

```
!pip install langchain openai faiss-cpu pypdf sentence-transformers
```

```

⇒ Requirement already satisfied: langchain in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: openai in /usr/local/lib/python3.11/dist-packages (1.9
Collecting faiss-cpu
  Downloading faiss_cpu-1.11.0.post1-cp311-cp311-manylinux_2_27_x86_64.manylinux_2_28
Collecting pypdf
  Downloading pypdf-5.9.0-py3-none-any.whl.metadata (7.1 kB)
Requirement already satisfied: sentence-transformers in /usr/local/lib/python3.11/dis
Requirement already satisfied: langchain-core<1.0.0,>=0.3.72 in /usr/local/lib/python
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.9 in /usr/local/l
Requirement already satisfied: langsmith>=0.1.17 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in /usr/local/lib/python3.11/di
Requirement already satisfied: SQLAlchemy<3,>=1.4 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: requests<3,>=2 in /usr/local/lib/python3.11/dist-packa
Requirement already satisfied: PyYAML>=5.3 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: anyio<5,>=3.5.0 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: distro<2,>=1.7.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: httpx<1,>=0.23.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: jiter<1,>=0.4.0 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: sniffio in /usr/local/lib/python3.11/dist-packages (fr
Requirement already satisfied: tqdm>4 in /usr/local/lib/python3.11/dist-packages (fro
Requirement already satisfied: typing-extensions<5,>=4.11 in /usr/local/lib/python3.1
Requirement already satisfied: numpy<3.0,>=1.25.0 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: transformers<5.0.0,>=4.41.0 in /usr/local/lib/python3.
Requirement already satisfied: torch>=1.11.0 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (from
Requirement already satisfied: huggingface-hub>=0.20.0 in /usr/local/lib/python3.11/d
Requirement already satisfied: Pillow in /usr/local/lib/python3.11/dist-packages (fro
Requirement already satisfied: idna>=2.8 in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: certifi in /usr/local/lib/python3.11/dist-packages (fr
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: h11>=0.16 in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (f
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.11/dist
Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in /usr/local/lib/pyth
Requirement already satisfied: jsonpatch<2.0,>=1.33 in /usr/local/lib/python3.11/dist
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in /usr/local/lib/python3.11/dis
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in /usr/local/lib/pyth
Requirement already satisfied: zstandard<0.24.0,>=0.23.0 in /usr/local/lib/python3.11
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: pydantic-core==2.33.2 in /usr/local/lib/python3.11/dis
Requirement already satisfied: typing-inspection>=0.4.0 in /usr/local/lib/python3.11/
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: greenlet>=1 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (f
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (fro
Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch>=1.11.0->sentence-transformer

```

```

Downloading nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.metadata
Collecting nvidia-cuda-runtime-cu12==12.4.127 (from torch>=1.11.0->sentence-transformer)
Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.metadata
Collecting nvidia-cuda-cupti-cu12==12.4.127 (from torch>=1.11.0->sentence-transformer)
Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.metadata
Collecting nvidia-cudnn-cu12==9.1.0.70 (from torch>=1.11.0->sentence-transformers)
Downloading nvidia_cudnn_cu12-9.1.0.70-py3-none-manylinux2014_x86_64.whl.metadata (
Collecting nvidia-cublas-cu12==12.4.5.8 (from torch>=1.11.0->sentence-transformers)

