iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** IV **Month of publication:** April 2019

DOI: https://doi.org/10.22214/ijraset.2019.4341

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue IV, Apr 2019- Available at www.ijraset.com

Gender Prediction by Voice using Logistic Regression

R Praveen Kumar[1], P Sree Varsha[2], L Sandhya Rani[3], G Bharadwaj[4]
[1]*Associate Professor, Department of Computer Science & Engineering, Vignan Institute of Technology and Science*
[2, 3, 4, 5]*B.Tech (CSE), Vignan Institute of Technology and Science*

Abstract: Gender Recognition is a python script that can detect the gender of a speaker from the voice given using Logistic Regression. It is trained using the voice kaggle set and uses pyaudio to record and extract the features from audio segments. Detecting the gender of a person (male or female) through their voice seems to be a very trivial task for humans. Our minds are trained over the course of time to detect the differences in voices of males and females. But it is a challenging problem for computers. So the prime motivation of this project is to detect the gender of the speaker that can be used for the Video Games and Mobile Applications.

I. INTRODUCTION

Gender identity is the personal sense of one's own gender. Gender identity can correlate with assigned sex at birth or can differ from it. All societies have a set of gender categories that can serve as the basis of the formation of a person's social identity in relation to other members of society. A lot of social interactions today depend on gender. Recently with the development of social media environments and smartphones, gender recognition applications have begun to grow and become important. In many fields like facial expression analysis, tracking and surveillance, human-computer interactions, biometry, gender recognition application can be seen. In detail the various uses in different applications are:

1) *Surveillance and Security Systems:* Classifying gender can contribute towards increasing intelligence of security and surveillance systems by the investigation of criminals who intentionally try to hide their identity information.
2) *Video Games:* In games, male and female players have different preferences, automatic gender classification would provide their preferred game characters or contents.
3) *Human-Computer Interaction:* In the area of HCI like virtual assistants, computers are needed to identify and verify gender in order to enhance the performance based on personal information.

II. RELATED WORK

Several types of research works are being carried out in field of Speech processing. A gender prediction system was modelled by extracting first formant and pitch using linear predictive analysis. And the work concluded that both features are high in case of female's voice than that of males. Somya Goeal and Pramit Gupta worked on the gender recognition which uses a stacked technique method. Also the other worked proposed a model to determine the gender using the K nearest neighbour algorithm, but because of the constant K value the model lacked in the accuracy. Che et al. worked on cepstral peak prominence, spectral magnitude and harmonic to noise ratio. The results varied with age group and gained 60% accuracy in 8-10 years and 94% accuracy in 16-17 years group of children. Also a paper focused on the pitch factor where energy entropy, short time energy and zero crossing rates were given as input to fuzzy logic and neural network and calculated percentage of male and female in the speech sample. The proposed technique achieved 65% accuracy, higher than the fuzzy and the neural networks which gave 50% and 60% accuracy.
In the paper the work proposed a model to determine the gender of the speaker using the Logistic regression model with better accuracy.

III. DATA DESCRIPTION

Our data was taken from kaggle which consists of 3,169 datasets with 21 attributes in which it contains 1 independent variable i.e male/female and 20 independent variables like mean frequency, Q25, Q75, skew, IQR, mode, centroid, kurt, spectral entropy, spectral flatness etc. Of all the attributes the attributes that are mainly focused to predict the gender are mode, minimum fundamental frequency, first quantile, third quantile, interquantile range, skew, maximum dominant frequency, mean fundamental frequency and the median.From the observations it is observed that the mean frequency plays a very important role in determining the gender. Also the IQR also plays a significant role.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue IV, Apr 2019- Available at www.ijraset.com

IV. PROPOSED METHODOLOGY

For recognizing the speaker the main attributes that are required to train our model are:

A. Mode
B. minimum fundamental frequency
C. Q25 (first quantile)
D. Q75 (third quantile)
E. Interquantile range
F. skew
G. maximum dominant frequency
H. mean fundamental frequency
I. median

1) Logistic Regression: also known as Logit Regression or Logit Model, is a mathematical model used in statistics to estimate (guess) the probability of an event occurring having been given some previous data. Logistic Regression works with binary data, where either the event happens (1) or the event does not happen (0). So given some feature x it tries to find out whether some event y happens or not. So y can either be 0 or 1. In the case where the event happens, y is given the value 1. If the event does not happen, then y is given the value of 0. For example, if y represents whether a sports team wins a match, then y will be 1 if they win the match or y will be 0 if they do not. This is known as Binomial Logistic Regression. There is also another form of Logistic Regression which uses multiple values for the variable y. This form of Logistic Regression is known as Multinomial Logistic Regression.

