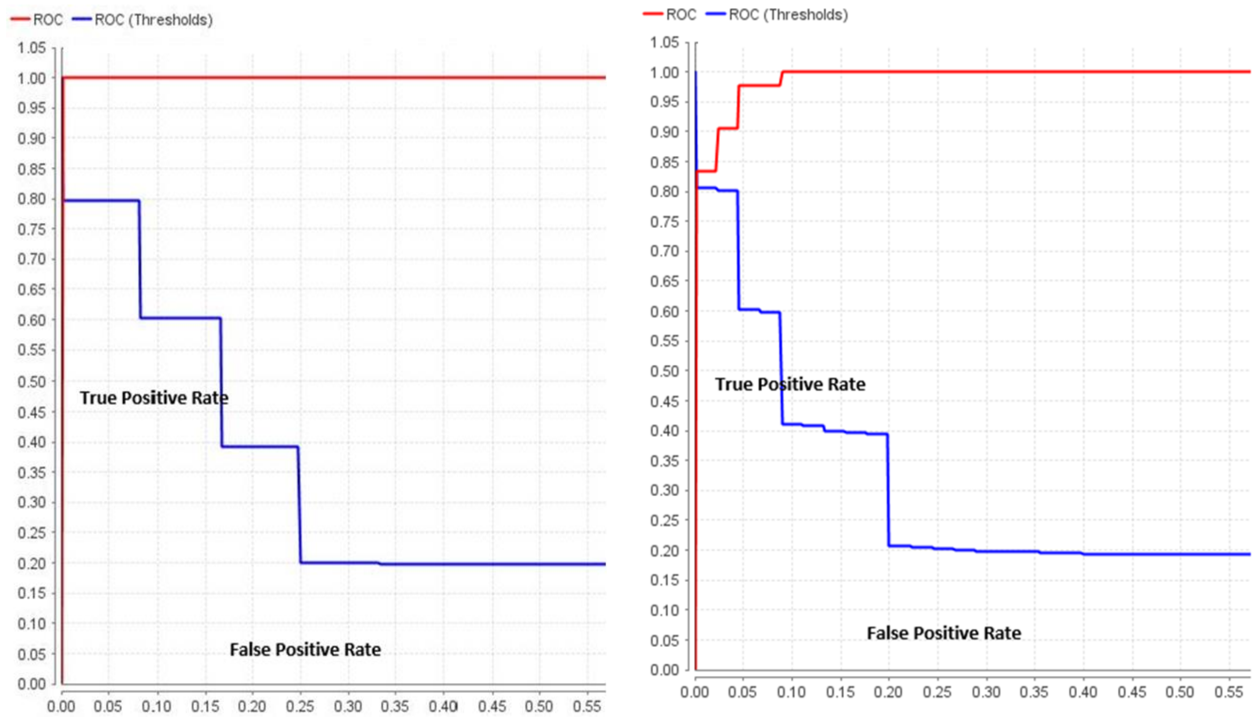
on dataset-2

Fig. 5: The ROC for the RF.



