Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 167 (2020) 994–1004

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques

Suman Raj[a] ,Sarfaraz Masood[b*]

[a, b]Department of Computer Engineering,Jamia Millia Islamia University,New Delhi-110025 ,India

Abstract

Autism Spectrum Disorder (ASD) is a neuro-disorder in which a person has a lifelong effect on interaction and communication with others. Autism can be diagnosed at any stage in once life and is said to be a "behavioral disease" because in the first two years of life symptoms usually appear. According to the ASD problem starts with childhood and continues to keep going on into adolescence and adulthood. Propelled with the rise in use of machine learning techniques in the research dimensions of medical diagnosis, in this paper there is an attempt to explore the possibility to use Naïve Bayes, Support Vector Machine, Logistic Regression, KNN, Neural Network and Convolutional Neural Network for predicting and analysis of ASD problems in a child, adolescents, and adults. The proposed techniques are evaluated on publicly available three different non-clinically ASD datasets. First dataset related to ASD screening in children has 292 instances and 21 attributes. Second dataset related to ASD screening Adult subjects contains a total of 704 instances and 21 attributes. Third dataset related to ASD screening in Adolescent subjects comprises of 104 instances and 21 attributes. After applying various machine learning techniques and handling missing values, results strongly suggest that CNN based prediction models work better on all these datasets with higher accuracy of 99.53%, 98.30%, 96.88% for Autistic Spectrum Disorder Screening in Data for Adult, Children, and Adolescents respectively.

Keywords: Autism Spectrum Disorder (ASD); Convolutional neural network (CNN); Artificial Neural Network (ANN); K- Nearest Neighbours (KNN); Logistic Regression (LR); Support Vector Machine (SVM);

Introduction

The problem of autism spectrum disorder (ASD) have been mounting swiftly nowadays among all ages of the human population. Early detection of this neurological disease can greatly assist in the maintenance of the subject's mental and physical health. With the rise of application of machine learning-based models in the predictions of various

Corresponding author E-mail: smasood@jmi.ac.in

1877-0509
10.1016/j.procs.2020.03.399

human diseases, their early detection based on various health and physiological parameter now seems possible. This factor motivated us to increase interest in the detection and analysis of ASD diseases to improve better treatment methodology. Detection of ASD becomes a challenge as there are several other mental disorders whose few symptoms are very similar to those with ASD symptoms, thereby makes this task a difficult one.

Autism Spectrum disorder is a problem that is related to human brain development. A person who has suffered from the Autism Spectrum Disorder is generally not able to do social interaction and communication with other persons [1] or [3]. In this, a person's life is usually affected for his or her entire lifetime. It is interesting to know that both environmental and genetic factors may turn out to be the causing factors for this disease. The symptoms of this problem may be started at the age of three years and may continue for the lifetime. It is not possible to complete treat the patient suffering from this disease, however its effects can be reduced for some time if the symptoms are early detected. By assuming that human genes are responsible for it, the exact causes of ASD have not been recognized by the scientist yet. The human genes affect the development by influencing the environment. There is some risk factor which influences ASD like as low birth weight, a sibling with ASD and having old parents, etc. Instead of this, there are some social interaction and communication problems like as:

- Inappropriate laughing and giggling
- No sensitivity of pain
- Not able to make eye contact properly
- No proper response to sound
- May not have a wish for cuddling
- Not able to express their gestures
- No interaction with others
- Inappropriate objects attachment
- Want to live alone
- Using echo words etc.

People with ASD also have difficulty with constrained interests and consistently repetition of behaviors. The following list presents specific examples of the types of behaviors.

- Repeating certain behaviors like repeating words or phrases much time.
- The Person will be upset when a routine is going to change.
- Having a little interest in certain matters of the topic like numbers, facts, etc.
- Less sensitive than another person in some cases like light, noise, etc.

Early detection and treatment are most important steps to be taken to decrease the symptoms of autism spectrum disorder problem and to improve the quality of life of ASD suffering people. However, there is no procedure of medical test for detection of autism. ASD Symptoms usually recognized by observation. In Older and adolescents who go to school, ASD symptoms are usually identified by their parents and teachers. After that ASD symptoms are evaluated by a special education team of the school. These school team suggested these children visit their health care doctor for required testing. In adults identifying ASD symptoms is very difficult than older children and adolescents because some symptoms of ASD may be overlap with other mental health disorders. It is easy to identify the behavioral changes in a child easily by observation because it can be seen early in the 6 months of age than Autism specific brain imaging because brain imaging can be identifying after 2 years of age.

