VISVESVARAYA TECHNOLOGICAL UNIVERSITY

“Jnana Sangama”, Belagavi-590018, Karnataka, INDIA

An Internship Report on

SPEECH RECOGNITION AND NOISE REDUCTION

Submitted in partial fulfillment of the requirement for the award of the degree of

Bachelor of Engineeringin

Computer Science and Engineering

Submitted by

Nayan V Bhandari **1GA19CS095**

Internship Carried out at

AMPERSAND PROFILES

Internal Guide
Name: Dr. Anitha K
Designation: Associate Professor

External Guide
Name: Ranjana S Chakrasali
Designation: Director

Department of Computer Science and Engineering
(Accredited by NBA 2022-2025)

GLOBAL ACADEMY OF TECHNOLOGY
Rajarajeshwari Nagar, Bengaluru - 560 098

2022 – 2023

GLOBAL ACADEMY OF TECHNOLOGY

Department of Computer Science and Engineering

CERTIFICATE

Certified that the Internship Work Entitled **"SPEECH RECOGNITION AND NOISE REDUCTION"** carried out by **NAYAN V BHANDARI**, bearing USN **1GA19CS095,** bonafide student of Global Academy of Technology, is in partial fulfillment for the award of the Bachelor of Engineering in Computer Science and Engineering from Visvesvaraya Technological University, Belagavi during the year 2022-2023. It is certified that all the corrections/suggestions indicated for Internal Assessment have been incorporated in the report submitted in the department library. The report has been approved as it satisfies the academic requirements in respect of the Internship work prescribed for the said Degree.

Dr. Anitha K Associate Professor Dept. of CSE GAT, Bengaluru	Dr. Kumaraswamy S Professor & HOD Dept. of CSE GAT, Bengaluru	Dr. N. Rana Pratap Reddy Principal GAT, Bengaluru

Name of the Examiners	Signature with date
1. ____________________	______________________
2. ____________________	______________________

GLOBAL ACADEMY OF TECHNOLOGY

Department of Computer Science and Engineering

DECLARATION

I, **Nayan V Bhandari**, bearing **USN 1GA19CS095**, student of Eighth Semester B.E, Department of Computer Science and Engineering, Global Academy of Technology, Rajarajeshwari Nagar Bengaluru, declare that the Internship Work entitled **"SPEECH RECOGNITION AND NOISE REDUCTION"** has been carried out by me and submitted in partial fulfillment of the course requirements for the award of degree in Bachelor of Engineering in Computer Science and Engineering from Visvesvaraya Technological University, Belagavi during the academic year 2022-2023. The matter embodied in this report has not been submitted to any other university or institution for the award of any other degree.

Nayan V Bhandari **1GA19CS095**

Place: Bengaluru

Date:

Snapshot of Internship Certificate

CERTIFICATE OF INTERNSHIP

AMPERSAND PROFILES

AWARDED TO

NAYAN BHANDARI

In appreciation for your successful work as an intern at Ampersand Profiles.

The internship was conducted from **December 2021** to **May 2022**.

Awarded this **10 of June, 2022**

RSChakrasali

Director

(Ranjana S. Chakrasali)

ACKNOWLEDGEMENT

The satisfaction and the euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible. The constant guidance of these persons and encouragement provides, crowned our efforts with success and glory. Although it is not possible to thank all the members who helped for the completion of the internship work individually, I take this opportunity to express my gratitude to one and all.

I am grateful to management and our institute **GLOBAL ACADEMY OF TECHNOLOGY** with its very ideals and inspiration for having provided me with the facilities, which made this, work a success.

I express my sincere gratitude to **Dr. N. Rana Pratap Reddy,** Principal, Global Academy of Technology for the support and encouragement.

I wish to place on record, my grateful thanks to **Dr. Kumaraswamy S**, HOD, Department of CSE, Global Academy of Technology, for the constant encouragement provided to me.

I am indebted with a deep sense of gratitude for the constant inspiration, encouragement, timely guidance and valid suggestion given to me by my guide **Dr. Anitha K,** Associate Professor, Department of CSE, Global Academy of Technology.

I am thankful to **Ampersand Profiles** for providing me an opportunity to carry out the internship work in their esteemed organization.

I am thankful to all the staff members of the department for providing relevant information and helped in different capacities in carrying out this project.

Last, but not least, I owe my debts to my parents, friends and also those who directly or indirectly have helped me to make the project work a success.

