



Featured Dataset

142

Gender Recognition by Voice

Identify a voice as male or female



Kory Becker • last updated a year ago

Overview

Data

Kernels

Discussion

Activity




Download (417 KB)

New Kernel

Tags

- linguistics
- gender
- small
- featured

Top Contributors >

	Niraj Ve...	1st
	Kai Sas...	2nd
	Umesh	3rd

Kernels >

Support Vector ...	57 votes
run 7 months ago	
Evaluation of gen...	27 votes
run a year ago	
TSNE Visualization	18 votes
run 8 months ago	

Discussion >

TSNE Visualizati...	1 reply
17 days ago	
Simple Benchm...	1 reply
4 months ago	
Evaluation of g...	3 replies
5 months ago	

Description

Voice Gender

Gender Recognition by Voice and Speech Analysis

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 voice samples are pre-processed by acoustic analysis in R using the seewave and tuneR packages, with an analyzed frequency range of 0hz-280hz ([human vocal range](#)).

The Dataset

The following acoustic properties of each voice are measured and included within the CSV:

- **meanfreq**: mean frequency (in kHz)
- **sd**: standard deviation of frequency
- **median**: median frequency (in kHz)
- **Q25**: first quantile (in kHz)
- **Q75**: third quantile (in kHz)
- **IQR**: interquantile range (in kHz)
- **skew**: skewness (see note in specprop description)
- **kurt**: kurtosis (see note in specprop description)
- **sp.ent**: spectral entropy
- **sfm**: spectral flatness
- **mode**: mode frequency
- **centroid**: frequency centroid (see specprop)
- **peakf**: peak frequency (frequency with highest energy)
- **meanfun**: average of fundamental frequency measured across acoustic signal
- **minfun**: minimum fundamental frequency measured across acoustic signal
- **maxfun**: maximum fundamental frequency measured across acoustic signal
- **meandom**: average of dominant frequency measured across acoustic signal
- **mindom**: minimum of dominant frequency measured across acoustic signal
- **maxdom**: maximum of dominant frequency measured across acoustic signal
- **dfrange**: range of dominant frequency measured across acoustic signal
- **modindx**: modulation index. Calculated as the accumulated absolute difference between adjacent measurements of fundamental frequencies divided by the frequency range
- **label**: male or female

Accuracy

Baseline (always predict male)

50% / 50%

Logistic Regression

97% / 98%

CART

96% / 97%

Random Forest

100% / 98%

SVM

100% / 99%

XGBoost

100% / 99%

Research Questions

An original analysis of the data-set can be found in the following article:

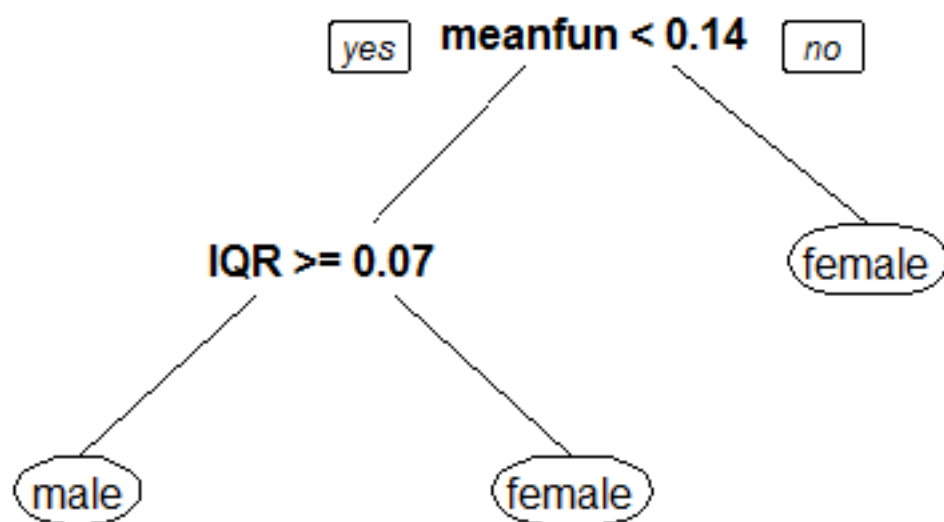
[Identifying the Gender of a Voice using Machine Learning](#)

The best model achieves 99% accuracy on the test set. According to a CART model, it appears that looking at the mean fundamental frequency might be enough to accurately classify a voice. However, some male voices use a higher frequency, even though their resonance differs from female voices, and may be incorrectly classified as female. To the human ear, there is apparently more than simple frequency, that determines a voice's gender.

Questions

- What other features differ between male and female voices?
- Can we find a difference in resonance between male and female voices?
- Can we identify falsetto from regular voices? (separate data-set likely needed for this)
- Are there other interesting features in the data?

CART Diagram



Mean fundamental frequency appears to be an indicator of voice gender, with a threshold of 140hz separating male from female classifications.

References











[The Harvard-Haskins Database of Regularly-Timed Speech](#)

[Telecommunications & Signal Processing Laboratory \(TSP\) Speech Database at McGill University, Home](#)

[VoxForge Speech Corpus, Home](#)

[Festvox CMU_ARCTIC Speech Database at Carnegie Mellon University](#)



		Kaan Can	Ran version 18 of kernel ROC Curve with k-Fold CV	11 hours ago
		Akshay Mewada	Ran version 2 of kernel Gender Recognition by Voice Acoustic Parameters	3 days ago
		ВикторияРябцева	Ran version 1 of kernel Breast Cancer Wisconsin (Diagnostic) Data Set	9 days ago
		Jaydeep Kulkarni	Ran version 4 of kernel Single Logistic Neuron from Scratch	11 days ago
		Andrés D'Ambrosio	Commented on kernel TSNE Visualization	17 days ago

59,598 views · 6,203 downloads · 22 topics

Explore More Datasets