Human Computer Interaction (In English)

Prof. Rajiv Ratn Shah

Department of Computer Science and Engineering
Institute IIT Madras

Human Centered AI for Autism Diagnosis

Today, I will be talking about our project on human-centered AI for autism diagnosis, designing context-aware assessment tools for Indian children. Autism spectrum disorder is a neurodevelopmental disorder characterized by the impairments in social interaction and the presence of repetitive behaviors. In the absence of biological markers, the diagnosis of autism is basically dependent on parent reporting and the observations made by the clinicians. This introduces lot of subjectivity in the diagnostic procedure, which is dependent on the expertise and experience of the experts. The prevalence of autism has been increasing tremendously over the years. It has increased 317% since early 2000s.

Center for Disease Control and Prevention says that 1 in 36 children are diagnosed with autism every year. And in India, the numbers are 1 in 82. It contributes to 1.8 million DALYs globally.

The increased prevalence in autism led to the increased burden in healthcare, especially in countries like India and other countries in global south where we face a lot of challenges in the healthcare sector. This contributed or this motivated us to look for solutions, technology assisted solutions, which may help the healthcare sector, which may help the clinicians in making better diagnosis and giving better services to the autistic community. Artificial intelligence tools have been found tremendously useful in areas such as education, entertainment, travel, administration and so on. Which motivated us to look into healthcare sector where AI-assisted tools could be an aid for the clinicians and thereby reduce their burden. When coming on to the mental health sector, there also AI-assisted tools and AI-assisted systems were used for diagnosing different neurodevelopmental conditions like depression, anxiety, autism, ADHD and so on.

While going through the literature pertaining to autism spectrum disorder and the use of AI-assisted tools in autism spectrum disorder, we identified a set of research gaps and this led us to the objectives of our project. First, most of this research has been limited to the global north and it doesn't cater to the needs of the socio-cultural diversity of countries like India. In this direction, we explored the acceptance of AI-assisted tools, in our case, especially social robots, in resource-contained settings of India. explored the perspectives of special educators as well as the responsiveness of children of Indian ethnicity towards

AA-assisted tools. Next, our focus was on how to identify the behavioral traits for autism from a diagnostic perspective.

And this led us to identify the challenge of comprehensive multimodal behavior analysis. In this direction, we developed interpretable machine learning tools or systems for characterizing the behavioral traits of autism. We developed modules for facial emotion recognition as well as speech analysis in children with autism. Finally, once the system was developed, we wanted to identify the challenges of establishing or implementing such a system in Indian society or Indian community. In this direction, we validated our robot-mediated AI-assisted system in the Indian context by doing a pilot study with children of Indian ethnicity.

Overall, our project is aimed at the development of an AI-assisted system for the multimodal behavior analysis of autism spectrum disorder in Indian children. First, we will look into the acceptance of AA-assisted system in Indian community. As involved in any HCI projects, ethical considerations are to be made before starting any human experiments. So, we have approval from the institutional review board of the research institute as well as the partnering institutes. First, we looked into understanding the special educators' perceptions towards AI-assisted autism care.

In our case, we specifically focused on the AI-assisted system in which the social robots were used as the task administrators or the procedure administrators. The study was focused on three main research questions. One is, what is the nature of the present ecosystem of technology-enabled interventions for autism in India? Second, What according to the Indian special educators are the perceived challenges and the benefits of using social robots as an administrator or as an assistant in autism care? And third, what initiatives could be taken to integrate social robots seamlessly into the Indian special education landscape? In this direction, we interviewed 25 special educators. We used a mixed method approach where a video presentation of the social robot interacting or administering the diagnostic task with the children were pre-recorded and this was used as an elicitation tool for the special educators. This set a background for the special educators to understand more about the project and how the social robots could be used in such a setup.

This was followed by a semi-structured interview. And finally, we conducted workshops and panel discussions with experts as well as special educators. These workshops and panel discussions were conducted including the participants who already participated in the semi-structured interviews as well. As a result of the study, there were few main findings that came out. First, highlighted the socio-technical context of the current ecosystem of autism care in India.

It highlighted the distribution hierarchies, distribution of resources and the role of hierarchy in the distribution of resources in the special education sector in India and details about the educators sentiments towards AI assisted or technology enabled solutions. The special educators expressed that the social robots could be a positive addition to the special education sector in India as they could be non-judgmental aids for the special educators in administering the task, in dealing with their day-to-day burden and giving emotional support as well. They also expressed the concern that these social robots or AI-assisted tools are expensive machines and they express their concern in the authority to use the system and to take responsibility if some damage has happened to these tools. Also, the existing system uses traditional, is adhering to the traditional intervention and diagnostic procedures and all the special educators are well trained in that. Training in new system or a new intervention technique may require them to translate all these traditional methods to the means in which the social robots could administer it or social robots are also involved in the process.

