Gender of a voice prediction

William Wei

Liyewei2019@u.northwestern.edu

EECS 349 Spring 2018 Northwestern University

Introduction

Determining a person's gender as male or female, based upon a sample of their voice seems to initially be an easy task. Often, the human ear can easily detect the difference between a male or female voice within the first few spoken words. However, designing a computer program to do this turns out to be a bit trickier.

My task was to predict the gender of users based on characteristics of voices provided in the dataset. While prior work has investigated male/female classification based on language use, we are specifically inclusive of certain attribute and corresponding importance. Thus, this project makes an important contribution to our understanding of the relationship between gender and voice.

Dataset

The dataset, *Gender Recognition by Voice*, is acquired by the team from the kaggle website. This database was created to identify a voice as male or female, based upon acoustic properties of the voice and speech. The dataset consists of 3,168 recorded voice samples, collected from male and female speakers. The 20% data was regarded as testing set while 80% data was set as training set. And there are 28 attributes overall, and standard deviation of frequency, first quantile, interquantile range, spectral flatness, mode frequency and average of fundamental frequency measured across acoustic signal are chosen as initial tasting attributes. The voice samples are pre-processed by acoustic analysis in R using the seewave and tuneR packages, with an analyzed frequency range of 0hz-280hz.

Methodology

In order to determine which exact properties indicate a target gender of male or female, we could initially guess that it likely one of the statistically significant features, but ultimately this decision breakdown is masked within the model. And to gain an understanding of a trained model, we can apply a classification via regression tree model to our dataset on Weka software to determine how these properties might correspond to a gender classification of male or female.

The classification via regression model achieves an accuracy of 81% on the training set and the mode frequency attribute serves as a root node for detecting

the gender as male or female. From there, it then checks the minimum fundamental frequency, followed by more specific properties.

Then the random forest model is also applied and it achieves an accuracy of 100% on the training set, which is further improvement over the classification via regression model.

Results

```
=== Summary ===

Correctly Classified Instances        3168               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances             3168
```

Future Work

The further plan would be expanding on how exploring new algorithms might be beneficial as well as adding novelty might be to enhance the dataset by computing additional features for the project.

In the meanwhile, it's important to keep in mind what the model is actually trained upon. In the case of voices, there is a large array of both male and female voices that lie within different androgynous zones of frequency and pitch. A dataset that includes a much larger number of samples from the general population would likely train a model that could achieve more accurate results.