BELLEVUE UNIVERISTY

Predictors for Parkinson’s Disease

A Study of Voice Recordings for Predictive Markers

Christine Hathaway

4/25/2020

Abstract

Parkinson’s disease (PD) is a degenerative neurological disorder that affects millions of people around the globe. Unfortunately, there is no cure, and no known tests to confirm it. The affliction costs thousands of dollars each year to help ease its symptoms and attempt to create a more comfortable life for those living with it. Finding a way to predict or test for PD is the first step in fighting back against this disease.

This study looks at the data compiled from measuring various aspects of the human voice. Analysis of speech and voice patterns has indicated it is possible to detect PD at earlier stages. Correlations of the variables in the study show that measuring certain aspects of the voice can indicate PD. These results could lead to earlier testing for PD, resulting in earlier treatment options that could slow the progress of the disease.

Intro/Background of the Problem

Parkinson’s disease (PD) is a progressive nervous system disorder that affects movement (Mayo Clinic, n.d.). Symptoms often start gradually, sometimes with a slight tremor in the hand. The disorder can also cause stiffness or slowing of movement. Symptoms can also include rigid muscles, impaired posture and balance, loss of automatic movements, speech changes, and writing changes (Mayo Clinic, n.d.). Typically, there are five stages to PD. The stages correspond both to the severity of movement symptoms and how much the disease affects a person’s daily activity. Doctors also use scales that help them understand the progression of the disease, which focus on motor symptoms (Parkinson's Foundation, n.d.).

PD is the second most common degenerative neurological disorder after Alzheimer’s disease (Downward & Johns Pool, 2019). It is estimated that 1% of the population over the age of 60 is affected by PD. It is difficult to determine exactly the number of people with PD, as many people do not get diagnosed in the early stages, but overall as many as 1 million Americans and 10 million people worldwide are living with PD. Approximately 60,000 Americans are diagnosed with PD every year. As life expectancy increases worldwide, it is expected that the burden of chronic diseases like PD could grow to between 8.7 million and 9.3 million by 2030 (Downward & Johns Pool, 2019). The combined direct and indirect cost of PD, including treatment, social security payments and lost income, is estimated to be nearly $52 billion per year in the United States alone (Parkinson's Foundation, n.d.).

Making an accurate diagnosis of PD can be complicated. Factors that doctors must carefully consider include not just symptoms but family history as well. The standard diagnosis is clinical, meaning there is no test that can give a conclusive result. Instead, doctors look for certain physical symptoms to be present to qualify a person’s condition as PD (Johns Hopkins Medicine, n.d.). A study conducted by three research centers are looking at techniques that analyze speech and vocal patterns that might be effective tools to diagnose PD disease, and possibly at earlier stages than is now possible (Holtzman, 2013). Their results build on past work that demonstrates that speech carries information relevant to an accurate and differential diagnosis of PD, and also shows that speech features of interest can be automated and assessed, with diagnostic reliability (Holtzman, 2013). Early detection, along with the start of treatment, would have a relevant effect on both the quality of life of patients and the healthcare system. This would allow the development of new therapies, and a better understanding of the disease and its evolution, according to Juan Ignacio Godino, a researcher in the study (Holtzman, 2013). This project seeks to address the question, are some speech indicators better than others at predicting PD?