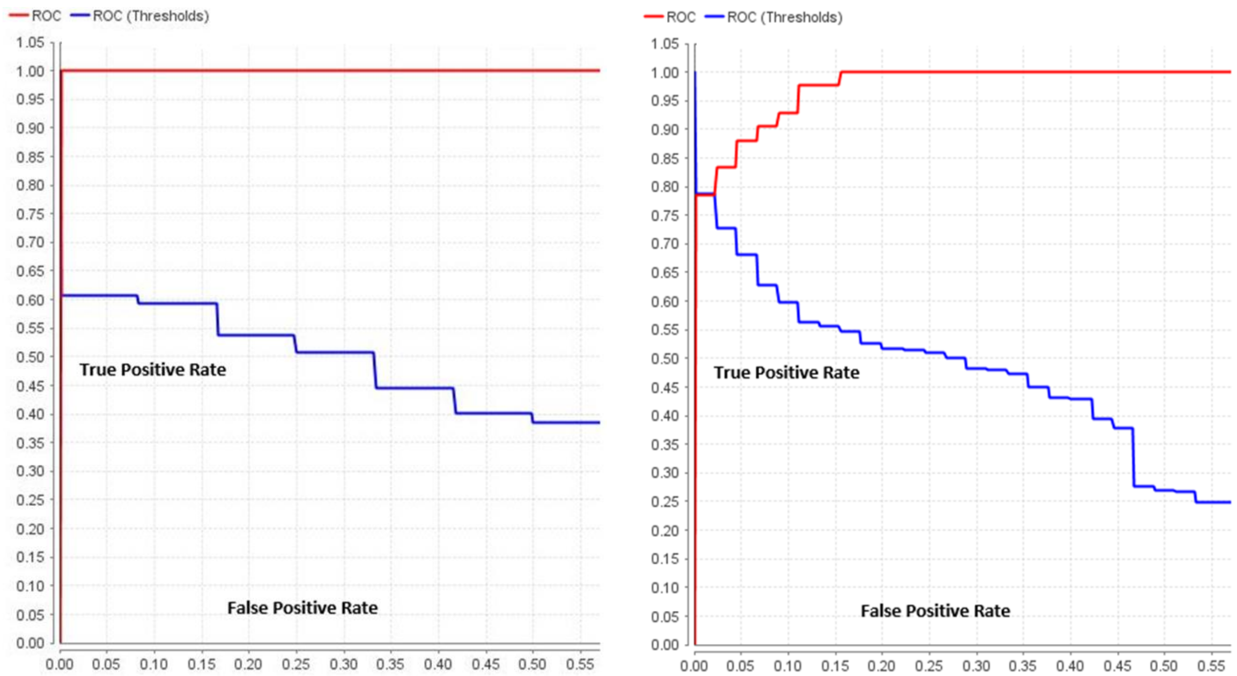
on dataset-1 on dataset-2

Fig. 6: The ROC for the DT.



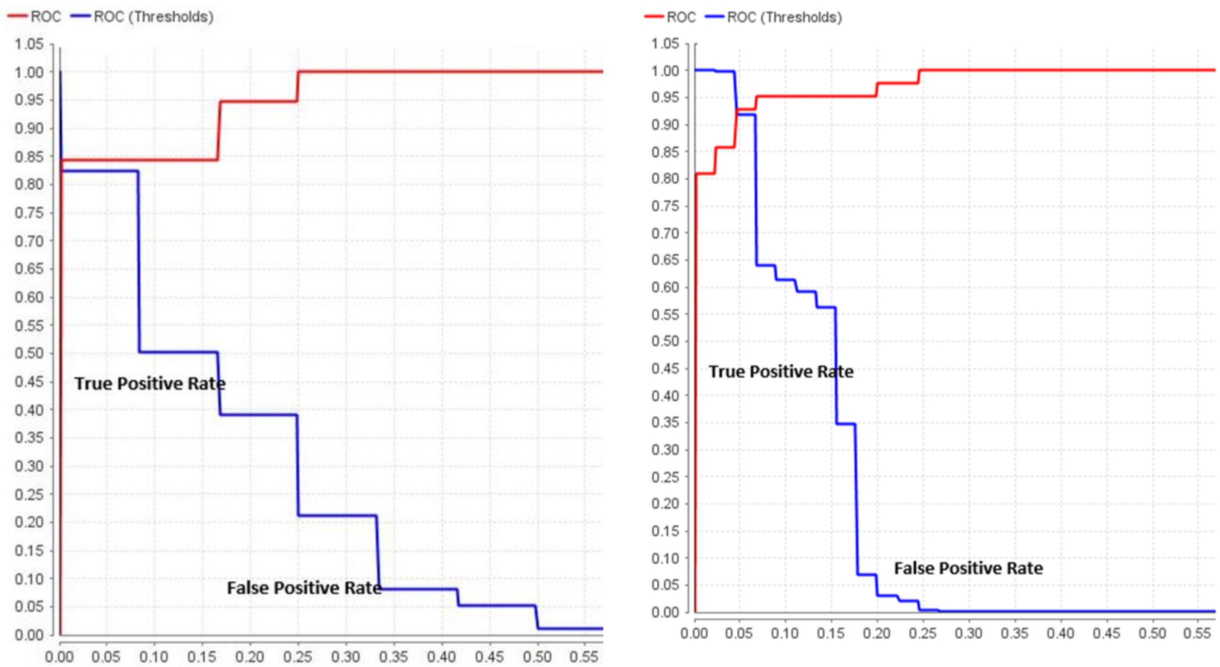
on dataset-1 on dataset-2

Fig. 7: The ROC for the KNN.



