

```

1 !pip install pysqlite3
2 !pip install pysqlite3-binary
3 !pip install pypdf
4 !pip install -U langchain-community

```

```

Requirement already satisfied: tenacity!=8.4.0,<10,>=8.1.0 in /usr/local/lib/python3.11/dist-packages (from langchain-
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-2.8.1-py3-none-any.whl.metadata (3.5 kB)
Requirement already satisfied: langsmith<0.4,>=0.1.125 in /usr/local/lib/python3.11/dist-packages (from langchain-comm
Collecting httpx_sse<1.0.0,>=0.4.0 (from langchain-community)
  Downloading httpx_sse-0.4.0-py3-none-any.whl.metadata (9.0 kB)
Requirement already satisfied: numpy<3,>=1.26.2 in /usr/local/lib/python3.11/dist-packages (from langchain-community)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->l
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.
Collecting marshmallow<4.0.0,>=3.18.0 (from dataclasses-json<0.7,>=0.5.7->langchain-community)
  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-community)
  Downloading typing_inspect-0.9.0-py3-none-any.whl.metadata (1.5 kB)
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.6 in /usr/local/lib/python3.11/dist-packages (from
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in /usr/local/lib/python3.11/dist-packages (from langchain<1.0.0
Requirement already satisfied: jsonpatch<2.0,>=1.33 in /usr/local/lib/python3.11/dist-packages (from langchain-core<1.
Requirement already satisfied: packaging<25,>=23.2 in /usr/local/lib/python3.11/dist-packages (from langchain-core<1.0
Requirement already satisfied: typing-extensions>=4.7 in /usr/local/lib/python3.11/dist-packages (from langchain-core<
Requirement already satisfied: httpx<1,>=0.23.0 in /usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.1
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in /usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in /usr/local/lib/python3.11/dist-packages (from langsm
Requirement already satisfied: zstandard<0.24.0,>=0.23.0 in /usr/local/lib/python3.11/dist-packages (from langsmith<0.
Collecting python-dotenv>=0.21.0 (from pydantic-settings<3.0.0,>=2.4.0->langchain-community)
  Downloading python_dotenv-1.0.1-py3-none-any.whl.metadata (23 kB)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2->lan
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2->lan
Requirement already satisfied: greenlet!=0.4.17 in /usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4->l
Requirement already satisfied: anyio in /usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->langsmith<0.4,
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->langsm
Requirement already satisfied: h11<0.15,>=0.13 in /usr/local/lib/python3.11/dist-packages (from httpcore==1.*->httpx<1
Requirement already satisfied: jsonpointer>=1.9 in /usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33-
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,
Requirement already satisfied: pydantic-core==2.27.2 in /usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>
Collecting mpy_extensions>=0.3.0 (from typing-inspect<1,>=0.4.0->dataclasses-json<0.7,>=0.5.7->langchain-community)
  Downloading mpy_extensions-1.0.0-py3-none-any.whl.metadata (1.1 kB)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.11/dist-packages (from anyio->httpx<1,>=0.23.0->
Downloading langchain_community-0.3.19-py3-none-any.whl (2.5 MB)
2.5/2.5 MB 47.0 MB/s eta 0:00:00
Downloading dataclasses_json-0.6.7-py3-none-any.whl (28 kB)
Downloading httpx_sse-0.4.0-py3-none-any.whl (7.8 kB)
Downloading pydantic_settings-2.8.1-py3-none-any.whl (30 kB)
Downloading marshmallow-3.26.1-py3-none-any.whl (50 kB)
50.9/50.9 kB 5.0 MB/s eta 0:00:00
Downloading python_dotenv-1.0.1-py3-none-any.whl (19 kB)
Downloading typing_inspect-0.9.0-py3-none-any.whl (8.8 kB)
Downloading mpy_extensions-1.0.0-py3-none-any.whl (4.7 kB)
Installing collected packages: python-dotenv, mpy_extensions, marshmallow, httpx-sse, typing-inspect, pydantic-settin
Successfully installed dataclasses-json-0.6.7 httpx-sse-0.4.0 langchain-community-0.3.19 marshmallow-3.26.1 mpy-exten

```

```

1 from langchain.document_loaders import PyPDFLoader
2 from langchain.text_splitter import RecursiveCharacterTextSplitter
3 from langchain.embeddings import OpenAIEmbeddings
4 import pysqlite3
5 import sys

```

```

1 # Load and split the pdf
2 pdf_loader = PyPDFLoader("/content/the_nestle_hr_policy_pdf_2012.pdf")
3 pdf_pages = pdf_loader.load_and_split()
4 print(len(pdf_pages))

```

```

5

```

```

1 #Split the PDF pages into smaller chunks and print the number of chunks.
2 doc_splitter = RecursiveCharacterTextSplitter(chunk_size=1024, chunk_overlap=64)
3 docs = doc_splitter.split_documents(pdf_pages)
4 doc_list = [doc.page_content for doc in docs]
5 print(doc_list[1])

```

Policy  
Mandatory  
September 20  
12  
Issuing departement  
Hum  
an Resources  
Target audience  
All  
employees  
Approver  
Executive Board, Nestlé S.A.  
Repository  
All Nestlé Principles and Policies, Standards and  
Guidelines can be found in the Centre online repository at:  
<http://intranet.nestle.com/nestledocs>  
Copyright and confidentiality  
All  
All rights belong to Nestec Ltd., Vevey, Switzerland.  
© 2012, Nestec Ltd.  
Design  
Nestec Ltd., Corporate Identity & Design,  
Vevey, Switzerland  
Production  
brain'print GmbH, Switzerland  
Paper  
This report is printed on BVS, a paper produced  
from well-managed forests and other controlled sources  
certified by the Forest Stewardship Council (FSC).

