Deep Learning based prediction and monitoring of Parkinson’s Disease using Voice Data

Sohom Sen   
Computer Science and Engineering  
Herit*age Institute Of Technology*Kolkata, India  
sensohom2002@gmail.com

Rituparna Mondal  
Department Of Computer Application   
Techno India UniversityKolkata, India  
mondal.rituparna@gmail.comSamaresh Maiti  
Computer Science and Engineering  
Herit*age Institute Of Technology*Kolkata, India  
samareshmaiti02@gmail.com

Alok Kole  
Electrical Engineering  
RCC Institute of Information TechnologyKolkata, India  
alok.kole@rcciit.org.inSumanta Manna  
Computer Science and Engineering  
Herit*age Institute Of Technology*Kolkata, India  
mannasumanta9903@gmail.com

Banashree Chatterjee  
Information Technology  
Dr. B.C. Roy Engineering College Durgapur, India  
banashree2014@gmail.com

*Abstract*—This paper focuses on Deep Learning based prediction and monitoring of Parkinson’s Disease (PD) using voice data. PD is a progressive neurological disorder which affects the Central Nervous System (CNS) preventing the proper functioning of the same. This results in tremors, stiffness, slow movements, difficulty in balance and coordination, speech disorders and many more. As per recent studies and surveys conducted by World Health Organization (WHO), the mortality rate (per 100,000 population) in 1994 was 1.76 and reached 5.67 in 2019 and that in women increased from 1.63 (per 100,000 population) in 1994 to 4.81 in 2019. The incorporation of Machine Learning for early detection and assessment of the severity of Parkinson's disease is deemed imperative in light of the current circumstances. PD diagnosis is mainly on the analysis of symptoms, so speech recognition can introduce a new methodology of investigation in the diagnosis and monitoring of PD. Machine Learning models with Deep Learning features are proposed in this research work for the automatic detection and the severity of PD. In this proposed research work, a dataset based on the speech of PD patients and healthy people has been gathered. The dataset has been analyzed using acoustic features such as jitter, shimmer, intensity, pitch, etc. A Deep Learning model with Residual Neural Network architecture has been implemented for the prediction and monitoring of PD using voice data. Comparative performance analysis of the proposed Residual Neural Network architecture is performed with different Machine Learning models namely K-Nearest Neighbors, Support Vector Classifier, Decision Tree, Random Forest, Naïve Bayes, Logistic Regression, Extreme Gradient Boost and Gradient Boosting along with modern neural network techniques like Artificial Neural Network and Multi-Layer Perceptron. The results demonstrate that the proposed Residual Neural Network architecture yields favourable performance superior to that of the rest of the implemented other standard Machine Learning models in terms of accuracy, F1-score, precision, recall, AUC-ROC and AUC-PR.

Keywords—Artificial Intelligence, Dimensionality reduction, Machine Learning models, Neural Networks, Parkinson’s Disease, Residual Neural Networks, Voice dataset.

# Introduction

Parkinson's disease is a chronic and progressive movement disorder affecting the brain and central nervous system. It results from the degeneration of dopamine-releasing neurons, responsible for motor control. The gradual loss of dopamine leads to tremors, impaired movements, and other symptoms.

Despite extensive research, the exact cause of Parkinson's remains unknown, and a complete cure is yet to be found. The disease progresses slowly, often remaining unnoticed until the advanced stages.

Modern medical science, with the aid of machine learning, can play a crucial role in early and automatic detection. By identifying Parkinson's disease at an early stage, patients can receive supportive treatments to manage symptoms and slow down the progression of the disease.