```

```
!pip install -U langchain-community
```



```

Collecting langchain-community
  Downloading langchain_community-0.3.27-py3-none-any.whl.metadata (2.9 kB)
Requirement already satisfied: langchain-core<1.0.0,>=0.3.66 in /usr/local/lib/python
Requirement already satisfied: langchain<1.0.0,>=0.3.26 in /usr/local/lib/python3.11/
Requirement already satisfied: SQLAlchemy<3,>=1.4 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: requests<3,>=2 in /usr/local/lib/python3.11/dist-packa
Requirement already satisfied: PyYAML<=5.3 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in /usr/local/lib/python3.11/dis
Requirement already satisfied: tenacity!=8.4.0,<10,>=8.1.0 in /usr/local/lib/python3.
Collecting dataclasses-json<0.7,>=0.5.7 (from langchain-community)
  Downloading dataclasses_json-0.6.7-py3-none-any.whl.metadata (25 kB)
Collecting pydantic-settings<3.0.0,>=2.4.0 (from langchain-community)
  Downloading pydantic_settings-2.10.1-py3-none-any.whl.metadata (3.4 kB)
Requirement already satisfied: langsmith>=0.1.125 in /usr/local/lib/python3.11/dist-p
Collecting httpx-sse<1.0.0,>=0.4.0 (from langchain-community)
  Downloading httpx_sse-0.4.1-py3-none-any.whl.metadata (9.4 kB)
Requirement already satisfied: numpy>=1.26.2 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.11/d
Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-pa
Collecting marshmallow<4.0.0,>=3.18.0 (from dataclasses-json<0.7,>=0.5.7->langchain-c
  Downloading marshmallow-3.26.1-py3-none-any.whl.metadata (7.3 kB)
Collecting typing-inspect<1,>=0.4.0 (from dataclasses-json<0.7,>=0.5.7->langchain-com
  Downloading typing_inspect-0.9.0-py3-none-any.whl.metadata (1.5 kB)
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.9 in /usr/local/l
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in /usr/local/lib/python3.11/di
Requirement already satisfied: jsonpatch<2.0,>=1.33 in /usr/local/lib/python3.11/dist
Requirement already satisfied: typing-extensions>=4.7 in /usr/local/lib/python3.11/di
Requirement already satisfied: packaging>=23.2 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: httpx<1,>=0.23.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in /usr/local/lib/python3.11/dis
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in /usr/local/lib/pyth
Requirement already satisfied: zstandard<0.24.0,>=0.23.0 in /usr/local/lib/python3.11
Collecting python-dotenv>=0.21.0 (from pydantic-settings<3.0.0,>=2.4.0->langchain-com
  Downloading python_dotenv-1.1.1-py3-none-any.whl.metadata (24 kB)
Requirement already satisfied: typing-inspection>=0.4.0 in /usr/local/lib/python3.11/
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-p

```

```
Requirement already satisfied: greenlet>=1 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: anyio in /usr/local/lib/python3.11/dist-packages (from
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: h11>=0.16 in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: jsonpointer>=1.9 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: pydantic-core==2.33.2 in /usr/local/lib/python3.11/dis
Collecting mypy_extensions>=0.3.0 (from typing-inspect<1,>=0.4.0->dataclasses-json<0.
  Downloading mypy_extensions-1.1.0-py3-none-any.whl.metadata (1.1 kB)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.11/dist-package
Downloading langchain_community-0.3.27-py3-none-any.whl (2.5 MB)
    2.5/2.5 MB 36.9 MB/s eta 0:00:00
Downloading dataclasses_json-0.6.7-py3-none-any.whl (28 kB)
Downloading httpx_sse-0.4.1-py3-none-any.whl (8.1 kB)
```

```
# import required models
```

```
from langchain.document_loaders import PyPDFLoader
from langchain.text_splitter import RecursiveCharacterTextSplitter
from langchain.embeddings import OpenAIEmbeddings
from langchain.vectorstores import FAISS
from langchain.chat_models import ChatOpenAI
from langchain.chains import RetrievalQA
import os
```

```
import pandas as pd
```

```
from langchain_community.document_loaders import PyPDFLoader
```

```
loader = PyPDFLoader("mlpdf.pdf")
documents = loader.load()
```

```
# View sample text
print(documents[0].page_content[:500])
```

```
Andreas C. Müller & Sarah Guido
Introduction to
Machine
Learning
with Python
A GUIDE FOR DATA SCIENTISTS
```

```
!pip install -q pymupdf
import fitz # PyMuPDF
```

```
import fitz  # PyMuPDF

doc = fitz.open("mlpdf.pdf")  # Replace with your file
text = ""
for page in doc:
    text += page.get_text()

print(text[:1000])  # Print first 1000 characters
```

