2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST) | 978-0-7381-3042-2/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICREST51555.2021.9331013

Improved Machine Learning based Classification Model for Early Autism Detection

Tania Akter
Dept. of CSE Jahangirnagar University Dhaka, Bangladesh
shilpy339cse@gmail.com

Md. Imran Khan
Dept. of CSE
Gono Bishwabidyalay Dhaka, Bangladesh
imrankhan770707@gmail.com

Mohammad Hanif Ali
Dept. of CSE
Jahangirnagar University
Dhaka, Bangladesh
hanif.ju03@yahoo.com

Md. Shahriare Satu
Dept. of MIS
NSTU
Noakhali, Bangladesh
shahriar.setu@gmail.com

Md. Jamal Uddin
Dept. of CSE
BSMRSTU
Gopalgonj, Bangladesh
jamal.bsmrstu@gmail.com

Mohammad Ali Moni
UNSW Digital Health
Uiversity of New South Wales
Sydney, Australia
m.moni@unsw.edu.au

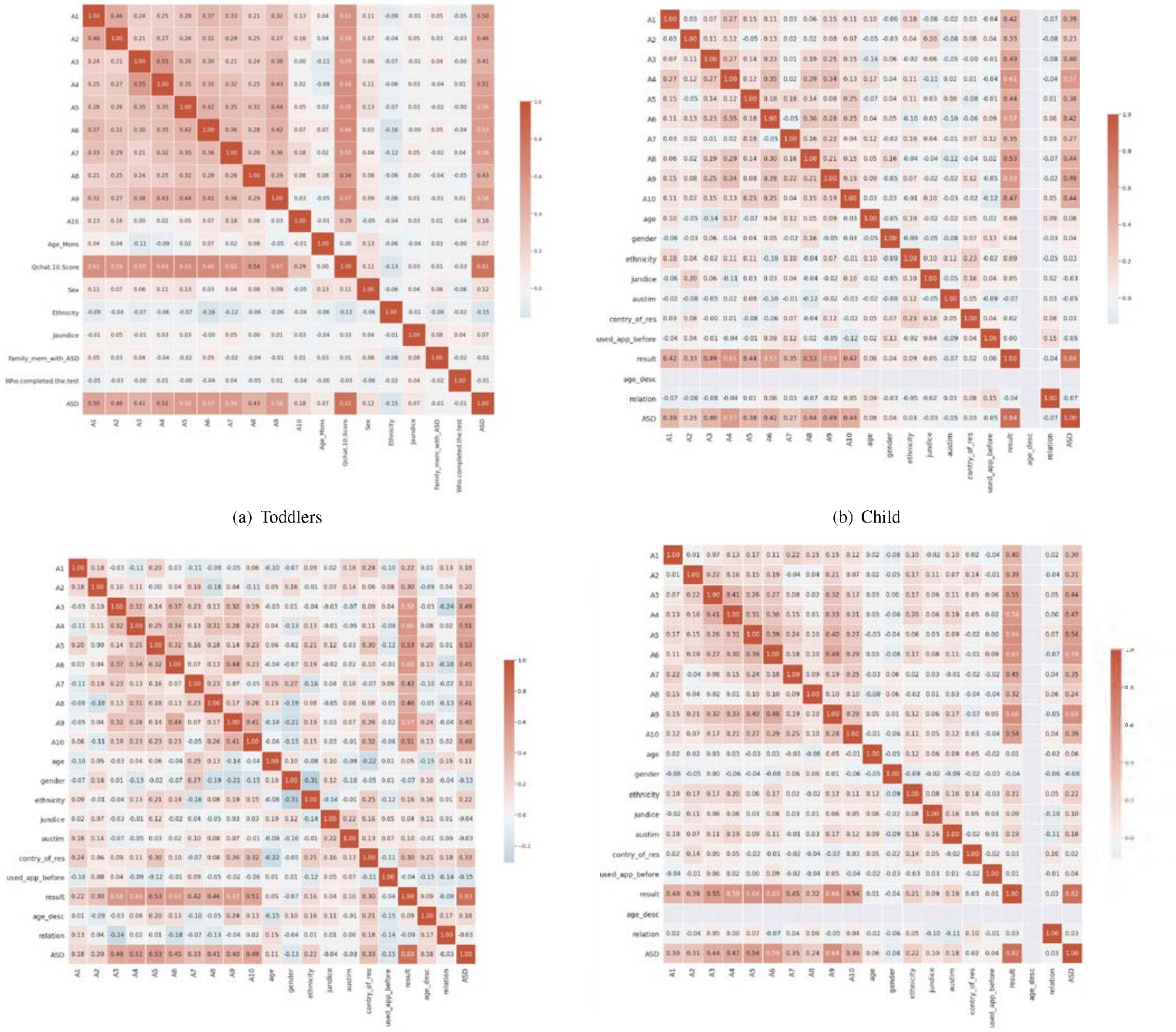
***Abstract*—Autism spectrum disorder is a complex, lifelong developmental disability where the affected people show repetitive behavior and faces abnormal communication challenges. The goal of this work is to propose an enhanced machine learning model that detects autism more accurately. Hence, we collected ASD datasets of toddler, child, adolescent, and adult from kaggle and UCI machine learning repository. The correlation among individual features was scrutinized and eliminated highly co-linear features in these datasets. Then, feature transformation methods including standardization and normalization were applied in these datasets. Different classifiers like artificial neural network, recurrent neural network, decision tree, extreme learning machine, gradient boost, k nearest neighbor, logistic regression, multilayer perceptron, naïve bayes, random forest, support vector machine, and xgboost were employed in these altered ASD datasets and determined their performances. Logistic regression shows the best result that outperforms other classifiers. This model is useful to extract significant traits and detect autism more precisely.**