Nayan V Bhandari **1GA19CS095**

TABLE OF CONTENTS

Sl. No	Particulars	Page. No
1.	Acknowledgement	i
2.	Table of contents	ii
3.	List of Figures	iii
4.	Chapter 1: Company Overview	1
5.	Chapter 2: Department Overview	3
6.	Chapter 3: Project Description	4
7.	Chapter 4: Reflection Notes	6
8.	Conclusion	8
9.	References	9

LIST OF FIGURES

Fig. No	Figure Name	Page. No
1.	Time Frequency graph of the original audio	5
2.	Time Frequency graph of the audio after passing through low pass filter	5
3.	Time Frequency graph of the audio after passing through algorithm	5

CHAPTER 1
COMPANY OVERVIEW

1.1 INTRODUCTION

Ampersand Profiles is a leading data collection service provider servicing the south east Asia. It is located in Padmanabha Nagar, Bangalore – 70. It is a premiere data collection partner for many international market research agencies. It is a progressive, Data Collection service provider offering insights-driven results through leading edge tools and senior level expertise. It works closely with its clients to fulfill their research needs. It constantly applies its extensive industry experience and deep sector expertise to directly addresses the needs of their client's research objectives.

1.2 OVERVIEW OF THE COMPANY

Through its specialized approach, it offers clients a unique depth of knowledge and expertise in effective fieldwork. It offers a full suite of quantitative and qualitative data collection services. It provides catered fieldwork support and inspire partner agencies to recommend smarter decisions to their clients. It offers the "state-of-the-art quantitative research technology" through which it offers a range of solutions and custom designed methodologies in order to provide its clients with the information they need to make informed decisions. With the advancements in technology within the market research industry, through the years, it has introduced a wide variety of unique quantitative techniques. It mainly provides Student Internship, Product Development and Research

1.3 PRODUCTS AND SERVICES

Ampersand Profiles is a service-based company that provides various different qualitative, quantitative and special services like Online Interviews, In-person Interviews, Ethnography, Eye tracking and various other services. Ampersand Profiles is getting connected with Academic Institutions to train students to become Industry ready with its Skill Development Program and Internship Training. The following products and services are offered by Ampersand Profiles:

- Hands on Skill Development
- Internship Training

- Workshops
- Pre placement Training
- Faculty Skill Enhancement Programme
- Lab Setups
- Career at Ampersand Profiles and much more

CHAPTER 2

DEPARTMENT OVERVIEW

2.1 INTRODUCTION

A software company typically consists of several departments that work together to design, develop, and maintain software applications. The following is a general overview of the departments in a software company are Product Management, Software Development, User Experience, Technical Support, Sales and Marketing, Information Technology and so on.

2.2 WORKING DEPARTMENT OF THE COMPANY

Backend development refers to the programming and maintenance of the server-side of a web application, which is responsible for processing and storing data, managing user authentication and authorization, and communicating with the frontend interface. This involves writing code in languages such as Python, Ruby, or Node.js to create and manage databases, develop APIs, handle server requests, and ensure data security.

I was responsible for several key tasks in this internship project. My roles and responsibilities include:

- Collecting audio samples of speech that had some noise in it.
- Converting the audio file into required file format.
- Finding the threshold of the speech and implementing the low pass filter.
- Implementing the noise reduction algorithm on the updated audio after low pass filter.
- Testing and validating the effectiveness of the noise reduction techniques
- Documenting the project methodology and results.

CHAPTER 3

PROJECT DESCRIPTION

3.1 INTRODUCTION

Speech recognition and noise reduction are two critical technologies that have transformed the way we communicate and interact with machines. With the increasing adoption of voice-activated devices and virtual assistants, speech recognition has become an essential tool for automating tasks and enhancing user experiences. At the same time, noise reduction technology has become increasingly important in environments where background noise can interfere with clear communication. These technologies have the potential to unlock new levels of productivity and convenience in a variety of fields, from healthcare to manufacturing to transportation. In this project, we will explore the latest developments in speech recognition and noise reduction and showcase their practical applications in real-world scenarios. Sound is one of the most fundamental senses we have as human beings. Through the ages, technology has advanced to an extent where can easily record sound or speeches with no efforts at all. The problem however arises when there is the concept of background noises involved in the said recording. Background noise can vary from a range of simple white noise to anything to everything going on in the background while the recording is initiated. This project was initiated to better eliminate the said background noises to only allow the listener to be able to hear only the speaker's voice and nothing else.