This may cause an additional burden to them. Of course, as I mentioned before, they express their concern for damages and the accountability of this damage. They have told that most often in this special education sector, often the special educators take care of these tools and they may have to give a part of their income towards the damage of these tools if something happened to them. Further, they expressed their concern of replacement by AI-assisted tools. We had to convince them that AI-assisted tools are supposed to be an aid to the special educators and in no way they will replace the special educators in near future.

Even though they expressed all these concerns, they were very positive in adapting these technologies in the benefit of the children. After seeing the positive effects on the effects of the socially assistive robots or in any other similar AI assisted tools for that matter, in the improvement or the development of skills in children with autism, they are willing to learn even this technology even if it is novel, even if it is new and even if it adds to their burden. In the benefit of the children, they are very positive towards incorporating such tools in their daily routine for autism care. From these findings, we have identified that The needs, the challenges and the expectations of the special educators are in the direction of the development of cost-effective and robust social robots or AI-assisted tools. Next, as I said, all these experts or all these special educators are well-versed and trained in the traditional system.

And a transition from this traditional system to any novel technology will add burden to their daily routine. So the tools that are developed should be developed in such a way that there is ease of integration of these tools with the existing methods or approaches. The special educators also highlighted the need for catering to the diverse population of children with autism spectrum disorder. As ASD is a spectrum of symptoms, its presentation also varies from one child to another or one individual to another who is affected by ASD. In that case, the tools which are developed should be customizable to the needs of each child, each individual and it should be catering to the diversity of the symptoms presented in these children.

Further, the current AA assisted tools are mostly based on the diagnostic tools that are developed and used commonly in Global North. Further, the experts highlighted the point that the existing AA-assisted tools that have been developed are based on tools or standards that are developed in Global North. So, they lack the cultural appropriateness and they lack the sensitivity towards the cultural norms in countries like India and hence the AI assisted tools or systems that are developed to cater this population should also be context aware and culture specific. In order to address their fear for replacement, they also Demanded for appropriate policies and ethical guidelines which rules the use of AI-assisted tools or social robots in autism care. Finally, they demanded for the training and support which helps them with the transition from the traditional tools to the AI-assisted tools.

Overall, our study highlighted that despite challenges, special educators have positive sentiments about integrating social robots or AI-assisted tools as aids for children with autism, expecting positive outcomes in the diagnosis and interventions. Once we understood the perspectives of special educators towards AI-assisted tools for autism care in India, we progressed to understand the responsiveness of children towards robot-assisted interventions or for that matter, AI-assisted interventions. Since social robots are a new concept for the Indian society, Even the responsiveness of typically developing or neurotypical children towards social robots are not available. In that case, our study proceeded with understanding the responsiveness of neurotypical or typically developing children towards robot-assisted or robot-mediated interventions. The major research questions of this study was focused on what will be the responsive behavior of typically developing children of Indian origin towards robot mediated interventions targeted on autism diagnostics.

We recruited 10 participants. who are children of the age 3 to 6 years, out of which 8 were female and 2 were male participants. We used a Cosmo robot, which is a toy-like robot for this study. The study was planned in three sessions. First, the introduction where the researcher introduced the social robot and the task to the children, followed by a response to name task where the child's name will be taken by the robot and the child is supposed to respond with a social cue.

And third is joint attention task where the child shares attention with the robot. For this, the tasks were selected from the diagnostic tool developed by All India Institute of Medical Sciences, which is the AIIMS-modified Inclin Diagnostic Tool for ASD. We used a survey questionnaire towards the end of the study to understand the parent perceptions about the interactions administered by robots with the children. Our experimental results show that all the children could successfully complete the task. In this table, the lower the score means the higher responsiveness and the higher the score shows lower responsiveness.

We could see that 9 out of 10 children could successfully complete the task with score either 0 or 1 which shows higher responsiveness of these 9 children out of 10. Further, our analysis of the engagement of the children showed that there was sustained attention or there was sustained engagement of these children towards the robot mediator task. This was evaluated using the eye contact metrics that the children made, the eye contact that they made with the robot during the interaction. which is an indirect measure to show the engagement of the children during the interactions. The major takeaway of this study was that all the children showed positive response towards robot assisted interactions.

