



Maestría en Inteligencia Artificial Aplicada

Proyecto Integrador

Resumen Ejecutivo

Alumno	Matrícula
Johanna Rodríguez Jaramillo	A01794010
Katherine Vanegas Salamanca	A01794113
Marcos Chávez Chávez	A01688507

Profesores Titulares

Dra. Grettel Barceló Alonso / Luis Eduardo Falcón Morales

Profesor Asistente

Horacio Martinez Alfaro

16 de Junio de 2024

1. Síntesis del Problema

Viakable, empresa de Xignux, enfrenta un desafío significativo con la gestión y análisis de cientos de documentos normativos relacionados con la fabricación de cables. La diversidad y complejidad de estos documentos dificultan la rápida localización de información crítica, lo que impacta negativamente en la eficiencia operativa y el cumplimiento normativo. El proyecto se centró en la implementación de una solución de inteligencia artificial que utilice la suite Azure OpenAI para mejorar la capacidad de búsqueda y recuperación de información en estos documentos, optimizando así los procesos de calidad y cumplimiento.

2. Hallazgos más Importantes del Análisis Exploratorio de Datos

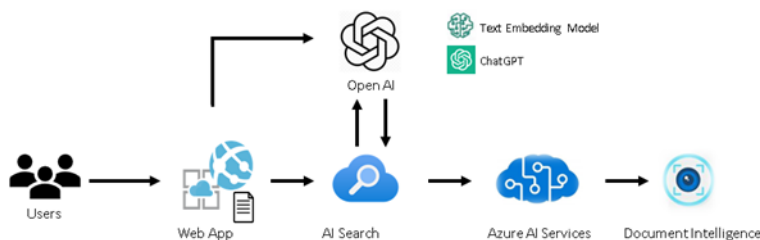
Durante el análisis exploratorio de datos, se identificaron patrones clave en los documentos normativos. La mayoría de los documentos varía significativamente en tamaño, desde 0.09 MB hasta más de 23 MB, y en longitud, con algunos documentos que tienen solo 5 páginas mientras que otros exceden las 150 páginas. La cantidad de imágenes y contenido gráfico también varía ampliamente, lo que resalta la necesidad de herramientas capaces de manejar tanto texto como datos tabulados.

Se observaron diversas estructuras en los documentos, incluyendo títulos, listas, tablas y párrafos, que son cruciales para la extracción y organización de la información. Este análisis permitió definir los criterios para la segmentación y extracción de entidades relevantes, como términos técnicos y pasos de fabricación, que facilitan la búsqueda y recuperación de información específica.

3. Modelos Generados y Razones de la Elección del Modelo Final

Se exploraron varios modelos para abordar el problema, incluyendo modelos pre entrenados de huggingface, Azure AI Studio y modelos avanzados de Azure OpenAI como GPT-4. Se decidió utilizar Azure AI Studio debido a su capacidad para crear y entrenar modelos personalizados, y por su integración robusta con otros servicios de Azure como Azure Machine Learning y Azure Data Lake. Esta plataforma ofreció un entorno colaborativo

y herramientas avanzadas para la gestión y optimización de modelos, lo que resultó ideal para las necesidades del proyecto.



El modelo final seleccionado fue un modelo de embeddings de Azure OpenAI, que permite la conversión de datos en vectores de alta dimensionalidad. Este modelo se eligió por su capacidad para capturar la semántica de los documentos y representar de manera eficiente la similitud entre términos y textos, facilitando la recuperación de información relevante de manera rápida y precisa.

4. Recomendaciones Clave para Implementar la Solución

Para implementar la solución de manera efectiva, se recomienda:

- Automatizar la Segmentación y Extracción de Datos: Utilizar herramientas como Azure AI Document Intelligence para la segmentación de documentos y la extracción de entidades y términos relevantes.
- Optimizar los Modelos de Búsqueda: Ajustar los parámetros del modelo de embeddings y mejorar la calidad de los prompts para asegurar respuestas más específicas y relevantes.
- Capacitación Continua del Modelo: Realizar entrenamientos periódicos del modelo con nuevos datos para mantener su precisión y relevancia.
- Implementar un Sistema de Feedback: Incluir un mecanismo para que los usuarios finales puedan proporcionar retroalimentación sobre la precisión y utilidad de las respuestas, lo que permitirá mejoras continuas.



Inicia una conversación

Consulta información en la base de conocimientos del bot

A screenshot of a chat interface. It shows a white input box with the placeholder text 'Escribe una pregunta...' and a blue arrow icon on the right. Above the input box is a small grey speech bubble icon. The entire input area is set against a light blue background.

5. Análisis Costo-Beneficio

La implementación de la solución en Azure ofrece beneficios significativos en términos de eficiencia y costo. La inversión inicial en la configuración de la infraestructura y el desarrollo del modelo se compensa con la reducción de tiempo y recursos necesarios para la búsqueda y análisis de documentos normativos. Esto no solo mejora la productividad, sino que también reduce el riesgo de incumplimiento y potenciales sanciones.

El costo de la infraestructura de Azure es escalable y puede ajustarse según la demanda, lo que ofrece flexibilidad y control sobre los gastos operativos. La automatización de procesos reduce la necesidad de intervención manual, lo que se traduce en ahorros adicionales a largo plazo. Para nuestro caso puntual y la implementación realizada el costo total mensual ascendió a USD42,69 por el uso de un recurso en específico dentro de la plataforma llamado azure search (búsqueda de texto completo tradicional y la búsqueda de similitud de vectores de próxima generación).

De acuerdo con una estimación inicial de los beneficios generados con esta solución, se puede lograr una reducción de 2 horas semanales por usuario (10 iniciales), es decir, 20 horas equivalente a 0.5 recurso.

Por otro lado, si se incorporarán otras 3,500 normas al desarrollo, se podría expandir la solución hasta 70 usuarios, lo que podría potenciar su valor a 140 horas semanales, equivalente a 3.5 recursos.

6. Riesgos y Desafíos de la Solución

Entre los principales riesgos y desafíos que se identificaron se incluyen:

Dependencia de la Infraestructura de Azure: La solución depende en gran medida de la disponibilidad y rendimiento de los servicios de Azure, lo que podría ser un riesgo si se producen interrupciones o cambios en la plataforma.

Manejo de Datos Sensibles: La gestión de documentos normativos puede implicar la manipulación de información sensible, por lo que se debe asegurar que se cumplan todas las normativas de privacidad y seguridad de datos.

Actualización Continua de Modelos: Es crucial mantener los modelos actualizados con los últimos datos y normativas para asegurar su precisión y relevancia.

Resistencia al Cambio: La implementación de nuevas tecnologías puede encontrar resistencia entre los usuarios finales, por lo que se debe planificar una estrategia de cambio adecuada para asegurar la adopción de la solución.