

```

# Create a 'models' directory first
!mkdir -p models

# Download the model file directly into the 'models' directory
!wget -O models/mistral-7b-instruct-v0.2.Q4_K_M.gguf
"https://huggingface.co/TheBloke/Mistral-7B-Instruct-v0.2-GGUF/resolve/main/mistral-7b-instruct-v0.2.Q4_K_M.gguf"

--2025-09-21 18:57:12-- https://huggingface.co/TheBloke/Mistral-7B-Instruct-v0.2-GGUF/resolve/main/mistral-7b-instruct-v0.2.Q4_K_M.gguf
Resolving huggingface.co (huggingface.co)... 3.171.171.6, 3.171.171.128, 3.171.171.65, ...
Connecting to huggingface.co (huggingface.co)|3.171.171.6|:443...
connected.
HTTP request sent, awaiting response... 302 Found
Location:
https://cdn-lfs-us-1.hf.co/repos/72/62/726219e98582d16c24a66629a4dec1b0761b91c918e15dea2625b4293c134a92/3e0039fd0273fcbabb49228943b17831aadd55cbcbf56f0af00499be2040ccf9?response-content-disposition=inline%3B+filename%3DUTF-8%27%27mistral-7b-instruct-v0.2.Q4_K_M.gguf%3B+filename%3D%22mistral-7b-instruct-v0.2.Q4_K_M.gguf%22%3B&Expires=1758484567&Policy=eyJTdGF0ZWl1bnQl0lt7IkNvbmlRpdGlubiI6eyJEYXRlTGZvc1RoYW4i0nsiQVdT0kVWb2NoVGltZSI6MTc1ODQ4NDU2N319LCJSZXNvdXJjZSI6Imh0dHBz0i8vY2RuLWxmcyl1cy0xLmhmLmNvL3JlcG9zLzcyLzYyLzcyNjIxOWU5ODU4MmQxNmMyNGE2NjYyOWE0ZGVjMWIwNzYxYjYkYzYkx0GUxNWRLYTI2MjViNDI5M2MxMzRhOTIvM2UwMDM5ZmQwMjcZmNiZWJiNDkyMjg5NDNiMTc4MzFhYWRLNTVjYmNiZjU2ZjBhZjAwNDk5YmUyMDQwY2NmOT9yZXNwb25zZS1jb250ZW50LWRpc3Bvc2l0aW9uPSoifV19&Signature=UIgUo7qw3zIJe9FBiS7giS%7ENY783XegJGh7iSXc6q0-dpe5sIYyWy83iw9gPQ36dy6T4WQ8aBGkbRiiqX6-luu9wx9Sqp55VxPugFclw9kxi4H1n44d0bMTBnKoAWbfc58odrf5mfZw4kmK7Ro-uqztLGERzVRz5J9H7kVDbHrnX535g-00RJZEeFga0jUPo-LstjzAGtJNZpig4xBji3PM4GZqheSr3rTAG1JXTdTCLYCH7o9wCd0I73oM8zMfZFvGGgoKMLTrx3en1RkdNVmL-NelbiuD5UQXfkUavxn%7E5gDuzCJqeJ-Lu0SubmjlabCZ04kSntkpJCFNNxTMow__&Key-Pair-Id=K24J24Z295AEI9[following]
--2025-09-21 18:57:12-- https://cdn-lfs-us-1.hf.co/repos/72/62/726219e98582d16c24a66629a4dec1b0761b91c918e15dea2625b4293c134a92/3e0039fd0273fcbabb49228943b17831aadd55cbcbf56f0af00499be2040ccf9?response-content-disposition=inline%3B+filename%3DUTF-8%27%27mistral-7b-instruct-v0.2.Q4_K_M.gguf%3B+filename%3D%22mistral-7b-instruct-v0.2.Q4_K_M.gguf%22%3B&Expires=1758484567&Policy=eyJTdGF0ZWl1bnQl0lt7IkNvbmlRpdGlubiI6eyJEYXRlTGZvc1RoYW4i0nsiQVdT0kVWb2NoVGltZSI6MTc1ODQ4NDU2N319LCJSZXNvdXJjZSI6Imh0dHBz0i8vY2RuLWxmcyl1cy0xLmhmLmNvL3JlcG9zLzcyLzYyLzcyNjIxOWU5ODU4MmQxNmMyNGE2NjYyOWE0ZGVjMWIwNzYxYjYkYzYkx0GUxNWRLYTI2MjViNDI5M2MxMzRhOTIvM2UwMDM5ZmQwMjcZmNiZWJiNDkyMjg5NDNiMTc4MzFhYWRLNTVjYmNiZjU2ZjBhZjAwNDk5YmUyMDQwY2NmOT9yZXNwb25zZS1jb250ZW50LWRpc3Bvc2l0aW9uPSoifV19&Signature=UIgUo7qw3zIJe9FBiS7giS%7ENY783XegJGh7iSXc6q0-dpe5sIYyWy83iw9gPQ36dy6T4WQ8aBGkbRiiqX6-luu9wx9Sqp55VxPugFclw9kxi4H1n44d0bMTBnKoAWbfc58odrf5mfZw4kmK7Ro-

