International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-7, Issue-6, March 2019

Predicting Autism Spectrum Disorder using Machine Learning Algorithms with Jaundice Symptomatic Analysis

Sharath Chandrika M B, Nallakaruppan M K, Siva Rama Krishnan S, Senthilkumar N C
VIT University, Vellore

***Abstract* We utilized a dataset identified with autism screening all age set of autism: toddler, child, adolescent, adult contained 20 attributes which are used for investigation particularly in deciding persuasive autistic traits, enhancing the order of ASD cases. With 10 social features in addition to 10 individual qualities that have ended up being successful in identifying the ASD cases, consequently applied RT to get the best clusters, process them through RF to get exactness. Primary objective of this work is to predict the correlation between the ASD with its symptoms by applying the machine learning techniques of the data science. The prescribed work is done to predict the correlation between the jaundice symptomatic patients, further progression of the same to ASD. This work also compares the chances of genetic influence which is the secondary classifier that leads to the disorder. To accomplish this objective, we applied our validated supervised Machine Learning, random tree, and random forest.**

Index Terms: Autism Spectrum Disorder (ASD), Machine Learning (ML), Random Tree (RT), Random Forest (RF), Correctly Classified Instances (CCI), Incorrectly Classified Instances (ICCI), Kappa Statistics(KS), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Relative Absolute Error (RAE), Root Relative Squared Error (RRSE), Total Number Of Instances (T. INSTANCES), Ignored Class Unknown Instances (ICUI), Aggressive Behaviour (AB), Autism Diagnostic Observation Schedule (ADOS), Autism Genetics Resource Exchange (AGRE), cross – sectional (CS).

I. INTRODUCTION

Autism Spectrum Disorder is a heterogeneous disorder at the behavioral ambit with children varying in level of cognitive skills, types of sensory, so on. The landscape of psychological, psychiatric study is logically interdisciplinary. ASD normally shows up before a child reaches the age of 3. ASD influences the advancement of a child's brain, bringing about debilitated communication, social working, connecting with others. The hazard components of ASD occur because of older parenting, exposure to valproate amid pregnancy, very low birth weight. ML is by all accounts a practical alternative for quickening these analytic endeavors by distinguishing the basic segments; deleting redundancy yet looking after exactness. One clear utility of ML in autism study is to make a productive, strong diagnostic algorithm dependent on human coded practices. The aims of the current study are: (1) Identifying the major cause by considering their age, gender, jaundice, genetic, (2) Also to identify whether boys with ASD are more than girls with ASD.

II. LITERATURE REVIEW

Reactive aggression among children with, without ASD in 2012, by Miia Kaartinen et al [3] proposed that each person is brought into the world with certain AB. Usually, the recurrence of AB crests peaks the initial 3 years [11]. The members were 27 boys, 8 girls with ASD, their controls coordinated with sex, age, last score insight. It appeared not to be a distinction in the recurrence of AB among boys, girls with ASD in new-born child, toddlerhood. This examination was to find out whether girls, boys with ASD , absolute IQ more than 70 indicated more exceptional receptive hostility than girls, boys without ASD , whether girls , boys with ASD neglected to utilize the situational signals (age , gender) of an aggressor as inhibitory prompts to weaken their responses to assaults in the Pulkkinen animosity machine (PAM). The purpose of this examination was to explore whether girls, boys with ASD demonstrated more extraordinary responsive enmity than girls, boys without ASD in the PAM that moves the strength of proactive threatening vibe, the physical quality, gender of the aggressor.In 2015, Applying ML to facilitate Autism Diagnostics by Daniel Bone et al[5] proposes ADOS comprises of four distant modules that shift contingent upon a person's age, verbal capacities. et al. (2012a) consider under examination. Module 1 joins 29 social conduct codes, 10 sub - tasks. The training data utilized in Wall et al. contained basically of ADOS relationship from the AGRE database. Their endeavour at replication with the AGRE information created extremely distant outcomes from the proposed 8 codes just 5 of the 9 distinguished codes in our replication cover, don't sum up crosswise over datasets as far as to order execution. The results suggest the values in Wall et al.

