_____

Speech Recognition System by Using Python

_____
Miss. Sayali Kiran Mande (MCS Pursuing)
RNC Arts, JDB Commerce & NSC Science College, Nashik Road.
Asst.Prof. **Ashwini P**. **Bawdhankar.**
Computer Science Department
RNC Arts, JDB Commerce & NSC Science College, Nashik Road.
_____

Abstract: The importance of Speech recognition system is very high in everyday human life as this software allows us to feed audio to the mobile devices to interact with them. This software converts the audio into various sound wave forms and inspects every form with the help of numerous algorithms. After analysing each sound form, it finds the most accurate word in that particular language and convert it into the text. As it works on the speech, it is very helpful for especially abled people with blindness or no forelimbs and it also important in healthcare, gaming, learning languages. This paper compiles the various studies of the popular existing system namely SIRI, GOOGLE ASSISTANT, BIXBY, ALEXA, CORTANA. This paper also analyses the concept of NLP (Natural processing) with speech recognition.

Introduction: Speech recognition lets people control electronic devices by talking instead of using keyboards or buttons. It changes spoken words into something the device can understand, so you can use voice commands to control it. The main idea is to make devices easy for anyone to use, even if they don't know much about technology. Speech recognition is used in areas like banking, marketing, healthcare, and learning languages. Important parts of speech processing include things like voice pitch, quality, loudness, and background noise.

_____

Working: working of speech recognition process [3] Speech Recognition System Fig .no:1

1. Feature Extraction:
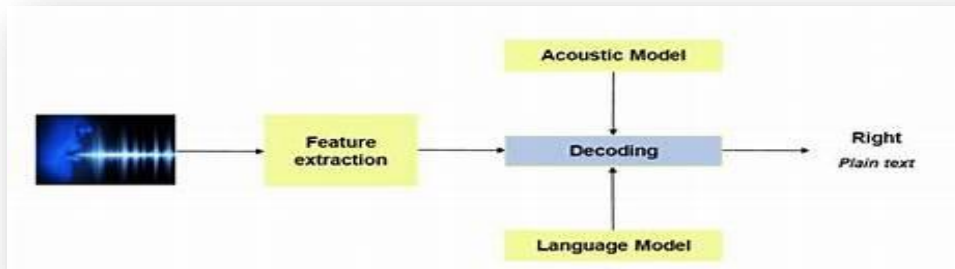
Speech recognition software analyses sound based on different factors like pitch, voice variation, and strength of the voice.

2. Acoustic Models:

An acoustic model is a file that contains different sounds that form words. It shows how audio signals relate to the basic units of speech. These representations are created using mathematical models, like HMMs (Hidden Markov Models).

3. Decoder:

A neural network breaks down speech into smaller units called neurons. Various algorithms are used to decode these neurons to ensure accurate results**.**

4. Language Model:

The language model helps distinguish between words and phrases that sound similar. By combining the information from the language model with pronunciation and acoustic models, the system can resolve confusion between similar-sounding words.

Existing system which uses speech recognition

- **APPLE (SIRI):** Siri is a virtual assistant is a part of Apple Inc. It is designed to offer you a multiple way of interaction with your phone by speak up. It will take the query through microphone and help you solve the query with in time. It has some features that make it different from other speech recognition, for instance, can activate low power mode, Enable do-not disturb mode as well it has non-English option. Most of them use it as entertaining purpose. On the other hand, Siri has one demerit point it only works on IOS devices. [3]

- **GOOGLE ASSISTANT:** Google Assistant is virtual assistant of google Inc's. Google assistant control your devices and smart phone. Some important features of

_____

Google assistant are, it controls your device and your smartphone and access information from calendar. As well as it can also handle your music system. It has some demerit points such as it uses more battery power due to this sometime it slows down the working of system**.[3]**

- **MICROSOFT CORTANA:** Like Siri and Google Assistant, Cortana is also a voice assistant developed and created by Microsoft. Basically, it is designed for window devices Nowadays, it is available in various devices. It can perform a multiple task for users, like remainder setting, as well as it can also scheduling the calendar events, even most of people who use Cortana to performing some computational data and many more Cortana has an API (application programming interface) and can work with a variety of windows app, as well as third-party apps such as Facebook and Twitter. Apart from that, it has several demerits such as Vulnerability found hit the listening button again and again and many more.[3]

- **AMAZON ALEXA:** Alexa is a virtual assistant technology designed and created by Amazon. This technology is based upon Machine learning, NLP (natural language processing). Alexa can perform various task such as it can acknowledge the user about Weather. Furthermore, it can handle your smart phone like when user give the instruction it will take and solve it with in time so that user can do another work. But it has several demerit points, for example, it cannot send a message and Email through voice command. Apart from that, if a person wants to access the heath and hobbies related data through Alexa, in that cases Alexa will not give the accurate result.[3]

