Classification of Autism Spectrum Disorder Across Age using Questionnaire and Demographic Information

SK Rahatul Jannat and Shaun Canavan

University of South Florida, Tampa FL, 33620, USA
{jannat, scanavan}@usf.edu

Abstract. Currently, diagnosis of Autism Spectrum Disorder (ASD) is a lengthy, subjective process and machine learning has been shown to be able to accurately classify ASD, which can help take some of the subjectivity out of the diagnosis. Considering this, we propose a machine learning-based approach to classification of ASD, across age, that make use of subject self-report and demographic information. We analyze the efficacy of the proposed approach on 3 classifiers: k-nearest neighbors (KNN), random forest, and a feed-forward neural network. Our results suggest that the proposed approach can accurately classify ASD in children, adolescents, and adults as it is comparable to or outperforms current state of the art on the publicly available AQ-10 dataset.

Keywords: ASD · Machine Learning · Classification

1 Introduction

Autism spectrum disorders (ASD) affect as many as 1 in 59 youth [5], with many higher-functioning children not diagnosed until school-age or later [18]. Significant impairment in social-communication, adaptive, and school functioning is common, and compared to other types of pediatric psychopathology, ASD is particularly severe and longstanding [4]. Currently, diagnosing ASD is a lengthy process that involves multiple experts, where the result can include subjective bias [17]. Machine-learning based approaches can provide an objective approach to diagnosis that has the potential to improve accuracy and reduce the time required for diagnosis. To determine high priority patients that should receive a referral for diagnosis, it has been proposed that those that have a high score on the Autism-Spectrum Quotient (AQ) questionnaire [2] should be referred. The AQ questionnaire is one of the main ways that patients are assessed for autistic traits [11]. Ashwood et al. [1] investigated whether the AQ could predict who would receive an ASD diagnosis later in life. They found that while the AQ scores had a high sensitivity, there were a lot of false negatives based on a threshold score (e.g. patient had ASD, but they were below threshold). Wakabayahsi et al. [16] investigated AQ scores across culture, more specifically the United Kingdom and Japan. The results suggest that autistic conditions are similar across cultures, as the results from Japan replicated those from the United Kingdom.

Omar et al. [12] have shown that machine learning can be applied to the AQ questionnaire to predict ASD. They used a random forest [3] along with AQ data to predict ASD in children, adolescents, and adults with 92.26%, 93.78%, and 97.10% accuracy, respectively. As Ashwood et al. [1] found a lot of false negatives with a threshold, this work is encouraging that machine learning classifiers can help improve the accuracy of diagnosis from the AQ questionnaire. Motivated by these works, we propose a machine-learning based approach to classify ASD from AQ data and demographic information across age. The contributions of this work can be summarized as follows:

1. To the best of our knowledge, this is the first work to propose a machine learning-based approach to classifying ASD with AQ data across age (e.g. train on child data and test on adult).
2. Accuracy of 3 machine learning classifiers, for classifying ASD from AQ questionnaire information, is compared. Namely, random forest [3], a feed-forward neural network [10], and k-nearest neighbor (KNN) [7].
3. Proposed approach outperforms state of the art on the publicly available AQ-10 dataset [14], which contains child, adolescent, and adult AQ information.

2 Experimental Design

2.1 Dataset

To conduct our experiments, we used the AQ-10 dataset [14], which consists of 3 datasets based on the AQ-10 screening tool [6]; 1 for children, adolescents, and adults. Each dataset has attributes including, but not limited to, age, gender, ethnicity, and answers from the AQ questionnaire. All available attributes can be seen in Table 1. There are 292 children with an age range of [4,11] in the child, 104 adolescents with an age range of [12-16], and 704 adults with an age range of [17-64]. Each subject is given a class of either ASD or no ASD.

Table 1. Details of all attributes in AQ-10 dataset [14].