After identifying the perceptions and exploring the responsiveness of children towards robot assisted interventions, we moved ahead towards our AA-assisted behavioral modeling for identifying the behavioral traits which are indicative of autism spectrum disorder. First, we looked into identifying the facial expressions of children. In the absence of enough dataset for training large-scale deep learning models, we formulated our problem in scarce data settings. Facial expressions are kind of like a mirror to our mind or they show what our intentions and emotions are to some extent. Towards this, we formulated a source-free multi-source transfer learning approach.

Our hypothesis was that the knowledge gained from multiple data sources of similar domains could be leveraged to inform a new target model on a related task. In the absence of sufficient data set for training large-scale deep learning models, formulated this problem as a multi-source transfer learning problem. We identified that even in the absence of large facial emotion data set for child's emotions, There are smaller data sets which are available and there are many pre-trained models which are available, which are trained on these emotion data sets. So why can't we leverage the information that these models have already learned, which led to our multi-source transfer learning approach. Our hypothesis also assumed that the knowledge gained in such a way will be more generalizable and can perform better on a new domain.

In this direction, we proposed a multivariate maximal correlation based multi-source transfer learning approach, which is built on an ensemble of pre-trained source networks to represent a target classifier. Every deep learning network could be considered as a combination of two functions. One which maps the input data to the feature function or feature maps. And second part is where the function takes in these features and convert it into the final outcome or the final label. Here, every pre-trained network which we have accessed, you have to note here that we don't have access to the data with which the pre-trained network has been trained on, we only have the pre-trained network which is trained on some data.

The data is not available and we don't have control over the training process of this pre-trained network. So as I said, this network could be considered as two part. One is the function that maps the data to the features, x to the features. And second part is where these features are further mapped to the output label. So in our approach, we have used the target samples along with this feature extraction part and extracted the features from each one of these pre-trained networks.

Now, we used maximal correlation on these features to find out the maximal correlation functions, gsy for source one, gsy for source two and gsy for source n. Now, These functions are formed in such a way that the maximal correlation coefficient rho of s1, rho of s2 and rho of sn, the aggregate maximal correlation is maximized in this ensemble. I will repeat once again, we use the features coming out from each of these pre-trained networks based on our target sample and we used maximal correlation analysis to generate the classification function which is GSY here in such a way that the aggregate maximal correlation coefficient of this entire example is maximized. Once this model is trained, the final prediction is based on the product of FS1 and GS1 weighted by the maximum correlation coefficient. We experimented this approach using Lennet and ResNet as pre-trained models.

We used a homogeneous models for as the pre-trained networks. And we could see that our model performed well with different kinds of, with an ensemble of different facial expression recognition data sets. We used RAFDB, FEA 2013, CAFE and JEFE data sets as our source data set. And we use different permutations of these data sets for the experiments. And we could see that in all the settings, our approach could learn better information from the ensemble of these pre-trained source networks.

Here, in the first setting, the RAFDB, CAFE and JEFE are considered as the source datasets and the FER2013 is considered as the target dataset. Further, in order to identify the generalizability of this approach towards other image datasets as well, performed the experiments on other image datasets. We performed the experiment on a lab-curated children's facial expression data, which consists of facial expression data of children of Indian ethnicity. With that as well, our MSDL-MCA approach performed relatively well with the state-of-the-art approaches. We could also find that our approach could learn from lesser number of data samples.

And further, we could see that the higher the correlation coefficient, the higher the contribution of a source task towards the target task. The key takeaway of this approach was that the MSDL-MCA approach could translate the knowledge learned from multiple sources to target task, and this has been further validated using benchmark dataset and lab curated datasets. We then proceeded towards the speech signal analysis for autism diagnosis. There are multiple approaches that have been explored for the automatic extraction of speech behaviors for autism. For example, social interactions were measured using voice characteristics, the intonation and the rhythm of speech, the different acoustic features, et cetera, were used for the identification of ASD behaviors in children.

And there were many automated language markers that were identified for the ASD behaviors. Our hypothesis was that, Many of these features are basically oriented on the prosodic and acoustic features, and there has been less focus on semantic and pragmatic features of speech. Based on this, we formulated our hypothesis that children with autism exhibit distinct semantic and pragmatic language features, which could help in better identification of the speech behaviors indicating autism. Further, the presentation of speech behaviors may differ according to the language variabilities. We first identified the significance of semantic and pragmatic features for distinguishing children with ASD and TD.

evaluated the automatic detection of features for Hindi language, and then performed a cross-linguistic analysis for speech features in Hindi and English. For English data, we used Child's dataset, and for Hindi, in the absence of any valid dataset for use, we collected the dataset from children who are neurotypical as well as who are under the spectrum. For collecting the data, we conducted a session, which consists of eight diagnostic activities, including emotion identification, pointing, imitation, conversation, response to name, joint attention, ball catching, drinking from a cup, handshaking, and so on. These diagnostic activities are selected from the Inclined INDT ASD tool for autism screening. There were 15 participants from ASD and 18 participants from TD group.

