Analysis and Detection of Autism Machine Learning

R .Karuppusamy, Dr. A. Narendrakumar
PG Scholar, Department of Electronics and Communication Engineering,
Hindusthan College of Engineering and Technology, Coimbatore, Tamil Nadu, India.
Associate Professor, Department of Electronics and Communication Engineering,
Hindusthan College of Engineering and Technology, Coimbatore, Tamil Nadu, India.

Abstract : ASD problems in a child, adolescents, and adults. The proposed Autism Spectrum Disorder (ASD) is a neuron-disorder in which a person has a lifelong effect on interaction and communication withothers.Autismcanbediagnosedatanystageinoncelifeandissaidtobea"behavioraldisease"becauseinthefirsttwoyears of life symptoms usually appear. According to the ASD problem starts with childhood and continues to keep going on into adolescence and adulthood. Propelled with the rise in use of machine learning techniques in the research dimensions of medical diagnosis, in this paper there is an attempt to explore the possibility to use Naïve Bays, Support Vector Machine, Logistic Regression, KNN, Neural Network and Convolution Neural Network for predicting and analysis techniques are evaluated on publicly available three different non-clinically ASD datasets. First dataset related to ASD screening in children has 292 instances and 21 attributes. Second dataset related to ASD screening Adult subjects contains a total of 704 instances and 21 attributes. Third dataset related to ASD screening in Adolescent subjects comprises of 104 instances and 21 attributes. 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% and 96.88% for Autistic Spectrum Disorder Screening in Data for Adult, Children and Adolescent respectively.

I. INTRODUCTION

The problem of autism spectrum disorder (ASD) have been mounting swiftly nowadays among all ages of the human population. Early detection of this neurological disease can greatly assist in the maintenance of the subject's mental and physical health. With the rise of application of machine learning-based models in the predictions of various human diseases, their early detection based on various health and physiological parameter now seems possible. This factor motivated us to increase interest in the detection and analysis of ASD diseases to improve better treatment methodology. Detection of ASD becomes a challenge as there are several other mental disorders whose few symptoms are very similar to those with ASD symptoms, thereby makes this task a difficult one. Autism Spectrum disorder is a problem that is related to human brain development. A person who has suffered from the Autism Spectrum Disorder is generally not able to do social interaction and communication with other persons. In this, a person's life is usually affected for his or her entire lifetime. It is interesting to know that both environmental and genetic factors may turn out to be the causing factors for this disease. The symptoms of this problem may be started at the age of three years and may continue for the lifetime. It is not possible to complete treat the patient suffering from this disease, however its effects can be reduced for some time if the symptoms are early detected. By assuming that human genes are responsible for it, the exact causes of ASD have not been recognized by the scientist yet. The human genes affect the development by influencing the environment. There is some risk factor which influences ASD like as low birth weight, a sibling with ASD and having old parents, etc. Instead of this, there are some social interaction and communication problems like as:

- Inappropriate laughing and giggling
- No sensitivity of pain
- Not able to make eye contact properly
- No proper response to sound
- May not have a wish for cuddling
- Not able to express their gestures
- No interaction with others
- Inappropriate objects attachment
- Want to live alone
- Using echo words etc.

People with ASD also have difficulty with constrained interests and consistently repetition of behaviors. The following list presents specific examples of the types of behaviors.

• Repeating certain behaviors like repeating words or phrases much time.
• The Person will be upset when a routine is going to change.
• Having a little interest in certain matters of the topic like numbers, facts, etc.
• Less sensitive than another person in some cases like light, noise, etc.

Early detection and treatment are most important steps to be taken to decrease the symptoms of autism spectrum disorder problem and to improve the quality of life of ASD suffering people. However, there is no procedure of medical test for detection of autism. ASD Symptoms usually recognized by observation. In Older and adolescents who go to school, ASD symptoms are usually identified by their parents and teachers. After that ASD symptoms are evaluated by a special education team of the school. These school team suggested these children visit their health care doctor for required testing. In adults identifying ASD symptoms is very difficult than older children and adolescents because some symptoms of ASD may be overlap with other mental health disorders. It is easy to identify the behavioral changes in a child easily by observation because it can be seen early in the 6 months of age than Autism specific brain imaging because brain imaging can be identifying after 2 years of age

II. LITERATURESURVEY

Literature review was carried out to gain knowledge and improve the skills needed to complete this project. This chapter shows the different techniques that have been implemented.

Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques

Autism Spectrum Disorder (ASD) is a neurological disorder which might have a lifelong impact on the language learning, speech, cognitive, and social skills of an individual. Its symptoms usually show up in the developmental stages, i.e., within the first two years after birth, and it impacts around 1% of the population globally. ASD is mainly caused by genetics or by environmental factors; however, its conditions can be improved by detecting and treating it at earlier stages. In the current times, clinical standardized tests are the only methods which are being used, to diagnose ASD. This not only requires prolonged diagnostic time but also faces a steep increase in medical costs. To improve the precision and time required for diagnosis, machine learning techniques are being used to complement the conventional methods. We have applied models such as Support Vector Machines (SVM), Random Forest Classifier (RFC), Naïve Bays (NB), Logistic Regression (LR), and KNN to our dataset and constructed predictive models based on the outcome. The main objective of our paper is to thus determine if the child is susceptible to ASD in its nascent stages, which would help streamline the diagnosis process. Based on our results, Logistic Regression gives the highest accuracy for our selected dataset.

Automated Detection of Autism Spectrum Disorder Using a Convolution Neural Network

Convolution neural networks (CNN) have enabled significant progress in speech recognition, image classification, automotive software engineering, and neuroscience. This impressive progress is largely due to a combination of algorithmic breakthroughs, computation resource improvements, and access to a large amount of data. In this paper, we focus on the automated detection of autism spectrum disorder (ASD) using CNN with a brain imaging dataset. We detected ASD patients using most common resting-state functional magnetic resonance imaging (fMRI) data from a multi-site dataset named the Autism Brain Imaging Exchange (ABIDE). The proposed approach was able to classify ASD and control subjects based on the patterns of functional connectivity.

Detection of Autism Spectrum Disorder using Machine Learning

Autistic Spectrum Disorder (ASD) is a severe neuron- logical condition that affects the entire brain system and which in turn impacts the cognitive, emotional, social, and physical health of the individual. They experience difficulty in socializing and communicating with others. They are always in need of support either from parents, relatives, or friends to guide them. Unfortunately, there is no cure for autism but early detection can help in better treatment. A person's behavior a behaviors can be used to diagnose autism disorder. This method of diagnosis is time-consuming and ineffective for early detection of autism.. Therefore, there is a need for time-efficient and low-cost ASD screening to help individuals to decide whether they should undergo a clinical diagnosis and seek treatment. Therefore, we propose a machine learning-based, time-efficient solution to detect autism.

Autism Spectrum Disorder Detection with Machine Learning Methods

Autistic Spectrum Disorder (ASD) is a disorder associated with genetic and neurological components leading to difficulties in social interaction and communication. According to statistics of WHO, the number of patients diagnosed with ASD is gradually increasing. Most of the current studies focus on clinical diagnosis, data collection and brain images analysis, but do not focus on the diagnosis of ASD based on machine learning. Three ASD datasets are used for children, adolescences and adults. To classify the ASD data, we used the k-Nearest Neighbors method (kNN), the Support Vector Machine method (SVM) and the Random Forests method (RF). In our experiments, the

data was randomly split into training and test sets. The parts of the data were randomly selected 100 times to test the classification methods.

ANALYSIS & DETECTION OF AUTISM SPECTRUM DISORDER USING MACHINE LEARNING TECHNIQUES

A neuron-disease known as autism spectrum disorder (ASD) affects a human's ability towards engage & communicate among others on a lifelong basis. Autism is referred towards as a "behavioral disorder" since signs typically develop in first two years about life, but it can be diagnosed at any point in one's life. According towards ASD theory, problems begin in childhood & persist into adolescence & maturity. This paper attempts towards investigate potential use about machine learning algorithms for predicting & analyzing ASD problems in children, adults & adolescents. On three separate publicly accessible, non-clinically relevant ASD datasets, suggested approaches are assessed. There are 292 instances & 21 attributes in first dataset relating towards screening for ASD in children. Adult individuals make up second dataset for ASD screening, which has a total about 704 instances & 21 attributes. There are 104 cases & 21 attributes in third dataset, which is focused on ASD screening in adolescent individuals. Convolutional neural network based models had higher accuracy about 99.53 percent, 98.30 percent, & 96.88 percent for the three datasets respectively.

Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques

Autism Spectrum Disorder (ASD) is a neuron-disorder in which a person has a lifelong effect on interaction and communication with others. Autism can be diagnosed at any stage in once life and is said to be a "behavioral disease" because in the first two years of life symptoms usually appear. According to the ASD problem starts with childhood and continues to keep going on into adolescence and adulthood. Propelled with the rise in use of machine learning techniques in the research dimensions of medical diagnosis, in this paper there is an attempt to explore the possibility to use Naïve Bays, 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. First dataset related to ASD screening in children has 292 instances and 21 attributes. Second dataset related to ASD screening Adult subjects contains a total of 704 instances and 21 attributes. Third dataset related to ASD screening in Adolescent subjects comprises of 104 instances and 21 attributes. 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%, and 96.88% for Autistic Spectrum Disorder Screening in Data for Adult, Children, and Adolescents respectively.

Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms

Autism spectrum disorder (ASD) is a type of mental illness that can be detected by using social media data and biomedical images. Autism spectrum disorder (ASD) is a neurological disease correlated with brain growth that later impacts the physical impression of the face. Children with ASD have dissimilar facial landmarks, which set them noticeably apart from typically developed (TD) children. Novelty of the proposed research is to design a system that is based on autism spectrum disorder detection on social media and face recognition. To identify such landmarks, deep learning techniques may be used, but they require a precise technology for extracting and producing the proper patterns of the face features. This study assists communities and psychiatrists in experimentally detecting autism based on facial features, by using an uncomplicated web application based on a deep learning system, that is, a convolutional neural network with transfer learning and the flask framework. Exception, Visual Geometry Group Network (VGG19), and NASNET Mobile are the pertained models that were used for the classification task. The dataset that was used to test these models was collected from the Kaggle platform and consisted of 2,940 face images. Standard evaluation metrics such as accuracy, specificity, and sensitivity were used to evaluate the results of the three deep learning models. The Exception model achieved the highest accuracy result of 91%, followed by VGG19 (80%) and NASNET Mobile (78%).

Identification of autism spectrum disorder using deep learning and the ABIDE dataset

The goal of the present study was to apply deep learning algorithms to identify autism spectrum disorder (ASD) patients from large brain imaging dataset, based solely on the patient's brain activation patterns. We investigated ASD patient's brain imaging data from a world-wide multi-site database known as ABIDE (Autism Brain Imaging Data Exchange). ASD is a brain-based disorder characterized by social deficits and repetitive behaviors. According to recent Centers for Disease Control data, ASD affects one in 68 children in the United States. We investigated patterns of functional connectivity that objectively identify ASD participants from functional brain imaging data, and attempted to unveil the neural patterns that emerged from the classification. The results improved the state-of-the-art by achieving 70% accuracy in identification of ASD versus control patients in the dataset. The patterns that emerged from the classification show an autocorrelation of brain function between anterior and posterior areas of the brain; the autocorrelation corroborates current empirical evidence of anterior-posterior disruption in brain connectivity in ASD. We present the results and identify the areas of the brain that contributed most to

differentiating ASD from typically developing controls as per our deep learning model.

Identification of Autism Spectrum Disorder using Deep Neural Network

One of the acute neuron developmental disorders throughout the world today is the Autism Spectrum disorder (ASD). It is lifelong disorder which affects the behavior and communication skill of an individual. According to world health organization 2019 report, the number of individuals diagnosed with ASD is increasing creating a threat as it is analogous to significant health care cost. Early recognition can considerably reduce the effect. In order to get rid of the time consuming and expensive diagnosis procedures for ASD, a mobile based ASD screening tool known as ASD Test app was developed. The app recorded over 1400 number of instances covering toddler, child, adolescent and adult. It is available publicly in haggle and UCI Machine Learning repository for research purpose. The paper gives a new approach for identification of ASD using a deep classifier. The ASD identification works in the following steps. Feature analysis explains ASD traits thereby improving the efficiency of screening process. Further, Machine Learning (ML) classifier models report ASD class type with evaluation parameters. In this analysis, an attempt is made for the incorporation of Principal Component Analysis (PCA) for feature dimension reduction followed by the usage of Deep Neural Network (DNN) for classification of ASD class type. The data upon which the techniques are applied are collected from Haggle and UCI ML repository. The experiment result indicates that, PCA in combination with DNN provide clinically acceptable output for effective ASD identification. *Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment*

