International Journal of Control and Automation
Vol. 13, No. 2s, (2020), pp. 202-216

Screening and classification of Autism Spectrum Disorder using Machine Learning

Marisa Ramesh Kumar [1] , Dr.Vamsidhar Enireddy [2], Dr.Kolla Bhanu Prakash [3]

[1]*Lecturer , Dept of CSE ,MVR Degree & PG College ,Vishakhapatanam,A.P, India.*

[2]*Associate Professor ,Dept of CSE, Koneru Lakshmaiah Education Foundation ,Guntur, A.P India.*

[3]*Professor , Dept of CSE, Koneru Lakshmaiah Education Foundation ,Guntur, A.P India..*

Abstract

The inability of the person to communicate with the others either verbally and non-verbally, also portraying certain difficulties in their social behavior or showing a repetitive behavior is a development disorder called as Autism Spectrum Disorder or commonly ASD. The assessment or screening of the ASD requires a countless things with a large number of reactions to be noted mostly conducted by clinical staff, guardian and some self assessment tools to identify it. Using the modern tools such as artificial intelligence, machine learning and computational intelligence a more potential methods can be designed to improve the efficiency and accuracy of the screening tools. Using these new technologies smaller platforms can be built to improve screening or having an option to help moving towards precise opinion. In this paper screening of disorder is built and also machine learning techniques are applied on the datasets related to the autism spectrum disorder. Autism Spectrum Disorder is a social behavioral disorder and the application of machine learning techniques helps in the better screening and diagnosis of the disease. This helps the medical practitioners, nursing staff or parents that deal with autism affected patients to better treat the patients, and also in the early detection of the disorder. The results obtained from the machine learning applications can be in turn used to develop more better and accurate screening tools. This helps in the ease of treatment.

Keywords: *Autism, artificial intelligence, computational intelligence, diagnosis, screening and machine learning*

1. Introduction

Autism Spectrum Disorder (ASD) is an unavoidable formative issue that thwarts a person's aptitudes in socialization, makes redundant practices, and effects expressive or verbal correspondence with disturbances running from moderate to extreme [1]. The side effects of mental imbalance are progressively obvious and simpler to distinguish in kids a few years old. Late evaluated commonness of ASD in India ranges from 0.15% to 1.01% in different examinations, contingent upon the screening strategy utilized, and the territories studied [4,5]. A study that was conducted by the INCLEN on the prevalence of ASD at that point named as PDD it was noted that for every 125 youngsters 1 person in the age of 3-6 years and for every 85 kids 1 person in the age of 6-9 years are affected. When it was looked at the prevalence in the various regions it was found that it is 0.6% in mountain areas, 0.90% in the country side or in rural areas, 0.61% in coastal areas, 1.01% in urban regions and with only 0.1% in the tribal region.

The study on the autism dates back to 1943, first characterized by the Kanner. Leo Kanner in their first research article where they observed such children since 1938. The authors mainly focused more on the psychological and behavioral aspects of the children as opposed to on the examinations of conceivable clinical issues fundamental formative relapse of these youngsters. [3]. The author had come up with the "refrigerator mother theory" accepting that mental imbalance has the neuropsychological reason[4] From that point forward, different screening techniques have been created by driving clinical specialists and therapists overall trying to recognize medically introverted characteristics in the crude stage in order to promptly give the important medication. Formal ASD determination is commonly led by specific doctors in a clinical put together condition with respect to detectable and quantifiable conduct indicators. Present autism screening strategies depend on ASD diagnosis techniques, so they generally set aside a long effort to direct because of the huge number of things that the client must experience while depending on static human implanted guidelines. This has required an adjustment in the manner diagnostics are coded and carry on inside ASD screening techniques during the time spent arranging cases. [5]. The screening tools that are used in the present time by and large utilize human created and centric rules to classify the cases and also the controls. The experts in the fields of psychiatric and behavioral sciences have structured these guidelines, and the nature of results and choices relies generously upon the emotional commitments of these experts and the translations of the particular clinical staff directing the appraisals. On the other hand, the determination of ASD may be enabled via mechanized choices produced by reliable and clever algorithms mostly that are used in machine learning. Machine learning can be defined as a technique which uses statistical methods to learns with the conveyance of information in order to settle on choices on new information. AI is a gathering of factual methods that learn with the conveyance of information in order to settle on choices on new information. It is utilized to devise complex applications to make exact classifications/predictions on different information. [6].

ML explore territory based around artificial intelligence, probability and statistics, databases, and other software engineering regions that intend to brilliantly find concealed information (models) from datasets [7]. Utilizing ML only here and there includes clients in the procedures of arrangement or model learning and for sure may support the characterization precision and productivity. All the more significantly, the models inferred utilizing ML won't supplant clinicians, rather these models offer the clinician direction to conceivably improve the referral choices of people experiencing ASD investigation. [5].

