

OpenSearch

Contents

- similarity_search using Approximate k-NN Search with Custom Parameters
- similarity_search using Script Scoring with Custom Parameters
- similarity_search using Painless Scripting with Custom Parameters

This notebook shows how to use functionality related to the OpenSearch database.

To run, you should have the opensearch instance up and running: [here](#) `similarity_search` by default performs the Approximate k-NN Search which uses one of the several algorithms like lucene, nmslib, faiss recommended for large datasets. To perform brute force search we have other search methods known as Script Scoring and Painless Scripting. Check [this](#) for more details.

```
from langchain.embeddings.openai import OpenAIEmbeddings
from langchain.text_splitter import CharacterTextSplitter
from langchain.vectorstores import OpenSearchVectorSearch
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()
```

```
docsearch = OpenSearchVectorSearch.from_texts(texts, embeddings,
opensearch_url="http://localhost:9200")

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

[Skip to main content](#)

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

similarity_search using Approximate k-NN Search with Custom Parameters

```
docsearch = OpenSearchVectorSearch.from_texts(texts, embeddings,
opensearch_url="http://localhost:9200", engine="faiss", space_type="innerproduct",
ef_construction=256, m=48)

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

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

similarity_search using Script Scoring with Custom Parameters

```
docsearch = OpenSearchVectorSearch.from_texts(texts, embeddings,
opensearch_url="http://localhost:9200", is_appx_search=False)

query = "What did the president say about Ketanji Brown Jackson"
docs = docsearch.similarity_search("What did the president say about Ketanji Brown
Jackson", k=1, search_type="script_scoring")
```

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

similarity_search using Painless Scripting with Custom Parameters

```
docsearch = OpenSearchVectorSearch.from_texts(texts, embeddings,
opensearch_url="http://localhost:9200", is_appx_search=False)
filter = {"bool": {"filter": {"term": {"text": "smuggling"}}}}
```

[Skip to main content](#)

```
Jackson", search_type="painless_scripting", space_type="cosineSimilarity",  
pre_filter=filter)
```

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