VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELAGAVI – 590018, Karnataka

INTERNSHIP REPORT

ON

"Voice Classification using ML"

Submitted in partial fulfillment for the award of degree(18CSI85)

BACHELOR OF ENGINEERING IN
ComputerScience

Submitted by:

ISAAC A

1AK19CS017

Conducted at
COMPSOFT TECHNOLOGIES

AKSHAYA INSTITUTE OF TECHNOLOGY

Department of Computer Science

Accredited by NBA, New Delhi

Oblapura Post,Kortgere Tumkur 572-106

ABC INSTITUTE OF TECHNOLOGY
Department of branch Name
Accredited by NBA, New Delhi
Oblapura Post,Kortgere Tumkur 572-106

CERTIFICATE

This is to certify that the Internship titled **"Voice Classification using ML"** carried out by **Mr. ISAAC,** a bonafide student of abc Institute of Technology, in partial fulfillment for the award of **Bachelor of Engineering**, in **ComputerScience** under Visvesvaraya Technological University, Belagavi, during the year 2022-2023. It is certified that all corrections/suggestions indicated have been incorporated in the report.

The project report has been approved as it satisfies the academic requirements in respect of Internship prescribed for the course Internship / Professional Practice (18CSI85)

Signature of Guide **Signature of HOD** **Signature of Principal**

External Viva:

Name of the Examiner Signature with Date

1)_____
_____

2)_____
_____

D E C L A R A T I O N

I, Isaac A, final year student of Branch, College Name - 560 082, declare that the Internship has been successfully completed, in **COMPSOFT TECHNOLOGIES**. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Branch name, during the academic year 2022-2023.

Date : :

Place :

USN : 1AK19CS017

NAME : ISAAC A

INTERNSHIP OFFER LETTER

Date: **23rd August, 2022**

Name: **Isaac A**
USN: **1AK19CS017**

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning With Python(ResearchBased)** Internship position with **Compsoft Technologies**, effective Start Date **23rd August,**
2022, All of us are excited about this opportunity provided to you!

This internship is viewed as being an educational opportunity for you, rather than a part-time job. As such, your internship will include training/orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts of **Machine Learning With Python(Research Based)** through hands-on application of the knowledge you learn while you train with the senior developers. You will be bound to follow the rules and regulations of the company during your internship duration.

Again, congratulations and we look forward to working

with you!.Sincerely,

Nithin K. S
Project Manager
COMPSOFT
TECHNOLOGIES *No.*
363, 19th main
road, 1st Block
Rajajinagar
Bangalore -
560010

A C K N O W L E D G E M E N T

This Internship is a result of accumulated guidance, direction and support of several important persons. We take this opportunity to express our gratitude to all who have helped us to complete the Internship.

We express our sincere thanks to our Principal, for providing us adequate facilities to undertake this Internship.

We would like to thank our Head of Dept – branch code, for providing us an opportunity to carry out Internship and for his valuable guidance and support.

We would like to thank our (Lab assistant name) Software Services for guiding us during the period of internship.

We express our deep and profound gratitude to our guide, Guide name, Assistant/Associate Prof, for her keen interest and encouragement at every step in completing the Internship.

We would like to thank all the faculty members of our department for the support extended during the course of Internship.

We would like to thank the non-teaching members of our dept, forhelping us during the Internship.

Last but not the least, we would like to thank our parents and friends without whose constant help, the completion of Internship would have not been possible.

ISAAC A
1AK19CS017

ABSTRACT

Communication is the key to express one's thoughts and ideas clearly. Amongst all forms of communication, speech is the most preferred and powerful form of communications in human. The era of the Internet of Things (IoT) is rapidly advancing in bringing more intelligent systems available for everyday use.

These applications range from simple wearables and widgets to complex self-driving vehicles and automated systems employed in various fields. Intelligent applications are interactive and require minimum user effort to function, and mostly function on voice-based input.

This creates the necessity for these computer applications to completely comprehend human speech.

A speech percept can reveal information about the speaker including gender, age, language, and emotion.

Several existing speech recognition systems used in IoT applications are integrated with an emotion detection system in order to analyze the emotional state of the speaker.

The performance of the emotion detection system can greatly influence the overall performance of the IoT application in many ways and can provide many advantages over the functionalities of these applications.

