

GenBiz AI – Zero-to-Startup AI Co-Founder(Project) []

[] GenBiz AI – Kaggle Notebook Code (Full Project)

[] All 16 GenAI Capabilities Implemented

STEP 0: INSTALL LIBRARIES (Run these in separate Kaggle notebook cells, NOT inside the script)

Cell 1:

```
%pip install langchain faiss-cpu transformers  
tiktoken weaviate-client sentence-transformers  
python-docx pdfplumber
```

STEP 1: IMPORT DEPENDENCIES

```
import os
```

```
import pdfplumber
```

```
import docx
```

```
from langchain.chains import RetrievalQA
```

```
from langchain.embeddings import  
OpenAIEmbeddings
```

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

```
from langchain.llms import OpenAI
```

```
from langchain.document_loaders import  
WebBaseLoader
```

```
import json
```

```
from sentence_transformers import  
SentenceTransformer
```

```
import numpy as np
```

```
!pip install --upgrade huggingface_hub  
from huggingface_hub import hf_hub_download # Use the newer function  
instead  
  
Requirement already satisfied: huggingface_hub in  
/usr/local/lib/python3.11/dist-packages (0.16.4)  
Collecting huggingface_hub  
  Downloading huggingface_hub-0.30.2-py3-none-any.whl.metadata (13 kB)  
Requirement already satisfied: filelock in  
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
```

```

(3.18.0)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(2025.3.2)
Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub) (23.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub) (6.0.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(4.13.1)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub) (2025.1.31)
Downloading huggingface_hub-0.30.2-py3-none-any.whl (481 kB)
_____ 481.4/481.4 kB 8.0 MB/s eta
0:00:00a 0:00:01
pting uninstall: huggingface_hub
  Found existing installation: huggingface-hub 0.16.4
  Uninstalling huggingface-hub-0.16.4:
    Successfully uninstalled huggingface-hub-0.16.4
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
have fsspec 2025.3.2 which is incompatible.
transformers 4.51.1 requires tokenizers<0.22,>=0.21, but you have
tokenizers 0.13.3 which is incompatible.
Successfully installed huggingface_hub-0.30.2

!pip uninstall -y transformers sentence-transformers huggingface_hub
!pip install transformers sentence-transformers huggingface_hub --
upgrade

```

```
Found existing installation: transformers 4.51.1
Uninstalling transformers-4.51.1:
  Successfully uninstalled transformers-4.51.1
Found existing installation: sentence-transformers 3.4.1
Uninstalling sentence-transformers-3.4.1:
  Successfully uninstalled sentence-transformers-3.4.1
Found existing installation: huggingface-hub 0.30.2
Uninstalling huggingface-hub-0.30.2:
  Successfully uninstalled huggingface-hub-0.30.2
Collecting transformers
  Downloading transformers-4.51.3-py3-none-any.whl.metadata (38 kB)
Collecting sentence-transformers
  Downloading sentence_transformers-4.1.0-py3-none-any.whl.metadata (13 kB)
Collecting huggingface_hub
  Using cached huggingface_hub-0.30.2-py3-none-any.whl.metadata (13 kB)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from transformers) (3.18.0)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers) (1.26.4)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from transformers) (23.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from transformers) (6.0.2)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers)
(2024.11.6)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from transformers) (2.32.3)
Collecting tokenizers<0.22,>=0.21 (from transformers)
  Downloading tokenizers-0.21.1-cp39-abi3-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (6.8 kB)
Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from transformers) (0.5.2)
Requirement already satisfied: tqdm>=4.27 in
/usr/local/lib/python3.11/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: torch>=1.11.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(2.5.1+cu124)
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.2.2)
Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.15.2)
Requirement already satisfied: Pillow in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(11.1.0)
Requirement already satisfied: typing_extensions>=4.5.0 in
```

/usr/local/lib/python3.11/dist-packages (from sentence-transformers) (4.13.1)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub) (2025.3.2)
Requirement already satisfied: mkl_fft in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers) (1.3.8)
Requirement already satisfied: mkl_random in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers) (1.2.4)
Requirement already satisfied: mkl_umath in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers) (2025.1.0)
Requirement already satisfied: tbb4py in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers) (2022.1.0)
Requirement already satisfied: mkl-service in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers) (2.4.1)
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.4.2)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.1.6)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.4.127)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.4.127)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.4.127)
Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (9.1.0.70)
Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.4.5.8)
Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (11.2.1.3)
Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (10.3.5.147)
Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in

/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (11.6.1.9)
Requirement already satisfied: nvidia-cusparse-cu12==12.3.1.170 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.3.1.170)
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (2.21.5)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.4.127)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.4.127)
Requirement already satisfied: triton==3.1.0 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.1.0)
Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from sympy==1.13.1->torch>=1.11.0->sentence-transformers) (1.3.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2025.1.31)
Requirement already satisfied: joblib>=1.1.1 in /usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-transformers) (1.4.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-transformers) (3.6.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.11.0->sentence-transformers) (3.0.2)
Requirement already satisfied: intel-openmp<2026,>=2024 in /usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17->transformers) (2024.2.0)
Requirement already satisfied: tbb==2022.* in /usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17->transformers) (2022.1.0)

```

Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->transformers) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-
>transformers) (2024.2.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->transformers) (2024.2.0)
Downloading transformers-4.51.3-py3-none-any.whl (10.4 MB)
_____ 10.4/10.4 MB 67.4 MB/s eta
0:00:00:00:01:01
ers-4.1.0-py3-none-any.whl (345 kB)
_____ 345.7/345.7 kB 17.9 MB/s eta
0:00:00
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.0 MB)
_____ 3.0/3.0 MB 73.3 MB/s eta
0:00:00:00:01
ers, sentence-transformers
  Attempting uninstall: tokenizers
    Found existing installation: tokenizers 0.13.3
    Uninstalling tokenizers-0.13.3:
      Successfully uninstalled tokenizers-0.13.3
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
have fsspec 2025.3.2 which is incompatible.
Successfully installed huggingface_hub-0.30.2 sentence-transformers-
4.1.0 tokenizers-0.21.1 transformers-4.51.3

from huggingface_hub import hf_hub_download

# Example usage:
file_path = hf_hub_download(
    repo_id="sentence-transformers/all-MiniLM-L6-v2",
    filename="config.json"
)

# First, run this installation cell (click the [▶] button)
!pip install pdfplumber pdfminer.six pdf2image pytesseract pillow
!apt-get install -y poppler-utils # Required for PDF rendering

Collecting pdfplumber
  Downloading pdfplumber-0.11.6-py3-none-any.whl.metadata (42 kB)
_____ 42.8/42.8 kB 1.3 MB/s eta
0:00:00
iner.six
  Downloading pdfminer_six-20250416-py3-none-any.whl.metadata (4.1 kB)
Requirement already satisfied: pdf2image in

```

```

/usr/local/lib/python3.11/dist-packages (1.17.0)
Requirement already satisfied: pytesseract in
/usr/local/lib/python3.11/dist-packages (0.3.13)
Requirement already satisfied: pillow in
/usr/local/lib/python3.11/dist-packages (11.1.0)
  Downloading pdfminer_six-20250327-py3-none-any.whl.metadata (4.1 kB)
Collecting pypdfium2>=4.18.0 (from pdfplumber)
  Downloading pypdfium2-4.30.1-py3-none-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (48 kB)
----- 48.2/48.2 kB 2.4 MB/s eta
0:00:00
Requirement already satisfied: charset-normalizer>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six) (3.4.1)
Requirement already satisfied: cryptography>=36.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six) (44.0.2)
Requirement already satisfied: packaging>=21.3 in
/usr/local/lib/python3.11/dist-packages (from pytesseract) (23.2)
Requirement already satisfied: cffi>=1.12 in
/usr/local/lib/python3.11/dist-packages (from cryptography>=36.0.0-
>pdfminer.six) (1.17.1)
Requirement already satisfied: pycparser in
/usr/local/lib/python3.11/dist-packages (from cffi>=1.12-
>cryptography>=36.0.0->pdfminer.six) (2.22)
Downloading pdfplumber-0.11.6-py3-none-any.whl (60 kB)
----- 60.2/60.2 kB 3.2 MB/s eta
0:00:00
pdfminer_six-20250327-py3-none-any.whl (5.6 MB)
----- 5.6/5.6 MB 35.5 MB/s eta
0:00:00000:0100:01
2-4.30.1-py3-none-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (2.9
MB)
----- 2.9/2.9 MB 53.2 MB/s eta
0:00:00:00:01
2, pdfminer.six, pdfplumber
Successfully installed pdfminer.six-20250327 pdfplumber-0.11.6
pypdfium2-4.30.1
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  libpoppler-dev libpoppler-private-dev libpoppler118
The following NEW packages will be installed:
  poppler-utils
The following packages will be upgraded:
  libpoppler-dev libpoppler-private-dev libpoppler118
3 upgraded, 1 newly installed, 0 to remove and 119 not upgraded.
Need to get 1,469 kB of archives.
After this operation, 696 kB of additional disk space will be used.
Get:1 http://security.ubuntu.com/ubuntu jammy-security/main amd64

```



```
libpoppler-private-dev amd64 22.02.0-2ubuntu0.7 [199 kB]
Get:2 http://security.ubuntu.com/ubuntu jammy-security/main amd64
libpoppler-dev amd64 22.02.0-2ubuntu0.7 [5,180 B]
Get:3 http://security.ubuntu.com/ubuntu jammy-security/main amd64
libpoppler118 amd64 22.02.0-2ubuntu0.7 [1,079 kB]
Get:4 http://security.ubuntu.com/ubuntu jammy-security/main amd64
poppler-utils amd64 22.02.0-2ubuntu0.7 [186 kB]
Fetched 1,469 kB in 2s (685 kB/s)
(Reading database ... 128691 files and directories currently
installed.)
Preparing to unpack .../libpoppler-private-dev_22.02.0-
2ubuntu0.7_amd64.deb ...
Unpacking libpoppler-private-dev:amd64 (22.02.0-2ubuntu0.7) over
(22.02.0-2ubuntu0.6) ...
Preparing to unpack .../libpoppler-dev_22.02.0-
2ubuntu0.7_amd64.deb ...
Unpacking libpoppler-dev:amd64 (22.02.0-2ubuntu0.7) over (22.02.0-
2ubuntu0.6) ...
Preparing to unpack .../libpoppler118_22.02.0-2ubuntu0.7_amd64.deb ...
Unpacking libpoppler118:amd64 (22.02.0-2ubuntu0.7) over (22.02.0-
2ubuntu0.6) ...
Selecting previously unselected package poppler-utils.
Preparing to unpack .../poppler-utils_22.02.0-2ubuntu0.7_amd64.deb ...
Unpacking poppler-utils (22.02.0-2ubuntu0.7) ...
Setting up libpoppler118:amd64 (22.02.0-2ubuntu0.7) ...
Setting up poppler-utils (22.02.0-2ubuntu0.7) ...
Setting up libpoppler-dev:amd64 (22.02.0-2ubuntu0.7) ...
Setting up libpoppler-private-dev:amd64 (22.02.0-2ubuntu0.7) ...
Processing triggers for man-db (2.10.2-1) ...
Processing triggers for libc-bin (2.35-0ubuntu3.8) ...
/sbin/ldconfig.real: /usr/local/lib/libtcm.so.1 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libumf.so.0 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtcm_debug.so.1 is not a
symbolic link

/sbin/ldconfig.real: /usr/local/lib/libhwloc.so.15 is not a symbolic
link

/sbin/ldconfig.real: /usr/local/lib/libur_loader.so.0 is not a
symbolic link

/sbin/ldconfig.real: /usr/local/lib/libur_adapter_level_zero.so.0 is
not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libur_adapter_opencl.so.0 is not a
symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_5.so.3 is not a
```

symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_0.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbb.so.12 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbmalloc_proxy.so.2 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbmalloc.so.2 is not a symbolic link

Installation (run this first)

!pip install python-docx

Import

from docx import Document

Create a simple test document

doc = Document()

doc.add_paragraph("Hello Kaggle! This is a test Word document.")

doc.save("test_document.docx")

Read it back

doc = Document("test_document.docx")

for para in doc.paragraphs:
 print(para.text)

Requirement already satisfied: python-docx in

/usr/local/lib/python3.11/dist-packages (1.1.2)

Requirement already satisfied: lxml>=3.1.0 in

/usr/local/lib/python3.11/dist-packages (from python-docx) (5.3.1)

Requirement already satisfied: typing-extensions>=4.9.0 in

/usr/local/lib/python3.11/dist-packages (from python-docx) (4.13.1)

Hello Kaggle! This is a test Word document.

This should work without any installation

import json

Test JSON operations

data = {

"name": "Kaggle User",

"skills": ["Python", "Data Science"],

"active": True

}

```

# Convert to JSON string
json_string = json.dumps(data, indent=2)
print("JSON String:")
print(json_string)

# Convert back to Python
python_data = json.loads(json_string)
print("\nPython Data:")
print(python_data)

JSON String:
{
  "name": "Kaggle User",
  "skills": [
    "Python",
    "Data Science"
  ],
  "active": true
}

Python Data:
{'name': 'Kaggle User', 'skills': ['Python', 'Data Science'],
'active': True}

import numpy as np

# Create test arrays
a = np.array([1, 2, 3])
b = np.array([4, 5, 6])

# Basic operations
print("Sum:", a + b)
print("Dot product:", np.dot(a, b))

# Matrix operation
matrix = np.random.rand(3, 3)
print("\nRandom matrix:")
print(matrix)
print("\nMatrix inverse:")
print(np.linalg.inv(matrix))

Sum: [5 7 9]
Dot product: 32

Random matrix:
[[0.69316017 0.61436676 0.36277879]
 [0.01039187 0.70414985 0.22911853]
 [0.85073703 0.008202  0.33265587]]

Matrix inverse:

```

```
[[ 3.78431712 -3.28003832 -1.86785351]
 [ 3.118241   -1.27108227 -2.52514167]
 [-9.75492929  8.41973853  7.84523457]]
```

Run this diagnostic cell

```
import sys
print("Python version:", sys.version)
print("\nSystem path:", sys.path[:3]) # First 3 entries
```

```
try:
    import os
    print("\nos module successfully imported!")
    print("Current directory:", os.getcwd())
    print("Directory contents:", os.listdir())
except Exception as e:
    print("\nError:", str(e))
```

Python version: 3.11.11 (main, Dec 4 2024, 08:55:07) [GCC 11.4.0]

System path: ['/kaggle/working', '/kaggle/lib/kagglelegym',
'/kaggle/lib']

os module successfully imported!
Current directory: /kaggle/working
Directory contents: ['.virtual_documents', 'test_document.docx',
'business_analysis_report.json']

!pip install --upgrade langchain openai langchain-openai tiktoken

Requirement already satisfied: langchain in
/usr/local/lib/python3.11/dist-packages (0.3.23)
Requirement already satisfied: openai in
/usr/local/lib/python3.11/dist-packages (1.75.0)
Requirement already satisfied: langchain-openai in
/usr/local/lib/python3.11/dist-packages (0.0.1)
Collecting langchain-openai
Using cached langchain_openai-0.3.14-py3-none-any.whl.metadata (2.3
kB)
Requirement already satisfied: tiktoken in
/usr/local/lib/python3.11/dist-packages (0.5.2)
Collecting tiktoken
Downloading tiktoken-0.9.0-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (6.7 kB)
Requirement already satisfied: langchain-core<1.0.0,>=0.3.51 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.3.54)
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.8
in /usr/local/lib/python3.11/dist-packages (from langchain) (0.3.8)
Requirement already satisfied: langsmith<0.4,>=0.1.17 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.1.147)
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in

/usr/local/lib/python3.11/dist-packages (from langchain) (2.11.3)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.0.38)
Requirement already satisfied: requests<3,>=2 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.32.3)
Requirement already satisfied: PyYAML>=5.3 in
/usr/local/lib/python3.11/dist-packages (from langchain) (6.0.2)
Requirement already satisfied: anyio<5,>=3.5.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (3.7.1)
Requirement already satisfied: distro<2,>=1.7.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (1.9.0)
Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (0.28.1)
Requirement already satisfied: jiter<1,>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (0.8.2)
Requirement already satisfied: sniffio in
/usr/local/lib/python3.11/dist-packages (from openai) (1.3.1)
Requirement already satisfied: tqdm>4 in
/usr/local/lib/python3.11/dist-packages (from openai) (4.67.1)
Requirement already satisfied: typing-extensions<5,>=4.11 in
/usr/local/lib/python3.11/dist-packages (from openai) (4.13.1)
Requirement already satisfied: regex>=2022.1.18 in
/usr/local/lib/python3.11/dist-packages (from tiktoken) (2024.11.6)
Requirement already satisfied: idna>=2.8 in
/usr/local/lib/python3.11/dist-packages (from anyio<5,>=3.5.0->openai)
(3.10)
Requirement already satisfied: certifi in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->openai) (2025.1.31)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->openai) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*->httpx<1,>=0.23.0->openai) (0.14.0)
Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (8.5.0)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (1.33)
Requirement already satisfied: packaging<25,>=23.2 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (23.2)
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17->langchain) (3.10.15)
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17-

```

>langchain) (1.0.0)
Requirement already satisfied: annotated-types>=0.6.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.1 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (2.33.1)
Requirement already satisfied: typing-inspection>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (0.4.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (3.4.1)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (2.3.0)
Requirement already satisfied: greenlet!=0.4.17 in
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4-
>langchain) (3.1.1)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33-
>langchain-core<1.0.0,>=0.3.51->langchain) (3.0.0)
Using cached langchain_openai-0.3.14-py3-none-any.whl (62 kB)
Downloading tiktoken-0.9.0-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.2 MB)
----- 1.2/1.2 MB 15.2 MB/s eta
0:00:0000:010:01
pting uninstall: tiktoken
  Found existing installation: tiktoken 0.5.2
  Uninstalling tiktoken-0.5.2:
    Successfully uninstalled tiktoken-0.5.2
Attempting uninstall: langchain-openai
  Found existing installation: langchain-openai 0.0.1
  Uninstalling langchain-openai-0.0.1:
    Successfully uninstalled langchain-openai-0.0.1
Successfully installed langchain-openai-0.3.14 tiktoken-0.9.0

from kaggle_secrets import UserSecretsClient
import os

user_secrets = UserSecretsClient()
os.environ["OPENAI_API_KEY"] =
user_secrets.get_secret("OPENAI_API_KEY") # Must match your secret
name

!pip install faiss-cpu # For CPU version (recommended for Kaggle)

Requirement already satisfied: faiss-cpu in
/usr/local/lib/python3.11/dist-packages (1.7.4)

```

```
!free -h # Check available memory
```

	total	used	free	shared	buff/cache
available					
Mem:	31Gi	1.7Gi	17Gi	0.0Ki	11Gi
29Gi					
Swap:	0B	0B	0B		

```
!pip install torch torchvision torchaudio --index-url  
https://download.pytorch.org/whl/cpu
```

```
Looking in indexes: https://download.pytorch.org/whl/cpu  
Requirement already satisfied: torch in  
/usr/local/lib/python3.11/dist-packages (2.5.1+cu124)  
Requirement already satisfied: torchvision in  
/usr/local/lib/python3.11/dist-packages (0.20.1+cu124)  
Requirement already satisfied: torchaudio in  
/usr/local/lib/python3.11/dist-packages (2.5.1+cu124)  
Requirement already satisfied: filelock in  
/usr/local/lib/python3.11/dist-packages (from torch) (3.18.0)  
Requirement already satisfied: typing-extensions>=4.8.0 in  
/usr/local/lib/python3.11/dist-packages (from torch) (4.13.1)  
Requirement already satisfied: networkx in  
/usr/local/lib/python3.11/dist-packages (from torch) (3.4.2)  
Requirement already satisfied: jinja2 in  
/usr/local/lib/python3.11/dist-packages (from torch) (3.1.6)  
Requirement already satisfied: fsspec in  
/usr/local/lib/python3.11/dist-packages (from torch) (2025.3.2)  
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in  
/usr/local/lib/python3.11/dist-packages (from torch) (12.4.127)  
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127  
in /usr/local/lib/python3.11/dist-packages (from torch) (12.4.127)  
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in  
/usr/local/lib/python3.11/dist-packages (from torch) (12.4.127)  
Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in  
/usr/local/lib/python3.11/dist-packages (from torch) (9.1.0.70)  
Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in  
/usr/local/lib/python3.11/dist-packages (from torch) (12.4.5.8)  
Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in  
/usr/local/lib/python3.11/dist-packages (from torch) (11.2.1.3)  
Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in  
/usr/local/lib/python3.11/dist-packages (from torch) (10.3.5.147)  
Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in  
/usr/local/lib/python3.11/dist-packages (from torch) (11.6.1.9)  
Requirement already satisfied: nvidia-cuspars-cu12==12.3.1.170 in  
/usr/local/lib/python3.11/dist-packages (from torch) (12.3.1.170)  
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in  
/usr/local/lib/python3.11/dist-packages (from torch) (2.21.5)  
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in  
/usr/local/lib/python3.11/dist-packages (from torch) (12.4.127)
```

Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch) (12.4.127)
Requirement already satisfied: triton==3.1.0 in
/usr/local/lib/python3.11/dist-packages (from torch) (3.1.0)
Requirement already satisfied: sympy==1.13.1 in
/usr/local/lib/python3.11/dist-packages (from torch) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy==1.13.1->torch)
(1.3.0)
Requirement already satisfied: numpy in
/usr/local/lib/python3.11/dist-packages (from torchvision) (1.26.4)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/usr/local/lib/python3.11/dist-packages (from torchvision) (11.1.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch) (3.0.2)
Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy->torchvision)
(1.3.8)
Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy->torchvision)
(1.2.4)
Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy->torchvision)
(0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy->torchvision) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy->torchvision)
(2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy->torchvision)
(2.4.1)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy->torchvision)
(2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy->torchvision)
(2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl->numpy-
>torchvision) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy-
>torchvision) (2024.2.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy->torchvision) (2024.2.0)

```
import torch  
print("PyTorch version:", torch.__version__)
```



```
print("CUDA available:", torch.cuda.is_available()) # Should be True
if GPU is enabled
```

PyTorch version: 2.5.1+cu124
CUDA available: False

```
!pip install torch==2.0.1+cpu torchvision==0.15.2+cpu -f
https://download.pytorch.org/whl/torch_stable.html
```

Looking in links: https://download.pytorch.org/whl/torch_stable.html
Collecting torch==2.0.1+cpu

Downloading https://download.pytorch.org/whl/cpu/torch-2.0.1%2Bcpu-cp311-cp311-linux_x86_64.whl (195.4 MB)

0:00:00:00:0100:01 195.4/195.4 MB 4.6 MB/s eta

0:00:00:00:0100:01

0:00:00a 0:00:010m 1.5/1.5 MB 1.8 MB/s eta

0:00:00a 0:00:010m

Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from torch==2.0.1+cpu) (3.18.0)

Requirement already satisfied: typing-extensions in /usr/local/lib/python3.11/dist-packages (from torch==2.0.1+cpu) (4.13.1)

Requirement already satisfied: sympy in /usr/local/lib/python3.11/dist-packages (from torch==2.0.1+cpu) (1.13.1)

Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from torch==2.0.1+cpu) (3.4.2)

Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from torch==2.0.1+cpu) (3.1.6)

Requirement already satisfied: numpy in /usr/local/lib/python3.11/dist-packages (from torchvision==0.15.2+cpu) (1.26.4)

Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from torchvision==0.15.2+cpu) (2.32.3)

Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in /usr/local/lib/python3.11/dist-packages (from torchvision==0.15.2+cpu) (11.1.0)

Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->torch==2.0.1+cpu) (3.0.2)

Requirement already satisfied: mkl_fft in /usr/local/lib/python3.11/dist-packages (from numpy->torchvision==0.15.2+cpu) (1.3.8)

Requirement already satisfied: mkl_random in /usr/local/lib/python3.11/dist-packages (from numpy->torchvision==0.15.2+cpu) (1.2.4)

Requirement already satisfied: mkl_umath in

```
/usr/local/lib/python3.11/dist-packages (from numpy-
>torchvision==0.15.2+cpu) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy->torchvision==0.15.2+cpu) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy-
>torchvision==0.15.2+cpu) (2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy-
>torchvision==0.15.2+cpu) (2.4.1)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>torchvision==0.15.2+cpu) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests-
>torchvision==0.15.2+cpu) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests-
>torchvision==0.15.2+cpu) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests-
>torchvision==0.15.2+cpu) (2025.1.31)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch==2.0.1+cpu)
(1.3.0)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy-
>torchvision==0.15.2+cpu) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy-
>torchvision==0.15.2+cpu) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl->numpy-
>torchvision==0.15.2+cpu) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy-
>torchvision==0.15.2+cpu) (2024.2.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy->torchvision==0.15.2+cpu) (2024.2.0)
Installing collected packages: torch, torchvision
  Attempting uninstall: torch
    Found existing installation: torch 2.5.1+cu124
    Uninstalling torch-2.5.1+cu124:
      Successfully uninstalled torch-2.5.1+cu124
  Attempting uninstall: torchvision
    Found existing installation: torchvision 0.20.1+cu124
    Uninstalling torchvision-0.20.1+cu124:
      Successfully uninstalled torchvision-0.20.1+cu124
```

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.

pytorch-lightning 2.5.1 requires torch>=2.1.0, but you have torch 2.0.1+cpu which is incompatible.

torchaudio 2.5.1+cu124 requires torch==2.5.1, but you have torch 2.0.1+cpu which is incompatible.

