A Predictive Model for Paediatric Autism Screening

Journal:	*Health Informatics Journal*
Manuscript ID	HIJ-19-0129.R2
Manuscript Type:	Original Article
Keywords:	decision-support systems, machine learning, autism spectrum disorder, data mining, computer aided screening tool
Abstract:	Autism Spectrum Disorder (ASD) is an umbrella term for a group of neurodevelopmental disorders that is associated with impairments to social interaction, communication, and behaviour. Typically, ASD is first detected with a screening tool (e.g. M-CHAT). However, the interpretation of ASD behavioural symptoms varies across cultures: the sensitivity of M-CHAT is as low as 25% in Sri Lanka. A culturally-sensitive screening tool called Pictorial Autism Assessment Schedule (PAAS) has overcome this problem. Low and Middle-Income Countries (LMIC) have a shortage of mental health specialists, which is a key barrier for obtaining an early ASD diagnosis. Early identification of ASD enables intervention before atypical patterns of behaviour and brain function become established. This article proposes a culturally sensitive ASD screening mobile application. The proposed application embeds an intelligent machine learning model and uses a clinically validated symptom checklist to monitor and detect ASD in LMIC for the first time. Machine learning models were trained on clinical PAAS data and their predictive performance evaluated, which demonstrated that the Random Forest was the optimal classifier (AUROC 0.98) for embedding into the mobile screening tool. Additionally, feature selection demonstrated that many PAAS questions are redundant, and can be removed to optimise the screening process.
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
hij_final.zip	

A Predictive Model for Paediatric Autism Screening

Journal Title
XX(X):1–12

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Benjamin Wingfield[1], Shane Miller[1], Pratheepan Yogarajah[1], Dermot Kerr[1], Bryan Gardiner[1], Sudarshi Seneviratne[2], Pradeepa Samarasinghe[3], and Sonya Coleman[1]

Abstract

Autism Spectrum Disorder (ASD) is an umbrella term for a group of neurodevelopmental disorders that is associated with impairments to social interaction, communication, and behaviour. Typically, ASD is first detected with a screening tool (e.g. M-CHAT). However, the interpretation of ASD behavioural symptoms varies across cultures: the sensitivity of M-CHAT is as low as 25% in Sri Lanka. A culturally-sensitive screening tool called Pictorial Autism Assessment Schedule (PAAS) has overcome this problem. Low and Middle-Income Countries (LMIC) have a shortage of mental health specialists, which is a key barrier for obtaining an early ASD diagnosis. Early identification of ASD enables intervention before atypical patterns of behaviour and brain function become established. This article proposes a culturally sensitive ASD screening mobile application. The proposed application embeds an intelligent machine learning model and uses a clinically validated symptom checklist to monitor and detect ASD in LMIC for the first time. Machine learning models were trained on clinical PAAS data and their predictive performance evaluated, which demonstrated that the Random Forest was the optimal classifier (AUROC 0.98) for embedding into the mobile screening tool. Additionally, feature selection demonstrated that many PAAS questions are redundant, and can be removed to optimise the screening process.

Keywords

Autism spectrum disorder, machine learning, decision support system

Introduction

Autism Spectrum Disorder (ASD) is a developmental disorder that affects social interaction, communication, and behaviour. ASD can be diagnosed at any age, but symptoms manifest within 24 months of birth, and ASD affects up to 3.8 children per 1,000 in the U.K.[1]. However, less evidence is available for ASD prevalence estimates in Low and Middle-Income Countries (LMIC)[2]. ASD detection is poor in LMIC compared with developed countries due to research and funding limitations in LMIC. To date, LMIC have only completed limited research to determine how many of their citizens are autistic, and health officials who say ASD is non-existent in their regions likely don't know how to identify it[3]. This is a deeply concerning issue and there is an urgent need for more support and services for these individuals living in LMIC.

Early identification of ASD in children enables intensive intervention before neuronal pruning is completed[4]. It has been reported that not addressing ASD at a young age has a major influence on development into adulthood and results in a high economic cost, exceeding the lifetime costs of asthma, intellectual disability and diabetes[5]. It should be noted that many LMIC provide medical and care home facilities for free or with reduced costs for their citizens. However, it can still cost a substantial amount of money and governments must allocate this in their annual budget. Additionally, people with mental health disorders cannot be included in the national workforce, which will negatively impact the economy of LMIC.

Typically ASD is first identified in children using a screening tool that implements a symptom checklist such as the Modified Checklist for Autism in Toddlers, revised with follow-up (M-CHAT-R/F)[6]. Screening tools are preferred over clinical observations at this early stage as atypical behaviour can be absent or amplified in busy outpatient clinics (e.g. due to anxiety in an unfamiliar environment), impairing ASD detection. M-CHAT-R/F is a pair of 20 items checklists of symptoms and is valid for children between 16 and 30 months old. M-CHAT and its derivatives are a popular choice for screening ASD in young children as they are quick to administer (taking 2 – 3 minutes to complete). As ASD is an inherently biological phenomenon the symptom checklists are consistent across different ethnic environments. However, the description and interpretation of ASD behavioural symptoms varies across different cultures. Therefore culture can affect the ability of screening tools to detect ASD. The sensitivity of M-CHAT is as low as 25%[7] in Sri Lanka. To overcome this a culturally sensitive

[1]Ulster University, UK
[2]University of Colombo, Sri Lanka
[3]Sri Lanka Institute of Information Technology, Sri Lanka

Corresponding author:
Benjamin Wingfield
Intelligent Systems Research Centre
Ulster University , L'Derry, BT48 7JL, UK
Email: b.wingfield@ulster.ac.uk

screening tool that incorporates written and pictoral content has been developed called Pictorial Autism Assessment Schedule (PAAS)[8]. PAAS has been found to be an effective screening tool for ASD in Sri Lanka, with a sensitivity of up to 88% when discriminating between ASD and neurotypical paediatric subjects.

