An Approach Towards Autism Spectrum Disorder Detection: A Review

Vyankatesh Rampurkar [1], Omkar Lokhande[2], Abhishek Lambate[2], Bhagyashri Dhage[2], Shruti Jare [2]

[1]Asst. Prof, Department of Computer Engineering, Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Savitribai Phule Pune University, Pune, India.

[2]Student, Department of Computer Engineering, Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Savitribai Phule Pune University, Pune, India.

Corresponding Author: omkarlokhande002001@gmail.com

Abstract: The autism spectrum disorder (ASD) has been prolife rating rapidly around the world. Autism Spectrum Disorder (ASD) is a neurological disorder which might have a lifelong impact on the language learning, speech, cognitive, and social skills of an individual. Its symptoms usually show up in the developmental stages. ASD is mainly caused by genetics or by environmental factors, however, its conditions can be improved by detecting and treating it at earlier stages. In many aspects of research, machine learning played an extensive role for development especially in terms of data analytics. Implement multiple machine learning techniques, to predict Autism Spectrum Disorder (ASD) over large scale datasets. Autism is also identified as a range of conditions categorized by various challenges such as social skills, repetitive behaviors, Speech and non-verbal communication and unique strengths and differences. The ultimate goal is to train the data model with the set of training data and then testing and evaluating the data model using the test data. As a result, the research will come up with a solution that applies machine learning to detect autism.

Key Words: — ***ASD-Autism Spectrum Disorder, ADOS- Autism Diagnostic Observation Schedule, SVM-Support Vector Machines.***

I. INTRODUCTION

Autism, first proposed in 1943 by Kanner. Autism is a neurodevelopmental disorder that affects a person's communication, interaction and learning skills. In recent years, the prevalence of autism spectrum disorder (ASD) has been enormously increasing. Although diagnosis of autism can be done at any age, its symptoms generally appear in the first two years of life and develops through time. Autism patients face different types of challenges such as difficulties with mental health problems such as anxiety, depression, concentration, learning disabilities, motor difficulties, sensory problems and many others. Current explosion rate of autism spectrum disorder around the world is numerous and it is increasing at a very high rate. The increase in the number of patients hints at a large group of people sufferers, which may also reflect people's increasing awareness of the disorder. Autism spectrum is order has a significant economic impact due to the increase in the number of autism spectrum disorder cases worldwide. Diagnosis of autism spectrum disorder requires significant amount of time and cost. Earlier detection of autism spectrum disorder can come to a great help by prescribing patients with proper medication at an early stage.

Using machine learning, Autism spectrum disorder can be detected more quickly. It can prevent the patient's condition from deteriorating further and would help to reduce long term costs associated with delayed diagnosis. When past data is used to predict autism, machine learning takes less time. Decision Tree, SVM and Logistic Regression algorithms are used to detect Autism spectrum disorder, and the datasets are

Manuscript revised December 16, 2022; accepted December 17, 2022. Date of publication December 18, 2022.

This paper available online at www.ijprse.com

ISSN (Online): 2582-7898; SJIF: 5.59

collected from online kaggle.com. In order to obtain accurate results that can be used to predict autism spectrum disorder at an early stage, these data are analyzed. The purpose of applying classification techniques is to obtain improved precision, recall and predictive accuracy on the results of the feature selection methods for better result. Thus a time efficient, accurate and easy screening test tool is required which would predict autism traits in an individual and identify whether autism or not.

II. LITERATURE SURVEY

2.1 Autistic Spectrum Disorder Screening: Prediction with Machine Learning Models

Recently in the research field, there is a high demand for automating processes to reduce the cost and time of any industry. Machine learning is important factor in current technologies. In various approaches in research, machine learning played an important role for development specially in terms of data analytics.

Healthcare is well known field where Machine Learning made big phases of enhancements because of the huge amount of data being processed and analyzed. It changed the diagnosing process for better accuracy based on the new technologies emerged in the past decade that became tools in every hospital. The purpose of the thesis is to research and implement multiple machine learning techniques, to predict Autism Spectrum Disorder over large scale datasets.

Autism Spectrum Disorder is a neurological and developmental disorder that starts early in stages and lasts throughout a person's life affecting information flow in the brain leading to other problems for the patient. Autism Spectrum Disorders bring about certain challenges in social, behavioral, communication and emotional understanding in an individual. Autism is also identified as a range of conditions like social skills, repetitive behaviors, Speech and non-verbal communication and unique strengths and differences.

People diagnosed with autism spectrum disorder have a range of symptoms. That is why it is termed as a ‘spectrum’ disorder. These disorders can be perceived in early childhood, there are some cases in which the symptoms are not diagnosed until adulthood. The main cause of Autism is not defiantly identified; research shows links between genetics and the environmental factors with Autism. Autism Spectrum Disorder currently has no standard treatment to cure it. An early diagnosis and a head start in therapies can potentially lead to better results.

