

```
%%time

! pip install -qq -U langchain tiktoken pypdf chromadb faiss-gpu
unstructured openai
! pip install -qq -U transformers InstructorEmbedding
sentence_transformers

huggingface/tokenizers: The current process just got forked, after
parallelism has already been used. Disabling parallelism to avoid
deadlocks...
To disable this warning, you can either:
  - Avoid using `tokenizers` before the fork if possible
  - Explicitly set the environment variable
TOKENIZERS_PARALLELISM=(true | false)
huggingface/tokenizers: The current process just got forked, after
parallelism has already been used. Disabling parallelism to avoid
deadlocks...
To disable this warning, you can either:
  - Avoid using `tokenizers` before the fork if possible
  - Explicitly set the environment variable
TOKENIZERS_PARALLELISM=(true | false)
CPU times: user 302 ms, sys: 129 ms, total: 431 ms
Wall time: 26.8 s
```

## Imports

```
import openai
### Multi-document retriever
from langchain.vectorstores import Chroma, FAISS
from langchain.text_splitter import RecursiveCharacterTextSplitter

from langchain.chains import RetrievalQA, VectorDBQA
from langchain.document_loaders import DirectoryLoader
from langchain.embeddings import SentenceTransformerEmbeddings

from langchain.callbacks.streaming_stdout import
StreamingStdOutCallbackHandler
from InstructorEmbedding import INSTRUCTOR
from langchain.embeddings import HuggingFaceInstructEmbeddings

directory = '/kaggle/input/donald-trumps-rallies'

def load_docs(directory):
    loader = DirectoryLoader(directory)
    documents = loader.load()
    return documents

documents = load_docs(directory)
len(documents)
```

35

```
def split_docs(documents, chunk_size=1000, chunk_overlap=20):  
    text_splitter =  
RecursiveCharacterTextSplitter(chunk_size=chunk_size,  
chunk_overlap=chunk_overlap)  
    docs = text_splitter.split_documents(documents)  
    return docs
```

```
docs = split_docs(documents)  
print(len(docs))
```

2786

### Embedding Text Using Langchain

```
embeddings = SentenceTransformerEmbeddings(model_name="all-MiniLM-L6-  
v2")
```

```
{"model_id": "6df58ea1459349b5bb1f6724f46de157", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "48b3162a7c18498b92c968e5bcdb03e1", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "ae2e2af18e8a4009af536901b52912e5", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "68a2e7cd1144420cacdd35aafa66024e", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "bb3893f3a43c48e896ea0e87035b30ef", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "45d5630886ce44bab9823a967f2121de", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "aa183f4b8f2d49ddb0cc0d85cecada89", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "dcbeed3001654c69a09545206539354c", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "94afe70a1ba54f7b9c822e8361098f2c", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "932928afdbec4c75a17773cc3225e8fb", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "c0dab0a6287e44f69764562da19fe729", "version_major": 2, "vers  
ion_minor": 0}
```

```
{"model_id": "3f7c7c3db5fa4a5cad358c38ae911902", "version_major": 2, "version_minor": 0}

{"model_id": "70e112389c8c4348a6a37e6b67292ff7", "version_major": 2, "version_minor": 0}

{"model_id": "9e20e5afeb054dcfbdc21b56b09a764f", "version_major": 2, "version_minor": 0}
```

## Creating Vector Store with Chroma DB

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

{"model_id": "1c94d059978f497d995494b52f309157", "version_major": 2, "version_minor": 0}
```

## Retrieving Semantically Similar Documents

```
query = "Who are considered allies of Donald Trump?"
matching_docs = db.similarity_search(query)

matching_docs[0]

{"model_id": "c14a3f25305b46ca8594e520a738579c", "version_major": 2, "version_minor": 0}

Document(page_content='very smart guy. We get along. Getting along with foreign powers is not a bad thing. We\'re trying to teach that to the media and these idiots that ran our government for so many years. These fools. These fools. You know, you got nuclear weapons, you got big armies, big navies, "He gets along with Putin. That\'s a terrible thing." No, it\'s a good thing. If I get along, that\'s good. He likes me. I like him. Not so bad. "He gets along with Kim Jong-un." ... Not so bad. He gets along with Kim Jong-Un.', metadata={'source': '/kaggle/input/donald-trumps-rallies/FreelandSep10_2020.txt'})
```

## Persistence in Chroma DB

```
persist_directory = "chroma_db"

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

```
vectoradb.persist()
```

```
{"model_id": "69d0df91029b44fa97a1c661c1fa6608", "version_major": 2, "version_minor": 0}
```