```

```
uqztLGERzVRz5J9H7kVDbHrnX535g-00RJZEeFga0jUPo-  
LstjzAGtJNZpig4xBJi3PM4GZqheSr3rTAG1JXTdTCLYCH7o9wCd0I73oM8zMfZFvGGgoK  
MLTrx3en1RkdNVmL-NelbiuD5UQXfkUavxn%7E5gDuzCJqeJ-  
Lu0SubmjlabCZ04kSntkpJCFNNxTMow__&Key-Pair-Id=K24J24Z295AEI9  
Resolving cdn-lfs-us-1.hf.co (cdn-lfs-us-1.hf.co)... 3.165.181.78,  
3.165.181.13, 3.165.181.95, ...  
Connecting to cdn-lfs-us-1.hf.co (cdn-lfs-us-1.hf.co)|  
3.165.181.78|:443... connected.  
HTTP request sent, awaiting response... 200 OK  
Length: 4368439584 (4.1G) [binary/octet-stream]  
Saving to: 'models/mistral-7b-instruct-v0.2.Q4_K_M.gguf'
```

```
models/mistral-7b-i 100%[=====>]    4.07G   264MB/s   in  
19s
```

```
2025-09-21 18:57:31 (222 MB/s) - 'models/mistral-7b-instruct-  
v0.2.Q4_K_M.gguf' saved [4368439584/4368439584]
```

```
pip install langchain_community
```

```
Collecting langchain_community
```

```
  Downloading langchain_community-0.3.29-py3-none-any.whl.metadata  
(2.9 kB)
```

```
Collecting langchain-core<2.0.0,>=0.3.75 (from langchain_community)
```

```
  Downloading langchain_core-0.3.76-py3-none-any.whl.metadata (3.7 kB)
```

```
Collecting langchain<2.0.0,>=0.3.27 (from langchain_community)
```

```
  Downloading langchain-0.3.27-py3-none-any.whl.metadata (7.8 kB)
```

```
Requirement already satisfied: SQLAlchemy<3,>=1.4 in
```

```
/usr/local/lib/python3.11/dist-packages (from langchain_community)  
(2.0.41)
```

```
Collecting requests<3,>=2.32.5 (from langchain_community)
```

```
  Downloading requests-2.32.5-py3-none-any.whl.metadata (4.9 kB)
```

```
Requirement already satisfied: PyYAML<=5.3 in
```

```
/usr/local/lib/python3.11/dist-packages (from langchain_community)  
(6.0.2)
```

```
Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in
```

```
/usr/local/lib/python3.11/dist-packages (from langchain_community)  
(3.12.13)
```

```
Requirement already satisfied: tenacity!=8.4.0,<10,>=8.1.0 in
```

```
/usr/local/lib/python3.11/dist-packages (from langchain_community)  
(8.5.0)
```

```
Requirement already satisfied: dataclasses-json<0.7,>=0.6.7 in
```

```
/usr/local/lib/python3.11/dist-packages (from langchain_community)  
(0.6.7)
```

```
Collecting pydantic-settings<3.0.0,>=2.10.1 (from langchain_community)
```

```
  Downloading pydantic_settings-2.10.1-py3-none-any.whl.metadata (3.4  
kB)
```

```
Requirement already satisfied: langsmith<=0.1.125 in
```

```
/usr/local/lib/python3.11/dist-packages (from langchain_community)
```

```
(0.4.1)
Collecting httpx-sse<1.0.0,>=0.4.0 (from langchain_community)
  Downloading httpx_sse-0.4.1-py3-none-any.whl.metadata (9.4 kB)
Requirement already satisfied: numpy>=1.26.2 in
/usr/local/lib/python3.11/dist-packages (from langchain_community)
(1.26.4)
Requirement already satisfied: aiohappyeyeballs>=2.5.0 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) (1.7.0)
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) (6.6.3)
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) (0.3.2)
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) (1.20.1)
Requirement already satisfied: marshmallow<4.0.0,>=3.18.0 in
/usr/local/lib/python3.11/dist-packages (from dataclasses-
json<0.7,>=0.6.7->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,>=0.6.7->langchain_community) (0.9.0)
Collecting langchain-text-splitters<1.0.0,>=0.3.9 (from
langchain<2.0.0,>=0.3.27->langchain_community)
  Downloading langchain_text_splitters-0.3.11-py3-none-
any.whl.metadata (1.8 kB)
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in
/usr/local/lib/python3.11/dist-packages (from
langchain<2.0.0,>=0.3.27->langchain_community) (2.11.7)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<2.0.0,>=0.3.75->langchain_community) (1.33)
Requirement already satisfied: typing-extensions>=4.7 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<2.0.0,>=0.3.75->langchain_community) (4.14.0)
Requirement already satisfied: packaging>=23.2 in
/usr/local/lib/python3.11/dist-packages (from langchain-
core<2.0.0,>=0.3.75->langchain_community) (25.0)
```

Requirement already satisfied: httpx<1,>=0.23.0 in  
/usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.125-  
>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>=0.1.125-  
>langchain\_community) (3.10.18)

Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in  
/usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.125-  
>langchain\_community) (1.0.0)

Requirement already satisfied: zstandard<0.24.0,>=0.23.0 in  
/usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.125-  
>langchain\_community) (0.23.0)

Requirement already satisfied: mkl\_fft in  
/usr/local/lib/python3.11/dist-packages (from numpy>=1.26.2-  
>langchain\_community) (1.3.8)

Requirement already satisfied: mkl\_random in  
/usr/local/lib/python3.11/dist-packages (from numpy>=1.26.2-  
>langchain\_community) (1.2.4)

Requirement already satisfied: mkl\_umath in  
/usr/local/lib/python3.11/dist-packages (from numpy>=1.26.2-  
>langchain\_community) (0.1.1)

Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-  
packages (from numpy>=1.26.2->langchain\_community) (2025.2.0)

Requirement already satisfied: tbb4py in  
/usr/local/lib/python3.11/dist-packages (from numpy>=1.26.2-  
>langchain\_community) (2022.2.0)

Requirement already satisfied: mkl-service in  
/usr/local/lib/python3.11/dist-packages (from numpy>=1.26.2-  
>langchain\_community) (2.4.1)

Collecting python-dotenv>=0.21.0 (from pydantic-  
settings<3.0.0,>=2.10.1->langchain\_community)

Downloading python\_dotenv-1.1.1-py3-none-any.whl.metadata (24 kB)

Requirement already satisfied: typing-inspection>=0.4.0 in  
/usr/local/lib/python3.11/dist-packages (from pydantic-  
settings<3.0.0,>=2.10.1->langchain\_community) (0.4.1)

Requirement already satisfied: charset\_normalizer<4,>=2 in  
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2.32.5-  
>langchain\_community) (3.4.2)

Requirement already satisfied: idna<4,>=2.5 in  
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2.32.5-  
>langchain\_community) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in  
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2.32.5-  
>langchain\_community) (2.5.0)

Requirement already satisfied: certifi>=2017.4.17 in  
/usr/local/lib/python3.11/dist-packages (from requests<3,>=2.32.5-  
>langchain\_community) (2025.6.15)

Requirement already satisfied: greenlet>=1 in  
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4-

```

>langchain_community) (3.2.3)
Requirement already satisfied: anyio in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>langsmith>=0.1.125->langchain_community) (4.9.0)
Requirement already satisfied: httpcore==1.* in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>langsmith>=0.1.125->langchain_community) (1.0.9)
Requirement already satisfied: h11>=0.16 in
/usr/local/lib/python3.11/dist-packages (from httpcore==1.*-
>httpx<1,>=0.23.0->langsmith>=0.1.125->langchain_community) (0.16.0)
Requirement already satisfied: jsonpointer>=1.9 in
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33-
>langchain-core<2.0.0,>=0.3.75->langchain_community) (3.0.0)
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<2.0.0,>=0.3.27->langchain_community) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.2 in
/usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4-
>langchain<2.0.0,>=0.3.27->langchain_community) (2.33.2)
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,>=0.6.7->langchain_community)
(1.1.0)
Requirement already satisfied: intel-openmp<2026,>=2024 in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.26.2-
>langchain_community) (2024.2.0)
Requirement already satisfied: tbb==2022.* in
/usr/local/lib/python3.11/dist-packages (from mkl->numpy>=1.26.2-
>langchain_community) (2022.2.0)
Requirement already satisfied: tcmlib==1.* in
/usr/local/lib/python3.11/dist-packages (from tbb==2022.*->mkl-
>numpy>=1.26.2->langchain_community) (1.4.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath-
>numpy>=1.26.2->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>=1.26.2->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>=0.1.125->langchain_community) (1.3.1)
Downloading langchain_community-0.3.29-py3-none-any.whl (2.5 MB)
_____ 2.5/2.5 MB 18.2 MB/s eta
0:00:00a 0:00:01
_____ 1.0/1.0 MB 34.2 MB/s eta
0:00:00
_____ 447.5/447.5 kB 18.1 MB/s eta
0:00:00

