International Journal on Emerging Technologies **10**(3): 422-427(2019)

ISSN No. (Print) : 0975-8364
ISSN No. (Online) : 2249-3255

A Survey on Predicting Autism Spectrum Disorder using Machine Learning Techniques

N. Priya[1] and C. Radhika[2]
[1]Associate Professor, Department of Computer Science,
S.D.N.B. Vaishnav College for Women, University of Madras, Chrompet, Chennai, (Tamil Nadu), India.
[2]Research Scholar, Department of Computer Science,
S.D.N.B. Vaishnav College for Women, University of Madras, Chrompet, Chennai, (Tamil Nadu), India.

(Corresponding author: N. Priya)
(Received 06 July 2019, Revised 17 September 2019 Accepted 05 October 2019)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Autism Spectrum Disorder (ASD) is a developmental disorder of an individual's behaviors, communication, learning skills, and social interaction. The nature of ASD varies among individuals and brain-based or neurobiological disorder of development. Recent studies show that around one in sixty-eight people is diagnosed with ASD and overall 18 million people suffering from autism in India. Objective of the study is to identify ASD in the early stage to improve brain development and also provide the awareness of ASD to the parents and the caretaker. Nowadays Machine Learning techniques play a vital role in predicting Autism Spectrum Disorder. It is an application of Artificial Intelligence (AI) that focuses on the development of computer programs that can access data and use it learn for themselves. Several Machine Learning techniques are proposed by many researchers for diagnosing the various types of ASD in a fast and accurate manner. This paper proposes to classify and review the various Machine Learning techniques and discuss the features of ASD and performance evaluation using different metrics and identify the promising direction for future research and this review shows the Random forest algorithm gives better results compare with other algorithms.

Keywords: Autism Spectrum Disorder, Data Mining, Machine Learning.

I. INTRODUCTION

A. Autism Spectrum Disorder

Autism is a chronic disability that affects individuals who fail to make normal relations with other individuals. Nowadays, ASD is predicated at eighteen months based on the kid's social interaction but many kids are detected only at their three years or the first step to the schools. The different types of Autistic disabilities are Autistic Disorder, Asperger's Disorder, Childhood Disintegrative Disorder, Rett's Disorder and Pervasive Developmental Disorder-Not Otherwise Specified [19,22].

Autistic Disorder: This type of individuals delay with social interaction, stereotyped patterns or behaviors interests and restricted repetitive. This disorder affects the individual's motor skills, lack of eye contact and may have a fixation on specific parts of an object that are viewed by the individuals.

Asperger's syndrome: Asperger's syndrome is characterized as a unique ASD because they have normal communication/language development, though they have difficulty in understanding normal conventional social rules or lack of understanding in other's behaviors.

Childhood Disintegrative Disorder (CDD): CDD is also called as regressive autism. It is characterized by 2 to 4 years of normal development followed by the onset of autism symptoms like severe and sudden reversals in language, motor skills, social interaction, and behaviors.

Rett's Disorder: Rett's Disorder is a neurological disorder that affects only girl children. The children with Rett's syndrome haveslow growth when they are between 12 to 18 months and their head is usually small in size. This disorder is named MICROCEPHALY by the doctors and slow growth shows that the brain doesn't develop properly. Depending on the child Rett's syndrome various and some sequence of symptoms are breathing, sleep problems, teeth grinding, stereotyped patterns and slowly they lose their abilities in each level of their age.

Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS): This type of Child have impairments in social communication such as interactions or they often have repetitive behaviors like Hand-flapping, rocking, twirling or jumping. They may have a lack of eye contact, trouble in controlling emotions and high-pitched.

B. Characteristic Features of ASD

Every ASD child has unique abilities and the most common characteristics of ASD [9] are Impaired Social Interaction, Behaviour Patterns, Cognitive Problems, and Sensory Aspects are shown in Fig. 1.

C. Machine Learning Techniques

Machine Learning techniques which embodies the principle of automatic correlations and learning from the new algorithm. The present study focused on the application of various Machine learning techniques over the Autism Spectrum Disorder diagnosis. Fig. 2. Presents a list of Machine Learning techniques used in current trends [1,2 5].