One of the primary psychiatric disorders is Autistic Spectrum Disorder (ASD). ASD is a mental disorder that limits the use of linguistic, communicative, cognitive, skills as well as social skills and abilities. Recently, ASD has been studied in the behavioral sciences using intelligent methods based around machine learning to speed up the screening time or to improve sensitivity, specificity or accuracy of the diagnosis process. Machine learning considers the ASD diagnosis problem as a classification task in which predictive models are built based on historical cases and controls. These models are supposed to be plugged into a screening tool to accomplish one or more of the aforementioned goals. In this paper, we shed light on recent studies that employ machine learning in ASD classification in order to discuss their pros and cons. Moreover, we highlight a noticeable problem associated with current ASD screening tools; the reliability of these tools using the DSM-IV rather than the DSM-5 manual. Hence the necessity to amend current screening tools to reflect the new imposed criteria of ASD classification in the DSM-5 particularly the diagnostic algorithms embedded within these methods.

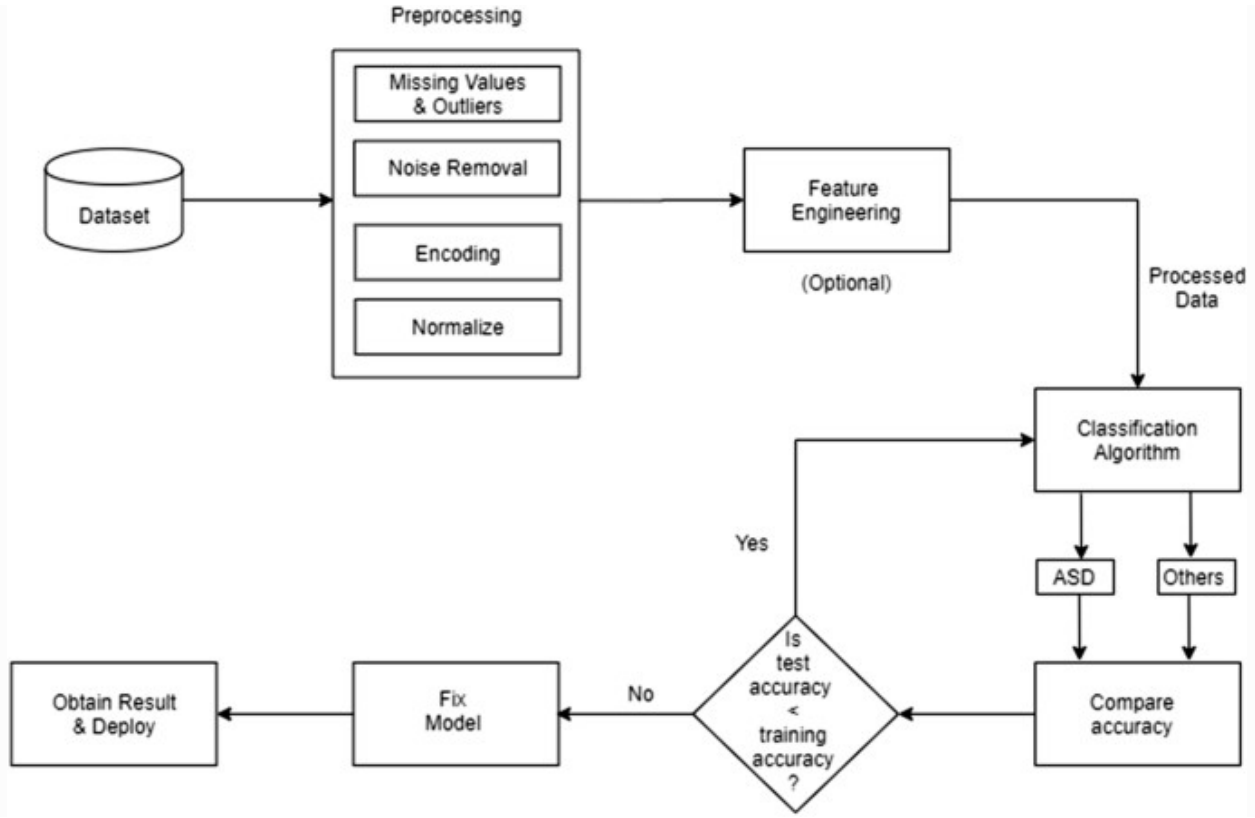
III. SYSTEMANALYSIS

EXISTINGSYSTEM

In the early stages Data pre-processing is a technique in which transform the raw data into a meaningful and understandable format. Real-world data is commonly incomplete and inconsistent because it contains lots of errors and null values. A goodpre-processeddataalwaysyieldstoagoodresult.VariousDatapre-processingmethodsareusedtohandleincompleteandinconsistentdatalikeashandlingmissingvalues,outlierdetection,data discretization,datareduction(dimensionand numerosity reduction), etc. The problems of missing values in these data set has been handled by imputation method.

