

Early Detection and Severity Assessment of Parkinson's Disease Using Speech-Based Machine Learning

Ranjeet Kumar*, Akkaraju Venkata Ambareesh, Shaik Mohammad Rabbani, Chatwada Muni Vijayesh Singh, and Rekkala Venkateswara Reddy

Department of Artificial Intelligence and Machine Learning, Mohan Babu University, Tirupati, India

*Corresponding author: ranjeetjha8484@gmail.com

Abstract. Parkinson's disease (PD) is a neurodegenerative disease characterized by progressive motor disability and early speech disorders. Early diagnosis and regular follow-up allow timely clinical intervention and a better quality of life. This paper introduces a machine learning framework that identifies PD and estimates its severity from acoustic features derived from speech. The data analyzed are from the Parkinson's Telemonitoring dataset of the UCI Machine Learning Repository, which consists of 5,875 voice recordings. Acoustic characteristics such as jitter, shimmer, harmonicity, and nonlinear vocal parameters capture the speech changes associated with PD. Supervised linear regression, support vector regression (SVR), random forest (RF) regression, and gradient boosting (GB) models, together with an RF and GB ensemble, were trained to predict the motor and total Unified Parkinson's Disease Rating Scale (UPDRS) scores, and the predicted scores were then grouped into discrete severity levels. Model performance was analyzed with accuracy, precision, recall, and F1 score. The RF model achieved the best accuracy, 95.49%, which indicates strong predictive ability.

Keywords. Parkinson's disease; speech signals; machine learning; UPDRS; random forest regression; telemonitoring; severity prediction

I. INTRODUCTION

Parkinson's disease (PD) is a progressive neurological disorder and the second most common neurodegenerative disorder after Alzheimer's disease [1]. Its symptoms are insidious and are hard to notice in a routine clinical examination. Speech impairment is one of the early symptoms, and the traditional means of diagnosis are subjective, expensive, and unsuitable for continuous monitoring. More than ten million people worldwide are affected by the disease [2]. As telemonitoring data grow, there is increasing demand for non-invasive and objective diagnostic methods. Machine learning applied to acoustic features of speech can estimate disease severity in terms of the clinically relevant Unified Parkinson's Disease Rating Scale (UPDRS) scores, which allows early intervention and better disease management. The rest of the paper reviews related work and then presents the methodology, the experimental results, and directions for future research.

II. LITERATURE SURVEY

As biomedical engineering and machine learning have evolved, non-invasive speech analysis has considerably improved the diagnosis and monitoring of PD. From acoustic features of speech such as jitter, shimmer, and the harmonic-to-noise ratio, researchers can predict the early development of PD and identify motor weakness through vocal patterns. A study of voice-based PD detection described the effectiveness of both traditional machine learn-

ing and newer deep neural network methods in identifying vocal abnormalities [1]. Early studies relied on feature engineering with classification algorithms such as support vector machines and decision trees to distinguish patients with PD from healthy people [2]-[4]. However, these models had limited ability to predict disease progression or severity. More recent studies have used regression and ensemble learning to forecast UPDRS scores from speech characteristics [2], [5]. Random forest and SVR models capture the nonlinear relationship between acoustic biomarkers and motor scores [5], [6]. Advanced methods such as deep learning, explainable artificial intelligence, and multimodal learning have the potential to improve prediction accuracy and make models more transparent [7]-[15].

III. PROPOSED METHODOLOGY

The proposed methodology uses acoustic speech variables from the Parkinson's Telemonitoring dataset to estimate disease severity with machine learning. The inputs include features such as jitter, shimmer, and harmonicity. Regression models predict the motor and total UPDRS scores, which are then grouped into clinical severity levels for monitoring.

A. Experimental Dataset

The experiments used the Parkinson's Telemonitoring dataset from the UCI Machine Learning Repository. It contains speech-based features extracted during remote monitoring of patients with PD, including jitter, shimmer, harmonicity, and nonlinear speech parameters that indicate

vocal instability. The motor and total UPDRS scores support severity prediction, remote monitoring, and clinical decision-making.

