Autism - Open Access
ISSN: 2165-7890
Mahmoudi et al., Autism Open Access 2018, 8:3
DOI: 10.4172/2165-7890.1000230
Research Article Open Access

An Autism Screening Expert System: Reliability, Validity and Factorial Structure

Maryam Mahmoudi[1]*, Saeed Akbari-Zardkhaneh[2], Amir Ali Bagher Zadeh[3], Bagher Ghobari-Bonab[1], Mohsen Shokoohi-Yekta[1], Hadi Moradi[4] and Hamid Reza Pouretemad[2]

[1]*Department of Psychology and Education, University of Tehran, Tehran, Iran*
[2]*Department of Psychology and Education, University of Shahid Beheshti, Tehran, Iran*
[3]*School of Computer Science, Language Technologies Institute, Carnegie Mellon University, USA*
[4]*Machine Intelligence and Robotics School, University of Tehran, Tehran, Iran*

Abstract

Purpose: The main aim of the current study was to develop an expert system for screening autism spectrum disorders. The statistical properties of the system were further examined.

Methods: To design an expert system, having a reliable and valid knowledge base (KL) is very important. To provide the knowledge base, items linked to autism diagnoses were collected from the literature and then reviewed by a group of psychologists and psychometrists experts. The questionnaire was completed by parents of children with autism (85), parents of normal children (65) and parents of children with Down syndrome (20). Next, some psychometric and machine learning methods were implemented to select the items having more power to discriminate children with autism from normal children and children with Down syndrome and evaluate its reliability and validity features.

Results: Findings yielded good reliability (0.96) and criterion validity (0.86) for the system. The accuracy was about 0.93 and .92, using Random Forest (RF) and Support Vector Machines (SVM)), respectively. In addition, specificity and sensitivity of the system using SVM is 84.1 and 98.5%, respectively, while RF is 73.4 specificity and 99.3% sensitivity.

Conclusion: This system can be considered as a reliable and valid system for screening ASDs.

Keywords: Children with autism; Expert system; Screening; Validity

Introduction

Autism is a developmental disorder that characterized by severe and continuous impairments in various areas, such as reciprocal social interactions, communication and restricted interests as well as repeated behaviors [1]. In many countries the prevalence of autism is higher than one per cent [2,3]. In Iran, the prevalence of these children based on the screening of five year old children before entering preschool was nearly 0.24% using Social Communication Questionnaire (SCQ). On the other hand, prevalence of Iranian children diagnosed with autism utilizing Autism Diagnostic Interview-Revised (ADI-R) was 0.06% [4]. This indicates measure dependency of the prevalence of autism. In order to provide the best therapeutic outcome, children must be diagnosed before age of six [5]. Early identification has many advantages and can help to provide effective intervention for these children [6,7]. Although a recent study by Momeni et al. shows that by examining blood, it is possible to determine autism at 86% sensitivity; it is not easy and would take a long time to become available globally [8]. Consequently, the diagnosis of this disorder is dependent on behavioural assessments. Normally for diagnosis of autism, experts use observation and historical reports which are collected from several sources including interview with parents, direct observation and using inventories. This process is complex and may last for four years or more between the first times that parents concern about their children until they receive final diagnosis [9].

Implementing the current diagnostic methods such as Autism Diagnostic Interview-Revised and Autism Diagnostic Observation Schedule are expensive, time consuming and require extensive and specialized training that can result in restricted using of these methods. Moreover, it is infeasible and unrealistic to have all children receiving the same level of extensive diagnosis. Hence, it is better to have a first level of screening to identify the possible autistic subjects and then refer those for more precise evaluation processes [10]. Although these problems can be solved by screening instruments, however due to the complexity of the process all children are assessed in formal assessment at higher ages (for e.g. In Iran the children are assessed when they are 6 years old) and it is very late to begin the intervention. Consequently, having widely accessible systems, such as online screening systems, allows parents to immediately check out their children's symptoms, especially if they were concerned about their developmental process. Such an online system provides wide access to all people around the country and the world to screen their children, received adequate responses to their questions, and learns more about possible symptoms. It should be noted that the system is used for initial screening and would recommend further evaluation for a child suspicious of autism.

