VOL. 14, NO. 10, MAY 2019 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

ENHANCE PREDICTION OF AUTISM SPECTRUM DISORDER USING ADAPTIVE BAYESIAN CLASSIFIER

D. Umanandhini[1] and G. Kalpana[2]
[1]Department of Computer Science, Kovai Kalaimagal College of Arts and Science, Coimbatore, India
[2]Department of Computer Science, Sri Ramakrishna College of Arts and Science for Women, Coimbatore, India
E-Mail: adipoli.umanandhini@gmail.com

ABSTRACT

Classification plays a major role in the medical field to predict diseases. The prediction analyzes the relation between the expected information and the available information. It's the duty of the classifier to make a classification in an efficient manner to predict diseases accurately. Misclassification may lead to a high risk to the individuals. In this paper, an adaptive Bayesian classifier is proposed to efficiently classify the autism spectrum disorder among children, where it is considered a serious and increasing medical problem among the children. Autism spectrum disorder cannot be detected like other diseases. The proposed classifier is designed to check the classification accuracy based on the threshold value, when the result did not meet the threshold value then the reclassification will be preceded. Also, the proposed classifier is designed to check the hidden patterns in the dataset to overcome the delay in classification. This research work uses the benchmark performance metrics to evaluate its performance. The result shows that the proposed classifier outperforms the baseline classifier by giving the better classification accuracy in low delay.

Keywords: autism spectrum disorder, ASD, sensitivity, specificity, classification.

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a type of neurodevelopment disorder that exist for a lifetime. In the year 1943 Leo Kanner described autism as a prototypical condition. ASD is described by the qualitative behavioral abnormalities in the way of communication with parents and others, where ASD affected persons have shorten and stereotyped interests and activities. Maximum amount of sophisticated behavioral features occurs only after a variation in environment. ASD is considered as heterogeneous among the affected individuals. 1 in 68 individuals are affected by ASD and 3:4 ratio of affected males to females. These kinds of insufficiency are determined in a pervasive manner, and it is normally found in the early childhood and possibly to prompt weaknesses in performing across different settings. Currently, ASD has grown from very mild to severe with many individuals and a lifelong support is needed. ASD is connected with physiological behavioral changes which can be non-invasively checked by utilizing the persons behavior. This behavioral change provides the aim of measuring ASD towards the treatment programs for one who has challenges with self-analyzing, communicating, and recognition the emotion.

In the current world, many researches started applying the data mining for analyzing the existence data, with the motto of finding the hidden information which is not found. Data mining can also be defined as set of methods applied to extract the potential information in a understandable manner, which id from a very large datasets or databases. One of the major objectives of data mining is to prediction, which is expected many field for various reason. In medical field, data mining is used to predict disease. For certain types of diseases, it's necessary to predict disease at the earlier stage (like ASD), and for some diseases detection or prediction at later stage might not have big affect. Data mining hold two types of models [14], which are, (i) Prediction model – It estimate values for the target variables under study based on the other variables in the database, (ii) Descriptive model - It identify patterns to explain or summarize the data. These patterns are used to explore the properties of the examined data.

1.1 Problem statement

Currently, ASD is increasing among children. It's necessary to detect it in early stage, where the failing of ASD detection may lead to issues in children health condition mentally. Current algorithms don't have much impact in giving the accuracy towards the classification for detection of ASD. Kalman filtering framework offers competitive accuracy of 85% but it is considerably lower, which consumes more time for classifying the data. Traditional classification algorithms have better performance in the dataset which have low number of records, but those algorithms have worst performance when the dataset size is large.

1.2 Objectives

This research work utilizes data mining to understand data distributions at all levels and its relationship to predict ASD. The main objective of this research work is to propose a unsupervised learning based classifier which can effectively detect ASD from the given dataset in short duration. This research work aims to overcome the barriers in providing the treatment for ASD at the early stage.

The remainder of this paper is organized as follows. Section 2 provides the summary of the related works as literature review. Section 3 discusses the proposed classifier towards the detection of ASD. Section 4 illustrates the chosen performance metrics along with the tool used. Section 5 confers the results. Finally, Section 6 concludes the paper with future dimensions.

