Parkinson's Disease Detection Using Machine Learning

Uma Konar, Department of Information Technology, SIES Graduate School of Technology,Navi Mumbai, India, konaruma28@gmail.com

Pooja Borgaonkar, Department of Information Technology, SIES Graduate School ofTechnology, Navi Mumbai, India, poojaborgaonkar11@gmail.com

Megha Kaliyaparumal, Department of Information Technology, SIES Graduate School of Technology, Navi Mumbai, India, meghakperumal@gmail.com

Prof. Savita Lohiya, Department of Information Technology, SIES Graduate School ofTechnology, Navi Mumbai, India, savita.lohiya@siesgst.ac.in

Abstract - Parkinson's Disease (PD) is a disorder of the nervous system that affects many people. The disease has progressive stages of stopping all neural activity of any person. The latest technology focuses on Machine Learning (ML) and Deep Learning (DL) techniques to analyze effective PD data to make an effective predictivesystem. However, existing systems lack access to more data and more data analysis to predict a PD. To obtain an analysis of the various Parkinson's symptoms, neural implantation systems should be included in other features. The main aim of our project is to develop the Deep Neural Network based model do detect the Parkinson disease at early stages using Speech samples. Detecting Parkinson's disease with speech samples and Deep Learning requires huge processing Power. Till recent time such a huge processing power was not available at low cost. However, with the development of GPU architecture; required processing power and memory is available at affordable cost, which encourages the use of Deep Learning based approaches with high accuracy. Graphics Processing Unit is taken into account as heart of Deep Learning, a neighborhood of AI. It is one chip processor used for a faster artificial neural network. It is wont to implement the matrix operation of a neural network to reinforce the time performance of a text detection system. Artificial neural network (ANN) may bea computational model that consists of several processing elements that receive inputs and deliver outputs supported their predefined activation functions.

***Keywords* — Parkinson Disease, Machine learning, Graphics Processing Unit, Artificial Neural Network.**

I. INTRODUCTION

PD (Parkinson's disease) is defined within the brain's centerby the destruction of dopaminergic neurons in nigra pars compact. Many kinds of symptoms contribute to this neurodegeneration, including connectivity problems, bradykinesia, tone changes, and stiffness. Dysarthria is also seen in patients having Parkinson's disease; the motor system is characterized by disability, weakness, and lack of communication: affects breathing, telephone, speech, and prosody. As the disease's symptoms and course vary, Parkinson's disease is often undiagnosed for many years. Therefore, we need a more accurate diagnostic method to diagnose PD because more symptoms arise as the disease progresses, making it more challenging to cure PD.

PD's main shortcomings are voice tone, voice loss & voice, decreased stress, inappropriate silence, shortness of speech, inconsistent speech, slurred speech, and harsh & breathless voice (dysphonia). As capturing voice information is not an offensive thing, and we can do it easily with mobile devices' aid, the selection of voice-related symbols promises a potential detection tool. Owing to the first hidden signs, it is hard to discern early on. Owing to diagnosis delays, there is a massive strain on patients in the existing health care program. Difficulties in early detection of PD have promptedresearchers/scholars to develop experimental methods based on specific automated algorithms to differentiate safe controls in people with Parkinson's disease.

II. LITERATURE SURVEY

[1]. Separation is a significant feature in computer vision, proposed by Srilatha et al. (2019). Image segregation refers to incorporating one of the predefined categories of an image label, including image sensors, image processing,object detection, object detection, extraction of features, andseparation of objects. Many split methods are designed for image classification. The highest concentration is used with various partitions combined with several partition algorithms to detect the tumor used in

 DOI : 10.35291/2454-9150.2021.0283

image processing.

[2]. Shraddha et al. (2019) have suggested that the authors' Performance parameters are true, accurate, and accurate. The authors use various classified classifiers to detect intrusion detection. All classifiers used the NSL KDD database to see intrusion detection.