```

```
45.2/45.2 kB 2.0 MB/s eta
0:00:00
64.7/64.7 kB 2.7 MB/s eta
0:00:00
munity
  Attempting uninstall: requests
    Found existing installation: requests 2.32.4
    Uninstalling requests-2.32.4:
      Successfully uninstalled requests-2.32.4
  Attempting uninstall: langchain-core
    Found existing installation: langchain-core 0.3.66
    Uninstalling langchain-core-0.3.66:
      Successfully uninstalled langchain-core-0.3.66
  Attempting uninstall: langchain-text-splitters
    Found existing installation: langchain-text-splitters 0.3.8
    Uninstalling langchain-text-splitters-0.3.8:
      Successfully uninstalled langchain-text-splitters-0.3.8
  Attempting uninstall: langchain
    Found existing installation: langchain 0.3.26
    Uninstalling langchain-0.3.26:
      Successfully uninstalled langchain-0.3.26
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.
bigframes 2.8.0 requires google-cloud-bigquery-storage<3.0.0,>=2.30.0,
which is not installed.
datasets 3.6.0 requires fsspec[http]<=2025.3.0,>=2023.1.0, but you
have fsspec 2025.5.1 which is incompatible.
google-colab 1.0.0 requires google-auth==2.38.0, but you have google-
auth 2.40.3 which is incompatible.
google-colab 1.0.0 requires notebook==6.5.7, 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.
google-colab 1.0.0 requires requests==2.32.3, but you have requests
2.32.5 which is incompatible.
google-colab 1.0.0 requires tornado==6.4.2, but you have tornado 6.5.1
which is incompatible.
pandas-gbq 0.29.1 requires google-api-core<3.0.0,>=2.10.2, but you
have google-api-core 1.34.1 which is incompatible.
gcsfs 2025.3.2 requires fsspec==2025.3.2, but you have fsspec 2025.5.1
which is incompatible.
google-cloud-storage 2.19.0 requires google-api-
core<3.0.0dev,>=2.15.0, but you have google-api-core 1.34.1 which is
incompatible.
dataproc-spark-connect 0.7.5 requires google-api-core>=2.19, but you
have google-api-core 1.34.1 which is incompatible.
bigframes 2.8.0 requires google-cloud-
bigquery[bqstorage,pandas]>=3.31.0, but you have google-cloud-bigquery
```

3.25.0 which is incompatible.  
bigframes 2.8.0 requires rich<14,>=12.4.4, but you have rich 14.0.0 which is incompatible.  
jupyter-kernel-gateway 2.5.2 requires jupyter-client<8.0,>=5.2.0, but you have jupyter-client 8.6.3 which is incompatible.  
Successfully installed httpx-sse-0.4.1 langchain-0.3.27 langchain-core-0.3.76 langchain-text-splitters-0.3.11 langchain\_community-0.3.29 pydantic-settings-2.10.1 python-dotenv-1.1.1 requests-2.32.5  
Note: you may need to restart the kernel to use updated packages.

```
pip install pypdf
```

Requirement already satisfied: pypdf in  
/usr/local/lib/python3.11/dist-packages (5.7.0)  
Note: you may need to restart the kernel to use updated packages.

```
pip install faiss-cpu
```

Collecting faiss-cpu  
 Downloading faiss\_cpu-1.12.0-cp311-cp311-manylinux\_2\_27\_x86\_64.manylinux\_2\_28\_x86\_64.whl.metadata (5.1 kB)  
Requirement already satisfied: numpy<3.0,>=1.25.0 in  
/usr/local/lib/python3.11/dist-packages (from faiss-cpu) (1.26.4)  
Requirement already satisfied: packaging in  
/usr/local/lib/python3.11/dist-packages (from faiss-cpu) (25.0)  
Requirement already satisfied: mkl\_fft in  
/usr/local/lib/python3.11/dist-packages (from numpy<3.0,>=1.25.0->faiss-cpu) (1.3.8)  
Requirement already satisfied: mkl\_random in  
/usr/local/lib/python3.11/dist-packages (from numpy<3.0,>=1.25.0->faiss-cpu) (1.2.4)  
Requirement already satisfied: mkl\_umath in  
/usr/local/lib/python3.11/dist-packages (from numpy<3.0,>=1.25.0->faiss-cpu) (0.1.1)  
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-packages (from numpy<3.0,>=1.25.0->faiss-cpu) (2025.2.0)  
Requirement already satisfied: tbb4py in  
/usr/local/lib/python3.11/dist-packages (from numpy<3.0,>=1.25.0->faiss-cpu) (2022.2.0)  
Requirement already satisfied: mkl-service in  
/usr/local/lib/python3.11/dist-packages (from numpy<3.0,>=1.25.0->faiss-cpu) (2.4.1)  
Requirement already satisfied: intel-openmp<2026,>=2024 in  
/usr/local/lib/python3.11/dist-packages (from mkl->numpy<3.0,>=1.25.0->faiss-cpu) (2024.2.0)  
Requirement already satisfied: tbb==2022.\* in  
/usr/local/lib/python3.11/dist-packages (from mkl->numpy<3.0,>=1.25.0->faiss-cpu) (2022.2.0)  
Requirement already satisfied: tcmlib==1.\* in  
/usr/local/lib/python3.11/dist-packages (from tbb==2022.\*->mkl-

```
>numpy<3.0,>=1.25.0->faiss-cpu) (1.4.0)
Requirement already satisfied: intel-cmplr-lib-rt in
/usr/local/lib/python3.11/dist-packages (from mkl_umath-
>numpy<3.0,>=1.25.0->faiss-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<3.0,>=1.25.0->faiss-cpu) (2024.2.0)
Downloading faiss_cpu-1.12.0-cp311-cp311-
manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl (31.4 MB)
31.4/31.4 MB 46.4 MB/s eta
```

```
0:00:00:00:0100:01
```

```
ay need to restart the kernel to use updated packages.
```

```
pip install ctransformers
```

```
Collecting ctransformers
```

```
  Downloading ctransformers-0.2.27-py3-none-any.whl.metadata (17 kB)
Requirement already satisfied: huggingface-hub in
/usr/local/lib/python3.11/dist-packages (from ctransformers) (0.33.1)
Requirement already satisfied: py-cpuinfo<10.0.0,>=9.0.0 in
/usr/local/lib/python3.11/dist-packages (from ctransformers) (9.0.0)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub-
>ctransformers) (3.18.0)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub-
>ctransformers) (2025.5.1)
Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub-
>ctransformers) (25.0)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub-
>ctransformers) (6.0.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub-
>ctransformers) (2.32.5)
Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub-
>ctransformers) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub-
>ctransformers) (4.14.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.2 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub-
>ctransformers) (1.1.5)
Requirement already satisfied: charset_normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub->ctransformers) (3.4.2)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
```



```
hub->ctransformers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub->ctransformers) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub->ctransformers) (2025.6.15)
Downloading ctranformers-0.2.27-py3-none-any.whl (9.9 MB)
_____ 9.9/9.9 MB 55.3 MB/s eta
```

```
0:00:0000:0100:01
```

```
ers
```

```
Successfully installed ctranformers-0.2.27
```