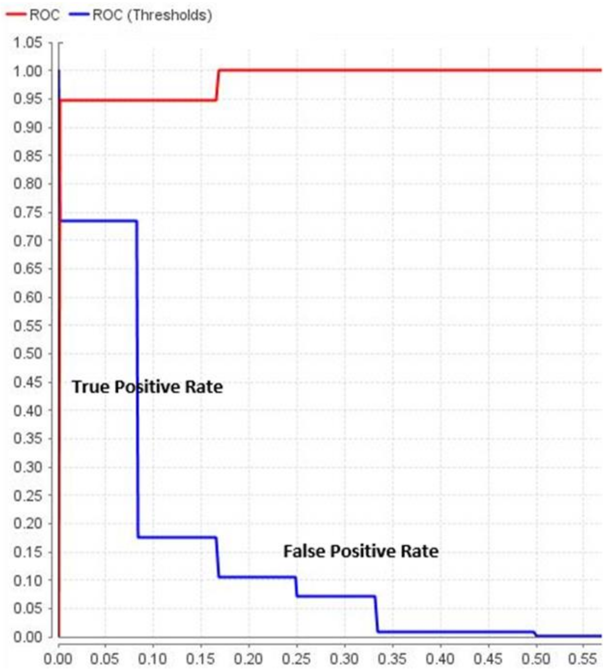
on dataset-1 on dataset-2

Fig. 8: The ROC for the SVM.

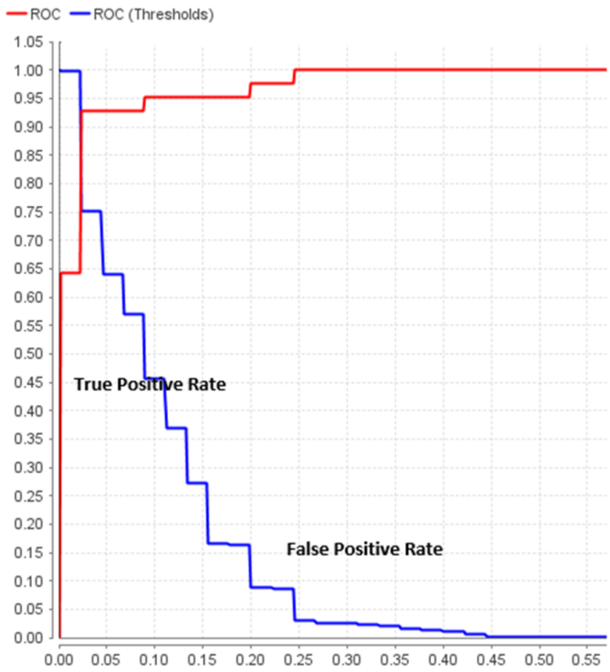


on dataset-1 on dataset-2

Fig. 9: The ROC for the NN.



on dataset-1



on dataset-2

Fig. 10: The ROC for the NB.

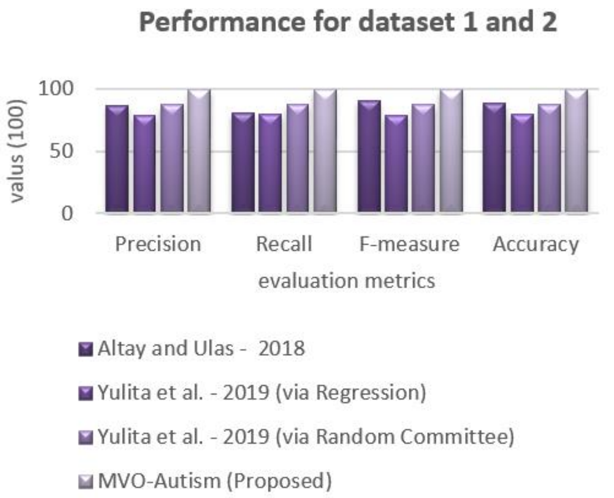


Fig. 11: A comparison among MVO-Autism and other works on datasets.