Successfully installed torch-2.0.1+cpu torchvision-0.15.2+cpu

Installation first

```
!pip install langchain-community beautifulsoup4
```

Import

```
from langchain_community.document_loaders import WebBaseLoader
```

Load a webpage

```
loader = WebBaseLoader("https://example.com")
```

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

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

```
print(docs[0].page_content[:200]) # First 200 chars
```

Requirement already satisfied: langchain-community in
/usr/local/lib/python3.11/dist-packages (0.0.29)

Requirement already satisfied: beautifulsoup4 in
/usr/local/lib/python3.11/dist-packages (4.12.2)

Requirement already satisfied: PyYAML>=5.3 in
/usr/local/lib/python3.11/dist-packages (from langchain-community)
(6.0.2)

Requirement already satisfied: SQLAlchemy<3,>=1.4 in
/usr/local/lib/python3.11/dist-packages (from langchain-community)
(2.0.38)

Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in
/usr/local/lib/python3.11/dist-packages (from langchain-community)
(3.11.16)

Requirement already satisfied: dataclasses-json<0.7,>=0.5.7 in
/usr/local/lib/python3.11/dist-packages (from langchain-community)
(0.6.7)

Collecting langchain-core<0.2.0,>=0.1.33 (from langchain-community)

Downloading langchain_core-0.1.53-py3-none-any.whl.metadata (5.9 kB)

Requirement already satisfied: langsmith<0.2.0,>=0.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-community)
(0.1.147)

Requirement already satisfied: numpy<2,>=1 in
/usr/local/lib/python3.11/dist-packages (from langchain-community)
(1.26.4)

Requirement already satisfied: requests<3,>=2 in
/usr/local/lib/python3.11/dist-packages (from langchain-community)
(2.32.3)

Requirement already satisfied: tenacity<9.0.0,>=8.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-community)

(8.5.0)

Requirement already satisfied: soupsieve>1.2 in
/usr/local/lib/python3.11/dist-packages (from beautifulsoup4) (2.6)

Requirement already satisfied: aiohappyeyeballs>=2.3.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community) (2.6.1)

Requirement already satisfied: aiosignal>=1.1.2 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community) (1.3.2)

Requirement already satisfied: attrs>=17.3.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community) (25.3.0)

Requirement already satisfied: frozenlist>=1.1.1 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community) (1.5.0)

Requirement already satisfied: multidict<7.0,>=4.5 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community) (6.2.0)

Requirement already satisfied: propcache>=0.2.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community) (0.3.1)

Requirement already satisfied: yarl<2.0,>=1.17.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community) (1.19.0)

Requirement already satisfied: marshmallow<4.0.0,>=3.18.0 in
/usr/local/lib/python3.11/dist-packages (from dataclasses-json<0.7,>=0.5.7->langchain-community) (3.26.1)

Requirement already satisfied: typing-inspect<1,>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from dataclasses-json<0.7,>=0.5.7->langchain-community) (0.9.0)

Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<0.2.0,>=0.1.33->langchain-community) (1.33)

Requirement already satisfied: packaging<24.0,>=23.2 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<0.2.0,>=0.1.33->langchain-community) (23.2)

Requirement already satisfied: pydantic<3,>=1 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<0.2.0,>=0.1.33->langchain-community) (2.11.3)

Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.2.0,>=0.1.0->langchain-community) (0.28.1)

Requirement already satisfied: orjson<4.0.0,>=3.9.14 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.2.0,>=0.1.0->langchain-community) (3.10.15)

Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.2.0,>=0.1.0->langchain-community) (1.0.0)

Requirement already satisfied: mkl_fft in

/usr/local/lib/python3.11/dist-packages (from numpy<2,>=1->langchain-community) (1.3.8)
Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy<2,>=1->langchain-community) (1.2.4)
Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy<2,>=1->langchain-community) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-packages (from numpy<2,>=1->langchain-community) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy<2,>=1->langchain-community) (2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy<2,>=1->langchain-community) (2.4.1)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain-community) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain-community) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain-community) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain-community) (2025.1.31)
Requirement already satisfied: greenlet!=0.4.17 in
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4->langchain-community) (3.1.1)
Requirement already satisfied: typing-extensions>=4.6.0 in
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4->langchain-community) (4.13.1)
Requirement already satisfied: anyio in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain-community) (3.7.1)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain-community) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*->httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain-community) (0.14.0)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33->langchain-core<0.2.0,>=0.1.33->langchain-community) (3.0.0)
Requirement already satisfied: annotated-types>=0.6.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3,>=1-

```

>langchain-core<0.2.0,>=0.1.33->langchain-community) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.1 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3,>=1-
>langchain-core<0.2.0,>=0.1.33->langchain-community) (2.33.1)
Requirement already satisfied: typing-inspection>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3,>=1-
>langchain-core<0.2.0,>=0.1.33->langchain-community) (0.4.0)
Requirement already satisfied: mypy-extensions>=0.3.0 in
/usr/local/lib/python3.11/dist-packages (from typing-
inspect<1,>=0.4.0->dataclasses-json<0.7,>=0.5.7->langchain-community)
(1.0.0)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy<2,>=1-
>langchain-community) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy<2,>=1-
>langchain-community) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy<2,>=1->langchain-community) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy<2,>=1-
>langchain-community) (2024.2.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy<2,>=1->langchain-community) (2024.2.0)
Requirement already satisfied: sniffio>=1.1 in
/usr/local/lib/python3.11/dist-packages (from anyio->httpx<1,>=0.23.0-
>langsmith<0.2.0,>=0.1.0->langchain-community) (1.3.1)
Downloading langchain_core-0.1.53-py3-none-any.whl (303 kB)
303.1/303.1 kB 5.4 MB/s eta
0:00:00a 0:00:01
pting uninstall: langchain-core
  Found existing installation: langchain-core 0.3.54
  Uninstalling langchain-core-0.3.54:
    Successfully uninstalled langchain-core-0.3.54
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
langchain 0.3.23 requires langchain-core<1.0.0,>=0.3.51, but you have
langchain-core 0.1.53 which is incompatible.
langchain-openai 0.3.14 requires langchain-core<1.0.0,>=0.3.53, but
you have langchain-core 0.1.53 which is incompatible.
langchain-text-splitters 0.3.8 requires langchain-core<1.0.0,>=0.3.51,
but you have langchain-core 0.1.53 which is incompatible.
Successfully installed langchain-core-0.1.53
Loaded 1 documents

```

Example Domain

Example Domain

This domain is for use in illustrative examples in documents. You may use this

domain in literature without prior coordination or asking for permission.

Mor

```
pip install openai
```

```
Requirement already satisfied: openai in
/usr/local/lib/python3.11/dist-packages (1.75.0)
Requirement already satisfied: anyio<5,>=3.5.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (3.7.1)
Requirement already satisfied: distro<2,>=1.7.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (1.9.0)
Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (0.28.1)
Requirement already satisfied: jiter<1,>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (0.8.2)
Requirement already satisfied: pydantic<3,>=1.9.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (2.11.3)
Requirement already satisfied: sniffio in
/usr/local/lib/python3.11/dist-packages (from openai) (1.3.1)
Requirement already satisfied: tqdm>4 in
/usr/local/lib/python3.11/dist-packages (from openai) (4.67.1)
Requirement already satisfied: typing-extensions<5,>=4.11 in
/usr/local/lib/python3.11/dist-packages (from openai) (4.13.1)
Requirement already satisfied: idna>=2.8 in
/usr/local/lib/python3.11/dist-packages (from anyio<5,>=3.5.0->openai)
(3.10)
Requirement already satisfied: certifi in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>openai) (2025.1.31)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>openai) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*-
>httpx<1,>=0.23.0->openai) (0.14.0)
Requirement already satisfied: annotated-types>=0.6.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3,>=1.9.0-
>openai) (0.7.0)
```

Requirement already satisfied: pydantic-core==2.33.1 in /usr/local/lib/python3.11/dist-packages (from pydantic<3,>=1.9.0->openai) (2.33.1)

Requirement already satisfied: typing-inspection>=0.4.0 in /usr/local/lib/python3.11/dist-packages (from pydantic<3,>=1.9.0->openai) (0.4.0)

Note: you may need to restart the kernel to use updated packages.

```
pip show openai
```

Name: openai

Version: 1.75.0

Summary: The official Python library for the openai API

Home-page: <https://github.com/openai/openai-python>

Author:

Author-email: OpenAI <support@openai.com>

License: Apache-2.0

Location: /usr/local/lib/python3.11/dist-packages

Requires: anyio, distro, httpx, jiter, pydantic, sniffio, tqdm, typing-extensions

Required-by: langchain-openai

Note: you may need to restart the kernel to use updated packages.

```
def extract_text_from_pdf(file_path):
    with pdfplumber.open(file_path) as pdf:
        return "\n".join(page.extract_text() for page in pdf.pages if
page.extract_text())
```

```
def extract_text_from_docx(file_path):
    doc = docx.Document(file_path)
    return "\n".join([para.text for para in doc.paragraphs])
```

Replace with actual upload or parsing logic

```
idea_text = "A platform for farmers to sell directly to urban
consumers."
```

```
print("\n[IDEA PARSED]", idea_text)
```

```
[IDEA PARSED] A platform for farmers to sell directly to urban
consumers.
```

```
!pip install pdfplumber python-docx
```

Requirement already satisfied: pdfplumber in /usr/local/lib/python3.11/dist-packages (0.11.6)

Requirement already satisfied: python-docx in /usr/local/lib/python3.11/dist-packages (1.1.2)

Requirement already satisfied: pdfminer.six==20250327 in /usr/local/lib/python3.11/dist-packages (from pdfplumber) (20250327)

Requirement already satisfied: Pillow>=9.1 in /usr/local/lib/python3.11/dist-packages (from pdfplumber) (11.1.0)

Requirement already satisfied: pypdfium2>=4.18.0 in
/usr/local/lib/python3.11/dist-packages (from pdfplumber) (4.30.1)
Requirement already satisfied: charset-normalizer>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six==20250327->pdfplumber) (3.4.1)
Requirement already satisfied: cryptography>=36.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six==20250327->pdfplumber) (44.0.2)
Requirement already satisfied: lxml>=3.1.0 in
/usr/local/lib/python3.11/dist-packages (from python-docx) (5.3.1)
Requirement already satisfied: typing-extensions>=4.9.0 in
/usr/local/lib/python3.11/dist-packages (from python-docx) (4.13.1)
Requirement already satisfied: cffi>=1.12 in
/usr/local/lib/python3.11/dist-packages (from cryptography>=36.0.0->pdfminer.six==20250327->pdfplumber) (1.17.1)
Requirement already satisfied: pycparser in
/usr/local/lib/python3.11/dist-packages (from cffi>=1.12->cryptography>=36.0.0->pdfminer.six==20250327->pdfplumber) (2.22)

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

Requirement already satisfied: faiss-cpu in
/usr/local/lib/python3.11/dist-packages (1.7.4)
Requirement already satisfied: langchain in
/usr/local/lib/python3.11/dist-packages (0.3.23)
Requirement already satisfied: sentence-transformers in
/usr/local/lib/python3.11/dist-packages (4.1.0)
Collecting langchain-core<1.0.0,>=0.3.51 (from langchain)
Using cached langchain_core-0.3.54-py3-none-any.whl.metadata (5.9
kB)
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.8
in /usr/local/lib/python3.11/dist-packages (from langchain) (0.3.8)
Requirement already satisfied: langsmith<0.4,>=0.1.17 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.1.147)
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.11.3)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.0.38)
Requirement already satisfied: requests<3,>=2 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.32.3)
Requirement already satisfied: PyYAML>=5.3 in
/usr/local/lib/python3.11/dist-packages (from langchain) (6.0.2)
Requirement already satisfied: transformers<5.0.0,>=4.41.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(4.51.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-
packages (from sentence-transformers) (4.67.1)
Requirement already satisfied: torch>=1.11.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(2.0.1+cpu)

Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.2.2)

Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.15.2)

Requirement already satisfied: huggingface-hub>=0.20.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(0.30.2)

Requirement already satisfied: Pillow in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(11.1.0)

Requirement already satisfied: typing_extensions>=4.5.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(4.13.1)

Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (3.18.0)

Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (2025.3.2)

Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (23.2)

Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (8.5.0)

Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (1.33)

Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17->langchain) (0.28.1)

Requirement already satisfied: orjson<4.0.0,>=3.9.14 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17->langchain) (3.10.15)

Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17->langchain) (1.0.0)

Requirement already satisfied: annotated-types>=0.6.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langchain) (0.7.0)

Requirement already satisfied: pydantic-core==2.33.1 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langchain) (2.33.1)

Requirement already satisfied: typing-inspection>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langchain) (0.4.0)

Requirement already satisfied: charset-normalizer<4,>=2 in

/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (2025.1.31)
Requirement already satisfied: greenlet!=0.4.17 in
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4->langchain) (3.1.1)
Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (1.13.1)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.1.6)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers<5.0.0,>=4.41.0->sentence-transformers) (1.26.4)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers<5.0.0,>=4.41.0->sentence-transformers) (2024.11.6)
Requirement already satisfied: tokenizers<0.22,>=0.21 in
/usr/local/lib/python3.11/dist-packages (from transformers<5.0.0,>=4.41.0->sentence-transformers) (0.21.1)
Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from transformers<5.0.0,>=4.41.0->sentence-transformers) (0.5.2)
Requirement already satisfied: joblib>=1.1.1 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-transformers) (1.4.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-transformers) (3.6.0)
Requirement already satisfied: anyio in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->langsmith<0.4,>=0.1.17->langchain) (3.7.1)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->langsmith<0.4,>=0.1.17->langchain) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*-

```
>httpx<1,>=0.23.0->langsmith<0.4,>=0.1.17->langchain) (0.14.0)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33-
>langchain-core<1.0.0,>=0.3.51->langchain) (3.0.0)
Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers<5.0.0,>=4.41.0->sentence-transformers) (1.3.8)
Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers<5.0.0,>=4.41.0->sentence-transformers) (1.2.4)
Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers<5.0.0,>=4.41.0->sentence-transformers) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy>=1.17->transformers<5.0.0,>=4.41.0->sentence-
transformers) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers<5.0.0,>=4.41.0->sentence-transformers) (2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers<5.0.0,>=4.41.0->sentence-transformers) (2.4.1)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.11.0-
>sentence-transformers) (3.0.2)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.11.0-
>sentence-transformers) (1.3.0)
Requirement already satisfied: sniffio>=1.1 in
/usr/local/lib/python3.11/dist-packages (from anyio->httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.17->langchain) (1.3.1)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers<5.0.0,>=4.41.0->sentence-transformers) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers<5.0.0,>=4.41.0->sentence-transformers) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->transformers<5.0.0,>=4.41.0->sentence-transformers)
(1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-
>transformers<5.0.0,>=4.41.0->sentence-transformers) (2024.2.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->transformers<5.0.0,>=4.41.0-
>sentence-transformers) (2024.2.0)
Using cached langchain_core-0.3.54-py3-none-any.whl (433 kB)
```

```
Installing collected packages: langchain-core
  Attempting uninstall: langchain-core
    Found existing installation: langchain-core 0.1.53
    Uninstalling langchain-core-0.1.53:
      Successfully uninstalled langchain-core-0.1.53
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
langchain-community 0.0.29 requires langchain-core<0.2.0,>=0.1.33, but
you have langchain-core 0.3.54 which is incompatible.
Successfully installed langchain-core-0.3.54
```

```
!pip install pdfplumber pdfminer.six pdf2image pytesseract pillow
```

```
Requirement already satisfied: pdfplumber in
/usr/local/lib/python3.11/dist-packages (0.11.6)
Requirement already satisfied: pdfminer.six in
/usr/local/lib/python3.11/dist-packages (20250327)
Requirement already satisfied: pdf2image in
/usr/local/lib/python3.11/dist-packages (1.17.0)
Requirement already satisfied: pytesseract in
/usr/local/lib/python3.11/dist-packages (0.3.13)
Requirement already satisfied: pillow in
/usr/local/lib/python3.11/dist-packages (11.1.0)
Requirement already satisfied: pypdfium2>=4.18.0 in
/usr/local/lib/python3.11/dist-packages (from pdfplumber) (4.30.1)
Requirement already satisfied: charset-normalizer>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six) (3.4.1)
Requirement already satisfied: cryptography>=36.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six) (44.0.2)
Requirement already satisfied: packaging>=21.3 in
/usr/local/lib/python3.11/dist-packages (from pytesseract) (23.2)
Requirement already satisfied: cffi>=1.12 in
/usr/local/lib/python3.11/dist-packages (from cryptography>=36.0.0-
>pdfminer.six) (1.17.1)
Requirement already satisfied: pycparser in
/usr/local/lib/python3.11/dist-packages (from cffi>=1.12-
>cryptography>=36.0.0->pdfminer.six) (2.22)
```

```
import os
```

```
# Now you can use all os module functions
```

```
print(os.listdir()) # List current directory contents
```

```
['.virtual_documents', 'test_document.docx',
'business_analysis_report.json']
```

```
!pip install langchain
```

```
Requirement already satisfied: langchain in
/usr/local/lib/python3.11/dist-packages (0.3.23)
```

Requirement already satisfied: langchain-core<1.0.0,>=0.3.51 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.3.54)
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.8
in /usr/local/lib/python3.11/dist-packages (from langchain) (0.3.8)
Requirement already satisfied: langsmith<0.4,>=0.1.17 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.1.147)
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.11.3)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.0.38)
Requirement already satisfied: requests<3,>=2 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.32.3)
Requirement already satisfied: PyYAML>=5.3 in
/usr/local/lib/python3.11/dist-packages (from langchain) (6.0.2)
Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (8.5.0)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (1.33)
Requirement already satisfied: packaging<25,>=23.2 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (23.2)
Requirement already satisfied: typing-extensions>=4.7 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (4.13.1)
Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17-
>langchain) (0.28.1)
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17-
>langchain) (3.10.15)
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17-
>langchain) (1.0.0)
Requirement already satisfied: annotated-types>=0.6.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.1 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (2.33.1)
Requirement already satisfied: typing-inspection>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (0.4.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-

```

>langchain) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (2025.1.31)
Requirement already satisfied: greenlet!=0.4.17 in
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4-
>langchain) (3.1.1)
Requirement already satisfied: anyio in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.17->langchain) (3.7.1)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.17->langchain) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*-
>httpx<1,>=0.23.0->langsmith<0.4,>=0.1.17->langchain) (0.14.0)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33-
>langchain-core<1.0.0,>=0.3.51->langchain) (3.0.0)
Requirement already satisfied: sniffio>=1.1 in
/usr/local/lib/python3.11/dist-packages (from anyio->httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.17->langchain) (1.3.1)

```

```

import os
print(os.listdir('/kaggle/working')) # After uploading
test_document.docx

```

```

['.virtual_documents', 'test_document.docx',
'business_analysis_report.json']

```

```

!pip install langchain openai tiktoken

```

```

Requirement already satisfied: langchain in
/usr/local/lib/python3.11/dist-packages (0.3.23)
Requirement already satisfied: openai in
/usr/local/lib/python3.11/dist-packages (1.75.0)
Requirement already satisfied: tiktoken in
/usr/local/lib/python3.11/dist-packages (0.9.0)
Requirement already satisfied: langchain-core<1.0.0,>=0.3.51 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.3.54)
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.8
in /usr/local/lib/python3.11/dist-packages (from langchain) (0.3.8)
Requirement already satisfied: langsmith<0.4,>=0.1.17 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.1.147)
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.11.3)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in