Both screening tools and ASD interventions are usually administered by mental health specialists in collaboration with the child's parents. LMIC have a shortage of mental health specialists, which is a key barrier for obtaining an ASD diagnosis and accessing services that improve the prognosis of autistic children. Long waiting times (often multiple years) are therefore common in LMIC and many children do not receive any diagnosis or treatment at all, with a treatment gap of up to 100% in some areas[5]. Increasing the number of mental health specialists is unrealistic for many LMIC due to insufficient resources. An alternative approach is to use non-specialist healthcare workers to administer the screening tool at home or in healthcare clinics. The benefit of this approach is two-fold: firstly, minimal additional resources are required. Secondly, there is a low awareness about ASD in LMIC. Parents will not be aware of developmental delays and their potential link to ASD and will instead consider the symptoms of developmental delay to be normal behavioural deviations (i.e. they may consider their child to be badly behaved rather than developmentally delayed). ASD screening can be integrated into standard paediatric checkups performed by non-specialist healthcare workers (e.g. vaccination procedures) to improve diagnosis rates and to increase parental awareness of developmental delays. To assist non-specialist healthcare workers it is essential to develop computer aided tools to conduct ASD screening and provide early intervention activities for autistic children.

To overcome the challenges described above, a mobile application that predicts ASD from a clinically validated culturally sensitive symptom checklist is proposed. Multiple machine learning algorithms are thoroughly evaluated on clinical PAAS data, and the best performing algorithm is embedded into the application. The proposed application can be administered by non-specialist healthcare workers in LMIC countries at home, to advise if a clinical referral is recommended. As more data are collected, the application can be refined and improved with software updates. An analysis of the PAAS checklist with feature selection algorithms can reveal which questions are superfluous, and can be used to refine the checklist in the future. In this article, an innovative mobile application for ASD screening in LMIC countries is developed. There are limited studies that have embedded an intelligent machine learning algorithm into a mobile screening application. An ASD screening application that incorporates intelligent decision making in combination with a culturally sensitive and clinically validated screening tool is novel, and this combined approach will reduce the burden of the shortage of mental health services in LMIC. Additionally, this will also enable the early detection of ASD in LMIC, improving clinical outcomes. Furthermore, valuable data can be collected about the prevalence of ASD in LMIC, which is severely lacking.

This article is structured as follows: the sections on 'Machine learning and intelligent methods for autism detection' and 'Autism screening mobile applications' critically reviews current machine learning algorithms applied to ASD prediction and mobile ASD screening applications. The methods section describes the data collection and analytics pipeline. The section 'Model evaluation' reviews the predictive performance of each model, discusses choosing a final model for embedding into the mobile screening an application, and investigates the importance of PAAS questions for ASD detection. In the final section conclusions, future work, and limitations are presented and considered.

Literature review

Machine learning and intelligent methods for autism detection

Machine learning algorithms have been broadly applied to detect ASD and to investigate its uncertain aetiology[17]. For ASD detection a variety of data types have been used as input to supervised learning algorithms, including screening tools such as the Autism Diagnosis Interview Revised (ADI-R[18]) and the Autism Diagnostic Observation Schedule—Generic (ADOS[19]). Popular supervised learning models included support vector machines[20] and decision tree variants[21,22]. Random forests[23], neural networks[24], least absolute shrinkage and selection operator (LASSO) regression[25], and ridge regression[25] were also popular. Other input data types such as functional MRI[26], eye tracking[27], and genetic data[28] have also been used for the detection of ASD.

A method of combining questionnaire and home video screening data has been shown to be a reliable method of detecting early autism at home[29]. The output of two independent machine learning classifiers was combined into a single screening outcome, with clinical results showing a sensitivity of up to 70% and a specificity of up to 67%. This approach included novel feature selection, feature engineering, and feature encoding methods to achieve this result.

A new machine learning method called Rules-Machine Learning has been developed to detect ASD from questionnaire data[30,31]. Many of the machine learning approaches discussed above, such as artificial neural networks, are known as black box models. The decision that a black box model makes cannot be interpreted[32] from the structure of the model (e.g. neuron weights). The Rules-Machine Learning approach induces rules to provide a knowledge base that domain experts can interpret to provide an understanding of why a decision has been reached.

Feature selection has often been applied to autism screening questionnaires to identify minimal feature subsets to speed up the diagnosis process[17]. Notably, a novel computational intelligence algorithm called Variable Analysis has been used to detect a small set of core relevant features while maintaining predictive performance[9,31].

An interesting approach that detects ASD from the presence of co-occurring conditions has been proposed[33]. For example, co-occurring conditions could include obesity, developmental delay, and speech problems. The approach demonstrated an accuracy of up to 86% for two classes (ASD

Table 1. Computer assisted screening tools for ASD. Top: automated screening tools. Bottom: augmented screening tools.

Name	Language	Description	Platform	LMIC target?	Ref.
ASDTests	11 Languages	Implements four different tests across multiple age cohorts	Android & iOS	No	9
AutismAI	English	Implements an intelligent machine learning algorithm to predict ASD from 10 questions	Android	No	10
Autism Barta	Bengali	Implements a pictorially augmented M-CHAT screening tool. Provides a framework for storing responses on an online database.	iOS	Yes	11
Autism & Beyond	English	Uses inbuilt camera to analyse facial expressions during visual stimuli	iOS	No	12
Naturalistic Observation Diagnostic Assessment	English	Service platform to record video of autistic children in certain scenarios at home. Videos are assessed remotely by clinicians according to DSM-IV criteria.	iOS	No	13
Gaze-Wasserstein	English	Using eye-tracking to quantitatively measure gaze and classify ASD	iOS	No	14 15

present or absent). The approach could not reliably predict the severity of ASD from the co-occurring conditions.

