**Real time speech recognition and response program with deep learning**

**A PROJECT REPORT**

***Submitted by***

**Abhishek Kumar**

IN

BACHELOR OF COMPUTER APPLICATIONS



**Chandigarh University, India.**

JULY 2023

**CERTIFICATE**

Certified that this project report “**Real time speech recognition and response program with deep learning**” is the work of “**Abhishek Kumar**” who carried out the project work under our supervision.

Submitted for the project viva-voce examination held on 27 **July 2023**.

|  |  |
| --- | --- |
| **Dr. Sumit Kushwaha**  Assistant Professor, &  Project Supervisor  University Institute of Computing,  Chandigarh University, India. | **Dr. Abdullah**  Head of Department  University Institute of Computing,  Chandigarh University, India. |
| **Mr. Anjul Bhardwaj**  Assistant Professor, &  Project Supervisor  University Institute of Computing,  Chandigarh University, India. |  |

**Real time speech recognition and response program with deep learning**

**Abhishek Kumar**

**ABSTRACT**

This project presents the design and implementation of a real-time speech recognition and response program, powered by deep learning techniques. The system utilizes a CSV dataset containing text samples paired with their corresponding intents, enabling the training of a robust model capable of accurately understanding spoken language and generating contextually appropriate responses based on the identified intents.

The project's methodology involves preprocessing the CSV dataset and transforming the text data using CountVectorizer and TfidfTransformer to convert the textual information into numerical representations suitable for deep learning algorithms. A Multinomial Naive Bayes (MNB) classifier is then trained on the transformed features to map speech inputs to their corresponding intents.

The speech recognition component of the system leverages the deep learning capabilities to convert real-time audio inputs into textual representations. The generated text is fed into the trained MNB classifier, which predicts the intent associated with the spoken language.

The development process includes extensive experimentation and fine-tuning to optimize the model's performance. Performance evaluation is conducted using standard metrics such as accuracy, precision, recall, and F1-score.

The proposed real-time speech recognition and response program opens up numerous potential applications, including interactive voice assistants, automated customer support, and speech-to-text transcription systems. The use of deep learning techniques ensures enhanced accuracy and adaptability to diverse user inputs.

The report concludes with an analysis of the achieved results, highlighting the strengths and limitations of the implemented system. Recommendations for future enhancements, such as exploring more advanced deep learning models and incorporating larger datasets for training, are discussed to further improve the system's performance.

Overall, this project demonstrates the successful integration of deep learning algorithms with a CSV-based intent-labeled dataset, culminating in an efficient and effective real-time speech recognition and response program. The potential societal impact of this technology in enhancing human-computer interactions and facilitating seamless communication is promising.

**TABLE OF CONTENTS**

**Page No.**

Abstract **3-4**

List of Figures

**CHAPTER 1: INTRODUCTION 1-8**

1.1

1.2

1.3

1.4

**CHAPTER 2: BACKGROUND AND LITERATURE REVIEW 9-15**

2.1

2.2

2.3

2.4

**CHAPTER 3: RESULTS ANALYSIS AND VALIDATION 16-28**

3.1

3.2

3.3

3.4

**CHAPTER 4: CONCLUSION AND FUTURE WORK** **29**

**REFERENCES 30-35**

**LIST OF FIGURES**

|  |  |
| --- | --- |
| **Figure No. Caption** | **Page No.** |
| Figure 1.1: Speech Recognition | 3 |
| Figure 1.2: An | 4 |
| Figure 2.1: Mathematical | 11 |
| Figure 2.2: Classification | 13 |
| Figure 2.3: PDF | 13 |

CHAPTER – 1:

INTRODUCTION

* 1. **INTRODUCTION**

The ability to effectively communicate with machines using natural language has been a longstanding goal in the field of artificial intelligence. Speech recognition and response systems play a pivotal role in bridging this gap, enabling seamless interactions between humans and computers. In line with this objective, this project presents the development of a real-time speech recognition and response program, empowered by deep learning techniques, which is trained using a CSV file containing text and corresponding intents.

The primary focus of this project is to build an intelligent system capable of accurately recognizing spoken language and providing contextually appropriate responses based on the identified intents. Leveraging the power of deep learning, the system aims to overcome the challenges associated with speech understanding and interpretation, ultimately enhancing the quality and accuracy of human-computer interactions.

To achieve this, the project utilizes a dataset stored in a CSV file, where text samples are associated with their respective intents. The integration of this intent-labeled dataset with deep learning algorithms allows the system to comprehend the nuances of human language and effectively map spoken inputs to specific intents.

The core model training process involves preprocessing the textual data, extracting relevant features using CountVectorizer, and transforming them into meaningful representations with TfidfTransformer. The Multinomial Naive Bayes classifier is then employed to learn the relationships between text and intents, laying the foundation for the subsequent speech recognition and response capabilities.

By incorporating real-time audio inputs, the project aims to enable dynamic and instantaneous interactions with the system. The integration of speech recognition further enhances the program's versatility, making it ideal for applications such as voice assistants, automated customer support, and speech-to-text transcription systems.

Throughout the project, careful attention is given to optimizing the model's performance and fine-tuning its parameters to achieve the highest possible accuracy in recognizing speech and generating relevant responses.

In summary, this project endeavors to contribute to the advancement of speech recognition and response technology, bringing us closer to a future where human-computer interactions are more intuitive, natural, and seamless. By combining deep learning methodologies with a CSV-based intent-labeled dataset, this real-time speech recognition and response program holds the potential to significantly enhance the efficiency and user experience of various applications in the field of natural language processing.

