Predicting Autism Spectrum Disorder Based On Gender Using Machine Learning Techniques

Tania Akter
Computer Science and Engineering
Jahangirnagar University, Gono Bishwabidyalay
Savar, Dhaka, Bangladesh
Email: shilpy339cse@gmail.com

Mohammad Hanif Ali
Computer Science and Engineering
Jahangirnagar University
Savar, Dhaka, Bangladesh
Email: hanif_ju03@juniv.edu

2021 3rd International Conference on Electrical & Electronic Engineering (ICEEE) | 978-1-6654-8281-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICEEE54059.2021.9718798

Abstract—**Autism is a set of complicated developmental disorders marked by social skills impairments, communication difficulties (verbal and nonverbal), and recurring behavior. Autistic children are frequently alienated as a result of these impairments. Rapid recognition of autism can help to establish a treatment strategy and lessen the burden on sufferers. As a result, effective methods to early diagnosis and treatment for ASD are necessary. The toddler, child, adolescent, and adult screening datasets are collected in this study and separated according to gender (male and female). By using random oversampling (ROS), these datasets are balanced. Next, different classifiers are applied to both the primary and balanced datasets. The MLP classifier produced the best results, and the hyperparameter for it was tuned to improve autism identification rate. However, the experimental outcome for the female dataset is better than the male dataset. The shapely adaptive explanation (SHAP) method is also employed to assess the significant features of male and female.**

keyword Autism, Machine Learning, ROS, Grid Search, SHAP Analysis

I. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that is found among all ages of people and hampers a patient's capacity to link and communicate with others. According to Center for Diseases Control and Prevention (CDC) monitoring statistics, there were held a nationwide survey of parents about the current ASD prevalence where 1 in 59 children were found ASD within eight years old [1]. This disorder often occurs in all ethnic or racial groups, although the numbers of boys are intended to affect four times than girls. Therefore, children should need to diagnose autism at early stage and mediate quickly [2].

Nowadays, some tools such as Autism Diagnostic Observation Schedule (ADOS), Autism Diagnostic Interview-Revised (ADI-R), and Diagnostic and Statistical Manual of Mental Disorders 5 (DSM-5) are widely used to diagnose autism. However, they take a long periods to finish the assessment and evaluate ASD as well. Therefore, an intelligent approach of machine learning was presented to solve this problem.

In this study, we used several machine learning algorithms to investigate ASD male and female individually. Then, the hyperparameters of the best classifier were optimized using grid search approach and determined SHapley Additive exPlanations (SHAP) values to rank individual characteristics of ASD. The main goals of this work is to reduce diagnostic time and recognizing the best ASD characteristics of male and female. From this work, various researchers and autism-related institutions will benefit. The significant achievements of this task are given as follows:

- To apply various machine learning algorithms for detecting ASD in boys and girls as early as possible.
- To make use of the random oversampling (ROS) approach to balance the model.
- To employ the grid search technique to optimize the hyperparameter of a model.
- To identify significant features of ASD using SHapley Additive exPlanations (SHAP).

II. Materials and Method

We proposed a machine learning framework to detect ASD of male and female individually. Figure 1 depicts a gradual sequence of our methodology and briefly described as follows:

A. Dataset Description

We used version 2 toddler, child, adolescent, and adult datasets in this work. These data are collected from a developed mobile application called ASDTests [3]. This app calculates the score limit from 0 to 10 according to answers to screening questions (A1 to A10). Each dataset contains 23 features with class label except toddler (toddler have 18 features). If the finishing score value is fewer than or equivalent to 7, it assigns a "No" class value for ASD; otherwise, it assigns "Yes". The toddler dataset contains 1,054 instances with the age of 18 to 36 months, and the child dataset contains 509 instances with 4 to 11 years, adolescent contains 248 instances with 12 to 15 years and adult contains 1,118 instances with the age of 17 years old and greater.

