

# Embed documents with watsonx's embedding

May 11, 2025

## 1 Embed documents using watsonx's embedding model

Estimated time needed: **30** minutes

### 1.1 Overview

Imagine you work in a company that handles a vast amount of text data, including documents, emails, and reports. Your task is to build an intelligent search system that can quickly and accurately retrieve relevant documents based on user queries. Traditional keyword-based search methods often fail to understand the context and semantics of the queries, leading to poor search results.

To address this challenge, you can use embedding models to convert documents into numerical vectors. These vectors capture the semantic meaning of the text, enabling more accurate and context-aware search capabilities. Document embedding is a powerful technique to convert textual data into numerical vectors, which can then be used for various downstream tasks such as search, classification, clustering, and more.

In this lab, you will learn how to use embedding models from watsonx.ai and Hugging Face to embed documents. By the end of this lab, you will be able to effectively use these embedding models to transform and utilize textual data in your projects.

### 1.2 Table of Contents

Objectives

Setup

Installing required libraries

Load data

Split data

Watsonx embedding model

Model description

Build model

Query embeddings

Document embeddings

HuggingFace embedding models

Model description

Build model

Query embeddings

Document embeddings

Exercises

Exercise 1. Using another watsonx embedding model

### 1.3 Objectives

After completing this lab, you will be able to:

- Prepare and preprocess documents for embedding
  - Use watsonx.ai and Hugging Face embedding models to generate embeddings for your documents
- 

### 1.4 Setup

For this lab, you will use the following libraries:

- [ibm-watson-ai](#) for using embedding models from IBM's watsonx.ai.
- [langchain](#), [langchain-ibm](#), [langchain-community](#) for using relevant features from LangChain.
- [sentence-transformers](#) for using embedding models from HuggingFace.

#### 1.4.1 Installing required libraries

The following required libraries are **not** preinstalled in the Skills Network Labs environment. **You need to run the following cell** to install them:

**Note:** The version is being pinned here to specify the version. It's recommended that you do this as well. Even if the library is updated in the future, the installed library could still support this lab work.

This might take around 1-2 minutes.

As `%%capture` is used to capture the installation, you won't see the output process. But after the installation completes, you will see a number beside the cell.

```
[1]: #After executing the cell, please RESTART the kernel and run all the cells.
!pip install --user "ibm-watsonx-ai==1.1.2"
!pip install --user "langchain==0.2.11"
!pip install --user "langchain-ibm==0.1.11"
!pip install --user "langchain-community==0.2.10"
!pip install --user "sentence-transformers==3.0.1"
```

Collecting ibm-watsonx-ai==1.1.2

Downloading ibm\_watsonx\_ai-1.1.2-py3-none-any.whl.metadata (6.4 kB)

Requirement already satisfied: requests in /opt/conda/lib/python3.12/site-

```

packages (from ibm-watsonx-ai==1.1.2) (2.32.3)
Requirement already satisfied: urllib3 in /opt/conda/lib/python3.12/site-
packages (from ibm-watsonx-ai==1.1.2) (2.3.0)
Collecting pandas<2.2.0,>=0.24.2 (from ibm-watsonx-ai==1.1.2)
  Downloading
pandas-2.1.4-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata
(18 kB)
Requirement already satisfied: certifi in /opt/conda/lib/python3.12/site-
packages (from ibm-watsonx-ai==1.1.2) (2024.12.14)
Collecting lomond (from ibm-watsonx-ai==1.1.2)
  Downloading lomond-0.3.3-py2.py3-none-any.whl.metadata (4.1 kB)
Collecting tabulate (from ibm-watsonx-ai==1.1.2)
  Downloading tabulate-0.9.0-py3-none-any.whl.metadata (34 kB)
Requirement already satisfied: packaging in /opt/conda/lib/python3.12/site-
packages (from ibm-watsonx-ai==1.1.2) (24.2)
Collecting ibm-cos-sdk<2.14.0,>=2.12.0 (from ibm-watsonx-ai==1.1.2)
  Downloading ibm-cos-sdk-2.13.6.tar.gz (58 kB)
  Preparing metadata (setup.py) ... done
Requirement already satisfied: importlib-metadata in
/opt/conda/lib/python3.12/site-packages (from ibm-watsonx-ai==1.1.2) (8.6.1)
Collecting ibm-cos-sdk-core==2.13.6 (from ibm-cos-sdk<2.14.0,>=2.12.0->ibm-
watsonx-ai==1.1.2)
  Downloading ibm-cos-sdk-core-2.13.6.tar.gz (1.1 MB)
                                1.1/1.1 MB
64.8 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Collecting ibm-cos-sdk-s3transfer==2.13.6 (from ibm-cos-
sdk<2.14.0,>=2.12.0->ibm-watsonx-ai==1.1.2)
  Downloading ibm-cos-sdk-s3transfer-2.13.6.tar.gz (139 kB)
  Preparing metadata (setup.py) ... done
Collecting jmespath<=1.0.1,>=0.10.0 (from ibm-cos-
sdk<2.14.0,>=2.12.0->ibm-watsonx-ai==1.1.2)
  Downloading jmespath-1.0.1-py3-none-any.whl.metadata (7.6 kB)
Requirement already satisfied: python-dateutil<3.0.0,>=2.9.0 in
/opt/conda/lib/python3.12/site-packages (from ibm-cos-sdk-core==2.13.6->ibm-cos-
sdk<2.14.0,>=2.12.0->ibm-watsonx-ai==1.1.2) (2.9.0.post0)
Collecting requests (from ibm-watsonx-ai==1.1.2)
  Downloading requests-2.32.2-py3-none-any.whl.metadata (4.6 kB)
Collecting numpy<2,>=1.26.0 (from pandas<2.2.0,>=0.24.2->ibm-watsonx-ai==1.1.2)
  Downloading
numpy-1.26.4-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata
(61 kB)
Requirement already satisfied: pytz>=2020.1 in /opt/conda/lib/python3.12/site-
packages (from pandas<2.2.0,>=0.24.2->ibm-watsonx-ai==1.1.2) (2024.2)
Collecting tzdata>=2022.1 (from pandas<2.2.0,>=0.24.2->ibm-watsonx-ai==1.1.2)
  Downloading tzdata-2025.2-py2.py3-none-any.whl.metadata (1.4 kB)
Requirement already satisfied: charset-normalizer<4,>=2 in
/opt/conda/lib/python3.12/site-packages (from requests->ibm-watsonx-ai==1.1.2)

```

(3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.12/site-packages (from requests->ibm-watsonx-ai==1.1.2) (3.10)

Requirement already satisfied: zipp>=3.20 in /opt/conda/lib/python3.12/site-packages (from importlib-metadata->ibm-watsonx-ai==1.1.2) (3.21.0)

Requirement already satisfied: six>=1.10.0 in /opt/conda/lib/python3.12/site-packages (from lomond->ibm-watsonx-ai==1.1.2) (1.17.0)

Downloading ibm\_watsonx\_ai-1.1.2-py3-none-any.whl (966 kB)

966.6/966.6 kB

71.3 MB/s eta 0:00:00

Downloading

pandas-2.1.4-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (11.7 MB)

11.7/11.7 MB

170.0 MB/s eta 0:00:00

Downloading requests-2.32.2-py3-none-any.whl (63 kB)

Downloading lomond-0.3.3-py2.py3-none-any.whl (35 kB)

Downloading tabulate-0.9.0-py3-none-any.whl (35 kB)

Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)

Downloading

numpy-1.26.4-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (18.0 MB)

18.0/18.0 MB

188.0 MB/s eta 0:00:00

Downloading tzdata-2025.2-py2.py3-none-any.whl (347 kB)

Building wheels for collected packages: ibm-cos-sdk, ibm-cos-sdk-core, ibm-cos-sdk-s3transfer

Building wheel for ibm-cos-sdk (setup.py) ... one

Created wheel for ibm-cos-sdk: filename=ibm\_cos\_sdk-2.13.6-py3-none-any.whl size=77287

sha256=5ca5c6d9c0366141e064b45dc39a1fab87bd7f00ea7f76e667d497f339cc7f7f

Stored in directory: /home/jupyterlab/.cache/pip/wheels/17/c5/52/64bd703cf3cc4a83ba0e39d9b3eedc3a15afced69bbe6cf306

Building wheel for ibm-cos-sdk-core (setup.py) ... one

Created wheel for ibm-cos-sdk-core:  
filename=ibm\_cos\_sdk\_core-2.13.6-py3-none-any.whl size=661517

sha256=879e8aa5f2f9f11efc3c4a1ff5b26950ac6ef97c278b2c4e043c48b715ea0b5d

Stored in directory: /home/jupyterlab/.cache/pip/wheels/3f/11/58/646f89b99861739fce853f8776e5037d0f704d5d9623a07f7c

Building wheel for ibm-cos-sdk-s3transfer (setup.py) ... one

Created wheel for ibm-cos-sdk-s3transfer:  
filename=ibm\_cos\_sdk\_s3transfer-2.13.6-py3-none-any.whl size=90257

sha256=89bc26fac638e0d1dd9f09044a5787b53cc46f46497027992dad9e08de27ad80

Stored in directory: /home/jupyterlab/.cache/pip/wheels/52/9d/54/09e4008d97d8a2e38ecff56eea0d60ef27a77d9a313cc77360

Successfully built ibm-cos-sdk ibm-cos-sdk-core ibm-cos-sdk-s3transfer

Installing collected packages: tzdata, tabulate, requests, numpy, lomond, jmespath, pandas, ibm-cos-sdk-core, ibm-cos-sdk-s3transfer, ibm-cos-sdk, ibm-

```

watsonx-ai
Successfully installed ibm-cos-sdk-2.13.6 ibm-cos-sdk-core-2.13.6 ibm-cos-
sdk-s3transfer-2.13.6 ibm-watsonx-ai-1.1.2 jmespath-1.0.1 lomond-0.3.3
numpy-1.26.4 pandas-2.1.4 requests-2.32.2 tabulate-0.9.0 tzdata-2025.2
Collecting langchain==0.2.11
  Downloading langchain-0.2.11-py3-none-any.whl.metadata (7.1 kB)
Requirement already satisfied: PyYAML>=5.3 in /opt/conda/lib/python3.12/site-
packages (from langchain==0.2.11) (6.0.2)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in
/opt/conda/lib/python3.12/site-packages (from langchain==0.2.11) (2.0.37)
Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in
/opt/conda/lib/python3.12/site-packages (from langchain==0.2.11) (3.11.18)
Collecting langchain-core<0.3.0,>=0.2.23 (from langchain==0.2.11)
  Downloading langchain_core-0.2.43-py3-none-any.whl.metadata (6.2 kB)
Collecting langchain-text-splitters<0.3.0,>=0.2.0 (from langchain==0.2.11)
  Downloading langchain_text_splitters-0.2.4-py3-none-any.whl.metadata (2.3 kB)
Collecting langsmith<0.2.0,>=0.1.17 (from langchain==0.2.11)
  Downloading langsmith-0.1.147-py3-none-any.whl.metadata (14 kB)
Requirement already satisfied: numpy<2.0.0,>=1.26.0 in
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain==0.2.11)
(1.26.4)
Requirement already satisfied: pydantic<3,>=1 in /opt/conda/lib/python3.12/site-
packages (from langchain==0.2.11) (2.10.6)
Requirement already satisfied: requests<3,>=2 in
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain==0.2.11)
(2.32.2)
Collecting tenacity!=8.4.0,<9.0.0,>=8.1.0 (from langchain==0.2.11)
  Downloading tenacity-8.5.0-py3-none-any.whl.metadata (1.2 kB)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in
/opt/conda/lib/python3.12/site-packages (from
aiohttp<4.0.0,>=3.8.3->langchain==0.2.11) (2.6.1)
Requirement already satisfied: aiosignal>=1.1.2 in
/opt/conda/lib/python3.12/site-packages (from
aiohttp<4.0.0,>=3.8.3->langchain==0.2.11) (1.3.2)
Requirement already satisfied: attrs>=17.3.0 in /opt/conda/lib/python3.12/site-
packages (from aiohttp<4.0.0,>=3.8.3->langchain==0.2.11) (25.1.0)
Requirement already satisfied: frozenlist>=1.1.1 in
/opt/conda/lib/python3.12/site-packages (from
aiohttp<4.0.0,>=3.8.3->langchain==0.2.11) (1.6.0)
Requirement already satisfied: multidict<7.0,>=4.5 in
/opt/conda/lib/python3.12/site-packages (from
aiohttp<4.0.0,>=3.8.3->langchain==0.2.11) (6.4.3)
Requirement already satisfied: propcache>=0.2.0 in
/opt/conda/lib/python3.12/site-packages (from
aiohttp<4.0.0,>=3.8.3->langchain==0.2.11) (0.3.1)
Requirement already satisfied: yarl<2.0,>=1.17.0 in
/opt/conda/lib/python3.12/site-packages (from
aiohttp<4.0.0,>=3.8.3->langchain==0.2.11) (1.20.0)

```

Requirement already satisfied: jsonpatch<2.0,>=1.33 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3.0,>=0.2.23->langchain==0.2.11) (1.33)

Requirement already satisfied: packaging<25,>=23.2 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3.0,>=0.2.23->langchain==0.2.11) (24.2)

Requirement already satisfied: typing-extensions>=4.7 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3.0,>=0.2.23->langchain==0.2.11) (4.12.2)

Requirement already satisfied: httpx<1,>=0.23.0 in /opt/conda/lib/python3.12/site-packages (from langsmith<0.2.0,>=0.1.17->langchain==0.2.11) (0.28.1)

Collecting orjson<4.0.0,>=3.9.14 (from langsmith<0.2.0,>=0.1.17->langchain==0.2.11)

Downloading orjson-3.10.18-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl.metadata (41 kB)

Collecting requests-toolbelt<2.0.0,>=1.0.0 (from langsmith<0.2.0,>=0.1.17->langchain==0.2.11)

Downloading requests\_toolbelt-1.0.0-py2.py3-none-any.whl.metadata (14 kB)

Requirement already satisfied: annotated-types>=0.6.0 in /opt/conda/lib/python3.12/site-packages (from pydantic<3,>=1->langchain==0.2.11) (0.7.0)