Logistic regression uses the concept of odds ratios to calculate the probability. This is defined as the ratio of the odds of an event happening to its not happening.
The odds can be defined as
Odds=p(y=1/x)/1-p(y=1/x)
The natural logarithm of the odds ratio is then taken in order to create the logistic equation.
The new equation is known as the logit:
Logit(p(x))=ln(p(y=1/x)/1-p(y=1-x))
In Logistic regression the Logit of the probability is said to be linear with respect to x, so the logit becomes:
Logit(p(x))=a+bx
Using the two equations together then gives the following:
P(y=1/x)/1-p(y=1/x) = e^(a+bx)
This then leads to the probability:
P(y=1/x)=e^(a+bx)/1+e^(a+bx) = 1/1+e^ -(a+bx)
This final equation is the logistic curve for Logistic regression. It models the non-linear relationship between x and y with an 'S'-like curve for the probabilities that y =1 - that event the y occurs. In this example a and b represent the gradients for the logistic function just like in linear regression. The logit equation can then be expanded to handle multiple gradients. This gives more freedom with how the logistic curve matches the data. The multiplication of two vectors can then be used to model more gradient values and give the following equation:
Logit(P(x))=w_(0)x^(0)+w_(1)x^(1)+.....w_(n)x^(n) =w^(T)x
In this equation w = [w0 , w1 , w2 , ... , wn] and represents the n gradients for the equation. The powers of x are given by the vector x = [1 , x , x2 , .. , xn] . These two vectors give the new logit equation with multiple gradients. The logistic equation then can then be changed to show this:
P(y=1|x)={1 /1+ e^-(w^(T))x)
This is then a more general logistic equation allowing for more gradient values.

2) Process Flow: The training data is used to train the model with the important attributes using the logistic regression. The whole dataset is considered as the training data and the voice given as input is considered as the test data. The background sounds that is considered as the noise will be eliminated while predicting the gender of the speaker.

©IJRASET: All Rights are Reserved

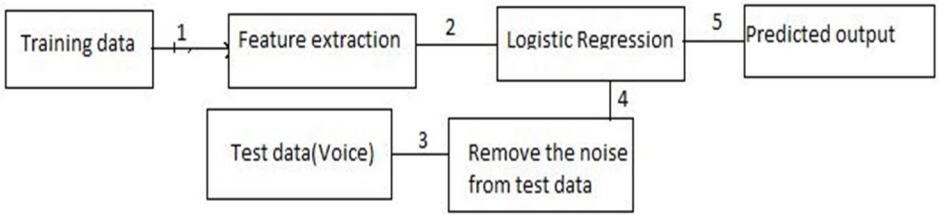


Figure 1. System Architecture

V. CONCLUSION

We proposed a methodology that detects the gender of the speaker with real time test data a new solution to detect the gender of the speaker. In order to achieve our goal, we conducted, feature extraction and applying the data to Logistic Regression. The test data(our voice) is recorded and converted into numerical values and also the background noise the subtracted from the test data. The main 9 features are used to train the model. So our proposed model with the Logistic Regression gives the 97% accuracy. Figure 2 shows the importance of individual feature in the model. From the figure Logistic Regression considers the Mean Fundamental Frequency to be the most useful feature. The IQR and the Standard Deviation are considered as the significant features in predicting the gender.