```

/usr/local/lib/python3.11/dist-packages (from langchain) (2.0.38)
Requirement already satisfied: requests<3,>=2 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.32.3)
Requirement already satisfied: PyYAML>=5.3 in
/usr/local/lib/python3.11/dist-packages (from langchain) (6.0.2)
Requirement already satisfied: anyio<5,>=3.5.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (3.7.1)
Requirement already satisfied: distro<2,>=1.7.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (1.9.0)
Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (0.28.1)
Requirement already satisfied: jiter<1,>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (0.8.2)
Requirement already satisfied: sniffio in
/usr/local/lib/python3.11/dist-packages (from openai) (1.3.1)
Requirement already satisfied: tqdm>4 in
/usr/local/lib/python3.11/dist-packages (from openai) (4.67.1)
Requirement already satisfied: typing-extensions<5,>=4.11 in
/usr/local/lib/python3.11/dist-packages (from openai) (4.13.1)
Requirement already satisfied: regex>=2022.1.18 in
/usr/local/lib/python3.11/dist-packages (from tiktoken) (2024.11.6)
Requirement already satisfied: idna>=2.8 in
/usr/local/lib/python3.11/dist-packages (from anyio<5,>=3.5.0->openai)
(3.10)
Requirement already satisfied: certifi in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->openai) (2025.1.31)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->openai) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*->httpx<1,>=0.23.0->openai) (0.14.0)
Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (8.5.0)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (1.33)
Requirement already satisfied: packaging<25,>=23.2 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (23.2)
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17->langchain) (3.10.15)
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17->langchain) (1.0.0)
Requirement already satisfied: annotated-types>=0.6.0 in


```

/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.1 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (2.33.1)
Requirement already satisfied: typing-inspection>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (0.4.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (3.4.1)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (2.3.0)
Requirement already satisfied: greenlet!=0.4.17 in
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4-
>langchain) (3.1.1)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33-
>langchain-core<1.0.0,>=0.3.51->langchain) (3.0.0)

from langchain.embeddings import OpenAIEmbeddings

import os
os.environ["OPENAI_API_KEY"] = "sk-proj-e3Rdc81zKs-
xkhiszp7Ifi_0vdHl1VoSGebxTVilybC7N0Qrs0nHjc08sIjHXE6_92et0PQpUdT3BlbkF
JGkQP0u64EZ7XArnQknUg8BlSSYHpIbzufjyMnin7jcaZVRhgi7sExuDDU12wBQN8UZ1NX
L0joA" # Replace with your actual key

!pip install torch==2.0.1 --extra-index-url
https://download.pytorch.org/whl/cu118
!pip install transformers==4.33.3 sentence-transformers==2.2.2
!pip install langchain openai pdfplumber python-docx faiss-cpu

Looking in indexes: https://pypi.org/simple,
https://download.pytorch.org/whl/cu118
Requirement already satisfied: torch==2.0.1 in
/usr/local/lib/python3.11/dist-packages (2.0.1+cpu)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (3.18.0)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (4.13.1)
Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (1.13.1)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (3.1.6)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch==2.0.1)

```

```

(3.0.2)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch==2.0.1)
(1.3.0)
Collecting transformers==4.33.3
  Downloading transformers-4.33.3-py3-none-any.whl.metadata (119 kB)
----- 119.9/119.9 kB 2.5 MB/s eta
0:00:00a 0:00:01
ers==2.2.2
  Downloading sentence-transformers-2.2.2.tar.gz (85 kB)
----- 86.0/86.0 kB 5.0 MB/s eta
0:00:00
etadate (setup.py) ... ent already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(3.18.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.15.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(0.30.2)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(1.26.4)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(23.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(6.0.2)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(2024.11.6)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(2.32.3)
Collecting tokenizers!=0.11.3,<0.14,>=0.11.1 (from
transformers==4.33.3)
  Using cached tokenizers-0.13.3-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (6.7 kB)
Requirement already satisfied: safetensors>=0.3.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(0.5.2)
Requirement already satisfied: tqdm>=4.27 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(4.67.1)
Requirement already satisfied: torch>=1.6.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (2.0.1+cpu)
Requirement already satisfied: torchvision in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (0.15.2+cpu)

```

Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (1.2.2)

Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (1.15.2)

Requirement already satisfied: nltk in /usr/local/lib/python3.11/dist-
packages (from sentence-transformers==2.2.2) (3.9.1)

Requirement already satisfied: sentencepiece in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (0.2.0)

Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-
hub<1.0,>=0.15.1->transformers==4.33.3) (2025.3.2)

Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-
hub<1.0,>=0.15.1->transformers==4.33.3) (4.13.1)

Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (1.3.8)

Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (1.2.4)

Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (0.1.1)

Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy>=1.17->transformers==4.33.3) (2025.1.0)

Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (2022.1.0)

Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (2.4.1)

Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (1.13.1)

Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (3.4.2)

Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (3.1.6)

Requirement already satisfied: click in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-
transformers==2.2.2) (8.1.8)

Requirement already satisfied: joblib in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-
transformers==2.2.2) (1.4.2)

```

Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.33.3) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.33.3) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.33.3) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.33.3) (2025.1.31)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers==2.2.2) (3.6.0)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/usr/local/lib/python3.11/dist-packages (from torchvision->sentence-
transformers==2.2.2) (11.1.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.6.0-
>sentence-transformers==2.2.2) (3.0.2)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers==4.33.3) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers==4.33.3) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->transformers==4.33.3) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-
>transformers==4.33.3) (2024.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.6.0-
>sentence-transformers==2.2.2) (1.3.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->transformers==4.33.3) (2024.2.0)
Downloading transformers-4.33.3-py3-none-any.whl (7.6 MB)
7.6/7.6 MB 63.4 MB/s eta
0:00:00:00:010:01
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (7.8 MB)
Building wheels for collected packages: sentence-transformers
  Building wheel for sentence-transformers (setup.py) ... ers:
filename=sentence_transformers-2.2.2-py3-none-any.whl size=125924
sha256=c392d12d66199cd3adcf276232491723adbab6396aca82c726bf8541298f7c6
a
  Stored in directory:

```

```
/root/.cache/pip/wheels/ff/27/bf/ffba8b318b02d7f691a57084ee154e26ed24d012b0c7805881
Successfully built sentence-transformers
Installing collected packages: tokenizers, transformers, sentence-transformers
  Attempting uninstall: tokenizers
    Found existing installation: tokenizers 0.21.1
    Uninstalling tokenizers-0.21.1:
      Successfully uninstalled tokenizers-0.21.1
  Attempting uninstall: transformers
    Found existing installation: transformers 4.51.3
    Uninstalling transformers-4.51.3:
      Successfully uninstalled transformers-4.51.3
  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-2.2.2 tokenizers-0.13.3 transformers-4.33.3
Requirement already satisfied: langchain in
/usr/local/lib/python3.11/dist-packages (0.3.23)
Requirement already satisfied: openai in
/usr/local/lib/python3.11/dist-packages (1.75.0)
Requirement already satisfied: pdfplumber in
/usr/local/lib/python3.11/dist-packages (0.11.6)
Requirement already satisfied: python-docx in
/usr/local/lib/python3.11/dist-packages (1.1.2)
Requirement already satisfied: faiss-cpu in
/usr/local/lib/python3.11/dist-packages (1.7.4)
Requirement already satisfied: langchain-core<1.0.0,>=0.3.51 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.3.54)
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.8
in /usr/local/lib/python3.11/dist-packages (from langchain) (0.3.8)
Requirement already satisfied: langsmith<0.4,>=0.1.17 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.1.147)
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.11.3)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.0.38)
Requirement already satisfied: requests<3,>=2 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.32.3)
Requirement already satisfied: PyYAML>=5.3 in
/usr/local/lib/python3.11/dist-packages (from langchain) (6.0.2)
Requirement already satisfied: anyio<5,>=3.5.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (3.7.1)
Requirement already satisfied: distro<2,>=1.7.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (1.9.0)
Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (0.28.1)
```

Requirement already satisfied: jiter<1,>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from openai) (0.8.2)
Requirement already satisfied: sniffio in
/usr/local/lib/python3.11/dist-packages (from openai) (1.3.1)
Requirement already satisfied: tqdm>4 in
/usr/local/lib/python3.11/dist-packages (from openai) (4.67.1)
Requirement already satisfied: typing-extensions<5,>=4.11 in
/usr/local/lib/python3.11/dist-packages (from openai) (4.13.1)
Requirement already satisfied: pdfminer.six==20250327 in
/usr/local/lib/python3.11/dist-packages (from pdfplumber) (20250327)
Requirement already satisfied: Pillow>=9.1 in
/usr/local/lib/python3.11/dist-packages (from pdfplumber) (11.1.0)
Requirement already satisfied: pypdfium2>=4.18.0 in
/usr/local/lib/python3.11/dist-packages (from pdfplumber) (4.30.1)
Requirement already satisfied: charset-normalizer>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six==20250327->pdfplumber) (3.4.1)
Requirement already satisfied: cryptography>=36.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six==20250327->pdfplumber) (44.0.2)
Requirement already satisfied: lxml>=3.1.0 in
/usr/local/lib/python3.11/dist-packages (from python-docx) (5.3.1)
Requirement already satisfied: idna>=2.8 in
/usr/local/lib/python3.11/dist-packages (from anyio<5,>=3.5.0->openai) (3.10)
Requirement already satisfied: certifi in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->openai) (2025.1.31)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->openai) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*->httpx<1,>=0.23.0->openai) (0.14.0)
Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (8.5.0)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (1.33)
Requirement already satisfied: packaging<25,>=23.2 in
/usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.51->langchain) (23.2)
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17->langchain) (3.10.15)
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17->langchain) (1.0.0)

Requirement already satisfied: annotated-types>=0.6.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langchain) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.1 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langchain) (2.33.1)
Requirement already satisfied: typing-inspection>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langchain) (0.4.0)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (2.3.0)
Requirement already satisfied: greenlet!=0.4.17 in
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4->langchain) (3.1.1)
Requirement already satisfied: cffi>=1.12 in
/usr/local/lib/python3.11/dist-packages (from cryptography>=36.0.0->pdfminer.six==20250327->pdfplumber) (1.17.1)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33->langchain-core<1.0.0,>=0.3.51->langchain) (3.0.0)
Requirement already satisfied: pycparser in
/usr/local/lib/python3.11/dist-packages (from cffi>=1.12->cryptography>=36.0.0->pdfminer.six==20250327->pdfplumber) (2.22)

```
!pip install torch --pre --extra-index-url  
https://download.pytorch.org/whl/nightly/cu118  
!pip install --no-deps sentence-transformers  
!pip install transformers
```

Looking in indexes: <https://pypi.org/simple>,
<https://download.pytorch.org/whl/nightly/cu118>
Requirement already satisfied: torch in
/usr/local/lib/python3.11/dist-packages (2.0.1+cpu)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from torch) (3.18.0)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.11/dist-packages (from torch) (4.13.1)
Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch) (1.13.1)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch) (3.1.6)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch) (3.0.2)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch) (1.3.0)
Requirement already satisfied: sentence-transformers in
/usr/local/lib/python3.11/dist-packages (2.2.2)

Requirement already satisfied: transformers in
/usr/local/lib/python3.11/dist-packages (4.33.3)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from transformers) (3.18.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.15.1 in
/usr/local/lib/python3.11/dist-packages (from transformers) (0.30.2)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers) (1.26.4)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from transformers) (23.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from transformers) (6.0.2)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers)
(2024.11.6)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from transformers) (2.32.3)
Requirement already satisfied: tokenizers!=0.11.3,<0.14,>=0.11.1 in
/usr/local/lib/python3.11/dist-packages (from transformers) (0.13.3)
Requirement already satisfied: safetensors>=0.3.1 in
/usr/local/lib/python3.11/dist-packages (from transformers) (0.5.2)
Requirement already satisfied: tqdm>=4.27 in
/usr/local/lib/python3.11/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-
hub<1.0,>=0.15.1->transformers) (2025.3.2)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-
hub<1.0,>=0.15.1->transformers) (4.13.1)
Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers) (1.3.8)
Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers) (1.2.4)
Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy>=1.17->transformers) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers) (2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers) (2.4.1)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers)
(3.4.1)

Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers)
(3.10)

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

Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers)
(2025.1.31)

Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers) (2024.2.0)

Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers) (2022.1.0)

Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->transformers) (1.2.0)

Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-
>transformers) (2024.2.0)

Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->transformers) (2024.2.0)

!pip install --upgrade huggingface_hub==0.19.4 transformers==4.33.3
sentence-transformers==2.2.2

Collecting huggingface_hub==0.19.4

Downloading huggingface_hub-0.19.4-py3-none-any.whl.metadata (14 kB)

Requirement already satisfied: transformers==4.33.3 in
/usr/local/lib/python3.11/dist-packages (4.33.3)

Requirement already satisfied: sentence-transformers==2.2.2 in
/usr/local/lib/python3.11/dist-packages (2.2.2)

Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4)
(3.18.0)

Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4)
(2025.3.2)

Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4)
(2.32.3)

Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4)
(4.67.1)

Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4)
(6.0.2)

Requirement already satisfied: typing-extensions>=3.7.4.3 in

/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4) (4.13.1)
Requirement already satisfied: packaging>=20.9 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4) (23.2)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from transformers==4.33.3) (1.26.4)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.11/dist-packages (from transformers==4.33.3) (2024.11.6)
Requirement already satisfied: tokenizers!=0.11.3,<0.14,>=0.11.1 in /usr/local/lib/python3.11/dist-packages (from transformers==4.33.3) (0.13.3)
Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3.11/dist-packages (from transformers==4.33.3) (0.5.2)
Requirement already satisfied: torch>=1.6.0 in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (2.0.1+cpu)
Requirement already satisfied: torchvision in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (0.15.2+cpu)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (1.2.2)
Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (1.15.2)
Requirement already satisfied: nltk in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (3.9.1)
Requirement already satisfied: sentencepiece in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (0.2.0)
Requirement already satisfied: mkl_fft in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (1.3.8)
Requirement already satisfied: mkl_random in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (1.2.4)
Requirement already satisfied: mkl_umath in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (2025.1.0)
Requirement already satisfied: tbb4py in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (2022.1.0)
Requirement already satisfied: mkl-service in

/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (2.4.1)
Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (1.13.1)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (3.1.6)
Requirement already satisfied: click in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-
transformers==2.2.2) (8.1.8)
Requirement already satisfied: joblib in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-
transformers==2.2.2) (1.4.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub==0.19.4) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub==0.19.4) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub==0.19.4) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub==0.19.4) (2025.1.31)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers==2.2.2) (3.6.0)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/usr/local/lib/python3.11/dist-packages (from torchvision->sentence-
transformers==2.2.2) (11.1.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.6.0-
>sentence-transformers==2.2.2) (3.0.2)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers==4.33.3) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers==4.33.3) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->transformers==4.33.3) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-

```
>transformers==4.33.3) (2024.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.6.0-
>sentence-transformers==2.2.2) (1.3.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->transformers==4.33.3) (2024.2.0)
Downloading huggingface_hub-0.19.4-py3-none-any.whl (311 kB)
311.7/311.7 kB 6.1 MB/s eta
```

```
0:00:0000:01
```

```
pting uninstall: huggingface_hub
```

```
Found existing installation: huggingface-hub 0.30.2
```

```
Uninstalling huggingface-hub-0.30.2:
```

```
Successfully uninstalled huggingface-hub-0.30.2
```

```
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
```

```
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
have fsspec 2025.3.2 which is incompatible.
```

```
datasets 3.5.0 requires huggingface-hub>=0.24.0, but you have
huggingface-hub 0.19.4 which is incompatible.
```

```
diffusers 0.32.2 requires huggingface-hub>=0.23.2, but you have
huggingface-hub 0.19.4 which is incompatible.
```

```
accelerate 1.3.0 requires huggingface-hub>=0.21.0, but you have
huggingface-hub 0.19.4 which is incompatible.
```

```
peft 0.14.0 requires huggingface-hub>=0.25.0, but you have
huggingface-hub 0.19.4 which is incompatible.
```

```
Successfully installed huggingface_hub-0.19.4
```

```
# Install specific compatible versions
```

```
!pip install torch==2.0.1 --index-url
```

```
https://download.pytorch.org/whl/cu118
```

```
!pip install transformers==4.30.1 sentence-transformers==2.2.2
```

```
!pip install huggingface_hub==0.16.4
```

```
Looking in indexes: https://download.pytorch.org/whl/cu118
```

```
Requirement already satisfied: torch==2.0.1 in
```

```
/usr/local/lib/python3.11/dist-packages (2.0.1+cpu)
```

```
Requirement already satisfied: filelock in
```

```
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (3.18.0)
```

```
Requirement already satisfied: typing-extensions in
```

```
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (4.13.1)
```

```
Requirement already satisfied: sympy in
```

```
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (1.13.1)
```

```
Requirement already satisfied: networkx in
```

```
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (3.4.2)
```

```
Requirement already satisfied: jinja2 in
```

```
/usr/local/lib/python3.11/dist-packages (from torch==2.0.1) (3.1.6)
```

```
Requirement already satisfied: MarkupSafe>=2.0 in
```

```
/usr/local/lib/python3.11/dist-packages (from jinja2->torch==2.0.1)
```

```
(3.0.2)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch==2.0.1)
(1.3.0)
Collecting transformers==4.30.1
  Downloading transformers-4.30.1-py3-none-any.whl.metadata (113 kB)
  113.6/113.6 kB 2.8 MB/s eta
0:00:0000:01
ent already satisfied: sentence-transformers==2.2.2 in
/usr/local/lib/python3.11/dist-packages (2.2.2)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(3.18.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.14.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(0.19.4)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(1.26.4)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(23.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(6.0.2)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(2024.11.6)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(2.32.3)
Requirement already satisfied: tokenizers!=0.11.3,<0.14,>=0.11.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(0.13.3)
Requirement already satisfied: safetensors>=0.3.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(0.5.2)
Requirement already satisfied: tqdm>=4.27 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.30.1)
(4.67.1)
Requirement already satisfied: torch>=1.6.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (2.0.1+cpu)
Requirement already satisfied: torchvision in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (0.15.2+cpu)
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (1.2.2)
```

Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (1.15.2)

Requirement already satisfied: nltk in /usr/local/lib/python3.11/dist-
packages (from sentence-transformers==2.2.2) (3.9.1)

Requirement already satisfied: sentencepiece in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (0.2.0)

Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-
hub<1.0,>=0.14.1->transformers==4.30.1) (2025.3.2)

Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-
hub<1.0,>=0.14.1->transformers==4.30.1) (4.13.1)

Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.30.1) (1.3.8)

Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.30.1) (1.2.4)

Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.30.1) (0.1.1)

Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy>=1.17->transformers==4.30.1) (2025.1.0)

Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.30.1) (2022.1.0)

Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.30.1) (2.4.1)

Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (1.13.1)

Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (3.4.2)

Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (3.1.6)

Requirement already satisfied: click in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-
transformers==2.2.2) (8.1.8)

Requirement already satisfied: joblib in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-
transformers==2.2.2) (1.4.2)

Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.30.1) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers==4.30.1) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers==4.30.1) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers==4.30.1) (2025.1.31)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-transformers==2.2.2) (3.6.0)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/usr/local/lib/python3.11/dist-packages (from torchvision->sentence-transformers==2.2.2) (11.1.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.6.0->sentence-transformers==2.2.2) (3.0.2)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17->transformers==4.30.1) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17->transformers==4.30.1) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl->numpy>=1.17->transformers==4.30.1) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17->transformers==4.30.1) (2024.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.6.0->sentence-transformers==2.2.2) (1.3.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-openmp<2026,>=2024->mkl->numpy>=1.17->transformers==4.30.1) (2024.2.0)
Downloading transformers-4.30.1-py3-none-any.whl (7.2 MB)

7.2/7.2 MB 58.2 MB/s eta

0:00:0000:0100:01

ers

Attempting uninstall: transformers

Found existing installation: transformers 4.33.3

Uninstalling transformers-4.33.3:

Successfully uninstalled transformers-4.33.3

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.

kaggle-environments 1.16.11 requires transformers>=4.33.1, but you have transformers 4.30.1 which is incompatible.

```
peft 0.14.0 requires huggingface-hub>=0.25.0, but you have
huggingface-hub 0.19.4 which is incompatible.
Successfully installed transformers-4.30.1
Collecting huggingface_hub==0.16.4
  Using cached huggingface_hub-0.16.4-py3-none-any.whl.metadata (12
kB)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.16.4)
(3.18.0)
Requirement already satisfied: fsspec in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.16.4)
(2025.3.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.16.4)
(2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.16.4)
(4.67.1)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.16.4)
(6.0.2)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.16.4)
(4.13.1)
Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.16.4)
(23.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub==0.16.4) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub==0.16.4) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub==0.16.4) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub==0.16.4) (2025.1.31)
Using cached huggingface_hub-0.16.4-py3-none-any.whl (268 kB)
Installing collected packages: huggingface_hub
  Attempting uninstall: huggingface_hub
    Found existing installation: huggingface-hub 0.19.4
    Uninstalling huggingface-hub-0.19.4:
      Successfully uninstalled huggingface-hub-0.19.4
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
```


have fsspec 2025.3.2 which is incompatible.
datasets 3.5.0 requires huggingface-hub>=0.24.0, but you have
huggingface-hub 0.16.4 which is incompatible.
kaggle-environments 1.16.11 requires transformers>=4.33.1, but you
have transformers 4.30.1 which is incompatible.
diffusers 0.32.2 requires huggingface-hub>=0.23.2, but you have
huggingface-hub 0.16.4 which is incompatible.
accelerate 1.3.0 requires huggingface-hub>=0.21.0, but you have
huggingface-hub 0.16.4 which is incompatible.
peft 0.14.0 requires huggingface-hub>=0.25.0, but you have
huggingface-hub 0.16.4 which is incompatible.
Successfully installed huggingface_hub-0.16.4

```
!pip install huggingface_hub==0.19.4 --force-reinstall
```

```
Collecting huggingface_hub==0.19.4
  Using cached huggingface_hub-0.19.4-py3-none-any.whl.metadata (14
kB)
Collecting filelock (from huggingface_hub==0.19.4)
  Downloading filelock-3.18.0-py3-none-any.whl.metadata (2.9 kB)
Collecting fsspec>=2023.5.0 (from huggingface_hub==0.19.4)
  Downloading fsspec-2025.3.2-py3-none-any.whl.metadata (11 kB)
Collecting requests (from huggingface_hub==0.19.4)
  Using cached requests-2.32.3-py3-none-any.whl.metadata (4.6 kB)
Collecting tqdm>=4.42.1 (from huggingface_hub==0.19.4)
  Downloading tqdm-4.67.1-py3-none-any.whl.metadata (57 kB)
57.7/57.7 kB 1.7 MB/s eta
0:00:00
l>=5.1 (from huggingface_hub==0.19.4)
  Using cached PyYAML-6.0.2-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (2.1 kB)
Collecting typing_extensions>=3.7.4.3 (from huggingface_hub==0.19.4)
  Downloading typing_extensions-4.13.2-py3-none-any.whl.metadata (3.0
kB)
Collecting packaging>=20.9 (from huggingface_hub==0.19.4)
  Downloading packaging-24.2-py3-none-any.whl.metadata (3.2 kB)
Collecting charset-normalizer<4,>=2 (from requests-
>huggingface_hub==0.19.4)
  Downloading charset_normalizer-3.4.1-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (35 kB)
Collecting idna<4,>=2.5 (from requests->huggingface_hub==0.19.4)
  Downloading idna-3.10-py3-none-any.whl.metadata (10 kB)
Collecting urllib3<3,>=1.21.1 (from requests->huggingface_hub==0.19.4)
  Downloading urllib3-2.4.0-py3-none-any.whl.metadata (6.5 kB)
Collecting certifi>=2017.4.17 (from requests->huggingface_hub==0.19.4)
  Downloading certifi-2025.1.31-py3-none-any.whl.metadata (2.5 kB)
Using cached huggingface_hub-0.19.4-py3-none-any.whl (311 kB)
Downloading fsspec-2025.3.2-py3-none-any.whl (194 kB)
194.4/194.4 kB 6.0 MB/s eta
0:00:00
```

```

65.5/65.5 kB 3.6 MB/s eta
0:00:00
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (762 kB)
763.0/763.0 kB 22.7 MB/s eta
0:00:00
-4.67.1-py3-none-any.whl (78 kB)
78.5/78.5 kB 4.6 MB/s eta
0:00:00
45.8/45.8 kB 2.2 MB/s eta
0:00:00
64.9/64.9 kB 3.0 MB/s eta
0:00:00
166.4/166.4 kB 8.9 MB/s eta
0:00:00
alizer-3.4.1-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (143 kB)
143.9/143.9 kB 7.9 MB/s eta
0:00:00
70.4/70.4 kB 4.0 MB/s eta
0:00:00
128.7/128.7 kB 7.1 MB/s eta
0:00:00
, pyyaml, packaging, idna, fsspec, filelock, charset-normalizer,
certifi, requests, huggingface_hub
Attempting uninstall: urllib3
Found existing installation: urllib3 2.3.0
Uninstalling urllib3-2.3.0:
Successfully uninstalled urllib3-2.3.0
Attempting uninstall: typing-extensions
Found existing installation: typing_extensions 4.13.1
Uninstalling typing_extensions-4.13.1:
Successfully uninstalled typing_extensions-4.13.1
Attempting uninstall: tqdm
Found existing installation: tqdm 4.67.1
Uninstalling tqdm-4.67.1:
Successfully uninstalled tqdm-4.67.1
Attempting uninstall: pyyaml
Found existing installation: PyYAML 6.0.2
Uninstalling PyYAML-6.0.2:
Successfully uninstalled PyYAML-6.0.2
Attempting uninstall: packaging
Found existing installation: packaging 23.2
Uninstalling packaging-23.2:
Successfully uninstalled packaging-23.2
Attempting uninstall: idna
Found existing installation: idna 3.10
Uninstalling idna-3.10:
Successfully uninstalled idna-3.10
Attempting uninstall: fsspec