In this article, eight (08) traditional machine learning (ML) and three (03) neural network models namely, K-Nearest Neighbor (KNN) Classifier, Support Vector Machine (SVM) Classifier, Decision Tree Classifier (DTC), Random Forest Classifier (RFC), Naive Bayes Classifier (NBC), Logistic Regression Classifier (LR), XGBoost Classifier (XGBC), Gradient Boost Classifier (GBC), Artificial Neural Network (ANN), Multi-Layer Perceptron (MLP) Classifier and Residual Neural Network (ResNet) have been implemented to classify whether a patient is suffering from Parkinson’s disease or not based on their speech patterns and thorough performance analysis of all the models implemented has been done by testing each of them using a dataset.

# Related Work

Agarwal et al. [1] proposed an efficient Extreme Machine Learning approach on a reliable UCI repository speech dataset of Parkinson's patients. They achieved 90.76% accuracy and 0.81 MCC, focusing on Neural Networks and Support Vector Machines.

Ouhmida et al. [2] used Convolution Neural Networks (CNN) and Artificial Neural Networks (ANN) on UCI datasets I and II, with 22 and 45 features respectively. CNN achieved 93.10% accuracy on database I.

Ogawa and Yang [3] detected Parkinson's Disease early using 10-layered 1-d CNN and novel ResNet on vocal features dataset. They achieved 0.888 accuracy, 0.928 F-measure, and 0.692 MCC in classification.

Aghzal and Mourhir [4] combined Histogram of Oriented Gradients with CNN to automate Parkinson's Disease detection based on handwriting patterns, achieving

87% accuracy and 83.21% F1-Score, surpassing clinical techniques.

Anand et al. [5] used machine learning and deep learning models with dimensional reduction techniques on UCI's Parkinson's Speech Dataset. Comparative analysis identified the best-performing model.

# Methodology

## Dataset

The dataset on which ResNet has been implemented is sourced from UCI Machine Learning Repository [6]. The dataset has been prepared with the help of 188 PD positive patients and 64 healthy individuals comprising 130 men and 122 women. Their age ranges from 33 to 87. The dataset has a dimension of 755 columns with 756 data. The dataset has 564 Parkinson’s positive data points and 192 Parkinson’s negative data points. The size of the dataset on disk is 5.3 mb.

The attributes involved in the classification of PD positive and negative patients consist of various speech signal processing algorithms such as Time-Frequency Features, MEL Frequency Cepstral Coefficients (MFCCs), Wavelet Transform based Features, Vocal Fold Features and TWQT Features.

Although the dataset is relatively small, the high amount of attributes although good for training the models made the implemented algorithm susceptible to overfitting, which has been taken care of by using Principal Component Analysis (PCA) to reduce the number of attributes without losing the training ability. This enabled the training of traditional ML algorithms and especially the Neural Network models to achieve a very high level of precision and accuracy. However, it is certain that with the inclusion of more data points, further improvement of the models would be feasible.

## Data Pre-Processing

Data pre-processing [7] is a very important step in machine learning. The goal of data preprocessing is to prepare and clean the raw dataset so that the efficiency and accuracy of the machine learning algorithms can be maximized.

The Processes gone through to prepare the dataset are:

1. ***Null Value Replacement****: Real World datasets can have missing data which needs to be replaced by the mean value of the missing attribute [8]. The dataset was checked for missing values and as the dataset did not have any nothing was needed to be done.*
2. ***Skewness Reduction****: Skewness measures attribute distribution asymmetry. High skewness can introduce bias in the model. In the code, attributes with absolute skewness greater than 1 are added to the "skewedCols" list. Table I lists attributes with highest and lowest skewness. Fig. 1 displays the distribution of ten(10) highly skewed attributes pre-skewness reduction[9].*

TABLE I.

