Spring 2017: CSCI 6990 Programming Assignment #3
Name: Franklin D. Worrell

Parkinson's disease (PD) is a degenerative condition of the nervous system. Its primary effects are on patients' motor skills. The disease is incurable, though some treatments in the early stages of PD have been shown to slow its progression, which is inevitable. One study whose data were used for the current experiment argues that we can use analysis of a patient's voice to help determine if they have PD, and the authors use pitch period entropy (PPE)—a novel metric of voice disorder that they have created—in combination with other features to produce a model that can correctly classify patients as having PD or as healthy with over ninety percent accuracy.[1] Surveying the literature, they explain that "Research has shown that approximately 90% of [people with Parkinson's disease] exhibit some form of vocal impairment…vocal impairment may also be one of the earliest indicators for the onset of the illness."[2] Specifically, one such test in which patients are asked to sustain a single vowel sound at a particular pitch for as long as possible has proven diagnostically useful.[3] Given that PD primarily effects motor control and voice modulation is a particularly sensitive motor feature, it seems that sustaining a single sound at a single pitch would require the kind of fine-tuned motor control that PD effects. Furthermore, since this action is very sensitive, it also makes sense that it could be used as an early detector.

Extracting features from voice samples to be used for classification is a trickier endeavor. Specifically, the authors of the above referenced study are looking for features that are robust in the face of very serious and significant confounding features, as they would like their tool to be used for remote monitoring via phone or internet. They used both traditional and novel features. All observations were made under the same recording conditions, and the recordings produced were analyzed to extract features.[4] All features were extracted based on analysis of only the first half of the sustained vowel sound—in order to control for patients running out of breath by the end of the test.[5] Both traditional and nontraditional features were extracted by close analyses of, e.g., baseline frequencies, cycles, and sound pressures and comparison with variations therein. The nontraditional features typically involved more significant statistical analysis. For example, the authors' PPE metric should ignore smooth variations in pitch as these are normal; hence, it "assess[es] abnormal variations in speech pitch [by] the perceptually relevant, *logarithmic* (*tonal*) *scale*, rather than the absolute frequency scale."[6] Such analyses allow for more fine-tuned, precise control when comparing variations with baselines and help to establish when deviations are significant and not the result of normal, healthy processes. In short, to extract features, careful and sensitive audio analysis of the recordings was required to meaningfully quantify features.

In this experiment, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and logistic regression (LR) prediction algorithms were implemented in GNU Octave. Ten-fold cross validation was applied, and the error rate for each predictor was recorded. This error rate was taken simply as the percent of test samples that were incorrectly classified. These errors were averaged across the folds to produce the results summarized in Table 1. These results are not entirely surprising. Since it produces a linear decision boundary, LDA produces a less responsive predictor than either of the other algorithms. No basis expansion was performed when LDA was applied; however, doing so might have

---

[1] Little, Max A., Patrick E. McSharry, Eric J. Hunter, Jennifer Spielman, and Lorraine O. Ramig. 2009. "Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease" *IEEE Transactions on Biomedical Engineering* 56.4: 1015-22.
[2] *Ibid*, 1015.
[3] *Ibid*.
[4] *Ibid*, 1016.
[5] *Ibid*, 1018.
[6] *Ibid*.

improved its performance. Minimal basis expansion was performed on the data before QDA was trained and tested. Since the decision boundary produced by QDA is quadratic, it is more responsive; hence, it allows for more accurate prediction during tests. In fact, the improvement here was rather dramatic and a bit surprising. Finally, an initial column of ones was added to the raw observations and LR trained and tested. Unsurprisingly, it outperformed the others.

The provided data set was modified in that the headers for the ARFF file were removed and the file saved as a CSV file. This file, data.csv, must be in the current working directory with the code when the code is run. To run the program, PDDetection.m, simply run it by name as a MATLAB/Octave script. At runtime, the outcome classes are modified from 1 (normal) and 2 (Parkinson's) to 0 (normal) and 1 (Parkinson's) by simply subtracting 1—this eases the use of the LR algorithm.

Table 1: Average Error Rate by Prediction Algorithm

	Error Rate (%)		
Cross-Validation	**LDA**	**QDA**	**LR**
CV1	84.21%	15.79%	0.00%
CV2	75.00%	25.00%	5.00%
CV3	75.00%	25.00%	20.00%
CV4	70.00%	30.00%	10.00%
CV5	65.00%	35.00%	35.00%
CV6	70.00%	30.00%	30.00%
CV7	68.42%	31.58%	15.79%
CV8	78.95%	21.05%	15.79%
CV9	84.21%	15.79%	10.53%
CV10	84.21%	15.79%	15.79%
Average	75.50%	24.50%	15.79%