

TEXT TO SPEECH CONVERSION USING PYTHON

1)S.SAKETH 2) CH.SAKSHITHA 3) K.SAMANTHKUMAR 4) SD.SAMEER 5) K.AKSHAY 6) G.SANDEEP

Assistant Professor Sai Teja

Head Of Department Dr.Thayaba Khatoon

Department of Artificial Intelligence and Machine Learning (AI&ML)

Malla Reddy University, Maisammaguda, Hyderabad

2111CS020469@mallareddyuniversity.ac.in, 2111CS020470@mallareddyuniversity.ac.in

2111CS020471@mallareddyuniversity.ac.in , 2111CS020472@mallareddyuniversity.ac.in

2111CS020473@mallareddyuniversity.ac.in, 2111CS020474@mallareddyuniversity.ac.in

ABSTRACT

This project is to translate text to voice using python. we will create how to convert the human language text into human-like speech using python. I think listening is more efficient than reading for learning to understand a language. Python provides many APIs to convert text to speech. The Google Text to Speech API is popular and commonly known as the GTTS API. it is a Python library to interface with Google Translate's text to speech API.

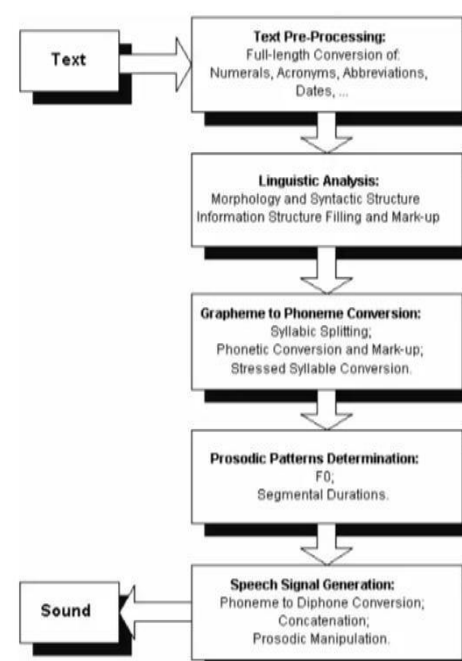
Tkinter is a standard GUI Python library that is one of the fastest and easiest ways to build GUI applications using tkinter. The playsound module is used to play audio files. With this module, we can play a sound file with a single line of code. It is very easy to use the tool and provides many built-in functions which used to save the text file as an mp3 file.

To install the required libraries, you can use pip install command : pip install Tkinter->pip install GTTS- >pip Install playsound .

In this project, we add a message which we want to convert into voice and click on play button to play the voice of that text message . First we import the modules ->create the display window- >define functions. Import libraries-initializing window-function to convert text to speech or voice-function to exit-function to reset-define buttons This is the process we will convert the text to voice or speech programmatically goggle GTTS API and python.

INTRODUCTION

GTTS (Google Text-to-Speech) is a Python library, which is a very easy library that converts the text into audio. The playsound module is used to play audio files. With this module, we can play a sound file with a single line of code. There are several APIs available to convert text to speech in Python. One of such APIs is the Google Text to Speech API commonly known as the gTTS API. gTTS is a very easy to use tool which converts the text entered, into audio which can be saved as a mp3 file. With this project, users can easily convert emails, articles, or any other written content into speech for listening on the go or for accessibility purposes. While there are some limitations to this technology, the text to speech project represents a significant advancement in the field of digital communication and accessibility.



