Visvesvaraya Technological University

Jnana Sangama, Belagavi - 590018

A Project Work Phase-2 (18CSP83)

Report on

"Machine Learning based model for prediction of Autism Spectrum Disorder"

Project Report submitted in partial fulfilment of the requirement for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

Dhanush Vasudevaraju	**1KG19CS025**
C.Ramesh Chandra	**1KG19CS015**
C.Sai Ranganath	**1KG19CS016**
Dinesh.M	**1KG19CS026**

Under the guidance of
Mrs. Sandhya A. KulkarniAsst.
Prof, CSE, KSSEM
Department of Computer Science & Engineering
KSSEM, Bengaluru-560109

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

K. S. School of Engineering and Management

#15, Mallasandra, off. Kanakapura Road, Bengaluru – 560109

2022 - 2023

K. S. School of Engineering and Management

#15, Mallasandra, off. Kanakapura Road, Bengaluru - 560109

Department of Computer Science & Engineering

KSSEM

CERTIFICATE

Certified that the Project Work Phase-II (18CSP83) entitled **"Machine Learning based model for prediction of Autism Spectrum Disorder"** is a bonafide work carried out by:

Dhanush Vasudevaraju	**1KG19CS025**
C.Ramesh Chandra	**1KG19CS015**
C.Sai Ranganath	**1KG19CS016**
Dinesh.M	**1KG19CS026**

in partial fulfilment for VIII semester B.E., Project Work in the branch of Computer Science and Engineering prescribed by **Visvesvaraya Technological University, Belagavi** during the period of February 2023 to May 2023. It is certified that all the corrections and suggestions indicated for internal assessment have been incorporated. The Project Work Phase-2 Report has been approved as it satisfies the academic requirements in report of project work prescribed for the Bachelor of Engineering degree.

Signature of the Guide	**Signature of the HOD**	**Signature of the Principal**
[Mrs. Sandhya A. Kulkarni]	[Dr. K Venkata Rao]	[Dr. K Rama Narasimha]

HOD
Department of Computer Science Engineering
K.S School of Engineering & Management
Bangalore-560109

Dr. K. RAMA NARASIMHA
Principal/Director
K S School of Engineering and Management
Bengaluru - 560 109

DECLARATION

We, the undersigned students of 7th semester, Computer Science & Engineering, KSSEM, declare that our Project Work Phase-II entitled "**Machine Learning based model for prediction of Autism Spectrum Disorder**", is a bonafide work of ours. Our project is neither a copy nor by means a modification of any other engineering project.

We also declare that this project was not entitled for submission to any other university in the past and shall remain the only submission made and will not be submitted by us to any other university in the future.

Place:

Date :

Name and USN	**Signature**
Dhanush Vasudevaraju (1KG19CS025)	Dhanush
C.Ramesh Chandra (1KG19CS015)	C Ramesh
C.Sai Ranganath (1KG19CS016)	Sai Ranganath
Dinesh.M (1KG19CS026)	Dinesh M

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task will be incomplete without the mention of the individuals, we are greatly indebted to, who through guidance and providing facilities have served as a beacon of light and crowned our efforts with success.

We would like to express our gratitude to our **MANAGEMENT,** K.S. School of Engineering and Management, Bengaluru, for providing a very good infrastructure and all the kindness forwarded to us in carrying out this project work in college.

We would like to express our gratitude to **Dr. K.V.A Balaji, CEO**, K.S. School of Engineering and Management, Bengaluru, for his valuable guidance.

We would like to express our gratitude to **Dr. K. Rama Narasimha, Principal**, K.S. School of Engineering and Management, Bengaluru, for his valuable guidance.

We like to extend our gratitude to **Dr. K Venkata Rao**, **Professor and Head**, Department of Computer Science & Engineering, for providing a very good facilities and all the support forwarded to us in carrying out this Project Work Phase-I successfully.

We also like to thank our Project Coordinators, **Mrs. Jayashubha J, Asst. Professor, Mrs. Supriya Suresh, Asst. Professor, Department of Computer Science & Engineering** for their help and support provided to carry out the Project Work Phase-I successfully.

Also, we are thankful to **Mrs. Sandhya A. Kulkarni , Assistant Professor** for being our Project Guide, under whose able guidance this project work has been carried out Project Work Phase-2 successfully.

We are also thankful to the teaching and non-teaching staff of Computer Science & Engineering, KSSEM for helping us in completing the Project Work Phase-2 work.

Dhanush

Vasudevaraju

C.Ramesh Chandra

C.Sai Ranganath

Dinesh.M

ABSTRACT

Autism spectrum disorder is a neurodevelopmental disorder that affects a person's interaction, communication and learning skills. Although diagnosis of autism can be done at any age, its symptoms generally appear in the first two years of life and develop through time. Autism patients face different types of challenges such as difficulties with concentration, learning disabilities, mental health problems such as anxiety, depression, motor difficulties, sensory problems, and many others. Diagnosis of autism requires significant amount of time and cost. Earlier detection of autism can come to a great help by prescribing patients with proper medication at an early stage. It can prevent the patient's condition from deteriorating further and would help to reduce long term costs associated with delayed diagnosis. Thus, an efficient, accurate and easy screening test tool is very much required which would predict autism traits in an individual. The main idea behind this project is to detect autism spectrum disorder in an individual (male/female). This project is implemented by making use of a Machine Learning model using parameters such as an individual's age, gender, ethnicity, Autism Quotient Tool. The detection derived from this project will help an individual to get required diagnosis in time to prevent further complications of developing Alzheimer's disease.

TABLE OF CONTENTS

Chapter No.	Title	Page No.
1.	**INTRODUCTION**	**1-4**
1.1	Overview	1
1.2	Purpose of the Project	2
1.3	Scope of the Project	2
1.4	Definitions	2
2.	**LITERATURE SURVEY**	**6 -16**
2.1	A Research Paper On Autism Spectrum Disorder	6
2.2	A Research Paper On Autism Spectrum Disorder	6
2.3	A Research Paper On Autism Spectrum Disorder	7
2.4	A Research Paper On Autism Spectrum Disorder	7
2.5	A Research Paper On Autism Spectrum Disorder	8
2.6	A Research Paper On Autism Spectrum Disorder	8
2.7	A Research Paper On Autism Spectrum Disorder	9
2.8	A Research Paper On Autism Spectrum Disorder	9
2.9	A Research Paper On Autism Spectrum Disorder	10
2.10	A Research Paper On Autism Spectrum Disorder	10

2.11	A Research Paper On Autism Spectrum Disorder	11
2.12	A Research Paper On Autism Spectrum Disorder	11
2.13	A Research Paper On Autism Spectrum Disorder	12
2.14	A Research Paper On Autism Spectrum Disorder	12
2.15	A Research Paper On Autism Spectrum Disorder	13
2.16	A Research Paper On Autism Spectrum Disorder	13
2.17	A Research Paper On Autism Spectrum Disorder	14
2.18	A Research Paper On Autism Spectrum Disorder	15
2.19	A Research Paper On Autism Spectrum Disorder	15
2.20	A Research Paper On Autism Spectrum Disorder	16
3.	**PROBLEM IDENTIFICATION**	**17**
3.1	Problem Statement	17
3.2	Project Scope	17
4.	**GOALS AND OBJECTIVES**	**18**
4.1	Project Goals	18
4.2	Project Objectives	18
5.	**SYSTEM REQUIREMENT SPECIFICATION**	**19**
5.1	Software Requirements	19
7.	**IMPLEMENTATION**	**22-45**
8.	**RESULTS AND SNAPSHOTS**	**46-49**

9.	**APPLICATIONS**	**50**
10.	**CONTRIBUTION TO SOCIETY AND ENVIRONMENT**	**51**
	REFERENCES	**52-58**
	APPENDIX - I CERTIFICATES OF PAPER PRESENTED	**59-62**
	APPENDIX- II PUBLISHED JOURNAL PAPER	**63-67**
	APPENDIX- III	**68-69**

LIST OF FIGURES

Fig. No.	Figure Name	Page No.
6.1	Flow chart of the application	**20**
6.2	Block Diagram of the user-interface	**21**
8.1	Accuracy results of Support Vector Classifier model for Toddler Dataset	**46**
8.2	Accuracy results of Random Forest model based on Toddler Dataset	**46**
8.3	Accuracy results of ANN model for Toddler Dataset	**47**
8.4	Models Accuracy comparison for Toddler Dataset	**47**
8.5	Accuracy results of Random Forest model on Child Dataset	**48**
8.6	Accuracy results of Support Vector Classifier model on Child Dataset	**48**
8.7	Accuracy Results of ANN model on child Dataset	**49**
8.8	Models Accuracy Comparison for Child Dataset	**49**

Chapter 1

INTRODUCTION

1.1 Overview

Autism spectrum disorder is a neurodevelopmental disorder that affects a person's interaction, communication and learning skills.Autism is generally understood to be a spectrum disorder, as it can manifest differently in each person: any autistic individual is likely to show some, but not all, of the characteristics associated with it, and the person may exhibit them to varying degrees and frequencies. There is large variation in the support needs of autistic people, and some are nonspeaking, while others have proficient spoken language. Although diagnosis of autism can be done at any age, its symptoms generally appear in the first two years of life and develop through time. Autism patients face different types of challenges such as difficulties with concentration, learning disabilities, mental health problems such as anxiety, depression etc, motor difficulties, sensory problems and many others.

Current explosion rate of autism around the world is numerous and it is increasing at a very high rate. According to WHO, about 1 out of every 160 children has ASD. Some people with this disorder can live independently, while others require life-long care and support.
Diagnosis of autism requires significant amount of time and cost. Earlier detection of autism can come to a great help by prescribing patients with proper medication at an early stage.It can prevent the patient's condition from deteriorating further and would help to reduce long term costs associated with delayed diagnosis. Thus a time efficient, accurate and easy screening test tool is very much required which would predict autism traits in an individual and identify whether or not they require comprehensive autism assessment.

The objective of this work is to propose an autism prediction model using ML techniques and to develop an user-interface that could effectively predict autism traits of an individual of any age. In other words, this work focuses on developing an autism screening application for predicting the ASD traits among people of age groups 0-2 years and 3-13 years.

1.2 Purpose of the project

The main purpose of this project is to develop an user-interface to implement a machine learning based solution that can detect Autism Spectrum Disorder in an individual(male/female) during early stages for early diagnosis using efficient algorithms to obtain results with required accuracy. Then, design and develop an user-interface for user to interact.

1.3 Scope of the project

Our proposed project focuses on a novel machine learning procedures for Autism spectrum disorder (ASD) classification and prediction, thus overcoming the existing problem. By utilizing Random Forest (RF), Support Vector Machine(SVM), AdaBooster algorithms we will make our model in order to increase the performance and accuracy.

1.4 Definition

1.4.1 Machine Learning

Machine Learning is a subset or specific application of Artificial intelligence that aims to create machines that can learn autonomously from data. Machine Learning is specific, not general, which means it allows a machine to make predictions or take some decisions on a specific problem using data. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, and computer vision,where it is difficult or unfeasible to develop conventional algorithmsto perform the needed tasks.

1.4.2 Random Forest Algorithm

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes or mean prediction of the individual trees.

1.4.3 Support Vector Machines

Support Vector Machines (SVMs), also support vector networks are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. It is formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples.

1.4.4 AdaBoost Algorithm

AdaBoost is short for Adaptive Boosting, is a statistical classification meta-algorithman and an ensemble learning method (also known as "meta-learning") which was initially created to increase the efficiency of binary classifiers. AdaBoost uses an iterative approach to learn from the mistakes of weak classifiers, and turn them into strong ones.

1.4.5 Artificial Neural Network

Artificial neural networks (ANNs), usually simply called neural networks (NNs) or neural nets are computing systems inspired by the biological neural networks that constitute animal brains.

An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons. An artificial neuron receives signals then processes them and can signal neurons connected to it. The "signal" at a connection is a real number, and the output of each neuron is computed by some non-linear function of the sum of its inputs. The connections are called *edges*. Neurons and edges typically have a *weight* that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a threshold such that a signal is sent only if the aggregate signal crosses that threshold.

1.4.6 Data Cleaning

Data cleansing or data cleaning is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data. Data cleansing may be performed interactively with data wrangling tools, or as batch processing through scripting or a data quality firewall.

After cleansing, a data set should be consistent with other similar data sets in the system. The inconsistencies detected or removed may have been originally caused by user entry errors, by corruption in transmission or storage, or by different data dictionary definitions of similar entities in different stores.

1.4.7 Data-Preprocessing

Data-Preprocessing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis. Therefore, certain steps are executed to convert the data into a small clean data set. This technique is performed before the execution of Iterative Analysis. The set of steps is known as Data Pre-processing.

1.4.8 Data Validation

Data validation is the process of ensuring data has undergone data cleansing to ensure they have data quality, that is, that they are both correct and useful. It uses routines, often called "validation rules", "validation constraints", or "check routines", that check for correctness, meaningfulness, and security of data that are input to the system. The rules may be implemented through the automated facilities of a data dictionary, or by the inclusion of explicit application program validation logic of the computer and its application.

1.4.9 Exploratory Data Analysis

Exploratory Data Analysis (EDA) is an approach of analyzing data sets to summarize their main characteristics, often using statistical graphics and other data visualization methods. A statistical model can be used or not, but primarily EDA is for seeing what the data can tell us beyond the formal modeling and thereby contrasts traditional hypothesis testing.

1.4.10 Python

Python is a popular programming language. It was created by Guido van Rossum, and released in 1991. Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for rapid applications development, as well as for use as a scripting or glue language to connect existing components together. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

1.4.11 Anaconda

Anaconda is a distribution of the Python and R programming languages for scientific computing, that aims to simplify package management and deployment. The distribution includes data-science packages suitable for Windows, Linux, and macOS.

1.4.12 Jupyter Notebook

The Jupyter Notebook is an open source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebook is maintained by the people at Project Jupyter. Jupyter Notebooks are a spin-off project from the IPython project, which used to have an IPython Notebook project itself. The name, Jupyter, comes from the core supported programming languages that it supports:Julia, Python, and R. Jupyter ships with the IPython kernel, which allows you to write your programs in Python, but there are currently over 100 other kernels that you can also use.

1.4.13 User-Interface

The user interface is the point at which human users interact with a computer, website or application. The goal of effective UI is to make the user's experience easy and intuitive, requiring minimum effort on the user's part to receive the maximum desired outcome.

Chapter 2

LITERATURE SURVEY

2.1 Title:- A Machine Learning Approach to Predict Autism Spectrum Disorder [1]

Abstract:- In present day Autism Spectrum Disorder (ASD) is gaining its momentum faster than ever. Detecting autism traits through screening tests is very expensive and time consuming. With the advancement of artificial intelligence and machine learning (ML), autism can be predicted at quite early stage. Though number of studies have been carried out using different techniques, these studies didn't provide any definitive conclusion about predicting autism traits in terms of different age groups. Therefore this paper aims to propose an effective prediction model based on ML technique and to develop a mobile application for predicting ASD for people of any age. As outcomes of this research, an autism prediction model was developed by merging Random Forest-CART (Classification and Regression Trees) and Random Forest-ID3 (Iterative Dichotomiser 3) and also a mobile application was developed based on the proposed prediction model. The proposed model was evaluated with AQ-10 dataset and 250 real dataset collected from people with and without autistic traits. The evaluation results showed that the proposed prediction model provide better results in terms of accuracy, specificity, sensitivity, precision and false positive rate (FPR) for both kinds of datasets.

2.2 Title:- Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders [2]

Abstract:- Autism Spectrum Disorder (ASD) is a group of neurodevelopmental disabilities that are not curable but may be ameliorated by early interventions. We gathered early-detected ASD datasets relating to toddlers, children, adolescents and adults, and applied several feature transformation methods, including log, Z-score and sine functions to these datasets. Various classification techniques were then implemented with these transformed ASD datasets and assessed for their performance. We found SVM showed the best performance for the toddler dataset, while Adaboost gave the best results for the children dataset, Glmboost for the adolescent and Adaboost for the adult datasets.

