Journal of Biomedical Informatics 118 (2021) 103797

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Original Research

Q-CHAT-NAO: A robotic approach to autism screening in toddlers

Rubén Romero-García [a,*], Rafael Martínez-Tomás [a,c], Pilar Pozo [b,c], Félix de la Paz [a,c], Encarnación Sarriá [b,c]

[a] Department of Artificial Intelligence, School of Computer Science, UNED, 28040 Madrid, Spain
[b] Department of Methodology of Behavioral Sciences, Faculty of Psychology, UNED, 28040 Madrid, Spain
[c] Joint Research Institute UNED and Health Institute Carlos III (IMIENS), 28040 Madrid, Spain

ARTICLE INFO

Keywords:
Autism spectrum disorder
Machine learning
Human-robot interaction

ABSTRACT

The use of humanoid robots as assistants in therapy processes is not new. Several projects in the past several years have achieved promising results when combining human–robot interaction with standard techniques. Moreover, there are multiple screening systems for autism; one of the most used systems is the *Quantitative Checklist for Autism in Toddlers* (Q-CHAT-10), which includes ten questions to be answered by the parents or caregivers of a child. We present Q-CHAT-NAO, an observation-based autism screening system supported by a NAO robot. It includes the six questions of the Q-CHAT-10 that can be adapted to work in a robotic context; unlike the original system, it obtains information from the toddler instead of from an indirect source. The detection results obtained after applying machine learning models to the six questions in the Autistic Spectrum Disorder Screening Data for Toddlers dataset were almost equivalent to those of the original version with ten questions. These findings indicate that the Q-CHAT-NAO could be a screening option that would exploit all the benefits related to human-robot interaction.

1. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social communication, restricted interests and repetitive behaviors [1]. The prevalence of ASD increased from 6.7 per 1000 (one in 150) children in the year 2000 [2] to 18.5 per 1000 (one in 54) children in the year 2016 [3]. This increase in the number of diagnosed cases increases the need for screening and early detection tools to increase the quality of life of affected individuals. Currently, the most commonly used screening system is the Q-CHAT-10 [4], which includes ten questions to be answered by parents or caregivers of a child. The Q-CHAT-10 works as a preliminary step preceding standardized diagnoses according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders [1,5].

In general, current state-of-the-art research shows that an environment based on human-robot interaction—more specifically, the NAO robot—positively influences evaluation and therapy, although the degree of improvement is highly variable [6,7]. Furthermore, there are certain recurring problems: the poor knowledge of the clinical domain makes some results difficult to replicate or directly invalid, and other studies with good metrics have poor generalizability [8]. Furthermore, a study of this type is very expensive because it requires access to a very narrowly defined population.

Therefore, why not adapt the Q-CHAT-10 to include NAO using a toddler's reactions as answers to the test questions? Hypothetically, using reactions as answers would lead to a better evaluation. However, the answer to the question is that it is not simple: not all the questions included in the test are suitable for robotic evaluation. Furthermore, there is an immediate follow-up question: Are suitable questions sufficient to obtain good results? The main objective of this project is to answer these two questions, which aims to show that it is possible to exploit of the benefits of human-robot interaction in children with ASD. In the new system, Q-CHAT-NAO, answers to the test were not provided by the children's caregivers, but by the children's behaviors. The system automatically classifies toddlers according to the presence of the early indicators of the risk of ASD, under the supervision of a therapist, with the goal of detecting ASD at an early stage. The detection results serve the following purposes: first, to establish a clinical prioritization for diagnostic evaluation; second, to guide the planning of psychoeducational intervention and the monitoring and evaluation of any improvement caused by the intervention.

A first step in that direction is proposed in this paper. A subset of

* Corresponding author.
E-mail address: ruromgar@dia.uned.es (R. Romero-García).

https://doi.org/10.1016/j.jbi.2021.103797
Received 7 November 2020; Received in revised form 24 April 2021; Accepted 26 April 2021
Available online 30 April 2021

questions, six out of ten, were adapted such that a NAO robot could perform the assessment by interacting with a child while a therapist simply controlled and observed the scene. The results were extracted from the child's actions, which were used as a source of truth. This did not exactly mirror the indirect approach; however, the automation component relieved the therapist from certain workloads, allowing her or him to provide greater attention to details. Additionally, the classification of the Q-CHAT-NAO was supported by machine learning models. The results showed that the initial hypothesis was valid: the data obtained with the six questions adapted to the Q-CHAT-NAO would be sufficient to determine the risk of autism in a child, with results similar to those from the original test. The early stage screening test was successful. Thus, the framework is suitable for further investigation.

