

# Prototipo 6 - Médico IA

Andrés Felipe Patarroyo Muñoz - 506221032

Santiago Jair Torres Rivera - 506221074

**Abstract**—La aplicación "Médico IA" es una aplicación web que combina tecnologías de reconocimiento de voz e inteligencia artificial para proporcionar respuestas médicas a los usuarios. Utilizando un backend desarrollado con FastAPI en Python, la aplicación ofrece una variedad de endpoints para permitir la comunicación fluida entre el usuario y la IA médica. A través de una interfaz de usuario diseñada con HTML, CSS y JavaScript, los usuarios pueden interactuar con el chatbot médico, haciendo preguntas y recibiendo respuestas de manera intuitiva. La captura de audio desde el micrófono del dispositivo permite una interacción verbal, mientras que la síntesis de voz convierte las respuestas del chatbot en audio para una experiencia más inmersiva. La aplicación demuestra el potencial de la inteligencia artificial en el campo de la atención médica, ofreciendo a los usuarios una forma conveniente y accesible de obtener información médica relevante y precisa.

**Keywords:** Aplicación Médica, Inteligencia Artificial, Reconocimiento de Voz, Backend FastAPI, Frontend HTML/CSS/JavaScript, Interacción Usuario-IA, Respuestas Médicas, Síntesis de Voz, Entorno Virtual, Librerías Python, Videollamada, Procesamiento de Datos, Seguridad de Datos.

## I. INTRODUCCIÓN

En la actualidad, la tecnología de inteligencia artificial (IA) y el reconocimiento de voz están revolucionando diversos campos, incluida la atención médica. El proyecto "Médico IA" es una aplicación web que utiliza estas tecnologías para ofrecer respuestas médicas de manera accesible y eficiente. La aplicación se compone de un backend desarrollado en Python utilizando el framework FastAPI y un frontend diseñado con HTML, CSS y JavaScript, lo que permite una interacción fluida y dinámica entre el usuario y la IA médica. La aplicación permite a los usuarios hacer preguntas médicas de forma verbal a través del micrófono de su dispositivo. Utilizando la biblioteca de reconocimiento de voz de Google, el audio se convierte en texto, que luego es procesado por un modelo de IA para generar respuestas médicas precisas. Estas respuestas se devuelven al usuario tanto en formato de texto como de audio, proporcionando una experiencia interactiva e inmersiva. "Médico IA" no solo demuestra la capacidad de la IA para analizar y responder consultas médicas, sino que también subraya la importancia de la accesibilidad y la conveniencia en la atención sanitaria. A través de una interfaz de usuario intuitiva y la implementación de avanzadas técnicas de procesamiento de lenguaje natural, esta aplicación tiene el potencial de mejorar la interacción entre pacientes y profesionales de la salud, ofreciendo una herramienta útil para obtener información médica relevante de manera rápida y efectiva.

## II. IMPLEMENTACIÓN

### *Entorno virtual y librerías*

Primeramente hablaremos de las librerías que tuvimos que instalar en nuestro entorno virtual, el cual será un entorno aislado para este proyecto, esta es la lista de librerías usadas: uvicorn, watchfiles, anyio, certifi, charset-normalizer, click, colorama, dnspython, email-validator, fastapi, fastapi-cli, h11, httpcore, httptools, httpx, idna, Jinja2, markdown-it-py, MarkupSafe, mdurl, numpy, opencv-python, orjson, PyAudio, pydantic, pydantic-core, Pygments, python-dotenv, python-multipart, PyYAML, requests, rich, shellingham, sniffio, SpeechRecognition, starlette, typer, typing-extensions, ujson, urllib3, websockets, para un plus de nuestro proyecto instalamos dos librerías las cuales son SpeechRecognition y pytsx3 (esta es de una rama opcional) para que la pregunta se haga de forma interactiva con la voz y así mismo se reproduzca la respuesta de Ollama

### *Backend*

El backend desarrollado con FastAPI está diseñado para simular una experiencia interactiva con una IA médica, ofreciendo una serie de endpoints que facilitan diferentes formas de comunicación entre el usuario y la inteligencia artificial.

El primer endpoint, de tipo GET, proporciona una transmisión en tiempo real de video desde la cámara del dispositivo. Este endpoint permite al usuario visualizar la IA médica, creando una experiencia similar a una videollamada, lo que facilita una comunicación más personalizada y efectiva. El segundo endpoint, también de tipo GET, permite capturar y reconocer voz a través del micrófono del dispositivo. Esta funcionalidad posibilita al usuario comunicarse verbalmente con la IA médica, proporcionando información sobre síntomas, consultas médicas o cualquier otra pregunta relacionada con su salud. El tercer endpoint, de tipo POST, recibe texto como entrada y lo convierte en voz sintetizada. Esto permite que la IA médica responda verbalmente al usuario, brindando información médica, consejos o instrucciones sobre el tratamiento adecuado según la consulta realizada.

Además de estos endpoints específicos de interacción, el backend también incluye soporte para archivos estáticos y utiliza plantillas Jinja2 para la generación dinámica de páginas HTML. Asimismo, se ha configurado el middleware CORS para permitir solicitudes desde cualquier origen, lo que facilita la integración con otras aplicaciones o servicios web.

