

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

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

Samaresh Maiti
Computer Science and Engineering
Heritage Institute Of Technology
Kolkata, India
samareshmaiti02@gmail.com

Sumanta Manna
Computer Science and Engineering
Heritage Institute Of Technology
Kolkata, India
mannasumanta9903@gmail.com

Rituparna Mondal
Department Of Computer Application
Techno India University
Kolkata, India
mondal.rituparna@gmail.com

Alok Kole
Electrical Engineering
RCC Institute of Information
Technology
Kolkata, India
alok.kole@rcciit.org.in

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

Abstract— This paper focuses on using Deep Learning for predicting and monitoring Parkinson's Disease (PD) through voice data. PD is a progressive neurological disorder that affects the Central Nervous System (CNS), leading to symptoms like tremors, stiffness, slow movements, balance and coordination difficulties, and speech disorders. As per recent studies by the World Health Organization (WHO), the mortality rate for PD has increased significantly over the years. Early detection and severity assessment of PD using Machine Learning is crucial. Speech recognition offers a new approach for diagnosis and monitoring. The research proposes using Deep Learning models with acoustic features like jitter, shimmer, intensity, and pitch for automatic detection and severity assessment. A dataset containing speech samples from PD patients and healthy individuals is used. A Residual Neural Network (ResNet) architecture is implemented and compared to other Machine Learning models like K-Nearest Neighbors, Support Vector Classifier, Decision Tree, Random Forest, Naïve Bayes, Logistic Regression, Extreme Gradient Boost, Gradient Boosting, as well as modern neural network techniques like Artificial Neural Network and Multi-Layer Perceptron. The results show that the proposed Residual Neural Network outperforms all other standard Machine Learning models in terms of accuracy, F1-score, precision, recall, AUC-ROC, and AUC-PR for PD prediction and monitoring using voice data.

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

I. INTRODUCTION

Parkinson's disease is a progressive movement disorder impacting the brain and central nervous system due to the degeneration of dopamine-releasing neurons, affecting motor control and causing tremors and impaired movements. The exact cause remains unknown, and a cure is yet to be discovered. Early detection is vital for managing symptoms and slowing down the disease's progression.

This article explores eight traditional machine learning (ML) models and three neural network models (K-Nearest Neighbour, Support Vector Machine, Decision Tree, Random Forest, Naive Bayes, Logistic Regression, XGBoost, Gradient Boost, Artificial Neural Network, Multi-Layer Perceptron, and Residual Neural Network). The goal is to classify whether a patient has Parkinson's disease based on their speech patterns. Thorough performance analysis using a

dataset has been conducted to evaluate the effectiveness of each model.

II. 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.

III. METHODOLOGY

A. Dataset

The ResNet implementation utilized a dataset sourced from UCI Machine Learning Repository [6]. It included 188 PD positive patients and 64 healthy individuals (130 men, 122 women) aged between 33 to 87. The dataset contained 755 columns with 756 data points, consisting of 564 Parkinson's positive and 192 Parkinson's negative cases. The dataset size on disk was 5.3 MB.

The classification of PD positive and negative patients involved 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. Despite the small size of the dataset, the abundance of attributes posed a risk of

overfitting. To address this, Principal Component Analysis (PCA) was employed to reduce the attributes without compromising the training capability. This enabled traditional ML algorithms and Neural Network models to achieve high levels of precision and accuracy. However, incorporating more data points could further enhance the models' performance.

B. 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_3_2	26.48	0.198
tqwt_TKEO_std_dec_32	26.06	0.019
tqwt_TKEO_mean_dec_3_3	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_2_9	-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

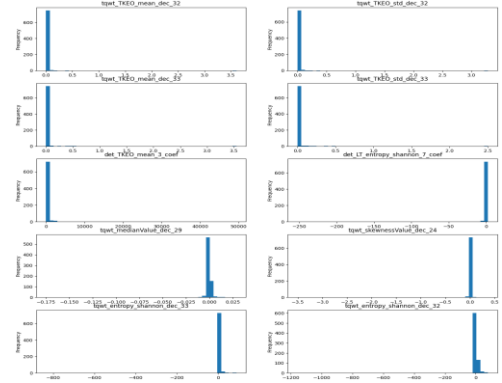


Fig. 1. Histogram of skewed attributes

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.