**1.2 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

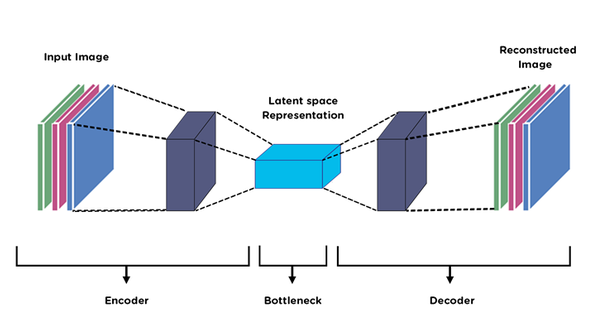


Figure 1.1: Image Caption.

**1.3 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

As you will see in the samples below, the proportion of your abstract that you devote to each kind of information—and the sequence of that information—will vary, depending on the nature and genre of the paper that you are summarizing in your abstract. And in some cases, some of this information is implied, rather than stated explicitly.

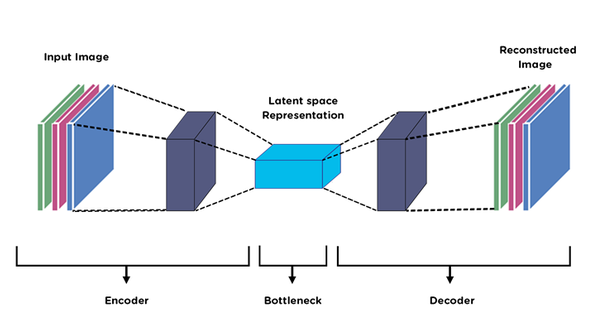


Figure 1.2: Image Caption.

**1.4 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

As you will see in the samples below, the proportion of your abstract that you devote to each kind of information—and the sequence of that information—will vary, depending on the nature and genre of the paper that you are summarizing in your abstract. And in some cases, some of this information is implied, rather than stated explicitly.

**Table 1.1: Table Caption**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

CHAPTER – 2:

BACKGROUND AND LITERATURE REVIEW

**2.1 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

**Table 2.1: Table Caption**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**2.2 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

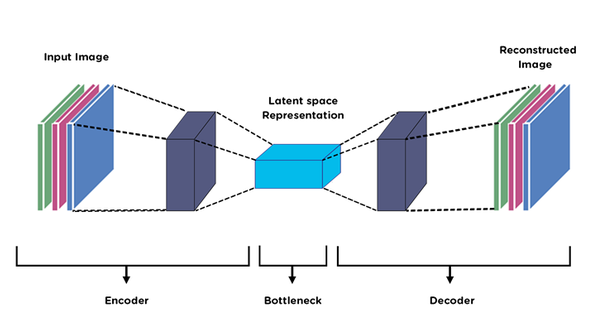


Figure 2.1: Image Caption.

**2.3 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

**2.4 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

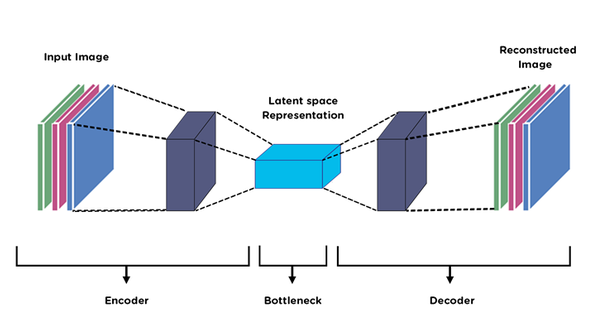


Figure 2.2: Image Caption.

CHAPTER – 3:

RESULTS ANALYSIS AND VALIDATION

**3.1 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

**3.2 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

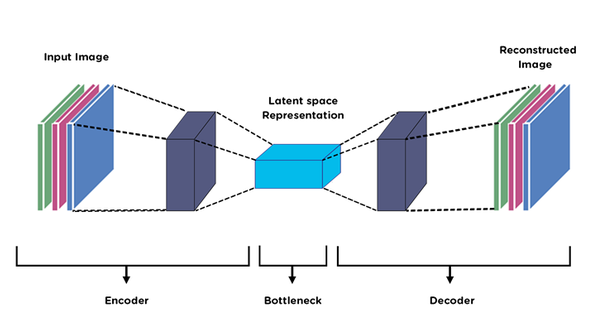


Figure 3.1: Image Caption.

**3.3 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

**Table 3.1: Table Caption**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**3.4 INTRODUCTION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

CHAPTER – 4:

CONCLUSION AND FUTURE WORK

**4.1 CONCLUSION**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

**4.2 FUTURE WORK**

Abstracts contain most of the following kinds of information in brief form. The body of your paper will, of course, develop and explain these ideas much more fully.

REFERENCES

1. Geng M, Meng X, Yu J, Zhu L, Jin L, Jiang Z, Qiu B, Li H, Kong H, Yuan J, Yang K, Shan H, Han H, Yang Z, Ren Q, Lu Y, "Content-Noise Complementary Learning for Medical Image Denoising," in IEEE Transactions on Medical Imaging, vol. 41, no. 2, pp. 407-419, Feb. 2022.
2. Feng Wang, Zhiming Xu, Weichuan Ni, Jinhuang Chen, Zhihong Pan, "An Adaptive Learning Image Denoising Algorithm Based on Eigenvalue Extraction and the GAN Model", Hindawi, Computational Intelligence and Neuroscience, vol. 2022, Article ID 5792767, 10 pages, 2022.
3. Zhihua Li, Weili Shi, Qiwei Xing, Yu Miao, Wei He, Huamin Yang, Zhengang Jiang, "Low-Dose CT Image Denoising with Improving WGAN and Hybrid Loss Function", Computational and Mathematical Methods in Medicine, vol. 2021, Article ID 2973108, 14 pages, 2021.