2. LITERATURE REVIEW

Brain network based feature representation [1] was proposed to utilize deep neural network (DNN) classifier in order to perform the classification of ASD. At first, individual brain network for each subject was constructed extract the connectivity features, where the features were selected by checking top ranked features. But the classification accuracy gets lack with a major difference. An evaluation of a generic framework [2] that leverages two different types of information was presented to analyze the brain for the prediction of ASD in large populations. The framework explores Graph Convolutional Networks (GCNs) and engages denoting the populations as a sparse graph, where it didn't considered edge weights in the classification of ASD. A initiative attempt [3] was made to use graph theory in deriving distinct features of resting-state functional networks in young adults with ASD. Global graph measures were utilized to tailor the dataset to specific hypotheses concerning the spatial distribution of abnormalities in connectivity among individuals with ASD.

Virtual Reality based Brain Computer Interface [4] paradigm was used in social cues to direct the focus of attention. Interactive immersive virtual reality technology which includes VRBCI signals was combined in a training tool and used to classify ASD, where it didn't created impact in detecting the false positive. Neuroimaging marker was developed [5] by implementing support vector machine as a core was used to differentiate ASD from ADHD (Attention-Deficit/Hyperactivity Disorder). Even after utilizing the difference between prefrontal activation in adults with ASDs and ADHD, it was found that the result was not able give the appropriate true positives and true negatives. Endogenous Posner paradigm task [6] was administered to classify 15 children with ASD and 16 typically developing (TD) children to make an investigation, and comparison of behavioral performance and event-related potentials (ERPs) measures. The result demonstrates that classifier wont able to provide more accurate results only based on behavior.

Multivariate pattern analysis approach based on voxel morphometry [7] was used with searchlight algorithm to classify the data of autism children and adults, where the impact of the algorithms became a drawback in detecting the ASM with gender. A methodology was proposed for the investigation of functional connectivity in patients with ASD using Fuzzy Synchronization Likelihood (Fuzzy SL) [8], while applying it in between and within different regions and different EEG sub-bands for distinguishing the autistic children from healthy, it showed the negative results. Four machine learning techniques [9] namely support vector machines, multilayer perceptrons, functional trees, and logistic model trees (LMTs) were employed to generate the diagnostic models for ASD, where it was based measurements extracted by using the classifier. A variable Analysis [10], computational intelligence method was proposed with the consideration of feature-to-class correlations and reduces feature-to-feature correlations. It was a made to identify features in ASD screening methods in order to achieve efficient screening as demands on evaluating the items influences on ASD.

Advantages of VR for individuals with ASD [11] was presented to identify the challenges and design issues for the upcoming years training applications, where the state of the art on virtual reality (VR) for individuals with autism spectrum disorder (ASD) with a focus on targeted intervention was discussed, but failed to discuss/deploy the algorithms with the classification accuracy results. A collaborative virtual environment (CVE) [12] based social interaction platform for ASD intervention was proposed, where it leads a way to the create low-cost intervention environment. The main drawback of this system was the force give to the ASD children to use the application in a mandatory manner resulting the children in a uncategorized way. A kalman filtering framework [13] based on the Kalman filtering theory was proposed for detection of physiological arousal based on cardiac activity, it's a an unsupervised and real-time arousal detection algorithm which lacks in taking too much time for classifying the data for ASD.

3. ADAPTIVE BAYESIAN CLASSIFIER FOR PREDICTION OF ASD

This section discusses about the proposed classifier namely Adaptive Bayesian Classifier (ABC) towards the prediction of ASD among children. ABC puts forwards a enhanced strategy to swap fitness and complexity of the data mining model. With the end goal to learn efficient patterns for Bayesian classifier, it not only tends to receive the specific information like dataset length and information. ABC utilizes the threshold value to proceed with reclassification to give more accurate prediction of ASD.

ABC, it limits the overall results of methods of encoding. To denote ABC, it's necessary to express the ancestor and probability table of every node. It is a consideration that every node Y^h, there are l^h ancestors, at that point we require $l^h \log[n]$ bits to represent the ancestor of the node. Let $s^h[1 <= h <= n]$ be the property of node of Y^h, $pa[Y^h]$ the ancestor of Y^h, the quantity of dependent probabilities for node Y^h is $s^h / \sum_{k \emptyset pa[Y^h]} s^k$. Since every single dependent probability for one node will be 1, and the encoding length for each restrictive probability is $logP * 2$, the aggregate depiction length for model A will be:

$$CR[A] = \prod_{h=1}^{n} \left[l^h \log n - \frac{log_n^P(P)}{[s^h + 1]} \div \sum_{k \emptyset pa[Y^h]} s^k \right] \quad (1)$$