Attribute	**Type**	**Description**
Age	Integer	Years
Gender	String	Male/Female
Ethnicity	String	e.g. Latino, Caucasian, Black, etc.
Born with jaundice	Boolean	Subject born with jaundice
Family member with PDD	Boolean	Immediate family member with PDD
Who is completing test	String	Parent, self, caregiver, medical staff
Country of residence	String	e.g. USA, Brazil, Palestine, etc.
Used screening app before	Boolean	Subject has used app before
Screening Method Type	Integer (0, 1, 2, 3)	toddler, child, adolescent, adult
Questions 1-10	Binary (0,1)	Answer to AQ questions
Screening Score	Integer	Final AQ score

2.2 Experiments

To classify ASD, we propose to use Autism-Spectrum Quotient questionnaire data along with demographic information, specifically from the AQ-10 dataset (Table 1). To evaluate using AQ data, we have selected the following 6 feature sets: (1) All available attributes except for the final screening score resulting in the 19-dimension feature vector $v_1 = [age, gender, ethnicity, jaundice, PDD, test, country, app, method, Q1, \ldots, Q10]$; (2) AQ questions 1-10 and family member with PDD resulting in the 11-dimension feature vector $v_2 = [PDD, Q1, \ldots, Q10]$; (3) the 10-dimension feature vector $v_3 = [Q1, \ldots, Q10]$ with AQ questions but not family history of Genes; (4) AQ questions, ethnicity and family member with PDD resulting in the 12-dimension feature vector $v_4 = [PDD, ethnicity, Q1, \ldots Q10]$; (5) also using all parameters in v_4 but not the genes $v_5 = [ethnicity, Q1, \ldots Q10]$; and (6) family member with PDD resulting in the 1-dimension feature vector $v_6 = [PDD]$. All subject-independent evaluations were conducted by randomly selecting 80% of the data for training and 20% for testing. Accuracy is the evaluation metric used in all experiments.

To evaluate the robustness of the feature sets to classify ASD across different classifiers, we evaluated a random forest [3], a feed-forward neural network, and k-nearest neighbors [7]. Random forest ensembles the results of the large number of decision tree it is made of. For random forest we used 100 trees as the depth.The Knn Algorithm predicts by using the information of data which exists near to each other. Here we have used this $k = 13$ for boundary pf this proximity. The neural networks has 3 hidden layers with 32, 16 an 16 neurons and 1 output layer with 1 output. We used Rectified linear activation functions(Relu) in the hidden layers and in output we have used sigmoid activation.

3 Results

3.1 Within-dataset Evaluation on Child, Adolescent, and Adult

To evaluate child, adolescent, and adult data we used feature vectors v_1, v_2, v_3, v_4, v_5, and v_6 with data from the AQ-10 dataset. As it can be seen in Table 2, random forest and our neural network architecture both performed well on all 3 datasets (child, adolescent, and adult) using v_1. The random forest performed best at 98.8% for all 3, while the neural network had an accuracy of 98.8% on the adolescent dataset as well, but performed slightly worse on child and adult (96.6% and 94.5%, respectively). While k-nearest neighbors performed reasonably well on child and adolescent (81.1% on both), it did not perform well on adult data. Interestingly, both k-nearest neighbors and the neural network had the lowest accuracy on the adult dataset (66.4% and 94.5%, respectively).

When we evaluated feature vector v_2, we found that the neural network had the best performance with 100% on both adolescent and adult datasets, and 98% on the child dataset (Table 3). Again, the adult dataset performed the worst with KNN with 50% accuracy. While the accuracy of random forest decreased by 6.8% across all datasets, it still performed reasonably well with

Table 2. Evaluation for within-dataset using all attributes (v_1).

Classifier	Train	Test	Accuracy
Random Forest	Child	Child	98.8%
	Adult	Adult	98.8%
	Adolescent	Adolescent	98.8%
Neural Network	Child	Child	96.6%
	Adult	Adult	94.5%
	Adolescent	Adolescent	98.8%
K-nearest Neighbors	Child	Child	81.1%
	Adult	Adult	66.4%
	Adolescent	Adolescent	81.1%

Table 3. Evaluation for within-dataset using AQ with PDD in family(PDD) (v_2) and without PDD in family(NPDD(v_3).

