IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 4, December 2023

A Machine Learning Framework for Early-Stage Detection of Autism Spectrum Disorders

Muhammed Danish P, Dr. Kavitha S M, Mr. Sreekanth S, Ms. Anusree C, Ms. Dilna VC
Department of Computer Science and Engineering
AWH Engineering College, Kuttikattor, Calicut, Kerala, India
Dr. APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala, India

Abstract: *Autism Spectrum Disorder (ASD) is a neuro developmental condition that significantly impacts the daily lives of those affected. While complete eradication remains challenging, early interventions can help alleviate its severity. This study presents a novel framework for assessing various Machine Learning (ML) techniques to detect ASD early. The framework incorporates four Feature Scaling (FS) methods—Quantile Transformer (QT), Power Transformer (PT), Normalizer, and Max Abs Scaler (MAS). Subsequently, the scaled datasets are subjected to classification using eight ML algorithms: Ada Boost (AB), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (GNB), Logistic Regression (LR), Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA). Experiments are conducted on four established ASD datasets categorizing individuals by age groups—Toddlers, Adolescents, Children, and Adults. By evaluating classification outcomes through diverse statistical metrics such as Accuracy, Receiver Operating Characteristic (ROC) curve, F1-score, Precision, Recall, Mathews Correlation Coefficient (MCC), Kappa score, and Log loss, optimal classification methods and FS techniques are determined for each dataset. Results show AB achieving the highest accuracy of 99.25% for Toddlers and 97.95% for Children, while LDA achieves 97.12% for Adolescents and 99.03% for Adults. Notably, using normalizer FS for Toddlers and Children, and QT FS for Adolescents and Adults yield the best accuracies. Furthermore, ASD risk factors are quantified, and attribute importance is ranked employing four Feature Selection Techniques (FSTs)—Info Gain Attribute Evaluator (IGAE), Gain Ratio Attribute Evaluator (GRAE), Relief F Attribute Evaluator (RFAE), and Correlation Attribute Evaluator (CAE). These comprehensive evaluations underscore the significance of fine-tuning ML methodologies for accurate ASD prediction across different age groups. The detailed analysis of feature importance presented herein can aid healthcare practitioners in ASD screening, offering promising advancements compared to existing detection approaches.*

Keywords: Autism spectrum disorder, machine learning, classification, feature scaling, feature selection technique

DOI: 10.48175/IJARSCT-15183
ISSN 2581-9429 IJARSCT