This research presents a speech emotion detection system with improvements over an existing system in terms of data, feature selection, and methodology that aims at classifying speech percepts based on emotions, more accurately.

Table of Contents

CHAPTER 1

COMPANY PROFILE

1. COMPANY PROFILE

A Brief History of Compsoft Technologies

Compsoft Technologies, was incorporated with a goal ”To provide high quality and optimal Technological Solutions to business requirements of our clients”. Every business is a different and has a unique business model and so are the technological requirements. They understand this and hence the solutions provided to these requirements are different as well. They focus on clients requirements and provide them with tailor made technological solutions. They also understand that Reach of their Product to its targeted market or the automation of the existing process into e-client and simple process are the key features that our clients desire from Technological Solution they are looking for and these are the features that we focus on while designing the solutions for their clients.

Sarvamoola Software Services. is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Sarvamoola Software Services. specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements.

Compsoft Technologies, strive to be the front runner in creativity and innovation in software development through their well-researched expertise and establish it as an out of the box software development company in Bangalore, India. As a software development company, they translate this software development expertise into value for their customers through their professional solutions.

They understand that the best desired output can be achieved only by understanding the clients demand better. Compsoft Technologies work with their clients and help them to defiine their exact solution requirement. Sometimes even they wonder that they have completely redefined their solution or new application requirement during the brainstorming session, and here they position themselves as an IT solutions consulting group comprising of high caliber consultants.

They believe that Technology when used properly can help any business to scale and achieve new heights of success. It helps Improve its efficiency, profitability, reliability; to put itin one sentence ” Technology helps you to Delight your Customers” and that is what we want to achieve.

CHAPTER 2

ABOUT THE COMPANY

2. ABOUT THE COMPANY

Compsoft Technologies is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Compsoft Technologies specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements. The organization where they have a right mix of professionals as a stakeholders to help us serve our clients with best of our capability and with at par industry standards.They have young, enthusiastic, passionate and creative Professionals to develop technologicalinnovations in the field of Mobile technologies, Web applications as well as Business and Enterprise solution. Motto of our organization is to "Collaborate with our clients to provide them with best Technological solution hence creating Good Present and Better Future for our client which will bring a cascading a positive effect in their business shape as well". Providing a Complete suite of technical solutions is not just our tag line, it is Our Vision for Our Clients and for Us, We strive hard to achieve it.

Products of Compsoft Technologies.

Android Apps

It is the process by which new applications are created for devices running the Android operating system. Applications are usually developed in Java (and/or Kotlin; or other such option) programming language using the Android software development kit (SDK), but other development environments are also available, some such as Kotlin support the exact same Android APIs (and bytecode), while others such as Go have restricted API access.

The Android software development kit includes a comprehensive set of development tools. These include a debugger, libraries, a handset emulator based on QEMU, documentation, sample code, and zutorials. Currently supported development platforms include computers running Linux (any modern desktop Linux distribution), Mac OS X 10.5.8 or later, and Windows 7 or later. As of March 2015, the SDK is not available on Android itself, but softwaredevelopment is possible by using specialized Android applications.

Web Application

It is a client–server computer program in which the client (including the user interface and client- side logic) runs in a web browser. Common web applications include web mail, online

retail sales, online auctions, wikis, instant messaging services and many other functions. web applications use web documents written in a standard format such as HTML and JavaScript,which are supported by a variety of web browsers. Web applications can be considered as a specifific variant of client–server software where the client software is downloaded to the client machine when visiting the relevant web page, using standard procedures such as HTTP. The Client web software updates may happen each time the web page is visited. During the session, the web browser interprets and displays the pages, and acts as the universal client for any web application. The use of web application frameworks can often reduce the number of errors in a program, both by making the code simpler, and by allowing one team to concentrate on the framework while another focuses on a specifified use case. In applications which are exposed to constant hacking attempts on the Internet, security-related problems can be caused by errors in the program.