Requirement already satisfied: pydantic-core==2.27.2 in /opt/conda/lib/python3.12/site-packages (from pydantic<3,>=1->langchain==0.2.11) (2.27.2)

Requirement already satisfied: charset-normalizer<4,>=2 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain==0.2.11) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain==0.2.11) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain==0.2.11) (2.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain==0.2.11) (2024.12.14)

Requirement already satisfied: greenlet!=0.4.17 in /opt/conda/lib/python3.12/site-packages (from SQLAlchemy<3,>=1.4->langchain==0.2.11) (3.1.1)

Requirement already satisfied: anyio in /opt/conda/lib/python3.12/site-packages (from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.17->langchain==0.2.11) (4.8.0)

Requirement already satisfied: httpcore==1.\* in /opt/conda/lib/python3.12/site-packages (from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.17->langchain==0.2.11) (1.0.7)

Requirement already satisfied: h11<0.15,>=0.13 in /opt/conda/lib/python3.12/site-packages (from httpcore==1.\*->httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.17->langchain==0.2.11) (0.14.0)

Requirement already satisfied: jsonpointer>=1.9 in /opt/conda/lib/python3.12/site-packages (from jsonpatch<2.0,>=1.33->langchain-core<0.3.0,>=0.2.23->langchain==0.2.11) (3.0.0)

Requirement already satisfied: sniffio>=1.1 in /opt/conda/lib/python3.12/site-packages (from anyio->httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.17->langchain==0.2.11) (1.3.1)

Downloading langchain-0.2.11-py3-none-any.whl (990 kB)  
990.3/990.3 kB

76.6 MB/s eta 0:00:00

Downloading langchain\_core-0.2.43-py3-none-any.whl (397 kB)

Downloading langchain\_text\_splitters-0.2.4-py3-none-any.whl (25 kB)

Downloading langsmith-0.1.147-py3-none-any.whl (311 kB)

Downloading tenacity-8.5.0-py3-none-any.whl (28 kB)

Downloading orjson-3.10.18-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (133 kB)

Downloading requests\_toolbelt-1.0.0-py2.py3-none-any.whl (54 kB)

Installing collected packages: tenacity, orjson, requests-toolbelt, langsmith, langchain-core, langchain-text-splitters, langchain

Successfully installed langchain-0.2.11 langchain-core-0.2.43 langchain-text-splitters-0.2.4 langsmith-0.1.147 orjson-3.10.18 requests-toolbelt-1.0.0 tenacity-8.5.0

Collecting langchain-ibm==0.1.11

Downloading langchain\_ibm-0.1.11-py3-none-any.whl.metadata (5.2 kB)

Requirement already satisfied: ibm-watsonx-ai<2.0.0,>=1.0.8 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-ibm==0.1.11) (1.1.2)

Requirement already satisfied: langchain-core<0.3,>=0.2.2 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-ibm==0.1.11) (0.2.43)

Requirement already satisfied: requests in /home/jupyterlab/.local/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (2.32.2)

Requirement already satisfied: urllib3 in /opt/conda/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (2.3.0)

Requirement already satisfied: pandas<2.2.0,>=0.24.2 in /home/jupyterlab/.local/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (2.1.4)

Requirement already satisfied: certifi in /opt/conda/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (2024.12.14)

Requirement already satisfied: lomond in /home/jupyterlab/.local/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (0.3.3)

Requirement already satisfied: tabulate in /home/jupyterlab/.local/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (0.9.0)

Requirement already satisfied: packaging in /opt/conda/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (24.2)

Requirement already satisfied: ibm-cos-sdk<2.14.0,>=2.12.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (2.13.6)

Requirement already satisfied: importlib-metadata in /opt/conda/lib/python3.12/site-packages (from ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (8.6.1)

Requirement already satisfied: PyYAML>=5.3 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11) (6.0.2)

Requirement already satisfied: jsonpatch<2.0,>=1.33 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11) (1.33)

Requirement already satisfied: langsmith<0.2.0,>=0.1.112 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11) (0.1.147)

Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11) (2.10.6)

Requirement already satisfied: tenacity!=8.4.0,<9.0.0,>=8.1.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11) (8.5.0)

Requirement already satisfied: typing-extensions>=4.7 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11) (4.12.2)

Requirement already satisfied: ibm-cos-sdk-core==2.13.6 in /home/jupyterlab/.local/lib/python3.12/site-packages (from ibm-cos-sdk<2.14.0,>=2.12.0->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (2.13.6)

Requirement already satisfied: ibm-cos-sdk-s3transfer==2.13.6 in /home/jupyterlab/.local/lib/python3.12/site-packages (from ibm-cos-sdk<2.14.0,>=2.12.0->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (2.13.6)

Requirement already satisfied: jmespath<=1.0.1,>=0.10.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from ibm-cos-sdk<2.14.0,>=2.12.0->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (1.0.1)

Requirement already satisfied: python-dateutil<3.0.0,>=2.9.0 in /opt/conda/lib/python3.12/site-packages (from ibm-cos-sdk-core==2.13.6->ibm-cos-sdk<2.14.0,>=2.12.0->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (2.9.0.post0)

Requirement already satisfied: jsonpointer>=1.9 in /opt/conda/lib/python3.12/site-packages (from jsonpatch<2.0,>=1.33->langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11) (3.0.0)

Requirement already satisfied: httpx<1,>=0.23.0 in /opt/conda/lib/python3.12/site-packages (from langsmith<0.2.0,>=0.1.112->langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11) (0.28.1)

Requirement already satisfied: orjson<4.0.0,>=3.9.14 in /home/jupyterlab/.local/lib/python3.12/site-packages (from



langsmith<0.2.0,>=0.1.112->langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11)  
(3.10.18)

Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from  
langsmith<0.2.0,>=0.1.112->langchain-core<0.3,>=0.2.2->langchain-ibm==0.1.11)  
(1.0.0)

Requirement already satisfied: numpy<2,>=1.26.0 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from  
pandas<2.2.0,>=0.24.2->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11)  
(1.26.4)

Requirement already satisfied: pytz>=2020.1 in /opt/conda/lib/python3.12/site-  
packages (from pandas<2.2.0,>=0.24.2->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-  
ibm==0.1.11) (2024.2)

Requirement already satisfied: tzdata>=2022.1 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from  
pandas<2.2.0,>=0.24.2->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11)  
(2025.2)

Requirement already satisfied: annotated-types>=0.6.0 in  
/opt/conda/lib/python3.12/site-packages (from pydantic<3.0.0,>=2.7.4->langchain-  
core<0.3,>=0.2.2->langchain-ibm==0.1.11) (0.7.0)

Requirement already satisfied: pydantic-core==2.27.2 in  
/opt/conda/lib/python3.12/site-packages (from pydantic<3.0.0,>=2.7.4->langchain-  
core<0.3,>=0.2.2->langchain-ibm==0.1.11) (2.27.2)

Requirement already satisfied: charset-normalizer<4,>=2 in  
/opt/conda/lib/python3.12/site-packages (from requests->ibm-watsonx-  
ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.12/site-  
packages (from requests->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11)  
(3.10)

Requirement already satisfied: zipp>=3.20 in /opt/conda/lib/python3.12/site-  
packages (from importlib-metadata->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-  
ibm==0.1.11) (3.21.0)

Requirement already satisfied: six>=1.10.0 in /opt/conda/lib/python3.12/site-  
packages (from lomond->ibm-watsonx-ai<2.0.0,>=1.0.8->langchain-ibm==0.1.11)  
(1.17.0)

Requirement already satisfied: anyio in /opt/conda/lib/python3.12/site-packages  
(from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.112->langchain-  
core<0.3,>=0.2.2->langchain-ibm==0.1.11) (4.8.0)

Requirement already satisfied: httpcore==1.\* in /opt/conda/lib/python3.12/site-  
packages (from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.112->langchain-  
core<0.3,>=0.2.2->langchain-ibm==0.1.11) (1.0.7)

Requirement already satisfied: h11<0.15,>=0.13 in  
/opt/conda/lib/python3.12/site-packages (from  
httpcore==1.\*->httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.112->langchain-  
core<0.3,>=0.2.2->langchain-ibm==0.1.11) (0.14.0)

Requirement already satisfied: sniffio>=1.1 in /opt/conda/lib/python3.12/site-  
packages (from anyio->httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.112->langchain-  
core<0.3,>=0.2.2->langchain-ibm==0.1.11) (1.3.1)

Downloading langchain\_ibm-0.1.11-py3-none-any.whl (19 kB)  
 Installing collected packages: langchain-ibm  
 Successfully installed langchain-ibm-0.1.11  
 Collecting langchain-community==0.2.10  
   Downloading langchain\_community-0.2.10-py3-none-any.whl.metadata (2.7 kB)  
 Requirement already satisfied: PyYAML>=5.3 in /opt/conda/lib/python3.12/site-packages (from langchain-community==0.2.10) (6.0.2)  
 Requirement already satisfied: SQLAlchemy<3,>=1.4 in /opt/conda/lib/python3.12/site-packages (from langchain-community==0.2.10) (2.0.37)  
 Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in /opt/conda/lib/python3.12/site-packages (from langchain-community==0.2.10) (3.11.18)  
 Collecting dataclasses-json<0.7,>=0.5.7 (from langchain-community==0.2.10)  
   Downloading dataclasses\_json-0.6.7-py3-none-any.whl.metadata (25 kB)  
 Requirement already satisfied: langchain<0.3.0,>=0.2.9 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-community==0.2.10) (0.2.11)  
 Requirement already satisfied: langchain-core<0.3.0,>=0.2.23 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-community==0.2.10) (0.2.43)  
 Requirement already satisfied: langsmith<0.2.0,>=0.1.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-community==0.2.10) (0.1.147)  
 Requirement already satisfied: numpy<2.0.0,>=1.26.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-community==0.2.10) (1.26.4)  
 Requirement already satisfied: requests<3,>=2 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-community==0.2.10) (2.32.2)  
 Requirement already satisfied: tenacity!=8.4.0,<9.0.0,>=8.1.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain-community==0.2.10) (8.5.0)  
 Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /opt/conda/lib/python3.12/site-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community==0.2.10) (2.6.1)  
 Requirement already satisfied: aiosignal>=1.1.2 in /opt/conda/lib/python3.12/site-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community==0.2.10) (1.3.2)  
 Requirement already satisfied: attrs>=17.3.0 in /opt/conda/lib/python3.12/site-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community==0.2.10) (25.1.0)  
 Requirement already satisfied: frozenlist>=1.1.1 in /opt/conda/lib/python3.12/site-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community==0.2.10) (1.6.0)  
 Requirement already satisfied: multidict<7.0,>=4.5 in /opt/conda/lib/python3.12/site-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community==0.2.10) (6.4.3)  
 Requirement already satisfied: propcache>=0.2.0 in

/opt/conda/lib/python3.12/site-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community==0.2.10) (0.3.1)

Requirement already satisfied: yarll<2.0,>=1.17.0 in /opt/conda/lib/python3.12/site-packages (from aiohttp<4.0.0,>=3.8.3->langchain-community==0.2.10) (1.20.0)

Collecting marshmallow<4.0.0,>=3.18.0 (from dataclasses-json<0.7,>=0.5.7->langchain-community==0.2.10)

Downloading marshmallow-3.26.1-py3-none-any.whl.metadata (7.3 kB)

Collecting typing-inspect<1,>=0.4.0 (from dataclasses-json<0.7,>=0.5.7->langchain-community==0.2.10)

Downloading typing\_inspect-0.9.0-py3-none-any.whl.metadata (1.5 kB)

Requirement already satisfied: langchain-text-splitters<0.3.0,>=0.2.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langchain<0.3.0,>=0.2.9->langchain-community==0.2.10) (0.2.4)

Requirement already satisfied: pydantic<3,>=1 in /opt/conda/lib/python3.12/site-packages (from langchain<0.3.0,>=0.2.9->langchain-community==0.2.10) (2.10.6)

Requirement already satisfied: jsonpatch<2.0,>=1.33 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3.0,>=0.2.23->langchain-community==0.2.10) (1.33)

Requirement already satisfied: packaging<25,>=23.2 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3.0,>=0.2.23->langchain-community==0.2.10) (24.2)

Requirement already satisfied: typing-extensions>=4.7 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3.0,>=0.2.23->langchain-community==0.2.10) (4.12.2)

Requirement already satisfied: httpx<1,>=0.23.0 in /opt/conda/lib/python3.12/site-packages (from langsmith<0.2.0,>=0.1.0->langchain-community==0.2.10) (0.28.1)

Requirement already satisfied: orjson<4.0.0,>=3.9.14 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langsmith<0.2.0,>=0.1.0->langchain-community==0.2.10) (3.10.18)

Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langsmith<0.2.0,>=0.1.0->langchain-community==0.2.10) (1.0.0)

Requirement already satisfied: charset-normalizer<4,>=2 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain-community==0.2.10) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain-community==0.2.10) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain-community==0.2.10) (2.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain-community==0.2.10) (2024.12.14)

Requirement already satisfied: greenlet!=0.4.17 in /opt/conda/lib/python3.12/site-packages (from SQLAlchemy<3,>=1.4->langchain-community==0.2.10) (3.1.1)

Requirement already satisfied: anyio in /opt/conda/lib/python3.12/site-packages (from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain-community==0.2.10) (4.8.0)

Requirement already satisfied: httpcore==1.\* in /opt/conda/lib/python3.12/site-packages (from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain-community==0.2.10) (1.0.7)

Requirement already satisfied: h11<0.15,>=0.13 in /opt/conda/lib/python3.12/site-packages (from httpcore==1.\*->httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain-community==0.2.10) (0.14.0)

Requirement already satisfied: jsonpointer>=1.9 in /opt/conda/lib/python3.12/site-packages (from jsonpatch<2.0,>=1.33->langchain-core<0.3.0,>=0.2.23->langchain-community==0.2.10) (3.0.0)

Requirement already satisfied: annotated-types>=0.6.0 in /opt/conda/lib/python3.12/site-packages (from pydantic<3,>=1->langchain<0.3.0,>=0.2.9->langchain-community==0.2.10) (0.7.0)

Requirement already satisfied: pydantic-core==2.27.2 in /opt/conda/lib/python3.12/site-packages (from pydantic<3,>=1->langchain<0.3.0,>=0.2.9->langchain-community==0.2.10) (2.27.2)

Collecting mpy-extensions>=0.3.0 (from typing-inspect<1,>=0.4.0->dataclasses-json<0.7,>=0.5.7->langchain-community==0.2.10)

Downloading mpy\_extensions-1.1.0-py3-none-any.whl.metadata (1.1 kB)

Requirement already satisfied: sniffio>=1.1 in /opt/conda/lib/python3.12/site-packages (from anyio->httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain-community==0.2.10) (1.3.1)

Downloading langchain\_community-0.2.10-py3-none-any.whl (2.3 MB)

2.3/2.3 MB

102.5 MB/s eta 0:00:00

Downloading dataclasses\_json-0.6.7-py3-none-any.whl (28 kB)

Downloading marshmallow-3.26.1-py3-none-any.whl (50 kB)

Downloading typing\_inspect-0.9.0-py3-none-any.whl (8.8 kB)

Downloading mpy\_extensions-1.1.0-py3-none-any.whl (5.0 kB)

Installing collected packages: mpy-extensions, marshmallow, typing-inspect, dataclasses-json, langchain-community

Successfully installed dataclasses-json-0.6.7 langchain-community-0.2.10 marshmallow-3.26.1 mpy-extensions-1.1.0 typing-inspect-0.9.0

Collecting sentence-transformers==3.0.1

Downloading sentence\_transformers-3.0.1-py3-none-any.whl.metadata (10 kB)

Collecting transformers<5.0.0,>=4.34.0 (from sentence-transformers==3.0.1)

Downloading transformers-4.51.3-py3-none-any.whl.metadata (38 kB)

Requirement already satisfied: tqdm in /opt/conda/lib/python3.12/site-packages (from sentence-transformers==3.0.1) (4.67.1)

Collecting torch>=1.11.0 (from sentence-transformers==3.0.1)

Downloading torch-2.7.0-cp312-cp312-manylinux\_2\_28\_x86\_64.whl.metadata (29 kB)

Requirement already satisfied: numpy in /home/jupyterlab/.local/lib/python3.12/site-packages (from sentence-transformers==3.0.1) (1.26.4)

Collecting scikit-learn (from sentence-transformers==3.0.1)

Downloading scikit\_learn-1.6.1-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl.metadata (18 kB)  
 Collecting scipy (from sentence-transformers==3.0.1)  
 Downloading scipy-1.15.3-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl.metadata (61 kB)  
 Collecting huggingface-hub>=0.15.1 (from sentence-transformers==3.0.1)  
 Downloading huggingface\_hub-0.31.1-py3-none-any.whl.metadata (13 kB)  
 Collecting Pillow (from sentence-transformers==3.0.1)  
 Downloading pillow-11.2.1-cp312-cp312-manylinux\_2\_28\_x86\_64.whl.metadata (8.9 kB)  
 Collecting filelock (from huggingface-hub>=0.15.1->sentence-transformers==3.0.1)  
 Downloading filelock-3.18.0-py3-none-any.whl.metadata (2.9 kB)  
 Collecting fsspec>=2023.5.0 (from huggingface-hub>=0.15.1->sentence-transformers==3.0.1)  
 Downloading fsspec-2025.3.2-py3-none-any.whl.metadata (11 kB)  
 Requirement already satisfied: packaging>=20.9 in /opt/conda/lib/python3.12/site-packages (from huggingface-hub>=0.15.1->sentence-transformers==3.0.1) (24.2)  
 Requirement already satisfied: pyyaml>=5.1 in /opt/conda/lib/python3.12/site-packages (from huggingface-hub>=0.15.1->sentence-transformers==3.0.1) (6.0.2)  
 Requirement already satisfied: requests in /home/jupyterlab/.local/lib/python3.12/site-packages (from huggingface-hub>=0.15.1->sentence-transformers==3.0.1) (2.32.2)  
 Requirement already satisfied: typing-extensions>=3.7.4.3 in /opt/conda/lib/python3.12/site-packages (from huggingface-hub>=0.15.1->sentence-transformers==3.0.1) (4.12.2)  
 Collecting hf-xet<2.0.0,>=1.1.0 (from huggingface-hub>=0.15.1->sentence-transformers==3.0.1)  
 Downloading hf\_xet-1.1.0-cp37-abi3-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl.metadata (494 bytes)  
 Requirement already satisfied: setuptools in /opt/conda/lib/python3.12/site-packages (from torch>=1.11.0->sentence-transformers==3.0.1) (75.8.0)  
 Collecting sympy>=1.13.3 (from torch>=1.11.0->sentence-transformers==3.0.1)  
 Downloading sympy-1.14.0-py3-none-any.whl.metadata (12 kB)  
 Collecting networkx (from torch>=1.11.0->sentence-transformers==3.0.1)  
 Downloading networkx-3.4.2-py3-none-any.whl.metadata (6.3 kB)  
 Requirement already satisfied: jinja2 in /opt/conda/lib/python3.12/site-packages (from torch>=1.11.0->sentence-transformers==3.0.1) (3.1.5)  
 Collecting nvidia-cuda-nvrtc-cu12==12.6.77 (from torch>=1.11.0->sentence-transformers==3.0.1)  
 Downloading nvidia\_cuda\_nvrtc\_cu12-12.6.77-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)  
 Collecting nvidia-cuda-runtime-cu12==12.6.77 (from torch>=1.11.0->sentence-transformers==3.0.1)  
 Downloading nvidia\_cuda\_runtime\_cu12-12.6.77-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl.metadata (1.5 kB)

Collecting nvidia-cuda-cupti-cu12==12.6.80 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_cuda\_cupti\_cu12-12.6.80-py3-none-manylinux2014\_x86\_64.whl.metadata (1.6 kB)

Collecting nvidia-cudnn-cu12==9.5.1.17 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_cudnn\_cu12-9.5.1.17-py3-none-manylinux\_2\_28\_x86\_64.whl.metadata (1.6 kB)

Collecting nvidia-cublas-cu12==12.6.4.1 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_cublas\_cu12-12.6.4.1-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)

Collecting nvidia-cufft-cu12==11.3.0.4 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_cufft\_cu12-11.3.0.4-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)

Collecting nvidia-curand-cu12==10.3.7.77 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_curand\_cu12-10.3.7.77-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)

Collecting nvidia-cusolver-cu12==11.7.1.2 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_cusolver\_cu12-11.7.1.2-py3-none-manylinux2014\_x86\_64.whl.metadata (1.6 kB)

Collecting nvidia-cusparse-cu12==12.5.4.2 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_cusparse\_cu12-12.5.4.2-py3-none-manylinux2014\_x86\_64.whl.metadata (1.6 kB)

Collecting nvidia-cusparselt-cu12==0.6.3 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_cusparselt\_cu12-0.6.3-py3-none-manylinux2014\_x86\_64.whl.metadata (6.8 kB)

Collecting nvidia-nccl-cu12==2.26.2 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_nccl\_cu12-2.26.2-py3-none-manylinux2014\_x86\_64.whl.metadata (2.0 kB)

Collecting nvidia-nvtx-cu12==12.6.77 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_nvtx\_cu12-12.6.77-py3-none-manylinux2014\_x86\_64.whl.metadata (1.6 kB)

Collecting nvidia-nvjitlink-cu12==12.6.85 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_nvjitlink\_cu12-12.6.85-py3-none-manylinux2010\_x86\_64.whl.metadata (1.5 kB)

Collecting nvidia-cufile-cu12==1.11.1.6 (from torch>=1.11.0->sentence-transformers==3.0.1)

Downloading nvidia\_cufile\_cu12-1.11.1.6-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)

Collecting triton==3.3.0 (from torch>=1.11.0->sentence-transformers==3.0.1)  
 Downloading triton-3.3.0-cp312-cp312-manylinux\_2\_27\_x86\_64.manylinux\_2\_28\_x86\_64.whl.metadata (1.5 kB)  
 Collecting regex!=2019.12.17 (from transformers<5.0.0,>=4.34.0->sentence-transformers==3.0.1)  
 Downloading regex-2024.11.6-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl.metadata (40 kB)  
 Collecting tokenizers<0.22,>=0.21 (from transformers<5.0.0,>=4.34.0->sentence-transformers==3.0.1)  
 Downloading tokenizers-0.21.1-cp39-abi3-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl.metadata (6.8 kB)  
 Collecting safetensors>=0.4.3 (from transformers<5.0.0,>=4.34.0->sentence-transformers==3.0.1)  
 Downloading safetensors-0.5.3-cp38-abi3-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl.metadata (3.8 kB)  
 Collecting joblib>=1.2.0 (from scikit-learn->sentence-transformers==3.0.1)  
 Downloading joblib-1.5.0-py3-none-any.whl.metadata (5.6 kB)  
 Collecting threadpoolctl>=3.1.0 (from scikit-learn->sentence-transformers==3.0.1)  
 Downloading threadpoolctl-3.6.0-py3-none-any.whl.metadata (13 kB)  
 Collecting mpmath<1.4,>=1.1.0 (from sympy>=1.13.3->torch>=1.11.0->sentence-transformers==3.0.1)  
 Downloading mpmath-1.3.0-py3-none-any.whl.metadata (8.6 kB)  
 Requirement already satisfied: MarkupSafe>=2.0 in /opt/conda/lib/python3.12/site-packages (from jinja2->torch>=1.11.0->sentence-transformers==3.0.1) (3.0.2)  
 Requirement already satisfied: charset-normalizer<4,>=2 in /opt/conda/lib/python3.12/site-packages (from requests->huggingface-hub>=0.15.1->sentence-transformers==3.0.1) (3.4.1)  
 Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.12/site-packages (from requests->huggingface-hub>=0.15.1->sentence-transformers==3.0.1) (3.10)  
 Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/conda/lib/python3.12/site-packages (from requests->huggingface-hub>=0.15.1->sentence-transformers==3.0.1) (2.3.0)  
 Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.12/site-packages (from requests->huggingface-hub>=0.15.1->sentence-transformers==3.0.1) (2024.12.14)  
 Downloading sentence\_transformers-3.0.1-py3-none-any.whl (227 kB)  
 Downloading huggingface\_hub-0.31.1-py3-none-any.whl (484 kB)  
 Downloading torch-2.7.0-cp312-cp312-manylinux\_2\_28\_x86\_64.whl (865.0 MB)  
865.0/865.0 MB  
 11.4 MB/s eta 0:00:0000:0100:01  
 Downloading nvidia\_cublas\_cu12-12.6.4.1-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (393.1 MB)  
393.1/393.1 MB  
 13.4 MB/s eta 0:00:0000:0100:01  
 Downloading nvidia\_cuda\_cupti\_cu12-12.6.80-py3-none-

manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (8.9 MB)  
8.9/8.9 MB  
2.1 MB/s eta 0:00:000:00:01  
Downloading nvidia\_cuda\_nvrtc\_cu12-12.6.77-py3-none-manylinux2014\_x86\_64.whl (23.7 MB)  
23.7/23.7 MB  
8.2 MB/s eta 0:00:000:00:01  
Downloading nvidia\_cuda\_runtime\_cu12-12.6.77-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (897 kB)  
897.7/897.7 kB  
64.1 MB/s eta 0:00:00  
Downloading nvidia\_cudnn\_cu12-9.5.1.17-py3-none-manylinux\_2\_28\_x86\_64.whl (571.0 MB)  
571.0/571.0 MB  
12.9 MB/s eta 0:00:0000:0100:01  
Downloading nvidia\_cufft\_cu12-11.3.0.4-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (200.2 MB)  
200.2/200.2 MB  
46.2 MB/s eta 0:00:0000:0100:01  
Downloading nvidia\_cufile\_cu12-1.11.1.6-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (1.1 MB)  
1.1/1.1 MB  
47.9 MB/s eta 0:00:00  
Downloading nvidia\_curand\_cu12-10.3.7.77-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (56.3 MB)  
56.3/56.3 MB  
48.6 MB/s eta 0:00:00:00:01  
Downloading nvidia\_cusolver\_cu12-11.7.1.2-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (158.2 MB)  
158.2/158.2 MB  
48.9 MB/s eta 0:00:0000:0100:01  
Downloading nvidia\_cusparsparse\_cu12-12.5.4.2-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (216.6 MB)  
216.6/216.6 MB  
14.2 MB/s eta 0:00:0000:0100:01  
Downloading nvidia\_cusparselt\_cu12-0.6.3-py3-none-manylinux2014\_x86\_64.whl (156.8 MB)  
156.8/156.8 MB  
32.1 MB/s eta 0:00:0000:0100:01  
Downloading nvidia\_nccl\_cu12-2.26.2-py3-none-manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (201.3 MB)  
201.3/201.3 MB  
39.0 MB/s eta 0:00:0000:0100:01  
Downloading nvidia\_nvjitlink\_cu12-12.6.85-py3-none-manylinux2010\_x86\_64.manylinux\_2\_12\_x86\_64.whl (19.7 MB)  
19.7/19.7 MB  
4.0 MB/s eta 0:00:000:00:01  
Downloading nvidia\_nvtx\_cu12-12.6.77-py3-none-



manylinux2014\_x86\_64.manylinux\_2\_17\_x86\_64.whl (89 kB)  
 Downloading  
 triton-3.3.0-cp312-cp312-manylinux\_2\_27\_x86\_64.manylinux\_2\_28\_x86\_64.whl (156.5 MB)  
 156.5/156.5 MB  
 25.5 MB/s eta 0:00:0000:0100:01  
 Downloading transformers-4.51.3-py3-none-any.whl (10.4 MB)  
 10.4/10.4 MB  
 58.9 MB/s eta 0:00:00  
 Downloading pillow-11.2.1-cp312-cp312-manylinux\_2\_28\_x86\_64.whl (4.6 MB)  
 4.6/4.6 MB  
 49.4 MB/s eta 0:00:00  
 Downloading  
 scikit\_learn-1.6.1-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (13.1 MB)  
 13.1/13.1 MB  
 49.3 MB/s eta 0:00:0000:00:01  
 Downloading  
 scipy-1.15.3-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (37.3 MB)  
 37.3/37.3 MB  
 48.7 MB/s eta 0:00:0000:00:01  
 Downloading fsspec-2025.3.2-py3-none-any.whl (194 kB)  
 Downloading  
 hf\_xet-1.1.0-cp37-abi3-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (53.6 MB)  
 53.6/53.6 MB  
 49.1 MB/s eta 0:00:0000:00:01  
 Downloading joblib-1.5.0-py3-none-any.whl (307 kB)  
 Downloading  
 regex-2024.11.6-cp312-cp312-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (796 kB)  
 796.9/796.9 kB  
 61.8 MB/s eta 0:00:00  
 Downloading  
 safetensors-0.5.3-cp38-abi3-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (471 kB)  
 Downloading sympy-1.14.0-py3-none-any.whl (6.3 MB)  
 6.3/6.3 MB  
 50.8 MB/s eta 0:00:00  
 Downloading threadpoolctl-3.6.0-py3-none-any.whl (18 kB)  
 Downloading  
 tokenizers-0.21.1-cp39-abi3-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (3.0 MB)  
 3.0/3.0 MB  
 57.9 MB/s eta 0:00:00  
 Downloading filelock-3.18.0-py3-none-any.whl (16 kB)  
 Downloading networkx-3.4.2-py3-none-any.whl (1.7 MB)  
 1.7/1.7 MB