***Index Terms*—ASD, AQ-10, Correlation, Feature Transformation, Classifier, Machine Learning.**

I. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that is usually characterized by difficulties in social interaction and communication. It often accompanies narrow interests and repetitive or stereotyped behaviors [1]. However, it is very hard to detect and diagnose ASD nevertheless these patients can be identified anytime. There are existing numerous screening tools to identify autism where most of them are maintained in monotonous ways. All procedures are depended on the behavioral analysis that contributes to more optimistic outcomes for ASD. However, Allison et al. [2] generated a short version of Quantitative Checklist for Autism in Toddlers (Q-CHAT) and Autism Quotient (AQ) where the number of items was reduced from 50 to 10 using discriminant indexing (DI) method and denoted them as Q-CHAT-10 and AQ-10 respectively. In this case, Q-CHAT-10 is used for toddler and AQ-10 is used for the child, adolescent, and adult. Moreover, numerous works were happened to early identify ASD using AQ-10 screening tool. Akyol et al. [3] identified some significant features of children, adolescents, and adults to detect ASD through recursive feature elimination and stability selection methods. Badel et al. [4] proposed a new semi-supervised machine learning model called Clustering-based Autism Trait Classification (CATC) that evaluated individual classification methods respectively. In 2020, Baranwal and Vanitha [5] investigated ASD screening datasets to compare and predict probable cases of children, adolescents, and adults. Thabtah et al. [6] developed a mobile app named ASDTests based on Q-CHAT and AQ-10 tools to detect autism as early as possible. Later, they [7] proposed a new rules-based machine learning (RML) method which detects autism traits and offers intelligent rules to explore reasons for classification. Akter et al. [8] implemented several feature transformation, classification, and feature selection methods to identify significant ASD risk factors for the toddlers, children, adolescents, and adults. Omar et al. [9] developed a mobile app where its data was scrutinized for predicting ASD by training AQ-10 screening and normal instances. Halibas et al. [10] analyzed the child, adolescent, and adult ASD screening datasets using five supervised models. Again, Thabtah et al. [11] proposed a computational intelligence (CI) method called Variable Analysis (VA) that extracted feature-to-class and feature-to-feature correlations and built a predictive model.

