A picture containing text

Description automatically generated

**[Autism Screening](https://www.kaggle.com/datasets/faizunnabi/autism-screening" \t "_blank)**

**Classifying autistic patients based upon the screening results.**

Last Updated: 4 years ago (Version 1)

**About this Dataset**

Name:

Autistic Spectrum Disorder Screening Data for ADS   
(Autistic Spectrum Disorder)

Autistic Spectrum Disorder (ASD) is a neurodevelopment condition associated with significant healthcare costs, and early diagnosis can significantly reduce these. Unfortunately, waiting times for an ASD diagnosis are lengthy and procedures are not cost effective. The economic impact of autism and the increase in the number of ASD cases across the world reveals an urgent need for the development of easily implemented and effective screening methods. Therefore, a time-efficient and accessible ASD screening is imminent to help health professionals and inform individuals whether they should pursue formal clinical diagnosis. The rapid growth in the number of ASD cases worldwide necessitates datasets related to behavior traits. However, such datasets are rare making it difficult to perform thorough analyses to improve the efficiency, sensitivity, specificity and predictive accuracy of the ASD screening process. Presently, very limited autism datasets associated with clinical, or screening are available and most of them are genetic in nature. Hence, we propose a new dataset related to autism screening of adults that contained 20 features to be utilized for further analysis especially in determining influential autistic traits and improving the classification of ASD cases. In this dataset, we record ten behavioral features (AQ-10-Adult) plus ten individuals characteristics that have proved to be effective in detecting the ASD cases from controls in behavior science.

Source:

Fadi Fayez Thabtah  
Department of Digital Technology  
Manukau Institute of Technology,  
Auckland, New Zealand  
[fadi.fayez@manukau.ac.nz](mailto:fadi.fayez@manukau.ac.nz)

**Data Type**:

Multivariate OR Univariate OR Sequential OR Time-Series OR Text OR Domain-Theory  
Nominal / categorical, binary and continuous

**Task**: Classification

**Attribute Type**: Categorical, continuous and binary

**Area**: Medical, health and social science

**Format** Type: Non-Matrix

**Number of Instances (records in your data set)**: 704

**Number of Attributes (fields within each record**): 21

**Relevant Information:** For Further information about the attributes/feature see below table.

**Attribute Information:**

Attribute Type Description

|  |  |
| --- | --- |
| Age | Age Number Age in years |
| Age  Gender  Ethnicity  Born/aundice Boolean (yes or no)  PDD  Who  COResidence  used a screening app  Screening Method  (0=toddler, 1=child, 2= adolescent, 3= adult)  Question 1  Question 2  Question  Question 4  Question 5  Question 6  Question 7  Question 8  Question 9  Question 10  Screening Score Integer | 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) Whether the case was 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, medical staff, clinician ,etc.  Country of residence String List of countries in text format  Used the screening app before Boolean (yes or no) Whether the user has used a screening app  Screening Method Type Integer (0,1,2,3) The type of screening methods chosen based on age category (0=toddler, 1=child, 2= adolescent, 3= adult)  Question 1 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 2 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 3 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 4 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 5 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 6 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 7 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 8 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 9 Answer Binary (0, 1) The answer code of the question based on the screening method used  Question 10 Answer Binary (0, 1) The answer code of the question based on the screening method used  Screening Score Integer The final score obtained based on the scoring algorithm of the screening method used. This was computed in an automated manner. |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Relevant Papers:

1) Tabtah, F. (2017). Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment. Proceedings of the 1st International Conference on Medical and Health Informatics 2017, pp.1-6. Taichung City, Taiwan, ACM.

2) Thabtah, F. (2017). ASDTests. A mobile app for ASD screening. [www.asdtests.com](http://www.asdtests.com/) [accessed December 20th, 2017].

3) Thabtah, F. (2017). Machine Learning in Autistic Spectrum Disorder Behavioural Research: A Review. To Appear in Informatics for Health and Social Care Journal. December, 2017 (in press)

Citation Request:

Please refer to the Machine Learning Repository's citation policy