B. Data Preprocessing

In this work, we split only the male and female datasets from combined datasets of the toddler, child, adolescent, and adult, respectively. Next, the missing values in the dataset are replaced with mean values. Then, we reduced some features such as "User (who completed the screening)", "Case", "Age",

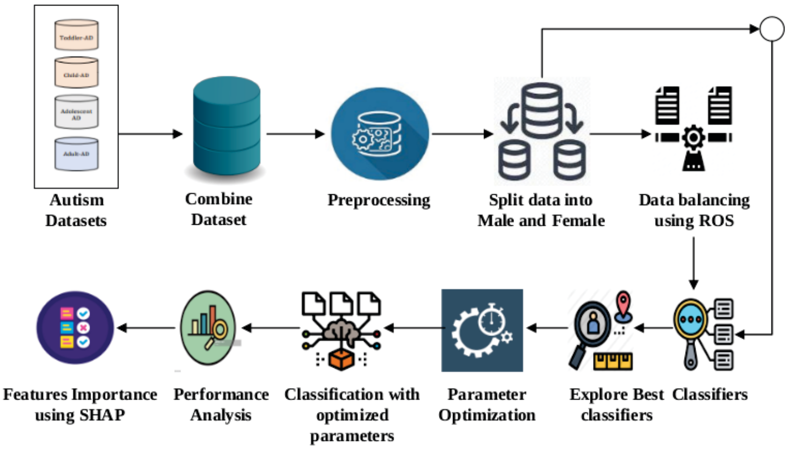


Figure 1: The proposed workflow of the work

"Used App Before", "Language", "Why taken the screening", "Screening Type", and "Score" which are meta information and are not associated adequately with ASD [4].

C. Data Balancing

When we separate the combined dataset by male and female, we observe that the class values of ASD and non-ASD are vastly different, which requires more research. As a result, we use Random OverSampling (ROS) to balance the dataset. ROS selects samples from the minority class randomly and concatenates them towards the training sample.

D. Data Transformation

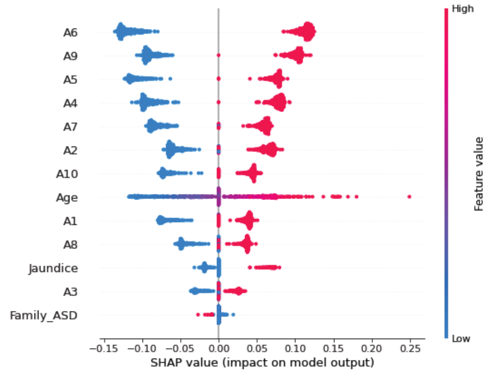
For further ML investigation, it is necessary to transform data from the original format to the required format. In this paper, we use a standard scalar transformation approach to translate both primary and ROS male and female data into an acceptable format. The Standard Scaler adjusts the numbers so that the mean is equal to zero as well as the variance is equal to one.

E. Classification Techniques

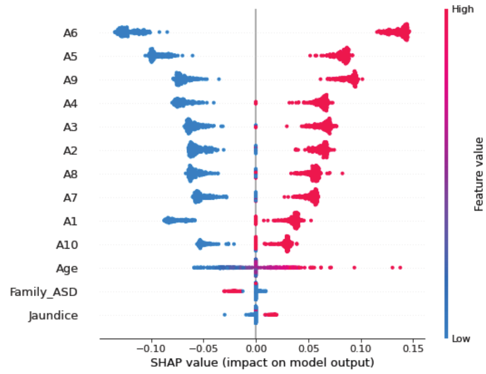
After that, we use several clinically applicable classification methods already used in previous works relating to detecting autism [5] [6] [7] [8]. These are Extreme Gradient Boosting (XGB), Decision Tree (DT), Naïve Bayes (NB), Random Forest (RF), K-Nearest Neighbor (KNN), Gradient Boost (GB), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), and Logistic Regression (LR) on male and female primary and ROS datasets [9]. For all classifiers, default parameters are employed.

F. Hyperparameter Optimization

All classifiers have distinct hyperparameters to regulate their performance. The Grid search approach is used in this study to modify the hyperparameters of a classifier. After setting the hyperparameter value limit, the classifier is built by scanning multiple hyperparameter permutations. The hyperparameters that produce the most significant results serve as the model's



(a) SHAP values analysis for male autism dataset



(b) SHAP values analysis for female autism dataset

Figure 2: SHAP value analysis using MLP classifier to rank the features of the dataset

identification[10]. In this paper, the hyperparameters (Activation, alpha, hidden layer sizes, learning rate, and optimizer) of the best performing algorithm (MLP) are optimized.

G. Performance Evaluation Metrics

All classification models are justified their result with an evaluation metric which quantifies the performance of a predictive model. In this work, we used accuracy, precision, recall, AUROC and F-measure to evaluate individuals classifiers performance.