67.0 MB/s eta 0:00:00

Downloading mpmath-1.3.0-py3-none-any.whl (536 kB)

536.2/536.2 kB

39.9 MB/s eta 0:00:00

Installing collected packages: nvidia-cusparselt-cu12, mpmath, triton, threadpoolctl, sympy, scipy, safetensors, regex, Pillow, nvidia-nvtx-cu12, nvidia-nvjitlink-cu12, nvidia-nccl-cu12, nvidia-curand-cu12, nvidia-cufile-cu12, nvidia-cuda-runtime-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-cupti-cu12, nvidia-cublas-cu12, networkx, joblib, hf-xet, fsspec, filelock, scikit-learn, nvidia-cusparse-cu12, nvidia-cufft-cu12, nvidia-cudnn-cu12, huggingface-hub, tokenizers, nvidia-cusolver-cu12, transformers, torch, sentence-transformers  
Successfully installed Pillow-11.2.1 filelock-3.18.0 fsspec-2025.3.2 hf-xet-1.1.0 huggingface-hub-0.31.1 joblib-1.5.0 mpmath-1.3.0 networkx-3.4.2 nvidia-cublas-cu12-12.6.4.1 nvidia-cuda-cupti-cu12-12.6.80 nvidia-cuda-nvrtc-cu12-12.6.77 nvidia-cuda-runtime-cu12-12.6.77 nvidia-cudnn-cu12-9.5.1.17 nvidia-cufft-cu12-11.3.0.4 nvidia-cufile-cu12-1.11.1.6 nvidia-curand-cu12-10.3.7.77 nvidia-cusolver-cu12-11.7.1.2 nvidia-cusparse-cu12-12.5.4.2 nvidia-cusparselt-cu12-0.6.3 nvidia-nccl-cu12-2.26.2 nvidia-nvjitlink-cu12-12.6.85 nvidia-nvtx-cu12-12.6.77 regex-2024.11.6 safetensors-0.5.3 scikit-learn-1.6.1 scipy-1.15.3 sentence-transformers-3.0.1 sympy-1.14.0 threadpoolctl-3.6.0 tokenizers-0.21.1 torch-2.7.0 transformers-4.51.3 triton-3.3.0

After you install the libraries, restart your kernel. You can do that by clicking the **Restart the kernel** icon.

## 1.5 Load data

A text file has been prepared as the source document for the downstream embedding task.

Now, let's download and load it using LangChain's TextLoader.

```
[1]: !pip install langchain_community
```

```
Requirement already satisfied: langchain_community in  
/home/jupyterlab/.local/lib/python3.12/site-packages (0.2.10)  
Requirement already satisfied: PyYAML>=5.3 in /opt/conda/lib/python3.12/site-  
packages (from langchain_community) (6.0.2)  
Requirement already satisfied: SQLAlchemy<3,>=1.4 in  
/opt/conda/lib/python3.12/site-packages (from langchain_community) (2.0.37)  
Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in  
/opt/conda/lib/python3.12/site-packages (from langchain_community) (3.11.18)  
Requirement already satisfied: dataclasses-json<0.7,>=0.5.7 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain_community)  
(0.6.7)  
Requirement already satisfied: langchain<0.3.0,>=0.2.9 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain_community)  
(0.2.11)  
Requirement already satisfied: langchain-core<0.3.0,>=0.2.23 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain_community)
```

(0.2.43)  
Requirement already satisfied: langsmith<0.2.0,>=0.1.0 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain\_community)  
(0.1.147)  
Requirement already satisfied: numpy<2.0.0,>=1.26.0 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain\_community)  
(1.26.4)  
Requirement already satisfied: requests<3,>=2 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain\_community)  
(2.32.2)  
Requirement already satisfied: tenacity!=8.4.0,<9.0.0,>=8.1.0 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from langchain\_community)  
(8.5.0)  
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in  
/opt/conda/lib/python3.12/site-packages (from  
aiohttp<4.0.0,>=3.8.3->langchain\_community) (2.6.1)  
Requirement already satisfied: aiosignal>=1.1.2 in  
/opt/conda/lib/python3.12/site-packages (from  
aiohttp<4.0.0,>=3.8.3->langchain\_community) (1.3.2)  
Requirement already satisfied: attrs>=17.3.0 in /opt/conda/lib/python3.12/site-  
packages (from aiohttp<4.0.0,>=3.8.3->langchain\_community) (25.1.0)  
Requirement already satisfied: frozenlist>=1.1.1 in  
/opt/conda/lib/python3.12/site-packages (from  
aiohttp<4.0.0,>=3.8.3->langchain\_community) (1.6.0)  
Requirement already satisfied: multidict<7.0,>=4.5 in  
/opt/conda/lib/python3.12/site-packages (from  
aiohttp<4.0.0,>=3.8.3->langchain\_community) (6.4.3)  
Requirement already satisfied: propcache>=0.2.0 in  
/opt/conda/lib/python3.12/site-packages (from  
aiohttp<4.0.0,>=3.8.3->langchain\_community) (0.3.1)  
Requirement already satisfied: yarl<2.0,>=1.17.0 in  
/opt/conda/lib/python3.12/site-packages (from  
aiohttp<4.0.0,>=3.8.3->langchain\_community) (1.20.0)  
Requirement already satisfied: marshmallow<4.0.0,>=3.18.0 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from dataclasses-  
json<0.7,>=0.5.7->langchain\_community) (3.26.1)  
Requirement already satisfied: typing-inspect<1,>=0.4.0 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from dataclasses-  
json<0.7,>=0.5.7->langchain\_community) (0.9.0)  
Requirement already satisfied: langchain-text-splitters<0.3.0,>=0.2.0 in  
/home/jupyterlab/.local/lib/python3.12/site-packages (from  
langchain<0.3.0,>=0.2.9->langchain\_community) (0.2.4)  
Requirement already satisfied: pydantic<3,>=1 in /opt/conda/lib/python3.12/site-  
packages (from langchain<0.3.0,>=0.2.9->langchain\_community) (2.10.6)  
Requirement already satisfied: jsonpatch<2.0,>=1.33 in  
/opt/conda/lib/python3.12/site-packages (from langchain-  
core<0.3.0,>=0.2.23->langchain\_community) (1.33)  
Requirement already satisfied: packaging<25,>=23.2 in

/opt/conda/lib/python3.12/site-packages (from langchain-core<0.3.0,>=0.2.23->langchain\_community) (24.2)  
 Requirement already satisfied: typing-extensions>=4.7 in /opt/conda/lib/python3.12/site-packages (from langchain-core<0.3.0,>=0.2.23->langchain\_community) (4.12.2)  
 Requirement already satisfied: httpx<1,>=0.23.0 in /opt/conda/lib/python3.12/site-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 /home/jupyterlab/.local/lib/python3.12/site-packages (from langsmith<0.2.0,>=0.1.0->langchain\_community) (3.10.18)  
 Requirement already satisfied: requests-toolbelt<2.0.0,>=1.0.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from langsmith<0.2.0,>=0.1.0->langchain\_community) (1.0.0)  
 Requirement already satisfied: charset-normalizer<4,>=2 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain\_community) (3.4.1)  
 Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain\_community) (3.10)  
 Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain\_community) (2.3.0)  
 Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.12/site-packages (from requests<3,>=2->langchain\_community) (2024.12.14)  
 Requirement already satisfied: greenlet!=0.4.17 in /opt/conda/lib/python3.12/site-packages (from SQLAlchemy<3,>=1.4->langchain\_community) (3.1.1)  
 Requirement already satisfied: anyio in /opt/conda/lib/python3.12/site-packages (from httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain\_community) (4.8.0)  
 Requirement already satisfied: httpcore==1.\* in /opt/conda/lib/python3.12/site-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 /opt/conda/lib/python3.12/site-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 /opt/conda/lib/python3.12/site-packages (from jsonpatch<2.0,>=1.33->langchain-core<0.3.0,>=0.2.23->langchain\_community) (3.0.0)  
 Requirement already satisfied: annotated-types>=0.6.0 in /opt/conda/lib/python3.12/site-packages (from pydantic<3,>=1->langchain<0.3.0,>=0.2.9->langchain\_community) (0.7.0)  
 Requirement already satisfied: pydantic-core==2.27.2 in /opt/conda/lib/python3.12/site-packages (from pydantic<3,>=1->langchain<0.3.0,>=0.2.9->langchain\_community) (2.27.2)  
 Requirement already satisfied: mypy-extensions>=0.3.0 in /home/jupyterlab/.local/lib/python3.12/site-packages (from typing-

```
inspect<1,>=0.4.0->dataclasses-json<0.7,>=0.5.7->langchain_community) (1.1.0)
Requirement already satisfied: sniffio>=1.1 in /opt/conda/lib/python3.12/site-
packages (from
anyio->httpx<1,>=0.23.0->langsmith<0.2.0,>=0.1.0->langchain_community) (1.3.1)
```

```
[2]: !wget "https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/
↪i5V3ACEyz6hnYpVq6MTSvg/state-of-the-union.txt"
```

```
--2025-05-11 17:16:35-- https://cf-courses-data.s3.us.cloud-object-
storage.appdomain.cloud/i5V3ACEyz6hnYpVq6MTSvg/state-of-the-union.txt
Resolving cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud (cf-
courses-data.s3.us.cloud-object-storage.appdomain.cloud)... 169.63.118.104
Connecting to cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud (cf-
courses-data.s3.us.cloud-object-storage.appdomain.cloud)|169.63.118.104|:443...
connected.
```

```
200 OKrequest sent, awaiting response...
```

```
Length: 39027 (38K) [text/plain]
```

```
Saving to: 'state-of-the-union.txt.1'
```

```
state-of-the-union. 100%[=====>] 38.11K --.-KB/s in 0s
```

```
2025-05-11 17:16:36 (84.6 MB/s) - 'state-of-the-union.txt.1' saved [39027/39027]
```

```
[3]: from langchain_community.document_loaders import TextLoader
```

```
[4]: loader = TextLoader("state-of-the-union.txt")
data = loader.load()
```

Let's take a look at the document.

```
[5]: data
```

```
[5]: [Document(metadata={'source': 'state-of-the-union.txt'}, page_content='Madam
Speaker, Madam Vice President, our First Lady and Second Gentleman. Members of
Congress and the Cabinet. Justices of the Supreme Court. My fellow Americans.
\n\nLast year COVID-19 kept us apart. This year we are finally together again.
\n\nTonight, we meet as Democrats Republicans and Independents. But most
importantly as Americans. \n\nWith a duty to one another to the American people
to the Constitution. \n\nAnd with an unwavering resolve that freedom will always
triumph over tyranny. \n\nSix days ago, Russia's Vladimir Putin sought to shake
the foundations of the free world thinking he could make it bend to his menacing
ways. But he badly miscalculated. \n\nHe thought he could roll into Ukraine and
the world would roll over. Instead he met a wall of strength he never imagined.
\n\nHe met the Ukrainian people. \n\nFrom President Zelenskyy to every
Ukrainian, their fearlessness, their courage, their determination, inspires the
world. \n\nGroups of citizens blocking tanks with their bodies. Everyone from
students to retirees teachers turned soldiers defending their homeland. \n\nIn
this struggle as President Zelenskyy said in his speech to the European
```

Parliament "Light will win over darkness." The Ukrainian Ambassador to the United States is here tonight. \n\nLet each of us here tonight in this Chamber send an unmistakable signal to Ukraine and to the world. \n\nPlease rise if you are able and show that, Yes, we the United States of America stand with the Ukrainian people. \n\nThroughout our history we've learned this lesson when dictators do not pay a price for their aggression they cause more chaos. \n\nThey keep moving. \n\nAnd the costs and the threats to America and the world keep rising. \n\nThat's why the NATO Alliance was created to secure peace and stability in Europe after World War 2. \n\nThe United States is a member along with 29 other nations. \n\nIt matters. American diplomacy matters. American resolve matters. \n\nPutin's latest attack on Ukraine was premeditated and unprovoked. \n\nHe rejected repeated efforts at diplomacy. \n\nHe thought the West and NATO wouldn't respond. And he thought he could divide us at home. Putin was wrong. We were ready. Here is what we did. \n\nWe prepared extensively and carefully. \n\nWe spent months building a coalition of other freedom-loving nations from Europe and the Americas to Asia and Africa to confront Putin. \n\nI spent countless hours unifying our European allies. We shared with the world in advance what we knew Putin was planning and precisely how he would try to falsely justify his aggression. \n\nWe countered Russia's lies with truth. \n\nAnd now that he has acted the free world is holding him accountable. \n\nAlong with twenty-seven members of the European Union including France, Germany, Italy, as well as countries like the United Kingdom, Canada, Japan, Korea, Australia, New Zealand, and many others, even Switzerland. \n\nWe are inflicting pain on Russia and supporting the people of Ukraine. Putin is now isolated from the world more than ever. \n\nTogether with our allies -we are right now enforcing powerful economic sanctions. \n\nWe are cutting off Russia's largest banks from the international financial system. \n\nPreventing Russia's central bank from defending the Russian Ruble making Putin's \$630 Billion "war fund" worthless. \n\nWe are choking off Russia's access to technology that will sap its economic strength and weaken its military for years to come. \n\nTonight I say to the Russian oligarchs and corrupt leaders who have bilked billions of dollars off this violent regime no more. \n\nThe U.S. Department of Justice is assembling a dedicated task force to go after the crimes of Russian oligarchs. \n\nWe are joining with our European allies to find and seize your yachts your luxury apartments your private jets. We are coming for your ill-begotten gains. \n\nAnd tonight I am announcing that we will join our allies in closing off American air space to all Russian flights - further isolating Russia - and adding an additional squeeze -on their economy. The Ruble has lost 30% of its value. \n\nThe Russian stock market has lost 40% of its value and trading remains suspended. Russia's economy is reeling and Putin alone is to blame. \n\nTogether with our allies we are providing support to the Ukrainians in their fight for freedom. Military assistance. Economic assistance. Humanitarian assistance. \n\nWe are giving more than \$1 Billion in direct assistance to Ukraine. \n\nAnd we will continue to aid the Ukrainian people as they defend their country and to help ease their suffering. \n\nLet me be clear, our forces are not engaged and will not engage in conflict with Russian forces in Ukraine. \n\nOur forces are not going to Europe to fight in Ukraine, but to defend our