The feature transformations resulting in the best classifications was sine function for toddler and Z-score for children and adolescent datasets. After these analysis, several feature selection techniques were used with these Z-score transformed datasets to identify the significant ASD risk factors for the toddler, child, adolescent and adult subjects. The results of these analytical approaches indicate that, when appropriately optimised, machine learning methods can provide good predictions of ASD status. This suggests that it may possible to apply these models for the detection of ASD in its early stages.

2.3 Title:- Autistic Spectrum Disorder Screening: Prediction with Machine Learning Models [3]

Abstract:- Autistic Spectrum Disorder (ASD) is a developmental disorder that can be observed in all age groups. This paper uses ASD screening dataset for analysis and prediction of probable cases in adults, children and adolescents. The dataset for each of the age groups are analyzed and inferences are drawn from them. Machine learning algorithms like Artificial Neural Networks (ANN), Random Forest, Logistic Regression, Decision Tree and Support Vector Machines (SVM) are used for prediction and comparison.

2.4 Title:- Estimating Autism Severity in Young Children From Speech Signals Using a Deep Neural Network [4]

Abstract:-Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that involves difficulties in social communication. Previous research has demonstrated that these difficulties are apparent in the way ASD children speak, indicating that it may be possible to estimate ASD severity using quantitative features of speech. Here, we extracted a variety of prosodic, acoustic, and conversational features from speech recordings of Hebrew speaking children who completed an Autism Diagnostic Observation Schedule (ADOS) assessment. Sixty features were extracted from the recordings of 72 children and 21 of the features were significantly correlated with the children's ADOS scores. Positive correlations were found with pitch variability and Zero Crossing Rate (ZCR), while negative correlations were found with the speed and number of vocal responses to the clinician, and the overall number of vocalizations.
Using these features, several Deep Neural Network (DNN) algorithms were built to estimate ADOS scores and compared their performance with Linear Regression and Support Vector

Regression (SVR) models. It was found that a Convolutional Neural Network (CNN) yielded the best results. This algorithm predicted ADOS scores with a mean RMSE of 4.65 and a mean correlation of 0.72 with the true ADOS scores when trained and tested on different sub-samples of the available data. Automated algorithms with the ability to predict ASD severity in a reliable and sensitive manner have the potential of revolutionizing early ASD identification, quantification of symptom severity, and assessment of treatment efficacy.

2.5 Title:- Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques [5]

Abstract:- In this paper there is an attempt to explore the possibility to use Naïve Bayes, Support Vector Machine, Logistic Regression, KNN, Neural Network and Convolutional Neural Network for predicting and analysis of ASD problems in a child, adolescents, and adults. The proposed techniques are evaluated on publicly available three different non-clinically ASD datasets. After applying various machine learning techniques and handling missing values, results strongly suggest that CNN based prediction models work better on all these datasets with higher accuracy of 99.53%, 98.30%, 96.88% for Autistic Spectrum Disorder Screening in Data for Adult, Children, and Adolescents respectively.

2.6 Title:- Smart autism — A mobile, interactive and integrated framework for screening and confirmation of autism [6]

Abstract:- Smart Autism is a cloud based, automated framework for autism screening and confirmation. In developing countries, due to lack of resources and expertise, autism is detected later than early ages which consequently delays timely intervention. Therefore a mobile, interactive and integrated framework is proposed to screen and confirm autism in different age group (0 to 17 years) with 3 layers of assessment process. Firstly, it screens by evaluating the responses of pictorial based screening questionnaire through mobile application. If autism is suspected, then in virtual assessment process, the child watches a video, its reaction is recorded and uploaded to the cloud for remote expert assessment. If autism is still suspected, then the child is referred to the nearest Autism Resource Center (ARC) for actual assessment. Analyzing these results, the integrated framework confirms autism automatically and reduce user's ARC

visit. It is expected that the proposed framework will bring changes in autism diagnosis process and create awareness.

2.7 Title:- Autism Barta — A smart device based automated autism screening tool for Bangladesh [7]

Abstract:- Autism is a neurodevelopmental disorder which is not fully curable. However, early intervention can improve the condition of the children which requires early detection of autism. For this purpose, screening tools have been immensely used in developed countries. Whereas in developing countries, people are not getting such benefits. In this paper, a new automated approach, Autism Barta is proposed to screen autism in children using smart devices. The application integrates the questions of Bengali version of M-CHAT screening tool with pictorial representation. Therefore, parents can easily understand the interactive questions and use it effectively. The app will automatically screen autism in children, inform the user, store the responses in an online database and will suggest nearby Autism resource center for confirmation and intervention. It is expected that, this system will help to identify and streamline autism for improving the condition in developing countries like Bangladesh.

2.8 Title:- Automatic detection and labeling of self-stimulatory behavioral patterns in children with Autism Spectrum Disorder [8]

Abstract:- An infrastructure to record, detect and label the behavioral patterns of children with Autism Spectrum Disorder (ASD) has been developed. The system incorporates 2 different sensor platforms which are wearable and static. The wearable system is based on accelerometer which detects behavioral patterns of a subject, while the static sensors are microphones and cameras which captures the sounds, images and videos of the subjects within a room. The video also provides ground truth for wearable sensor data analysis. The system labels the segment of video data upon detection of the autistic behavior. That is, it stores the time of the video when the activities are detected. Time-Frequency methods are used to extract features and Hidden Markov Model (HMM) are used for analyzing the accelerometer signal. Using these methods, it is able to achieve 91.5% of classification rate for behavioral patterns studied in this paper which is used to label and save data.

2.9 Title:- The Effect of Sound Manipulation to Know Response Rate in Autism Children Using FFT [9]

Abstract:- Autism children have an inability to read the expressions of others, have difficulty recognizing certain emotions and difficulty controlling their emotions. Music has many health benefits and the brain one of which music can trigger brain development. Kitaro instrumental music is included in the line of music that is able to bring someone in a quiet condition. The author takes the type of music therapy so that children can perceive through his hearing, then activated in his brain, then connected to the nerve centers associated with emotions, imagination, and tranquility. Research carried out in the center's autism and tested try to son autism instrument by listening to the song to know response children autism parameters are used in the frequency of the voice of, the amplitude and sound intensity. The results of the analysis determine response children autism the voice of it can be concluded that the lower a value sound intensity (dB) and they will be obtained is a positive emotion the increase of your autism and the higher the test scores of the sound intensity (dB) and they will be obtained is the increase in your negative autism emotions deepening sadness or fear.

2.10 Title:- Genetic variant analysis of boys with Autism: A pilot study on linking facial phenotype to genotype [10]

Abstract:- This work examines the validity of facial phenotypes as Autism Spectrum Disorders (ASD) biomarkers in boys with essential autism. A family-based association analysis framework is presented that uses previously identified facially-delineated (FD) clusters to examine relationship between FD clusters and known ASD genes. The hypothesis is that there are certain genetic variants, single nucleotide polymorphisms (SNP), specific to the FD clusters. Although statistical significance was not established, the results identified some candidate SNPs unique to each of the FD clusters that could indicate an underlying etiological difference. Further, recommendations are provided for larger-scale studies that could utilize the analysis framework presented.

2.11 Title:- A new computational intelligence approach to detect autistic features for autism screening [11]

Abstract:- Autism Spectrum Disorder (ASD) is one of the fastest growing developmental disability diagnosis. General practitioners (GPs) and family physicians are typically the first point of contact for patients or family members concerned with ASD traits observed in themselves or their family member. Unfortunately, some families and adult patients are unaware of ASD traits that may be exhibited and as a result do not seek out necessary diagnostic services or contact their GP. Therefore, providing a quick, accessible, and simple tool utilizing items related to ASD to these families may increase the likelihood they will seek professional assessment and is vital to the early detection and treatment of ASD. This study aims at identifying fewer, albeit influential, features in common ASD screening methods in order to achieve efficient screening as demands on evaluating the items' influences on ASD within existing tools is urgent. To achieve this aim, a computational intelligence method called Variable Analysis (Va) is proposed that considers feature-to-class correlations and reduces feature-to-feature correlations. The results of the Va have been verified using two machine learning algorithms by deriving automated classification systems with respect to specificity, sensitivity, positive predictive values (PPVs), negative predictive values (NPVs), and predictive accuracy. Experimental results using cases and controls related to items in three common screening methods, along with features related to individuals, have been analysed and compared with results obtained from other common filtering methods. The results exhibited that Va was able to derive fewer numbers of features from adult, adolescent, and child screening methods yet maintained competitive predictive accuracy, sensitivity, and specificity rates.

2.12 Title:- Machine Learning Based Automated Speech Dialog Analysis Of Autistic Children [12]

Abstract:- Children with autism spectrum disorder (ASD) have altered behaviors in communication, social interaction, and activity, out of which communication has been the most prominent disorder among many. Despite the recent technological advances, limited attention has been given to screening and diagnosing ASD by identifying the speech deficiencies (SD) of autistic children at early stages. This research focuses on bridging the gap in ASD screening by

developing an automated system to distinguish autistic traits through speech analysis. Data was collected from 40 participants for the initial analysis and recordings were obtained from 17 participants. We considered a three-stage processing system; first stage utilizes thresholding for silence detection and Vocal Activity Detection for vocal isolation, second stage adopts machine learning technique neural network with frequency domain representations in developing a reliant utterance classifier for the isolated vocals and stage three also adopts machine learning technique neural network in recognizing autistic traits in speech patterns of the classified utterances. The results are promising in identifying SD of autistic children with the utterance classifier having 78% accuracy and pattern recognition 72% accuracy.

2.13 Title:- Identification of Autism Based on SVM-RFE and Stacked Sparse Auto-Encoder [13]

Abstract:- In order to improve the classification accuracy of patients with autism based on the full Autism Brain Imaging Data Exchange dataset, a total of 501 subjects with autism and 553 subjects with typical control across 17 sites were involved in the study. Firstly, we applied the resting-state functional magnetic resonance imaging data to calculate the functional connectivity (FC) based on the automated anatomical labeling atlas with 116 brain regions. Secondly, we adopted the support vector machine-recursive feature elimination algorithm to select top 1000 features from the primitive FC features. Thirdly, we trained a stacked sparse auto-encoder with two hidden layers to extract the high-level latent and complicated features from the 1000 features. Finally, the optimal features obtained were fed into the softmax classifier. Experimental results demonstrate that the proposed classification algorithm is able to identify the autism with a state-of-the-art accuracy of 93.59% (sensitivity 92.52%, specificity 94.56%).

2.14 Title:- Identify autism spectrum disorder via dynamic filter and deep spatiotemporal feature extraction [14]

Abstract:- Early intervention and treatment are crucial for individuals with autism spectrum disorder (ASD). However, it is challenging to identify individuals with ASD at an early age, *i.e.* under 3 years old, due to the lack of an effective and objective identification method. The mainstream clinical diagnosis relies on long-term observation of children’s behaviors, which is time-consuming and expensive, and thus how to accurately and quickly distinguish children with ASD in early childhood has become a critical issue. In this paper, we propose an eye movement based model to identify children with ASD. Specifically, children are required to freely observe some images. At the same time, their eye movements are recorded to analyze. Both the observed image and eye movements are input into our model. The input data are processed by the embedding layer, dynamic filters and LSTM block, respectively. Eventually, the spatiotemporal features are extracted to identify the eye movements belonging to a child with ASD or a typically developed child. Experiments on the Saliency4ASD dataset demonstrate that the proposed model achieves state-of-the-art performance in identifying children with ASD.

2.15 Title:-Early detection of children with Autism Spectrum Disorder based on visual exploration of images [15]

Abstract:- Autism Spectrum Disorder is a developmental disorder characterized by a deficit in social behaviour and specific interactions such as reduced eye contact and body gestures. Recent advancements in software and hardware multimedia technologies provide the tools for early detecting the presence of this disorder. In this paper we present an approach based on the combined use of machine learning and eye tracking information. More specifically, features are extracted from image content and viewing behaviour, such as the presence of objects and fixations towards the centre of a scene. Those features are used to train a machine learning-based classifier. The obtained results show that the considered features allow to identify children affected by autism spectrum disorder and typically developing ones.

2.16 Title:-Diagnostic prediction of autism spectrum disorder using complex network measures in a machine learning framework [16]

Abstract:- Objective imaging-based biomarker discovery for psychiatric conditions is critical for accurate diagnosis and treatment. Using a machine learning framework, this work investigated the utility of brain's functional network topology (complex network features) extracted from functional magnetic resonance imaging (fMRI) functional connectivity (FC) as viable biomarker of autism spectrum disorder (ASD). To this end, we utilized resting-state fMRI data from the publicly available ABIDE dataset consisting of 432 ASD patients and 556

matched healthy controls. Upon standard pre-processing, 3D + time fMRI data were parcellated into 200 functionally homogenous regions, and whole-brain FC network using Pearson's correlation was obtained from corresponding regional mean time series. A battery of complex network features were computed from the FC network using graph theoretic techniques. Recursive-Cluster-Elimination Support Vector Machine algorithm was employed to compare the predictive performance of three independent feature sets, (i) FC, (ii) complex network measures, and (iii) both combined. The study found that FC could diagnose ASD with 67.3 % accuracy and graph measures with 64.5 % accuracy, while the combined feature set could diagnose with 70.1 % accuracy (all accuracies were significantly different, $p < 10-30$). The most discriminative imaging features were mainly from lateral temporal, occipital, precuneus (all reduced in ASD) and orbito-frontal (elevated in ASD) regions. We concluded that network topology (graph measures) carried some unique information about ASD pathology not available in bivariate connectivity (FC), and that using both together provided better prediction than using individual measures. Future prediction studies could incorporate multiple fMRI analysis strategies within their framework to achieve superior prediction performances.

2.17 Title:-Complex Network Measures in Autism Spectrum Disorders [17]

Abstract:- Recent studies have suggested abnormal brain network organization in subjects with Autism Spectrum Disorders (ASD). Here we applied spectral clustering algorithm, diverse centrality measures (betweenness (BC), clustering (CC), eigenvector (EC), and degree (DC)), and also the network entropy (NE) to identify brain sub-systems associated with ASD. We have found that BC increases in the following ASD clusters: in the somatomotor, default-mode, cerebellar, and fronto-parietal. On the other hand, CC, EC, and DC decrease in the somatomotor, default-mode, and cerebellar clusters. Additionally, NE decreases in ASD in the cerebellar cluster. These findings reinforce the hypothesis of under-connectivity in ASD and suggest that the difference in the network organization is more prominent in the cerebellar system. The cerebellar cluster presents reduced NE in ASD, which relates to a more regular organization of the networks. These results might be important to improve current understanding about the etiological processes and the development of potential tools supporting diagnosis and therapeutic interventions.

2.18 Title:-Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and K-nearest neighbor in children [18]

Abstract:- Autism Spectrum Disorder (ASD) negatively affects the whole life of people. The main indications of ASD are seen as lack of social interaction and communication, repetitive patterns of behavior, fixed interests and activities. It is very important that ASD is diagnosed at an early age. In this study, the classification method for ASD diagnosis was used in children aged 4-11 years. The Linear Discriminant Analysis (LDA) and The K-Nearest Neighbor (KNN) algorithms are used for classification. To test the algorithms, 30 percent of the data set was selected as test data and 70 percent as training data. As a result of the work done; In the LDA algorithm, the accuracy is 90.8%, whereas the accuracy of the KNN algorithm is 88.5%. For the LDA algorithm, sensitivity and specificity values are calculated as 0.9524 and .08667, respectively. For KNN algorithm, these values are calculated as 0.9762 and 0.80. F-measure values are calculated as 0.9091 for the LDA algorithm and 0.8913 for the KNN algorithm.