The contents of this paper are organized as follows: Section 1 presents the introduction to the Autism Spectrum Disorder problem and the challenges faced by the subjects. Section 2 presents the review of various recent literature, where some models for ASD detection have been developed. Section 3 describes the datasets used in this study, which is followed by description of each component of the methodology used in this work in section 4. The results obtained after various experiments are presented and discussed in Section 5 which is finally followed by the conclusion in section 6.

*Corresponding author. Tel.: +0-965-477-4604.
E-mail: smasood@jmi.ac.in

Literature Survey

Vaishali R, Sasikala R. et al. [3] have proposed a method to identify Autism with optimum behavior sets. In this work, an ASD diagnosis dataset with 21 features obtained from the UCI machine learning repository experimented with swarm intelligence based binary firefly feature selection wrapper. The alternative hypothesis of the experiment claims that it is possible for a machine learning model to achieve a better classification accuracy with minimum feature subsets. Using Swarm intelligence based single-objective binary firefly feature selection wrapper it is found that 10 features among 21 features of ASD dataset are sufficient to distinguish between ASD and non-ASD patients. The results obtained with this approach justifies the hypothesis by producing an average accuracy in the range of 92.12%-97.95% with optimum feature subsets which are approximately equal to the average accuracy produced by the entire ASD diagnosis dataset.

Fadi Thabtah et al. [8] have proposed an ASD screening model using Machine Learning Adaption and DSM-5. A screening tool has been used to realize one or more goals in ASD screening. In this paper, the researcher discussed the ASD Machine Learning classification with their pros and cons. The researcher tried to highlight the problem accompanying with existing ASD screening tools and the consistency of such tools using the DSM-IV instead of the DSM-5 manual.

M. S. Mythili, A. R. Mohamed Shanavas et al. [13] have a study on ASD using Classification Techniques. The main aim of this paper was to detect the autism problem and the levels of autism. In this Neural Network, SVM and Fuzzy techniques with WEKA tools are used to analyze the student's behavior and their social interaction.

J. A. Kosmicki1, V. Sochat, M. Duda and D.P. Wall Et al. [14] have supposed a searching method for a least set of traits for autism detection. In this, the authors used a machine learning approach to evaluate the clinical assessment of ASD. The ADOS was performed on the subset of behaviors of children based on the autism spectrum. ADOS having four modules. In this work, 8 different machine learning algorithms were employed, involving stepwise backward feature identification on score sheets from 4540 individuals. It uses 9 out of the 28 behaviors from module 2 and 12 out of the 28 behaviors from module 3 to identify an ASD risk with an overall accuracy of 98.27% and 97.66% respectively.

Li B, A. Sharma, J Meng, S. Purushwalkam, E. Gowen (2017) et al. [11] have used machine learning classifiers to detect autistic adults by imitation method. The goal of this study was to investigate the fundamental problem related to discriminative test conditions and kinematic parameters. The dataset contains 16 ASC participants who have a series of hand movements. In this 40 kinematic constraints from 08 imitation conditions has been extracted by using machine learning methods. This research shows that for a small sample, there is a feasibility of applying machine learning methods to analyze high-dimensional data and the diagnostic classification of autism. The sensitivity rates achieved by RIPPER which have the features Va (87.30 %), CHI (80.95%), IG (80.95%), Correlation (84.13%), CFS (84.13%), and "no feature selection"(80.00%) on the AQ-Adolescent dataset.

It is evident from the above discussed section that there is definitely a need to explore the possibility of applying deep learning based models for the detection of ASD in human population. Most of the work discussed above use conventional machine learning approaches and hence are limited in their performance. In this work, performance of several machine learning models have been compared to that of the deep learning model for this purpose. Separate models have been prepared for separate population set (discussed in section below) and compared individually.