Classifier	Train	Test	Accuracy(PDD)	Accuracy (NPDD)
Random Forest	Child	Child	92%	94%
	Adult	Adult	92%	94%
	Adolescent	Adolescent	92%	94%
Neural Network	Child	Child	98%	98%
	Adult	Adult	100%	100%
	Adolescent	Adolescent	100%	98%
K-nearest Neighbors	Child	Child	88%	86%
	Adult	Adult	50%	50%
	Adolescent	Adolescent	88%	86%

92% accuracy. It is interesting that by removing some of the features such as age, gender, and whether the subject had jaundice, the neural network was able to classify ASD with a high degree of accuracy. This suggests that AQ questions along with family history are a strong indicator for classifying ASD, however, to further investigate this, we analyzed feature vector v_3. We removed family history of ASD from the v_2 feature vector and the accuracies remained largely the same compared to v_2, where the results from random forest increased by 2%. Now the most intriguing question came to this, whether the family history really impacts classifying ASD or not. To answer this, we have done some other experiment as well. In Table 4 we have used feature vector v_4 which includes AQ questions, ethnicity and family history and feature vector v_5 which discards the family history information form v_4. From this experiment we can tell the result of v_4 is similar to the result of v_5. This ultimately suggests that the AQ questions and ethnicity are a stronger indicator of ASD compared to family history when automatically classifying ASD with machine learning algorithms.

For our final within-dataset evaluation, we investigated whether family history alone (v_6) can classify ASD. As can be seen in Table 5 the random forest and KNN give the best result among the 3 algorithms but still it is 50% accuracy and for Neural network it is 40%. While it is not as common to try to classify data with 1 feature, this experiment is justified by the heritability of ASD being high with studies finding anywhere from 50%-90% [13]. Although family history can

Table 4. Evaluation for within-dataset using AQ and Ethnicity with PDD in family(PDD) (v_4) and without PDD in family(NPDD(v_5).

Classifier	Train	Test	Accuracy(PDD)	Accuracy (NPDD)
Random Forest	Child	Child	92%	96%
	Adult	Adult	92%	96%
	Adolescent	Adolescent	92%	92%
Neural Network	Child	Child	96.6%	96%
	Adult	Adult	94.9%	91%
	Adolescent	Adolescent	93.2%	94%
K-nearest Neighbors	Child	Child	90%	89%
	Adult	Adult	50%	50%
	Adolescent	Adolescent	90%	89%

Table 5. Evaluation for within-dataset using PDD in family only (v_6).

Classifier	Train	Test	Accuracy
Random Forest	Child	Child	50%
	Adult	Adult	50%
	Adolescent	Adolescent	50%
Neural Network	Child	Child	40.6%
	Adult	Adult	40.6%
	Adolescent	Adolescent	40.6%
K-nearest Neighbors	Child	Child	50%
	Adult	Adult	50%
	Adolescent	Adolescent	50%

be a strong indicator of ASD, our results again suggest that it is not sufficient for use in machine learning classifiers as seen in Tables 3 - 5.

3.2 Cross-dataset Evaluation on Child, Adolescent, and Adult

Along with within-dataset experiments, we also evaluated cross-dataset experiments (e.g. train on child, test on adult). We performed an exhaustive combination of cross-dataset experiments (Table 6). When all features were used (v_1), similar results are obtained, across all 3 classifiers, compared to within-dataset. For the random forest an accuracy of 98.8% was achieved for all experiments, the neural network had an average accuracy of 95.95%, and KNN had an average accuracy of 76.12%. These results suggest that AQ questionnaire data along with demographic information can be used to classify ASD across age. More importantly, it also suggests that this information can be used to predict ASD as there are features in the child dataset that are similar to both adolescents and adults. This is an open question that requires further investigation.

