Gender Recognition using Voice

Kexian Wu
CWID: 10408805
kwu5@stevens.edu

Yingfei Li
CWID: 10429076
yli224@stevens.edu

I. PROBLEM DESCRIPTION

The speech signal processing has a wide range of applications in most of technical fields. In speech processing, gender identification plays an important role.

A human ear can easily distinguish between a male and female voice. What if a machine want to do the same? What features of a voice would be required for a machine to classify voices by gender?

The voice of an adult male can range between 85 to 180 Hz and that of adult female ranges between 165 to 255 Hz. From figure 1, we can observe that the fundamental frequencies exhibited by male voices are much lower than those exhibited by females.

Although the frequency ranges quite differently, there is a mid range where the voice frequencies seem to overlay. This is why differentiating voices only on the basis of voice frequency is not adequate, it does require additional feature detections.

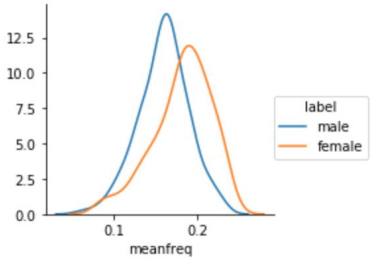

Figure 1: Data Visualization of voice frequencies by genders.

Our goal for this project is to build different models to identify a voice as male or female. This project focuses on the general process of approaching a classification problem, through data exploration, model creation, and model improvement.

II. RELATED WORK

A. Dataset

Acoustic analysis of the voice depends upon parameter settings specific to sample characteristics such as intensity, duration, frequency and filtering. The acoustic properties of the voice and speech can be used to detect the gender of speaker.

We will use the dataset provided by kaggle to train our models. The dataset consists of 3,168 recorded voice samples, collected from male and female speakers. We will be using 80% of the original data for training and the rest for testing.

B. Models

We propose to train the conventional modes to learn the task of classifying male and female voice samples. The models we chose are as following:

1) Decision tree

2) K Nearest Neighbors

3) Logistic Regression

4) Naive Bayes

5) Random Forest

7) Neural Network

8) SVM

C. Techniques

1) Full Feature Set

We will train all the models mentioned above on the dataset containing all the 20 features.

2) PCA

We will reduce the dimension of our data set to fewer dimensions and then trained all the models to observe the changes in accuracy.

3) Voting

We will conduct voting on all the models and use the mode of all results as our prediction to reduce bias.

III. PROJECT MILESTONES

Projected Project Milestones are given in the Chart 1. The most of the work for our project will be training the conventional models to learn the task of classifying male and female voice samples we collected. We will implement various classifiers in Python 3.

After obtaining a solid understanding of the machine learning algorithms, we will write our own functions for most of the machine learning methods we used for this process, and we will compare our results to the ones obtained by using Sklearn/Keras package. We anticipate this will take a large swathe of time and represent that accordingly on our timeline - once this is complete we will then propose some techniques like PCA algorithm to reduce the variance by reducing the dimensionality and improve classification accuracy.

2018/03/27	Finish project proposal and contact TA
2018/04/03	Finalize Project Goals
2018/04/10	Collect data and extract features from audio samples. Build 8 conventional and deep learning models for gender classification
2018/04/17	Continue working on our conventional models and deep learning models
2018/04/24	Use feature extraction techniques such as PCA to reduce dimensionality and improve the accuracies
2018/05/01	Finalize Project
2018/05/08	Finalize report and project submission

Chart 1: Tentative Project Milestones

IV. REFERENCES

[1] Kory Becker, Gender Recognition by Voice and Speech Analysis, http://www.kaggle.com.

[2] L. Breiman, “Random forests”, Machine Learning, Springer US, 45:5– 32, 2001.

[3] M. Abadi, A. Agarwal, TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.