B. Data Preprocessing

Data preprocessing was used to ensure effective model performance. The selected acoustic features and target variables were organized and normalized so that every feature contributed equally to training. The dataset was split into training and testing sets to measure performance on unseen data and to reduce overfitting. Because the dataset contains clean numerical features with no missing values, little data cleaning was needed before the machine learning models were trained to predict PD severity.

C. Feature Selection and Representation

PD causes subtle changes in voice frequency, stability, and volume. This study used the acoustic features of the Parkinson's Telemonitoring dataset, such as jitter, shimmer, harmonicity, and nonlinear speech parameters, which are signs of neuromotor impairment. Because the dataset provides extracted parameters, no raw audio processing was needed for severity prediction.

D. Machine Learning Prediction Models

Several supervised machine learning models were used for prediction and comparative analysis.

Random forest (RF): RF uses an ensemble of decision trees, and the final prediction is the average of the individual tree predictions,

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T h_t(x), \quad (1)$$

where T is the total number of trees, $h_t(x)$ is the prediction of the t th tree, and $\hat{y}$ is the aggregated prediction.

Linear regression (LR): LR models the relationship as a linear combination of the input features,

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n, \quad (2)$$

where $\hat{y}$ is the predicted value (for example, the UPDRS score), x_i are the input acoustic features, β_i are the coefficients (weights) learned from the data, and β_0 is the intercept term.

Support vector regression (SVR): SVR fits the best line (or hyperplane) within a margin of tolerance defined by the ε -insensitive loss,

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b, \quad (3)$$

where α_i and α_i^* are Lagrange multipliers, $K(x_i, x)$ is the kernel function (for example, the radial basis function $\exp(-\gamma \|x_i - x\|^2)$), b is the bias term, and $f(x)$ is the predicted output. SVR minimizes

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (4)$$

subject to

$$\begin{aligned} y_i - w^T x_i - b &\leq \varepsilon + \xi_i, \\ w^T x_i + b - y_i &\leq \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* &\geq 0. \end{aligned} \quad (5)$$

Gradient boosting (GB): GB is a tree-based ensemble in which decision trees are added sequentially, each fitted to the residual errors of the current ensemble.

Tree-based ensemble (RF + GB): The RF and GB models were also combined into a tree-based ensemble, whose results are reported in Table 1.

All models were trained on the same training set, and their performance was measured and compared to determine the most suitable algorithm for severity prediction.

E. Model Evaluation

The trained models were assessed with accuracy, precision, recall, F1 score, and the confusion matrix. This comparative evaluation highlights the strengths of each model and supports informed decision-making.

IV. RESULTS AND DISCUSSION

Machine learning analysis of speech helps to determine the relationship between a speaker's speech features and the severity of PD. Acoustic features were used to predict UPDRS scores with LR, SVR, RF, and GB, and with the RF + GB ensemble. After preprocessing and feature reduction, the data were split into training and testing sets in an 80:20 ratio.

Fig. 1 shows a correlation heatmap of the relationships between the speech-derived acoustic features and the UPDRS scores. The strong correlations among the jitter, shimmer, and harmonicity features highlight their importance in predicting disease severity.

Fig. 2 is a bar chart of the distribution of the detected PD severity levels. The distribution is asymmetric: most samples are moderate, followed by mild, and severe cases are the least common.

As Fig. 3 and Table 1 show, speech features are well suited to predicting PD severity. Ensemble predictions from speech features assess PD severity more stably than individual regression models do, and the RF model gives the best precision.

The confusion matrix of the RF model is shown in Fig. 4. RF and the ensemble model give consistent and reliable predictions across the full range of PD severity levels, which makes balanced and clinically useful classification possible in remote health monitoring.

TABLE 1. PERFORMANCE COMPARISON OF MODELS FOR PARKINSON'S DISEASE SEVERITY CLASSIFICATION

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Random forest	95.49	95.75	95.49	95.46
Tree-based ensemble (RF + GB)	89.79	91.14	89.79	89.57

V. CONCLUSION

This paper applied machine learning to predict the severity of PD from a dataset acquired through a telemonitoring application. The motor and total UPDRS scores were predicted with LR, SVR, RF, and GB models and with an RF + GB ensemble. Among these models, RF performed best, with a classification accuracy of 95.49%. The proposed approach is a reliable, non-invasive way of measuring PD severity that supports monitoring and timely response.

