Audio & text conversion Document

20 January 2024   05:09 PM

1. Audio to Text Conversion using whisper :

```python
# step: 01: set up ffmpeg in windows : as :
# need to install it as a python module (with pip install ffmpeg) and from here (https://ffmpeg.org/download.html)
# while following these instructions: https://www.geeksforgeeks.org/how-to-install-ffmpeg-on-windows/

# Step: 02: !pip install -U openai-whisper


# step: 03: select the model:
import whisper

model = whisper.load_model("small") #769M Parameter

# WER : word error rate
```

```python
# Play the video that we have downloaded

from IPython.display import Audio, display

display(Audio('2_input_audio.mp3', autoplay=True))
```

i/p

```python
result = model.transcribe("./2_input_audio.mp3")
print(result["text"])
```

o/p

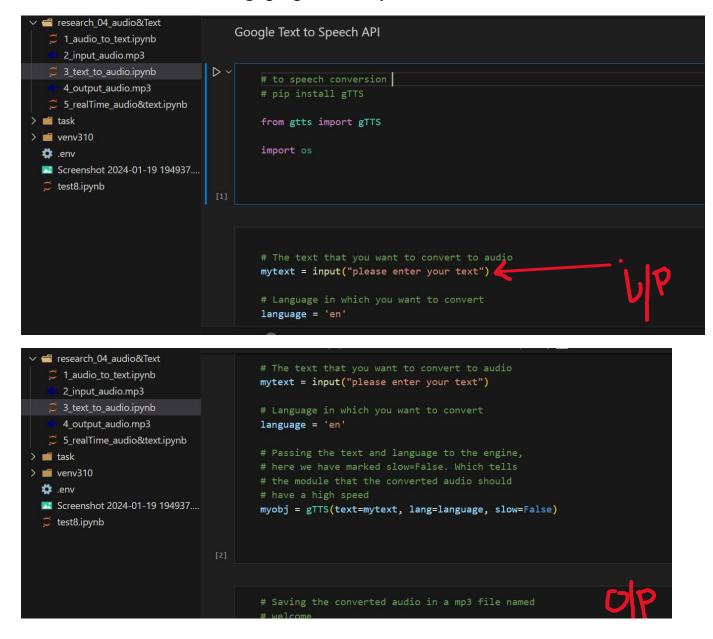leading management consulting firm focusing on translating the power of technology into realizable business strategies

2. Audio to Text Conversion using "assemblyai" :

The first screenshot shows a VS Code notebook with the following code:

```python
# `pip3 install assemblyai` (macOS)
# `pip install assemblyai` (Windows)

import assemblyai as aai

aai.settings.api_key = "bc72a669f0364df28edbfe616afbc00d"

config = aai.TranscriptionConfig(auto_highlights=True)

transcriber = aai.Transcriber()
transcript = transcriber.transcribe("2_input_audio.mp3",config=config)   # i/p

# for result in transcript.auto_highlights.results:
#   print(f"Highlight: {result.text}, Count: {result.count}, Rank: {result.rank}")

print(transcript.text)   # o/p
```

Output: `Leading management consulting firm, focusing on translation the power of technology into realizable business strategies`

3. Text To Audio Conversion using "google text to speech API":

Google Text to Speech API

```python
# to speech conversion
# pip install gTTS

from gtts import gTTS

import os
```

```python
# The text that you want to convert to audio
mytext = input("please enter your text")   # i/p

# Language in which you want to convert
language = 'en'
```

```python
# The text that you want to convert to audio
mytext = input("please enter your text")

# Language in which you want to convert
language = 'en'

# Passing the text and language to the engine,
# here we have marked slow=False. Which tells
# the module that the converted audio should
# have a high speed
myobj = gTTS(text=mytext, lang=language, slow=False)
```

```python
# Saving the converted audio in a mp3 file named
# welcome
```
o/p

```python
# Saving the converted audio in a mp3 file named
# welcome
myobj.save("4_output_audio.mp3")

# Playing the converted file
```

O|P