2. Related work

Several studies have investigated the interactions of autistic children with a NAO robot by comparing the degree of social interactions in a human-robot environment to that in a human–human environment [9,10]. The results showed very high variability, indicating a slight overall improvement when interacting with a robot versus interacting with humans [6]. The hypothesis that the human-robot interactions between a child with ASD and a robot could increase the child's interest in communicating, and thus, decrease the communicational characteristics associated with autism [11,7] has been partially probed in a study in which a NAO robot manages to attract the attention of a child and teaches them concepts about emotions through an approach based on games and songs [12]. The conclusions provided hopeful results regarding the influence of the robot on the rehabilitation of children with ASD. In addition, various studies have shown positive effects on two-way communication and the evolution of learning and social interactions [13] owing to the use of the NAO robot as a teaching tool for emotion recognition in a mobile environment [14] and to improve social skills [15]. Studies agree that an environment based on the NAO robot positively influences the evaluation and therapy with different degrees of improvement.

Furthermore, numerous studies have sought a confluence between the detection, diagnosis, and evaluation of autism and artificial intelligence; in 2012 a study applying machine learning methods to a dataset from *Autism Diagnostic Observation Schedule-Generic* (ADOS-G), the most commonly used instrument for diagnosing autism spectrum disorders, indicated that a small percentage of characteristics from the total test are sufficient to classify autism [16]. Successive studies slightly lowered expectations, and in 2015, another study sought to replicate the results without success, indicating that *although machine learning has immense potential to improve diagnosis and intervention, its use in the absence of the knowledge of the clinical domain could lead to erroneous conclusions* [17]; the study proposes a series of good practices when investigating autism using automatic methods. Certain studies warned that the set of individuals analyzed was small and highly defined and that they did not study whether the method had good generalizability for other types of patients [18]. A recent study analyzed the different state-of-the-art machine learning methods applied to the detection and diagnosis of ASD [19,8,20] to address the problem that many are outdated in terms of domain knowledge as they use the DSM-IV instead of the more recent DSM-5, proposing methods to follow regarding conceptualization, implementation and data.

Regarding detection, the 2012 article by Allison et al. defining a method that can prediagnose boys and girls aged 18 months of age and older [4] is of special importance for this study. The authors' idea is to define a *red flag*, a sign that can be used as an early decision tool to either completely rule out the presence of ASD or to mark the subject for an exhaustive subsequent review. This quantitative list of elements denoting ASD in children, or *the Quantitative Checklist for Autism in Toddlers* (Q-CHAT), is a set of questions addressed to the parents or caregivers of children that serve that purpose. Starting from 50 questions, the authors refined the method statistically until they arrived at ten questions that carried the most weight. This list serves as the basis for our study.

Currently, there is no consensus in the academic world on how to address the confluence between machine learning and the detection, diagnosis and evaluation of ASD. It can be affirmed that there is considerable variability in the results because the population with ASD is very heterogeneous, and there is generally a cautious environment when announcing generalizable results. In summary, the main problems that currently exist are the poor generalizability of the results obtained, narrowly defined study populations and a lack of clinical knowledge.

3. Early detection system

This process can be summarized as follows: implement a miniaturized version of the Q-CHAT-10 in the simplest possible way using a NAO robot as the main actor for interacting with a child, store the user's reaction data, and apply automatic methods to classify the user regarding the presence of autism symptoms. To accomplish this, a simple web service is developed; the code is accessible at https://github.com/ruromgar/q-chat-nao. The web service allows access to an implementation of the adapted questions.

The implementation can obtain the necessary data from a child through an in situ evaluation by the therapist who operates with the system. The original Q-CHAT-10 questions are single-selection multiple-choice questions with five different options regarding the frequency of the behavior described in the item. Each question is binarized: the first two options are scored as zero, and the last three options are scored as one. The tenth question, an exception to this rule, inverts the values. However, in the system proposed in this study, the extrapolation is not valid: a therapist, without prior exposure to the child cannot know if the result of one of the tests occurs *always* or only *sometimes*. The solution used was to binarize the tests at their source: the therapist only marked yes or no depending on whether the child responded as expected or not; ideally, each boy or girl will participate in several sessions at different moments to obtain a significant sample of behaviors.

3.1. Adaptation to the robotic interaction of the question set

Not all questions were adaptable. In four of the ten questions, the information had to be provided by the mother, father or responsible caregiver. Therefore, the Q-CHAT-NAO included only six out of the ten original questions, more precisely, 1, 2, 3, 4, 7, and 9: "Does your child look at you when you call her/his name?", "How easy is it for you to make eye contact with your child?", "Does your child point to indicate that she/he wants something?", "Does your child point to share interest with you?", "If you or someone else in the family is visibly upset, does your child show signs of wanting to provide comfort?" and "Does your child use simple gestures?" The starting point for all the activities was common: the child was transferred to a room accompanied by a therapist to make contact with the NAO robot. The objective of this step was to establish a climate of trust in the therapeutic context, which is convenient for the therapist to introduce the child to the NAO robot. The therapist had to be close enough to observe the child and monitor the activity, if necessary. The necessary materials included a chair for the child and a table where the NAO robot would be placed.

1. Does your child look at you when you call her/his name? The robot was located behind the child to better discern whether the objective was met or not. Once the child was seated, the robot addressed the child by calling her/him by her/his name and, after a few seconds, greeted her/him by saying: *"[Child's name]…hello, [child's name]! I'm NAO, how are you?"* A short period of silence was required between the first call to the child by her/his name and NAO's greeting sentence to give the therapist an opportunity to check whether the child responded to the call. Once the

reaction (or its absence) was marked, the activity ended. Because the activity was short, no reinforcement or a farewell was preferred.