Conclusion

Autism is the most common disease that occurs in any age people. For the diagnosis of this disease, the techniques used to improve the classification performance gave better results, and other computerized technologies were applied. These technologies compete for performance standards. However, poor quality leads most therapeutic datasets to unreliability and inaccurate results.

The failure to use pretreatment techniques on these datasets prevents producing effective models for predicting autism. This article remains on a compound of pre-treatment methods consist of lost and value substitution characteristics. The missing values are replaced by using the Replace Missing Value through mean and detect outlier through KNN. Methods for selecting a subgroup during classification on known autism. The dataset is the autism dataset. Experiments were conducted to evaluate the impact of missing value analysis and choosing a set of algorithms, including DT, NB, NN, SVM, and KNN classifiers to achieve the best performance for our model. We observed that Combine substitution with the mean (e.g. Missing Value Analysis Method) and improvement selection with RF technology and DT technology. The accuracy of the two technologies has reached 100%, and Precision 100% outperforms their counterparts. The obtained results reveal that the suggested model can increase classification performance and exceeds autism prediction. For future work, we want to work on ensemble heuristic methods and other pre-treatment techniques in this context.

Author Contributions

K. Ali Mohsin Alhameedawi designed the experiments and collected the data and carried out the data analysis. R. Asgarnezhad and K. Ali Mohsin Alhameedawi evaluated the results and rewrote the manuscript.

Acknowledgment

We thank the editor and all anonymous reviewers.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

I	Samples
J	Variables
P	Probability
c	Class
S	Data
T	Number of trees
d	Current data to be assigned
Value	The various values that the variable j can have
P(c|x)	The following possibility of class
P(c)	The previous possibility of class
P(x|c)	The probability is the possibility of predictor provided class
P(x)	The previous likelihood of predictor

References

[1] M.L. Pennington, D. Cullinan, L.B. Southern, "Defining autism: variability in state education agency definitions of and evaluations for autism spectrum disorders," Autism Res. Treat., 2014: 1-8, 2018.

[2] P.O. Towle, P.A. Patrick, "Autism spectrum disorder screening instruments for very young children: a systematic review," Autism Res. Treat., 2016: 1-29, 2016.

[3] M. Legerstee, D. Anderson, A. Schaffer, "Five-and eight-month-old infants recognize their faces and voices as familiar and social stimuli," Child Dev., 69: 37-50, 1998.

[4] G. Dawson, A.N. Meltzoff, J. Osterling, J. Rinaldi, E. Brown, "Children with autism fail to orient to naturally occurring social stimuli," J. Autism Dev. Disord., 28: 479-485, 1998.

[5] J. Osterling, G. Dawson, "Early recognition of children with autism: A study of first birthday home videotapes," J. Autism Dev. Disord., 24: 247-257, 1994.

[6] D.J. Grelotti, I. Gauthier, R.T. Schultz, "Social interest and the development of cortical face specialization: What autism teaches us about face processing," Dev. Psychobiol., 40: 213-225, 2002.

[7] K. Pelphrey, R. Adolphs, J.P. Morris, "Neuroanatomical substrates of social cognition dysfunction in autism," Ment. Retard. Dev. Disabil. Res. Rev., 10: 259-271, 2004.

[8] A. Roman-Urrestarazu, C. Yáñez, C. López-Garí, C. Elgueta, C. Allison, C. Brayne, et al., "Autism screening and conditional cash transfers in Chile: Using the Quantitative Checklist (Q-CHAT) for early autism detection in a low resource setting," Autism, 25: 932-945, 2021.

[9] G. Divan, S. Bhavnani, K. Leadbitter, C. Ellis, J. Dasgupta, A. Abubakar, et al., "Annual Research Review: Achieving universal health coverage for young children with autism spectrum disorder in low-and middle-income countries: A review of reviews," J. Child Psychol. Psychiatry, 62: 514-535, 2021.

