Predicting Total UPDRS for Parkinson's Disease based on Disease Progression

Introduction:

Parkinson's disease (PD) is a progressive neurodegenerative disorder that affects more than ten million people globally (Parkinson's Foundation, 2025). Pathologically, PD is primarily characterized by intraneural aggregates of misfolded alpha-synuclein protein, also known as Lewy bodies, and the gradual loss of dopaminergic neurons in the midbrain (Gomez-Benito, 2020). Patients with PD present motor and non-motor symptoms, such as resting tremor, muscle rigidity, and postural instability, as well as the loss of smell, sleep disturbances, and mood and cognitive changes (Zafar, 2023). PD patients are usually treated with Levodopa, dopamine agonists, deep brain stimulation (DBS), and other physical therapies (Mayo Clinic, 2023).

The Movement Disorder Society sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) remains the standard instrument for quantifying the severity of PD in clinical and research settings. The scale categorizes the disease progression based on (I) behavior and mood, (II) activities of daily living, (III) motor examination, and (IV) complications of therapy. Each section comprises multiple items rated on a 0–4 ordinal scale, with higher scores indicating greater impairment. The overall scale totals to 272 points and the baseline scores are typically around 32 ± 13 points (Goetz, 2008). However, current PD measurements are mostly done in clinical settings and require trained physicians; this limitation can result in delayed diagnosis because of infrequent hospital visits and available resources. As a result, many patients present in later stages, when care is more symptomatic relief rather than treatment (Raty, 2025). Thus, developing a model that estimates the total UPDRS from accessible features could potentially identify patients that are at risk, guide treatment adjustments, and support remote monitoring between visits.

Voice impairment is prevalently found in PD, reported in approximately 70–90% of patients. This suggests that speech is a strong benchmark for disease measurement (Rusz, 2011). Many patients show reduced voice volume, monotone, and imprecise consonant pronunciations before even obvious changes in limb motor function. Since vocal changes are critical in early PD detection, it is a reliable biomarker that supports timely risk stratification and diagnosis. Features, such as jitter, shimmer, harmonics-to-noise ratio, formant dynamics, and articulation rate, provide objective, repeatable measurements. In research settings, acoustic models show strong diagnostic accuracy and sensitivity to medication state, reflecting their importance in early detection and treatment optimization (Skodda, 2013). Therefore, the overall goal of the project is to develop an interpretable linear baseline, which is refuted through ordinary least squares (OLS) on selected features that are most significantly correlated with the total UPDRS, to generate reliable predictions of the total UPDRS scores.

Methods:

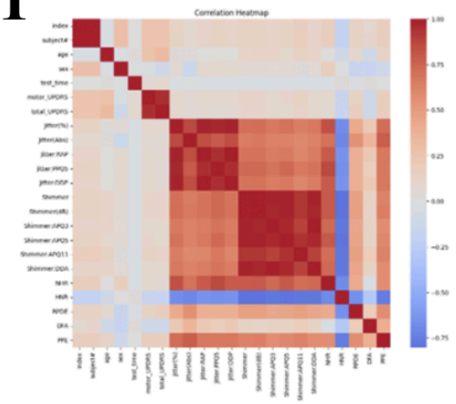
The dataset of the recorded speech signals is sourced from Max Little at the University of Oxford in collaboration with the National Center for Voice and Speech in Denver. It originates from the study "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," comprising 5874 datapoints collected from 31 patients (Little, 2008). An initial audit shows no missing values across the 23 columns, eliminating the need for additional data cleaning. Then, a ranked correlation table and a heatmap summarize the linear relationship between predictors and outcomes, with darker red symbolizing higher association with greater total_UPDRS (Figure 1). A multivariable linear regression is trained on all predictors and is used to estimates the total UPDRS. Its performance is then evaluated by both R^2 and root mean squared error (RMSE), which are 0.91 and 3.09, respectively. These metrics reflect that the model explains approximately 91% of the variance in the total UPDRS with a prediction error of 3.1 points, thus supporting a relatively accurate prediction model of the total UPDRS.

The 23 predictors' importance is ranked by the calculated, signed coefficients: larger magnitudes indicate greater contribution, positive coefficients mean higher feature values are associated with higher total_UPDRS, and vice versa. The "True vs. Predicted total_UPDRS" scatter plot demonstrates a strong, linear relationship between the true and predicted total UPDRS, indicating a well-adapted calibration across the observed range (Figure 2). Guided by the correlation map, coefficient screening, and domain knowledge, weak or redundant variables are removed and an OLS model was refit on six features, which are motor_UPDRS, harmonics-to-noise ratio (HNR), recurrence period density entropy (RPDE), pitch period entropy (PPE), eleven-point amplitude perturbation quotient eleven (Shimmer:APQ11), and detrended fluctuation analysis (DFA). Bivariate fits further show the expected directions, positive for RPDE/PPE/APQ11 and negative for HNR/DFA (Figure 3). Restricting to six predictors focuses the study on the most critical factors and allows greater model efficiency with only a negligible loss of fit, thus reducing the R^2 value from 0.91 to 0.90, while preserving practical deployability.

