Comprehensive Analysis of Parkinson Disease Prediction using Vocal Parameters

Saish Shinde[1], Shreya Satav[2], Unnati Shirole[3], Sujata Oak[4]

[1,2,3,4]*Department of Information Technology Ramrao Adik Institute of Technology*

Nerul, Navi Mumbai

[1] saishshinde786@gmail.com

[2] satav2011shreya@gmail.com

[3] unnatishirole12@gmail.com

[4] sujata.oak@rait.ac.in

Abstract— **There are lot of neurological disease which affect nervous system but Parkinson disease (PD) is one of the very current neurological disease. There has been a challenge of accurate diagnosis of Parkinson's Disease largely due to the close association of PD in other neurological diseases. Parkinson's disorder (PD) is a common neurodegenerative disorder that reasons harm to dopaminergic neurons withinside the Nigrostriatal pathway and launch of Dopamine into the striatum. Machine learning algorithms had been used as a device for predicting and diagnosing diseases. Some of those algorithms have received recognition because of their excessive recognition and overall performance. In PD recognition, research have established the effectiveness of numerous reputation and this systematic overview of the overall performance of machine learning of algorithms. The early stages of Parkinson's can be predicted using vocal or voice features . This are the frequency parameters which can predict the severity of Parkinsons'. Predict Parkinson's using person's vocal parameters.**

Keywords—**Machine Learning, Parkinson's, Voice, Vocal parameters, Data Science**

I. INTRODUCTION

Parkinson's disease may be a neurological ailment that ends in tremors, stiffness, and problem in on foot, stability, and coordination. Parkinson's symptoms normally begin slowly and collect worse over the years. due to the fact the ailment progresses, human beings may also have trouble walking and speaking. Parkinson's disorder can be a modern systema nervosum sickness that impacts movement. signs begin step by step, occasionally with diffused tremors. Tremors are common, however the sickness additionally causes stiffness or moderate motion. within the early stages of brain ailment, your face may additionally show very little appearance. Your fingers may not move after you stroll. Your speech is also slurred or distorted. The signs and symptoms of Parkinson's syndrome degenerate as your situation progresses through the years. for the duration of this venture we are going to predict the primary stage and whether mind ailment is gift or not. it were 100 years earlier than main brain changes have been first observed inside the brains of people with brain ailment, and every other 50 years later before many professionals recounted that these modifications have been absolutely a

system of the ailment itself. in the early 1960's knowledge the importance of the mind chemical, dopamine (di-ortho-phenyl-alanine), became quickly observed by the number one effective treatment for encephalopathy, L-Dopa. remedy for Parkinson's syndrome turned into worse before this. Parkinson himself encouraged that the neck be amputated on the rear which the cork be inserted to forestall the wound from restoration. He thought that the pus from the contamination represented the diseased fluid that brought on the ailment so inhaling the substance should enhance the patient. As you will be able to imagine, this wasn't L-Dopa become also much less popular first of all, even though it had a manner higher clinical history, as it brought on nausea and vomiting. Carbodopa was delivered quickly thereafter, to forestall vomiting, ensuing in a mix of medication, carbidopa / levodopa, on the market like Sinemet (sine = externally, imesis = cleaning). This drug is frequently one in all of the only tablets we've got to treat the symptoms of Parkinsonism. once we discuss treatment Parkinsonism presently simplest way treating the symptoms. the vital ailment is an assault of nerve cells in the mind, and, to a lesser volume, outdoor the brain. Our cutting-edge medicinal drugs are shown to decorate signs, which means by means of helping to restore ordinary analytical balance in the brain, we improve tremors, firmness, slowing down, movement, and so forth., but we have were given no longer truely reversed the method that reasons the damage. it is very kind of like treating a cold. we are taking medication that produces the inflammatory sickness less painful, the cough much less excessive and greater painful, however we've done not anything to save you the virus this is inflicting the problem. Our frame fights it in the next few days, even as we take medicinal drug for more relief. regrettably, for PD, the contamination isn't always improving.