Various tools for screening and diagnosis of autism are available. The first tool is Checklist for Autism in Toddlers (CHAT) [11]. Specificity of this tool is high, however its sensitivity is low [12]. The tool was standardized by Robins, Fein, Barton & Green (2001) in America and renamed to Modified Checklist for Autism in Toddlers

***Corresponding author:** Maryam Mahmoudi, Department of Psychology and Education, University of Tehran, Tehran, Iran, Tel: +98-21-6111-7451; E-mail: mahmoody_m@alumni.ut.ac.ir

Received: September 18, 2018; **Accepted:** September 25, 2018; **Published:** October 02, 2018

Citation: Mahmoudi M, Akbari-Zardkhaneh S, Zadeh AAB, Ghobari-Bonab B, Shokoohi-Yekta M, et al. (2018) An Autism Screening Expert System: Reliability, Validity and Factorial Structure. Autism Open Access 8: 230. doi:10.4172/2165-7890.1000230

Copyright: © 2018 Mahmoudi M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

[13]. Sensitivity, Specificity and positive prediction value may be lower than what reported [14]. Autism Behaviour Checklist (ABC) was introduced by Krug, Arick & Almond [15]. Marteleto & Pedromônico showed that Autism Behaviour Checklist tool has specificity and sensitivity coefficient of 0.81, 0.92, but it failed in diagnosis of high function children [16,17]. Social Communication Questionnaire (SCQ) was developed based on DSM-IV-TR [18]. Allen, Silove, Williams & Hutchins reported specificity of 0.58 and sensitivity 0.93 in children aged 2 to 6 years-old [19]. They also reported specificity of 0.62 and sensitivity 1 in children ages 3 to 5 which has low specificity compared to other related tools. Another instrument in this field is Gilliam Autism Rating Scale (GARS) with internal consistency between 0.88 - 0.93 for four sub-categories. However, GARS has a good reliability and validity for children older than three years. However, some professionals reported a questionable sensitivity for this scale [20].

In Iran, SCQ is used for screening autism disorder in children up to 5 years old. Internal consistency of this scale was estimated around 0.81 in a sample size of 712 children 6 to 13 years old. The other diagnostic tool used in Iran is Autism Diagnostic Interview-Revised (ADI-R) which was translated by Sasanfar and Toloui, and is standardized for children between 4 and 14 years of age [4]. This instrument has internal consistency of 0.86 for questions about previous behaviours of children and 0.85 for current level of functioning/behaviours. It should be noted that the designed instruments for other cultures may not be valid or reliable to use in Iran due to cultural differences. Furthermore, in Iran, the typical age in which the children are formally screened is four years of age which is too late to provide early intervention and treatment. Consequently, in order to facilitate the early screening of children with autism in Iran we need a widely accessible to parents and paraprofessionals, easy to use, and an inexpensive system for evaluation. That is why the current online system has been proposed that can be accessed widely all over the country.

In addition, different tools for diagnosis of autism have different advantages and disadvantages. It is necessary to collect good items from all devices and compile them in one valid and reliable instrument that has advantages of all existing instruments in the field. That is why we have proposed to create a comprehensive set of questions and determine the best set of questions through an expert system using machine learning techniques. Expert systems are software programs that can simulate an expert decision and help problem solving. The system is based on the rules in the knowledge base of the program that tries to emulate an expert knowledge. The inference engine, which is another part of the system, has rules to conduct and simulate the expert's reasoning process [21]. Expert systems have lots of advantages such as being accessible through a computer, can provide explanation on how it has reached a conclusion, learn over time, and it is less costly compared to a human expert. Expert systems automatically keep the information of each client which can be used in the future for improving the system or track a desired client. Moreover, expert systems are normally fast and emotion free, i.e. is not affected by emotions involved in the screening process. It should be noted that despite many advantages of expert systems, they cannot be used as a sole screener. They should provide the first level of screening or give a second opinion to human expert increasing its reliability [21-23].

Expert systems are used in psychology and special education for several purposes. For instance, ASSESS is the first and oldest system in psychology. This program can be used to help professionals through simulated clinical judgment and the interpretation of psychological assessment tests. After several revisions, in 1996, ASSESS expanded into series of shorter tests with several norms [24]. After that, lots of expert systems have been made for evaluation, diagnosis and treatment. For instance, facial action coding system (FACS) is designed for measuring facial movements in order to identify human facial emotions [25]. Lau developed another system similar to FACS for emotion detection in individuals with disabilities [26]. Tentori & Hayes reported that, by Ubicomb technology, they could improve social skills. In their study they investigated innovative computer technology for social training showing its efficacy; including increased social skills. There are expert systems for diagnosis of different psychological and behavioral problems such an expert system for diagnosis of eating disorder [27], depression [28], sleep disorder [29], epilepsy [30] and aphasia [31]. Casado-Lumbreras, Rodriguez-Gonzalez, Alvarez-Rodriguez & Colomo-Palacios made an expert system for diagnosis of mental disorders and reported its considerable accuracy [32]. An expert system by Marakakis, Vassilakis & Papadakis for diagnosis of epilepsy in childhood was developed which its initial evaluation showed accuracy of 83.3% as successful diagnosis [33]. An expert system was designed for discriminating diagnosis in specific language impairments from dyslexia as well as autism based on fuzzy cognitive maps [34].

