FAISS

Contents

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

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

```
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()
```

```
db = FAISS.from_documents(docs, embeddings)

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
```

Skip to main content

```
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

There are some FAISS specific methods. One of them is `similarity_search_with_score`, which allows you to return not only the documents but also the similarity score of the query to them.

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

```
docs_and_scores[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),
 0.3914415)
```

It is also possible to do a search for documents similar to a given embedding vector using `similarity_search_by_vector` which accepts an embedding vector as a parameter instead of a string.

```
embedding_vector = embeddings.embed_query(query)
docs_and_scores = db.similarity_search_by_vector(embedding_vector)
```

Skip to main content

Saving and loading

You can also save and load a FAISS index. This is useful so you don’t have to recreate it everytime you use it.

```
db.save_local("faiss_index")
```

```
new_db = FAISS.load_local("faiss_index", embeddings)
```

```
docs = new_db.similarity_search(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)
```

Merging

You can also merge two FAISS vectorstores

```
db1 = FAISS.from_texts(["foo"], embeddings)
db2 = FAISS.from_texts(["bar"], embeddings)
```

Skip to main content

```
db1.docstore._dict
```

```
{'e0b74348-6c93-4893-8764-943139ec1d17': Document(page_content='foo',
lookup_str='', metadata={}, lookup_index=0)}
```

```
db2.docstore._dict
```

```
{'bdc50ae3-a1bb-4678-9260-1b0979578f40': Document(page_content='bar',
lookup_str='', metadata={}, lookup_index=0)}
```

```
db1.merge_from(db2)
```

```
db1.docstore._dict
```

```
{'e0b74348-6c93-4893-8764-943139ec1d17': Document(page_content='foo',
lookup_str='', metadata={}, lookup_index=0),
 'd5211050-c777-493d-8825-4800e74cfdb6': Document(page_content='bar',
lookup_str='', metadata={}, lookup_index=0)}
```