A Review Paper on Autism Detection using Machine Learning

Rolapati Sushmitha[1], Golla Shiva Kumar[2]

[1] Student, Information Technology, B V Raju Institute of Technology, Telangana, India
[2] student, Information Technology, B V Raju Institute of Technology, Telangana, India

ABSTRACT

Health is an important sector which stands at the center to move forward in life. Children between the ages of 6 and 17 may develop autism spectrum disorder(ASD), a form of neurological condition that affects social interaction and communication are impacted by ASD. It makes patients engage in repetitive actions. Early detection in childhood is crucial because it can help children with ASD with their social skills and communication issues, improving their quality of life. In this paper we made an attempt to explore Logistic Regression for predicting and analysis of Autism Spectrum Disorder problems in an early stage

Keyword: *Autism Spectrum disorder(ASD); Logistic Regression; K Nearest – neighbor(KNN);Random Forest*

1. INTRODUCTION

WHAT IS AUTISM?

Variations in the brain cause the developmental disability known as autism disorder. ASD sufferers may also move, pay attention, or learn in different ways.

What is Machine Learning?

The field of study known as machine learning enables computers to learn without being explicitly programmed. Using sample data or prior knowledge, machine learning involves programming computers to optimize a performance criterion. We have a model that is defined up to a certain point, and learning is the application of a computer program to optimize the model's parameters using training data or prior knowledge.

Types of Machine Learning

Supervised Machine Learning

To train the supervised learning models, we use the labelled dataset. The model is tested by supplying a sample of test data after training and processing to see if it predicts the proper output.

Unsupervised Machine Learning

Unsupervised models can be trained using a dataset that has no classifications or categorizations and no labels, and the algorithm must act on that data without any supervision.

Reinforcement Learning

In reinforcement learning, an agent creates interactions with its surroundings and absorbs information from feedback. The agent gets feedback in the form of rewards; for instance, he gets a good reward for every good action and a terrible reward for every bad action.

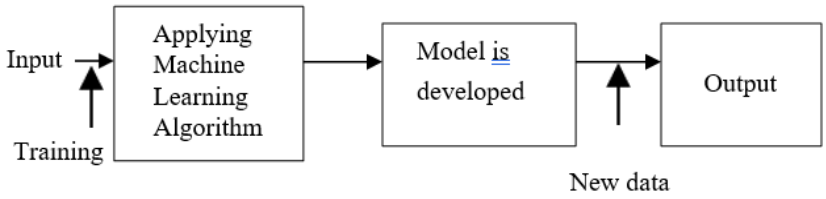

Fig-1*:* Machine Learning Block Diagram

2. LITERATURE REVIEW

Machine learning in ASD behavioral research: A review:

Autism spectrum disorder (ASD) patients struggle to develop language, communication, cognitive, and social skills.. Some individuals with ASD have remarkable intellectual, non-academic, and artistic talent, which makes it challenging for researchers to develop solutions in these circumstances. Now, researchers investigating social & computational intelligence have begun studying ASD, employing cutting-edge methods like machine learning to improve diagnosis timeliness, accuracy, & quality. In the diverse field of machine learning, smart techniques are utilized to uncover significant hidden patterns that are forecasted to be used in decision-making. By applying ML techniques like SVM, decision trees, LR, and others to datasets relevant to the illness, predictive models for autism have been developed. These models aim to enhance doctors' abilities to provide comprehensive diagnoses and prognoses about ASD. However, studies on the use of machine learning for ASD diagnosis and treatment are hampered by conceptual, implementation, and data issues, such as the way diagnostic codes, feature selection, evaluation methods, and class imbalances in data are chosen, among other things. A more serious claim made in recent studies is the development of a brand-new machine learning-based approach for detecting ASDs. This work critically analyses prior investigative studies on autism and makes recommendations for future research that will enhance conception, implementation, and data concerning machine learning in ASD in addition to addressing the issues raised above in these investigations. These suggestions will significantly help future research on machine learning in autism studies.

3. PROPOSED METHODOLOGY

People with ASD also struggle with their limited interests and repetitive behaviors. In the list below, behavior categories are shown with examples from specific situations.

- Repetition of specific actions, such as words or phrases, over and over again
- A person will become unhappy when a routine is likely to change.
- Possessing a little of an interest in particular topics, such facts, figures, etc.
- Less sensitive than another person in some circumstances, such as those involving light, noise, etc.