[3]. Mallikarjuna et al. (2020) presented a comparison based on the general and unconventional response in the form of back distribution. In the training phase, the outline sequence of standardized and non- normal walking, three classes A, B, C, D standard, Parkinson gait, Hemiplegic gait, and neuropathic gait data sets are compared standard data sets.

[4]. Abós et al. (2017) described that data feature without a priori model. Using the XGBoost algorithm for partitioning, the XGBoost algorithm benefits from regular reading or retraining does not guarantee customization/retransmission. However, when trained and maintained, the XGBoost learning approach is far more potent than Logistic retreat in solving real-world problems. An earlier report on using the XGBoost procedure to diagnose Parkinson's disease is determined based on their understanding. XGBoost provided 96% accuracy of data classification, and asset disposal provided 79% accuracy. This PD predictive system was developed comparing the accuracy of LR and XGBoost on train and test data.

III. NEED OF PROJECT

With the help of machine learning technology, we can quickly improve diagnosing Parkinson's disease. In recent days techniques of deep learning have got major success in the medical science domain. Not much investigation is to been done for applying deep learning techniques for Parkinson's disease detection. Therefore, in the earlier stage itself, there is a need to enhance the identification rate or reliability of Parkinson's diseases to reduce the false prediction rate.

HARDWARE REQUIREMENTS:

Processor: Intel i3 RAM: 4GB or more

Hard disk: 16 GB or more

SOFTWARE REQUIREMENTS:

Windows OSPython Python IDE

IV. OBJECTIVE

The Objective of the Proposal is to develop the Deep NeuralNetwork (2D convolution neural network) based model do detect the Parkinson disease at early stages using Speech samples.

V. IMPLEMENTATION

The developed framework is used in a device with just an Intel Core i5-5200U CPU @ 2.20GHz and 8 GB RAM. To use the Deep Neural Network, Python Library, TensorFlow (tf. estimator) has been used.

Collection of Data

We all used the UCI Machine Learning Database Parkinson Telemonitoring from Voice Data Collection. There are bio- medical voice assessments for 42 patients in the database. Numbers of subjects, years of study, gender by topic, travel time, Total UPDRS, Motor UPDRS, & biomedical voice measurements are various data features. There are 5,875 audio samples of all these patients in the database. The details are in CSV format with ASCII. On average, approximately 200 patients are collected (recognition can be made using first adjective number).

Execution of Data

We used the min-max norm to render the database at 0-1 distance. The normalization is performed wisely by the column using equation: when x = column value, min (x) = column value minimum, and maximum value (x) = column value maximum.

Building the DNN

The cumulative total UPDRS for the data ranges from 5,0377 to 54,992, while for car-UPDRS scores, at least5,0377 to 39,511. We have constructed train and test data sets by dividing the standard database into 80% (train database) and 20% (test database) components. The train and test set sets are also divided into two points: the UPDRS motor and the full UPDRS scale, keeping each of these schools as an outputvariant in their corresponding files. Standard values for seventeen live voice steps namely Jitter RAP, Shimmer: APQ3 Jitter (%), Jitter (Abs), PPQ5, Shimmer (dB), DDP, Shimmer: APQ11, Shimmer DDA, APQ5, HNR, DFA, RPDE, NHR, PPE selected as segregation features. Release categories are ‘not difficult’ and ‘difficult.’ We have defined the range of metrics for the problematic and innocent classes as shown in Table 1 due tothe price limit on the database.

Table 1. Severity Class range

Metric	Severe	Non-severe
Total-UPDRS	Above 25	0-25
Motor-UPDRS	Above 20	0-20

Fig.1. Severity Class range

The algorithm needs the dataset's input, creates input pipelines, and defines iterators to search the data. The algorithm mentioned also offers the function to shape a data set randomly. Stage 2 is to inject the input data into the testing set after the input pipeline is specified using lambda.After data collection, the model offers preparation, assessment, and prediction. Training is done only in the processing system by defining the hidden layer arrays with the pre-initialized value to the layer, model development,

 DOI : 10.35291/2454-9150.2021.0283

and saving. Finally, it tests the resulting DNN classifier. Forthe creation of a DNN Classifier, TensorFlow using Keras as a reference. In each of the three hidden layers of our neural network, seventeen input layer units, 10, 20, and 10 neurons. The network has been further trained in 1000 and 2000 phases.