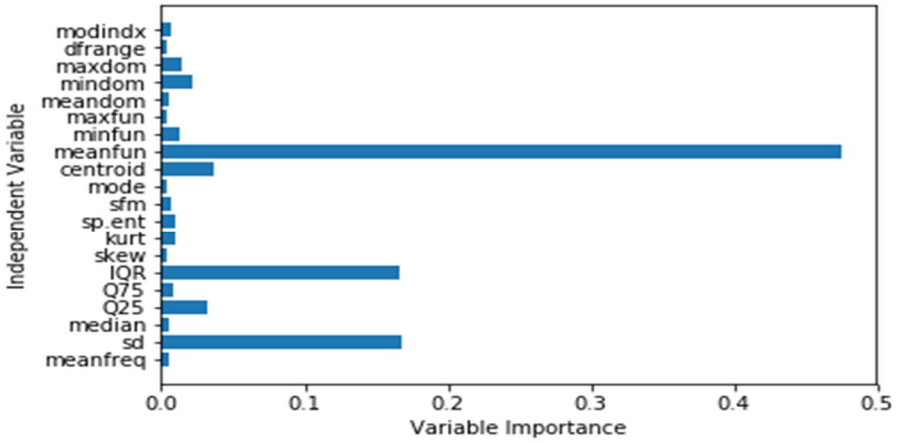
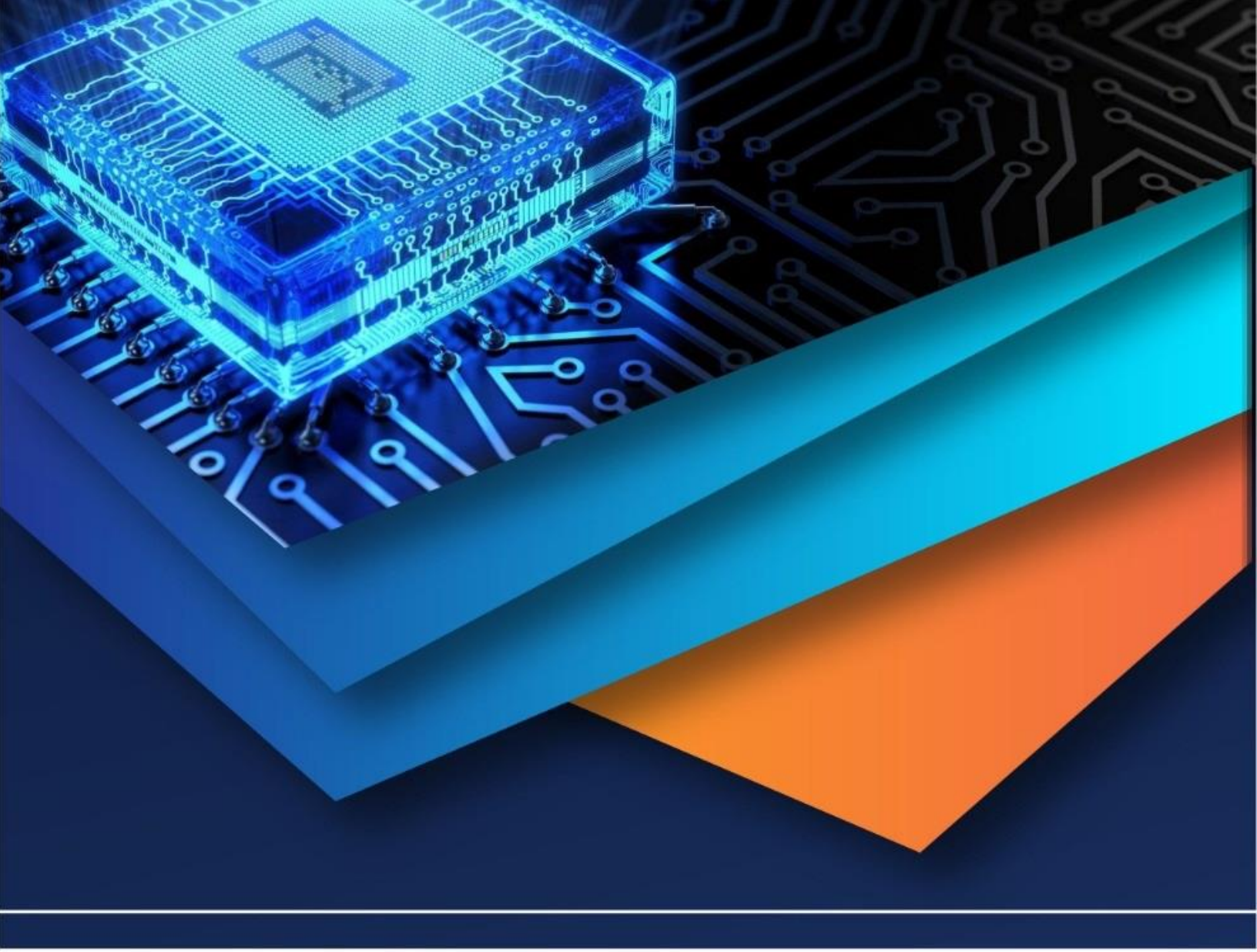


Figure 2. Importance of individual feature in the model.

REFERENCES

[1] Whiteside S P. 1996. Temporal-Based Acoustic-phonetic Patterns in read speech. International Phonetic Association 26 23-40
[2] Zeng Y M,Wn Z Y,Falk T and Chan W Y. 2006. Robust GMM based gender classification using pitch and RASTA-PLP Parameters of speech. Proceeding of the International Conference on Machine Learning and Cybernetics 3376-3379
[3] Udry J. R., "The nature of gender," Demography, vol. 31, pp. 561-573, 1994.
[4] Ting, H, Yingchun, Zhaohui, W. 2006. Combining MFCC and Pitch to Enhance the Performance of the Geder Recognition, IEEE.
[5] E.S. Parris and M.J. Carey, Language Independent Gender Identification, ICASSP, pp 685-688, 1996. □
[6] P. Alku, “Parameterisation methods of the glottal flow estimated by inverse filtering,” in Proc ISCA Workshop on Voice Quality(VOQUAL03), Switzerland 2003, pp. 81–87. □
[7] Gaikwad S,GGawali B,and Mehrotra S C. 2012. Gender Identification using Svm with the combination of MFCC, Advances in Computational Research. 4 67-73
[8] Smith D R and patterson R D. 2005. The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size,sex and age. The genral of the Acoustical Society of America, 118-5. 3177-3186
[9] A.P. Vogel, P. Maruff, P. J. Snyder, J.C. Mundt, Standardization of pitch- range settings in voice acoustic analysis, Behavior Research Methods, v.41, n.2, p.318-324, 2009.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)