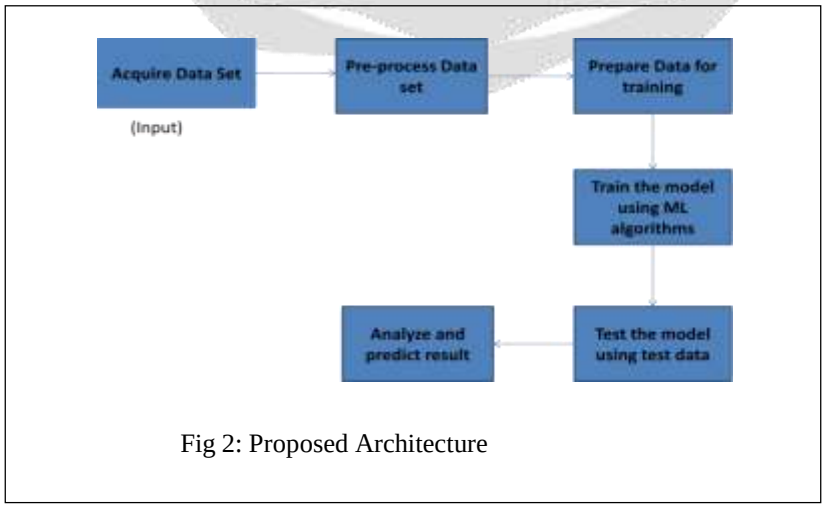

Fig 2: Proposed Architecture

The proposed architecture's modules include:

1) Upload the ASD Dataset: This module is used to upload the dataset for the application

Name of the dataset	:	Toddler Autism Dataset
Source	:	Kaggle
Types of Attributes	:	continuous and binary
No of Attributes	:	18 including class variable
Instances	:	1054

List of Attributes:

Attribute ID	Description
1	S No
2-11	Based on answers for 10 questions
12	Age months
13	Qchat-10-Score
14	Gender
14	Ethinicity
16	Jaundice
17	Family with ASD or not
18	Who completed the test
19	Class/ASD traits

2) Pre-process the data:

There are 19 attributes present int the dataset. By using Information Gain technique the required attributes have been selected. Information gain is used to determine how much entropy will be lost as a result of a dataset update.. By assessing each variable's information gain in relation to the target variable, it can be used for feature selection.

Selected attributes are:

S No
2-11 : Based on answers for 10 questions
Age
Family with ASD or not

Now the entire data is been splitted into training and testing dataset.80 percent of data is splitted for training and 20

3) Run RANDOM FOREST algorithm: now that the training data has been analyzed, the RANDOM FOREST algorithm will use it to create a trained prediction model. This model will then be used to 20% of the test data to calculate the RANDOM FOREST prediction accuracy

4) Run KNN algorithm: now that the train data has been analyzed, the trained prediction model will be applied to 20% of the test data to determine the KNN prediction accuracy.

5)Run Logistic Regression: Now that the train data has been analyzed, the trained prediction model will be applied to 20% of the test data to determine the accuracy of the LOGISTIC REGRESSION algorithm.

6) Detect Autism from Test Data: using this module we will upload test data for each algorithm and each algorithm will predict weather test data is normal or contains Autism disorder.

7) Visualization of the data: This module will be used to create an accuracy graph for all algorithms.

4. RESULTS

By using a confusion matrix, the models' performance is assessed in terms of accuracy. The model's training accuracy determines the outcomes. The effectiveness and application of the classification models on the test dataset are assessed using these indicators. For toddler data with ASD, the results of multiple ML approaches using the chosen features have been displayed.

Table 1: Results for ASD Data for children

Algorithm	Accuracy
Logistic Regression	100
Random Forest	97.15
KNN	90.99

5. CONCLUSION

With minimal behavior sets chosen from each diagnosis dataset, we created an automated ASD prediction model. We have used various supervised learning techniques like Logistic Regression, K-nearest neighbor, Random Forest out of which logistic regression has given high accuracy standard for producing a better estimating result.

6. FUTURE WORK

The proposed system takes lot of time in observing the features in child. Hence we can directly take a video of a child and then convert the video into frames. Later extract the Autism related features in child. Then apply Machine Learning algorithms to predict the Autism in child.

7. REFERENCES

[1] McCarty, Patrick, and Richard E. Frye. "Early detection and diagnosis of autism spectrum disorder: why is it so difficult?." In Seminars in Pediatric Neurology, vol. 35, p. 100831. WB Saunders, 2020.

[2] Thabtah, Fadi. "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward." Informatics for Health and Social Care 44, no. 3 (2019): 278-297.

[3] Constantino, John N., Patricia D. Lavesser, Y. I. Zhang, Anna M. Abbacchi, Teddi Gray, & Richard D. Todd. (2007) "Rapid quantitative assessment about autistic social impairment through classroom teachers." Journal about American Academy about Child & Adolescent Psychiatry 46(12): 1668-167