

AUTISM SPECTRUM DISORDER USING MACHINE LEARNING TECHNIQUES

NUKAMREDDY SRINADHREDDY¹, JARUGUMALLI MADHURI²

**#1Assistant Professor, Department of CSE, PBR Visvodaya Institute of Technology and Science,
Kavali**

**#2Assistant Professor, Department of CSE, PBR Visvodaya Institute of Technology and Science,
Kavali**

Abstract_ Autism Spectrum Disorder (ASD) is a neuro-disorder that has a lifetime impact on a person's relationship and communication with others. Autism can be diagnosed at any age and is referred to as a "behavioural disease" since symptoms typically develop in the first two years of life. According to the ASD problem, it begins in childhood and extends into adolescence and maturity. With the increased use of machine learning techniques in medical diagnosis research, this paper attempts to investigate the possibility of using Nave Bayes, Support Vector Machine, Logistic Regression, KNN, Neural Network, and Convolutional Neural Network for predicting and analysing ASD problems in children, adolescents, and adults. The proposed strategies are tested on three publicly available non-clinically ASD datasets. The first dataset on ASD screening in children has 292 cases and 21 attributes. The second dataset contains 704 cases and 21 attributes linked to ASD screening adult participants. The third dataset contains 104 incidences and 21 attributes related to ASD screening in adolescent individuals. Following the application of various machine learning techniques and the handling of missing values, the results strongly suggest that MLP-based prediction models perform better on all of these datasets, with higher accuracy of 99.53 percent, 98.30 percent, and 96.88 percent for Autistic Spectrum Disorder Screening in Data for Adults, Children, and Adolescents, respectively.

1.INTRODUCTION

The trouble of autism spectrum sickness (ASD) have been mounting unexpectedly presently amongst all a long time of the human population. Early detection of this neurological sickness can noticeably aid in the protection of the subject's intellectual and bodily health. With the upward jostle of utility of desktop learning-based fashions in the predictions of various.

human diseases, their early detection primarily based on a range of fitness and physiological parameter now looks possible. This element inspired us to enlarge activity in the detection and evaluation of ASD illnesses to improve higher therapy methodology. Detection of ASD will become a project as there are numerous different intellectual issues whose few signs are very comparable to these with ASD symptoms, thereby makes this mission a tough one. Autism Spectrum ailment is a trouble that is associated to human Genius development. A individual

who has suffered from the Autism Spectrum Disorder is normally no longer capable to do social interplay and verbal exchange with different men and women [1] or [3]. In this, a person's lifestyles is normally affected for his or her whole lifetime. It is fascinating to recognize that each environmental and genetic elements may additionally flip out to be the inflicting elements for this disease. The signs and symptoms of this trouble may also be began at the age of three years and may additionally proceed for the lifetime. It is no longer feasible to whole deal with the affected person struggling from this disease, then again its consequences can be decreased for some time if the signs and symptoms are early detected. By assuming that human genes are accountable for it, the genuine motives of ASD have now not been diagnosed by using the scientist yet. The human genes have an effect on the improvement through influencing the environment. There is some danger component which influences ASD like as low delivery weight, a sibling with ASD and having ancient parents, etc. Instead of this, there are some social interplay and verbal exchange troubles like as:

- Inappropriate laughing and guffawing
- No sensitivity of ache
- Not capable to make eye contact precise
- No perfect response to sound
- May now not have a desire for cuddling
- Not capable to specific their gestures
- No interplay with others
- Inappropriate objects attachment
- Want to stay on my own
- Using echo phrases etc.

People with ASD additionally have issue with limited hobbies and persistently repetition of behaviors. The following listing affords unique examples of the sorts of behaviors.

- Repeating sure behaviors like repeating phrases or phrases tons time.
- The Person will be upset when a hobbies is going to change.
- Having a little hobby in positive things of the subject like numbers, facts, etc.
- Less touchy than every other character in some instances like light, noise, etc.