NATO Allies - in the event that Putin decides to keep moving west. \n\nFor that purpose we've mobilized American ground forces, air squadrons, and ship deployments to protect NATO countries including Poland, Romania, Latvia, Lithuania, and Estonia. \n\nAs I have made crystal clear the United States and our Allies will defend every inch of territory of NATO countries with the full force of our collective power. \n\nAnd we remain clear-eyed. The Ukrainians are fighting back with pure courage. But the next few days weeks, months, will be hard on them. \n\nPutin has unleashed violence and chaos. But while he may make gains on the battlefield - he will pay a continuing high price over the long run. \n\nAnd a proud Ukrainian people, who have known 30 years of independence, have repeatedly shown that they will not tolerate anyone who tries to take their country backwards. \n\nTo all Americans, I will be honest with you, as I've always promised. A Russian dictator, invading a foreign country, has costs around the world. \n\nAnd I'm taking robust action to make sure the pain of our sanctions is targeted at Russia's economy. And I will use every tool at our disposal to protect American businesses and consumers. \n\nTonight, I can announce that the United States has worked with 30 other countries to release 60 Million barrels of oil from reserves around the world. \n\nAmerica will lead that effort, releasing 30 Million barrels from our own Strategic Petroleum Reserve. And we stand ready to do more if necessary, unified with our allies. \n\nThese steps will help blunt gas prices here at home. And I know the news about what's happening can seem alarming. \n\nBut I want you to know that we are going to be okay. \n\nWhen the history of this era is written Putin's war on Ukraine will have left Russia weaker and the rest of the world stronger. \n\nWhile it shouldn't have taken something so terrible for people around the world to see what's at stake now everyone sees it clearly. \n\nWe see the unity among leaders of nations and a more unified Europe a more unified West. And we see unity among the people who are gathering in cities in large crowds around the world even in Russia to demonstrate their support for Ukraine. \n\nIn the battle between democracy and autocracy, democracies are rising to the moment, and the world is clearly choosing the side of peace and security. \n\nThis is a real test. It's going to take time. So let us continue to draw inspiration from the iron will of the Ukrainian people. \n\nTo our fellow Ukrainian Americans who forge a deep bond that connects our two nations we stand with you. \n\nPutin may circle Kyiv with tanks, but he will never gain the hearts and souls of the Ukrainian people. \n\nHe will never extinguish their love of freedom. He will never weaken the resolve of the free world. \n\nWe meet tonight in an America that has lived through two of the hardest years this nation has ever faced. \n\nThe pandemic has been punishing. \n\nAnd so many families are living paycheck to paycheck, struggling to keep up with the rising cost of food, gas, housing, and so much more. \n\nI understand. \n\nI remember when my Dad had to leave our home in Scranton, Pennsylvania to find work. I grew up in a family where if the price of food went up, you felt it. \n\nThat's why one of the first things I did as President was fight to pass the American Rescue Plan. \n\nBecause people were hurting. We needed to act, and we did. \n\nFew pieces of legislation have done more in a critical moment in our history to lift us out of crisis. \n\nIt fueled our efforts to vaccinate the nation and combat COVID-19.

It delivered immediate economic relief for tens of millions of Americans.  
Helped put food on their table, keep a roof over their heads, and cut the cost of health insurance. And as my Dad used to say, it gave people a little breathing room. And unlike the \$2 Trillion tax cut passed in the previous administration that benefitted the top 1% of Americans, the American Rescue Plan helped working people-and left no one behind. And it worked. It created jobs. Lots of jobs. In fact-our economy created over 6.5 Million new jobs just last year, more jobs created in one year than ever before in the history of America. Our economy grew at a rate of 5.7% last year, the strongest growth in nearly 40 years, the first step in bringing fundamental change to an economy that hasn't worked for the working people of this nation for too long. For the past 40 years we were told that if we gave tax breaks to those at the very top, the benefits would trickle down to everyone else. But that trickle-down theory led to weaker economic growth, lower wages, bigger deficits, and the widest gap between those at the top and everyone else in nearly a century. Vice President Harris and I ran for office with a new economic vision for America. Invest in America. Educate Americans. Grow the workforce. Build the economy from the bottom up and the middle out, not from the top down. Because we know that when the middle class grows, the poor have a ladder up and the wealthy do very well. America used to have the best roads, bridges, and airports on Earth. Now our infrastructure is ranked 13th in the world. We won't be able to compete for the jobs of the 21st Century if we don't fix that. That's why it was so important to pass the Bipartisan Infrastructure Law-the most sweeping investment to rebuild America in history. This was a bipartisan effort, and I want to thank the members of both parties who worked to make it happen. We're done talking about infrastructure weeks. We're going to have an infrastructure decade. It is going to transform America and put us on a path to win the economic competition of the 21st Century that we face with the rest of the world-particularly with China. As I've told Xi Jinping, it is never a good bet to bet against the American people. We'll create good jobs for millions of Americans, modernizing roads, airports, ports, and waterways all across America. And we'll do it all to withstand the devastating effects of the climate crisis and promote environmental justice. We'll build a national network of 500,000 electric vehicle charging stations, begin to replace poisonous lead pipes-so every child-and every American-has clean water to drink at home and at school, provide affordable high-speed internet for every American-urban, suburban, rural, and tribal communities. 4,000 projects have already been announced. And tonight, I'm announcing that this year we will start fixing over 65,000 miles of highway and 1,500 bridges in disrepair. When we use taxpayer dollars to rebuild America - we are going to Buy American: buy American products to support American jobs. The federal government spends about \$600 Billion a year to keep the country safe and secure. There's been a law on the books for almost a century to make sure taxpayers' dollars support American jobs and businesses. Every Administration says they'll do it, but we are actually doing it. We will buy American to make sure everything from the deck of an aircraft carrier to the



steel on highway guardrails are made in America. \n\nBut to compete for the best jobs of the future, we also need to level the playing field with China and other competitors. \n\nThat's why it is so important to pass the Bipartisan Innovation Act sitting in Congress that will make record investments in emerging technologies and American manufacturing. \n\nLet me give you one example of why it's so important to pass it. \n\nIf you travel 20 miles east of Columbus, Ohio, you'll find 1,000 empty acres of land. \n\nIt won't look like much, but if you stop and look closely, you'll see a "Field of dreams," the ground on which America's future will be built. \n\nThis is where Intel, the American company that helped build Silicon Valley, is going to build its \$20 billion semiconductor "mega site". \n\nUp to eight state-of-the-art factories in one place. 10,000 new good-paying jobs. \n\nSome of the most sophisticated manufacturing in the world to make computer chips the size of a fingertip that power the world and our everyday lives. \n\nSmartphones. The Internet. Technology we have yet to invent. \n\nBut that's just the beginning. \n\nIntel's CEO, Pat Gelsinger, who is here tonight, told me they are ready to increase their investment from \$20 billion to \$100 billion. \n\nThat would be one of the biggest investments in manufacturing in American history. \n\nAnd all they're waiting for is for you to pass this bill. \n\nSo let's not wait any longer. Send it to my desk. I'll sign it. \n\nAnd we will really take off. \n\nAnd Intel is not alone. \n\nThere's something happening in America. \n\nJust look around and you'll see an amazing story. \n\nThe rebirth of the pride that comes from stamping products "Made In America." The revitalization of American manufacturing. \n\nCompanies are choosing to build new factories here, when just a few years ago, they would have built them overseas. \n\nThat's what is happening. Ford is investing \$11 billion to build electric vehicles, creating 11,000 jobs across the country. \n\nGM is making the largest investment in its history-\$7 billion to build electric vehicles, creating 4,000 jobs in Michigan. \n\nAll told, we created 369,000 new manufacturing jobs in America just last year. \n\nPowered by people I've met like JoJo Burgess, from generations of union steelworkers from Pittsburgh, who's here with us tonight. \n\nAs Ohio Senator Sherrod Brown says, "It's time to bury the label "Rust Belt." \n\nIt's time. \n\nBut with all the bright spots in our economy, record job growth and higher wages, too many families are struggling to keep up with the bills. \n\nInflation is robbing them of the gains they might otherwise feel. \n\nI get it. That's why my top priority is getting prices under control. \n\nLook, our economy roared back faster than most predicted, but the pandemic meant that businesses had a hard time hiring enough workers to keep up production in their factories. \n\nThe pandemic also disrupted global supply chains. \n\nWhen factories close, it takes longer to make goods and get them from the warehouse to the store, and prices go up. \n\nLook at cars. \n\nLast year, there weren't enough semiconductors to make all the cars that people wanted to buy. \n\nAnd guess what, prices of automobiles went up. \n\nSo-we have a choice. \n\nOne way to fight inflation is to drive down wages and make Americans poorer. \n\nI have a better plan to fight inflation. \n\nLower your costs, not your wages. \n\nMake more cars and semiconductors in America. \n\nMore infrastructure and innovation in America. \n\nMore goods moving faster and cheaper in America. \n\nMore jobs

where you can earn a good living in America. \n\nAnd instead of relying on foreign supply chains, let's make it in America. \n\nEconomists call it "increasing the productive capacity of our economy." \n\nI call it building a better America. \n\nMy plan to fight inflation will lower your costs and lower the deficit. \n\n17 Nobel laureates in economics say my plan will ease long-term inflationary pressures. Top business leaders and most Americans support my plan. And here's the plan: \n\nFirst - cut the cost of prescription drugs. Just look at insulin. One in ten Americans has diabetes. In Virginia, I met a 13-year-old boy named Joshua Davis. \n\nHe and his Dad both have Type 1 diabetes, which means they need insulin every day. Insulin costs about \$10 a vial to make. \n\nBut drug companies charge families like Joshua and his Dad up to 30 times more. I spoke with Joshua's mom. \n\nImagine what it's like to look at your child who needs insulin and have no idea how you're going to pay for it. \n\nWhat it does to your dignity, your ability to look your child in the eye, to be the parent you expect to be. \n\nJoshua is here with us tonight. Yesterday was his birthday. Happy birthday, buddy. \n\nFor Joshua, and for the 200,000 other young people with Type 1 diabetes, let's cap the cost of insulin at \$35 a month so everyone can afford it. \n\nDrug companies will still do very well. And while we're at it let Medicare negotiate lower prices for prescription drugs, like the VA already does. \n\nLook, the American Rescue Plan is helping millions of families on Affordable Care Act plans save \$2,400 a year on their health care premiums. Let's close the coverage gap and make those savings permanent. \n\nSecond - cut energy costs for families an average of \$500 a year by combatting climate change. \n\nLet's provide investments and tax credits to weatherize your homes and businesses to be energy efficient and you get a tax credit; double America's clean energy production in solar, wind, and so much more; lower the price of electric vehicles, saving you another \$80 a month because you'll never have to pay at the gas pump again. \n\nThird - cut the cost of child care. Many families pay up to \$14,000 a year for child care per child. \n\nMiddle-class and working families shouldn't have to pay more than 7% of their income for care of young children. \n\nMy plan will cut the cost in half for most families and help parents, including millions of women, who left the workforce during the pandemic because they couldn't afford child care, to be able to get back to work. \n\nMy plan doesn't stop there. It also includes home and long-term care. More affordable housing. And Pre-K for every 3- and 4-year-old. \n\nAll of these will lower costs. \n\nAnd under my plan, nobody earning less than \$400,000 a year will pay an additional penny in new taxes. Nobody. \n\nThe one thing all Americans agree on is that the tax system is not fair. We have to fix it. \n\nI'm not looking to punish anyone. But let's make sure corporations and the wealthiest Americans start paying their fair share. \n\nJust last year, 55 Fortune 500 corporations earned \$40 billion in profits and paid zero dollars in federal income tax. \n\nThat's simply not fair. That's why I've proposed a 15% minimum tax rate for corporations. \n\nWe got more than 130 countries to agree on a global minimum tax rate so companies can't get out of paying their taxes at home by shipping jobs and factories overseas. \n\nThat's why I've proposed closing loopholes so the very wealthy don't pay a lower tax rate than a teacher or a firefighter. \n\nSo that's my plan. It will

grow the economy and lower costs for families. \n\nSo what are we waiting for? Let's get this done. And while you're at it, confirm my nominees to the Federal Reserve, which plays a critical role in fighting inflation. \n\nMy plan will not only lower costs to give families a fair shot, it will lower the deficit. \n\nThe previous Administration not only ballooned the deficit with tax cuts for the very wealthy and corporations, it undermined the watchdogs whose job was to keep pandemic relief funds from being wasted. \n\nBut in my administration, the watchdogs have been welcomed back. \n\nWe're going after the criminals who stole billions in relief money meant for small businesses and millions of Americans. \n\nAnd tonight, I'm announcing that the Justice Department will name a chief prosecutor for pandemic fraud. \n\nBy the end of this year, the deficit will be down to less than half what it was before I took office. \n\nThe only president ever to cut the deficit by more than one trillion dollars in a single year. \n\nLowering your costs also means demanding more competition. \n\nI'm a capitalist, but capitalism without competition isn't capitalism. \n\nIt's exploitation-and it drives up prices. \n\nWhen corporations don't have to compete, their profits go up, your prices go up, and small businesses and family farmers and ranchers go under. \n\nWe see it happening with ocean carriers moving goods in and out of America. \n\nDuring the pandemic, these foreign-owned companies raised prices by as much as 1,000% and made record profits. \n\nTonight, I'm announcing a crackdown on these companies overcharging American businesses and consumers. \n\nAnd as Wall Street firms take over more nursing homes, quality in those homes has gone down and costs have gone up. \n\nThat ends on my watch. \n\nMedicare is going to set higher standards for nursing homes and make sure your loved ones get the care they deserve and expect. \n\nWe'll also cut costs and keep the economy going strong by giving workers a fair shot, provide more training and apprenticeships, hire them based on their skills not degrees. \n\nLet's pass the Paycheck Fairness Act and paid leave. \n\nRaise the minimum wage to \$15 an hour and extend the Child Tax Credit, so no one has to raise a family in poverty. \n\nLet's increase Pell Grants and increase our historic support of HBCUs, and invest in what Jill-our First Lady who teaches full-time-calls America's best-kept secret: community colleges. \n\nAnd let's pass the PRO Act when a majority of workers want to form a union-they shouldn't be stopped. \n\nWhen we invest in our workers, when we build the economy from the bottom up and the middle out together, we can do something we haven't done in a long time: build a better America. \n\nFor more than two years, COVID-19 has impacted every decision in our lives and the life of the nation. \n\nAnd I know you're tired, frustrated, and exhausted. \n\nBut I also know this. \n\nBecause of the progress we've made, because of your resilience and the tools we have, tonight I can say \n\nwe are moving forward safely, back to more normal routines. \n\nWe've reached a new moment in the fight against COVID-19, with severe cases down to a level not seen since last July. \n\nJust a few days ago, the Centers for Disease Control and Prevention-the CDC-issued new mask guidelines. \n\nUnder these new guidelines, most Americans in most of the country can now be mask free. \n\nAnd based on the projections, more of the country will reach that point across the next couple of weeks. \n\nThanks to the progress we have made this past year,

