

1. INTRODUCTION

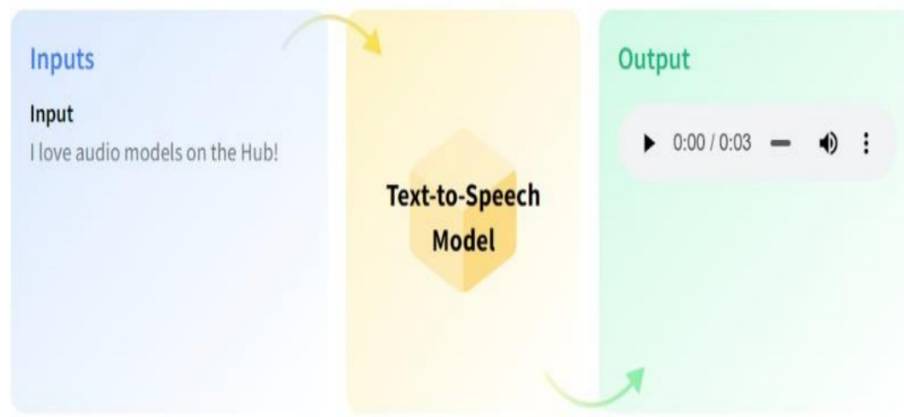
Text-to-speech translator is the automatic conversion of a text into speech that resembles as closely as possible, a native speaker of the language reading that text. Text-to-speech translator (TTS) is the technology which lets computer speak to you. The TTS system gets the text as the input and then a computer algorithm which called TTS engine analyses the text, pre-processes the text and synthesizes the speech with some mathematical models. The TTS engine usually generates sound data in an audio format as the output.

Natural language processing (NLP) is one of the fields of artificial intelligence, computer science and computational linguistics concerned with the interactions between computers and human (natural) languages, and, in particular, concerned with programming computers to process large natural language corpora. Singaravelan et al., Challenges in natural language processing frequently involve speech recognition, natural language understanding, natural language generation (frequently from formal, machine-readable logical forms), connecting language and machine perception, dialog systems, or some combination thereof.

Speech process has become one among the foremost vital analysis areas because of its importance in varied applications. In order to improve communication via network effectively various techniques are applied in the concern field. D Arun Shunmugam et al., In the last twenty years, Text to speech conversion encounters excellent development especially TTS systems are applied in the great extent in various real time application for automatic speech. Rui Ren et al. Remarkably the recent trend in speech process is sentiment analysis and which is the upcoming potential area for research. It is mainly focusing the analysis of people's expression which may be either in text or speech. Sentiment analysis is otherwise called as opinion mining since it analyse opinion, emotions and attitudes of people.

Text-to-Speech (TTS)

Text-to-Speech (TTS) is the task of generating natural sounding speech given text input. TTS models can be extended to have a single model that generates speech for multiple speakers and multiple languages.



Applications:

Text-to-Speech (TTS) models can be used in any speech-enabled application that requires converting text to speech.

- Voice Assistants

TTS models are used to create voice assistants on smart devices. These models are a better alternative compared to concatenative methods where the assistant is built by recording sounds and mapping them, since the outputs in TTS models contain elements in natural speech such as emphasis