The aim of this work is to investigate autism datasets with various machine learning methods and generated better results to early detect autism more accurately. To detect autism in the early stages is a challenging task where many works had happened throughout the world. Hence, we gathered toddler, child, adolescent, and adult ASD datasets that followed Q-CHAT-10 and AQ-10 screening criteria. Then, correlation based analysis and transformation methods were implemented



(a) Toddlers

(b) Child

(c) Adolescent

(d) Adult

Fig. 1. Correlations of Different Datasets

to manipulate these datasets. Several state-of-art classifiers were implemented into generated datasets that can identify ASD with high achievement and explored the best approach. This work is significant that can detect ASD at any stage of age more accurately. As a result, various organizations who are worked with ASD can be interested to utilize this proposed model. Medical researchers can be easily detected significant traits and accelerated the investigation about ASD to cure autism.

II. Materials and Methods

We proposed a gradual sequence of work to detect ASD respectively. Figure 2 reveals the working procedure by analyzing various cases and exploring ASD with high accuracy. This working methodology had been partitioned into several steps which are given as follows:

A. Data Collection

Autism screening datasets for toddlers, child, adolescent, and adults were used 20 features to explore autism traits and improve the performance of ASD cases respectively. Each dataset contains 10 behavioral features that had been used as efficient criteria to detect ASD cases. Using Q-CHAT-10 and AQ-10 tools, Thabtah et al. [6] developed ASDTests app for screening, collecting data, and identifying ASD risk factors respectively. This app is taken point limits from 0 to 10. If the

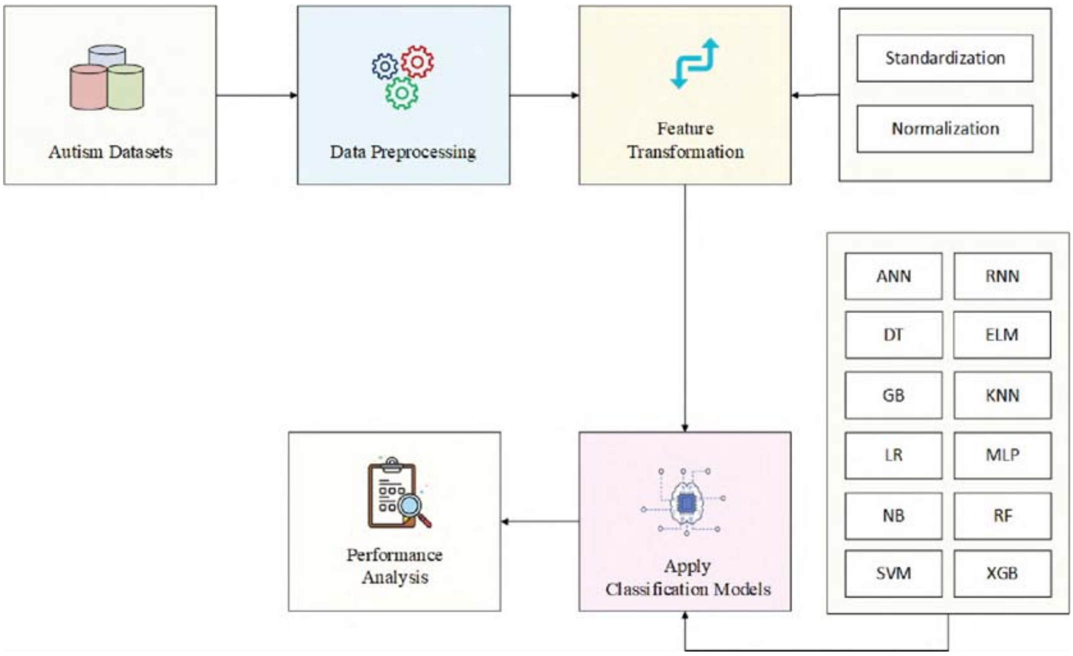


Fig. 2. Proposed Methodology

point is greater than 6, it indicates an affirmative result about ASD otherwise negative. We gathered children, adolescent, and adult datasets from the University of California Irvine (UCI) and toddlers dataset from Kaggle machine learning repository [12], [13], [14], [15]. The toddlers dataset contains 1054 samples (319 female and 735 male) and 248 records are found from the child dataset along with 74 male and 174 female samples. Again, 98 records are explored from the adolescent dataset where both female and male contain 49 records. However, 609 samples are found from the adult dataset that consists of 288 female and 321 male records.