[10] P. Mazumdar, G. Arru, F. Battisti, "Early detection of children with autism spectrum disorder based on visual exploration of images," Signal Process. Image Commun., 94: 116184, 2021.

[11] R. Asgarnezhad, A. Monadjemi, M. Soltanaghaei, "A high-performance model based on ensembles for twitter sentiment classification," J. Electr. Comput. Eng. Innovations, 8(1): 41-52, 2020.

[12] R. Asgarnezhad, S.A. Monadjemi, M. Soltanaghaei, "An application of MOGW optimization for feature selection in text classification," J. Supercomp, 77: 5806-5839, 2021.

[13] R. Asgarnezhad, A. Monadjemi, M. Soltanaghaei, "NSE-PSO: Toward an effective model using optimization algorithm and sampling methods for text classification," J. Electr. Comput. Eng. Innovations 8(2): 183-192, 2020.

[14] R. Asgarnezhad, S.A. Monadjemi, "NB VS. SVM: AContrastive study for sentiment classification on two text domains," J. Appl. Intell. Syst. Inform. Sci., 2: 1-12, 2021.

[15] L. Zhang, A.Z. Amat, H.Zhao, A. Swanson, A.S. Weitlauf, Z. Warren, et al., "Design of an intelligent agent to measure collaboration and verbal-communication skills of children with autism spectrum disorder in collaborative puzzle games," IEEE Trans. Learn. Technol., 14(3): 338-352, 2020.

[16] S. Kopp, L. Gesellensetter, N.C. Krämer, I. Wachsmuth, "A conversational agent as museum guide–design and evaluation of a real-world application," in Proc. International Workshop on Intelligent Virtual Agents: 329-343, 2005.

[17] A. Tewari, T. Brown, J. Canny, "A question-answering agent using speech driven non-linear machinima," in Proc. International Workshop on Intelligent Virtual Agents: 129-138, 2013.

[18] R. Yaghoubzadeh, K. Pitsch, S. Kopp, "Adaptive grounding and dialogue management for autonomous conversational assistants for elderly users," in Proc. International Conference on Intelligent Virtual Agents: 28-38, 2015.

[19] C.-H. Min, "Automatic detection and labeling of self-stimulatory behavioral patterns in children with Autism Spectrum Disorder," in Proc. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): 279-282, 2017.

[20] T. Westeyn, K. Vadas, X. Bian, T. Starner, G.D. Abowd, "Recognizing mimicked autistic self-stimulatory behaviors using hmms," in Proc. Ninth IEEE International Symposium on Wearable Computers (ISWC'05): 164-167, 2005.

[21] E. Linstead, R. German, D. Dixon, D. Granpeesheh, M. Novack, A. Powell, "An application of neural networks to predicting mastery of learning outcomes in the treatment of autism spectrum disorder," in Proc. 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA): 414-418, 2015.

[22] F.D. Foresee, M.T. Hagan, "Gauss-Newton approximation to Bayesian learning," in Proc. International Conference on Neural Networks (ICNN'97): 1930-1935, 1997.

[23] O. Altay, M. Ulas, "Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and K-nearest neighbor in children," in Proc. 2018 6th International Symposium on Digital Forensic and Security (ISDFS): 1-4, 2018.

[24] M.J. Maenner, M. Yeargin-Allsopp, K. Van Naarden Braun, D.L. Christensen, L.A. Schieve, "Development of a machine learning algorithm for the surveillance of autism spectrum disorder," PloS one, 11(12): 1-11,, 2016.

[25] W. Liu, M. Li, L. Yi, "Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework," Autism Res., 9: 888-898, 2016.

[26] Y. Jiao, R. Chen, X. Ke, K. Chu, Z. Lu, E.H. Herskovits, "Predictive models of autism spectrum disorder based on brain regional cortical thickness," Neuroimage, 50: 589-599, 2010.

[27] C. Ecker, V. Rocha-Rego, P. Johnston, et al., "Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach," Neuroimage, 49: 44-56, 2010.

[28] I.N. Yulita, M.I. Fanany, A.M. Arymurthy, "Comparing classification via regression and random committee for automatic sleep stage classification in autism patients," J. Phys. Conf. Ser., 1230: 1-9, 2019.