Early detection and remedy are most necessary steps to be taken to limit the signs of autism spectrum sickness trouble and to enhance the excellent of lifestyles of ASD struggling people. However, there is no process of clinical take a look at for detection of autism. ASD Symptoms generally diagnosed by means of observation. In Older and young people who go to school, ASD signs are commonly recognized by using their mother and father and teachers. After that ASD signs are evaluated through a different training group of the school. These college crew advised these youngsters go to their fitness care health practitioner for required testing. In adults figuring out ASD signs is very hard than older adolescents and youth due to the fact some signs of ASD can also be overlap with different intellectual fitness disorders. It is effortless to become aware of the behavioral modifications in a baby without difficulty by means of statement due to the fact it can be viewed early in the 6 months of age than Autism particular Genius imaging due to the fact Genius imaging can be figuring out after two years of age

2.LITERATURE SURVEY

Vaishali R, Sasikala R. et al. [3] have proposed a approach to discover Autism with greatest conduct sets. In this work, an ASD analysis dataset with 21 points received from the UCI computer gaining knowledge of repository experimented with swarm Genius primarily based binary firefly function resolution wrapper. The choice speculation of the scan claims that it is viable for a computer getting to know mannequin to reap a higher classification accuracy with minimal characteristic subsets. Using Swarm talent based totally single-objective binary firefly function resolution wrapper it is observed that 10 aspects amongst 21 points of ASD dataset

are adequate to distinguish between ASD and non-ASD patients. The outcomes brought with this strategy justifies the speculation through producing an common accuracy in the vary of 92.12%- 97.95% with foremost characteristic subsets which are about equal to the common accuracy produced with the aid of the complete ASD prognosis dataset. Fadi Thabtah et al. [8] have proposed an ASD screening mannequin the usage of Machine Learning Adaption and DSM-5. A screening device has been used to comprehend one or greater dreams in ASD screening. In this paper, the researcher mentioned the ASD Machine Learning classification with their execs and cons. The researcher tried to spotlight the hassle accompanying with present ASD screening equipment and the consistency of such equipment the usage of the DSM-IV alternatively of the DSM-5 manual.

M. S. Mythili, A. R. Mohamed Shanavas et al. [13] have a find out about on ASD the use of Classification Techniques. The major purpose of this paper used to be to discover the autism trouble and the stages of autism. In this Neural Network, SVM and Fuzzy strategies with WEKA equipment are used to analyze the student's conduct and their social interaction.

J. A. Kosmicki¹, V. Sochat, M. Duda and D.P. Wall Et al. [14] have supposed a looking technique for a least set of characteristics for autism detection. In this, the authors used a computer mastering strategy to consider the medical evaluation of ASD. The ADOS was once carried out on the subset of behaviors of teens primarily based on the autism spectrum. ADOS having 4 modules. In this work, eight extraordinary computing device gaining knowledge of algorithms have been employed, involving stepwise backward characteristic identification on rating sheets from 4540 individuals. It makes use of 9 out of the 28 behaviors from module two and 12 out of the 28 behaviors from module three to become aware of an ASD threat with an ordinary accuracy of 98.27% and 97.66% respectively.

Li B, A. Sharma, J Meng, S. Pustaka, E. Gowen (2017) et al. [11] have used desktop studying classifiers to notice autistic adults via imitation method. The purpose of this find out about was once to check out the crucial trouble associated to discriminative check prerequisites and kinematic parameters. The dataset incorporates sixteen ASC members who have a sequence of hand movements. In these forty kinematic constraints from eight imitation stipulations has been extracted via the use of computer mastering methods. This lookup suggests that for a small sample, there is a feasibility of making use of desktop mastering techniques to analyze high-dimensional facts and the diagnostic classification of autism. The sensitivity quotes accomplished by way of RIPPER which have the aspects Va (87.30 %), CHI (80.95%), IG (80.95%), Correlation (84.13%), CFS (84.13%), and "no function selection"(80.00%) on the AQ-Adolescent dataset. It is evident from the above mentioned area that there is surely a want to discover the opportunity of making use of deep mastering primarily based fashions for the detection of ASD in human population. Most of the work mentioned above use traditional laptop getting to know processes and subsequently are restrained in their performance. In this work, overall performance of numerous computer

mastering fashions have been in contrast to that of the deep getting to know mannequin for this purpose. Separate fashions have been organized for separate populace set (discussed in part below) and in contrast individually.