We have also conducted other cross-dataset experiments to learn which attributes have more impact on classifying ASD. Similar to within-dataset, as can be seen in Table 7, only using AQ question answers and family history resulted in a slight decrease in the accuracy for random forest (4.4%), the neural network had the best performance, with an average accuracy of 97.17%, and KNN again

Table 6. Evaluation for cross-dataset using all attributes (v_1).

Classifier	Train	Test	Accuracy
Random Forest	Child	Adult	98.8%
	Child	Adolescent	98.8%
	Adult	Child	98.8%
	Adult	Adolescent	98.8%
	Adolescent	Adult	98.8%
	Adolescent	Child	98.8%
Neural Network	Child	Adult	93%
	Child	Adolescent	98%
	Adult	Child	98.3%
	Adult	Adolescent	94.9%
	Adolescent	Adult	93.2%
	Adolescent	Child	98.3%
K-nearest Neighbors	Child	Adult	66%
	Child	Adolescent	81%
	Adult	Child	81.1%
	Adult	Adolescent	81.1%
	Adolescent	Adult	66.4%
	Adolescent	Child	81.1%

performed the worst with an average accuracy of 76.97%. Similar to the previous experiment which has been done for within-dataset (Tables 3 and 4), the same experiments have been done using feature vectors v_2 and v_3 (Table 7) and also using v_4 and v_5 (Table 8). In each case, we can see that family history can decrease performance in cross-dataset experiments, which similarly replicates the results from our within-dataset experiments. Finally, we have again conducted the experiment with only family history. Here, the result is the same as Table 5, for all cross-dataset experiments: random forest and KNN have an accuracy of 50% and the neural network has an accuracy of 40%.

3.3 Comparison to State of the Art

We also compared our proposed approach to current state of the art. As can be seen in Table 9, our proposed approach is comparable to or outperform state of the art across all datasets (child, adolescent, and adult).

4 Conclusion

We proposed an approach to classifying ASD across age using Autism-Spectrum Quotient questionnaire data along with demographic information. We evaluated random forest, neural network, and KNN classifiers. Results suggest this data is robust to multiple machine learning classifiers and can accurately classify children, adolescents, and adults with ASD. The results are comparable to or outperform state of the art on the AQ-10 dataset. To the best of our knowledge, this is the first work to propose using AQ and demographic information for

Table 7. Evaluation for cross-dataset using AQ with PDD in family(PDD) (v_2) and without PDD in family(NPDD(v_3).

Classifier	Train	Test	Accuracy(PDD)	Accuracy(NPDD)
Random Forest	Child	Adult	94.4%	94.4%
	Child	Adolescent	94.4%	94.4%
	Adult	Child	94.4%	94.4%
	Adult	Adolescent	94.4%	94.4%
	Adolescent	Adult	94.4%	94.4%
	Adolescent	Child	94.4%	94.4%
Neural Network	Child	Adult	96.6%	98%
	Child	Adolescent	98.3%	98%
	Adult	Child	100%	100%
	Adult	Adolescent	96.6%	100%
	Adolescent	Adult	94.9%	100%
	Adolescent	Child	96.6%	98%
K-nearest Neighbors	Child	Adult	50%	50%
	Child	Adolescent	88.1%	86%
	Adult	Child	88.1%	86%
	Adult	Adolescent	88.1%	86%
	Adolescent	Adult	50%	50%
	Adolescent	Child	88.1%	86%

Table 8. Evaluation for cross-dataset using AQ and Ethnicity with PDD in family(PDD) (v_4) and without PDD in family(NPDD(v_5).