1.1 Motivation

Parkinson's syndrome may be a neurological disease that finally ends up in tremors, stiffness, and problem in on foot, stability, and coordination. Parkinson's signs normally begin slowly and discover worse over the years. because the disorder progresses, people may also have trouble taking walks and speakme. they will actually have mental and behavioral adjustments, sleep problems, depression, reminiscence troubles

and fatigue. both men and girls can get Parkinson's. however, the disorder impacts about 50 percent of fellows over women. the commonplace lifespan of an individual with PD is commonly the same as that of individuals with out the ailment. thankfully, there are many treatment options available to humans with PD. however, in latest tiers, PD may not solution medication and will be associated with extra extreme issues like choking, pneumonia, and falls. PD could be a gradual-shifting trouble. it is impossible to predict what direction the sickness will absorb every character. So all through this project we're going to predict paralysis agitans in some methods. within the early stages of encephalopathy, your face may additionally show little or no appearance. Your arms won't circulate after you stroll. Your speech could also be slurred or distorted. The symptoms of Parkinson's disease become worse as your situation progresses over the years.

1.2. Scope

Assist a consultant in predicting whether a affected person has Parkinsons or not. educate the model to induce the most accurate result the use of the concept of sound and acoustic parameters. to use differing kinds of inputs and soar to guessing the use of the maximum current actual database and switch as enter. using Voice statistics that creates system results a great deal in the direction of real consequences that will increase accuracy. To evaluate the data of wholesome and infected sufferers to induce the perfect effects.

II. LITERATURE SURVEY

The research was conducted on the techniques for obtaining PD with help of therapeutical and using research approach. [1] plan and apply the approach of Artificial Neural Network for predicting Parkinson Disease manifestations. Artificial Neural Network approach are used to identify patient gestures that are ongoing. The worker used Artificial Neural Network-based sample coordinating approach. Various movement patterns of patients are registered and trained for complete diagnosis. The work has yielded acceptable results for PD analysis in similar clinical settings. At the same time, this work has experienced limitations in analyzing various aspects. [2] suggested a determined forecast method and stages for PD evaluation. This study is done on data processing techniques, folding validations, and ML methods. Neuro data features and seismic data features were analyzed to predict symptoms. The methods used in this work provided significant outcomes but insufficient in PD forecast and response level. A lot of testing activities were inherited as a result of the different theories of forecasting PD. In this way, [3] found a result using a evaluation system. The system is built to detect voice disorders. Patient voice recordings are identified as the main symptoms of the disease. Voice variability was assessed using Random Forest, Decision Tree and Reinforcement Learning approaches. There where a lot of concurrent crises faced. [4] based on the research of inclusion/exclusion criteria and summarizing information from SID, MagIran, PubMed, ProQuest, ScienceDirect and Google Scholar) we get that from 10980 studies 82 of them were selected. In order to diagnose the 'disease 52 of them used clinical indicators, 2 of them had used genetic characteristics, 12 used MRI, 2 used PET, 5 used SPECT ,2 used laboratory markers. [5] created a model for forecasting PD parameters on a website. That step uses a lot of Machine Learning and insights are achieved in an good way. That approach had worked good for the use of the PD. But very least results were generated as compared to Deep Learning-based PD approach. [6] studied and applied Parkinson's techniques for the solution of non-motorized values. That study suggested the work can be used to create and improve future solutions. For that reason, the study help to gather evidence of a lot of medicines and created a lot of solutions. [7] the new approaches for achieving Synuclein protein degradation to detect symptoms of PD. The protein synthesis techniques are being used detecting Parkinson's. There was no approach taken form ML an DL models to train protein parameters. Different classifiers were used such as random forest, support vector machine, KNN, Naive Bayes. [9] 96 subjects (PD-MCI = 45; Parkinson's disease with normal cognition (PD-NC) = 51 subjects) were analyzed. The random forest method achieved a higher sensitivity than the decision tree model. [10] the system helps to forecast D signals with the help of voice parameters. In the research, the approaches of Support Vector Machine and Naïve Bayes were discussed with help of voice parameters. The study creates ML-based vocal forecast of models to predict PD manifestations. Study has been done to produce very authenticate forecast. This operation did not had the features of a limited data set.