3.PROPOSED SYSTEM

Autism spectrum disorder (ASD) classification and prediction, thus overcoming the existing problem.

By utilizing Logistic Regression, Support Vector Machine (SVM), Naive bayes algorithms we will make our model in order to increase the performance and accuracy.

This project is developed using CNN, MLP, decission tree, random forest algorithms accuracy of each algorithm is compared and most accurate result is used for prediction

3.1 IMPLEMNETATION

3.1.1 Data Collection:

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

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

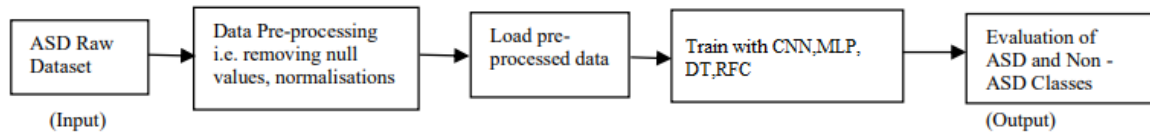


Figure. 1. Steps in the proposed ASD detection solution

3.1.2 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.

3.1.3 Train-Test Split And Model Fitting:

Now, we divide our dataset into training and testing data. Our objective for doing this split is to assess the performance of our model on unseen data and to determine how well our model has generalized on training data. This is followed by a model fitting which is an essential step in the model building process.

4.DATASET

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	age	gender	ethnicity	jundice	autism	relation	contry_of	used_app	age_desc	A1_Score	A2_Score	A3_Score	A4_Score	A5_Score	A6_Score	A7_Score	A8_Score	A9_Score	A10_Score	result	Class
2	6 m	Others	no	no	Parent	Jordan	no	'4-11 year:	1	1	0	0	1	1	0	1	0	0	0	0	5 NO
3	6 m	'Middle Ea	no	no	Parent	Jordan	no	'4-11 year:	1	1	0	0	1	1	0	1	0	0	0	0	5 NO
4	6 m	?	no	no	?	Jordan	yes	'4-11 year:	1	1	0	0	0	1	1	1	1	0	0	0	5 NO
5	5 f	?	yes	no	?	Jordan	no	'4-11 year:	0	1	0	0	1	1	0	0	0	0	1	1	4 NO
6	5 m	Others	yes	no	Parent	'United Stc	no	'4-11 year:	1	1	1	1	1	1	1	1	1	1	1	1	10 YES
7	4 m	?	no	yes	?	Egypt	no	'4-11 year:	0	0	1	0	1	1	0	1	0	1	0	1	5 NO
8	5 m	'White-Eur	no	no	Parent	'United Kir	no	'4-11 year:	1	0	1	1	1	1	1	0	1	0	1	0	7 YES
9	5 f	'Middle Ea	no	no	Parent	Bahrain	no	'4-11 year:	1	1	1	1	1	1	1	1	1	1	0	0	8 YES
10	11 f	'Middle Ea	no	no	Parent	Bahrain	no	'4-11 year:	1	1	1	1	1	1	1	1	0	0	0	0	7 YES
11	11 f	?	no	yes	?	Austria	no	'4-11 year:	0	0	1	1	1	1	0	1	1	0	0	0	5 NO
12	10 m	'White-Eur	yes	no	Self	'United Kir	no	'4-11 year:	1	0	0	0	0	1	1	1	1	1	1	1	7 YES
13	5 f	?	no	no	?	Kuwait	no	'4-11 year:	0	1	0	0	0	1	0	0	0	0	0	1	3 NO
14	4 m	'White-Eur	yes	no	Parent	'United Stc	no	'4-11 year:	0	1	1	1	1	1	1	1	1	1	1	1	9 YES
15	4 f	Black	no	no	Parent	'United Ari	no	'4-11 year:	1	0	0	0	0	0	0	1	0	0	0	0	2 NO
16	6 m	'White-Eur	no	no	Parent	Europe	no	'4-11 year:	1	1	1	1	1	1	1	1	1	1	1	1	10 YES
17	8 m	'White-Eur	no	no	Parent	Malta	no	'4-11 year:	1	1	1	1	1	1	1	1	1	1	1	1	10 YES
18	4 m	'South Asi	no	no	Parent	Bulgaria	no	'4-11 year:	1	1	1	1	1	1	1	0	1	1	1	1	9 YES
19	7 m	Others	no	no	Parent	'United Stc	no	'4-11 year:	0	0	0	0	0	0	1	0	0	0	0	0	1 NO
20	11 m	'White-Eur	no	yes	Parent	'United Stc	no	'4-11 year:	1	0	1	1	1	1	0	1	1	1	1	1	8 YES
21	5 m	?	no	no	?	Egypt	no	'4-11 year:	1	1	1	1	1	1	1	0	1	0	1	0	8 YES
22	5 m	'White-Eur	yes	no	Parent	'South Afri	no	'4-11 year:	1	1	1	1	1	1	1	1	0	1	0	1	8 YES
23	9 f	?	no	no	?	Egypt	no	'4-11 year:	0	0	1	1	0	1	0	1	1	1	0	0	5 NO
24	4 m	Asian	no	no	Parent	India	no	'4-11 year:	1	1	0	1	0	0	0	0	0	0	0	0	3 NO
25	6 f	'South Asi	no	no	Parent	India	no	'4-11 year:	1	0	1	1	0	1	0	0	0	1	0	0	5 NO
26	11 m	?	no	no	?	Egypt	no	'4-11 year:	1	0	1	1	1	1	1	0	1	1	1	1	8 YES
27	6 m	'White-Eur	no	yes	Relative	'United Kir	no	'4-11 year:	0	0	1	1	1	1	0	1	1	1	1	0	6 NO