24.1/24.1 MB 30.6 MB/s eta 0:00:00

Andreas C. Müller & Sarah Guido  
 Introduction to  
 Machine  
 Learning  
 with Python  
 A GUIDE FOR DATA SCIENTISTS  
 Andreas C. Müller and Sarah Guido  
 Introduction to Machine Learning  
 with Python  
 A Guide for Data Scientists  
 Boston  
 Farnham  
 Sebastopol  
 Tokyo  
 Beijing  
 Boston  
 Farnham  
 Sebastopol  
 Tokyo  
 Beijing  
 978-1-449-36941-5  
 [LSI]  
 Introduction to Machine Learning with Python  
 by Andreas C. Müller and Sarah Guido  
 Copyright © 2017 Sarah Guido, Andreas Müller. All rights reserved.  
 Printed in the United States of America.  
 Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 954  
 O'Reilly books may be purchased for educational, business, or sales promotional use.  
 also available for most titles (<http://safaribooksonline.com>). For more information,  
 institutional sales department: 800-998-9938 or [corporate@oreilly.com](mailto:corporate@oreilly.com).  
 Editor: Dawn Schanafelt  
 Production Editor: Kristen Brown  
 Copyeditor: Rachel Head  
 Proofreader: Jasmine Kwityn  
 Indexe

```
# text splitter
text_splitter = RecursiveCharacterTextSplitter(
    chunk_size=500,
    chunk_overlap=100,
)
```

```
chunks = text_splitter.split_text(text)
print(f"Total Chunks: {len(chunks)}")
print(chunks[0])
```

```
Total Chunks: 1793
Andreas C. Müller & Sarah Guido
Introduction to
Machine
Learning
with Python
A GUIDE FOR DATA SCIENTISTS
Andreas C. Müller and Sarah Guido
Introduction to Machine Learning
with Python
A Guide for Data Scientists
Boston
Farnham
Sebastopol
Tokyo
Beijing
Boston
Farnham
Sebastopol
Tokyo
Beijing
978-1-449-36941-5
[LSI]
Introduction to Machine Learning with Python
by Andreas C. Müller and Sarah Guido
Copyright © 2017 Sarah Guido, Andreas Müller. All rights reserved.
```

```
# chunks to embeddings
!pip install -U sentence-transformers
```

```
Requirement already satisfied: sentence-transformers in /usr/local/lib/python3.11/dis
Collecting sentence-transformers
  Downloading sentence_transformers-5.0.0-py3-none-any.whl.metadata (16 kB)
Requirement already satisfied: transformers<5.0.0,>=4.41.0 in /usr/local/lib/python3.
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-packages (from
Requirement already satisfied: torch>=1.11.0 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (from
Requirement already satisfied: huggingface-hub>=0.20.0 in /usr/local/lib/python3.11/d
Requirement already satisfied: Pillow in /usr/local/lib/python3.11/dist-packages (fro
Requirement already satisfied: typing_extensions>=4.5.0 in /usr/local/lib/python3.11/
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (f
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: packaging>=20.9 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (f
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.11/dist
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (f
```

Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (fro  
 Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in /usr/local/lib/pyt  
 Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127 in /usr/local/lib/p  
 Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in /usr/local/lib/pyt  
 Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in /usr/local/lib/python3.  
 Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in /usr/local/lib/python3  
 Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in /usr/local/lib/python3.  
 Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in /usr/local/lib/pytho  
 Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in /usr/local/lib/pytho  
 Requirement already satisfied: nvidia-cusparselt-cu12==0.6.2 in /usr/local/lib/python  
 Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in /usr/local/lib/python3.11/  
 Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in /usr/local/lib/python3.1  
 Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in /usr/local/lib/pyth  
 Requirement already satisfied: triton==3.2.0 in /usr/local/lib/python3.11/dist-packag  
 Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.11/dist-packag  
 Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.11/dist-p  
 Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages  
 Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.11/dist-pa  
 Requirement already satisfied: tokenizers<0.22,>=0.21 in /usr/local/lib/python3.11/di  
 Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.11/dist-p  
 Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/python3.11/dist-packag  
 Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.11/dist  
 Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-pack  
 Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/  
 Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-package  
 Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-p  
 Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-p  
 Downloading sentence\_transformers-5.0.0-py3-none-any.whl (470 kB)

470.2/470.2 kB 11.0 MB/s eta 0:00:00

Installing collected packages: sentence-transformers

Attempting uninstall: sentence-transformers

Found existing installation: sentence-transformers 4.1.0

Uninstalling sentence-transformers-4.1.0:

Successfully uninstalled sentence-transformers-4.1.0

Successfully installed sentence-transformers-5.0.0

```
from sentence_transformers import SentenceTransformer
import numpy as np
```

```
model = SentenceTransformer('all-MiniLM-L6-v2')
```

```
embeddings = model.encode(chunks)
print(embeddings[0].shape) # Should be (384,)
```

