

Ideation Phase

Literature Survey & Information Gathering

Parkinson's Disease Detection using ML

Introduction to Parkinson's disease:

The recent report of the World Health Organization shows a visible increase in the number and health burden of Parkinson's disease patients increases rapidly. In China, this disease is spreading so fast and estimated that it reaches half of the population in the next 10 years. Classification algorithms are mainly used in the medical field for classifying data into different categories according to the number of characteristics. Parkinson's disease is the second most dangerous neurological disorder that can lead to shaking, shivering, stiffness, and difficulty walking and balance. It caused mainly due by the breaking down of cells in the nervous system. Parkinson's can have both motor and non-motor symptoms. The motor symptoms include slowness of movement, rigidity, balance problems, and tremors. If this disease continues, the patients may have difficulty walking and talking. The non-motor symptoms include anxiety, breathing problems, depression, loss of smell, and change in speech. If the above-mentioned symptoms are present in the person then the details are stored in the records. In this paper, the author considers the speech features of the patient, and this data is used for predicting whether the patient has Parkinson's disease or not. Neurodegenerative disorders are the results of progressive tearing and neuron loss in different areas of the nervous system. Neurons are functional units of the brain. They are contiguous rather than continuous. A good healthy looking neuron as shown in fig 1 has extensions called dendrites or axons, a cell body, and a nucleus that contains our DNA. DNA is our genome and a hundred billion neurons contain our entire genome which is packaged into it. When a neuron gets sick, it loses its extension and hence its ability to communicate which is not good for it and its metabolism becomes low so it starts to accumulate junk and it tries to contain the junk in the little packages in little pockets. When things become worse and if the neuron is a cell culture it completely loses its extension, becomes round and full of vacuoles.

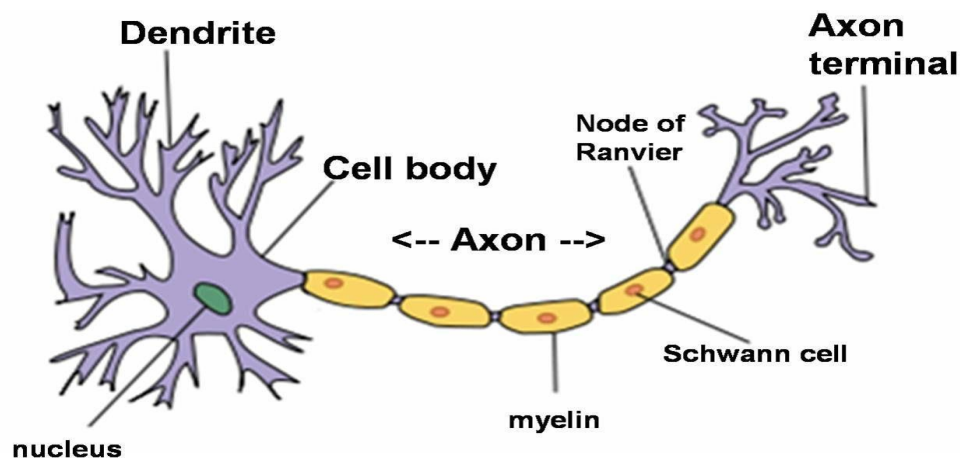


Fig-1.1 Structure of Neuron

This work deals with the prediction of Parkinson's disorder which is now a day is tremendously increasing incurable disease. Parkinson's disease is a most spreading disease which gets its name from James Parkinson who earlier described it as a paralysis agitans and later gave his surname was known as PD. It generally affects the neurons which are responsible for overall body movements. The main chemicals are dopamine and acetylcholine which affect the human brain. There is a various environmental factor which has been implicated in PD below are the listed factor which caused Parkinson's disease in an individual.

• **Environmental factors:** Environment is defined as the surroundings or the place in which an individual lives. So the environment is the major factor that will not only affects the human's brain but also affects all the living organism who lives in the vicinity of it. Many types of research and evidence have proved that the environment has a big hand in the development of neurodegenerative disorders mainly Alzheimer's and Parkinson's. There are certain environmental factors that are influencing neurodegenerative disorder with high pace are:-

- Exposure to heavy metals (like lead and aluminum) and pesticides.
- Air Quality: Pollution results in respiratory diseases.
- Water quality: Biotic and Abiotic contaminants present in water lead to water pollution.
- Unhealthy lifestyle: It leads to obesity and a sedentary lifestyle.
- Psychological stress: It increases the level of stress hormone that depletes the functions of neurons.