VI. METHODOLOGY

As the Machine learning techniques can vastly improve the process for accurate diagnosis of Parkinson disease. As we know in the recent days deep learning techniques have achieved major success in the Medical domain. Not much investigation is been done so to applying deep learning techniques for Parkinson disease detection. So there is need to improve the detection rate or accuracy of Parkinson's diseases in the earlier stage itselfso that we can decrease the false prediction rate.

The methodology proposed using deep learning to predict Parkinson's disease's seriousness is defined in Fig. 2. The first stage is to obtain voice data for analysis from patients with PD. The data is then normalized by min-max normalization. The following deep stage consists of the input, hidden, and output layers in the neural network. In the input layer, the total number of neurons is the same as the total number of input data attributes. The output layer includes two neurons, the "severe" or "non- severe" classes. For training and testing, normalized data are transferred to the deep neural network created.

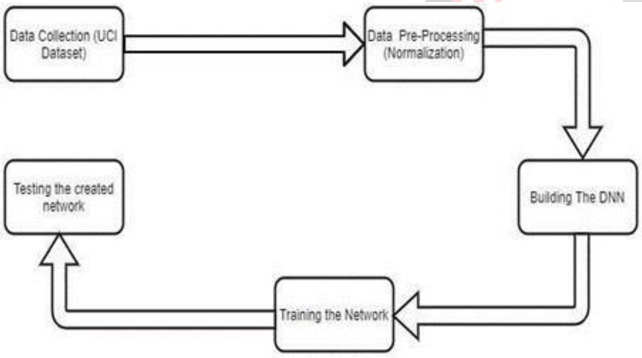

FIG.2.Proposed System

Algorithm	Primary Problem	Predictors	Power	Raw implementation	Inter-pretability	Regression also	Normalization
Logistic Regression	Binary	Numeric	Low	Easy	Good	No	No
Random Forest	Multiclass or Binary	Numeric or Categorical	High	Difficult	Good	Yes	No
SVM	Binary	Numeric or Categorical	High	Very Difficult	Medium	No	Yes
Neural Network	Multiclass or Binary	Numeric and Categorical	Very High	Very Difficult	Weak	Yes	Yes

Fig.3. Comparisons of Machine Learning Algorithms

Block DiagramDFD Level 0

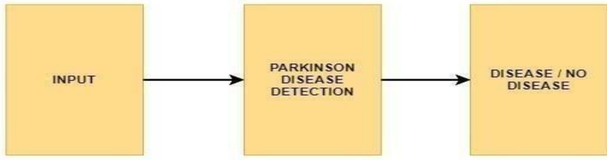

Fig.4.Data Flow Diagram L0

DFD Level 1

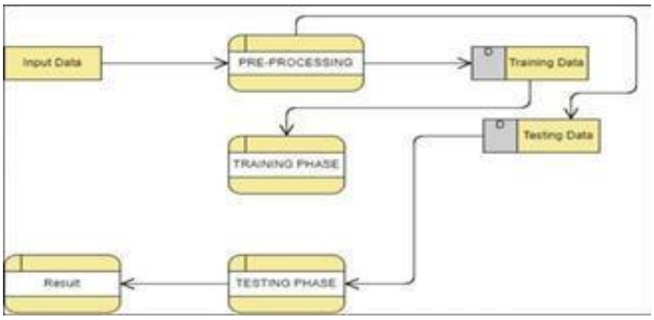

Fig.5. Data Flow Diagram L1

VII. CODING RESULT AND DISCUSSION

Disease identification and prediction is feasible through machine-driven machine learning architectures exploitation solely noninvasive voice biomarkers as options. Our analysis provides a comparison of the effectiveness of varied machine learning classifiers in sickness identification with clattery and high dimensional information.

