2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)
University of Malaya, Kuala Lumpur, Malaysia. Sep 24-26, 2021

Autism Screening Disorder : Early Prediction

Tamanna Yesmin Rashme
Dept. of Computer Science & Engineering
Uttara University
Dhaka, Bangladesh
tamanna1991.cse@gmail.com

Linta Islam
Dept. of Computer Science & Engineering
Jagannath University
Dhaka, Bangladesh
lislamcsejnu@gmail.com

Ayesha Aziz Prova
Dept. of Computer Science & Engineering
Central Women's University
Dhaka, Bangladesh
ayeshaaziz.csejnu@gmail.com

Sohely Jahan
Dept. of Computer Science & Engineering
University of Brishal
Dhaka, Bangladesh
sohely.cse@gmail.com

2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON) | 978-1-7281-9951-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/GUCON50781.2021.9573547

***Abstract*— Autism Spectrum Disorder (ASD), which is a developmental disability due to neuro-development disorder. This disorder is a brain growth disorder that affects how an individual perceives and interacts with others, resulting in social contact and communication difficulties. Though this disorder is the result of genetic or in nature but earlier detection allows earlier intervention, which is generally more effective. Based on ASD history research, This ASD problem begins with childhood and as time goes this disorder progresses into adolescence and adulthood. Thus, in this paper, a model has been proposed for the early prediction and analysis of ASD problems in children, adolescents, and adults. Principal Component Analysis (PCA) is used for the feature reduction method and then various machine learning algorithms have been applied to three ASD datasets relating to Children, adolescents, and adults. 10-fold cross-validation is used to evaluate the performance of applied machine learning algorithms. The result produces from our proposed model shows that Logistic Regression shows outperform comparing to all other benchmark machine learning algorithms in terms of accuracy, precision, recall, and F1-score during the prediction of ASD cases.**

***Keywords*— ASD, PCA, MACHINE LEARNING, LOGISTIC REGRESSION**

I. Introduction

Autism Spectrum Disorder (ASD) is referred to as neurodevelopmental inability that states as slow growth of expressive capacity, repetitious actions, and lack of passion. ASD is referred to as disabilities that can be improved with early detection but cannot be cured [1]. People with ASD can speak, contact, behave, and learn in ways that are different from most other people, despite the fact that their appearance is often unremarkable. People with ASD vary from talented to seriously disable in their understanding, reasoning, and problem-solving skills. In daily life, some people with ASD need a great deal of assistance where some others may need less [2]. Among all cultural, national, and economic groups, ASD can happen. Early treatment and facilities enhance the symptoms and abilities of an ASD person even though this is a lifelong syndrome. Minshew and William defined autism as polygenetic development and neurobiological disorder that can identify by conventional action with a lack of social reciprocity. There is some potential factor that impacts ASD, such as the prematurity, the ASD sibling, the elderly, etc[3]. Also some types of social dealings and communication problems like failure to make eye contact clearly, laughing and giggling improperly, lack of interest in interacting with others, for any kind of sound not appropriate answer, etc. also responsible that impact ASD. The aspects both environmental and genetic may play the roles in this disease. For a person who suffers from autism, the symptoms may be started from his/her early age as two or three years and can last a lifetime because this is not curable [4].

Now pre-screening technology has been promoted as this issue has become more widespread. Symptoms of ASD are normally detected by observation, though no medical evaluation is available for the detection of autism [5]. ASD signs are typically detected by parents and teachers in older and adolescent students who attend school. Also, ASD signs are typically detected by the school's child care team by recommending the children who go to their doctor for the necessary testing. As people grow older with time, the symptoms of ASD overlap with other mental health conditions, and then it becomes difficult to identify ASD in adults [6]. For speeding the diagnosis times in Autism cases, it's needed to early detection of this disorder. Machine Learning techniques have recently developed several applications to support artificial intelligence at the human level. Machine learning technology allows building a model to predict and speeding up complex and time-consuming diagnostic and treatment procedures.

