International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 7 July 2023, pp: 213-216 www.ijaem.net ISSN: 2395-5252

Incoming Text Message Recognizer to Speech

Jacksonkp, Jeeva Vigneshr, Kameshn, Karthikrajvn
Vel Tech High Tech,Chennai,TamilNadu,India

Date of Submission: 01-07-2023 Date of Acceptance: 10-07-2023

ABSTRACT
Speech recognition is one of the most rapidly expanding technological advancements. In order to do this, significant effort has been made and this has been developed and constructed with that fact in mind. In numerous fields, it can be useful and may even be advantageous. A large percentage of those with disabilities—nearly 20% of the world's population—are blind or unable to use their arms effectively. Some blind people who have trouble reading but can listen to an accessible device's recording of a research paper. Speech recognition software is useful in these circumstances because it enables users of voice-operated devices to communicate with others.The conversion of speech to text using a speech to text system (STT). TTS also converts text into speech in a human-readable way. The suggested device is a hardware option for speech synthesis and voice access to digital content. Our role is to be able to recognize speech and convert text into audio from audio input, as well as recognize text and convert audio into audio from text input.

Keywords: Speech Recognition, Automatic, Speech to text system, Text to Speech system.

I. INTRODUCTION

The purpose of this study is to give a brief overview of enterprise software, voice recognition technology, and business procedures. A spoken word or dictation can be converted into text using speech recognition (SR).Speech recognition, sometimes called "automatic speech recognition" (ASR) or "speech to text," is a technique that transcribes spoken language into written language.Text to speech is a technique for speaking digital text after being converted into voice. A desktop tool called a text to speech synthesizer can automatically read the text, no matter how it may be presentedan input stream from a computer. A speech synthesizer can be created using both hardware and software. This field has had remarkable growth over the last few decades, including a range of rapid advancement, additionally, many top-notch TTS systems are now available for business use. Everyone in culture, whether a human being or an animal, wants to interact with others and spread their own message. Depending on how the communication is received, the recipient may understand the sender completely, partially, or not at all. It might take place in circumstances when there is a communication barrier (for example, when a baby sends a message, some mothers quickly grasp it and some do not).

The first section explores speech recognition technology and its many uses. Sectors, shortcomings, and, in the end, the future of the technology.

Speech recognition systems can be categorized into a number of classes based on their capacity to recognize words and the list of words they have.

Depending on their ability to recognize distinct words, speech recognition systems can be grouped into a number of categories.

Isolated speech is comparable to attached words, also referred to as connected speech.

It is also known as computer dictation,continuous speech allows the user to talk nearly naturally.

In its simplest form, it can be described as un-staged, natural-sounding speech. An ASR device with spontaneous speech functionality can handle run-on words, pauses, and small utterances.It is also the method that, in its most advanced forms, enables humans to interact with a computing device using their voices in a manner that is similar to that of conventional human contact.

The most advanced ASR methods now in use are based on natural language processing (NLP).

Some people find it challenging to type due to physical impairments like multiple sclerosis, repetitive strain injuries (RSI), or other conditions.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 7 July 2023, pp: 213-216 www.ijaem.net ISSN: 2395-5252

For instance, those with hearing impairments could utilize a gadget connected to their phone to convert spoken words from a caller to text.Speech recognition's shortcomings and constraints.

A perfect voice recognition system is challenging to construct, despite all of these benefits and features. There are several things that can affect how accurate and effective a voice recognition program is.

Although speech recognition is a simple operation for a human, it is a challenging challenge for a computer system. Speech recognition programs seem less intelligent when compared to a human mind. The reason for this is that a human brain is a gift from God, andIt is a challenging task for a computer program to think, comprehend, and respond in the same way that humans do it naturally. In order to achieve a balance between phrases, sound, and spaces, it must first understand spoken words in terms of their meanings. While a person is able to separate speech from other noises on their own, computers require instruction and aid in doing so.

II. METHODOLOGY

Algorithms and Description

The platform-independent pyttsx3 library is used to develop the Text-to-Speech converter. The fact that this library may be used offline is its main benefit for text-to-speech conversion. However only Python 2.x is supported by pyttsx. We will now examine pyttsx3, which has been adjusted to run on Python 2.x and Python 3.x.

1. In order to give the synthetic voice a naturalness character, speech synthesis techniques will be applied.
2. The English language's structure may be
3. used as the fundamental building block for voice synthesis.
4. A phoneme-based speech database for the English language will be created.
5. To create the synthetic output speech, phonemes will be searched for in the database and their associated phoneme sounds will be concatenated.

Even though it takes only a few minutes to master the basics of speech formation and speech recognition (and many other people start talking and hearing by the age of 2), computerized speech offers dramatic and powerful capabilities that coders may want to realize and employ.

Despite a 40-year investment in research, voice synthesizer and detection systems still have significant shortcomings.

Overall, voice technology does not always live up to the high standards set by customers who are used to regular human interaction.

To effectively integrate speech outputs and inputs in a user experience and to understand some of Python Speech's more complicated capabilities, it is essential to be aware of both the disadvantages and the advantages.

If you're deciding whether or not todevelopers should be aware of the potential and constraints of speech intelligence when using voice input and output in a particular application.