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

```
!pip install -q langchain langchain-cohere langchain_groq cohere
openai
```

```
_____ 42.3/42.3 kB 1.2 MB/s eta
```

```
0:00:00
```

```
_____ 295.4/295.4 kB 6.7 MB/s eta
```

```
0:00:00ta 0:00:01
```

```
_____ 3.4/3.4 MB 44.7 MB/s eta
```

```
0:00:0000:0100:01
```

```
_____ 134.9/134.9 kB 5.5 MB/s eta
```

```
0:00:00
```

```
import os
```

```
from kaggle_secrets import UserSecretsClient
```

```
# Securely load API Keys from Kaggle Secrets
```

```
os.environ['COHERE_API_KEY'] =
```

```
UserSecretsClient().get_secret("COHERE_API_KEY")
```

```
os.environ['GROQ_API_KEY'] =
```

```
UserSecretsClient().get_secret("GROQ_API_KEY")
```

```
os.environ['OPENAI_API_KEY'] =
```

```
UserSecretsClient().get_secret("OPENAI_API_KEY")
```

```
print("✅ API keys loaded securely.")
```

```
✅ API keys loaded securely.
```

```
pip install ragas
```

```
Collecting ragas
```

```
Downloading ragas-0.3.5-py3-none-any.whl.metadata (21 kB)
```

```
Requirement already satisfied: numpy in
```

```
/usr/local/lib/python3.11/dist-packages (from ragas) (1.26.4)
```

```
Collecting datasets>=4.0.0 (from ragas)
```

```
Downloading datasets-4.1.1-py3-none-any.whl.metadata (18 kB)
```

```
Requirement already satisfied: tiktoken in
```

```
/usr/local/lib/python3.11/dist-packages (from ragas) (0.9.0)
```

Requirement already satisfied: pydantic>=2.0.0 in  
/usr/local/lib/python3.11/dist-packages (from ragas) (2.11.7)  
Requirement already satisfied: nest-asyncio in  
/usr/local/lib/python3.11/dist-packages (from ragas) (1.6.0)  
Collecting appdirs (from ragas)  
 Downloading appdirs-1.4.4-py2.py3-none-any.whl.metadata (9.0 kB)  
Collecting diskcache>=5.6.3 (from ragas)  
 Downloading diskcache-5.6.3-py3-none-any.whl.metadata (20 kB)  
Requirement already satisfied: typer in  
/usr/local/lib/python3.11/dist-packages (from ragas) (0.16.0)  
Requirement already satisfied: rich in /usr/local/lib/python3.11/dist-  
packages (from ragas) (14.0.0)  
Requirement already satisfied: openai>=1.0.0 in  
/usr/local/lib/python3.11/dist-packages (from ragas) (1.91.0)  
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-  
packages (from ragas) (4.67.1)  
Collecting instructor (from ragas)  
 Downloading instructor-1.11.3-py3-none-any.whl.metadata (11 kB)  
Requirement already satisfied: gitpython in  
/usr/local/lib/python3.11/dist-packages (from ragas) (3.1.44)  
Requirement already satisfied: pillow>=10.4.0 in  
/usr/local/lib/python3.11/dist-packages (from ragas) (11.2.1)  
Requirement already satisfied: networkx in  
/usr/local/lib/python3.11/dist-packages (from ragas) (3.5)  
Collecting scikit-network (from ragas)  
 Downloading scikit\_network-0.33.3-cp311-cp311-  
manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl.metadata (4.5 kB)  
Requirement already satisfied: langchain in  
/usr/local/lib/python3.11/dist-packages (from ragas) (0.3.27)  
Requirement already satisfied: langchain-core in  
/usr/local/lib/python3.11/dist-packages (from ragas) (0.3.76)  
Requirement already satisfied: langchain-community in  
/usr/local/lib/python3.11/dist-packages (from ragas) (0.3.29)  
Collecting langchain\_openai (from ragas)  
 Downloading langchain\_openai-0.3.33-py3-none-any.whl.metadata (2.4  
kB)  
Requirement already satisfied: filelock in  
/usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas)  
(3.18.0)  
Collecting pyarrow>=21.0.0 (from datasets>=4.0.0->ragas)  
 Downloading pyarrow-21.0.0-cp311-cp311-  
manylinux\_2\_28\_x86\_64.whl.metadata (3.3 kB)  
Requirement already satisfied: dill<0.4.1,>=0.3.0 in  
/usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas)  
(0.3.8)  
Requirement already satisfied: pandas in  
/usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas)  
(2.2.3)  
Requirement already satisfied: requests>=2.32.2 in

/usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas) (2.32.5)  
Requirement already satisfied: xxhash in /usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas) (3.5.0)  
Requirement already satisfied: multiprocessing<0.70.17 in /usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas) (0.70.16)  
Requirement already satisfied: fsspec<=2025.9.0,>=2023.1.0 in /usr/local/lib/python3.11/dist-packages (from fsspec[http]<=2025.9.0,>=2023.1.0->datasets>=4.0.0->ragas) (2025.5.1)  
Requirement already satisfied: huggingface-hub>=0.24.0 in /usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas) (0.33.1)  
Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas) (25.0)  
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from datasets>=4.0.0->ragas) (6.0.2)  
Requirement already satisfied: mkl\_fft in /usr/local/lib/python3.11/dist-packages (from numpy->ragas) (1.3.8)  
Requirement already satisfied: mkl\_random in /usr/local/lib/python3.11/dist-packages (from numpy->ragas) (1.2.4)  
Requirement already satisfied: mkl\_umath in /usr/local/lib/python3.11/dist-packages (from numpy->ragas) (0.1.1)  
Requirement already satisfied: mkl in /usr/local/lib/python3.11/dist-packages (from numpy->ragas) (2025.2.0)  
Requirement already satisfied: tbb4py in /usr/local/lib/python3.11/dist-packages (from numpy->ragas) (2022.2.0)  
Requirement already satisfied: mkl-service in /usr/local/lib/python3.11/dist-packages (from numpy->ragas) (2.4.1)  
Requirement already satisfied: anyio<5,>=3.5.0 in /usr/local/lib/python3.11/dist-packages (from openai>=1.0.0->ragas) (4.9.0)  
Requirement already satisfied: distro<2,>=1.7.0 in /usr/local/lib/python3.11/dist-packages (from openai>=1.0.0->ragas) (1.9.0)  
Requirement already satisfied: httpx<1,>=0.23.0 in /usr/local/lib/python3.11/dist-packages (from openai>=1.0.0->ragas) (0.28.1)  
Requirement already satisfied: jiter<1,>=0.4.0 in /usr/local/lib/python3.11/dist-packages (from openai>=1.0.0->ragas) (0.10.0)  
Requirement already satisfied: sniffio in /usr/local/lib/python3.11/dist-packages (from openai>=1.0.0->ragas) (1.3.1)  
Requirement already satisfied: typing-extensions<5,>=4.11 in /usr/local/lib/python3.11/dist-packages (from openai>=1.0.0->ragas)

(4.14.0)  
Requirement already satisfied: annotated-types>=0.6.0 in  
/usr/local/lib/python3.11/dist-packages (from pydantic>=2.0.0->ragas)  
(0.7.0)  
Requirement already satisfied: pydantic-core==2.33.2 in  
/usr/local/lib/python3.11/dist-packages (from pydantic>=2.0.0->ragas)  
(2.33.2)  
Requirement already satisfied: typing-inspection>=0.4.0 in  
/usr/local/lib/python3.11/dist-packages (from pydantic>=2.0.0->ragas)  
(0.4.1)  
Requirement already satisfied: gitdb<5,>=4.0.1 in  
/usr/local/lib/python3.11/dist-packages (from gitpython->ragas)  
(4.0.12)  
Requirement already satisfied: aiohttp<4.0.0,>=3.9.1 in  
/usr/local/lib/python3.11/dist-packages (from instructor->ragas)  
(3.12.13)  
Requirement already satisfied: docstring-parser<1.0,>=0.16 in  
/usr/local/lib/python3.11/dist-packages (from instructor->ragas)  
(0.16)  
Requirement already satisfied: jinja2<4.0.0,>=3.1.4 in  
/usr/local/lib/python3.11/dist-packages (from instructor->ragas)  
(3.1.6)  
Requirement already satisfied: tenacity<10.0.0,>=8.2.3 in  
/usr/local/lib/python3.11/dist-packages (from instructor->ragas)  
(8.5.0)  
Requirement already satisfied: markdown-it-py>=2.2.0 in  
/usr/local/lib/python3.11/dist-packages (from rich->ragas) (3.0.0)  
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in  
/usr/local/lib/python3.11/dist-packages (from rich->ragas) (2.19.2)  
Requirement already satisfied: click>=8.0.0 in  
/usr/local/lib/python3.11/dist-packages (from typer->ragas) (8.2.1)  
Requirement already satisfied: shellingham>=1.3.0 in  
/usr/local/lib/python3.11/dist-packages (from typer->ragas) (1.5.4)  
Requirement already satisfied: langchain-text-splitters<1.0.0,>=0.3.9  
in /usr/local/lib/python3.11/dist-packages (from langchain->ragas)  
(0.3.11)  
Requirement already satisfied: langsmith>=0.1.17 in  
/usr/local/lib/python3.11/dist-packages (from langchain->ragas)  
(0.4.1)  
Requirement already satisfied: SQLAlchemy<3,>=1.4 in  
/usr/local/lib/python3.11/dist-packages (from langchain->ragas)  
(2.0.41)  
Requirement already satisfied: jsonpatch<2.0,>=1.33 in  
/usr/local/lib/python3.11/dist-packages (from langchain-core->ragas)  
(1.33)  
Requirement already satisfied: dataclasses-json<0.7,>=0.6.7 in  
/usr/local/lib/python3.11/dist-packages (from langchain-community-  
>ragas) (0.6.7)  
Requirement already satisfied: pydantic-settings<3.0.0,>=2.10.1 in

```
/usr/local/lib/python3.11/dist-packages (from langchain-community-
>ragas) (2.10.1)
Requirement already satisfied: httpx-sse<1.0.0,>=0.4.0 in
/usr/local/lib/python3.11/dist-packages (from langchain-community-
>ragas) (0.4.0)
Collecting openai>=1.0.0 (from ragas)
  Downloading openai-1.108.1-py3-none-any.whl.metadata (29 kB)
Requirement already satisfied: regex>=2022.1.18 in
/usr/local/lib/python3.11/dist-packages (from tiktoken->ragas)
(2024.11.6)
Requirement already satisfied: scipy>=1.7.3 in
/usr/local/lib/python3.11/dist-packages (from scikit-network->ragas)
(1.15.3)
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.9.1-
>instructor->ragas) (2.6.1)
Requirement already satisfied: aiosignal>=1.1.2 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.9.1-
>instructor->ragas) (1.3.2)
Requirement already satisfied: attrs>=17.3.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.9.1-
>instructor->ragas) (25.3.0)
Requirement already satisfied: frozenlist>=1.1.1 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.9.1-
>instructor->ragas) (1.7.0)
Requirement already satisfied: multidict<7.0,>=4.5 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.9.1-
>instructor->ragas) (6.6.3)
Requirement already satisfied: propcache>=0.2.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.9.1-
>instructor->ragas) (0.3.2)
Requirement already satisfied: yarl<2.0,>=1.17.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.9.1-
>instructor->ragas) (1.20.1)
Requirement already satisfied: idna>=2.8 in
/usr/local/lib/python3.11/dist-packages (from anyio<5,>=3.5.0-
>openai>=1.0.0->ragas) (3.10)
Requirement already satisfied: marshmallow<4.0.0,>=3.18.0 in
/usr/local/lib/python3.11/dist-packages (from dataclasses-
json<0.7,>=0.6.7->langchain-community->ragas) (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,>=0.6.7->langchain-community->ragas) (0.9.0)
Requirement already satisfied: smmap<6,>=3.0.1 in
/usr/local/lib/python3.11/dist-packages (from gitdb<5,>=4.0.1-
>gitpython->ragas) (5.0.2)
Requirement already satisfied: certifi in
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0-
>openai>=1.0.0->ragas) (2025.6.15)
```