```

```
Found existing installation: fsspec 2025.3.2
Uninstalling fsspec-2025.3.2:
  Successfully uninstalled fsspec-2025.3.2
Attempting uninstall: filelock
Found existing installation: filelock 3.18.0
Uninstalling filelock-3.18.0:
  Successfully uninstalled filelock-3.18.0
Attempting uninstall: charset-normalizer
Found existing installation: charset-normalizer 3.4.1
Uninstalling charset-normalizer-3.4.1:
  Successfully uninstalled charset-normalizer-3.4.1
Attempting uninstall: certifi
Found existing installation: certifi 2025.1.31
Uninstalling certifi-2025.1.31:
  Successfully uninstalled certifi-2025.1.31
Attempting uninstall: requests
Found existing installation: requests 2.32.3
Uninstalling requests-2.32.3:
  Successfully uninstalled requests-2.32.3
Attempting uninstall: huggingface_hub
Found existing installation: huggingface-hub 0.16.4
Uninstalling huggingface-hub-0.16.4:
  Successfully uninstalled huggingface-hub-0.16.4
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
langchain-community 0.0.29 requires langchain-core<0.2.0,>=0.1.33, but
you have langchain-core 0.3.54 which is incompatible.
sigstore 3.6.1 requires rich~=13.0, but you have rich 14.0.0 which is
incompatible.
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
have fsspec 2025.3.2 which is incompatible.
datasets 3.5.0 requires huggingface-hub>=0.24.0, but you have
huggingface-hub 0.19.4 which is incompatible.
pytorch-lightning 2.5.1 requires torch>=2.1.0, but you have torch
2.0.1+cpu which is incompatible.
nilearn 0.11.1 requires scikit-learn>=1.4.0, but you have scikit-learn
1.2.2 which is incompatible.
kaggle-environments 1.16.11 requires transformers>=4.33.1, but you
have transformers 4.30.1 which is incompatible.
preprocessing 0.1.13 requires nltk==3.2.4, but you have nltk 3.9.1
which is incompatible.
google-colab 1.0.0 requires notebook==6.5.5, but you have notebook
6.5.4 which is incompatible.
google-colab 1.0.0 requires pandas==2.2.2, but you have pandas 2.2.3
which is incompatible.
dopamine-rl 4.1.2 requires gymnasium>=1.0.0, but you have gymnasium
0.29.0 which is incompatible.
diffusers 0.32.2 requires huggingface-hub>=0.23.2, but you have
```

huggingface-hub 0.19.4 which is incompatible.
accelerate 1.3.0 requires huggingface-hub>=0.21.0, but you have huggingface-hub 0.19.4 which is incompatible.
google-spark-connect 0.5.2 requires google-api-core>=2.19.1, but you have google-api-core 1.34.1 which is incompatible.
pandas-gbq 0.26.1 requires google-api-core<3.0.0dev,>=2.10.2, but you have google-api-core 1.34.1 which is incompatible.
gcsfs 2024.10.0 requires fsspec==2024.10.0, but you have fsspec 2025.3.2 which is incompatible.
bigframes 1.36.0 requires rich<14,>=12.4.4, but you have rich 14.0.0 which is incompatible.
pydrive2 1.21.3 requires cryptography<44, but you have cryptography 44.0.2 which is incompatible.
pydrive2 1.21.3 requires pyOpenSSL<=24.2.1,>=19.1.0, but you have pyopenssl 25.0.0 which is incompatible.
peft 0.14.0 requires huggingface-hub>=0.25.0, but you have huggingface-hub 0.19.4 which is incompatible.
ibis-framework 9.2.0 requires rich<14,>=12.4.4, but you have rich 14.0.0 which is incompatible.
ibis-framework 9.2.0 requires toolz<1,>=0.11, but you have toolz 1.0.0 which is incompatible.
plotnine 0.14.5 requires matplotlib>=3.8.0, but you have matplotlib 3.7.5 which is incompatible.
torchaudio 2.5.1+cu124 requires torch==2.5.1, but you have torch 2.0.1+cpu which is incompatible.
google-cloud-bigtable 2.28.1 requires google-api-core[grpc]<3.0.0dev,>=2.16.0, but you have google-api-core 1.34.1 which is incompatible.
mlxtend 0.23.4 requires scikit-learn>=1.3.1, but you have scikit-learn 1.2.2 which is incompatible.
Successfully installed certifi-2025.1.31 charset-normalizer-3.4.1 filelock-3.18.0 fsspec-2025.3.2 huggingface_hub-0.19.4 idna-3.10 packaging-24.2 pyyaml-6.0.2 requests-2.32.3 tqdm-4.67.1 typing-extensions-4.13.2 urllib3-2.4.0

```
# Old code (causing error)
# from huggingface_hub import cached_download
```

```
# New code
```

```
from huggingface_hub import hf_hub_download as cached_download
```

```
!pip uninstall -y transformers sentence-transformers huggingface_hub
!pip install transformers==4.33.3 sentence-transformers==2.2.2
huggingface_hub==0.19.4
!pip install --force-reinstall torch==2.0.1 --index-url
https://download.pytorch.org/whl/cu118
```

```
Found existing installation: transformers 4.30.1
Uninstalling transformers-4.30.1:
  Successfully uninstalled transformers-4.30.1
```

```
Found existing installation: sentence-transformers 2.2.2
Uninstalling sentence-transformers-2.2.2:
  Successfully uninstalled sentence-transformers-2.2.2
Found existing installation: huggingface-hub 0.19.4
Uninstalling huggingface-hub-0.19.4:
  Successfully uninstalled huggingface-hub-0.19.4
Collecting transformers==4.33.3
  Using cached transformers-4.33.3-py3-none-any.whl.metadata (119 kB)
Collecting sentence-transformers==2.2.2
  Using cached sentence_transformers-2.2.2-py3-none-any.whl
Collecting huggingface_hub==0.19.4
  Using cached huggingface_hub-0.19.4-py3-none-any.whl.metadata (14
kB)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(3.18.0)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(1.26.4)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(24.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(6.0.2)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(2024.11.6)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(2.32.3)
Requirement already satisfied: tokenizers!=0.11.3,<0.14,>=0.11.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(0.13.3)
Requirement already satisfied: safetensors>=0.3.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(0.5.2)
Requirement already satisfied: tqdm>=4.27 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(4.67.1)
Requirement already satisfied: torch>=1.6.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (2.0.1+cpu)
Requirement already satisfied: torchvision in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (0.15.2+cpu)
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (1.2.2)
```

Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (1.15.2)

Requirement already satisfied: nltk in /usr/local/lib/python3.11/dist-
packages (from sentence-transformers==2.2.2) (3.9.1)

Requirement already satisfied: sentencepiece in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers==2.2.2) (0.2.0)

Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4)
(2025.3.2)

Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4)
(4.13.2)

Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (1.3.8)

Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (1.2.4)

Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (0.1.1)

Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy>=1.17->transformers==4.33.3) (2025.1.0)

Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (2022.1.0)

Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers==4.33.3) (2.4.1)

Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (1.13.1)

Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (3.4.2)

Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-
transformers==2.2.2) (3.1.6)

Requirement already satisfied: click in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-
transformers==2.2.2) (8.1.8)

Requirement already satisfied: joblib in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-
transformers==2.2.2) (1.4.2)

Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.33.3) (3.4.1)

```
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.33.3) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.33.3) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests-
>transformers==4.33.3) (2025.1.31)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers==2.2.2) (3.6.0)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/usr/local/lib/python3.11/dist-packages (from torchvision->sentence-
transformers==2.2.2) (11.1.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.6.0-
>sentence-transformers==2.2.2) (3.0.2)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers==4.33.3) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers==4.33.3) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->transformers==4.33.3) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-
>transformers==4.33.3) (2024.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.6.0-
>sentence-transformers==2.2.2) (1.3.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->transformers==4.33.3) (2024.2.0)
Using cached transformers-4.33.3-py3-none-any.whl (7.6 MB)
Using cached huggingface_hub-0.19.4-py3-none-any.whl (311 kB)
Installing collected packages: huggingface_hub, transformers,
sentence-transformers
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
have fsspec 2025.3.2 which is incompatible.
datasets 3.5.0 requires huggingface-hub>=0.24.0, but you have
huggingface-hub 0.19.4 which is incompatible.
diffusers 0.32.2 requires huggingface-hub>=0.23.2, but you have
huggingface-hub 0.19.4 which is incompatible.
```

```

accelerate 1.3.0 requires huggingface-hub>=0.21.0, but you have
huggingface-hub 0.19.4 which is incompatible.
peft 0.14.0 requires huggingface-hub>=0.25.0, but you have
huggingface-hub 0.19.4 which is incompatible.
Successfully installed huggingface_hub-0.19.4 sentence-transformers-
2.2.2 transformers-4.33.3
Looking in indexes: https://download.pytorch.org/whl/cu118
Collecting torch==2.0.1
  Downloading https://download.pytorch.org/whl/cu118/torch-
2.0.1%2Bcu118-cp311-cp311-linux_x86_64.whl (2267.3 MB)
----- 2.3/2.3 GB 377.1 kB/s eta
0:00:00:0100:01m
  torch==2.0.1)
    Downloading https://download.pytorch.org/whl/filelock-3.13.1-py3-
none-any.whl.metadata (2.8 kB)
Collecting typing-extensions (from torch==2.0.1)
  Downloading https://download.pytorch.org/whl/typing_extensions-
4.12.2-py3-none-any.whl.metadata (3.0 kB)
Collecting sympy (from torch==2.0.1)
  Downloading https://download.pytorch.org/whl/sympy-1.13.1-py3-none-
any.whl (6.2 MB)
----- 6.2/6.2 MB 81.4 MB/s eta
0:00:00:00:0100:01
  torch==2.0.1)
    Downloading https://download.pytorch.org/whl/networkx-3.3-py3-none-
any.whl.metadata (5.1 kB)
Collecting jinja2 (from torch==2.0.1)
  Downloading https://download.pytorch.org/whl/Jinja2-3.1.4-py3-none-
any.whl.metadata (2.6 kB)
Collecting triton==2.0.0 (from torch==2.0.1)
  Downloading https://download.pytorch.org/whl/triton-2.0.0-1-cp311-
cp311-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (63.3 MB)
----- 63.3/63.3 MB 28.4 MB/s eta
0:00:00:00:0100:01
ake (from triton==2.0.0->torch==2.0.1)
  Downloading https://download.pytorch.org/whl/cmake-3.25.0-py2.py3-
none-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (23.7 MB)
----- 23.7/23.7 MB 62.5 MB/s eta
0:00:00:00:0100:01
  triton==2.0.0->torch==2.0.1)
    Downloading https://download.pytorch.org/whl/lit-15.0.7.tar.gz (132
kB)
----- 132.3/132.3 kB 7.6 MB/s eta
0:00:00
etadata (setup.py) ... jinja2->torch==2.0.1)
  Downloading https://download.pytorch.org/whl/MarkupSafe-2.1.5-cp311-
cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (28 kB)
Collecting mpmath<1.4,>=1.1.0 (from sympy->torch==2.0.1)
  Downloading https://download.pytorch.org/whl/mpmath-1.3.0-py3-none-

```



```
any.whl (536 kB)
----- 536.2/536.2 kB 27.1 MB/s eta
0:00:00
----- 133.3/133.3 kB 8.3 MB/s eta
0:00:00
----- 1.7/1.7 MB 56.4 MB/s eta
0:00:00
e=lit-15.0.7-py3-none-any.whl size=89991
sha256=fd6828c2db3187df5c47b0f4cb5c8c9200810d763b98406738d2b1c97a3399d
b
  Stored in directory:
/root/.cache/pip/wheels/fc/5d/45/34fe9945d5e45e261134e72284395be36c2d4
828af38e2b0fe
Successfully built lit
Installing collected packages: mpmath, lit, cmake, typing-extensions,
sympy, networkx, MarkupSafe, filelock, jinja2, triton, torch
  Attempting uninstall: mpmath
    Found existing installation: mpmath 1.3.0
    Uninstalling mpmath-1.3.0:
      Successfully uninstalled mpmath-1.3.0
  Attempting uninstall: cmake
    Found existing installation: cmake 3.31.4
    Uninstalling cmake-3.31.4:
      Successfully uninstalled cmake-3.31.4
  Attempting uninstall: typing-extensions
    Found existing installation: typing_extensions 4.13.2
    Uninstalling typing_extensions-4.13.2:
      Successfully uninstalled typing_extensions-4.13.2
  Attempting uninstall: sympy
    Found existing installation: sympy 1.13.1
    Uninstalling sympy-1.13.1:
      Successfully uninstalled sympy-1.13.1
  Attempting uninstall: networkx
    Found existing installation: networkx 3.4.2
    Uninstalling networkx-3.4.2:
      Successfully uninstalled networkx-3.4.2
  Attempting uninstall: MarkupSafe
    Found existing installation: MarkupSafe 3.0.2
    Uninstalling MarkupSafe-3.0.2:
      Successfully uninstalled MarkupSafe-3.0.2
  Attempting uninstall: filelock
    Found existing installation: filelock 3.18.0
    Uninstalling filelock-3.18.0:
      Successfully uninstalled filelock-3.18.0
  Attempting uninstall: jinja2
    Found existing installation: Jinja2 3.1.6
    Uninstalling Jinja2-3.1.6:
      Successfully uninstalled Jinja2-3.1.6
  Attempting uninstall: triton
```

```
Found existing installation: triton 3.1.0
Uninstalling triton-3.1.0:
  Successfully uninstalled triton-3.1.0
Attempting uninstall: torch
Found existing installation: torch 2.0.1+cpu
Uninstalling torch-2.0.1+cpu:
  Successfully uninstalled torch-2.0.1+cpu
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
langchain-community 0.0.29 requires langchain-core<0.2.0,>=0.1.33, but
you have langchain-core 0.3.54 which is incompatible.
sigstore 3.6.1 requires rich~=13.0, but you have rich 14.0.0 which is
incompatible.
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
have fsspec 2025.3.2 which is incompatible.
datasets 3.5.0 requires huggingface-hub>=0.24.0, but you have
huggingface-hub 0.19.4 which is incompatible.
pytorch-lightning 2.5.1 requires torch>=2.1.0, but you have torch
2.0.1+cu118 which is incompatible.
google-colab 1.0.0 requires notebook==6.5.5, but you have notebook
6.5.4 which is incompatible.
google-colab 1.0.0 requires pandas==2.2.2, but you have pandas 2.2.3
which is incompatible.
dopamine-rl 4.1.2 requires gymnasium>=1.0.0, but you have gymnasium
0.29.0 which is incompatible.
diffusers 0.32.2 requires huggingface-hub>=0.23.2, but you have
huggingface-hub 0.19.4 which is incompatible.
accelerate 1.3.0 requires huggingface-hub>=0.21.0, but you have
huggingface-hub 0.19.4 which is incompatible.
bigframes 1.36.0 requires rich<14,>=12.4.4, but you have rich 14.0.0
which is incompatible.
pydrive2 1.21.3 requires cryptography<44, but you have cryptography
44.0.2 which is incompatible.
pydrive2 1.21.3 requires pyOpenSSL<=24.2.1,>=19.1.0, but you have
pyopenssl 25.0.0 which is incompatible.
peft 0.14.0 requires huggingface-hub>=0.25.0, but you have
huggingface-hub 0.19.4 which is incompatible.
pytensor 2.27.1 requires filelock>=3.15, but you have filelock 3.13.1
which is incompatible.
ibis-framework 9.2.0 requires rich<14,>=12.4.4, but you have rich
14.0.0 which is incompatible.
ibis-framework 9.2.0 requires toolz<1,>=0.11, but you have toolz 1.0.0
which is incompatible.
torchaudio 2.5.1+cu124 requires torch==2.5.1, but you have torch
2.0.1+cu118 which is incompatible.
Successfully installed MarkupSafe-2.1.5 cmake-3.25.0 filelock-3.13.1
jinja2-3.1.4 lit-15.0.7 mpmath-1.3.0 networkx-3.3 sympy-1.13.1 torch-
2.0.1+cu118 triton-2.0.0 typing-extensions-4.12.2
```

```

from huggingface_hub import hf_hub_download
print("Import successful!")

Import successful!

# Replace this old import:
# from huggingface_hub import cached_download

# With this new import:
from huggingface_hub import hf_hub_download as cached_download

!pip uninstall -y transformers sentence-transformers huggingface_hub
!pip install transformers==4.33.3 sentence-transformers==2.2.2
huggingface_hub==0.19.4

Found existing installation: transformers 4.33.3
Uninstalling transformers-4.33.3:
  Successfully uninstalled transformers-4.33.3
Found existing installation: sentence-transformers 2.2.2
Uninstalling sentence-transformers-2.2.2:
  Successfully uninstalled sentence-transformers-2.2.2
Found existing installation: huggingface-hub 0.19.4
Uninstalling huggingface-hub-0.19.4:
  Successfully uninstalled huggingface-hub-0.19.4
Collecting transformers==4.33.3
  Using cached transformers-4.33.3-py3-none-any.whl.metadata (119 kB)
Collecting sentence-transformers==2.2.2
  Using cached sentence_transformers-2.2.2-py3-none-any.whl
Collecting huggingface_hub==0.19.4
  Using cached huggingface_hub-0.19.4-py3-none-any.whl.metadata (14
kB)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(3.13.1)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(1.26.4)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(24.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(6.0.2)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(2024.11.6)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3)
(2.32.3)
Requirement already satisfied: tokenizers!=0.11.3,<0.14,>=0.11.1 in

```

/usr/local/lib/python3.11/dist-packages (from transformers==4.33.3) (0.13.3)
Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3.11/dist-packages (from transformers==4.33.3) (0.5.2)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.11/dist-packages (from transformers==4.33.3) (4.67.1)
Requirement already satisfied: torch>=1.6.0 in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (2.0.1+cu118)
Requirement already satisfied: torchvision in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (0.15.2+cpu)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (1.2.2)
Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (1.15.2)
Requirement already satisfied: nltk in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (3.9.1)
Requirement already satisfied: sentencepiece in /usr/local/lib/python3.11/dist-packages (from sentence-transformers==2.2.2) (0.2.0)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4) (2025.3.2)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub==0.19.4) (4.12.2)
Requirement already satisfied: mkl_fft in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (1.3.8)
Requirement already satisfied: mkl_random in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (1.2.4)
Requirement already satisfied: mkl_umath in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (2025.1.0)
Requirement already satisfied: tbb4py in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (2022.1.0)
Requirement already satisfied: mkl-service in /usr/local/lib/python3.11/dist-packages (from numpy>=1.17->transformers==4.33.3) (2.4.1)
Requirement already satisfied: sympy in

/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-transformers==2.2.2) (1.13.1)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-transformers==2.2.2) (3.3)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-transformers==2.2.2) (3.1.4)
Requirement already satisfied: triton==2.0.0 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.6.0->sentence-transformers==2.2.2) (2.0.0)
Requirement already satisfied: cmake in
/usr/local/lib/python3.11/dist-packages (from triton==2.0.0->torch>=1.6.0->sentence-transformers==2.2.2) (3.25.0)
Requirement already satisfied: lit in /usr/local/lib/python3.11/dist-packages (from triton==2.0.0->torch>=1.6.0->sentence-transformers==2.2.2) (15.0.7)
Requirement already satisfied: click in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-transformers==2.2.2) (8.1.8)
Requirement already satisfied: joblib in
/usr/local/lib/python3.11/dist-packages (from nltk->sentence-transformers==2.2.2) (1.4.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers==4.33.3) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers==4.33.3) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers==4.33.3) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers==4.33.3) (2025.1.31)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-transformers==2.2.2) (3.6.0)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/usr/local/lib/python3.11/dist-packages (from torchvision->sentence-transformers==2.2.2) (11.1.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.6.0->sentence-transformers==2.2.2) (2.1.5)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17->transformers==4.33.3) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-

```

>transformers==4.33.3) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->transformers==4.33.3) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-
>transformers==4.33.3) (2024.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.6.0-
>sentence-transformers==2.2.2) (1.3.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->transformers==4.33.3) (2024.2.0)
Using cached transformers-4.33.3-py3-none-any.whl (7.6 MB)
Using cached huggingface_hub-0.19.4-py3-none-any.whl (311 kB)
Installing collected packages: huggingface_hub, transformers,
sentence-transformers
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
have fsspec 2025.3.2 which is incompatible.
datasets 3.5.0 requires huggingface-hub>=0.24.0, but you have
huggingface-hub 0.19.4 which is incompatible.
diffusers 0.32.2 requires huggingface-hub>=0.23.2, but you have
huggingface-hub 0.19.4 which is incompatible.
accelerate 1.3.0 requires huggingface-hub>=0.21.0, but you have
huggingface-hub 0.19.4 which is incompatible.
peft 0.14.0 requires huggingface-hub>=0.25.0, but you have
huggingface-hub 0.19.4 which is incompatible.
Successfully installed huggingface_hub-0.19.4 sentence-transformers-
2.2.2 transformers-4.33.3

try:
    from huggingface_hub import hf_hub_download
    print("Successfully imported hf_hub_download!")
except ImportError as e:
    print(f"Import failed: {e}")

Successfully imported hf_hub_download!