2. How easy is it for you to make eye contact with your child? The robot was placed in front of the child to establish visual contact. Once the child was seated, the robot addressed the child by calling her/him by her/his name: *"Hello, [child's name]! I'm NAO. What is your favorite food? Mine is spaghetti with tomato and cheese."* The phrase had to last long enough for the therapist to assess whether the child kept her/his gaze on the NAO robot.
3. Does your child point to indicate that s/he wants something? The robot was placed in front of the child to establish visual contact. There were several objects, one of which was something that the child liked, such as a specific toy. Once the child was seated, the robot addressed the child by calling her/him by her/his name: *"Hello, [child's name]! Let's play for a while. Would you like to play? What do you want to play with? (…)"* The therapist noted whether or not the child pointed to the toy or another chosen object. The therapist then dedicated some time to play, and the activity ended. When finished, the NAO robot thanked the child for her/his help: *"Thank you very much, [child's name]!"*
4. Does your child point to share interest with you? In addition to the chair and table, a small speaker was located near (but not next to) the robot. The robot was placed in front of the child to establish visual contact. There were several objects, one of which was something that the child liked, such as a specific toy. Once the child was seated, the robot addressed the child by calling her/him by her/his name: *"Hello, [child's name]! I'm NAO. Let's listen to a song!"* Then, a song was heard from the speaker. At the climax, the song was stopped abruptly; primarily, the cut had to be clear and unmistakable, with a natural finish. The therapist checked whether the child was pointing to the NAO robot or the speaker, indicating that something had happened. At the end, the NAO robot said goodbye and other expressions like *"Whoa, I don't know what happened! Sorry"*
5. Does your child pretend? In their article, Allison et al. mentioned that pretending is the most critical component [4]; however, mainly because of its complexity, pretending has not been adapted to a robotic environment. When the child was discharged, the child's caretakers were asked whether the child pretended, using the original five responses as options: how many times a day, a few times a day, a few times a week, less than once a month, and never. The answers provided were binarized according to the article's criteria: many/a few times a day was coded as 0, and any other option was coded as 1.
6. Does your child follow where you are looking? As in the previous activity, the child's caretakers had to provide this information with the five original answers as options: many times a day, a few times a day, a few times a week, less than once a month, and never. The answers given were binarized according to the article's criteria: many/a few times a day was coded as 0, and any other option was coded as 1.
7. If you or someone else in the family is visibly upset, does your child show signs of wanting to provide comfort? Holding a toy, the NAO robot was placed in front of the child to establish visual contact. Once the child was seated, the robot addressed the child by calling her/him by her/his name: *"Hello, [child's name]! I'm NAO. Look what toy I have! It's new, a friend of mine gave it to me."* The NAO robot then dropped the toy onto the table and immediately showed grief: *"Oh no! What a pity! It's broken!"* The therapist checked whether, in her/his opinion, the child made any gesture to comfort the NAO robot.
8. Would you describe your child's first words as…? For obvious reasons, the child's caretakers had to provide this information using the five original responses as options: very common, more or less common, uncommon, very infrequent or my child does not speak. The answers provided were binarized according to the article's criteria: very/more or less common was coded as 0, and any other option was coded as 1.
9. Does your child use simple gestures? The robot was placed in front of the child to establish visual contact. Once the child was seated, the robot addressed the child by calling her/him by her/his name: *"Hello, [child's name]! I'm NAO. How are you?"* Then, the NAO robot said goodbye to the child with its hand, making the gesture of a farewell. The therapist observed whether the child responded with a goodbye.
10. Does your child stare at nothing with no apparent purpose? Again, because of the complexity of creating a context in which this trait could be measured quickly, the child's caretakers had to provide the information using the original five responses as options: many times a day, a few times a day, a few times a week, less than once a month, and never. The answers given were binarized according to the article's criteria: many/a few times a day/week was coded as 1, and any other option was coded as 0.

3.2. Evaluating the Q-CHAT-NAO

3.2.1. Dataset

The *Autistic Spectrum Disorder Screening Data for Toddlers* is a dataset collected by Fadi Thabtah of the University of Auckland during his investigations, and it was released in July 2018 [21]. The dataset contains 1054 responses to the Q-CHAT-10, without any missing values, and the answers are binarized following the same method as that of the authors of the original article. In the first nine questions, the first two options are scored as 0 and the last three options are scored as 1. In the tenth question, the first three options are scored as 0 and the last two options are scored as 1. The complete list of attributes is as follows:

- Binary responses to Q-CHAT-10: ten binary [0, 1] columns.
- Age of the child in months: numeric and integer.
- Q-CHAT-10 score: points scored; numeric and integer.
- Subject gender: binary [0, 1].
- Subject's ethnicity: text string.
- Born with jaundice: binary [0, 1].
- Relatives with ASD: binary [0, 1].
- Who answers the test (mother, father, caregiver, other): text string.
- Autistic traits: target column, binary [0, 1].