Requirement already satisfied: httpcore==1.\* in  
/usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->openai>=1.0.0->ragas) (1.0.9)

Requirement already satisfied: h11>=0.16 in  
/usr/local/lib/python3.11/dist-packages (from httpcore==1.\*->httpx<1,>=0.23.0->openai>=1.0.0->ragas) (0.16.0)

Requirement already satisfied: hf-xet<2.0.0,>=1.1.2 in  
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.24.0->datasets>=4.0.0->ragas) (1.1.5)

Requirement already satisfied: MarkupSafe>=2.0 in  
/usr/local/lib/python3.11/dist-packages (from jinja2<4.0.0,>=3.1.4->instructor->ragas) (3.0.2)

Requirement already satisfied: jsonpointer>=1.9 in  
/usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33->langchain-core->ragas) (3.0.0)

Requirement already satisfied: orjson<4.0.0,>=3.9.14 in  
/usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.17->langchain->ragas) (3.10.18)

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

Requirement already satisfied: zstandard<0.24.0,>=0.23.0 in  
/usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.17->langchain->ragas) (0.23.0)

Requirement already satisfied: mdurl~=0.1 in  
/usr/local/lib/python3.11/dist-packages (from markdown-it-py>=2.2.0->rich->ragas) (0.1.2)

Requirement already satisfied: python-dotenv>=0.21.0 in  
/usr/local/lib/python3.11/dist-packages (from pydantic-settings<3.0.0,>=2.10.1->langchain-community->ragas) (1.1.1)

Requirement already satisfied: charset\_normalizer<4,>=2 in  
/usr/local/lib/python3.11/dist-packages (from requests>=2.32.2->datasets>=4.0.0->ragas) (3.4.2)

Requirement already satisfied: urllib3<3,>=1.21.1 in  
/usr/local/lib/python3.11/dist-packages (from requests>=2.32.2->datasets>=4.0.0->ragas) (2.5.0)

Requirement already satisfied: greenlet>=1 in  
/usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4->langchain->ragas) (3.2.3)

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

Requirement already satisfied: tbb==2022.\* in  
/usr/local/lib/python3.11/dist-packages (from mkl->numpy->ragas) (2022.2.0)

Requirement already satisfied: tcmlib==1.\* in  
/usr/local/lib/python3.11/dist-packages (from tbb==2022.\*->mkl->numpy->ragas) (1.4.0)

Requirement already satisfied: intel-cmplr-lib-rt in

```

/usr/local/lib/python3.11/dist-packages (from mkl_umath->numpy->ragas)
(2024.2.0)
Requirement already satisfied: python-dateutil>=2.8.2 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets>=4.0.0-
>ragas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets>=4.0.0-
>ragas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets>=4.0.0-
>ragas) (2025.2)
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->ragas) (2024.2.0)
Requirement already satisfied: six>=1.5 in
/usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2-
>pandas->datasets>=4.0.0->ragas) (1.17.0)
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,>=0.6.7->langchain-community-
>ragas) (1.1.0)
Downloading ragas-0.3.5-py3-none-any.whl (284 kB)
----- 284.3/284.3 kB 6.0 MB/s eta
0:00:00ta 0:00:01
----- 503.6/503.6 kB 15.1 MB/s eta
0:00:00
----- 45.5/45.5 kB 1.7 MB/s eta
0:00:00
----- 155.5/155.5 kB 6.2 MB/s eta
0:00:00
----- 75.0/75.0 kB 2.6 MB/s eta
0:00:00
----- 948.4/948.4 kB 23.5 MB/s eta
0:00:0000:01
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (8.1 MB)
----- 8.1/8.1 MB 57.8 MB/s eta
0:00:0000:01:00:01
anylinux_2_28_x86_64.whl (42.8 MB)
----- 42.8/42.8 MB 27.3 MB/s eta
0:00:00:00:0100:01
pting uninstall: pyarrow
  Found existing installation: pyarrow 19.0.1
  Uninstalling pyarrow-19.0.1:
    Successfully uninstalled pyarrow-19.0.1
Attempting uninstall: openai
  Found existing installation: openai 1.91.0
  Uninstalling openai-1.91.0:
    Successfully uninstalled openai-1.91.0
Attempting uninstall: datasets

```

Found existing installation: datasets 3.6.0

Uninstalling datasets-3.6.0:

Successfully uninstalled datasets-3.6.0

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.

bigframes 2.8.0 requires google-cloud-bigquery-storage<3.0.0,>=2.30.0, which is not installed.

pylibcudf-cu12 25.2.2 requires pyarrow<20.0.0a0,>=14.0.0;

platform\_machine == "x86\_64", but you have pyarrow 21.0.0 which is incompatible.

cudf-cu12 25.2.2 requires pyarrow<20.0.0a0,>=14.0.0; platform\_machine == "x86\_64", but you have pyarrow 21.0.0 which is incompatible.

pandas-gbq 0.29.1 requires google-api-core<3.0.0,>=2.10.2, but you have google-api-core 1.34.1 which is incompatible.

bigframes 2.8.0 requires google-cloud-bigquery[bqstorage,pandas]>=3.31.0, but you have google-cloud-bigquery 3.25.0 which is incompatible.

bigframes 2.8.0 requires rich<14,>=12.4.4, but you have rich 14.0.0 which is incompatible.

Successfully installed appdirs-1.4.4 datasets-4.1.1 diskcache-5.6.3

instructor-1.11.3 langchain\_openai-0.3.33 openai-1.108.1 pyarrow-21.0.0 ragas-0.3.5 scikit-network-0.33.3

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

*# Import all necessary libraries*

from datasets import Dataset

from ragas import evaluate

from ragas.metrics import context\_precision, context\_recall

from ragas.llms import LangchainLLMWrapper

from langchain.chains import RetrievalQA

from langchain.retrievers import ContextualCompressionRetriever #

*Note: We will replace the Cohere-based reranker*

from langchain.prompts import ChatPromptTemplate

from langchain\_core.output\_parsers import StrOutputParser

from langchain\_core.documents import Document

from langchain.load import dumps, loads

from langchain\_groq import ChatGroq # <-- IMPORT GROQ

from langchain\_cohere import CohereRerank # *We still use Cohere for its high-quality reranker*

from langchain\_community.vectorstores import FAISS

from langchain\_community.embeddings import HuggingFaceEmbeddings

print("□ Step 1: Setup Complete.")