2. Literature Survey

Author in [5]proposed that the absence of incorporating innovation, for example, ML with existing screening strategies, may add to current constraints, for example, openness, absence of conduct datasets, unwavering quality of straightforward scoring capacity and hand-made guidelines, and subjectivity of a ultimate conclusion. The ML advancements to be inspected and produced for screening tools are proposed to make the classification procedure of ASD automated, as opposed to static. These progressions may viably supplant prior human-created rules and strategies bringing about unmistakable and effective

points of interest: upgraded proficiency (time taken to play out the screening) with ASD order, decrease in the quantity of highlights of ASD appraisals to insignificant levels while keeping up evaluation respectability (distinguishing proof of key parts that produce exact findings), expanding availability particularly when utilizing a versatile domain, giving new datasets that can be used for additional examination, and all the more significantly improving the arrangement precision, affectability and explicitness. The author in [8] composed the article with the large gaps found in upward of a year between beginning doubt and analysis, important time where medicines and conduct intercessions could be applied is lost as these disarranges stay undetected. Techniques to rapidly and precisely evaluate hazard for these, and other, formative issue are important to streamline the procedure of conclusion and give families access to much-required treatments sooner. In [9] authors showed that as of late, the Autism Brain Imaging Data Exchange (ABIDE) has assembled information from more than 1,000 subjects and made it accessible for the ASD examine network. Tolerate dataset has been exceptionally adding to the improvement of different AI procedures in the ongoing occasions. Different analytic devices and research works have been created with an assistance from the ABIDE dataset. A genuine model for an diagnostic device that has been created is 'Aquabot' [10]. Aquabot is a Diagnostic Chatbot for Achluophobia and Autism. Aquabot explicitly center around choosing whether the client has the dread of obscurity or not – or is influenced by Achluophobia or not. Aquabot utilizes choice tree (DFS traversal) for its presentation improvement. The use of choice tree improves the reaction time and normality in reactions of aquabot. Aquabot likewise consolidates Natural Language Processing into it. It works by taking contribution from clients dependent on a progression of inquiries. Based on answers client gives it is chosen whether client is casualty of achluophobia or not.

Another device called 'FACE' tests the non-verbal connection in kids. In [11] authors had proposed in their paper that FACE is a social trustworthy antique ready to connect with the outer condition, deciphering and passing on feelings through on verbal correspondence. In the structure of a social treatment, face can go about as an interface between a patient and a prepared advisor in an extraordinarily prepared room. The procedure of clinical determination of Autism Spectrum Disorder can be frequently exceptionally protracted. Different screening strategies are proposed for chemical imbalance screening dependent on various parameters. [1], examined shortening the time connected with self-administrated ASD pre conclusion in clinical family facilities. Their point was to empower clinical consideration staff, including doctors, medical attendants, and other clinical staff, to use all things considered ten highlights/inquiries as a structure for brisk clinical referrals of potential ASD cases. The creators at that point broke down various forms of current self-directed or parent helped ASD screening instruments, which included : Quantitative Checklist for Autism Toddlers (Q-CHAT), Autism Spectrum Quotient (AQ) in three versions Child, Adolescent and adult

2.1 Motivation:

In United States alone it was found by Centers for Disease Control and Prevention that 1 in every 59 adolescents are having a issue of the Autism[12]. The girls are

less prone to this disorder compared to boys as 1 in 37 boys and 1 in 151 girls are having this autism spectrum disorder. It is also found that early detection of this disorder can help the child. If the disorder is found at the early age of two then the valuable time is available for the doctors to treat the disorder rather than checking at the later ages. It was learnt that in 31% of kids with this disorder on testing for the IQ levels these kids are having inability to read, write or understand the things, having IQ<70, where as 44% in them had shown IQ<85 and remaining in range of 71-85[]. Early detection helps a better life expectancy. There are about 50% of these suffering with the security issue due to the mental imbalance. It was found that 66% of the young children in the age group of 6-15 suffering with the disorder facing the harassments and 28% having self damaging practices which includes scratching, biting and head hammering. The drowning has become a major factor of death for these children [13-15].

The issue with the autism is that medicines are not available to cure the disorder but in detecting the disorder at early stage helps to solve the other issues related to the autism. Other problems that arise due to this effect is mainly the gastrointestinal problem that too in the youngsters [16].

So in this paper a method for screening the disease is implemented which would assist the doctors and specialist in early detection of the disease and also machine learning techniques are applied on the available datasets for classification of the disorder.