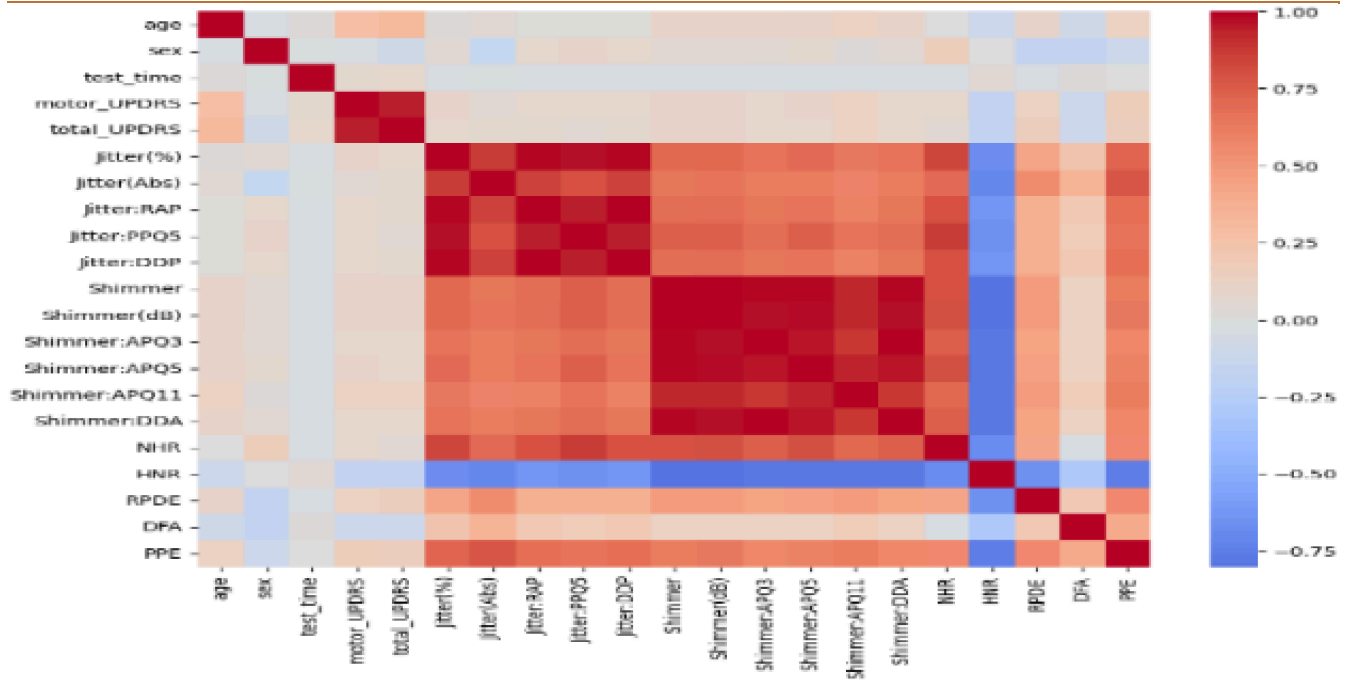


Fig. 1. Correlation matrix of speech features and UPDRS scores.