Sajjad, Qamar, Tariq & Bano developed an expert system for diagnosis of autism named PECADEX (Pakistan Childhood Autism Diagnostic Expert System) that they reported successful results, though they did not reported any reliability coefficients [35]. Kannappan, Tamilarasi and Papageorgiou designed an expert system for predicting autism through fuzzy cognitive maps [36]. Classification accuracy of this system was 79.9%. In a comprehensive project, Mabry & Frye-williams developed a tool, called SPECTRUM-agents providing comprehensive information to help professionals, teachers and researchers in decision making combining three components [37].

1. Screening tool that investigates probability of ASD.

2. Autism Spectrum Informatics (ASIST) that is analytical model of follow-up making connection between achievement and treatment performed periodically.

3. An extensive diagnostic evaluation model that provides more complex diagnosis and facilitates identification of a disorder and its severity. The authors have not reported any accuracy coefficient yet.

As mentioned earlier, there are lots of expert systems that can be helpful in the diagnosis and treatment of different disorders. Autism is also a pervasive disorder that should be diagnosed as soon as possible before losing critical years of intervention. In addition, in Iran, it is necessary to develop a tool widely accessible for early diagnosis. This requirement necessitates the design of a system which can be accessible on the web. Our system is based on selecting the best set of questions, i.e. features, from a large set of questions that can be used for highest correct diagnosis rate.

Methods

Participants

The Three groups of children participated in the study: children with Down syndrome (20), autism (85), and normal children (65). The age of participants ranged from two to six years old. The questionnaire was filled by their parents. The reason for including the Down syndrome group in the study was that this group has an average IQ in the range of intellectual disability and has similar morphology as those with intellectual disability. The children with autism were independently diagnosed by two experts (a psychologist and a psychiatrist) using interviews, observation, and ADI-R based on DSM-IV-TR criteria in two different situations [38]. The mean age was 4.67 years in the autistic group that was less than the children with Down syndrome (6.1) and a little more than the normal children group (4.5). Sampling was drawn by means of accessible sampling procedure, however, normal group sampled by random sampling.

Procedure

This section refers to two parts.

System development: In collecting data, different sources, were integrated under supervision of professionals and thesis supervising committee members. These procedures were elaborated in our previous article by Shokoohi Yekta et al. [39]. After collecting needed items, a bulk item pool (1385 items) was gathered and the following procedures were used to screen and organize the data: First, we removed repetitive items and classified the remaining items based on DSM-IV-TR. Then, three experts studied the items pool, revised them and selected the items with discriminant power suitable for diagnosis. Finally the revised items were proof read and finally 242 items were selected. Then a few questionnaires were completed and a few ambiguous items were removed from the pool leaving 238 items in four categories (social interaction, Communication, Stereotyping behaviours and Developmental delay) [39].

The aim of the current study was to develop a valid and reliable device to assess autism and differentiate it from normal children. To calculate reliability of the expert system and its sub-categories, Cronbach's alpha coefficient was used. Chi-square was also calculated to show discrimination power of each item. First, a three group chi-square compared the overall group in terms of whether there was a significant difference between the three groups, and then the two groups were compared. Since comparing groups two by two is not possible in follow-up among three groups in the chi square procedure, proportion test was used. Items with higher power were selected to discriminate between autism and the other two groups at 0.05 alpha levels. The item selection phase led to 167 (out of 238) items in the long form, and 81 items in the short form [39].

Psychometric properties: After completing items analysis, we analysed the sub-categories and categories of the instrument. At first, stepwise discriminant analysis was used for the purpose of putting together sub-categories resulting in discrimination between autism and the other two groups. For data analysis, SPSS-18 was used. Before this analysis, missing data was analysed for systematic no-response patterns. Because the missing data proportion was averagely under five per cent in each of the items, the missing data were replaced using linear interpolation.

After selecting appropriate items based on statistical methods, the long form containing 167 items became the base for further interpretations. Next, in order to develop an expert system, two classification methods of machine learning were used, i.e. Random Forest (RF) and Support Vector Machines (SVM). In the two methods, randomly 80% of data used for training and 20% for testing. It should be noted that the data of children with Down syndrome was not used for training data because of low number of cases involved. For the validity of the expert system, Kappa coefficient of agreement was used. This method is considered as criterion validity, measuring the percentage of agreement between GARS with the expert system's decision and the percentage of agreement between human expert and the expert system.

Results

In continue the psychometric properties have been presented.

Item analysis: Diagnostic items, prepared for the knowledge base (KB), included four categories, i.e. social interaction, communication, stereotypic behaviours, and developmental delays. Item analysis results helped to reduce the numbers of sub-categories from 30 to 15 and total number of items to 167, from 279. Internal consistency for the remaining items was between 0.85 to 0.96 for the categories. Item-total correlation of items in any category was between 0.21 and 0.72 and most of them were more than 0.3 indicating high sensitivity of items in identifying autism. In other words, this shows the capability of each item in identifying autism.