Classifier	Train	Test	Accuracy(PDD)	Accuracy(NPDD)
Random Forest	Child	Adult	92%	96%
	Child	Adolescent	94%	96%
	Adult	Child	92%	94%
	Adult	Adolescent	94%	96%
	Adolescent	Adult	92%	96%
	Adolescent	Child	92%	92%
Neural Network	Child	Adult	98%	94%
	Child	Adolescent	93%	98%
	Adult	Child	94%	98%
	Adult	Adolescent	96%	93%
	Adolescent	Adult	89%	94%
	Adolescent	Child	94%	96%
K-nearest Neighbors	Child	Adult	50%	50%
	Child	Adolescent	90%	89%
	Adult	Child	90%	89%
	Adult	Adolescent	90%	89%
	Adolescent	Adult	50%	50%
	Adolescent	Child	90%	89%

Table 9. Comparisons to state of the art across child, adolescent, and adult datasets.

	Child Dataset	**Adolescent Dataset**	**Adult Dataset**
Proposed Approach	98.8%	**100%**	**100%**
Erkan et al. [8]	**100%**	**100%**	**100%**
Omar et al. [12]	92.26%	93.78%	97.10%
Thabtah et al. [15]	N/A	97.58%	99.91%

cross-dataset classification. We performed cross-dataset experiments where we achieved a max accuracy of 98.8%. These results suggest that this data can be used to predict ASD into adulthood from child data. There are some limitations to our study as well. First, we only used one dataset and while our results suggest family history does not classify ASD well with machine learning, these results are inconclusive as it has been shown that it can be a strong indicator due to heritability [13]. Larger and more varied longitudinal studies are required to further investigate this. Secondly, it has been shown that features such as gaze [9] can classify ASD. It is important to compare the AQ and demographic features to other types of features to learn the best features are to classify ASD.

References

1. Ashwood, K., et al.: Predicting the diagnosis of autism in adults using the aq questionnaire. Psychological medicine **46**(12), 2595–2604 (2016)
2. Baron-Cohen, S., et al.: The aq: Evidence from asperger syndrome/high-function autism, males, females, scientists and mathematicians. J of autism and DD (2001)
3. Breiman, L.: Random forests. Machine learning (2001)
4. Burke, L., et al.: Medical and health problems in adults with high-function autism and asperger syndrome. In: Adolescents and adults with ASD. Springer (2014)
5. CDC: Prevalence and characteristics of asd among children aged 8 years–autism and developmental disabilities monitoring network. Surveillance summaries (2016)
6. Dua, D., et al.: UCI ML repository (2017), http://archive.ics.uci.edu/ml
7. Duda, R.O., et al.: Pattern classification and scene analysis. Wiley New York (1973)
8. Erkan, U., et al.: Asd detection with machine learning methods. Current Psychiatry Reviews **15**(4), 297–308 (2019)
9. Fabiano, D., et al.: Gaze-based classification of asd. PRL (2020)
10. Fine, T.: Feedforward network method. Springer Science & Business Media (2006)
11. Lundqvist, L., et al.: Is the aq a valid measure of traits assoc with the autism spec? a rasch validation in adults with and without asd. J of Autism and DD (2017)
12. Omar, K.S., et al.: A machine learning approach to predict ASD. In: ECCE. pp. 1–6
13. Sandin, S., et al.: The familial risk of autism. Jama **311**(17), 1770–1777 (2014)
14. Thabtah, F.: Autism spectrum disorder screening: machine learning adaptation and dsm-5 fulfillment. In: MHI. pp. 1–6 (2017)
15. Thabtah, F., et al.: A machine learning autism classification based on logistic regression analysis. Health information science and systems **7**(1), 12 (2019)
16. Wakabayashi, A., et al.: The aq in japan: a cross-cultural comparison. Journal of autism and developmental disorders **36**(2), 263–270 (2006)
17. zeldovich, L.: Why the definition of autism needs refined. Spectrum, Autism Research News (2018)
18. Zwaigenbaum, L., et al.: Early identification of asd: recommendations for practice and research. Pediatrics **136**(Supplement 1), S10–S40 (2015)