In this paper, for early detection of autism, a model has been proposed using three ASD screening datasets of children, adolescents, and adults from the UCI machine learning repository. These datasets contain ten behavioral features (AQ-10-Adult) along with ten individual characteristics to prove efficacy in detecting ASD cases. Several machine learning algorithms are evaluated on datasets for assessing the effectiveness for the prediction of the result of ASD treatment.

This paper is presented as follows. The second section discusses literature review of this work. In third section the proposed methodology of this paper is discussed. The experimental outcomes with discussion where the result of proposed model comparison has showed with many existence works are detailed in section four. Finally, we conclude the paper in section five.

II. Literature Review

Suman Raj et.al [7] has used various machine learning algorithms and deep learning techniques for the detection of Autism Spectrum Disorder. In this work, to detect ASD three different non-clinically dataset was used based on age groups; child, adolescent, adult. 6different algorithms were employed for detection purposes and Cross-validation was applied to the training dataset. Among them, the CNN classifier gave the best result after handling missing values. Also, the SVM classifier gave the same accuracy as CNN for child ASD prediction. CNN achieved more than 98% accuracy rate in the prediction of autism spectrum disorder on different datasets.

D. Varshini G and Chinnaiyan R [8] proposed a system using optimized machine learning algorithms to identify early traits of autism among adults and toddlers. In this work, 3 different algorithms are used for prediction purposes; Logistic regression, K-Nearest Neighbor and Random Forest. Missing data has been checked by heatmap for both datasets. Using dataset attributes different types of the graph have plotted to select required attributes of features for the prediction of ASD traits in both the datasets. The evaluation metrics F1 score and precision are estimated to categorize which classifier is more efficient for predicting ASD traits in the given data set. The accuracy rate achieved by KNN, Logistic regression, and the random forest is 69.2%, 68.601%, and 67.78% respectively. Out of three classifiers, KNN provided the best accuracy comparing the two others.

The research conducted by Abdullah et al.[9] used Chi-square and Least Absolute Shrinkage and Selection Operator (LASSO) for feature selection purposes. And important 13 features have selected from the Chi-square method have run on three different machine learning algorithms name: Logistic regression, Random Forest, and K-nearest neighbor. The researchers used an adult Dataset from the ASD screening dataset to perform their experiment. Dataset used in this work contained 13.5% missing data. Using pandas commands missing data were removed by dropping the empty column. During the evaluation of algorithms performance instead of using the common split method K fold cross-validation has been used to split the dataset into 5 pieces and run the algorithms 5 times. As a result, Logistic regression gives the highest accuracy with more than 97% for the imminent recognition of ASD for adults.

A time conserving approach to display screen possible ASD entities for self-screening has been proposed in [10]. The study was conducted on three different optimal models based on the three autism screening datasets: child, adolescent, and adult. For feature selection, Lasso regression was used where 14 important features were selected from the adolescent dataset, 18 variables from the adult dataset. In this work Five different machine learning classification models are used: Logistic Regression(LR), Decision Tree, Support Vector Classifier, Random Forest(RF), Artificial Neural Network (ANN) on these three datasets to calculate different evaluation metrics: Accuracy, recall, specificity, ROC and AUC curve and F1 score. Among these classifiers the outcome of the decision tree was poor. For the adolescent autism screening dataset, LR gives optimal outcome where ANN and Support vector classifier gives the best outcome on adult and child autism dataset respectively.

A framework based on machine learning algorithm for predictive analysis of autism has proposed on [11] . This paper used datasets related to the autism screening of adults and adolescent where ten behavioral features has taken by using AQ screening approach also with individual's physiognomies. Information gain(IG) and chi square these two feature selection methods used for depth analysis of dataset's features. Among 10 features obtained by using AQ screening approaches, four important features (A4, A5, A6, A9) and five important features (A3, A4, A5, A6, A9) has selected from adult screening dataset and adolescent screening dataset respectively by using those two feature selection methods. After feature selection logistic regression algorithm applied to the dataset for evaluation purpose in terms of sensitivity, accuracy and specificity.

