International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Comparative Analysis of Machine Learning Algorithms for Early Detection of Neurological Disorders: Autism and Parkinson's

Asst. Prof. Pavan korlaharlli[1], Mrunalini Patil[2], Anjali Rankalakoppa[3], Tushar Kutre[4], Pratiksha Dhamnekar[5]

[1,2,3,4,5]Department of Computer Science and Engineering, Angadi Institute of Technology and Management, Belagavi-590009, India

A B S T R A C T

This paper provides a comprehensive comparative analysis of machine learning algorithms including CNNs, SVMs, Random Forests, KNN, Decision Trees, Naive Bayes, GBM and logistic regression for the detection of autism spectrum disorder (ASD) and Parkinson's disease (PD). The study uses specific data sets and different evaluation metrics to evaluate precision, accuracy, recall and F1 score. The goal is to identify the most effective algorithms for early detection and accurate classification of these neurological conditions, potentially improving diagnostic tools and intervention strategies. In addition, the paper explores the use of machine learning and entropy measures in the automatic detection of neurological disorders such as autism disorder and parkinson's, disease emphasizing the generality of support vector machines and specific entropy measures for different disorders.

Keywords: Autism Spectrum Disorder (ASD), Parkinson's disease (PD), Neurological Disorders, Machine Learning, Comparative Analysis,Diagnostic Tools.

1. INTRODUCTION

In the realm of healthcare, timely detection of neurological diseases such as Autism Spectrum Disorder (ASD) and Parkinson's disease (PD) is pivotal for effective intervention and improved patient outcomes. This comparative study delves into the algorithms utilized for the detection of these disorders, unraveling their nuances, strengths, and limitations.

ASD, a developmental disorder characterized by challenges in social interaction, repetitive behaviors, and communication difficulties, has witnessed a transformative impact from machine learning algorithms. Decision trees and ensemble methods, among others, scrutinize datasets encompassing behavioral observations and speech patterns for ASD detection. Natural Language Processing (NLP) algorithms, examining speech through prosody and vocabulary analysis, contribute valuable insights. This study contributes to improving the lives of individuals affected by Autism and Parkinson's through more accurate and timely detection. On the other hand, PD, a neurodegenerative disease marked by motor symptoms like tremors, bradykinesia, and postural instability, demands a distinct set of algorithms. Analyses of neurological assessments, including electromyography (EMG) and movement patterns, form the core of PD detection algorithms.

The ultimate goal is to enhance the lives of individuals affected by ASD and PD through more precise and timely detection. Early diagnosis not only facilitates better management of symptoms but also alleviates the social and economic burdens on families. With the prevalence of ASD on the rise, early detection becomes crucial, and machine learning models like the Extreme Gradient Intensify algorithm offer a cost-effective pre-screen test, empowering families and caregivers in the diagnostic process. This research advances diagnostic methodologies for Autism Spectrum Disorder and Parkinson's disease, promising improved outcomes. By scrutinizing machine learning algorithms and their applications, the study seeks to usher in a brighter future for individuals facing neurological disorders, offering timely interventions and enhancing overall quality of life.

2. METHODOLOGY

Developing model to detecting neurological disorders such as autism and Parkinson's that involves several key steps. Below figure shows the flowchart of the proposed approach.