III. PROPOSAL
3.1 Module Description
So in this project we are going to use Jupyter Notebook for implementing Machine Learning algorithms and Google Colab for implementing Deep Learning and image processing examples. For purpose of deployment, we are going to use Heroku cloud and deploy the model. We are going to use Node JS or React JS using Flask as backend.
The next section focuses on the proposed methodology and algorithms used. Fig 1 show the workflow of the system
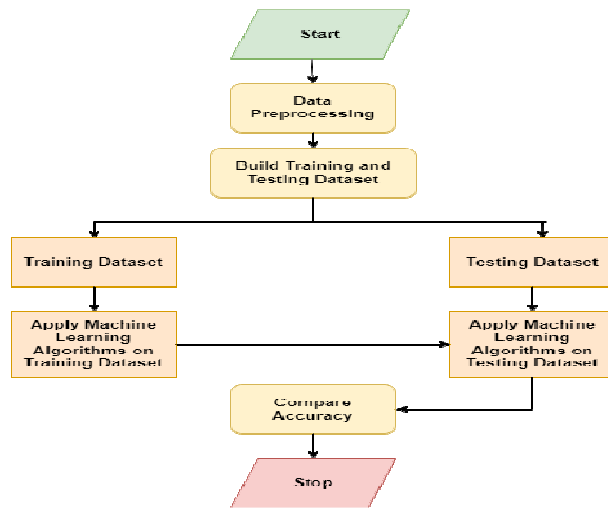
370

Fig. 1. Workflow of the proposed system

IV. PROPOSED METHODOLOGY

4.1 Details of Works Accomplished

Many machine learning algorithms where tested on the dataset from Kaggle based on Parkinson's Disease dataset. The parameters that where considered are average vocal fundamental frequency, maximum vocal fundamental frequency, minimum vocal fundamental frequency several measures of variation in fundamental frequency , several measures of variation in amplitude , Two measures of ratio of noise to tonal components in the voice, health status of the subject (one) - Parkinson's infected , (zero) – healthy, two nonlinear dynamical complexity measures, Signal fractal scaling exponent and three nonlinear measures of fundamental frequency variation. The various algorithms implemented include:

NAÏVE Bayes

It's a tribe of algorithms and not one algorithm. The Thomas Bayes classifier might be the straight probabilistic classifier that supports the use of theorems with strong naïve assumptions of independence (from Bayesian statistics). The Bayes theorem with associated grades of independence between the predictors was supported by classification method. Simply put, the presence of a special feature in an extremely class is assumed by a Naive Thomas Bayes category to be distinct from that of the other. For example, if this is red, spherical, and around a few inches' diameter, the fruit may also be considered as an association grade apple. Even though those characteristics are interdependent and dependent on other characteristics, a naïve classifier in Bayes would regard all these attributes separately to contribute to the likelihood that this fruit is an apple. The class label for each training data set is predicted by these learners. The class label which most of the models predict is voted using the majority voting procedure and the class label is agreed on in the training dataset. Regulations are produced from the ensemble models.

Decision Tree

The DT algorithm is a part of supervised algorithms. It is not like other algorithms, but it has a set of rules tree rules can be used to solve setbacks and classroom problems. a classy label that we start at the base of the tree. We compare the characteristics of the primary element with the report. In a comparative idea, we see a door that matches that price and jump to the next spot.

Random Forest

Random forest square measures the associated technique for classifications, regression and various tasks of learning that works by constructing an unpleasant call tree throughout coaching time and by producing the category in which individual trees are categorized or mean predictive. Random forests correctly match your coaching set to the custom of the call tree. Tin Kam Ho's random space methodology, which in the Ho formula could be a result of the 'stochastic discrimination' approach to categorization proposed by Eughe Kleinberg, was established as the fundamental guideline for random call forests. Leo Breiman and Adele Bargainer's have devised a partnership rule, and their trade mark is Random Forests. The addition combines the 'bagging' strategy and random alternatives presented by Breiman first by Ho and then by Amit and Geman in order to build a group of controlled variances called trees. Although random forests are intended naturally to only contain third-dimensional information, it was proven that they may be used for random objects using similarities just in pairs of objects.

KNN Classifier

It is utilized for regression and classification. The input includes the closest training instances in the data set in both situations. Input The outcome relies on how k-NN is used to classify or regress. k-NN is a classification type in that the function is locally approximated and all calculation is delayed until function evaluation is performed. Since this method is classified based on distance, it may be greatly improved by standardizing training data when the features represent distinct physical components or in vastly different sizes.