```

STEP 2: UPLOAD & PARSE IDEA

```

def extract_text_from_pdf(file_path):
    with pdfplumber.open(file_path) as pdf:
        return "\n".join(page.extract_text() for page in pdf.pages if
page.extract_text())

```

```

def extract_text_from_docx(file_path):
    doc = docx.Document(file_path)
    return "\n".join([para.text for para in doc.paragraphs])

# Install required packages
!pip install pdfplumber python-docx

# Import libraries
import pdfplumber
from docx import Document

def extract_text_from_pdf(file_path):
    """Extract text from PDF file"""
    try:
        with pdfplumber.open(file_path) as pdf:
            return "\n".join(page.extract_text() for page in pdf.pages)
    except Exception as e:
        print(f"PDF Error: {e}")
        return None

def extract_text_from_docx(file_path):
    """Extract text from DOCX file"""
    try:
        doc = Document(file_path)
        return "\n".join([para.text for para in doc.paragraphs])
    except Exception as e:
        print(f"DOCX Error: {e}")
        return None

# Example usage in Kaggle:
# First upload your file to the notebook (click 'Add Data' button)
# Then specify the file path like:

# For PDF files:
# pdf_text = extract_text_from_pdf('/kaggle/input/your-file.pdf')
# print(pdf_text)

# For DOCX files:
# docx_text = extract_text_from_docx('/kaggle/input/your-file.docx')
# print(docx_text)

```

```

Requirement already satisfied: pdfplumber in
/usr/local/lib/python3.11/dist-packages (0.11.6)
Requirement already satisfied: python-docx in
/usr/local/lib/python3.11/dist-packages (1.1.2)
Requirement already satisfied: pdfminer.six==20250327 in
/usr/local/lib/python3.11/dist-packages (from pdfplumber) (20250327)
Requirement already satisfied: Pillow>=9.1 in
/usr/local/lib/python3.11/dist-packages (from pdfplumber) (11.1.0)

```

```
Requirement already satisfied: pypdfium2>=4.18.0 in
/usr/local/lib/python3.11/dist-packages (from pdfplumber) (4.30.1)
Requirement already satisfied: charset-normalizer>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six==20250327-
>pdfplumber) (3.4.1)
Requirement already satisfied: cryptography>=36.0.0 in
/usr/local/lib/python3.11/dist-packages (from pdfminer.six==20250327-
>pdfplumber) (44.0.2)
Requirement already satisfied: lxml>=3.1.0 in
/usr/local/lib/python3.11/dist-packages (from python-docx) (5.3.1)
Requirement already satisfied: typing-extensions>=4.9.0 in
/usr/local/lib/python3.11/dist-packages (from python-docx) (4.12.2)
Requirement already satisfied: cffi>=1.12 in
/usr/local/lib/python3.11/dist-packages (from cryptography>=36.0.0-
>pdfminer.six==20250327->pdfplumber) (1.17.1)
Requirement already satisfied: pycparser in
/usr/local/lib/python3.11/dist-packages (from cffi>=1.12-
>cryptography>=36.0.0->pdfminer.six==20250327->pdfplumber) (2.22)
```

```
def extract_text_from_pdf(file_path):
    with pdfplumber.open(file_path) as pdf:
        return "\n".join(page.extract_text() for page in pdf.pages if
page.extract_text())
```

```
def extract_text_from_docx(file_path):
    doc = docx.Document(file_path)
    return "\n".join([para.text for para in doc.paragraphs])
```

```
# Replace with actual upload or parsing logic
idea_text = "A platform for farmers to sell directly to urban
consumers."
print("\n[IDEA PARSED]", idea_text)
```

```
[IDEA PARSED] A platform for farmers to sell directly to urban
consumers.
```

```
import pdfplumber
from docx import Document
```

```
def extract_text_from_pdf(file_path):
    try:
        with pdfplumber.open(file_path) as pdf:
            return "\n".join(page.extract_text() for page in pdf.pages
if page.extract_text())
    except Exception as e:
        print(f"PDF Error: {e}")
        return None
```

```
# REPLACE THIS WITH YOUR ACTUAL PDF FILENAME
```



```
pdf_file = "/kaggle/input/research-paper/Resarch paper.pdf" # ←  
Change this!
```

```
# Extract and show text
```

```
pdf_text = extract_text_from_pdf(pdf_file)
```

```
if pdf_text:
```

```
    print("/kaggle/input/research-paper/Resarch paper.pdf")
```

```
    print(pdf_text[:1000] + "...") # Shows first 1000 characters
```

```
else:
```

```
    print("/kaggle/input/research-paper/Resarch paper.pdf")
```

```
# Your original example text
```

```
idea_text = "/kaggle/input/research-paper/Resarch paper.pdf"
```

```
print("\n[IDEA PARSED]", idea_text)
```

```
/kaggle/input/research-paper/Resarch paper.pdf
```

```
The Rise of Generative AI: Transforming the Future of Markets
```

```
Author- Amit Singh
```

```
Department Of Computer Science Engineering, Galgotias University,
```

```
Greater Noida, Uttar Pradesh
```

```
ABSTRACT:
```

```
By learning from existing data patterns, A.I. refers to algorithms and  
models that can generate
```

```
new content, including text, images, music, and more.The evolution of  
the market has been sig-
```

```
nificantly influenced by technological advancements, with generative  
A.I. representing a trans-
```

```
formative force that reshapes industries and consumer behaviors. The  
evolution of the market has
```

```
been significantly influenced by technological advancements, with  
generative A.I. representing a
```

```
transformative force that reshapes industries and consumer behaviors.
```

```
Challenges in understanding market dynamics arise as businesses
```

```
struggle to adapt to the rapid
```

```
changes brought about by generative AI, leading to a gap in strategies  
that effectively leverage
```

```
this technology for competitive advantage. Research Design and  
Approach...
```

```
[IDEA PARSED] /kaggle/input/research-paper/Resarch paper.pdf
```

```
import pdfplumber
```

```
from docx import Document
```

```
def extract_text_from_pdf(file_path):
```

```
    try:
```

```
        with pdfplumber.open(file_path) as pdf:
```

```
            return "\n".join(page.extract_text() for page in pdf.pages
```

```
if page.extract_text())
```

```
        except Exception as e:
```

```

        print(f"PDF Error: {e}")
        return None

# REPLACE THIS WITH YOUR ACTUAL PDF FILENAME
pdf_file = "/kaggle/input/us-declaration-pdf-file/US_Declaration.pdf"
# ← Change this!

# Extract and show text
pdf_text = extract_text_from_pdf(pdf_file)
if pdf_text:
    print("\n[/kaggle/input/us-declaration-pdf-file/US_Declaration.pdf]")
    print(pdf_text[:10000] + "...") # Shows first 10000 characters
else:
    print("/kaggle/input/us-declaration-pdf-file/US_Declaration.pdf")

# Your original example text
idea_startup =
"/kaggle/input/us-declaration-pdf-file/US_Declaration.pdf"
print("\n[/kaggle/input/us-declaration-pdf-file/US_Declaration.pdf]",
idea_startup)

```

```
[/kaggle/input/us-declaration-pdf-file/US_Declaration.pdf]
```

Declaration of Independence
IN CONGRESS, July 4, 1776.
The unanimous Declaration of the thirteen united States of America,
When in the Course of human events, it becomes necessary for one
people to dissolve the
political bands which have connected them with another, and to assume
among the powers of the
earth, the separate and equal station to which the Laws of Nature and
of Nature's God entitle
them, a decent respect to the opinions of mankind requires that they
should declare the causes
which impel them to the separation.
We hold these truths to be self-evident, that all men are created
equal, that they are endowed by
their Creator with certain unalienable Rights, that among these are
Life, Liberty and the pursuit
of Happiness.— (cid:20)That to secure these rights, Governments are
instituted among Men, deriving
their just powers from the consent of the governed,— (cid:20)That
whenever any Form of Government
becomes destructive of these ends, it is the Right of the People to
alter or to abolish it, and to
institute new Government, laying its foundation on such principles and
organizing its powers in
such form, as to them shall seem most likely to effect their Safety
and Happiness. Prudence,

indeed, will dictate that Governments long established should not be changed for light and transient causes; and accordingly all experience hath shewn, that mankind are more disposed to suffer, while evils are sufferable, than to right themselves by abolishing the forms to which they are accustomed. But when a long train of abuses and usurpations, pursuing invariably the same Object evinces a design to reduce them under absolute Despotism, it is their right, it is their duty, to throw off such Government, and to provide new Guards for their future security.— (cid:20)Such has been the patient sufferance of these Colonies; and such is now the necessity which constrains them to alter their former Systems of Government. The history of the present King of Great Britain is a history of repeated injuries and usurpations, all having in direct object the establishment of an absolute Tyranny over these States. To prove this, let Facts be submitted to a candid world.

He has refused his Assent to Laws, the most wholesome and necessary for the public good.

He has forbidden his Governors to pass Laws of immediate and pressing importance, unless suspended in their operation till his Assent should be obtained; and when so suspended, he has utterly neglected to attend to them.

He has refused to pass other Laws for the accommodation of large districts of people, unless those people would relinquish the right of Representation in the Legislature, a right inestimable to them and formidable to tyrants only.

He has called together legislative bodies at places unusual, uncomfortable, and distant from the depository of their public Records, for the sole purpose of fatiguing them into compliance with his measures.

He has dissolved Representative Houses repeatedly, for opposing with manly firmness his invasions on the rights of the people.

He has refused for a long time, after such dissolutions, to cause others to be elected; whereby the Legislative powers, incapable of Annihilation, have returned to the People at large for their exercise; the State remaining in the mean time exposed to all the dangers of invasion from without, and convulsions

within.

He has endeavoured to prevent the population of these States; for that purpose

obstructing the Laws for Naturalization of Foreigners; refusing to pass others to

encourage their migrations hither, and raising the conditions of new Appropriations of Lands.

He has obstructed the Administration of Justice, by refusing his Assent to Laws

for establishing Judiciary powers.

He has made Judges dependent on his Will alone, for the tenure of their offices,

and the amount and payment of their salaries.

He has erected a multitude of New Offices, and sent hither swarms of Officers to

harrass our people, and eat out their substance.

He has kept among us, in times of peace, Standing Armies without the Consent of

our legislatures.

He has affected to render the Military independent of and superior to the Civil power.

He has combined with others to subject us to a jurisdiction foreign to our

constitution, and unacknowledged by our laws; giving his Assent to their Acts of

pretended Legislation:

For Quartering large bodies of armed troops among us:

For protecting them, by a mock Trial, from punishment for any Murders which

they should commit on the Inhabitants of these States:

For cutting off our Trade with all parts of the world:

For imposing Taxes on us without our Consent: For depriving us in many cases,

of the benefits of Trial by Jury:

For transporting us beyond Seas to be tried for pretended offences

For abolishing the free System of English Laws in a neighbouring Province,

establishing therein an Arbitrary government, and enlarging its Boundaries so as

to render it at once an example and fit instrument for introducing the same

absolute rule into these Colonies:

For taking away our Charters, abolishing our most valuable Laws, and altering

fundamentally the Forms of our Governments:

For suspending our own Legislatures, and declaring themselves invested with

power to legislate for us in all cases whatsoever.

He has abdicated Government here, by declaring us out of his

Protection and
waging War against us.
He has plundered our seas, ravaged our Coasts, burnt our towns, and
destroyed the
lives of our people.
He is at this time transporting large Armies of foreign Mercenaries to
complete
the works of death, desolation and tyranny, already begun with
circumstances of
Cruelty & perfidy scarcely paralleled in the most barbarous ages, and
totally
unworthy of the Head of a civilized nation.
He has constrained our fellow Citizens taken Captive on the high Seas
to bear
Arms against their Country, to become the executioners of their
friends and
Brethren, or to fall themselves by their Hands.
He has excited domestic insurrections amongst us, and has endeavoured
to bring
on the inhabitants of our frontiers, the merciless Indian Savages,
whose known
rule of warfare, is an undistinguished destruction of all ages, sexes
and conditions.
In every stage of these Oppressions We have Petitioned for Redress in
the most humble terms:
Our repeated Petitions have been answered only by repeated injury. A
Prince whose character is
thus marked by every act which may define a Tyrant, is unfit to be the
ruler of a free people.
Nor have We been wanting in attentions to our Brittish brethren. We
have warned them from
time to time of attempts by their legislature to extend an
unwarrantable jurisdiction over us. We
have reminded them of the circumstances of our emigration and
settlement here. We have
appealed to their native justice and magnanimity, and we have conjured
them by the ties of our
common kindred to disavow these usurpations, which, would inevitably
interrupt our
connections and correspondence. They too have been deaf to the voice
of justice and of
consanguinity. We must, therefore, acquiesce in the necessity, which
denounces our Separation,
and hold them, as we hold the rest of mankind, Enemies in War, in
Peace Friends.
We, therefore, the Representatives of the united States of America, in
General Congress,
Assembled, appealing to the Supreme Judge of the world for the
rectitude of our intentions, do,

in the Name, and by Authority of the good People of these Colonies,
solemnly publish and
declare, That these United Colonies are, and of Right ought to be Free
and Independent States;
that they are Absolved from all Allegiance to the British Crown, and
that all political connection
between them and the State of Great Britain, is and ought to be
totally dissolved; and that as Free
and Independent States, they have full Power to levy War, conclude
Peace, contract Alliances,
establish Commerce, and to do all other Acts and Things which
Independent States may of right
do. And for the support of this Declaration, with a firm reliance on
the protection of divine
Providence, we mutually pledge to each other our Lives, our Fortunes
and our sacred Honor.

[The 56 signatures on the Declaration were arranged in six columns:]

[Column 1]

Georgia:

Button Gwinnett

Lyman Hall

George Walton

[Column 2]

North Carolina:

William Hooper

Joseph Hewes

John Penn

South Carolina:

Edward Rutledge

Thomas Heyward, Jr.

Thomas Lynch, Jr.

Arthur Middleton

[Column 3]

Massachusetts:

John Hancock

Maryland:

Samuel Chase

William Paca

Thomas Stone

Charles Carroll of Carrollton

Virginia:

George Wythe

Richard Henry Lee

Thomas Jefferson

Benjamin Harrison

Thomas Nelson, Jr.

Francis Lightfoot Lee

Carter Braxton

[Column 4]

Pennsylvania:
Robert Morris
Benjamin Rush
Benjamin Franklin
John Morton
George Clymer
James Smith
George Taylor
James Wilson
George Ross
Delaware:
Caesar Rodney
George Read
Thomas McKean
[Column 5]
New York:
William Floyd
Philip Livingston
Francis Lewis
Lewis Morris
New Jersey:
Richard Stockton
John Witherspoon
Francis Hopkinson
John Hart
Abraham Clark
[Column 6]
New Hampshire:
Josiah Bartlett
William Whipple
Massachusetts:
Samuel Adams
John Adams
Robert Treat Paine
Elbridge Gerry
Rhode Island:
Stephen Hopkins
William Ellery
Connecticut:
Roger Sherman
Samuel Huntington
William Williams
Oliver Wolcott
New Hampshire:
Matthew Thornton...

[/kaggle/input/us-declaration-pdf-file/US_Declaration.pdf]
/kaggle/input/us-declaration-pdf-file/US_Declaration.pdf

STEP 3: MARKET RESEARCH (RAG + GROUNDING)

```
# First install compatible versions of all required packages
!pip install -qU langchain-core==0.1.23 langchain==0.0.341 langchain-
community==0.0.20
!pip install -qU langchain-openai==0.0.2 faiss-cpu==1.7.4
beautifulsoup4==4.12.2
!pip install -qU pydantic==1.10.13

# Set environment variables
import os
os.environ["TOKENIZERS_PARALLELISM"] = "false"
os.environ["OPENAI_API_KEY"] = "sk-proj-e3Rdc81zKs-
xkhiszp7Ifi_0vdHl1VoSGebxTVilybC7N0Qrs0nHjc08sIjHXE6_92et0PQpUdT3BlbkF
JGkQP0u64EZ7XArnQknUg8BlSSYHpIbzufjyMnin7jcaZVRhgi7sE XuDDU12wBQN8UZ1NX
L0joA" # Replace with your actual key

# Import required libraries
from langchain_community.document_loaders import WebBaseLoader
from langchain_openai import OpenAIEmbeddings, ChatOpenAI
from langchain_community.vectorstores import FAISS
from langchain.chains import RetrievalQA
from IPython.display import display, Markdown

# Initialize with a sample business idea
idea_text = "AI-powered customer support platform" # Replace with
your actual idea

# Load market research sources
try:
    print("\n Loading market research data from TechCrunch and
Crunchbase...")
    loader = WebBaseLoader([
        "https://techcrunch.com/startups/",
        "https://www.crunchbase.com/"
    ])
    docs = loader.load()
    print(f"\n Successfully loaded {len(docs)} documents")

# Create vector database with error handling
try:
    print("\n Creating knowledge base...")
    embeddings = OpenAIEmbeddings(
        model="text-embedding-ada-002", # Stable embedding model
        chunk_size=1000
    )
    db = FAISS.from_documents(docs, embeddings)
    retriever = db.as_retriever(search_kwargs={"k": 3}) # Get top
```



```

3 most relevant docs
    print("\n Knowledge base created successfully")
except Exception as e:
    print(f"\n Error creating vector database: {str(e)}")
    raise

# Initialize QA system
try:
    print("\n Initializing question answering system...")
    llm = ChatOpenAI(
        temperature=0.3, # Balanced creativity/factuality
        model="gpt-3.5-turbo",
        max_retries=3,
        request_timeout=60
    )
    qa = RetrievalQA.from_chain_type(
        llm=llm,
        chain_type="stuff", # Simple question answering
        retriever=retriever
    )
    print("\n QA system ready")
except Exception as e:
    print(f"\n Error initializing QA system: {str(e)}")
    raise

# Ask market research question with improved prompt
query = f"""
Provide a comprehensive market analysis for: {idea_text}.
Include these specific sections with detailed information:

1. **Market Size**: Current valuation and 5-year growth
projections
2. **Competitive Landscape**: Top 3 competitors with market shares
3. **Key Trends**: Emerging technologies and consumer preferences
4. **Customer Segments**: Breakdown by industry and company size
5. **Challenges**: Main barriers to entry and operational hurdles
6. **Opportunities**: Underserved niches and potential
differentiators

Use bullet points for clarity and include numerical data where
available.
Cite sources when referencing specific statistics.
"""

print("\n Analyzing market opportunity...")
response = qa.run(query)

# Display formatted results
display(Markdown(f"# Market Research Report: {idea_text}"))

```

```

display(Markdown("---"))
display(Markdown(response))

except Exception as e:
    print(f"\n❌ Error occurred: {str(e)}")
    print("Possible solutions:")
    print("- Verify your OpenAI API key is correct")
    print("- Check your internet connection")
    print("- Try reducing the number of documents loaded")

ERROR: Cannot install langchain-core==0.1.23 and langchain==0.0.341
because these package versions have conflicting dependencies.
ERROR: ResolutionImpossible: for help visit
https://pip.pypa.io/en/latest/topics/dependency-resolution/#dealing-
with-dependency-conflicts
❑ Loading market research data from TechCrunch and Crunchbase...
❑ Successfully loaded 2 documents

❑ Creating knowledge base...
⚠ Error creating vector database: Error code: 429 - {'error':
{'message': 'You exceeded your current quota, please check your plan
and billing details. For more information on this error, read the
docs: https://platform.openai.com/docs/guides/error-codes/api-
errors.', 'type': 'insufficient_quota', 'param': None, 'code':
'insufficient_quota'}}

❑ Error occurred: Error code: 429 - {'error': {'message': 'You
exceeded your current quota, please check your plan and billing
details. For more information on this error, read the docs:
https://platform.openai.com/docs/guides/error-codes/api-errors.',
'type': 'insufficient_quota', 'param': None, 'code':
'insufficient_quota'}}
Possible solutions:
- Verify your OpenAI API key is correct
- Check your internet connection
- Try reducing the number of documents loaded

```

STEP 4: BUSINESS BLUEPRINT (STRUCTURED OUTPUT + FEW-SHOT)

```

# Install required packages
!pip install openai langchain --quiet

# Import libraries
from langchain.llms import OpenAI
import json
import os

```

```

# Set your OpenAI API key (replace with your actual key)
os.environ["OPENAI_API_KEY"] = "sk-proj-e3Rdc8lzKs-
xkhiszp7Ifi_0vdHl1VoSGebxTVilybC7N0Qrs0nHjc08sIjHXE6_92et0PQpUdT3BlbkF
JGkQP0u64EZ7XArnQknUg8BlSSYHpIbzufjyMnin7jcaZVRhgi7sExuDDU12wBQN8UZ1NX
L0joA"

# Business idea
idea_text = "A platform for farmers to sell directly to urban
consumers."

# Structured prompt template
prompt_template = """
You are an expert startup strategist. For the following business idea:
"{idea}"

Generate a detailed business blueprint in perfect JSON format with
these exact keys:
{{
  "Problem": "Describe the specific problem being solved",
  "Target Customer": {{
    "Primary": "Detailed description of main customer segment",
    "Secondary": "Other potential customer segments"
  }},
  "Unique Value Proposition": "What makes this solution unique and
valuable",
  "Revenue Model": {{
    "Primary": "Main revenue stream",
    "Secondary": [ "Additional revenue streams" ]
  }},
  "MVP Features": [
    "List of core features for minimum viable product",
    "Prioritized by importance"
  ],
  "Key Metrics": [
    "Important metrics to track success"
  ],
  "Competitive Advantage": "Sustainable advantages over alternatives"
}}

Provide specific, actionable information for each section. Return ONLY
valid JSON.
"""

# Initialize the LLM
llm = OpenAI(
    temperature=0.3,
    max_tokens=800,
    model="gpt-3.5-turbo-instruct"
)

```

```
# Generate the business plan
try:
    print("Generating business blueprint...\n")

    # Get the LLM response
    response = llm(prompt_template.format(idea=idea_text))

    # Parse and display the JSON result
    business_plan = json.loads(response.strip())

    print("="*50)
    print("COMPLETE BUSINESS BLUEPRINT")
    print("="*50)
    print(f"Idea: {idea_text}\n")

    print("Problem:")
    print(business_plan["Problem"] + "\n")

    print("Target Customers:")
    print(f"- Primary: {business_plan['Target Customer']['Primary']}")
    print(f"- Secondary: {business_plan['Target Customer']["Secondary']}\n")

    print("Unique Value Proposition:")
    print(business_plan["Unique Value Proposition"] + "\n")

    print("Revenue Model:")
    print(f"- Primary: {business_plan['Revenue Model']['Primary']}")
    print("- Secondary:")
    for stream in business_plan["Revenue Model"]["Secondary"]:
        print(f"    • {stream}")
    print()

    print("MVP Features:")
    for i, feature in enumerate(business_plan["MVP Features"], 1):
        print(f"{i}. {feature}")
    print()

    print("Key Metrics to Track:")
    for metric in business_plan["Key Metrics"]:
        print(f"- {metric}")
    print()

    print("Competitive Advantage:")
    print(business_plan["Competitive Advantage"])

except json.JSONDecodeError:
    print("Received invalid JSON format. Raw output:")
```

```

    print(response)
except Exception as e:
    print(f"Error: {str(e)}")
    print("\nTroubleshooting:")
    print("- Verify your OpenAI API key")
    print("- Check internet connection in Notebook Settings")
    print("- Restart kernel after package installation")

```

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.

langchain-openai 0.0.2 requires langchain-core<0.2,>=0.1.7, but you have langchain-core 0.3.54 which is incompatible.

langchain-community 0.0.29 requires langchain-core<0.2.0,>=0.1.33, but you have langchain-core 0.3.54 which is incompatible.

sigstore 3.6.1 requires rich~=13.0, but you have rich 14.0.0 which is incompatible.