Literature Review:

In N. K. P. S. Shashank Tripathi proposes a system that enables visually impaired, blind and people to use email facility as efficiently as some normal user. The dependency of the system on mouse or keyboard is almost diminished and it work on STT and TTS processes. Face Recognition is also used for authenticating the user identity. They suggested a number of speech representation and classification methods. A number of feature extraction techniques were also deployed by them along with database evaluation and performance. The analysed the various concerns related to Automatic-Speech Recognition and proposed methods to resolve them. The various methods to speech recognition addressed by them are: the AI Approach, the pattern recognition Approach and acoustic phonetic approach. In N. K. P. S. Shashank Tripathi proposes that systems are trained by the individual who will be using the system. These systems are capable of achieving a high command count and better than 95% accuracy for word recognition. The drawback to this approach is that the system only responds accurately only to the individual who trained the system. This is the most common approach employed in software for personal computers

In Van Santen, J.P.H., Sproat, Olive, J.P., and Hirschberg suggested that the quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood. An intelligible text-to-speech program allows people with visual impairments or reading disabilities to listen to written works on a home computer. A text-to-speech system (or "engine") is composed of two parts: a front-end and a back-end.

In Kaveri Kamble, Ramesh Kagalkar suggested to create a TTS system for native languages like Hindi. The system involves of 2 main steps: Text Pre-Processing and Speech Generation. A Concatenative synthesis-based approach is considered for obtaining the speech from the text. A spellchecker module is also implemented for checking the correctness of words for native languages like Hindi. By analysing the various papers, we have concluded that there is vast scope of evolution in the domain of Text-to-speech and Speech-to-text conversion. In the next section we have analysed various TTS and TTS synthesis methods

In S. R. Mache suggested that Test-to-Speech synthesizer is developing rapidly from past few years to gain the current shape. The most suitable methods for TTS are Formant, Articulator and concatenative synthesis. Even in India some research organizations are also working on Text-to-Speech in regional languages like Marathi, Hindi,

Telugu, Punjabi, Kannada, so on. A vast scope of improvement can be achieved in TSS synthesis to obtain a good amount of natural and emotion aspect.

Problem Statement:

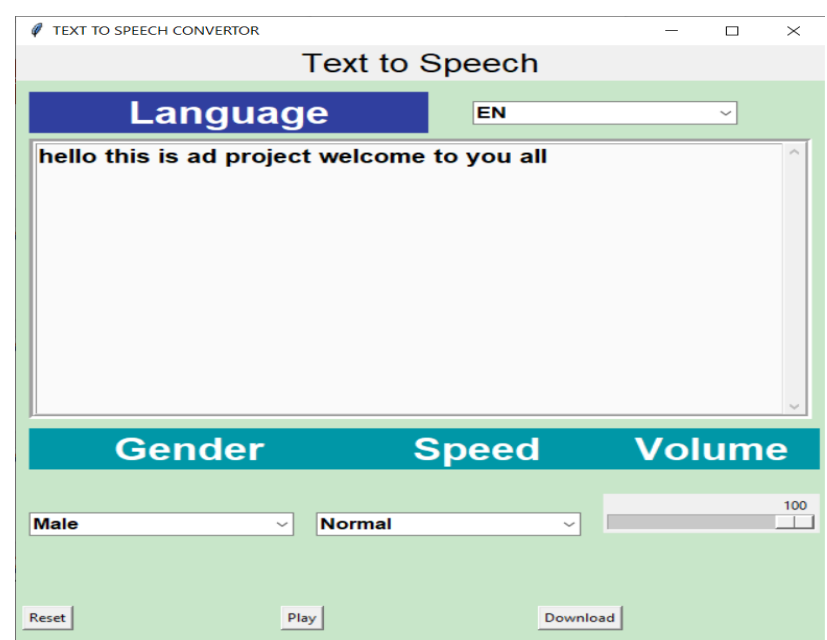
This section provides a clear and concise statement of the problem, this should include a description of the data used in the project and the research questions and hypotheses that guided the project. This problem statement should clearly identify the problem that the research paper is trying to solve and how it will be addressed in the project.

