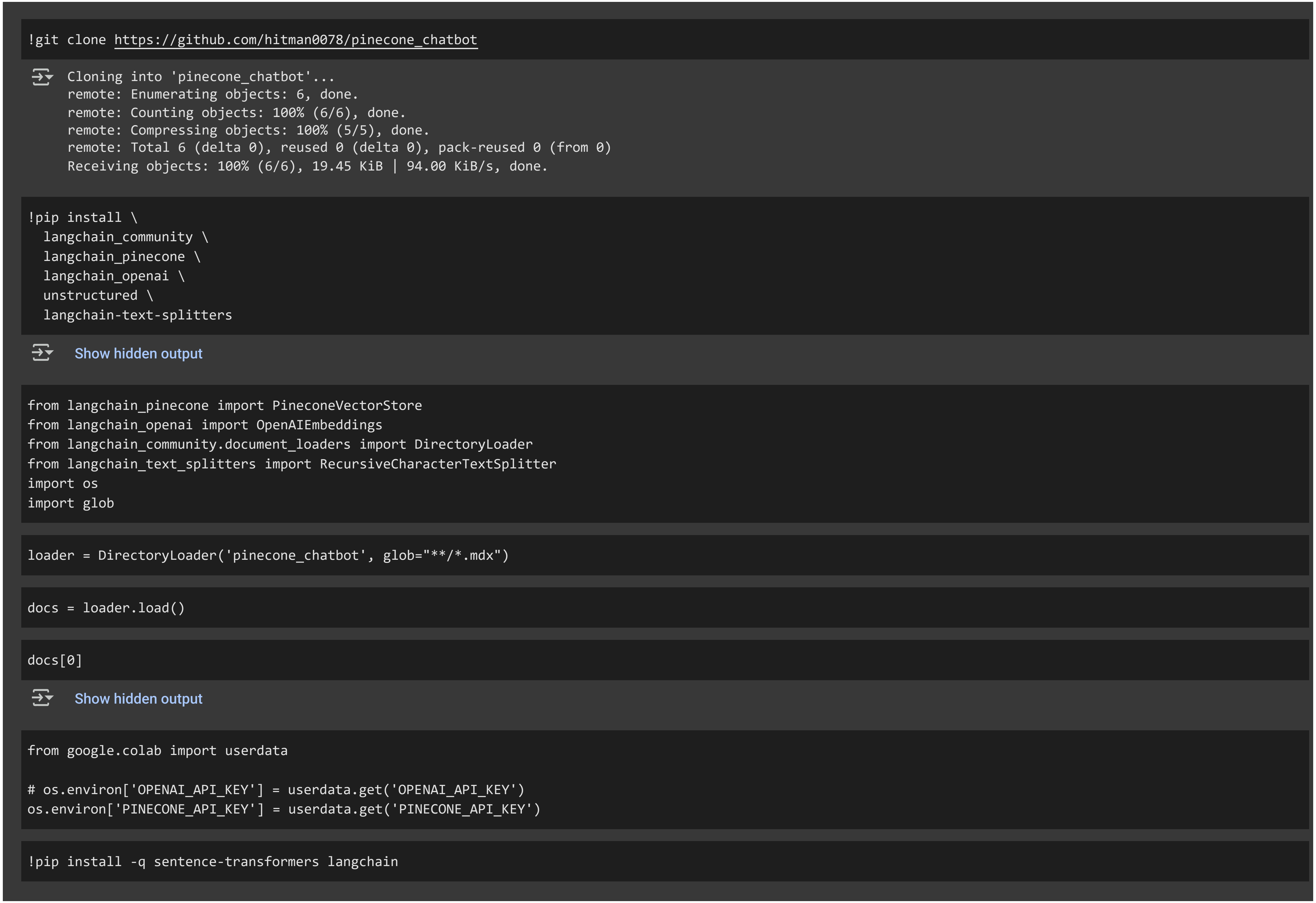


```
!git clone https://github.com/hitman0078/pinecone_chatbot
```

```
Cloning into 'pinecone_chatbot'...
remote: Enumerating objects: 6, done.
remote: Counting objects: 100% (6/6), done.
remote: Compressing objects: 100% (5/5), done.
remote: Total 6 (delta 0), reused 0 (delta 0), pack-reused 0 (from 0)
Receiving objects: 100% (6/6), 19.45 KiB | 94.00 KiB/s, done.