• **Brain injuries or Biochemical Factors:** The brain is the control center of our complete body. Due to certain trauma, people have brain injuries which leads some biochemical enzymes to come into the picture which provides neurons stability and provides support to some chromosomes and genes in maintenance.

• **Aging Factor:** Aging is one of the reasons for the development of Parkinson's disease. According to the author in India, 11,747,102 people out of 1,065,070,6072 are affected by Parkinson's disease.

• **Genetic factors:** Genetic factor is considered as the main molecular physiological cause which leads to neurodegenerative disorders. The size, depth, and effect of actions of different genes define the status or level of neurodegenerative disease which increases itself gradually over time. Mainly the genetic factors which lead to Neurodegenerative disorders are categorized into pharmacodynamics and pharmacokinetics.

• **Speech Articulation factors:** Due to the condition associated with Parkinson's disease (rigidity and bradykinesia), some speech-language pathology such as voice, articulation and swallowing alterations are found. There are various ways in which Parkinson's disease (PD) might affect the individual.

1. The voice get breathy and softer.
2. Speech may be smeared.
3. The person finds difficulty in finding the right words due to which speech becomes slower.

Parkinson's disease symptoms

The symptoms of Parkinson's disease broadly divided into two categories.

• **Motor symptoms:** This is a symptom where any voluntary action involved. It indicates the movement-related disorders like tremors, rigidity, freezing, Bradykinesia or any voluntary muscle movement.

• **Non-Motor symptoms:** Non motor symptoms include disorders of mood and affect with apathy, cognitive dysfunction as well as complex behavioral disorders. There are two other categories of PD which are divided by doctors: Primary symptom and Secondary symptom.

• **Primary symptoms:** It is the most important symptom. Primary symptoms are rigidity, tremor and slowness of movement.

• **Secondary symptoms:** It is a symptom that directly impacts the life of an individual. These can be either motor or non-motor. Its effect depends on person to person. A very wide range of symptoms is associated with Parkinson's,. Besides these symptoms, there are some other symptoms found that lead to Parkinson's disease. These symptoms are micrographic, decreased olfaction &

postural instability, slowing of the digestive system, constipation, fatigue, weakness, and Hypotension. Speech difficulties i.e. dysphonia (impaired speech production) and dysarthria (speech articulation difficulties) are found in patients with Parkinson's

Parkinson's disease symptoms:

The symptoms of Parkinson's disease broadly divided into two categories.

Motor symptoms: This is a symptom where any voluntary action involved. It indicates the movement-related disorders like tremors, rigidity, freezing, Bradykinesia or any voluntary muscle movement.

Non-Motor symptoms: Non motor symptoms include disorders of mood and affect with apathy, cognitive dysfunction as well as complex behavioral disorders. There are two other categories of PD which are divided by doctors:

Primary symptom and Secondary symptom.

Primary symptoms: It is the most important symptom. Primary symptoms are rigidity, tremor and slowness of movement.

Secondary symptoms: It is a symptom that directly impacts the life of an individual. These can be either motor or non-motor. Its effect depends on person to person.

A very wide range of symptoms is associated with Parkinson's. Besides these symptoms, there are some other symptoms found that lead to Parkinson's disease. These symptoms are micrographic, decreased olfaction & postural instability, slowing of the digestive system, constipation, fatigue, weakness, and Hypotension. Speech difficulties i.e. dysphonia (impaired speech production) and dysarthria (speech articulation difficulties) are found in patients with Parkinson's

Literature Survey:

Speech or voice data is assumed to be 90% helpful to diagnose a person for identifying the presence of disease. It is one of the most important problems that have to be detected in the early stages so that the progression rate of the disease is reduced. Many of the researchers work on different datasets to predict the disease more efficiently. In general, Persons with PD suffer from speech problems, which can be categorized into two: hypophonia and dysarthria.

Hypophonia indicates a very soft and weak voice from a person and dysarthria indicates slow speech or voice, that can hardly be understood at one time and this causes damage to the central nervous system. So, most of the clinicians who treat PD patients observe dysarthria and check out to rehabilitate with specific treatments to improvise vocal intensity. Lots of researchers did work on the pre-processing data and feature selection in the past.

Anila M and Dr G Pradeepini proposed the paper titled "Diagnosis of Parkinson's disease using Artificial Neural network". The main objective of this paper is that the detection of the disease is performed by using the voice analysis of the people affected with Parkinson's disease. For this purpose, various machine learning techniques like ANN, Random Forest, KNN, SVM, XG Boost are used to classify the best model, error rates are calculated, and the performance metrics are evaluated for all the models used.