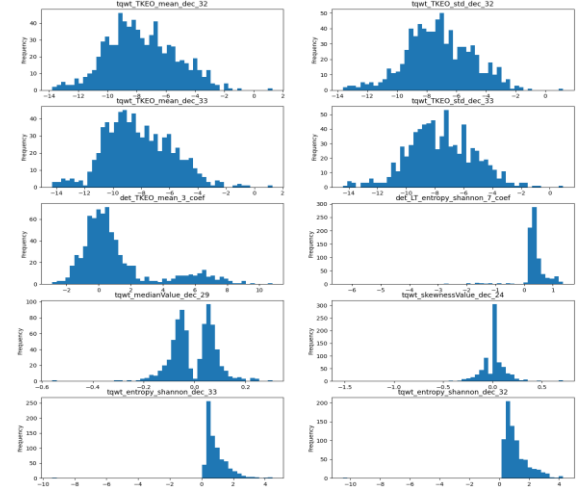


Fig. 2. Attributes with the highest skewness after skewness Reduction

- 3) **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.

- 4) **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.

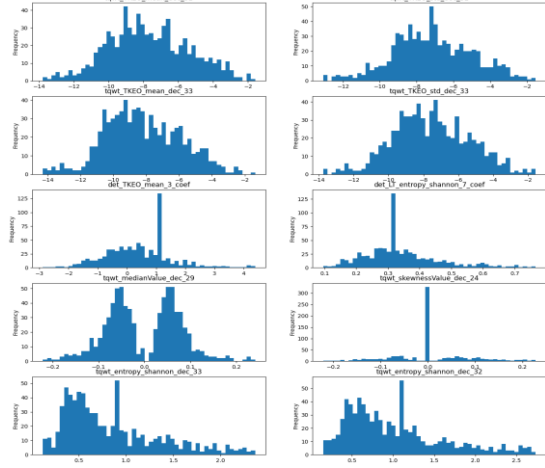


Fig. 3. Attributes with the highest skewness after outlier Reduction

- 5) **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.

C. Model Selection

In this research, Residual Neural Network (ResNet) is utilized as a Supervised Deep Learning Model, an extension of Convolution Neural Network (CNN) commonly used for Computer Vision tasks, particularly in image processing.

CNN faced limitations in handling a specific number of hidden layers, leading to the vanishing gradient problem when updating weights through Backpropagation. This issue resulted in performance saturation. To overcome this, ResNet was introduced with "skip connections," which stack identity mappings, accelerating training by reusing activations from previous layers, and compressing the network.

In retraining, ResNet expands and allows residual parts to enhance the input image's feature space. Skipping two or three layers at a time with nonlinearity and batch normalization in between. Advanced versions, like HighwayNets, introduce "skip weights" dynamically determining the number of layers to skip, thereby improving flexibility and performance.

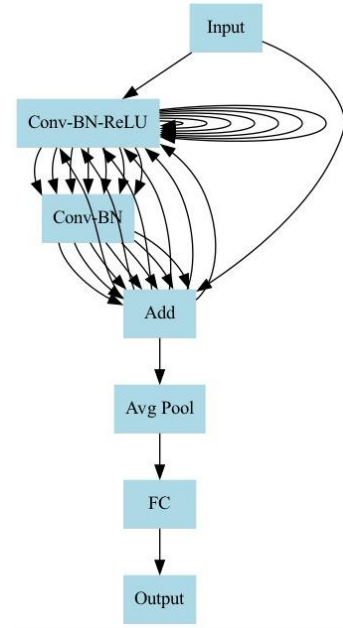


Fig. 4. Residual Block of ResNet

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

D. 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.

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

Table II: Confusion matrix of ResNet.