B. Data Preprocessing

These datasets were contained several missing values that were replaced by mean values. Thus, we investigated the correlation among individual variables and measured how strongly they are related. When one variable is highly correlated with other variables, it can be shown unstable classification result and one of them is needed to omit. Hence, some attributes were removed from the dataset. In the toddlers dataset, "QChat10_Score" is highly correlated with other variables. Again, "age_desc" and "result" are extremely correlated with others in the child and adolescent dataset. Consequently, "result" is highly correlated with other much more variables in the adult dataset. Therefore, we excluded "QChat10_Score", "age_ desc", and "result" variables from these datasets. Figure 1 shows the correlations among all features.

C. Feature Transformation and Classification

Various feature transformation methods are used to convert features from one to another. In this work, these datasets contain numerous categorical variables that alter categorical variables into a suitable format so that the performance of individual classifiers were improved. So, two feature transformation models such as standardization and normalization were employed (see Table I) into these autism datasets. Besides, different machine learning classifiers such as Artificial Neural Network (ANN), Recurrent Neural Network (RNN), Decision Tree (DT), Extreme Learning Machine (ELM), Gradient Boost (GB), K Nearest Neighbor (KNN), Logistic Regression (LR), Multilayer Perceptron (MLP), Naïve Bayes (NB), Random Forest (RF), Support Vector Machine (SVM) and XGBoost (XGB) were employed into autism datasets to explore the best performing classifier for detecting autism more accurately. However, these classifiers were widely used to analyze numerous types of datasets [16], [17], [18], [19], [20], [21]. All classifiers were executed standardized datasets except RNN. Instead, only RNN was manipulated by the normalization dataset in this work.

TABLE I
FEATURE TRANSFORMATION METHODS

Metrics	Details	Formula
Standardization	Re-scales features in order to distribute them into 0 and 1.	$X_{new} = \frac{X_i - X_{mean}}{\sigma}$
Normalization	Shifted and re-scaled features where they ranges them between 0 and 1.	$X = \frac{X - X_{min}}{X_{max} - X_{min}}$

D. Evaluation Metrics

In classification approach, individual values are generated like true positive (TP), true negative (TN), false positive (FP), and false negative (FN). These values is represented in a suitable structure which is called confusion matrix. Thus, various evaluation metrics such as accuracy, f-measure, g-mean, AUC, sensitivity, specificity, fall out, and miss rate were used to verify the performance of these classifiers from these values.

III. Experimental Result

In this work, we used several machine learning classifiers like ANN, RNN, DT, ELM, GB, KNN, LR, MLP, NB, RF, SVM, and XGB in toddler, child, adolescent, and adult ASD datasets. These methods were followed by 10 fold cross validation technique using sci-kit learn and Keras library in python. Different evaluation metrics were used to manipulate their performance. The experiment result for the Toddlers dataset is shown in Table II. All classifiers perform their results above 90.00% in all evaluation metrics. However, LR demonstrates the best outcomes where it represents 100% accuracy, 100% f-measure, 100% g-mean, 100% AUC, 100% sensitivity, and 100% specificity, along with 0% fall-out and 0% miss-rate. Then, MLP, RNN, ANN, XGB, GB, SVM, RF, KNN, NB, ELM, and DT provide better outcomes sequentially.