2.2 Methodology

2.2.1 Screening

This module focuses on the development of a screening tool for detection Autism Spectrum Disorder in Toddlers. The chatbot also takes an inspiration from the screening tool developed by Dr. Fadi Fayez which he states that is only for gathering data for research uses. The tool should not be mistaken as something equivalent to a medical tool. The application is only developed as an early detection tool. The results from the tool are not equivalent to a diagnosis from a medical examiner. The tool contributes by helping in an early detection of Autism Spectrum Disorder in toddlers. The result from the application might be considered as a slight alarm to the users that the toddler might be suffering from Autism Spectrum Disorder (ASD). The screening tool is implemented as a mobile application. The application works by asking the user a series of questions and a score is calculated based on the answers from the user. A screening questionnaire is developed as qauntative checklist for detecting the autism in toddlers, it consists of a check list and a series of questions that are put to the parents and the care takers to detect the autistic features in the toddlers and in these features they calculate the threshold and sub-threshold features to increase the efficiency in detecting the disorder. The application is designed in such a way that the questions being asked for the detection of ASD vary according to the age of the toddler. The application specially focuses on the detection of ASD in toddlers whose age ranges between 3 and 24 months. The questions selected in the application for detecting ASD are resulted from a research in various web resources that are related to the traits toddlers with ASD might exhibit. The diagnosis in the

application can be performed by a parent or a medical practitioner or nursing staff on behalf of the toddler. The application works by asking various questions about the behavior of the toddler. Hence a person that has a better idea about the toddler would be helpful for more accurate results.

The proposed architecture of the application is that a user agrees to take the test and gives details about the age of the toddler. Then, a series of questions will be asked about the traits the toddler may exhibit to which the user has to select yes or no as answers. The result will be displayed and the result is majorly classified into High, medium or low symptoms the toddler is exhibiting. Fig.1 shows the proposed architecture of the mobile application.

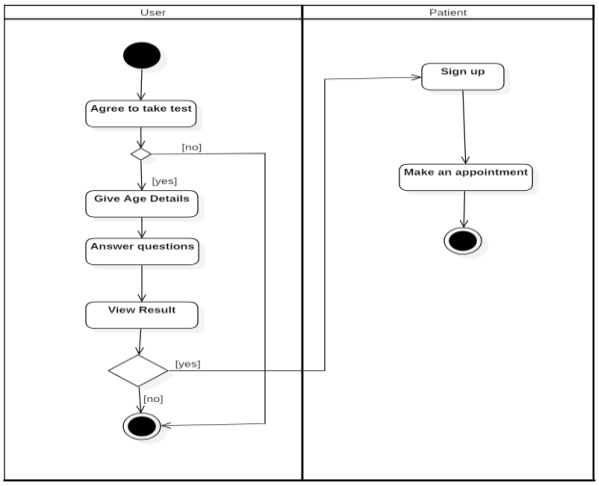


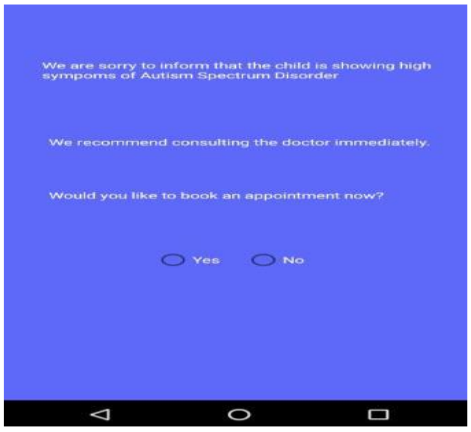
Fig.1 Proposed architecture of mobile application

The figure below shows different screens of the mobile application developed to screen the autism.

Fig 2. Login Screen

Fig 3.Readinessprompt Screen

Fig 4.Screen for age of Toddler



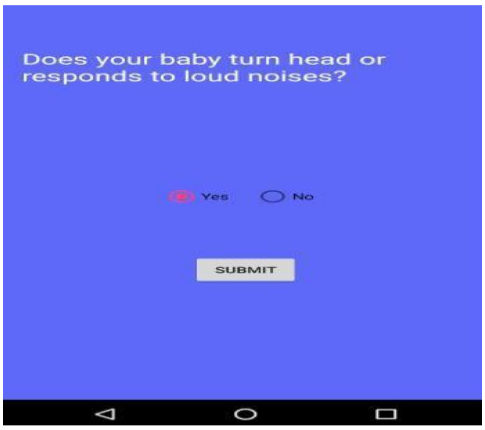


Fig 5: Screen showing an example of a question in the test

Fig 6 : Showing result of the test screen

3. Datasets

The datasets utilized in the task are an after effect of the ASD Tests Screening application. There are 3 variations of the dataset to be explicit. They are dataset for babies (Toddler Dataset), dataset for kids (Children Dataset), dataset for grown-ups (Adult Dataset). These datasets are created remembering about the information on the AQ-10 Toddler, AQ-10 Children and AQ-10 Adult scores. The screening of chemical imbalance instrument Q-CHAT is the most established model created by [13], it is a quantitative model with an agenda that is observed by clinical pros and the report produced by it is seen as in fortuitous event with a report that is put together by the guardians dependent on the translation of the kid's activity.

