*Proceedings of the 11th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'25)*
*Paris, France - August, 2025*
*Paper No. ICBES 170*
*DOI: 10.11159/icbes25.170*

# Comparison of Machine Learning Models for At-Home Diagnosis of Parkinson's Disease Progression Using Speech Biomarkers

**Minghui Yao[1,*], Vijay Ravi[2], Abeer Alwan[2]**
[1]The Bishop's School
7607 La Jolla Blvd, La Jolla, USA
grace.yao.26@bishops.com; vijaysumaravi@ucla.edu
[2]Department of Electrical and Computer Engineering, University of California, Los Angeles
420 Westwood Plaza, Los Angeles, USA
alwan@ee.ucla.edu
*Participant in UCLA California State Summer School for Mathematics & Science (COSMOS) program

## Extended Abstract

Parkinson's Disease (PD) is an incurable neurodegenerative disorder characterized by progressive motor and speech impairments, and early diagnosis enables therapeutic interventions that decelerate disease progression. Speech deficits such as word slurring, mumbling, stuttering, and reduced vocal intensity manifest early in PD, serving as disease biomarkers [1]. The Unified Parkinson's Disease Rating Scale (UPDRS) quantifies PD severity on a scale from 0-176 [2]. While there is currently no cure for PD, accurate treatment regimens tailored to the disease stage and severity are critical for delaying disease progression and the onset of more severe motor and cognitive symptoms. Machine-learning-driven analysis of speech biomarkers such as jitter and shimmer metrics can expedite initial diagnosis of progression using the UPDRS, allowing patients to seek clinical evaluation and tailored treatment sooner.

The aim of this study was to use speech biomarkers from processed in-home vocal recordings by patients with PD to train different machine learning models, including linear regression, decision tree regressors, and a multilayer perceptron (MLP) neural network, to predict UPDRS scores within a moderate clinically important difference (CID). The "Oxford Parkinson's Disease Telemonitoring Dataset" was used to obtain a range of biomedical voice measurements from 42 people with early-stage PD who recorded their voices at home on a telemonitoring device over a six-month trial [3]. Patient information such as age, gender, and true motor and total UPDRS scores were recorded, and the speech biomarkers of multiple numerical measures of jitter (variability in frequency in the acoustic signal) and shimmer (variation in amplitude) were extracted from the patient recordings. We evaluated two approaches for predicting UPDRS scores from biomarkers. First, we implemented linear regression and decision tree models with optimization techniques including L1/L2 regularization, cross-validation, and SHAP (SHapley Additive exPlanations)-based feature selection. Accuracy was assessed by clinical usefulness, where a validation mean absolute error (MAE) within the threshold of 8.5 to 10.3 points on the UPDRS total score was considered a moderate CID and therefore "correct," and parameters were optimized accordingly [4]. These methods achieved a validation MAE of 9.3, within our target clinical importance threshold of 8.5 to 10.3 UPDRS points. We then developed a custom-built neural network solution using z-score normalized speech biomarkers. Our MLP architecture consisted of an input layer, two hidden layers (64/32 neurons with ReLU activation), and a linear output neuron. Using Adam optimization with adaptive learning rate, this model achieved a superior validation MAE of 7.6 UPDRS points after 200 epochs. The findings of this study demonstrate the potential for machine learning models in Parkinson's Disease monitoring. With clinical oversight, such models could enable clinically useful, non-invasive, real-time, and continuous tracking of disease progression through telemonitoring platforms, facilitating stage-specific and personalized treatment adjustments.

## References


[1] M. Kumaresan and S. Khan, “Spectrum of Non-Motor Symptoms in Parkinson’s Disease,” *Cureus*, vol. 13, no. 2, Feb. 2021, doi: https://doi.org/10.7759/cureus.13275.

[2] “Table 18, Summary of Unified Parkinson’s Disease Rating Scale,” *Nih.gov*, Sep. 2018. https://www.ncbi.nlm.nih.gov/books/NBK539553/table/cl5.tab14/ (accessed May 09, 2025).

[3] A. Tsanas and M. Little. "Parkinson’s Telemonitoring," UCI Machine Learning Repository, 2009. [Online]. Available: https://doi.org/10.24432/C5ZS3N.

[4] L. M. Shulman, A. L. Gruber-Baldini, K. E. Anderson, P. S. Fishman, S. G. Reich, and W. J. Weiner, “The Clinically Important Difference on the Unified Parkinson’s Disease Rating Scale,” *Archives of Neurology*, vol. 67, no. 1, Jan. 2010, doi: https://doi.org/10.1001/archneurol.2009.295.