

Survey on Early Detection of Autism Using Data Mining Techniques

B. Ida Seraphim, Lavi Samuel Rao, Shiwani Joshi

Computer Science and Engineering, S.R.M Institute of Science and Technology, Chennai, India

*Corresponding Author Email: idaseraphim.b@ktr.srmuniv.ac.in,

Abstract

The children of today are the future of the nation and there are many hurdles in their development like ASD. ASD (Autism Spectrum Disorder) is a neurological disorder which has a lifetime impact on the basic skills and talents of a human being. The foremost goal of the paper is to know the number of people suffering from autism and the various symptoms of autism.

Keywords: Autism Spectrum Disorder, J48, Decision Tree, SVM, IB1

1. Introduction

Data mining is the interdisciplinary field of extracting patterns from a given data set with the help of intersection of machine learning, statistics and database systems. It is a “buzzword” which is often used in cases involving large data sets for information processing (collection, extraction, warehouse, etc.)

And as well as in computer applications (such as business intelligence, artificial intelligence and machine learning).

Data mining is a multidisciplinary field which is semi-automatic analysis of large sets of data anomalies, interesting patterns and dependencies. Data mining techniques can be classified as supervised and unsupervised. Supervised learning is a type of machine learning algorithm which uses a known dataset. Supervised learning has two types of algorithms:

- 1) Classification (for categorical values)
- 2) Regression (for continuous values)

Autism Spectrum Disorder (ASD) is a name for a range of diseases having similar conditions including Asperger Syndrome, that affect a person’s social interaction, communication, interests and behaviour. In case of children the symptoms are present before the age of three. Although there is no “cure” for ASD but certain therapies can be used to improve the victim’s condition. The result of the analysis can help us in understanding the health condition, the skills and education styles to be adopted for such people. It can also help us in taking precautions against certain practices which can worsen the victim’s condition. Data mining provides us with various techniques and in this paper we would be focussing on the classification technique.

2. Autism Spectrum Disorder

Autism spectrum disorder is a group of complex disorders of the brain developments-ESE disorders include difficulty in social interactions, repetitive behaviours and communications.

These disorders root at very young age of 2-3 years and hence, symptoms of ASD may be seen at a very young age. Autism Spectrum Disorders are common in boys than in girls. There are the following types of ASD:

A) Autistic Disorder

Also known as classic “autism”. People suffering from this disorder have significant language delays, social and communication challenges, unusual behaviours and interest and may also have intellectual disability.

B) Asperger Syndrome

People suffering from this disorder have milder symptoms of autistic disorder and have social challenges and unusual behaviours. Although they do not suffer from language or intellectual disability.

C) Pervasive Developmental Disorder-Not Otherwise Specified

Also known as atypical autism. The people who meet only some of the symptoms of autism or Asperger Syndrome are classified under this category.

Most of the common behaviours shown by children suffering from autism are:

- Aggression (hitting, kicking, scratching, etc.)
- Self-injury (head-banging, hand-biting, hitting walls, etc.)
- Disruption (interrupting, yelling, knocking things over, etc.)
- Elopement (wandering, bolting, etc.)
- Stereotypy (hand-flapping, rocking, toe-walking, etc.)
- Tantrums (crying, screaming, defiant behaviour, etc.)
- Non-compliance (disobeying directions, whining, etc.)
- Obsession (repeatedly talking about the same topic, perseveration, etc.)

3. Related works

The vital goal of the paper is to review the autism problem, to detect the levels of autism with the help of data mining, classification algorithms and also to study the autism student’s performance using data mining methodologies.

Libero et al[3] has utilized structured MRI Diffusion Tensor Imaging, proton magnetic resonance .Information on cortical thickness can be determined from the structural MRI,

White matter connectivity can be determined from DTI and neurochemical concentration can be determined from 1H-MRS which is used to classify ASD and TD subjects. They identified fractional anisotropy, radial diffusivity and cortical thickness as significant predictors of the disorder. They identified fractional anisotropy, radial diffusivity and cortical thickness as significant predictors of the disorder. They applied decision tree classification on the predictors obtained from the three modalities and they achieved the highest accuracy of 91.9%.

Sumi Simon et al., [7] performed an empirical evaluation on autistic data and used many classification algorithms like SVM, J48, MLP and IB1 on the behavioural data to classify the autistic children. The author collected autistic data from National Institute for the Mentally Handicapped and extracted relevant features from the data using Relief and Runs Filtering (Feature Selection algorithms). Of all the classifiers SVM showed to be the best because it provided the highest accuracy at low error rate.

Gondy A. Leroy et al., [1] worked on therapy success with autistic children. He made use of ID3 for decision tree building for classification and also used association rule to predict the levels of appropriate or inappropriate behaviour .The decision tree gave 69% accuracy when high or low levels of appropriate behaviour could occur before treatment (Pre Treatment) and gave an increased accuracy (76%) during treatment and then he found it to decreased again after treatment (Post Treatment).The decision tree was unable to predict post treatment accurately. The decision tree provided same accuracy in the levels of inappropriate behaviour also. Since the decision tree was not able to provide accuracy post treatment, he made use of association rule mining. The rules found in association rule mining confirm the strongest results for predicted levels of appropriate behaviour and also add details for predicting levels of inappropriate behaviour.