2.2 A machine learning based approach to classify autism with optimum behavior sets

The proposed system aimed to design an automated ASD prediction model with minimum behaviour sets selected from ASD diagnosis dataset feature selection. Machine Learning based behavioural analytics emphasis the need to develop accurate prediction models for detecting the risk of autism faster than the traditional diagnostic methods. Quality of prediction rely on the accuracy of the supplied dataset and the machine learning model. To improve accuracy of prediction, dimensionality reduction with feature selection is applied to eliminate noisy features from a dataset. In the proposed system works with an ASD diagnosis dataset with 21 features obtained from UCI machine learning repository is experimented with swarm intelligence based binay firefly feature selection wrapper.

The feature selection algorithm is a recent and fast performer that has outperformed benchmark algorithms such as on 40 datasets. The dataset can produce 220 feature subsets for evaluation. Exhaustive search based feature subset selection algorithms will face exponential increase in time complexity as feature subselection is classified as an np-hard problem. Stochastic search algorithms with objective evaluation function and feature elimination algorithms with candidate evaluation are best solutions to overcome the np-hard search problems. According to features selected by ranking approach are highly prone to inter-feature correlation bias and thus leads to redundancy among features which inversely affect the performance of machine learning model.

The ASD children dataset is trained with 8 different machine learning algorithms using 10 fold cross validation. The results of module are compared with results of machine learning models obtained after feature selection with Binary firefly feature selection algorithm. Binary firefly algorithm is accelerated with a logistic chaotic map to boost attractiveness. The local and global search strategy of feature selection is enhanced by simulated annealing. Thus, the algorithm converges towards global best solution within minimum iterations. The binary firefly feature selection algorithm is classified as a swarm intelligence optimizer based wrapper feature selection algorithm with a single objective function. The hypothesis is to find whether machine learning models trained with minimum behaviour sets are capable of better performance or not. In order to select features aswarm intelligence-based wrapper is considered as a better alternative to Ranking based feature elimination algorithms. From the above results and discussions, the hypothesis is validated.

2.3 Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning

The Autism Spectrum Disorders (ASD) is now on the rise affecting 1 in 68 individuals in the United States. Despite an important advance in understanding the genetics of ASD, ASD remains diagnosed by behavioral testing. ASD currently performed using tools designed to measure impairment in the two main domains of ASD as defined according to the Diagnostic and Statistical Manual of Mental Disorders, communication and social interaction and restricted interests and repetitive behaviors. Autism diagnostic Observation Schedule (ADOS) is one of the most widely used uses tools to help diagnose ASD. ADOS consists of a series of semi-structured activities designed to elicit specific social interaction behavior, communication, imaginative use objects, limited interests, and repetitive behaviors.

Diagnostic test is divided into four modules, each of which is adapted to a specific individuals based on their language and developmental level ensure coverage of a diverse set of behavioral manifestations. A certified professional in a clinical setting first administers the ADOS examination and then scores the individual based on his or her observation to establish a final .Initial Rating - the evaluation may take 30 to 60 minutes increasing the total time to 60 to 90 minutes. Due to differences in interater reliability, other experts may re-evaluate an individual score, further increasing the time between testing and receiving an official clinical diagnosis. Even ignoring geographic and logistical obstacles to finding a certified professional to administer the ADOS at the required time to the test and the increase in the number of children at risk of ASD have contributed to the increase of bottlenecks in the healthcare sector system.

Therefore, risk assessment and sorting tools that can to reach families earlier and allow them to receive such care needs are very needed. Given the promising findings from our previous work on the first module of ADOS and ADI-R, we assumed that we could get similar results by examining records from the other two ADOS modules that apply to a large part population suspected of ASD. Improving our previous work, here we using the best estimate of the clinical diagnosis where possible, and incorporated a gradual reverse function selection into our machine learning pipeline quantitatively select an optimal set of significant behavioral features that can to accurately detect ASD risk in a large population of individuals.

Behavioral surveys and questionnaires are time-consuming currently the primary methods used in the diagnosis of PAS.By using machine learning, we created classifiers from two modules of one of the most universally administered behavioral tests, the ADOS. A logistic regression classifier based on archival analysis records from ADOS Module 2 consisted of nine items, 67.86 % less than the entire ADOS 2 module and is performed with 98.81% sensitivity and 89.39 % specificity in independent testing. SVM 3module classifier based on the analysis of archived ADOS Module 3 entries consisted of 12 items,57.14 % less than complete ADOS 3 module and execution with more than 97 % sensitivity and specificity in testing. Moreover, these results can help encourage the future efforts to develop screening-based ASD detection tools and mobile health approaches that ultimately empower individuals to receive more effective care than is currently possible paradigms.