2.19 Title:-Cognitive Load Measurement in a Virtual Reality-Based Driving System for Autism Intervention [19]

Abstract:- Autism Spectrum Disorder (ASD) is a highly prevalent neurodevelopmental disorder with enormous individual and social cost. In this paper, a novel virtual reality (VR)-based driving system was introduced to teach driving skills to adolescents with ASD. This driving system is capable of gathering eye gaze, electroencephalography, and peripheral physiology data in addition to driving performance data. The objective of this paper is to fuse multimodal information to measure cognitive load during driving such that driving tasks can be individualized for optimal skill learning. Individualization of ASD intervention is an important criterion due to the spectrum nature of the disorder. Twenty adolescents with ASD participated in our study and the data collected were used for systematic feature extraction and classification of cognitive loads based on five well-known machine learning methods. Subsequently, three information fusion schemes-feature level fusion, decision level fusion and hybrid level fusion-were explored. Results indicate that multimodal information fusion can be used to measure

cognitive load with high accuracy. Such a mechanism is essential since it will allow individualization of driving skill training based on cognitive load, which will facilitate acceptance of this driving system for clinical use and eventual commercialization.

2.20 Title:-Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence [20]

Abstract:- According to a recent influential proposal, several phenotypic features of autism spectrum disorder (ASD) may be accounted for by differences in predictive skills between individuals with ASD and neurotypical individuals. In this systematic review, we describe results from 47 studies that have empirically tested this hypothesis. We assess the results based on two observable aspects of prediction: learning a pairing between an antecedent and a consequence and responding to an antecedent in a predictive manner. Taken together, these studies suggest distinct differences in both predictive learning and predictive response. Studies documenting differences in learning predictive pairings indicate challenges in detecting such relationships especially when predictive features of an antecedent have low salience or consistency, and studies showing differences in habituation and perceptual adaptation suggest low-level predictive processing differences in ASD. These challenges may account for the observed differences in the influence of predictive priors, in spontaneous predictive movement or gaze, and in social prediction. An important goal for future research will be to better define and constrain the broad domain-general hypothesis by testing multiple types of prediction within the same individuals. Additional promising avenues include studying prediction within naturalistic contexts and assessing the effect of prediction-based intervention on supporting functional outcomes for individuals with ASD.

Chapter 3

PROBLEM IDENTIFICATION

3.1 Problem Statement

Autism Spectrum Disorder (ASD) brings about certain challenges in social, behavioral, communication and emotional understanding in an individual regardless of their age. Also, adults with ASD are more likely to develop Alzheimer's disease. Therefore, this project aims at Early-Stage Detection of Autism Spectrum Disorder in any individual using a Machine Learning model.

3.2 Project Scope

Our proposed strategy focuses on a novel machine learning procedures for Autism spectrum disorder (ASD) classification and prediction, thus overcoming the existing problem. By utilizing Random Forest (RF), Support Vector Machine(SVM), AdaBooster algorithms we will make our model in order to increase the performance and accuracy. This method has multiple advantages such as - we don’t have to figure out the features ahead of time, it is more effective, fault tolerant and scalable. Different existing data mining procedures and its application were considered or explored. Utilization of machine learning algorithms was connected in various medical data sets. Machine learning strategies have diverse power in different medical data sets. Previously mentioned conventional machine learning techniques gave less exact outcome and results additionally shifts in light of the procedures has been utilized for the prediction.

Chapter 4
GOALS AND OBJECTIVES

4.1 Project Goals

The goal of this project is to develop an application to implement a machine learning based solution that can detect Autism Spectrum Disorder in an individual (male/female) during early stages for early diagnosis using efficient algorithms to obtain results with required accuracy.

4.2 Project Objectives

1. To carry out literature survey.
2. To obtain the dataset to be fed to the model.
3. To perform data-preprocessing on collected datasets.
4. To implement Machine Learning models.
5. To perform prediction with Comparative Analysis.

Chapter 5

SYSTEM REQUIREMENT SPECIFICATION

5.1 Software Requirements Analysis

A software requirements definition is an abstract description of the services, which the system should provide, and the constraints under which the system must operate. It should only specify only the external behavior of the system and is not concerned with system design characteristics. It is a solution, in a natural language plus diagrams, of what services the system is expected to provide and the constraints under which it must operate.

Software Requirements

- **Operating system** : Windows 7 / 8 / 10
- **Coding Language** : Python
- **Software** : Anaconda
- **IDE** : Jupyter Notebook

5.2 Hardware Requirements Analysis

Hardware Requirements Analysis is to define and analyze a complete set of functional, operational, performance, interface, quality factors, and design, criticality and test requirements.

Hardware Requirements

- **System** : Pentium IV 2.4 GHz.
- **Hard Disk** : 500 GB.
- **Ram** : 4 GB
- Any desktop / Laptop system with above configuration or higher level

Chapter 6

METHODOLOGY

6.1 Working-Flow of Application

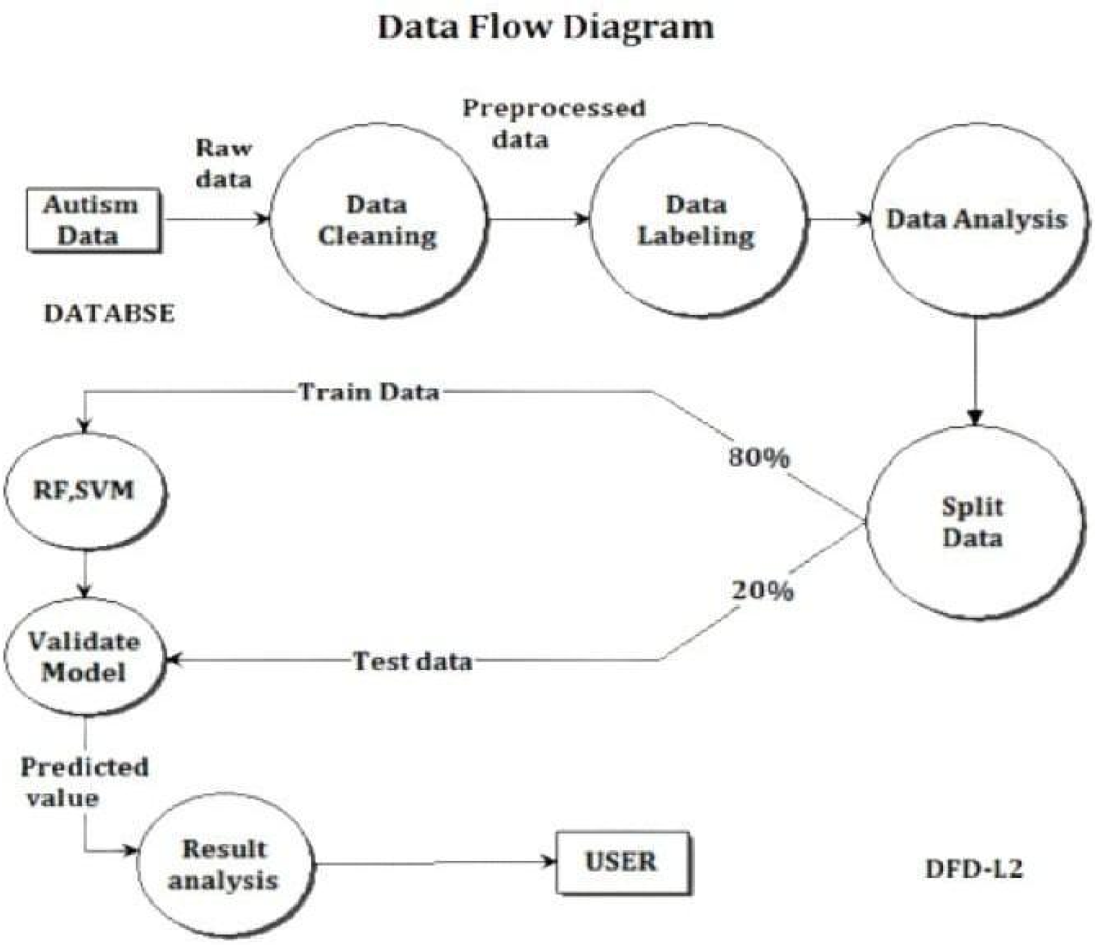

Fig 6.1 Flow chart of the application

The flowchart above shows the flow of our interface. Firstly the interface starts by feeding datasets on to the application. Then, data-preprocessing is performed on the dataset. Then machine learning model is implemented to provide prediction to the user.

6.2 Working

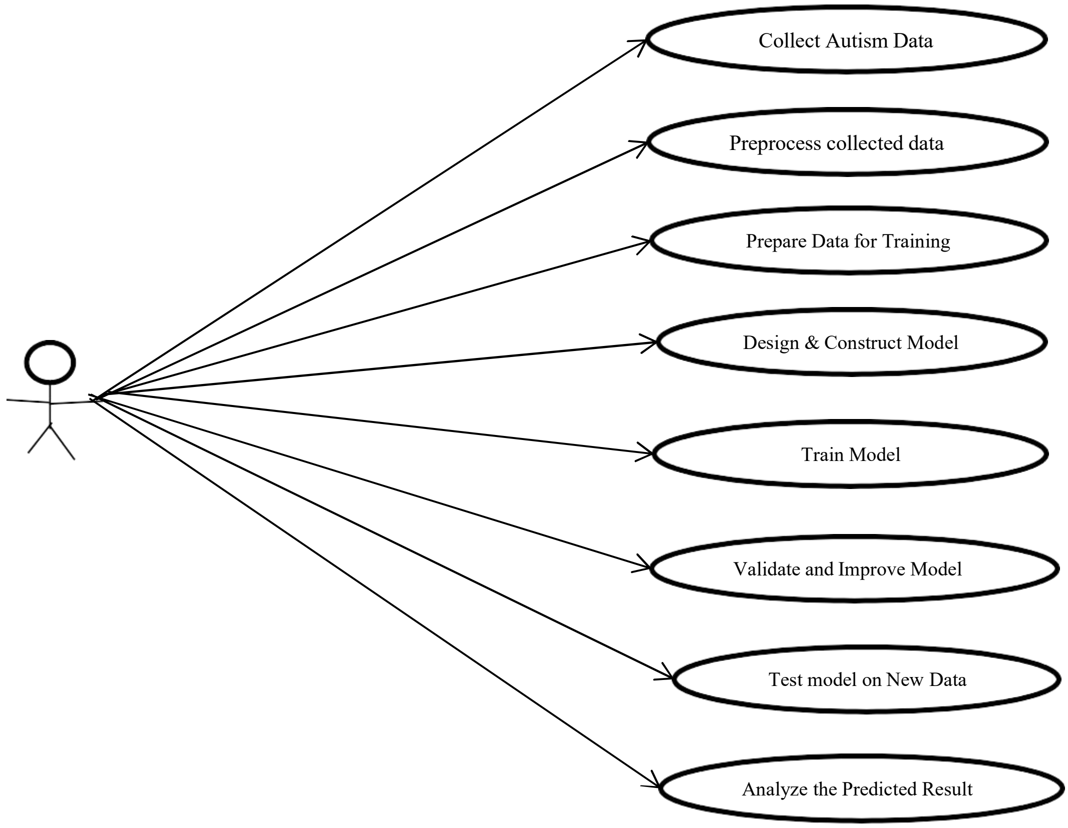

Fig 6.2 Block Diagram of the user-interface

Once user logs on to the interface, the user can feed the necessary datasets for the machine learning model to perform data-preprocessing, testing, validate and improve, predict and analyze. The predicted results will be displayed on the interface.

Chapter 7
IMPLEMENTATION

7.1 FILES USED

- **Jupyter Notebook:**
  - Toddler_ModelTraining.ipynb
  - Child_ModelTraining.ipynb

7.2 MODULES AND THEIR ROLES

7.2.1 PYTHON CODE FOR TODDLER DATASET

Import Libraries

```
import warnings
warnings.filterwarnings("always")
warnings.filterwarnings("ignore")

import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score,classification_report,confusion_matrix
import pickle
```

Data Loading

```
os.listdir()

df=pd.read_csv(filepath_or_buffer="Toddler.csv")

df.head()

print(df.shape)
```

Data-Preprocessing

```
print(df.columns)

***dropping unwanted columns***
```

```
df=df.drop(labels=["Case_No","Ethnicity"],axis=1)

df.head()

***renaming the columns***

df=df.rename(columns={"Qchat-10-Score":"Q_Score"})

df.head()

***Visualizing Age_Mons Column***

plt.figure(figsize=(15,7))
plt.rcParams["font.size"]=20
sns.countplot(x="Age_Mons",data=df)
plt.show()

***Visualizing Sex Column***

df["Sex"].value_counts()

plt.figure(figsize=(10,7))
plt.rcParams["font.size"]=20
plt.bar(x=["Male","Female"],height=df["Sex"].value_counts(),color=["Silver","green"])
plt.show()

df["Sex"].replace(to_replace=["m","f"],value=[0,1],inplace=True) df.head()

***Visualizing Jaundice***

df["Jaundice"].unique()

df["Jaundice"].value_counts()

plt.figure(figsize=(10,6))
plt.rcParams["font.size"]=20
plt.pie(x=df["Jaundice"].value_counts(),labels=["No","Yes"],colors=["deeppink","orange"],autopct=
"%.2f")
plt.show()

df["Jaundice"].replace(to_replace=["no","yes"],value=[0,1],inplace=True)

df.head()
```

Visualizing Family_mem_with_ASD column

```
df["Family_mem_with_ASD"].unique()

df["Family_mem_with_ASD"].value_counts()

plt.figure(figsize=(10,6))
plt.rcParams["font.size"]=20
plt.barh(y=["No","Yes"],width=df["Family_mem_with_ASD"].value_counts(),color=["gold","dimgr
ay"])
plt.show()

df["Family_mem_with_ASD"].replace(to_replace=["no","yes"],value=[0,1],inplace=True)

df.head()
```

visualizing Who completed the test column

```
df["Who completed the test"].unique()

df["Who completed the test"].replace(to_replace="Health care professional",value="Health Care
Professional",inplace=True)

df["Who completed the test"].unique()

df["Who completed the test"].value_counts()

df["Who completed the test"].replace(to_replace=["family member","Health Care
Professional","Self","Others"],value=[0,1,2,3],inplace=True)

df.head()
```

Analyzing and Visualizing Class column

```
df["Class"].unique()

df["Class"].value_counts()

plt.figure(figsize=(10,6))
plt.rcParams["font.size"]=20
sns.countplot(x="Class",data=df)
plt.show()

df["Class"].replace(to_replace=["No","Yes"],value=[0,1],inplace=True)

df.head()

df.isnull().sum()
```

```
df.info()

X=df.drop(labels="Class",axis=1)
y=df["Class"]

X.head()

y.head()

X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=42)

print(X_train.shape,X_test.shape,y_train.shape,y_test.shape)
```

Model-1 Support Vector Classifier

```
from sklearn.svm import SVC
svc_model=SVC(tol=3)
svc_model.fit(X_train,y_train)

svc_pred=svc_model.predict(X_test)

print(svc_pred)

Not_Autism=0
Autism=0
for predicted_label in svc_pred:
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"SupportVectorClassifier model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(svc_pred)} test samples.")
```

Result Analysis

Accuracy Score

```
svc_accuracy=accuracy_score(y_true=y_test,y_pred=svc_pred)
print("SupportVectorClassifier Accuracy is {:.2f}".format(svc_accuracy*100))
```

Classification Report

```
print(classification_report(y_true=y_test,y_pred=svc_pred,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
```

```
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_test,y_pred=svc_pred),
            cbar=False,
            annot=True,
            fmt="d",
            xticklabels=labels,
            yticklabels=labels,
            cmap="YlGnBu",
            linecolor="black",
            linewidths=0.2)
plt.show()
```

Model Saving

```
with open(file="SVM_model_toddler.pkl",mode="wb") as file:
    pickle.dump(obj=svc_model,file=file)
```

Model-2 RandomForestClassifier

```
from sklearn.ensemble import RandomForestClassifier
RF_model=RandomForestClassifier(n_estimators=1,min_samples_split=10,min_samples_leaf=10)
RF_model.fit(X=X_train,y=y_train)

RF_pred=RF_model.predict(X=X_test)

print(RF_pred)

Not_Autism=0
Autism=0
for predicted_label in RF_pred:
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"RandomForestClassifier model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(RF_pred)} test samples.")
```