|  |  |  |
| --- | --- | --- |
| **Attribute Name** | **Skew value before reducing** | **Skew value after reducing** |
| tqwt\_TKEO\_mean\_dec\_32 | 26.48 | 0.198 |
| tqwt\_TKEO\_std\_dec\_32 | 26.06 | 0.019 |
| tqwt\_TKEO\_mean\_dec\_33 | 24.94 | 0.310 |
| tqwt\_TKEO\_std\_dec\_33 | 24.28 | 0.075 |
| det\_TKEO\_mean\_3\_coef | 20.87 | 1.443 |
| det\_LT\_entropy\_shannon\_7\_coef | -21.41 | -4.780 |
| tqwt\_medianValue\_dec\_29 | -21.62 | -0.315 |
| tqwt\_skewnessValue\_dec\_24 | -22.68 | -1.69 |
| tqwt\_entropy\_shannon\_dec\_33 | -25.06 | -2.023 |
| tqwt\_entropy\_shannon\_dec\_32 | -25.67 | -2.12 |

Table I: Skewness of 10 most skewed attributes

*Attributes in "skewedCols" are classified based on positive, zero, or negative values into separate lists. For "skewedCols\_PositiveVals," Box-Cox Transformation[10] reduces skewness by transforming into a normal distribution. "skewedCols\_ZeroVals" and "SkewedCols\_NegativeVals" use Cube Root Transformation, making attributes more like normal distributions, by taking the cube root of each value of the attribute. Fig. 2 shows the distribution change after skewness reduction.*

1. ***Kurtosis Reduction****: Kurtosis measures peakedness/flatness relative to the normal distribution. High kurtosis can bias the model thus kurtosis reduction [11] is required. The kurtosis of the normal distribution is considered zero. Dataset attributes with kurtosis greater than 3 were checked but none required reduction.*
2. ***Outlier Detection****: Outliers [12] are data points significantly different from others and can bias machine learning models. Each attribute is checked for points below the 25th or above the 75th percentile. Outliers are replaced by the attribute mean. Fig. 3 shows the distribution plot after outlier reduction.*
3. ***Principal Component Analysis****: One approach to reduce the dimensionality of a dataset while retaining most of the variability is Principal Component Analysis (PCA). For the 148 principal components [13] were obtained after looping through them which gave us the optimal results.*

## Model Selection

In this proposed research work, Residual Neural Network (ResNet) is used as a Supervised Deep Learning Model.

Residual Neural Network (ResNet) is an extension of the Convolution Neural Network (CNN). ResNet and CNN are primarily deployed for image processing as an application of Computer Vision.

The CNN was able to handle a particular number of hidden layers. For updating the weights, the Backpropagation method is used. After going back, a certain number of layers, there is a shift down of loss function. Over a significant number of layers, the gradient “vanishes” leading to the vanishing gradient problem, leading to a saturation in the performance of the CNN model. To overcome this issue, the ResNet model is deployed.

ResNet's "skip connections" are a unique solution stacking identity mappings to speed up training. These connections ignore certain layers, utilizing the activations of previous layers. This compression helps in the initial training process.

During retraining, the network expands, allowing the residual parts to elaborate on the input image's feature space. ResNet typically skips two or three layers at a time with nonlinearity and batch normalization in between.

HighwayNets, a more advanced ResNet model, introduce "skip weights" that dynamically determine the number of layers to skip, enhancing the network's flexibility and performance.

Fig. 4. shows the working of the Residual Block, which is responsible for skip connections in ResNet.

## Experimental Result

Residual Neural Network outperformed the Traditional Machine Learning models, Artificial Neural Network and Multi-Layered Perceptron Classifier. ResNet was able to achieve an accuracy of 98%, PD positive precision of 0.98, PD negative precision of 0.94, PD positive recall of 0.98, PD negative recall of 0.94, PD positive F1-score of 0.98 and PD negative F1-score of 0.94.

TABLE II

|  |  |  |
| --- | --- | --- |
| Actual Values | Positive | Negative |
| Predicted values |
| Positive | 15 (TP) | 1 (FP) |
| Negative | 0 (FN) | 60 (TN) |

Table II: Confusion matrix of ResNet.

