

Desarrollo de una Herramienta Avanzada para la Consulta de Desapariciones basada en Modelos LLM Multimodales y Fuentes de Datos Geográficos

Objetivo del Proyecto: El propósito de esta solución es desarrollar una herramienta de consulta avanzada que utiliza modelos de lenguaje de gran tamaño (LLM), como ChatGPT, Llama o Gemini, para procesar relatos de desapariciones. La herramienta está orientada a brindar información detallada sobre personas, hechos o situaciones asociadas con desapariciones, a partir de datos integrados de diversas fuentes como información geográfica, tablas y bases de datos relacionadas con el universo de desapariciones.

Fase 1: Identificación del problema

En los últimos años, la información relacionada con desapariciones ha aumentado en volumen y complejidad. Muchas veces, las personas o familiares que buscan información no cuentan con un sistema centralizado y accesible para encontrar datos relevantes sobre estos hechos. Además, los relatos de desapariciones tienden a ser subjetivos y pueden carecer de detalles precisos que faciliten su análisis.

El reto principal es diseñar una solución que permita integrar diferentes fuentes de información, procesar relatos y ofrecer respuestas coherentes basadas en datos estructurados (tablas, mapas geográficos) y no estructurados (relatos).

Fase 2: Generación y conceptualización de ideas

Para abordar este problema, se consideraron varias soluciones, desde herramientas convencionales de búsqueda hasta la implementación de **Modelos LLM Multimodales**. Tras evaluar las opciones, se optó por esta última debido a su capacidad para procesar grandes volúmenes de datos en lenguaje natural y combinar información de diversas fuentes (relatos, tablas, geografía).

Los pasos clave en esta fase incluyen:

1. **Desarrollo de un modelo LLM** personalizado que pueda procesar textos largos y relatos sobre desapariciones.
2. **Integración de fuentes de datos** como bases de datos de personas desaparecidas, información geográfica de sitios de interés y tablas con información estructurada.
3. **Implementación de consultas interactivas** a través de un front-end en React que permita al usuario ingresar un relato y recibir una respuesta coherente.

Fase 3: Evaluación de viabilidad

- **Viabilidad técnica:** Se utilizó una combinación de tecnologías como **Python**, **Flask** para la API, y **React** para el front-end, lo que permite desarrollar el proyecto localmente para reducir costos iniciales.

- **Costos:** Dado que el despliegue en la nube implica altos costos, el MVP se ejecutará inicialmente de manera **local** para pruebas y mejoras. Esto permite ahorrar gastos significativos mientras se asegura la funcionalidad.
- **Recursos de datos:** El desafío aquí es integrar fuentes de datos relevantes, asegurando su calidad y coherencia. Se han definido sets de datos que incluyen relatos de desapariciones, tablas de personas y coordenadas geográficas de los hechos, lo que permite ofrecer respuestas precisas.

Fase 4: Prototipado y prueba

El **MVP** ya está en funcionamiento a nivel local, permitiendo que los usuarios ingresen un relato sobre un hecho de desaparición. La API, conectada a una base de datos MySQL, procesa la información y el modelo LLM responde con coincidencias de datos existentes. Los próximos pasos incluyen:

- **Pruebas de usuario** para mejorar la interfaz y la calidad de las respuestas.
- **Iteraciones** sobre el modelo LLM para mejorar la precisión y relevancia de las respuestas generadas.

Fase 5: Iteración y mejora

A medida que se recoja retroalimentación, se seguirán ajustando tanto el modelo como las fuentes de datos. Se planea:

- **Mejorar la integración de datos geográficos** para que las respuestas incluyan ubicaciones precisas en mapas interactivos.
- **Añadir más datos** de fuentes externas para enriquecer la calidad de las respuestas.

Conclusión:

El **MVP** ha logrado demostrar su viabilidad técnica y utilidad, proporcionando una herramienta de consulta avanzada para desapariciones, aunque aún en su fase de desarrollo. Con un enfoque en la integración de fuentes de datos múltiples y la implementación de modelos LLM multimodales, se está en el camino de desarrollar una solución innovadora que aporte valor en la búsqueda de información relevante sobre desapariciones.