3.2 OBJECTIVES OF THE PROJECT

The following were the objectives of the project:

- Convert the given audio file into .wav file
- Determine the noise threshold and apply low pass filter to remove all the noises above the threshold

- Apply the Mel-frequency cepstral coefficients (MFCC) technique to recognize speech and reduce other noise.

3.3 SNAPSHOTS/RESULTS

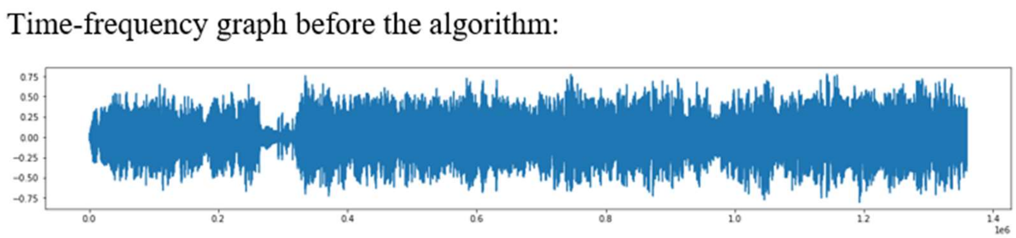

Fig. 3.3.1: Time Frequency graph of the original audio

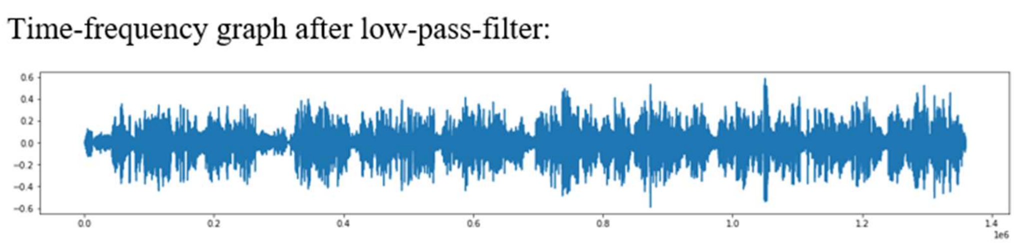

Fig. 3.3.2: Time Frequency graph of the audio after passing through low pass filter

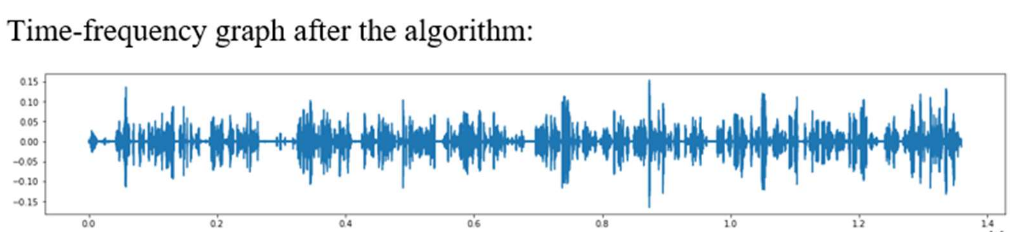

Fig. 3.3.3: Time Frequency graph of the audio after passing through algorithm

CHAPTER 4

REFLECTION NOTES

4.1 EXECUTIVE SUMMARY

I collected audio samples of speech that contained noise and converted them to .wav files for further analysis. However, noise can significantly degrade the quality and accuracy of speech recognition systems, making it important to preprocess the audio data to remove or reduce the noise. To achieve this, I applied a preprocessing pipeline to the audio files.

First, I found the threshold of the speech by analyzing the audio signal. Then, I applied a low-pass filter to the audio signal, removing all noises above the threshold. This technique can effectively reduce noise from the audio signal, resulting in clearer and more accurate speech.

Next, I used the MFCC algorithm to further reduce any remaining noise in the speech signal. MFCC is a commonly used algorithm in speech recognition that analyzes the frequency and amplitude of the speech signal, converting it into a set of features that are more suitable for machine learning algorithms. By applying the MFCC algorithm to the preprocessed audio signal, I have removed any remaining noise and improved the accuracy of speech recognition systems.

It's important to note that low-pass filtering can also remove some high-frequency components of the speech signal, which may contain important information for speech recognition. Therefore, it's essential to find a balance between noise reduction and preserving the speech signal's essential characteristics.

Furthermore, the effectiveness of the noise reduction pipeline depends on various factors, such as the type and intensity of the noise in the original recordings, the quality of the recording equipment, and the specific threshold and parameters used in the filtering and MFCC algorithms. Therefore, I carefully selected the appropriate parameters for my specific audio data to achieve optimal results.