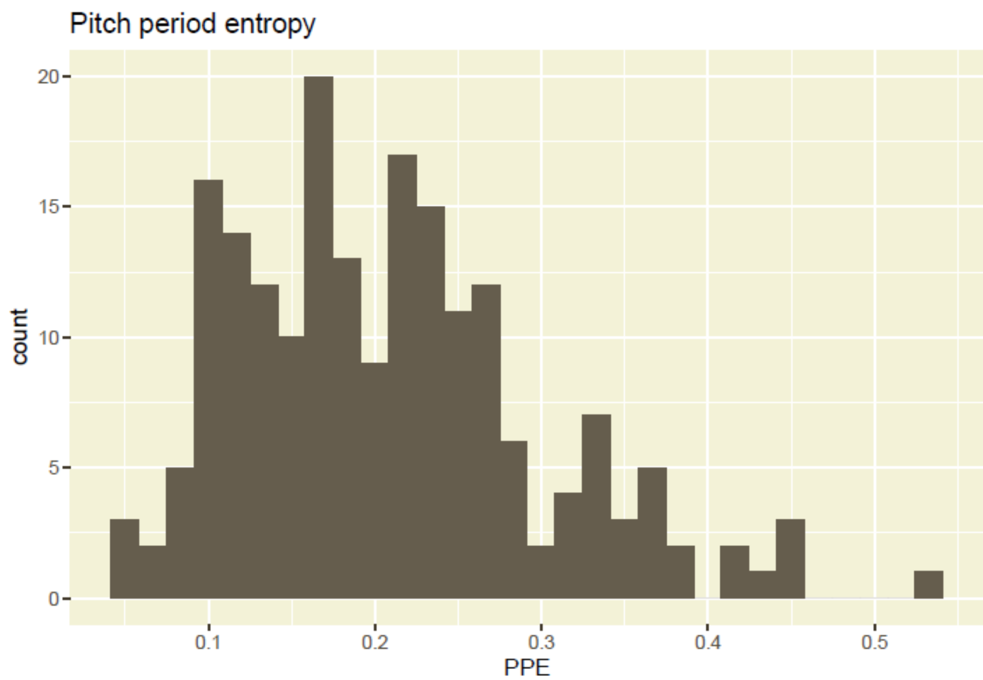
Data Understanding

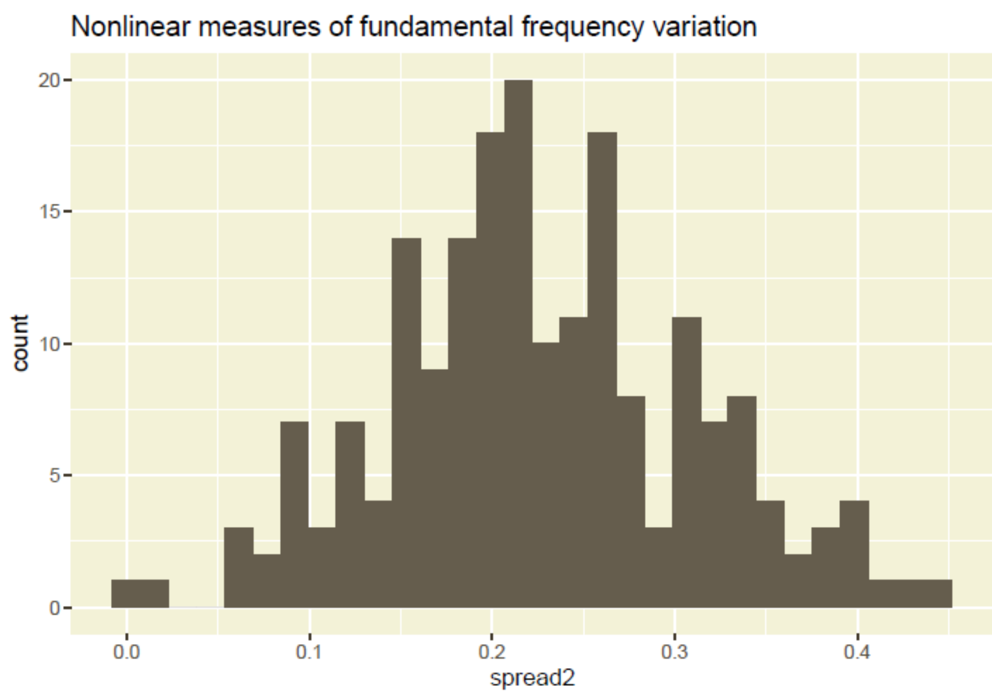
This dataset is composed of a range of biomedical voice measurements from 31 people, 23 with PD. Each column in the table is a particular voice measure, and each row corresponds to one of 195 voice recording from these individuals. The main aim of the data is to discriminate healthy people from those with PD, according to the "status" column which is set to 0 for healthy and 1 for PD. Each row of the CSV file contains an instance corresponding to one voice recording, and there are approximately six recordings per patient (Little, McSharry, Roberts, Costella, & Moroz, 2007).

The variables are different measurements of speech. These include three measures of vocal fundamental frequency, five measures of variations in fundamental frequency, six measures of variation in amplitude, two measures of ratio of noise to tonal components in the voice, two nonlinear dynamical complexity measures, signal fractal scaling exponent, and three nonlinear measures of fundamental frequency variation (Parkinsons Data Set, 2008).

Methods

The data was examined using Python and R software programs. The data set was clean with no missing values. Histograms performed with R resulted in no outliers being discarded or in need of transformation. Two graphs of note include the Pitch Period Entropy (PPE) and the Spread2 variable, which was one of the nonlinear measures of fundamental frequency variation. PPE is a new measure of dysphonia, or voice disorder, that the study used to measure sensitive changes in speech. See Appendix A for further details regarding the PPE measurement (Little, McSharry, Roberts, Costella, & Moroz, 2007).