Frameworks can also promote the use of best practices such as GET after POST. There are some who view a web application as a two-tier architecture. This can be a “smart” client that performs all the work and queries a “dumb” server, or a “dumb” client that relies on a “smart” server. The client would handle the presentation tier, the server would have the database (storage tier), and the business logic (application tier) would be on one of them or on both. While this increases the scalability of the applications and separates the display and the database, it still doesn‟t allow for true specialization of layers, so most applications will outgrow this model. An emerging strategy for application software companies is to provide web access to software previously distributed as local applications. Depending on the type of application, it may require the development of an entirely different browser-based interface, or merely adapting an existing application to use different presentation technology. These programs allow the user to pay a monthly or yearly fee for use of a software application without having to install it on a local hard drive. A company which follows this strategy is known as an application service provider (ASP), and ASPs are currently receiving much attention in the software industry.

Security breaches on these kinds of applications are a major concern because it can involve both enterprise information and private customer data. Protecting these assets is an important part of any web application and there are some key operational areas that must be included in the development process. This includes processes for authentication, authorization, asset handling, input, and logging and auditing. Building security into the applications from the beginning can be more effective and less disruptive in the long run.

Web design

It is encompasses many different skills and disciplines in the production and maintenance of websites. The different areas of web design include web graphic design; interface design; authoring, including standardized code and proprietary software; user experience design; and

search engine optimization. The term web design is normally used to describe the design process relating to the front-end (client side) design of a website including writing mark up. Web design partially overlaps web engineering in the broader scope of web development. Web designers are expected to have an awareness of usability and if their role involves creating mark up then they are also expected to be up to date with web accessibility guidelines. Web design partially overlaps web engineering in the broader scope of web development.

Departments and services offered

Compsoft Technologies plays an essential role as an institute, the level of education, development of student's skills are based on their trainers. If you do not have a good mentor then you may lag in many things from others and that is why we at Compsoft Technologies gives you the facility of skilled employees so that you do not feel unsecured about the academics. Personality development and academic status are some of those things which lie on mentor's hands. If you are trained well then you can do well in your future and knowing its importance of Compsoft Technologies always tries to give you the best.

They have a great team of skilled mentors who are always ready to direct their trainees in the best possible way they can and to ensure the skills of mentors we held many skill development programs as well so that each and every mentor can develop their own skills with the demands of the companies so that they can prepare a complete packaged trainee.

Services provided by Compsoft Technologies.

• Core Java and Advanced Java

• Web services and development

• Dot Net Framework

• Python

• Selenium Testing

• Conference / Event Management Service

• Academic Project Guidance

• On The Job Training

• Software Training

CHAPTER 3

INTRODUCTION

3. INTRODUCTION

Introduction to ML

According to a report from the United Nations (22), an increasing number of people will interact with a voice assistance machine than with their partners in the next five years. With proliferation of Virtual Personal Assistants (VPA) such as Siri, Alexa and Google Assistant in our day-to-day interactions, they fill a role of answering our questions and fulfilling our requests quickly and accurately. Though these assistants understand our commands, they are not proficient enough in recognizing our mood and reacting accordingly. Therefore, it is pertinent to develop an efficient emotion recognition system which can enhance the capabilities of these assistants and revolutionize the whole industry. Speech is a rich, dense form of communication that can convey information effectively. It contains two types of information, namely linguistic and paralinguistic. The former refers to the verbal content, the underlying language code, while the latter refers to the implicit information such as body language, gestures, facial expressions, tone, pitch, emotion etc. Paralinguistic characteristics can help understand the mental state of the person (emotion), gender, attitude, dialect, and more (24). Recorded speech has key features that can be leveraged to extract information, such as emotion, in a structured way. To get such information would be invaluable in facilitating more natural conversations between the virtual assistant and the user since emotion color everyday human interactions. There are two widely used representations of emotion: continuous and discrete. In the continuous representation, the emotion of an utterance can be expressed as continuous values along multiple psychological dimensions. According to Ayadi, Kamel, & Karray (2011) (1), “emotion can be characterized in two dimensions: activation and valence.” Activation is the “amount of energy required to express a certain emotion” (p. 573) and research has shown that joy, anger, and fear can be linked to high energy and pitch in speech, whereas sadness can be linked to low energy and slow speech. Valence gives more nuance and helps distinguish between emotions like being angry and happy since increased activation can indicate both (p. 573). In the discrete representation, emotions can be discretely expressed as specific categories, such as angry, sad, happy, etc.