[29] R. Anirudh, J.J. Thiagarajan, "Bootstrapping graph convolutional neural networks for autism spectrum disorder classification," in Proc. 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): 3197-3201, 2019.

[30] X.A. Bi, Y. Wang, Q. Shu, Q. Sun, Q. Xu, "Classification of autism spectrum disorder using random support vector machine cluster," Front. Genet., 9: 1-18, 2018.

[31] Y. Kong, J. Gao, Y. Xu, Y. Pan, J. Wang, J. Liu, "Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier," Neurocomputing, 324: 63-68, 2019.

[32] S. Raj, S. Masood, "Analysis and detection of autism spectrum disorder using machine learning techniques," Procedia Comput. Sci., 167: 994-1004, 2020.

[33] H. Cheng, J. Yu, L. Xu, J. Li, "Power spectrum of spontaneous cerebral homodynamic oscillation shows a distinct pattern in autism spectrum disorder," Biomed. Opt. Express, 10: 1383-1392, 2019.

[34] A. Kazeminejad, R.C. Sotero, "Topological properties of resting-state fMRI functional networks improve machine learning-based autism classification," Front. Neurosci., 12: 10-18, 2019.

[35] Z. Sherkatghanad, M. Akhondzadeh, S. Salari, M. Zomorodi-Moghadam, M. Abdar, U.R. Acharya, et al., "Automated detection of autism spectrum disorder using a convolutional neural network," Front. Neurosci., 13: 13-25, 2020.

[36] J. Han, M. Kamber, J. Pei, "Data mining concepts and techniques third edition," The Morgan Kaufmann Series in Data Management Systems, 5: 83-124, 2011.

[37] V. Vapnik, The nature of statistical learning theory: Springer science & business media, 2013.

[38] T.K. Ho, "Random decision forests," in Proc. 3rd international conference on document analysis and recognition: 278-282, 1995.

[39] T.K. Ho, "The random subspace method for constructing decision forests," IEEE Trans. Pattern Anal. Mach. Intell., 20: 832-844, 1998.

[40] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," in Proc. European conference on machine learning: 137-142, 1998.

[41] S. Zhang, "KNN-CF Approach: Incorporating certainty factor to kNN classification," IEEE Intell. Informatics Bull., 11: 24-33, 2010.

[42] R. Asgarnezhad, K. Ali Mohsin Alhameedawi, "Weka vs. Rapid Miner: Toward a comprehensive comparison for diagnosis of autism using ensembles," Int. J. Software Eng. Appl., 6: 1-17, 2021.

[43] R. Asgarnezhad, S.A. Monadjemi, M. Soltanaghaei, "A new hierarchy framework for feature engineering through multi-objective evolutionary algorithm in text classification," Concurrency Comput. Pract. Ex., 2021.

Biography

Karrar Ali Mohsin Alhameedawi received his B.Sc. degrees in Computer Engineering from the Al-Rafidain University of Baghdad in 2018. He received a certified international development trainer specialized in human development from Germany in 2019. After one year, he received the title of the best humanitarian personality in Iraq in 2020. He has experience as a website developer, programmer, the highest educational social networking site in Iraq, and an electronic contest program. He is a member of seven international institutions and has received 200 certificates and a letter of thanks from European and local countries. He is currently an MSc. student at Azad Isfahan University. His current researches include Data Mining and Big Data.

Razieh Asgarnezhad received her B.Sc. and MSc. degrees in Computer Engineering from Kashan Azad University in 2009 and Arak Azad University in 2012, respectively. She received Ph.D. degree from Isfahan Azad University in 2020. Her current researches include Data Mining, Text Mining, Learning Automata, Recommendation System, Sentiment Analysis, and Wireless Sensor Network.

How to cite this paper:
R. Asgarnezhad, K. Ali Mohsin Alhameedawi, “MVO-Autism: An effective pre-treatment with high performance for improving diagnosis of autism mellitus,” J. Electr. Comput. Eng. Innovations, 10(1): 209-220, 2022.

DOI: 10.22061/JECEI.2021.8109.480

URL: https://jecei.sru.ac.ir/article_1605.html