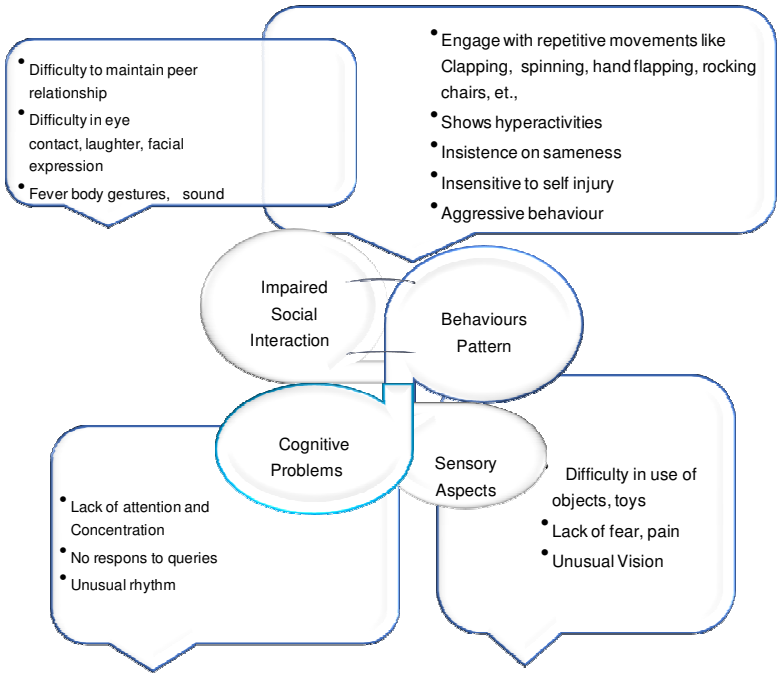

Fig. 1.

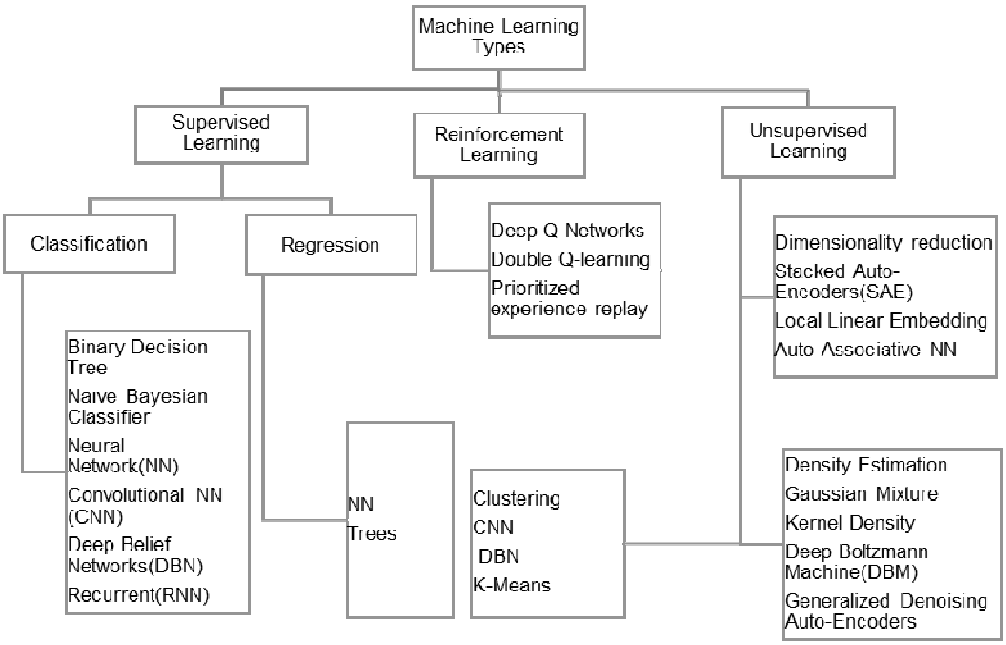

Fig. 2. List of Machine Learning techniques.

II. LITERATURE REVIEW

The study aims to identify Autism Spectrum Disorder for children with various levels, different motors and using various Machine Learning techniques. The issues are related to overlapping notions in Data Mining and all techniques are described by Shhab, that is used to solve specific Data Mining problems as highlighted by scientific papers published in the past 10 years. The experimental results of four different AI techniques: Instance-based Learning algorithms, Decision Trees, Rule Induction and Artificial Neural Networks, ANN such as Multi-Layer Perceptron (MLP) algorithm are competitive methods for predictive toxicology data mining [1]. Related works are proposed by Bishop-Fitzpatrick, *et al.* [2] Autism children had higher rates of health issues like diabetics, cancer, cardiovascular, digestive, urinary, motor problems and respiratory in their health records. Machine Learning classifier (Random Forest) is applied and to classify participants into two groups based on their ICD- codes, V-codes, E-codes, and the Random Forest algorithm compacts overfitting by learning multiple decision trees.10-fold cross-validation is used for testing and evaluation. Wekatool is used to trained and tested the model with