Afterwards, the classification outcomes for the child dataset are shown in Table III. Consequently, eight classifiers produce their results greater than 90% and four classifiers represent their performance between 85% and 90%. Again, LR performs the highest accuracy (99.32%), AUC (99.29%), f-measure (99.31%), g-mean (99.29%), sensitivity (99.32%), specificity (99.27%), along with lowest fall-out (0.73%) and miss-rate (0.68%). Besides, MLP generates the second-best result for all evaluation metrics. Then, RNN, ANN, XGB, SVM, GB, and RF make obvious better outcomes than other classifiers except for LR and MLP. The ELM, DT, KNN, and NB have manipulated the moderate results in this work. The classification results for the adolescent dataset are represented at Table IV. LR, MLP, XGB, and ANN produce their accuracy above 90%. Besides, RF, SVM, RNN, GB, NB, ELM, and KNN compute their result greater than 80% and less than 90%. However, DT generates the lowest 75.96% accuracy among all of the classifiers. LR shows the best 95.19% accuracy, 94.33% AUC, 95.16% f-measure, 94.32% g-mean, 95.19% sensitivity, 93.46% specificity including 6.54% fall-out and 4.81% miss-rate than all models.

The outcomes of different classifiers for the Adult dataset are represented in Table V. All of them provide their accuracy greater than 90%. LR shows the maximum 99.86% accuracy, 99.74% AUC, 99.86% f-measure, 99.74% g-mean, 99.86% sensitivity, and the minimum 14% miss-rate. However, ANN produces the highest 99.64% specificity and the lowest 0.36% fall-out. MLP also generates the second highest result showing 99.57% accuracy, 99.54% AUC, 99.57% f-measure, 99.54% g-mean, 99.57% sensitivity and 99.51% specificity, 0.49% fall-out and 0.43% miss-rate. Then, other classifiers also represent excellent results likewise LR and MLP.

IV. Discussions

In this work, several machine learning classifiers were used to investigate these datasets and detect ASD as early as possible. Most of them were shown their best performance according to previous relevant works. Many researchers were investigated ASD screening datasets to identify autism and typical developed children more accurately. Therefore, they used various data analytical procedure to get excellent outcomes. But, a high performing model was not found that can detect ASD/non-ASD more precisely. Thus, we proposed a machine learning model where it provided more improved results to detect ASD. To generate better result, data exploration is a significant tool that properly represented the characteristics of the dataset. In this case, we scrutinized individual variables of each ASD dataset and explored highly co-linear factors of them. Besides, various transformation methods were used to convert these variables into suitable forms and got more appropriate results. Moreover, different types of possible evaluation metrics were used to verify the performance of classifiers comparing to existing studies. Besides, these classifiers were widely used to diagnose ASD in previous works by the researchers. So, the experimental outcomes were critically justified based on various metrics and finally explored more significant model precisely.

In this work, LR gives the best result than other classifiers for toddlers, child, adolescents, and adults respectively (see Table II, III, IV, V). Therefore, we observed that LR and neural network-based methods are shown the best result than others. LR also associates with neural network model because a small modification of it can be used in the ANN. However, most of our classifiers were selected as the neural network basis. In previous work, many machine learning methods had been used where most of the classifiers produce more false positive and negative instances for the redundant features and inappropriate analysis. Besides, they provided a more unstable or garbage result for unseemly values. In the proposed model, we reduced highly correlated variables, transformed them, applied various neural networks and showed better results than others. Table VI provides the comparison of existing works with the proposed model where it shows the highest outcomes than previous works for various evaluation metrics.

V. Conclusion and Future Works

In summary, we preprocessed autism datasets into a transformed format and implemented classifiers to detect ASD and typically normal people as well. Then, the best classification method had been detected that produces more improved results. However, LR provides the best result for all datasets. Therefore, physicians and psychiatrists can gather the most significant behavioral features and design essential treatment policy for ASD from this analysis. This work also accelerates the research on ASD drug engineering. Further, social awareness about early detection can be arisen by following individual factors. Some limitations are found for instance, no screening tools are not perfectly detected autism throughout