```

```
!pip install \
  langchain_community \
  langchain_pinecone \
  langchain_openai \
  unstructured \
  langchain-text-splitters
```

```
from langchain_pinecone import PineconeVectorStore
from langchain_openai import OpenAIEmbeddings
from langchain_community.document_loaders import DirectoryLoader
from langchain_text_splitters import RecursiveCharacterTextSplitter
import os
import glob
```

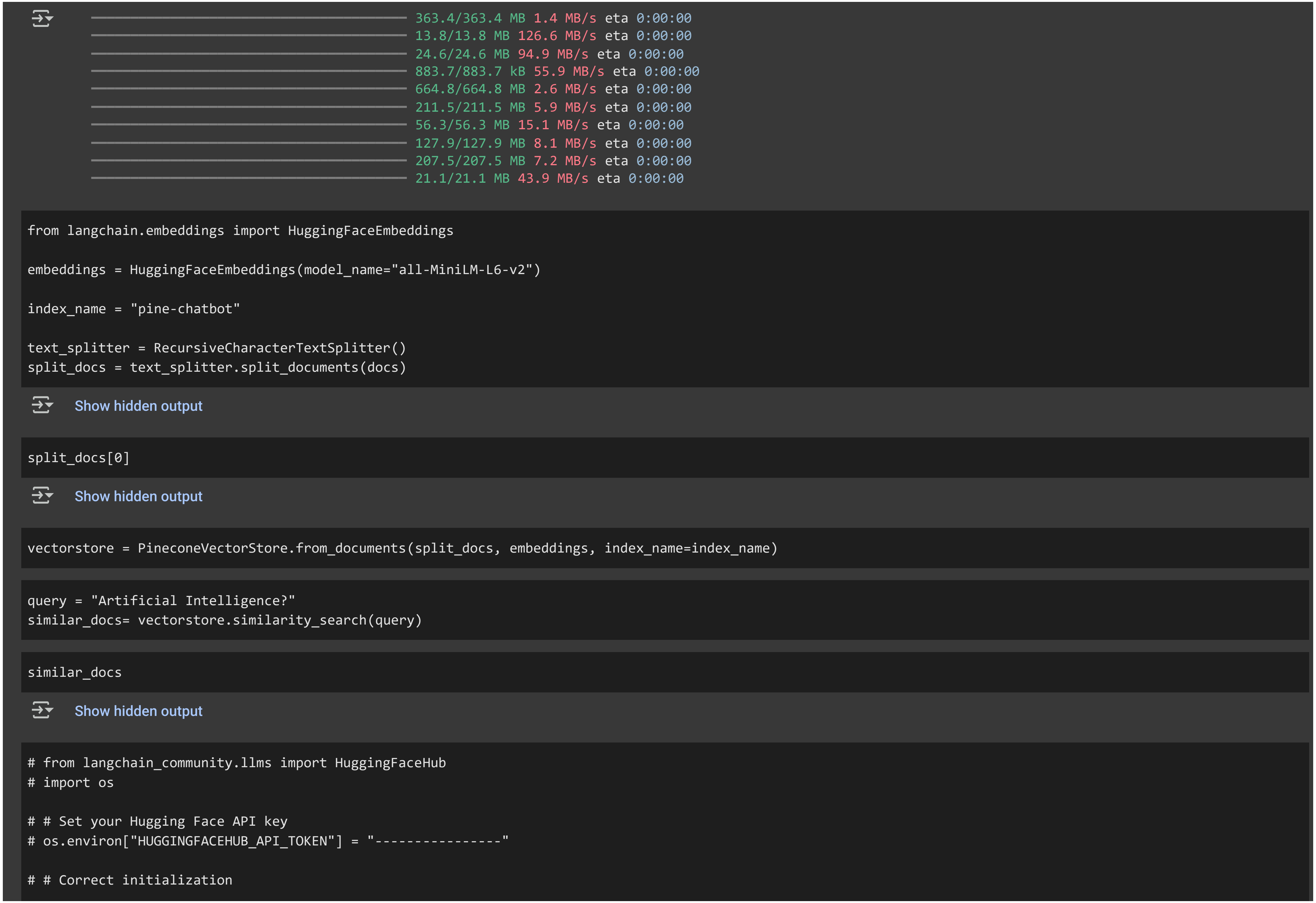
```
loader = DirectoryLoader('pinecone_chatbot', glob="**/*.mdx")
```

```
docs = loader.load()
```

```
docs[0]
```

```
from google.colab import userdata

# os.environ['OPENAI_API_KEY'] = userdata.get('OPENAI_API_KEY')
os.environ['PINECONE_API_KEY'] = userdata.get('PINECONE_API_KEY')
```

```
!pip install -q sentence-transformers langchain
```



```
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 363.4/363.4 MB 1.4 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 13.8/13.8 MB 126.6 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 24.6/24.6 MB 94.9 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 883.7/883.7 kB 55.9 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 664.8/664.8 MB 2.6 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 211.5/211.5 