



Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

“Yo, como integrante de la comunidad estudiantil del Tecnológico de Monterrey, soy consciente de que la trampa y el engaño afectan mi dignidad como persona, mi aprendizaje y mi formación, por ello me comprometo a actuar honestamente, respetar y dar crédito al valor y esfuerzo con el que se elaboran las ideas propias, las de los compañeros y de los autores, así como asumir mi responsabilidad en la construcción de un ambiente de aprendizaje justo y confiable”

“Inteligencia artificial avanzada para la ciencia de datos II”

Transcripción de audio y resumen con GPT-3

Frida Cano Falcón

A01752953

Profesor:

Juan Arturo Nolasco Flores

Fecha de entrega: 23 de noviembre de 2023

Descripción

En este proyecto se busca desarrollar una aplicación web simple desarrollada con Streamlit con el objetivo de transcribir un archivo de audio y generar un resumen del mismo utilizando el modelo GPT-3.

Importación de bibliotecas:

Se importan las bibliotecas necesarias, incluyendo Streamlit, OpenAI, y Whisper (para transcripción de audio). También se importa la biblioteca *os* para manipulación de archivos.

Configuración de la clave de API de OpenAI:

Se establece la clave de API de OpenAI para autenticar las solicitudes al servicio GPT-3.

Carga del modelo Whisper:

Se carga el modelo Whisper para transcripción de audio utilizando la biblioteca Whisper.

Definición de funciones:

- `transcribe_audio(model, file_path)`: Esta función toma un modelo y la ruta de un archivo de audio como entrada y devuelve la transcripción del audio utilizando el modelo Whisper.
- `CustomChatGPT(user_input)`: Esta función utiliza el modelo GPT-3.5-turbo de OpenAI para generar un resumen basado en el input del usuario.

Streamlit App:

- Se define una función principal (main) para la aplicación Streamlit. Se utiliza **`st.title`** para mostrar el título de la aplicación. Se utiliza **`st.markdown`** para mostrar información adicional, incluyendo el nombre del desarrollador y una descripción de la aplicación.
- Se proporciona un botón de carga de archivos de audio (**`st.file_uploader`**) para que el usuario pueda subir su archivo.

- Se guarda el archivo en la carpeta local y se muestra un mensaje de éxito utilizando `st.success`. Esto para que se pueda aplicar la función de transcripción en el audio.
- Se utiliza `st.audio` para mostrar un reproductor de audio para el archivo subido.
- Se realiza la transcripción del audio utilizando la función ***transcribe_audio*** y se muestra el resultado con **`st.write`**.
- Se genera un resumen utilizando la función ***CustomChatGPT*** y se muestra en la interfaz.

Ejecución de la aplicación:

La aplicación se ejecuta cuando el script se ejecuta directamente (`__name__ == "__main__"`). El usuario puede cargar un archivo de audio, y la aplicación transcribe el audio y genera un resumen utilizando GPT-3, mostrando los resultados en la interfaz de Streamlit.