- **SAMSUNG BIXIBY:** Bixby is a virtual assistant developed by Samsung Electronics. With the help of this system, you can send message from one device to another, as well as you can check the cricket and any other game score. A hardware button on the side of the device to bring up Bixby. Bixby also supports some fancy features. But Bixby has some limitation, for illustrating, as it supports some limited language due to this, sometimes it fails to provide the direct answer of query [3]
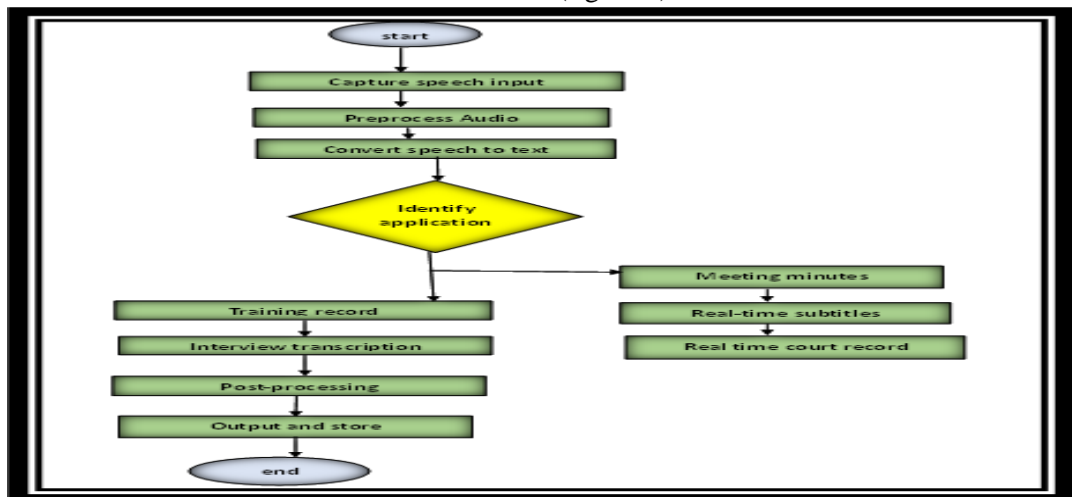
USE CASE SPEECH RECOGNITION: [1]

Automatic speech recognition system(fig no:2)

_____
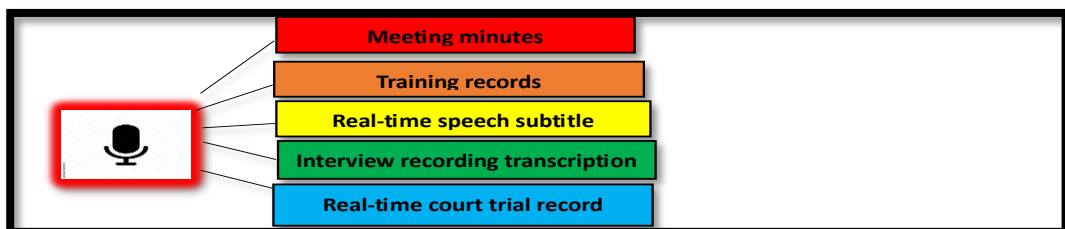
UTILIZATION OF SPEECH RECOGNITION IN INDUSTRY:
FLOW CHART OF THIS:

Workflow (fig no:3)

1. Virtual Assistants:

Virtual assistants like Amazon Alexa, Google Assistant, and Apple Siri use speech recognition to understand and follow spoken commands. Users can ask questions, set reminders, control smart home devices, and complete tasks by speaking to the assistant.

2. Accessibility:

Speech recognition helps people with disabilities, such as those with mobility or vision issues, use computers, mobile devices, and smart home devices by speaking instead of typing or pressing buttons.

3. Customer Service and Support:

Many companies use speech recognition in customer service. Interactive Voice Response (IVR) systems can listen to what customers say and help them by providing information or directing them to the right department.

4. Transcription Services:

Speech recognition is used to turn spoken words in audio or video recordings into written text. This is helpful for making interviews, meetings, and podcasts easy to search and read.

5. Dictation Software:

Dictation software lets users speak instead of typing. This is useful for professionals like writers, journalists, and doctors who need to quickly turn their spoken thoughts or notes into text.

6. Language Learning:

Speech recognition can be used in language learning apps to help learners improve their pronunciation. Learners speak into the app, and it gives feedback on how accurate their pronunciation is.

7. Voice-Controlled Devices:

_____

Speech recognition lets users control devices like smart speakers, TVs, and cars by speaking. Users can give voice commands to play music, change settings, or navigate menus.

8. Medical Documentation:

In healthcare, speech recognition helps doctors and healthcare workers turn their spoken notes into written records. This helps create medical documents more efficiently.

9. Security and Authentication:

Speech recognition can be used for security purposes, like identifying someone by their voice to allow access to secure systems or complete transactions

10. Automated Translations:

Speech recognition can be combined with translation systems to provide real-time translations of spoken words. This is useful in multilingual environments, such as international conferences or while traveling.

1Conclusion:

In future we can implement emotion detection into a speech recognition system, you can create an emotionally intelligent assistant capable of recognizing the speaker's feelings and responding accordingly. This is especially useful in customer service, therapy, or human-computer interaction, enhancing user experience significantly

References:

Basireddy, Maheswara Reddy. "Implementation of Speech Recognition System Using Python." Review Article.

Washani, Nitin, and Sandeep Sharma. "Speech Recognition System."

Sharma, Dr. Schin. "Speech Recognition System."