*# --- 2. INITIALIZE BASE COMPONENTS (Economics PDF) ---*

from langchain\_community.document\_loaders import PyPDFLoader

from langchain.text\_splitter import RecursiveCharacterTextSplitter



```

embedding_model = HuggingFaceEmbeddings(model_name="BAAI/bge-small-en-
v1.5")

# Load your Economics PDF
loader = PyPDFLoader("/kaggle/input/eco-book/Chap01_Eco.pdf")
documents = loader.load()

# Split into smaller chunks
text_splitter = RecursiveCharacterTextSplitter(chunk_size=300,
chunk_overlap=150)
chunked_docs = text_splitter.split_documents(documents)

# Build FAISS Vector Store from Economics content
vector_store = FAISS.from_documents(chunked_docs, embedding_model)
base_retriever = vector_store.as_retriever(search_kwargs={"k": 15})
print("✅ Step 2: Economics PDF Loaded and Vector Store Initialized.")

# --- 3. BUILD THE ADVANCED RAG-FUSION + RERANK RETRIEVER ---
# a. RAG-Fusion: Chain to generate multiple queries using Groq
template = """You are a helpful assistant that generates multiple
search queries based on a single input query.
Generate 4 search queries, one on each line, related to the following
input query:
Query: {original_query}
Queries: """
prompt_rag_fusion = ChatPromptTemplate.from_template(template)
# Use a fast Groq model for query generation
groq_llm = ChatGroq(model_name="llama-3.1-8b-instant", temperature=0)

generate_queries = (
    prompt_rag_fusion
    | groq_llm
    | StrOutputParser()
    | (lambda x: x.strip().split("\n"))
)

# b. Chain to retrieve documents for each query and flatten the list
retrieval_chain = generate_queries | base_retriever.map() | (lambda x:
[doc for sublist in x for doc in sublist])

# c. Cohere Reranker: We still use Cohere's excellent reranker for
best results
# Note: You can skip this step for a 100% Groq-only pipeline, but the
results will be better with it.
os.environ['COHERE_API_KEY'] =
UserSecretsClient().get_secret("COHERE_API_KEY")
compressor = CohereRerank(model="rerank-english-v3.0", top_n=4)

```

```

advanced_retriever = ContextualCompressionRetriever(
    base_compressor=compressor, base_retriever=retrieval_chain
)
print("📌 Step 3: Advanced RAG-Fusion + Rerank Retriever Built (Query
Generation by Groq).")

# --- 4. BUILD THE FINAL QA CHAIN USING GROQ ---
qa_chain = RetrievalQA.from_chain_type(
    llm=groq_llm, # <-- Use the fast Groq LLM for the final answer
    chain_type="stuff",
    retriever=advanced_retriever,
    return_source_documents=True
)
print("📌 Step 4: Final QA Chain Built with Groq.")

# --- 5. RUN THE RAGAS EVALUATION USING GROQ ---
eval_dataset = [
    {
        'question': "Briefly summarize the Economics PDF book.",
        'ground_truth': ""The first chapter introduces economics as
the study of scarcity, choice,
opportunity cost, and the functioning of markets. It explains the
division
into microeconomics and macroeconomics, with examples such as airline
pricing and consumer decisions.""
    },
    {
        "question": "What is economics?",
        "ground_truth": ""Economics is the social science that
studies how people, businesses, and governments make choices when
resources are
limited (scarcity) to satisfy their needs and wants. It
analyzes the production, distribution, and consumption of goods and
services,
ranging from the decisions of individual consumers and firms
(microeconomics) to the workings of entire economies and global trade
(macroeconomics). ""
    },
    {
        "question": "What are some basic concepts in economics?",
        "ground_truth": ""Basic economics concepts include scarcity
(unlimited wants, limited resources), supply and demand (how prices
are set),
opportunity cost (the value of the next best alternative given
up), utility (satisfaction from consuming goods/services), and factors
of
production (land, labor, capital, and entrepreneurship). Other
key ideas are markets, equilibrium, elasticity (responsiveness to

```

```

price changes),
    and economic growth. Scarcity, opportunity cost, and market
systems are fundamental concepts in economics."""
    }
]

questions = [ex['question'] for ex in eval_dataset]
ground_truths = [ex['ground_truth'] for ex in eval_dataset]

results = qa_chain.batch([{'query': q} for q in questions])

# Show outputs for inspection
for i, res in enumerate(results):
    print(f"\n❑ Original Question: {questions[i]}")
    print(f"❑ Model Answer: {res['result']}")
    print(f"❑ Retrieved Contexts (first 200 chars each):
{[doc.page_content[:200] for doc in res['source_documents']]}")

answers = [res['result'] for res in results]
contexts = [[doc.page_content for doc in res['source_documents']] for
res in results]

response_dataset = Dataset.from_dict({
    "question": questions,
    "answer": answers,
    "contexts": contexts,
    "ground_truth": ground_truths
})

# Use the Groq model for the evaluation task itself
ragas_llm_eval = LangchainLLMWrapper(groq_llm)

evaluation_result = evaluate(
    dataset=response_dataset,
    metrics=[context_precision, context_recall],
    llm=ragas_llm_eval # Pass the Groq wrapper
)
print("\n--- RAGAS Evaluation Results (Groq-Powered Pipeline) ---")
print(evaluation_result)

```

❑ Step 1: Setup Complete.

```

/tmp/ipykernel_36/3711923461.py:25: LangChainDeprecationWarning: The
class `HuggingFaceEmbeddings` was deprecated in LangChain 0.2.2 and
will be removed in 1.0. An updated version of the class exists in
the :class:`~langchain-huggingface` package and should be used instead.
To use it run `pip install -U :class:`~langchain-huggingface` and
import as `from :class:`~langchain_huggingface import
HuggingFaceEmbeddings`.

```

```
embedding_model = HuggingFaceEmbeddings(model_name="BAAI/bge-small-
en-v1.5")
2025-09-21 18:59:08.954327: E
external/local_xla/xla/stream_executor/cuda/cuda_fft.cc:477] Unable to
register cuFFT factory: Attempting to register factory for plugin
cuFFT when one has already been registered
WARNING: All log messages before absl::InitializeLog() is called are
written to STDERR
E0000 00:00:1758481149.352247      36 cuda_dnn.cc:8310] Unable to
register cuDNN factory: Attempting to register factory for plugin
cuDNN when one has already been registered
E0000 00:00:1758481149.447666      36 cuda_blas.cc:1418] Unable to
register cuBLAS factory: Attempting to register factory for plugin
cuBLAS when one has already been registered
```

```
{"model_id":"2402d9ee60a845e9bd1e0ff134fb86c2","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"ef3720185c734ca4b3e2827dcb573be1","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"911e4942ea794627b735ba74c98c94cd","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"abcf86e31b894c86af5c8ca479db55b8","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"391a871ea9984df3bd39ecef7f7c2cd","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"3ddb6047fec240e9a2e6ba9e3a68ede9","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"cdbbeladdd6940b3aebefb0565fd92aa","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"e84579b2a6b4462faabe12d222f0bbbb","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"7d92a9e4a61943e49a44360bccdd4f81","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"19eb456debd479d943c8be2864fe45a","version_major":2,"vers
ion_minor":0}
```

```
{"model_id":"2bbdd96471864b35950e0e33090d714d","version_major":2,"vers
ion_minor":0}
```

- Step 2: Economics PDF Loaded and Vector Store Initialized.
- Step 3: Advanced RAG-Fusion + Rerank Retriever Built (Query Generation by Groq).
- Step 4: Final QA Chain Built with Groq.