An adaptive Bayesian classifier (ABC) has been used to predict ASD[34]. An advantage of the ABC system is that the classifier will reclassify data if a confidence threshold is not met. This would be particularly useful for clinical applications. Combining the output of multiple weak predictors to form a single strong predictor (ensembling) is a common technique in machine learning[35]. A proposed system called Ensemble Classification for Autism Screening (ECAS)[36] implements an ensemble approach to predict ASD from data collected via the ASDTests Android application[9,31] — described further in the next subsection — in children. The ensemble system performed better in benchmarks compared with other common machine learning algorithms. However, ensemble systems are generally considered to be black boxes, and it is difficult to understand why certain features contribute to the overall decision. A logistic regression analysis of the same data has been undertaken[37] to screen for ASD and identify influential features for ASD detection. The system was capable of accurately predicting ASD in adults and children whilst identifying key features of interest.

More advanced Computational Intelligence approaches such as Fuzzy Set Theory have been applied for ASD detection in adults. A hybrid system based on fuzzy rules has been proposed to predict ASD with a high degree of accuracy[38]. Fuzzy set theory can model the uncertainty present in real-world data in a transparent way, which is particularly useful for biomedical and clinical applications[39].

Autism screening mobile applications

A variety of computer aided screening tools for ASD detection have been developed. The computer aided tools fall into two broad categories: automating current paper-based screening tools (e.g. M-CHAT) or augmenting screening tools (e.g. using gaze tracking to collect a quantitative measure for classification[14]). A significant number of these computer aided tools have been implemented as a mobile application (see Table 1), as smart devices provide an ideal platform for mental health disorder screening[40]. Smart devices have a large amount of sensors and are widely available (including in LMIC): 36% of the world's population owned a smartphone in 2018[41]. The automation tools implement a variety of standardised symptom checklists (e.g. M-CHAT) and sometimes incorporate other data modalities to improve comprehension (e.g. pictoral data in addition to textual data). The augmentation tools typically record quantitative data (e.g. from eye gaze) and perform classification using supervised learning algorithms to identify ASD in subjects.

ASDTests[9,31] is an Android application that contains four different screening tests for toddlers, children, adolescents and adults. It is proposed that the application can be used by health professionals to signpost individuals towards a formal autism diagnosis. The application also provides the opportunity to collect valuable data from the four different age groups to improve the efficiency and accuracy of the screening process. However, the application does not incorporate an intelligent machine learning decision model to make the recommendation, and the question sets have not been developed in collaboration with clinicians from LMIC as PAAS has. AutismAI[10] is an Android application that incorporates intelligent machine learning decisions from 10 questions according to an age cohort. However, these 10 questions have not been validated in a clinical setting in LMIC as PAAS has.

Methods

Data collection

The PAAS checklist was developed from a mixture of different sources, including Diagnostic and Statistical Manual of Mental Disorders (DSM-V[42]) and M-CHAT and modified to include cultural considerations[8]. The checklist contains 21 items (see Table 3), and each item is written in English, Sinhala, and Tamil (Sinhala and Tamil are the

two main languages spoken in Sri Lanka). Each item on the checklist is associated with a picture or a video to aid comprehension. Acceptable responses for each item include "yes" or "no" (e.g. "Does your child bring over things to show you?"). The child's age and gender is also recorded but not used for assessment purposes. ASD is detected by PAAS if four or more positive indicators are identified. Positive indicators for ASD are defined as a positive response to questions 15, 16, and 21, and a negative response to any other questions (see Table 3).

PAAS responses from 228 children (see top section of Table 4) were collected and used to train and evaluate a variety of different supervised learning models, including 33 control subjects and 195 ASD subjects. Up to 5% of PAAS were missing for each subject, but no imputation was performed to replace the missing values. The rationale for this was that the lack of a response was caused by uncertainty (i.e. the parent is unsure if their child performs a certain behaviour) and therefore missing answers contained useful information. Each child had their diagnosis (i.e. ASD or neurotypical) confirmed by clinical observation independent of the decision provided by PAAS.

Prediction models

A learning pipeline that included data resampling, a variety of classification models, and feature ranking was applied to train a model and evaluate its performance and overall suitability.

Imbalanced data The distribution of the two classes is approximately 20% (control) to 80% (ASD). This class imbalance can have an impact on many classification algorithms, typically by introducing a performance bias in favour of the majority class[43]. For example, if a classification algorithm classified all samples as ASD by default then it will have an accuracy of approximately 80%. Class imbalance can therefore affect the ability of classification algorithms to generalise well to unseen data and care must be taken when evaluating the performance of the classifier (e.g. by using evaluation metrics that are insensitive to class imbalance). Resampling the data is a popular method of mitigating class imbalance[43]. We applied the Synthetic Minority Oversampling Technique[44] (SMOTE) in Weka to the data (see bottom section of Table 4, random seed $S = 42$). SMOTE operates by oversampling the minority class to generate interpolated synthetic instances, thus reducing the data imbalance. The complete dataset including synthetic samples was input to the classification algorithms.

Classification algorithms A variety of popular data mining techniques[45] were evaluated to test their suitability for the ASD prediction task. The evaluated algorithms include adaptive boosting, decision trees, neural networks, naïve bayes, rule based models, and random forests. A description of the applied algorithms is provided in Table 2. The models were implemented using the `Weka v3.8.2` data mining software[46]. Performance metrics were generated using 10 iterations of 10-fold cross validation to understand the generalisation ability of each model. The generated performance metrics of each model were compared using a paired t-test, and the optimal model was determined from the results of the t-tests.

Feature selection Correlation based feature selection (CFS[47]) and minimum redundancy maximum relevance (mRMR[48]) were applied to the dataset with 10-fold cross validation, to verify the usefulness of the features. The feature subsets identified were not used for data collection or during the training process, but is useful for optimising clinical applications of PAAS (mobile application or paper-based) in the future. Removing questions with no predictive value can speed up the screening process, which would be useful during a home visit by a non-specialist healthcare worker.