*PROPOSEDSYSTE*M

We begin by pre-processing the dataset to eliminate missing values and outliers, remove noise, and encode categorical attributes. We also employ feature engineering to choose the most beneficial features out of all the features present in the data set. This reduces data dimensionality to improve speed and efficiency during training. Once the data set has been pre-processed, classification algorithms like Logistic Regression, Naïve Bays, Support Vector Machine, K-Nearest Neighbours, CNN and ANN Classifiers are used to predict the output label (ASD or no ASD). The accuracy of each classifier is observed and compared. Furthermore, metrics like the F1 score and precision-recall values have also been computed for better evaluation of each classifier. If the classifier performs well, then the training accuracy will be higher than its test accuracy. This model can then be deemed to be the best model and hence be used for further training and classification.



IV. METHODS AND MODELS

Thefollowingmodelshavebeenusedinthispredictionprocessandhassomeaccuracylevelsforprediction.

Data Pre-Processing

Data pre-processing is a technique in which transform the raw data into a meaningful and understandable format. Real-world data is commonly incomplete and inconsistent because it contains lots of errors and null values. A good pre-processed data always yields to a good result. Various Data pre-processing methods are used to handle incomplete and inconsistent data like as handling missing values, outlier detection, data discretization, data reduction (dimension and tuberosity reduction), etc. The problems of missing values in these dataset has been handled by imputation method.

Training and Testing Model

The whole dataset has been split into two parts i.e. one part is training the dataset and the other one is testing dataset with a ratio of 80:20 respectively. For cross-validation purposes again training data has been split into two parts. One part is the training dataset and another part is the validation dataset into an 80:20 ratio respectively.

Support Vector Machine (SVM)

SVM is a linear supervised machine learning approach that is used for classification and regression. It is a pattern recognition problem solver. It does not cause the problem of over fitting. SVM separates the classes by defining a decision boundary

Naïve Bayes (NB)

A naive Bayes classifier is a supervised learning algorithm. It is a generative model and is based on joint probability distribution. The Naive Bays concept based on independence assumptions. It exhibits less training time as compared to SVM and ME model. It calculates the posterior probability for a dataset using the prior probability and likelihood.

Convolution Neural Network (CNN)

CNN is one of the deep learning techniques known to build models for various problems. It is a feed-forward neural network that is inspired by the human brain. A CNN model contains one input layer, one output layer, and many other different layers i.e. convolution layers, max pooling, fully connected layers, and normalization layers. Their activation functions can be computed with Matrix Multiplication, which is followed by a bias offset.

Logistic Regression (LR)

LR is a regression tool that is used to analyze the binary dependent variables. Its output value lies in either the 0 or 1 form. It is used for the continuous value dataset. It tells the relationship between one dependent binary variable, and one nominal or ordinary variable. It can be represented by the sigmoid function.

K- Nearest Neighbor (KNN)

KNN is a supervised learning approach and is the simplest of all. It is used for classification as well as regression problems. It assumes that similar data exist nearby. The 'K' part indicates the number of seed point that is to be selected. It should be chosen carefully to reduce the error. Thus it is based on the idea of similarity which can be in terms of distance, closeness or proximity. The most common distance measure is Euclidean distance.

Artificial Neural Network

ANN is a neural network that has a connection with multiple neurons. Each neuron cell having a group of input values and associated weights. The most common artificial 0 neural network feeds forward neural network. In this network, the flow of information moves in the only forward direction. This type of network contains three main layers, first is the input layer, the second is a hidden layer and last is the output layer. There is no cycle or loop in the network

V. IMPLEMENTATIONDETAILS

Thischapterbriefsabouttheimplementationdetailsoftheprojectinadetailedmannerand also results as well a performance analysis.

