

## Portafolio de evidencias

Este proyecto de Procesamiento de Lenguaje Natural (NLP) se centra en la transcripción de un archivo de audio, seguido por la generación de un resumen que se presenta en forma de viñetas. La implementación se realiza completamente en Python, utilizando bibliotecas y herramientas especializadas para el procesamiento de audio y texto. Para visualizar y compartir los resultados de manera interactiva, se utiliza Streamlit, una biblioteca que simplifica la creación de aplicaciones web interactivas. Este enfoque proporciona una solución eficiente y amigable para explorar y comprender fácilmente los resultados del análisis de NLP.

Lo primero que se hace es instalar las librerías correspondientes en el environment en el cual se está trabajando.

```
Anaconda Prompt (anaconda)
act1 D:\anaconda\envs\act1
retoTC2004B D:\anaconda\envs\retoTC2004B
tda D:\anaconda\envs\tda

(base) C:\Users\gaspe>conda activate act1

(act1) C:\Users\gaspe>conda list
# packages in environment at D:\anaconda\envs\act1:
#
# Name Version Build Channel
altair 5.1.2 pypi_0 pypi
attrs 23.1.0 pypi_0 pypi
blinker 1.7.0 pypi_0 pypi
bzip2 1.0.8 he774522_0
ca-certificates 2023.08.22 haa95532_0
cachetools 5.3.2 pypi_0 pypi
certifi 2023.7.22 pypi_0 pypi
charset-normalizer 3.3.2 pypi_0 pypi
click 8.1.7 pypi_0 pypi
colorama 0.4.6 pypi_0 pypi
gitdb 4.0.11 pypi_0 pypi
gitpython 3.1.40 pypi_0 pypi
idna 3.4 pypi_0 pypi
importlib-metadata 6.8.0 pypi_0 pypi
jinja2 3.1.2 pypi_0 pypi
jsonschema 4.19.2 pypi_0 pypi
jsonschema-specifications 2023.7.1 pypi_0 pypi
libffi 3.4.4 hd77b12b_0
markdown-it-py 3.0.0 pypi_0 pypi
```

Posteriormente se hacen ajustes al código que hace el transcript para que se pueda ejecutar en streamlit.

```
import streamlit as st
import openai
import whisper
import numpy as np

# Configura la clave de la API de OpenAI
openai.api_key = 'sk-JDQ4we5lxs0jIjpu0Ww8T3B1bkFJAKVsxxhz4f4JLDyFC7sI7'

# Carga el modelo de Whisper
model = whisper.load_model("base")

def transcribe_audio(model, file_path):
    transcript = model.transcribe(file_path)
    return transcript['text']

def CustomChatGPT(user_input):
    messages = [{"role": "system", "content": "You are an office administrator, summarize the text in key points"}]
    messages.append({"role": "user", "content": user_input})
    response = openai.ChatCompletion.create(
        model="gpt-3.5-turbo",
        messages=messages
    )
    ChatGPT_reply = response["choices"][0]["message"]["content"]
    return ChatGPT_reply

def load_audio_file(file_path):
    # Esta función carga el archivo de audio y lo convierte en un array de NumPy
    # Aquí asumimos que el formato es WAV
    import librosa

    audio, _ = librosa.load(file_path, sr=None)
    return audio.astype(np.float32)
```

```
def main():
    st.title("Resumen de Audio con Streamlit")

    # Audio predefinido (puedes cambiarlo según tus necesidades)
    file_path = 'MA1.m4a'

    st.audio(file_path, format="audio")

    # Transcribe el audio
    st.subheader("Transcripción:")
    transcription = transcribe_audio(model, file_path)
    st.write(transcription)

    # Genera el resumen con ChatGPT
    st.subheader("Resumen Generado:")
    summary = CustomChatGPT(transcription)
    st.write(summary)

if __name__ == "__main__":
    main()
```

