Speech Recognition using Python

**[1]Velcia Mendonca, [1]Sanjana M. Rao, [1]Anjana R., [1]Teena S.,
[2] Dr. Pavithra G., [3] Dr. Sindhu Sree M., [4]Dr. T.C.Manjunath*,
[5]Rajashekher Koyyeda, [6]Aditya T.G.**

*[1]First Semester BE (ECE) Students, Dept. of Electronics & Communication Engg.,
Dayananda Sagar College of Engineering, Bangalore, Karnataka
[2]Associate Prof., Electronics & Communication Engg. Dept.,
Dayananda Sagar College of Engineering, Bangalore, Karnataka
[3]Assistant Prof., Electronics & Communication Engg. Dept.,
Dayananda Sagar College of Engineering, Bangalore, Karnataka
[4] Professor & HOD, Electronics & Communication Engg. Dept.,
Dayananda Sagar College of Engineering, Bangalore, Karnataka
[5]Asst. Prof., EEE Dept., Tatyasaheb Kore Inst. of Engg. & Tech., Warananagar, Kolhapur
[6] Fifth Sem Student, CSE Dept., PES University, Bangalore*

Abstract

In this paper, the Speech Recognition using Python is presented. Speech recognition involves the conversion of spoken words into text and is utilized in various applications such as virtual assistants and speech-to-text dictation. Python is a favored programming language for developing speech recognition systems due to its flexibility, simplicity, and availability of libraries and tools. Various packages and libraries such as Speech Recognition, Py Audio, and Pocket Sphinx are used for implementing speech recognition in Python. Recording audio and converting it into text using a speech recognition engine are the basic steps involved in building a speech recognition system using Python. Natural language processing (NLP) can then be employed to further process the text. The ease-of-use and wide-ranging applications make speech recognition using Python an exciting field. The work done & presented in this paper is the result of the mini-project work that has been done by the first sem engineering students of the college and as such there is little novelty in it and the references are being taken from various sources from the internet, the paper is being written by the students to test their writing skills in the starting of their engineering career and also to test the presentation skills during their mini-project presentation. The work done & presented in this paper is the report of the assignment / alternate assessment tool as a part and parcel of the academic assignment of the first year subject on nanotechnology & IoT.

Keywords: Speech, Python, Recognition.
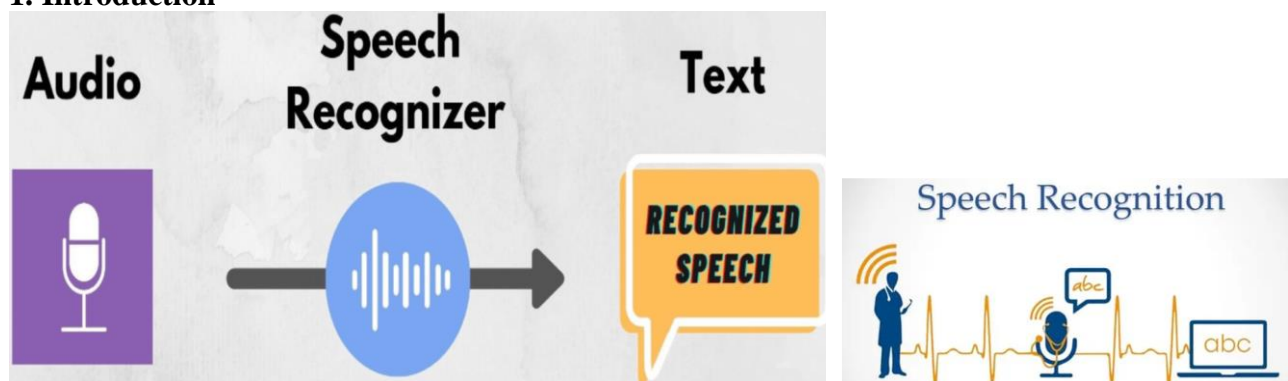
1. Introduction

Fig. 1 : Block diagram

```
ef main():

    path = "C:/Program Files (x86)/Google/Chrome/Application/chrome.exe %s"


    r = sr.Recognizer()
    with sr.Microphone() as source:
        r.adjust_for_ambient_noise(source)

        print("Please say something ")

        audio = r.listen(source)

        print("Reconizing Now ... ")



        try:
            dest = r.recognize_google(audio)
            print("You have said : " + dest)

            web.get(path).open(dest)

        except Exception as e:
            print("Error : " + str(e))
```

Fig. 2 : Python scripting

2. Background review

Speech recognition, the technology that enables machines to understand and interpret human speech, has gained tremendous popularity and practicality in recent years. From voice assistants to transcription services, speech recognition has found applications in various fields, including communication, automation, and accessibility. Python, a powerful and versatile programming language, provides a range of tools and libraries that make speech recognition implementation accessible and efficient. In this three-page introduction, we will explore the concept, benefits, and key components of speech recognition using Python, highlighting its potential to revolutionize human-machine interaction and enable innovative applications across industries.

