



Tecnológico de Monterrey
Escuela de Ingeniería y Ciencias

Tecnológico de Monterrey

ESCUELA DE INGENIERÍA Y CIENCIAS

Inteligencia Artificial Avanzada para la Ciencia de Datos

Procesamiento de lenguaje natural

04/12/2023

Profesor:

Juan Arturo Nolasco Flores

Autor:

Leonardo Ramírez Ramírez - A01351715



1. Introducción

El desarrollo cada vez más palpable de tecnologías de IA ha abierto camino a un nuevo paradigma en el desarrollo tecnológico de diferentes aplicaciones, así como un cambio importante en el acercamiento utilizado para la solución de problemas.

Dentro del procesamiento de lenguaje natural (NLP) el desarrollo no solamente ha tenido grandes avances sino que ha demostrado ser una de las áreas con aplicaciones más importantes para resolver problemas actuales. Por ello, en el presente documento se explica de forma general el proyecto implementado para la clase de NLP utilizando aplicación de procesamiento de lenguaje no estructurado a lenguaje estructurado. El objetivo principal de este proyecto es explorar las posibilidades que brindan las herramientas de NLP y las APIs para transformar datos no estructurados, como el habla humana, en información procesable y comprensible. Nos enfocamos en dos tareas clave: transcripción de archivos de audio y generación de resúmenes mediante modelos de lenguaje avanzados.

2. Desarrollo

El proyecto se desarrolló utilizando Streamlit, un framework de desarrollo web para proyectos de ciencia de datos e inteligencia artificial. Simultáneamente, se implementaron dos herramientas de NLPa través de su API. La primera de ellas fue Whisper la cual, al ser un modelo especializado en transcripción de voz para convertir señales acústicas en texto, realizó la transcripción precisa de archivos de audio. La segunda fue GPT-3.5-turbo de OpenAI la cual, al ser un modelo de lenguaje de última generación, fue capaz de comprender y generar texto de manera contextual y resumida, ofreciendo respuestas coherentes y relevantes según la transcripción hecha por Whisper. Cabe mencionar que fue necesario utilizar la API key de OpenAI para poder acceder al modelo GPT-3.5-turbo.

El proceso requiere que el usuario proporcione un archivo de audio de tipo m4a para que Whisper y GPT-3.5-turbo puedan hacer su función de transcripción y resumen respectivamente. Si el archivo se cargó correctamente, se desplegará un aviso de éxito. Posteriormente el usuario deberá presionar el botón *Transcribir y resumir* para ejecutar la instrucción para que Whisper y GPT-3.5-turbo hagan su función sobre el archivo adjunto. Una vez haya sido procesada la información, se desplegará el resultado debajo del botón *Transcribir y resumir*.

2.1. Sobre la interfaz en Streamlit

La interfaz de Streamlit ofrece dos secciones: **Inicio** y **Transcripción y resumen**. En la primera se desglosa un resumen del objetivo y finalidad del proyecto, además de mostrar información de la autoría de este trabajo. En la segunda sección se lleva a cabo el proceso de transcripción y resumen como ya se mencionó anteriormente.

2.2. Demostración

A continuación se muestran capturas de pantalla del funcionamiento de la interfaz.

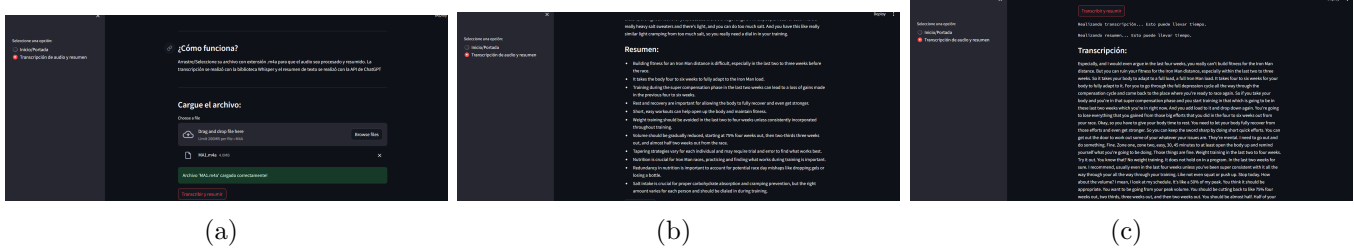


Figura 1: Capturas de pantalla de Streamlit en cada paso: a) Antes de ejecutar las APIs de NLP. b) Después de ejecutar ambas APIs se muestra el resultado de la transcripción. c) Se muestra el resultado del resumen

3. Comentarios finales

La evidencia/actividad en cuestión fue retadora pues exigió entender la interconexión de herramientas para la solución de problemas de procesamiento de lenguaje natural así como su ajuste con el framework de trabajo de Streamlit.

4. Anexo: Código de Python

Se anexa el link de repositorio de GitHub donde se encuentra el código implementado en formato .py. El repositorio se encuentra en la siguiente liga: Github link

Referencias