The most punctual rendition of Q-CHAT was utilized to identify mental imbalance in little children matured somewhere in the range of 18 and two years as it were. Other than the primary highlights that are identified with the screening of ASD, an objective class variable has been made with a Boolean incentive to decide if the individual experiencing the test has ASD or not. The class variable worth is appointed naturally dependent on the last score acquired by the individual taking the ASD test. The Fig 5 shows the route of the ASD Tests screening application. [5]

ISSN: 2005-4297 IJCA

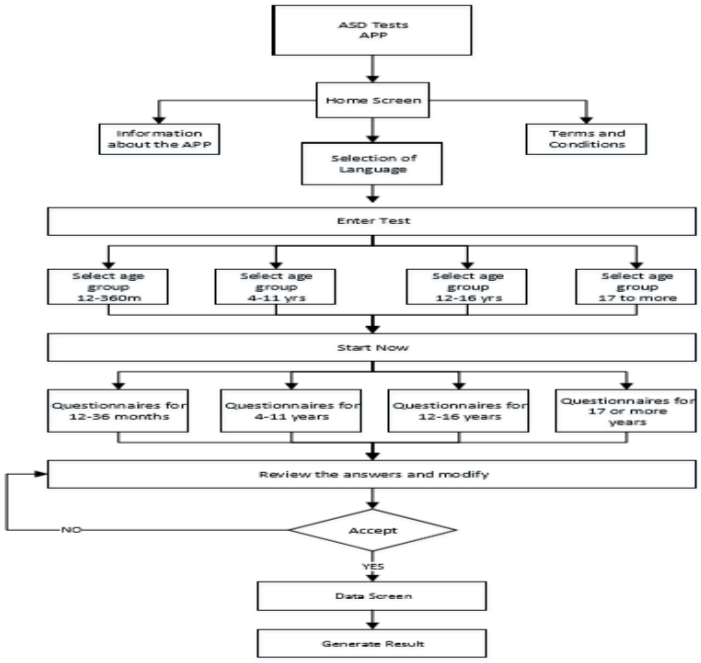


Fig.7. ASD Test screening application

.

No	Independent Variable	Data Type	Comments
1.	A1	Binary	Refer next table
2.	A2	Binary	Refer next table
3.	A3	Binary	Refer next table
4.	A4	Binary	Refer next table
5.	A5	Binary	Refer next table
6.	A6	Binary	Refer next table
7.	A7	Binary	Refer next table
8.	A8	Binary	Refer next table
9.	A9	Binary	Refer next table
10.	A10	Binary	Refer next table
11.	Age	Continuous	Age of participant
12.	Gender	Binary	Male or Female
13.	Ethnicity	Categorical Data	Chosen from a list of predefined values
14.	Jaundice	Binary	Yes or No
15.	Family History	Binary	Whether any family members diagnosed with autism
16.	User type	Categorical	Person taking the test (parent, self, relative, caregiver, etc)
17.	Language	Categorical	used for taken the test such as English, French, Spanish, Portuguese, Farsi, Arabic, Turkish, Russian, etc.
18.	Taken the test before	Binary	If the individual had taken the test before
19.	Score	Continuous	The score generated by the screening method function
20.	Test_type	Categorical	Based on the user age four types exist (toddler, child, adolescent, adult)
21.	Country_Residence	Categorical	The country of residence of the individual (drop list)
22.	why_are_you_taken_the_screening	Categorical	Users insert the reason(s) of why are they taking the screening
23.	Target class	Binary	0 (No ASD traits) or 1 (ASD traits)

Table 1. A gives the complete description of the attributes and their types present in the Datasets

The AQ-chat score is calculated as a result of all the above attributes. Amongst them, the attributes A1 to A10 are questions that are based on the traits the user suffering with autism spectrum disorder might exhibit.

Variable in Dataset	Corresponding AQ-10-Adult Features	Corresponding AQ-10-Adolescent Features	Corresponding AQ-10-Child Features
A1	I often notice small sounds when others do not	S/he notices patterns in things all the time	S/he often notices small sounds when others do not
A2	I usually concentrate more on the whole picture rather than the small details	S/he usually concentrates more on the whole picture rather than the small details	S/he usually concentrates more on the whole picture rather than the small details
A3	I find it easy to do more than one thing at once	In a social group, s/he can easily keep track of several different people's conversations	In a social group, s/he can easily keep track of several different people's conversations
A4	If there is an interruption, I can switch back to what I was doing very quickly	If there is an interruption, s/he can switch back to what s/he was doing very quickly	S/he finds it easy to go back and forth between different activities
A5	I find it easy to 'read between the lines' when someone is talking to me	S/he frequently finds that s/he doesn't know how to keep a conversation going	S/he doesn't know how to keep a conversation going with his/her peers
A6	I know how to tell if someone listening to me is getting bored	S/he is good at social chit-chat	S/he is good at social chit-chat
A7	When I'm reading a story I find it difficult to work out the characters' intentions	When s/he was younger, s/he used to enjoy playing games involving pretending with other children	When s/he is read a story, s/he finds it difficult to work out the character's intentions or feelings
A8	I like to collect information about categories of things (e.g. types of car, types of bird, types of train, types of plant, etc)	S/he finds it difficult to imagine what it would be like to be someone else	When s/he was in preschool, s/he used to enjoy playing pretending games with other children
A9	I find it easy to work out what someone is thinking or feeling just by looking at their face	S/he finds social situations easy	S/he finds it easy to work out what someone is thinking or feeling just by looking at their face
A10	I find it difficult to work out people's intentions	S/he finds it hard to make new friends	S/he finds it hard to make new friends

Table 2: Shows the list of questions A1 to A10 with respect to the users age

By far most of the occasions that have a place with little children have been related with a "no ASD" class name, making such a gathering of information totally imbalanced. To be precise, 96% of the cases who took part in the screening test for the Q-CHAT-10 (little children) have not been related with ASD, and in this manner the babies occurrences are isolated from different occasions just as precluded from further examination.