For the information of depiction length, likelihood technique is utilized to encode which is as follows:

$$CR[C:A] = \prod_{h=1}^{P} logN^{A}[Y^{h}] = P\prod_{k=1}^{n}\prod_{Y^{k},pa[Y^{k}]} N^{C}\left[Y^{k},pa[Y^{k}]\right] logN^{A}[Y^{k}:pa[Y^{k}]] \quad (2)$$

where N^{A} is the probability of distribution over model A and N^{C} the genuine probability distribution.

$$N^{A}[Y^{1},\ldots.,Y^{n}] = \sum_{h=1}^{n} N^{A}\left[Y^{h}:pa[Y^{h}]\right] \quad (3)$$

Clearly $CR[C:A]$ is limited when

$$N^{C}\left[Y^{k}:pa[Y^{k}]\right] = N^{A}\left[Y^{k}:pa[Y^{k}]\right] \quad (4)$$

Applying Eq. (4) to Eq. (2), $CR[C:A]$ is composed as

$$CR[C:A] = P\prod_{h=1}^{n} I\left[Y^{h}:pa[Y^{h}]\right] \quad (5)$$

As indicated by the information hypothesis, the dependent entropy will be

$$I[Y:pa[Y]] = \prod_{Y,pa[Y]} N^{C}[y,pa[y]] logN^{C}[y:pa[y]] \quad (6)$$

It is outstanding that the common information of $Y, pa\ [Y]$ is

$$H[Y,pa[Y]] = \prod_{Y,pa[Y]} N^{C}[y,pa[y]] log\left(\frac{N^{C}[y,pa[y]]}{N^{C}[y]N^{C}[pa[y]]}\right) \quad (7)$$

Thus, it is anything but difficult to verification that

$$I[Y:pa[Y]] = I[Y] + H[Y,pa[Y]] \quad (8)$$

In the wake of applying Eq. (8) to Eq. (5), we get

$$CR[C:A] = P\prod_{h=1}^{n} I[Y^{h}] - P\prod_{h=1}^{n} H\left[Y^{h},pa[Y^{h}]\right] \quad (9)$$

At long last the MDL for a ABC model will be the difference between $CR[A]$ and $CR[C:A]$, that is

$$NCR[A:C] = \prod_{h=1}^{n}\left[l^{h} log\, n - \frac{\left(\frac{(\log P)}{[s^{h}+1]}\right)}{:\sum_{k\emptyset pa[Y^{h}]} s^{k}}\right] - P\prod_{h=1}^{n} I[Y^{h}] + P\prod_{h=1}^{n} H[Y^{h},pa[Y^{h}]] \quad (10)$$

For ABC, dependency that exist between patterns are given full importance for detecting the ASD, else it may lead to misdetection of ASD. If suppose, dependency are not preferred and given importance then the ancestor node will be considered as the only class attribute. Likewise, an indifferent technique is utilized to discover patterns for every achievable class mark $[s^{h} = 1]$ so the common information $H\left[Y^{h},pa[Y^{h}]\right]$ will be 0 in light of the fact that the opinion of ancestor node is steady. The estimated length L of the patterns of ABC is

$$L^{PDR}[A:C] = \prod_{h=1}^{n} \frac{\binom{logP}{2}}{[s^{h}+1]} - P\prod_{h=1}^{n} I[Y^{h}] \quad (11)$$

where PDR denotes the pattern discovery rate, n the quantity of patterns, s^{h} the cardinality of pattern Y^{h}, P the quantity of records for one exceptional class.

4. ABOUT PERFORMANCE METRICS AND DATASET

4.1 Performance metrics

Bench mark Performance metrics are chosen to provide the better classification accuracy.

ARPN Journal of Engineering and Applied Sciences

Table-1. Performance Metrics.

Metrics	Description	Formula
Sensitivity	Proportion of actual positives that are correctly identified.	$\frac{True\ Positive}{(True\ Positive + False\ Negative)}$
Specificity	Proportion of actual negatives that are correctly identified.	$\frac{True\ Negative}{(False\ Positive + True\ Negative)}$
Disease Prevalence	Proportion of disease found in the total populace of the dataset.	$\frac{(True\ Positive + False\ Negative)}{(True\ Positive + False\ Negative + False\ Positive + True\ Negative)}$
Positive Likelihood Ratio	Proportion of ratio between the probabilities of a positive test result given the presence of the disease and positive test result given the absence of the disease.	$\frac{Sensitivity}{(1 - Specificity)}$
Negative Likelihood Ratio	Proportion of ratio between the probabilities of a negative test result given the presence of the disease and negative test result given the absence of the disease.	$\frac{(1 - Sensitivity)}{Specificity}$
Positive Predictive Value	Proportion of probability of disease that is present when the test is positive.	$\frac{True\ Positive}{(True\ Positive + False\ Positive)}$
Negative Predictive Value	Proportion of probability of disease that is not present when the test is negative.	$\frac{True\ Negative}{(False\ Negative + True\ Negative)}$
Accuracy	Proportion of true results (both true positives and true negatives) among the total number of cases examined.	$\frac{(True\ Positive + True\ Negative)}{\left((True\ Positive + False\ Negative + False\ Positive + True\ Negative)\right)}$