Logistic Regression

A mathematical model of a particular class or event, for example, passing / failing, winning / rebelling whether you are alive / dead or healthy / sick, is used to model the possibility of a particular class / event. This may be extended to illustrate other event classes such as a picture of a cat, a dog, a lion etc.
 it is possible. Chances between 0 and 1 are given for each item found in the picture at the total value of one.
Subtraction is a mathematical model that uses the function of order in its basic form to model a binary dependent variable, but there are many more complex variations. Object regression [1] (or log retrospective) analyzes the input model parameters (binary type reversal). Mathematically, a binary entry model contains two potential

values such as variability, for example pass / fail shown and index variance, where two values are marked '0' and '1.' In the order model, log-ods of the label label "1" are a combination of one or more independent variables ("predictors"), each of which can be binary (2 classes encrypted with index variables) or continuous variables (any real value). The probability relative to the value labeled "1," may range from 0 (yes value "0") to 1 (yes value "1"), hence labeling: order function, hence the name, function that translates log entry be possible. The next section, focuses on the benefits of project work. In the next section, we will focus on the results of using the various algorithms discussed.

4.2 Result and Analysis

Each value in the dataset is a different frequency or voice parameters. According to our research KNN Classifier gives the maximum accuracy of 94%. So it is been used in the proposed system. Below is the diagram of UI implementation.

Table 1: Accuracy Table

Algorithm	Accuracy
KNN Classifier	94%
Random Forest	92%
Gradient Boosting	91%
Decision Tree	91%
Logistic Regression	85%

Fig 2.Website UI for Parkinson Prediction

Fig 3. Predicted Output

Below is the accuracy table of different algorithms:

Table 2. Classification Report – KNN Classifier

Classification Report				
	Precision	Recall	f1-score	support
False	0.88	0.88	0.88	17
True	0.96	0.96	0.96	48
Accuracy			0.94	65
Macro avg	0.92	0.92	0.92	65
Weighted avg	0.94	0.94	0.94	65

Table 3. Classification Report – Logistic Regression

Classification Report				
	Precision	Recall	f1-score	support
False	0.77	0.59	0.67	17
True	0.87	0.94	0.90	48
Accuracy			0.85	65
Macro avg	0.82	0.76	0.78	65
Weighted avg	0.84	0.85	0.84	65

Table 4. Classification Report – Gradient Boosting

Classification Report				
	Precision	Recall	f1-score	support
False	0.87	0.76	0.81	17
True	0.92	0.96	0.94	48
Accuracy			0.91	65
Macro avg	0.89	0.86	0.88	65
Weighted avg	0.91	0.91	0.91	65

Table 5. Classification Report – Decision Tree

Classification Report				
	Precision	Recall	f1-score	support
False	0.87	0.76	0.81	17
True	0.92	0.96	0.94	48
Accuracy			0.91	65
Macro avg	0.89	0.86	0.88	65
Weighted avg	0.91	0.91	0.91	65

Table 6. Classification Report – Random Forest

Classification Report				
	Precision	Recall	f1-score	support
False	0.93	0.76	0.84	17
True	0.92	0.98	0.95	48
Accuracy			0.92	65
Macro avg	0.93	0.87	0.89	65
Weighted avg	0.92	0.92	0.92	65

V. Conclusion

In this system, we proposed a study of early Parkinsons disease prediction and would make it more accurate by various tests performed to predict the disease. We have applied various machine learning algorithms and made a predictive model to generate a final report on the persons disease health and would determine whether Parkinsons is present or not. In future we are going to implement direct application and voice as input which would automatically get converted into these parameters and detect it. Next, we would try to predict Parkinson's by various-by-various other methods to increase the accuracy. Considering certain parameters may also be useful in the future for further analysis as an EMG tremor analysis can be used to diagnose Parkinson's disease as well as significant tremors and improved seizures. It is readily available, non-invasive, and a cost-effective diagnostic tool. Further focus on patient history can be considered.