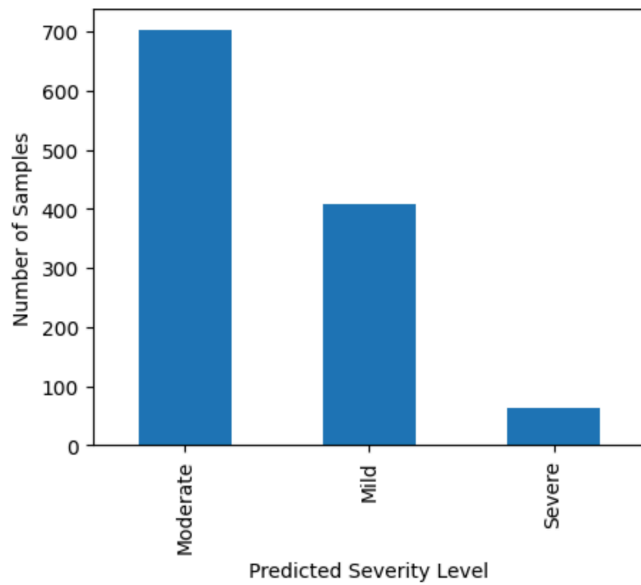


Fig. 2. Distribution of detected Parkinson's disease severity levels.

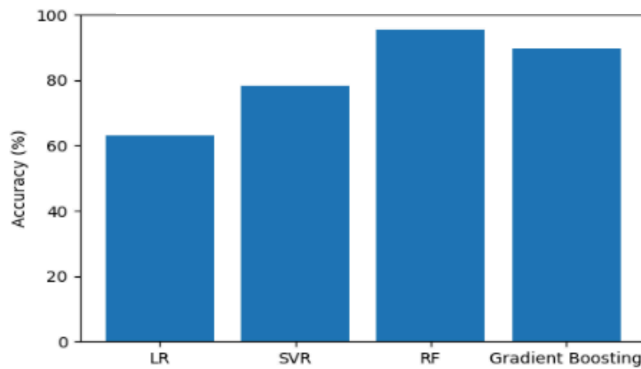


Fig. 3. Severity classification accuracy of the LR, SVR, RF, and GB models.

Actual	Mild	407	40	0
	Moderate	2	652	1
	Severe	0	10	63
		Predicted		
		Mild	Moderate	Severe

Fig. 4. Confusion matrix of the random forest model.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception and preparation of this work and approved the final version of the manuscript.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

DECLARATION OF GENERATIVE AI USE

The authors declare that no generative artificial intelligence tools were used in the preparation of this manuscript. The authors take full responsibility for the content of the publication.

DATA AVAILABILITY

This study used the publicly available Parkinson's Telemonitoring dataset (UCI Machine Learning Repository).