4. Observation From Dataset

The average age of users in years in the children, adolescent and adult datasets is 6.39, 14.04 and 30.14 respectively. More instances of male users have occurred than female instances. The number of male and female instances in the combined data of 1875 are 1076 and 799 respectively. From the combined dataset, it is observed that 883 instances have family history of ASD and 1563 instances have no family history of ASD.

The top ethnicities that participated in the screening are Asian, middle eastern, latino, Hispanic, aboriginal, black, white. A few instances of South-Asian ethnicity are also found in the children version of the dataset. For children, the screening test was performed majorly by parents which is indicated by 438 instances out of the 510 instances.

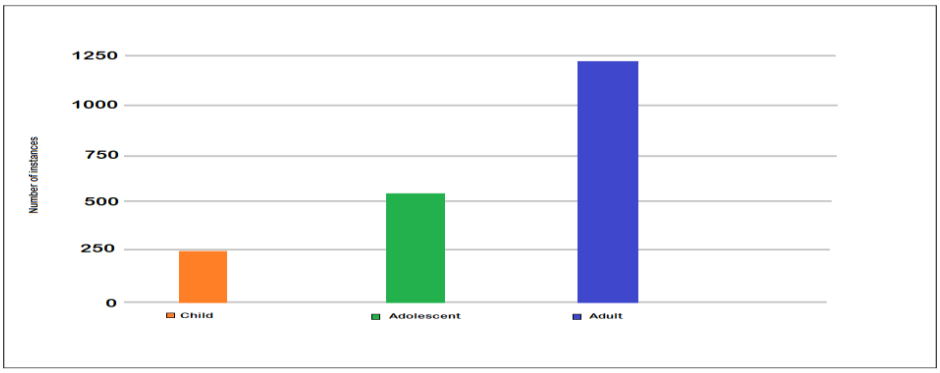


Fig 8 : shows the distribution of instances of the datasets with respect to the number of instances.

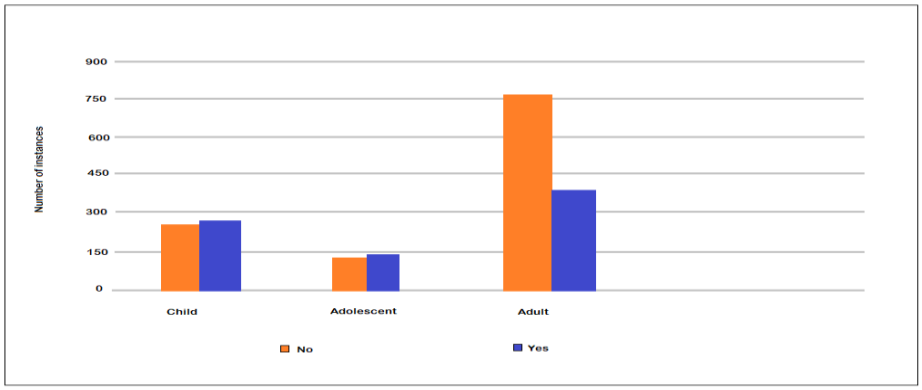


Fig 9 :Shows the class distribution with respect to the number of instances

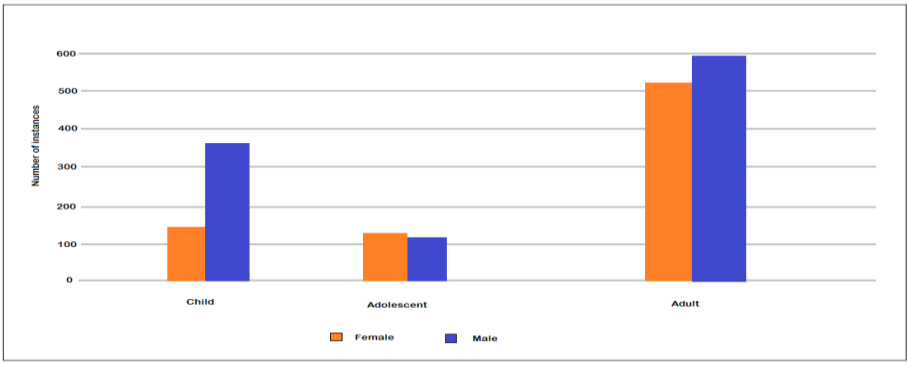


Fig 10 : Shows the relation between the number of the instances and the gender of the user.