Revised Manuscript Received on 30 March 2019.
* Correspondence Author
Sarath Chandrika M B*, M.S.(Master of Software), VIT University, Vellore, India.
Nallakaruppan M K, School Of Information Technology and Engineering, VIT University, Vellore, India.
Siva Rama Krishnan S, School Of Information Technology and Engineering, VIT University, Vellore, India.
Senthilkumar N C, School Of Information Technology and Engineering, VIT University, Vellore, India.

are not reproducible. The results proved the better classification performance of ASD/Non-ASD classes in BID used all 29 codes. In 2015, Use of Machine Learning to identify children with autism, their motor abnormalities by Aless et al [4] embraced a proof-of-idea concentrate to decide if a basic development could be helpful to precisely group low-working children with ASD aged 2-4. Data utilized are 15 preschool – children with 15 regularly creating (TD) children who were coordinated by mental age. They used pattern classification method based on supervised ML algorithm SVM was utilized to group ASD versus TD by expanding the separation between the 2 gatherings of datasets. They checked the execution of the classification algorithm by using a strategy, cross-validation. 2 datasets: cross verifying the training dataset with the testing dataset. The paper guarantees that machine-learning strategy could effectively arrange members by diagnosis, the grouping exactness achieved a most extreme precision of 96.7 % (affectability 100 %, explicitness 93.8 %) by utilizing 7 features chosen by the FDR-based procedure.

Identifying Autism with a Brief , Low-Cost Screening Instrument—OERA: Construct Validity, Invariance Testing, Agreement between Judges in 2018, proposed another assessment instrument OERA [7] to screen children discernible social direct with no tremendous information on ASD required. Data utilized in this examination is 99 children aged from 3 to 10: without ASD - 23, with ASD - 76. The test understood a cut-off of five or higher achieved the most critical affectability (92.75), refinement (90.91) rates. An audit by Falkmer watching out on ASD saw 3 tools for observational screenings to review the related cases with ASD: (ADOS),(ADI--R) Autism Diagnostic Interview-Revised,(CARS) Childhood Autism rating scale - none of them gave validity. The OERA scores exhibited a mean of 0.3, 0.1 among common children to a mean of 4.6, 3.7 among non–ASD with ID children, (M=15.8, M=16.4) among ASD children.

Autism Spectrum Disorder Decision Tree Subgroups predict adaptive behaviour, autism severity trajectories in children with ASD in 2018, by Michael J. Flory et al [10] experienced a few endeavours to manage this heterogeneous ASD by distinguishing subgroup by connecting CS watched social conduct profiles to biological factors, eg, genotype. A CS examination of PDDBI Data broke down with order, regression tree yielded a decision tree (ASD -- DT) distinguished 3 behaviourally particular ASD subgroups: verbal, atypical, minimally verbal. A total of 110 participants between the ages of 1.5, 6.9 years, who had complete parent PDDBI. 49 were seen - 3, 18 - 4, 4 – 5; 1 child was seen - 6 times. MINIVERBAL: the one that showed the most impairment with decreases in adaptive standard scores across age. ATYPICAL: showed no change in standard scores, but this was restricted to the two domains most linked to ASD: COM, SOC. Finally, the VERBAL: the one showed the most enhancements, predictable with past research.

III. METHODS

DATASET:

Very limited datasets are available for autism related to clinical or screening, most of them are genetic in nature. Dataset [1] contains 21 features: 10 behavioral features, 10 individual's characteristics with 704 instances. The attributes are age, gender (M/F), born with jaundice(y/n), genetic problem(y/n), used screening app, who is completing the test, 10 more individual's characteristics. 52% of the participants were male. Dataset contains some missing values which weren't provided by the participants. In this manner, a period effective, available ASD screening is fast moving toward that helps wellbeing experts, advises people whether they should seek after formal clinical analysis.

ATTRIBUTE SELECTION:

Jaundice is likewise connected to a higher hazard for autism. In uncommon cases, it can result in cerebrum harm, cerebral paralysis, even demise. Babies who developed jaundice were 67% bound to be determined to have autism amid early childhood. Measurements state that babies created jaundice, 532 were later determined to have ASD. The family history of an ASD can build the hazard for an autism diagnosis. Gender – the point of the present examination is to discover which gender has a higher probability of having ASD.

IV. ALGORITHMS

RANDOM TREE:

A random tree is a tree/arborescence which is formed by a random process, can deal with both classification, regression problems. The classifier takes the input feature vector, classifies it with every tree in the forest, and outputs the class label that received the majority of votes. All the trees are trained with the same parameters but on training sets.

RANDOM FOREST:

Random forests or random decision forests are ML method for classification, regression, other operations run by constructing a multiple of decision [12] trees at the training time, the class that is the mode of the classes (classification). Here the multiple decision trees are nothing but the trees generated from random tree algorithm we used for each parameter, supplied as input to the random forest.