SOFTWAREDETAILS

WINDOWS

Windows is a graphical operating system developed by Microsoft. It allows users to view and store files, run the software, play games, watch videos, and provides a way to connect to the internet. It was released for both home computingandprofessionalworks.Windowsstageismostappropriateforgame and programming engineers. Window have a huge number crowd so designer's want to make utilities, games and programming for windows OS.

PYTHON

Python is an interpreter, 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 Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. 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.

Python Features

Python has few keywords, simple structure, and a clearly defined syntax. Python code is more clearly defined and visible to the eyes. Python's source code is fairly easy-to-maintaining. Python's bulk of the library is very portable and cross-platform compatible on UNIX, Windows, and Macintosh. Python has support for an interactive mode which allows interactive testing and debugging of snippets of code.

Portable

Python can run on a wide variety of hardware platforms and has the same interface on all platforms.

Extendable

It allows to add low-level modules to the Python interpreter. These modules enable programmers to add to or customize their tools to be more efficient.

Databases

Python provides interfaces to all major commercial databases.

GUI Programming

Python supports GUI applications that can be created and ported to many system calls, libraries and windows systems, such as Windows MFC, Macintosh, and the X Window system of UNIX.

Scalable

Python provides a better structure and support for large programs than shell scripting.

Object-Oriented Approach

One of the key aspects of Python is its object-oriented approach. This basically means that Python recognizes the concept of class and object encapsulation thus allowing programs to be efficient in the long run.

Highly Dynamic

Python is one of the most dynamic languages available in the industry today. There is no need to specify the type of the variable during coding, thus saving time and increasing efficiency.

Extensive Array of Libraries

Python comes inbuilt with many libraries that can be imported at any instance and be used in a specific program.

Open Source and Free

Python is an open-source programming language which means that anyone can create and contribute to its development. Python is free to download and use in any operating system, like Windows, Mac or Linux.

ANACONDA

Anaconda is a free and open-source distribution of the Python and R programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management and deployment. Package versions are managed by the package management system .The Anaconda distribution includes data-science packages suitable for Windows, Linux, and MacOS.

Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda distribution that allows users to launch applications and manage condo packages, environments and channels without using command-line commands. Navigator can search for packages on Anaconda Cloud or in a local Anaconda Repository, install them in an environment, run the packages and update them. It is available for Windows, Mac OS and Linux.

JUPYTER NOTEBOOK

"Jupiter" is a loose acronym meaning Julia, Python, and R. These programming languages were the first target languages of the Jupiter application. As a server-client application, the Jupiter Notebook App allows you to edit and run your notebooks via a web browser. The application can be executed on a PC without Internet access, or it can be installed on a remote server and it can access through the Internet.

A kernel is a program that runs and introspects the user's code. The Jupiter Notebook App has a kernel for Python code. "Notebook" or "Notebook documents" denote documents that contain both code and rich text elements, such as figures, links, and equations. The mix of code and text elements, these documents are the ideal place to bring together an analysis description, and can be executed to perform the data analysis in real time.

The Jupiter Notebook App is a server-client application that allows editing and running notebook documents via a web browser. The Jupiter 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. In addition to displaying/editing/running notebook documents, the Jupiter Notebook App has a "Dashboard" (Notebook Dashboard), a "control panel" showing local files and allowing to open notebook documents or shutting down their kernels Jupiter Notebook document is a JSON file, following a versioned schema, usually ending with the ".ipynb" extension. The main parts of the Jupiter Notebooks are: Metadata, Notebook format and Lis to fcells.

Jupiter Notebook contains two components such as web application and notebook documents.

A web application is a browser-based tool for interactive authoring of documents which combine explanatory text, mathematics, computations and their rich media output. Notebook documents is a representation of all content visible in the web application, including inputs and outputs of the computations, explanatory text, mathematics, images, and rich media representations of objects.

Structure of a notebook document

The notebook consists of a sequence of cells. A cell is a multiline text input field, and its contents can be executed by using Shift-Enter, or by clicking either the "Play" button the toolbar. The execution behavior of a cell is determined by the cell's type. There are three types of cells namely code cells, markdown cells, and raw cells. Every cell starts off being a code cell, but its type can be changed by using a drop-down on the toolbar.

Code cells

A code cell allows you to edit and write new code, with full syntax highlighting and tab completion. The programming language you use depends on the kernel, and the default kernel (I Python) runs Python code.