COVID-19 need no longer control our lives. \n\nI know some are talking about "living with COVID-19". Tonight - I say that we will never just accept living with COVID-19. \n\nWe will continue to combat the virus as we do other diseases. And because this is a virus that mutates and spreads, we will stay on guard. \n\nHere are four common sense steps as we move forward safely. \n\nFirst, stay protected with vaccines and treatments. We know how incredibly effective vaccines are. If you're vaccinated and boosted you have the highest degree of protection. \n\nWe will never give up on vaccinating more Americans. Now, I know parents with kids under 5 are eager to see a vaccine authorized for their children. \n\nThe scientists are working hard to get that done and we'll be ready with plenty of vaccines when they do. \n\nWe're also ready with anti-viral treatments. If you get COVID-19, the Pfizer pill reduces your chances of ending up in the hospital by 90%. \n\nWe've ordered more of these pills than anyone in the world. And Pfizer is working overtime to get us 1 Million pills this month and more than double that next month. \n\nAnd we're launching the "Test to Treat" initiative so people can get tested at a pharmacy, and if they're positive, receive antiviral pills on the spot at no cost. \n\nIf you're immunocompromised or have some other vulnerability, we have treatments and free high-quality masks. \n\nWe're leaving no one behind or ignoring anyone's needs as we move forward. \n\nAnd on testing, we have made hundreds of millions of tests available for you to order for free. \n\nEven if you already ordered free tests tonight, I am announcing that you can order more from covidtests.gov starting next week. \n\nSecond - we must prepare for new variants. Over the past year, we've gotten much better at detecting new variants. \n\nIf necessary, we'll be able to deploy new vaccines within 100 days instead of many more months or years. \n\nAnd, if Congress provides the funds we need, we'll have new stockpiles of tests, masks, and pills ready if needed. \n\nI cannot promise a new variant won't come. But I can promise you we'll do everything within our power to be ready if it does. \n\nThird - we can end the shutdown of schools and businesses. We have the tools we need. \n\nIt's time for Americans to get back to work and fill our great downtowns again. People working from home can feel safe to begin to return to the office. \n\nWe're doing that here in the federal government. The vast majority of federal workers will once again work in person. \n\nOur schools are open. Let's keep it that way. Our kids need to be in school. \n\nAnd with 75% of adult Americans fully vaccinated and hospitalizations down by 77%, most Americans can remove their masks, return to work, stay in the classroom, and move forward safely. \n\nWe achieved this because we provided free vaccines, treatments, tests, and masks. \n\nOf course, continuing this costs money. \n\nI will soon send Congress a request. \n\nThe vast majority of Americans have used these tools and may want to again, so I expect Congress to pass it quickly. \n\nFourth, we will continue vaccinating the world. \n\nWe've sent 475 Million vaccine doses to 112 countries, more than any other nation. \n\nAnd we won't stop. \n\nWe have lost so much to COVID-19. Time with one another. And worst of all, so much loss of life. \n\nLet's use this moment to reset. Let's stop looking at COVID-19 as a partisan dividing line and see it for what it is: A God-awful disease. \n\nLet's stop seeing each other as enemies, and start seeing each other for who we really are:

Fellow Americans. \n\nWe can't change how divided we've been. But we can change how we move forward-on COVID-19 and other issues we must face together. \n\nI recently visited the New York City Police Department days after the funerals of Officer Wilbert Mora and his partner, Officer Jason Rivera. \n\nThey were responding to a 9-1-1 call when a man shot and killed them with a stolen gun. \n\nOfficer Mora was 27 years old. \n\nOfficer Rivera was 22. \n\nBoth Dominican Americans who'd grown up on the same streets they later chose to patrol as police officers. \n\nI spoke with their families and told them that we are forever in debt for their sacrifice, and we will carry on their mission to restore the trust and safety every community deserves. \n\nI've worked on these issues a long time. \n\nI know what works: Investing in crime prevention and community police officers who'll walk the beat, who'll know the neighborhood, and who can restore trust and safety. \n\nSo let's not abandon our streets. Or choose between safety and equal justice. \n\nLet's come together to protect our communities, restore trust, and hold law enforcement accountable. \n\nThat's why the Justice Department required body cameras, banned chokeholds, and restricted no-knock warrants for its officers. \n\nThat's why the American Rescue Plan provided \$350 Billion that cities, states, and counties can use to hire more police and invest in proven strategies like community violence interruption-trusted messengers breaking the cycle of violence and trauma and giving young people hope. \n\nWe should all agree: The answer is not to Defund the police. The answer is to FUND the police with the resources and training they need to protect our communities. \n\nI ask Democrats and Republicans alike: Pass my budget and keep our neighborhoods safe. \n\nAnd I will keep doing everything in my power to crack down on gun trafficking and ghost guns you can buy online and make at home-they have no serial numbers and can't be traced. \n\nAnd I ask Congress to pass proven measures to reduce gun violence. Pass universal background checks. Why should anyone on a terrorist list be able to purchase a weapon? \n\nBan assault weapons and high-capacity magazines. \n\nRepeal the liability shield that makes gun manufacturers the only industry in America that can't be sued. \n\nThese laws don't infringe on the Second Amendment. They save lives. \n\nThe most fundamental right in America is the right to vote - and to have it counted. And it's under assault. \n\nIn state after state, new laws have been passed, not only to suppress the vote, but to subvert entire elections. \n\nWe cannot let this happen. \n\nTonight. I call on the Senate to: Pass the Freedom to Vote Act. Pass the John Lewis Voting Rights Act. And while you're at it, pass the Disclose Act so Americans can know who is funding our elections. \n\nTonight, I'd like to honor someone who has dedicated his life to serve this country: Justice Stephen Breyer-an Army veteran, Constitutional scholar, and retiring Justice of the United States Supreme Court. Justice Breyer, thank you for your service. \n\nOne of the most serious constitutional responsibilities a President has is nominating someone to serve on the United States Supreme Court. \n\nAnd I did that 4 days ago, when I nominated Circuit Court of Appeals Judge Ketanji Brown Jackson. One of our nation's top legal minds, who will continue Justice Breyer's legacy of excellence. \n\nA former top litigator in private practice. A former federal public defender. And from a family of public school educators and police

officers. A consensus builder. Since she's been nominated, she's received a broad range of support-from the Fraternal Order of Police to former judges appointed by Democrats and Republicans. \n\nAnd if we are to advance liberty and justice, we need to secure the Border and fix the immigration system. \n\nWe can do both. At our border, we've installed new technology like cutting-edge scanners to better detect drug smuggling. \n\nWe've set up joint patrols with Mexico and Guatemala to catch more human traffickers. \n\nWe're putting in place dedicated immigration judges so families fleeing persecution and violence can have their cases heard faster. \n\nWe're securing commitments and supporting partners in South and Central America to host more refugees and secure their own borders. \n\nWe can do all this while keeping lit the torch of liberty that has led generations of immigrants to this land-my forefathers and so many of yours. \n\nProvide a pathway to citizenship for Dreamers, those on temporary status, farm workers, and essential workers. \n\nRevise our laws so businesses have the workers they need and families don't wait decades to reunite. \n\nIt's not only the right thing to do-it's the economically smart thing to do. \n\nThat's why immigration reform is supported by everyone from labor unions to religious leaders to the U.S. Chamber of Commerce. \n\nLet's get it done once and for all. \n\nAdvancing liberty and justice also requires protecting the rights of women. \n\nThe constitutional right affirmed in Roe v. Wade-standing precedent for half a century-is under attack as never before. \n\nIf we want to go forward-not backward-we must protect access to health care. Preserve a woman's right to choose. And let's continue to advance maternal health care in America. \n\nAnd for our LGBTQ+ Americans, let's finally get the bipartisan Equality Act to my desk. The onslaught of state laws targeting transgender Americans and their families is wrong. \n\nAs I said last year, especially to our younger transgender Americans, I will always have your back as your President, so you can be yourself and reach your God-given potential. \n\nWhile it often appears that we never agree, that isn't true. I signed 80 bipartisan bills into law last year. From preventing government shutdowns to protecting Asian-Americans from still-too-common hate crimes to reforming military justice. \n\nAnd soon, we'll strengthen the Violence Against Women Act that I first wrote three decades ago. It is important for us to show the nation that we can come together and do big things. \n\nSo tonight I'm offering a Unity Agenda for the Nation. Four big things we can do together. \n\nFirst, beat the opioid epidemic. \n\nThere is so much we can do. Increase funding for prevention, treatment, harm reduction, and recovery. \n\nGet rid of outdated rules that stop doctors from prescribing treatments. And stop the flow of illicit drugs by working with state and local law enforcement to go after traffickers. \n\nIf you're suffering from addiction, know you are not alone. I believe in recovery, and I celebrate the 23 million Americans in recovery. \n\nSecond, let's take on mental health. Especially among our children, whose lives and education have been turned upside down. \n\nThe American Rescue Plan gave schools money to hire teachers and help students make up for lost learning. \n\nI urge every parent to make sure your school does just that. And we can all play a part-sign up to be a tutor or a mentor. \n\nChildren were also struggling before the pandemic. Bullying, violence, trauma, and the harms of social media. \n\nAs Frances Haugen, who is here with

us tonight, has shown, we must hold social media platforms accountable for the national experiment they're conducting on our children for profit. \n\nIt's time to strengthen privacy protections, ban targeted advertising to children, demand tech companies stop collecting personal data on our children. \n\nAnd let's get all Americans the mental health services they need. More people they can turn to for help, and full parity between physical and mental health care. \n\nThird, support our veterans. \n\nVeterans are the best of us. \n\nI've always believed that we have a sacred obligation to equip all those we send to war and care for them and their families when they come home. \n\nMy administration is providing assistance with job training and housing, and now helping lower-income veterans get VA care debt-free. \n\nOur troops in Iraq and Afghanistan faced many dangers. \n\nOne was stationed at bases and breathing in toxic smoke from "burn pits" that incinerated wastes of war-medical and hazard material, jet fuel, and more. \n\nWhen they came home, many of the world's fittest and best trained warriors were never the same. \n\nHeadaches. Numbness. Dizziness. \n\nA cancer that would put them in a flag-draped coffin. \n\nI know. \n\nOne of those soldiers was my son Major Beau Biden. \n\nWe don't know for sure if a burn pit was the cause of his brain cancer, or the diseases of so many of our troops. \n\nBut I'm committed to finding out everything we can. \n\nCommitted to military families like Danielle Robinson from Ohio. \n\nThe widow of Sergeant First Class Heath Robinson. \n\nHe was born a soldier. Army National Guard. Combat medic in Kosovo and Iraq. \n\nStationed near Baghdad, just yards from burn pits the size of football fields. \n\nHeath's widow Danielle is here with us tonight. They loved going to Ohio State football games. He loved building Legos with their daughter. \n\nBut cancer from prolonged exposure to burn pits ravaged Heath's lungs and body. \n\nDanielle says Heath was a fighter to the very end. \n\nHe didn't know how to stop fighting, and neither did she. \n\nThrough her pain she found purpose to demand we do better. \n\nTonight, Danielle-we are. \n\nThe VA is pioneering new ways of linking toxic exposures to diseases, already helping more veterans get benefits. \n\nAnd tonight, I'm announcing we're expanding eligibility to veterans suffering from nine respiratory cancers. \n\nI'm also calling on Congress: pass a law to make sure veterans devastated by toxic exposures in Iraq and Afghanistan finally get the benefits and comprehensive health care they deserve. \n\nAnd fourth, let's end cancer as we know it. \n\nThis is personal to me and Jill, to Kamala, and to so many of you. \n\nCancer is the #2 cause of death in America-second only to heart disease. \n\nLast month, I announced our plan to supercharge \nthe Cancer Moonshot that President Obama asked me to lead six years ago. \n\nOur goal is to cut the cancer death rate by at least 50% over the next 25 years, turn more cancers from death sentences into treatable diseases. \n\nMore support for patients and families. \n\nTo get there, I call on Congress to fund ARPA-H, the Advanced Research Projects Agency for Health. \n\nIt's based on DARPA-the Defense Department project that led to the Internet, GPS, and so much more. \n\nARPA-H will have a singular purpose-to drive breakthroughs in cancer, Alzheimer's, diabetes, and more. \n\nA unity agenda for the nation. \n\nWe can do this. \n\nMy fellow Americans-tonight , we have gathered in a sacred space-the citadel of our democracy. \n\nIn this Capitol, generation after

generation, Americans have debated great questions amid great strife, and have done great things. \n\nWe have fought for freedom, expanded liberty, defeated totalitarianism and terror. \n\nAnd built the strongest, freest, and most prosperous nation the world has ever known. \n\nNow is the hour. \n\nOur moment of responsibility. \n\nOur test of resolve and conscience, of history itself. \n\nIt is in this moment that our character is formed. Our purpose is found. Our future is forged. \n\nWell I know this nation. \n\nWe will meet the test. \n\nTo protect freedom and liberty, to expand fairness and opportunity. \n\nWe will save democracy. \n\nAs hard as these times have been, I am more optimistic about America today than I have been my whole life. \n\nBecause I see the future that is within our grasp. \n\nBecause I know there is simply nothing beyond our capacity. \n\nWe are the only nation on Earth that has always turned every crisis we have faced into an opportunity. \n\nThe only nation that can be defined by a single word: possibilities. \n\nSo on this night, in our 245th year as a nation, I have come to report on the State of the Union. \n\nAnd my report is this: the State of the Union is strong-because you, the American people, are strong. \n\nWe are stronger today than we were a year ago. \n\nAnd we will be stronger a year from now than we are today. \n\nNow is our moment to meet and overcome the challenges of our time. \n\nAnd we will, as one people. \n\nOne America. \n\nThe United States of America. \n\nMay God bless you all. May God protect our troops.'])]