3. Dataset

Dataset for this research purpose has been collected from the UCI Repository which is publicly available [12] or [15] or [16]. In this research mainly three types of the dataset have been used. The detailed summary of the dataset is given below:

Table 1: List of ASD datasets

Sr. No.	Dataset Name	Sources	Attribute Type	Number of Attributes	Number of Instances
1	ASD Screening Data for Adult	UCI Machine Learning Repository [12]	Categorical, continuous and binary	21	704
2	ASD Screening Data for Children	UCI Machine Learning Repository [15]	Categorical, continuous and binary	21	292
3	ASD Screening Data for Adolescent	UCI Machine Learning Repository [16]	Categorical, continuous and binary	21	104

These datasets have 20 common attributes that are used for prediction. These attributes are listed below:

Table 2: List of Attributes in the dataset

Attribute Id	Attributes Description
1	Patient age
2	Sex
3	Nationality
4	The patient suffered from Jaundice problem by birth
5	Any family member suffered from pervasive development disorders
6	Who is fulfilment the experiment
7	The country in which the user lives
8	Screening Application used by the user before or not?
9	Screening test type
10-19	Based on the screening method answers of 10 questions
20	Screening Score

4. Proposed Methodology

Figure 1 shows the steps in the proposed workflow which involves the pre-processing of data, training, and testing with specified models, evaluation of results and prediction of ASD. This work is implemented in Python 3.

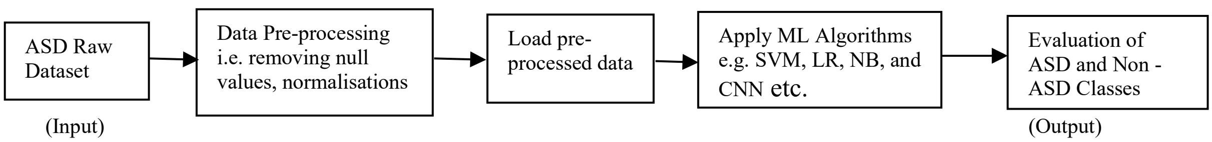


Figure. 1. Steps in the proposed ASD detection solution

4.1. Data Pre-Processing

Data pre-processing is a technique in which transform the raw data into a meaningful and understandable format. Real-world data is commonly incomplete and inconsistent because it contains lots of errors and null values. A good pre-processed data always yields to a good result. Various Data pre-processing methods are used to handle incomplete and inconsistent data like as handling missing values, outlier detection, data discretization, data reduction (dimension and numerosity reduction), etc. The problems of missing values in these dataset has been handled by imputation method.

4.2. Training and Testing Model

The whole dataset has been split into two parts i.e. one part is training the dataset and the other one is testing dataset with a ratio of 80:20 respectively. For cross-validation purposes again training data has been split into two parts. One part is the training dataset and another part is the validation dataset into an 80:20 ratio respectively. Figure 2 shows the final training, testing and validation sets on which classification has been performed.

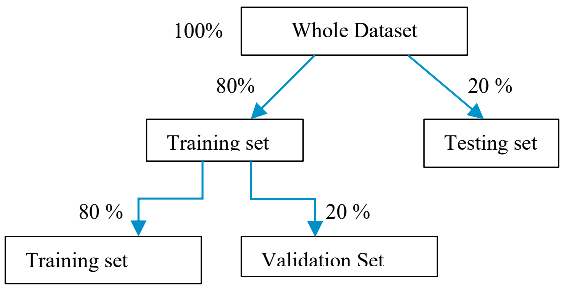


Figure. 2. Final Training, Testing and Validation Sets

4.2.1 Support Vector Machine (SVM)

SVM is a linear supervised machine learning approach that is used for classification and regression. It is a pattern recognition problem solver. It does not cause the problem of overfitting. SVM separates the classes by defining a decision boundary [19].

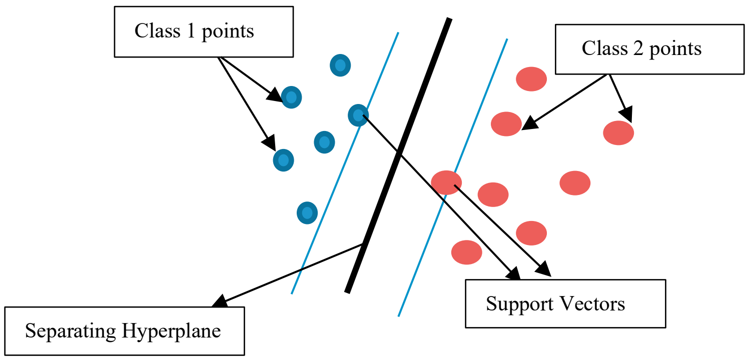


Figure. 3. An SVM classifier

4.2.2 Naïve Bayes (NB)

A naive Bayes classifier is a supervised learning algorithm. It is a generative model and is based on joint probability distribution. The Naive Bayes concept based on independence assumptions. It exhibits less training time as compared to SVM and ME model. It calculates the posterior probability for a dataset using the prior probability and likelihood [17].

