iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.44347

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: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Gender Voice Recognition Using Machine Learning Algorithms

Arun Pal Bommaji[1], Neeraj Shiwani[2], Arushi Shandilya[3]
[1, 2, 3]*Dept.Of CSE, CMR Technical Campus*

Abstract: ***Gender identification is one of the major problems in the area of signal processing. The system deals with finding the gender of a person using oral features. One of the most persnickety problems faced is feature selection from wide range of features, which is distinguishing factor in classifying the gender of aperson. The ideal of this design is to design a system that determines the speaker gender using the pitch of the speaker's voice. relating the gender from the plots of voice data set i.e., pitch, median, frequency etc. can be possible by using machine learning. In this design, we're trying to classify gender into male or female based on the data set containing varied attributes related to voice like pitch, frequency etc. The data set have features with explanation data points recorded samples of male and female voices. The data set can be trained with different machine learning algorithms. The proposed system can determine the gender of the speaker with real time test data a new result to discover the gender of the speaker using Fast Fourier Transform with Logistic Regression***
Keywords: Voice recognition, machine learning- Frequency Cepstrum Coefficients, Gaussain mixture model.

I. INTRODUCTION

Human based gender voice recognition system is used to recognize the gender of the person whether the person is male or female. These gender voice recognition machines have currently received symbolic attention in computer vision community. The capability to do automatic recognition of human gender is essential for several systems that process human- source information like information retrieval, human- robot intercommunication etc. Gender voice recognition has great significance in games, business intelligence, demographic scan and forensics.

II. LITERATURE SURVEY

An automatic male-female voice discrimination system has two steps extraction of audio features like ZCR, STE, Pitch, Tempo, MFCC etc from the input speech signal and discrimination based on the uprooted feature. After extracting features, using supervised classifier(like k- NN classifier) voice division will be erected which is computationally affordable and able of differentiating a male and female voice. The performance of the systems will be estimated for a wide range of speech quality. As our main ideal is to determine to which class or gender a particular speech sample belongs to, with the help of point birth and posterior bracket. We propose a methodology for answering the problem. Raw data collected would be pre-processed for missing data, anomalies and outliers. furthermore an algorithm would be trained on this data to produce a model. This model would be used for forecasting the final results. ETL stands for Extract, Transform and load. It's a tool which is a combination of three functions. It's used to get data 8 from one database and transform it into a suitable format. Data prepossessing is a data mining technique used to transform sample raw data into an accessible format. Real world collected data may be inconsistent, incomplete or contains an error and hence data preprocessing is required.

III. PROPOSED SYSTEM

Human based gender voice recognition system is used to recognize the gender of the person whether the person is male or female. Where as in proposed system we're adding the(GMM) Gaussian mixture models, A Gaussian Mixture Model(GMM) is a parametric probability consistence function represented as a weighted sum of Gaussian component consistence. GMMs are generally used as a parametric model of the probability distribution of nonstop measures or features in a biometric system, similar as spoken- tract related spectral features in a speaker recognition system. GMM parameters are estimated from training data using the iterative Expectation- Maximization(EM) algorithm or Maximum A Posteriori(MAP) estimation from a well- trained previous model.

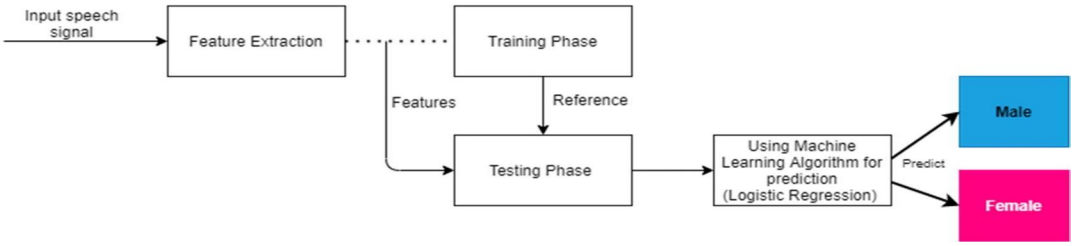


Figure1.1 Architecture of the Model.

IV. RESULT

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of factors,sub-assemblies, assemblies and/ or a finished product It's the process of exercising software with the intent of icing that the

1) The system results in a 95 preciseness of gender discovery.
2) The law can be further optimized usingmulti-threading, acceleration libs and multi- processing.
3) The delicacy can be further bettered using GMM normalization aka a UBM- GMM system.

```
vector = extract_features('/content/drive/MyDrive/TestingData/males/m0003_us_m0003_00011.wav')
winner = identify_gender(vector)
print(winner)

WARNING:root:frame length (800) is greater than FFT size (512), frame will be truncated. Increase NFFT to avoid.
male
```

FIG 1.2 Result

```
vector = extract_features('/content/drive/MyDrive/TestingData/females/f0001_us_f0001_00018.wav')
winner = identify_gender(vector)
print(winner)

WARNING:root:frame length (800) is greater than FFT size (512), frame will be truncated. Increase NFFT to avoid.
female
```