(384,)

```
!pip install faiss-cpu
```

Requirement already satisfied: faiss-cpu in /usr/local/lib/python3.11/dist-packages (  
 Requirement already satisfied: numpy<3.0,>=1.25.0 in /usr/local/lib/python3.11/dist-p  
 Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-packages (

```
import faiss
dimension = embeddings[0].shape[0]
index = faiss.IndexFlatL2(dimension)
index.add(np.array(embeddings))
```

```
# create retrieval function
def retrieve(query, top_k=3):
    query_vector = model.encode([query])
    distances, indices = index.search(query_vector, top_k)
    return [chunks[i] for i in indices[0]]
```

```
# answer questions using rag logic
from transformers import pipeline
```

```
qa = pipeline("text-generation", model="gpt2")
```

```
query = "What is supervised learning?"
context = " ".join(retrieve(query))
```

```
response = qa(context + "\nQuestion: " + query, max_new_tokens=100)[0]['genera
print(response)
```

```
# faiss + text chunks

def retrieve(query):
    from langchain.embeddings import HuggingFaceEmbeddings
    from langchain.vectorstores import FAISS

    # Load embeddings and FAISS index
    embeddings = HuggingFaceEmbeddings()
    vectorstore = FAISS.load_local("faiss_index", embeddings)

    # Perform similarity search
    docs = vectorstore.similarity_search(query, k=3)
    return [doc.page_content for doc in docs]
```

```
from langchain.text_splitter import CharacterTextSplitter
from langchain.embeddings import HuggingFaceEmbeddings
from langchain.vectorstores import FAISS

# Sample text (you can use PDF text, CSV text, etc.)
text = "Supervised learning is a type of machine learning where a model is tra

# Split into chunks
text_splitter = CharacterTextSplitter(chunk_size=500, chunk_overlap=50)
texts = text_splitter.split_text(text)

# Create embeddings
```



```
embeddings = HuggingFaceEmbeddings()
db = FAISS.from_texts(texts, embeddings)

# Save the vector store
db.save_local("faiss_index")
```

```
# run full rag pipeline
response = qa(context + "\nQuestion: " + query, max_new_tokens=100)[0]['genera
print(response)
```

Setting `pad\_token\_id` to `eos\_token\_id`:50256 for open-end generation.  
flowers.

Remember that supervised learning is used whenever we want to predict a certain outcome from a given input, and we have examples of input/output pairs. We build a machine learning model from these input/output pairs, which comprise our training set. Our goal is to make accurate predictions for new, never-before-seen data. Supervised learning often requires human effort to build the training set, but afterward automates and often speeds up an otherwise laborious or infeasible task. email, the a spam.

Machine learning algorithms that learn from input/output pairs are called supervised learning algorithms because a “teacher” provides supervision to the algorithms in the form of the desired outputs for each example that they learn from. While creating a dataset of inputs and outputs is often a laborious manual process, supervised learnin

iii

## 2. Supervised Learning. . . . .

### Classification and Regression

#### Generalization, Overfitting, and Underfitting

Question: What is supervised learning? Answer: Well, as we discussed below, supervise a trained system to learn from a dataset. In both the supervised learning and the tra the