### *Frontend*

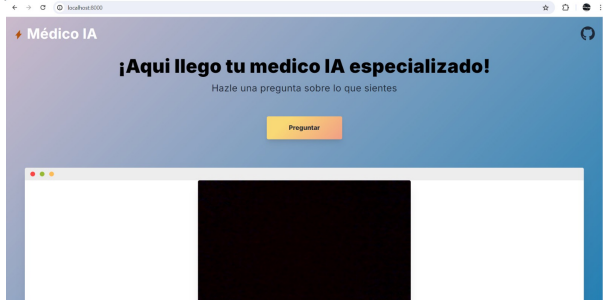
El código HTML proporciona la estructura visual y la disposición de elementos para interactuar con un chatbot

médico. Incluye un encabezado con un enlace al repositorio del proyecto en GitHub para referencia adicional. La sección principal presenta un título llamativo que invita al usuario a realizar una consulta médica al chatbot, junto con un mensaje motivador. Además, se encuentra un botón de acción claramente visible que invita al usuario a iniciar la interacción con el chatbot. El área de visualización del video en tiempo real, aunque se deja comentada, está presente para permitir futuras implementaciones visuales.

La sección de respuesta está diseñada para mostrar las respuestas del chatbot médico. Aquí, un título proporciona una indicación visual de que se ha recibido una respuesta, y un párrafo vacío espera llenarse con el texto de la respuesta del chatbot.

Por otro lado, en el código JavaScript se maneja la interacción dinámica del usuario con el chatbot. Cuando el usuario hace clic en el botón "Preguntar", se ejecuta una función definida. Esta función inicia una serie de acciones secuenciales: primero, se realiza una solicitud GET al endpoint /habla del backend para capturar la entrada de voz del usuario. Una vez que se captura el audio, se genera una pregunta basada en el texto obtenido y se envía como una solicitud POST al endpoint /api/generate del backend para obtener una respuesta del modelo de generación de lenguaje.

Luego, la respuesta obtenida del chatbot se envía como una solicitud POST al endpoint /responde del backend para que la IA médica responda verbalmente. Finalmente, la respuesta del chatbot se muestra en la interfaz de usuario, actualizando el título y el párrafo correspondiente para reflejar el mensaje recibido.



### III. RESULTADOS

La aplicación "Médico IA" es una plataforma que utiliza tecnologías de reconocimiento de voz e inteligencia artificial para proporcionar respuestas médicas a los usuarios. La aplicación se inicia mediante un backend desarrollado en Python utilizando el framework FastAPI. Este backend inicializa un servidor web que proporciona una API para la comunicación con el frontend y realiza diversas tareas de procesamiento de datos.

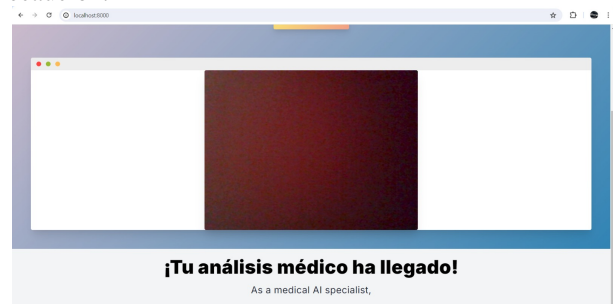
El frontend de la aplicación está construido utilizando HTML, CSS y JavaScript. Proporciona una interfaz de usuario intuitiva que permite a los usuarios interactuar con la aplicación. La interfaz incluye un título, un área para mostrar respuestas médicas y un botón para iniciar la interacción.

Cuando un usuario hace clic en el botón de "Preguntar", se activa la captura de audio desde el micrófono de su dispositivo.

Este audio se procesa utilizando la API de reconocimiento de voz de Google, convirtiéndolo en texto. El texto reconocido se envía al backend, donde se utiliza como "prompt" para solicitar a un modelo de IA que genere una respuesta médica. Este modelo de IA utiliza datos previamente entrenados para generar una respuesta precisa y relevante. La respuesta generada se envía de vuelta al frontend.

La respuesta del modelo se envía nuevamente al backend, donde se utiliza una librería de síntesis de voz para convertirla en audio. Este audio se devuelve al frontend y se reproduce para el usuario, proporcionando una respuesta hablada a su pregunta.

Al finalizar el proceso, el usuario recibe una respuesta hablada a su pregunta médica. Esta respuesta se genera utilizando tecnologías de inteligencia artificial y se adapta a la consulta específica del usuario. La aplicación "Médico IA" demuestra el potencial de la inteligencia artificial y el reconocimiento de voz en el campo de la atención médica. Proporciona a los usuarios una forma conveniente y accesible de obtener información médica relevante y precisa. Sin embargo, es importante tener en cuenta que esta aplicación no reemplaza el consejo médico profesional y debe utilizarse con precaución.



### IV. CONCLUSIÓN

La aplicación "Médico IA" demuestra el poder y la versatilidad de las tecnologías de reconocimiento de voz e inteligencia artificial en el ámbito de la atención médica. A través de una combinación eficiente de backend desarrollado con FastAPI y un frontend intuitivo, la aplicación proporciona una plataforma accesible y efectiva para obtener respuestas médicas. Permite a los usuarios interactuar verbalmente con un chatbot médico, lo que hace que la experiencia sea más natural y cómoda. El uso de modelos de IA para procesar y responder consultas médicas asegura que las respuestas sean precisas y relevantes, mejorando la accesibilidad a la información médica. Además, la síntesis de voz ofrece una experiencia interactiva y envolvente, transformando las respuestas en audio y mejorando la usabilidad para los usuarios.