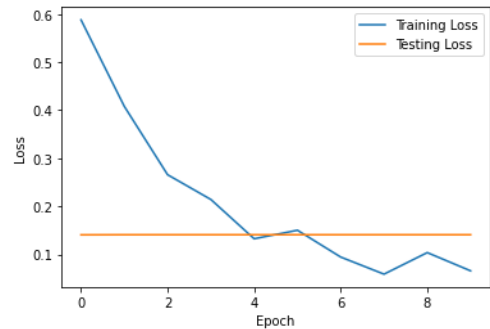


Fig. 5. Error graph for ResNets

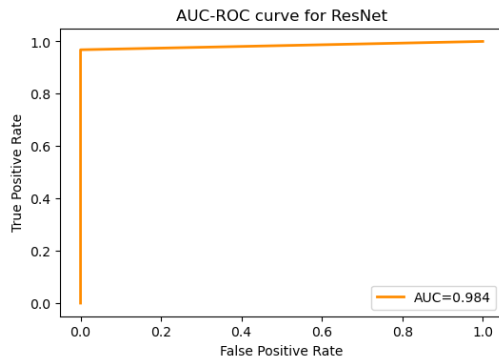


Fig. 6. AUC-ROC curve for ResNet

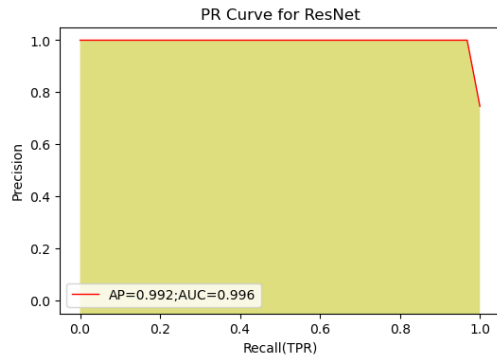


Fig. 7. AUC-PR curve for ResNet

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.

IV. 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.

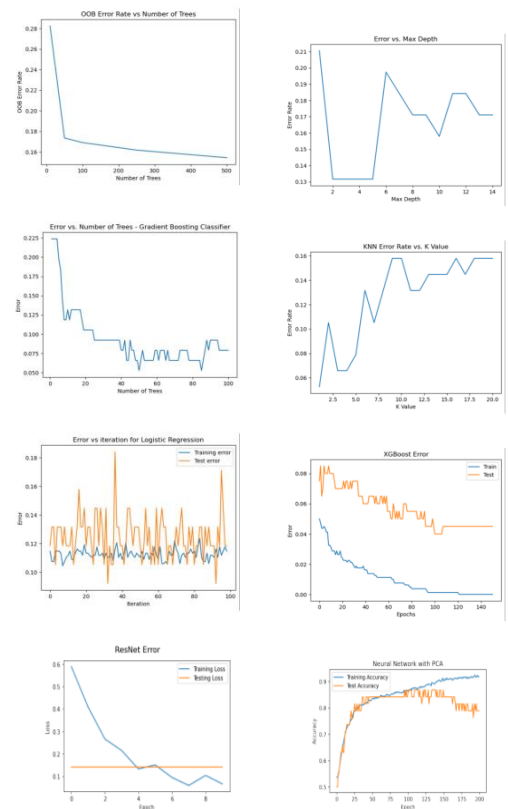
TABLE III.

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

Name Of Model	True Positive	False Positive	False Negative	True Negative
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.



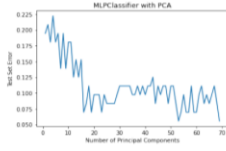


Fig. 8. Error Curve

- **Precision** - In Table IV, Precision is the fraction of correctly predicted positive instances among all instances classified as positive. The Residual Neural Network (ResNet) achieved the highest precision, accurately predicting positive instances in its classification.
- **F1-Score** - Table IV displays the F1-Scores for the training split, which consider both Precision and recall. A high F1-Score implies accurate predictions and fewer False Negatives. The Residual Neural Network achieved the highest F1-Score, establishing it as 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

