Chroma

Contents

- Similarity search with score
- Persistance
- Retriever options

This notebook shows how to use functionality related to the Chroma vector database.

```python
from langchain.embeddings.openai import OpenAIEmbeddings
from langchain.text_splitter import CharacterTextSplitter
from langchain.vectorstores import Chroma
from langchain.document_loaders import TextLoader
```

```python
from langchain.document_loaders import TextLoader
loader = TextLoader('../../../state_of_the_union.txt')
documents = loader.load()
text_splitter = CharacterTextSplitter(chunk_size=1000, chunk_overlap=0)
docs = text_splitter.split_documents(documents)

embeddings = OpenAIEmbeddings()
```

```python
db = Chroma.from_documents(docs, embeddings)

query = "What did the president say about Ketanji Brown Jackson"
docs = db.similarity_search(query)
```

```
Using embedded DuckDB without persistence: data will be transient
```

```python
print(docs[0].page_content)
```

Tonight. I call on the Senate to: Pass the Freedom to Vote Act. Pass the John

Skip to main content

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.

Similarity search with score

```
docs = db.similarity_search_with_score(query)
```

```
docs[0]
```

```
(Document(page_content='In state after state, new laws have been passed, not only
to suppress the vote, but to subvert entire elections. \n\nWe cannot let this
happen. \n\nTonight. 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. \n\nTonight, 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. \n\nOne of the most
serious constitutional responsibilities a President has is nominating someone to
serve on the United States Supreme Court. \n\nAnd 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.', lookup_str='', metadata={'source': '../../state_of_the_union.txt'},
lookup_index=0),
 0.3913410007953644)
```

Persistance

The below steps cover how to persist a ChromaDB instance

Skip to main content

Initialize PeristedChromaDB

Create embeddings for each chunk and insert into the Chroma vector database. The persist_directory argument tells ChromaDB where to store the database when it's persisted.

```python
# Embed and store the texts
# Supplying a persist_directory will store the embeddings on disk
persist_directory = 'db'

embedding = OpenAIEmbeddings()
vectordb = Chroma.from_documents(documents=docs, embedding=embedding,
persist_directory=persist_directory)
```

```
Running Chroma using direct local API.
No existing DB found in db, skipping load
No existing DB found in db, skipping load
```

Persist the Database

In a notebook, we should call persist() to ensure the embeddings are written to disk. This isn't necessary in a script - the database will be automatically persisted when the client object is destroyed.

```python
vectordb.persist()
vectordb = None
```

```
Persisting DB to disk, putting it in the save folder db
PersistentDuckDB del, about to run persist
Persisting DB to disk, putting it in the save folder db
```

Load the Database from disk, and create the chain

Be sure to pass the same persist_directory and embedding_function as you did when you instantiated the database. Initialize the chain we will use for question answering.

Skip to main content

```
# Now we can load the persisted database from disk, and use it as normal.
vectordb = Chroma(persist_directory=persist_directory,
embedding_function=embedding)
```

```
Running Chroma using direct local API.
loaded in 4 embeddings
loaded in 1 collections
```

Retriever options

This section goes over different options for how to use Chroma as a retriever.

MMR

In addition to using similarity search in the retriever object, you can also use `mmr`.

```
retriever = db.as_retriever(search_type="mmr")
```

```
retriever.get_relevant_documents(query)[0]
```

```
Document(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. \n\nTonight, 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.
\n\nOne of the most serious constitutional responsibilities a President has is
nominating someone to serve on the United States Supreme Court. \n\nAnd 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.', metadata={'source': '../../../state_of_the_union.txt'})
```