CJSJ

Evaluation of Gender Effects in Predicting Parkinson's Disease from Voice: A Random Forest Approach

Shreya Sreekantham

***Abstract*— Parkinson's Disease (PD) is the second most prevalent neurodegenerative disease in the world, affecting more than 10 million people. It has no cure, but early diagnosis of PD can help slow down the progression and improve the patient's life quality. However, the diagnosis of PD is often subjective and inaccurate because its presentation varies widely between individuals. This study focuses on early PD diagnosis and evaluates biomedical voice parameters variation by gender using a novel Random Forest Algorithm (RFA). The study utilized a multivariate PD dataset extracted from the UCI Machine Learning data repository that consisted of 5,875 voice recordings from 42 subjects. The novel RFA introduced in this study both improves the accuracy for PD detection and consistently performs well across gender. In addition, this study identifies gender-based differences in the expression profiles of voice parameters that can be useful in future clinical applications.**

I. INTRODUCTION

Parkinson's Disease (PD) is a condition characterized by a degeneration of dopaminergic neurons in the substantia nigra, impacting motor control, sensory systems, and cognition. [1] Some specific motor symptoms include dysarthria (weakness and coordination issues), akinesia (impairment of voluntary movement), tremors, and vocal changes. PD has no cure, so early diagnosis is critical to monitor, minimize, and control disease progression. However, as symptoms and progression vary widely, PD often goes undiagnosed for many years. In addition, factors like gender differences, aging, and immune status tend to complicate PD diagnosis. Since no diagnostic lab tests for PD exist, traditional diagnosis for PD relies on costly in-clinic tests and are not very accurate to date. For example, a recent study finds that early diagnosis of PD, with the patient having symptoms for five years or less, is only 58% accurate. [2] Therefore, there is an urgent medical need to develop more sensitive tools for PD diagnosis.

Given this subjectivity in diagnosing PD, machine learning approaches that target specific symptoms below the range of human observation have gained popularity. [3] Since vocal impairments are one of the earliest symptoms of the disease, this study focuses on detecting PD from voice. Additionally, despite different presentations in symptoms, there are relatively few studies investigating how gender differences affect the accuracy of PD diagnosis. One significant gender-based difference in Parkinson's disease is that although women demonstrate higher mortality and faster progression rates, PD is twice as prevalent in men than women. [4] This study seeks to understand how gender differences in symptoms affect the voice-based diagnosis of Parkinson's and proposes a novel Random-Forest Algorithm (RFA) that consistently performs well in detecting PD across gender. Below are the accuracies of PD diagnosis achieved by the algorithm:

- 99.2% for females without gender-impacted parameters
- 96.6% for males without gender-impacted parameters
- 99.3% for females with gender-impacted parameters
- 97.15% for males with gender-impacted parameters

In comparison, previous machine learning approaches for voice-based PD diagnosis when accounting for differences across these genders achieved a highest accuracy of 82.14%. This novel RFA improves PD diagnostic accuracy and performs well across gender. This study is an important step towards creating robust diagnostic models that account for demographic variation.

II. METHODS

Dataset Description

The dataset utilized in this study was generated by Athanasios Tsanas and Max Little of the University of Oxford12. This dataset has 16 biomedical early-stage PD voice parameters from 42 subjects. [5] The data has a total of 5,875 voice recordings during a six-month trial of a telemonitoring device. About 200 recordings were collected from each patient. The data was categorized as follows for each subject: Sex (Male: 0 and Female:1), Jitter (Five measures of variation in fundamental frequency), Shimmer (Six measures of variation in amplitude), NHR (Two measures of noise to tonal components ratio), RPDE (nonlinear dynamical complexity measure), DFA (Signal fractal scaling exponent), and PPE (a nonlinear measure of fundamental frequency variation).

Random Forest Algorithm

The goal of the machine learning algorithm is to predict Unified PD Rating Scale (UPDRS) scores, a scale of PD progression, from the parameters extracted from the voice recordings. Conducting initial exploratory analysis revealed no clear linear, logistic, or planar relationships between variables. To bypass these limitations, a random-forest algorithm (RFA) was utilized. With multiple decision trees rather than just one, Random Forests enables the exploration of the dependencies between variables in the multivariate dataset.

The data was sorted by gender for further analysis. This distribution was then utilized to identify gender-specific patterns of one or a combination of parameters using simpler statistical analysis. Following this Random Forest algorithm (Fig 1) was utilized to further evaluate the prediction parameters.

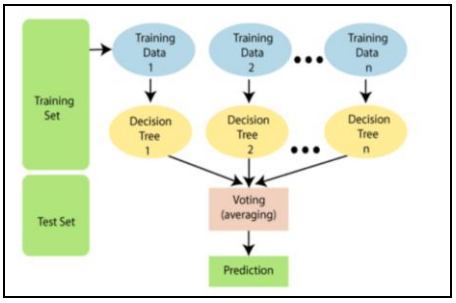

Fig 1. Depiction of Random Forest Cluster

The process for generating the RFA primarily involved the following steps: selecting random data points from the training set, building the decision trees with the selected data points, choosing the number of decision trees, repeating the steps, and finding predictions of each decision tree. Differences based on gender were identified in the data set, and the RFA was applied to the data to compare accuracies in predicting UPDRS scores.

