



BASES DE DATOS PARA IA

TP n° 4 Bases de Vectores

Ejercicio 1: Crear y Consultar

Crea un documento nuevo con el texto: "La gravedad es una fuerza fundamental en el universo".

Realiza una consulta para encontrar documentos similares a la frase: "Fuerzas naturales que gobiernan el cosmos".

¿Qué tipo de operación CRUD usaste primero?

¿Cuál es el ID del documento creado? (Usa `collection.peek()`).

Explica por qué el documento aparece como coincidencia en la consulta.

Ejercicio 2: Actualizar y Eliminar

Crea un documento con el texto: "Redes neuronales imitan procesos cerebrales".

Encuentra su ID, luego actualízalo por: "Deep Learning es un subcampo de IA con redes neuronales profundas".

Elimina el documento original usando el mismo ID.

¿Qué ocurre si intentas actualizar un documento eliminado?

¿Cómo se garantiza la integridad de los datos en las operaciones?

Ejercicio 3: Errores y Validación

Intenta eliminar un documento con el ID "no_existe".

Crea un nuevo documento sin usar `get_embeddings()` (ejemplo: enviar un array vacío).

¿Qué errores ocurren en cada caso?

¿Cómo prevenir estos problemas?

Ejercicio 4: Consultas por Tema

Crea tres documentos:

"Python es ideal para ciencia de datos"

"JavaScript maneja interacciones web dinámicas"

"Java se usa en aplicaciones empresariales"

Realiza una consulta por la frase: "Lenguajes backend y frontend".

¿Cuáles documentos aparecen como coincidencias?

Analiza cómo las palabras clave influyen en los resultados.

Ejercicio 5: Eliminaciones Selectivas

Crea dos documentos con tópicos opuestos (ejemplo: uno sobre "Ecología" y otro sobre "Minería de datos").

Elimina el documento sobre ecología usando su ID.

Verifica que solo quede un documento.

¿Cómo usar `collection.count()` para confirmar?



Ejercicio 6: Actualización en Cadena

Crea un documento con el texto "IA y ética son temas actuales".

Actualízalo dos veces:

Primera actualización: "Ética en IA aborda dilemas morales"

Segunda actualización: "El uso de IA debe estar regulado por principios éticos".

¿Cómo sigue evolucionando el contenido del documento con cada update?

Ejercicio 7: Consultas Ambiguas

Crea dos documentos similares pero no idénticos:

"Blockchain garantiza transacciones seguras"

"Criptomonedas usan tecnología blockchain".

Realiza una consulta por "Seguridad en finanzas digitales".

¿Ambos documentos aparecen? Explica la similitud semántica.

Ejercicio 8: Reiniciar Base de Datos

Elimina todos los documentos usando delete() con el parámetro correcto.

Vuelve a cargar la base desde cero con crear_base_datos().

¿Cuántos documentos hay ahora? (Usa collection.count()).

Ejercicio 9: Errores de Inserción

Intenta crear un documento sin texto (ejemplo: cadena vacía).

Corrige el error usando la función get_embeddings().

¿Qué lección aprendemos sobre los requisitos de ChromaDB?

Ejercicio opcional: Retrieval-Augmented Generation

Intente levantar un llm en un endpoint local con ollama o LMStudio. Utilice el llm para generar respuestas aumentadas con contexto similar a una consulta dada.

