International Journal of Research in Engineering, Science and Management
Volume-3, Issue-6, June-2020
www.ijresm.com | ISSN (Online): 2581-5792

Gender Detection via Non-Appearance Method (Voice)

Abhinav Shivam[1*], Archit Bansal[2], Aman Kumar Singh[3], Abhishek Kumar[4]
[1,2,3,4]*Student, Department of Computer Science and Engineering, IMS Engineering College, Ghaziabad, India*
*Corresponding author: abhinav.shivam2721996@gmail.com

***Abstract*: Acoustic signal, speech (voice), having a property for identifying the gender of an actual speaker. This is also known as Gender Detection (GD). In this research paper, we will find the gender from the given acoustic data i.e. pitch, median, frequency etc. Machine learning gives promising and best results for classification problem in all the field of research domains. There are certain performance metrics to evaluate algorithms of an area for ny project. Our Comparative model algorithm for evaluating 4 different machine learning algorithms based on eight different metrics in gender classification from acoustic data. Location of the person and gender of the person have become very crucial in economic markets in the form of AdSense. Here with this comparative model algorithm, we are also trying to assess the different Machine Learning algorithms and find the best fit for gender classification of acoustic data.**

***Keywords*: Gender detection, Non-appearance.**

1. Introduction

Determining and identifying a person's gender as male, female or transgender, based upon the sample of their voice seems to initially be an easy task but it is not. Often, the human ear can easily detect and identify the difference between a male or female voice within the first spoken sentence. However, designing a computer program or project to do this turns out to be a bit trickier and difficult. This research paper describes the design of a computer program to model acoustic analyse of voices and speech for determining gender as male or female. Diamorphism is the property of voice that is highly observed in human beings both in male and female. Intonation, speech rate, and duration are certain characteristics that distinguish human voices, mainly male and female voices. The perceived dimorphism accounts for 98% which primarily consists of the gender of the speaker and the respective frequencies. Variation in gender, however, cannot be predicted by vocal speech. Some voice pitch may vary between female and male so it is difficult to predict female and male accurately. The samples or data are processed using acoustic analysis and then applied to an artificial intelligence (AI) /machine learning (ML) algorithm to learn gender-specific traits. The resulting program achieves 90% accuracy on the test of dataset.

In this research paper, we propose a comparative model algorithm which classifies gender based on various prediction algorithms such as EM Algorithm. Prediction is done majorly on the dataset whose values are processed from speech. Results that are obtained are compared with previous research and project results and it is evaluated and matched with other algorithms to conclude which algorithm results in better and the best performance for classifying gender based on certain given parameters. Prediction of how accurately this comparative model algorithm detects the gender based on these algorithms is also found.

2. Algorithm

EM Algorithms (Expectation Maximization Algorithms)

An expectation–maximization (EM) algorithm is an repetitive method to find maximum likelihood or maximum of a posteriori (MAP) estimates a parameter in statistical models, where the model or project depends on an unobserved latent variable. The EM iteration alternates to find between performing an Expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimated for the parameter, and a maximization (M) step, which computes parameter maximizing and increasing the expected log-likelihood found on the E step. These parameter-estimates are then used to determine the distribution of the latent variables in the next E step.

A. Application

EM (Expectation Maximization Algorithms) Algorithms is often used as data clustering in machine learning (ML) and computer vision. In natural language processing (NLP), two prominent instances of the algorithm are the Baum–Welch algorithm for hidden Markov models, and the inside-outside algorithm for unsupervised induction of probabilistic context-free grammars.

3. Voice Gender

A. Gender Recognition by Voice and Speech Analysis

The database which was created to identify the voice as female or male, based upon acoustic properties of their voice and speech. The dataset consists of 3,169 recorded voice samples, collected from female and male speakers. The voice samples in dataset 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).

Each voice sample is stored as a .WAV file, which is then pre-processed for acoustic analysis using the specan function from the WarbleR R package. Specan measures 22 acoustic parameters on acoustic signals for which the start and end times are provided.

4. 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

5. Accuracy

Baseline (always predict as male)
50% / 50%

Logistic Regression
96% / 97%

CART
95% / 96%

Random Forest
100% / 97%
SVM
100% / 98%

XGBoost
100% / 98%

6. Testing the Models

For testing the models of our project that we made as an interface for needs to be created for recording human voice. After the human voice has been recorded and then the other noises of the surrounding have been cleared out. Then this recorded input is given to the best model (model with high accuracy compared to others) and the gender is detected.

7. Conclusion

Contrasting the afore mentioned 5 different algorithms with a combined total of eight parameters for example plots and tests such as Box, Density, Parallel, Dot, Pair, and tests like statistical gravity, using Comparative algorithm, we have come to the following wind up.

The diagnosis of the outcome we come to see is that using Support Vector Machine(SVM) algorithm, we get better results in terms of classification and reduced rate of error. The outcome which we got is solely based on the gender dataset we have taken into account and things may differ for unalike dataset. Support Vector Machine(SVM) proves to have more precision over other algorithms when it comes to classifying gender regardless of variations in pitch and frequency. Subsequent work to include more algorithms to this Comparative model and to contrast the effectiveness with this project and to spot which algorithm whether linear or non-linear proves efficient in the classification of gender in vocal gender dataset.

References

[1] Byrd D 1992 Preliminary results on speaker-dependent variation in the
[2] TIMIT database J Acoust Soc Am 92 593-596
[3] Henton C G 1989 Fact and fiction in the description of male and female pitch Language and Communication, 9, 299-311.
[4] Bishop J and Keating P 2012 Perception of pitch location within a speaker's range: Fundamental Frequency, voice quality and speaker sex, The Journal of the Acoustical Society of America 32-2 1100-1112.
[5] A Raahul, R Sapthagiri, K Pankaj and V. Vijayarajan, “Voice based gender classification using machine learning,” IOP Science
[6] https://sarthakahuja.org/public/docs/report_he_or_she.pdf
[7] T. Vogt and E. André, “Improving automatic emotion recognition from speech via gender differentiation,” in Proc. Language Resources and Evaluation Conference (LREC 2006), Genoa, 2006.
[8] H. Harb and L. Chen, “Voice-based gender identification in multimedia applications,” Journal of Intelligent Information Systems, vol. 24, no. 2-3, pp. 179 – 198, 2005.