III. PROPOSED METHODOLOGY

The objective of this paper is to early prediction of ASD problem among children, adolescent and adults. For this purpose, a model is proposed that is shown in "Fig. 1", which involves data pre-processing, splitting data and evaluation of results using different machine learning algorithms.

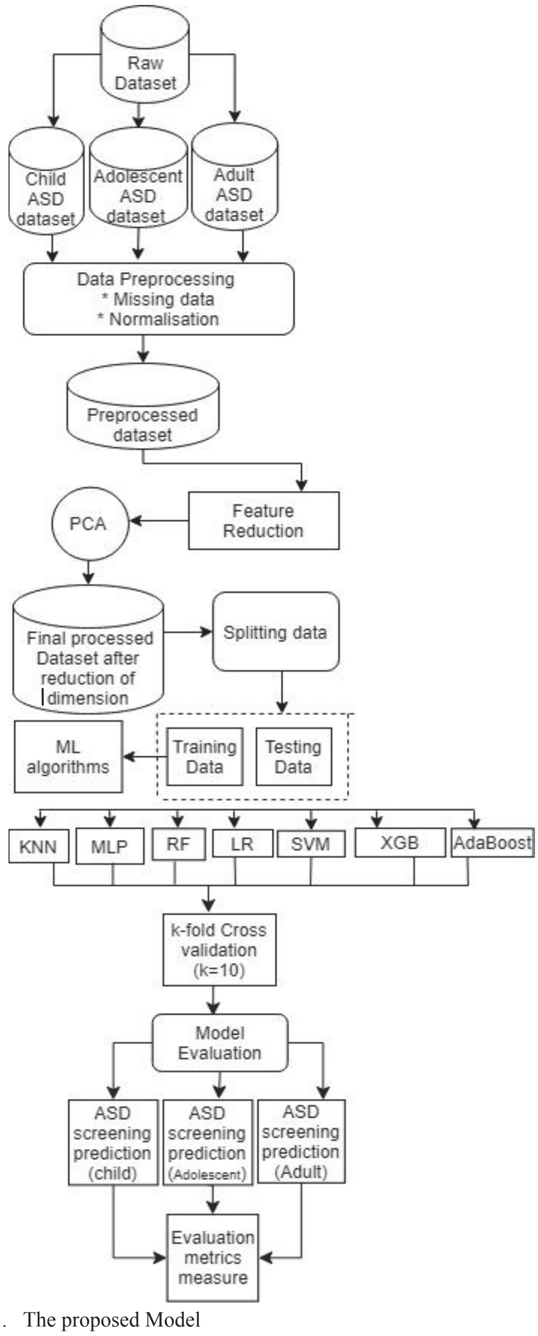

Fig. 1. The proposed Model

A. Dataset Description

In this paper, three types of advanced level ASD screening datasets: Adult, Adolescent, and Children have been used for research purposes. These datasets have been collected from the UCI repository which is freely accessible [12] [13] [14]. ASD screening test data for 704 adults, 104 adolescents, and 292 children were used in this advanced level dataset. All the datasets contain 21 attributes includes demographics of test takers such as Age, Gender, Ethnicity, etc. There are 10 questions (A1 to A10) represent as attributes in all datasets that were answered by the survey participants in the screening test. The dataset attributes and their values are presented in Table 1.

B. Data Preprocessing

Preprocessing means the conversion applied before algorithms are fed on data. Different data preprocessing techniques are being applied for incomplete and inconsistent data like outlier finding, missing values, discretization of data, null values, etc. In our work, two attributes Ethnicity and relation contain missing values for all datasets wherein the Children dataset the rate of missing values in ethnicity is 15% and relation is 15%, Adolescent dataset the rate of missing values in ethnicity is 6% and relation is 6%, an Adult dataset the rate of missing values in ethnicity is 13% and relation is 13%. The problems of missing values in these datasets have been replaced by the mean of the data values of that particular data column.

C. Training & Testing Dataset

Data splitting is the process of splitting data into a train, test, or validation set and it is the technique that finds the model hyper-parameter and also approximating the simplification performance. In this paper, all datasets have been split into three ratios to compare the values of accuracy, precision, recall, and f1-score among different machine learning algorithms. "Fig. 2" shows data splits ratios used in this work.