TABLE II
EXPERIMENT RESULT OF INDIVIDUAL CLASSIFIERS FOR TODDLERS

Classifier	Accuracy	AUC	F-Measure	G-Mean	Sensitivity	Specificity	Fall-out	Miss-rate
ANN	0.9896	0.9891	0.9896	0.9891	0.9896	0.9886	0.0114	0.0104
RNN	0.9943	0.9933	0.9943	0.9933	0.9943	0.9924	0.0076	0.0057
DT	0.9175	0.9030	0.9174	0.9029	0.9175	0.8885	0.1115	0.0825
ELM	0.9231	0.9046	0.9227	0.9044	0.9231	0.8860	0.1140	0.0769
GB	0.9782	0.9723	0.9781	0.9723	0.9782	0.9665	0.0335	0.0218
KNN	0.9488	0.9443	0.9490	0.9443	0.9488	0.9398	0.0602	0.0512
LR	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**0.0000**	**0.0000**
MLP	0.9991	0.9993	0.9991	0.9993	0.9991	0.9996	0.0004	0.0009
NB	0.9431	0.9291	0.9428	0.9290	0.9431	0.9152	0.0848	0.0569
RF	0.9592	0.9468	0.9590	0.9467	0.9592	0.9343	0.0657	0.0408
SVM	0.9753	0.9661	0.9752	0.9660	0.9753	0.9568	0.0432	0.0247
XGB	0.9820	0.9793	0.9820	0.9793	0.9820	0.9767	0.0233	0.0180

TABLE III
EXPERIMENT RESULT OF INDIVIDUAL CLASSIFIERS FOR CHILD

Classifier	Accuracy	AUC	F-Measure	G-Mean	Sensitivity	Specificity	Fall-out	Miss-rate
ANN	0.9589	0.9591	0.9589	0.9591	0.9589	0.9593	0.0407	0.0411
RNN	0.9726	0.9723	0.9726	0.9723	0.9726	0.9721	0.0279	0.0274
DT	0.8938	0.8938	0.8938	0.8938	0.8938	0.8938	0.1062	0.1062
ELM	0.8973	0.8969	0.8972	0.8969	0.8973	0.8965	0.1035	0.1027
GB	0.9315	0.9317	0.9315	0.9317	0.9315	0.9318	0.0682	0.0685
KNN	0.8904	0.8929	0.8901	0.8929	0.8904	0.8953	0.1047	0.1096
LR	**0.9932**	**0.9929**	**0.9931**	**0.9929**	**0.9932**	**0.9927**	**0.0073**	**0.0068**
MLP	0.9863	0.9861	0.9863	0.9861	0.9863	0.9858	0.0142	0.0137
NB	0.8664	0.8650	0.8661	0.8650	0.8664	0.8635	0.1365	0.1336
RF	0.9110	0.9104	0.9109	0.9104	0.9110	0.9098	0.0902	0.0890
SVM	0.9452	0.9449	0.9452	0.9449	0.9452	0.9446	0.0554	0.0548
XGB	0.9555	0.9555	0.9555	0.9555	0.9555	0.9556	0.0444	0.0445

TABLE IV
EXPERIMENT RESULT OF INDIVIDUAL CLASSIFIERS FOR ADOLESCENT

Classifier	Accuracy	AUC	F-Measure	G-Mean	Sensitivity	Specificity	Fall-out	Miss-rate
ANN	0.9038	0.8993	0.9038	0.8993	0.9038	0.8948	0.1052	0.0962
RNN	0.8846	0.8835	0.8851	0.8835	0.8846	0.8823	0.1177	0.1154
DT	0.7596	0.7207	0.7482	0.7196	0.7596	0.6817	0.3183	0.2404
ELM	0.8269	0.8316	0.8285	0.8316	0.8269	0.8363	0.1637	0.1731
GB	0.8750	0.8542	0.8725	0.8540	0.8750	0.8335	0.1665	0.1250
KNN	0.8077	0.7604	0.7927	0.7589	0.8077	0.7130	0.2870	0.1923
LR	**0.9519**	**0.9433**	**0.9516**	**0.9432**	**0.9519**	**0.9346**	**0.0654**	**0.0481**
MLP	0.9423	0.9311	0.9417	0.9310	0.9423	0.9199	0.0801	0.0577
NB	0.8558	0.8469	0.8554	0.8468	0.8558	0.8380	0.1620	0.1442
RF	0.8942	0.8786	0.8928	0.8785	0.8942	0.8630	0.1370	0.1058
SVM	0.8942	0.8744	0.8921	0.8741	0.8942	0.8545	0.1455	0.1058
XGB	0.9135	0.8945	0.9117	0.8943	0.9135	0.8755	0.1245	0.0865