The age groups were four to 14 for ASD and four to 10 for TD. The range for ASD participants were kept slightly broader in order to equate them with the verbal capabilities with the TD kids. This is roughly the architecture that we have used. The child partner conversation was recorded using microphone. And then the speech signals were pre-processed to extract the audio signals as well as the transcripts.

And then the different speech acoustic and linguistic speech features were extracted. And finally, using a classifier, we distinguished the ASD kits from the TD kits. These are the different features that we have used for this classification, which were broadly classified under linguistic and acoustic features. We used majority classifier as the baseline classifier and We identified that the linguistic and acoustic features together gave an F1 score of 91.3%, which is quite good for an AI-assisted diagnostic system.

Further, we evaluated and analyzed the contribution of the linguistic and acoustic features towards the identification or towards the classification of ASD versus TD. This analysis was done between the English speaking group and the Hindi speaking group. For this we leverage the SHAP analysis. In our analysis, we found that in Hindi-speaking children, features like unintelligent propagation, average length of utterances, POS tags which represents the ad positions, discourse coherence, etc.

contributed more towards the classification. And in Hindi speaking participants, the adverbs, non-distinct word roots, number of distinct word roots, POS tags showing the conjunction, etc. contributed more towards the identification of ASD kits. These variations were indeed contributed by the specific nature of the language and the and the grammatical structure of the language. For example, English places greater emphasis on the word order and the word choices, whereas Hindi exhibits more flexible word order. For example, the adjoins or adpositions differ widely in Hindi with respect to English.

The key finding of this work was that whenever we develop AI-assisted tools, we have to cater to the needs of the linguistic diversity of that population. Linguistic diversity is indicative of their cultural diversity and hence it has to be taken care while developing any AI-assisted tools. Finally, we wanted to validate the AA-assisted system and find its performance and evaluate its performance in children of Indian ethnicity. In this direction, we conducted a pilot study. And in this pilot study, we collected the facial expression as well as speech behaviors of children of autism who are of Indian origin.

In this, the robot assistant took the role of administering the assessment task and the AI-based behavior analysis system detected the behaviors throughout the assessment process. There were 21 participants recruited for this study, out of which 9 were ASD. There was a female-male ratio of 7 is to 14. The age group was kept between 4 to 14. We used the NAO robot, which is a humanoid robot, as the social robot assistant for administering the task.

And the diagnostic tool was the AIMS-modified inclined diagnostic tool for ASD. As used in our previous study, we used the different diagnostic activities extracted from INDT-ASD, which are conversation, response to name, response to joint attention, functional limitation, anticipation of objects, and vocabulary and emotions. Each child engaged with both the human administrator and a robot assistant. The participants were divided into two groups. In one group, they interacted with the human administrator first, followed by the robot.

And the second group, they interacted with the robot and then the human administrator. We wanted to avoid the bias of novelty when the children are interacting with the robot first. That is why we used two groups. This is the architecture of the system used. The child interacted with the expert or the robot mediator and during which the audio and video data were collected and this audio data was sent to the speech recognition module or speech analysis module and the video data where the frames were extracted, were sent to the facial expression recognition module, which was based on MSDL-MCA approach.

Finally, the predictions were made based on the features extracted from speech modules and the video module. And a report was generated, which lists out the speech behaviors as well as the facial expression behaviors exhibited by the children during the interaction. This report could be shared with the experts so that they can use this to cross verify their observations during interventions. We found that In both the groups, the robot mediated system using the AI-assisted analytical tool gave comparable accuracies with the human administrator.

The mean accuracy is 76.19 with an F1 score of 78.26. This shows positive evidence towards AI-enabled robot-assisted assessment systems. In this work, the main contribution was towards the development of an AI-assisted system for the assessment of ASD in children of Indian ethnicity. The traditional system being highly reliable on the expertise and experience of the clinicians, our tool gives an aid to these clinicians by reducing their burden and by making more engaging interactions possible for the children during the diagnostic procedures. Our system is a good example where the HCI and AI hold hands and develop a solution that is in the benefit of children or in specific case of children with autism. These are some of the references that you can follow to understand more about this domain.