Table II describes the confusion matrix of ResNet showing 15 True Positive (TP), 1 False Positive (FP), 0 False Negative (FN) and 60 True Negative (TN) classifications. The negative sloping Error Vs Epoch Curve is illustrated in Fig. 5. The AUC-ROC curve with an AUC of 0.984 and AUC-PR curve with an AUC of 0.996 is depicted in Fig. 6 and 7 respectively.

# Comparative Performance Analysis

After preprocessing, each model was trained on a dataset split into a training set (85%) and a testing set. Traditional ML models and Neural Networks were trained on the training set. ResNet's superior performance was demonstrated through comparative analysis on the testing set.

* **Confusion Matrix:** Confusion Matrix (CM) evaluates binary classification algorithms with True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN). Table II shows CM for ML and Deep Learning models. ResNet had the highest TP and TN, and the lowest FP and FN values.

TABLE III.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name Of Model** | **True Positive** | **False Positive** | **False Negative** | **True Negative** |
| KNN | 14 | 3 | 1 | 58 |
| SVC | 11 | 6 | 1 | 58 |
| DTC | 12 | 5 | 5 | 54 |
| RFC | 10 | 7 | 1 | 58 |
| NBC | 11 | 6 | 3 | 56 |
| LR | 12 | 5 | 5 | 54 |
| XGBC | 15 | 2 | 5 | 54 |
| GBC | 12 | 5 | 4 | 55 |
| ANN | 14 | 2 | 1 | 59 |
| MLP | 14 | 3 | 1 | 54 |
| ResNet | 15 | 1 | 0 | 60 |

Table III: Confusion matrix of ML models

* **Accuracy -** Table IV displays model accuracies for the training split. Residual Neural Network achieved the highest accuracy of 98% in classifying Parkinson's disease presence based on attributes.
* Error - Error is the difference between true and predicted results produced by the models. Fig. 8 shows how error changes with parameter tweaks. ResNet exhibits the most improvement over iterations.
* **Recall -** The recall values of the different models have been accrued for the training split in the table above. Recall measures the proportion of actual positive cases that are correctly identified by the model as positive. From Table IV, it is evident that Residual Neural Network is the best-performing model in terms of recall.
* **Precision -** Precision is the measure of the fraction of correctly predicted positive instances out of all the instances the model has predicted as positive. Precision values for the training split for all the implemented models are in Table IV. Residual Neural Network shows the highest precision, correctly predicting positive instances out of all instances classified as positive.
* **F1-Score -** Table IV contains F1-Scores for the training split. F1-Score considers both Precision and recall. High F1-Score indicates accurate predictions and reduced False Negatives. Residual Neural Network has the highest F1-Score, making it the optimal model for both classes.

TABLE IV

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Name of Model | Acc | Precision | | Recall | | F1-Score | |
| PD  -ve | PD  +ve | PD  -ve | PD  +ve | PD  -ve | PD  +ve |
| KNN | 95% | 0.93 | 0.95 | 0.82 | 0.98 | 0.87 | 0.97 |
| SVM | 91% | 0.92 | 0.91 | 0.65 | 0.98 | 0.76 | 0.94 |
| DTC | 87% | 0.71 | 0.92 | 0.71 | 0.92 | 0.71 | 0.92 |
| RFC | 89% | 0.91 | 0.89 | 0.59 | 0.98 | 0.71 | 0.94 |
| NBC | 88% | 0.79 | 0.90 | 0.65 | 0.95 | 0.71 | 0.93 |
| LR | 87% | 0.71 | 0.92 | 0.71 | 0.92 | 0.71 | 0.92 |
| XGBC | 91% | 0.75 | 0.96 | 0.88 | 0.92 | 0.81 | 0.94 |
| GBC | 88% | 0.75 | 0.92 | 0.71 | 0.93 | 0.73 | 0.92 |
| ANN | 96% | 0.94 | 0.97 | 0.90 | 0.98 | 0.92 | 0.97 |
| MLP | 94% | 0.82 | 0.98 | 0.93 | 0.95 | 0.87 | 0.96 |
| ResNet | 98% | 0.94 | 0.98 | 0.94 | 0.98 | 0.94 | 0.98 |