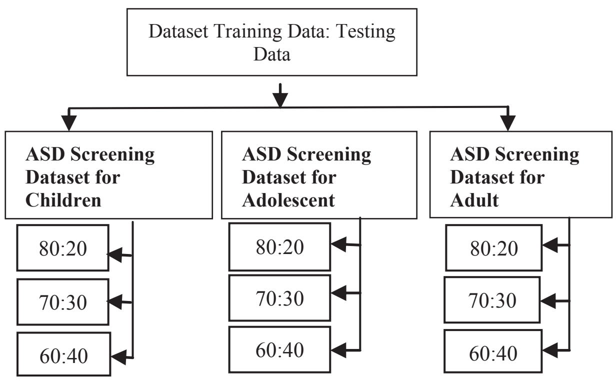

Fig. 2. Data Splitting Ratios

D. Feature Extraction

Feature reduction is the method of reducing the number of features without losing the irrelevant and important information for data processing. Many experiments have been clarified that the result of feature extraction method can increase the predictive accuracy as well as improve interpretation and generality of the model that has been developed. Feature extraction is useful when a more concise statement needs to produce from a data that has lot of dimension.

1) Principal Component Analysis(PCA)

In machine learning, Principle Components Analysis (PCA) is an unsupervised approach for dimensionality reduction. PCA is used to reduce the dimensionality of a dataset that includes a vast number of interrelated attributes or variables while preserving as much of the variance present in the data set as possible [15]. This is accomplished by converting to a new set of uncorrelated variables known as principal components (PCs), which are ordered so that the first few preserve the majority of the variance found in all of the initial variables.It is also recognized as a general factor analysis where regression decides the optimal fit.

E. K-Fold Cross-validation

Cross-validation is a method for assessing a machine learning model. Cross-validation is widely employed in machine learning tasks mostly where over-fitting needs to be mitigated. There are many types of cross-validation where some are used only for general purpose and rest others for theory. K-fold cross-validation is the simple form of cross-validation. In k-fold cross-validation, firstly k-equally size of data segments or folds is divided [16]. Successively during k iterations of training and validation, validation is performed within each iteration when an altered fold of the data is held out, and at the same time remaining $k - 1$ fold is used for learning. In this experiment for each machine learning algorithm, 10-fold cross-validation was employed for comparing and picking an appropriate algorithm for the autism prediction where the dataset will split into 10 pieces and each algorithm will run 10 times.

F. Machine Learning Algorithms

1) Logistic Resgression

Logistic regression is an algorithm used for classification problems for machine learning. Logistic regression is also linear classifier, where linear function $(\mathbf{x}) = a_0 + a_1x_1 + \cdots + a_rx_r$, is referred as "logit". Here,The variables a_0, a_1, …, a_r are the the regression coefficients estimators , which are also called the predicted weights or only coefficients.

2) K-Nearest Neighbor

KNN is a machine earning technique which is used for both regression and classification. The adjacency matrix is generated firstly where according to KNN principle edges are set and secondly, the edge weights are calculated. In order to identify the similarities KNN use with distance where the higher the K factor, the more robust the results get.

3) Multilayer Perception(MLP)

Multilayer Perceptron (MLP): A Multilayer Perceptron is referred as supervised learning method used to classify the dataset which are not linearly separable. MLP consists of multiple hidden layers between input layer and output layer. For training data, its uses backpropagation technique.