ETHICS STATEMENT

This study used publicly available, de-identified data and did not involve direct interaction with human participants or animals; ethical approval was therefore not required.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- [1] Guilherme C. Oliveira, Nemuel D. Pah, Quoc C. Ngo, Arissa Yoshida, Nicolás B. Gomes, João P. Papa, and Dinesh Kumar, "A pilot study for speech assessment to detect the severity of Parkinson's disease: An ensemble approach," *Comput. Biol. Med.*, vol. 185, Art. no. 109565, 2025, doi: 10.1016/j.compbio.2024.109565.
- [2] Huangqing Xu, Wei Xie, Mingzhen Pang, Ya Li, Luhua Jin, Fangliang Huang, and Xian Shao, "Non-invasive detection of Parkinson's disease based on speech analysis and interpretable machine learning," *Front. Aging Neurosci.*, vol. 17, Art. no. 1586273, 2025, doi: 10.3389/fnagi.2025.1586273.
- [3] Tabea Thies, Elisa Mallick, Johannes Tröger, Ebru Baykara, Doris Mücke, and Michael T. Barbe, "Automatic speech analysis combined with machine learning reliably predicts the motor state in people with Parkinson's disease," *npj Parkinson's Dis.*, vol. 11, Art. no. 105, 2025, doi: 10.1038/s41531-025-00959-4.
- [4] Daniel Escobar-Grisales, Tomás Arias-Vergara, Cristian David Ríos-Urrego, Elmar Nöth, Adolfo M. García, and Juan Rafael Orozco-Arroyave, "An automatic multimodal approach to analyze linguistic and acoustic cues on Parkinson's disease patients," in *Proc. Interspeech 2023*, 2023, pp. 1703-1707, doi: 10.21437/Interspeech.2023-2287.
- [5] Hadi Sedigh Malekroodi, Nuwan Madusanka, Byeong-il Lee, and Myunggi Yi, "Speech-based Parkinson's detection using pre-trained self-supervised automatic speech recognition (ASR) models and supervised contrastive learning," *Bioengineering*, vol. 12, no. 7, Art. no. 728, 2025, doi: 10.3390/bioengineering12070728.
- [6] Muna Alqam, Abdulnasir Hossen, Salima Al-Abri, Rami Al-Hmouz, and Muhammet Deveci, "Classification of the severity of speech disorders in patients with Parkinson's disease using machine learning techniques," *Int. J. Comput. Intell. Syst.*, vol. 19, no. 1, Art. no. 31, 2026, doi: 10.1007/s44196-025-01109-7.
- [7] Farhad Abedinzadeh Torghabeh, Seyyed Abed Hosseini, and Elham Ahmadi Moghadam, "Enhancing Parkinson's disease severity assessment through voice-based wavelet scattering, optimized model selection, and weighted majority voting," *Med. Novel Technol. Devices*, vol. 20, Art. no. 100266, 2023, doi: 10.1016/j.medntd.2023.100266.
- [8] Mohammad Amran Hossain and Francesco Amenta, "Machine learning-based classification of Parkinson's disease patients using speech biomarkers," *J. Parkinson's Dis.*, vol. 14, no. 1, pp. 95-109, 2024, doi: 10.3233/JPD-230002.
- [9] Benedetta Perrone, Federica Amato, and Gabriella Olmo, "Voice classification in Parkinson's disease: A deep learning approach using transformers and error rate metrics," *Biomed. Signal Process. Control*, vol. 113, Art. no. 108954, 2026, doi: 10.1016/j.bspc.2025.108954.
- [10] Matthew Shen, Pouria Mortezaagha, and Arya Rahgozar, "Explainable artificial intelligence to diagnose early Parkinson's disease via voice analysis," *Sci. Rep.*, vol. 15, no. 1, Art. no. 11687, 2025, doi: 10.1038/s41598-025-96575-6.
- [11] Lorenzo Simone, Mauro Giuseppe Camporeale, Vito Marco Rubino, Vincenzo Gervasi, and Giovanni Dimauro, "Interpretable early detection of Parkinson's disease through speech analysis," in *Artificial Intelligence in Medicine (AIME 2025)*, R. Bellazzi, J. M. Juarez Herrero, L. Sacchi, and B. Zupan, Eds., Lecture Notes in Artificial Intelligence, vol. 15735. Cham, Switzerland: Springer, 2025, pp. 373-378, doi: 10.1007/978-3-031-95841-0_69.
- [12] Tariq Adnan, Abdelrahman Abdelkader, Zipei Liu, Ekram Hossain, Sooyong Park, Md Saiful Islam, and Ehsan Hoque, "A novel fusion architecture for detecting Parkinson's disease using semi-supervised speech embeddings," *npj Parkinson's Dis.*, vol. 11, no. 1, Art. no. 176, 2025, doi: 10.1038/s41531-025-00956-7.
- [13] Athanasios Sarlas, Alexandros S. Kalafatis, Georgios Alexandridis, Michail-Alexandros Kourtis, and Panagiotis Trakadas, "Exploring federated learning for speech-based Parkinson's disease detection," in *Proc. 18th Int. Conf. Availability, Reliability and Security (ARES 2023)*, Benevento, Italy, 2023, doi: 10.1145/3600160.3605088.
- [14] Wee Shin Lim, Shu-I Chiu, Pei-Ling Peng, Jyh-Shing Roger Jang, Sol-Hee Lee, Chin-Hsien Lin, and Han-Joon Kim, "A cross-language speech model for detection of Parkinson's disease," *J. Neural Transm.*, vol. 132, no. 4, pp. 579-590, 2025, doi: 10.1007/s00702-024-02874-z.
- [15] Anna Favaro, Tianyu Cao, Thomas Thebaud, Jesus Villalba, Ankur Butala, Najim Dehak, and Laureano Moro-Velazquez, "Do phonatory features display robustness to characterize Parkinsonian speech across corpora?" in *Proc. Interspeech 2023*, 2023, pp. 2388-2392, doi: 10.21437/Interspeech.2023-1784.