4.2.3 Convolutional Neural Network (CNN)

CNN is one of the deep learning techniques known to build models for various problems [24] or [25] or [26]. It is a feed-forward neural network that is inspired by the human brain. A CNN model contains one input layer, one output layer, and many other different layers i.e. convolution layers, max pooling, fully connected layers, and normalization layers. Their activation functions can be computed with Matrix Multiplication, which is followed by a bias offset. A simple diagram of CNN is given below:

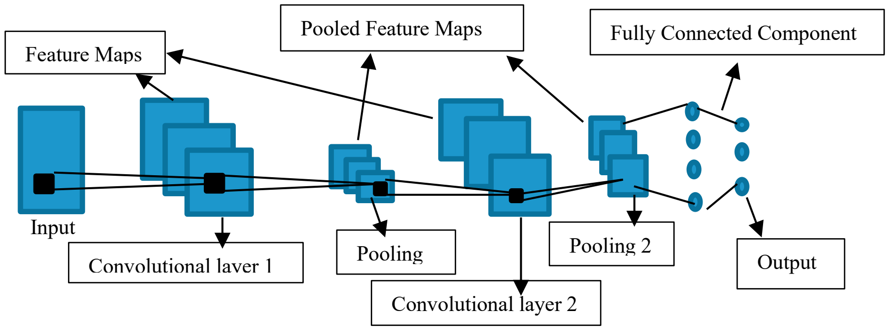


Figure. 4. Basic Structure of a CNN model

4.2.4 Logistic Regression (LR)

LR is a regression tool that is used to analyse the binary dependent variables. Its output value lies in either the 0 or 1 form. It is used for the continuous value dataset. It tells the relationship between one dependent binary variable, and one nominal or ordinary variable. It can be represented by the sigmoidal function.

4.2.5 K- Nearest Neighbour (KNN)

KNN is a supervised learning approach and is the simplest of all. It is used for classification as well as regression problems. It assumes that similar data exist nearby. The 'K' part indicates the number of seed point that is to be selected. It should be chosen carefully to reduce the error. Thus it is based on the idea of similarity which can be in terms of distance, closeness or proximity. The most common distance measure is Euclidean distance.

4.2.6 Artificial Neural Network

ANN is a neural network that has a connection with multiple neurons. Each neuron cell having a group of input values and associated weights. The most common artificial 0 neural network feeds forward neural network. In this network, the flow of information moves in the only forward direction. This type of network contains three main layers, first is the input layer, the second is a hidden layer and last is the output layer. There is no cycle or loop in the network [21].

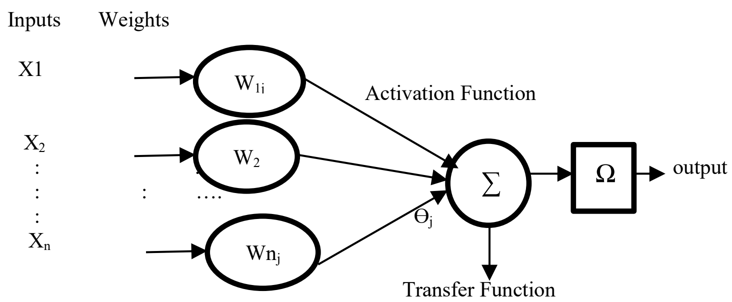


Figure. 5. Artificial Neural Network

5. Result and Discussion

The result is measured in terms of specificity, sensitivity, and accuracy by using the confusion matrix and classification report. The result depends on how accurate the model is trained.

5.1 Performance Evaluation metrics

Measuring performance is key to check how well a classification model work to achieve a target. Performance evaluation metrics are used to evaluate the effectiveness and performance of the classification model on the test dataset. It is important to choose the correct metrics to evaluate the model performance such as confusion matrix, accuracy, specificity, sensitivity, etc. Following formulas are used to find the performance metrics:

