# 向量存储的原理和使用

#### 向量存储的原理

在处理非结构化数据的存储和检索过程中，最常见的方式之一是将其嵌入，并存储生成的嵌入向量。然后在查询时，将非结构化查询嵌入，并检索与嵌入查询”最相似”的嵌入向量。向量存储就是用来存储嵌入数据和执行向量搜索的。

#### 如何使用向量存储

下面我们将通过以下步骤介绍如何使用向量存储。

#### 1.环境设置

本文展示了与向量存储相关的基本功能。使用向量存储的关键部分是创建要放入其中的向量，这通常是通过嵌入来创建的。因此，在深入了解这个内容之前，建议你先熟悉文本嵌入模型的接口。

本文使用的是FAISS向量数据库，该数据库使用了Facebook AI Similarity Search (FAISS)库。

pip install faiss-cpu

我们需要使用OpenAIEmbeddings，所以我们需要获取OpenAI API Key。

import os  
import getpass  
  
os.environ['OPENAI\_API\_KEY'] = getpass.getpass('OpenAI API Key:')

from langchain.document\_loaders import TextLoader  
from langchain.embeddings.openai import OpenAIEmbeddings  
from langchain.text\_splitter import CharacterTextSplitter  
from langchain.vectorstores import FAISS  
  
  
raw\_documents = TextLoader('../../../state\_of\_the\_union.txt').load()  
text\_splitter = CharacterTextSplitter(chunk\_size=1000, chunk\_overlap=0)  
documents = text\_splitter.split\_documents(raw\_documents)  
  
embeddings = OpenAIEmbeddings()  
db = FAISS.from\_documents(documents, embeddings)

#### 2.相似度搜索

query = "What did the president say about Ketanji Brown Jackson"  
docs = db.similarity\_search(query)  
print(docs[0].page\_content)

Tonight. I call on the Senate to: Pass the Freedom to Vote Act. Pass the John Lewis Voting Rights Act. And while you’re at it, pass the Disclose Act so Americans can know who is funding our elections.  
  
 Tonight, I’d like to honor someone who has dedicated his life to serve this country: Justice Stephen Breyer—an Army veteran, Constitutional scholar, and retiring Justice of the United States Supreme Court. Justice Breyer, thank you for your service.  
  
 One of the most serious constitutional responsibilities a President has is nominating someone to serve on the United States Supreme Court.  
  
 And I did that 4 days ago, when I nominated Circuit Court of Appeals Judge Ketanji Brown Jackson. One of our nation’s top legal minds, who will continue Justice Breyer’s legacy of excellence.

通过向量进行相似度搜索：你也可以使用similarity\_search\_by\_vector方法来搜索与给定的嵌入向量相似的文档，这个方法接受一个嵌入向量作为参数，而不是一个字符串。

embedding\_vector = embeddings.embed\_query(query)  
docs = db.similarity\_search\_by\_vector(embedding\_vector)

#### 3.异步操作

向量存储通常作为一个需要一些IO操作的单独服务运行，因此可能会异步调用。这有助于提升性能，因为你不需要浪费时间等待外部服务的响应。如果你使用的是异步框架，例如FastAPI，这可能非常重要。

Langchain支持在向量存储上进行异步操作。所有的方法都可以使用其异步对应方法调用，这些方法前缀为”a”，表示异步。

Qdrant是一个向量存储，支持所有的异步操作，因此将在本文中使用。

pip install qdrant-client

from langchain.vectorstores import Qdrant

#### 4.异步创建向量存储

db = await Qdrant.afrom\_documents(documents, embeddings, "http://localhost:6333")

#### 5.相似度搜索

通过向量进行相似度搜索。

query = "What did the president say about Ketanji Brown Jackson"  
docs = await db.asimilarity\_search(query)  
print(docs[0].page\_content)

Tonight. I call on the Senate to: Pass the Freedom to Vote Act. Pass the John Lewis Voting Rights Act. And while you’re at it, pass the Disclose Act so Americans can know who is funding our elections.  
  
 Tonight, I’d like to honor someone who has dedicated his life to serve this country: Justice Stephen Breyer—an Army veteran, Constitutional scholar, and retiring Justice of the United States Supreme Court. Justice Breyer, thank you for your service.  
  
 One of the most serious constitutional responsibilities a President has is nominating someone to serve on the United States Supreme Court.  
  
 And I did that 4 days ago, when I nominated Circuit Court of Appeals Judge Ketanji Brown Jackson. One of our nation’s top legal minds, who will continue Justice Breyer’s legacy of excellence.

#### 6.最大边际相关性搜索（MMR）

最大边际相关性是为查询的相似性和所选文档的多样性进行优化。在异步API中也支持这个功能。

query = "What did the president say about Ketanji Brown Jackson"  
found\_docs = await qdrant.amax\_marginal\_relevance\_search(query, k=2, fetch\_k=10)  
for i, doc in enumerate(found\_docs):  
 print(f"{i + 1}.", doc.page\_content, "\n")

1. Tonight. I call on the Senate to: Pass the Freedom to Vote Act. Pass the John Lewis Voting Rights Act. And while you’re at it, pass the Disclose Act so Americans can know who is funding our elections.  
  
Tonight, I’d like to honor someone who has dedicated his life to serve this country: Justice Stephen Breyer—an Army veteran, Constitutional scholar, and retiring Justice of the United States Supreme Court. Justice Breyer, thank you for your service.  
  
One of the most serious constitutional responsibilities a President has is nominating someone to serve on the United States Supreme Court.  
  
And I did that 4 days ago, when I nominated Circuit Court of Appeals Judge Ketanji Brown Jackson. One of our nation’s top legal minds, who will continue Justice Breyer’s legacy of excellence.  
  
2. We can’t change how divided we’ve been. But we can change how we move forward—on COVID-19 and other issues we must face together.  
  
I recently visited the New York City Police Department days after the funerals of Officer Wilbert Mora and his partner, Officer Jason Rivera.  
  
They were responding to a 9-1-1 call when a man shot and killed them with a stolen gun.  
  
Officer Mora was 27 years old.  
  
Officer Rivera was 22.  
  
Both Dominican Americans who’d grown up on the same streets they later chose to patrol as police officers.  
  
I spoke with their families and told them that we are forever in debt for their sacrifice, and we will carry on their mission to restore the trust and safety every community deserves.  
  
I’ve worked on these issues a long time.  
  
I know what works: Investing in crime preventionand community police officers who’ll walk the beat, who’ll know the neighborhood, and who can restore trust and safety.

以上便是使用LangChain进行向量存储的基本步骤和方法，相信通过以上的介绍，你已经对如何使用LangChain的向量存储有了基本的理解和掌握。