Random Forest. Elizabeth *et al.,* [3] proposed a K-means cluster to detect behavior, learning outcomes, providing personalized therapeutic interventions with maximum efficacy and minimum time and cost. Chan, *et al.*, [4] ANOVA test is used to analyze the long-term therapeutic effects of a patented herbal formula that is measured by neuropsychological tests and daily behavioral checklists. This analyst proves that the intranasal application of the herbal formula may have an improvement in the executive functions of ASD children. Thabtah, proposed Machine Learning techniques to tackled the ASD problem as a classification problem and critically analyzed their advantages and disadvantages. The development of new intelligent diagnostic tools based on machine learning by replacing the handcrafted rules inside the ASD diagnostic tools will fulfill the new criteria of the DSM [5]. Dawson *et al.* in this case study author proposed a lack of social interaction began to emerge during the first 6^{th} months to 24^{th} months of infant autism [6].

III. COMPARISON OF CLASSIFICATION ALGORITHMS

A. Identifying autism using the Random Forest algorithm
Ramya, *et al.*, describes the new DRN (Deep Learning, Random Forest, Naïve Bayes) model to predicted different level of ASD, DRN models implemented in Rapid Miner to find the accuracy, classification error and executed time with other hybrid models like Ada Boost, Bagging, Vote, Stacking and Bayesian Boosting. DRN with accuracy 98.6% [7]. Geetha Ramaniand & Sivaselvi proposed a supervised learning technique for detecting ASD and TD (Typically Developing), describes new centrality measure (weighted leverage variants), Sparsity thresholding, Random Tree classifier with Fisher feature selection achieved a higher accuracy 88.46% in the detection of ASD and TD. Table 1 describes the Comparative study of the Random Forest algorithm in different datasets [8].

B. Identifying autism using Support Vector Machines
Sumi *et al.* compares the existing methods for data collection, pre-processing and different classifiers for predicting levels of autism. Empirical evaluation proves that SVM, J48, BVM, and Decision tree are the best suit to analysis behavioral and learning skills of the autistic data and SV provides high accuracy 95-97% and low error rate. They mainly deal with social relationships and reciprocity, emotional responsiveness, speech-language and communication, behavior patterns, cognitive components, and sensory aspects [9]. Bi *et al*. [10] proposed a random SVM cluster and several graph metrics of brain functional connectivity to classify ASD and Typical Control(TC). This method has the highest performance and classification accuracy based on the optimal feature set-Inferior Frontal Gyrus (IFG), hippocampus, and praecuneus with an accuracy of 96.15%. Diagnosis of the ASD status is achieved using Support Vector Machines (SVM), KNN, random forest machine learning methods using with ASD adolescent scan data and the performance of these methods are compared. Performance and Accuracy rates-100% are achieved as a result of binary classification with 10 fold cross-validation (cv) [11]. Table 2 describes the performance analysis of Support Vector Machines in different Autism Dataset.

Table 1: Comparative study of Random Forest algorithm in Autism Dataset.

Author	Dataset	Pre-processing /techniques Tool	Algorithm	Metrics used
Ramya, *et al.* [7]	VAERS	Null values, Redundant Values, and Missing Values Split Validation Tool-Rapid Miner	DRN	Classification Error, Kappa Statistic, Mean Absolute and Root Mean Squared Error
Geetha Ramaniand & Sivaselvi [8]	Data set collected from UCLA's Centre for Autism Research and Treatment	DISCO	Random tree	Fisher, Runs, Relief-(Feature selection), New centrality measure, Neuroimaging techniques, RS-FMRI

Table 2: Performance of Support Vector Machines in Autism Dataset.