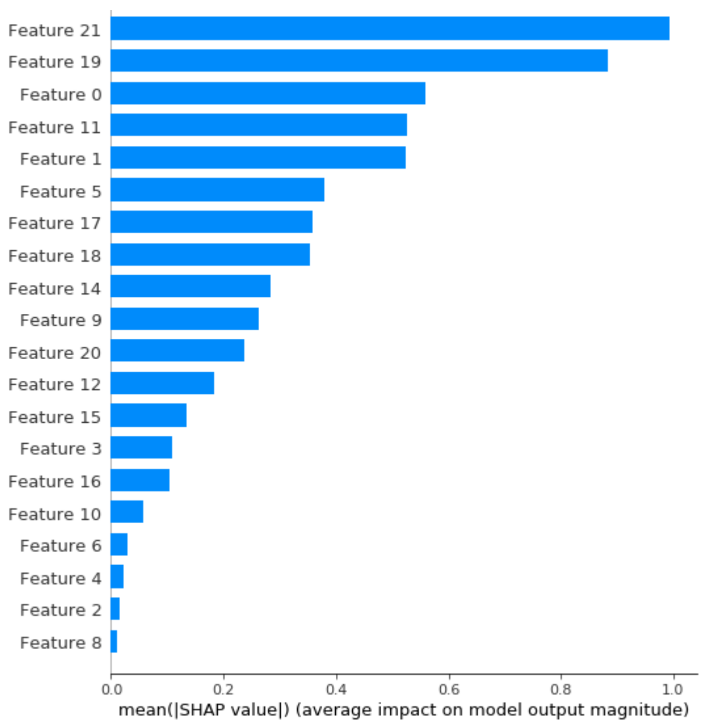


Modeling and Results

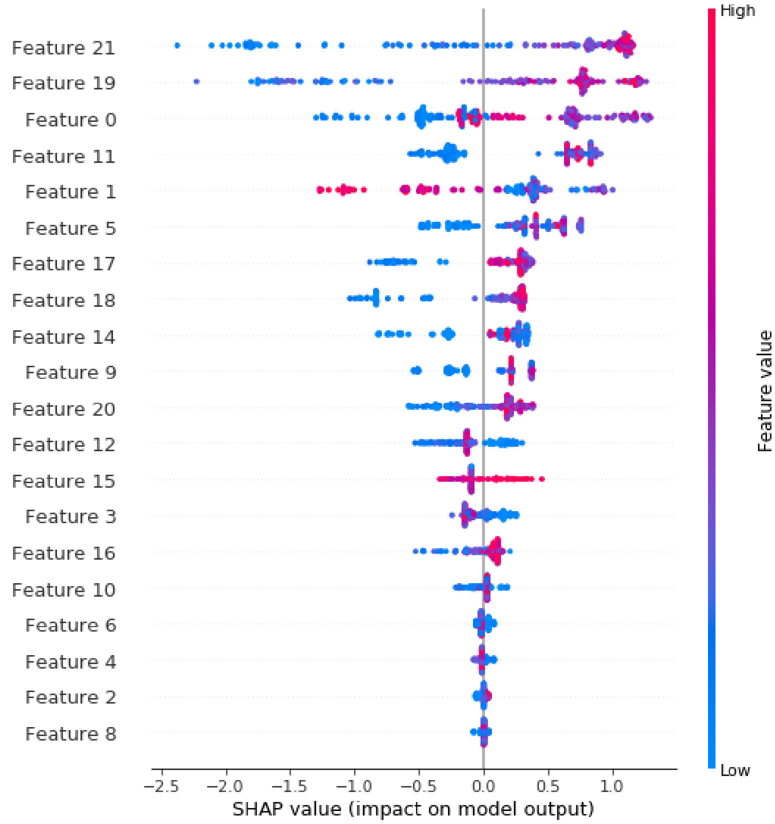
Regression models were completed using R software, but the results were not impressive. With so many variables, there were not enough measurements to test all variables at the same time in a regression model. Individual models were created to test selected variables together, with no significant results. The decision was made to look at alternative modeling algorithms.

Gradient boosting is a technique for regression problems that produces a prediction model in the form of decision trees. These algorithms work well for small-to-medium structured or tabular data, so it was decided to try it with this dataset.

Python has an algorithm called xgboost that uses gradient boosting. This algorithm was applied to the dataset with significant results. There are various ways to run the model, which gives slightly different results. Python also supports the shap package that is beneficial to interpret results. Using shap, the results indicated that feature 21, which is variable PPE, is the most significant in predicting PD.



The shap value plot further shows the positive and negative relationships of the predictors with the target variable. The variables are ranked in descending order, with the horizontal locations showing whether the effect of the value is associated with a higher or lower prediction. Color also shows whether that variable is high (in red) or low (in blue) for that observation. Therefore, Feature 21, or PPE, has a high positive correlation with the status of PD.



