

CHAPTER 5

Machine Learning Techniques to Predict Autism in Children

*Norliza Mohd Noor, Syaidatul Amaleena Rossli,
Nurul Amirah Mashudi, and Norulhusna Ahmad*

5.1 INTRODUCTION

According to World Health Organization (WHO), one out of 160 children suffer from Autism Spectrum Disorder (ASD) (Zeidan et al., 2022); other research recorded a higher number. ASD is a neurodevelopmental disorder that usually develops in early childhood and continues on an individual's lifespan. Furthermore, ASD is identified by social problems, communication issues, and repetitive or limited behaviours, interests, or activities (Iandolo et al., 2020). Autism-related issues can provide particular problems in education, career, relationships, and independent living as individuals age from childhood to adolescence and adulthood. With the proper support and accommodations, however, many autistic individuals can live productive and meaningful lives. Genetic factors significantly influence autism. Recent research shows that the emergence of ASD is significantly influenced by genetic factors involving families, twins, and siblings that consistently revealed a higher rate of ASD among kin (Otaru et al., 2022).

Therapies, behavioural treatment, and parent skills training programs can reduce communication and social behaviour difficulties, positively impacting the well-being and quality of life of persons with ASD and their caregivers. WHO also stated that

there is no known cure for ASD. The diagnostic process for ASD is labour-intensive and expensive (Thabtah, 2019). Hence, the development of autistic children can benefit from early diagnosis and treatment. The treatment and therapeutic measures had a beneficial effect on enhancing attention in children diagnosed with mild ASD.

5.2 RELATED WORKS

Various researchers employed machine learning algorithms for the ASD classification, including Naïve Bayesian, which has proven an efficient classifier due to its easy-to-build nature (Reeta et al., 2018). Moreover, the authors showed that the classifiers' performance improved with larger training sets. A study by Talabani and Avci (2018) focused on using different Support Vector Machines (SVM) kernels for the same purpose. The finding concludes that the Polynomial Kernel (PK) and Pearson VII Universal Kernel (PUK) managed to obtain 100% accuracy; meanwhile, normalized polynomial kernel (NP) and radial basis function (RBF) produced 95.547% and 99.315% accuracy, respectively.

Tyagi et al. proposed kNN, Linear Discriminant Analysis (LDA), SVM, Classification and Regression Tree (CART), Naïve Bayes, and linear regression algorithms to classify ASD features for early ASD diagnosis (Tyagi et al., 2018). The methods obtained not more than 80% accuracy, which are 67.56%, 72.20%, 70.60%, 69.17%, 70.78% and 72.02%, respectively. The study employed a ratio of 70:30 for training and testing datasets. Research is being conducted to effectively diagnose autism in early childhood due to its high implications on children with ASD (Wu et al., 2022). However, the symptoms are hardly distinguishable for a child. Hence, it is challenging to

identify the symptoms in a child. Thus, a mobile application was developed to classify if a child showing ASD symptoms in order to facilitate early diagnosis.

In the study by Thabtah et al. (2019), subjects were instructed to respond to ten questions using a mobile application. The application allocated a score to each response, which indicated the severity level of autism based on ethnicity, gender, and age. The dataset collected through the utilization of the mobile application has the potential to improve the development of Machine Learning (ML) models that can predict autism in children. Multiple ML algorithms were employed to classify ASD and early diagnose ASD in children accurately.

This study aims to analyse the performance of classification methods for children with ASD. The performance was evaluated based on accuracy, sensitivity, specificity, and F-measure. We employed several ML methods, including k-Nearest Neighbour (kNN), Linear Support Vector Machine (SVM), Decision Tree (J48), AdaBoost, Bagging, and Stacking. This study also aims to analyze Bayesian Network as a classifier and Directed Acrylic Graph (DAG).

5.3 MATERIALS AND METHODS

5.3.1 Dataset

The datasets were obtained from the Department of Digital Technologies, Manukau Institute of Technology, Auckland, New Zealand. The author developed a mobile app for ASD screening. In addition, the author adopted ML in identifying ASD patients. A total of 292 children which were from age group 4 to 11 years old, were screened for ASD in this study. The following result was tabulated in Table 5.1.