Categories and sub-categories analysis: As mentioned earlier, 15 sub-categories were selected from the pool of items by eliminating the redundant, irrelevant, and not important items. The descriptive characteristics of these sub-categories are presented in (Table 1). As shown in this table, the means of scores for children with autism were higher than the other two groups in all categories except social interaction in which children with autism got lower scores. This is due

Categories	Sub-categories	Autistic		Down		Normal	
		Mean	SD	Mean	SD	Mean	SD
Social Interaction	Nonverbal relation	8.05	3.37	11.18	3.33	15.04	1.25
	Relation with Peer	5.42	3.58	10.8	3.82	14.05	2.03
	Sharing enjoyment	11.87	5.52	16.68	4.08	20.62	1.75
	Social emotional reciprocal	4.3	2.4	6.73	2.01	8.74	1.18
	Theory of mind	3.21	2.37	4.4	2.76	7.15	1.93
communication	Delay in verbal communication	12.58	3.3	7.75	3.58	3.97	2.17
	Initiate or maintain conversation	6.53	2.38	3.83	2.85	0.96	1.12
	Stereotyped language	5.01	1.81	2.5	2.33	2.16	2.01
	Play	6.51	1.89	4.62	2.14	2.36	1.33
Stereotyped behaviors	Occupation	2.42	2.37	2.18	1.59	2.03	1.85
	Following the rituals	2.32	1.37	1.68	1.23	1.39	1.13
	Repetitive movements	2.41	1.61	1.1	1.42	1.79	1.48
Developmental delays	Delay	5.79	1.5	7.25	1	3	0.91
	Self-help skills	2.37	1.41	2.92	1.67	0.63	0.81
	Regulation factors	5.74	2.63	3.6	1.85	2.06	2.01

Table 1: Means and standard deviations of the categories and sub-categories for each group.

to the fact that this category assesses the sound social interaction in which children with autism typically fail.

The stepwise discriminant analysis showed that two linear functions, based on the sub-category scores of each category, can discriminate the three groups. To determine the discriminant functions, first the standardized coefficients are calculated for each sub-category (Table 2). The first function in all categories discriminates between autism and Down syndrome, whereas the second discriminates between Down syndrome and normal children. Therefore, the first function was more compatible with our purpose in the current study. Next, to clarify the accuracy of the discrimination function, predicted position of individuals based on discriminant functions were compared with real position of them (autism, normal and Down). Machine learning based analysis. In this way, the two classifier methods were included RF and SVM. SVM is a classifier that uses the best hyper plane, from a set of possible hyper planes, to separates between two classes. On the other hand, RF is an ensemble method in which a lot of trees are built, and then, are voted from all of trees and checked out their average votes. The result shows that about 150 trees will give good results. By using these two methods, the set of 167 items was reduced to 68 final items.

In order to diagnose autism, there is a trade-off between detecting children with autism correctly (True Positive) or falsely including non-autistic children as autistic (False Positive). It can be imagined that the cost of identifying a child with autism as normal is higher than diagnosing normal child as autistic. If a child with autism identified as normal (false negative), it can be losing critical ages to intervention and consequently, lead to poor improvements and outcomes. To prevent this error, we applied 2nu-SVM method to use different misclassification costs. Indeed, the best classification value is the one that has the least false negatives and maintain correct classification ratio in a good rate. As Figure 1 shows a point at which the false negatives are near zero and false positives are acceptable. In this case, false negatives begin to increase and false positives decrease (Figure 2). Illustrates the 2 dimensions representation of the data using RF approach in which the same misclassification cost as SVM is used. In this method the purpose was keeping amount of false negatives near zero with maintaining least false positives. So in this way RF can extract more important items that even can be result in having fewer items with the same accuracy. Figure 3 shows the accuracy and the misclassification cost in RF. In the classification performances for 2nu-SVM and RF are given (Table 3). Thus, it should be considered whether system is valid. The validity and reliability analysis of these methods are given in the next section. Table 4 shows that both SVM and RF have good accuracy. In addition, both SVM and RF have shown a very acceptable misclassification rate.

Validation findings: In this section we are going to validate the findings of the machine learning methods using statistical analysis methods. In other words, we want to show that these 68 items extracted by the machine learning methods are as good as it is claimed. The following shows the validation findings.

Validity: Two types of validity evidences were gathered. First, content validity approved by four psychologists and psychometrists Second, the experimental validity achieved by the correlation coefficient between the human expert plus GARS and the expert system was 0.97 on RF and 0.96 on SVM, that indicate a good concurrent validity for the expert system.

Reliability: To calculate the reliability, Cronbach's alpha coefficient was utilized and was equal to 0.96 showing good reliability of the expert system.

Sensitivity and specificity: Receiver Operating Characteristic (ROC) analysis was performed to determine discriminating power of 68-items form of the expert system. Table 5 shows the sensitivity and specificity for a few recommended cut-offs from the ROC analysis (Figure 4). The best cut-off is the one that its coefficients (specificity and sensitivity) be equal. In this analysis, at the cut-off score of 22.17, the coefficients are equal and were selected as a cut-off for this instrument.