H. Exploring Significant Features for Male and Female

To explore the significant features of males and females individually, we used the SHapley Additive exPlanations (SHAP) technique [11] to detect the precedence of separate attribute of the male and female primary datasets using the best classifier. Next, by analyzing the correlation of individual characteristics in the male and female primary datasets, discriminating factors of male and female autism are identified.

III. Result and Discussions

We used Random OverSampling (ROS) to balance the dataset in this study. Several classifiers were implemented

Table I: Performance Analysis of Male and Female dataset

Male dataset									
Classifier	**DT**	**NB**	**KNN**	**SVM**	**LR**	**MLP**	**RF**	**XGB**	**GB**
Accuracy	0.841	0.8520	0.8520	0.8796	0.8642	**0.8995**	0.8769	0.8896	0.8796
Precision	0.8414	0.8676	0.8453	0.8846	0.8683	**0.8936**	0.8768	0.8875	0.8806
Recall	0.8591	0.8475	0.8791	0.8864	0.8738	**0.9180**	0.8906	0.9043	0.8917
AUROC	0.8400	0.8523	0.8506	0.8793	0.8637	**0.8985**	0.8761	0.8888	0.8790
F-Measure	0.8502	0.8574	0.8619	0.8855	0.8711	**0.9056**	0.8837	0.8958	0.8861
Female Dataset									
Classifier	**DT**	**NB**	**KNN**	**SVM**	**LR**	**MLP**	**RF**	**XGB**	**GB**
Accuracy	0.8506	0.8810	0.8766	0.9106	0.8945	**0.9141**	0.9061	0.8900	0.9007
Precision	0.8438	0.8587	0.8483	0.9006	0.8734	**0.9075**	0.8935	0.8808	0.8778
Recall	0.8324	0.8902	0.8940	0.9075	0.9037	0.9075	0.9056	0.8825	**0.9133**
AUROC	0.8494	0.8816	0.8777	0.9104	0.8951	**0.9137**	0.9060	0.8895	0.9016
F-Measure	0.8380	0.8742	0.8705	0.9040	0.8883	**0.9075**	0.8995	0.8816	0.8952

Table II: Performance Analysis of Male and Female ROS dataset

Male Dataset									
Classifier	**DT**	**NB**	**KNN**	**SVM**	**LR**	**MLP**	**RF**	**XGB**	**GB**
Accuracy	0.8617	0.8512	0.8554	0.8917	0.8649	**0.9022**	0.8896	0.8970	0.8943
Precision	0.8598	0.8568	0.8456	0.8992	0.8630	**0.9135**	0.8929	0.8970	0.8890
Recall	0.8644	0.8433	0.8696	0.8822	0.8675	0.8885	0.8854	0.8970	**0.9012**
AUROC	0.8617	0.8512	0.8554	0.8917	0.8649	**0.9022**	0.8896	0.8970	0.8943
F-Measure	0.8621	0.8500	0.8574	0.8907	0.8652	**0.9009**	0.8891	0.8970	0.8950
Female dataset									
Classifier	**DT**	**NB**	**KNN**	**SVM**	**LR**	**MLP**	**RF**	**XGB**	**GB**
Accuracy	0.8748	0.8823	0.8815	0.9115	0.8923	**0.9140**	0.9098	0.9082	0.9023
Precision	0.8662	0.8706	0.8621	0.9061	0.8730	**0.9133**	0.8992	0.8976	0.8875
Recall	0.8865	0.8982	0.9082	0.9182	0.9182	0.9149	**0.9232**	0.9215	0.9215
AUROC	0.8748	0.8823	0.8815	0.9115	0.8923	**0.9140**	0.9098	0.9082	0.9023
F-Measure	0.8762	0.8841	0.8846	0.9121	0.8950	**0.9141**	0.9110	0.9094	0.9042

Table III: Performance analysis of tuned MLP model

	Accuracy	Precision	Recall	AUROC	F-Measure
Male	0.9727	0.9787	0.9664	0.9727	0.9725
Female	**0.9825**	**0.9833**	**0.9816**	**0.9825**	**0.9825**

with 10 fold cross-validation into the primary and balanced male and female autism datasets. To justify the performance of different classifiers, seven evaluation metrics are utilized. SHapley Additive exPlanations (SHAP) is used to calculate the degree of involvement of each attribute to the performance of a classification model. All of the work was executed to Google Colaboratory on the cloud platform.