Author	Dataset	Pre-processing /techniques Tool	Algorithm	Metrics used
Sumi *et al.* [9]	National Institute for the Mentally Handicapped	Missing & Redundant values, Fisher Filtering, Step disc	SVM, J48, BVM.	relief, Runs filtering
Bi *et al.* [10]	The Autism Brain Imaging Data Exchange -2014	Slicing timing, Normalizing with the echo-planar imaging template, Smoothing, Temporal filtering, Removing covariates	Random SVM Cluster	graph metrics like local efficiency, shortest path
Demirhan [11]	ASD adolescent scan data from UCI Machine Learning Repository is used -2017	MATLAB	KNN, SVM, RF	Sigmoid, polynomial, radial basis function (RBF)kernel function,10fold CV

C. Identifying autism using smart devices

A Machine Learning algorithm to practical and fully automatic ASD screening solution and the result of the proposed solution makes it potentially very helpful to ASD children. They mainly deal with different features set such as Interaction Recording, Monitoring Recording, Picture, Sentence Story Word-[Acoustic and Linguistic Joint Analysis]. They applied the SVM algorithm with an accuracy of 88.9% [12]. Machine Learning techniques determine a set of conditions that together prove to be predictive of Autism Spectrum Disorder, which then can used by Physicians for their formal screening for ASD. In this study, author deals with different features set such as Learning disability, Developmental delay, Speech or problems, Birth weight, Prematurely born, Physical activity, Attendance to religious events, Body mass index and achieved accuracy of 90.2% using the J48 algorithm [13].

A fuzzy unordered rule induction algorithm (FURIA) that has been evaluated for ASD traits detection. This FURIA developed screening models and then utilizes the models to detect the possibility of autistic traits in new individuals. Empirical results demonstrate the high performance of the fuzzy data mining model concerning predictive accuracy rates of 91.35% [14]. A system for monitoring autistic child behaviors by analyzing accelerometer data collected from wearable mobile devices. They mainly deal with three communication and interaction behaviors (goodbye, drinking, clapping) and one stereotypical activity (hand flapping). In this study authors analyzing behaviors by using a novel algorithm called DTWDir and it is based on calculating displacement and direction between two signals. Machine Learning classifier KNN, $1 recognition, DTW(Dynamic Time Warping)and DTWDir applied on the dataset and they concluded.

DTWDir with an accuracy of 93% is better than $1, KNN, DTW recognition [15]. Table 3 describes the performance analysis of smart devices used in different Autism Dataset [15].

Table 3: Performance of smart device used in Autism Dataset.

Author	Dataset	Tool/ Language	Algorithm	Metrics used
Bram Van den Bekerom [13]	National Survey of Children Health (NCSH)-2011 to 2012	Java	J48 algorithm	1-away method, k-Fold cross-validation
Al-diabat [14]	Mobile app for ASD Screening-at UCI data repository - 2017	ASD Tests tool	FURIA	ten-fold cross-validation, fuzzy rules
Gong *et al* [12]	Data were collected from 35 children in the USA that can easily be implemented as an app on modern child-specific devices from Apr 2016 to Sep 2017	Correlation Coefficient	SVM,CFS	COVERED toolkit, statistic functions, confusion matrix
Salwa O Slim *et al* [15]	Dataset is collected by 'Lenovo's smartphone sensor which its sampling rate is approximately 98 samples per second. We recorded 6 sec for each try of four behaviors	C++	DTWDir	K-fold cross-validation

D. Identifying autism disorders with Classification Based on Association (CBA)

Manimekalai, *et al.* [16], according to author the resulting classifier can distinguish behavior patterns between Autism and PDD-NOS with a relatively high accuracy-85.27%. Table 4 describes the different features and DSM-IV in CBA.

Table 4: Performance of CBA.

Dataset	Feature	Tool	Algorithm	Metrics used
Dataset used actual patient profile from 2 hospitals in Thailand	Social interaction, Communication and Restricted repetitive/stereotyped (32 criteria)	DSM-IV	CBA	K-Fold, DSM-IV, ADOS

E. Identifying autism disorders with FEAST

Mythili, *et al.* [17], proposed a Machine learning technique Fuzzy Cognitive Map and Feature extraction methods (FEAST) to improve social skills, cognitive methods and functioning in Autism. FEAST with an accuracy of 87.24%. Table 5 describes the performance of the FEAST algorithm.

Table 5: Performance of FEAST algorithm.

Dataset	Feature	Tool	Algorithm	Metrics used
The data is Collected from Autism and developmental disabilities monitoring network US (2010 to 2012)	Face recognition, Speech/Voice recognition, Gestures recognition, Object recognition, Avoid eye contact, Forgets letters, numbers, No face emotions, No society linkages, Unexpected attitude, Dis ambiguity, lack of social interaction	Rapid Miner	FEAST	Fuzzy Cognitive Map

IV. RESULT AND DISCUSSION

In this review, several algorithms are compared with various metrics. Table 6 shows the algorithms and accuracy of the autism dataset.

Fig. 3 describes the performance of classification algorithms and the Random forest algorithm gives better results compare with other algorithms.