Display de la aplicación

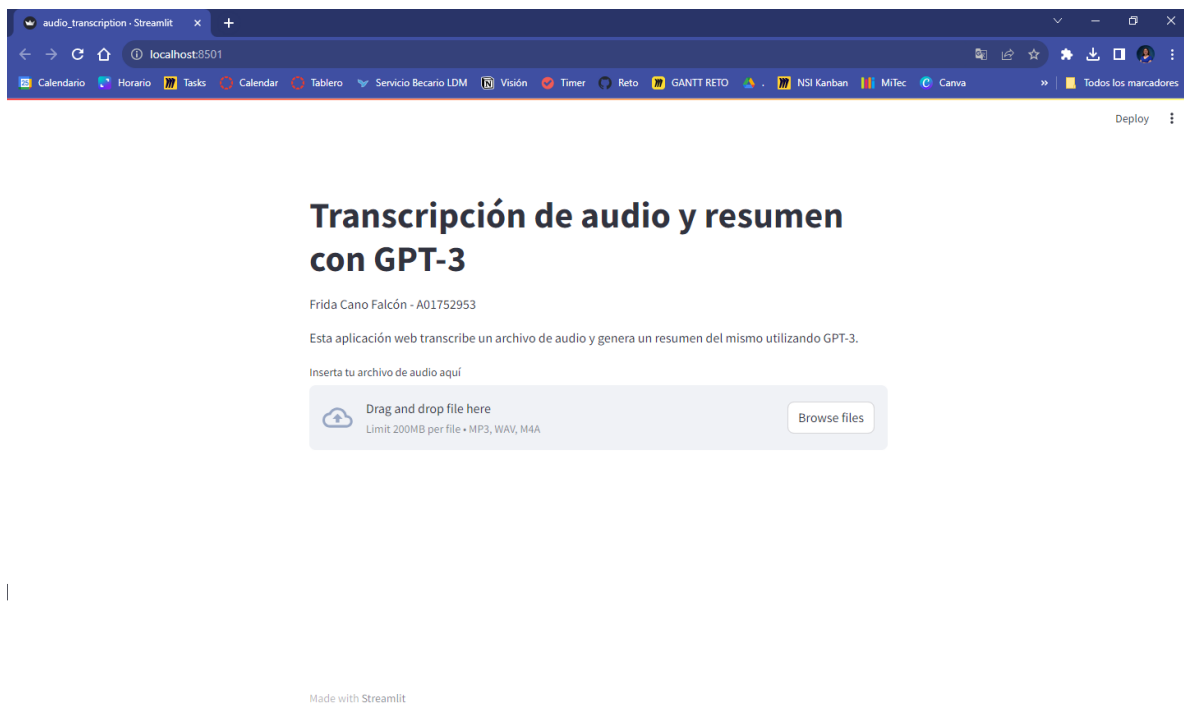


Fig 1. Página inicial

Transcripción de audio y resumen con GPT-3

Frida Cano Falcón - A01752953

Esta aplicación web transcribe un archivo de audio y genera un resumen del mismo utilizando GPT-3.

Inserta tu archivo de audio aquí

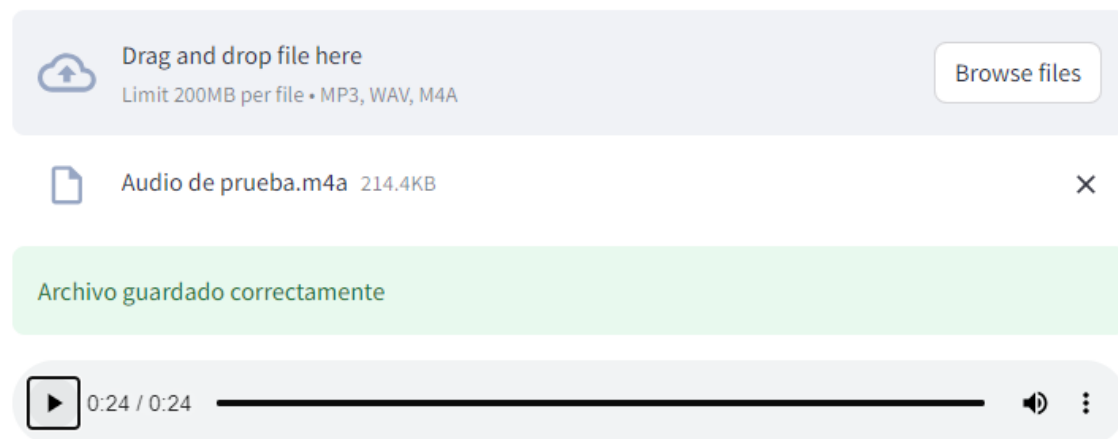


Fig 2. Cuando se arrastra el archivo de texto

Transcripción:

Estoy realizando una actividad de logística, utilizando el neto de barrido para gestionar cuándo de 8 puntos de entrega del nozen, debido a restricciones que nos ponen. El problema de entregas, este cantidad de camiones disponibles y pues horarios y demás.

Resumen:

- El autor está llevando a cabo una actividad de logística para gestionar ocho puntos de entrega del producto Nozen.
- El autor menciona que hay restricciones que se les imponen, lo cual puede afectar las entregas.
- Hay un problema con la cantidad de camiones disponibles para realizar las entregas.
- También menciona que hay problemas con los horarios y otros factores relacionados con las entregas.

Fig 3. Resultados de la transcripción y resumen