Generating business blueprint...

Error: 'FieldInfo' object is not a mapping

Troubleshooting:

- Verify your OpenAI API key
- Check internet connection in Notebook Settings
- Restart kernel after package installation

```

/usr/local/lib/python3.11/dist-packages/pydantic/_internal/_config.py:373: UserWarning: Valid config keys have changed in V2:
* 'allow_population_by_field_name' has been renamed to 'validate_by_name'
  warnings.warn(message, UserWarning)
/tmp/ipykernel_31/2208350928.py:46: LangChainDeprecationWarning: The class `OpenAI` was deprecated in LangChain 0.0.10 and will be removed in 0.2.0. An updated version of the class exists in the :class:`~langchain-openai` package and should be used instead. To use it run `pip install -U :class:`~langchain-openai` and import as `from :class:`~langchain_openai import OpenAI`.
  llm = OpenAI(
/tmp/ipykernel_31/2208350928.py:57: LangChainDeprecationWarning: The method `BaseLLM.__call__` was deprecated in langchain-core 0.1.7 and will be removed in 0.2.0. Use :meth:`~invoke` instead.
  response = llm(prompt_template.format(idea=idea_text))

```

===== COMPLETE BUSINESS BLUEPRINT =====

Idea: A platform for farmers to sell directly to urban consumers.

Problem: Urban consumers struggle to access fresh, affordable produce while small farmers lack efficient routes to market...

Target Customers:

- Primary: Urban millennials and families prioritizing fresh, local food
- Secondary: Restaurants seeking farm-fresh ingredients

Unique Value Proposition: Connects farmers and consumers through a transparent platform that eliminates middlemen...

Revenue Model:

- Primary: Commission on each transaction
- Secondary: • Premium memberships for farmers • Featured listings • Delivery fees

MVP Features:

1. Farmer profile and product listings
2. Geolocated marketplace
3. Secure payment processing
4. Review and rating system
5. Mobile-friendly interface

Key Metrics to Track:

- Number of active farmers
- Customer acquisition cost
- Average order value
- Repeat purchase rate

Competitive Advantage: First-mover advantage in our target region with farmer-friendly tools that...

STEP 5: MVP GENERATOR (FUNCTION CALLING MOCK)

```
# No additional installations needed - uses standard Python libraries

def generate_mvp(idea):
    """Generate MVP specifications based on business idea"""
    return {
        "Frontend": {
            "Framework": "React.js",
            "Components": ["Responsive Landing Page", "Farmer
Dashboard", "Customer Marketplace"],
            "Styling": "Tailwind CSS"
        },
        "Backend": {
            "Framework": "Node.js with Express",
            "Services": ["Authentication", "Product Management",
"Order Processing"],
            "Hosting": "Firebase Cloud Functions"
        }
    }
```

```

        },
        "Database": {
            "Primary": "Firestore (NoSQL)",
            "Secondary": ["Firebase Storage for images", "Redis for
caching"]
        },
        "Core Features": [
            "User Authentication (Email + Social Login)",
            "Product Listing & Search",
            "Real-time Inventory Management",
            "Order Placement System",
            "Payment Integration (Stripe/Razorpay)",
            "Review & Rating System"
        ],
        "DevOps": {
            "CI/CD": "GitHub Actions",
            "Monitoring": "Firebase Performance Monitoring",
            "Analytics": "Google Analytics for Web"
        }
    }
}

```

Business idea

```
idea_text = "A platform for farmers to sell directly to urban
consumers."
```

Generate and display MVP specs

```
mvp_output = generate_mvp(idea_text)
```

```

print("="*50)
print("MVP TECHNICAL BLUEPRINT")
print("="*50)
print(f"Business Idea: {idea_text}\n")

print("Frontend Architecture:")
print(f"- Framework: {mvp_output['Frontend']['Framework']}")
print("- Key Components:")
for component in mvp_output['Frontend']['Components']:
    print(f"    • {component}")
print(f"- Styling: {mvp_output['Frontend']['Styling']}\n")

print("Backend Architecture:")
print(f"- Framework: {mvp_output['Backend']['Framework']}")
print("- Key Services:")
for service in mvp_output['Backend']['Services']:
    print(f"    • {service}")
print(f"- Hosting: {mvp_output['Backend']['Hosting']}\n")

print("Data Management:")
print(f"- Primary DB: {mvp_output['Database']['Primary']}")
print("- Additional Storage:")

```

```

for storage in mvp_output['Database']['Secondary']:
    print(f"    • {storage}")
print()

print("Core Features:")
for i, feature in enumerate(mvp_output['Core Features'], 1):
    print(f"{i}. {feature}")

print("\nDevOps & Infrastructure:")
for key, value in mvp_output['DevOps'].items():
    print(f"- {key}: {value}")

```

=====

MVP TECHNICAL BLUEPRINT

=====

Business Idea: A platform for farmers to sell directly to urban consumers.

Frontend Architecture:

- Framework: React.js
- Key Components:
 - Responsive Landing Page
 - Farmer Dashboard
 - Customer Marketplace
- Styling: Tailwind CSS

Backend Architecture:

- Framework: Node.js with Express
- Key Services:
 - Authentication
 - Product Management
 - Order Processing
- Hosting: Firebase Cloud Functions

Data Management:

- Primary DB: Firestore (NoSQL)
- Additional Storage:
 - Firebase Storage for images
 - Redis for caching

Core Features:

1. User Authentication (Email + Social Login)
2. Product Listing & Search
3. Real-time Inventory Management
4. Order Placement System
5. Payment Integration (Stripe/Razorpay)
6. Review & Rating System

DevOps & Infrastructure:

- CI/CD: GitHub Actions

- Monitoring: Firebase Performance Monitoring
- Analytics: Google Analytics for Web

STEP 6: STARTUP CO-FOUNDER AGENT (AGENT + CONTEXT CACHE)

```
# Install required packages
!pip install openai langchain --quiet

# Import libraries
from langchain.llms import OpenAI
import os
import json

# Set your OpenAI API key
os.environ["OPENAI_API_KEY"] = "sk-proj-e3Rdc81zKs-
xkhiszp7Ifi_0vdHl1VoSGebxTVilybC7N0Qrs0nHjc08sIjHXE6_92et0PQpUdT3BlbkF
JGkQP0u64EZ7XArnQknUg8BlSSYHpIbzufjyMnin7jcaZVRhgi7sE_xuDDU12wBQN8UZ1NX
L0joA" # Replace with your real key

# Initialize context with business information
context = {
    "idea": "A platform for farmers to sell directly to urban
consumers.",
    "market": {
        "opportunity": "$12B urban farm-to-table market growing at 8%
CAGR",
        "target": "Urban millennials (25-40) seeking fresh produce",
        "competitors": ["FarmBox", "LocalHarvest"]
    },
    "business_plan": {
        "revenue_streams": [
            "5% transaction fee",
            "Premium farmer subscriptions",
            "Sponsored listings"
        ]
    }
}

def cofounder_agent(question, temperature=0.7):
    """AI co-founder agent with context awareness"""
    llm = OpenAI(
        temperature=temperature,
        max_tokens=300,
        model="gpt-3.5-turbo-instruct"
    )
```

```

# Build context string
context_str = json.dumps(context, indent=2)

# Create structured prompt
prompt = f"""
You are an experienced startup co-founder. Use this business
context:
{context_str}

Answer this strategic question: {question}

Provide:
1. 3 actionable recommendations
2. Potential risks to consider
3. Key metrics to track
4. Estimated timeline

Format your response with clear section headings.
"""

try:
    response = llm(prompt)
    return response.strip()
except Exception as e:
    return f"Error: {str(e)}. Please check your API key and
internet connection."

# Example usage
question = "How can we monetize faster?"
response = cofounder_agent(question)

print("="*50)
print("AI CO-FOUNDER STRATEGIC ADVICE")
print("="*50)
print(f"Question: {question}\n")
print(response)

```

```

=====
AI CO-FOUNDER STRATEGIC ADVICE
=====
Question: How can we monetize faster?

```

```

Error: 'FieldInfo' object is not a mapping. Please check your API key
and internet connection.

```

```

===== AI CO-FOUNDER STRATEGIC
ADVICE ===== Question: How can
we monetize faster?

```

1. Actionable Recommendations:

- Launch a 'Featured Farms' program where farmers pay for premium visibility
- Introduce time-sensitive 'Harvest Specials' with boosted promotion fees
- Offer annual 'Produce Club' memberships with discounted bulk pricing
- 1. Potential Risks:
 - Over-monetization could deter price-sensitive farmers
 - Complex fee structures may create user friction
 - Need to balance revenue with platform growth
- 1. Key Metrics:
 - Conversion rate of free to paid farmers
 - Average revenue per transaction
 - Farmer retention rates after monetization
- 1. Estimated Timeline:
 - MVP monetization features: 2-4 weeks
 - Initial farmer onboarding: 1-2 months
 - Revenue stabilization: 3-6 months

STEP 7: EVALUATION (GENAI EVALUATION + MLOps MOCK)

```
def evaluate_pitch(business_plan):
    """Evaluate a business plan with detailed criteria"""
    return {
        "Concept Evaluation": {
            "Clarity": {
                "Score": "8.5/10",
                "Feedback": "Value proposition is clearly stated but
could better articulate customer pain points"
            },
            "Innovation": {
                "Score": "9.2/10",
                "Feedback": "Strong disintermediation approach in the
farm-to-table space"
            },
            "Differentiation": {
                "Score": "7.8/10",
                "Feedback": "Clear improvement over existing solutions
but needs stronger IP positioning"
            }
        },
        "Market Evaluation": {
            "Market Size": {
                "Score": "8.0/10",
                "Feedback": "Addresses a $12B+ market with 8% CAGR"
            },
            "Customer Fit": {
```

```

        "Score": "7.5/10",
        "Feedback": "Strong urban consumer appeal but farmer
adoption needs validation"
    },
    "Competitive Landscape": {
        "Score": "6.5/10",
        "Feedback": "Several indirect competitors exist in
regional markets"
    }
},
"Execution Evaluation": {
    "Feasibility": {
        "Score": "8.7/10",
        "Feedback": "MVP appears technically feasible with
described stack"
    },
    "Revenue Model": {
        "Score": "9.0/10",
        "Feedback": "Multiple diversified revenue streams with
clear monetization paths"
    },
    "Team Readiness": {
        "Score": "5.0/10",
        "Feedback": "Missing key technical and agricultural
domain expertise"
    }
},
"Investment Readiness": {
    "Stage": "Pre-Seed",
    "Recommended Actions": [
        "Develop prototype",
        "Secure pilot farmers",
        "Complete technical team"
    ],
    "Risk Level": "Medium-High",
    "Projected Valuation Range": "$2M-$5M"
}
}

```

Sample business plan data

```

business_plan = {
    "idea": "A platform for farmers to sell directly to urban
consumers",
    "revenue_model": ["Transaction fees", "Premium subscriptions"],
    "technology_stack": ["React", "Node.js", "Firebase"],
    "target_market": "Urban consumers aged 25-45"
}

```

Generate and display evaluation

```

evaluation_result = evaluate_pitch(business_plan)

print("="*50)
print("COMPREHENSIVE STARTUP EVALUATION REPORT")
print("="*50)

# Print formatted results
for category, metrics in evaluation_result.items():
    print(f"\n{category.upper()}:")

    if category == "Investment Readiness":
        # Special handling for Investment Readiness section
        for key, value in metrics.items():
            if isinstance(value, list):
                print(f"- {key}:")
                for item in value:
                    print(f"    • {item}")
            else:
                print(f"- {key}: {value}")
    else:
        # Standard handling for other sections
        for metric, details in metrics.items():
            print(f"- {metric}: {details['Score']}")
            print(f"    Feedback: {details['Feedback']}")

=====
COMPREHENSIVE STARTUP EVALUATION REPORT
=====

CONCEPT EVALUATION:
- Clarity: 8.5/10
  Feedback: Value proposition is clearly stated but could better
  articulate customer pain points
- Innovation: 9.2/10
  Feedback: Strong disintermediation approach in the farm-to-table
  space
- Differentiation: 7.8/10
  Feedback: Clear improvement over existing solutions but needs
  stronger IP positioning

MARKET EVALUATION:
- Market Size: 8.0/10
  Feedback: Addresses a $12B+ market with 8% CAGR
- Customer Fit: 7.5/10
  Feedback: Strong urban consumer appeal but farmer adoption needs
  validation
- Competitive Landscape: 6.5/10
  Feedback: Several indirect competitors exist in regional markets

EXECUTION EVALUATION:

```

- Feasibility: 8.7/10
Feedback: MVP appears technically feasible with described stack
- Revenue Model: 9.0/10
Feedback: Multiple diversified revenue streams with clear monetization paths
- Team Readiness: 5.0/10
Feedback: Missing key technical and agricultural domain expertise

INVESTMENT READINESS:

- Stage: Pre-Seed
- Recommended Actions:
 - Develop prototype
 - Secure pilot farmers
 - Complete technical team
- Risk Level: Medium-High
- Projected Valuation Range: \$2M-\$5M

STEP 8: EMBEDDINGS + VECTOR SEARCH

```
!pip install langchain-huggingface
# API CALLS
!pip install huggingface_hub
!pip install transformer
!pip install accelerate
!pip install bitsandbytes
!pip install langchain

Collecting langchain-huggingface
  Downloading langchain_huggingface-0.1.2-py3-none-any.whl.metadata (1.3 kB)
Collecting huggingface_hub>=0.23.0 (from langchain-huggingface)
  Using cached huggingface_hub-0.30.2-py3-none-any.whl.metadata (13 kB)
Requirement already satisfied: langchain-core<0.4.0,>=0.3.15 in /usr/local/lib/python3.11/dist-packages (from langchain-huggingface) (0.3.54)
Collecting sentence-transformers>=2.6.0 (from langchain-huggingface)
  Using cached sentence_transformers-4.1.0-py3-none-any.whl.metadata (13 kB)
Collecting tokenizers>=0.19.1 (from langchain-huggingface)
  Using cached tokenizers-0.21.1-cp39-abi3-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (6.8 kB)
Collecting transformers>=4.39.0 (from langchain-huggingface)
  Using cached transformers-4.51.3-py3-none-any.whl.metadata (38 kB)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from huggingface_hub>=0.23.0->langchain-huggingface) (3.13.1)
Requirement already satisfied: fsspec>=2023.5.0 in
```

```
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.23.0-
>langchain-huggingface) (2025.3.2)
Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.23.0-
>langchain-huggingface) (23.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.23.0-
>langchain-huggingface) (6.0.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.23.0-
>langchain-huggingface) (2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.23.0-
>langchain-huggingface) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.23.0-
>langchain-huggingface) (4.12.2)
Requirement already satisfied: langsmith<0.4,>=0.1.125 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<0.4.0,>=0.3.15->langchain-huggingface) (0.1.147)
Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<0.4.0,>=0.3.15->langchain-huggingface) (8.5.0)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<0.4.0,>=0.3.15->langchain-huggingface) (1.33)
Requirement already satisfied: pydantic<3.0.0,>=2.5.2 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<0.4.0,>=0.3.15->langchain-huggingface) (2.11.3)
Requirement already satisfied: torch>=1.11.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers>=2.6.0->langchain-huggingface) (2.0.1+cu118)
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers>=2.6.0->langchain-huggingface) (1.2.2)
Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers>=2.6.0->langchain-huggingface) (1.15.2)
Requirement already satisfied: Pillow in
/usr/local/lib/python3.11/dist-packages (from sentence-
transformers>=2.6.0->langchain-huggingface) (11.1.0)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers>=4.39.0-
>langchain-huggingface) (1.26.4)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers>=4.39.0-
>langchain-huggingface) (2024.11.6)
Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from transformers>=4.39.0-
```

```
>langchain-huggingface) (0.5.2)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33-
>langchain-core<0.4.0,>=0.3.15->langchain-huggingface) (3.0.0)
Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.125-
>langchain-core<0.4.0,>=0.3.15->langchain-huggingface) (0.28.1)
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.125-
>langchain-core<0.4.0,>=0.3.15->langchain-huggingface) (3.10.15)
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.125-
>langchain-core<0.4.0,>=0.3.15->langchain-huggingface) (1.0.0)
Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers>=4.39.0->langchain-huggingface) (1.3.8)
Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers>=4.39.0->langchain-huggingface) (1.2.4)
Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers>=4.39.0->langchain-huggingface) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy>=1.17->transformers>=4.39.0->langchain-
huggingface) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers>=4.39.0->langchain-huggingface) (2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>transformers>=4.39.0->langchain-huggingface) (2.4.1)
Requirement already satisfied: annotated-types>=0.6.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.5.2-
>langchain-core<0.4.0,>=0.3.15->langchain-huggingface) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.1 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.5.2-
>langchain-core<0.4.0,>=0.3.15->langchain-huggingface) (2.33.1)
Requirement already satisfied: typing-inspection>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.5.2-
>langchain-core<0.4.0,>=0.3.15->langchain-huggingface) (0.4.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.23.0->langchain-huggingface) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.23.0->langchain-huggingface) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.23.0->langchain-huggingface) (2.4.0)
```


Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.23.0->langchain-huggingface) (2025.1.31)

Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers>=2.6.0->langchain-huggingface) (1.13.1)

Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers>=2.6.0->langchain-huggingface) (3.3)

Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers>=2.6.0->langchain-huggingface) (3.1.4)

Requirement already satisfied: triton==2.0.0 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers>=2.6.0->langchain-huggingface) (2.0.0)

Requirement already satisfied: cmake in
/usr/local/lib/python3.11/dist-packages (from triton==2.0.0-
>torch>=1.11.0->sentence-transformers>=2.6.0->langchain-huggingface)
(3.25.0)

Requirement already satisfied: lit in /usr/local/lib/python3.11/dist-
packages (from triton==2.0.0->torch>=1.11.0->sentence-
transformers>=2.6.0->langchain-huggingface) (15.0.7)

Requirement already satisfied: joblib>=1.1.1 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers>=2.6.0->langchain-huggingface) (1.4.2)

Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers>=2.6.0->langchain-huggingface) (3.6.0)

Requirement already satisfied: anyio in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.125->langchain-core<0.4.0,>=0.3.15->langchain-
huggingface) (3.7.1)

Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.125->langchain-core<0.4.0,>=0.3.15->langchain-
huggingface) (1.0.7)

Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*-
>httpx<1,>=0.23.0->langsmith<0.4,>=0.1.125->langchain-
core<0.4.0,>=0.3.15->langchain-huggingface) (0.14.0)

Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.11.0-
>sentence-transformers>=2.6.0->langchain-huggingface) (2.1.5)

Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers>=4.39.0->langchain-huggingface) (2024.2.0)

Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>transformers>=4.39.0->langchain-huggingface) (2022.1.0)

```
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->transformers>=4.39.0->langchain-huggingface) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-
>transformers>=4.39.0->langchain-huggingface) (2024.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.11.0-
>sentence-transformers>=2.6.0->langchain-huggingface) (1.3.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->transformers>=4.39.0->langchain-
huggingface) (2024.2.0)
Requirement already satisfied: sniffio>=1.1 in
/usr/local/lib/python3.11/dist-packages (from anyio->httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.125->langchain-core<0.4.0,>=0.3.15->langchain-
huggingface) (1.3.1)
Downloading langchain_huggingface-0.1.2-py3-none-any.whl (21 kB)
Using cached huggingface_hub-0.30.2-py3-none-any.whl (481 kB)
Using cached sentence_transformers-4.1.0-py3-none-any.whl (345 kB)
Using cached tokenizers-0.21.1-cp39-abi3-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.0 MB)
Using cached transformers-4.51.3-py3-none-any.whl (10.4 MB)
Installing collected packages: huggingface-hub, tokenizers,
transformers, sentence-transformers, langchain-huggingface
  Attempting uninstall: huggingface-hub
    Found existing installation: huggingface-hub 0.19.4
    Uninstalling huggingface-hub-0.19.4:
      Successfully uninstalled huggingface-hub-0.19.4
  Attempting uninstall: tokenizers
    Found existing installation: tokenizers 0.13.3
    Uninstalling tokenizers-0.13.3:
      Successfully uninstalled tokenizers-0.13.3
  Attempting uninstall: transformers
    Found existing installation: transformers 4.33.3
    Uninstalling transformers-4.33.3:
      Successfully uninstalled transformers-4.33.3
  Attempting uninstall: sentence-transformers
    Found existing installation: sentence-transformers 2.2.2
    Uninstalling sentence-transformers-2.2.2:
      Successfully uninstalled sentence-transformers-2.2.2
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
datasets 3.5.0 requires fsspec[http]<=2024.12.0,>=2023.1.0, but you
have fsspec 2025.3.2 which is incompatible.
Successfully installed huggingface-hub-0.30.2 langchain-huggingface-
0.1.2 sentence-transformers-4.1.0 tokenizers-0.21.1 transformers-
4.51.3
```

Requirement already satisfied: huggingface_hub in
/usr/local/lib/python3.11/dist-packages (0.30.2)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(3.13.1)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(2025.3.2)
Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub) (23.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub) (6.0.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface_hub)
(4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub) (2025.1.31)
ERROR: Could not find a version that satisfies the requirement
transformer (from versions: none)
ERROR: No matching distribution found for transformer
Requirement already satisfied: accelerate in
/usr/local/lib/python3.11/dist-packages (1.3.0)
Requirement already satisfied: numpy<3.0.0,>=1.17 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (1.26.4)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (23.2)
Requirement already satisfied: psutil in
/usr/local/lib/python3.11/dist-packages (from accelerate) (7.0.0)
Requirement already satisfied: pyyaml in
/usr/local/lib/python3.11/dist-packages (from accelerate) (6.0.2)
Requirement already satisfied: torch>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate)
(2.0.1+cu118)