Discussion

The aim of the study was to investigate the validity and reliability of the expert system designed for screening purposes in autism. At first a big item pool was prepared and edited repeatedly by the experts; then, some statistical methods were utilized to prepare items psychometrically for two forms of the questionnaire (short and long), which the later was used for machine learning analysis. Two machine learning methods, i.e. support vector machine and RF, were used which had a high accuracy that could discriminate children with autism from normal ones. In addition, specificity and sensitivity of the system using SVM is 84.1 and 98.5%, respectively, while RF is 73.4 specificity and 99.3% sensitivity. In addition, sensitivity and specificity were calculated by ROC analysis. In this study, the cut-off score of ≥ 22.17, in which the specificity and sensitivity are the same and were equal with 0.98, was the best cut-off. It should be acknowledged that performance of the system were high in the both the expert system approach and the statistical method approaches. Studies suggested that the range of 0.70 and above is an acceptable value for specificity and sensitivity [40].

Categories	Eigen value	Wilk's Lambda	Chi-square	df	Sig	Functions
Social Interaction	6.01	0.109	258.38	8	0.005	Function 1 = - 0.829 (theory of mind) + 0.782 (Social emotional reciprocal) + 0.901(sharing enjoyment with others) - 0.042 (relation with peer)
	0.31	0.763	31.5	3	0.005	Function 2 = 1.24 (theory of mind) - 1.92 (Social emotional reciprocal) - 0.616 (*sharing enjoyment with others*) + 1.84 (relation with peer)
communication	4.2	0.154	218.89	6	0.005	Function 1 = 1.472 (Delay in verbal communication) + 0.135 (Stereotyped language) - 0.635 (play)
	0.248	0.801	25.91	2	0.005	Function 2 = - 0.939 (Delay in verbal communication) - 1.05 (Stereotyped language) + 2.07 (play)
Stereotyped behaviors	4.59	0.069	312.77	6	0.005	Funtion 1 = 0.639 (Occupation) + 0.312 (Following the rituals) + 0.18 (Repetitive movements)
	1.58	0.38	111.24	2	0.005	Function 2 = - 0.059 (Occupation) - 0.788 (Following the rituals) +1.09 (Repetitive movements)
Developmental delays	3.52	0.087	283.87	4	0.005	Function 1= 0.235 (delay) + 0.846 (regulation factors)
	1.52	0.39	108	1	0.005	Function 2 = 1.2 (delay) - 0.886 (regulation factors)

Table 2: Discriminant analysis of the four categories.

Citation: Mahmoudi M, Akbari-Zardkhaneh S, Zadeh AAB, Ghobari-Bonab B, Shokoohi-Yekta M, et al. (2018) An Autism Screening Expert System: Reliability, Validity and Factorial Structure. Autism Open Access 8: 230. doi:10.4172/2165-7890.1000230

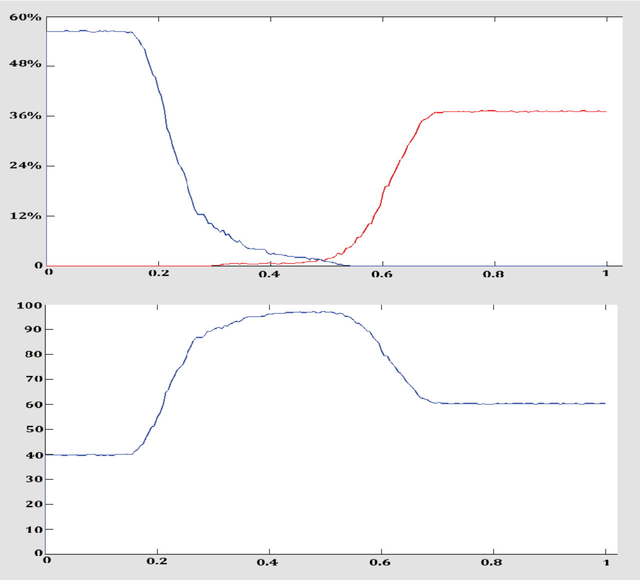

Figure 1: Upper graph shows the occurrence of false positives (Blue line) and false negatives (Red line) in 50-times random sub-sampling cross-validation for each different cost on 2nu-svm classifier. X-axis is the classification cost of FPs. Y-Axis is the probability that the observation will be misclassified. It can be seen as the misclassification cost of FPs increase, the number of FNs rise. The number of FNs tends to rise from about the cost of 0.3 for FPs and 0.7 for FNs. Lower graph shows the correct classification rate. Just around 0.3 on the X-axis seems to be a good point for misclassification costs between FPs and FNs.