TABLE I. DATA SET DESCRIPTION

Attributes/ Features	Type	Narrative
Age	Number	Age in years
Gender	String	Male or Female
Ethnicity	String	List of common ethnicities in text format
Born with jaundice	Boolean (yes or no)	In the case of born with jaundice
Family member with PDD	Boolean (yes or no)	Whether any immediate family member has a PDD
Who is completing the test	String	Parent, self, caregiver, clinician , medical staff etc.
Country of residence	String	Countries list is shown in the text form
Used the screening app before	Boolean (yes or no)	Whether the user has used a screening app
Method of Screening Type	Integer (0,1,2,3)	Depend on age category screening methods types chosen (0=toddler, 1=child, 2= adolescent, 3= adult)
A1_Score	Binary (0, 1)	The question's response code relying on the screening process used
A2_Score	Binary (0, 1)	The question's response code relying on the screening process used
A3_Score	Binary (0, 1)	The question's response code relying on the screening process used
A4_Score	Binary (0, 1)	The question's response code relying on the screening process used
A5_Score	Binary (0, 1)	The question's response code relying on the screening process used
A6_Score	Binary (0, 1)	The question's response code relying on the screening process used
A7_Score	Binary (0, 1)	The question's response code relying on the screening process used
A8_Score	Binary (0, 1)	The question's response code relying on the screening process used
A9_Score	Binary (0, 1)	The question's response code relying on the screening process used
A10_Score	Binary (0, 1)	The question's response code relying on the screening process used
Class/ASD	Yes/No	The final score obtained based on the scoring algorithm of the screening method used. This was computed in an automated manner

4) *AdaBoost Classifier*

AdaBoost algorithm is a boosting technique that is used as an Ensemble method for machine learning. This algorithm also known as Adaptive boosting. For each instance weights are reallocated with higher weight to improperly classified instances. Maintaining a distribution or sequence of weights over the training set is one of the key concepts of the AdaBoost algorithm.

5) *Random Forest*

Random forest is an ensemble method that is used to explain machine learning problems like classification and regression. This algorithm employs a method known as bootstrap aggregation to generate multiple decision trees. It produces the class that the maximum of the decision trees estimated in the forest for classification. It produces the class with the mean projections of the different trees for regression.

$$\text{Entropy} = \sum_{k=1}^{p} x_k \log_2 (x_k)$$

6) *Support Vector Machine*

This algorithm is used to solve a two-class problem in which samples can be separated using a hyper-plane. This is an imbalance classifier that is entirely influenced by support vectors. SVM can construct the hyper-plane in an iterative manner with the intention of limiting the error. The aim of SVM is to divide datasets into groups in order to find the most extreme peripheral hyper-plane.

7) *Extreme Gradient Boosting Classifiers (XGB)*

The XGBoost is an advanced and optimized version of gradient boosting decision tree based algorithm. Using various regularization methods, XGBoost improves efficiency by managing the complexity of the trees. Shrinkage proposed by Friedman [15] is implemented by XGBoost. The shrinkage variables measure the weight value of a feature by using a factor η. This factor η is also known as learning rate.

IV. RESULT ANALYSIS & DISCUSSION

This part shows the experimental results applied on three datasets by using seven well-known machine learning algorithms and 10-fold cross-validation. Also comparing the proposed model with other research papers have shown here.

A. *Experimental & Result analysis*

A quad core-15 system with 4 GB RAM, Scipy, pandas tool, Ipython was used in Colaboratory web application environment. The experimental analysis takes place in three levels. Different evaluation metrics such as confusion matrix, precision, accuracy, recall F1-score etc. are used to evaluate the model performance and to find out the best algorithm. A confusion matrix is a table with dimensions: "Actual" and "Predicted," as well as "True Positives (TP)," "True Negatives (TN)," "False Positives (FP)," and "False Negatives (FN)" on both dimensions.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad \text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN} \qquad \text{F1-score} = 2 * \frac{\text{Precision} * \text{recall}}{\text{Precision} + \text{recall}}$$

Using proposed model the overall performance measures of all machine learning classifiers with 10-fold cross-validation in all ASD screening dataset viz. Children, adolescent and adults have been shown in Table [II] [III] [IV] respectively. On Children ASD screening dataset discovered the range of accuracy is (91.78% to 100.0%) and to predict autism Logistic Regression, Multilayer perception, AdaBoost classifier, and XGB classifier performance best. As we split our dataset into three forms to conduct the analysis, these four algorithms produce 100% accuracy, precision, recall, and f1-score to predict the disorder. Similarly, in adolescent ASD screening dataset the range of accuracy is (89.25% to 100.0%) and Logistic Regression algorithm performs best with 100% for accuracy, precision, recall and f1-score comparing to others algorithms. But dataset that split into 60:40 ratio, logistic regression showed less accuracy rate, precision rate, and F1-score comparing to other split ratios. others algorithms give 90% above detection rate for all performance metrics .Similarly two other datasets, on Adult ASD screening dataset the range of accuracy is (98.16% to 100.0%) and also here logistic Regression algorithm performs best with 100% for accuracy, precision, recall and f1-score comparing to others algorithms. All other algorithms show performance metrics above 95% to detect this disorder.