Discussion/Conclusion

The results of the xgboost algorithm, through the shap package interpretations clearly indicate that the measure of PPE is strongly correlated to having PD. Identifying this measurement is a key step in finding a test for PD. Currently there are no definitive tests for PD, it is usually diagnosed through clinical exams. Therefore, measuring PPE could be a start in developing tests that use voice measurements to identify those with PD. The earlier the diagnosis, the better the chances of stemming the physical ramifications of PD. With an aging population, and predictions that more and more will suffer from the degenerative disease, this is a crucial discovery in the fight against and finding a cure for PD.

Acknowledgments

I would like to acknowledge Max Little and his team for creating this dataset at the University of Oxford, along with the National Centre for Voice and Speech in Denver Colorado, who recorded the speech signals. Their work in the field of PD will be invaluable in the fight against PD.

References

Downward, E., & Johns Pool, J. (2019, September). *How Common Is Parkinson’s Disease?* Retrieved from Parkinsons Disease.net: https://parkinsonsdisease.net/basics/statistics/

Holtzman, J. (2013, March 24). *The Parkinson’s Voice Initiative: Early Diagnosis for Parkinson’s Disease through Speech Recognition*. Retrieved from Stanford.edu: https://web.stanford.edu/group/sjph/cgi-bin/sjphsite/the-parkinsons-voice-initiative-early-diagnosis-for-parkinsons-disease-through-speech-recognition/

Johns Hopkins Medicine. (n.d.). *How Parkinson's Disease Is Diagnosed*. Retrieved from Johns Hopkins Medicine: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/how-parkinson-disease-is-diagnosed

Little, M. A., McSharry, P. E., Roberts, S. J., Costella, D. A., & Moroz, I. M. (2007, June 26). 'Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection. *BioMedical Engineering OnLine, 6*(23).

Mayo Clinic. (n.d.). *Parkinson's disease*. Retrieved from Mayo Clinic: https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/symptoms-causes/syc-20376055

Parkinsons Data Set. (2008, June 26). *Oxford Parkinson's Disease Detection Dataset*. Retrieved from UCI Machine Learning Repository: http://archive.ics.uci.edu/ml/datasets/parkinsons

Parkinson's Foundation. (n.d.). *Parkinson's Foundation*. Retrieved from Parkinson.org: https://www.parkinson.org/Understanding-Parkinsons/Statistics

Appendix A

A New Measure of PD Dysphonia: Pitch Period Entropy (PPE)

“All healthy voices exhibit natural pitch (F0) variation characterized by smooth vibrato and microtremor [41], and this is detected in traditional jitter measures, for example. However, one common dysphonic PD symptom is impaired control of stationary voice pitch (F0) during sustained phonation [21]. Thus, with traditional measures it is difficult to separate natural, healthy pitch variations from dysphonic variations due to PD.

“Similarly, the extent of this natural variation is related to the average voice pitch of the subject; speakers with naturally high-pitched voices will have much larger vibrato and microtremor than those with lower-pitched voices, when these variations are measured on an absolute frequency (Hertz) scale. Therefore, measurements of abnormal speech pitch variation need to take into account these two important effects: healthy, smooth vibrato and microtremor, and the logarithmic nature of speech pitch in speech production (and perception).

“These observations suggest that a more relevant scale on which to assess abnormal variations in speech pitch is the perceptually-relevant, logarithmic (tonal) scale, rather than the absolute frequency scale [42]. It also suggests that in order to better capture pitch period variation due to PD-related dysphonia independent of these natural variations, smooth variations should be removed prior to measuring the extent of such variations.

“To implement these two insights algorithmically, we first obtain the pitch sequence of the phonation and convert to the logarithmic semitone scale, $p(t)$, where p is the semitone pitch at time t. We next analyze the roughness of variations in this sequence over and above any healthy, smooth variations, by first removing linear temporal correlations in this semitone sequence with a standard linear whitening filter (coefficients of which are estimated using linear prediction by the covariance method [43]), to produce the relative semitone variation sequence $r(t)$. This filtering effectively flattens the spectrum of the semitone time series, and removes the effect of the mean semitone (which depends on the individual preferences and gender).

Subsequently, we construct a discrete probability distribution of occurrence of relative semitone variations, P(r). Finally, we calculate the entropy of this probability distribution [44] which then characterizes the extent of (non-Gaussian) fluctuations in the sequence of relative semitone pitch period variations.

“An increase in this entropy measure reflects better the variations over and above natural healthy variations in pitch observed in healthy speech production.” (Little, McSharry, Roberts, Costella, & Moroz, 2007)