Result Analysis

Accuracy score

```
RF_accuracy=accuracy_score(y_true=y_test,y_pred=RF_pred)
print("RandomForestClassifier accuracy is {:.2f}".format(RF_accuracy*100))
```

Classification report

```
print(classification_report(y_true=y_test,y_pred=RF_pred,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_test,y_pred=RF_pred),
        cbar=False,
        annot=True,
        fmt="d",
        xticklabels=labels,
        yticklabels=labels,
        cmap="YlGnBu",
        linecolor="black",
        linewidths=0.2)
plt.show()

with open(file="RF_model_toddler.pkl",mode="wb") as file:
   pickle.dump(obj=RF_model,file=file)
```

Model-3 ANN(Artificial Neural Network)

```
from tensorflow.keras import Sequential
from tensorflow.keras.layers import Dense,Flatten
from tensorflow.keras import optimizers

hidden_units=100
hidden_layer_act='tanh'
output_layer_act='sigmoid'
no_epochs=10

model = Sequential()
model.add(Dense(hidden_units, input_dim=16, activation=hidden_layer_act))
model.add(Dense(hidden_units, activation=hidden_layer_act))
model.add(Dense(1, activation=output_layer_act))

model.compile(loss='binary_crossentropy',optimizer="adam", metrics=['accuracy'])

model.summary()

history=model.fit(x=X_train,y=y_train,epochs=no_epochs,
batch_size=32,validation_data=(X_test,y_test))
```

Model Accuracy and Loss Plot-Graphs

```
plt.figure(figsize=(10,5))
plt.plot(history.history["accuracy"])
plt.plot(history.history["val_accuracy"])
plt.title(label="plot-graphs for accuracy and validated acccuracy")
```

```
plt.legend(["Accuracy","Val_accuracy"])
plt.show()

plt.figure(figsize=(10,5))
plt.plot(history.history["loss"])
plt.plot(history.history["val_loss"])
plt.title(label="plot-graphs for loss and validated loss")
plt.legend(["Loss","Val_Loss"])
plt.show()
```

Model Saving

```
model.save("model/Toddler/ANN_model_toddler.h5")

ann_pred=model.predict(x=X_test,batch_size=10,verbose=1)

print(ann_pred)

rounded = [int(round(x[0])) for x in ann_pred]
print(rounded)

y_true=list(y_test)

print(y_true)

Not_Autism=0
Autism=0
for predicted_label in rounded :
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"ArtificialNeuralNetwork model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(rounded )} test samples.")
```

Result Analysis

Accuracy Score

```
ann_accuracy=accuracy_score(y_true=y_true,y_pred=rounded)
print("ANN Accuracy is {:.2f}".format(ann_accuracy*100))
```

Accuracy Comparision for Toddler Dataset

```
data=[svc_accuracy,RF_accuracy,ann_accuracy]
labels=["SVC","RFC","ANN"]
colors=["gold","blue","lime"]
plt.figure(figsize=(10,5))
```

```
plt.grid()
plt.bar(x=labels,height=data,color=colors)
plt.title(label="models Accuracy Comparision for Toddler Dataset")
plt.show()
```

Comparative Analysis

```
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.3,random_state=42)

print(X_train.shape,X_test.shape,y_train.shape,y_test.shape)
```

Model-1 Support Vector Classifier

```
from sklearn.svm import SVC
svc_model=SVC(tol=3)
svc_model.fit(X_train,y_train)

svc_pred=svc_model.predict(X_test)

print(svc_pred)

Not_Autism=0
Autism=0
for predicted_label in svc_pred:
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"SupportVectorClassifier model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(svc_pred)} test samples.")
```

Result Analysis

Accuracy Score

```
svc_accuracy2=accuracy_score(y_true=y_test,y_pred=svc_pred)
print("SupportVectorClassifier Accuracy is {:.2f}".format(svc_accuracy*100))
```

Classification Report

```
print(classification_report(y_true=y_test,y_pred=svc_pred,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_test,y_pred=svc_pred),
```

```
     cbar=False,
      annot=True,
      fmt="d",
      xticklabels=labels,
      yticklabels=labels,
      cmap="YlGnBu",
      linecolor="black",
      linewidths=0.2)
plt.show()
```

Model Saving

```
with open(file="model/Toddler/SVM_model_toddler.pkl",mode="wb") as file:
   pickle.dump(obj=svc_model,file=file)
```

Model-2 RandomForestClassifier

```
from sklearn.ensemble import RandomForestClassifier
RF_model=RandomForestClassifier(n_estimators=1,min_samples_split=10,min_samples_leaf=10)
RF_model.fit(X=X_train,y=y_train)

RF_pred=RF_model.predict(X=X_test)

print(RF_pred)

Not_Autism=0
Autism=0
for predicted_label in RF_pred:
   if predicted_label==0:
      Not_Autism+=1
   else:
      Autism+=1
print(f"RandomForestClassifier model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(RF_pred)} test samples.")
```

Result Analysis

Accuracy score

```
RF_accuracy2=accuracy_score(y_true=y_test,y_pred=RF_pred)
print("RandomForestClassifier accuracy is {:.2f}".format(RF_accuracy*100))
```

Classification report

```
print(classification_report(y_true=y_test,y_pred=RF_pred,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_test,y_pred=RF_pred),
        cbar=False,
        annot=True,
        fmt="d",
        xticklabels=labels,
        yticklabels=labels,
        cmap="YlGnBu",
        linecolor="black",
        linewidths=0.2)
plt.show()

with open(file="model/Toddler/RF_model_toddler.pkl",mode="wb") as file:
    pickle.dump(obj=RF_model,file=file)
```

Model-3 ANN(Artificial Neural Network)

```
from tensorflow.keras import Sequential
from tensorflow.keras.layers import Dense,Flatten
from tensorflow.keras import optimizers

hidden_units=100
hidden_layer_act='tanh'
output_layer_act='sigmoid'
no_epochs=10

model = Sequential()
model.add(Dense(hidden_units, input_dim=16, activation=hidden_layer_act))
model.add(Dense(hidden_units, activation=hidden_layer_act))
model.add(Dense(1, activation=output_layer_act))

model.compile(loss='binary_crossentropy',optimizer="adam", metrics=['accuracy'])

model.summary()

history=model.fit(x=X_train,y=y_train,epochs=no_epochs,
batch_size=32,validation_data=(X_test,y_test))
```

Model Accuracy and Loss Plot-Graphs

```
plt.figure(figsize=(10,5))
plt.plot(history.history["accuracy"])
plt.plot(history.history["val_accuracy"])
plt.title(label="plot-graphs for accuracy and validated acccuracy")
plt.legend(["Accuracy","Val_accuracy"])
plt.show()
```

```
plt.figure(figsize=(10,5))
plt.plot(history.history["loss"])
plt.plot(history.history["val_loss"])
plt.title(label="plot-graphs for loss and validated loss")
plt.legend(["Loss","Val_Loss"])
plt.show()
```

Model Saving

```
model.save("model/Toddler/ANN_model_toddler.h5")

ann_pred=model.predict(x=X_test,batch_size=10,verbose=1)

print(ann_pred)

rounded = [int(round(x[0])) for x in ann_pred]
print(rounded)

y_true=list(y_test)

print(y_true)

Not_Autism=0
Autism=0
for predicted_label in rounded :
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"ArtificialNeuralNetwork model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(rounded )} test samples.")
```

Result Analysis

Accuracy Score

```
ann_accuracy2=accuracy_score(y_true=y_true,y_pred=rounded)
print("ANN Accuracy is {:.2f}".format(ann_accuracy*100))
```

Accuracy Comparision for Toddler Dataset

```
data=[svc_accuracy,RF_accuracy,ann_accuracy]
labels=["SVC","RFC","ANN"]
colors=["gold","blue","lime"]
plt.figure(figsize=(10,5))
plt.grid()
plt.bar(x=labels,height=data,color=colors)
plt.title(label="Accuracy Comparision for Toddler Dataset")
```

```
plt.show()
```

Accuracy Comparision for different test size of Toddler Dataset

```
data=[svc_accuracy,svc_accuracy2,RF_accuracy,RF_accuracy2,ann_accuracy,ann_accuracy2]
labels=["SVC","SVC2","RFC","RFC2","ANN","ANN2"]
colors=["gold","blue","lime"]
plt.figure(figsize=(10,5))
plt.grid()
plt.bar(x=labels,height=data,color=colors)
plt.title(label="Accuracy Comparision for Comparartive Analysis")
plt.show()

print("Middle aged adults with autism are 2.6 times more likely to be diagnosed with Alzheimer's disease")
```

7.2.2 PYTHON CODE FOR CHILD DATASET

Import Libraries

```
import warnings
warnings.filterwarnings("always")
warnings.filterwarnings("ignore")

import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score,classification_report,confusion_matrix
import pickle
pd.set_option("display.max_columns",None)
```

Data Loading

```
os.listdir()

df=pd.read_csv(filepath_or_buffer="Child.csv")

df.head()

print(df.shape)
```

Data-Preprocessing

```
print(df.columns)
```

dropping unwanted columns

```
df=df.drop(labels=["id","ethnicity","contry_of_res","used_app_before","austim","age_desc"],axis=1)

df.head()
```

renaming the columns

```
df=df.rename(columns={"Class/ASD":"Class"})

df.head()
```

Visualizing Age Column

```
plt.figure(figsize=(15,7))
plt.rcParams["font.size"]=20
sns.countplot(x="age",data=df)
plt.show()
```

Visualizing gender Column

```
df["gender"].value_counts()

plt.figure(figsize=(10,7))
plt.rcParams["font.size"]=20
plt.bar(x=["Male","Female"],height=df["gender"].value_counts(),color=["orange","red"])
plt.show()

df["gender"].replace(to_replace=["m","f"],value=[0,1],inplace=True)

df.head()
```

Visualizing jundice

```
df["jundice"].unique()

df["jundice"].value_counts()

plt.figure(figsize=(10,6))
plt.rcParams["font.size"]=20
plt.pie(x=df["jundice"].value_counts(),labels=["No","Yes"],colors=["darkgray","lightgray"],autopct
="%.2f")
plt.show()

df["jundice"].replace(to_replace=["no","yes"],value=[0,1],inplace=True)

df.head()
```

Visualizing result column

```
pos_result=df[df["result"]>=5]
neg_result=df[df["result"]<5]

print("positive results count is {}".format(len(pos_result)))
print("negative results count is {}".format(len(neg_result)))

plt.figure(figsize=(10,6))
plt.rcParams["font.size"]=20
plt.barh(y=["Positive","Negative"],width=[len(pos_result),len(neg_result)],color=["khaki","olive"])
plt.show()
```

visualizing Who completed the test column

```
df["relation"].unique()

df["relation"].value_counts()

df["relation"].replace(to_replace=["?","self"],value=["Parent","Self"],inplace=True)

df["relation"].value_counts()

df["relation"].replace(to_replace=["Parent","Relative","Health care
professional","Self"],value=[0,1,2,3],inplace=True)

df.head()

plt.figure(figsize=(10,6))
plt.rcParams["font.size"]=10
plt.bar(x=["Parent","Relative","Health care
professional","Self"],height=df["relation"].value_counts(),color=["firebrick","maroon","darkred","re
d"])
plt.show()
```

Analyzing and Visualizing Class column

```
df["Class"].unique()

df["Class"].value_counts()

plt.figure(figsize=(10,6))
plt.rcParams["font.size"]=20
sns.countplot(x="Class",data=df)
plt.show()

df["Class"].replace(to_replace=["NO","YES"],value=[0,1],inplace=True)

df.head()

df.isnull().sum()
```

```
df.info()

df["age"].unique()

df["age"].replace(to_replace="?",value="11",inplace=True)

df["age"].unique()

df.info()

df["age"]=df["age"].astype(dtype="int64")

df.info()

X=df.drop(labels="Class",axis=1)
y=df["Class"]

X.head()

y.head()

X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=42)

print(X_train.shape,X_test.shape,y_train.shape,y_test.shape)
```

Model-1 Support Vector Classifier

```
from sklearn.svm import SVC
svc_model=SVC(tol=50,kernel='linear')
svc_model.fit(X_train,y_train)

svc_pred=svc_model.predict(X_test)

print(svc_pred)

Not_Autism=0
Autism=0
for predicted_label in svc_pred:
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"SupportVectorClassifier model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(svc_pred)} test samples.")
```

Result Analysis

Accuracy Score

```
svc_accuracy=accuracy_score(y_true=y_test,y_pred=svc_pred)
print("SupportVectorClassifier Accuracy is {:.2f}".format(svc_accuracy*100))

***Classification Report***

print(classification_report(y_true=y_test,y_pred=svc_pred,target_names=["No","Yes"]))

***Confusion Matrix***

plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_test,y_pred=svc_pred),
            cbar=False,
            annot=True,
            fmt="d",
            xticklabels=labels,
            yticklabels=labels,
            cmap="YlGnBu",
            linecolor="black",
            linewidths=0.2)
plt.show()
```

Model Saving

```
with open(file="SVM_model_Child.pkl",mode="wb") as file:
    pickle.dump(obj=svc_model,file=file)
```

Model-2 RandomForestClassifier

```
from sklearn.ensemble import RandomForestClassifier
RF_model=RandomForestClassifier(n_estimators=1,min_samples_split=20,min_samples_leaf=20)
RF_model.fit(X=X_train,y=y_train)

RF_pred=RF_model.predict(X=X_test)

print(RF_pred)

Not_Autism=0
Autism=0
for predicted_label in RF_pred:
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"RandomForestClassifier model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(RF_pred)} test samples.")
```

Result Analysis

Accuracy Score

```
RF_accuracy=accuracy_score(y_true=y_test,y_pred=RF_pred)
print("RandomForestClassifier accuracy is {:.2f}".format(RF_accuracy*100))
```

Classification Report

```
print(classification_report(y_true=y_test,y_pred=RF_pred,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_test,y_pred=RF_pred),
        cbar=False,
        annot=True,
        fmt="d",
        xticklabels=labels,
        yticklabels=labels,
        cmap="YlGnBu",
        linecolor="black",
        linewidths=0.2)
plt.show()

with open(file="RF_model_Child.pkl",mode="wb") as file:
   pickle.dump(obj=RF_model,file=file)
```

Model-3 ANN(Artificial Neural Network)

```
from tensorflow.keras import Sequential
from tensorflow.keras.layers import Dense,Flatten
from tensorflow.keras import optimizers

hidden_units=100
hidden_layer_act='tanh'
output_layer_act='sigmoid'
no_epochs=20

model = Sequential()
model.add(Dense(hidden_units, input_dim=15, activation=hidden_layer_act))
model.add(Dense(hidden_units, activation=hidden_layer_act))
model.add(Dense(1, activation=output_layer_act))

model.compile(loss='binary_crossentropy',optimizer="adam", metrics=['accuracy'])
```

```
model.summary()

history=model.fit(x=X_train,y=y_train,epochs=no_epochs,
batch_size=128,validation_data=(X_test,y_test))
```

Model Accuracy and Loss Plot-Graphs

```
plt.figure(figsize=(10,5))
plt.plot(history.history["accuracy"])
plt.plot(history.history["val_accuracy"])
plt.title(label="plot-graphs for accuracy and validated acccuracy")
plt.legend(["Accuracy","Val_accuracy"])
plt.show()

plt.figure(figsize=(10,5))
plt.plot(history.history["loss"])
plt.plot(history.history["val_loss"])
plt.title(label="plot-graphs for loss and validated loss")
plt.legend(["Loss","Val_Loss"])
plt.show()
```

Model Saving

```
model.save("ANN_model_Child.h5")

ann_pred=model.predict(x=X_test,batch_size=10,verbose=1)

print(ann_pred)

rounded = [int(round(x[0])) for x in ann_pred]
print(rounded)

y_true=list(y_test)

print(y_true)

Not_Autism=0
Autism=0
for predicted_label in rounded :
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"ArtificialNeuralNetwork model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(rounded )} test samples.")
```