**Create vector representations for text chunks using Chroma dB and OpenAI's embeddings.**

```
1 !pip install chromadb openai
2 !pip install tiktoken
```



Requirement already satisfied: certifi>=2017.4.1/ in /usr/local/lib/python3.11/dist-packages (from requests>=2.26.0->t  
 Downloading tiktoken-0.9.0-cp311-cp311-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (1.2 MB)  
 1.2/1.2 MB 19.6 MB/s eta 0:00:00  
 Installing collected packages: tiktoken  
 Successfully installed tiktoken-0.9.0

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

```
1 __import__('pysqlite3')
2 import sys
3 sys.modules['sqlite3'] = sys.modules.pop('pysqlite3')
```


```
1 import chromadb
2 from langchain.vectorstores import Chroma
3 from langchain.chains import RetrievalQA
4 from langchain.chat_models import ChatOpenAI
```

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

```
1 # Initilize the OpenAI Embedding model
2 embedding_model = OpenAIEmbeddings()
```


```
1 # Create a Chroma vector store from text chunks
2 vector_store = Chroma.from_texts(texts=doc_list, embedding=embedding_model, persist_directory="./chroma_db")
```

```
1 vector_store.persist()
```

 <ipython-input-14-261e6ea3744e>:1: LangChainDeprecationWarning: Since Chroma 0.4.x the manual persistence method is no longer supported. Please use the new persistence methods instead.

```
1 ## Create my own query
2 query = "What is working conditions?"
3
4 #Get the similar embedding from the vector_store
5 results = vector_store.similarity_search(query, k=3)
```

```
1 for doc in results:
2     print(doc.page_content)
3     print("\n.....")
```


 of mutual trust. HR ensures that a respectful dialogue is present and the voice of the employees is heard.  
 Corporate policy:  
 Policy on Conditions of Work and Employment  
 Employment and working conditions

.....  
 and Wellness Company.  
 As such, we provide flexible working conditions whenever possible and encourage our employees to have outside interests especially community involvement.  
 Those with line management responsibilities are required to take personal ownership of safety and health within their area of responsibility and are encouraged to develop their capability in this area.  
 Nestlé's commitment however goes beyond its own employees. We care about all people working inside or outside our premises under contractual obligations with service providers and we insist that they also take steps so that adequate working conditions are made available to them.  
 We believe that it is essential to build a relationship based on trust and respect of employees at all levels. We do not tolerate any form of harassment or discrimination.  
 Therefore, managers are committed to build and sustain, with their teams, an environment of mutual trust. HR ensures that a respectful


.....  
 important matters such as Creating Shared Value, the health and safety in the workplace and our concern for the environment.  
 The Company and employee representatives are expected to make all necessary efforts to develop fair and constructive dialogues,

overcome the difficulties that they might encounter, reach sustainable agreements and implement them.  
 Corporate policy:  
 The Nestlé Employee Relations Policy  
 Employee relations  
 The Human Resources management described in this document requires and supports an organisation "on the move".  
 Nestlé is committed to continue the journey to establishing flat and flexible structures with minimal levels of management and broad spans of control, which enable people development, increase efficiency, and ease implementation of our "Nestlé Management and Leadership Principles".  
 Less hierarchical layers call for increased cooperation between colleagues. This is what


```
1 #Define the chat-gpt model
2 llm = ChatOpenAI(model_name="gpt-3.5-turbo", temperature=0)
3
4 # Retrieve similar embedding from chroma db i.e vector_store
5 retriever = vector_store.as_retriever(search_type="similarity", search_kwargs={"k": 3})
6
7 # Define RetrievalQA which takes user query and find the similar content from pdfs and send to chat-gpt model to get the
8 # answer
9 qa_chain = RetrievalQA.from_chain_type(
10     llm=llm,
11     chain_type="stuff",
12     retriever=retriever,
13     return_source_documents=True
14 )
```

 <ipython-input-17-01102e2510f0>:2: LangChainDeprecationWarning: The class `ChatOpenAI` was deprecated in LangChain 0.0.1  
 llm = ChatOpenAI(model\_name="gpt-3.5-turbo", temperature=0)

```
1 response = qa_chain({"query": query})
```

 <ipython-input-18-7e13e1e84d4b>:1: LangChainDeprecationWarning: The method `Chain.\_\_call\_\_` was deprecated in langchain  
 response = qa\_chain({"query": query})

```
1 response['result']
```

 'Working conditions refer to the environment in which employees carry out their daily tasks. This includes factors such as the physical work environment, hours of work, safety regulations, benefits, and overall atmosphere in the workplace. In the context provided, working conditions are emphasized as an important aspect of Nestlé's corporate policy, ensuring that employees have a safe, respectful, and supportive environment to work in.'