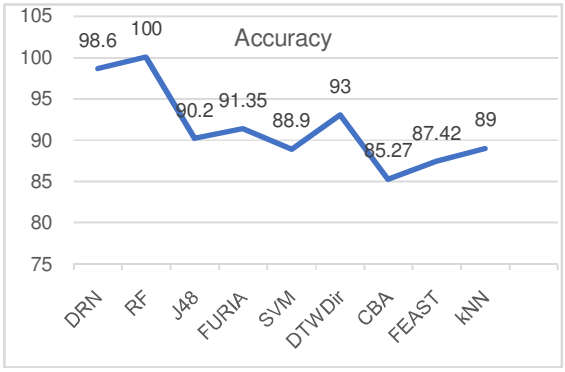

Fig. 3. Performance of Classification algorithms.

Table 6: Comparison of classification algorithms.

Algorithm	Accuracy level	Tools/ Language	Metrics
DRN	98.6%	Rapid Miner	Classification Error, Kappa Statistic, Mean Absolute and Root Mean Squared Error
RF	88.46%	DISCO	Fisher, Runs, Relief-
J48	90.2	Java	1-away method, k-Fold cross-validation
FURIA	91.35%	ASD Tests tool	ten-fold cross-validation, fuzzy rules
SVM	88.9%	Correlation Coefficient	COVERED tool kit, statistic functions, confusion matrix
DTWDir	93%	C++	K-fold cross-validation
CBA	85.27%	DSM-IV	K-fold, DSM-IV, ADOS
FEAST	87.42%	Rapid Miner	Fuzzy Cognitive Map
RF	100%	MATLAB	Sigmoid, polynomial, radial basis function (RBF) kernel function,10-fold CV

V. CONCLUSION

Based on this study, Machine Learning algorithms applied for analysis Impaired Social Interaction, Behaviour Patterns, Cognitive Problems, and Sensory Aspects and improve *developmental disorder problems in autism.* Predicting ASD in an early stage is not easy because lots of questions arise to the parents or caretakers and Physicians or Therapists, etc., The diagnosis of ASD is difficult to understand until the clear signs of the disease appear. This normally happens 16-18 months of life, when the lack of communication challenges, a tenderer to engage in repetitive behaviors and social deficiencies, etc. The research work shows the new way of analysis in autism screening and diagnostics, which can be used by Therapists, Physicians to help them with their formal screenings.

VI. FUTURE SCOPE

In the future, by creating an application for a better understanding of the children with ASD and identifying the disorder in an early stage using a new Machine Learning algorithm which will make physicians, therapists better knowledge and with their formal screening.

Conflict of Interest: Authors declare that they have no conflict of interest.

REFERENCES

[1]. Shhab, A., Guo, G., & Neagu, D. (2005). A study on applications of Machine Learning Techniques in Data Mining. In *Proc. of the 22nd BNCOD workshop on Data Mining and Knowledge Discovery in Databases, Sunderland, UK.*

[2]. Bishop-Fitzpatrick, L., Movaghar, A., Greenberg, J. S., Page, D., DaWalt, L. S., Brilliant, M. H., & Mailick, M. R. (2018). Using machine learning to identify patterns of lifetime health problems in decedents with autism spectrum disorder. *Autism Research*, *11*(8): 1120-1128.

[3]. Elizabeth Stevens, Abigail Atchison, Laura Stevens, Esther Hong, Doreen Granpeesheh, Dennis Dixon, Erik Linstead. (Dec 2017). A Cluster Analysis of Challenging Behaviors in Autism Spectrum Disorder. *16th IEEE International Conference on Machine Learning and Applications,* pp.00-85.

[4]. Chan, A. S., Sze, S. L., & Han, Y. M. (2014). An intranasal herbal medicine improves executive functions and activates the underlying neural network in children with autism. *Research in Autism Spectrum Disorders*, *8*(6): 681-691.

[5]. Thabtah, F. (2017). Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment. *Proceedings of the 1st International Conference on Medical and Health Informatics 2017* (pp. 1-6). ACM.

[6]. Dawson, G., Osterling, J., Meltzoff, A. N., & Kuhl, P. (2000). Case study of the development of an infant with autism from birth to two years of age. *Journal of applied developmental psychology*, *21*(3): 299-313.

[7]. Ramya, R., & Zoraida, B. S. E. (2017). DRN Hybrid Model for Predicting Autism Using Rapid Miner Tool. *International Journal of Advanced Research in Computer Science*, *8*(8): 111-115.