A. Comparison of Classification Results

Table I shows the experimental results for the male and female primary autism datasets. All classifiers produce results that are more than 80 percent. For the male dataset, the MLP model presented the highest accuracy of 89.95%, precision of 89.36%, recall of 91.80%, AUROC of 89.85% and F-measure of 90.56%. On the other hand, for the female dataset, the MLP model again evaluated the highest accuracy of 91.41%, precision of 90.75%, AUROC of 91.37% and F-measure of 90.75%.

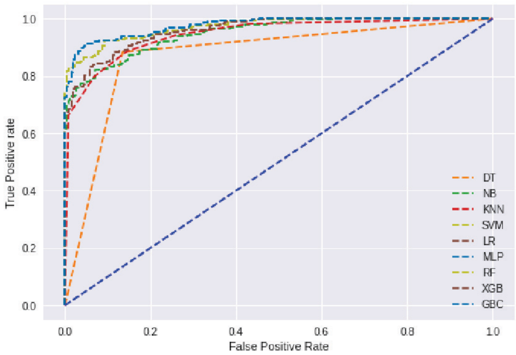
Afterward, the experimental outcomes for the male and female ROS datasets are shown in Table II. Overall, each classifier's experimental outcomes in the balanced datasets outperformed those in the original datasets. The MLP classifier demonstrated the maximum accuracy (90.22%), precision (91.35%), AUROC (90.22%) and F-measure (90.09%) in the male ROS dataset. All classifiers calculated the accuracy above 85% for oversampling dataset. On the other hand, the MLP classifier again generated the utmost accuracy (91.40%), precision (91.33%), AUROC (91.40%) and F-measure (91.41%) in the female ROS dataset. Different classifiers' receiver operating characteristic (ROC) curves are depicted in Figure 3. In these curves, the MLP model exceeds other analyzers to both male and female datasets.

B. Parameter Tuning of the Best Classifier

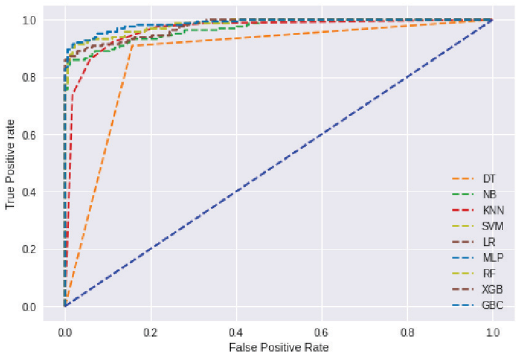
The hyperparameters of MLP are tuned because this algorithm showed the best result previously, and the grid search technique is employed for tuning. For the female balanced dataset, the optimum values of hyperparameters in MLP are relu for activation, 0.05 for alpha, $(10, 30, 10)$ for hidden layer sizes, adaptive for learning rate, and adam for optimizer. Conversely, the relu for activation, 0.0001 for alpha, $(10, 30, 10)$ for hidden layer sizes, constant for learning rate, and adam for optimizer are the tuned values for MLP hyperparameters in male balanced dataset. The experimental result for MLP after hyperparameter tuning is shown in Table III. It is observed that MLP performed the highest accuracy (98.25%), precision (98.33%), Recall (98.16%), AUROC (98.25%) and F-measure (98.25%) in the female dataset.

C. Identifying Significant Features of Male and Female

Figure 2 depicts the results of the SHAP value analysis using MLP. The peak features have a higher predictive ability than the base attributes since they contribute more to the system. For the male autism dataset, the most significant features are social chit-chat (A6), social situation (A9), and interpreting conversation (A5), whereas the least essential attributes are



(a) Classification ROC Curve Analysis for Male dataset



(b) Classification ROC Curve Analysis for Female dataset

Figure 3: Classification ROC curve analysis of used machine learning algorithms

Family ASD, track conversation (A3), and Jaundice [shown in Figure 2a]. The most noteworthy features are social chit-chat (A6), interpreting conversation (A5), and social situation (A9), whereas the minor vital attributes are Jaundice, Family ASD, and Age for the female autism dataset [shown in Figure 2b].