Table 3: Elements of a Confusion Matrix

	Predictive ASD values	
Actual ASD values	True positive (TP)	False Positive (FP)
	False Negative (FN)	True Negative (TN)

$$\text{Specificity} = \frac{TN}{(TN+TP)} \tag{1}$$

$$\text{True Positive Rate or Sensitivity} = \frac{TN}{(TN+FN)} \tag{2}$$

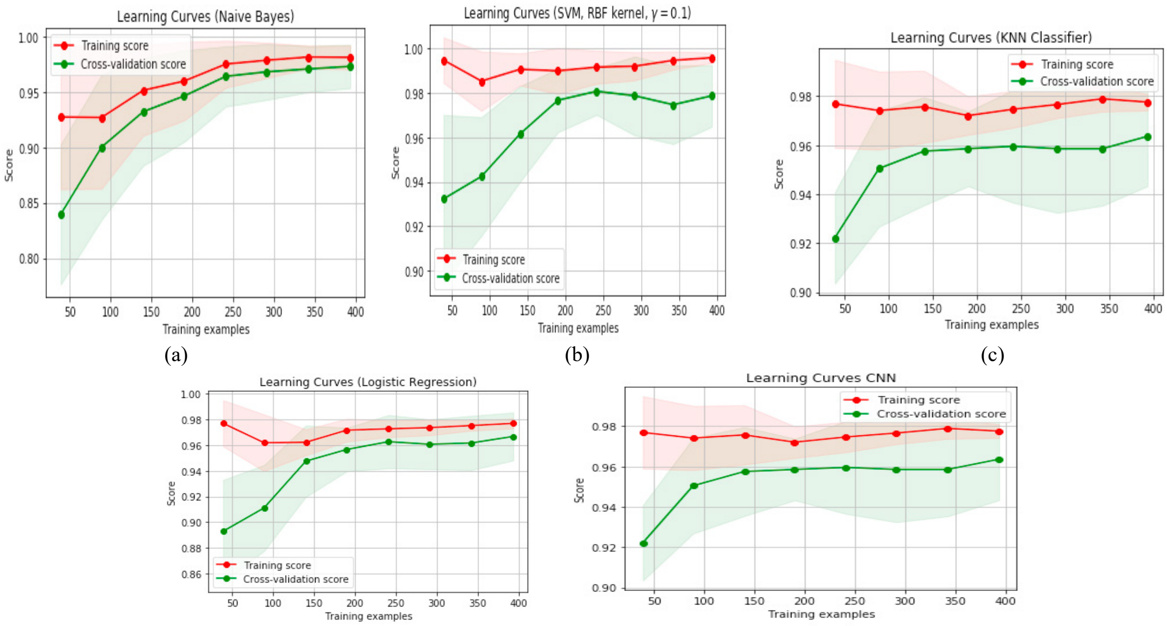
$$\text{Accuracy} = \frac{TP+TN}{(TN+TP+FP+FN)} \tag{3}$$

Experimental results of various machine learning algorithms approach with all features selection have been shown for ASD screening data for Adults, children, and adolescents. In this, all 21 features are selected to find the specificity, sensitivity, and accuracy of the predicted model. For the implementation of Naïve Bias algorithm Gaussian NB has been used. For SVM, RBF Kernel has been used with 0.1 gamma value. In KNN, N=5 has been used. In ANN, Adam Optimizer with 0.01 learning rate and 100 epoch has been used. In CNN, Relu activation Function, Adam Optimizer, binary cross-entropy loss function, 16 & 32 filters and 0.5 dropouts with 150 epoch has been used. The Overall performance measures of all machine learning classifiers with all 3 datasets have been shown below in details:

Table 4: Overall Results for Autistic Spectrum Disorder Screening Data for Adult

Classifier	Specificity	Sensitivity	Accuracy
Logistic Regression	0.9575	0.9696	96.69
SVM	0.9574	0.88888	98.11
Naive Bayes	0.9361	96.96	96.22
KNN	0.9148	0.9696	95.75
ANN	0.9787	0.9757	97.64
CNN	1.0	0.9939	99.53

Evaluation of various machine learning models on ASD adult diagnosis dataset observed an accuracy in the range of (95.75% to 99.53 %) on the original dataset. K-NN classifier with K=5 has produced the least accuracy of 95.75%. CNN produced 99.53 % prediction accuracy on the original dataset. The learning curves of all Machine Learning algorithms also describe the results of the prediction model.