3. Concept and Functionality

Speech recognition, also known as automatic speech recognition (ASR) or voice recognition, refers to the technology that converts spoken language into written text or other machine-readable formats. It involves the analysis and processing of audio signals to extract meaningful linguistic information. Python, with its rich ecosystem of libraries and tools, offers a comprehensive platform for developing speech recognition systems. The functionality of speech recognition using Python typically involves several key steps. First, the audio input, such as spoken words or phrases, is recorded or obtained from a source. Next, the recorded audio is preprocessed to enhance its quality and remove background noise or artifacts. Python libraries like PyAudio or Soundfile are commonly used for audio input and preprocessing.

4. Audio streams

Once the audio is ready, Python-based speech recognition libraries, such as Speech Recognition or Google Cloud Speech-to-Text API, are utilized to transcribe the speech into text. These libraries

leverage advanced machine learning algorithms and models trained on vast amounts of speech data to accurately recognize and convert spoken words into written form.

After the speech recognition process, the resulting text can be further processed, analyzed, or used to trigger specific actions or responses in applications. Python's extensive natural language processing (NLP) capabilities, along with libraries like NLTK or spaCy, allow for advanced text processing and language understanding.

5. Benefits

Speech recognition using Python offers several notable benefits that contribute to its widespread adoption and success. First and foremost, speech recognition technology enhances human-machine interaction by enabling voice-based commands and interactions. This hands-free and intuitive approach improves user experience, making applications more accessible and efficient. Voice-enabled interfaces, such as voice assistants or voice-controlled automation systems, provide convenience and ease of use, particularly in scenarios where manual input or interaction may not be practical or feasible.

Moreover, speech recognition has the potential to improve accessibility for individuals with disabilities. It allows people with mobility or dexterity limitations to interact with technology and perform tasks through voice commands. By leveraging Python's capabilities, developers can create inclusive applications that cater to diverse user needs. Furthermore, speech recognition in Python facilitates automation and efficiency in various domains. Voice-controlled systems can streamline tasks in industries like healthcare, customer service, or logistics, where hands-free operation and quick information retrieval are crucial. Automated transcription services powered by speech recognition offer significant time savings and accuracy improvements in fields like journalism, legal documentation, and content creation.

6. Key Components

Implementing speech recognition using Python involves leveraging various key components and libraries. These include:

a) Audio Input and Preprocessing: Python libraries like PyAudio or Soundfile assist in capturing audio input from different sources, such as microphones or audio files, and applying preprocessing techniques for noise reduction or signal enhancement.

b) Speech Recognition Libraries: Python provides powerful speech recognition libraries such as Speech Recognition, Google Cloud Speech-to-Text API, or CMUSphinx, which offer robust speech-to-text conversion capabilities. These libraries leverage machine learning algorithms and models for accurate recognition of spoken words.

c) Natural Language Processing (NLP) Libraries: Python's rich NLP ecosystem, including libraries like NLTK, spaCy, or TextBlob, enables developers to process and analyze the transcribed text further. This includes tasks such as language understanding, sentiment analysis, or entity recognition.

d) Integration and Application Development: Python's flexibility allows for seamless integration of speech recognition capabilities into existing applications or the development of standalone voice-controlled systems. Integration with other libraries and frameworks, such as GUI frameworks like Tkinter or web frameworks like Flask, enables the creation of user-friendly interfaces and applications.

7. Conclusions

Speech recognition using Python has revolutionized the way machines understand and interpret human speech, opening up opportunities for improved human-machine interaction, automation, and accessibility. Python's extensive ecosystem of libraries and tools empowers developers to implement speech recognition systems efficiently and effectively. With its ability to convert spoken language into text or actionable commands, speech recognition using Python has transformed industries and

applications across sectors, ranging from voice assistants to transcription services. As technology continues to evolve, speech recognition using Python holds immense potential to further enhance communication, productivity, and user experiences in diverse fields and scenarios.