□ Original Question: Briefly summarize the Economics PDF book.  
□ Model Answer: The provided PDF book is an excerpt from "Macroeconomics: an Introduction" by Charles R. Nelson, specifically Chapter 1: "An Overview of Economics".

The chapter provides a brief introduction to economics and macroeconomics. It defines what "the economy" and "economics" are, and discusses the standard of living. However, the provided text does not contain a comprehensive summary of the book.

To provide a more detailed summary, I would need access to the full book. However, based on the provided chapter, here is a brief summary:

The book likely covers the basics of macroeconomics, including definitions, concepts, and theories related to the economy. It may discuss topics such as:

- The definition of economics and macroeconomics
- The standard of living and its measurement
- The economy as a system
- Macroeconomic variables such as GDP, inflation, and unemployment
- The role of government in the economy
- International trade and finance

Please note that this is a very brief summary, and the actual content of the book may be more comprehensive and detailed.

□ Retrieved Contexts (first 200 chars each): ['1\n \n \nMacroeconomics: an Introduction \n \nChapter 1 \n \nAn Overview of Economics \n \nInternet Edition 2009 (as of Dec. 12, 2008) \nCopyright © 2005-2009 by Charles R. Nelson \nAll rights reserved. \n\*\*\*\*\*', '1\n \n \nMacroeconomics: an Introduction \n \nChapter 1 \n \nAn Overview of Economics \n \nInternet Edition 2009 (as of Dec. 12, 2008) \nCopyright © 2005-2009 by Charles R. Nelson \nAll rights reserved. \n\*\*\*\*\*', '1\n \n \nMacroeconomics: an Introduction \n \nChapter 1 \n \nAn Overview of Economics \n \nInternet Edition 2009 (as of Dec. 12, 2008) \nCopyright © 2005-2009 by Charles R. Nelson \nAll rights reserved. \n\*\*\*\*\*', '1\n \n \nMacroeconomics: an Introduction \n \nChapter 1 \n \nAn Overview of Economics \n \nInternet Edition 2009 (as of Dec. 12, 2008) \nCopyright © 2005-2009 by Charles R. Nelson \nAll rights reserved. \n\*\*\*\*\*']

□ Original Question: What is economics?

□ Model Answer: Economics is the social science that studies the production, distribution, and consumption of goods and services. It examines how individuals, businesses, governments, and societies allocate resources to meet their unlimited wants and needs. Economics looks at how these activities are influenced by various factors, such as supply and demand, market forces, government policies, and technological advancements.

Economics is a broad field that can be divided into two main branches: microeconomics and macroeconomics. Microeconomics focuses on individual economic units, such as households and firms, and examines how they make decisions about resource allocation. Macroeconomics, on the other hand, looks at the economy as a whole and studies issues such as economic growth, inflation, and unemployment.

□ Retrieved Contexts (first 200 chars each): ['economics is a very broad one. Just as the study of the physical world is \ndivided into fields such as physics and chemistry, economics is likewise \ndivided into fields comprised of closely related to', 'economics is a very broad one. Just as the study of the physical world is \ndivided into fields such as physics and chemistry, economics is likewise \ndivided into fields comprised of closely related to', '2\nPreview \nEconomics is one of the oldest and most influential of intellectual \ndisciplines. Practically all of the great thinkers, from Aristotle to Einstein, \nhave tried their hand at it, and the gr', '2\nPreview \nEconomics is one of the oldest and most influential of intellectual \ndisciplines. Practically all of the great thinkers, from Aristotle to Einstein, \nhave tried their hand at it, and the gr']

□ Original Question: What are some basic concepts in economics?

□ Model Answer: Based on the provided context, it seems that economics is the study of the economy, focusing on the production and distribution of goods and services. Some basic concepts in economics include:

1. **\*\*Scarcity\*\***: The fundamental problem of economics, which is the limited availability of resources to meet unlimited human wants and needs.
2. **\*\*Opportunity Cost\*\***: The value of the next best alternative that is given up when a choice is made.
3. **\*\*Supply and Demand\*\***: The interaction between the quantity of a good or service that producers are willing and able to produce (supply) and the quantity that consumers are willing and able to buy (demand).
4. **\*\*Market Equilibrium\*\***: A state in which the supply and demand for a good or service are equal, resulting in a stable market price.
5. **\*\*Economic Systems\*\***: Different ways in which a society organizes the production, distribution, and exchange of goods and services, such as capitalism, socialism, and communism.
6. **\*\*Gross Domestic Product (GDP)\*\***: A measure of the total value of goods and services produced within a country's borders.
7. **\*\*Inflation\*\***: A sustained increase in the general price level of goods and services in an economy over time.
8. **\*\*Unemployment\*\***: A situation in which people are able and willing to work but are unable to find employment.
9. **\*\*Fiscal Policy\*\***: The use of government spending and taxation to influence the overall level of economic activity.
10. **\*\*Monetary Policy\*\***: The actions of a central bank to control the

money supply and interest rates to promote economic growth and stability.

These are just a few of the many basic concepts in economics.

```
▯ Retrieved Contexts (first 200 chars each): ['production and distribution of these goods and services. \nEconomics, as the study of the economy, seeks to address three basic \nquestions: \n \n• Are there fundamental principles that help us understand', 'production and distribution of these goods and services. \nEconomics, as the study of the economy, seeks to address three basic \nquestions: \n \n• Are there fundamental principles that help us understand', 'production and distribution of these goods and services. \nEconomics, as the study of the economy, seeks to address three basic \nquestions: \n \n• Are there fundamental principles that help us understand', 'production and distribution of these goods and services. \nEconomics, as the study of the economy, seeks to address three basic \nquestions: \n \n• Are there fundamental principles that help us understand']
```

```
{"model_id":"1e03bbb1b6054dd686991a29c8c825d8","version_major":2,"version_minor":0}
```

```
--- RAGAS Evaluation Results (Groq-Powered Pipeline) ---  
{'context_precision': 0.6667, 'context_recall': 0.6667}
```