The main drawback of this paper is that it is limited to ANN with only two hidden layers. And this type of neural networks with two hidden layers are sufficient and efficient for simple datasets. They used only one technique for feature selection which reduces the number of features.

Arvind Kumar Tiwari Proposed the paper titled “Machine Learning-based Approaches for Prediction of Parkinson’s Disease”. In this paper, minimum redundancy maximum relevance feature selection algorithms were used to select the most important feature among all the features to predict Parkinson diseases. Here, it was observed that the random forest with 20 number of features selected by minimum redundancy maximum relevance feature selection algorithms provide the overall accuracy 90.3%, precision 90.2%, Mathews 12 correlation coefficient values of 0.73 and ROC values 0.96 which is better in comparison to all other machine learning based approaches such as bagging, boosting, random forest, rotation forest, random subspace, support vector machine, multilayer perceptron, and decision tree based methods.

Mohamad Alissa Proposed the paper titled “Parkinson’s Disease Diagnosis Using Deep Learning”. This project mainly aims to automate the PD diagnosis process using deep learning, Recursive Neural Networks (RNN) and Convolutional Neural Networks (CNN), to differentiate between healthy and PD patients. Besides that, since different datasets may capture different aspects of this disease, this project aims to explore which PD test is more effective in the discrimination process by analysing different imaging and movement datasets (notably cube and spiral pentagon datasets). In general, the main aim of this paper is to automate the PD diagnosis process in order to discover this disease as early as possible.

If we discover this disease earlier, then the treatments are more likely to improve the quality of life of the patients and their families.

There are some limitations to this paper namely:

- They used the validation set only to investigate the model performance during the training and this reduced the number of samples in the training set.
- RNN training is too slow and this is not flexible in practice work.
- Disconnecting and resource exhaustion: working with cloud services like Google Collaboratory causes many problems like disconnecting suddenly. And because it is shareable service by the world zones, this leads to resource exhaustion error many times.

Afzal Hussain Shahid and Maheshwari Prasad Singh proposed the paper titled “A deep learning approach for prediction of Parkinson’s disease progression”. This paper proposed a deep neural network (DNN) model using the reduced input feature space of Parkinson’s telemonitoring dataset to predict Parkinson’s disease (PD) progression and also proposed a PCA based DNN model for the prediction of Motor-UPDRS and Total-UPDRS in Parkinson's Disease progression. The DNN model was evaluated on a real-world PD dataset taken from UCI. Being a DNN model, the performance of the proposed model may improve with the addition of more data points in the datasets.

T. J. Wroge, Y. Özkanca, C. Demiroglu, D. Si, D. C. Atkins and R. H. Ghomi, proposed the paper titled “Parkinson’s Disease Diagnosis Using Machine Learning and Voice” is that it explores the effectiveness of using supervised classification algorithms, such as deep neural networks, to accurately diagnose individuals with the disease. Historically, PD has been difficult to quantify and doctors have tended to focus on some symptoms while ignoring others, relying primarily on subjective rating scales. The analysis of this paper provides a comparison of the effectiveness of

various machine learning classifiers in disease diagnosis with noisy and high dimensional data. Their peak accuracy of 85% provided by the machine learning models exceeds the average clinical diagnosis accuracy of non-experts (73.8%) and average accuracy of movement disorder specialists (79.6% without follow-up, 83.9% after follow-up) with pathological post-mortem examination as ground truth.

Siva Sankara Reddy Donthi Reddy and Udaya Kumar Ramanadham proposed the paper “Prediction of Parkinson’s Disease at Early Stage using Big Data Analytics”. This paper describes mainly various Big Data Analytical techniques that may be used in diagnosing of right disease in the right time. The main intention is to verify the accuracy of prediction algorithms. Their future study aims to propose an efficient method to diagnose this type of neurological disorder by some symptoms at the early stage with better accuracy using different Big Data Analytical techniques like Hadoop, Hive, R Programming, MapReduce, PIG, Zookeeper, HBase, Cassandra, Mahout etc...

