

# A Comparative Study on Autism Among Children Using Machine Learning Classification



Ainie Hayati Binti Noruzman, Ngahzaifa Ab Ghani,  
and Nor Saradatul Akmar Zulkiffi

**Abstract** Autism Spectrum Disorder (ASD) is a neurodevelopment that affects communication and behavior in humans. It is a condition associated with a complex brain disorder, leading to significant changes in a human being's social interaction and behavior. Typically to detect toddlers who have ASD through screening tests is very expensive and time-consuming. Typically, detecting toddlers who have ASD through screening tests is very expensive and time-consuming. However, with machine learning technology today, autism can be diagnosed efficiency and accuracy. This study aims to analyze and make a comparison on which prediction model that gives a high accuracy after the feature selection. The importance of attributes is investigated using correlation and the predictive models are constructed for the detection of this disorder in children. The dataset consists of 1054 instances and each instance includes 19 attributes. Experimental results clearly show that using feature selection with 10 attributes can lead the impact of accuracy with predictive model of Random Forest (RF) returns the highest accuracy with 94.78%. The findings also indicated that the number of questions in screening tools can be reduced and give an impact with the good results.

**Keywords** Machine learning · Classification · Feature selection

---

A. H. B. Noruzman (✉) · N. A. Ghani · N. S. A. Zulkiffi  
Faculty of Computing, University Malaysia Pahang, Pekan, Malaysia  
e-mail: [pcp20001@student.ump.edu.my](mailto:pcp20001@student.ump.edu.my)

N. A. Ghani  
e-mail: [zaifa@ump.edu.my](mailto:zaifa@ump.edu.my)

N. S. A. Zulkiffi  
e-mail: [saradatulakmar@ump.edu.my](mailto:saradatulakmar@ump.edu.my)

A. H. B. Noruzman · N. A. Ghani  
Centre for Software Development and Integrated Computing, University Malaysia Pahang,  
Pekan, Malaysia