The ensemble form of RFAs is expressed as Forest F = f1, f2, ..., fn. Each F yields a distribution that was averaged prior to applying the algorithm. The predictors were then combined by using the mean of the continuous target variables. The prediction performance of the proposed model was finally confirmed with a 5-fold cross-validation.

III. RESULTS

The dataset utilized in this research article has 5875 recordings, 1867 from females and 4008 from males. Exploratory analysis was conducted using the data from the training instances for all the inputs described in Table 1. Gender based data distributions were compared for both the target variables (motor UPDRS and Total UPDRS). The results showed that both target variables reflected a normal distribution profile (Fig 2).

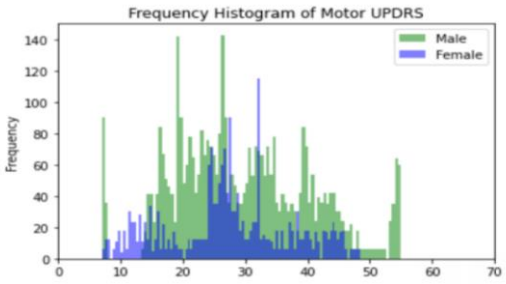

Fig 2. Gender-based frequency histogram profiles for motor UPDRS. Similar patterns were observed for Total UPDRS (not shown)

The biomedical voice parameter measurements were next evaluated individually for distribution trends between male and female subjects. For simplistic analysis, the data was normalized to the male measurements. The analysis established two of the sixteen parameters (Jitter Abs and NHR) showed differences across genders (Fig 3).

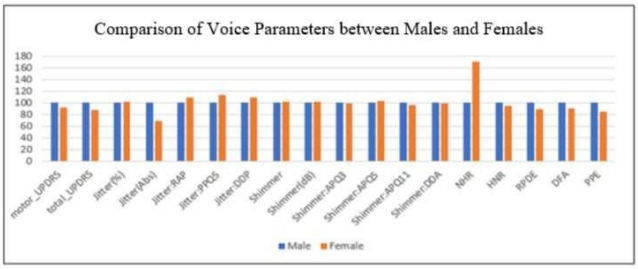

Fig 3. Comparison of individual biomedical parameters establishes possible differences in two of the sixteen parameters. (Data normalized to males)

This data indicates there are gender-based differences in the individual parameters, however it is unclear if these differences affect the prediction of UPDRS scores. Therefore, in order to obtain unbiased results, the RFA was run on the test cohort, to calculate accuracy with and without the gender dependent features (Jitter (Abs) and NHR). These prediction accuracies were then validated with a 5-fold cross-validation. Table 1 below summarizes the accuracies of the RFA.

Table 1: Accuracy Predictions with and without the identified gender-impacted parameters (Jitter Abs and NHS).

Gender	Accuracy Output	W/gender bias parameters	W/o gender bias parameters
Males	Motor UPDRS	97.3	96.7
	Total UPDRS	97.0	96.5
Females	Motor UPDRS	99.4	99.3
	Total UPDRS	99.2	99.1

The RFA averages a 98.0% with or without the gender-impacted parameters. Despite the gender-based differences in the individual parameters, this RFA maintains high accuracy across genders in the dataset.

IV. DISCUSSION

This study introduced a novel RFA which enhanced the accuracy of PD diagnosis compared to existing methods. It also helps conclude that variation in symptoms by gender does not affect the accuracy of voice-based diagnosis of PD. Furthermore, the gender-specific data were evaluated on one generalized model, which reduced the complexity of both the analysis and the prediction. For future work, applying this model to additional datasets and investigating the effects of other demographic differences will be a useful step towards implementing widely applicable voice-based screenings for Parkinson's.

V. ACKNOWLEDGEMENT

I would like to thank my mentor, Dr. Avani Wildani at Emory University, for guiding me through this research.

VI. REFERENCES

[1] Lill, Christina M., Johannes T. Roehr, Matthew B. McQueen, Fotini K. Kavvoura, Sachin Bagade, Brit-Maren M. Schjeide, Leif M. Schjeide, et al. “Comprehensive Research Synopsis and Systematic Meta-Analyses in Parkinson’s Disease Genetics: The PDGene Database.”
[2] Beach, Thomas G., and Charles Adler. “Importance of Low Diagnostic Accuracy for Early Parkinson’s Disease: Low Diagnostic Accuracy For Early PD.”
[3] Solla, Paolo, Antonino Cannas, Federica Carla Ibba, Federico Loi, Marta Corona, Gianni Orofino, Maria Giovanna Marrosu, and Francesco Marrosu. “Gender Differences in Motor and Non-Motor Symptoms among Sardinian Patients with Parkinson’s Disease.”
[4]Baldereschi, M., A. Di Carlo, W. A. Rocca, P. Vanni, S. Maggi, E. Perissinotto, F. Grigoletto, L. Amaducci, and D. Inzitari. “Parkinson’s Disease and Parkinsonism in a Longitudinal Study: Two-Fold Higher Incidence in Men.”
[5]“UCI Machine Learning Repository: Parkinsons Telemonitoring Data Set.” Accessed December 11, 2021.