Table IV: Evaluation Metrics for ML models

* **AUC-ROC Curve -** Area Under the Receiver Operating Characteristics Curve (AUC-ROC) evaluates classification models graphically. It represents the model's ability to distinguish positive and negative classes across various threshold values. A high AUC value (close to 1) indicates precise discrimination between Parkinson's Positive and Negative patients. Fig. 9 shows the AUC-ROC curves for all models. Residual Neural Network stands out with the highest AUC-ROC of 0.984, making it the optimal model compared to others.
* AUC-PR Curve - The AUC-PR curve is made by plotting the model’s precision against the recall of the model. It is a particularly important metric to consider in medical use as it can provide a more accurate evaluation of the performance model when the class distribution is imbalanced. The same is true for this dataset as well. The AUC-PR curves of all the implemented Machine Learning and Neural Network Models are illustrated in Fig.10.

From the above graphs, it is evident that the Residual Neural Network is the optimal model as they have the highest AUC-PR value of 0.996 in comparison to others.

# Conclusion and scope for future work

The research article focuses on the Deep Learning Based prediction and monitoring of Parkinson’s Disease using the Voice Data sourced from UCI Machine Learning Repository. To accomplish this objective, Residual Neural Network architecture was used as a Supervised Classification Model.

The factors responsible for choosing Residual Neural Network over other standard models is the concept of “skip connections”. The Residual Neural Network brings forth an innovative solution to the vanishing gradient problem in the form of skip connections. ResNet reuses the activation of previous layers, speeding up the initial training by compressing the network layers.

From the comparative analysis of the selected dataset, it can be concluded, Residual Neural Network (ResNet) has worked exceptionally better than all the traditional Machine Learning Models with an accuracy of 98%, a precision of 0.98, a recall of 0.98 and an F1-Score of 0.98. Standing next to the ResNet is ANN with an accuracy of 96%, a precision of 0.97, a recall of 0.98 and an F1-Score of 0.97. Thus, in conclusion, for the chosen dataset, the Residual Neural Network has performed exceptionally better in terms of the evaluation metrics. With the advent of more precise attributes in the future, the accuracy of the ML Models and Neural Networks can be boosted to a great extent using Feature Selection, Dimension Reduction and all the other effective and essential Data Pre-processing Techniques.

Further advancements in this field of study are certain to occur. A few of the possible enhancements can be utilising the MRI images of the human brain in predicting Parkinson’s disease using various advanced networks for optimal results. Another approach to the prediction of PD can be done using the datasets comprising Spiral writing test images [15] and Speech [16]. Also rather than depending on only one source of input dataset for prediction, a cumulated form of training data comprising not only numerical inputs but also images and sounds [17] can be prepared for even more accurate and precise results.