Fig 2:Dataset Information

5.RESULTS AND DISCUSSION

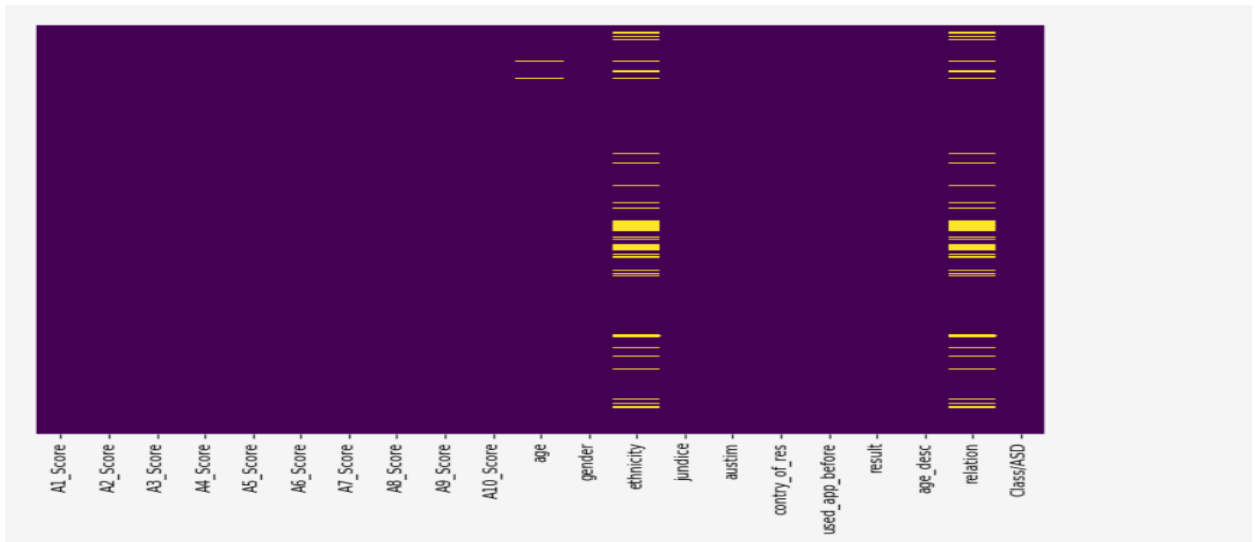