Daiga Heisters proposed the paper titled “Parkinson’s: symptoms, treatments and research”. This paper initially says that Current treatments can help to ease the symptoms but none can repair the damage in the brain or slow the progress of the condition; now, Parkinson’s UK researchers are working to develop new treatments that can and finally worked together to build on existing discoveries and explore these innovative areas of research, it is hoped that a cure for Parkinson’s will be found. Parkinson’s UK offers support for everyone affected,, including people with the condition, their family, friends and careers, researchers and professionals working in this area.

T. Swapna, Y. Sravani Devi proposed a paper and titled “Performance Analysis of Classification algorithms on Parkinson’s Dataset with Voice Attributes”. This paper deals with the application of seven classification algorithms on the acquired data set and then drawing out a comparison of the results to one another and also predicting the outcome whether the person is healthy or Parkinson disease effected from the given data. The results of the selected algorithms namely Naïve Bayes, Random Forest, Neural Networks, Decision Trees, AdaBoost, SVM, KNN were compared and tabulated. According to the outputs derived with the help of python, implementing Scikit Libraries. Final accuracy was calculated using these parameters. Random Forest algorithm gives with optimum accuracy of 78.56% which is closely followed by Decision Tree Algorithm with the optimal accuracy of 77.63%. Following the Decision Tree Algorithm is the MLP Classifier with an optimal accuracy of 76.72%, and lastly the Naïve Bayes Algorithm which has the optimal accuracy of 70.82%. Finally, these algorithms can help in classifying whether a person get effected with Parkinson’s disease or not.

M. Abdar and M. Zomorodi-Moghadam proposed a paper “Impact of Patients’ Gender on Parkinson’s disease using Classification Algorithms” [10]. In this paper, the author chooses the UCI PD dataset for finding the accuracy of Parkinson’s using SVM and Bayesian Network algorithms. The author chooses the most ten important features in the dataset to predict PD. The output variable is Sex and other factors are input, the author provides an approach for finding relationships between genders. The result obtained is SVM algorithm gives better performance than Bayesian Network with 90.98% accuracy. Dragana Miljkovic, et al, proposed a paper “Machine Learning and Data Mining Methods for Managing Parkinson’s Disease”. In this paper, the author concluded that based on the medical tests taken by the patients the Predictor part was able to predict the 15 different Parkinson’s symptoms separately. The machine learning and data mining techniques are applied on different symptoms separately and gives an accuracy range between 57.1% and 77.4% where tremor detection has the highest accuracy.

Sriram, T. V., et al. proposed a paper “Intelligent Parkinson Disease Prediction Using Machine Learning Algorithms”. In this paper, the author used voice measures of the patients to check whether the patient has Parkinson’s or not. The author applied the dataset to various machine learning algorithms and find the maximum accuracy. To analyse the models the author used the ROC curve and sieve graph. The random forest results with more accuracy i.e. are 90.26%.

Dr. R.GeethaRamani, G.Sivagami, and ShomonaGraciajacob proposed a paper “Feature Relevance Analysis and Classification of Parkinson’s Disease TeleMonitoring data Through Data Mining”. In this paper, the author used thirteen classification algorithms to diagnose the disease. The author used the Tele-monitoring dataset which contains biomedical voice features for evaluating the system. The aim of this paper is to predict motor UPDRS and total UPDRS from the voice measures.

A. Ozcift, proposed a paper “SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease” [1]. In this paper, the author summarizes that improve the PD diagnosis accuracy with the use of support vector machine feature selection. To evaluate the performances the author used accuracy, kappa statistics, and area under the curve of the classification algorithms. The rotation Forest ensemble of these classifiers used to increase the performance of the system.