3.2.2. Preprocessing

Although the Q-CHAT-NAO collected age and gender data, it did not collect race, jaundice, family medical history or test purpose data; therefore these columns were removed before feeding the data to the model. Additionally, Thabtah recommends removing the Q-CHAT-10 score column as it caused overfitting, therefore, it was also removed. To simplify the test according to the adapted system, questions 5, 6, 8, and 10 ("Does your child pretend?", "Does your child follow where you look?", "Would you describe your child's first words as…?" and "Does your child stare at nothing with no apparent purpose?") were removed from the original dataset. Thus, the dataset was as follows:

- Binary responses to six Q-CHAT-10 adapted questions: six binary [0, 1] columns.
- Age of the child in months: numeric and integer.
- Subject gender: binary [0, 1].
- Autistic traits: target column, binary [0, 1].

The illustration of the data correlation, shown in Fig. 1, shows no signs of high correlation.

Regarding the target column, "ASD Traits", the data were imbalanced: 30–70% of children had ASD traits. The distribution of the data of the input features presented in Table 1 indicated that—generally—the distribution of responses was balanced, at approximately 50%. The only

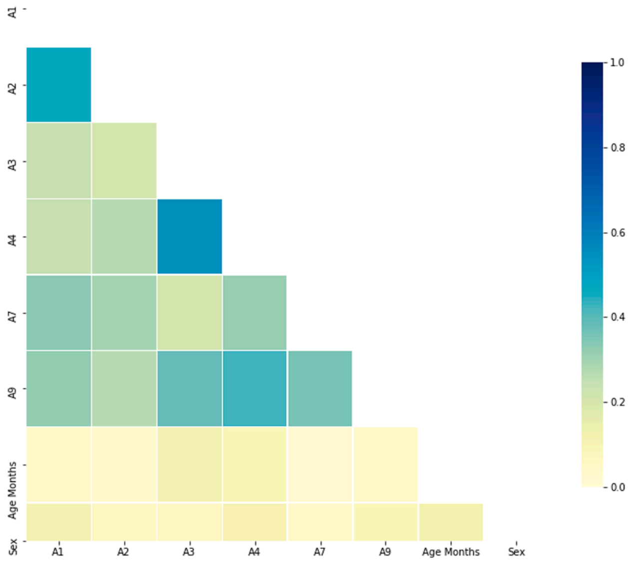

Fig. 1. Correlation of the input variables.

Table 1
Analysis of the data balance.

Feature	Percentage answering yes	Percentage answering no
Q1	56.36	43.64
Q2	44.88	55.12
Q3	40.13	59.87
Q4	51.23	48.77
Q7	64.99	35.01
Q9	48.96	51.04
ASD Traits	69.07	30.93

exception appeared to be question 7, "If you or someone else in the family is visibly upset, does your child show signs of wanting to comfort them?", which had a vast majority of positive responses. Regarding gender and age, there were more samples from male children and a very large majority of children aged 36 and 12 months compared to the other values.

4. Methods

because the interpretability of the model is a critical factor for this problem, three tree-based models were tested: the decision tree, random forest and boosted tree. Eighty percent of the dataset were assigned to a training set, and the remaining 20% were assigned to an evaluation set. Data were stratified according to the target variable. A search for hyperparameters was then conducted using grid search with ten-fold cross-validation. The optimal complexity of the models was determined using the maximum depth parameter. A maximum depth of six was set for the decision tree, a maximum depth of five with 200 estimators was set for the random forest and a maximum depth of one with 200 estimators was set for the boosted tree.

After optimizing the hyperparameters of each model and assessing the feature importance, the specificity (ACC), precision (PPV), F-score (F1), and sensitivity (REC) metrics were calculated. These measures were chosen for the evaluation because the balance between true negatives and true positives (precision-recall tradeoff) was particularly important in the context of this study. The sensitivity metric was more important than the specificity metric, and both were more important than the others (for instance, the F1 metric); therefore, we assumed the same relative importance for both.

Consider the objectives of the model: Is it possible to achieve maximum precision? Are the true positive and true negative ratios equally important? Should a perfect balance be maintained between the two? The answer to these questions is No. This model is a part of a prediagnosis system. The results will be used to conduct a more detailed follow-up, a subsequent in-depth analysis, or an expert study. Therefore, all users with ASD must be correctly classified, even if this involves a certain number of false positives. Otherwise, some users with ASD may be incorrectly classified as neurotypical, thus depriving them of the opportunity for further diagnosis. Therefore, the critical metric is the ratio of true negatives (sensitivity or recall), and it must be as high as possible.

5. Results

The relative importance of the variables was analyzed for each model as shown in Fig. 2. Questions 7 and 9 unanimously were the strongest predictors, whereas physical characteristics, such as age and sex, were relatively less important. The metrics are shown in Fig. 3. the critical metric is the ratio of true negatives (sensitivity or recall). This metric indicated that the best model was, by a slight margin, the boosted tree.