VI.REFERENCES

1. Salari, N., Kazeminia, M., Sagha, H. et al. The performance of various machine learning methods for Parkinson's disease recognition: a systematic review. Curr Psychol (2022). https://doi.org/10.1007/s12144-022-02949-8
2. Byeon H. Is the Random Forest Algorithm Suitable for Predicting Parkinson's Disease with Mild Cognitive Impairment out of Parkinson's Disease with Normal Cognition? International Journal of Environmental Research and Public Health. 2020; 17(7):2594. https://doi.org/10.3390/ijerph17072594
3. Aich, S.; Youn, J.; Chakraborty, S.; Pradhan, P.M.; Park, J.-h.; Park, S.; Park, J. A Supervised Machine Learning Approach to Detect the On/Off State in Parkinson's Disease Using Wearable Based Gait Signals. Diagnostics 2020, 10, 421. https://doi.org/10.3390/diagnostics10060421
4. Mei J, Desrosiers C, Frasnelli J. Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature. *Front Aging Neurosci.* 2021;13:633752. Published 2021 May 6. doi:10.3389/fnagi.2021.633752
5. Suppa, Antonio & Costantini, Giovanni & Asci, Francesco & Di Leo, Pietro & Alwardat, Mohammad & Di Lazzaro, Giulia & Scalise, Simona & Pisani, Antonio & Saggio, Giovanni. (2022). Voice in Parkinson's Disease: A Machine Learning Study. Frontiers in Neurology. 831428. 10.3389/fneur.2022.831428.
6. Rahman A, Rizvi SS, Khan A, Afzaal Abbasi A, Khan SU, Chung TS. Parkinson's disease diagnosis in cepstral domain using MFCC and dimensionality reduction with svm classifier. Mobile Inform Syst. (2021) 2021:e8822069. doi: 10.1155/2021/8822069
7. Cavallieri F, Budriesi C, Gessani A, Contardi S, Fioravanti V, Menozzi E, et al. Dopaminergic treatment effects on dysarthric speech: acoustic analysis in a cohort of patients with advanced Parkinson's disease. Front Neurol. (2020) 11:616062. doi: 10.3389/fneur.2020.616062
8. Asci F, Costantini G, Di Leo P, Saggio G, Suppa A. Reply to: Reproducibility of voice analysis with machine learning. Mov Disord.
(2021) 36:1283– 4. doi: 10.1002/mds.28601
9. Rusz J, Tykalova T, Ramig LO, Tripoliti E. Guidelines for speech recording and acoustic analyses in dysarthrias of movement disorders. Mov Disord. (2020) 36:803–14. doi: 10.1002/mds.28465
10. Saggio G, Costantini G. Worldwide healthy adult voice baseline parameters: A comprehensive review. J Voice. (2020) S0892-1997(20)30328-3. doi: 10.1016/j.jvoice.2020.08.028. [Epub ahead of print].
11. Tripoliti E. Voice tremor and acoustic analysis: finding harmony through the waves. Clin Neurophysiol. (2020) 131:1144– 5. doi: 10.1016/j.clinph.2020.02.017
12. Norel R, Agurto C, Heisig S, Rice JJ, Zhang H, Ostrand R, et al. Speech-based characterization of dopamine replacement therapy in people with Parkinson's disease. NPJ Parkinsons Dis. (2020) 6:12. doi: 10.1038/s41531-020-0 113-5
13. Lechien JR, Delsaut B, Abderrakib A, Huet K, Delvaux V, Piccaluga M, et al. Orofacial strength and voice quality as outcome of levodopa challenge test in Parkinson disease. Laryngoscope. (2020) 130:E896– 903. doi: 10.1002/lary.28645
14. Karapinar Senturk Z. Early diagnosis of Parkinson's disease using machine learning algorithms. Med Hypoth. (2020) 138:109603. doi: 10.1016/j.mehy.2020.109603
15. Ma A, Lau KK, Thyagarajan D. Voice changes in Parkinson's disease: what are they telling us? J Clin Neurosci. (2020) 72:1–7. doi: 10.1016/j.jocn.2019.12.029
16. Suppa A, Asci F, Saggio G, Marsili L, Casali D, Zarezadeh Z, et al. Voice analysis in adductor spasmodic dysphonia: objective diagnosis and response to botulinum toxin. Parkinsonism Relat Disord. (2020) 73:23– 30. doi: 10.1016/j.parkreldis.2020.03.012
17. Suppa A, Asci F, Saggio G, Di Leo P, Zarezadeh Z, Ferrazzano G, et al. Voice analysis with machine learning: one step closer to an objective diagnosis of essential tremor. Mov Disord. (2021) 36:1401–10. doi: 10.1002/mds.2 8508
18. Asci F, Costantini G, Di Leo P, Zampogna A, Ruoppolo G, Berardelli A, et al. Machine-learning analysis of voice samples recorded through smartphones: the combined effect of ageing and gender. Sensors. (2020) 20:5022. doi: 10.3390/s20185022