4.2 Dataset

Actual time consumed for building the dataset is around six months. The dataset contains 1499 patients records. For the confidentiality reasons, the name of the patients is not obtained. Dataset holds the information of both gender children with the age ranging from 3 to 11. Out of 1499 records, 998 children are having the possibility of getting ASD and remaining don't have. This research uses the MATLAB tool version R2013a for evaluating the algorithms.

5. PERFORMANCE EVALUATION

5.1 Analysis of TP, TN, FP, and FN

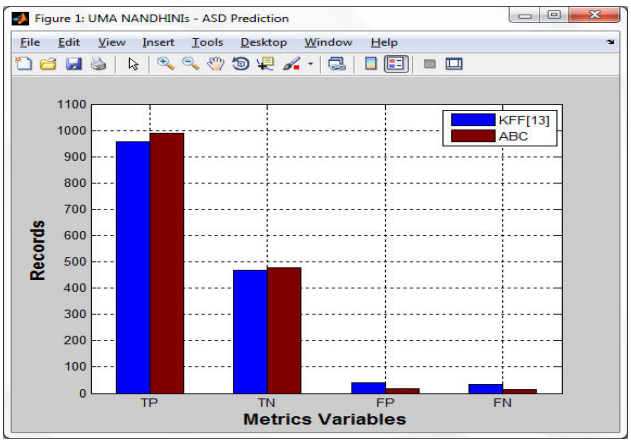

Figure-1. TP, TN, FP, FN Analysis.

The metric variables TP, TN, FP, and FN are plotted in x-axis, where the y-axis is plotted with number of records. Figure-1 compares the TP, TN, FP, and FN of the proposed classifier ABC with the existing classifier KFF [13]. In the MATLAB execution, the medical dataset containing 1499 records are given as input to both the classifier. During the assessment of the proposed classifier (ABC) and KFF [13], the multiple records of the dataset are fed in random manner in a different attempt. The proposed classifier is designed to process the input in whatever order it is given, where the existing classifier KFF [13] takes the input only in the sequential manner with limited number of records. Due to this, the proposed classifier attempts to give better results than the KFF [13].

5.2 Analysis of sensitivity and specificity

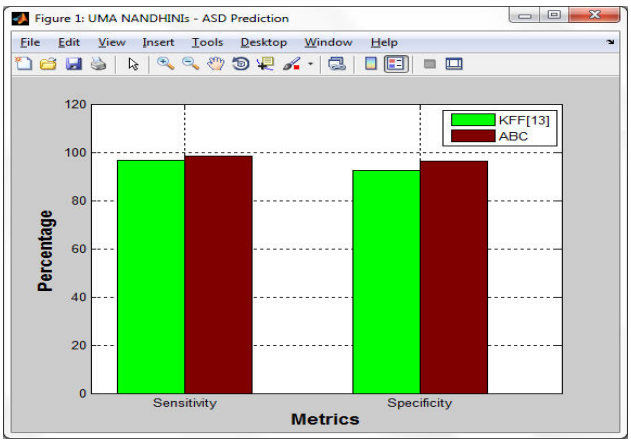

Figure-2. Sensitivity and Specificity Analysis.

In Figure-2, the metrics sensitivity and specificity are plotted in x-axis, where the result percentages are

plotted in y-axis. KFF [13] and ABC are having little difference in the result, where the proposed classifier is outperforming than the exisiting. ABC is designed to reclassify the reclassify the records if it has not met the threshold value. Like this, records are classified multiple times for true and false values. In KFF [13], if record is classified once, then it is not considered for the next time. Due to this KFF [13] lacks. When comparing KFF [13] and ABC for sensitivity and specificity, it is evident that the proposed classifier ABC outperforms than KFF [13].