# Reference

1. A. Agarwal, S. Chandrayan and S. S. Sahu, "Prediction of Parkinson's disease using speech signal with Extreme Learning Machine," *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Chennai, India, 2016, pp. 3776-3779, doi: 10.1109/ICEEOT.2016.7755419.
2. A. Ouhmida, O. Terrada, A. Raihani, B. Cherradi and S. Hamida, "Voice-Based Deep Learning Medical Diagnosis System for Parkinson's Disease Prediction," 2021 International Congress of Advanced Technology and Engineering (ICOTEN), Taiz, Yemen, 2021, pp. 1-5, doi: 10.1109/ICOTEN52080.2021.9493456.
3. M. Ogawa and Y. Yang, "Residual-Network -Based Deep Learning for Parkinson's Disease Classification using Vocal Datasets," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), Nara, Japan, 2021, pp. 275-277, doi: 10.1109/LifeTech52111.2021.9391925.
4. M. Aghzal and A. Mourhir, "Early Diagnosis of Parkinson’s Disease based on Handwritten Patterns using Deep Learning," 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), Fez, Morocco, 2020, pp. 1-6, doi: 10.1109/ICDS50568.2020.9268738
5. A. Anand, M. A. Haque, J. S. R. Alex and N. Venkatesan, "Evaluation of Machine learning and Deep learning algorithms combined with dimensionality reduction techniques for classification of Parkinson’s Disease," 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Louisville, KY, USA, 2018, pp. 342-347, doi: 10.1109/ISSPIT.2018.8642776.
6. Sakar, C.O., Serbes, G., Gunduz, A., Tunc, H.C., Nizam, H., Sakar, B.E., Tutuncu, M., Aydin, T., Isenkul, M.E. and Apaydin, H., 2018. A comparative analysis of speech signal processing algorithms for Parkinsonâ€™s disease classification and the use of the tunable Q-factor wavelet transform. Applied Soft Computing, DOI: <https://doi.org/10.1016/j.asoc.2018.10.022>
7. Sen, Sohom; GHOSH, ANKIT (2022): Analysis and Prediction of Parkinson's Disease using Machine Learning Algorithms. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.20005703.v1>
8. G. A. Mirza, "Null Value Conflict: Formal Definition and Resolution," 2015 13th International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan, 2015, pp. 132-137, doi: 10.1109/FIT.2015.32.
9. P. L. Kevin Ding, S. Martin and B. Li, "Improving Batch Normalization with Skewness Reduction for Deep Neural Networks," 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021, pp. 7165-7172, doi: 10.1109/ICPR48806.2021.9412949.
10. X. Li, Y. Sun, X. Chen, Y. Li, T. Jiang and Z. Liang, "Saline-Sodic Soil EC Retrieval Based on Box-Cox Transformation and Machine Learning," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 1692-1700, 2022, doi: 10.1109/JSTARS.2022.3145874.
11. T. Thaiupathump and R. Chompu-inwai, "Impact of kurtosis on performance of mixture control chart patterns recognition using Independent Component Analysis and neural networks," 2015 4th International Conference on Advanced Logistics and Transport (ICALT), Valenciennes, France, 2015, pp. 94-99, doi: 10.1109/ICAdLT.2015.7136600.
12. H. Du, Q. Ye, Z. Sun, C. Liu and W. Xu, "FAST-ODT: A Lightweight Outlier Detection Scheme for Categorical Data Sets," in IEEE Transactions on Network Science and Engineering, vol. 8, no. 1, pp. 13-24, 1 Jan.-March 2021, doi: 10.1109/TNSE.2020.3022869..
13. Y. Li and J. Zhang, "Principal Component Analysis Based on Different Starting Points," 2022 IEEE/ACIS 22nd International Conference on Computer and Information Science (ICIS), Zhuhai, China, 2022, pp. 245-251, doi: 10.1109/ICIS54925.2022.9882417.
14. Pal, Sankar K., and Dinabandhu Bhandari. "Selection of optimal set of weights in a layered network using genetic algorithms." *Information Sciences—Informatics and Computer Science, Intelligent Systems, Applications: An International Journal* 80.3-4 (1994): 213-234.
15. A. Talitckii et al., "Comparative Study of Wearable Sensors, Video, and Handwriting to Detect Parkinson’s Disease," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-10, 2022, Art no. 2509910, doi: 10.1109/TIM.2022.3176898.
16. M. K. Reddy and P. Alku, "Exemplar-Based Sparse Representations for Detection of Parkinson's Disease From Speech," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 1386-1396, 2023, doi: 10.1109/TASLP.2023.3260709.
17. J. P. Devarajan, V. R. Sreedharan and G. Narayanamurthy, "Decision Making in Health Care Diagnosis: Evidence From Parkinson's Disease Via Hybrid Machine Learning," in IEEE Transactions on Engineering Management, doi: 10.1109/TEM.2021.3096862.