TABLE II. RESULT ANALYSIS TABLE OF CHILDREN AUTISM DATASET

ML Algorithms	Data Split	Recall	Precision	Accuracy	F1 Score
LR	80:20	100.0 %	100.0 %	100.0 %	100.0 %
	70:30	100.0 %	100.0 %	100.0 %	100.0 %
	60:40	100.0 %	100.0 %	100.0 %	100.0 %
RF	80:20	90.95%	92.97%	94.87%	93.45%
	70:30	100.0%	98.00%	98.96%	98.69%
	60:40	90.7619	94.6891	91.78%	92.7624
SVM	80:20	100.0 %	98.08%	98.97%	98.86%
	70:30	100.0 %	98.00%	98.96%	98.85%
	60:40	100.0 %	98.05%	98.95%	98.88%
KNN	80:20	98.57%	96.56%	97.58%	97.68%
	70:30	98.57%	96.55%	97.59%	97.64%
	60:40	98.56%	96.52%	97.57%	97.66%
MLP	80:20	100.0 %	100.0 %	100.0 %	100.0 %
	70:30	100.0 %	100.0 %	100.0 %	100.0 %
	60:40	100.0 %	100.0 %	100.0 %	100.0 %
AdaBoost	80:20	100.0 %	100.0 %	100.0 %	100.0 %
	70:30	100.0 %	100.0 %	100.0 %	100.0 %
	60:40	100.0 %	100.0 %	100.0 %	100.0 %
XGB	80:20	100.0 %	100.0 %	100.0 %	100.0 %
	70:30	100.0 %	100.0 %	100.0 %	100.0 %
	60:40	100.0 %	100.0 %	100.0 %	100.0 %

TABLE III. RESULT ANALYSIS TABLE OF ADOLESCENT AUTISM DATASET

ML Algorithms	Data Split	Recall	Precision	Accuracy	F1 Score
LR	80:20	100.0 %	100.0 %	100.0 %	100.0 %
	70:30	100.0%	100.0%	100.0%	100.0%
	60:40	100 .0%	97.55 %	98.18%	98.04%
RF	80:20	93.81%	90.93%	92.45%	90.74%
	70:30	91.90%	98.00%	90.54%	92.94%
	60:40	92.1428	94.285	97.09%	94.69%
SVM	80:20	100.0 %	92.38%	94.09%	92.88%
	70:30	100.0 %	92.38%	94.08%	92.86%
	60:40	100.0 %	92.35%	94.05%	92.87%
KNN	80:20	98.33%	87.32%	89.27%	87.42%
	70:30	98.33%	87.31%	89.25%	87.41%
	60:40	98.32%	87.30%	89.27%	87.40%
MLP	80:20	98.34%	100.0 %	97.09%	96.95%
	70:30	98.33%	97.14%	97.08%	96.86%
	60:40	98.33%	97.13%	97.07%	96.84%
AdaBoost	80:20	98.43	97.14%	98.09%	96.95%
	70:30	98.43%	97.04%	98.00%	97.89%
	60:40	98.39%	97.08%	98.00%	96.94%
XGB	80:20	98.33%	97.16%	97.09%	94.75%
	70:30	98.33%	97.14%	97.00%	96.95%
	60:40	98.23%	97.16%	97.01%	96.54%

TABLE IV. RESULT ANALYSIS TABLE OF ADULT AUTISM DATASET

ML Algorithms	Data Split	Recall	Precision	Accuracy	F1 Score
LR	80:20	100.0%	100.0%	100.0%	100.0%
	70:30	100.0 %	100.0 %	100.0 %	100.0 %
	60:40	100.0 %	100.0 %	100.0 %	100.0 %
RF	80:20	98.39%	97.47%	99.01%	97.42%
	70:30	95.263	98.54%	98.16%	97.81%
	60:40	95.23%	91.90%	98.58%	97.40%
SVM	80:20	97.34%	99.50%	99.15%	98.89%
	70:30	97.30%	99.49%	99.13%	98.85%
	60:40	97.30%	99.45%	99.10%	98.82%
KNN	80:20	88.92%	98.02%	96.45%	95.34%
	70:30	88.92%	98.02%	96.45%	95.34%
	60:40	88.92%	98.02%	96.45%	95.34%
MLP	80:20	98.94%	100.0 %	99.58%	99.27%
	70:30	99.45%	100.0 %	99.71%	98.93%
	60:40	98.92%	100.0 %	99.55%	99.63%
AdaBoost	80:20	98.39%	94.83%	97.58%	95.18%
	70:30	98.38%	100.0 %	99.71%	99.81%
	60:40	99.44%	100.0 %	99.73%	99.63%
XGB	80:20	98.92%	100.0 %	99.73%	99.63%
	70:30	98.91%	100.0 %	99.71%	99.81%
	60:40	99.44%	100.0 %	99.86%	99.44%

B. *Discussion*

In this section, we compare the performance of our proposed method with the existing research works.

TABLE V. COMPARISON TABLE OF ADOLESCENT AUTISM PREDICTION.

Approach	Year	ASD Dataset Types	Accuracy (%)
[7]	2020	Child	98.30
		Adult	96.88
		Adolescent	99.53
[10]	2020	Child,	96.78,
		Adolescent,	92.00
		Adult	96.78
[11]	2019	Adolescent,	99.91
		Adult	97.58
[17]	2018	Child	97.95
[8]	2020	Adults	69.2
Proposed Model		Child,	100
		Adolescent,	100
		Adult	100

From Table V, we can see that best accuracy for the prediction of ASD has obtained from our proposed model is 100% using logistic Regression when comparing the result with recent works. The recent works has achieved the highest accuracy for child dataset below 98.3%, adult dataset below 69.2% whereas below 99.91% achieved from adolescent dataset. In our proposed model feature reduction method, PCA is used while other models used feature selection for analysis. For this reason, we achieved better accuracy than the other reference models. Fig. 3 shows the comparison graph of proposal model with existing benchmark research work discussed above. Here Blue lines shows accuracy of child ASD prediction, Green lines *shows* accuracy of adolescent ASD prediction and Lavender lines *shows* accuracy of adults ASD prediction, among different approach with our proposed model.

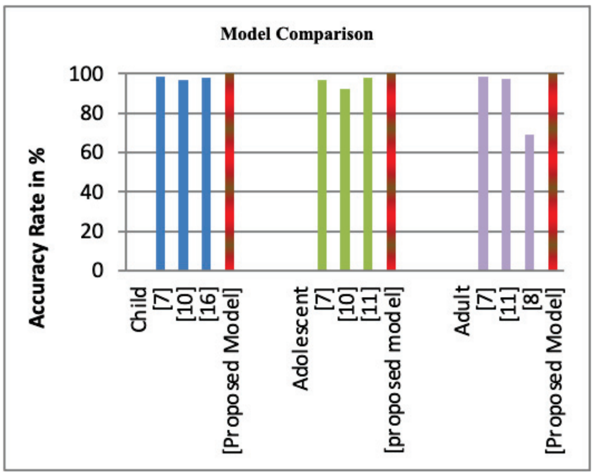

Fig. 3. Model Comparison graph on ASD Screening Data with proposed model

V. Conclusion

Early identification of ASD can be helped by recruiting appropriate treatment and prescription or medication required for both a patient and health services provider, it is possible to minimize the long-term costs of delayed diagnosis. In this paper, we have analyzed three ASD datasets for Children, Adolescents, and Adults aimed at early prediction of autism using various machine learning algorithms. To find the efficiency performance of the classification algorithms on these datasets, 10-fold cross-validation was also used after each algorithm applied on datasets. The outcome of this paper shows that Logistic regression has the highest score in all datasets to detect this Autism Spectrum Disorder. Also, our proposed model shows better results in this problem when compared to existing work.