REFERENCES

- [1] A. Ozcift, “SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease” *Journal of medical systems*, vol-36, no. 4, pp. 2141-2147, 2012.
- [2] Anila M Department of CS1, Dr G Pradeepini Department of CSE, “DIAGNOSIS OF PARKINSON’S DISEASE USING ARTIFICIAL NEURAL NETWORK”, *JCR*, 7(19): 7260-7269, 2020.
- [3] Arvind Kumar Tiwari, “Machine Learning based Approaches for Prediction of Parkinson’s Disease” *Machine Learning and Applications: An International Journal (MLAU)* vol. 3, June 2016.
- [4] Carlo Ricciardi, et al, “Using gait analysis’ parameters to classify Parkinsonism: A data mining approach” *Computer Methods and Programs in Biomedicine* vol. 180, Oct. 2019.
- [5] Dr. Anupam Bhatia and Raunak Sulekh, “Predictive Model for Parkinson’s Disease through Naive Bayes Classification” *International Journal of Computer Science & Communication* vol-9, Dec. 2017, pp. 194- 202, Sept 2017 - March 2018.
- [6] Dr. R.GeethaRamani, G.Sivagami, ShomonaGraciajacob “Feature Relevance Analysis and Classification of Parkinson’s Disease TeleMonitoring data Through Data Mining” *International Journal of Advanced Research in Computer Science and Software Engineering*, vol-2, Issue 3, March 2012.
- [7] Dragana Miljkovic et al, “Machine Learning and Data Mining Methods for Managing Parkinson’s Disease” *LNAI 9605*, pp. 209-220, 2016.
- [8] FarhadSoleimanianGharehehopogh, PeymenMohammadi, “A Case Study of Parkinson’s Disease Diagnosis Using Artificial Neural Networks” *International Journal of Computer Applications*, Vol-73, No.19, July 2013.
- [9] Heisters. D, “Parkinson’s: symptoms, treatments and research”. *British Journal of Nursing*, 20(9), 548–554. doi:10.12968/bjon.2011.20.9.548, 2011.
- [10] M. Abdar and M. Zomorodi-Moghadam, “Impact of Patients’ Gender on Parkinson’s disease using Classification Algorithms” *Journal of AI and Data Mining*, vol-6, 2018.
- [11] M. A. E. Van Stiphout, J. Marinus, J. J. Van Hilten, F. Lobbezoo, and C. De Baat, “Oral health of Parkinson’s disease patients: a case-control study” *Parkinson’s disease*, vol-67

2018, Article ID 9315285, 8 pages, 2018.

- [12] Md. Redone Hassan, et al, "A Knowledge Base Data Mining based on Parkinson's Disease" International Conference on System Modelling & Advancement in Research Trends, 2019.
- [13] Mandal, Indrajit, and N. Sairam. "New machine-learning algorithms for prediction of Parkinson's disease" International Journal of Systems Science 45.3: 647-666, 2014.
- [14] Mohamad Alissa, "Parkinson's Disease Diagnosis Using Deep Learning", August 2018.
- [15] Peyman Mohammadi, Abdolreza Hatamlou and Mohammed Msdaris "A Comparative Study on Remote Tracking of Parkinson's Disease Progression Using Data Mining Methods" International Journal in Foundations of Computer Science and Technology(IJFCST), vol-3, No.6, Nov 2013.
- [16] R. P. Duncan, A. L. Leddy, J. T. Cavanaugh et al., "Detecting and predicting balance decline in Parkinson disease: a prospective cohort study" Journal of Parkinson's Disease, vol-5, no. 1, pp. 131–139, 2015.
- [17] Ramzi M. Sadek et al., "Parkinson's Disease Prediction using Artificial Neural Network" International Journal of Academic Health and Medical Research, vol-3, Issue 1, January 2019.
- [18] Satish Srinivasan, Michael Martin & Abhishek Tripathi, "ANN based Data Mining Analysis of Parkinson's Disease" International Journal of Computer Applications, vol-168, June 2017.
- [19] Shahid, A.H., Singh, M.P. A deep learning approach for prediction of Parkinson's disease progression, <https://doi.org/10.1007/s13534-020-00156-7>, Biomed. Eng. Lett. 10, 227–239, 2020.
- [20] Shubham Bind, et al, "A survey of machine learning based approaches for Parkinson disease prediction" International Journal of Computer Science and Information Technologies vol-6, Issue 2, pp. 1648- 1655, 2015.
- [21] Siva Sankara Reddy Donthi Reddy and Udaya Kumar Ramanadham "Prediction of Parkinson's Disease at Early Stage using Big Data Analytics" ISSN: 2249 – 8958, Volume-9 Issue-4, April 2020
- [22] Sriram, T. V., et al. "Intelligent Parkinson Disease Prediction Using Machine Learning 68 Algorithms" International Journal of Engineering and Innovative Technology, vol-3, Issue 3, September 2013.
- [23] T. Swapna, Y. Sravani Devi, "Performance Analysis of Classification algorithms on Parkinson's Dataset with Voice Attributes". International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 2 pp. 452-458, 2019.
- [24] T. J. Wroge, Y. Özkanca, C. Demiroglu, D. Si, D. C. Atkins and R. H. Ghomi, "Parkinson's Disease Diagnosis Using Machine Learning and Voice," IEEE Signal Processing in Medicine and Biology Symposium (SPMB), pp.1-7, doi: 10.1109/SPMB.2018.8615607, 2018.