5. Feature Selection

It is the procedure of naturally or physically choosing those highlights which contribute most to the forecast variable or yield of intrigue. The feature selection techniques help in the mission to make a precise prescient model. With the less number of attributes that are important and that gives a better accuracy and precision are chosen in this method. These methods remove the repetitive traits and unnecessary data that do not have any contribution in the model and it also

decreases the performance of the model. Using less number of input variables also decreases the complexity of the model and increases the scope of understanding the model. There are number of feature selection methods are available but Information gain and Correlation are chosen here for the feature selection. One of the most popular method for the feature selection is Information Gain, it chooses the attributes having the maximum information. The attributes with the lowest score are removed and attributes with the high score are selected. Information gain technique works by calculation information gain value i.e, entropy for each attribute.

The second method is Correlation. It can also be referred as Pearson's Correlation. This technique gives the most relevant attributes from the dataset. Correlation technique works by calculating the correlation between each attribute and the target or class attribute. From these calculated correlation values, attributes that have moderate to high-positive or negative (-1 or 1) correlation values are only selected and attributes that have correlation values closer to zero can be discarded. The feature selection technique applied in this project is an amalgamation of the above 2 described technique. Feature selection is applied to each of the dataset (children, adolescent, adult) using both the techniques separately. Each of the techniques yields a set of features. Then an intersection of the first n selected attributes is chosen as the result of feature selection. The observation of the amalgamation technique is that both feature selection techniques return almost the same attributes, with at most difference of 1 attribute when a set of first 4 attributes are considered.

The following tables show the result of each feature selection technique applied to each of the datasets (children, adolescent, adult)

Children Dataset		Description in the screening method
Score	Attribute	Description
1	Item 20	Score
0.26	Item 5	S/he finds it easy to go back and forth between different activities
0.168	Item 17	Residence
0.163	Item 7	S/he is good at social chit-chat

Table 3.Information Gain on Children Dataset

Adolescent Dataset		Description in the screening method
Score	Attribute	Description
0.9996	Item 19	Score
0.2948	Item 7	S/he is good at social chit-chat
0.2627	Item 17	Residence
0.2578	Item 4	In a social group, s/he can easily keep track of several different people's conversations

Table 4.Information Gain on Adolescent Dataset

Adolescent Dataset		Description in the screening method
Score	Attribute	Description
0.9996	Item 19	Score
0.2948	Item 7	S/he is good at social chit-chat
0.2627	Item 17	Residence
0.2578	Item 4	In a social group, s/he can easily keep track of several different people's conversations

Table.5 Information Gain on Adult Dataset

Adult Dataset		Description in the screening method
Score	Attribute	Description
0.9046	Item 19	Score
0.274	Item 7	know how to tell if someone listening to me is getting bored
0.2736	Item 6	I find it easy to 'read between the lines' when someone is talking to me
0.2695	Item 10	I find it easy to work out what someone is thinking or feeling just by looking at their face

Table. 6 Correlation on Children Dataset

Children Dataset		Description in the screening method
Score	Attribute	Description
0.83	Item 20	Score
0.581	Item 5	S/he finds it easy to go back and forth between different activities
0.458	Item 7	S/he is good at social chit-chat
0.451	Item 10	S/he finds it easy to work out what someone is thinking or feeling just by looking at their face

Table. 7. Correlation on Adolescent Dataset

Adolescent Dataset		Description in the screening method
Score	Attribute	Description
0.84	Item 19	Score
0.608	Item 7	S/he is good at social chit-chat
0.578	Item 4	In a social group, s/he can easily keep track of several different people's conversations
0.536	Item 5	If there is an interruption, s/he can switch back to what s/he was doing very quickly

Table. 8. Correlation on Adult Dataset

6.Algorithms

6.1 K-Nearest-Neighbours (KNN)

It is a non-parametric classification method. In order to classify a record, it retrieves the nearest neighbors of that record. It does not consider the distance based weighting process, but chooses one of the data record by majority voting. In any case, to apply KNN we have to pick a suitable incentive for k, and the achievement of order is particularly reliant on this worth. In any case, to apply KNN we have to pick a suitable incentive for k, and the achievement of order is particularly reliant on this worth. It could be said, the KNN strategy is one-sided by k. There are numerous methods for picking the k esteem, however a basic one is to run the calculation ordinarily with various k esteems and pick the one with the best execution. It is a case-based learning technique, which saves all the preparation information for grouping. KNN calculation is one of the least complex arrangement calculations and it is one of the most utilized learning calculations. KNN is a non-parametric, lethargic learning calculation. Its motivation is to utilize a database wherein the information focuses are isolated into a few classes to anticipate the grouping of another example point. It likewise requires just a couple of parameters to tune: separation metric and k.

	Accuracy	Precision	Recall	F1-Score
Children	98.1	96.9	98.1	97.8
Adolescent	93.7	97.8	88.5	89.6
Adult	97.2	98.2	95.9	97.8

Table 9 : Output of KNN Algorithm:

KNN algorithm gives an accuracy of 98.1% and a sensitivity of 96.9%, which shows that knn works pretty good for the given data set.