supervised inputs, and the trained system learns. If you have an input that is the most significant of the input pairs in the training set, you can choose the worst exception at random from two or more of these input pairs. For example, if you want t

```
from langchain.embeddings import HuggingFaceEmbeddings
```

```
embeddings = HuggingFaceEmbeddings(model_name="sentence-transformers/all-MiniL
```

```
from langchain.vectorstores import FAISS
```

```
vectorstore = FAISS.load_local(
    folder_path="faiss_index",
    embeddings=embeddings,
    allow_dangerous_deserialization=True # Only if the index is trusted
)
```

```
def rag_chat(query):
    # Step 1: Retrieve top relevant chunks for the query
    docs = vectorstore.similarity_search(query, k=3) # Top 3 chunks
    context = " ".join([doc.page_content for doc in docs])

    # Step 2: Generate answer using GPT2 or other model
    from transformers import pipeline
    qa = pipeline("text-generation", model="gpt2")

    response = qa(context + "\nQuestion: " + query, max_new_tokens=100)[0]['ge
    return response
```

```
from langchain.vectorstores import FAISS
from langchain.docstore.document import Document
```

```
# Sample documents
```

```
texts = ["LangChain is a framework for developing applications powered by lang
        "It provides tools for retrieval, generation, agents, and chains."]
```

```
docs = [Document(page_content=text) for text in texts]

# Build vectorstore
vectorstore = FAISS.from_documents(docs, embeddings)

# Save it for future use
vectorstore.save_local("rag_faiss_index")
```

```
response = rag_chat("What is LangChain?")
print(response)
```

```
Device set to use cpu
Setting `pad_token_id` to `eos_token_id`:50256 for open-end generation.
LangChain is a framework for developing applications powered by language models. It p
Question: What is LangChain?
Question: What is the lang-chain, and how does it work?
Question: How does it work?
Answer: The lang-chain is a command line tool that can be used to create and run nati
Question: Is it compatible with Open
```

```
print("Vectorstore dimension:", vectorstore.index.d) # FAISS index dimension
print("Test embedding dimension:", len(embeddings.embed_query("Test")))
```

```
Vectorstore dimension: 384
Test embedding dimension: 384
```

```
print(rag_chat("What is LangChain?"))
print(rag_chat("What are use cases of LangChain?"))
print(rag_chat("Explain Retrieval-Augmented Generation."))
```

```
Device set to use cpu
Setting `pad_token_id` to `eos_token_id`:50256 for open-end generation.
LangChain is a framework for developing applications powered by language models. It p
Question: What is LangChain?
Answer: LangChain is a package for generating and storing data in a language-independ
Question: What is the language?
Answer: The language is a set of languages designed for generating, storing, and retr
Question: What is the syntax?
Answer: The syntax is a set of languages designed for
Device set to use cpu
Setting `pad_token_id` to `eos_token_id`:50256 for open-end generation.
LangChain is a framework for developing applications powered by language models. It p
Question: What are use cases of LangChain?
Answer: Using a language model is an important way to get a sense of the language's c
LangChain offers a number of features to help you get a sense of the language and its
```

It provides a simple way to search through the languages of the language, and to get The parser, an interface that allows you to convert, parse, and export all of your na Device set to use cpu

Setting `pad\_token\_id` to `eos\_token\_id`:50256 for open-end generation.

It provides tools for retrieval, generation, agents, and chains. LangChain is a frame Question: Explain Retrieval-Augmented Generation.

LangChain is a framework for the retrieval of data from a database by a user. The fra Question: Explain the use of Likert-based methods with the use of the Likert class.

Likert-based methods are an approach for the retrieval of data from a database. They

```
# enhances rag output
def rag_chat(query):
    docs = vectorstore.similarity_search(query, k=3)
    context = " ".join([doc.page_content for doc in docs])

    from transformers import pipeline
    qa = pipeline("text-generation", model="gpt2")

    response = qa(context + "\nQuestion: " + query, max_new_tokens=100)[0]['ge

    return {
        "answer": response,
        "sources": docs
    }
```

```
result = rag_chat("What is LangChain?")
print("Answer:\n", result["answer"])
print("\nSources:\n", result["sources"])
```

Device set to use cpu

Setting `pad\_token\_id` to `eos\_token\_id`:50256 for open-end generation.

Answer:

LangChain is a framework for developing applications powered by language models. It Question: What is LangChain?

Answer: LangChain is a framework for developing applications powered by language mode

Question: What is a LangChain?

Answer: A language model is a data format that can be used to implement any kind of f

Question: How does LangChain work?

Answer: LangChain is a framework for developing applications powered by language mode

Question:

Sources:

[Document(id='1b44a9b9-18c8-4f08-9512-8c8105edb977', metadata={}, page\_content='Lang

```
# use better llm
qa = pipeline("text2text-generation", model="google/flan-t5-base")
```

```
!pip install gradio
```

```
Requirement already satisfied: gradio in /usr/local/lib/python3.11/dist-packages (5.3
Requirement already satisfied: aiofiles<25.0,>=22.0 in /usr/local/lib/python3.11/dist
Requirement already satisfied: anyio<5.0,>=3.0 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: brotli>=1.1.0 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: fastapi<1.0,>=0.115.2 in /usr/local/lib/python3.11/dis
Requirement already satisfied: ffmpeg in /usr/local/lib/python3.11/dist-packages (from
Requirement already satisfied: gradio-client==1.11.0 in /usr/local/lib/python3.11/dis
Requirement already satisfied: groovy~=0.1 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: httpx<1.0,>=0.24.1 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: huggingface-hub>=0.28.1 in /usr/local/lib/python3.11/d
Requirement already satisfied: jinja2<4.0 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: markupsafe<4.0,>=2.0 in /usr/local/lib/python3.11/dist
Requirement already satisfied: numpy<3.0,>=1.0 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: orjson~=3.0 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: pandas<3.0,>=1.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: pillow<12.0,>=8.0 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: pydantic<2.12,>=2.0 in /usr/local/lib/python3.11/dist-
Requirement already satisfied: pydub in /usr/local/lib/python3.11/dist-packages (from
Requirement already satisfied: python-multipart>=0.0.18 in /usr/local/lib/python3.11/
Requirement already satisfied: pyyaml<7.0,>=5.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: ruff>=0.9.3 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: safehttpx<0.2.0,>=0.1.6 in /usr/local/lib/python3.11/d
Requirement already satisfied: semantic-version~=2.0 in /usr/local/lib/python3.11/dis
Requirement already satisfied: starlette<1.0,>=0.40.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: tomlikit<0.14.0,>=0.12.0 in /usr/local/lib/python3.11/d
Requirement already satisfied: typer<1.0,>=0.12 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: typing-extensions~=4.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: uvicorn>=0.14.0 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: fsspec in /usr/local/lib/python3.11/dist-packages (fro
Requirement already satisfied: websockets<16.0,>=10.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: idna>=2.8 in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.11/dist-package
```

```
Requirement already satisfied: certifi in /usr/local/lib/python3.11/dist-packages (fr
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: h11>=0.16 in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (f
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (f
Requirement already satisfied: tqdm=4.42.1 in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.11/dist
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/di
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packa
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: pydantic-core==2.33.2 in /usr/local/lib/python3.11/dis
Requirement already satisfied: typing-inspection>=0.4.0 in /usr/local/lib/python3.11/
Requirement already satisfied: click>=8.0.0 in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: shellingham>=1.3.0 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: rich>=10.11.0 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (f
Requirement already satisfied: markdown-it-py>=2.2.0 in /usr/local/lib/python3.11/dis
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/d
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: mdurl~=0.1 in /usr/local/lib/python3.11/dist-packages
```

Start coding or [generate](#) with AI.

```
import gradio as gr

def chat_interface(query):
    return rag_chat(query)['answer']

gr.Interface(fn=chat_interface, inputs="text", outputs="text").launch()
```

```
pip install streamlit
```

```
Collecting streamlit
```

```
  Downloading streamlit-1.47.1-py3-none-any.whl.metadata (9.0 kB)
```

```
Requirement already satisfied: altair<6,>=4.0 in /usr/local/lib/python3.11/dist-packa
Requirement already satisfied: blinker<2,>=1.5.0 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: cachetools<7,>=4.0 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: click<9,>=7.0 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: numpy<3,>=1.23 in /usr/local/lib/python3.11/dist-packa
Requirement already satisfied: packaging<26,>=20 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: pandas<3,>=1.4.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: pillow<12,>=7.1.0 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: protobuf<7,>=3.20 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: pyarrow>=7.0 in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: requests<3,>=2.27 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: tenacity<10,>=8.1.0 in /usr/local/lib/python3.11/dist-
Requirement already satisfied: toml<2,>=0.10.1 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: typing-extensions<5,>=4.4.0 in /usr/local/lib/python3.
Collecting watchdog<7,>=2.1.5 (from streamlit)
```

```
  Downloading watchdog-6.0.0-py3-none-manylinux2014_x86_64.whl.metadata (44 kB)
```

```
44.3/44.3 kB 3.7 MB/s eta 0:00:00
```

```
Requirement already satisfied: gitpython!=3.1.19,<4,>=3.0.7 in /usr/local/lib/python3
Collecting pydeck<1,>=0.8.0b4 (from streamlit)
```

```
  Downloading pydeck-0.9.1-py2.py3-none-any.whl.metadata (4.1 kB)
```

```
Requirement already satisfied: tornado!=6.5.0,<7,>=6.0.3 in /usr/local/lib/python3.11
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (fro
Requirement already satisfied: jsonschema>=3.0 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: narwhals>=1.14.2 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: gitdb<5,>=4.0.1 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/di
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packa
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-p
```

```

Requirement already satisfied: smmap<6,>=3.0.1 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: attrs>=22.2.0 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /usr/local/lib
Requirement already satisfied: referencing>=0.28.4 in /usr/local/lib/python3.11/dist-
Requirement already satisfied: rpds-py>=0.7.1 in /usr/local/lib/python3.11/dist-packa
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (f
Downloading streamlit-1.47.1-py3-none-any.whl (9.9 MB)
-----
9.9/9.9 MB 68.9 MB/s eta 0:00:00
Downloading pydeck-0.9.1-py2.py3-none-any.whl (6.9 MB)
-----
6.9/6.9 MB 90.3 MB/s eta 0:00:00
Downloading watchdog-6.0.0-py3-none-manylinux2014_x86_64.whl (79 kB)
-----
79.1/79.1 kB 8.3 MB/s eta 0:00:00
Installing collected packages: watchdog, pydeck, streamlit
Successfully installed pydeck-0.9.1 streamlit-1.47.1 watchdog-6.0.0