(a) (b) (c)

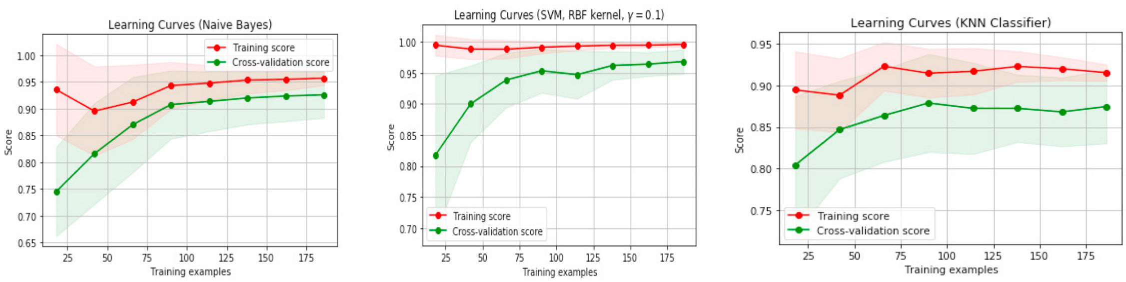
(d) (e)

Figure 6: Learning Curve of (a) Naïve Bayes; (b) SVM; (c) KNN; (d) Logistic Regression; (e) CNN for adult's dataset

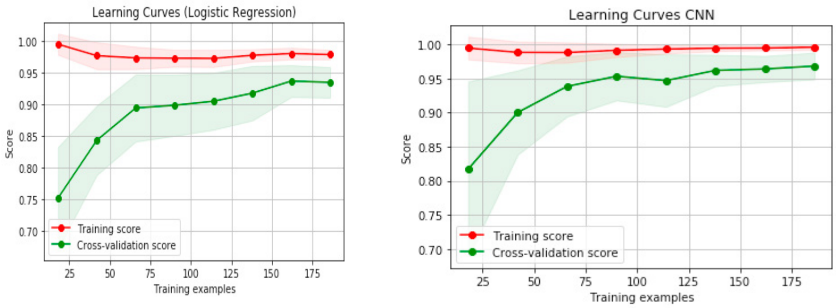
Table 5: Overall Results for Autistic Spectrum Disorder Screening Data for children

Classifier	Specificity	Sensitivity	Accuracy (%)
Logistic Regression	1.0	0.9677	98.30
SVM	1.0	0.9679	98.30
Naive Bayes	0.9642	0.9354	94.91
KNN	0.9642	0.8064	88.13
ANN	0.9642	1.0	98.30
CNN	1.0	0.9678	98.30

Evaluation of various machine learning models on ASD children's diagnosis dataset observed an accuracy in the range of (88.13% to 98.30 %) on the original dataset. K-NN classifier with K=5 has produced the least accuracy of 88.13%. CNN, SVM, ANN, and LR produced 98.30 % prediction accuracy on the original dataset. The learning curves of all Machine Learning algorithms also describe the results of the prediction model.



(a) (b) (c)



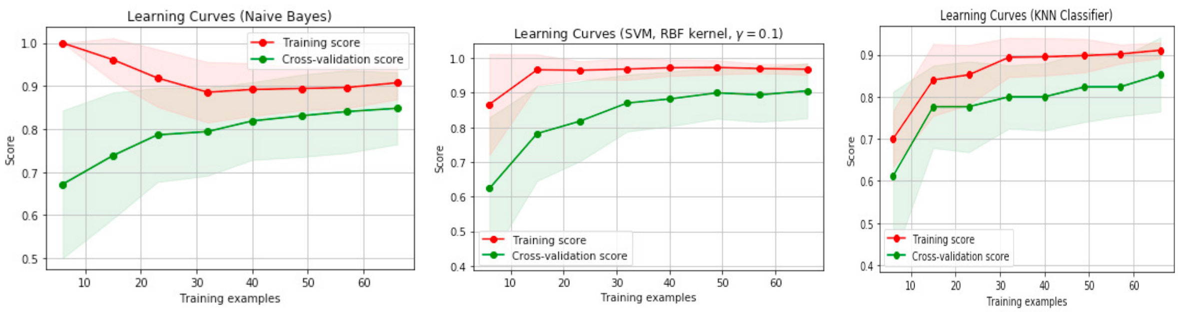
(d) (e)

Figure 7: Learning Curve of (a) Naïve Bayes; (b) SVM; (c) KNN; (d) Logistic Regression; (e) CNN for children's dataset.