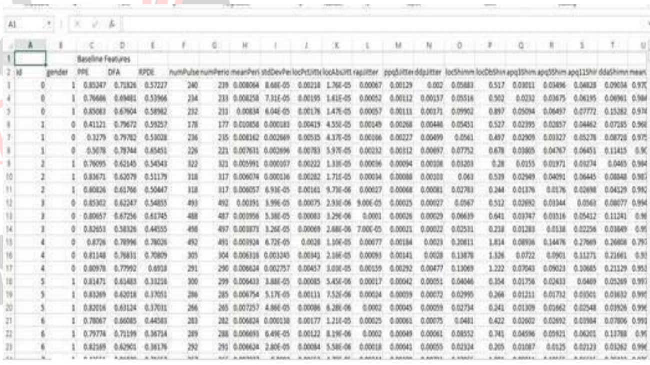

Fig.6.pd_feature Dataset

This is a big data which is used for implementation of PD and we are used this data as a input data. This dataset is composed of a range of biomedical voice measurements. The parameters are classified into 6 categories. It has755 rows. Each column in the table is particular voice measure. The class parameter is the most important among all parameter which will differentiate healthy people from those with Parkinson's disease. 0 state that person is healthy1 state that person has Parkinson's disease.

 DOI : 10.35291/2454-9150.2021.0283

Fig.7.Prediction Result

When thorough feature choice, clinical level accuracy is feasible. These results are promising as a result of they will introduce novel suggests that to assess patient health and neurologic diseases exploitation voice information. Thanks tothe high accuracy performed by the models with these short audio clips there's reason to believe denser feature sets with word, or different modalities would aid in sickness prediction and clinical validation of identification within the future.

VIII. CONCLUSION

Artificial Neural Network based mostly diagnosing of medical diseases has been taken into nice thought in recent years. ANNs are used to classify effective diagnosing of Parkinson's sickness. ANNs were accustomed differentiate between clinical variables of samples World Health Organization were littered with degenerative disorder and World Health Organization weren't. For this purpose, degenerative disorder knowledge set, taken from UCI machine learning info was used. Mean sq. normalized error perform was accustomed live the standard of our networks throughout trainings and direct performance calculations. it had been as certained that random forest is that the best set. This system will assist neurologists to create their final selections while not hesitation and additional shrewdly.

The proposed deep neural network model obtained superior reliability as compared to other strategies. It is also identified that the division based on the UPDRS motor marks is only the division based on the total points of the UPDRS and, therefore, can be concluded as the best Severity guessing metric. Although we have used a 5875-case database, our method's accuracy can be improved by using it in an extensive database, having many cases for each difficulty category, and an integrated database of the patient voice and another patient as handwriting and handwriting features.

REFERENCES

[1] Shu Lih Oh, Yuki Hagiwara, U. Raghavendra, Rajamanickam Yuvaraj, N. Arunkumar, M.Murugappan, U. Rajendra Acharya. "A deep learning approach for Parkinson's disease diagnosis from EEG signals." Computer Aided Medical Diagnosis. The Natural ComputingApplications Forum 2018. (9 August 2018).

[2] Kamal Nayan Reddy Challa, Venkata SasankPagolu, GanapatiPanda, Babita Majhi. "An Improved Approach for Prediction of Parkinson's Disease using Machine Learning Technique. International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)-2016. 26 Oct 2016.

[3] Kamagata, Koji, et al. "Connectome analysis with diffusion MRI inidiopathic Parkinson's disease: Evaluation using multi-shell, multi-tissue,constrained spherical deconvolution." NeuroImage: Clinical 17 (2018):518-529.

[4] Tripathi, Ashish Kumar, Kapil Sharma, and Manju Bala. "A novelclustering method using enhancedgrey wolf optimizer and mapreduce." Big data research 14 (2018): 93-100.

[5] Tripathi, Ashish Kumar, Kapil Sharma, and Manju Bala."Dynamic frequency based parallel k-bat algorithm for massive data clustering (DFBPKBA)." International Journal of System Assurance Engineering and Management 9.4 (2018) 866-87.

 DOI : 10.35291/2454-9150.2021.0283