



universidad
de león



Escuela de Ingenierías Industrial, Informática y Aeroespacial

GRADO EN INGENIERÍA INFORMÁTICA

SISTEMAS DE INFORMACIÓN DE GESTION Y BUSINESS INTELLIGENCE MEMORIA

Implementación de un chatGPT con datos propios

Autor: Jose María Villafañe Marcos

(enero, 2024)

Contenido

Introducción	3
Descripción del problema	3
Motivos por los que he elegido esta elección de estudio	3
Herramientas utilizadas.....	4
Proyecto realizado.....	5
DAFO.....	6
Lecciones aprendidas	7
Desarrollos a futuro.....	7
BIBLIOGRAFIA.....	7

Introducción

Esta asignatura nos ha servido para profundizar de lleno en conceptos relacionados con la inteligencia artificial generativa, ya que no es que vaya a ser el futuro, sino que estas herramientas son el presente en cualquier organización sea pública o privada a la hora de tomar decisiones que mejor se adapten a las circunstancias de cada ente.

A través de lecturas, videos, tweets y conversaciones que he tenido con profesionales que utilizan estas tecnologías, he aprendido muchos conceptos con los que antes de cursar esta asignatura nunca había tenido contacto, así como entrenar un sistema con datos propios.

En mi caso particular, intenté implementar un chatgpt utilizando llama-index, utilizando como bases de datos, documentos relacionados con licitadores públicos.

La idea consistía en realizar un recomendador de licitadores públicos para la administración pública, de tal forma que analizase las relaciones que tienen las organizaciones que acceden a un contrato público con las personas que trabajan en la administración pública.

También he tenido la oportunidad de conocer los grafos de conocimiento, y la manera en la que interactúan con el sistema recomendador.

Descripción del problema

A lo largo del recorrido de esta asignatura, he tenido claro en que ámbito quería aplicar estas herramientas de inteligencia artificial generativa. Debido a que trabajo en la administración autonómica de la Junta de Castilla Y León, tenía la oportunidad de realizar una aplicación cercana, desde el punto de vista laboral, a mi desempeño.

Sin duda, los ámbitos de aplicación de estas tecnologías son variados, tales como en selección de personal, selección de viajes, o bien como otro ejemplo, como recomendador de contraseñas, muy en auge hoy en día en ciberseguridad, tal y como se ha podido escuchar en diversas charlas a las que he asistido, como por ejemplo en el encuentro internaciones de seguridad de la información (enise), acontecido a mediados de octubre de 2023 en la ciudad de León.

Como conclusión a este punto, mi objetivo era crear un recomendador de licitadores públicos con datos propios, o bien con datos que se pudieran obtener abiertamente de las administraciones.

Motivos por los que he elegido esta elección de estudio

Quizá los recomendadores a los que más acostumbrados estemos sean recomendadores de series, de música, de productos a la hora de comprar por internet o de coches, pero se me ocurrió este objeto de estudio primero porque me gusta la temática y porque trabajo en la administración pública y también

porque hay muchos casos de corrupción dentro de la misma, tal y como se ha podido comprobar a lo largo de estos años.

Escucho a mucha gente, familiares, amigos que siempre tienen dudas a la hora de fiarse de la administración pública, cuando se tiene que hacer un trámite o bien cuando se tiene que formalizar un pago. Después de investigar la manera en la que se elige un licitador público a la hora de formalizar un contrato público menor, muchas veces me he preguntado cómo no se plantea la posibilidad de implementar un sistema de recomendación, creo que es algo que aporta mucho valor y a la vez ofrece un mejor servicio a la administración

Herramientas utilizadas

- Bases de datos

Debido a que trabajo en la administración autonómica de la jcyl, me dediqué a buscar información de contratos menores y licitadores públicos a través del sitio web de la Junta de Castilla y León, pero únicamente encontré datos relacionados con el sacyl.

También decidí buscar datos en el sitio web de kaggle, pero la búsqueda no fue satisfactoria, ya que no había ninguna base de datos que se adaptara a lo que yo buscaba.



Para la implementación del proyecto, he usado el entorno de visual studio code y Python.