Model evaluation

The predictive performance of models was evaluated using the true positive rate (TPR; also known as sensitivity), false positive rate (FPR), true negative rate (TNR; also known as specificity), false negative rate (FNR), and Area Under the Receiver Operating Characteristic (AUROC). By evaluating a range of metrics any bias introduced by class imbalance can be mitigated, and a more thorough understanding of the model's performance can be gained.

Predictive performance

Models were trained using the yes/no responses to 21 PAAS items and additionally the child's age and gender. Although PAAS does not use demographic information to predict ASD the rationale for including both features as predictors lies in the biological pathophysiology of ASD. Firstly, as ASD is a developmental disorder, age could be useful in order to determine if behaviours are absent because of atypical development or because of immaturity. For example, an 18 month old child that is missing a key behaviour that is predictive for autism may begin to spontaneously demonstrate that behaviour at 2 years of age. This is less likely to occur in a 5 year old, and thus the behaviour could be a better predictor for an older child. This is particularly true for questions that test the ability of the child to perform social interactions and communicate (and less important for questions regarding repetitive behaviour). Understanding if this is the case could help improve PAAS and other symptom checklist based screening tools in the future. Secondly, there is a variety of sex differences in ASD. Approximately four times more males than females are diagnosed with ASD on average[49]. Additionally, behaviours could be inconsistent across gender as children's brains develop differently depending on gender (e.g. females typically have better social mimicry skills which is thought to contribute to female underdiagnosis of ASD[50]).

The performance of the models was assessed using the `Weka Experimenter` environment. Only post-SMOTE data was input to the Weka experimenter environment. 10 iterations of 10-fold cross validation were used to generate sufficient performance metrics for valid statistical comparison (i.e. each model was trained and tested 100 times). All significance testing was conducted using a corrected paired t-test with a significance level of 0.05. The Naive Bayes (0.95) and Adaboost (0.92) algorithms had the highest TPR, while the remaining algorithms were not significantly different (see Figure 2). The Random Forest

Table 2. Description of applied classification algorithms.

Classification algorithm	Description	Parameters
Adaptive Boosting (AdaBoostM1)	Adaptive Boosting (AdaBoost) is an ensemble learning algorithm that combines the output of multiple weak learners is combined via weighted majority voting.. Adaptation occurs during training as weights are updated to focus on misclassified instances.	-P 100 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump
Decision Tree (J48)	A C4.5 Decision Tree (DT) uses a top-down divide-and-conquer strategy. C4.5 improves on Quinlan's ID3 algorithm by accepting continuous values, missing data, and introducing pruning to mitigate overfitting.	-C 0.25 -M 2
Neural Network (Multilayer Perceptron)	Neural Networks are biologically inspired collections of interconnected artificial neurons. The architecture of the network is defined by the designer of the neural network. The multiplayer Perceptron used here is a fully connected feedforward neural network that uses backpropagation to transform a set of inputs into a defined output to classify instances.	number of neurons in hidden layer $= (\frac{\text{no. attributes}+\text{no. classes})}{2}$; -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a
Naïve Bayes	Naive Bayes is a probabilistic classifier based on Bayes' theorem and has strong (naive) independence assumptions between features.	
Rule based model (PART)	Generates a set of rules from partial C4.5 decision trees.	-M 2, -C 0.25, -Q 1
Random Forest	An ensemble of decorrelated decision trees: the mean or mode is used to combine the output of DTs into a single decision. Random feature subsets are used during training.	-P 100 -I 100 –num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Sequential Minimal Optimisation (SMO) Support Vector Machine (SVM)	An SVM classifier trained with Platt's sequential minimal optimisation algorithm [16]. In classification an SVM creates a hyperplane separating the two classes. During training the SVM aims to identify the maximum margin hyperplane between the two classes,	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4

Table 3. PAAS screening questions and feature selection results.

Attribute	Description	Core domain	CFS no. of folds (%)	mRMR no. of folds (%)
age	Child age in years	NA	20%	20%
sex	Child gender	NA	30%	30%
P1	Does your child enjoy being thrown up and down on your lap?	Social interaction	0%	0%
P2	Does your child join in a play of another child?	Social interaction	0%	0%
P3	Does your child enjoy playing hide and seek?	Social interaction	10%	0%
P4	Does your child attempt to imitate your actions?	Communication	100%	100%
P5	Does your child show pretend play?	Social interaction	30%	90%
P6	Does your child point to request?	Communication	40%	100%
P7	Does your child point and show something that interests him?	Communication	80%	100%
P8	Does your child play with toys appropriately rather than mostly mouth or break them?	Social interaction	100%	100%
P9	Does your child bring over things to show you?	Social interaction	10%	60%
P10	If you point at something far away, does your child look in that direction?	Communication	40%	90%
P11	Does your child close his ears as he dislikes the sounds around him?	Communication	10%	10%
P12	If you point at something across the room, does your child look t it?	Communication	100%	100%
P13	If you point at something nearby, does your child look in that direction?	Communication	20%	10%
P14	Does your child show repetitive purposeless finger movements?	Repetitive behaviour	0%	0%
P15	Does your child watch rotating objects such as a fan or wheels for long periods?	Repetitive behaviour	80%	90%
P16	Does your child reciprocate affectionate gestures from you?	Social interaction	0%	0%
P17	Does your child imitate your facial gestures?	Communication	100%	100%
P18	Does your child respond when called by name?	Communication	100%	100%
P19	Does your child look at your face when you hold an object in front of you?	Social interaction	80%	100%
P20	Does your child show willingness to share toys with others?	Social interaction	100%	100%
P21	Does your child often appear as if he is in his own world?	Social interaction	100%	100%