Problem Statement:

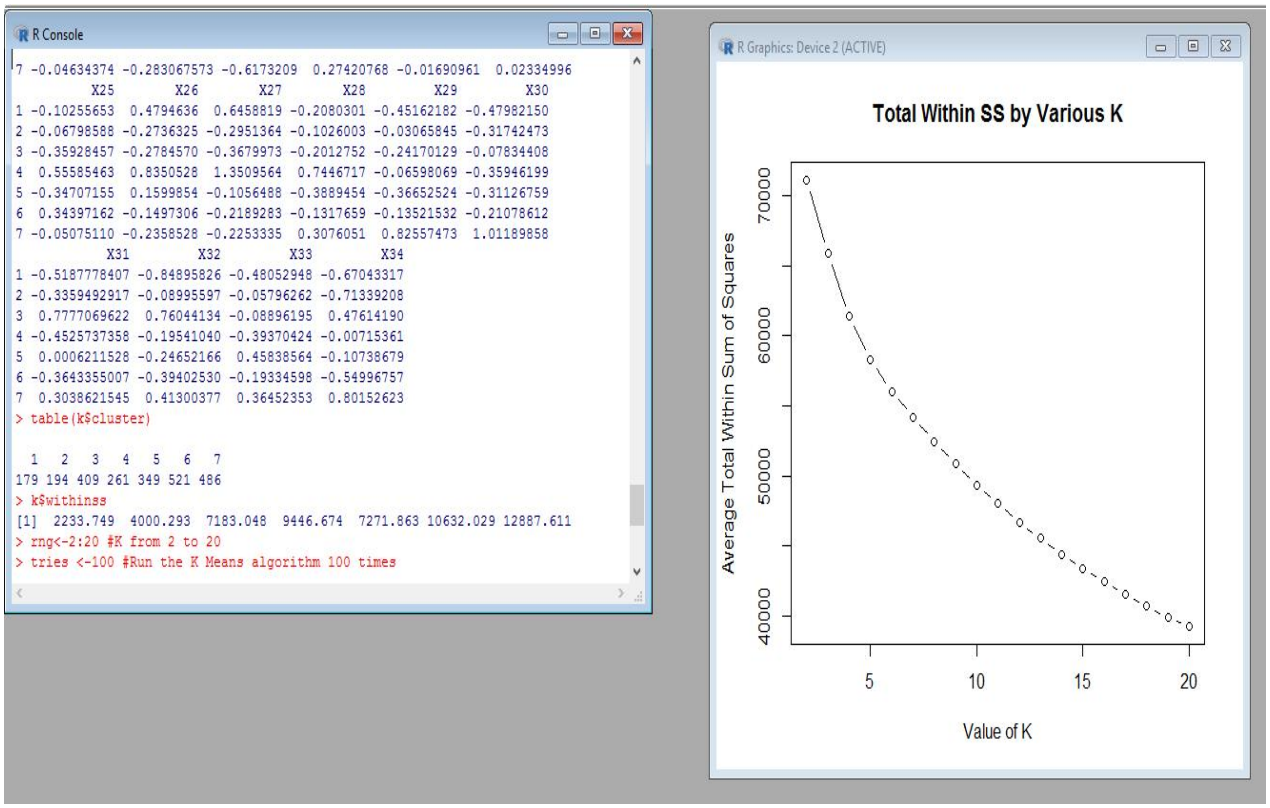
Voice Classification using ML that analyses the sentiment behind the tone of the voice and predicts the sentiment involved

CHAPTER 4

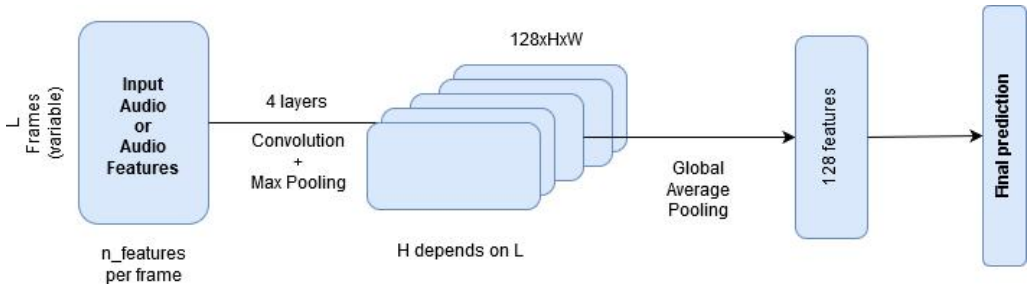
SYSTEM ANALYSIS

4. SYSTEM ANALYSIS

1. Existing System



2. Proposed System



3. Objective of the System

	Predicted class		
Actual Class		Class = Yes	Class = No
	Class = Yes	True Positive	False Negative
	Class = No	False Positive	True Negative