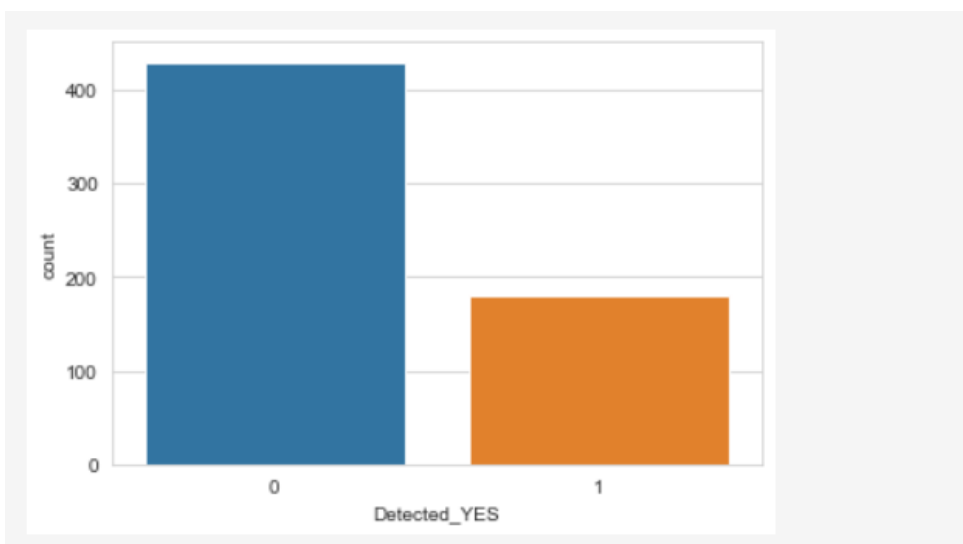
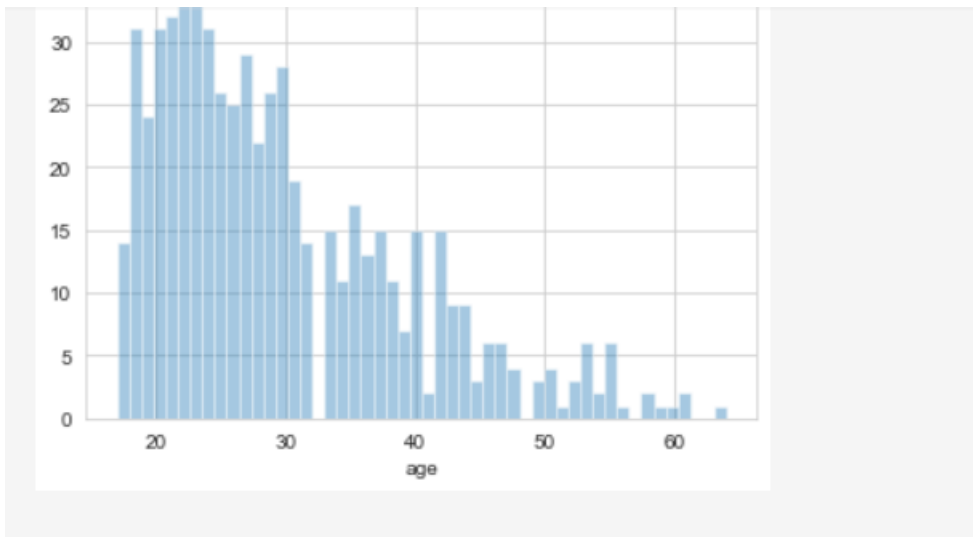
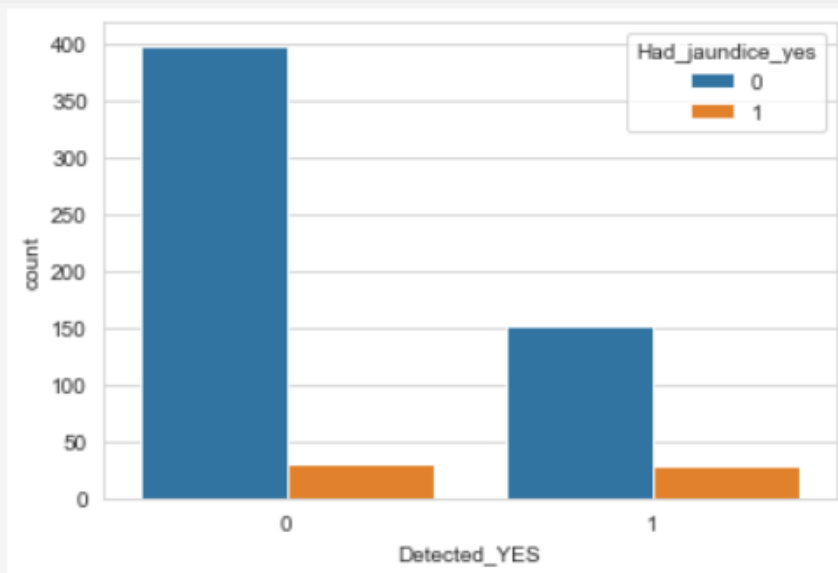


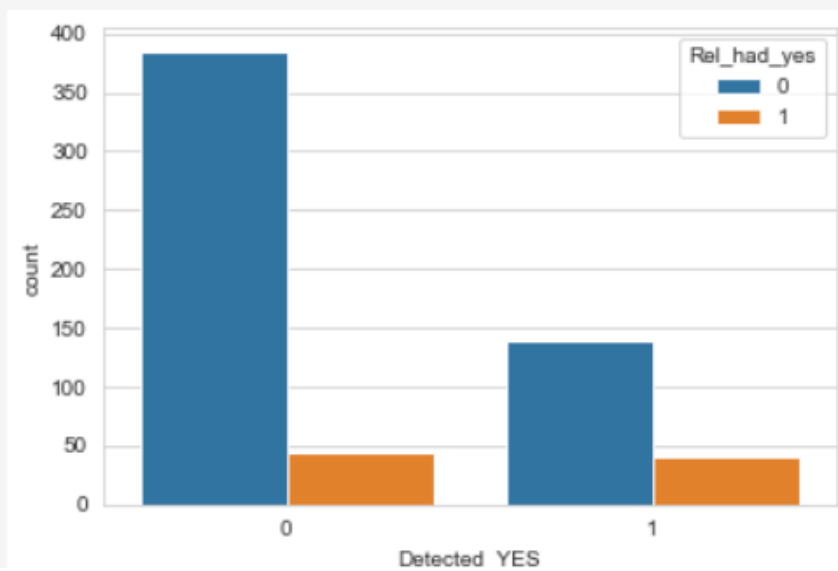
Fig 3: Feature Analysis





```
sns.countplot(x='Detected_YES',hue="Rel_had_yes",data=data_featured)
```

<matplotlib.axes._subplots.AxesSubplot at 0x1406f9c2588>



	precision	recall	f1-score	support
0	1.00	1.00	1.00	132
1	1.00	1.00	1.00	51
accuracy			1.00	183
macro avg	1.00	1.00	1.00	183
weighted avg	1.00	1.00	1.00	183

Fig 5:graphs

6.CONCLUSION

The detection of autism spectrum condition was tried in this study utilising a variety of machine learning and deep learning techniques. Various performance evaluation measures were utilised to assess the performance of the models implemented for ASD detection using non-clinical datasets from three age groups, namely children, adolescents, and adults. When compared to another recent study [3] on same subject, the MLP classifier performed better than the SVM with all of its feature properties included after handling missing values. After accounting for missing values, both the RFC and MLP-based models in this study achieve the same prediction accuracy of around 98.30 percent for the ASD Child dataset. However, for the remaining two datasets, the MLP-based model achieved the highest accuracy result compared to all other considered model building techniques. These findings strongly suggest that an MLP-based model can be implemented for detection of Autism Spectrum Disorder instead of the other conventional machine learning classifier suggested in previous studies.

REFERENCE

- [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 observationbased 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.