An objective of this study was to analyze whether the adapted version with only six characteristics obtained similar results as the full version with ten characteristics. The results obtained by the adapted version and the results of the method with all ten features are presented in Table 2. As recall was the most important metric, the ensemble models stood out. In both scenarios (with six and ten features), the results obtained by the random forest and the boosted tree were very similar: 91.30% vs. 91.93% with six features and 94.88% vs. 94.44% with ten features, respectively. As expected, all the metrics were better with ten features than with six features; however, the drop in the results was not considerable, and the recall remained above 91%.

6. Discussion

The present study is an exploratory study. The results appear promising, presenting a path to continue the research with a real-life environment interaction between a NAO robot and a toddler, to guarantee the validity of the Q-CHAT-NAO as a useful instrument for the early detection of ASD. In addition, we are trying to contribute to the confluence between the detection, diagnosis, and evaluation of autism and artificial intelligence.

One potential objection to our main purpose can be that we are transforming an indirect ten-items questionnaire into an interaction human robot observation-based six-items system. The answer to this objection has three components.

First, in the diagnosis or detection of ASD, it is not unusual to use both forms of assessment interchangeably to identify certain features of ASD. In reality, the ADOS-2 [22] protocol, the most widely used ASD diagnostic instrument, proposes situations for the observation of the child's interaction with the therapist to evaluate the objective behaviors of the Q-CHAT-10 items 1, 2, 3, 4, and 9 (response to call her/his name, eye contact, imperative/declarative use of the pointing gesture, use of simple gestures). Instead of asking parents the frequency of engagement of the child in these behaviors, the ADOS instructions specify situations that the therapist must address and that are intended to elicit such behaviors. This same idea supported the adaptation of the indirect questions of the Q-CHAT-10 to an observation-based system, but in this case, the observation of the child's interaction with the robot.

Second, there is currently considerable evidence regarding the positive effects of humanoid robot use in the psychoeducational intervention of children with ASD [13–15] and regarding the possible benefits of using the robot in situations that aim to elicit social interaction behaviors [6,9], which are required to assess ASD features. We described the keys to adapting six of the ten questions of the Q-CHAT-10 in the context of child—robot interactions that exploited the possible attractiveness of the NAO to elicit target child's behaviors.

Third, supporting the possibility of reducing the number of items in

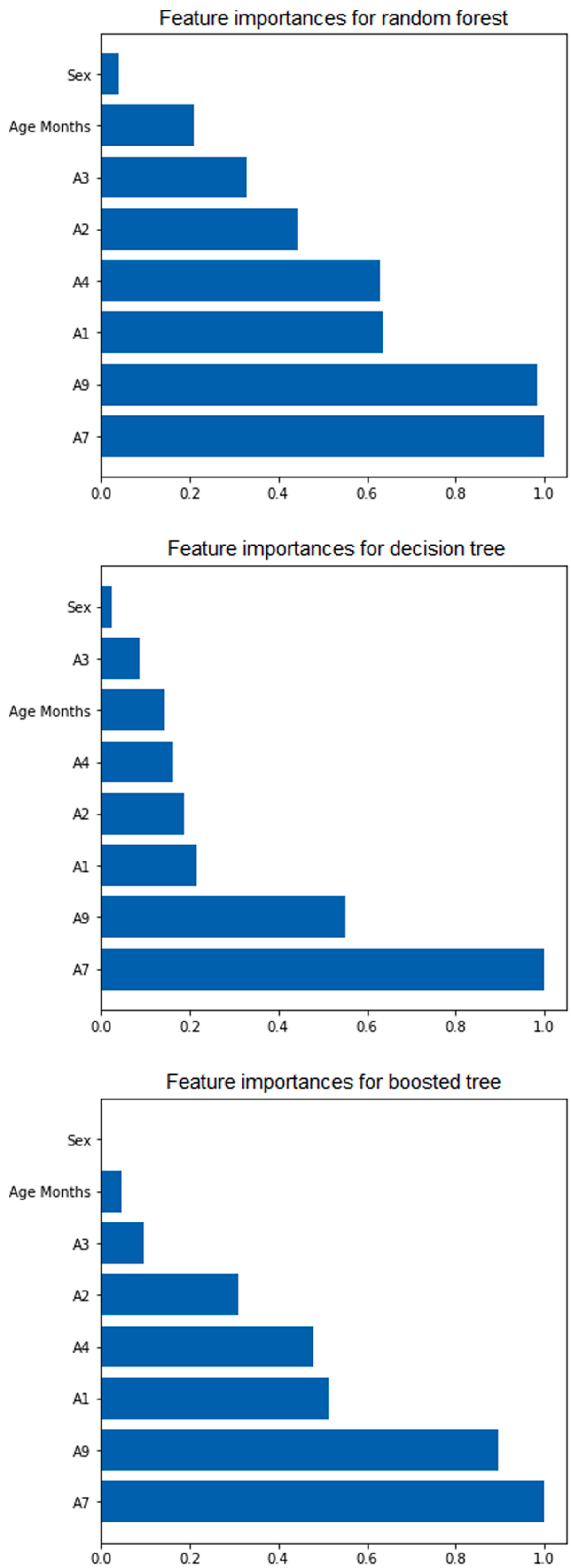

Fig. 2. Feature importances of the models.