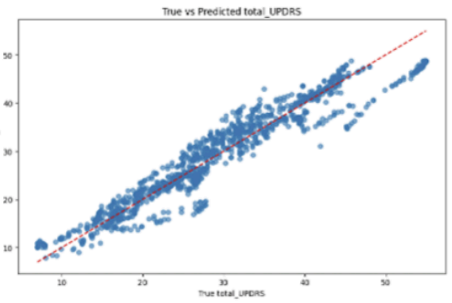
A two-layer neural network is built on the six selected features with an 80:20 train to test split. The network uses ReLU as the activation function and Mean Squared Error loss. Training and validation loss drop rapidly and stabilize, indicating convergence and minimal overfitting (Figure 5). The true vs. predicted scatter plot shows the dots clustering tightly around the 1:1 line, reflecting a strong correlation as confirmed by R^2 and RMSE, which are 0.94 and 2.49, respectively (Figure 6). Overall, the model seems well-calibrated.

Clearer and bigger figure images are all available in ipynb

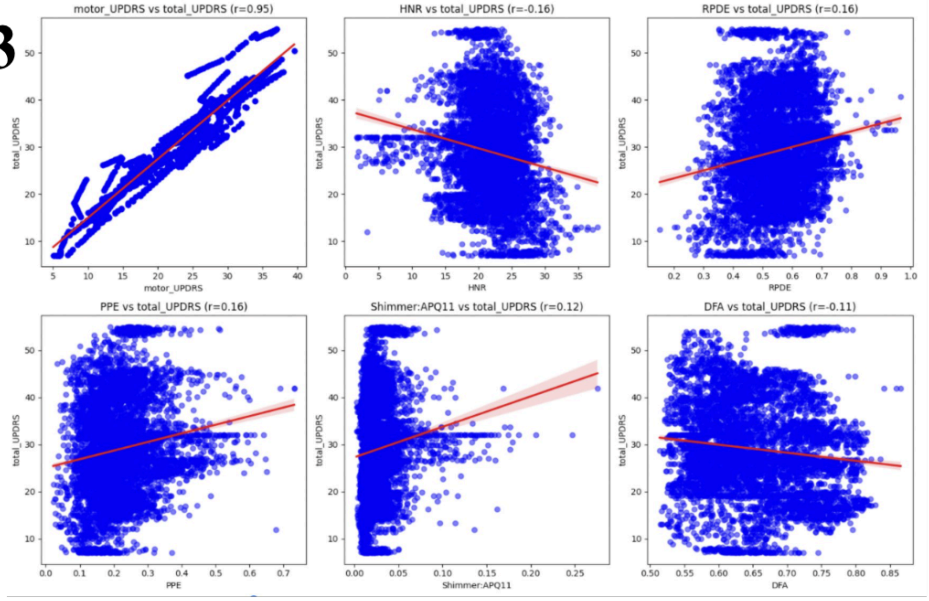
1



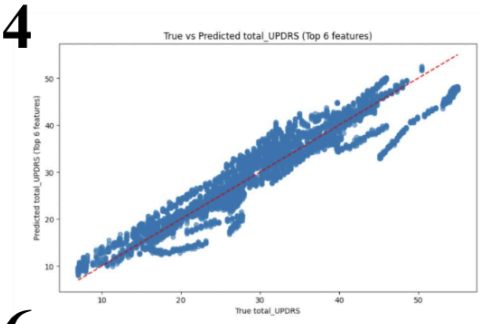
2



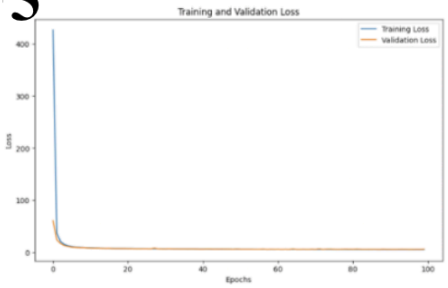
3



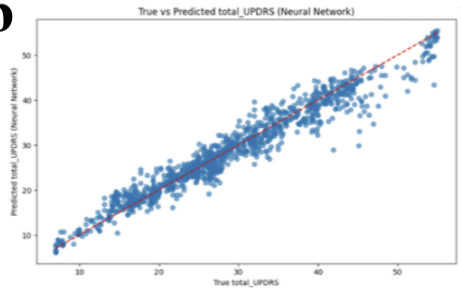
4



5



6



Results and Discussions:

The linear model initially uses 23 features to predict the total UPDRS, in which the performance closely aligns with the identity line ($R^2 = 0.91$, RMSE = 3.09). Correlations and coefficients identified motor_UPDRS as dominant, with additional signal from positive association for RPDE, PPE, and Shimmer: APQ11, and negative associations for HNR and DFA. To enhance efficiency and narrow down the number of important features, an OLS refit is based on the six features, resulting in a yielded with negligible fit loss ($R^2 = 0.90$), reducing redundancy from highly collinear acoustic families while preserving calibration. A two-layer neural network trained on the same six features further improved accuracy ($R^2 = 0.94$, RMSE = 2.49). Training and validation losses fell rapidly and plateaued together, indicating minimal overfitting. The true vs. predicted plot showed tight clustering around the 1:1 line, confirming a well-adapted prediction model for total UPDRS.