V. ARCHITECTURE

By classifying the attributes, we were able to find the main traits which cause ASD. The attributes classified are gender, jaundice, genetic with different classifiers. Clustered attributes are tested with ML algorithms RT, RF. The tables show the detailed information how the attributes classified by undergoing random tree, followed by random forest.

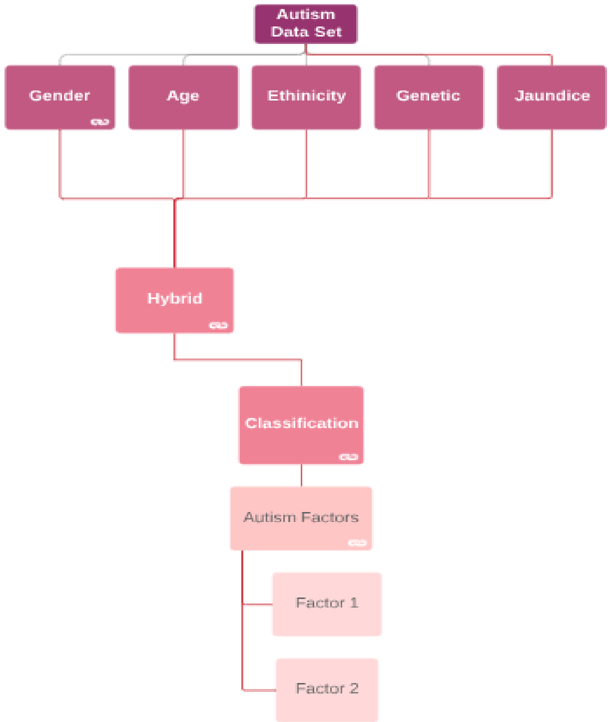

Figure 1. Architecture of the System

VI. RESULTS

A. RANDOM TREE

1 .GENDER:

Table 1. Summary of Gender Classification

CCI	693	98.4375 %
ICCI	11	1.5625%
KS	0.9687	
MAE	0.0471	
RMSE	0.1264	
RAE	9.4307 %	
RRSE	25.303 %	
T. INSTANCES	704	

2. JAUNDICE

Table 2. Summary of Jaundice based Classification

CCI	698	99.1477%
ICCI	6	0.8523%
KS	0.9505	
MAE	0.0131	
RMSE	0.0725	
RAE	7.3834%	
RRSE	24.3751%	
T. INSTANCES	704	

3. GENETIC:

Table 3.Summary of Results based on Genetic Factors

CCI	704	100%
ICCI	0	0
KS	1	
MAE	0.0015	
RMSE	0.0168	
RAE	0.6678%	
RRSE	4.994%	
T. INSTANCES	704	

Values from the table (jaundice, gender, genetic) gave less error compared other values. The trees have sent to RF. The below tables shows the detailed information of each trait's classified by undergoing RF by using Weka tool

B.RANDOM FOREST

1. GENDER:

Table 4 .Summary of Gender Classification

CCI	703	99.858%
ICCI	1	0.142%
KS	0.9979	
MAE	0.1835	
RMSE	0.2002	
RAE	36.7731%	
RRSE	40.0861%	
T. INSTANCES	704	

2. JAUNDICE

Table 5. Summary of Jaundice based Classification

CCI	701	99.1477%
ICCI	3	0.8523%
KS	0.9754	
MAE	0.0653	
RMSE	0.119	
RAE	36.7265%	
RRSE	40.0238%	
T. INSTANCES	704	

3. GENETIC:

Table 3.Summary of Results based on Genetic Factors

CCI	704	100%
ICCI	0	0
KS		1
MAE		0.0783
RMSE		0.1302
RAE		34.6519%
RRSE		38.8148%
T. INSTANCES		704

Results applying after RF showed much lesser error percentage (99.858%) than the error showed in RT (98.4375 %). The above shown results are run using weka tool so that we can able to identify the algorithm which will produce an accurate result. From the results, we were able to identify the trait which causes ASD. As predicted, genetics is also the possibility of ASD. 57% of boys are said to have ASD from the data supplied. About 60% of all babies are born with jaundice. Considering it is as a trait proves the possibility of having ASD. Other than figuring the exactness of the ML algorithms, we were especially keen on distinguishing which attribute contributed towards the classification

VII. CONCLUSION

Our outcomes give proof that ASD brought about by jaundice. We utilized a blended – sexual orientation dataset for characterization, saw that girls with ASD are not exactly boys with ASD.The critical prescient estimation of our characterization approach may be significant to help the clinical routine with regards to diagnosing ASD, thus empowering a PC supported diagnosis point of view. The current aims of the study are found to be (1) Genetic, jaundice are found to be the main cause for ASD, (2) it is identified that males with ASD are more than females with ASD. This may guide further exploration of the neuropathology of the disorder with jaundice, gender analysis.