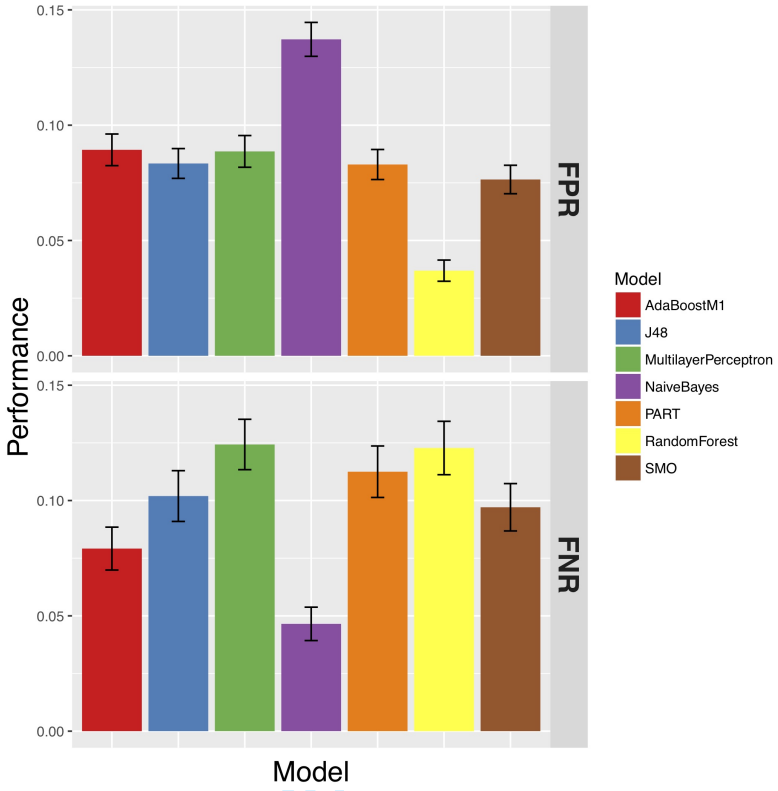

Figure 1. False negative rate (FNR) and false positive rate (FPR) of the models.

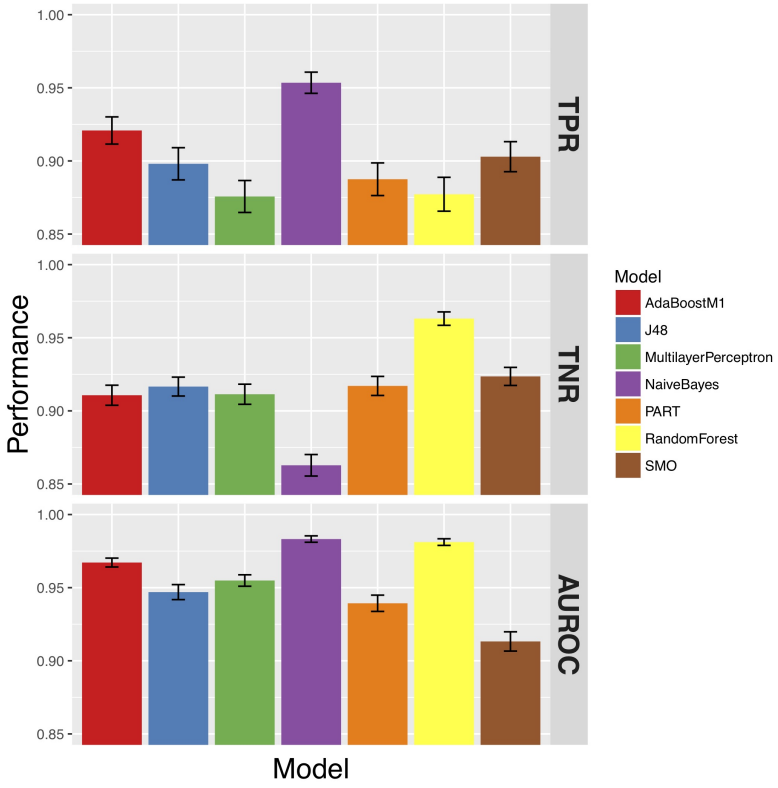

Figure 2. True positive rate (TPR), True negative rate (TNR), and Area Under the Receiver Operator Curve (AUROC) of the models

Table 4. Dataset demographics.

	Variable	Control	ASD
Before SMOTE	n	33	195
	Median age (years)	2.30	2.60
	Sex		
	Male	28	156
	Female	5	39
After SMOTE	n	86	195
	Median age (years)	2.6	3
	Sex		
	Male	76	156
	Female	10	39

Table 5. Dataset sample, n features excluding class: 23.

Age	Sex	Question 1	...	Question 21	Class
2.6	male	no	...	no	control
...	...	...	...	...	...
4	female	yes		yes	autism