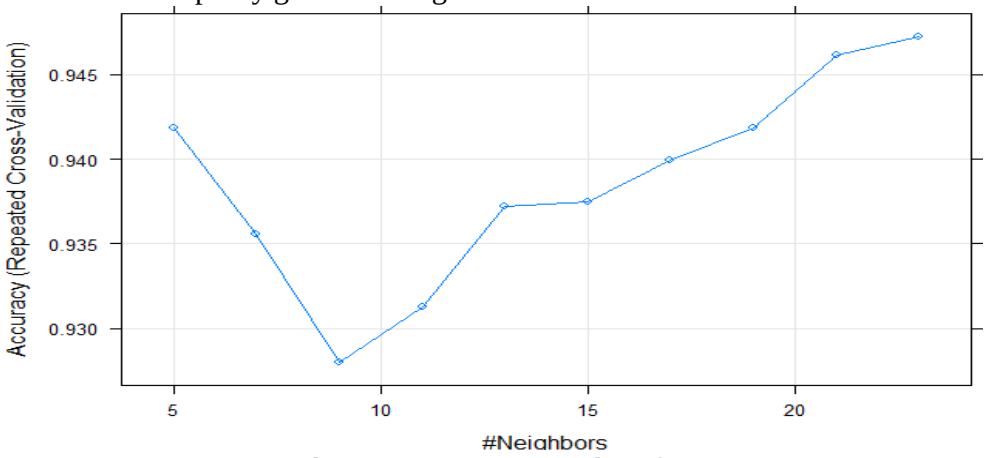


Fig 11: Accuracy Works of KNN

The plotted graph in Fig 5.5B between the accuracy and the K value shows that the highest accuracy of 98.1% is produced for a value of k around 23

7. Support Vector Machine

It is a discriminative classifier officially characterized by an isolating hyperplane. At the end of the day, given marked preparing information (administered learning), the calculation yields an ideal hyperplane which orders new models. In two-dimensional space this hyperplane is a line partitioning a plane in two

sections where in each class lay in either side. This algorithm can be linear or non-linear kernel. In this model a non-linear RBF Kernel is used for better results. It is a successful algorithm when applied on a high dimensional spaces. Still viable in situations where number of measurements is more prominent than the quantity of tests. Utilizations a subset of preparing focuses in the choice capacity , so it is likewise memory proficient. It is flexible i.e, diverse Kernel capacity can be determined for the choice capacity. Basic portions are given, however it is additionally conceivable to indicate custom kernels.

	Accuracy	Precision	Recall	F1-Score
Children	98.8	99.4	98.5	98.3
Adolescent	94.5	98.1	89.5	90.1
Adult	98.3	99.1	96.8	98.7

Table 10 : Output of SVM(Non-Linear) Algorithm

SVM(Non-Linear) algorithm gives an accuracy of 98.8% and a sensitivity of 96.1%, which shows that SVM(Non-Linear) algorithm also works pretty good for the given data set. The constant value of 0.02444254 is chosen as a sigma tuning parameter and C (cost) is equal to 8

7. Naïve Bayes

It uses Bayes theorem in this algorithm in order to classify the objects. It uses the independence, naïve or strong between the data points. Naïve Bayes is used in many applications such as medical diagnosis, spam filters etc., It is a simple classifier to learn and implement. This algorithm can make predictions in real time and can also be used in multi class prediction

	Accuracy	Precision	Recall	F1-Score
Children	98.4	98.6	97.5	97.3
Adolescent	92	97.1	87.5	87.1
Adult	97.3	98.2	95.2	97.8

Table 11 : Output of Naïve-Bayes Algorithm

The output of the Naïve bayes algorithm as shown in Table 5.5C gives maximum accuracy when applied to the children dataset amongst the 3 datasets.

8. Neural Network:

It is a mathematical model developed by the way the human brain process the data. As neurons are present in the brain that helps to process the information here also these neurons are arranged in different layers and interconnecting them with each other having weights allied through them. The neural networks are of many types Feed forward networks, Recurrent networks, Hopfield networks etc., and have many wide applications. Neural Networks can learn without anyone else and produce the yield that isn't restricted to the information gave to them. The information is put away in its own systems rather than a database, subsequently the loss of information doesn't influence it's working. These systems can gain from models and apply them when a comparable occasion emerges, making them ready to work even at real times.

Output of neural network:

This neural network can be converted to code for further implementation. The above neural network is built using the 'neural net' function in R studio.

	Accuracy	Precision	Recall	F1-Score
Children	50.8	51	100	67
Adolescent	48	48	100	65
Adult	67.4	67	100	81

Table 12: Output of Neural Network using Keras

As shown in table 5.5D, the neural network when applied to all the 3 datasets hives a high accuracy of 67.4% on the adult dataset

9. Conclusion

The main focus of the project is the application of machine learning techniques to the Autism Dataset. Machine Learning Techniques such as Feature Selection, Classification, Clustering and Neural Networks have been implemented. It can be observed from the above results that each algorithm gives a different a different accuracy and accuracies again vary depend upon the variant of the dataset i.e Children, Adolescent or Adult.