- Announcement Systems

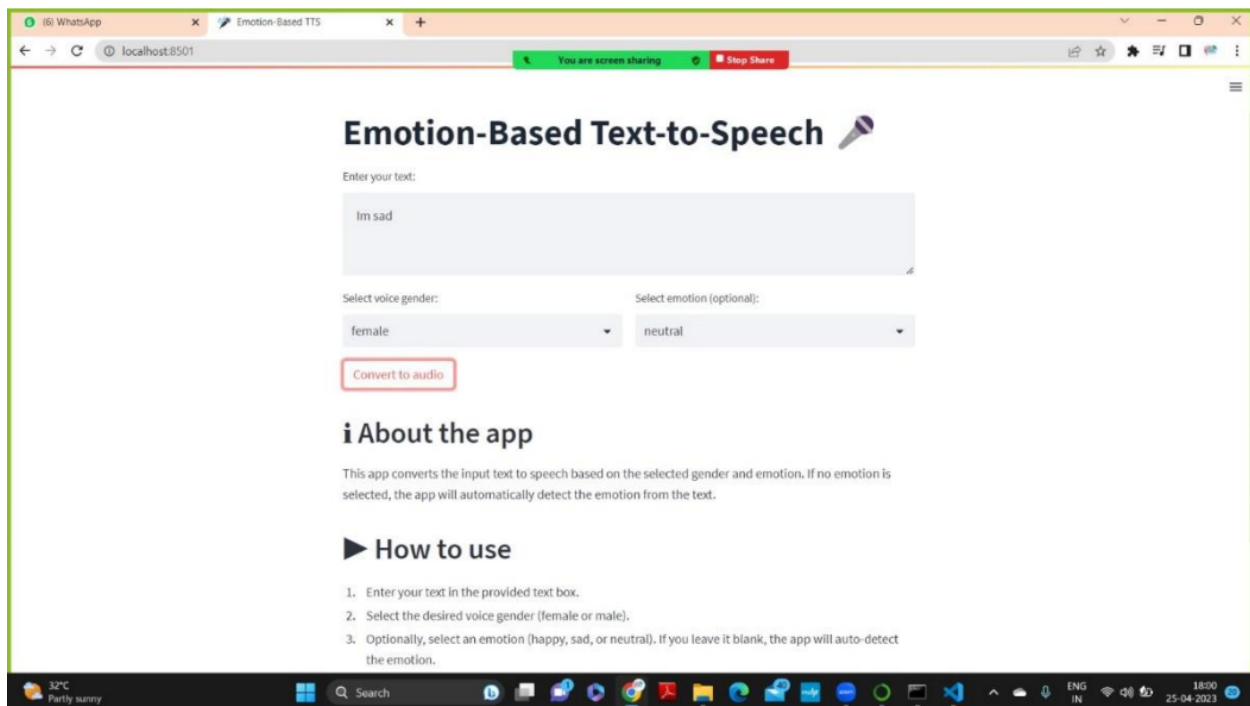
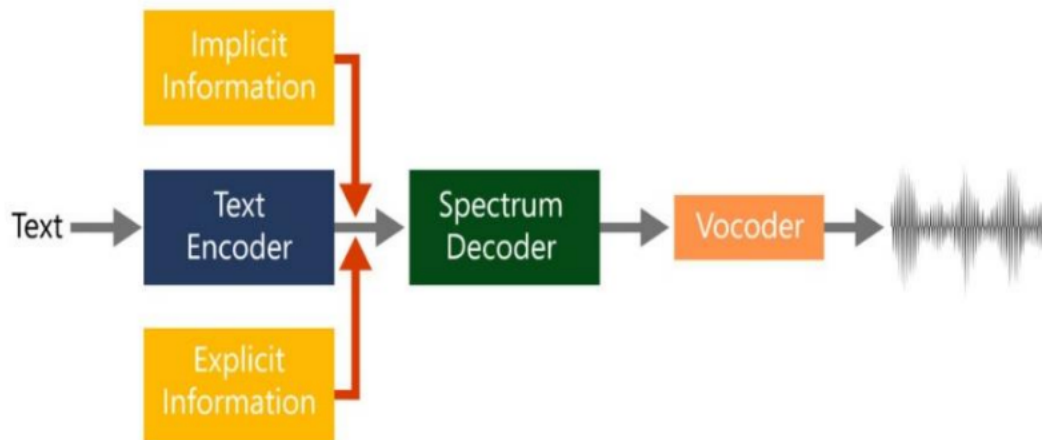
TTS models are widely used in airport and public transportation announcement systems to convert the announcement of a given text into speech.

- Inference

The Hub contains over 100 TTS models that you can use right away by trying out the widgets directly in the browser or calling the models as a service using the Inference API.

EMOTIONAL TEXT TO SPEECH TRANSLATION

3





Emotion-Based Text-to-Speech

Enter your text:

Select voice gender:

female



Select emotion (optional):

auto



Convert to audio

i About the app

This app converts the input text to speech based on the selected gender and emotion. If no emotion is selected, the app will automatically detect the emotion from the text.

► How to use

1. Enter your text in the provided text box.
2. Select the desired voice gender (female or male).
3. Optionally, select an emotion (happy, sad, or neutral). If you leave it blank, the app will auto-detect the emotion.
4. Click on the "Convert to audio" button to listen to the synthesized speech.



Search



EMOTIONAL TEXT TO SPEECH TRANSLATION

5

An audio generated with happy emotion



outputhappy.mp3

An audio generated with sad emotion



outputsad2.mp3

An audio generated with neutral emotion



outputneutral.mp3

female auto

Convert to audio

female
male

This app converts the input text to speech based on the selected gender and emotion. If no emotion is selected, the app will automatically detect the emotion from the text.

Text-to-speech (TTS) synthesis technology enhances user interaction across various platforms, and we have developed a user-friendly interface for converting typed text to speech using an English TTS engine. Future implementations include integrating TTS in telephony, ATMs, and video games. Additionally, we are developing an Emotion-Based Text-to-Speech (EBTTS) system using LSTM networks to extract and convey emotions in speech, which will benefit applications like storytelling, audiobooks, and assistive technologies for the visually impaired.