Markdown cells

Document the computational process in a literate way, alternating descriptive text with code, using rich text. In I Python this is accomplished by marking up text with the Markdown language. The corresponding cells are called Markdown cells. The Markdown language provides a simple way to perform this text mark-up, to specify which parts of the text should be emphasized (italics), bold, form lists, etc.

Raw cells

Raw cells provide a place in which you can write output directly. Raw cells are not evaluated by the notebook. When passed through nb convert, raw cells arrive in the destination format unmodified.

NUMPY

NumPy can be used to perform a wide variety of mathematical operations on arrays .It adds powerful data structures to Python that guarantee efficient calculationswitharraysandmatricesanditsuppliesanenormouslibraryofhigh-levelmathematical functions that operate on these arrays and matrices. NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, Fourier transform, and matrices. NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra. Fourier transform, and matrices.

MICROSOFT EXCEL

Microsoft Excel is a spreadsheet developed by Microsoft for Windows, Mac OS, Android and iOS. It features calculation, graphing tools, pivot tables and a macro programming language called Visual Basic for applications.

FEATURES

Basic Operation

Microsoft Excel has the basic features of all spread sheets using a grid of cells arranged in numbered rows and letter-named columns to organize data manipulations like arithmetic operations. It has a battery of supplied functions to answer statistical, engineering and financial needs. In addition, it can display data as line graphs, histograms and charts, and with a very limited three-dimensional graphical display.

VBA programming

The Windows version of Excel supports programming through Microsoft's Visual Basic for Applications (VBA), which is a dialect of Visual Basic. Programmers may write code directly using the Visual Basic Editor (VBE), which includes a window for writing code, debugging code, and code module organization environment. The user can implement numerical methods as well as automating tasks such as formatting or data organization in VBA and guide the calculation using any desired intermediate results reported back to the spreadsheet.

Charts

Excel supports charts, graphs, or histograms generated from specified groups of cells. The generated graphic component can either be embedded within the current sheet, or added as a separate object. These displays are dynamically updated if the content of cells change. For example, suppose that the important design requirements are displayed visually; then, in response to a user's change in trial values for parameters, the curves describing the design change shape, and their points of intersection shift, assisting the selection of the best design.

Data storage and communication

Number of rows and columns

Versions of Excel up to 7.0 had a limitation in the size of their data sets of 16K ($2^{14} = 16384$) rows. Versions 8.0 through 11.0 could handle 64K ($2^{16} = 65536$) rows and 256 columns (2^{8} as label 'IV'). Version 12.0 onwards, including the current Version 16.x, can handle over 1M ($2^{20} = 1048576$) rows, and 16384 (2^{14} as label 'XFD') columns.

File formats

Microsoft Excel up until 2007 version used a proprietary binary file format called Excel Binary File Format (.XLS) as its primary format. Excel 2007 uses Office Open XML as its primary file format, an XML-based format that followed after a previous XML-based format called "XML Spreadsheet" ("XMLSS"), first introduced in Excel 2002. In addition, most versions of Microsoft Excel can read CSV, DBF, SYLK, DIF, and other legacy formats. Support for some older file formats was removed in Excel 2007. The file formats were mainly from DOS-based programs.

Binary

OpenOffice.org has created documentation of the Excel format. Since then Microsoft made the Excel binary format specification available to freely download.

Export and migration of spreadsheets

Programmers have produced APIs to open Excel spreadsheets in a variety of applications and environments other than Microsoft Excel. These include opening Excel documents on the web using either ActiveX controls, or plugging like the Adobe Flash Player. The Apache POI open source project provides Java libraries for reading and writing Excel spreadsheet files. Excel Package is another open-source project that provides server-side generation of Microsoft Excel 2007 spreadsheets. PHP Excel is a PHP library that converts Excel5, Excel 2003, and Excel 2007 formats into objects for reading and writing within a web application. Excel Services is a current .NET developer

tool that can enhance Excel's capabilities. Excel spreadsheets can be accessed from Python with xlrd and openpyxl.

CSV File

A comma-separated values (CSV) file is a delimited text file that uses a comma to separate values. Each line of the file is a data record. Each record consists of one or more fields, separated by commas. The use of the comma as a field separator is the source of the name for this file format. A CSV file typically stores tabular data (numbers and text) in plain text, in which case each line will have the same number of fields. These files serve a few different business purposes. They help companies export a high volume of data to a more concentrated database.

