

## A machine learning approach to predict autism spectrum disorder (ASD) for both children and adults using feature optimization

Khandaker Mohammad Mohi Uddin<sup>1</sup>, Hasibur Rahman<sup>1</sup>, Mahadi Hasan<sup>1</sup>, Fatema Akter<sup>2</sup>, Suman Chandra Das<sup>3</sup>

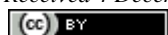
<sup>1</sup>Department of Computer Science and Engineering, Dhaka International University, Dhaka- 1205, Bangladesh

<sup>2</sup>Department of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh

<sup>3</sup>Bangabandhu Sheikh Mujib Medical College, Faridpur, Bangladesh

E-mail: jilanicejnu@gmail.com, Itzhasib1@gmail.com, Mehedi036@gmail.com, fatema.jnu.cse@gmail.com, sumanfmc21@gmail.com

Received 4 December 2022; Accepted 10 January 2023; Published online 28 January 2023; Published 1 June 2023



### Abstract

A central nervous system known as an Autism Spectrum Disorder (ASD) has long-term effects on a person's capacity for engagement and interaction with others. Since its symptoms often manifest in the first two years of life, autism is considered to be a behavioral condition that can be identified at any point in a person's life. This study investigated the potentiality of machine learning techniques such as Logistic Regression, Random Forest, Multinomial Naive Bayes (MNB), Bernoulli Naive Bayes (BNB), Support Vector Machine (SVM), and Gaussian Naive Bayes (GNB) to predict ASD using some health parameters. There are 292 instances and 21 attributes in the first dataset linked to the screening for ASD in children. The adult individuals in the second dataset had a total of 704 occurrences and 21 characteristics related to ASD detection. In order to achieve the highest accuracy possible from the machine learning models, feature optimization is used in this study along with other preprocessing approaches. The findings overwhelmingly support the notion that Random Forest performs better on all of these datasets, with the greatest accuracy (100%) for data on Autistic Spectrum Disorder (ASD) in children and adults, respectively.

**Keywords** Autism Spectrum Disorder (ASD); Machine Learning; Feature Optimization; Random Forest; Logistic Regression.

Network Biology  
ISSN 2220-8879  
URL: <http://www.iaees.org/publications/journals/nb/online-version.asp>  
RSS: <http://www.iaees.org/publications/journals/nb/rss.xml>  
E-mail: [networkbiology@iaees.org](mailto:networkbiology@iaees.org)  
Editor-in-Chief: Wenjun Zhang  
Publisher: International Academy of Ecology and Environmental Sciences

### 1 Introduction

ASD, a neurodevelopmental disorder, has a significant influence on many people's day-to-day life especially human interaction (Wing, 1997). The impacts of ASD symptoms degrade social and interaction skills (Vaishali, 2018). As per studies, 1 in 160 children in the world suffers from autism (World Health Organization, 2019). The signs of autism are more noticeable and easier to spot in children between the ages of two and three (Raj,































- Vakadkar K, Purkayastha D, Krishnan D. 2021. Detection of autism spectrum disorder in children using machine learning techniques. *SN Computer Science*, 2(5): 1-9
- Vaishali R, Sasikala, R. 2018. A machine learning based approach to classify autism with optimum behaviour sets. *International Journal of Engineering and Technology*, 7(4): 18
- Wing L. 1997. The autistic spectrum. *The Lancet*, 350(9093): 1761-1766
- Wall DP, Kosmicki J, Deluca TF, Harstad E, Fusaro VA. 2012. Use of machine learning to shorten observation-based screening and diagnosis of autism. *Translational psychiatry*, 2(4): e100-e100
- Wu W, Mallet Y, Walczak B, Penninckx W, Massart DL, Heurding S, Erni F. 1996. Comparison of regularized discriminant analysis linear discriminant analysis and quadratic discriminant analysis applied to NIR data. *Analytica Chimica Acta*, 329(3): 257-265
- Zayed KMES, Ibrahim MA, Farid MN, El-Shourbagy OESO, Tarkan RS. 2022. Zinc Levels Assay in Children with Autism Spectrum Disorder by Quantum Magnetic Resonance Analyzer and Direct Colorimetry. Europe PMC, preprint