FIG 1.3 Result

```
class FeaturesExtractor:
    def __init__(self):
        pass

    def extract_features(self, audio_path):
        """
        Extract voice features including the Mel Frequency Cepstral Coefficient (MFCC)
        from an audio using the python_speech_features module, performs Cepstral Mean
        Normalization (CMS) and combine it with MFCC deltas and the MFCC double
        deltas.

        Args:
            audio_path (str) : path to wave file without silent moments.
        Returns:
            (array) : Extracted features matrix.
        """
        rate, audio  = read(audio_path)
        mfcc_feature = mfcc(# The audio signal from which to compute features.
                            audio,
                            # The samplerate of the signal we are working with.
                            rate,
                            # The length of the analysis window in seconds.
                            # Default is 0.025s (25 milliseconds)
                            winlen      = 0.05,
                            # The step between successive windows in seconds.
                            # Default is 0.01s (10 milliseconds)
                            winstep     = 0.01,
                            # The number of cepstrum to return.
                            # Default 13.
                            numcep      = 5,
                            # The number of filters in the filterbank.
                            # Default is 26.
                            nfilt       = 30,
                            # The FFT size. Default is 512.
                            nfft        = 512,
                            # If true, the zeroth cepstral coefficient is replaced
                            # with the log of the total frame energy.
                            appendEnergy = True)


        mfcc_feature  = preprocessing.scale(mfcc_feature)
        deltas        = delta(mfcc_feature, 2)
        double_deltas = delta(deltas, 2)
        combined      = np.hstack((mfcc_feature, deltas, double_deltas))
        return combined
```

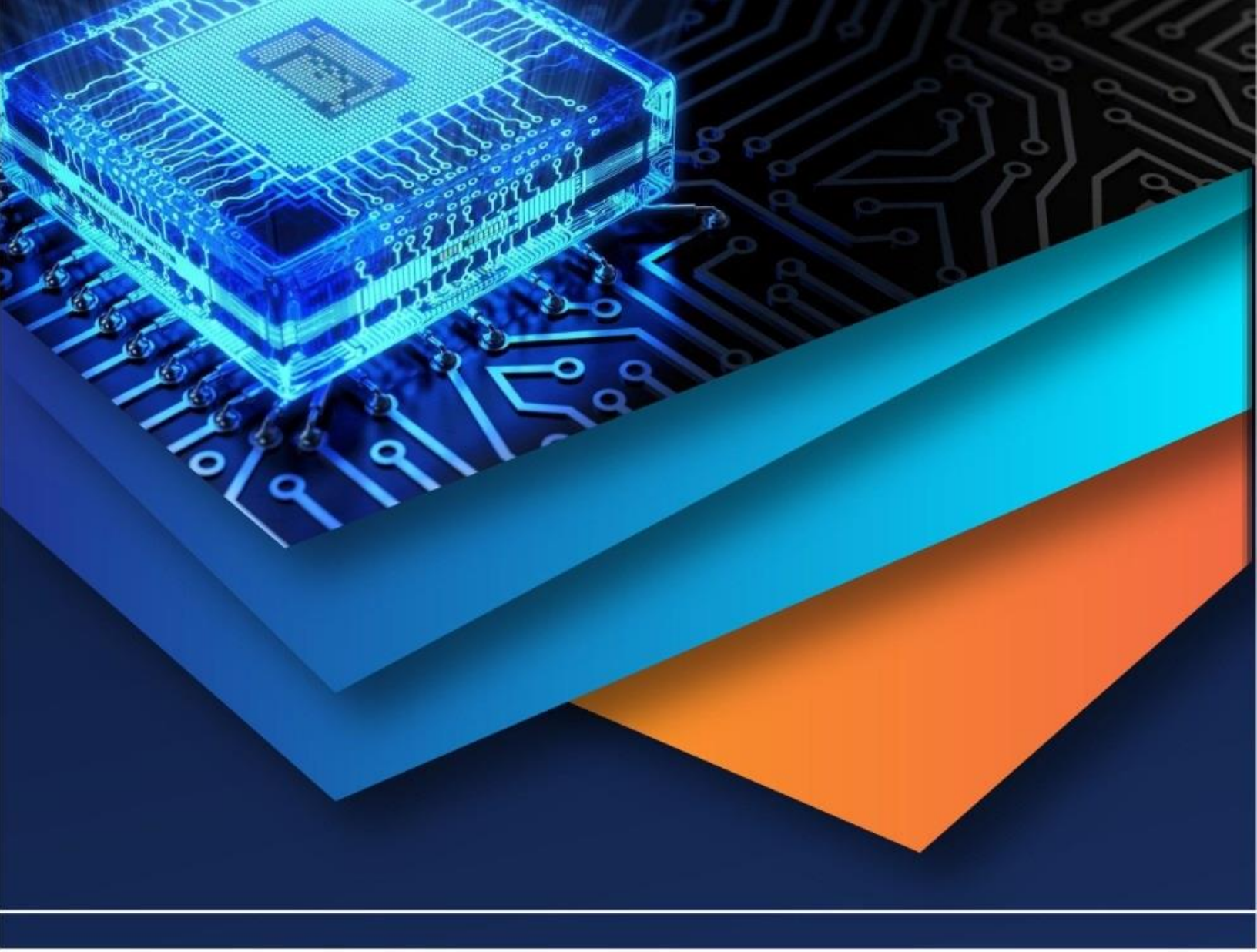
FIG 1.4 Code Execution

V. CONCLUSION

The proposed methodology is to descry the gender of the speaker with real time test data a new result to descry the gender of the speaker. The Fast Fourier transfigure with Machine Learning model has achieved delicacy of 95 on test data which is more as compared to MFCC gender voice recognition model

REFERENCES

[1] J.M. Hilbe, Logistic Regression Models, CRC Press, 2009.
[2] M. Araya- Salas,G. Smith- Vidaurre, warbleR an R package to streamline analysis of beast aural signals. styles Ecol Evolution, 2016, doi10.1111/ 2041-210X.12624
[3] Voice Gender Recognizer Recognition of Gender from Voice using Deep Neural Networks By Lakhan Jasuja, Akhtar Rasool, Gaurav Hajela at 2020 International Conference on Smart Electronics and Communication(ICOSEC).
[4] Gender recognition system using speech signal by Md. Sadek Ali, Md. Shariful Islam and Md. Alamgir Hossain



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY