https://doi.org/10.5281/zenodo.12726444

ISSN: 2096-3246
Volume 53, Issue 02, June 2021

Deep learning-based Parkinson's Disease prediction using Recurrent Neural Networks

KALWAKURTHI SRI SANDHYA, K BALAJI SUNIL CHANDRA, G BALA RENUKA
Assistant Professor [1,2,3]
k.srisandhya26@gmail.com, hod.cse@svitatp.ac.in, GOLLA.BALARENUKA@GMAIL.COM
Department of CSE, Sri Venkateswara Institute of Technology,
N.H 44, Hampapuram, Rapthadu, Anantapuramu, Andhra Pradesh 515722

Keywords:
Deep learning, Parkinson's Disease prediction, Recurrent Neural Networks, machine learning, neural network**.**

ABSTRACT

Biomarkers produced from human language may be used to study neurological illnesses like Parkinson's Disease (PD). About one million individuals are impacted by PD, a neurodegenerative illness that progresses over time. The severity of Parkinson's disease has been evaluated by doctors using subjective grading systems in the past. Finding and diagnosing Parkinson's disease via articulation is made feasible by difficulties in motor manipulation. Medical practitioners should reap the benefits of more affordable and precise diagnoses as a result of technology advancements and the widespread usage of sound storage devices in everyday life. utilising a decision-tree, logistic-regression, and Naive Bayes dataset as well as deep learning rule sets such Recurrent Neural Networks (RNN) for use predictions with pricing, we provide evidence to support this concept here utilising an audio dataset obtained from individuals with or without PD. Reliable and thorough evaluation of every deep learning and machine learning algorithm.

https://doi.org/10.5281/zenodo.12726444

ISSN: 2096-3246
Volume 53, Issue 02, June 2021

I. INTRODUCTION

Damage to cells in the brain that produce dopamine is the underlying cause of the progressive loss of movement abilities seen in Parkinson's disease, a neurodegenerative condition. Its effects may be devastating, including tremors, behavioural issues, shift difficulties, dementia, and a general lack of hope. People with these main motor disorders are known as "Parkinson's" or "Parkinsonian patients."

Studies on afflicted individuals' speech have shown that changes to their voice are among the most infrequent symptoms. As the condition worsens, the patient's speech grows weaker and stutters more. The analysis of beeps and other unstructured data is becoming more popular with the use of deep learning. In deep neural networks, many neurons are often used, type and function selection models stack these levels together. 'Excessive' and 'not intense' categories of patient voice data are created in this study using deep dating. The criteria for evaluation in this research were the total and motor UPDRS scores, which are part of the Unified Parkinson's Disease Rating Scale. On a scale from 0 to 108, the UPDRS motor evaluates the subject's motor potential; on a scale from 0 to 176, the total UPDRS evaluates the patient's mean ability. As a result of difficulties in early diagnosis, researchers have devised screening approaches that use automated algorithms to differentiate between healthy controls and individuals with Parkinson's disease. Unfortunately, the release does not include a differential analysis method that may distinguish between Parkinson's disease and other conditions that mimic its symptoms, such as essential tremor and Lewy body dementia. Because of the diagnostic backlog, the fitness instructor and her patients are under a lot of pressure. Scientists have developed computerised screening algorithms to differentiate between healthy individuals and people with PD since early diagnosis of PD is notoriously difficult. It is possible to differentiate between control and contamination using these biomarkers. Even so, they aren't considered at all anymore while making a differential diagnosis. is PD a genetic disorder or one of various illnesses that might provide symptoms similar to PD. Recent research have shown promising results, which might pave the way for the eventual provision of doctors with an optional set of helpful screening rules for Parkinson's disease. For these analyses, we train the system to differentiate between PD and controls using a variety of model- and algorithm-acquired facts from the Power Voice dataset. Digital vital signs and fitness statistics were gathered using an iPhone app from participants with and without Parkinson's disease as part of Sage Bionetworks' Power Scientific Monitoring Research. To facilitate the linking of statistics between data sets and to safeguard individual privacy, each participant is given a unique fitness code.