CHAPTER 5

REQUIREMENT ANALYSIS

5. REQUIREMENT ANALYSIS

ALGORITHM	APPROACH 1	APPROACH 2	APPROACH 3
SVM	**83%**	**77%**	**90%**
Decision Tree	**71%**	**65%**	**68%**
KNN	**80%**	**80%**	**87%**
Logistic Regression	**70%**	**77%**	**86%**
Random Forest	**76%**	**69%**	**72%**
Gaussian Naïve Bayes	**74%**	**73%**	**75%**
Gradient Boosting Trees	**77%**	**69%**	**75%**

```
train_img, test_img, train_lbl, test_lbl = train_test_split( x, y, test_size=1/7.0, random_state=0)
```

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
# Fit on training set only.
scaler.fit(train_img)
# Apply transform to both the training set and the test set.
train_img = scaler.transform(train_img)
test_img = scaler.transform(test_img)
```

```
from sklearn.decomposition import PCA
# Make an instance of the Model
pca = PCA(.95)
```

```
pca.fit(train_img)
```

```
PCA(copy=True, iterated_power='auto', n_components=0.95, random_state=None,
  svd_solver='auto', tol=0.0, whiten=False)
```

```
print(pca.n_components_ )
```

```
25
```

```
train_img = pca.transform(train_img)
test_img = pca.transform(test_img)
```

CHAPTER 6

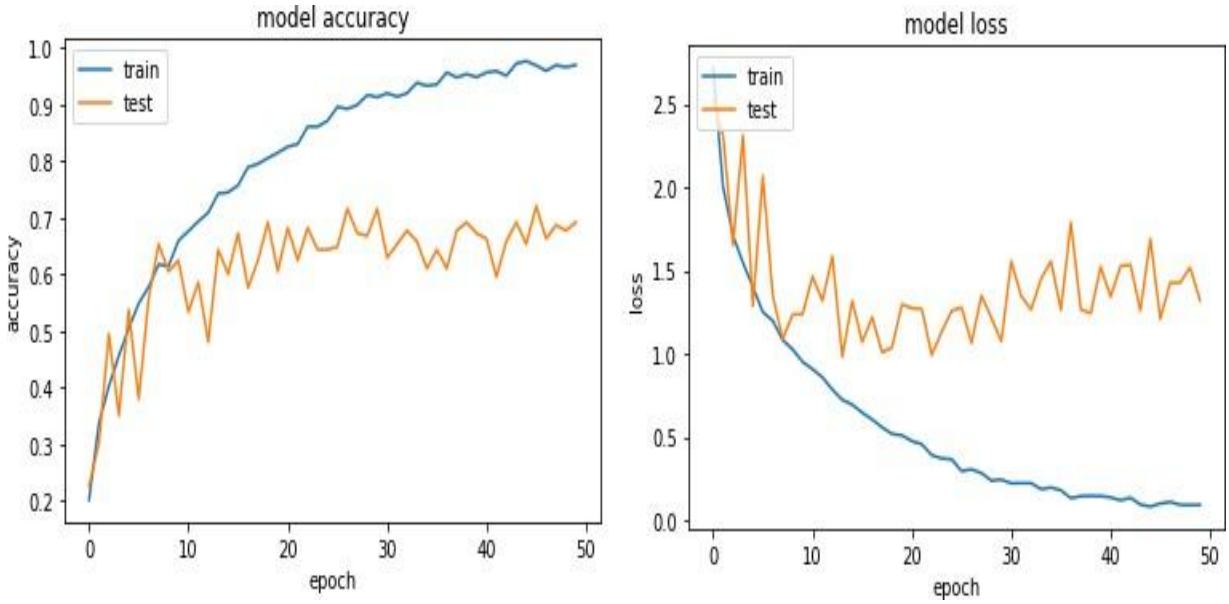
DESIGN ANALYSIS

6. DESIGN & ANALYSIS

1D CNNs and 2D CNNs were implemented on 6 emotion class classification and 12 gender+emotion class classification parallelly and it was observed that including gender gave better performance. Since CNNs are natural feature extractors, 1D CNN and 1D CNN-LSTM architectures were trained on the raw audio input.

2D CNNs were implemented on the engineered features such as MFCCs and Log-mel spectrogram. The training of 2D CNNs started with 2 convolutional layers with 3 3 filters and max pooling with
2 2 filters with stride 2. They were tuned by adding more convolutional layers and increasing the filter sizes in the initial layers. It was found that increasing the depth beyond 4 layers did not improve performance. Also the champion model on 14 class prediction was obtained with 12×12 filters and

7×7 filters in the first and second layers respectively. Also, the final 2 layers had 3×3 filters.



CHAPTER 7

IMPLEMENTATION

7. IMPLEMENTATION

Implementation is the stage where the theoretical design is turned into a working system. The most crucial stage in achieving a new successful system and in giving confidence on the new system for the users that it will work efficiently and effectively.

The system can be implemented only after thorough testing is done and if it is found to work according to the specification. It involves careful planning, investigation of the current system and it constraints on implementation, design of methods to achieve the change over and an evaluation of change over methods a part from planning.

Two major tasks of preparing the implementation are education and training of the users and testing of the system. The more complex the system being implemented, the more involved will be the system analysis and design effort required just for implementation.

The implementation phase comprises of several activities. The required hardware and software acquisition is carried out. The system may require some software to be developed. For this, programs are written and tested. The user then changes over to his new fully tested system and the old system is discontinued.

TESTING

The testing phase is an important part of software development. It is the Information zed system will help in automate process of finding errors and missing operations and also a complete verification to determine whether the objectives are met and the user requirements are satisfied. Software testing is carried out in three steps:

1. The first includes unit testing, where in each module is tested to provide its correctness, validity and also determine any missing operations and to verify whether theobjectives have been met. Errors are noted down and corrected immediately.
2. Unit testing is the important and major part of the project. So errors are rectified easily in particular module and program clarity is increased. In this project entire system is divided into several modules and is developed individually. So unit testing is conducted to individual modules.
3. The second step includes Integration testing. It need not be the case, the software whose modules when run individually and showing perfect results, will also show perfect results when run as a whole.

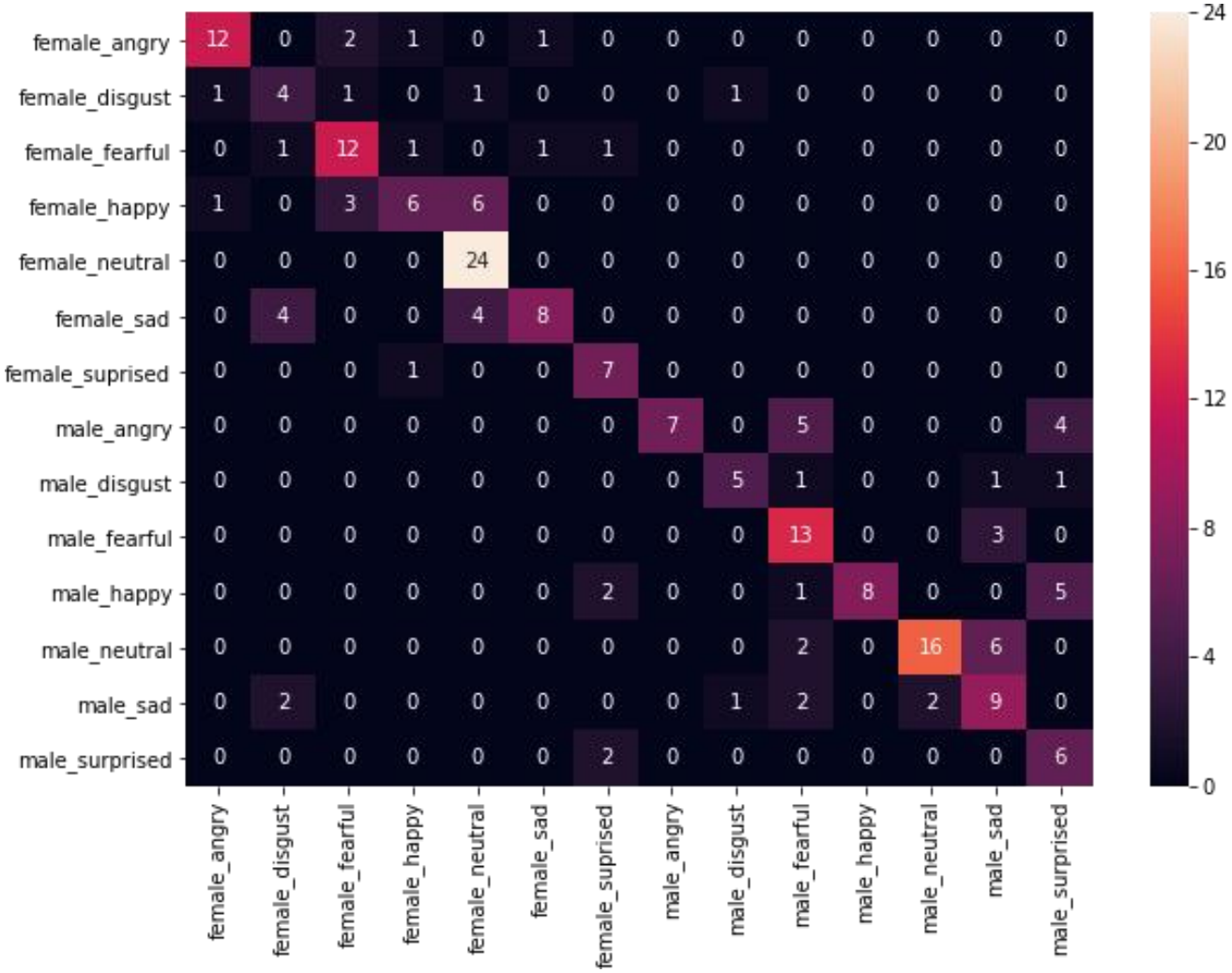
CHAPTER 8

SNAPSHOTS

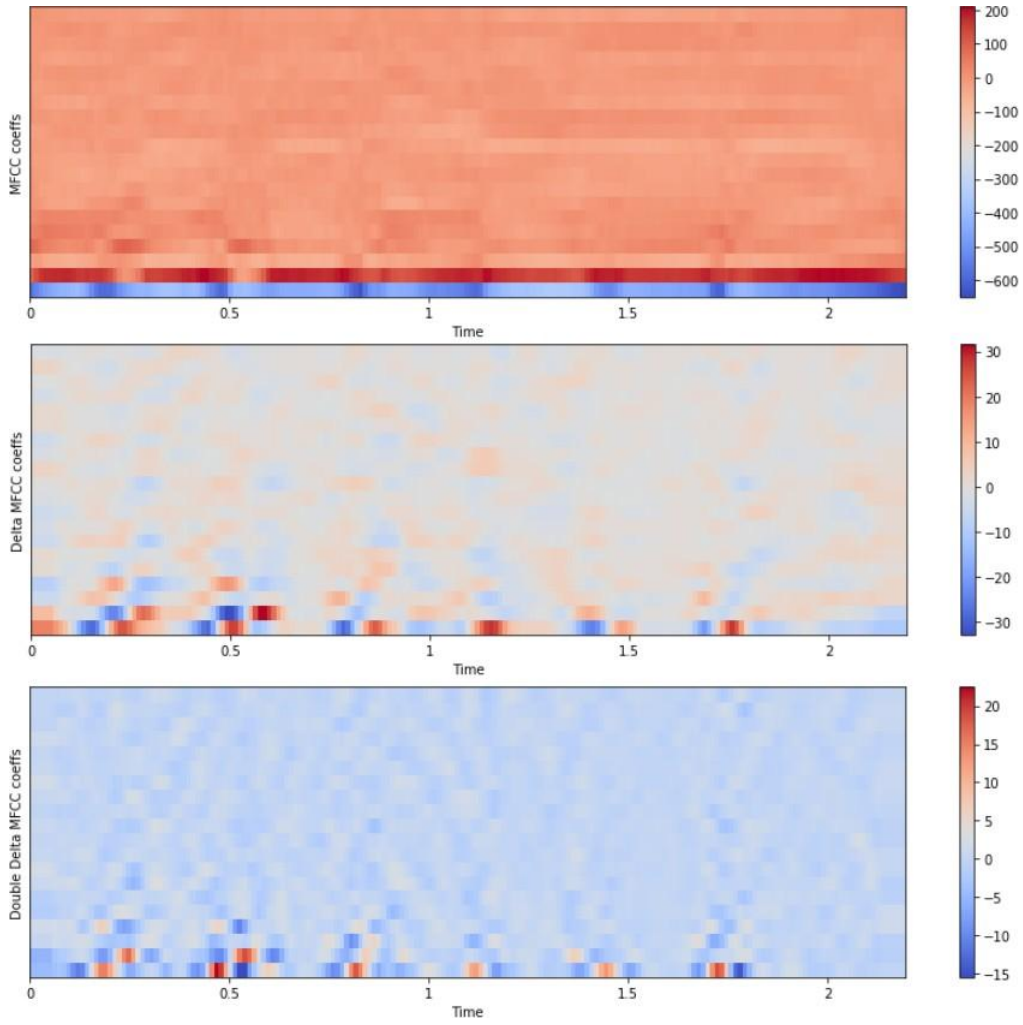
8. SNAPSHOTS

Voice Classification Using ML and Emotion Recognization
Compsoft Technology

Team : Isaac, Jahnavi ,Irfan , Hemanth