References

[1] Johnny L. Matson, Jonathan Wilkins, Melissa González, "Early identification and diagnosis in autism spectrum disorders in young children and infants: How early is too early?" Research in Autism Spectrum Disorders ,Volume 2, Issue 1,2008.

[2] K. S. Omar, P. Mondal, N. S. Khan, M. R. K. Rizvi and M. N. Islam, "A Machine Learning Approach to Predict Autism Spectrum Disorder," 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1-6, doi: 10.1109/ECACE.2019.8679454.

[3] Shen, Mark D, and Joseph Piven. "Brain and behavior development in autism from birth through infancy." Dialogues in clinical neuroscience vol. 19,4 (2017): 325-333. doi:10.31887/DCNS.2017.19.4/mshen

[4] Helt, M., Kelley, E., Kinsbourne, M. et al. "Can Children with Autism Recover? If So, How?." Neuropsychol Rev 18, 339–366 (2008). https://doi.org/10.1007/s11065-008-9075-9

[5] Thabtah, Fadi; Peebles, David. 2019. "Early Autism Screening: A Comprehensive Review" Int. J. Environ. Res. Public Health 16, no. 18: 3502. https://doi.org/10.3390/ijerph16183502

[6] Küpper, C., Stroth, S., Wolff, N. et al. "Identifying predictive features of autism spectrum disorders in a clinical sample of adolescents and adults using machine learning". Sci Rep 10, 4805 (2020). https://doi.org/10.1038/s41598-020-61607-w.

[7] Suman Raj, Sarfaraz Masood,"Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques", Procedia Computer Science,Volume 167,2020.ISSN 1877-0509,https://doi.org/10.1016/j.procs.2020.03.399.

[8] Devika Varshini G, Chinnaiyan R. "Optimized Machine Learning Classification Approaches for Prediction of Autism Spectrum Disorder". Ann Autism Dev Disord. 2020; 1(1): 1001.

[9] Abdullah, Azian & Rijal, Saroja & Dash, Satya. (2019). "Evaluation on Machine Learning Algorithms for Classification of Autism Spectrum Disorder (ASD)." Journal of Physics Conference Series. 1372. 012052. 10.1088/1742-6596/1372/1/012052.

[10] A. Baranwal and M. Vanitha, "Autistic Spectrum Disorder Screening: Prediction with Machine Learning Models," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-7, doi: 10.1109/ic-ETITE47903.2020.186.

[11] Thabtah, Fadi & Abdelhamid, Neda & Peebles, David. (2019). "A machine learning autism classification based on logistic regression analysis. Health Information Science and Systems."7. 10.1007/s13755-019-0073-5.

[12] Fadi Thabtah, "Autistic Screening Adult Data Set," UCI machine learning repository, 2017. [Online].Available: https://archive.ics.uci.edu/ml.

[13] Fadi Thabtah, "Autistic Spectrum Disorder Screening Data for Adlolescent Data Set," *UCI machine learning repository*, 2017. [Online].Available: https://archive.ics.uci.edu/ml .

[14] Fadi Thabtah, "Autistic Spectrum Disorder Screening Data for Children Data Set," *UCI machine learning repository*, 2017. [Online].Available: https://archive.ics.uci.edu/ml.

[15] Jerome H. Friedman. "Greedy Function Approximation: A Gradient Boosting Machine". In: The Annals of Statistics 29.5 (Oct. 2001), pp. 1189–1232. ISSN: 00905364.

[16] Anand Singh Rajawat, Romil Rawat, Kanishk Barhanpurkar, Rabindra Nath Shaw, AnkushGhosh, "Depression detection for elderly people using AI robotic systems leveraging the Nelder–Mead Method", Artificial Intelligence for Future Generation Robotics (2021)Pages 55-70. https://doi.org/10.1016/B978-0-323-85498-6.00006-X

[17] 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