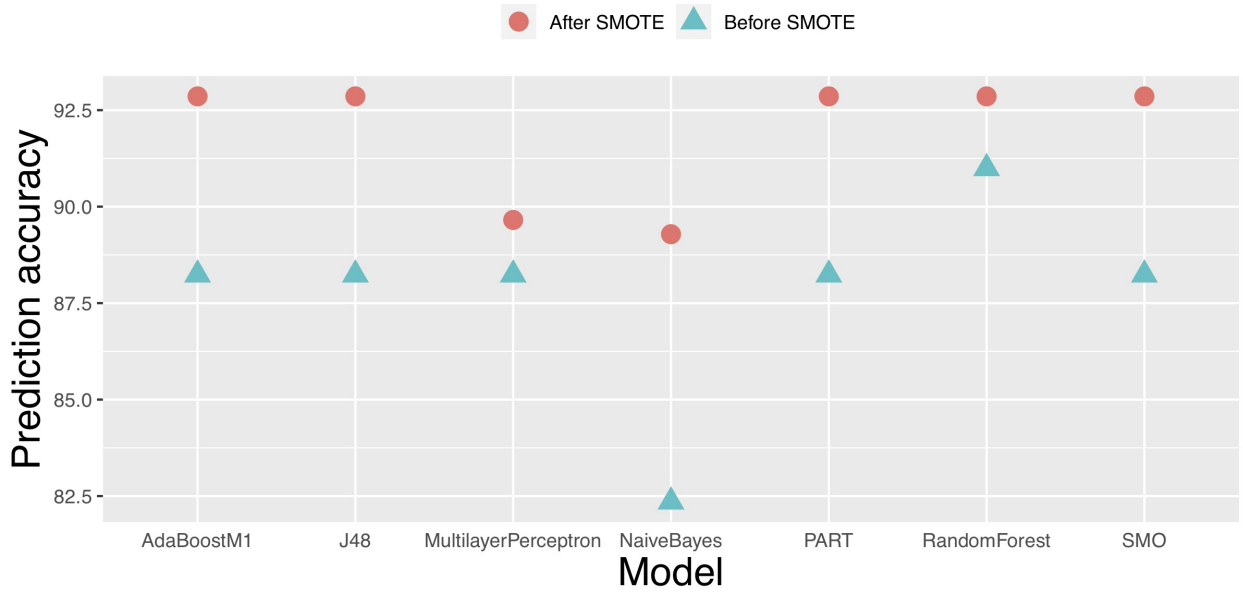

Figure 3. Effect of SMOTE on model classification accuracy.

(0.96) had a significantly better TNR compared with the remaining algorithms. The Random Forest (0.98) and Naive Bayes (0.98) algorithms had the best AUROC. The Random Forest (0.03) had a significantly better FPR compared with all other algorithms, and the Naive Bayes (0.05) had the best FNR (see Figure 1).

Due to the class imbalance in the dataset it is sensible to place greater importance on the AUROC when evaluating the models, which is not sensitive to class imbalance. From this evaluation it is clear that the Random Forest had the best AUROC, FPR, and TNR. The TPR and FNR were similar to all other algorithms. From this, the Random Forest was chosen to be embedded into the final mobile screening application.

The application of SMOTE improved the predictive performance for all of the models except AdaBoostM1 and ZeroR, and the Random Forest and Multilayer Perceptron improved the most (see Figure 3).

The predictive ability of the Random Forest exceeded the performance of the paper based PAAS system. The standard PAAS checklist has reported a TPR of 88% and a TNR of 93%[8]. The Random Forest model reported a TPR of 96% and a TNR of 97%. One of the limitations of PAAS was that it used an arbitrary cutoff score rather than a standardised one. The improved performance shows that the Random Forest has learned to classify from the data better than the arbitrary cut-off designed by the developers of PAAS.

As smart devices provide an ideal platform for mental health disorder assessment, the Random Forest model is integrated into an Android tablet-based application to assist non-specialist healthcare workers to screen for ASD. A sample result from the developed tablet-based autism screening tool is shown in Figure 4. To conduct the tablet-based autism screening, a user needs to respond to 21 PAAS items and additionally the child's age and gender. After completing this information, the user will be redirected to the Review screen to confirm all the questions and answers are provided correctly. Once the user clicks on the submit button on the Review screen (see Figure 4-left), they will be automatically redirected to the Screening result (see Figure 4-right). The Screening result screen provides suggests whether a further screening needed with a healthcare specialist or not.

Selected features

The selected features provide insight into what questions are most important for predicting ASD in a paediatric cohort (see Table 3). Questions were separated into four broad areas according to their core domain: communication, social interaction, repetitive behaviour, and demographic (demographic questions are not part of the standard PAAS questionnaire but are recorded separately). Only around half of the questions across all core domains were retained consistently over multiple folds. This suggests that PAAS could be significantly reduced in size while maintaining predictive power, saving time and resources. Interestingly, although age and gender were hypothesised to be important for predicting ASD, they were not present in the majority of folds. As ASD is a developmental disorder it was thought that age could be useful to determine if absent behaviours were absent because of physical immaturity or developmental delay caused by ASD. Additionally, many aspects of ASD and normal behaviours are different across genders. This could be because the behaviours described in PAAS are identical across genders or other undetected cultural factors.

Conclusion and future work

Due to cultural reasons, ASD awareness is low in LMIC. Resource constraints meant that when ASD is identified, patients are often left untreated for long periods of time. Early identification and diagnosis is important to improve clinical outcomes of young children with ASD. Smart devices represent an ideal platform for a computer aided tool, as they are highly accessible and prevalent across the world. Most existing screening tools automate standard screening checklists such as M-CHAT-R. Only the AutismAI application embeds an intelligent machine learning model to arrive at a decision. However, AutismAI does not incorporate a clinically validated and culturally specific symptom checklist. This article proposed a novel mobile application for ASD screening using a culturally sensitive symptom checklist and embedded machine learning model. A variety of supervised learning models were trained on PAAS data collected clinically and the best performing model — the Random Forest — was chosen to be embedded in the tablet-based mobile application. The proposed application has shown greater predictive performance than current paper-based methods (PAAS). The new application is important to improve ASD awareness and detection, by enabling non-specialist healthcare workers to screen for ASD during home visits. Furthermore, valuable data can be collected about the prevalence of ASD in LMIC (which is currently scarce) and resources allocated correctly to decrease treatment delays.

The predictive performance was analysed using multiple metrics to ensure a thorough evaluation of predictive power was conducted, including TPR, TNR, FPR, FNR, AUROC, and accuracy. We found the Random Forest model performed better (TPR: 88%, TNR: 96%) than the standard paper-based approach (TPR 88%, TNR: 93%): the model has learned from the data how to classify ASD, whereas PAAS uses an arbitrary scale to make this decision. In addition, feature selection revealed the most important questions were related to the communication and social interaction core domain, and approximately a quarter of the questions included by PAAS were irrelevant for ASD prediction. Therefore the efficiency of PAAS could be improved in the future by removing these irrelevant questions, which is valuable in environments such as busy outpatient clinics.

