Detecting Parkinson's with Multidimensional Voice Program Analysis Parameters

Damien Clark

damien.clark-ricard@ncf.edu | https://github.com/kadify/ | https://www.linkedin.com/in/damienclark/

Background

Parkinson's Disease is an autosomal disease characterized by muscular rigidity, tremors, and postural instability.[1]

In particular, patients with Parkinson's often have differences in their speech due to a degradation in the central nervous system.[2] Symptoms exhibited include:

- *Hypophonia - A softer, breathier or hoarser voice (caused by weakening muscles)*
- *Dysarthria - Slurred or slow speech*
- *Tachyphemia - Rapid speaking and may include stammered speech*
- *A more monotone style of speech*

Objectives

Multidimensional Voice Program (MDVP) analysis is an advanced system which allows for the measurement of 33 quantitative voice parameters. Measurements of vocal stability (jitter) and amplitude perturbation (shimmer) have been shown to be representative of pathologies. Vocal stability can directly be measured and used to model pathologies affecting the voice.

Methods

195 MDVP analysis data records were used to build a model capable of detecting vocal attributes possibly linked with tremors associated with Parkinson's Disease. From these 195 recordings, supplied by 31 people. 23 had Parkinson's and 8 people did not. Of the 33 parameters MDVP extrapolates, 22 were supplied in the dataset.

A comparison of several different statistical and machine learning models were used. Using a combination of upsampling/downsampling and bootstrapping, models were evaluated based upon recall to determine the fidelity the model represented in correctly detecting if a patient did indeed have Parkinson's Disease.

Question

Is it possible to detect whether someone has Parkinson's Disease from MDVP parameters?

Data

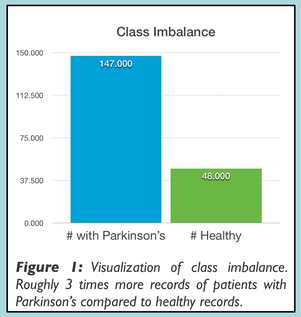

Figure 1: *Visualization of class imbalance. Roughly 3 times more records of patients with Parkinson's compared to healthy records.*

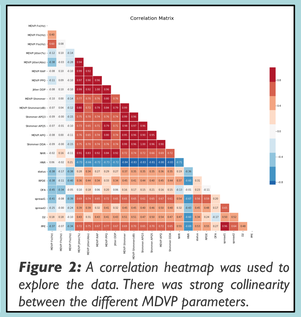

Figure 2: *A correlation heatmap was used to explore the data. There was strong collinearity between the different MDVP parameters.*

*From **Figure 1** & **2** there was a noticeable class inbalance. In addition, many of the parameters are transformations of other parameters. I manually selected 3 parameters to test a logistic regression model. This model was then compared to one with all parameters and using Lasso Regularization, the coefficients utilized and their importance in the model prediction equation was used to determine which parameters were instrumental in predicting whether someone had Parkinson's.*

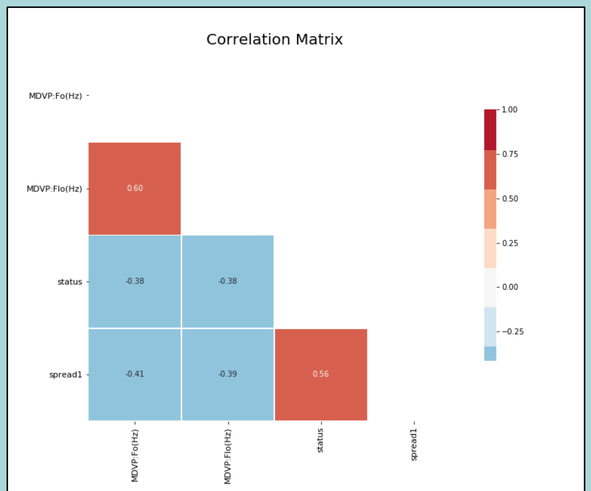

Figure 3: *A correlation heat map of manually selected features.*

Results

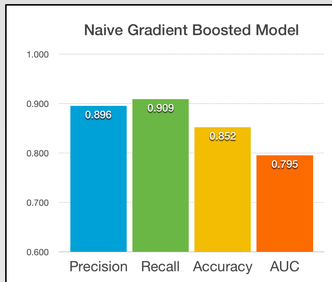

Figure 3: *A naive gradient boosted classifier model on the manually selected features. Using the MDVP: Fo(Hz), MDVP:Flo(Hz), and spread1 features, a precision of 0.896, recall of 0.909, accuracy of 0.852, and a ROC Area Under the Curve value of 0.795 represents the ability of the model to generalize well to unseen data.*

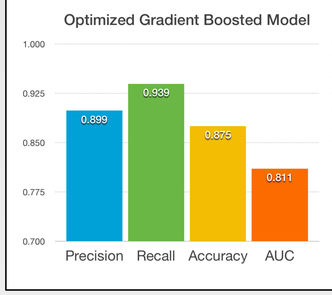

Figure 4: *An optimized gradient boosted classifier model on the manually selected features. Using the MDVP: Fo(Hz), MDVP:Flo(Hz), and spread1 features, a precision of 0.899, recall of 0.939, accuracy of 0.875, and a ROC Area Under the Curve value of 0.811 represents the ability of the model to generalize well to unseen data.*

*From **Figure 3** & **4**, it is apparent that the optimized gradient boosted model generalizes better than the naive model and represents a higher recall score over the naive model.*

Discussion

From Figure 3 & 4, there was a distinct improvement in Precision, Recall, Accuracy and the ROC AUC score between the naive and optimized gradient boosted models. Remarkably there was a nearly 3% increase in the recall score between the optimized model compared to the naive model.

Of note, the model was fit on the manually selected features and not the full feature set due to overfit problems. The model predictions were remarkably worse when fitting the model with the entire feature set and this was reasoned to be due to overfitting of the training dataset. The data was upsampled using SMOTEtomek in combination with a training bootstrap of 100,000 samples due to the limited data available. This resulted in the best recall scores.

Conclusion

This model was based upon a previous study performed by Arvind Kumar Tiwari. In his study, using a Random Forest Classifier he achieved an accuracy score of 0.903 and a precision score of 0.902.[3]

My goal in this investigation was to prioritize recall, as not detecting a case is worse than improperly classifying a case that did not exist. In this case, missing a patient with Parkinson's.

With my optimized model, and using the features MDVP: Fo(Hz), MDVP:Flo(Hz), and spread 1, I achieved a recall score of 0.939 which represents a model that incorrectly misses a diagnosis only ~6% of the time.

Stack

References

1. "Parkinson's Disease vs. Parkinsonism" (PDF). National Parkinson Foundation. Archived (PDF) from the original on 30 August 2017. Retrieved 22 June 2017.

2. Parkinson's Disease. NIH Publication No. 15-139. Dec 2014. National Institute of Neurological Disorders and Stroke, National Institutes of Health.

3. Tiwari, Arvind Kumar. "Machine Learning Based Approaches for Prediction of Parkinson's Disease." Machine Learning and Applications: An International Journal, vol. 3, no. 2, 2016, pp. 33–39., doi:10.5121/mlaij.2016.3203.