The highest accuracy of 98.8% is resulted in the application of Support Vector Machines using a non-linear RBF kernel.The accuracy of the Autism Spectrum Disorder screening tools primarily depends upon the extent of the knowledge the person that is performing the test has about the traits and behaviors of the toddler. In majority of the cases, parents may fail to identify the ASD traits the toddler might be exhibiting. The screening tool helps in such cases for detecting the traits and also as a primary alarm that the toddler might be suffering from Autism Spectrum Disorder.

The screening tool that is developed as apart of this project can be further enhanced to collect and store data from the users. This data can be made into datasets and can be made available online. Such datasets can act as cases for classification problems. They might also come in handy for research purposes for the domain experts. They can also help in improving the prediction, efficiency, accuracy of the screening tools of Autism Spectrum Disorder.

References

[1] M. D. C. a. L. S. Pennington, " Variability in state education agency definitions of and evaluations for Autsim Spectrum Disorders," 2014.
[2] M. J. a. S. Sairam1, ""Autism Spectrum Disorder - An Indian Perspective."," 2018.
[3] M. Cubała-Kucharska, "The review of the most frequently occuring medical disorders related to aetiology of autism and methods of treatment," 2010.
[4] E. L. Kanner L, ""Childhood schizophrenia;symposium, 1955. VI. Early infantile autism, 1943–1955," *American Journal of Orthopsychiatry*, pp. 556-566, 1956.

[5] Jayasuruthi L,Shalini A,**Vinoth Kumar V.**,(2018) " Application of rough set theory in data mining market analysis using rough sets data explorer" *Journal of Computational and Theoretical Nanoscience*, 15(6-7), pp. 2126-2130

[6] S. Velliangiri, P. Karthikeyan & V. Vinoth Kumar (2020) Detection of distributed denial of service attack in cloud computing using the optimization-based deep networks, *Journal of Experimental & Theoretical Artificial Intelligence*, DOI: 10.1080/0952813X.2020.1744196

[7] T. F. C. F. Qabajeh I., "A dynamic rule-induction method for classification in data mining," Journal of Management Analytics , pp. 233-253, 2015.

[8] M. M. R. H. N. a. W. D. P. Duda, "Use of machine learning for behavioral distinction of autism and ADHD," Transl. Psychiatry 6:e732. doi:10.1038/tp.2015.221, 2016.

[9] V.Vinoth Kumar, Ramamoorthy S (2017), "A Novel method of gateway selection to improve throughput performance in MANET", *Journal of Advanced Research in Dynamical and Control Systems*,9(Special Issue 16), pp. 420-432

[10] Katragadda, T., Srinivas, M., Prakash, K. B., & Pavan Kumar, T. (2020). Heart disease diagnosis using ANN, RNN and CNN. International Journal of Advanced Science and Technology, 29(4 Special Issue), 2232-2239. Retrieved from www.scopus.com

[11] Midhila, G., & Sheela Rani, C. M. (2020). A combined use of SIFT-CNN method for ship detection in optical satellite imagery. Test Engineering and Management, 83, 1488-1494. Retrieved from www.scopus.com

[12] M. I. M. S. M. F. R. J. F. M. A. A. D. D. R. G. Pioggial, "Human- Robot Interaction in Autism: FACE, an Android-based Social Therapy.," in 16th IEEE International Conference on Robot & Human Interactive Communication, 2007.

[13] Priya, S. S., Gladance, L. M., Valarmathi, A., & Joy, J. M. S. (2020). Brain tumor detection and disease prediction using CNN algorithm. Test Engineering and Management, 82(1-2), 13376-13381. Retrieved from www.scopus.com

[14] Karthikeyan, T., Sekaran, K., Ranjith, D., Vinoth kumar, V., Balajee, J.M. (2019) "Personalized Content Extraction and Text Classification Using Effective Web Scraping Techniques", International Journal of Web Portals (IJWP), 11(2), pp.41-52

[15] Shalini A, Jayasuruthi L, Vinoth Kumar V, "Voice Recognition Robot Control using Android Device" *Journal of Computational and Theoretical Nanoscience*, 15(6-7), pp. 2197-2201

[16] Kouser, R.R., Manikandan, T., Kumar, V.V (2018), "Heart disease prediction system using artificial neural network, radial basis function and case based reasoning" Journal of Computational and Theoretical Nanoscience, 15, pp. 2810-2817

[17] Dhilip Kumar V, Vinoth Kumar V, Kandar D (2018), "Data Transmission Between Dedicated Short-Range Communication and WiMAX for Efficient Vehicular Communication" Journal of Computational and Theoretical Nanoscience, Vol.15, No.8, pp.2649-2654

[18] Maithili, K , Vinothkumar, V, Latha, P (2018). "Analyzing the security mechanisms to prevent unauthorized access in cloud and network security" Journal of Computational and Theoretical Nanoscience, Vol.15, pp.2059-2063.