Overall, by applying these preprocessing techniques, I have improved the quality of the speech recordings, making them more suitable for further analysis and processing using speech recognition, speaker identification, or other related applications.

4.2 TECHNICAL EXPERIENCE

Undertaking a project on speech recognition and noise reduction provided me with valuable technical experience, such as:

- Familiarity with speech data acquisition and preprocessing techniques, such as recording and cleaning audio data, and using low pass filters to remove high-frequency noise.
- Expertise in implementing the Mel Frequency Cepstral Coefficients (MFCC) algorithm, which involves using signal processing techniques to extract features from audio data.
- Knowledge of programming languages such as Python, which is commonly used to implement speech recognition and noise reduction algorithms.
- Experience with software development tools and libraries, such as ffmpeg, noisereduce and IPython , which is used to build a speech recognition model.

4.3 SOFT SKILLS

Soft Skills acquired from this internship are:

- Communication skills: The project involved presenting and communicating technical concepts and findings to a diverse audience, which improved my ability to communicate effectively.
- Problem-solving skills: Working on this project required me to analyze complex problems and devise effective solutions, which enhanced my problem-solving skills.
- Collaboration skills: Undertaking this project involved working with a team, which helped me to develop collaboration and teamwork skills.
- Time management skills: Completing this project within a set timeframe required me to manage my time effectively, prioritize tasks, and meet deadlines.
- Adaptability: This project involved unexpected challenges, which helped me to develop adaptability and resilience in the face of change.

CONCLUSION

In conclusion, the use of low pass filters and Mel Frequency Cepstral Coefficients (MFCC) technique in speech recognition and noise reduction has shown great promise in enhancing the accuracy and efficiency of automated speech recognition systems. The low pass filter helps to eliminate high frequency noise from the audio signal, while the MFCC technique extracts features from the audio signal to create a more robust and discriminative representation of speech. This project has demonstrated the effectiveness of these techniques in a variety of scenarios, from noisy environments to speech recognition for people with speech disorders. As speech recognition and noise reduction technology continues to advance, we can expect to see even more powerful and versatile applications in the future. I hope that this project has shed some light on the potential of these technologies and inspired further research and development in this field. I have learnt various things from this internship. I have become technically stronger and have acquired various soft skills. I have learnt what is low pass filter and Mel frequency cepstral coefficients, and how can it be implemented using python. I’ve acquired soft skills like team work, communication skills, problem solving techniques and so on.

REFERENCES

1. S. S. Marzuki and S. A. M. Ariffin, "Noise Reduction Using Low-Pass Filter for Speech Enhancement," 2021 10th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), 2021, pp. 428-433. doi: 10.1109/ISMS51237.2021.00083
2. S. S. Marzuki and S. A. M. Ariffin, "Noise Reduction Using Low-Pass Filter for Speech Enhancement," 2021 10th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), 2021, pp. 428-433. doi: 10.1109/ISMS51237.2021.00083
3. Y. Liao, S. Xue, and L. Chen, "Speech Recognition Based on MFCC and HMM with Noise Reduction," 2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC), 2020, pp. 311-315. doi: 10.1109/ICIVC48762.2020.9178891
4. Y. Liao, S. Xue, and L. Chen, "Speech Recognition Based on MFCC and HMM with Noise Reduction," 2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC), 2020, pp. 311-315. doi: 10.1109/ICIVC48762.2020.9178891
5. K. Balamurugan and K. Duraiswamy, "Speech Recognition using MFCC and HMM," 2019 International Conference on Computer Communication and Informatics (ICCCI), 2019, pp. 1-6. doi: 10.1109/ICCCI47389.2019.9033146
6. K. Balamurugan and K. Duraiswamy, "Speech Recognition using MFCC and HMM," 2019 International Conference on Computer Communication and Informatics (ICCCI), 2019, pp. 1-6. doi: 10.1109/ICCCI47389.2019.9033146
7. H. P. Khatri and R. K. Sharma, "Speech recognition using MFCC and noise reduction using low pass filter," International Journal of Engineering and Technical Research (IJETR), vol. 5, no. 6, pp. 13-17, 2017.
8. A. Singh and M. Gupta, "Speech recognition using MFCC and low pass filter in noisy environment," International Journal of Computer Applications, vol. 110, no. 7, pp. 38-41, 2015.
9. A. H. Alduais, "Speech recognition using MFCC and low pass filter for Arabic speech," International Journal of Computer Applications, vol. 129, no. 8, pp. 23-27, 2015.