5.3 Analysis of likelihood ratios

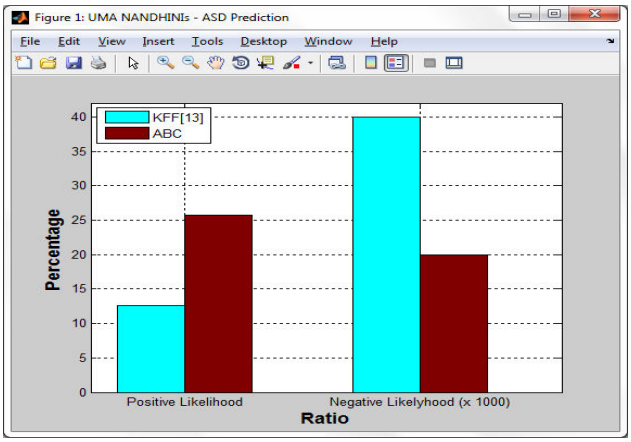

Figure-3. Likelihood Ratio Analysis.

In Figure-3, the ratios positive likelihood and negative likelihood are plotted in the x-axis and its percentages are plotted in y-axis. When comparing the results of KFF [13] and ABC for the likelihood ratio, there exist a major difference. The proposed classifier checks the pattern of the records and proceed with classification, where patterns are not considered and given importance in KFF [13]. Due to this reason, KFF[13] is resulting with lowest positive likelihood and highest negative likelihood, which is not appreciated in prediction of ASD. The proposed classifier ABC gives the full preference for patterns and gets the positive results in both the likelihood, and gives a confirmation that the prediction method is proceeding in a trustable way.

5.4 Analysis of prediction value

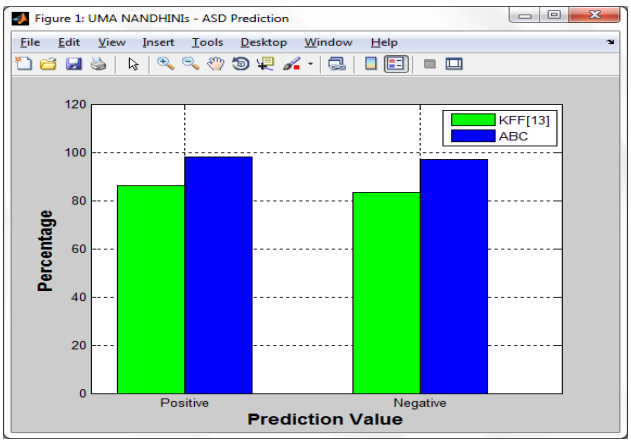

Figure-4. Prediction Value Analysis.

In Figure-4, the prediction value metrics positive and negative are plotted in x-axis and its percentages are plotted in y-axis. Figure-4 compares the positive prediction value and negative prediction value of the existing classifier KFF [13] and the proposed classifier ABC. It is evident that the proposed classifier is outperforming than KFF [13] in both the prediction value, this is due to the considering the reclassification based on the threshold value. KFF [13] performs the classification in a first come first serve basis in which the record received from the dataset; it may be the reason for its poor result.

5.5 Analysis of disease prevalence and accuracy

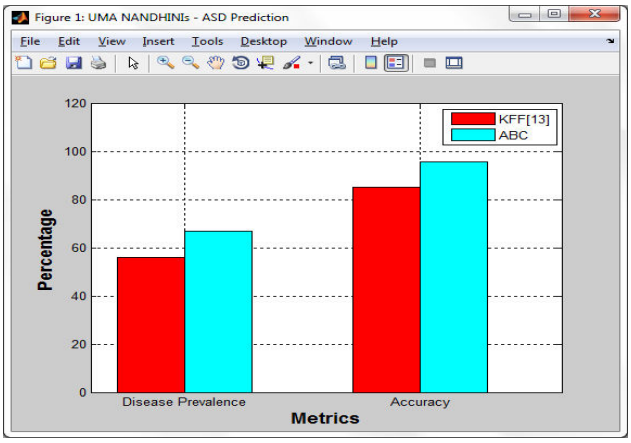

Figure-5. Disease Prevalence and Accuracy Analysis.

In Figure-5, disease prevalence and accuracy are plotted in x-axis and its percentage of results are plotted in y-axis. Figure-5 check compares the results of disease prevalence and accuracy of classification by KFF [13] and ABC. The disease prevalence of the proposed classifier ABC is better than KFF [13], and this is due to giving importance to patterns in the dataset which is most expected in the prediction of ASD. Finally, the accuracy of classification is compared for KFF [13] and the proposed

classifier ABC. The various aspects like reclassification, threshold value setting and pattern are considered and given preference in ABC and it results in giving the better accuracy of classification for ASD.