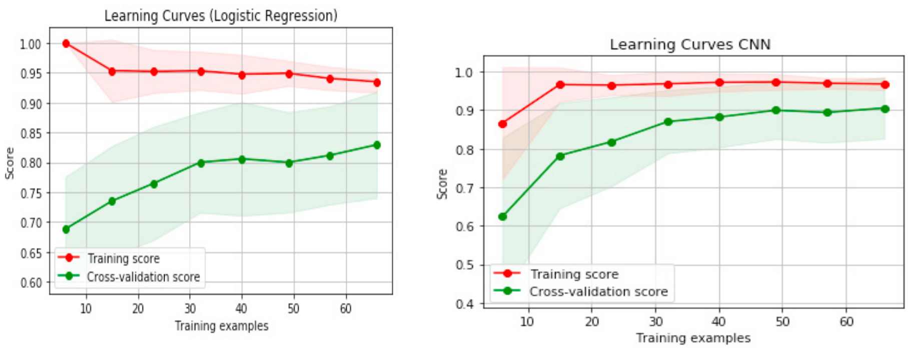
Table 6: Overall Results for Autistic Spectrum Disorder Screening Data for Adolescent

Classifier	Specificity	Sensitivity	Accuracy (%)
Logistic Regression	1.0	0.6666	85.71
SVM	1.0	0.8888	95.23
Naive Bayes	0.9166	0.8888	90.47
KNN	1.0	0.5555	80.95
ANN	1.0	0.7777	90.47
CNN	1.0	0.9335	96.88

Evaluation of various machine learning models on ASD Adolescent diagnosis dataset observed an accuracy in the range of (80.95 % to 96.88 %) on the original dataset. K-NN classifier with K=5 has produced the least accuracy of 80.95%.CNN classifiers produced the highest 96.88 % prediction accuracy on the original dataset.



(a) (b) (c)



(d) (e)

Figure 8: Learning Curve of (a) Naïve Bayes; (b) SVM; (c) KNN; (d) Logistic Regression; (e) CNN for adolescent's dataset.

6. Conclusion

In this work, detection of Autism Spectrum Disorder was attempted using various machine learning and deep learning techniques. Various performance evaluation metrics were used to analyze the performance of the models implemented for ASD detection on non-clinical dataset from three sets of age groups viz. Child, Adolescents and the Adult. When comparing the result with another recent study [3] on this problem got a better result of the CNN classifier instead of SVM with including all its features attributes after handling missing values. In this work after handling missing value, both the SVM and CNN based models show the same accuracy of prediction of about 98.30 % for ASD Child dataset. However for the remaining two other datasets, the CNN based model was able to achieve highest accuracy result than all the other considered model building techniques, These results strongly suggest that a CNN based model can be implemented for detection of Autism Spectrum Disorder instead of the other conventional machine learning classifier suggested in earlier researches.

Table 7:Comparison of results with existing methods [3] on Autistic Spectrum Disorder Screening Data for Children

Model	Accuracy before handling missing values	Accuracy after handling missing values
Support Vector Machine	97.95	98.30
Artificial Neural Network	97.60	98.30
Convolutional Neural Network	Not implemented	98.30

References

[1] Thabtah, Fadi. "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward. (2018) " Informatics for Health and Social Care : 1-20.

[2] Thabtah, Fadi, Firuz Kamalov, and Khairan Rajab. (2018) "A new computational intelligence approach to detect autistic features for autism screening." International journal of medical informatics **117**: 112-124.

[3] Vaishali, R., and R. Sasikala. "A machine learning based approach to classify Autism with optimum behaviour sets. (2018) " International Journal of Engineering & Technology **7(4)**: 18.

[4] Constantino, John N., Patricia D. Lavesser, Y. I. Zhang, Anna M. Abbacchi, Teddi Gray, and Richard D. Todd. (2007) "Rapid quantitative assessment of autistic social impairment by classroom teachers." Journal of the American Academy of Child & Adolescent Psychiatry **46(12)**: 1668-1676.

[5] Daniel Bone, Matthew S. Goodwin, Matthew P. Black, Chi-Chun Lee, Kartik Audhkhasi, and Shrikanth Narayanan. (2015) "Applying machine learning to facilitate autism diagnostics: pitfalls and promises." *Journal of autism and developmental disorders* **45(5)**: 1121-1136.

[6] Dennis Paul Wall, J. Kosmicki, T. F. Deluca, E. Harstad, and Vincent Alfred Fusaro. (2012) "Use of machine learning to shorten observation-based screening and diagnosis of autism." *Translational psychiatry*, ***2*(4):** e100.