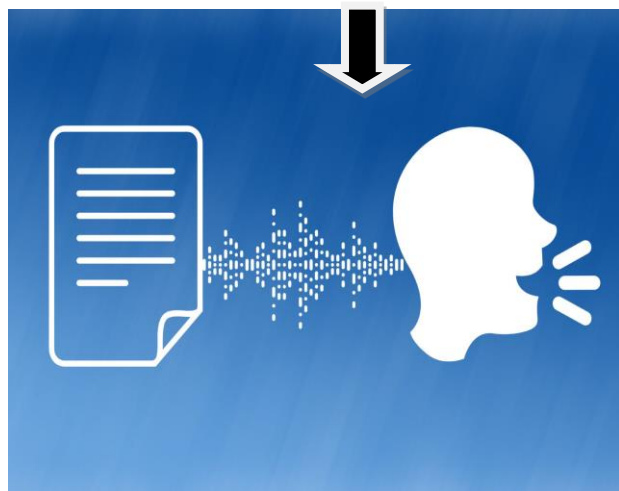
Methodology:

For building the Text-to-Speech converter pyttsx3 library is used which is platform independent. The major advantage of using this library for text-to-speech conversion is that it works offline. However, pyttsx supports only Python 2.x. Hence, we will see pyttsx3 which is modified to work on both Python 2.x and Python 3.x.

- 1 Speech synthesis techniques will be used in order to get the naturalness quality in the synthetic speech.
- 2 The process of the English language can be Used as the basic unit for speech synthesis.
- 3 Speech database for the English language will be developed using phoneme.
- 4 Phonemes will be searched in the database and corresponding phonemes sounds will be Concatenated to generate synthesized output Speech

Experimental Results:





Conclusion:

In conclusion, the text to speech converter project has been successfully completed, and it has provided an effective and user-friendly tool for converting text into speech. The project aimed to provide a solution for people who have difficulty reading or who prefer listening to text instead of reading it. The project has achieved its goals by using the Pyttsx3 library to convert the text to speech and providing a simple and intuitive user interface. Overall, the project has been a success and could be further developed with additional features such as language translation and voice customization.

Future Work:

The future enhancements for the text to speech converter project are numerous and exciting. One possible enhancement could be the addition of more languages to the system. Currently, the project only supports a limited number of languages, which could limit its usefulness in certain regions or for users who speak different languages. By expanding the system to include more languages, the project could become more inclusive and accessible to a wider range of users. Another possible enhancement could be the incorporation of more advanced artificial intelligence and natural language processing algorithms. This would allow the system to better understand and interpret the nuances of human speech, making it more accurate and responsive. Additionally, the system could be integrated with other

technologies such as virtual assistants, allowing users to interact with the system in a more intuitive and natural way. Finally, the project could be expanded to include more features and capabilities, such as the ability to save and store previous speech conversions, or to customize the voice and accent used for the speech output. These enhancements would make the system more versatile and useful for a wider range of applications, and could potentially lead to the development of new and innovative technologies in the field of speech synthesis

References:

- [1] S. R. Mache, "Review on Text-To-Speech Synthesizer," International Journal of Advanced Research in Computer and Communication Engineering, 2015.
- [2] N. K. P. S. Shashank Tripathi, "Voice based Email System for Visually Impaired and Differently abled," International Journal of Engineering Research & Technology (IJERT), 2019
- [3] Rubin, P., Baer, T., and Mermelstein, P., 1981. An articulatory synthesizer for perceptual research. Journal of the Acoustical Society of America 70: 321–328.
- [4] Van Santen, J.P.H., Sproat, Olive, J.P., and Hirschberg, J., 1997. Progress in Speech Synthesis. Springer.
- [5] Mattingly I. G., Speech Synthesis for Phonetic and Phonological Models, T.A. Sebeok (Ed.) Current Trends in Linguistics, Vol. 12, (1974) p. 2451-2487.
- [6] Kaveri Kamble, Ramesh Kagalkar, "A Review: Translation of Text to Speech Conversation for Hindi Language," International Journal of Science and Research (IJSR), 2