## 1.6 Split data

Since the embedding model has a maximum input token limit, you cannot input the entire document at once. Instead, you need to split it into chunks.

The following code shows how to use LangChain's `RecursiveCharacterTextSplitter` to split the document into chunks. - Use the default separator list, which is `["\n\n", "\n", " ", ""]`. - Chunk size is set to 100. This should be set to less than the model's maximum input token. - Chunk overlap is set to 20. - The length function is `len`.

```
[6]: from langchain.text_splitter import RecursiveCharacterTextSplitter
```

```
[7]: text_splitter = RecursiveCharacterTextSplitter(
    chunk_size=100,
    chunk_overlap=20,
    length_function=len,
)
```

```
[8]: chunks = text_splitter.split_text(data[0].page_content)
```

Let's see how many chunks you get.

```
[9]: len(chunks)
```

```
[9]: 570
```

Let's also see what these chunks looks like.



[10]: chunks

[10]: ['Madam Speaker, Madam Vice President, our First Lady and Second Gentleman. Members of Congress and',  
'of Congress and the Cabinet. Justices of the Supreme Court. My fellow Americans.',  
'Last year COVID-19 kept us apart. This year we are finally together again.',  
'Tonight, we meet as Democrats Republicans and Independents. But most importantly as Americans.',  
'With a duty to one another to the American people to the Constitution.',  
'And with an unwavering resolve that freedom will always triumph over tyranny.',  
'Six days ago, Russia's Vladimir Putin sought to shake the foundations of the free world thinking he',  
'world thinking he could make it bend to his menacing ways. But he badly miscalculated.',  
'He thought he could roll into Ukraine and the world would roll over. Instead he met a wall of',  
'he met a wall of strength he never imagined.',  
'He met the Ukrainian people.',  
'From President Zelenskyy to every Ukrainian, their fearlessness, their courage, their',  
'courage, their determination, inspires the world.',  
'Groups of citizens blocking tanks with their bodies. Everyone from students to retirees teachers',  
'retirees teachers turned soldiers defending their homeland.',  
'In this struggle as President Zelenskyy said in his speech to the European Parliament "Light will',  
'"Light will win over darkness." The Ukrainian Ambassador to the United States is here tonight.',  
'Let each of us here tonight in this Chamber send an unmistakable signal to Ukraine and to the',  
'Ukraine and to the world.',  
'Please rise if you are able and show that, Yes, we the United States of America stand with the',  
'stand with the Ukrainian people.',  
'Throughout our history we've learned this lesson when dictators do not pay a price for their',  
'a price for their aggression they cause more chaos.',  
'They keep moving. \n\nAnd the costs and the threats to America and the world keep rising.',  
'That's why the NATO Alliance was created to secure peace and stability in Europe after World War 2.',  
'after World War 2.',  
'The United States is a member along with 29 other nations.',  
'It matters. American diplomacy matters. American resolve matters.',  
'Putin's latest attack on Ukraine was premeditated and unprovoked.',

'He rejected repeated efforts at diplomacy.',  
'He thought the West and NATO wouldn't respond. And he thought he could divide us at home. Putin was',  
'at home. Putin was wrong. We were ready. Here is what we did.',  
'We prepared extensively and carefully.',  
'We spent months building a coalition of other freedom-loving nations from Europe and the Americas',  
'and the Americas to Asia and Africa to confront Putin.',  
'I spent countless hours unifying our European allies. We shared with the world in advance what we',  
'in advance what we knew Putin was planning and precisely how he would try to falsely justify his',  
'falsely justify his aggression.',  
'We countered Russia's lies with truth.',  
'And now that he has acted the free world is holding him accountable.',  
'Along with twenty-seven members of the European Union including France, Germany, Italy, as well as',  
'Italy, as well as countries like the United Kingdom, Canada, Japan, Korea, Australia, New Zealand,',  
'New Zealand, and many others, even Switzerland.',  
'We are inflicting pain on Russia and supporting the people of Ukraine. Putin is now isolated from',  
'now isolated from the world more than ever.',  
'Together with our allies -we are right now enforcing powerful economic sanctions.',  
'We are cutting off Russia's largest banks from the international financial system.',  
'Preventing Russia's central bank from defending the Russian Ruble making Putin's \$630 Billion "war',  
'\$630 Billion "war fund" worthless.',  
'We are choking off Russia's access to technology that will sap its economic strength and weaken its',  
'and weaken its military for years to come.',  
'Tonight I say to the Russian oligarchs and corrupt leaders who have bilked billions of dollars off',  
'of dollars off this violent regime no more.',  
'The U.S. Department of Justice is assembling a dedicated task force to go after the crimes of',  
'after the crimes of Russian oligarchs.',  
'We are joining with our European allies to find and seize your yachts your luxury apartments your',  
'apartments your private jets. We are coming for your ill-begotten gains.',  
'And tonight I am announcing that we will join our allies in closing off American air space to all',  
'air space to all Russian flights - further isolating Russia - and adding an additional squeeze -on',  
'squeeze -on their economy. The Ruble has lost 30% of its value.',

'The Russian stock market has lost 40% of its value and trading remains  
 suspended. Russia's economy',  
 'Russia's economy is reeling and Putin alone is to blame.',  
 'Together with our allies we are providing support to the Ukrainians in their  
 fight for freedom.',  
 'fight for freedom. Military assistance. Economic assistance. Humanitarian  
 assistance.',  
 'We are giving more than \$1 Billion in direct assistance to Ukraine.',  
 'And we will continue to aid the Ukrainian people as they defend their country  
 and to help ease',  
 'and to help ease their suffering.',  
 'Let me be clear, our forces are not engaged and will not engage in conflict  
 with Russian forces in',  
 'Russian forces in Ukraine.',  
 'Our forces are not going to Europe to fight in Ukraine, but to defend our NATO  
 Allies - in the',  
 'Allies - in the event that Putin decides to keep moving west.',  
 'For that purpose we've mobilized American ground forces, air squadrons, and  
 ship deployments to',  
 'ship deployments to protect NATO countries including Poland, Romania, Latvia,  
 Lithuania, and',  
 'Lithuania, and Estonia.',  
 'As I have made crystal clear the United States and our Allies will defend  
 every inch of territory',  
 'inch of territory of NATO countries with the full force of our collective  
 power.',  
 'And we remain clear-eyed. The Ukrainians are fighting back with pure courage.  
 But the next few days',  
 'the next few days weeks, months, will be hard on them.',  
 'Putin has unleashed violence and chaos. But while he may make gains on the  
 battlefield - he will',  
 '- he will pay a continuing high price over the long run.',  
 'And a proud Ukrainian people, who have known 30 years of independence, have  
 repeatedly shown that',  
 'shown that they will not tolerate anyone who tries to take their country  
 backwards.',  
 'To all Americans, I will be honest with you, as I've always promised. A  
 Russian dictator, invading',  
 'dictator, invading a foreign country, has costs around the world.',  
 'And I'm taking robust action to make sure the pain of our sanctions is  
 targeted at Russia's',  
 'at Russia's economy. And I will use every tool at our disposal to protect  
 American businesses and',  
 'businesses and consumers.',  
 'Tonight, I can announce that the United States has worked with 30 other  
 countries to release 60',  
 'to release 60 Million barrels of oil from reserves around the world.',

'America will lead that effort, releasing 30 Million barrels from our own Strategic Petroleum',  
'Strategic Petroleum Reserve. And we stand ready to do more if necessary, unified with our allies.',  
'These steps will help blunt gas prices here at home. And I know the news about what's happening can',  
'happening can seem alarming.',  
'But I want you to know that we are going to be okay.',  
'When the history of this era is written Putin's war on Ukraine will have left Russia weaker and the',  
'weaker and the rest of the world stronger.',  
'While it shouldn't have taken something so terrible for people around the world to see what's at',  
'to see what's at stake now everyone sees it clearly.',  
'We see the unity among leaders of nations and a more unified Europe a more unified West. And we see',  
'West. And we see unity among the people who are gathering in cities in large crowds around the',  
'crowds around the world even in Russia to demonstrate their support for Ukraine.',  
'In the battle between democracy and autocracy, democracies are rising to the moment, and the world',  
'and the world is clearly choosing the side of peace and security.',  
'This is a real test. It's going to take time. So let us continue to draw inspiration from the iron',  
'from the iron will of the Ukrainian people.',  
'To our fellow Ukrainian Americans who forge a deep bond that connects our two nations we stand with',  
'we stand with you.',  
'Putin may circle Kyiv with tanks, but he will never gain the hearts and souls of the Ukrainian',  
'of the Ukrainian people.',  
'He will never extinguish their love of freedom. He will never weaken the resolve of the free world.',  
'of the free world.',  
'We meet tonight in an America that has lived through two of the hardest years this nation has ever',  
'nation has ever faced.',  
'The pandemic has been punishing.',  
'And so many families are living paycheck to paycheck, struggling to keep up with the rising cost of',  
'the rising cost of food, gas, housing, and so much more.',  
'I understand.',  
'I remember when my Dad had to leave our home in Scranton, Pennsylvania to find work. I grew up in a',  
'I grew up in a family where if the price of food went up, you felt it.',  
'That's why one of the first things I did as President was fight to pass the

American Rescue Plan.',  
'Because people were hurting. We needed to act, and we did.',  
'Few pieces of legislation have done more in a critical moment in our history to lift us out of',  
'to lift us out of crisis.',  
'It fueled our efforts to vaccinate the nation and combat COVID-19. It delivered immediate economic',  
'immediate economic relief for tens of millions of Americans.',  
'Helped put food on their table, keep a roof over their heads, and cut the cost of health insurance.',  
'health insurance.',  
'And as my Dad used to say, it gave people a little breathing room.',  
'And unlike the \$2 Trillion tax cut passed in the previous administration that benefitted the top 1%',  
'the top 1% of Americans, the American Rescue Plan helped working people-and left no one behind.',  
'And it worked. It created jobs. Lots of jobs.',  
'In fact-our economy created over 6.5 Million new jobs just last year, more jobs created in one year',  
'created in one year',  
'than ever before in the history of America.',  
'Our economy grew at a rate of 5.7% last year, the strongest growth in nearly 40 years, the first',  
'40 years, the first step in bringing fundamental change to an economy that hasn't worked for the',  
'worked for the working people of this nation for too long.',  
'For the past 40 years we were told that if we gave tax breaks to those at the very top, the',  
'the very top, the benefits would trickle down to everyone else.',  
'But that trickle-down theory led to weaker economic growth, lower wages, bigger deficits, and the',  
'deficits, and the widest gap between those at the top and everyone else in nearly a century.',  
'Vice President Harris and I ran for office with a new economic vision for America.',  
'Invest in America. Educate Americans. Grow the workforce. Build the economy from the bottom up',  
'and the middle out, not from the top down.',  
'Because we know that when the middle class grows, the poor have a ladder up and the wealthy do very',  
'the wealthy do very well.',  
'America used to have the best roads, bridges, and airports on Earth.',  
'Now our infrastructure is ranked 13th in the world.',  
'We won't be able to compete for the jobs of the 21st Century if we don't fix that.',  
'That's why it was so important to pass the Bipartisan Infrastructure Law-the most sweeping',

'most sweeping investment to rebuild America in history.',  
 'This was a bipartisan effort, and I want to thank the members of both parties  
 who worked to make it',  
 'worked to make it happen.',  
 'We're done talking about infrastructure weeks. \n\nWe're going to have an  
 infrastructure decade.',  
 'It is going to transform America and put us on a path to win the economic  
 competition of the 21st',  
 'of the 21st Century that we face with the rest of the world-particularly with  
 China.',  
 'As I've told Xi Jinping, it is never a good bet to bet against the American  
 people.',  
 'We'll create good jobs for millions of Americans, modernizing roads, airports,  
 ports, and waterways',  
 'and waterways all across America.',  
 'And we'll do it all to withstand the devastating effects of the climate crisis  
 and promote',  
 'crisis and promote environmental justice.',  
 'We'll build a national network of 500,000 electric vehicle charging stations,  
 begin to replace',  
 'begin to replace poisonous lead pipes-so every child-and every American-has  
 clean water to drink at',  
 'water to drink at home and at school, provide affordable high-speed internet  
 for every',  
 'internet for every American-urban, suburban, rural, and tribal communities.',  
 '4,000 projects have already been announced.',  
 'And tonight, I'm announcing that this year we will start fixing over 65,000  
 miles of highway and',  
 'of highway and 1,500 bridges in disrepair.',  
 'When we use taxpayer dollars to rebuild America - we are going to Buy  
 American: buy American',  
 'buy American products to support American jobs.',  
 'The federal government spends about \$600 Billion a year to keep the country  
 safe and secure.',  
 'There's been a law on the books for almost a century',  
 'to make sure taxpayers' dollars support American jobs and businesses.',  
 'Every Administration says they'll do it, but we are actually doing it.',  
 'We will buy American to make sure everything from the deck of an aircraft  
 carrier to the steel on',  
 'to the steel on highway guardrails are made in America.',  
 'But to compete for the best jobs of the future, we also need to level the  
 playing field with China',  
 'field with China and other competitors.',  
 'That's why it is so important to pass the Bipartisan Innovation Act sitting in  
 Congress that will',  
 'Congress that will make record investments in emerging technologies and  
 American manufacturing.',