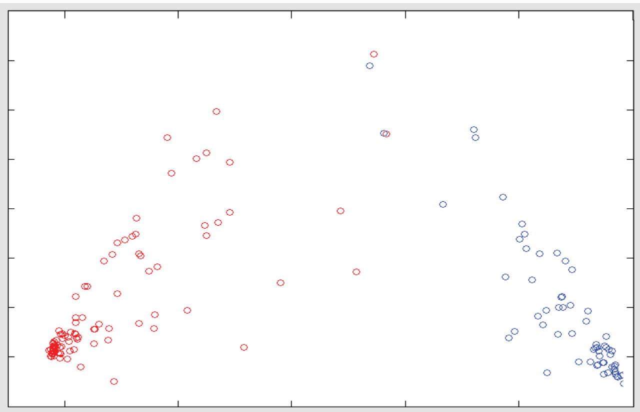

Figure 2: The data clusters obtained from Multidimensional Scaling using RF proximity measure as distance. Red circles are autistic cases and blue ones are normal.

According to this criterion can be stated that the all of these coefficients were acceptable. Given to comparison with literature, some studies are presented. For e.g. Martelto and Pedromonico reported specificity (0.81) and sensitivity (0.92) in Autism Behaviour Checklist (ABC) [16]. In regard to the study of Allen et al., specificity (0.58) and sensitivity (0.93) presented for the Social Communication Questionnaire in children aging 2 to 6 and specificity (0.62) and sensitivity of 1 in children aging 3 to 5 years old [19]. Our results seem can be with a high confidence, because the both of specificity and sensitivity can be kept in an acceptable point. In this way, we do not miss probably any child with autism and the most of normal children identified correctly. Unfortunately, there is no available information about the accuracy and sensitivity of the standardized instrument in Iran. Thus, the system cannot be compared to a nationally standardized instrument.

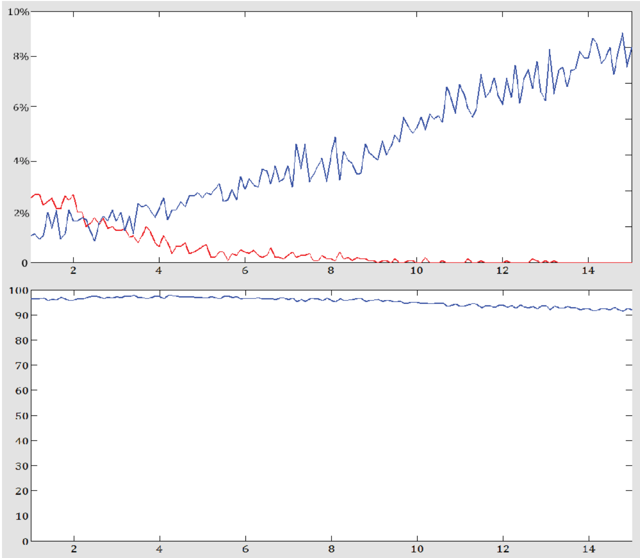

Figure 3: Upper graph shows the misclassification cost of the false negatives on the x-axis. The lower graph is the accuracy of the RF for different FN costs. It can be seen that with higher FNs cost, the rate of FNs reduces. Around 10 seem to be a good point.

Classifier	Accuracy	Specificity	Sensitivity	Misclassification Ratio	Baseline
2nu-SVM	0.924	0.841	0.985	0.17	0.375
RF	0.931	0.734	0.993	0.057	

Table 3: The classifiers used in the online Autism Screening Expert System.

Categories	Autism	Normal	Down
Social interaction	96.2	83.7	85
Communication	94.2	65.3	75
Stereotyped behaviors	90.4	93.9	100
Developmental delays	98.1	93.9	90

Table 4: Level Comparison of power of prediction in the three groups (autism, normal and Down).

	ROC AUC	Cut-off	Sensitivity	Specificity
68-item-form	0.99 (0.99-1)	≥ 20.00 P	0.981	0.94
	0.99 (0.99-1)	≥ 21.17 P	0.98	0.96
	0.99 (0.99-1)	≥ 22.17 P	0.98	0.98
	0.99 (0.99-1)	≥ 23.42 P	0.96	0.98

Table 5: Level ROC-AUC, sensitivity and specificity of 68-items-form between autism and normal.

In order to measure the concurrent validity of the expert system, one of the methods used was calculating the amount of agreement of expert system with human expert and the results of GARS. Findings showed inter-rater agreement of 0.97 on RF and 0.96 on SVM between the expert system with both human expert and GARS, indicating to be a reliable instrument in the area of human sciences. Furthermore, it indicates that this expert system can be used successfully for screening autism, because it has a good agreement with current criteria of human