MB 5.9 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 56.3/56.3 MB 15.1 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 127.9/127.9 MB 8.1 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 207.5/207.5 MB 7.2 MB/s eta 0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 21.1/21.1 MB 43.9 MB/s eta 0:00:00
```

```python
from langchain.embeddings import HuggingFaceEmbeddings

embeddings = HuggingFaceEmbeddings(model_name="all-MiniLM-L6-v2")

index_name = "pine-chatbot"

text_splitter = RecursiveCharacterTextSplitter()
split_docs = text_splitter.split_documents(docs)
```

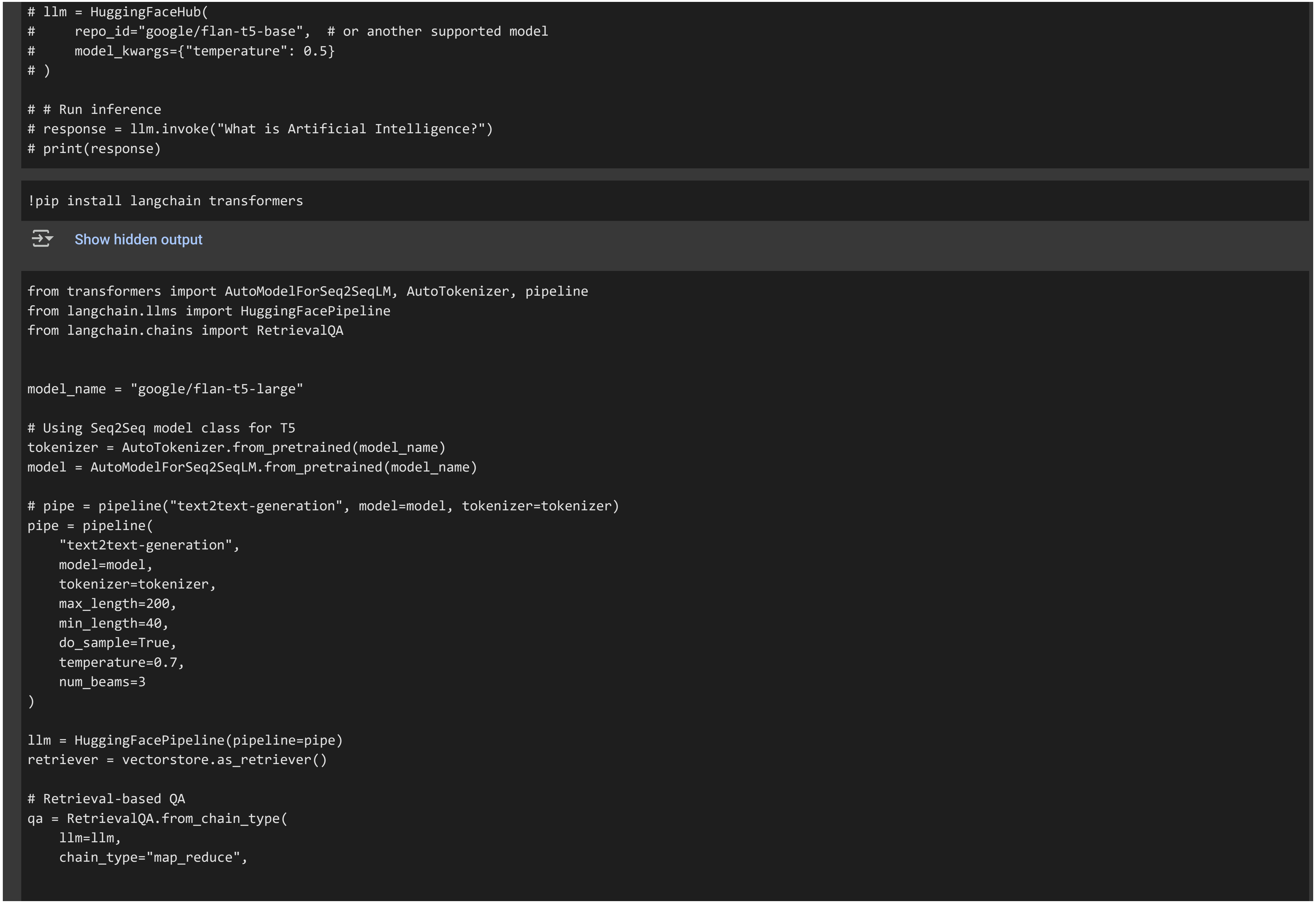
```python
split_docs[0]
```

```python
vectorstore = PineconeVectorStore.from_documents(split_docs, embeddings, index_name=index_name)
```

```python
query = "Artificial Intelligence?"
similar_docs= vectorstore.similarity_search(query)
```

```python
similar_docs
```

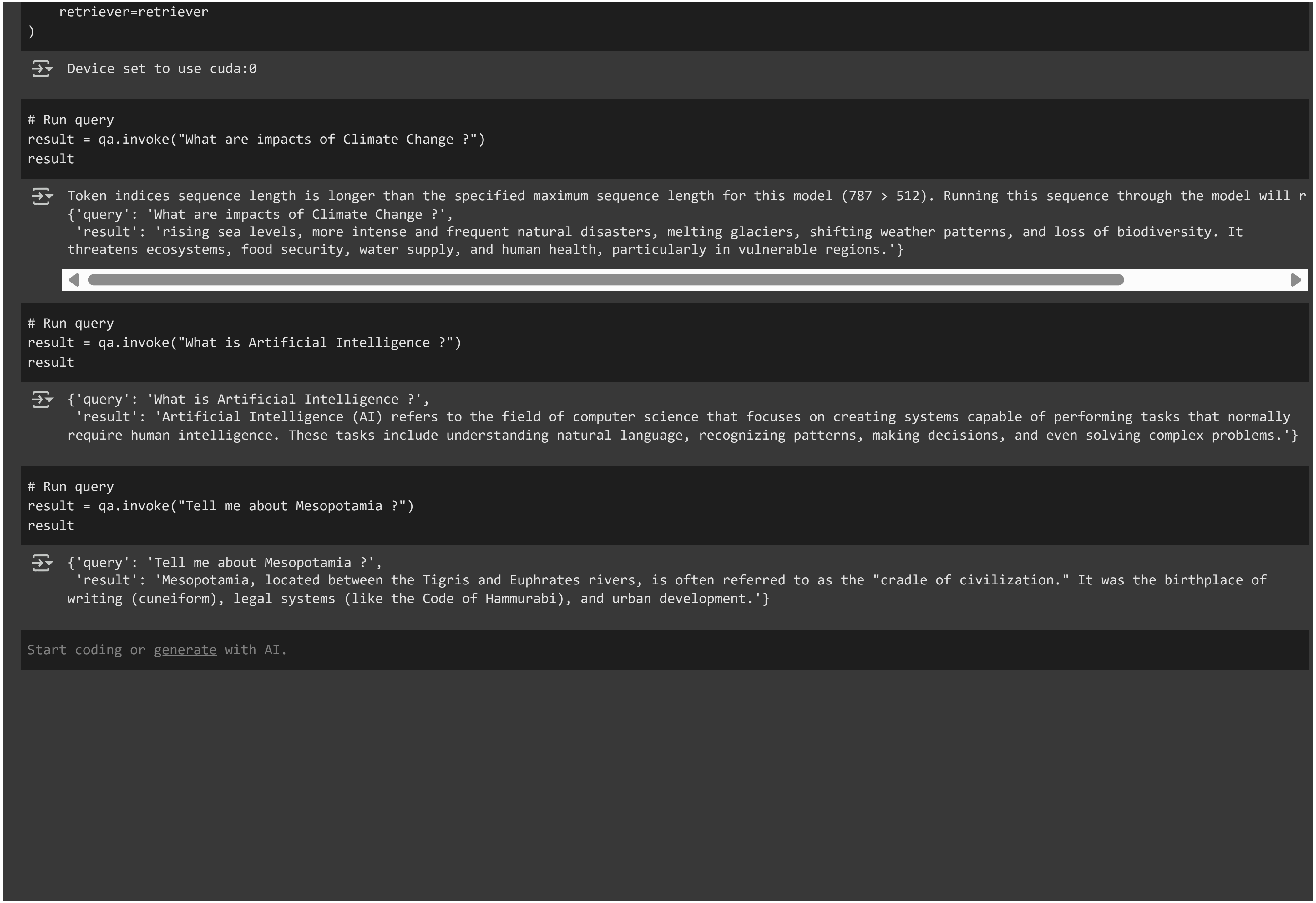
```python
# from langchain_community.llms import HuggingFaceHub
# import os

# # Set your Hugging Face API key
# os.environ["HUGGINGFACEHUB_API_TOKEN"] = "----------------"

# # Correct initialization
```



```
