

```
!pip install sentence-transformers faiss-cpu PyPDF2 transformers torch
```

```
Requirement already satisfied: sentence-transformers in /usr/local/lib/python3.12/dist-packages (5.1.1)
Requirement already satisfied: faiss-cpu in /usr/local/lib/python3.12/dist-packages (1.12.0)
Requirement already satisfied: PyPDF2 in /usr/local/lib/python3.12/dist-packages (3.0.1)
Requirement already satisfied: transformers in /usr/local/lib/python3.12/dist-packages (4.57.0)
Requirement already satisfied: torch in /usr/local/lib/python3.12/dist-packages (2.8.0+cu126)
Requirement already satisfied: tqdm in /usr/local/lib/python3.12/dist-packages (from sentence-transformers) (4.67.1)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.12/dist-packages (from sentence-transformers) (1.6.1)
Requirement already satisfied: scipy in /usr/local/lib/python3.12/dist-packages (from sentence-transformers) (1.16.2)
Requirement already satisfied: huggingface-hub>=0.20.0 in /usr/local/lib/python3.12/dist-packages (from sentence-transformers) (0.35.3)
Requirement already satisfied: Pillow in /usr/local/lib/python3.12/dist-packages (from sentence-transformers) (11.3.0)
Requirement already satisfied: typing_extensions>=4.5.0 in /usr/local/lib/python3.12/dist-packages (from sentence-transformers) (4.15.0)
Requirement already satisfied: numpy<3.0,>=1.25.0 in /usr/local/lib/python3.12/dist-packages (from faiss-cpu) (2.0.2)
Requirement already satisfied: packaging in /usr/local/lib/python3.12/dist-packages (from faiss-cpu) (25.0)
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from transformers) (3.20.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.12/dist-packages (from transformers) (6.0.3)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.12/dist-packages (from transformers) (2024.11.6)
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from transformers) (2.32.4)
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.22.1)
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.12/dist-packages (from transformers) (0.6.2)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from torch) (75.2.0)
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.12/dist-packages (from torch) (1.13.3)
Requirement already satisfied: networkx in /usr/local/lib/python3.12/dist-packages (from torch) (3.5)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from torch) (3.1.6)
Requirement already satisfied: fsspec in /usr/local/lib/python3.12/dist-packages (from torch) (2025.3.0)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch) (12.6.77)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch) (12.6.77)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in /usr/local/lib/python3.12/dist-packages (from torch) (12.6.80)
Requirement already satisfied: nvidia-cudnn-cu12==9.10.2.21 in /usr/local/lib/python3.12/dist-packages (from torch) (9.10.2.21)
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in /usr/local/lib/python3.12/dist-packages (from torch) (12.6.4.1)
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in /usr/local/lib/python3.12/dist-packages (from torch) (11.3.0.4)
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in /usr/local/lib/python3.12/dist-packages (from torch) (10.3.7.77)
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in /usr/local/lib/python3.12/dist-packages (from torch) (11.7.1.2)
Requirement already satisfied: nvidia-cusparselt-cu12==0.7.1 in /usr/local/lib/python3.12/dist-packages (from torch) (0.7.1)
Requirement already satisfied: nvidia-cusparse-cu12==12.5.4.2 in /usr/local/lib/python3.12/dist-packages (from torch) (12.5.4.2)
Requirement already satisfied: nvidia-cusparselt-cu12==0.7.1 in /usr/local/lib/python3.12/dist-packages (from torch) (0.7.1)
Requirement already satisfied: nvidia-nccl-cu12==2.27.3 in /usr/local/lib/python3.12/dist-packages (from torch) (2.27.3)
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in /usr/local/lib/python3.12/dist-packages (from torch) (12.6.77)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in /usr/local/lib/python3.12/dist-packages (from torch) (12.6.85)
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in /usr/local/lib/python3.12/dist-packages (from torch) (1.11.1.6)
Requirement already satisfied: triton==3.4.0 in /usr/local/lib/python3.12/dist-packages (from torch) (3.4.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface-hub>=0.20.0->sentence-tra)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.12/dist-packages (from sympy>=1.13.3->torch) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->torch) (3.0.3)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (3.4.3)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (2.5.0)
```

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->transformers) (2025.10.5)  
Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/python3.12/dist-packages (from scikit-learn->sentence-transformers) (1.5.2)  
Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.12/dist-packages (from scikit-learn->sentence-transformers)

```
from google.colab import files
```

```
uploaded = files.upload() # Select PDFs from your PC
for filename in uploaded.keys():
    print('Uploaded file:', filename)
```