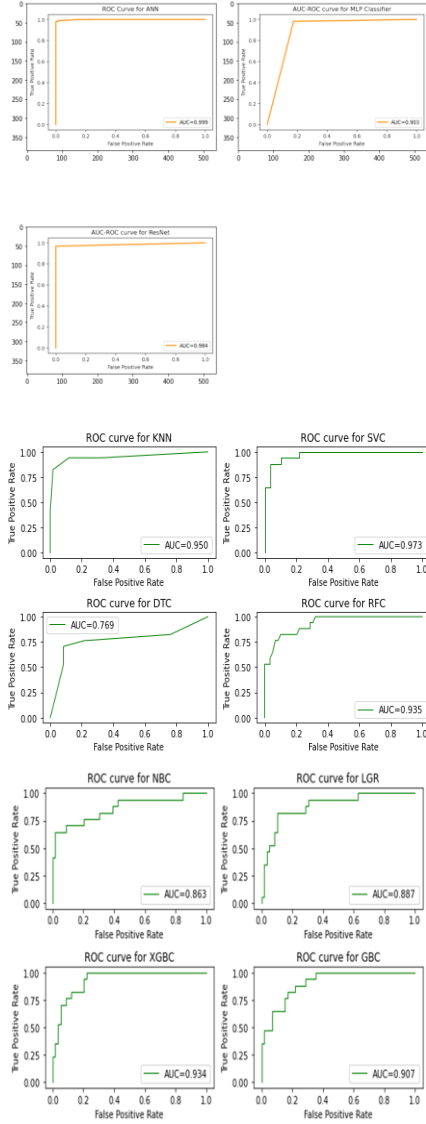
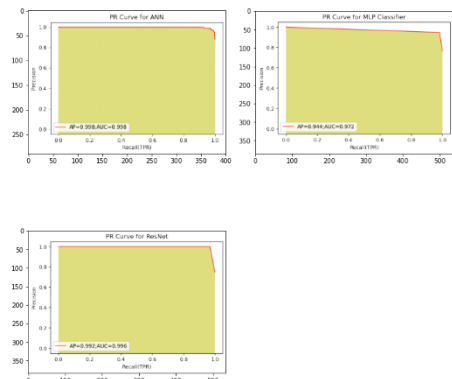


Fig. 9. Graphs of AUC-ROC for different ML and Neural Network

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 crucial for evaluating the model's performance, especially in medical applications with imbalanced class distributions. In this study, the AUC-PR curves for all implemented Machine Learning and Neural Network Models are shown 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.



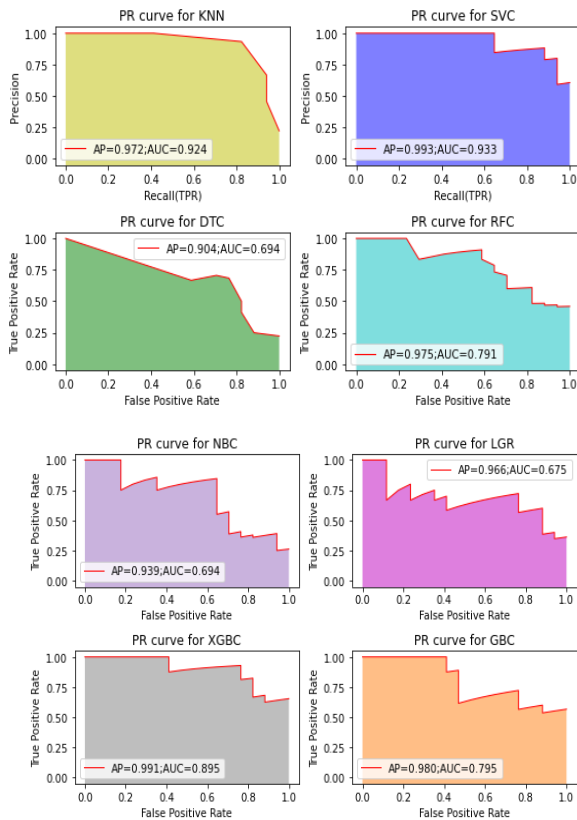


Fig. 10. Graph of AUC-PR for different ML models

V. CONCLUSION AND SCOPE FOR FUTURE WORK

The research article focused on using Deep Learning to predict and monitor Parkinson's Disease using voice data from the UCI Machine Learning Repository. The chosen model was the Residual Neural Network (ResNet) due to its innovative skip connections, which address the vanishing gradient problem and accelerate training. Comparing the dataset, ResNet outperformed all traditional Machine Learning Models with 98% accuracy, 0.98 precision, 0.98 recall, and 0.98 F1-Score. ANN followed closely with 96% accuracy, 0.97 precision, 0.98 recall, and 0.97 F1-Score. ResNet demonstrated superior performance in all evaluation metrics. Future enhancements in this field may involve incorporating MRI brain images for improved predictions, using advanced networks. Additionally, leveraging datasets with spiral writing test images [15] and speech [16] could be explored. Combining numerical, image, and sound data [17] could lead to even more accurate results.

VI. 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 Parkinson'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, Sohomi; 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.