2.4 Prediction of Autism Spectrum Disorder Using Supervised Machine Learning Algorithms

Autism is neuro developmental disorder that is visible in toodlers, childrens and adults in early years. Its symptoms can vary from person to person.it is found that one in 68 peoples have autism. Detection of autism is a complicated procedure because autism patient has many symptoms can be mixed with other mental health conditions, therefore early detection of autism is important. Machine learning is a powerful computing tool that learn complex relationships from large data sets to draw accurate conclusions. Disease assessment can be done using predictive health data analysis .Supervised learning is an important technique of machine learning that uses rule-based approach by examining datasets build accurate predictive models. Decision tree, Random Forest and SVM algorithms can be used to detect autism. An analysis can be performed on them accurate results that will be useful for making the right decisions and prediction of autism spectrum disorder (ASD) in the early stages. Early detection of autism using machine learning techniques open a new path for autistic individuals to develop the potential to lead a better life by improving one's abilities, behavioral and emotional skills.

ASD can be caused by chemicals used in various foods or environmental factors. Diagnos is of the autism the process is a manual process carried out by several streams doctors by observing the child's behavior. ASD is detected at the age of 3 years or above 3 years with manual diagnosis. Early detection of ASD can be done at an early age which improves the curing mechanism. machine learning is one that is efficient and reliable detect ASD with shorter processing time. Machine learning can train the system with past data and then the machine predict ASD. There are many supervised machine

learning algorithms like decision tree, random forest, SVM that can be used to predict autism.

Machine learning is known as training an algorithm to use on different test data. The main goal of machine learning algorithms is to train the system and work automatically without human intervention. Learning process it starts with preprocessing the data and removing noisy data from the dataset such as missing values, outliers.

Supervised Learning is one part of machine learning and artificial intelligence. ASD prediction is done by generating a model with an existing data set that is used for future prediction of ASD with shorter processing time. Supervised learning uses existing input and output data generate a model for future data processing or prediction. ASD prediction process using supervised. Learning algorithms involve two steps. In the first step the training data is used to generate the model classifier. In second step the generated classifier model is applied with the test data. Decision trees, SVM and Random Forest algorithms predict ASD with optimal accuracy.

Comparison of supervised learning algorithms applied to ASD dataset. A comparison of supervised learning algorithms is used in choosing a classification algorithm for effective prediction of ASD. Accuracy of model helps to decide more efficient algorithm for prediction of ASD.

Table.1. Comparison Table

Paper Name	**Methods**	**Accuracy**
Autistic Spectrum Disorder Screening Prediction with Machine Learning models	Random Forest, SVM, ANN, Decision Tree	99%
A machine learning based approach to classify Autism with optimum behavior sets	Binary furefly	92%
Applying machine learning to identify autistic adults using imitation	Imitation method with ML classifier (SVM Linear)	86%
Searching for minimal set of behaviors for autism detection through feature selection-based machine learning	ADOS Algorithm	97%
Prediction of autism spectrum disorder using supervised learning algorithm	Random Forest, Desicion Tree	98%

III. CONCLUSION

An Autism Spectrum Disorder needs early treatment. The sooner Autism Spectrum Disorder is diagnosed, the better are the results in long term. The outcome of the research provides an effective and efficient approach to detect autism traits for different age groups. Data analysis is done to figure out the relation between the attributes. Decision Tree results in an overfitted model in all the datasets. Logistic Regression gives the optimal result for adult autism dataset. Support Vector is the best for child autism screening dataset. With the help of autism screening application, an individual can be guided at an early stage that will prevent the situation from getting any worse and reduce costs associated with delayed diagnosis.

REFERENCES

[1]. Autistic Spectrum Disorder Screening: Prediction with Machine Learning Models Astha Baranwal Vanitha.M 2020.

[2]. Vaishali, R., and R. Sasikala. "A machine learning based approach to classify Autism with optimum behaviour sets. (2018)" International Journal of Engineering and Technology.

[3]. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning JA Kosmicki, V Sochat, M Duda and DP Wall 2015.

[4]. Prediction of Autism Spectrum Disorder Using Supervised Machine Learning Algorithms T. Lakshmi Praveena and N. V. Muthu Lakshmi.

[5]. Applying machine learning to identify autistic adults using imitation: An exploratory study Baihua Li, Arjun Sharma James Meng Senthil Purushwalkam, Emma Gowen 2017.

[6]. F. Thabtah, "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward," Informatics for Health and Social Care, pp. 1–20, 2018.

[7]. R. GeethaRamani and K. Sivaselvi, "Autism Spectrum Disorder Identification Using Data Mining Techniques", International Journal of Pure and Applied Mathematics.

[8]. Maria Nobile, Massimo Molteni, and Isabella Castiglioni, Use of Machine Learning to Identify Children with Autism and Their Motor Abnormalities.

[9]. L. Zhang, K. Mistry, C. P. Lim, and S. C. Neoh, "Feature selection using firefly optimization for classification and regression models," Decis. Support Syst., 2017.

[10]. Kakarwal, S. N., Paithane, P. M., &Khedikar, S. S. Natural Language Processing and Speech Signal for Mental Health Research and Care Application.