The rules should be followed to format CSV file

- Each record (row of data) is to be located on a separate line, delimited by a line break.
- The last record in the file may or may not have an ending line break.
- There may be an optional header line appearing as the first line of the file with the same format as normal record lines.
- It should contain the same number of fields as the records in the rest of the file.
- The header contains names corresponding to the fields in the file.
- In the header and each record, there may be one or more fields, separated by commas.
- Each line should contain the same number of fields throughout the file. Spaces are considered part of a field and should not be ignored.
- The last field in the record must not be followed by a comma.
- Each field may or may not be enclosed in double quotes.
- If fields are not enclosed with double quotes, then double quotes may not appear inside the fields.
- Fields containing line breaks (CRLF), double quotes, and commas should be enclosed in double quotes.
- If double quotes are used to enclose fields, then a double quote appearing inside a field must be escaped by preceding it with another double quote.

VI. CONCLUSION AND FUTURE ENHANCEMENT

In this work, detection of Autism Spectrum Disorder was attempted using various machine learning and deep learning techniques. Various performance evaluation metrics were used to analyze the performance of the models implemented for ASD detection on non-clinical dataset from three sets of age groups viz. Child, Adolescents and the Adult. When comparing the result with another recent study on this problem got a better result of the CNN classifier instead of SVM with including all its features attributes after handling missing values. In this work after handling missing value, both the SVM and CNN based models show the same accuracy of prediction of about 98.30 % for ASD Child dataset. However for the remaining two other datasets, the CNN based model was able to achieve highest accuracy result than all the other considered model building techniques, These results strongly suggest that a CNN based model can be implemented for detection of Autism Spectrum Disorder instead of the other conventional machine learning classifier suggested in earlier researches.

REFERENCES

[1] Uğur Erkan and Dang N.H. Thanhs, ‘‘Autism Spectrum Disorder Detection with Machine Learning Methods”, 2019.ning Techniques”, 2022.

[2] Suman Raja ,Sarfaraz Masood, ‘‘Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques”, 2020.

[3] Fawaz WaselallahAlsaadeand Mohammed Saeed Alzahrani, ‘‘Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms”, 2022.

[4] AnibalSólon Heinsfeld, AlexandreRosa Franco, R.Cameron Craddock, Augusto Buchweitz , Felipe Meneguzzi, ‘‘Identification of autism spectrum disorder using deep learning and the ABIDE dataset”, 2018.

[5] Sindhu Mohanty, Ashima ; Parida, Priyadarsan ; Patra, K. C., ‘‘Identification of Autism Spectrum Disorder using Deep Neural

Network", 2021.
[6] [10] FadiThabtah, ''Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment", 2017.
[7] Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
[8] C. Nagarajan, G.Neelakrishnan, R. Janani, S.Maithili, G. Ramya "Investigation on Fault Analysis for Power Transformers Using Adaptive Differential Relay"Asian Journal of Electrical Science, Vol.11 No.1
[9] P. Kiran Kumar , Dr. K. SanthiSree, "Analysis & Detection Of Autism Spectrum Disorder Using Machine LearDeepa Krishnan , Kaushik Vakadkar, Diya Purkayastha ''Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques, ''2021.
[10] C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of ELECTRICAL ENGINEERING, Vol.63 (6), pp.365-372, Dec.2012.
[11] G.Neelakrishnan, K.Anandhakumar, A.Prathap, S.Prakash "Performance Estimation of cascaded h-bridge MLI for HEV using SVPWM"SurajPunj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:750-756
[12] Zeinab Sherkatghanad, MohammadsadeghAkhondzadeh, ''Automated Detection of Autism Spectrum Disorder Using a Convolutional Neural Network", 2019.
[13] C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011.
[14] G.Neelakrishnan, S.N.Pruthika, P.T.Shalini, S.Soniya, "Perfromance Investigation of T-Source Inverter fed with Solar Cell" SurajPunj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp: 744-749
[15] R. Bhuvaneshwari , N. Mathubaala , Pranusha S Bavan , Dr. P. Lakshmi Harika, Dr. M R Sumalatha, ''Detection of Autism Spectrum Disorder using Machine Learning", 2022.
[16] C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor & Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011.
[17] G.Neelakrishnan, P.Iraianbu, T.Abishek, G.Rajesh, S.Vignesh, "IOT Based Monitoring in Agricultural" International Journal of Innovative Research in Science, Engineering and Technology, March 2020, Volume 9, Issue 3, pp: 814-819