Luego desde la consola se ejecuta el código

```

- Anaconda Prompt (anaconda) - conda install -c conda-forge ffmpeg - streamlit run whisper_summarize.py
File "D:\Descargas 2\whisper-chatGPT-Audio\whisper_summarize.py", line 43, in transcribe_audio
    transcript = model.transcribe(file_path)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "D:\anaconda\envs\act1\Lib\site-packages\whisper\transcribe.py", line 122, in transcribe
    mel = log_mel_spectrogram(audio, model.dims.n_mels, padding=N_SAMPLES)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "D:\anaconda\envs\act1\Lib\site-packages\whisper\audio.py", line 141, in log_mel_spectrogram
    audio = torch.from_numpy(audio)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
TypeError: expected np.ndarray (got UploadedFile)
Stopping...

(act1) D:\Descargas 2\whisper-chatGPT-Audio>streamlit run whisper_summarize.py

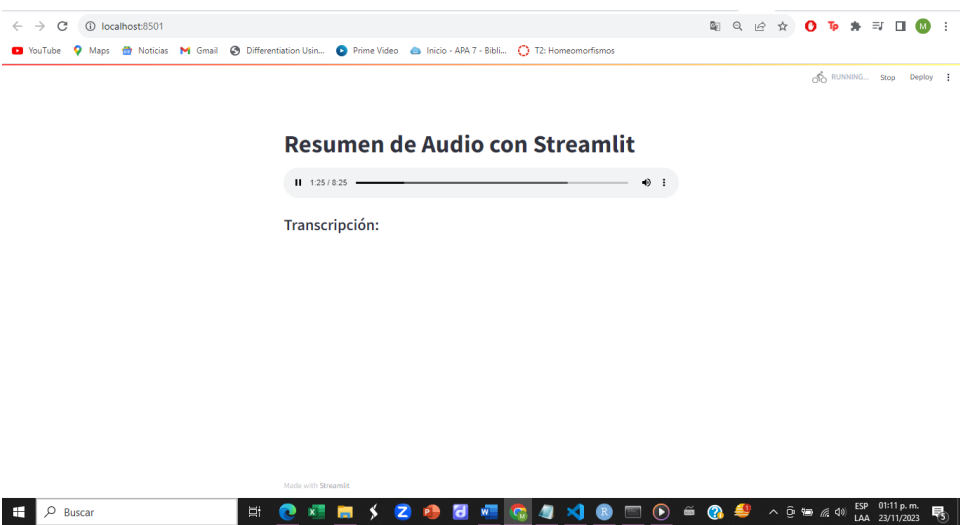
You can now view your Streamlit app in your browser.

Local URL: http://localhost:8501
Network URL: http://10.22.130.218:8501

D:\anaconda\envs\act1\Lib\site-packages\whisper\transcribe.py:115: UserWarning: FP16 is not supported on CPU; using FP32
instead
  warnings.warn("FP16 is not supported on CPU; using FP32 instead")

```

Finalmente, en una página web podemos ver el resultado. Primero sólo se muestra la opción de reproducir el audio.



Una vez que se reproduce el audio, se ve el transcript y al final se muestra el resumen.

localhost:8501/#f2585bfb

YouTube Maps Noticias Gmail Differentiation Usin... Prime Video Inicio - APA 7 - Bibli... T2: Homeomorfismos

RUNNING... Stop Deploy

## Resumen de Audio con Streamlit

▶ 3:41 / 8:25 🔊

**Transcripción:**

Especially, and I would even argue in the last four weeks, you really can't build fitness for the Iron Man distance. But you can ruin your fitness for the Iron Man distance, especially within the last two to three weeks. So it takes your body to adapt to a full load, a full Iron Man load. It takes four to six weeks for your body to fully adapt to it. For you to go through the full depression cycle all the way through the compensation cycle and come back to the place where you're ready to race again. So if you take your body and you're in that super compensation phase and you start training in that which is going to be in these last two weeks which you're in right now. And you add load to it and drop down again. You're going to lose everything that you gained from those big efforts that you did in the four to six weeks out from your race. Okay, so you have to give your body time to rest. You need to let your body fully recover from those efforts and even get stronger. So you can keep the sword sharp by doing short quick efforts. You can get out the door to work out some of your whatever your issues are. They're mental. I need to go out and do something. Fine. Zone one, zone two, easy, 30, 45 minutes to at least open the body up and remind yourself what you're going to be doing. Those things are fine. Weight training in the last two to four weeks. Try it out. You know that? No weight training. It does not hold on in a program. In the last two weeks for sure. I recommend, usually even in the last four weeks unless you've been super consistent with it all the way through your all the way through your training. Like not even squat or push up. Stop today. How about the volume? I mean, I look at my schedule. It's like a 50% of my peak. You think it should be appropriate. You want to be going from your peak volume. You should be cutting back to like 75% four

Buscar

01:13 p.m. ESP LAA 23/11/2023

localhost:8501/#f2585bfb

YouTube Maps Noticias Gmail Differentiation Usin... Prime Video Inicio - APA 7 - Bibli... T2: Homeomorfismos

Deploy

dialent, the right amount for you, because there's a huge range on what people need for salt. There's really heavy salt sweaters and there's light, and you can do too much salt. And you have this like really similar light cramping from too much salt, so you really need a dial in in your training.

**Resumen Generado:**

- Building fitness for Iron Man distance is not possible within the last four weeks before the race.
- Training during the super compensation phase in the last two weeks can lead to a loss of gains made in the previous four to six weeks.
- Rest and full recovery are essential during the tapering period, allowing the body to adapt and become stronger.
- Weight training should be avoided during the last two to four weeks, unless consistent throughout the entire training period.
- Volume should be reduced gradually, cutting back to 75% four weeks out and almost half two weeks out.
- Tapering strategies may differ for each individual, and it may take some trial and error to find the right approach.
- Proper nutrition is crucial for Iron Man races, and it is essential to practice and find what works during training.
- Having redundancies in nutrition during the race can help prevent potential GI issues.
- Salt intake should be balanced to ensure proper absorption of carbohydrates and prevent stomach discomfort and cramping.
- Proper hydration and fueling during the running portion of the race is crucial to maintain energy levels and avoid bonking.
- Dialing in the right amount of salt and finding individual needs through training is important.

Buscar

01:13 p.m. ESP LAA 23/11/2023