[7] Dennis P. Wall, Rebecca Dally, Rhiannon Luyster, Jae-Yoon Jung, and Todd F. DeLuca. (2012) "Use of artificial intelligence to shorten the behavioral diagnosis of autism." *PloS one*, ***7*(8)**: e43855.

[8] Fadi Thabtah. (2017). "Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment." In *Proceedings of the 1st International Conference on Medical and Health Informatics*, pp. 1-6. ACM.

[9] Daniel Bone, Chi-Chun Lee, Matthew P. Black, Marian E. Williams, Sungbok Lee, Pat Levitt, and Shrikanth Narayanan.(2014). "The psychologist as an interlocutor in autism spectrum disorder assessment: Insights from a study of spontaneous prosody." *Journal of Speech, Language, and Hearing Research*, ***57*(4)**: pp.1162-1177.

[10] Fadi Thabtah. (2017) "ASD Tests. A mobile app for ASD screening." www.asdtests.com [accessed December 20th, 2017].

[11] Baihua Li, Arjun Sharma, James Meng, Senthil Purushwalkam, and Emma Gowen. (2017) "Applying machine learning to identify autistic adults using imitation: An exploratory study." *PloS one*, ***12*(8)**: e0182652.

[12] Fadi Fayez Thabtah (2017), "Autistic Spectrum Disorder Screening Data for Adult"., https://archive.ics.uci.edu/ml/machine-learning-databases/00426/.

[13] M. S. Mythili, and AR Mohamed Shanavas. (2014) "A study on Autism spectrum disorders using classification techniques." *International Journal of Soft Computing and Engineering (IJSCE)*, ***4***: 88-91.

[14] J. A. Kosmicki, V. Sochat, M. Duda, and D. P. Wall. (2015) "Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning." *Translational psychiatry*, ***5*(2)**: e514.

[15] Fadi Fayez Thabtah (2017), "Autistic Spectrum Disorder Screening Data for children," https://archive.ics.uci.edu/ml/machine-learning-databases/00419/ ,2017

[16] Fadi Fayez Thabtah (2017), "Autistic Spectrum Disorder Screening Data for Adolescent", https://archive.ics.uci.edu/ml/machine-learning-databases/00420/.

[17] John, George H., and Pat Langley. (1995). "Estimating continuous distributions in Bayesian classifiers." In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence* (pp. 338-345). Morgan Kaufmann Publishers Inc.

[18] Quinlan, J. R. (1993) "Program for machine learning." *C4. 5*.

[19] Keerthi, S. Sathiya, Shirish Krishnaj Shevade, Chiranjib Bhattacharyya, and Karuturi Radha Krishna Murthy. (2001) "Improvements to Platt's

SMO algorithm for SVM classifier design. " *Neural computation*, ***13*****(3):**637-649.
[20] J. C. Platt (1999) "12 fast training of Support Vector Machines using sequential minimal optimization," Adv. Kernel method, pp. 185-208,
[21] Sankar K Pal., and Sushmita Mitra. (1992) "Multilayer perceptron, fuzzy sets, and classification." *IEEE Transactions on neural networks*, ***3*****(5):** 683-697.
[22] David W Aha., Dennis Kibler, and Marc K. Albert. (1991) "Instance-based learning algorithms. *Machine learning*," ***6*****(1)**: 37-66.
[23] Katherine Gotham, Susan Risi, Andrew Pickles, and Catherine Lord. (2007) "The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity." *Journal of autism and developmental disorders* ***37*****(4)**: p.613.
[24] Sarfaraz Masood, Abhinav Rai, Aakash Aggarwal, Mohammad Najmud Doja, and Musheer Ahmad. (2018) "Detecting distraction of drivers using convolutional neural network." *Pattern Recognition Letters*.
[25] Sarfaraz Masood, Adhyan Srivastava, Harish Chandra Thuwal, and Musheer Ahmad. (2018). "Real-time sign language gesture (word) recognition from video sequences using CNN and RNN." In *Intelligent Engineering Informatics* (pp. 623-632). Springer, Singapore.
[26] Sarfaraz Masood, Harish Chandra Thuwal, and Adhyan Srivastava. (2018). "American Sign Language character recognition using convolution neural network. " In *Smart Computing and Informatics* (pp. 403-412). Springer, Singapore.