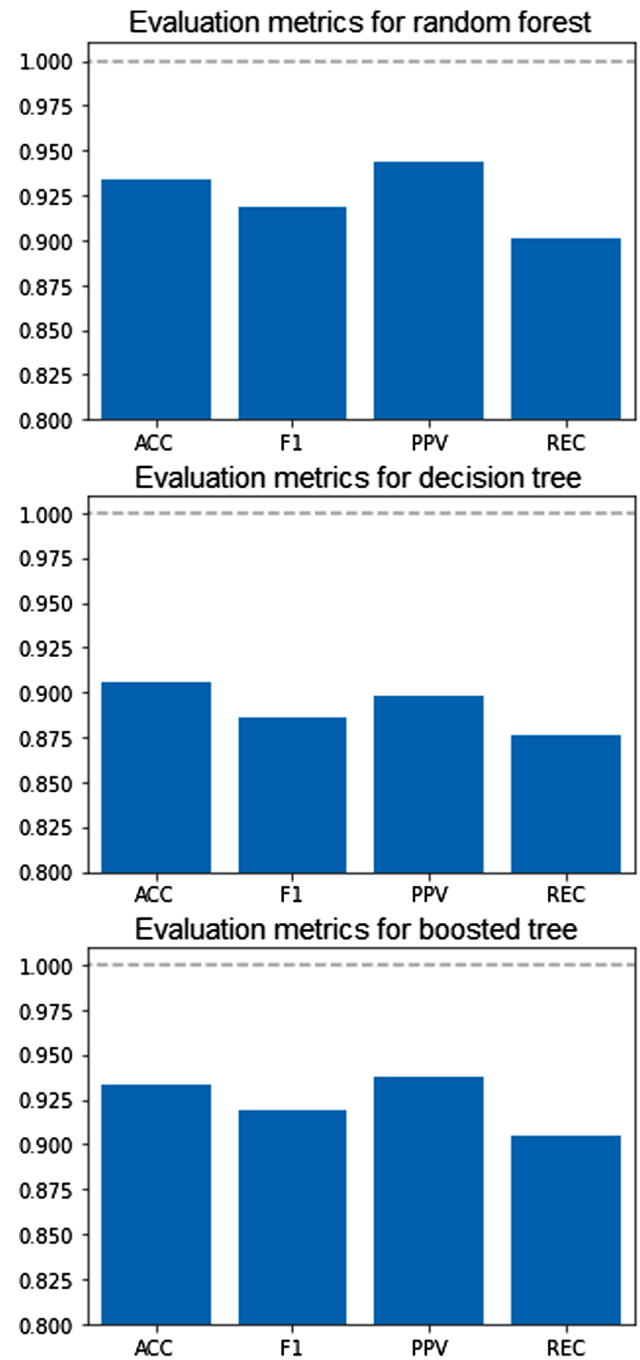

Fig. 3. Metrics of the models.

Table 2
Results per model.

Model	Accuracy	Precision	Recall	FScore
Decision tree (six features)	90.52	89.79	87.60	88.58
Random forest (six features)	92.89	93.45	89.74	91.30
Boosted tree (six features)	93.36	93.80	90.51	91.93
Decision tree (ten features)	91.94	90.20	91.20	90.66
Random forest (ten features)	95.73	96.02	93.93	94.88
Boosted tree (ten features)	95.26	94.44	94.44	94.44

the evaluation instrument from ten to six without losing sensitivity for the early detection of ASD is precisely the objective of the data analysis presented. The findings showed that the information obtained with binary responses to the six questions that could be adapted to the Q-CHAT-NAO would be sufficient to determine the risk of autism with results similar to those from the original ten items test. Consequently, we consider that the investigation path initiated is well-founded and promising.

This study has several limitations. We could have used more complex machine learning techniques. However, there was no real need for that because the main objective, that is, showing that the results with six binarized questions were similar to those with ten questions, was already achieved. The main limitation is the exploratory nature of the study, and, particularly, that it constitutes only the first step. More research is needed before the Q-CHAT-NAO can be considered as a validated instrument for the early detection of ASD. Future research could be developed in parallel in several ways. It would be necessary to test the convergent validity of the Q-CHAT-NAO by applying it to a large sample of toddlers with ASD, who are already diagnosed by an expert. Equally important is conducting a study with a sample of children with typical development to test its ability to discriminate. Finally, comparing the results obtained by the application of both instruments, the Q-CHAT-10 and the Q-CHAT-NAO, in children at the risk of ASD, and the subjective evaluation of professionals will provide valuable information on the possible advantages of Q-CHAT-NAO for early ASD detection.

However, we consider that this study is a first and an essential step for a promising line of work that will facilitate the exploitation of the child-robot interactions in caring for a group that may particularly benefit from it. The findings can also be considered as a relevant contribution thus far as they lay the foundations for an approach to evaluation based on a selection of six observable behaviors that can be elicited in different interactive contexts, whether with humans or humanoid robots other than the NAO.