[8]. Geetha Ramaniand, R. & Sivaselvi, K. (2017). Autism Spectrum Disorder Identification using Data mining techniques. *International journal of pure and applied mathematics, 117*(16): 427-436.

[9]. Sumi, S., Chandra, J. and Saravanan, N. (2016). Empirical Evaluation of Data Mining Classification Methods for Autistic Children. *International Journal of Trend in Research and Development(IJTRD),* pp.7-10.

[10]. Bi, X. A., Wang, Y., Shu, Q., Sun, Q., & Xu, Q. (2018). Classification of autism spectrum disorder using random support vector machine cluster. *Frontiers in genetics*, *9*, 18.

[11]. Demirhan, A. (2018). Performance of machine learning methods in determining the autism spectrum disorder cases. *Mugla J. Sci. Technol.*, *4*(1): 79-84.

[12]. Gong, Y., Yatawatte, H., Poellabauer, C., Schneider, S., & Latham, S. (2018, August). Automatic Autism Spectrum Disorder Detection Using Everyday Vocalizations Captured by Smart Devices. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics* (pp. 465-473). ACM.

[13]. Bram Van den Bekerom (2017). Using Machine Learning for Detection of Autism Spectrum Disorder. *Proc. 20th Student Conf. IT, p*p.1-7.

[14]. Al-Diabat, M. (2018). Fuzzy data mining for autism classification of children. *Int J Adv Comput Sci Appl.*, *9*(7): 11-17.
[15]. Salwa O Slim, Ayman Atia, and Mostafa-Sami M. Mostafa (2016). DTWDIR: An Enhanced Dtw Algorithm for Autistic Child Behaviour Monitoring. *International Journal of UbiComp.,* Vol. *7*(2): 1-8.
[16]. Manimekalai, M., Aarthi, A.E. Priya and Brind, S. (2014). Evaluating the Behavioral and Developmental Interventions for Autism Spectrum Disorder. *International Journal of Information Science and Application*, Vol. *6*(1): 1-10.
[17]. Mythili, M. S., & Shanavas, A. R. Mohamed (2016). An Improved Autism Predictive Mechanism Among Children Using Fuzzy Cognitive Map. *ARPN Journal of Engineering and Applied Sciences,* Vol. *11*(3): 1451-1456.
[18]. Madni, H. A., Anwar, Z., & Shah, M. A. (2017, September). Data mining techniques and applications—A decade review. In *2017 23rd International Conference on Automation and Computing (ICAC)* (pp. 1-7). IEEE.
[19]. Uma Rani, R. and Suguna, R. (2018). Exploratory Data Analysis of Autism Data. *IOSR Journal of Engineering, pp. 05-10.*
[20]. Padmapriya, S., & Murugan, S. Performance Analysis of Elman and Modified Elman Neural Network To Predict Asd Sub Groub From Gnome Data Sequence.
[21]. Katare, A., & Dubey, S. (2017). A Study of Various Techniques for Predicting student Performance under Educational Data Mining. *International Journal of Electrical, Electronics and Computer Engineering*, Vol. *6*(1): 24-28.
[22]. B. Ida Seraphim, Lavi Samuel Rao, Shiwani Joshi. (2018). Survey on Early Detection of Autism Using Data Mining Techniques. *International Journal of Engineering & Technology,* Vol. *7*(2.24): 79-80.
[23]. Verma, K. & Malviya, P. (2017). A literature review on feature Selection in Big Data. *International Journal of Electrical, Electronics and Computer Engineering,* Vol. *6*(1): 37-44.
[24]. Hadioui, A., Touimi, Y. B., & Bennani, S. (2017). Machine Learning Based On Big Data Extraction of Massive Educational Knowledge. *International Journal of Emerging Technologies in Learning (iJET)*, *12*(11): 151-167.
[24]. Stedman, A., Taylor, B., Erard, M., Peura, C., & Siegel, M. (2019). Are children severely affected by autism spectrum disorder underrepresented in treatment studies? An analysis of the literature. *Journal of autism and developmental disorders*, *49*(4): 1378-1390.
[25]. Guangyuan Piao (2018). Machine & Deep Learning for Network Management: An Overview with Benchmarks. *https://goo.gl/gp7gBb*.
[26]. https://www.aol.com/video/view/symptoms-of-autism-spectrum-disorder/518232164/

How to cite this article: Priya, N. and Radhika, C. (2019). A Survey on Predicting Autism Spectrum Disorder using Machine Learning techniques. *International Journal on Emerging Technologies*, **10**(3): **10**(3): 422–427.