Result Analysis

Accuracy Score

```
ann_accuracy=accuracy_score(y_true=y_true,y_pred=rounded)
print("ANN Accuracy is {:.2f}".format(ann_accuracy*100))
```

Classification Report

```
print(classification_report(y_true=y_true,y_pred=rounded,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_true,y_pred=rounded),
        cbar=False,
        annot=True,
        fmt="d",
        xticklabels=labels,
        yticklabels=labels,
        cmap="YlGnBu",
        linecolor="black",
        linewidths=0.2)
plt.show()
```

Accuracy Comparision for Toddler Dataset

```
data=[svc_accuracy,RF_accuracy,ann_accuracy]
labels=["SVC","RFC","ANN"]
colors=["yellow","lawngreen","teal"]
plt.figure(figsize=(10,7))
plt.grid()
plt.bar(x=labels,height=data,color=colors)
plt.title(label="models Accuracy Comparision for Child Dataset")
plt.show()
```

Comparative Analysis

```
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.3,random_state=42)

print(X_train.shape,X_test.shape,y_train.shape,y_test.shape)
```

Model-1 Support Vector Classifier

```
from sklearn.svm import SVC
svc_model=SVC(tol=50,kernel='linear')
```

```
svc_model.fit(X_train,y_train)
svc_pred=svc_model.predict(X_test)

print(svc_pred)

Not_Autism=0
Autism=0
for predicted_label in svc_pred:
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"SupportVectorClassifier model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(svc_pred)} test samples.")
```

Result Analysis

Accuracy Score

```
svc_accuracy2=accuracy_score(y_true=y_test,y_pred=svc_pred)
print("SupportVectorClassifier Accuracy is {:.2f}".format(svc_accuracy*100))
```

Classification Report

```
print(classification_report(y_true=y_test,y_pred=svc_pred,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_test,y_pred=svc_pred),
            cbar=False,
            annot=True,
            fmt="d",
            xticklabels=labels,
            yticklabels=labels,
            cmap="YlGnBu",
            linecolor="black",
            linewidths=0.2)
plt.show()
```

Model Saving

```
with open(file="SVM_model_Child.pkl",mode="wb") as file:
    pickle.dump(obj=svc_model,file=file)
```

Model-2 RandomForestClassifier

```
from sklearn.ensemble import RandomForestClassifier
RF_model=RandomForestClassifier(n_estimators=1,min_samples_split=20,min_samples_leaf=20)
RF_model.fit(X=X_train,y=y_train)

RF_pred=RF_model.predict(X=X_test)

print(RF_pred)

Not_Autism=0
Autism=0
for predicted_label in RF_pred:
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"RandomForestClassifier model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(RF_pred)} test samples.")
```

Result Analysis

Accuracy Score

```
RF_accuracy2=accuracy_score(y_true=y_test,y_pred=RF_pred)
print("RandomForestClassifier accuracy is {:.2f}".format(RF_accuracy*100))
```

Classification Report

```
print(classification_report(y_true=y_test,y_pred=RF_pred,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_test,y_pred=RF_pred),
            cbar=False,
            annot=True,
            fmt="d",
            xticklabels=labels,
            yticklabels=labels,
            cmap="YlGnBu",
            linecolor="black",
            linewidths=0.2)
plt.show()

with open(file="RF_model_Child.pkl",mode="wb") as file:
```

```
  pickle.dump(obj=RF_model,file=file)
```

Model-3 ANN(Artificial Neural Network)

```
from tensorflow.keras import Sequential
from tensorflow.keras.layers import Dense,Flatten
from tensorflow.keras import optimizers

hidden_units=100
hidden_layer_act='tanh'
output_layer_act='sigmoid'
no_epochs=20

model = Sequential()
model.add(Dense(hidden_units, input_dim=15, activation=hidden_layer_act))
model.add(Dense(hidden_units, activation=hidden_layer_act))

model.add(Dense(1, activation=output_layer_act))

model.compile(loss='binary_crossentropy',optimizer="adam", metrics=['accuracy'])

model.summary()

history=model.fit(x=X_train,y=y_train,epochs=no_epochs,
batch_size=128,validation_data=(X_test,y_test))
```

Model Accuracy and Loss Plot-Graphs

```
plt.figure(figsize=(10,5))
plt.plot(history.history["accuracy"])
plt.plot(history.history["val_accuracy"])
plt.title(label="plot-graphs for accuracy and validated acccuracy")
plt.legend(["Accuracy","Val_accuracy"])
plt.show()

plt.figure(figsize=(10,5))
plt.plot(history.history["loss"])
plt.plot(history.history["val_loss"])
plt.title(label="plot-graphs for loss and validated loss")
plt.legend(["Loss","Val_Loss"])
plt.show()
```

Model Saving

```
model.save("ANN_model_Child.h5")

ann_pred=model.predict(x=X_test,batch_size=10,verbose=1)
```

```
print(ann_pred)

rounded = [int(round(x[0])) for x in ann_pred]
print(rounded)

y_true=list(y_test)

print(y_true)

Not_Autism=0
Autism=0
for predicted_label in rounded :
    if predicted_label==0:
        Not_Autism+=1
    else:
        Autism+=1
print(f"ArtificialNeuralNetwork model predicted {Autism} samples as AUTISM and {Not_Autism}
samples as NOT_AUTISM from {len(rounded )} test samples.")
```

Result Analysis

Accuracy Score

```
ann_accuracy2=accuracy_score(y_true=y_true,y_pred=rounded)
print("ANN Accuracy is {:.2f}".format(ann_accuracy*100))
```

Classification Report

```
print(classification_report(y_true=y_true,y_pred=rounded,target_names=["No","Yes"]))
```

Confusion Matrix

```
plt.figure(figsize=(5,5))
plt.rcParams["font.size"]=20
labels=["No","Yes"]
sns.heatmap(data=confusion_matrix(y_true=y_true,y_pred=rounded),
        cbar=False,
        annot=True,
        fmt="d",
        xticklabels=labels,
        yticklabels=labels,
        cmap="YlGnBu",
        linecolor="black",
        linewidths=0.2)
plt.show()
```

Accuracy Comparision for Child Dataset

```
data=[svc_accuracy,RF_accuracy,ann_accuracy]
labels=["SVC","RFC","ANN"]
colors=["yellow","lawngreen","teal"]
plt.figure(figsize=(10,7))
plt.grid()
plt.bar(x=labels,height=data,color=colors)
plt.title(label="Accuracy Comparision for Child Dataset")
plt.show()
```

Accuracy Comparision for different test size of Child Dataset

```
data=[svc_accuracy,svc_accuracy2,RF_accuracy,RF_accuracy2,ann_accuracy,ann_accuracy2]
labels=["SVC","SVC2","RFC","RFC2","ANN","ANN2"]
colors=["gold","blue","lime"]
plt.figure(figsize=(10,5))
plt.grid()
plt.bar(x=labels,height=data,color=colors)
plt.title(label="Accuracy Comparision for Comparartive Analysis")
plt.show()

print("Middle aged adults with autism are 2.6 times more likely to be diagnosed with Alzheimer's disease")
```