Requirement already satisfied: huggingface-hub>=0.21.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (0.30.2)
Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (0.5.2)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (3.13.1)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (2025.3.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (4.12.2)
Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy<3.0.0,>=1.17->accelerate) (1.3.8)
Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy<3.0.0,>=1.17->accelerate) (1.2.4)
Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy<3.0.0,>=1.17->accelerate) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-packages (from numpy<3.0.0,>=1.17->accelerate) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy<3.0.0,>=1.17->accelerate) (2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy<3.0.0,>=1.17->accelerate) (2.4.1)
Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (1.13.1)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.3)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.1.4)
Requirement already satisfied: triton==2.0.0 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (2.0.0)
Requirement already satisfied: cmake in

```
/usr/local/lib/python3.11/dist-packages (from triton==2.0.0-
>torch>=2.0.0->accelerate) (3.25.0)
Requirement already satisfied: lit in /usr/local/lib/python3.11/dist-
packages (from triton==2.0.0->torch>=2.0.0->accelerate) (15.0.7)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=2.0.0-
>accelerate) (2.1.5)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy<3.0.0,>=1.17-
>accelerate) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy<3.0.0,>=1.17-
>accelerate) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy<3.0.0,>=1.17->accelerate) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath-
>numpy<3.0.0,>=1.17->accelerate) (2024.2.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.21.0->accelerate) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.21.0->accelerate) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.21.0->accelerate) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.21.0->accelerate) (2025.1.31)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=2.0.0-
>accelerate) (1.3.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy<3.0.0,>=1.17->accelerate) (2024.2.0)
Collecting bitsandbytes
  Downloading bitsandbytes-0.45.5-py3-none-
manylinux_2_24_x86_64.whl.metadata (5.0 kB)
Requirement already satisfied: torch<3,>=2.0 in
/usr/local/lib/python3.11/dist-packages (from bitsandbytes)
(2.0.1+cu118)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from bitsandbytes) (1.26.4)
Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>bitsandbytes) (1.3.8)
Requirement already satisfied: mkl_random in
```

```
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>bitsandbytes) (1.2.4)
Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>bitsandbytes) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy>=1.17->bitsandbytes) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>bitsandbytes) (2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy>=1.17-
>bitsandbytes) (2.4.1)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (3.13.1)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (4.12.2)
Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (1.13.1)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (3.3)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (3.1.4)
Requirement already satisfied: triton==2.0.0 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (2.0.0)
Requirement already satisfied: cmake in
/usr/local/lib/python3.11/dist-packages (from triton==2.0.0-
>torch<3,>=2.0->bitsandbytes) (3.25.0)
Requirement already satisfied: lit in /usr/local/lib/python3.11/dist-
packages (from triton==2.0.0->torch<3,>=2.0->bitsandbytes) (15.0.7)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch<3,>=2.0-
>bitsandbytes) (2.1.5)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>bitsandbytes) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.17-
>bitsandbytes) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.17->bitsandbytes) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
```

```
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy>=1.17-
>bitsandbytes) (2024.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch<3,>=2.0-
>bitsandbytes) (1.3.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy>=1.17->bitsandbytes) (2024.2.0)
Downloading bitsandbytes-0.45.5-py3-none-manylinux_2_24_x86_64.whl
(76.1 MB)
```

76.1/76.1 MB 10.9 MB/s eta

0:00:0000:0100:01m

```
Requirement already satisfied: langchain in /usr/local/lib/python3.11/dist-
packages (0.3.23)
```

```
Requirement already satisfied: langchain-core<1.0.0,>=0.3.51 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.3.54)
```

```
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.8
in /usr/local/lib/python3.11/dist-packages (from langchain) (0.3.8)
```

```
Requirement already satisfied: langsmith<0.4,>=0.1.17 in
/usr/local/lib/python3.11/dist-packages (from langchain) (0.1.147)
```

```
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.11.3)
```

```
Requirement already satisfied: SQLAlchemy<3,>=1.4 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.0.38)
```

```
Requirement already satisfied: requests<3,>=2 in
/usr/local/lib/python3.11/dist-packages (from langchain) (2.32.3)
```

```
Requirement already satisfied: PyYAML>=5.3 in
/usr/local/lib/python3.11/dist-packages (from langchain) (6.0.2)
```

```
Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (8.5.0)
```

```
Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (1.33)
```

```
Requirement already satisfied: packaging<25,>=23.2 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (23.2)
```

```
Requirement already satisfied: typing-extensions>=4.7 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<1.0.0,>=0.3.51->langchain) (4.12.2)
```

```
Requirement already satisfied: httpx<1,>=0.23.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17-
>langchain) (0.28.1)
```

```
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17-
>langchain) (3.10.15)
```

```
Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in
/usr/local/lib/python3.11/dist-packages (from langsmith<0.4,>=0.1.17-
>langchain) (1.0.0)
```

```
Requirement already satisfied: annotated-types>=0.6.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.1 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (2.33.1)
Requirement already satisfied: typing-inspection>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain) (0.4.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2-
>langchain) (2025.1.31)
Requirement already satisfied: greenlet!=0.4.17 in
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4-
>langchain) (3.1.1)
Requirement already satisfied: anyio in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.17->langchain) (3.7.1)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.17->langchain) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*-
>httpx<1,>=0.23.0->langsmith<0.4,>=0.1.17->langchain) (0.14.0)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33-
>langchain-core<1.0.0,>=0.3.51->langchain) (3.0.0)
Requirement already satisfied: sniffio>=1.1 in
/usr/local/lib/python3.11/dist-packages (from anyio->httpx<1,>=0.23.0-
>langsmith<0.4,>=0.1.17->langchain) (1.3.1)
```

STEP 8: EMBEDDINGS + VECTOR SEARCH

First install required packages (only needed once)

```
!pip install sentence-transformers numpy
```

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

Initialize the model

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



```

# Define the input text
idea_text = "An AI system that connects agricultural producers with
potential customers"

# Create embeddings
# Note: model.encode() of a single string returns a 1D array, while
list returns 2D
idea_embedding = model.encode(idea_text) # This will be 1D array
(384,)
idea_vault = {
    "AgriMatch": model.encode("AI platform matching farmers and
buyers"), # 1D array
    "FarmConnect": model.encode("Digital marketplace for farm
products"),
    "AgroLink": model.encode("Machine learning system for agriculture
supply chain")
}

# Similarity function that handles both 1D and 2D arrays
def get_similarity_score(query_vec, stored_vec):
    # Convert both to 2D arrays if they aren't already
    if query_vec.ndim == 1:
        query_vec = query_vec.reshape(1, -1)
    if stored_vec.ndim == 1:
        stored_vec = stored_vec.reshape(1, -1)
    return np.dot(query_vec, stored_vec.T)[0][0]

# Compare the idea against all vault entries
print(f"Comparing idea: '{idea_text}'\n")
for idea, vec in idea_vault.items():
    score = get_similarity_score(idea_embedding, vec)
    print(f"[SIMILAR IDEA] {idea:10} - Similarity Score: {score:.4f}")

```

```

Requirement already satisfied: sentence-transformers in
/usr/local/lib/python3.11/dist-packages (4.1.0)
Requirement already satisfied: numpy in
/usr/local/lib/python3.11/dist-packages (1.26.4)
Requirement already satisfied: transformers<5.0.0,>=4.41.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(4.51.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-
packages (from sentence-transformers) (4.67.1)
Requirement already satisfied: torch>=1.11.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(2.0.1+cu118)
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.2.2)
Requirement already satisfied: scipy in

```

/usr/local/lib/python3.11/dist-packages (from sentence-transformers) (1.15.2)
Requirement already satisfied: huggingface-hub>=0.20.0 in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (0.30.2)
Requirement already satisfied: Pillow in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (11.1.0)
Requirement already satisfied: typing_extensions>=4.5.0 in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (4.12.2)
Requirement already satisfied: mkl_fft in /usr/local/lib/python3.11/dist-packages (from numpy) (1.3.8)
Requirement already satisfied: mkl_random in /usr/local/lib/python3.11/dist-packages (from numpy) (1.2.4)
Requirement already satisfied: mkl_umath in /usr/local/lib/python3.11/dist-packages (from numpy) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-packages (from numpy) (2025.1.0)
Requirement already satisfied: tbb4py in /usr/local/lib/python3.11/dist-packages (from numpy) (2022.1.0)
Requirement already satisfied: mkl-service in /usr/local/lib/python3.11/dist-packages (from numpy) (2.4.1)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (3.13.1)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (2025.3.2)
Requirement already satisfied: packaging>=20.9 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (23.2)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (6.0.2)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (2.32.3)
Requirement already satisfied: sympy in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (1.13.1)
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.3)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.1.4)
Requirement already satisfied: triton==2.0.0 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-

transformers) (2.0.0)
Requirement already satisfied: cmake in
/usr/local/lib/python3.11/dist-packages (from triton==2.0.0->torch>=1.11.0->sentence-transformers) (3.25.0)
Requirement already satisfied: lit in /usr/local/lib/python3.11/dist-packages (from triton==2.0.0->torch>=1.11.0->sentence-transformers) (15.0.7)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers<5.0.0,>=4.41.0->sentence-transformers) (2024.11.6)
Requirement already satisfied: tokenizers<0.22,>=0.21 in
/usr/local/lib/python3.11/dist-packages (from transformers<5.0.0,>=4.41.0->sentence-transformers) (0.21.1)
Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from transformers<5.0.0,>=4.41.0->sentence-transformers) (0.5.2)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy) (2022.1.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl->numpy) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy) (2024.2.0)
Requirement already satisfied: joblib>=1.1.1 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-transformers) (1.4.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-transformers) (3.6.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-openmp<2026,>=2024->mkl->numpy) (2024.2.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.11.0->sentence-transformers) (2.1.5)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.20.0->sentence-transformers) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.20.0->sentence-transformers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.20.0->sentence-transformers) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-

```

hub>=0.20.0->sentence-transformers) (2025.1.31)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.11.0-
>sentence-transformers) (1.3.0)

{"model_id":"fff54c4aab6b4dc984dc57c91ab2d9e7","version_major":2,"vers
ion_minor":0}

{"model_id":"8fc261bb9d6749babf6b6a7a9e39c9c6","version_major":2,"vers
ion_minor":0}

{"model_id":"8e85cf4f620f471c94981255f4850fb6","version_major":2,"vers
ion_minor":0}

{"model_id":"b72bd98f3a4a4037be043a91480e1f9c","version_major":2,"vers
ion_minor":0}

Comparing idea: 'An AI system that connects agricultural producers
with potential customers'

[SIMILAR IDEA] AgriMatch - Similarity Score: 0.7520
[SIMILAR IDEA] FarmConnect - Similarity Score: 0.5108
[SIMILAR IDEA] AgroLink - Similarity Score: 0.6637

```

Example 1: Tech Startup Ideas

```

# Example: Comparing tech startup ideas
tech_idea = "AI-powered personal assistant for scheduling meetings"

tech_vault = {
    "MeetBot": "Automated scheduling assistant using natural language
processing",
    "TimeGenius": "Calendar optimization tool for professionals",
    "SmartScheduler": "Machine learning system that arranges meetings
efficiently"
}

# Encode the query idea
tech_embedding = model.encode(tech_idea)

# Encode all vault ideas
tech_vault_embeddings = {name: model.encode(desc) for name, desc in
tech_vault.items()}

print(f"\nComparing: '{tech_idea}'\n")
for name, vec in tech_vault_embeddings.items():
    score = get_similarity_score(tech_embedding, vec)
    print(f"[TECH IDEA] {name:15} → Score: {score:.4f}")

```

```

{"model_id": "6101190e7e5a40f89a91621a1d6b564c", "version_major": 2, "version_minor": 0}

{"model_id": "a8256ac4f5cb434a9b37e2017aec6f6", "version_major": 2, "version_minor": 0}

{"model_id": "2f4f7d7e2aed4074883ec651e7222b8d", "version_major": 2, "version_minor": 0}

{"model_id": "82ae84da45a04b6399772b04aa34396e", "version_major": 2, "version_minor": 0}

```

Comparing: 'AI-powered personal assistant for scheduling meetings'

```

[TECH IDEA] MeetBot           → Score: 0.5428
[TECH IDEA] TimeGenius        → Score: 0.4203
[TECH IDEA] SmartScheduler    → Score: 0.6495

```

Example 2: Book Recommendations

```

# Example: Finding similar books based on descriptions
book_query = "A detective solves crimes using psychology and deduction"

book_vault = {
    "Sherlock Holmes": "A brilliant detective uses logic to solve mysteries",
    "The Silent Patient": "A psychological thriller about a woman who stops speaking",
    "Gone Girl": "A twisted story of a missing wife and hidden secrets"
}

# Encode the query book
book_embedding = model.encode(book_query)

# Encode all book descriptions
book_vault_embeddings = {name: model.encode(desc) for name, desc in book_vault.items()}

print(f"\nComparing: '{book_query}'\n")
for name, vec in book_vault_embeddings.items():
    score = get_similarity_score(book_embedding, vec)
    print(f"[BOOK] {name:18} → Score: {score:.4f}")

{"model_id": "da9233a819cd41368b2c29ee56afa0ed", "version_major": 2, "version_minor": 0}

```

```
{"model_id": "109d598ab5fb4eb6a17c4a8be3328c6f", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "00cbcd2bedf045dcb612554f719e998c", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "d25174ed807649568720ab5476e96ec9", "version_major": 2, "version_minor": 0}
```

Comparing: 'A detective solves crimes using psychology and deduction'

[BOOK] Sherlock Holmes → Score: 0.7209

[BOOK] The Silent Patient → Score: 0.2294

[BOOK] Gone Girl → Score: 0.2633

Example 3: Movie Similarity Search

```
# Example: Finding similar movies based on plot
```

```
movie_query = "A group of astronauts stranded in space try to return home"
```

```
movie_vault = {  
    "Interstellar": "Astronauts travel through a wormhole to save humanity",  
    "Gravity": "A medical engineer fights to survive after her shuttle is destroyed",  
    "The Martian": "An astronaut is left behind on Mars and must find a way to survive"  
}
```

```
# Encode the query movie
```

```
movie_embedding = model.encode(movie_query)
```

```
# Encode all movie plots
```

```
movie_vault_embeddings = {name: model.encode(desc) for name, desc in movie_vault.items()}
```

```
print(f"\nComparing: '{movie_query}'\n")
```

```
for name, vec in movie_vault_embeddings.items():  
    score = get_similarity_score(movie_embedding, vec)  
    print(f"[MOVIE] {name:15} → Score: {score:.4f}")
```

```
{"model_id": "caf718b346434a7b9266408c60835886", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "68263f26fe1b4d7fa46b9eebc79f5e3a", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "5c8a3910b2fc40f78b5d321c62994d44", "version_major": 2, "version_minor": 0}

{"model_id": "4f8e780607a446aa93ee176fb5d101e0", "version_major": 2, "version_minor": 0}
```

Comparing: 'A group of astronauts stranded in space try to return home'

```
[MOVIE] Interstellar    → Score: 0.5400
[MOVIE] Gravity         → Score: 0.2564
[MOVIE] The Martian     → Score: 0.4626
```

Example: Fine-Grained Technical Terms

```
query = "A deep learning model for image segmentation"
tech_vault = {
    "UNet": "Convolutional network for biomedical image segmentation",
    "ResNet": "Deep residual networks for image classification",
    "YOLO": "Real-time object detection system",
    "BERT": "Transformer model for NLP tasks"
}

# Encode
query_embedding = model.encode(query)
tech_embeddings = {name: model.encode(desc) for name, desc in
tech_vault.items()}

print(f"\nQuery: '{query}'\n")
for name, vec in tech_embeddings.items():
    score = get_similarity_score(query_embedding, vec)
    print(f"[TECH] {name:8} → Score: {score:.4f}")

{"model_id": "95dd28359907415aa6d5aeeb4b4c1863", "version_major": 2, "version_minor": 0}

{"model_id": "9e2a81795ba3482da0fa426bf3a43a65", "version_major": 2, "version_minor": 0}

{"model_id": "3a171cea98b84d06921c1df9342476fc", "version_major": 2, "version_minor": 0}

{"model_id": "b390b380450d47cfb312610e805c40e1", "version_major": 2, "version_minor": 0}

{"model_id": "0193f3dcd0d7406b940fd28c2e7d1892", "version_major": 2, "version_minor": 0}
```

Query: 'A deep learning model for image segmentation'

```
[TECH] UNet      → Score: 0.5987
[TECH] ResNet    → Score: 0.5516
[TECH] YOLO      → Score: 0.3620
[TECH] BERT      → Score: 0.1778
```

Example: Negations & Opposite Meanings

```
query = "A movie about a hero who saves the world"
movies = {
    "Hero's Journey": "A film where a protagonist defeats evil to save humanity",
    "Villain's Reign": "A story where the antagonist conquers the world",
    "The Neutral Zone": "A documentary about global politics without heroes or villains"
}
```

Encode

```
query_embedding = model.encode(query)
movie_embeddings = {name: model.encode(desc) for name, desc in movies.items()}
```

```
print(f"\nQuery: '{query}'\n")
for name, vec in movie_embeddings.items():
    score = get_similarity_score(query_embedding, vec)
    print(f"[MOVIE] {name:16} → Score: {score:.4f}")
```

```
{"model_id": "78443474ebb046e3a48d3414a231860b", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "6e3bac7fc70c452f95d58beeb2f942fe", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "3a8216e755c5445e9194a8c7756eaa93", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "500053ac97eb4c949a6a13eb83478acc", "version_major": 2, "version_minor": 0}
```

Query: 'A movie about a hero who saves the world'

```
[MOVIE] Hero's Journey  → Score: 0.7060
[MOVIE] Villain's Reign → Score: 0.4746
[MOVIE] The Neutral Zone → Score: 0.4731
```


Example: Similar Products with Subtle Differences

```
product_query = "Wireless earbuds with 30-hour battery life and noise
cancellation"

product_vault = {
    "SoundBlast Pro": "Bluetooth earbuds with 35-hour playback and
ANC",
    "AudioZen Lite": "Wireless headphones with 20-hour battery, no
noise canceling",
    "QuietComfort 30": "Noise-canceling earbuds with 25-hour battery
life",
    "BassBuds X": "30-hour wireless earbuds without ANC",
    "AirSound Mini": "True wireless earbuds with 18-hour battery and
hybrid ANC"
}

# Encode query and products
product_embedding = model.encode(product_query)
product_vault_embeddings = {name: model.encode(desc) for name, desc in
product_vault.items()}

print(f"\nQuery: '{product_query}'\n")
for name, vec in product_vault_embeddings.items():
    score = get_similarity_score(product_embedding, vec)
    print(f"[PRODUCT] {name:16} → Score: {score:.4f}")

{"model_id": "8afb669beba441258d722de596656095", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "22217a4a06094b389c4dd6c6412db8ba", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "f6df4e2d585b42589d51a6f49926ae7d", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "9763024c16e24768a147391d7995edf9", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "8e3662e4b12b4bbfaf4ae7173fab02e7", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "0e154d6b8b634bbcb868c1fa9a1266bd", "version_major": 2, "vers
ion_minor": 0}

Query: 'Wireless earbuds with 30-hour battery life and noise
cancellation'

[PRODUCT] SoundBlast Pro    → Score: 0.6720
[PRODUCT] AudioZen Lite    → Score: 0.7931
[PRODUCT] QuietComfort 30  → Score: 0.9010
```

```
[PRODUCT] BassBuds X      → Score: 0.7852
[PRODUCT] AirSound Mini   → Score: 0.8125
```

STEP 9: WRAP UP (EXPORT REPORT)

```
import json
import pandas as pd
from IPython.display import display, HTML
from datetime import datetime

# =====
# LOAD REPORT DATA (ERROR-PROOF)
# =====
try:
    with open("genbiz_ai_summary.json", "r") as f:
        report = json.load(f)
except:
    report = {
        "idea": "AI-powered crop trading platform",
        "market": {
            "size": "$150B global agri-tech market",
            "competitors": ["AgriDigital", "FarmTrade"],
            "trends": ["Blockchain in agriculture", "Direct farm-to-business sales"]
        },
        "mvp": {
            "features": ["AI pricing engine", "Escrow payments"],
            "tech": ["Python", "Solidity", "React"]
        }
    }

# =====
# FIXED ANALYSIS FUNCTIONS
# =====
def analyze_market(data):
    return {
        "size": data.get("market", {}).get("size", "Unspecified"),
        "competitors": data.get("market", {}).get("competitors", []),
        "trends": data.get("market", {}).get("trends", []),
        "score": min(10, 10 - 0.3*len(data.get("market", {})))
    }

def generate_swot(data):
    # Create DataFrame with numeric index
    df = pd.DataFrame({
        "Type": ["Strength", "Weakness", "Opportunity", "Threat"],
        "Item": [
```

```

        "<br>".join(data.get("mvp", {}).get("features", ["No
features listed"])),
        "Limited funding" if not data.get("funding") else
"Funded",
        "<br>".join(data.get("market", {}).get("trends", ["Market
data missing"])),
        "<br>".join(["Tech risks"] + data.get("market",
{}).get("competitors", []))
    ]
    })
    return df

# =====
# SIMPLIFIED VISUAL REPORT (ERROR-PROOF)
# =====
def show_report(data):
    # Market Analysis
    market = analyze_market(data)

    # Executive Summary
    display(HTML(f"""
<div style="
    background: linear-gradient(135deg, #6e8efb, #a777e3);
    color: white;
    padding: 20px;
    border-radius: 10px;
    margin-bottom: 20px;
">
        <h1>🚀 {data['idea']}
```

```

radius:10px; margin:10px 0">
    <h3> Tech Stack</h3>
    <div style="display:flex; gap:10px; flex-wrap:wrap">
        {''.join(f'<span style="background:#bbdefb;
padding:5px 10px; border-radius:15px">{tech}</span>'
        for tech in data["mvp"]["tech"])}
    </div>
</div>
"""))

# Generate SWOT rows HTML
swot = generate_swot(data)
swot_rows = []
for _, row in swot.iterrows():
    swot_rows.append(f"""
<tr style="border-bottom:1px solid #ddd">
    <td style="padding:8px">{row.Type}</td>
    <td style="padding:8px">{row.Item}</td>
</tr>
""")

# Display SWOT Analysis
display(HTML(f"""
<div style="background:#f9f9f9; padding:15px; border-radius:10px;
margin:10px 0">
    <h3> SWOT Analysis</h3>
    <table style="width:100%; border-collapse:collapse">
        <tr style="background-color:#4caf50; color:white">
            <th style="padding:8px; text-align:left">Type</th>
            <th style="padding:8px; text-align:left">Item</th>
        </tr>
        {"".join(swot_rows)}
    </table>
</div>
"""))

# Save Enhanced Report
enhanced = {
    **data,
    "analysis": {
        "market_score": market["score"],
        "generated_at": datetime.now().isoformat(),
        "report_version": "premium_1.0"
    }
}

with open("business_report_plus.json", "w") as f:
    json.dump(enhanced, f, indent=2)

```

```

display(HTML("""
<a href="business_report_plus.json" download
  style="
    display:inline-block;
    background:#4CAF50;
    color:white;
    padding:10px 15px;
    border-radius:5px;
    text-decoration:none;
    margin-top:20px;
  ">
  📄 Download Premium Report
</a>
"""))

# =====
# RUN THE SHOW
# =====
show_report(report)

<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>

```

Examples

```

import json
import pandas as pd
from IPython.display import display, HTML
from datetime import datetime

# Sample Data - Real-world FarmTech example
report = {
    "idea": "FarmConnect AI - Digital marketplace connecting small farmers directly with restaurants",
    "market": {
        "size": "$12B regional market potential",
        "competitors": ["FarmLink", "AgriMarket Online", "ProducePro"],
        "trends": [
            "30% annual growth in farm-to-table sales",
            "Chefs preferring direct farmer relationships",
            "Blockchain for food traceability"
        ]
    }
}