6. CONCLUSIONS

Data mining is emerging in the medical field to over the barrier of diseases because to detect and classify certain diseases medical concepts alone is not enough. This research work has proposed adaptive Bayesian classifier to efficiently detect the ASD. The proposed classifier is designed to reclassify the data if the threshold value is not met; also it detects the hidden patterns to give better result without any delay. The MATLAB results show that the proposed classifier outperforms the baseline scheme namely kalman filtering framework in terms of chosen performance metrics.

REFERENCES

[1] Y. Kong, J. Gao, Y. i. Xu, Y. Pan, J. Wang, J. Liu. 2019. Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier. Neuro computing. 324: 63-68.

[2] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, D. Rueckert. 2018. Disease prediction using graph convolutional networks: Application to Autism Spectrum Disorder and Alzheimer's disease. Medical Image Analysis. 48: 117-130.

[3] V. Tsiaras, P. G. Simos, R. Rezaie, B. R. Sheth, E. Garyfallidis, E. M. Castillo, A. C. Papanicolaou. 2011. Extracting biomarkers of autism from MEG resting-state functional connectivity networks. Computers in Biology and Medicine. 41(12): 1166-1177.

[4] C. P. Amaral, M. A. Simões, S.a Mouga, J. Andrade, M. Castelo-Branco. 2017. A novel Brain Computer Interface for classification of social joint attention in autism and comparison of 3 experimental setups: A feasibility study. Journal of Neuroscience Methods. 290: 105-115.

[5] A. Ishii-Takahashi, R. Takizawa, Y. Nishimura, Y. Kawakubo, H. Kuwabara, J. Matsubayashi, K. Hamada, S. Okuhata, N. Yahata, T. Igarashi, S. Kawasaki, H. Yamasue, N. Kato, K. Kasai, Y. Kano. 2014. Prefrontal activation during inhibitory control measured by near-infrared spectroscopy for differentiating between autism spectrum disorders and attention deficit hyperactivity disorder in adults. NeuroImage: Clinical. 4: 53-63.

[6] C. L. Tsai, C. Y. Pan, C. H. Wang, Y. T. Tseng, K.W. Hsieh. 2011. An event-related potential and behavioral study of impaired inhibitory control in children with autism spectrum disorder. Research in Autism Spectrum Disorders. 5(3): 1092-1102.

[7] L. Q. Uddin, V. Menon, C. B. Young, S. Ryali, T. Chen, A. Khouzam, N. J. Minshew, A. Y. Hardan. 2011. Multivariate Searchlight Classification of Structural Magnetic Resonance Imaging in Children and Adolescents with Autism. Biological Psychiatry. 70(9): 833-841.

[8] M. Ahmadlou, H. Adeli, A. Adeli. 2012. Fuzzy Synchronization Likelihood-wavelet methodology for diagnosis of autism spectrum disorder. Journal of Neuroscience Methods. 211(2): 203-209.

[9] Y. Jiao, R. Chen, X. Ke, K. Chu, Z. Lu, E. H. Herskovits. 2010. Predictive models of autism spectrum disorder based on brain regional cortical thickness. Neuro Image. 50(2): 589-599.

[10] F. Thabtah, F. Kamalov, K. Rajab. 2018. A new computational intelligence approach to detect autistic features for autism screening. International Journal of Medical Informatics. 117: 112-124.

[11] L. Bozgeyikli, A. Raij, S. Katkoori and R. Alqasemi. 2018. A Survey on Virtual Reality for Individuals with Autism Spectrum Disorder: Design Considerations. in IEEE Transactions on Learning Technologies. 11(2): 133-151, 1 April-June.

[12] H. Zhao, A. R. Swanson, A. S. Weitlauf, Z. E. Warren and N. Sarkar. 2018. Hand-in-Hand: A Communication-Enhancement Collaborative Virtual Reality System for Promoting Social Interaction in Children with Autism Spectrum Disorders. in IEEE Transactions on Human-Machine Systems. 48(2): 136-148.

[13] A. Kushki, A. Khan, J. Brian and E. Anagnostou. 2015. A Kalman Filtering Framework for Physiological Detection of Anxiety-Related Arousal in Children with Autism Spectrum Disorder. in IEEE Transactions on Biomedical Engineering. 62(3): 990-1000.

[14] L. D. Olio, A. Ibeas, J. D. Ona, R. D. Ona. 2018. Public Transportation Quality of Service, Elsevier. pp. 155-179.