Chapter 8

RESULTS AND SNAPSHOTS

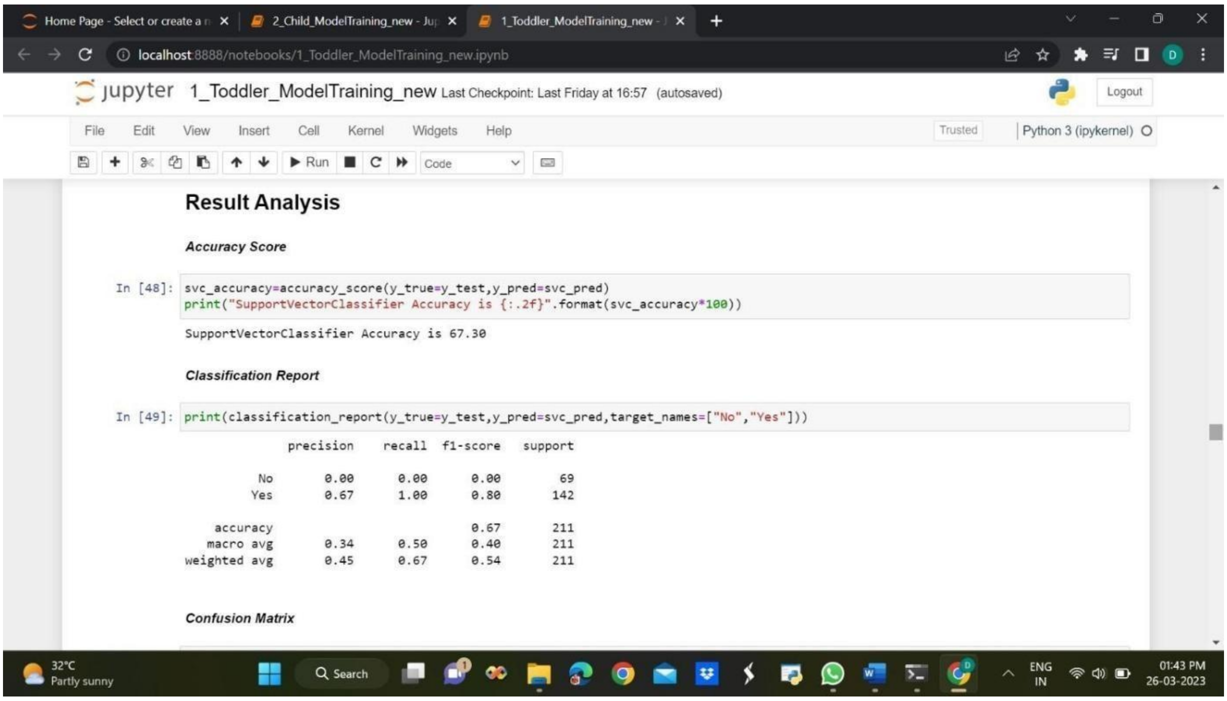

Fig 8.1: Accuracy results of Support Vector Classifier model for Toddler Dataset

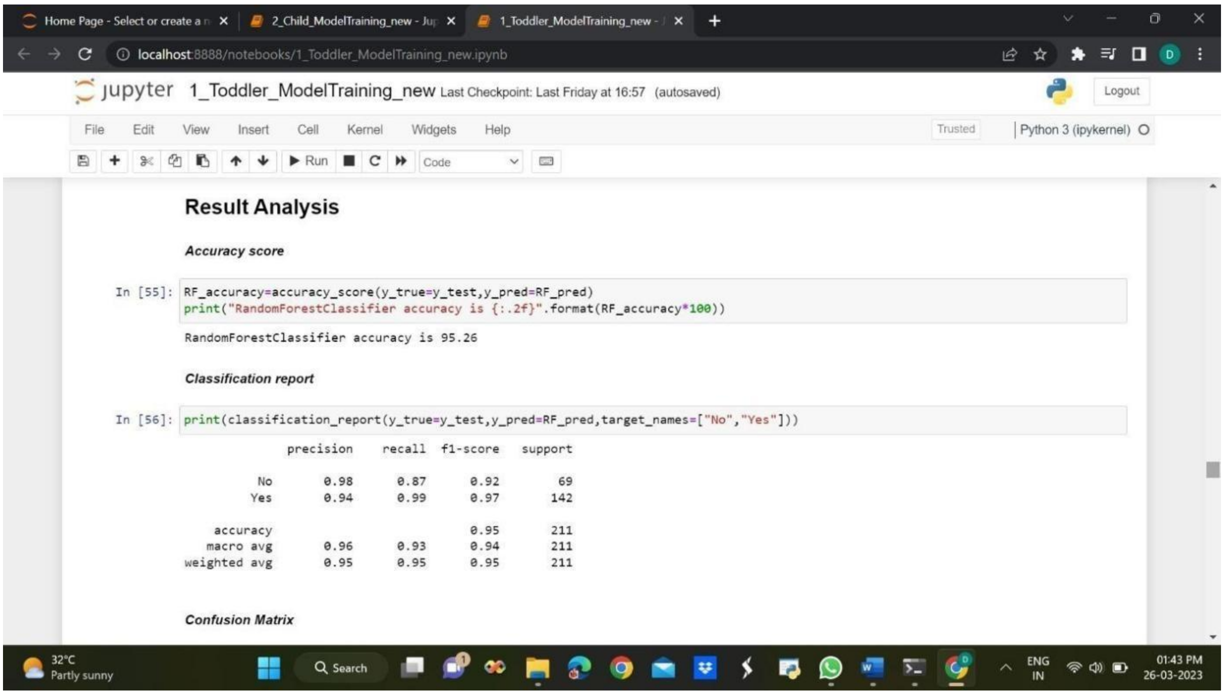

Fig 8.2: Accuracy results of Random Forest model based on Toddler Dataset

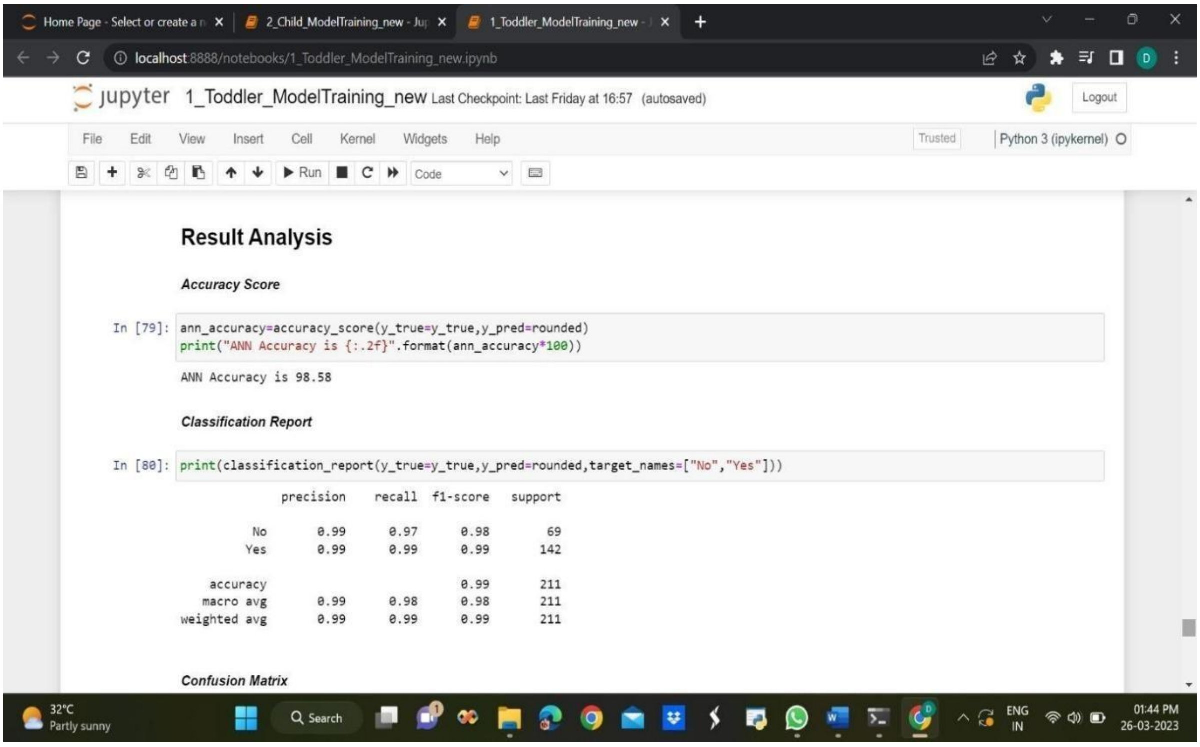

Fig 8.3: Accuracy results of ANN model for Toddler Dataset

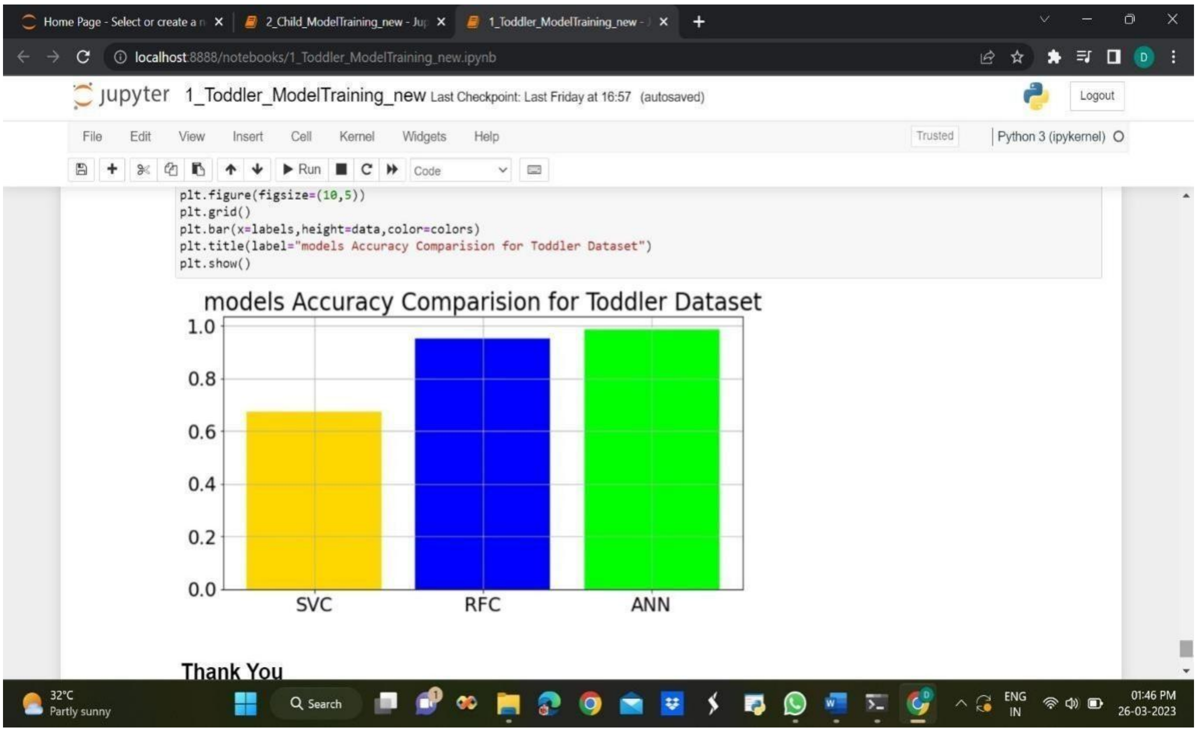

Fig 8.4: Models Accuracy comparison for Toddler Dataset

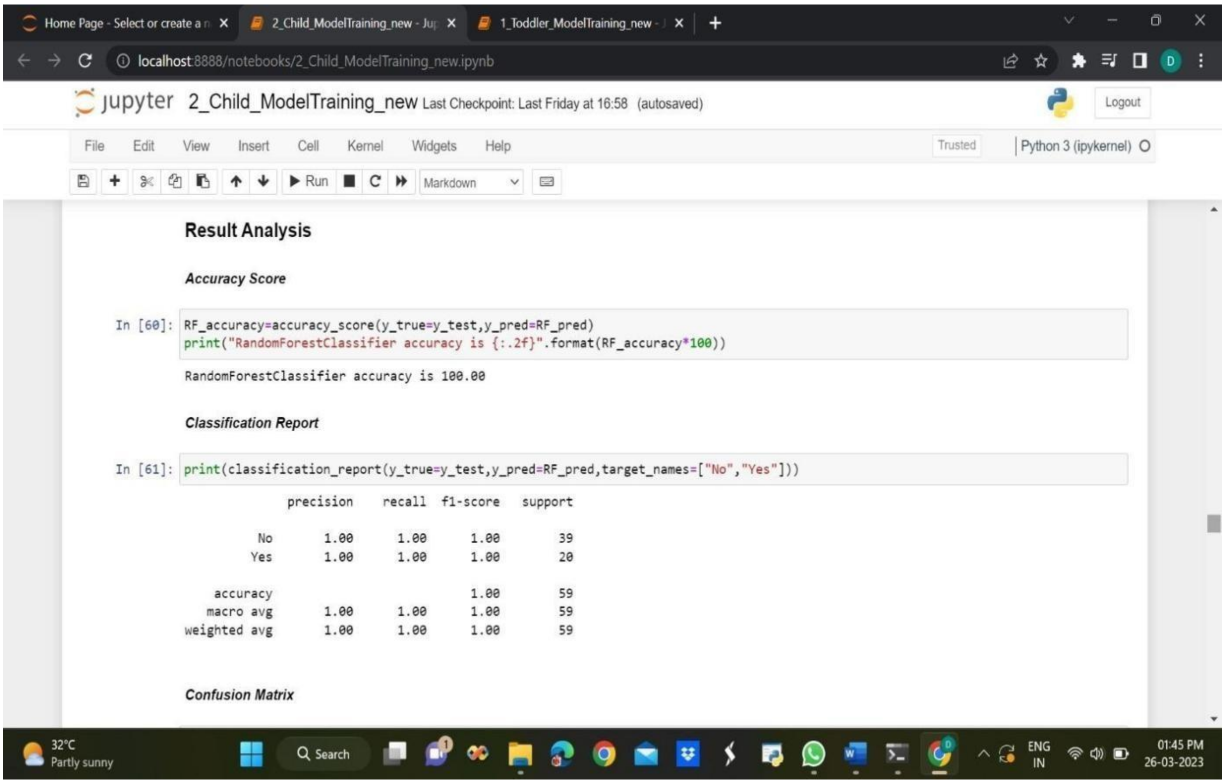

Fig 8.5: Accuracy results of Random Forest model on Child Dataset

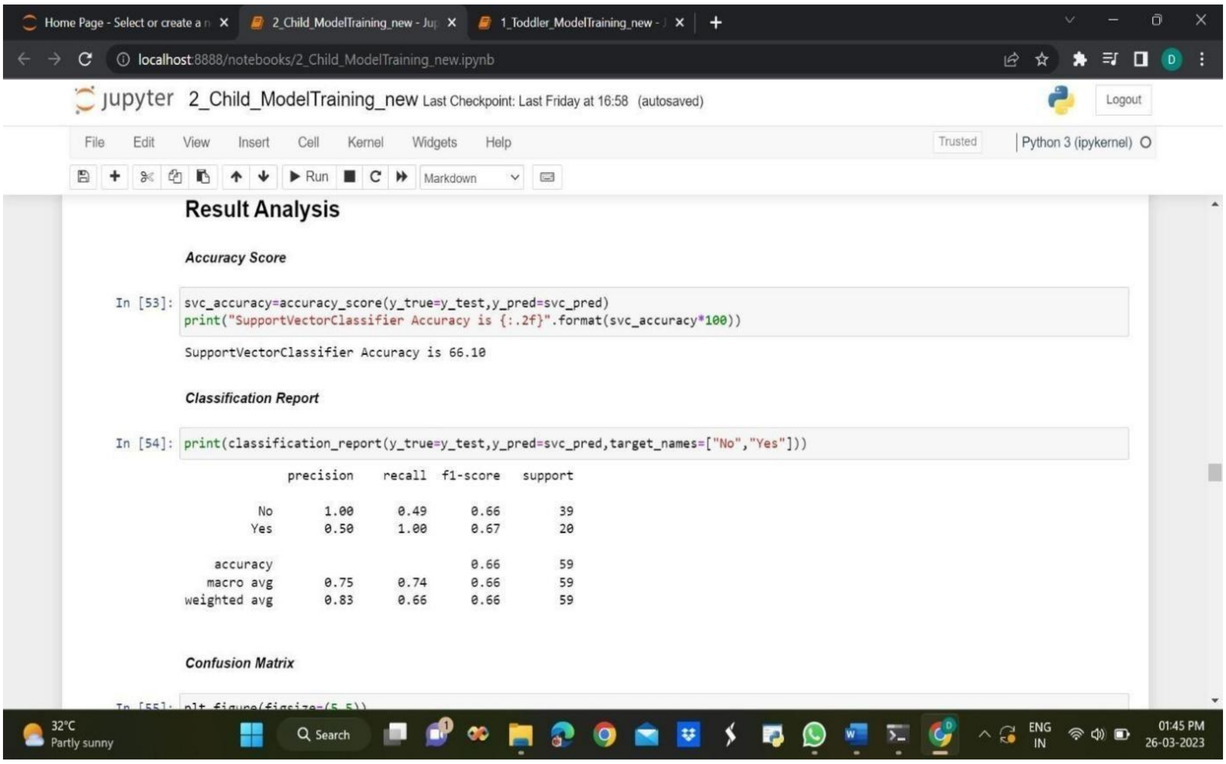

Fig 8.6: Accuracy results of Support Vector Classifier model on Child Dataset

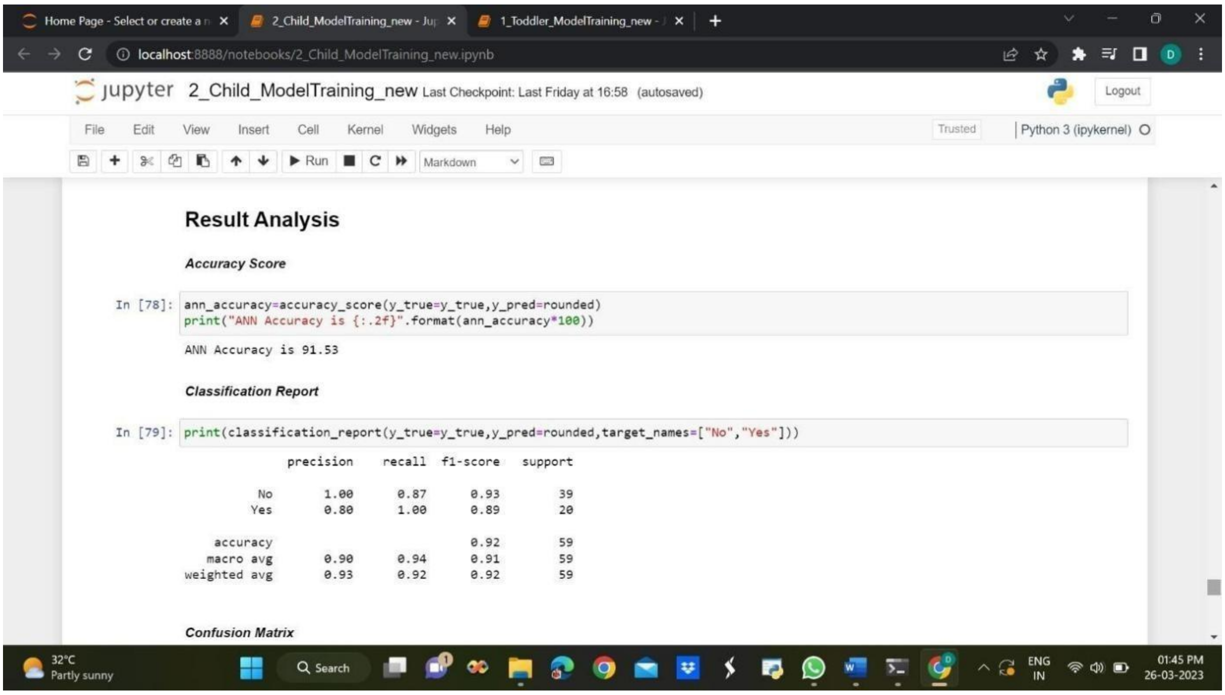

Fig 8.7: Accuracy Results of ANN model on child Dataset

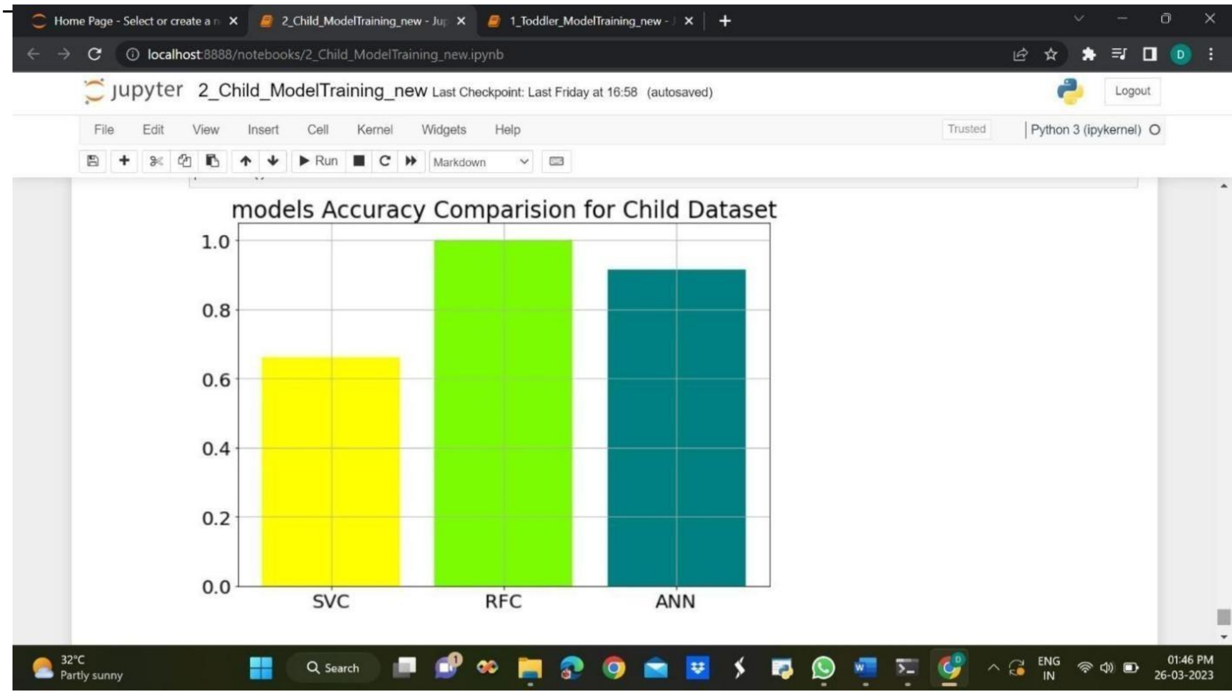

Fig 8.8: Models Accuracy Comparison for Child Dataset

Chapter 9

APPLICATIONS

The objective of this work is to propose an autism prediction model using ML techniques and to develop a mobile application that could effectively predict autism traits of an individual of any age. In other words, this work focuses on developing an autism screening application for predicting the ASD traits among people of age groups 4-11 years, 12-17 years and for people of age 18 and more.

- Can be used for medical and health care
- Can be used to provide cost-effective diagnosis.
- Can be used to provide pediatric care.
- Can be used to provide early detections.
- Can be used to spread psychological awareness.

Chapter 10

CONTRIBUTION TO SOCIETY AND ENVIRONMENT

Current explosion rate of autism around the world is numerous and it is increasing at a very high rate. According to WHO, about 1 out of every 160 children has ASD. Some people with this disorder can live independently, while others require life-long care and support.

Diagnosis of autism requires significant amount of time and cost. Earlier detection of autism can come to a great help by prescribing patients with proper medication at an early stage.It can prevent the patient’s condition from deteriorating further and would help to reduce long term costs associated with delayed diagnosis. Thus a time efficient, accurate and easy screening test tool is very much required which would predict autism traits in an individual and identify whether or not they require comprehensive autism assessment.

The objective of this work is to propose an autism prediction model using ML techniques and to develop a user-interface that could effectively predict autism traits of an individual of any age. In other words, this work focuses on developing an autism screening application for predicting the ASD traits among people of age groups 4-11 years, 12-17 years and for people of age 18 and more.

REFERENCES

[1] Omar, Kazi Shahrukh, Prodipta Mondal, Nabila Shahnaz Khan, Md Rezaul Karim Rizvi, and Md Nazrul Islam. "A machine learning approach to predict autism spectrum disorder." In 2019 International conference on electrical, computer and communication engineering (ECCE), pp. 1-6. IEEE, 2019.