7. Conclusions

The objective of this study was to explore the idea that it was possible to exploit the benefits of human—robot interaction with children for an early detection of ASD without losing any measurement accuracy. The Q-CHAT-NAO framework proposed here used an adapted subset of questions, six out of ten, of the Q-CHAT-10, to classify toddlers; furthermore, the answers to the test did not come from the caregivers but from the observation of the children's behaviors. The results obtained by applying machine learning models to the six questions in the toddler dataset indicated that the results obtained with these six predictors were very similar to those obtained with the original ten predictors, implying that there was no considerable amount of information loss. Therefore, it is concluded that the Q-CHAT-NAO presented in this study could be sufficient for generating a *red flag* for autism risk that can define the requirement for a diagnosis evaluation and a subsequent psychoeducational intervention.

This study is only a first, but essential, step and will help future research on human-robot interaction between toddlers with ASD and a NAO robot. Future research will recruit participants to test the validity of the Q-CHAT-NAO as an autism screening tool with its own dataset.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] American Psychiatric Association, Diagnostic and statistical manual of mental disorders (DSM-5), American Psychiatric Pub, 2013.
[2] Centers for Disease Control and Prevention, et al., Prevalence of autism spectrum disorders-autism and developmental disabilities monitoring network, MMWR Surveillance Summaries 56 (2007) 12–28.
[3] M.J. Maenner, K.A. Shaw, J. Baio, et al., Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, MMWR Surveillance Summaries 69 (4) (2020) 1.
[4] C. Allison, B. Auyeung, S. Baron-Cohen, Toward brief red flags for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls, Journal of the American Academy of Child & Adolescent Psychiatry 51 (2) (2012) 202–212.
[5] C.A. Mazefsky, D.P. Oswald, The discriminative ability and diagnostic utility of the ADOS-G, ADI-R, and GARS for children in a clinical setting, Autism 10 (6) (2006) 533–549.
[6] L.I. Ismail, S. Shamsudin, H. Yussof, F.A. Hanapiah, N.I. Zahari, Estimation of concentration by eye contact measurement in robot–based intervention program with autistic children, Procedia Engineering 41 (2012) 1548–1552.
[7] S. Shamsuddin, H. Yussof, L.I. Ismail, S. Mohamed, F.A. Hanapiah, N.I. Zahari, Humanoid robot NAO interacting with autistic children of moderately impaired intelligence to augment communication skills, Procedia Engineering 41 (2012) 1533–1538.
[8] F. Thabtah, Machine learning in autistic spectrum disorder behavioral research: A review and ways forward, Informatics for Health and Social Care (2018) 1–20.
[9] A. Tapus, A. Peca, A. Aly, C. Pop, L. Jisa, S. Pintea, A.S. Rusu, D.O. David, Children with autism social engagement in interaction with NAO, an imitative robot: A series of single case experiments, Interaction Studies 13 (3) (2012) 315–347.
[10] L.I. Ismail, S. Shamsudin, H. Yussof, F.A. Hanapiah, N.I. Zahari, Robot-based intervention program for autistic children with humanoid robot NAO: initial response in stereotyped behavior, Procedia Engineering 41 (2012) 1441–1447.
[11] S. Shamsuddin, H. Yussof, L. Ismail, F.A. Hanapiah, S. Mohamed, H.A. Piah, N. I. Zahari, Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO, 2012 IEEE 8th International Colloquium on Signal Processing and its Applications, IEEE 2012 (2012) 188–193.
[12] S. Shamsuddin, H. Yussof, M.A. Miskam, A.C. Hamid, N.A. Malik, H. Hashim, Humanoid robot NAO as HRI mediator to teach emotions using game-centered approach for children with autism, in: HRI 2013 Workshop on Applications for Emotional Robots, 2013.
[13] M.A. Miskam, M.A.C. Hamid, H. Yussof, S. Shamsuddin, N.A. Malik, S.N. Basir, Study on social interaction between children with autism and humanoid robot NAO, in: Applied Mechanics and Materials, Vol. 393, Trans Tech Publ, 2013, pp. 573–578.
[14] M.A. Miskam, S. Shamsuddin, M.R.A. Samat, H. Yussof, H.A. Ainudin, A.R. Omar, Humanoid robot NAO as a teaching tool of emotion recognition for children with autism using the android app, 2014 International Symposium on Micro-NanoMechatronics and Human Science (MHS) IEEE (2014) 1–5.
[15] M.A. Miskam, N.F.S. Masnin, M.H. Jamhuri, S. Shamsuddin, A.R. Omar, H. Yussof, Encouraging children with autism to improve social and communication skills through the game-based approach, Procedia Computer Science 42 (2014) 93–98.
[16] D.P. Wall, J. Kosmicki, T. Deluca, E. Harstad, V.A. Fusaro, Use of machine learning to shorten observation-based screening and diagnosis of autism, Translational Psychiatry 2 (4) (2012) e100.
[17] D. Bone, M.S. Goodwin, M.P. Black, C.-C. Lee, K. Audhkhasi, S. Narayanan, Applying machine learning to facilitate autism diagnostics: pitfalls and promises, J. Autism Dev. Disord. 45 (5) (2015) 1121–1136.
[18] A. Crippa, C. Salvatore, P. Perego, S. Forti, M. Nobile, M. Molteni, I. Castiglioni, Use of machine learning to identify children with autism and their motor abnormalities, J. Autism Dev. Disord. 45 (7) (2015) 2146–2156.
[19] F. Thabtah, Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment, in: Proceedings of the 1st International Conference on Medical and Health Informatics 2017, ACM, 2017, pp. 1–6.
[20] F. Thabtah, F. Kamalov, K. Rajab, A new computational intelligence approach to detect autistic features for autism screening, Int. J. Med. Informatics 117 (2018) 112–124.
[21] F. Thabtah, An accessible and efficient autism screening method for behavioural data and predictive analyses, Health informatics journal 25 (4) (2019) 1739–1755.
[22] C. Lord, M. Rutter, P.C. DiLavore, S. Risi, K. Gotham, S. Bishop, Autism diagnostic observation schedule, second edition, Torrance, CA: Western Psychological Services.