Resume.pdf

**Resume.pdf**(application/pdf) - 86314 bytes, last modified: 7/9/2025 - 100% done

Saving Resume.pdf to Resume.pdf

Uploaded file: Resume.pdf

```
import os
from PyPDF2 import PdfReader

doc_folder = './' # Change to your folder if using Drive
documents = []

for file in os.listdir(doc_folder):
    if file.endswith(".pdf"):
        reader = PdfReader(os.path.join(doc_folder, file))
        text = ""
        for page in reader.pages:
            page_text = page.extract_text()
            if page_text:
                text += page_text + " "
        documents.append(text)

print(f"Loaded {len(documents)} documents")
```

Loaded 1 documents

```
def chunk_text(text, chunk_size=500, overlap=50):
    words = text.split()
    chunks = []
    for i in range(0, len(words), chunk_size - overlap):
        chunk = " ".join(words[i:i + chunk_size])
        chunks.append(chunk)
```

```
        return chunks

all_chunks = []
for doc in documents:
    all_chunks.extend(chunk_text(doc))

print(f"Total chunks: {len(all_chunks)}")
```

Total chunks: 1

```
from sentence_transformers import SentenceTransformer

model = SentenceTransformer('all-MiniLM-L6-v2')
embeddings = model.encode(all_chunks, show_progress_bar=True)
```

Batches: 100% 1/1 [00:00<00:00, 6.19it/s]

```
import faiss
import numpy as np

dimension = embeddings.shape[1]
index = faiss.IndexFlatL2(dimension)
index.add(np.array(embeddings))

print(f"FAISS index has {index.ntotal} vectors")
```

FAISS index has 1 vectors

```
def search(query, k=5):
    query_vec = model.encode([query])
    distances, indices = index.search(np.array(query_vec), k)
    return [all_chunks[i] for i in indices[0]]

# Example:
query = "What are the symptoms of plant disease X?"
results = search(query)
for i, chunk in enumerate(results):
    print(f"Chunk {i+1}: {chunk[:200]}...") # Show first 200 chars
```

Chunk 1: K. KUNDANA [kundanasri2004@gmail.com](mailto:kundanasri2004@gmail.com) ◊LinkedIn ◊GitHub OBJECTIVE Being ambitious and hardworking, I am seeking a position in a com  
 Chunk 2: K. KUNDANA [kundanasri2004@gmail.com](mailto:kundanasri2004@gmail.com) ◊LinkedIn ◊GitHub OBJECTIVE Being ambitious and hardworking, I am seeking a position in a com  
 Chunk 3: K. KUNDANA [kundanasri2004@gmail.com](mailto:kundanasri2004@gmail.com) ◊LinkedIn ◊GitHub OBJECTIVE Being ambitious and hardworking, I am seeking a position in a com  
 Chunk 4: K. KUNDANA [kundanasri2004@gmail.com](mailto:kundanasri2004@gmail.com) ◊LinkedIn ◊GitHub OBJECTIVE Being ambitious and hardworking, I am seeking a position in a com  
 Chunk 5: K. KUNDANA [kundanasri2004@gmail.com](mailto:kundanasri2004@gmail.com) ◊LinkedIn ◊GitHub OBJECTIVE Being ambitious and hardworking, I am seeking a position in a com

```
import ipywidgets as widgets
from IPython.display import display

query_box = widgets.Text(
    value='',
    placeholder='Ask a question...',
    description='Query:',
    layout=widgets.Layout(width='100%')
)

display(query_box)
```

Query:

🔍 Query: W  
 🌸 Answer: Eclipse Projects Plant Disease

🔍 Query: What are pr  
 🌸 Answer: Java

🔍 Query: What are projects?  
 🌸 Answer: Eclipse Projects Plant Disease Detection

```
from transformers import pipeline

qa_model = pipeline("question-answering")
```

No model was supplied, defaulted to distilbert/distilbert-base-cased-distilled-squad and revision 564e9b5 (<https://huggingface.co/distilbert>)  
 Using a pipeline without specifying a model name and revision in production is not recommended.  
 Device set to use cpu

```
# 🌸 Function to generate an answer
def generate_answer(query):
    context = " ".join(search(query))
    # Pass the question and context separately to the QA model
    answer = qa_model(question=query, context=context, max_length=200)['answer']
    return answer

# 🌸 Function to handle query input and generate answer
```

```
.. 🔗 Function to handle query input and generate answer
```

```
def on_submit(change):  
    query = change['new']  
    if query:  
        print("🔍 Query:", query)  
        answer = generate_answer(query)  
        print("💡 Answer:", answer)  
  
# 🔗 Connect the query box to the handler  
query_box.observe(on_submit, names='value')
```

Could not connect to the reCAPTCHA service. Please check your internet connection and reload to get a reCAPTCHA challenge.