```

```
!pip install langchain_community
```

```

Collecting langchain_community
  Downloading langchain_community-0.3.27-py3-none-any.whl.metadata (2.9 kB)
Requirement already satisfied: langchain-core<1.0.0,>=0.3.66 in /usr/local/lib/python
Requirement already satisfied: langchain<1.0.0,>=0.3.26 in /usr/local/lib/python3.11/
Requirement already satisfied: SQLAlchemy<3,>=1.4 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: requests<3,>=2 in /usr/local/lib/python3.11/dist-packa
Requirement already satisfied: PyYAML>=5.3 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in /usr/local/lib/python3.11/dis
Requirement already satisfied: tenacity!=8.4.0,<10,>=8.1.0 in /usr/local/lib/python3.
Collecting dataclasses-json<0.7,>=0.5.7 (from langchain_community)
  Downloading dataclasses_json-0.6.7-py3-none-any.whl.metadata (25 kB)
Collecting pydantic-settings<3.0.0,>=2.4.0 (from langchain_community)
  Downloading pydantic_settings-2.10.1-py3-none-any.whl.metadata (3.4 kB)
Requirement already satisfied: langsmith>=0.1.125 in /usr/local/lib/python3.11/dist-p
Collecting httpx-sse<1.0.0,>=0.4.0 (from langchain_community)
  Downloading httpx_sse-0.4.1-py3-none-any.whl.metadata (9.4 kB)
Requirement already satisfied: numpy>=1.26.2 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.11/d
Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-pa
Collecting marshmallow<4.0.0,>=3.18.0 (from dataclasses-json<0.7,>=0.5.7->langchain_c
  Downloading marshmallow-3.26.1-py3-none-any.whl.metadata (7.3 kB)
Collecting typing-inspect<1,>=0.4.0 (from dataclasses-json<0.7,>=0.5.7->langchain_com
  Downloading typing_inspect-0.9.0-py3-none-any.whl.metadata (1.5 kB)
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.9 in /usr/local/l
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in /usr/local/lib/python3.11/di
Requirement already satisfied: jsonpatch<2.0,>=1.33 in /usr/local/lib/python3.11/dist
Requirement already satisfied: typing-extensions>=4.7 in /usr/local/lib/python3.11/di
Requirement already satisfied: packaging>=23.2 in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: httpx<1,>=0.23.0 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in /usr/local/lib/python3.11/dis
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in /usr/local/lib/pyth

```



```
Requirement already satisfied: zstandard<0.24.0,>=0.23.0 in /usr/local/lib/python3.11
Collecting python-dotenv>=0.21.0 (from pydantic-settings<3.0.0,>=2.4.0->langchain_com
  Downloading python_dotenv-1.1.1-py3-none-any.whl.metadata (24 kB)
Requirement already satisfied: typing-inspection>=0.4.0 in /usr/local/lib/python3.11/
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-package
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: greenlet>=1 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: anyio in /usr/local/lib/python3.11/dist-packages (from
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.11/dist-packag
Requirement already satisfied: h11>=0.16 in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: jsonpointer>=1.9 in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: pydantic-core==2.33.2 in /usr/local/lib/python3.11/dis
Collecting mypy-extensions>=0.3.0 (from typing-inspect<1,>=0.4.0->dataclasses-json<0.
  Downloading mypy_extensions-1.1.0-py3-none-any.whl.metadata (1.1 kB)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.11/dist-package
Downloading langchain_community-0.3.27-py3-none-any.whl (2.5 MB)
  2.5/2.5 MB 50.5 MB/s eta 0:00:00
Downloading dataclasses_json-0.6.7-py3-none-any.whl (28 kB)
Downloading httpx_sse-0.4.1-py3-none-any.whl (8.1 kB)
```

Start coding or [generate](#) with AI.

Start coding or [generate](#) with AI.