II. LITERATURE SURVEY

III. Interest and scientific indicators, including as the presence of motor signs and symptoms, are often used to diagnose Parkinson's disease (PD). Conventional methods of diagnosis are prone to self-disaggregation and misclassification because to their reliance on the assessment of motions that are inherently indistinct to the human sight. Consequently, these red flags often go undetected, which makes early identification of

Parkinson's disease is challenging. To circumvent these issues and improve PD diagnostic and evaluation procedures, machine learning techniques were used to categorise PD and healthy controls or patients exhibiting comparable medical symptoms. The major goal of the suggested device is to develop a system that can accurately identify early-stage Parkinson's disease, therefore

ISSN: 2096-3246
Volume 53, Issue 02, June 2021
https://doi.org/10.5281/zenodo.12726444

overcoming the shortcomings of the current standard. By using spectrogram pictures and artificial neural networks (ANNs) in audio recordings, a large-scale photographic categorization of gait indicators was transformed into a CNN (Convolutional Neural Network) for accurate illness prediction. This is a very early stage of Parkinson's disease. [2]. Our goal is to create prediction models that may identify early cases of Parkinson's disease using healthy, natural medical practices by analysing data collected from the Patient Questionnaire (PQ) component of the Parkinson's Disease Rating Scale (MDS-UPDRS), which is extensively utilised by the Standardised Movement Disorder Society. Statistical analysis using boosted trees, support vector machines, random forests, and regression. To evaluate several machine learning methods, we used problem-based and protocol-by-protocol validation. These

in both healthy standard grading and early PD, the techniques are quite accurate and near to the ROC curve (>95% each). A statistically high level of agreement between the logistic model and the data suggests that it is a good prediction model. We conclude that by using machine learning to associate questionnaire questions, these prediction models may be able to aid doctors in diagnostic procedures [3]. For the purpose of identifying Parkinson's disease mostly via facial expressions, we have presented an unorthodox situation. To begin, we analysed the characteristics that would differentiate a healthy manipulating group from one affected by Parkinson's disease. According to statistical analysis, the most crucial skill is the ability to depict changes in fear expression over time. With a balanced accuracy between 0 and 69, the XGBoost classifier surpassed many.

IV. PROPOSED WORK

As shown in Figure 1 below, a user in the early step gathers the necessary data set to get a Parkinson's disease diagnosis. The computer can pre-process datasets and extract features after the facts are loaded. The computer then compares the characteristics to the model after extracting the abilities needed for the prediction, and

At last, the forecast may be put out as the ultimate verdict. All of the ML and DL algorithms that were taken into account in our study had their performance metrics visualised.

https://doi.org/10.5281/zenodo.12726444
ISSN: 2096-3246
Volume 53, Issue 02, June 2021

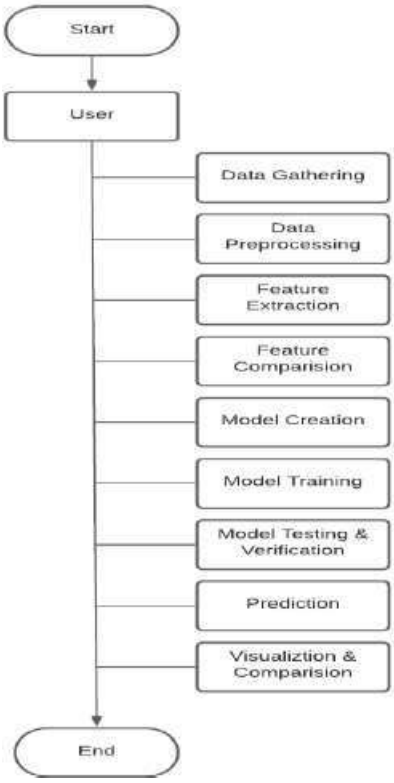

Fig.1 Proposed work architecture

V. IMPLEMENTATION

a) Data Collection

VI. Organoid acoustic (PD) measures were reported by 31 Parkinson's disease individuals in this cohort. There are 195 audio recordings of these people, with each row representing a different degree of audio (the 'connection' column). 'Reputation' column, which

is the principal metric for differentiating between healthy individuals and those with Parkinson's disease (PD) in the dataset, and it approaches 0 for health and 1. Information pertaining to this assignment is often stored in an ASCII-CSV format. Every line of the CSV file displays an audio recording. It would be difficult for six or eight bullets to lock the step, according to Miles.

a) Data Loading

Data analysis with a large number of remarkable assets is known as "fact- gathering." As an example, it is essential that we acquire and preserve records in a manner that is conducive to business logic. Data set mapping as a reference approach partitions the combined dataset as follows: 80% for instructional purposes and 20% for fact-finding. For the purpose of mapping fact points to the prior set and final points to the last set, the data sets are separated into training sets and examination sets for the modelling data set. In this way, the model is trained on a training set before being applied to a test set. The total performance of our utility may also be the primary metric by which it is assessed. Assuming they perform as expected, the next sixteen columns are considered secondary indicators in the dataset, whereas the first eight columns are considered the control group (main indicators). Since

ISSN: 2096-3246
Volume 53, Issue 02, June 2021
https://doi.org/10.5281/zenodo.12726444

parameters are associated with initial indications, they are unable to

remain a single integer. It is a collection of methods for determining the relative relevance of each item in the prediction process by scoring the input functions in the prediction version. A data-set-appropriate prediction model is used to ascertain the necessary maximum ratings.

Attribute Information

Matrix column entries (attributes)

name - ASCII subject name and recording number.

MDVP:Fo(Hz) - Average vocal fundamental frequency.

MDVP:Fhi(Hz) - Maximum vocal fundamental frequency.

MDVP:Flo(Hz) - Minimum vocal fundamental frequency.

MDVP:Jitter(%), MDVP:Jitter(Abs),MDVP:RAP, MDVP:PPQ,

Jitter:DDP - Several measures of variation in fundamental frequency. MDVP:Shimmer, MDVP:Shimmer(dB),Shimmer:APQ3,Shi mmer:APQ5,

MDVP:APQ,Shimmer:DDA - Several measures of variation in amplitude.

NHR, HNR - Two measures of ratio of noise to tonal components in the voice. status - Health status of the subject (one) - Parkinson's, (zero) – healthy.

RPDE, D2 - Two nonlinear dynamical complexity measures.

DFA - Signal fractal scaling exponent. spread1, spread2, PPE - Three nonlinear measures of fundamental frequency variation.

b) Data pre-processing

VII. Data that contains null or irrelevant facts may be filtered out in this module, and any unnecessary columns can be removed. In the log mining process, data preservation plays a crucial role. "Garbage in, garbage out" is a common adage in many projects, but it is especially true in platform mining and system learning projects. Data gathering practices are not always strictly controlled, which might lead to inaccurate results with missing or out-of-range statistics.

a) Model Creation

VIII. For the purpose of detecting fraudulent interests, this module targets clients according to their geographical areas and transactions. This machine learning technique sorts unlabeled input into predefined categories. "Data clustering is a method of arranging related fact factors into separate groups," would be a definition of this. Similar patterns in the unlabeled space include shape, length, shading, behaviour, etc.

information repository that sorts facts according to the existence or nonexistence of various trends.

a) Training and Testing

When using statistics planning, the data set collected divided into two components: 80 percent educational statistics and 20 percent test records. Facts were separated into school and departure units to map the former and latter information factors in the modelling data set. Thus, the model is taught to use the training set and then applied to the test set. Our utility can be evaluated in this way.

b) Prediction and Comparison

The model equipped to detect and predict Parkinson's disease is based on a specific data set. The truth function we got is the comparison. The easiest way to compare machine mastery algorithms is to check them for comparable data. When testing algorithms, we can also push them to be evaluated against a consistent testing tool.

c) Visualization

The generated graphical visualization provides a human-friendly way to examine and identify statistical traits, outliers, and patterns within statistics.

IX. EXPERMENTAL ANALYSIS

It mostly depends on how precise each set of rules detects a disease that affects the final results. Figure 2 and Figure 3 below show the variance plot for all ML and DL algorithms used in our task, indicating the slight difference between the algorithms in their accuracy and timestamp.

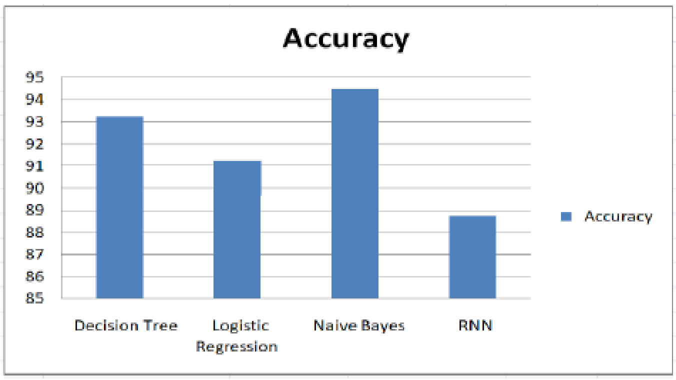

Fig.1 Accuracy Comparison Chart of all Algorithms

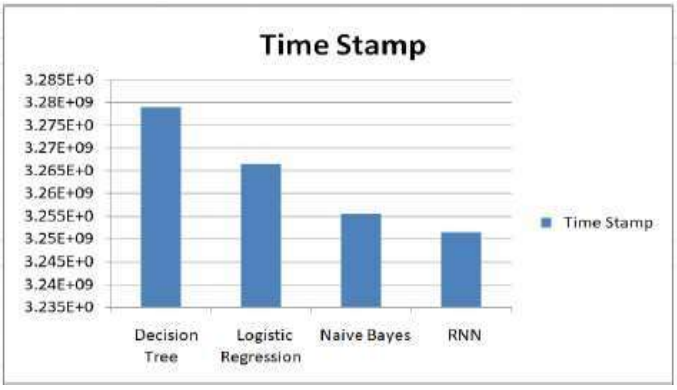

Fig. 2 Performance Comparison Chart of all Algorithms

X. CONCLUSION

- One of the great research articles on the topic of Parkinson's disease was based on the efficient invention of the disease by employing novel machine learning and deep learning computations for business subjects in a classic, elegant, and suspicious manner, with an eye towards special appearances. After doing thorough research and testing various methodologies, we have determined that deep learning is the best solution for speech impairment. Our findings also show that the suggested model is more efficient and produces superior outcomes in terms of accuracy. Here are the outcomes of the algorithms that prioritise precision:
- The decision tree is 93%.
- The logistic regression is 91.25%.
- The Bayesian clock is 94.5 %.
- The ratio is 88.75 %.

REFERENCES

Section of Computer Science at the Jaypee Institute of Information Technology in Noida, India, authored the paper "Parkinson Disease Detection Using Deep Neural Networks" with co-authors Ashish Tripathi and Anubhav Johri.
In their 2018 article published in the International Journal of Medical Informatics, Prashanth and Dutta Roy discuss the use of patient questionnaires and predictive modelling for the early detection of Parkinson's disease. The link to the article is https://doi.org/10.1016/j.ijmedinf.2018.09. 008.
the thirdRadim Burget and Justyna Skibińska's work on Parkinson's disease detection using

The 12th International Congress on Ultra-Modern Telecommunications and Control Systems and Workshops (ICUMT) will be held in Brno, Czech Republic in 2020 and will focus on "Changes of Emotions During Speech."
[4] "Predicting Parkinson's Disease Severity from Patient Voice Features." Genain, Huberth, and Vidyashankar. year 2014.
[5] "UPDRS tracking using linear regression and neural network for Parkinson's disease prediction" (Benmalek, Elmhamdi, & Jilbab, 2017). Paper published in 2015 in the International Journal of Emerging

A ES
ISSN: 2096-3246
Volume 53, Issue 02, June 2021
https://doi.org/10.5281/zenodo.12726444
Trends and Technology in Computer Science, volume 4, issues 1–19.
Warwick R. Adams. "Multi-characteristics of typing-related finger movement for early Parkinson's disease detection with high accuracy." 2017.
[7] "Dynamical Learning and Tracking of Tremor and Dyskinesia from Wearable Sensors." authored by Cole, Roy, De Luca, and Nawab. Journal of Rehabilitation Engineering and Neural Systems, 2018; 22:982-991.
Machine learning and deep learning models for Parkinson's disease diagnosis: a performance analysis, pp. 381-388, by P. Mounika and S. G. Rao [8].

In their 2018 paper titled "An Intelligent Computing Based Approach for Parkinson Disease Detection," authors A. Ranjan and A. Swetapadma discuss [9] on pages 1-3.
[10] "Building a Machine-Learning Framework to Remotely Assess Parkinson's Disease Using Smartphones," by O. Y. Chén et al., with reference to pages 3491-3500.