References

[1]Dr. T.C.Manjunath, Pavithra G., R. Jagadisha, "Design & development of an efficient path planning mechanism for a Micro-Robot", IEAE's International Journal of Emerging Research & Management (IJERMT), ISSN : 2278-9359, DOI : 10.23956/ijermt, Vol. 6, Issue 5, pp. 727-733, Impact Factor: 3.969, IF by ISRA 1.492, UGC Approved Journal, May 2017.

[2]Pavithra G., Pratheeksha Shetty M., Dr. T.C.Manjunath, "A brief review of wireless LANs for how to prevent the intrusions in the networking protocols", Int. Jr. of Adv. Res. in Sci. & Engg. (IJARSE), ISBN : 978-93-86171-60-3, ISSN(O) : 2319-8354, ISSN (P) : 2319-83446, IF : 2.83, Sl. No. 147, Vol. 6, Issue 8, pp. 1282-1284, Aug. 2017.

[3]Dr. T.C.Manjunath, Pavithra G., "Design of algorithms for diagnosis of primary glaucoma through estimation of CDR in different types of fundus images using IP techniques", Int. Journal of Innovative Research in Information Security (IJIRIS), Vol. 4, Special Issue 5, Impact Factor Value IJIRIS: Impact Factor Value – SJIF: Innospace, Morocco (2016): 4.346 ISRAJIF (2016): 3.318 & Indexcopernicus ICV (2015):73.48, Innospace : 4.615 (2016) ISRAJIF : 3.318 (2016), Index Copernicus 2015 (73.48), Indexed in Google Scholar, AMP Publications, India ISSN : 2349-7009 (Print), ISSN : 2349-7017 (On-line), Paper id MYISSP 10135, pp. 12-19, May 2017.

[4]Dr. T.C.Manjunath, Pavithra G., Rashmi J., "Design of a controller for a dynamical system by developing a transfer function", International Journal of Engineering Research-Online (IJOER), A Peer Reviewed International Journal, KY Publications, ISSN : 2321-7758, Index Copernicus Value ICV:72.81, Vol. 5, Issue S1, pp. 464-468, Jun. 2017.

[5]Dr. T.C.Manjunath, Pavithra G., R. Jagadisha, "A brief review of the advances in molecular nano technology for the bio-medical engineering applications", International Journal of Engineering Research-Online (IJOER), A Peer Reviewed International Journal, KY Publications, India, ISSN : 2321-7758, Index Copernicus Value ICV:72.81 (2015), Impact Factor 5.8701, Vol. 5, Issue S1, pp. 424-429, June 2017.

[6]Shashikanth Patil, Dr. T.C.Manjunath, Pavithra G., "Guidance Control and Tracking of a Missile", IFERP's International Journal of Engineering Research in Mechanical and Civil Engineering (IJERMCE), DOI : 01.1617/vol4/iss5/pid93174, ISSN (Online) 2456-1290, ISSN (Online) 2456-1290, indexed by Google Scholar, Thomson Reuters ID, ORCiD Research ID, Impact Factor 3.8, Vol. 2, Issue 5, pp. 967-970, May 2017.

[7]Dr. T.C.Manjunath, Pavithra G., "Design & Development of Algorithms to Diagnose Secondary Diabetic Retinopathy Glaucoma through the Analysis of Fundus Images Using IP Techniques", International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE), ISSN (Online): 2320-9801, ISSN (Print): 2320-9798, An ISO 3297: 2007 Certified Organization, ISSN (Online): 2320-9801, ISSN (Print): 2320-9798, Impact Factor 7.194, ISO 3297: 2007, Vol. 5, Special Issue 6, pp. 40-43, July 2017.

[8]Dr. K.N. Vijaykumar, Pavithra G., Dr. T.C.Manjunath, "Microcontroller based control of devices using a sophisticated control system", Journal of Applied Engg. & Technologies (JAET), ISSN 2278-1722, Vol. 5, Issue 1, pp. 98-100, April 2017.

[9]Dr. Suhasini V., Pavithra G., Dr. T.C.Manjunath, "Novel method of noise elimination of digital signals", Journal of Applied Engg. & Technologies (JAET), ISSN 2278-1722, Vol. 5, Issue 1, pp. 169-170, April 2017.

[10] Ankita V., Pavithra G., Dr. T.C.Manjunath, "Design, analysis, implementation of a force base sensor for aerospace engineering", Journal of Applied Engg. & Technologies (JAET), ISSN 2278-1722, Vol. 5, Issue 1, pp. 101-103, April 2017.