A limitation of this work is the quantity of data available, particularly for the control subjects. The decision of PAAS for each child must be independently confirmed by an experienced mental health specialist in a clinical setting. Therefore confirming that a child does not have ASD with a mental health specialist is difficult due to limited resources, particularly when many other children with other developmental disorders are in need of referral. In future work a new (class-balanced) cohort will be recruited to validate the predictive model on independent data. Additionally, the decision made by the Random Forest algorithm cannot be examined and interpreted by domain

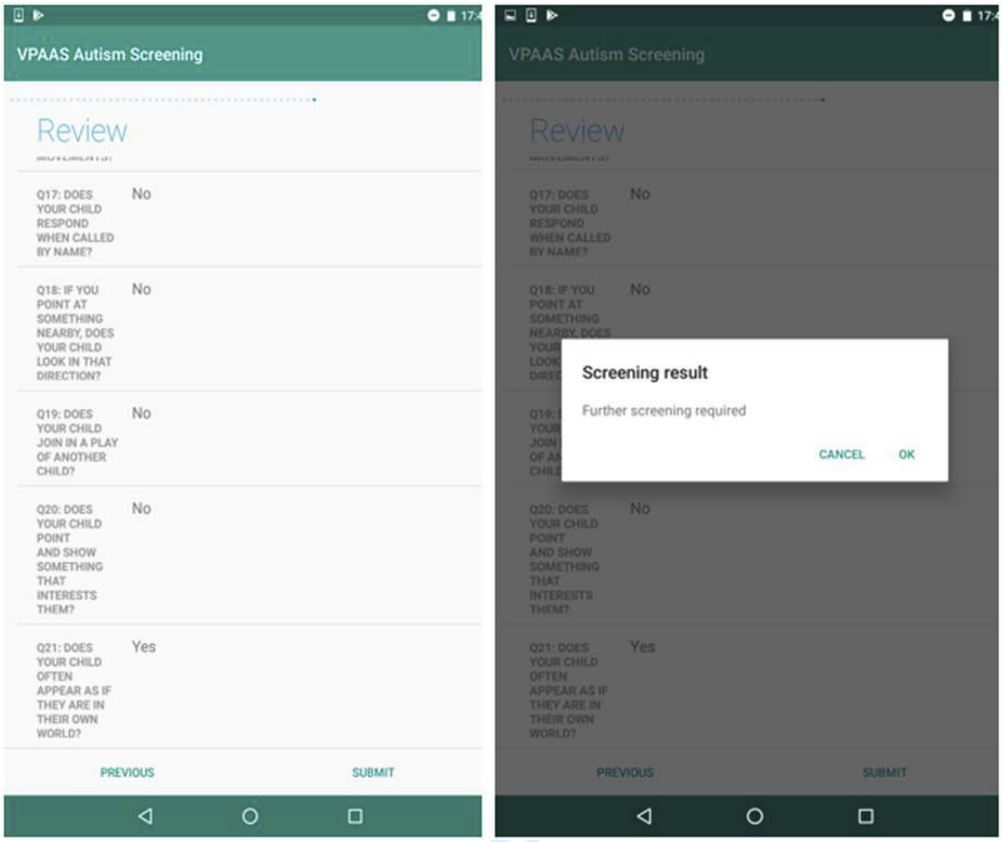

Figure 4. A sample of screen-shots from the tablet-based autism screening tool.

experts. In future work it may prove to be advantageous to use a fully transparent (e.g. rule-based) supervised learning algorithm.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research outlined here was supported by Department of Economy under the Global Challenge Research Fund grants. The funding sources had no role in the design, analysis, or interpretation of data or in the preparation of the report or decision to publish.

References

1. Taylor B, Jick H and MacLaughlin D. Prevalence and incidence rates of autism in the uk: time trend from 2004–2010 in children aged 8 years. *BMJ open* 2013; 3(10): e003219.
2. Elsabbagh M, Divan G, Koh YJ et al. Global prevalence of autism and other pervasive developmental disorders. *Autism Research* 2012; 5(3): 160–179.
3. Elsabbagh M. Perspectives from the common ground. *Autism Research* 2012; 5(3): 153–155.
4. Zwaigenbaum L, Bauman ML, Stone WL et al. Early identification of autism spectrum disorder: recommendations for practice and research. *Pediatrics* 2015; 136(Supplement 1): S10–S40.
5. Rahman A, Divan G, Hamdani SU et al. Effectiveness of the parent-mediated intervention for children with autism spectrum disorder in south asia in india and pakistan (pass): a randomised controlled trial. *The Lancet Psychiatry* 2016; 3(2): 128–136.
6. Robins DL, Casagrande K, Barton M et al. Validation of the modified checklist for autism in toddlers, revised with follow-up (m-chat-r/f). *Pediatrics* 2014; 133(1): 37–45.
7. Perera H, Wijewardena K and Aluthwelage R. Screening of 18–24-month-old children for autism in a semi-urban community in sri lanka. *Journal of tropical pediatrics* 2009; 55(6): 402–405.
8. Perera H, Jeewandara KC, Seneviratne S et al. Culturally adapted pictorial screening tool for autism spectrum disorder: A new approach. *World journal of clinical pediatrics* 2017; 6(1): 45.
9. Thabtah F, Kamalov F and Rajab K. A new computational intelligence approach to detect autistic features for autism screening. *International journal of medical informatics* 2018; 117: 112–124.
10. Shahamiri R and Thabtah F. AutismAI, 2019. URL https://play.google.com/store/apps/details?id=com.rezanet.intelligentasdscreener&hl=en_GB. Accessed 2019–07–23.
11. Bardhan S, Mridha GMM, Ahmed E et al. Autism barta—a smart device based automated autism screening tool for bangladesh. In *Informatics, Electronics and Vision (ICIEV), 2016 5th International Conference on*. IEEE, pp. 602–607.
12. Apple and University D. Autism & beyond, 2018. URL https://itunes.apple.com/us/app/autism-beyond/id1025327516?mt=8. Accessed 2019-03–27.