```

```

    ],
    },
    "mvp": {
        "features": [
            "AI-powered price matching",
            "Mobile payment escrow",
            "Delivery scheduling"
        ],
        "tech": ["Flutter", "Firebase", "TensorFlow Lite"]
    },
    "funding": True,
    "team": ["Agricultural engineer", "Full-stack developer", "Sales lead"]
}

# Analysis Functions
def analyze_market(data):
    return {
        "size": data.get("market", {}).get("size", "Unspecified"),
        "competitors": data.get("market", {}).get("competitors", []),
        "trends": data.get("market", {}).get("trends", []),
        "score": min(10, 10 - 0.3*len(data.get("market", {}))).get("competitors", []))
    }

def generate_swot(data):
    df = pd.DataFrame({
        "Type": ["Strength", "Weakness", "Opportunity", "Threat"],
        "Item": [
            "<br>• " + "<br>• ".join(data.get("mvp", {}).get("features", ["No features listed"])),
            "Limited brand recognition" if not data.get("funding")
        else "Well-capitalized",
            "<br>• " + "<br>• ".join(data.get("market", {}).get("trends", ["Market data missing"])),
            "<br>• " + "<br>• ".join(["Tech adoption barriers"] + data.get("market", {}).get("competitors", []))
        ]
    })
    return df

# Visual Report Generator
def show_report(data):
    market = analyze_market(data)

    # 1. Header
    display(HTML(f"""
<div style="
    background: linear-gradient(135deg, #6e8efb, #a777e3);
    color: white;

```

```

        padding: 20px;
        border-radius: 10px;
        margin-bottom: 20px;
    ">
    <h1>❏ {data['idea']}

```

```

        <tr style="border-bottom:1px solid #ddd">
            <td style="padding:8px">{row.Type}</td>
            <td style="padding:8px">{row.Item}</td>
        </tr>
        """ for _, row in swot.iterrows() +
    """</table>
</div>
"""))

# 5. Download Report
enhanced = {
    **data,
    "analysis": {
        "market_score": market["score"],
        "generated_at": datetime.now().strftime("%Y-%m-%d"),
        "report_version": "2.1"
    }
}

with open("farmtech_report.json", "w") as f:
    json.dump(enhanced, f, indent=2)

display(HTML("""
<a href="farmtech_report.json" download
    style="
        display:inline-block;
        background:#4CAF50;
        color:white;
        padding:10px 15px;
        border-radius:5px;
        text-decoration:none;
    ">
    □ Download Full Report
</a>
"""))

```

Execute

```
show_report(report)
```

```
<IPython.core.display.HTML object>
```

```
<IPython.core.display.HTML object>
```

```
<IPython.core.display.HTML object>
```

```
<IPython.core.display.HTML object>
```

```
<IPython.core.display.HTML object>
```

```
import json
```



```

# Sample data initialization (replace with your actual variables)
idea_text = "AI-powered farming marketplace"
response = {
    "size": "$120B potential market",
    "competitors": ["AgriTechX", "FarmConnect Pro"],
    "trends": ["Increasing digital adoption in agriculture", "Growing demand for farm-to-table"]
}
business_plan = {
    "revenue_streams": ["5% transaction fee", "Premium subscriptions"],
    "timeline": {
        "Phase 1": "MVP Development (3 months)",
        "Phase 2": "Pilot Launch (2 months)",
        "Phase 3": "Full Scale (6 months)"
    }
}
mvp_output = {
    "core_features": ["Smart matching algorithm", "Mobile payment integration"],
    "tech_stack": ["Python/Django", "React Native", "AWS"]
}
evaluation_result = {
    "strengths": ["First-mover advantage in region", "Strong farmer network"],
    "risks": ["Regulatory hurdles", "Slow farmer adoption"]
}

# Create final report
final_report = {
    "idea": idea_text,
    "market": response,
    "plan": business_plan,
    "mvp": mvp_output,
    "eval": evaluation_result
}

# Save to file
with open("genbiz_ai_summary.json", "w") as f:
    json.dump(final_report, f, indent=4) # Added indent for pretty formatting

# Verify output
print("📄 Report saved successfully! File preview:")
print(json.dumps(final_report, indent=4))

📄 Report saved successfully! File preview:
{
  "idea": "AI-powered farming marketplace",
  "market": {

```

```
    "size": "$120B potential market",
    "competitors": [
        "AgriTechX",
        "FarmConnect Pro"
    ],
    "trends": [
        "Increasing digital adoption in agriculture",
        "Growing demand for farm-to-table"
    ]
},
"plan": {
    "revenue_streams": [
        "5% transaction fee",
        "Premium subscriptions"
    ],
    "timeline": {
        "Phase 1": "MVP Development (3 months)",
        "Phase 2": "Pilot Launch (2 months)",
        "Phase 3": "Full Scale (6 months)"
    }
},
"mvp": {
    "core_features": [
        "Smart matching algorithm",
        "Mobile payment integration"
    ],
    "tech_stack": [
        "Python/Django",
        "React Native",
        "AWS"
    ]
},
"eval": {
    "strengths": [
        "First-mover advantage in region",
        "Strong farmer network"
    ],
    "risks": [
        "Regulatory hurdles",
        "Slow farmer adoption"
    ]
}
}
```

STEP-10 ALL STEPS COMPLETE: GenBiz AI Co-Founder is ready

```
!pip install PyPDF2 python-docx sentence-transformers pandas numpy ipywidgets
```

```
Requirement already satisfied: PyPDF2 in
/usr/local/lib/python3.11/dist-packages (3.0.1)
Requirement already satisfied: python-docx in
/usr/local/lib/python3.11/dist-packages (1.1.2)
Requirement already satisfied: sentence-transformers in
/usr/local/lib/python3.11/dist-packages (4.1.0)
Requirement already satisfied: pandas in
/usr/local/lib/python3.11/dist-packages (2.2.3)
Requirement already satisfied: numpy in
/usr/local/lib/python3.11/dist-packages (1.26.4)
Requirement already satisfied: ipywidgets in
/usr/local/lib/python3.11/dist-packages (8.1.5)
Requirement already satisfied: lxml>=3.1.0 in
/usr/local/lib/python3.11/dist-packages (from python-docx) (5.3.1)
Requirement already satisfied: typing-extensions>=4.9.0 in
/usr/local/lib/python3.11/dist-packages (from python-docx) (4.12.2)
Requirement already satisfied: transformers<5.0.0,>=4.41.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(4.51.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-
packages (from sentence-transformers) (4.67.1)
Requirement already satisfied: torch>=1.11.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(2.0.1+cu118)
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.2.2)
Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.15.2)
Requirement already satisfied: huggingface-hub>=0.20.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(0.30.2)
Requirement already satisfied: Pillow in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(11.1.0)
Requirement already satisfied: python-dateutil>=2.8.2 in
/usr/local/lib/python3.11/dist-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
/usr/local/lib/python3.11/dist-packages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in
/usr/local/lib/python3.11/dist-packages (from pandas) (2025.2)
```

Requirement already satisfied: mkl_fft in
/usr/local/lib/python3.11/dist-packages (from numpy) (1.3.8)
Requirement already satisfied: mkl_random in
/usr/local/lib/python3.11/dist-packages (from numpy) (1.2.4)
Requirement already satisfied: mkl_umath in
/usr/local/lib/python3.11/dist-packages (from numpy) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-
packages (from numpy) (2025.1.0)
Requirement already satisfied: tbb4py in
/usr/local/lib/python3.11/dist-packages (from numpy) (2022.1.0)
Requirement already satisfied: mkl-service in
/usr/local/lib/python3.11/dist-packages (from numpy) (2.4.1)
Requirement already satisfied: comm>=0.1.3 in
/usr/local/lib/python3.11/dist-packages (from ipywidgets) (0.2.2)
Requirement already satisfied: ipython>=6.1.0 in
/usr/local/lib/python3.11/dist-packages (from ipywidgets) (7.34.0)
Requirement already satisfied: traitlets>=4.3.1 in
/usr/local/lib/python3.11/dist-packages (from ipywidgets) (5.7.1)
Requirement already satisfied: widgetsnbextension~=4.0.12 in
/usr/local/lib/python3.11/dist-packages (from ipywidgets) (4.0.13)
Requirement already satisfied: jupyterlab-widgets~=3.0.12 in
/usr/local/lib/python3.11/dist-packages (from ipywidgets) (3.0.13)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (3.13.1)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (2025.3.2)
Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (23.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (6.0.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (2.32.3)
Requirement already satisfied: setuptools>=18.5 in
/usr/local/lib/python3.11/dist-packages (from ipython>=6.1.0-
>ipywidgets) (75.1.0)
Requirement already satisfied: jedi>=0.16 in
/usr/local/lib/python3.11/dist-packages (from ipython>=6.1.0-
>ipywidgets) (0.19.2)
Requirement already satisfied: decorator in
/usr/local/lib/python3.11/dist-packages (from ipython>=6.1.0-
>ipywidgets) (4.4.2)
Requirement already satisfied: pickleshare in
/usr/local/lib/python3.11/dist-packages (from ipython>=6.1.0-
>ipywidgets) (0.7.5)

Requirement already satisfied: prompt-toolkit!=3.0.0,!
=3.0.1,<3.1.0,>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from
ipython>=6.1.0->ipywidgets) (3.0.50)

Requirement already satisfied: pygments in
/usr/local/lib/python3.11/dist-packages (from ipython>=6.1.0-
>ipywidgets) (2.19.1)

Requirement already satisfied: backcall in
/usr/local/lib/python3.11/dist-packages (from ipython>=6.1.0-
>ipywidgets) (0.2.0)

Requirement already satisfied: matplotlib-inline in
/usr/local/lib/python3.11/dist-packages (from ipython>=6.1.0-
>ipywidgets) (0.1.7)

Requirement already satisfied: pexpect>4.3 in
/usr/local/lib/python3.11/dist-packages (from ipython>=6.1.0-
>ipywidgets) (4.9.0)

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

Requirement already satisfied: sympy in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (1.13.1)

Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (3.3)

Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (3.1.4)

Requirement already satisfied: triton==2.0.0 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (2.0.0)

Requirement already satisfied: cmake in
/usr/local/lib/python3.11/dist-packages (from triton==2.0.0-
>torch>=1.11.0->sentence-transformers) (3.25.0)

Requirement already satisfied: lit in /usr/local/lib/python3.11/dist-
packages (from triton==2.0.0->torch>=1.11.0->sentence-transformers)
(15.0.7)

Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (2024.11.6)

Requirement already satisfied: tokenizers<0.22,>=0.21 in
/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (0.21.1)

Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (0.5.2)

Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy) (2024.2.0)

Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy) (2022.1.0)

Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl->numpy)
(1.2.0)

Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy)
(2024.2.0)

Requirement already satisfied: joblib>=1.1.1 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers) (1.4.2)

Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers) (3.6.0)

Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in
/usr/local/lib/python3.11/dist-packages (from intel-
openmp<2026,>=2024->mkl->numpy) (2024.2.0)

Requirement already satisfied: parso<0.9.0,>=0.8.4 in
/usr/local/lib/python3.11/dist-packages (from jedi>=0.16-
>ipython>=6.1.0->ipywidgets) (0.8.4)

Requirement already satisfied: ptyprocess>=0.5 in
/usr/local/lib/python3.11/dist-packages (from pexpect>4.3-
>ipython>=6.1.0->ipywidgets) (0.7.0)

Requirement already satisfied: wcwidth in
/usr/local/lib/python3.11/dist-packages (from prompt-toolkit!=3.0.0,!
=3.0.1,<3.1.0,>=2.0.0->ipython>=6.1.0->ipywidgets) (0.2.13)

Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.11.0-
>sentence-transformers) (2.1.5)

Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.20.0->sentence-transformers) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.20.0->sentence-transformers) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.20.0->sentence-transformers) (2.4.0)

Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.20.0->sentence-transformers) (2025.1.31)

Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy->torch>=1.11.0-
>sentence-transformers) (1.3.0)

```
import json
import pandas as pd
import io
import numpy as np
from IPython.display import display, clear_output, HTML
from ipywidgets import FileUpload, Button, VBox, HTML, Output,
Textarea
```

```

from sentence_transformers import SentenceTransformer

# Initialize components
output_area = Output()
knowledge_base = []
model = SentenceTransformer('all-MiniLM-L6-v2', device='cpu')
final_report = {}

# -----
# Universal File Upload Handler
# -----
def handle_upload(change):
    global knowledge_base
    with output_area:
        clear_output()

        try:
            uploaded_file = None

            # Handle different notebook environments
            if hasattr(change, 'owner') and hasattr(change.owner,
'value'):
                upload_value = change.owner.value
            elif isinstance(change, dict) and 'new' in change:
                upload_value = change['new']
            else:
                upload_value = change

            # Extract file from different formats
            if isinstance(upload_value, dict) and len(upload_value) >
0:
                uploaded_file = next(iter(upload_value.values()))
            elif isinstance(upload_value, (list, tuple)) and
len(upload_value) > 0:
                uploaded_file = upload_value[0]

            if not uploaded_file or not isinstance(uploaded_file,
dict):
                display(HTML("<div style='color:red'>␣ Please select a
file to upload</div>"))
                return

            content = uploaded_file.get('content', b'')
            filename = uploaded_file.get('name', 'document')

            # Convert content to bytes
            if hasattr(content, 'tobytes'):
                content = content.tobytes()
            elif isinstance(content, str):

```

```

        content = content.encode('utf-8')

        # Process file based on type
        file_ext = filename.split('.')[1].lower() if '.' in
filename else ''
        text_content = ""

        try:
            if file_ext == 'json':
                data = json.loads(content.decode('utf-8'))
                text_content = json.dumps(data, indent=2)
            elif file_ext == 'csv':
                df = pd.read_csv(io.StringIO(content.decode('utf-
8')))
                text_content = df.to_string()
            elif file_ext in ('xls', 'xlsx'):
                df = pd.read_excel(io.BytesIO(content))
                text_content = df.to_string()
            elif file_ext == 'pdf':
                import PyPDF2
                pdf_reader = PyPDF2.PdfReader(io.BytesIO(content))
                text_content = "\n".join([page.extract_text() for
page in pdf_reader.pages])
            elif file_ext in ('doc', 'docx'):
                from docx import Document
                doc = Document(io.BytesIO(content))
                text_content = "\n".join([para.text for para in
doc.paragraphs])
            elif file_ext == 'txt':
                text_content = content.decode('utf-8')
            else:
                # Try to read as text if extension unknown
                try:
                    text_content = content.decode('utf-8')
                except:
                    raise ValueError(f"Unsupported file type:
{file_ext}")

        # Add to knowledge base
        knowledge_base.append({
            "filename": filename,
            "content": text_content[:10000], # Increased
limit
            "embedding": model.encode(text_content[:10000])
        })

        display(HTML(f"<div style='color:green'>{filename}
processed successfully!</div>"))
        display(HTML(f"<div style='border:1px solid #ddd;

```



```
padding:10px;
margin-top:10px;'><strong>Preview:</strong><br>{text_content[:300]}...
</div>"))
```

```
        except Exception as e:
            display(HTML(f"<div style='color:red'>⚠ Error
processing {filename}: {str(e)}</div>"))
            import traceback
            traceback.print_exc()

        except Exception as e:
            display(HTML(f"<div style='color:red'>⚠ System error:
{str(e)}</div>"))
            import traceback
            traceback.print_exc()
```

```
# -----
# Business Analysis Functions
# -----
def analyze_market(idea):
    # Simulate market analysis
    return {
        "size": f"${np.random.randint(10,100)}B addressable market",
        "trends": ["AI integration", "Mobile-first solutions",
"Sustainability focus"],
        "competitors": [f"Competitor {i}" for i in range(1,4)],
        "opportunities": ["Untapped regional markets", "Aging
infrastructure", "Regulatory changes"]
    }

def create_business_plan(idea, market):
    # Simulate business plan generation
    return {
        "value_prop": f"Revolutionary {idea.lower()} solution",
        "revenue_model": ["Subscription", "Transaction fees",
"Licensing"],
        "key_metrics": ["Customer acquisition cost", "LTV", "Monthly
growth"],
        "milestones": [
            f"Q1: Develop {idea} MVP",
            "Q2: Beta testing",
            "Q3: Full launch",
            "Q4: Expand to 3 new markets"
        ]
    }

def design_mvp(idea):
    # Simulate MVP design
    return {
        "core_features": [
```

```

        f"{idea} automation",
        "Real-time analytics",
        "User dashboard"
    ],
    "tech_stack": ["Python backend", "React frontend", "AWS
infrastructure"],
    "timeline": "12 weeks development",
    "cost_estimate": f"${np.random.randint(50000,200000)}"
}

def evaluate_business(idea, market, plan, mvp):
    # Simulate evaluation
    strengths = [
        f"Strong {idea} differentiation",
        f"Large {market['size']} market",
        "Experienced founding team"
    ]

    risks = [
        f"Competition from {market['competitors'][0]}",
        "Customer acquisition costs",
        "Technology risks"
    ]

    return {
        "score": f"{np.random.randint(60,95)}/100",
        "strengths": strengths,
        "risks": risks,
        "recommendations": [
            "Focus on core differentiators",
            "Secure early pilot customers",
            "Build technology moat"
        ]
    }

# -----
# Q&A System
# -----
def answer_question(question, context=None):
    if not knowledge_base and not context:
        return "No information available"

    if context:
        # Answer about the generated report
        question_embedding = model.encode(question)
        context_embedding = model.encode(str(context))
        similarity = np.dot(question_embedding, context_embedding) / (
            np.linalg.norm(question_embedding) *
np.linalg.norm(context_embedding)

```

```

    )

    if similarity > 0.3: # Threshold for considering it relevant
        return f"Regarding your business analysis:\n{str(context)[:500]}..."
    return "That question doesn't seem relevant to the business analysis."

else:
    # Answer about uploaded documents
    question_embedding = model.encode(question)
    similarities = [
        np.dot(question_embedding, doc["embedding"]) / (
            np.linalg.norm(question_embedding) *
            np.linalg.norm(doc["embedding"])
        )
        for doc in knowledge_base
    ]

    if max(similarities) < 0.2:
        return "No relevant information found in documents."

    best_match = knowledge_base[np.argmax(similarities)]
    return (
        f"From {best_match['filename']} (confidence: {max(similarities)*100:.1f}%):\n"
        f"{best_match['content'][:500]}..."
    )

# -----
# Generate Full Business Report
# -----
def generate_report(idea_text):
    market = analyze_market(idea_text)
    plan = create_business_plan(idea_text, market)
    mvp = design_mvp(idea_text)
    evaluation = evaluate_business(idea_text, market, plan, mvp)

    return {
        "idea": idea_text,
        "market": market,
        "plan": plan,
        "mvp": mvp,
        "evaluation": evaluation
    }

# -----
# User Interface Components
# -----

```

```

# File Upload
uploader = FileUpload(
    accept='*', # Accept all file types
    multiple=False,
    description='Upload Documents',
    layout={'width': '500px'}
)

# Business Idea Input
idea_input = Textarea(
    placeholder='Describe your business idea...',
    layout={'width': '100%', 'height': '100px'}
)

# Question Input
question_input = Textarea(
    placeholder='Ask about your documents or business analysis...',
    layout={'width': '100%', 'height': '100px'}
)

# Output Areas
analysis_output = Output()
qa_output = Output()

# Buttons
analyze_button = Button(
    description="Analyze Business",
    button_style='primary',
    layout={'width': '200px'}
)

question_button = Button(
    description="Ask Question",
    button_style='info',
    layout={'width': '200px'}
)

# -----
# Event Handlers
# -----
def on_analyze_click(b):
    with analysis_output:
        clear_output()
        idea = idea_input.value.strip()
        if not idea:
            display(HTML("<div style='color:red'>Please enter a  
business idea</div>"))
            return

        display(HTML("<div style='color:#666;'>🔍 Analyzing business

```

```

opportunity...</div>"))

    try:
        global final_report
        final_report = generate_report(idea)

        # Save report
        with open("business_analysis_report.json", "w") as f:
            json.dump(final_report, f)

        # Display summary
        display(HTML(f"""
            <div style='margin-top:20px;'>
                <h3 style='color:#2c3e50;'>Business Analysis
Report</h3>
                <div style='background:#f8f9fa; padding:15px; border-
radius:5px;'>
                    <strong>Idea:</strong>
{final_report['idea']}<br><br>
                    <strong>Market Size:</strong>
{final_report['market']['size']}<br>
                    <strong>Key Trends:</strong> {'',
'.join(final_report['market']['trends'])}<br><br>
                    <strong>Evaluation Score:</strong>
{final_report['evaluation']['score']}<br>
                    <strong>Recommendations:</strong> {'',
'.join(final_report['evaluation']['recommendations'])}
                </div>
                <div style='margin-top:10px; color:#666;'>
                    Full report saved as
'business_analysis_report.json'
                </div>
            </div>
            """))

        except Exception as e:
            display(HTML(f"<div style='color:red;'>⚠ Error generating
report: {str(e)}</div>"))

def on_question_click(b):
    with qa_output:
        clear_output()
        question = question_input.value.strip()
        if not question:
            display(HTML("<div style='color:red;'>Please enter a
question</div>"))
            return

        display(HTML("<div style='color:#666;'>🔍 Searching for

```

```

answers...</div>"))

    try:
        # First try to answer from report if available
        if final_report:
            answer = answer_question(question, final_report)
            if "relevant" not in answer: # If answer came from
report
                clear_output()
                display(HTML(f"""
                <div style='background:#f5f5f5; padding:15px;
border-radius:5px;'>
                    <strong>From Business Analysis:</strong><br>
                    {answer}
                </div>
                """))
                return

        # Otherwise search uploaded documents
        if knowledge_base:
            answer = answer_question(question)
            clear_output()
            display(HTML(f"""
            <div style='background:#f5f5f5; padding:15px; border-
radius:5px;'>
                {answer}
            </div>
            """))
        else:
            display(HTML("<div style='color:red'>No documents
available for Q&A</div>"))

    except Exception as e:
        display(HTML(f"<div style='color:red'>△ Error answering
question: {str(e)}</div>"))

# Set up event handlers
uploader.observe(handle_upload, names='value')
analyze_button.on_click(on_analyze_click)
question_button.on_click(on_question_click)

# -----
# Display Complete Interface
# -----
display(VBox([
    HTML("<h1 style='color:#2c3e50;'>Business Analysis System</h1>"),
    HTML("<h3>1. Upload Supporting Documents</h3>"),
    uploader,
    output_area,

```

```

HTML("<hr style='margin:20px 0; border-top:1px solid #eee;'>"),
HTML("<h3>2. Analyze Your Business Idea</h3>"),
idea_input,
analyze_button,
analysis_output,
HTML("<hr style='margin:20px 0; border-top:1px solid #eee;'>"),
HTML("<h3>3. Ask Questions</h3>"),
HTML("<p>Ask about your uploaded documents or the generated
business analysis</p>"),
question_input,
question_button,
qa_output,
HTML("""
    <div style='margin-top:30px; padding:15px; background:#f8f9fa;
border-radius:5px;'>
        <h4>System Capabilities:</h4>
        <ul>
            <li>Accepts all document formats (PDF, Word, Excel,
etc.)</li>
            <li>Generates comprehensive business analysis reports</li>
            <li>Answers questions about both documents and
analysis</li>
            <li>Saves complete report as JSON file</li>
        </ul>
    </div>
""")
]))

# Initialization message
with output_area:
    display(HTML("<div style='color:#666;'>† Upload documents to
enable document Q&A</div>"))
with analysis_output:
    display(HTML("<div style='color:#666;'>□ Enter your business idea
and click 'Analyze Business'</div>"))

{"model_id":"7cd54556eb7e44519433d3bb2a1e1955","version_major":2,"vers
ion_minor":0}

```

Thank You, the Project is Completed