TABLE V
EXPERIMENT RESULT OF INDIVIDUAL CLASSIFIERS FOR ADULT

Classifier	Accuracy	Auc	F-Measure	G-Mean	Sensitivity	Specificity	Fall-out	Miss-rate
ANN	0.9901	0.9932	0.9901	0.9932	0.9901	0.9964	0.0036	0.0099
RNN	0.9673	0.9392	0.9666	0.9387	0.9673	0.9110	0.0890	0.0327
DT	0.9062	0.8756	0.9058	0.8751	0.9062	0.8450	0.1550	0.0938
ELM	0.9190	0.8860	0.9181	0.8854	0.9190	0.8531	0.1469	0.0810
GB	0.9659	0.9482	0.9656	0.9481	0.9659	0.9306	0.0694	0.0341
KNN	0.9432	0.9360	0.9436	0.9360	0.9432	0.9289	0.0711	0.0568
LR	**0.9986**	**0.9974**	**0.9986**	**0.9974**	**0.9986**	**0.9961**	**0.0039**	**0.0014**
MLP	0.9957	0.9954	0.9957	0.9954	0.9957	0.9951	0.0049	0.0043
NB	0.9418	0.9284	0.9419	0.9283	0.9418	0.9150	0.0850	0.0582
RF	0.9588	0.9350	0.9583	0.9347	0.9588	0.9112	0.0888	0.0412
SVM	0.9716	0.9555	0.9713	0.9553	0.9716	0.9393	0.0607	0.0284
XGB	0.9659	0.9533	0.9658	0.9532	0.9659	0.9406	0.0594	0.0341

TABLE VI
COMPARATIVE ANALYSIS WITH PREVIOUS WORKS

Dataset	Authors	Accuracy	AUC	F-Measuse	G-Mean	Sensitivity	Specificity	Fall Out	Miss-Rate
Toddler	Omar et al. [9]								
	Thabtah et al. [6]								
	Akter et al. [8]	98.77				99.39	99.39		
	Proposed model	**100**	**100**	**100**	**100**	**100**	**100**	**0**	**0**
Child	Omar et al. [9]	92.26							
	Thabtah et al. [6]	97.8				98	97.35		
	Akter et al. [8]	97.2				98.4	98.46		
	Proposed model	**99.32**	**99.29**	**99.31**	**99.29**	**99.32**	**99.27**	**0.73**	**0.68**
Adolescent	Omar et al. [9]	93.78							
	Thabtah et al. [6]	94.23				92.2	92.68		
	Akter et al. [8]	93.89				97.5	98.33		
	Proposed model	**95.19**	**94.33**	**95.16**	**94.32**	**95.19**	**93.46**	**6.54**	**4.81**
Adult	Omar et al. [9]	97.1							
	Thabtah et al. [6]	99.85				99.9	99.7		
	Akter et al. [8]	98.36				99.3	96.11		
	Proposed model	**99.86**	**99.74**	**99.86**	**99.74**	**99.86**	**99.64**	**0.36**	**0.14**

the world. Therefore, collected behavioral instances and their analyzed outcomes cannot surely explain ASD cases. In addition, there were used a small amount of instances which were not sufficient to resolve these studies. In future, we will gather ASD datasets from more resources, scrutinize them, and find similar significant features from individual datasets. Besides, experimental results will be critically validated with existing works to explore better outcomes for different aspects of ASD.

REFERENCES

[1] K. L. Goh, S. Morris, S. Rosalie, C. Foster, T. Falkmer, and T. Tan, "Typically developed adults and adults with autism spectrum disorder classification using centre of pressure measurements," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 844–848.