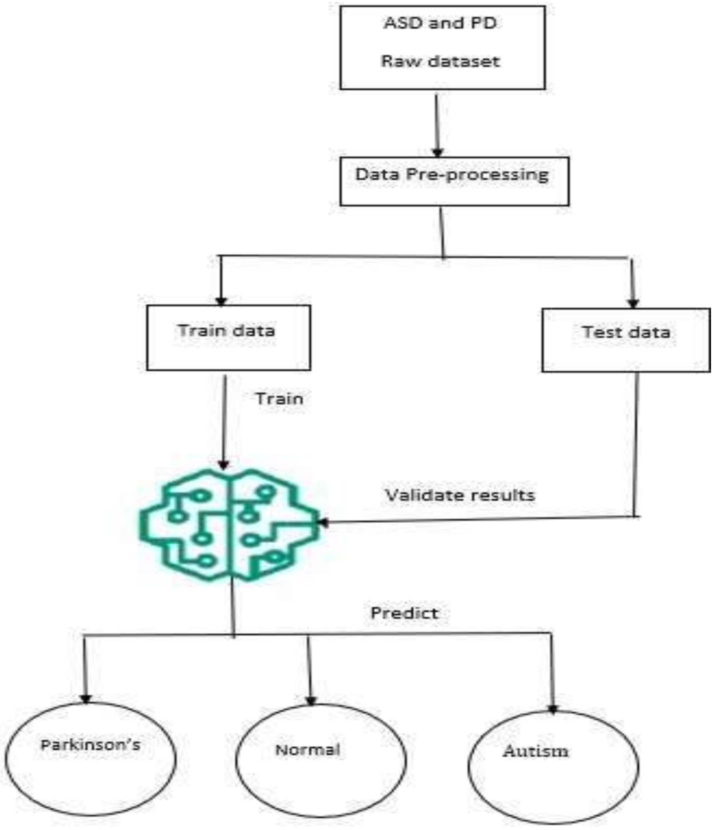

Fig. 1 - flowchart of proposed model

2.1 Dataset

The methodology proposed in this study involves the acquisition of audio data from the PPMI and UCI databases, focusing on voice modulations in individuals with Parkinson's and Autism Spectrum Disorder (ASD). The dataset includes crucial features such as jitter, shimmer, and MDVP obtained from vowel phonations. Following data collection, a rigorous preprocessing phase is implemented, accompanied by in-depth analysis and visualization to extract meaningful insights from these attributes.

The dataset is then partitioned, with 75% of the samples designated for training four distinct machine learning models: Logistic Regression, Support Vector Machines (SVM), Random Forest Regressor, and K Nearest Neighbors. These models are specifically tailored to classify the provided audio data into two categories: individuals diagnosed with either Parkinson's disease or Autism Spectrum Disorder. The training process leverages the observed frequency variations in the data to enhance the models' ability to accurately categorize patients based on their voice modulations.

2.2 Data Preprocessing

Data wrangling and pre-processing techniques are crucial stages in preparing datasets for analysis. Data wrangling involves cleaning the dataset and managing missing attributes, ensuring a more robust analysis. Figures depicting the noise-to-harmonic tone (NHR) ratio and harmonic tone to noise ratio (HNR) for individuals with Parkinson's disease (PWP) showcase how speech noise increases with the progression of the disease, resulting in higher NHR values. This increase in noise is visually apparent, with a notable skew in the data and instances of low NHR values (such as 0.3), indicating poor voice quality.

In data pre-processing, raw data is transformed into a more meaningful and comprehensible format. Real-world data often contains errors and null values, making it incomplete and inconsistent. Effective pre-processing methods, such as handling missing values through imputation, outlier detection, data discretization, and reduction techniques, are employed to address these issues and ensure the reliability of subsequent analyses.

2.3 Training and Testing Model

This study examines the performance of different machine learning models - logistic regressions, random forest classifiers, support vector classifiers and k-nearest neighbors - using three different approaches: using the full dataset of 195 records and 22 attributes. Uses a dataset reduced to 195 characters and 5 attributes after principal component analysis (PCA). Uses a balanced data set consisting of 109 records and 22 attributes.

- ***Logistic Regression (LR)***

Logistic Regression is employed for detecting Autism and Parkinson's disease based on relevant features in the dataset. The model is trained on data from PPMI and UCI databases, utilizing parameters such as tremor, flicker, and MDVP. Logistic Regression demonstrates efficacy in classifying individuals with these neurological disorders, contributing to diagnostic advancements.

- ***Random Forest (RF)***

Random Forest is used to detect autism and Parkinson's disease using features from the PPMI and UCI databases. This ensemble learning model, which includes multiple decision trees, proves effective in accurately classifying individuals with these neurological disorders. Multiple perspectives of individual trees increase the overall diagnostic power of the Random Forest model.

- ***Support vector Mechanism(SVM)***

Support vector machines (SVM) are used to detect autism and Parkinson's disease using data from the PPMI and UCI databases. SVM, a powerful classification algorithm, efficiently extracts and classifies individuals based on relevant features such as tremor, flicker and MDVP. The model shows the ability to recognize patterns associated with these neurological diseases.

- ***K Nearest Neighbors (KNN)***

K Nearest Neighbors is used to detect autism and Parkinson's disease using PPMI and UCI database data. The model classifies individuals based on proximity in feature space, showing efficiency in distinguishing patterns associated with their neurological diseases. KNN provides a localized and data-driven approach to classification.