Ruben Romero-Garcia was born in Madrid, Spain in 1991. He received the B.S. degree in electronics and automation engineering and the M.S. degree in industrial engineering from the Carlos III University of Madrid (UC3M), in 2012 and 2014, and the M.S. degree in artificial intelligence from the Universidad Nacional de Educación a Distancia (UNED), in 2019.

He started his career as a Researcher in the UC3M, when he co-autored several conference papers in biometrics, and has then held several positions within the software engineering industry, both in Spain and UK. He is currently Technical Leader in Intelygenz, Madrid.

Rafael Martínez-Tomás received her degree in Physics from the University of Valencia, Spain, in 1983, and the Ph.D. degree from the Department of Artificial Intelligence of the National University for Distance Education (UNED), Spain, in 2000, receiving the Extraordinary Ph.D Award.

He has held different positions as a researcher and professor at the Carlos III University of Madrid and UNED, but currently is Full Professor with the Department of Artificial Intelligence and Director of the Computing Science School of the UNED. His research interests are in knowledge engineering, semantic web and semantic technologies, and semantic recognition of human behavior, publishing research papers related to these areas in various international journals and in major international conferences. His research activity has focused especially on artificial intelligence applied to diagnostic tasks in different areas of medicine, particularly behavior-based diagnosis, using different techniques and strategies, such as modeling and semantic representation of human behavior and machine learning. He has published multiple research articles in different international journals in the area of Artificial Intelligence, as well as in the main international conferences. He is currently working on the integration of humanoid robots in psychoeducational interventions with children with special needs and in the diagnosis and intelligent monitoring of the elderly.

Pilar Pozo received the B.S. and Ph.D. degrees in Psychology, from the National University of Distance Education (UNED), Madrid, Spain, in 2002 and 2010 respectively.

She is currently Associate Professor with the Methodology of Behavioral Sciences Department, UNED, Madrid, Spain. Her research focuses on autism spectrum disorders (ASD) and its impact on family context, including stress, wellbeing, social support, coping and marital satisfaction. She has co-authored several papers published in international journals about developmental disorders and intellectual disability research. Her main aim is to improve the quality of life of children with ASD and their families.

Dr. Pozo received the UNED International Doctoral School Award for Extraordinary PhD. and she was two times, in 2010 and 2018, recipient of the Spanish Association of Professionals Award for Research, as co-author of research reports about psychological adaptation of parents of people with ASD and implications for applied intervention.

Félix de la Paz received the B.S. degree in Physics from the Universidad Complutense de Madrid, Spain, in 1995 and the Ph.D. degree from Universidad Nacional de Educación a Distancia (UNED), Spain in 2003. He is currently Associate Professor with the Departament of Artificial Intelligence of Universidad Nacional de Educación a Distancia (UNED).

He has co-autored several technical papers and conference papers in robotics. Also, he is coorganizer of the biennal IWINAC meetings and editor of these conference proceedings. His current interest includes autonomous robots, social robotics, human-robot interaction and applications of robotics in education.

Encarnación Sarriá was born in Sevilla, Spain, in 1960. She received the B.S. degree in Psychology from the Complutense University of Madrid (UCM), in 1982 and the Ph.D. degree in Psychology from National University of Distance Education (UNED), in 1989.

From 1983–1985 was Research Fellow with the Social Psychology Department, UCM, Madrid. Since 1985, she has been member of staff of Faculty of Psychology, UNED, and currently she is Full Professor with the Methodology of Behavioral Sciences Department and Director of the Postgraduate Program in Autism Spectrum Disorder, UNED, Madrid, Spain. She has co-authored several books about methodology and research reports published in international journals about experimental psychology, public health and developmental disorders. Her main research interests include intentional communication and social understanding development in typically developing children and children with Autism Spectrum Disorder.

Prof. Sarriá awards and honors include the Lafourcade-Ponce Research Award in 2000, the Good practises Award (UNED) in 2008, the Faculty of Psychology Medal (UNED) in 2013, and the Angel Riviére Research Award, in 2010 and 2018.