[2] C. Allison, B. Auyeung, and S. Baron-Cohen, "Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 51, no. 2, pp. 202–212, 2012.

[3] K. Akyol, "Assessing the importance of autistic attributes for autism screening," *Expert Systems*, p. e12562, 2020.

[4] S. Baadel, F. Thabtah, and J. Lu, "A clustering approach for autistic trait classification," *Informatics for Health and Social Care*, pp. 1–18, 2020.

[5] A. Baranwal and M. Vanitha, "Autistic spectrum disorder screening: Prediction with machine learning models," in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*. IEEE, 2020, pp. 1–7.

[6] F. Thabtah, F. Kamalov, and K. Rajab, "A new computational intelligence approach to detect autistic features for autism screening," *International journal of medical informatics*, vol. 117, pp. 112–124, 2018.

[7] F. Thabtah and D. Peebles, "A new machine learning model based on induction of rules for autism detection," *Health informatics journal*, vol. 26, no. 1, pp. 264–286, 2020.

[8] T. Akter, M. S. Satu, M. I. Khan, M. H. Ali, S. Uddin, P. Lio, J. M. Quinn, and M. A. Moni, "Machine learning-based models for early stage detection of autism spectrum disorders," *IEEE Access*, vol. 7, pp. 166 509–166 527, 2019.

[9] 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," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, 2019, pp. 1–6.

[10] A. S. Halibas, L. B. Reazol, E. G. T. Delvo, and J. C. Tibudan, "Performance analysis of machine learning classifiers for asd screening," in *2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*. IEEE, 2018, pp. 1–6.

[11] F. Thabtah, "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward," *Informatics for Health and Social Care*, vol. 44, no. 3, pp. 278–297, 2019.

[12] "Autism screening data for toddlers," https://www.kaggle.com/fab-delja/autism-screening-for-toddlers, (Accessed on 09/10/2018).

[13] "UCI machine learning repository: Autistic spectrum disorder screening data for children data set," https://archive.ics.uci.edu/-ml/datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++, (Accessed on 09/10/2018).

[14] "UCI machine learning repository: Autistic spectrum disorder screening data for adolescent data set," https://archive.ics.uci.edu/-ml/datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++, (Accessed on 09/10/2018).

[15] "UCI machine learning repository: Autism screening adult data set," https://archive.ics.uci.edu/-ml/datasets/Autism+Scree-ning+Adult, (Accessed on 09/10/2018).

[16] M. S. Satu, F. F. Sathi, M. S. Arifen, M. H. Ali, and M. A. Moni, "Early detection of autism by extracting features: A case study in bangladesh," in *2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*. IEEE, 2019, pp. 400–405.

[17] M. S. Satu, T. Akter, and M. J. Uddin, "Performance analysis of classifying localization sites of protein using data mining techniques and artificial neural networks," in *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, 2017, pp. 860–865.

[18] M. S. Satu, S. Rahman, M. I. Khan, M. Z. Abedin, M. S. Kaiser, and M. Mahmud, "Towards improved detection of cognitive performance using bidirectional multilayer long-short term memory neural network," in *International Conference on Brain Informatics*. Springer, 2020, pp. 297–306.

[19] K. Ahammed, M. S. Satu, M. I. Khan, and M. Whaiduzzaman, "Predicting infectious state of hepatitis c virus affected patient's applying machine learning methods," in *2020 IEEE Region 10 Symposium (TENSYMP)*. IEEE, 2020, pp. 1371–1374.

[20] M. S. Satu, S. T. Atik, and M. A. Moni, "A novel hybrid machine learning model to predict diabetes mellitus," in *International Joint Conference on Computational Intelligence*. Springer, 2019, pp. 453–465.

[21] M. S. Satu, F. Tasnim, T. Akter, and S. Halder, "Exploring significant heart disease factors based on semi supervised learning algorithms," in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*. IEEE, 2018, pp. 1–4.