# llm = HuggingFaceHub(
#     repo_id="google/flan-t5-base",  # or another supported model
#     model_kwargs={"temperature": 0.5}
# )

# # Run inference
# response = llm.invoke("What is Artificial Intelligence?")
# print(response)
```

```
!pip install langchain transformers
```

```
from transformers import AutoModelForSeq2SeqLM, AutoTokenizer, pipeline
from langchain.llms import HuggingFacePipeline
from langchain.chains import RetrievalQA


model_name = "google/flan-t5-large"

# Using Seq2Seq model class for T5
tokenizer = AutoTokenizer.from_pretrained(model_name)
model = AutoModelForSeq2SeqLM.from_pretrained(model_name)

# pipe = pipeline("text2text-generation", model=model, tokenizer=tokenizer)
pipe = pipeline(
    "text2text-generation",
    model=model,
    tokenizer=tokenizer,
    max_length=200,
    min_length=40,
    do_sample=True,
    temperature=0.7,
    num_beams=3
)

llm = HuggingFacePipeline(pipeline=pipe)
retriever = vectorstore.as_retriever()

# Retrieval-based QA
qa = RetrievalQA.from_chain_type(
    llm=llm,
    chain_type="map_reduce",
```



```
    retriever=retriever
)
```

```
Device set to use cuda:0
```

```
# Run query
result = qa.invoke("What are impacts of Climate Change ?")
result
```

```
Token indices sequence length is longer than the specified maximum sequence length for this model (787 > 512). Running this sequence through the model will r
{'query': 'What are impacts of Climate Change ?',
 'result': 'rising sea levels, more intense and frequent natural disasters, melting glaciers, shifting weather patterns, and loss of biodiversity. It
threatens ecosystems, food security, water supply, and human health, particularly in vulnerable regions.'}
```

```
# Run query
result = qa.invoke("What is Artificial Intelligence ?")
result
```

```
{'query': 'What is Artificial Intelligence ?',
 'result': 'Artificial Intelligence (AI) refers to the field of computer science that focuses on creating systems capable of performing tasks that normally
require human intelligence. These tasks include understanding natural language, recognizing patterns, making decisions, and even solving complex problems.'}
```

```
# Run query
result = qa.invoke("Tell me about Mesopotamia ?")
result
```

```
{'query': 'Tell me about Mesopotamia ?',
 'result': 'Mesopotamia, located between the Tigris and Euphrates rivers, is often referred to as the "cradle of civilization." It was the birthplace of
writing (cuneiform), legal systems (like the Code of Hammurabi), and urban development.'}
```