[2] Akter, Tania, Md Shahriare Satu, Md Imran Khan, Mohammad Hanif Ali, Shahadat Uddin, Pietro Lio, Julian MW Quinn, and Mohammad Ali Moni. "Machine learning-based models for early stage detection of autism spectrum disorders." IEEE Access 7 (2019): 166509-166527.

[3] Baranwal, Astha, and M. Vanitha. "Autistic spectrum disorder screening: prediction with machine learning models." In 2020 International conference on emerging trends in information technology and engineering (ic-ETITE), pp. 1-7. IEEE, 2020.

[4] Eni, M, Dinstein, I, Ilan, M., Menashe, I, Meiri, G, & Zigel, Y. (2020). Estimating Autism Severity in Young Children From Speech Signals Using a Deep Neural Network. IEEE Access, 8, 139489–139500. doi:10.1109/access.2020.301253

[5] Raj, S, & Masood, S. (2020). Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques. Procedia Computer Science, 167, 994-1004.doi:10.1016/j.procs.2020.03.39

[6] Mamun, K. A. A., Bardhan, S., Ullah, M. A., Anagnostou, E., Brian, J., Akhter, S., & Rabbani, M. G. (2016). Smart autism — A mobile, interactive and integrated framework for screening and confirmation of autism. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). doi:10.1109/embc.2016.7592093

[7] Bardhan, S., Mridha, G. M. M. M., Ahmed, E., Ullah, M. A., Ahmed, H. U., Akhter, S., … Mamun, K. A. A. (2016). Autism Barta — A smart device based automated autism screening tool for Bangladesh. 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV). doi:10.1109/iciev.2016.7760073

[8] Min, C.-H. (2017). Automatic detection and labeling of self-stimulatory behavioral patterns in children with Autism Spectrum Disorder. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). doi:10.1109/embc.2017.8036816

[9] Kirana, M. C., Lubis, M. Z., & Amalia, E. D. (2018). The Effect of Sound Manipulation to Know Response Rate in Autism Children Using FFT. 2018 International Conference on Applied Engineering(ICAE).doi:10.1109/incae.2018.8579418

[10] Obafemi-Ajayi, T., Settles, L., Su, Y., Germeroth, C., Olbricht, G. R., Wunsch, D. C., … Miles, J. (2017). Genetic variant analysis of boys with Autism: A pilot study on linking facial phenotype to genotype. 2017 IEEE International Conference on Bioinformatics and Biomedicine(BIBM).doi:10.1109/bibm.2017.8217658

[11] Thabtah, F., Kamalov, F., & Rajab, K. (2018). A new computational intelligence approach to detect autistic features for autism screening. International Journal of Medical Informatics, 117, 112–124. doi:10.1016/j.ijmedinf.2018.06.009

[12] Liaqat, S., Wu, C., Duggirala, P. R., Cheung, S. S., Chuah, C.-N., Ozonoff, S., & Young, G. (2021). Predicting ASD diagnosis in children with synthetic and image-based eye gaze data. Signal Processing: Image Communication, 94, 116198. doi:10.1016/j.image.2021.116198

[13] Wijesinghe, A., Samarasinghe, P., Seneviratne, S., Yogarajah, P., & Pulasinghe, K. (2019). Machine Learning Based Automated Speech Dialog Analysis Of Autistic Children. 2019 11th International Conference on Knowledge and Systems Engineering (KSE). doi:10.1109/kse.2019.8919266

[14] Wang, C., Xiao, Z., Wang, B., & Wu, J. (2019). Identification of Autism Based on SVM-RFE and Stacked Sparse Auto-Encoder. IEEE Access, 7, 118030–118036. doi:10.1109/access.2019.2936639.

[15] Wei, W., Liu, Z., Huang, L., Wang, Z., Chen, W., Zhang, T., Xu, L. (2021). Identify autism spectrum disorder via dynamic filter and deep spatiotemporal feature extraction. Signal Processing: Image Communication, 94, 116195. doi:10.1016/j.image.2021.116195

[16] Mazumdar, P., Arru, G., & Battisti, F. (2021). Early detection of children with Autism Spectrum Disorder based on visual exploration of images. Signal Processing: Image Communication, 94, 116184. doi:10.1016/j.image.2021.116184

[17] Chaitra, N., Vijaya, P. A., & Deshpande, G. (2020). Diagnostic prediction of autism spectrum disorder using complex network measures in a machine learning framework. Biomedical Signal Processing and Control, 62, 102099. doi:10.1016/j.bspc.2020.102099

[18] Sato, J. R., Calebe Vidal, M., de Siqueira Santos, S., Brauer Massirer, K., & Fujita, A. (2018). Complex Network Measures in Autism Spectrum Disorders. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 15(2), 581–587. doi:10.1109/tcbb.2015.2476787

[19] Altay, O., & Ulas, M. (2018). Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and K-nearest neighbor in children. 2018 6th International Symposium on Digital Forensic and Security (ISDFS). doi:10.1109/isdfs.2018.8355354

[20] Abraham, A., Milham, M. P., Di Martino, A., Craddock, R. C., Samaras, D., Thirion, B., & Varoquaux, G. (2017). Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. NeuroImage, 147, 736–745. doi:10.1016/j.neuroimage.2016.10

[21] Milne, M., Powers, D., & Leibbrandt, R. (2009). Development of a software-based social tutor for children with autism spectrum disorders. Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group on Design: Open 24/7 - OZCHI '09. doi:10.1145/1738826.1738870

[22] Zhang, L., Wade, J., Bian, D., Fan, J., Swanson, A., Weitlauf, A., … Sarkar, N. (2017). Cognitive Load Measurement in a Virtual Reality-Based Driving System for Autism Intervention. IEEE Transactions on Affective Computing, 8(2), 176–189. doi:10.1109/taffc.2016.2582490

[23] Kundu, M. R., & Suranjan Das, M. (2019). Predicting Autism Spectrum Disorder in Infants Using Machine Learning. Journal of Physics: Conference Series, 1362, 012018. doi:10.1088/1742-6596/1362/1/012018

[24] Cannon, J., O'Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence. Autism Research, 14(4), 604–630. doi:10.1002/aur.2482

[25] Skafidas, E., Testa, R., Zantomio, D., Chana, G., Everall, I. P., & Pantelis, C. (2012). Predicting the diagnosis of autism spectrum disorder using gene pathway analysis. Molecular Psychiatry, 19(4), 504–510. doi:10.1038/mp.2012.126

[26] Howsmon, D. P., Kruger, U., Melnyk, S., James, S. J., & Hahn, J. (2017). Classification and adaptive behavior prediction of children with autism spectrum disorder based upon multivariate data analysis of markers of oxidative stress and DNA methylation. PLOS Computational Biology, 13(3), e1005385. doi:10.1371/journal.pcbi.1005385

[27] Liu, W., Yu, X., Raj, B., Yi, L., Zou, X., & Li, M. (2015). Efficient autism spectrum disorder prediction with eye movement: A machine learning framework. 2015 International Conference on Affective Computing and Intelligent Interaction (ACII). doi:10.1109/acii.2015.7344638

[28] Gök, M. (2018). A novel machine learning model to predict autism spectrum disorders risk gene. Neural Computing and Applications. doi:10.1007/s00521-018-3502-5

[29] Oh, D. H., Kim, I. B., Kim, S. H., & Ahn, D. H. (2017). Predicting Autism Spectrum Disorder Using Blood-based Gene Expression Signatures and Machine Learning. Clinical Psychopharmacology and Neuroscience, 15(1), 47–52. doi:10.9758/cpn.2017.15.1.47

[30] Xu, G., Strathearn, L., Liu, B., & Bao, W. (2018). Prevalence of Autism Spectrum Disorder Among US Children and Adolescents, 2014-2016. JAMA, 319(1), 81. doi:10.1001/jama.2017.17812

[30] Xu, G., Strathearn, L., Liu, B., & Bao, W. (2018). Prevalence of Autism Spectrum Disorder Among US Children and Adolescents, 2014-2016. JAMA, 319(1), 81. doi:10.1001/jama.2017.17812

[31] Kosmicki, J. A., Sochat, V., Duda, M., & Wall, D. P. (2015). Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. Translational Psychiatry, 5(2), e514–e514. doi:10.1038/tp.2015.7

[32] Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., … Newschaffer, C. (2017). The Changing Epidemiology of Autism Spectrum Disorders. Annual Review of Public Health, 38(1), 81–102. doi:10.1146/annurev-publhealth-031816-044318

[33] Ejlskov, L., Wulff, J. N., Kalkbrenner, A., Ladd-Acosta, C., Fallin, M. D., Agerbo, E., … Schendel, D. (2021). Prediction of Autism Risk From Family Medical History Data Using Machine Learning: A National Cohort Study From Denmark. Biological Psychiatry Global Open Science, 1(2), 156–164. doi:10.1016/j.bpsgos.2021.04.007

[34] Bar-Haim, Y., Shulman, C., Lamy, D., & Reuveni, A. (2006). Attention to Eyes and Mouth in High-Functioning Children with Autism. Journal of Autism and Developmental Disorders, 36(1), 131–137. doi:10.1007/s10803-005-0046-1

[35] Brown, C. J., Kawahara, J., & Hamarneh, G. (2018). Connectome priors in deep neural networks to predict autism. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). doi:10.1109/isbi.2018.8363534

[36] Åsberg Johnels, J., Hovey, D., Zürcher, N., Hippolyte, L., Lemonnier, E., Gillberg, C., & Hadjikhani, N. (2016). Autism and emotional face-viewing. Autism Research, 10(5), 901–910. doi:10.1002/aur.1730

[37] Åsberg Johnels, J., Hovey, D., Zürcher, N., Hippolyte, L., Lemonnier, E., Gillberg, C., & Hadjikhani, N. (2016). Autism and emotional face-viewing. Autism Research, 10(5), 901–910. doi:10.1002/aur.1730

[38] Åsberg Johnels, J., Hovey, D., Zürcher, N., Hippolyte, L., Lemonnier, E., Gillberg, C., & Hadjikhani, N. (2016). Autism and emotional face-viewing. Autism Research, 10(5), 901–910. doi:10.1002/aur.1730

[39] Duan, H., Min, X., Fang, Y., Fan, L., Yang, X., & Zhai, G. (2019). Visual Attention Analysis and Prediction on Human Faces for Children with Autism Spectrum Disorder. ACM Transactions on Multimedia Computing, Communications, and Applications, 15(3s), 1–23. doi:10.1145/3337066

[40] Eman, D., & Emanuel, A. W. . (2019). Machine Learning Classifiers for Autism Spectrum Disorder: A Review. 2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering

[39] Stevens, E., Dixon, D. R., Novack, M. N., Granpeesheh, D., Smith, T., & Linstead, E. (2019). Identification and analysis of behavioral phenotypes in autism spectrum disorder via unsupervised machine learning. International Journal of Medical Informatics, 129, 29–36. doi:10.1016/j.ijmedinf.2019.05.006

[40] Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., & Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. NeuroImage:Clinical, 17, 16–23. doi:10.1016/j.nicl.2017.08.017

[41] Ghosh, T., Banna, M. H. A., Rahman, M. S., Kaiser, M. S., Mahmud, M., Hosen, A. S. M. S., & Cho, G. H. (2021). Artificial intelligence and internet of things in screening and management of autism spectrum disorder. Sustainable Cities and Society,74,103189. doi:10.1016/j.scs.2021.103189

[42] Bussu, G., Jones, E. J. H., Charman, T., Johnson, M. H., & Buitelaar, J. K. (2018). Prediction of Autism at 3 Years from Behavioural and Developmental Measures in High-Risk Infants: A Longitudinal Cross-Domain Classifier Analysis. Journal of Autism and Developmental Disorders, 48(7), 2418–2433. doi:10.1007/s10803-018-3509-x

[43] Chowdhury, K., & Iraj, M. A. (2020). Predicting Autism Spectrum Disorder Using Machine Learning Classifiers. 2020 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT). doi:10.1109/rteict49044.2020.9315

[44] Selvaraj, S., Palanisamy, P., Parveen, S., & Monisha. (2020). Autism Spectrum Disorder Prediction Using Machine Learning Algorithms. Advances in Intelligent Systems and Computing, 496–503. doi:10.1007/978-3-030-37218-7_56

[45] Sharma, A., & Tanwar, P. (2020). Deep Analysis of Autism Spectrum Disorder Detection Techniques. 2020 International Conference on Intelligent Engineering and Management (ICIEM). doi:10.1109/iciem48762.2020.9160123
10.1109/ICIEM48762.2020.9160123

[46] Nor, M. N. M., Jailani, R., & Tahir, N. M. (2016). Analysis of EMG signals of TA and GAS muscles during walking of Autism Spectrum Disorder (ASD) children. 2016 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). doi:10.1109/iscaie.2016.7575068

[47] Thabtah, Fadi; Peebles, David (2019). A new machine learning model based on induction of rules for autism detection. Health Informatics Journal, (), 146045821882471–. doi:10.1177/1460458218824711

[48] Thabtah, Fadi; Peebles, David (2019). Early Autism Screening: A Comprehensive Review. International Journal of Environmental Research and Public Health, 16(18), 3502–. doi:10.3390/ijerph16183502

[49] Kirana, M. C., Lubis, M. Z., & Amalia, E. D. (2018). The Effect of Sound Manipulation to Know Response Rate in Autism Children Using FFT. 2018 International Conference on Applied Engineering(ICAE).doi:10.1109/incae.2018.8579418

APPENDIX-I

CERTIFICATES OF PAPER PRESENTED (if Any)

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

Dhanush Vasudevaraju

In recognition of the publication of the paper entitled

MACHINE LEARNING BASED MODEL FOR PREDICTION OF NEURODEVELOPMENTAL DISORDER

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 4 April 2023 , Date of Publication: 04-April-2023

UGC Approved Journal No: 49023 (18)

PAPER ID : IJCRT2304190

Registration ID : 233933

EDITOR IN CHIEF

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool), Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

C Ramesh Chandra

In recognition of the publication of the paper entitled

MACHINE LEARNING BASED MODEL FOR PREDICTION OF NEURODEVELOPMENTAL DISORDER

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 4 April 2023 , Date of Publication: 04-April-2023

UGC Approved Journal No: 49023 (18)

PAPER ID : IJCRT2304190

Registration ID : 233933

EDITOR IN CHIEF

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

Dinesh M

In recognition of the publication of the paper entitled

MACHINE LEARNING BASED MODEL FOR PREDICTION OF NEURODEVELOPMENTAL DISORDER

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 4 April 2023 , Date of Publication: 04-April-2023

UGC Approved Journal No: 49023 (18)

PAPER ID : IJCRT2304190

Registration ID : 233933

ISSN 2320-2882 IJCRT

International Journal of Creative Research Thoughts(IJCRT)

EDITOR IN CHIEF

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

C Sai Ranganath

In recognition of the publication of the paper entitled

MACHINE LEARNING BASED MODEL FOR PREDICTION OF NEURODEVELOPMENTAL DISORDER

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 4 April 2023 , Date of Publication: 04-April-2023

UGC Approved Journal No: 49023 (18)

EDITOR IN CHIEF

PAPER ID : IJCRT2304190

Registration ID : 233933

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

APPENDIX-II

Journal Published Paper attachment (if any)

www.ijcrt.org © 2023 IJCRT | Volume 11, Issue 4 April 2023 | ISSN: 2320-2882

IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

MACHINE LEARNING BASED MODEL FOR PREDICTION OF NEURODEVELOPMENTAL DISORDER

1Dhanush Vasudevaraju, 2Sandhya A Kulkarni, 3C Ramesh Chandra, 4Dinesh M, 5C Sai Ranganath
1Student, 2Assistant Professor, 3Student, 4Student, 5Student
1K S School of Engineering and Management,
2K S School of Engineering and Management,
3K S School of Engineering and Management,
4K S School of Engineering and Management,
5K S School of Engineering and Management

Abstract* — *Autism spectrum disorder is a neurodevelopmental disorder that affects a person's interaction, communication and learning skills. Although diagnosis of autism can be done at any age, its symptoms generally appear in the first two years of life and develop through time. Autism patients face different types of challenges such as difficulties with concentration, learning disabilities, mental health problems such as anxiety, depression, motor difficulties, sensory problems, and many others. Diagnosis of autism requires significant amount of time and cost. Earlier detection of autism can come to a great help by prescribing patients with proper medication at an early stage. It can prevent the patient's condition from deteriorating further and would help to reduce long term costs associated with delayed diagnosis. Thus, an efficient, accurate and easy screening test tool is very much required which would predict autism traits in an individual. The main idea behind this project is to detect autism spectrum disorder in an individual (male/female). This project is implemented by making use of a Machine Learning model using parameters such as an individual's age, gender, ethnicity, Autism Quotient Tool. The detection derived from this project will help an individual to get required diagnosis in time to prevent further complications of developing Alzheimer's disease.