### Implement the Gradio to show the demo

```
1 !pip install gradio
2 !pip install --upgrade gradio
```



Requirement already satisfied: tomli<0.14.0,>=0.12.0 in /usr/local/lib/python3.11/dist-packages (from gradio) (0.13.0)

Requirement already satisfied: typer<1.0,>=0.12 in /usr/local/lib/python3.11/dist-packages (from gradio) (0.15.2)

Requirement already satisfied: typing-extensions~=4.0 in /usr/local/lib/python3.11/dist-packages (from gradio) (4.12.2)

Requirement already satisfied: uvicorn>=0.14.0 in /usr/local/lib/python3.11/dist-packages (from gradio) (0.34.0)

Requirement already satisfied: fsspec in /usr/local/lib/python3.11/dist-packages (from gradio-client==1.7.2->gradio) (2024.10.0)

Requirement already satisfied: websockets<16.0,>=10.0 in /usr/local/lib/python3.11/dist-packages (from gradio-client==1.7.2->gradio) (13.1)

Requirement already satisfied: idna>=2.8 in /usr/local/lib/python3.11/dist-packages (from anyio<5.0,>=3.0->gradio) (3.10.1)

Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.11/dist-packages (from anyio<5.0,>=3.0->gradio) (1.3.1)

Requirement already satisfied: certifi in /usr/local/lib/python3.11/dist-packages (from httpx>=0.24.1->gradio) (2025.1.1)

Requirement already satisfied: httpcore==1.\* in /usr/local/lib/python3.11/dist-packages (from httpx>=0.24.1->gradio) (1.0.7)

Requirement already satisfied: h11<0.15,>=0.13 in /usr/local/lib/python3.11/dist-packages (from httpcore==1.\*->httpx>=0.24.1->gradio) (0.14.0)

Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.28.1->gradio) (3.16.1)

Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.28.1->gradio) (2.32.3)

Requirement already satisfied: tqdm>=4.42.1 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.28.1->gradio) (4.67.1)

Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas<3.0,>=1.0->gradio) (2.9.0.post0)

Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas<3.0,>=1.0->gradio) (2024.2)

Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas<3.0,>=1.0->gradio) (2024.2)

Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/dist-packages (from pydantic>=2.0->gradio) (0.7.0)

Requirement already satisfied: pydantic-core==2.27.2 in /usr/local/lib/python3.11/dist-packages (from pydantic>=2.0->gradio) (2.27.2)

Requirement already satisfied: click>=8.0.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0,>=0.12->gradio) (8.1.8)

Requirement already satisfied: shellingham>=1.3.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0,>=0.12->gradio) (1.5.4)

Requirement already satisfied: rich>=10.11.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0,>=0.12->gradio) (13.9.0)

Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2->pandas<3.0,>=1.0->gradio) (1.17.0)

Requirement already satisfied: markdown-it-py>=2.2.0 in /usr/local/lib/python3.11/dist-packages (from rich>=10.11.0->typer<1.0,>=0.12->gradio) (3.0.0)

Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/dist-packages (from rich>=10.11.0->typer<1.0,>=0.12->gradio) (2.19.1)

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.28.1->gradio) (3.4.0)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.28.1->gradio) (2.3.0)

Requirement already satisfied: mdurl==0.1 in /usr/local/lib/python3.11/dist-packages (from markdown-it-py>=2.2.0->rich>=10.11.0->typer<1.0,>=0.12->gradio) (0.1.2)

```
1 import gradio as gr
```

```
1 def greet(query):
2     response = qa_chain({"query": query})
3     return response['result']
4
5
6 chat_demo = gr.Interface(
7     fn=greet,
8     inputs=['text'],
9     outputs=['text']
10 )
```

```
1 chat_demo.launch(share=True)
```

Colab notebook detected. To show errors in colab notebook, set debug=True in launch()  
 \* Running on public URL: <https://ef628e6efceb812379.gradio.live>

This share link expires in 72 hours. For free permanent hosting and GPU upgrades, run `gradio deploy` from the terminal

query

What about the training and learning part in the Nestlé company?

Clear

Submit

output

At Nestlé, learning is a significant part of the company culture. Employees are encouraged to continuously upgrade their knowledge and skills. The company determines training and development priorities, and the responsibility for implementing these actions is shared between employees, line managers, and Human Resources. Experience and on-the-job training are primary sources of learning, with managers guiding and coaching employees to succeed in their current roles. Nestlé also offers a comprehensive range of training activities and methodologies to support employee learning and growth. Additionally,

Flag

Use via API  · Built with Gradio  · Settings 

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