```
[+] Number of training samples: 504
[+] Number of testing samples: 168
[+] Number of features: 180
[*] Training the model...
Accuracy: 75.00%
```



CHAPTER 9

CONCLUTION

9. CONCLUSION

The emerging growth and development in the field of AI and machine learning have led to the new era of automation. Most of these automated devices work based on voice commands from the user. Many advantages can be built over the existing systems if besides recognizing the words, the machines could comprehend the emotion of the speaker (user). Some applications of a speech emotion detection system are computer-based tutorial applications, automated call center conversations, a diagnostic tool used for therapy and automatic translation system.

In this thesis, the steps of building a speech emotion detection system were discussed in detail and some experiments were carried out to understand the impact of each step. Initially, the limited number of publically available speech database made it challenging to implement a well-trained model. Next, several novel approaches to feature extraction had been proposed in the earlier works, and selecting the best approach included performing many experiments. Finally, the classifier selection involved learning about the strength and weakness of each classifying algorithm with respect to emotion recognition. At the end of the experimentation, it can be concluded that an integrated feature space will produce a better recognition rate when compared to a single feature.

For future advancements, the proposed project can be further modeled in terms of efficiency, accuracy, and usability. Additional to the emotions, the model can be extended to recognize feelings such as depression and mood changes. Such systems can be used by therapists to monitor the mood swings of the patients. A challenging product of creating machines with emotion is to incorporate a sarcasm detection system. Sarcasm detection is a more complex problem of emotion detection since sarcasm cannot be easily identified using only the words or tone of the speaker.

10. REFERENCE

[1] Soegaard, M. and Friis Dam, R. (2013). The Encyclopedia of Human-Computer Interaction. 2nd ed.

[2] Developer.amazon.com. (2018). Amazon Alexa. [online] Available at:https://developer.amazon.com/alexa

[3] Store.google.com. (2018). Google Home Tips & Tricks – Google Store. [online] Available at:https://store.google.com/product/google_home_learn

[4] Apple. (2018). iOS - Siri. [online] Available at: https://www.apple.com/ios/siri/

[5] The Official Samsung Galaxy Site. (2018). What is S Voice?. [online] Available at: http://www.samsung.com/global/galaxy/what-is/s-voice/ [Accessed 2 May 2018].

[6] Gartner.com. (2018). Gartner Says 8.4 Billion Connected. [online] Available at:https://www.gartner.com/newsroom/id/3598917.

[7] H. Cao, R. Verma, and A. Nenkova, "Speaker-sensitive emotion recognition via ranking: Studies onacted and spontaneous speech," Comput. Speech Lang., vol. 28, no. 1, pp. 186–202, Jan. 2015.

[8] L. Chen, X. Mao, Y. Xue, and L. L. Cheng, "Speech emotion recognition: Features and classificationmodels," Digit. Signal Process., vol. 22, no. 6, pp. 1154–1160, Dec. 2012.

[9] T. L. Nwe, S. W. Foo, and L. C. De Silva, "Speech emotion recognition using hidden Markovmodels," Speech Commun., vol. 41, no. 4, pp. 603–623, Nov. 2003.