D. Comparative Analysis with Existing Work

Several researchers have already worked with QCHAT and AQ-10 screening datasets. They employed a variety of machine learning approaches to diagnose autism with high accuracy in that study. However, no one has attempted to identify autism in men and women individually. Furthermore, most of the studies discovered essential characteristics responsible for autism, but not for males and females individually. In this study, we categorize male and female autism individually and compare their results. Furthermore, we attempt to detect autism more precisely and identify the most critical features responsible for autism. Thus, it improves physicians' technical abilities and management approach for detecting actual characteristics. Furthermore, patients' financial costs are decreased as a result of more accurate ASD detection.

IV. Conclusion and Future Work

In this work, we consolidate, preprocess, and split the dataset into male and female groups. Next, we balance the male and female datasets employing ROS and apply various classifiers. Finally, the hyperparameter of the best performing model (MLP) is tuned applying grid search technique where MLP represented the best outcome (accuracy (98.25%), precision (98.33%), Recall (98.16%), AUROC (98.25%), and F-measure (98.25%) for the female dataset. It was observed that the experimental findings for all female datasets produced positive efficacy than all male datasets. As a result, professionals and therapists can discover appropriate therapy for sufferers by quickly diagnosing autism and identifying the most relevant characteristics. Furthermore, this effort will raise public awareness regarding early detection of autism and will expedite research towards ASD drug manufacture. In the future, we plan to gather additional autism data and use deep learning techniques to improve the accuracy of autism detection.

References

[1] S. J. Moon, J. Hwang, R. Kana, J. Torous, and J. W. Kim, "Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder: Systematic review and meta-analysis of brain magnetic resonance imaging studies," *JMIR mental health*, vol. 6, no. 12, p. e14108, 2019.

[2] M. F. Rabbi, S. M. Hasan, A. I. Champa, and M. A. Zaman, "A convolutional neural network model for early-stage detection of autism spectrum disorder," in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*. IEEE, 2021, pp. 110–114.

[3] F. Thabtah, "An accessible and efficient autism screening method for behavioural data and predictive analyses," *Health informatics journal*, vol. 25, no. 4, pp. 1739–1755, 2019.

[4] M. D. Hossain, M. A. Kabir, A. Anwar, and M. Z. Islam, "Detecting autism spectrum disorder using machine learning techniques," *Health Information Science and Systems*, vol. 9, no. 1, pp. 1–13, 2021.

[5] F. Thabtah, N. Abdelhamid, and D. Peebles, "A machine learning autism classification based on logistic regression analysis," *Health information science and systems*, vol. 7, no. 1, pp. 1–11, 2019.

[6] K. K. Hyde, M. N. Novack, N. LaHaye, C. Parlett-Pelleriti, R. Anden, D. R. Dixon, and E. Linstead, "Applications of supervised machine learning in autism spectrum disorder research: a review," *Review Journal of Autism and Developmental Disorders*, vol. 6, no. 2, pp. 128–146, 2019.

[7] T. Akter, M. I. Khan, M. H. Ali, M. S. Satu, M. J. Uddin, and M. A. Moni, "Improved machine learning based classification model for early autism detection," in *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*. IEEE, 2021, pp. 742–747.

[8] M. M. Ahamad, S. Aktar, M. J. Uddin, M. Rashed-Al-Mahfuz, A. Azad, S. Uddin, S. A. Alyami, I. H. Sarker, P. Liò, J. M. Quinn *et al.*, "Adverse effects of covid-19 vaccination: machine learning and statistical approach to identify and classify incidences of morbidity and post-vaccination reactogenicity," *medRxiv*, 2021.

[9] M. M. Ahamad, S. Aktar, M. Rashed-Al-Mahfuz, S. Uddin, P. Liò, H. Xu, M. A. Summers, J. M. Quinn, and M. A. Moni, "A machine learning model to identify early stage symptoms of sars-cov-2 infected patients," *Expert systems with applications*, vol. 160, p. 113661, 2020.

[10] Z. Qu, J. Xu, Z. Wang, R. Chi, and H. Liu, "Prediction of electricity generation from a combined cycle power plant based on a stacking ensemble and its hyperparameter optimization with a grid-search method," *Energy*, vol. 227, p. 120309, 2021.

[11] T. Akter, M. H. Ali, M. Satu, M. Khan, M. Mahmud *et al.*, "Towards autism subtype detection through identification of discriminatory factors using machine learning," in *International Conference on Brain Informatics*. Springer, 2021, pp. 401–410.