Mythili et al. [4] detects the level of autism with data mining classification algorithms like neural network, SVM and fuzzy logic to classify the autistic children.

Mythili et al. [5] proposed an improved feature selection algorithm to predict the learning skills of the autism affected kids. SVM, J48, multilayer perceptron and IB1 classifiers along with the filters and wrappers are used to build the improved feature selection model to achieve better accuracy and to detect the learning skills of the autistic children.

Joao F. Santos et al. [2] proposed an early prediction of ASD based on the analysis of pre-verbal vocalizations of the autistic children. The acoustic prosodic features are extracted and provided as a training data to the SVM and probabilistic neural network classifiers and it is concluded that classification of autism is achieved by analysing the pre-vocalization of the autism affected children. The SVM and neural network classifier achieves ninety seven percent accuracy.

Siriwan Sunsirikul et al. [6] used associative classification mining to detect the behavioural patterns of the autism affected kids. The dataset used is categorized by the doctors as – autism and pervasive development disorder. The methodology included data pre-processing, rules generation, building the classifier and cross validation of the test data with the trained data. The classifier was effectively able to classify the behavioural patterns between the autism and pervasive developed disorder children with relatively high accuracy.

4. Conclusions

Autism Spectrum Disorder being an incurable neuro- development disorder, needs to be detected at an early stage .Its detection at an early stage will not only prepare the parents but will also help in devising education tools and tricks to help them educate and nurture their children. Classification is relatively the most suitable

technique for detecting the disease. Although, there have been other techniques like clustering which has been applied in four fields like environment, biotic, irrigated area, crop yields. There has also been further improvement in early detection and finding causes of this disease of this disease by using algorithms like neural network, SVM and fuzzy logic. For better and accurate results in the future, certain other techniques may also be employed so that the disease can be detected during the pregnancy stage itself and hence, further tools could be developed very much earlier to help such children.

References

- [1] G. Leroy, A. Irscher, and M.H. Charlop-Christy, "Data Mining Techniques to Study Therapy Success with Autistic Children", 2006 International Conference on Data Mining, 26 - 29 June 2006, Monte Carlo Resort, Las Vegas, USA.
- [2] Joao F. Santos, Nirit Brosh, Tiago H. Falk, Lonnie Zwaigenbaum, Susan E. Bryson, Wendy Roberts, Isabel M. Smith, Peter Szatmari and Jessica A. Brian, " Very Early Detection of Autism Spectrum Disorders Based on Acoustic Analysis of Pre-Verbal Vocalizations of 18-Month Old Toddlers," IEEE International Conference on Acoustics Speech and Signal Processing, pp. 7567-7571, 2013.
- [3] L.E. Libero, T.P. DeRamus, A.C. Lahti, G.Deshpande , and R.K. Kana, 'Multimodal neuroimaging based classification of autism spectrum disorder using anatomical, neurochemical, and white matter correlates' .Cortex, vol. 66, 2015, pp. 46-59.
- [4] M. S. Mythili & A. R. Mohamed Shanavas, "A Study on Autism Spectrum Disorder using Classification Techniques," International Journal of Soft Computing and Engineering, vol.04, pp. 88-91, 2014.
- [5] M. S. Mythili & A. R. Mohamed Shanavas, "An Improved Feature Selection (IFS) Algorithm for Detecting Autistic Children Learning Skills," Bioscience Biotechnology Research Asia, vol.12, pp. 499-505, 2015.
- [6] Siriwan Sunsirikul & Tiranee Achalakul, " Associative Classification Mining in the Behavior Study of Autism Spectrum Disorder," IEEE International Conference on Computer and Automation Engineering, vol.03, pp. 279-283, 2010.s
- [7] Sumi Simon, Chandra J and Saravanan N, "Empirical Evaluation of Data Mining Classification Methods for Autistic Children".
- [8] T. Padmapriya, V.Saminadan, "Performance Improvement in long term Evolution-advanced network using multiple input multiple output technique", Journal of Advanced Research in Dynamical and Control Systems, Vol. 9, Sp-6, pp: 990-1010, 2017.
- [9] S.V.Manikanthan and K.Baskaran "Low Cost VLSI Design Implementation of Sorting Network for ACSFD in Wireless Sensor Network", CiiT International Journal of Programmable Device Circuits and Systems,Print: ISSN 0974 – 973X & Online: ISSN 0974 – 9624, Issue : November 2011, PDCS112011008.
- [10] T. Padmapriya and V. Saminadan, "Distributed Load Balancing for Multiuser Multi-class Traffic in MIMO LTE-Advanced Networks", Research Journal of Applied Sciences, Engineering and Technology (RJASET) - Maxwell Scientific Organization , ISSN: 2040-7459; e-ISSN: 2040-7467, vol.12, no.8, pp:813-822, April 2016.
- [11] S.V.Manikanthan and K.srividhya "An Android based secure access control using ARM and cloud computing", Published in: Electronics and Communication Systems (ICECS), 2015 2nd International Conference on 26-27 Feb. 2015, Publisher:IEEE, DOI:10.1109/ECS.2015.7124833.