I. INTRODUCTION

Autism spectrum disorder is a neurodevelopmental disorder that affects a person's interaction, communication and learning skills. Although diagnosis of autism can be done at any age, its symptoms generally appear in the first two years of life and develop through time. Autism patients face different types of challenges such as difficulties with concentration, learning disabilities, mental health problems such as anxiety, depression etc, motor difficulties, sensory problems and many others.

Diagnosis of autism requires significant amount of time and cost. Earlier detection of autism can come to a great help by prescribing patients with proper medication at an early stage.It can prevent the patient's condition from deteriorating further and would help to reduce long term costs associated with delayed diagnosis. Thus a time efficient, accurate and easy screening test tool is very much required which would predict autism traits in an individual and identify whether or not they require comprehensive autism assessment.

The objective of this work is to propose an autism prediction model using ML techniques and to develop a mobile application that could effectively predict autism traits of an individual of any age. In other words, this work focuses on developing an autism screening application for predicting the ASD traits among people of age groups 0-3 years and 4-13 years.

II. RELATED WORK

Current explosion rate of autism around the world is numerous and it is increasing at a very high rate. According to WHO, about 1 out of every 160 children has ASD. Some people with this disorder can live independently, while others require life-long care and support. Diagnosis of autism requires significant amount of time and cost.

Earlier detection of autism can come to a great help by prescribing patients with proper medication at an early stage.It can prevent the patient's condition from deteriorating further and would help to reduce long term costs associated with delayed diagnosis. Thus, a time efficient, accurate and easy screening test tool is very much required which would predict autism traits in an individual and identify whether or not they require comprehensive autism assessment.

The main purpose of this project is to develop an application to implement a machine learning based solution that can detect Autism Spectrum Disorder in an individual(male/female) during early stages for early diagnosis using efficient algorithms to obtain results with required accuracy. Then, design and develop an user-interface for user to interact.

www.ijcrt.org © 2023 IJCRT | Volume 11, Issue 4 April 2023 | ISSN: 2320-2882

III. Incorporated Packages

A. Python

Python is a computer programming language often used to build websites and software, automate tasks, and conduct data analysis. Python is a general-purpose language, meaning it can be used to create a variety of different programs and isn't specialized for any specific problems. The python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

B. Jupyter Notebook

The Jupyter Notebook App is a server-client application that allows editing and running notebook documents via a web browser. The Jupyter Notebook App can be executed on a local desktop requiring no internet access (as described in this document) or can be installed on a remote server and accessed through the internet.

C. User-Interface

The user interface is the point at which human users interact with a computer, website or application. The goal of effective UI is to make the user's experience easy and intuitive, requiring minimum effort on the user's part to receive the maximum desired outcome.

D. Support Vector Machines

Support Vector Machines (SVMs), also support vector networks are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. It is formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples.

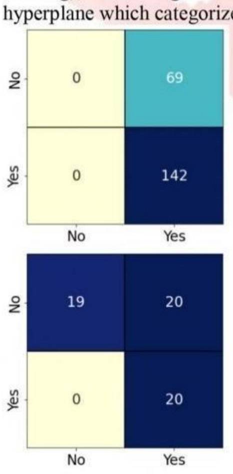

Figure 1: Confusion Matrix for SVM

E. Artificial Neural Network

An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons. An artificial neuron receives signals then processes them and can signal neurons connected to it. The "signal" at a connection is a real number, and the output of each neuron is computed by some non-linear function of the sum of its inputs. The connections are called *edges*. Neurons and edges typically have a *weight* that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a threshold such that a signal is sent only if the aggregate signal crosses that threshold.

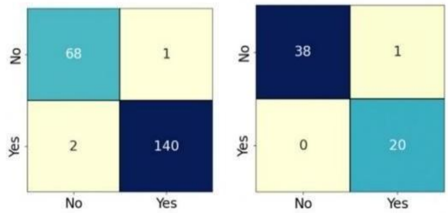

Figure 2: Confusion matrix for ANN

F. Random Forest Algorithm

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes or mean prediction of the individual trees.

Figure 3: Confusion Matrix for Random Forest Algorithm

IV. The Proposed Method

Our proposed strategy focuses on a novel machine learning procedures for Autism spectrum disorder (ASD) classification and prediction, thus overcoming the existing problem. By utilizing Random Forest (RF), Support Vector Machine(SVM), ANN algorithms we will make our model in order to increase the performance and accuracy. This method has multiple advantages such as - we don't have to figure out the features ahead of time, it is more effective, fault tolerant and scalable.

Different existing data mining procedures and its application were considered or explored. Utilization of machine learning algorithms was connected in various medical data sets. Machine learning strategies have diverse power in different medical data sets. In this work, a business intelligent model has been developed specific business structure deal with Autism classification using a suitable machine learning

www.ijcrt.org © 2023 IJCRT | Volume 11, Issue 4 April 2023 | ISSN: 2320-2882

technique. The model was evaluated by a scientific approach to measure accuracy. We are using Random forests, Support Vector Machines to build our model. Then, we make use of ANN algorithm. ANNs are composed of artificial neurons which are conceptually derived from biological neurons. Each artificial neuron has inputs and produces a single output which can be sent to multiple other neurons. After model construction it is time for model training. We were able to build an Random Forest, Support Vector Machine and ANN to recognize Autism Spectrum Disorder(ASD). In the next step, we will split the dataset into train dataset and test dataset. Then, we will build and train the model using training dataset. Finally, we will design an user-interface for any user to get the predicted results.

www.ijcrt.org © 2023 IJCRT | Volume 11, Issue 4 April 2023 | ISSN: 2320-2882

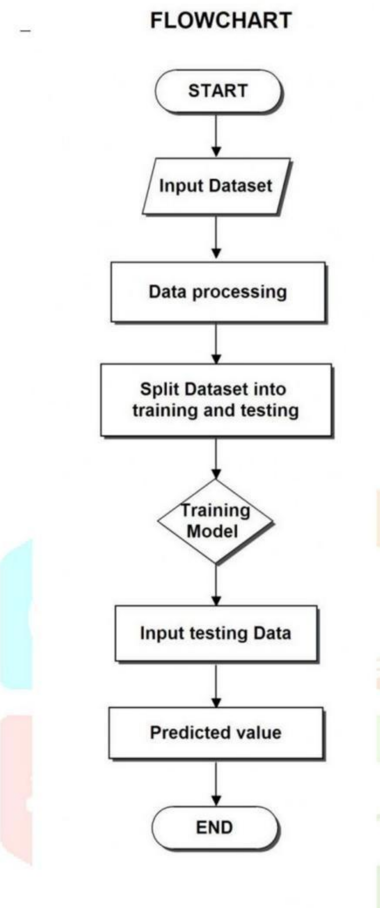

Fig 4: Proposed system

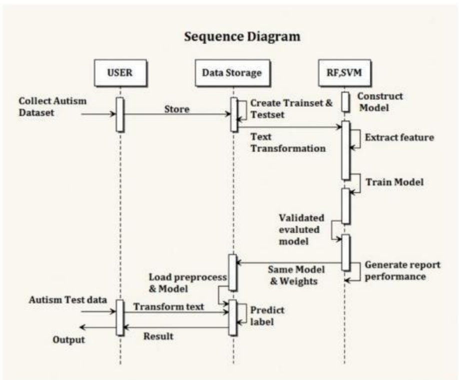

Fig 5: Sequence Diagram

V. Result And Analysis

In this final phase, we will test our classification model on our prepared image dataset and also measure the performance on our dataset. To evaluate the performance of our created classification and make it comparable to current approaches, we use accuracy to measure the effectiveness of classifiers.

After model building, knowing the power of model prediction on a new instance, is very important issue. Once a predictive model is developed using the historical data, one would be curious as to how the model will perform on the data that it has not seen during the model building process. One might even try multiple model types for the same prediction problem, and then, would like to know which model is the one to use for the real-world decision making situation, simply by comparing them on their prediction performance (e.g., accuracy). To measure the performance of a predictor, there are commonly used performance metrics, such as accuracy, recall etc. First, the most commonly used performance metrics will be described, and then some famous estimation methodologies are explained and compared to each other. "Performance Metrics for Predictive Modeling In classification problems, the primary source of performance measurements is a coincidence matrix (**classification matrix or a contingency table**)".

Below figure shows a coincidence matrix for a two-class classification problem.

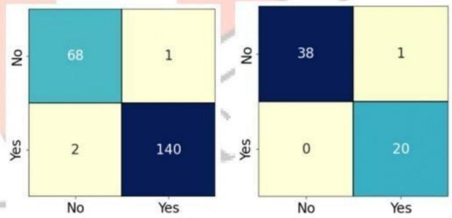

Figure 6: Confusion matrix

As being seen in above figure, the numbers along the diagonal from upper-left to lower-right represent the correct decisions made, and the numbers outside this diagonal represent the errors. "The true positive rate (also called hit rate or recall) of a classifier is estimated by dividing the correctly classified positives (the true positive count) by the total positive count. The false positive rate (also called a false alarm rate) of the classifier is estimated by dividing the incorrectly classified negatives (the false negative count) by the total negatives. The overall accuracy of a classifier is estimated by dividing the total correctly classified positives and negatives by the total number of samples.

VI. Conclusions

This system comes under machine learning which is advanced technique at present. RF, ANN, is more suitable for numerical processing especially in medical classification. We conclude the experimental result what we are getting from developed system is 99% and 98% accurate for child and toddler datasets

IJCRT2304190 | International Journal of Creative Research Thoughts (IJCRT) | b583

respectively. We have also compared all three algorithms as shown below.

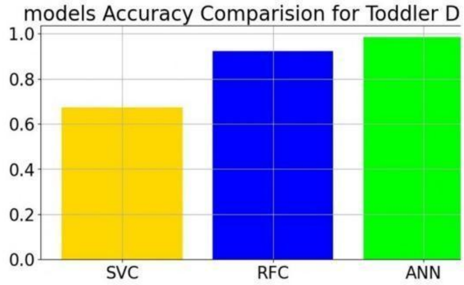

Figure 7: Model Comparison

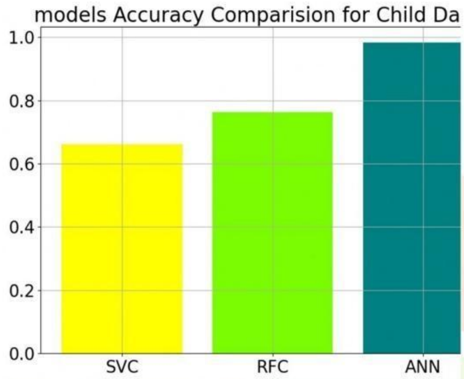

Figure 8: Model Comparison

REFERENCES

[1] Omar, Kazi Shahrukh, Prodipta Mondal, Nabila Shahnaz Khan, Md Rezaul Karim Rizvi, and Md Nazrul Islam. "A machine learning approach to predict autism spectrum disorder." In 2019 International conference on electrical, computer and communication engineering (ECCE), pp. 1-6. IEEE, 2019.doi:10.1109/ecace.2019.8679454

[2] Akter, Tania, Md Shahriare Satu, Md Imran Khan, Mohammad Hanif Ali, Shahadat Uddin, Pietro Lio, Julian MW Quinn, and Mohammad Ali Moni. "Machine learning-based models for early stage detection of autism spectrum disorders." IEEE Access 7 (2019): 166509-166527.doi:10.1109/access.2019.2952609

[3] Baranwal, Astha, and M. Vanitha. "Autistic spectrum disorder screening: prediction with machine learning models." In 2020 International conference on emerging trends in information technology and engineering (ic-ETITE), pp. 1-7. IEEE, 2020. . doi:10.1109/ic-etite47903.2020.186

[4] Eni, M, Dinstein, I, Ilan, M., Menashe, I, Meiri, G, & Zigel, Y. (2020). Estimating Autism Severity in Young Children From Speech Signals Using a Deep Neural Network. IEEE Access, 8, 139489–139500. doi:10.1109/access.2020.301253

[5] Raj, S, & Masood, S. (2020). Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques. Procedia Computer Science, 167, 994-1004.doi:10.1016/j.procs.2020.03.39

[6] Mamun, K. A. A., Bardhan, S., Ullah, M. A., Anagnostou, E., Brian, J., Akhter, S., & Rabbani, M. G. (2016). Smart autism — A mobile, interactive and integrated framework for screening and confirmation of autism. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). doi:10.1109/embc.2016.7592093

[7] F. Thabtah, F. Kamalov, and K. Rajab, ``A new computational intelligence approach to detect autistic features for autism screening," Int. J. Med. Inform., vol. 117, pp. 112_124,Sep.2018.https://doi.org/10.1016/j.ijmedinf.2018.06.009

[8] F. Thabtah and D. Peebles, ``A new machine learning model based on induction of rules for autism detection," Health Inform. J., 2019, Art. no. 1460458218824711, doi: 10.1177/1460458218824711.

[8] M. S. Satu, F. F. Sathi, M. S. Arifen, M. H. Ali, and M. A. Moni, ``Early detection of autism by extracting features: A case study in Bangladesh," in Proc. 1st Int. Conf. Robot., Elect. Signal Process. Techn. (ICREST), Jan. 2019, pp. 87_90. doi: 10.1109/ICREST.2019.8644357

[9] H. Abbas, F. Garberson, E. Glover, and D. P. Wall, ``Machine learning approach for early detection of autism by combining questionnaire and home video screening," J. Amer. Med. Informat. Assoc., vol. 25, no. 8, pp. 1000_1007, 2018. doi:10.1093/jamia/ocy039

[10] F. Thabtah, ``Machine learning in autistic spectrum disorder behavioral research: A review and ways forward," Informat. Health Social Care, vol. 44, no. 3, pp. 278_297, 2018. doi:10.1080/17538157.2017.1399132

APPENDIX-III

VARIABLE	AQ-10 TODDLER FEATURES (18-36 months)
AQ1	Does your child look at you when you call his/her name?
AQ2	How easy is it for you to get eye contact with your child?
AQ3	Does your child point to indicate that he/she wants something?
AQ4	Does your child point to share interest with you?
AQ5	Does your child pretend?
AQ6	Does your child follow where you are looking?
AQ7	If you or someone else in the family is visibly upset, does your child show signs of wanting to comfort them?
AQ8	Would you describe your child's first words as:
AQ9	Does your child use simple gestures?
AQ10	Does your child stare at nothing with no apparent purpose?

VARIABLE	AQ-10 CHILD FEATURES (4-13 years)
AQ1	Does he/she often notices small sounds when others do not?
AQ2	Does he/she usually concentrates more on the whole picture rather than small details?
AQ3	In a social group, can he/she easily keep track of several different people's conversation?
AQ4	Does he/she finds it easy to go back and forth between different activities?
AQ5	He/she does not know how to keep a conversation going with his/her peers?
AQ6	Is he/she good at social chit-chat?
AQ7	When he/she read a story, he/she finds it difficult to work out the character's intentions or feelings?
AQ8	When he/she was in preschool, he/she used to enjoy playing pretending games with other children?
AQ9	Does he/she finds it easy to work out what someone is thinking or feeling just by looking at their face?
AQ10	Does he/she finds it difficult to make new friends?