## Using OpenAI Large Language Models (LLM) with Chroma DB

```
import os
```

```
os.environ["OPENAI_API_KEY"] = "key"
```

```
from langchain.chat_models import ChatOpenAI
```

```
model_name = "gpt-3.5-turbo"
```

```
llm = ChatOpenAI(model_name=model_name)
```

```
import os
```

```
os.environ["OPENAI_API_KEY"] = ""
```

```
from langchain.chat_models import ChatOpenAI
```

```
model_name = "gpt-3.5-turbo"
```

```
llm = ChatOpenAI(model_name=model_name)
```

## Extracting Answers from Documents

```
from langchain.chains.question_answering import load_qa_chain  
chain = load_qa_chain(llm, chain_type="stuff", verbose=True)
```

```
query = "What were the foreign policies of Donald Trump?"
```

```
matching_docs = db.similarity_search(query)
```

```
answer = chain.run(input_documents=matching_docs, question=query)  
answer
```

```
{"model_id": "0feff72e9cea4fda80c7518b02f218bb", "version_major": 2, "version_minor": 0}
```

```
> Entering new StuffDocumentsChain chain...
```

```
> Entering new LLMChain chain...
```

```
Prompt after formatting:
```

```
System: Use the following pieces of context to answer the users
```

question.

If you don't know the answer, just say that you don't know, don't try to make up an answer.

-----

He never did anything about it. Biden's supported every globalist attack on the American worker let's make every country of the world rich but ourselves, including NAFTA, the disaster of NAFTA. He wanted to go and TPP would have ruined our automobile industry. Korea, I renegotiated the deal. I took a horrible deal. That was a Hillary Clinton special. That was a Hillary Clinton special. She said, "We have to make this deal with South Korea. It's going to mean 250,000 jobs." And she was right for South Korea, 250,000. And China's entry into the World Trade Organization was supported by sleepy Joe Biden. That was one of the worst deals economically in the history of our country. When China joined the World Trade Organization, they were flat line for decades. They became like a rocket ship. He voted for the war in Iraq. He voted for mass amnesty for illegal aliens. He supports sanctuary cities and now Biden wants to end immigration enforcement and he wants to require you to provide free

went to all parts of the world, but a lot went to Beijing. He shattered the plants in Chicago and sent them to Shanghai. Unlike the previous administration, under my administration we put America first. And actually, if we weren't interrupted with constant investigations that went nowhere. And nobody has been through what I've been through as president of the United States, nobody. They used to say Abraham Lincoln got the worst press. I said, "There's no way he got worse than me."

the killing of Soleimani. You remember that one? He oversaw the rise of ISIS and he cheered the rise of China as a positive development for America and for the world. That's not a positive development. Before they went into the WTO, China was flat lined. For years and years and decades, flat lined. And then they got into the WTO and a lot of bad things happened. First China's stole our jobs, plundered our intellectual property and unleashed the virus. And now China is trying to just hug and kiss Joe Biden.

he thought it would be good politically, that didn't work out too well. Just as he opposed the raid that killed Osama bin Laden. He opposed it, you know that. Biden is always on the wrong side of history, as said by people that are with him and worked with him. He never made a correct foreign policy decision. Biden is a puppet for China. Son walked out with 1.5 billion. I think it's a little bit... Allowing them to rip off America for many years. Now they're paying us billions and billions of dollars.

Human: What were the foreign policies of Donald Trump?

> Finished chain.

> Finished chain.

"Donald Trump's foreign policies focused on putting America first and prioritizing American workers. He criticized globalist policies such as NAFTA and the Trans-Pacific Partnership (TPP), which he believed were detrimental to the American workforce. He renegotiated trade deals, such as the one with South Korea, in an effort to benefit the United States. Trump also criticized China's entry into the World Trade Organization and accused China of stealing American jobs and intellectual property. He took a tough stance on China, implementing tariffs and engaging in a trade war. Trump also highlighted his opposition to the Iraq War and criticized Joe Biden's support for it. Overall, Trump's foreign policies were characterized by a focus on protecting American interests and challenging what he saw as unfair trade practices."

```
from langchain.chains import RetrievalQA
retrieval_chain = RetrievalQA.from_chain_type(llm, chain_type="stuff",
retriever=db.as_retriever())
retrieval_chain.run(query)
```

```
{"model_id": "3ea523decc9a42e18060bd227cab17b3", "version_major": 2, "version_minor": 0}
```

"From the given context, it is clear that Donald Trump focused on putting America first in his foreign policies. He criticized globalist approaches like NAFTA and the Trans-Pacific Partnership (TPP), which he believed harmed American workers. He renegotiated deals like the one with South Korea to protect American jobs. Trump also criticized China's entry into the World Trade Organization and accused them of stealing American jobs and intellectual property. He implemented tariffs on Chinese goods and aimed to hold China accountable for its actions. Additionally, Trump opposed the war in Iraq and supported strong immigration enforcement, while criticizing Biden's stance on amnesty and sanctuary cities. These are some of the foreign policy positions and actions associated with Donald Trump based on the provided context."