2.4 Model Evaluation

Various metrics are used to evaluate autism and Parkinson's detection models, including accuracy, precision, recall, and F1 score. The datasets from the PPMI and UCI databases are shared for training and testing logistic regression, SVM, Random Forest and KNN models. These metrics evaluate models and#039; efficiency in accurately classifying individuals based on features such as tremor, flicker and MDVP. The goal of the evaluation is to identify a model that shows better performance, thus improving the diagnostic accuracy of autism and Parkinson's disease**..**

3. COMPARISION TABLE

TABLE 1: Comparison of displacement of all 6 cases

SN.	Paper Title	Authors	Year of publisl	Algorithms	Dataset	Accuracy
1	Application of Entropy for Automated Detection of Neurological Disorders with EEG signals	Ravinesh DEO & Aruna Devi	2023	SVM	1	93%
2	Early Detection of Parkinsor Disease using machine learnir	Aditi Govindu& Sushila Palwe	2023	Random Forest, Logistic Regression	3	Random Forest -> 91.83%
3	Disorder(ASD)in children a Adults using machine learning	Muhammad Shoaib Farooq & RabiaTehseer	2023	SVM, LR	2	Children SVM: 99% LR: 98% Adults SVM: 81 LR: 78%

4	Efficient Machine Learning Model For Early Stage Detection Of Autism Spectrum Disorder	Mousumi Bala, Mohammad Hanif Ali&Md Shahriare Satu	2022	Logistic Regression, SVM, NB, CNN, K-means	4	(For child dataset) 97.82(For toddler dataset), 95.87(For Adolescent SVM-> 99.61% dataset), 96.82% (For adult dataset)
5	using Machine Learni Techniques	Suman Raj & Sarfaraz Masood	2020	SVM, CNN, KNN, Logisti Regression	3	99.53%
6	Deep Learning-Based Parkinsons Disease Classification Using Vocal Feature Sets	Zehra cataltepe Yuvuz & hakan Gunduz	2019	CNN, SVM, KNN, DNN	2	Accuracy: >93.82% - KNN- CNN->88.25

4. CONCLUSION

The comprehensive study on machine learning algorithms for autism and Parkinson's disease detection provides crucial insights for researchers and practitioners. While offering valuable knowledge for informed decision-making in algorithm selection, it emphasizes the need to recognize potential variations in performance influenced by diverse factors. The study on machine learning algorithms for autism and Parkinson's detection provides vital insights, guiding algorithm selection. It stresses the need to recognize performance variations influenced by factors like dataset size. Emphasizing tailored approaches for disorder-specific features, ongoing research aims to enhance accuracy, promising improved diagnostic tools for optimal disease detection and treatment.

References

[1] Ravinesh DEO & Aruna Devi, Application of Entropy for Automated Detection of Neurological Disorders with EEG signals, Application of Entropy for Automated Detection of NDs With Electroencephalogram Signals, vol 11, July 9 2023.

[2] Suman Raj & Sarfaraz Masood, “Analysis and detection of Autism Spectrum Disorder using Machine Learning Techniques”, International Journal of Computer Science and Data Science, 2020.

[3] Muhammad Shoaib Farooq & RabiaTehseen, “Detection of Autism Spectrum Disorder (ASD) in children and Adults using machine learning”,” scientific reports”, 2023.

[4] Yuvuz, hakan Gunduz & Zehra cataltepe, “Deep Learning-Based Parkinson’s Disease Classification using Vocal Feature Sets,” International Journal of Economics and Finance, Vol. 7, Issue. 7, pp. 178-188, 2015.

[5] Aditi Govindu& Sushila Palwe ,“ Early Detection of Parkinsons Disease using machine learning”, International Journal of Engineering Trends and Technology (IJETT) – Volume 4, Issue 6, June 2013, Pg.2330-2337.

[6] Mousumi Bala, Mohammad Hanif Ali&Md. Shahriare Satu, “Efficient Machine Learning Models For early Stage Detection of Autism Spectrum Disorder”, International Journal of Civil, Structural, Environmental and Infrastructure Engineering Research and Development (IJCSEIERD), Vol. 3, Issue 1, Mar 2013, 59-66