13. Nazneen N, Rozga A, Smith CJ et al. A novel system for supporting autism diagnosis using home videos: iterative development and evaluation of system design. *JMIR mHealth and uHealth* 2015; 3(2).
14. Cho KW, Lin F, Song C et al. Gaze-wasserstein: a quantitative screening approach to autism spectrum disorders. In *Wireless Health*. pp. 14–21.
15. Potter G. Smartphone app for early autism detection being developed by ub undergrad, 2016. URL http://www.buffalo.edu/ubnow/stories/2016/11/autism-app.html. Accessed 2019–03–27.
16. Platt J. Sequential minimal optimization: A fast algorithm for training support vector machines. Technical Report MSR-TR-98-14, 1998. URL https://www.microsoft.com/en-us/research/publication/sequential-minimal-optimization-a-fast-algorithm-for-training-support-vector-machines/
17. Hyde KK, Novack MN, LaHaye N et al. Applications of supervised machine learning in autism spectrum disorder research: a review. *Review Journal of Autism and Developmental Disorders* 2019; 6(2): 128–146.
18. Lord C, Rutter M and Le Couteur A. Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of autism and developmental disorders* 1994; 24(5): 659–685.
19. Lord C, Risi S, Lambrecht L et al. The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders* 2000; 30(3): 205–223.
20. Bone D, Bishop SL, Black MP et al. Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. *Journal of Child Psychology and Psychiatry* 2016; 57(8): 927–937.
21. Wall DP, Dally R, Luyster R et al. Use of artificial intelligence to shorten the behavioral diagnosis of autism. *PloS one* 2012; 7(8): e43855.
22. Wall DP, Kosmicki J, Deluca T et al. Use of machine learning to shorten observation-based screening and diagnosis of autism. *Translational psychiatry* 2012; 2(4): e100.
23. Maenner MJ, Yeargin-Allsopp M, Braun KVN et al. Development of a machine learning algorithm for the surveillance of autism spectrum disorder. *PloS one* 2016; 11(12): e0168224.
24. Florio T, Einfeld S, Tonge B et al. Providing an independent second opinion for the diagnosis of autism using artificial intelligence over the internet. *Couns, Psycho Health Use Technol Mental Health* 2009; 5: 232–248.
25. Duda M, Ma R, Haber N et al. Use of machine learning for behavioral distinction of autism and adhd. *Translational psychiatry* 2016; 6(2): e732.
26. Plitt M, Barnes KA and Martin A. Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *NeuroImage: Clinical* 2015; 7: 359–366.
27. Liu W, Li M and Yi L. Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. *Autism Research* 2016; 9(8): 888–898.
28. Jiao Y, Chen R, Ke X et al. Single nucleotide polymorphisms predict symptom severity of autism spectrum disorder. *Journal of autism and developmental disorders* 2012; 42(6): 971–983.
29. Abbas H, Garberson F, Glover E et al. Machine learning approach for early detection of autism by combining questionnaire and home video screening. *Journal of the American Medical Informatics Association* 2018; 25(8): 1000–1007.
30. Thabtah F and Peebles D. A new machine learning model based on induction of rules for autism detection. *Health informatics journal* 2019; : 1460458218824711.
31. Abdeljaber F. *Detecting Autistic Traits using Computational Intelligence & Machine Learning Techniques*. Master's Thesis, University of Huddersfield, 2019.
32. Vellido A, Martín-Guerrero JD and Lisboa PJ. Making machine learning models interpretable. In *ESANN*, volume 12. Citeseer, pp. 163–172.
33. van den Bekerom B. Using machine learning for detection of autism spectrum disorder. In *Proc. 20th Student Conf. IT*. pp. 1–7.
34. Umanandhini D and Kalpana G. Enhance prediction of autism spectrum disorder using adaptive bayesian classifier. *ARPN Journal of Engineering and Applied Sciences* 2019; 14(10): 1997–2002.
35. Polikar R. Ensemble based systems in decision making. *IEEE Circuits and systems magazine* 2006; 6(3): 21–45.
36. Al Diabat M and Al-Shanableh N. Ensemble learning model for screening autism in children. *International Journal of Computer Science & Information Technology* 2019; 11(2): 45–62.
37. Thabtah F, Abdelhamid N and Peebles D. A machine learning autism classification based on logistic regression analysis. *Health information science and systems* 2019; 7(1): 12.
38. Guimarães AJ, Araujo VJS, Araujo VS et al. A hybrid model based on fuzzy rules to act on the diagnosed of autism in adults. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, pp. 401–412.
39. Torres A and Nieto JJ. Fuzzy logic in medicine and bioinformatics. *BioMed Research International* 2006; 2006.
40. Gravenhorst F, Muaremi A, Bardram J et al. Mobile phones as medical devices in mental disorder treatment: an overview. *Personal and Ubiquitous Computing* 2015; 19(2): 335–353.
41. Statista. Number of smartphone users worldwide from 2014 to 2020 (in billions), 2018. URL https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/. Accessed 2019–03–27.
42. Association AP et al. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.
43. He H and Garcia EA. Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering* 2009; 21(9): 1263–1284. DOI:10.1109/TKDE.2008.239.
44. Chawla NV, Bowyer KW, Hall LO et al. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research* 2002; 16: 321–357.
45. Wu X, Kumar V, Quinlan JR et al. Top 10 algorithms in data mining. *Knowledge and information systems* 2008; 14(1): 1–37.
46. Hall M, Frank E, Holmes G et al. The weka data mining software: an update. *ACM SIGKDD explorations newsletter* 2009; 11(1): 10–18.
47. Hall MA. *Correlation-based feature selection for machine learning*. PhD Thesis, 1999.

48. Ding C and Peng H. Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology* 2005; 3(02): 185–205.
49. Fombonne E. Epidemiology of pervasive developmental disorders. *Pediatric research* 2009; 65(6): 591.
50. Gould J and Ashton-Smith J. Missed diagnosis or misdiagnosis? girls and women on the autism spectrum. *Good Autism Practice (GAP)* 2011; 12(1): 34–41.