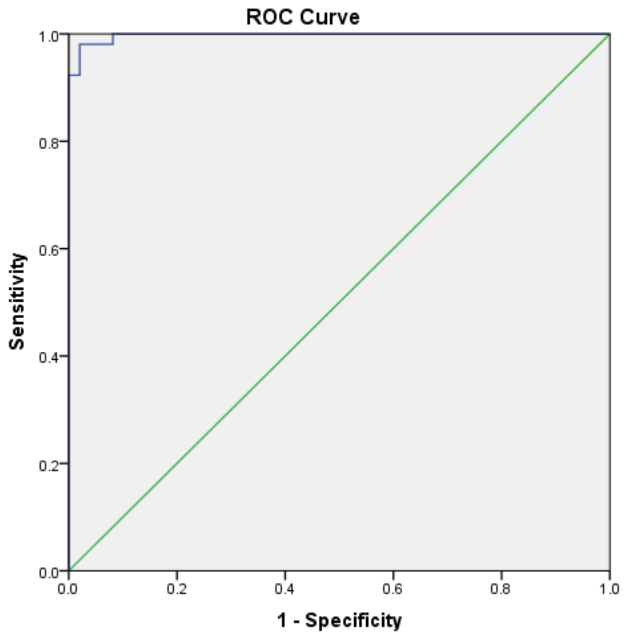

Figure 4: Level ROC Curve Receiver operating characteristic curves for autism versus normal. In sum, the expert system showed a good agreement with human expert and the screening autism instrument (GARS), suggesting a good validity for this system.

expert and GARS as screening tools. The success of expert system for screening autism has been reported in many studies.

For instance, Sajjad et al. developed an expert system for diagnosis of autism that showed successful accuracy, although not referred to as a definite coefficient [35]. In another study Kannappan et al., applied the expert system designed based on fuzzy cognitive maps for predicting autism and they could identify autism with accuracy of 0.86 correct classification rates [36]. Casado-Lambers et al. designed an expert system for the diagnoses of psychological disorders which was evaluated by two studies; results showing a considerable accuracy [32]. In sum, a number of studies indicated the validity and reliability of expert systems in order to be helpful for decision making by professionals and specialists [41]. However, in comparison with these works, our results are more promising.

In addition to calculating validity, reliability should be also measured. In order to measure reliability, the Cronbach's alpha coefficient is calculated. This was equal to 0.96 that indicates a high internal consistency. Therefore, it seems that the current expert system can be applied to help individuals in clinical affairs and diagnosis. In addition to using statistical methods, artificial intelligence procedures were used in the training and test of the data. The findings showed promising results, therefore, this expert system can be undoubtedly employed as an assistant to help clinicians in identifying autism in early ages. This expert system will be placed on the web and parents can check it freely to use for assessing their children's symptoms. In regards to the limitations of this study, the following points can be capitalized. Sample size is limited and we could trust the result with more confident if the sample size was bigger. Also, there was a considerable difference between the sample sizes of down group and the other two groups which may influence the results. Another limitation was that the system cannot discriminate varieties of autism subtypes; though these subtypes are removed in DSM-V. Further, the current expert system is in Farsi; therefore, Iranian only can use this. Our plan is to translate the questions into English so than all parents can implement it. Researchers, who are interested in this expert system, can use it for further research such as screening ADHD cases. This system can be also helpful for young clinicians and psychologists to follow the processes of diagnosis under a correct structure.

References

1. Volkmar FR, Lord C, Bailey A, Schultz RT, Klin A (2004) Autism and pervasive developmental disorders. J Child Psychol Psychiatry 45: 135-170.
2. Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, et al. (2006) Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: The Special Needs and Autism Project (SNAP). Lancet 368: 210-215.
3. Baron CS, Scott FJ, Allison C, Williams J, Bolton P, et al. (2009) Prevalence of autism spectrum conditions: UK school-based population study. Br J Psychiatry 194: 500-509.
4. Samadi SA, Mahmoodizadeh A, Mcconkey R (2011) A national study of the prevalence of autism among five-year-old children in Iran. SAGE Publications 1: 1362-3613.
5. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.555.6647&rep=rep1&type=pdf
6. Turner LM, Stone WL, Pozdol SL, Coonrod EE (2006) Follow up of children with autism spectrum disorders from age 2 to age 9. Autism 10: 243-265.
7. First L, Palfrey J (1994) The infant or young child with developmental delay. N Engl J Med 330: 478-483.
8. Momeni N, Bergquist J, Brudin L, Behnia F, Sivberg B, et al. (2012) A novel blood-based biomarker for detection of autism spectrum disorders. Transl Psychiatry 2: e91.
9. Howlin P, Asgharian A (1999) The diagnosis of autism and Asperger syndrome: Findings from a survey of 770 families. Dev Med Child Neurol 41: 834-839.
10. Norris M, Lecavalier L (2010) Screening Accuracy of Level 2 Autism Spectrum Disorder Rating Scales. A review of selected instruments. Autism 14: 263-284.
11. Baron CS, Allen J, Gillberg C (1992) Can autism be detected at 18 months? The needle, the haystack, and the CHAT. Br J Psychiatry 161: 839-843.
12. Baron CS, Wheelwright S, Cox A, Baird G, Charman T, et al. (2000) Early identification of autism by the Checklist for Autism in Toddlers (CHAT) J R Soc Med 93: 521-525.
13. Robins DL, Fein D, Barton ML, Green JA (2001) The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. J Autism Dev Disord 31: 131-144.
14. Kleinman JM, Robins DL, Ventola PE, Pandey J, Boorstein HC, et al.)2007). The Modified Checklist for Autism in Toddlers: A Follow-up Study Investigating the Early Detection of Autism Spectrum Disorders. J Autism Dev Disord 38: 827-839.
15. Krug DA, Arick JR, Almond PJ (1980) Behavior checklist for identifying severely handicapped individuals with high levels of autistic behaviors. J Child Psychol Psychiatry 21: 221-229.
16. Martelto MRF, Pedromonico MRM (2005) Validity of Autism Behavior Checklist (ABC): Preliminary study. Rev Bras Psiquiatr 27: 295-301.
17. Myles B, Simpson R, Johnson S (1995) Focus on Autistic Behavior 9: 1- 12.
18. Rutter M, Bailey A, Berument SK, Lord C, Pickles A (2003) The social communication questionnaire. Western Psychological Services.
19. Allen CW, Silove N, Williams K, Hutchins P (2007) Validity of the Social communication questionnaire in assessing risk of autism in preschool children with developmental problems. J Autism Dev Disord 37: 1272-1278.
20. Spence SJ, Sharifi P, Wiznitzer M (2004) Autism spectrum disorder: Screening, Diagnosis, and Medical Evaluation. Seminars Pediatric Neurology.
21. https://www.pearson.com/us/higher-education/product/Turban-Decision-Support-Systems-and-Intelligent-Systems-7th-Edition/9780130461063.html
22. Blonk AM, Van Den B, De Bruyn EJ (1996) Evaluation of DYSLEXPERT: A comparison of a knowledge-based system with experienced clinicians in the diagnosis of dyslexia. Computers in Human Behavior 12: 567-586.

23. Durkin J (1994) Expert systems. United States of America: Macmillan publications.

24. http://www.managertesting.com/pdfs/Validation.pdf

25. Ekman P, Friesen W, Hager JC (1978) Facial action coding system: A technique for the measurement of facial movement.

26. Lau BT (2010) Portable real time emotion detection system for the disabled. Expert Systems with Applications 37: 6561-6566.

27. Todd LK (1996) A computer-assisted expert sys-tem for clinical diagnosis of eating disorders: A potential learning tool for practitioners. Prof Psychol Res Prac 27: 184-187.

28. http://medind.nic.in/imvw/imvw12750.html

29. Fred A, Filipe J, Partinen M (2000) PSG-EXPERT: An experts system for the diagnosis of sleep disorders. Stud Health Technol Inform 78: 127-147.

30. Brasil LM, De Azevedo FM, Barreto JM (2001) Hybrid expert system for decision supporting in the medical area: Complexity and cognitive computing international. J Med Inform 63: 19-30.

31. Tsakonas A, Dounias G, Jantzen J, Axer H, Bjerregaard B, et al. (2004) Evolving rule based systems in two medical domains using genetic programming. Artif Intell Med 32: 195-216.

32. Casado LC, Rodriguez GA, Rodriguez JM, Colomo PR (2012) PsyDis: Towards a diagnosis support system for psychological disorders. Expert Systems with Applications 39: 11391-11403.

33. Marakakis E, Vassilakis K, Papadakis N (2010) Meta-rules and uncertain reasoning for diagnosis of epilepsy in childhood. Expert Systems with Applications 37: 6979-6985.

34. Georgopoulos VC, Malandraki GA, Stylios CD (2001) A fuzzy cognitive map approach to differential diagnosis of specific language impairment. Artif Intell Med 29: 261-278.

35. Sajjad S, Qamar H, Tariq K, Bano S (2011) Development of a diagnostic expert system for autism disorder-PCADEX.

36. Kannappan A, Tamilarasi A, Papageorgiou EI (2010) Analyzing the performance of fuzzy cognitive maps with non-linear hebbian learning algorithm in predicting autistic disorder. Expert Systems with Applications 38: 1282-1292.

37. Mabry SL, Williams BF (2010) Spectrum agents: Decision support for autism spectrum disorders.

38. https://books.google.co.in/books/about/Diagnostic_and_Statistical_Manual_of_Men.html?id=qRj7lwEACAAJ&redir_esc=y

39. https://www.researchgate.net/publication/284729588_Online_expert_system_for_screening_autism_an_item_analysisOnline_expert_system_for_screening_autism_an_item_analysis

40. Glascoe FP (1996) Disorders of development and learning: A practical guide.

41. Hile MG, Campbell DM, Ghobary B (1994) Automation for clinicians in the field: The validity of a performance support system 26: 205-208.