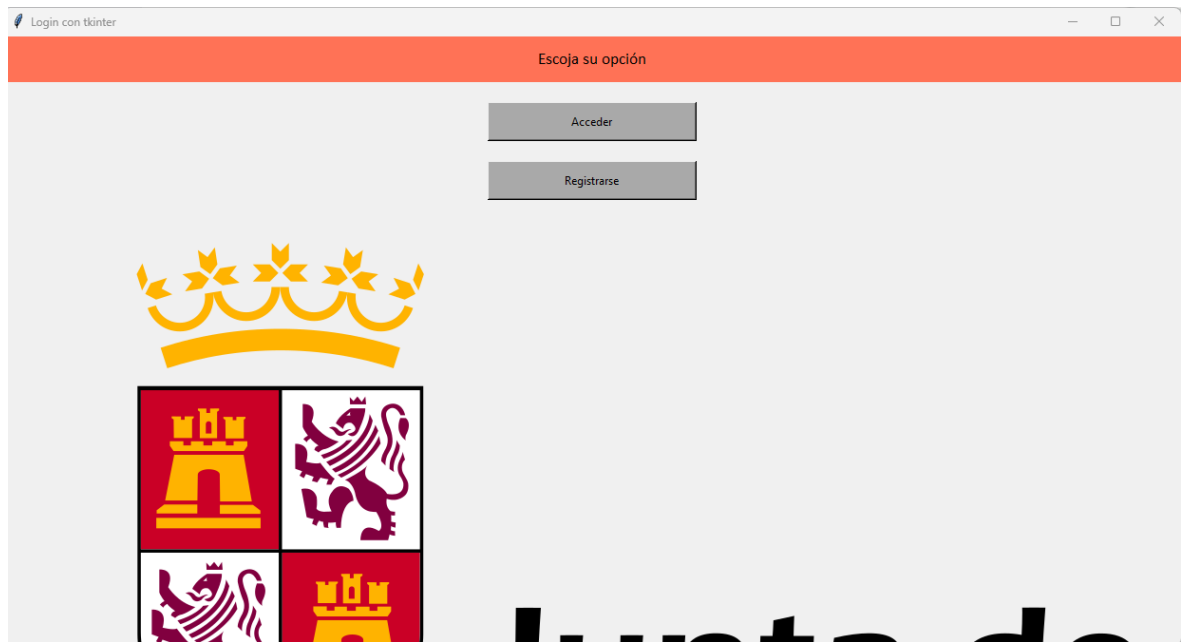


Acerca de Python, he tenido que mirarme tutoriales, ya que mis conocimientos en este lenguaje de programación eran muy limitados.

Proyecto realizado

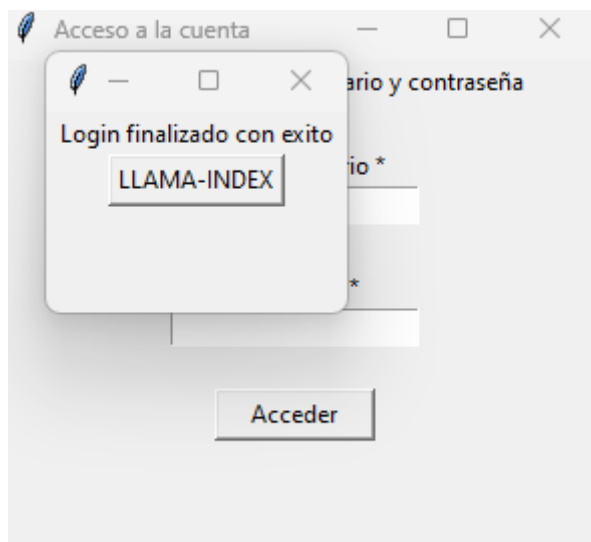
Lamentablemente no he podido alcanzar mi objetivo deseado.

Pantalla inicial



En esta primera pantalla, se puede acceder login para acceder a la aplicación, o bien registrarse como un nuevo usuario en el sistema.

Una vez, que se ha accedido correctamente al sistema, aparece la siguiente ventana, en la cual se hará click en el botón llama_index, para acceder al recomendador propiamente dicho





Chatbot con informacion propia

responde preguntas relacionadas con la informacion

Código fuente donde utilizo llama-index

```
def llamaindex():
    borrar_exito_login()
    borrar_pantalla_inicio()

    st.header("Chatbot con informacion propia")
    reader = SimpleDirectoryReader(input_dir='./data', recursive = True)
    docs = reader.load_data()
    service_context = ServiceContext.from_defaults(llm=OpenAI(model="gpt-3.5-turbo", temperature=0.5, system_prompt = "eres una machine learning y tu objetivo es responder preguntas"))
    index = VectorStoreIndex.from_documents(docs, service_context=service_context)
    #print(index)
    query = st.text_input("responde preguntas relacionadas con la informacion")

    if query:
        chat_engine = index.as_chat_engine(chat_mode="condense_question", verbose = True)
        response = chat_engine.chat(query)
        st.write(response.response)
```

DAFO



Lecciones aprendidas

Desde el inicio del curso en esta asignatura, se me ha enseñado que la constancia es fundamental.

El campo de la inteligencia artificial generativa está en constante cambio, por ese motivo la constancia, perseverancia y no caer en el desánimo, son actitudes importantes que se han de tener en cuenta.

Por otra parte, también me he dado cuenta que el fracaso no se debe tomar como un castigo, sino como una fase del aprendizaje.

Desarrollos a futuro

Convertir la aplicación local, a aplicación web.

Ser capaz de leer datos, no solo propios sino de internet.

BIBLIOGRAFIA

Kaggle: <https://www.kaggle.com/>

Python: <https://docs.python.org/3/>

Datos de Castilla y León: <https://datosabiertos.jcyl.es/web/es/datos-abiertos-castilla-leon.html>

Arxiv: <https://arxiv.org/abs/2401.03605>

Mis enlaces:

Github: <https://github.com/infjmv00/SistemasInformacionBI.git>

OSF: <https://osf.io/74cj6/>