REFERENCES

1. Fadi Thabtah, Nelson Marlborough Institute of Technology, Autism Spectrum Disorder Screening: Machine Learning Adaptation , DSM-5 Fulfilment.
2. Margaret McGonigle-Chalmers, Ben Alderson-Day (2010) Free Classification as a Window on Executive Functioning in Autism Spectrum Disorders, Journal of autism , developmental disorders. doi: 10.1007/s10803-010-0947-5.
3. Miia Kaartinen, Kaija Puura, Mika Helminen, Raili Sallmelin, Erja Pelkonen, , Petri Juujarvi (2014), Reactive Aggression among Children With , Without Autism Spectrum Disorder, Journal of autism , developmental disorders. doi: 10.1007/s10803-012-1743-1.
4. Aless,ro Crippa, Cheistian Salvatore, Paolo Perego, Sara Forti, Maria Nobile, Massimo Molteni , Isabella Castiglioni (2015), Use Of Machine Learning to Identify Children With Autism , their Motor Abnormalities, Journal of autism , developmental disorders. doi: 10.1007/s10803-015-2379-8.
5. Daniel Bone, Matthew S. Goodwin, Matthew P. Black, Chi-Chun Lee, Kartik Audhkhasi, Shrikanth Narayanan (2015), Applying Machine Learning to Facilitate Autism Diagnostics: Pitfalls , Promises, Journal of autism , developmental disorders. doi: 10.1007/s10803-014-2268-6.C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.
6. Ira L. Cohen, Xudong Liu, Melissa Hudson, Jennifer Gillis, Rachel N. S. Cavalari, Raymond G. Romanczyk, Bernard Z. Karmel, Judith M. Gardner (2016), Using The PDD Behaviour Inventory As A Level 2 Screener: A Classification , Regression Tree Analysis. doi: 10.1007/s10803-016-2843-0.M. Young, *The Techincal Writers Handbook.* Mill Valley, CA: University Science, 1989.
7. Cristiane S. Paula, Graccielle Rodrigues Cunha, Daniela Bordini, Decio Brunoni, Ana Claudia Moya, Cleonice Alves Bosa, Jair J. Mari, Hugo Cogo-Moreira (2017), Identifying Autism with a Brief , Low-Cost Screening Instrument—OERA: Construct Validity, Invariance Testing, , Agreement Between Judges. doi: 10.1007/s10803-017-3440-6.
8. John F. Strang, Laura G. Anthony, Benjamin E. Yerys, Kristina K. Hardy, Gregory L. Wallace, Anna C. Armourl, Katerina Dudley, Lauren Kenworthy (2017), The Flexibility Scale: Development , Preliminary Validation of a Cognitive Flexibility Measure in Children with Autism Spectrum Disorders. Journal of autism , developmental disorders. doi: 10.1007/s10803-017-3152-y.
9.] Linda R. Watson, Elizabeth R. Crais, Grace T. Baranek, Lauren Turner-Brown, John Sideris, Linn Wakeford, Jessica Kinard, J. Steven Reznick, Katrina L. Martin, Sallie W. Nowell (2017), Parent-Mediated Intervention for One-Year-Olds Screened as At- Risk for Autism Spectrum Disorder: A R,omized Controlled Trial. doi: 10.1007/s10803-017-3268-0.
10. Ira L. Cohen · Michael J. Flory (2018), Autism Spectrum Disorder Decision Tree Subgroups Predict Adaptive Behaviour , Autism Severity Trajectories in Children with ASD. doi: 10.1007/s10803-018-3830-4.
11. G. Bussu, E. J. H. Jones, T. Charman, M. H. Johnson, J. K. Buitelaar, the BASIS Team (2018), Prediction of Autism at 3 Years from Behavioural , Developmental Measures in High-Risk Infants: A Longitudinal Cross-Domain Classifier Analysis. Journal of autism , developmental disorder., doi: 0.1007/s10803-018-3509-x
12. Nallakaruppan, M.K., Kumaran, U.S(2018),Quick fix for obstacles emerging in management recruitment measure using IOT-based c,idate selection .Service Oriented Computing , Applications, 12 (3-4), pp. 275-284.