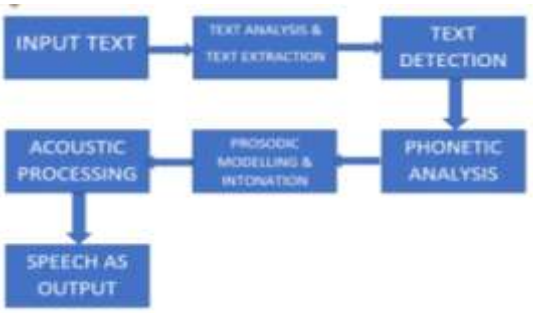

Fig: Text to Speech System

Fig: Method of Speech Recognition

III. DESCRIPTION FOR MODULE

The "Incoming Text Message Recognizer" module is a crucial component of the larger "Text Message to Speech" project. This module serves as the core engine responsible for recognizing and processing incoming text messages received on a user's mobile device and converting them into natural-sounding speech.

Key features and functionalities of the "Incoming Text Message Recognizer" module include:

1. Text message interception: The module intercepts incoming text messages as they arrive on the user's mobile device. It interfaces with the device's operating system or messaging application to capture and process the incoming message data.

2. Natural Language Processing (NLP): NLP algorithms play a vital role in this module. They are responsible for parsing and understanding the textual content of the incoming messages. The NLP component breaks down the text into meaningful units, such as sentences or phrases, and performs language-specific processing, including

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 7 July 2023, pp: 213-216 www.ijaem.net ISSN: 2395-5252

tokenization, part-of-speech tagging, and sentiment analysis.

3. Language detection: The module incorporates language detection mechanisms to identify the language of the incoming text. This ensures that the subsequent speech synthesis accurately renders the text in the appropriate language, supporting multilingual functionality.

4. Text-to-Speech (TTS) synthesis: After the NLP processing, the recognized text is sent to the Text-to-Speech synthesis component. This advanced TTS engine converts the processed text into natural and human-like speech output. It takes into account intonation, pronunciation, andcadence, making the speech sound as close to human speech as possible.

5. Voice selection and customization: Users are often given the option to choose from different voice options for the synthesized speech. Additionally, the module may provide customization settings such as pitch, speed, or accent to tailor the speech output to the user's preference.

6. Accessibility integration: The module seamlessly integrates with accessibility features on the user's mobile device, ensuring that users with visual impairments or reading difficulties can easily access and utilize the text-to-speech functionality.

7. Real-time processing: The module operates efficiently to ensure real-time processing of incoming text messages. This way, users receive immediate and timely speech outputs, enabling them to stay informed and respond promptly to messages.

8. Error handling and feedback: The module incorporates error handling mechanisms to address issues like misinterpretation or language ambiguity. It may provide options for users to provide feedback on the accuracy and quality of the speech output, which can be used to improve the system over time.

Overall, the "Incoming Text Message Recognizer" module is a sophisticated and powerful component that leverages advanced language processing and text-to-speech technologies. Its seamless integration with the user's mobile device allows for effortless and inclusive communication, catering to individuals with diverse needs and enhancing the overall user experience.

IV. CONCLUSION

In conclusion, the Incoming Text Message Recognizer to Speech project is a valuable and innovative solution that aims to enhance accessibility and convenience for users. By leveraging advanced technology, this project allows users to convert incoming text messages into speech, enabling them to listen to their messages instead of reading them.

The project's primary objective is to assist individuals who may have visual impairments, dyslexia, or other reading difficulties. By transforming text messages into speech, it empowers these individuals to engage in effortless and efficient communication through their mobile devices. Moreover, it promotes inclusivity by ensuring that everyone can access and understand text-based information.

The project incorporates cutting-edge natural language processing (NLP) techniques and speech synthesis technologies to accurately recognize and convert text messages into high-quality speech. This combination of NLP and speech synthesis guarantees accurate interpretation and articulation of the content, maintaining the intended meaning of the messages.

Furthermore, the Incoming Text Message Recognizer to Speech project can be seamlessly integrated into various platforms and devices, such as smartphones, tablets, and smartwatches. This versatility ensures widespread accessibility, enabling users to utilize the system across multiple devices and in different contexts.

By implementing this project, users can benefit from improved productivity, as they can multitask and listen to their messages while engaged in other activities. It also promotes safety by reducing the need for users to read messages while driving or performing tasks that require visual attention. In summary, the Incoming Text Message Recognizer to Speech project is a transformative solution that bridges the gap between written communication and auditory comprehension. It empowers individuals with reading difficulties and enhances accessibility for all users, promoting inclusivity and efficiency in communication.

REFERENCES

[1]. Xu Tan, Tao Qin, Frank Soong, and Tie-Yan Liu. A survey on neural speech synthesis. ArXiv preprint arXiv:2106.15561, 2021.

[2]. Baevski, A.; Schneider, S.; and Auli, M. 2020. vq-wav2vec: Self-Supervised Learning of Discrete Speech Representations. In 8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020. OpenReview.net.

[3]. Marta R Costa-jussa, James Cross, Onur C¸ elebi, Maha El- ` bayad,

Kenneth Heafield, Kevin Heffernan, ElaheKalbassi, Janice Lam, Daniel Licht, Jean Maillard, et al., "No language left behind: Scaling human-centered machine translation," arXiv preprint arXiv:2207.04672, 2022.

[4]. M. Afif, R. Ayachi, E. Pissaloux, Y. Said, M. Atri Indoor objects detection and recognition for an ict mobility assistance of visual Multimedia Tools and Applications, 79 (2020).

[5]. M. Praneesh, Dr. D. Napoleon "Extraction of Text from Compound Images using RidlerCalvard Based Median Filtering Algorithm" 17-03-2020.