Nonetheless, limitations of the model still exist. The present work has weak predictive power as it relies primarily on vocal measurements. Despite their importance in early disease measurement, the correlation between voice and total UPDRS remains modest relative to motor impairment. Thus, the model should incorporate other indicators of earlier PD progression, such as tremor, constipation, and loss of smell (Parkinson's Foundation, 2025). In addition, all the acoustic results come from one clinical source, likely recorded with specific microphones, environments, instructions, and patient states. Since the differences in recording hardware, background noise, native language/accent, or how the task was prompted to each patient can all alter the acoustic data, future work should collect data from multiple clinical settings, reducing such biases.

Conclusions:

This study presents a preliminary model for predicting the total UPDRS score, which measures the severity of PD progression in patients. The model is tested on a small cohort of 31 individuals with voice impairments and shows consistent estimation of the UPDRS scores with high accuracy, demonstrated by strong R^2 and low RMSE. These results suggest the model could be useful for routine clinical assessments and tracking disease progression over time. Since model uses only six selected predictors, it can be easily incorporated into existing clinic workflows without adding significant administrative burden.

For future work, the model's predictive capability can be expanded by incorporating motor subscores and other early indicators of Parkinson's disease, including tremor, constipation, and sense dysfunction (Parkinson's Foundation, 2025). Testing the model across different clinical sites, with appropriate local recalibration, will be essential for ensuring its reliability in detecting early disease progression in diverse patient populations.

References:

1. Goetz, Christopher G., et al. “Movement Disorder Society-Sponsored Revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS): Scale Presentation and Clinimetric Testing Results.” *Movement Disorders*, vol. 23, no. 15, 15 Nov. 2008, pp. 2129–2170, movementdisorders.onlinelibrary.wiley.com/doi/10.1002/mds.22340, https://doi.org/10.1002/mds.22340.
2. Gómez-Benito, Mónica, et al. “Modeling Parkinson’s Disease with the Alpha-Synuclein Protein.” *Frontiers in Pharmacology*, vol. 11, no. 356, 23 Apr. 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7191035/, https://doi.org/10.3389/fphar.2020.00356.
3. Max A. Little, Patrick E. McSharry, Eric J. Hunter, Lorraine O. Ramig. “Suitability of dysphonia measurements for telemonitoring of Parkinson's disease.” *IEEE Transactions on Biomedical Engineering*, 2008, https://pmc.ncbi.nlm.nih.gov/articles/PMC3051371/
4. Mayo Clinic. “Parkinson’s Disease - Diagnosis and Treatment.” *Mayoclinic.org*, 2023, www.mayoclinic.org/diseases-conditions/parkinsons-disease/diagnosis-treatment/drc-20376062.
5. Parkinson's Foundation. “Statistics.” *Parkinson's Foundation*, 2025, www.parkinson.org/Understanding-Parkinsons/Statistics.
6. Parkinson's Foundation. “10 Early Signs of Parkinson’s Disease | Parkinson’s Foundation.” *Www.parkinson.org*, www.parkinson.org/library/fact-sheets/10-signs.
7. Räty, Valtteri, et al. “Stability and Accuracy of a Diagnosis of Parkinson Disease over 10 Years.” *Neurology*, vol. 104, no. 9, 13 May 2025, https://doi.org/10.1212/wnl.0000000000213499.
8. Rusz, J., et al. “Quantitative Acoustic Measurements for Characterization of Speech and Voice Disorders in Early Untreated Parkinson’s Disease.” *The Journal of the Acoustical Society of America*, vol. 129, no. 1, Jan. 2011, pp. 350–367, asa.scitation.org/doi/abs/10.1121/1.3514381, https://doi.org/10.1121/1.3514381.
9. Skodda, S., et al. “Progression of Voice and Speech Impairment in the Course of Parkinson’s Disease: A Longitudinal Study.” *Parkinson's Disease*, vol. 2013, 2013, pp. 1–8, https://doi.org/10.1155/2013/389195.
10. Zafar, Saman, and Sridhara S Yaddanapudi. “Parkinson Disease.” *Nih.gov*, StatPearls Publishing, 2023, www.ncbi.nlm.nih.gov/books/NBK470193/.