```
!pip install langchain
!pip install openai
!pip install PyPDF2
!pip install faiss-cpu
!pip install tiktoken
```

```
Collecting langchain
  Downloading langchain-0.1.0-py3-none-any.whl (797 kB)
                                      ─────── 798.0/798.0 kB 6.3 MB/s eta 0:00:00
Requirement already satisfied: PyYAML>=5.3 in /usr/local/lib/python3.10/dist-packages (from langchain) (6.0.1)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in /usr/local/lib/python3.10/dist-packages (from langchain) (2.0.24)
Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in /usr/local/lib/python3.10/dist-packages (from langchain) (3.9.1)
Requirement already satisfied: async-timeout<5.0.0,>=4.0.0 in /usr/local/lib/python3.10/dist-packages (from langchain) (4.0.3)
Collecting dataclasses-json<0.7,>=0.5.7 (from langchain)
  Downloading dataclasses_json-0.6.3-py3-none-any.whl (28 kB)
Collecting jsonpatch<2.0,>=1.33 (from langchain)
  Downloading jsonpatch-1.33-py2.py3-none-any.whl (12 kB)
Collecting langchain-community<0.1,>=0.0.9 (from langchain)
  Downloading langchain_community-0.0.12-py3-none-any.whl (1.6 MB)
                                      ─────── 1.6/1.6 MB 11.8 MB/s eta 0:00:00
Collecting langchain-core<0.2,>=0.1.7 (from langchain)
  Downloading langchain_core-0.1.10-py3-none-any.whl (216 kB)
                                      ─────── 216.6/216.6 kB 16.2 MB/s eta 0:00:00
Collecting langsmith<0.1.0,>=0.0.77 (from langchain)
  Downloading langsmith-0.0.80-py3-none-any.whl (48 kB)
                                      ─────── 48.3/48.3 kB 3.8 MB/s eta 0:00:00
Requirement already satisfied: numpy<2,>=1 in /usr/local/lib/python3.10/dist-packages (from langchain) (1.23.5)
Requirement already satisfied: pydantic<3,>=1 in /usr/local/lib/python3.10/dist-packages (from langchain) (1.10.13)
Requirement already satisfied: requests<3,>=2 in /usr/local/lib/python3.10/dist-packages (from langchain) (2.31.0)
Requirement already satisfied: tenacity<9.0.0,>=8.1.0 in /usr/local/lib/python3.10/dist-packages (from langchain) (8.2.3)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain) (
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.8.3->langch
Requirement already satisfied: yarl<2.0,>=1.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchai
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain
Collecting marshmallow<4.0.0,>=3.18.0 (from dataclasses-json<0.7,>=0.5.7->langchain)
  Downloading marshmallow-3.20.2-py3-none-any.whl (49 kB)
                                      ─────── 49.4/49.4 kB 4.1 MB/s eta 0:00:00
Collecting typing-inspect<1,>=0.4.0 (from dataclasses-json<0.7,>=0.5.7->langchain)
  Downloading typing_inspect-0.9.0-py3-none-any.whl (8.8 kB)
Collecting jsonpointer>=1.9 (from jsonpatch<2.0,>=1.33->langchain)
  Downloading jsonpointer-2.4-py2.py3-none-any.whl (7.8 kB)
Requirement already satisfied: anyio<5,>=3 in /usr/local/lib/python3.10/dist-packages (from langchain-core<0.2,>=0.1.7->langchain
Requirement already satisfied: packaging<24.0,>=23.2 in /usr/local/lib/python3.10/dist-packages (from langchain-core<0.2,>=0.1.7-
Requirement already satisfied: typing-extensions>=4.2.0 in /usr/local/lib/python3.10/dist-packages (from pydantic<3,>=1->langchai
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2->langchai
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2->langchain) (3.6)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2->langchain) (2.
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2->langchain) (20
Requirement already satisfied: greenlet!=0.4.17 in /usr/local/lib/python3.10/dist-packages (from SQLAlchemy<3,>=1.4->langchain) (
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.10/dist-packages (from anyio<5,>=3->langchain-core<0.2,>=0.
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from anyio<5,>=3->langchain-core<0.2,>=
Collecting mypy-extensions>=0.3.0 (from typing-inspect<1,>=0.4.0->dataclasses-json<0.7,>=0.5.7->langchain)
  Downloading mypy_extensions-1.0.0-py3-none-any.whl (4.7 kB)
Installing collected packages: mypy-extensions, marshmallow, jsonpointer, typing-inspect, langsmith, jsonpatch, langchain-core, d
Successfully installed dataclasses-json-0.6.3 jsonpatch-1.33 jsonpointer-2.4 langchain-0.1.0 langchain-community-0.0.12 langchain
Collecting openai
  Downloading openai-1.7.2-py3-none-any.whl (212 kB)
                                      ─────── 212.1/212.1 kB 2.5 MB/s eta 0:00:00
Requirement already satisfied: anyio<5,>=3.5.0 in /usr/local/lib/python3.10/dist-packages (from openai) (3.7.1)
Requirement already satisfied: distro<2,>=1.7.0 in /usr/lib/python3/dist-packages (from openai) (1.7.0)
Collecting httpx<1,>=0.23.0 (from openai)
  Downloading httpx-0.26.0-py3-none-any.whl (75 kB)
```

```python
from PyPDF2 import PdfReader
from langchain.embeddings.openai import OpenAIEmbeddings
from langchain.text_splitter import CharacterTextSplitter
from langchain.vectorstores import FAISS


import os
os.environ["OPENAI_API_KEY"] = "your_api_key"


pdfreader = PdfReader('dass-42.pdf')


from typing_extensions import Concatenate

raw_text = ''
for i, page in enumerate(pdfreader.pages):
  content = page.extract_text()
  if content:
    raw_text += content
```

raw_text

```python
text_splitter = CharacterTextSplitter(
    separator = "\n",
    chunk_size = 800,
    chunk_overlap = 200,
    length_function = len,
)
```

```python
texts = text_splitter.split_text(raw_text)
```

```python
len(texts)
```

    54

```python
embeddings = OpenAIEmbeddings()
```

```python
document_search = FAISS.from_texts(texts, embeddings)
```

```python
document_search
```

    <langchain_community.vectorstores.faiss.FAISS at 0x79fa18222980>

```python
from langchain.chains.question_answering import load_qa_chain
from langchain.llms import OpenAI
```

```python
chain = load_qa_chain(OpenAI(), chain_type="stuff")
```

```python
query = "What is DASS"
docs = document_search.similarity_search(query)
chain.run(input_documents=docs, question=query)
```

    ' DASS is the abbreviation for the Depression, Anxiety and Stress Scales, which is a widely used instrument for assessing negative emotions in adults. It has two versions - one with 42 questions and one with 21 questions, and has been found to be valid and relia ble in both clinical and general populations. It has been translated and used in Iran, where its psychometric properties have been studied in high school students. '