'Let me give you one example of why it's so important to pass it.',  
'If you travel 20 miles east of Columbus, Ohio, you'll find 1,000 empty acres of land.',  
'It won't look like much, but if you stop and look closely, you'll see a "Field of dreams," the',  
'of dreams," the ground on which America's future will be built.',  
'This is where Intel, the American company that helped build Silicon Valley, is going to build its',  
'going to build its \$20 billion semiconductor "mega site".',  
'Up to eight state-of-the-art factories in one place. 10,000 new good-paying jobs.',  
'Some of the most sophisticated manufacturing in the world to make computer chips the size of a',  
'chips the size of a fingertip that power the world and our everyday lives.',  
'Smartphones. The Internet. Technology we have yet to invent. \n\nBut that's just the beginning.',  
'Intel's CEO, Pat Gelsinger, who is here tonight, told me they are ready to increase their',  
'to increase their investment from',  
'\$20 billion to \$100 billion.',  
'That would be one of the biggest investments in manufacturing in American history.',  
'And all they're waiting for is for you to pass this bill.',  
'So let's not wait any longer. Send it to my desk. I'll sign it. \n\nAnd we will really take off.',  
'And Intel is not alone. \n\nThere's something happening in America.',  
'Just look around and you'll see an amazing story.',  
'The rebirth of the pride that comes from stamping products "Made In America." The revitalization of',  
'revitalization of American manufacturing.',  
'Companies are choosing to build new factories here, when just a few years ago, they would have',  
'they would have built them overseas.',  
'That's what is happening. Ford is investing \$11 billion to build electric vehicles, creating 11,000',  
'creating 11,000 jobs across the country.',  
'GM is making the largest investment in its history-\$7 billion to build electric vehicles, creating',  
'vehicles, creating 4,000 jobs in Michigan.',  
'All told, we created 369,000 new manufacturing jobs in America just last year.',  
'Powered by people I've met like JoJo Burgess, from generations of union steelworkers from',  
'steelworkers from Pittsburgh, who's here with us tonight.',  
'As Ohio Senator Sherrod Brown says, "It's time to bury the label "Rust Belt." \n\nIt's time.',  
'But with all the bright spots in our economy, record job growth and higher

wages, too many families',  
 'too many families are struggling to keep up with the bills.',  
 'Inflation is robbing them of the gains they might otherwise feel.',  
 'I get it. That's why my top priority is getting prices under control.',  
 'Look, our economy roared back faster than most predicted, but the pandemic meant that businesses',  
 'that businesses had a hard time hiring enough workers to keep up production in their factories.',  
 'The pandemic also disrupted global supply chains.',  
 'When factories close, it takes longer to make goods and get them from the warehouse to the store',  
 'to the store, and prices go up.',  
 'Look at cars.',  
 'Last year, there weren't enough semiconductors to make all the cars that people wanted to buy.',  
 'And guess what, prices of automobiles went up. \n\nSo-we have a choice.',  
 'One way to fight inflation is to drive down wages and make Americans poorer.',  
 'I have a better plan to fight inflation. \n\nLower your costs, not your wages.',  
 'Make more cars and semiconductors in America. \n\nMore infrastructure and innovation in America.',  
 'More goods moving faster and cheaper in America.',  
 'More jobs where you can earn a good living in America.',  
 'And instead of relying on foreign supply chains, let's make it in America.',  
 'Economists call it "increasing the productive capacity of our economy."',  
 'I call it building a better America.',  
 'My plan to fight inflation will lower your costs and lower the deficit.',  
 '17 Nobel laureates in economics say my plan will ease long-term inflationary pressures. Top',  
 'pressures. Top business leaders and most Americans support my plan. And here's the plan:',  
 'First - cut the cost of prescription drugs. Just look at insulin. One in ten Americans has',  
 'ten Americans has diabetes. In Virginia, I met a 13-year-old boy named Joshua Davis.',  
 'He and his Dad both have Type 1 diabetes, which means they need insulin every day. Insulin costs',  
 'day. Insulin costs about \$10 a vial to make.',  
 'But drug companies charge families like Joshua and his Dad up to 30 times more. I spoke with',  
 'more. I spoke with Joshua's mom.',  
 'Imagine what it's like to look at your child who needs insulin and have no idea how you're going to',  
 'how you're going to pay for it.',  
 'What it does to your dignity, your ability to look your child in the eye, to be the parent you',  
 'be the parent you expect to be.',



'Joshua is here with us tonight. Yesterday was his birthday. Happy birthday, buddy.',

'For Joshua, and for the 200,000 other young people with Type 1 diabetes, let's cap the cost of',

'cap the cost of insulin at \$35 a month so everyone can afford it.',

'Drug companies will still do very well. And while we're at it let Medicare negotiate lower prices',

'lower prices for prescription drugs, like the VA already does.',

'Look, the American Rescue Plan is helping millions of families on Affordable Care Act plans save',

'Care Act plans save \$2,400 a year on their health care premiums. Let's close the coverage gap and',

'coverage gap and make those savings permanent.',

'Second - cut energy costs for families an average of \$500 a year by combatting climate change.',

'Let's provide investments and tax credits to weatherize your homes and businesses to be energy',

'to be energy efficient and you get a tax credit; double America's clean energy production in solar,',

'in solar, wind, and so much more; lower the price of electric vehicles, saving you another \$80 a',

'you another \$80 a month because you'll never have to pay at the gas pump again.',

'Third - cut the cost of child care. Many families pay up to \$14,000 a year for child care per',

'for child care per child.',

'Middle-class and working families shouldn't have to pay more than 7% of their income for care of',

'income for care of young children.',

'My plan will cut the cost in half for most families and help parents, including millions of women,',

'millions of women, who left the workforce during the pandemic because they couldn't afford child',

'afford child care, to be able to get back to work.',

'My plan doesn't stop there. It also includes home and long-term care. More affordable housing. And',

'housing. And Pre-K for every 3- and 4-year-old.',

'All of these will lower costs.',

'And under my plan, nobody earning less than \$400,000 a year will pay an additional penny in new',

'penny in new taxes. Nobody.',

'The one thing all Americans agree on is that the tax system is not fair. We have to fix it.',

'I'm not looking to punish anyone. But let's make sure corporations and the wealthiest Americans',

'Americans start paying their fair share.',

'Just last year, 55 Fortune 500 corporations earned \$40 billion in profits and

paid zero dollars in',  
'zero dollars in federal income tax.',  
'That's simply not fair. That's why I've proposed a 15% minimum tax rate for corporations.',  
'We got more than 130 countries to agree on a global minimum tax rate so companies can't get out of',  
'can't get out of paying their taxes at home by shipping jobs and factories overseas.',  
'That's why I've proposed closing loopholes so the very wealthy don't pay a lower tax rate than a',  
'tax rate than a teacher or a firefighter.',  
'So that's my plan. It will grow the economy and lower costs for families.',  
'So what are we waiting for? Let's get this done. And while you're at it, confirm my nominees to the',  
'my nominees to the Federal Reserve, which plays a critical role in fighting inflation.',  
'My plan will not only lower costs to give families a fair shot, it will lower the deficit.',  
'The previous Administration not only ballooned the deficit with tax cuts for the very wealthy and',  
'very wealthy and corporations, it undermined the watchdogs whose job was to keep pandemic relief',  
'pandemic relief funds from being wasted.',  
'But in my administration, the watchdogs have been welcomed back.',  
'We're going after the criminals who stole billions in relief money meant for small businesses and',  
'businesses and millions of Americans.',  
'And tonight, I'm announcing that the Justice Department will name a chief prosecutor for pandemic',  
'for pandemic fraud.',  
'By the end of this year, the deficit will be down to less than half what it was before I took',  
'was before I took office.',  
'The only president ever to cut the deficit by more than one trillion dollars in a single year.',  
'Lowering your costs also means demanding more competition.',  
'I'm a capitalist, but capitalism without competition isn't capitalism.',  
'It's exploitation-and it drives up prices.',  
'When corporations don't have to compete, their profits go up, your prices go up, and small',  
'go up, and small businesses and family farmers and ranchers go under.',  
'We see it happening with ocean carriers moving goods in and out of America.',  
'During the pandemic, these foreign-owned companies raised prices by as much as 1,000% and made',  
'as 1,000% and made record profits.',  
'Tonight, I'm announcing a crackdown on these companies overcharging American businesses and',

'businesses and consumers.',  
 'And as Wall Street firms take over more nursing homes, quality in those homes  
 has gone down and',  
 'has gone down and costs have gone up.',  
 'That ends on my watch.',  
 'Medicare is going to set higher standards for nursing homes and make sure your  
 loved ones get the',  
 'loved ones get the care they deserve and expect.',  
 'We'll also cut costs and keep the economy going strong by giving workers a  
 fair shot, provide more',  
 'shot, provide more training and apprenticeships, hire them based on their  
 skills not degrees.',  
 'Let's pass the Paycheck Fairness Act and paid leave.',  
 'Raise the minimum wage to \$15 an hour and extend the Child Tax Credit, so no  
 one has to raise a',  
 'one has to raise a family in poverty.',  
 'Let's increase Pell Grants and increase our historic support of HBCUs, and  
 invest in what Jill-our',  
 'in what Jill-our First Lady who teaches full-time-calls America's best-kept  
 secret: community',  
 'secret: community colleges.',  
 'And let's pass the PRO Act when a majority of workers want to form a  
 union-they shouldn't be',  
 'shouldn't be stopped.',  
 'When we invest in our workers, when we build the economy from the bottom up  
 and the middle out',  
 'and the middle out together, we can do something we haven't done in a long  
 time: build a better',  
 'build a better America.',  
 'For more than two years, COVID-19 has impacted every decision in our lives and  
 the life of the',  
 'and the life of the nation.',  
 'And I know you're tired, frustrated, and exhausted. \n\nBut I also know  
 this.',  
 'Because of the progress we've made, because of your resilience and the tools  
 we have, tonight I can',  
 'have, tonight I can say',  
 'we are moving forward safely, back to more normal routines.',  
 'We've reached a new moment in the fight against COVID-19, with severe cases  
 down to a level not',  
 'down to a level not seen since last July.',  
 'Just a few days ago, the Centers for Disease Control and Prevention-the  
 CDC-issued new mask',  
 'CDC-issued new mask guidelines.',  
 'Under these new guidelines, most Americans in most of the country can now be  
 mask free.',  
 'And based on the projections, more of the country will reach that point across

the next couple of',  
'the next couple of weeks.',  
'Thanks to the progress we have made this past year, COVID-19 need no longer control our lives.',  
'I know some are talking about "living with COVID-19". Tonight - I say that we will never just',  
'we will never just accept living with COVID-19.',  
'We will continue to combat the virus as we do other diseases. And because this is a virus that',  
'is a virus that mutates and spreads, we will stay on guard.',  
'Here are four common sense steps as we move forward safely.',  
'First, stay protected with vaccines and treatments. We know how incredibly effective vaccines are.',  
'vaccines are. If you're vaccinated and boosted you have the highest degree of protection.',  
'We will never give up on vaccinating more Americans. Now, I know parents with kids under 5 are',  
'kids under 5 are eager to see a vaccine authorized for their children.',  
'The scientists are working hard to get that done and we'll be ready with plenty of vaccines when',  
'of vaccines when they do.',  
'We're also ready with anti-viral treatments. If you get COVID-19, the Pfizer pill reduces your',  
'pill reduces your chances of ending up in the hospital by 90%.',  
'We've ordered more of these pills than anyone in the world. And Pfizer is working overtime to get',  
'overtime to get us 1 Million pills this month and more than double that next month.',  
'And we're launching the "Test to Treat" initiative so people can get tested at a pharmacy, and if',  
'a pharmacy, and if they're positive, receive antiviral pills on the spot at no cost.',  
'If you're immunocompromised or have some other vulnerability, we have treatments and free',  
'treatments and free high-quality masks.',  
'We're leaving no one behind or ignoring anyone's needs as we move forward.',  
'And on testing, we have made hundreds of millions of tests available for you to order for free.',  
'Even if you already ordered free tests tonight, I am announcing that you can order more from',  
'can order more from covidtests.gov starting next week.',  
'Second - we must prepare for new variants. Over the past year, we've gotten much better at',  
'much better at detecting new variants.',  
'If necessary, we'll be able to deploy new vaccines within 100 days instead of many more months or',  
'many more months or years.'

'And, if Congress provides the funds we need, we'll have new stockpiles of tests, masks, and pills',  
 'masks, and pills ready if needed.',  
 'I cannot promise a new variant won't come. But I can promise you we'll do everything within our',  
 'within our power to be ready if it does.',  
 'Third - we can end the shutdown of schools and businesses. We have the tools we need.',  
 'It's time for Americans to get back to work and fill our great downtowns again. People working',  
 'People working from home can feel safe to begin to return to the office.',  
 'We're doing that here in the federal government. The vast majority of federal workers will once',  
 'workers will once again work in person.',  
 'Our schools are open. Let's keep it that way. Our kids need to be in school.',  
 'And with 75% of adult Americans fully vaccinated and hospitalizations down by 77%, most Americans',  
 '77%, most Americans can remove their masks, return to work, stay in the classroom, and move forward',  
 'and move forward safely.',  
 'We achieved this because we provided free vaccines, treatments, tests, and masks.',  
 'Of course, continuing this costs money. \n\nI will soon send Congress a request.',  
 'The vast majority of Americans have used these tools and may want to again, so I expect Congress to',  
 'expect Congress to pass it quickly.',  
 'Fourth, we will continue vaccinating the world.',  
 'We've sent 475 Million vaccine doses to 112 countries, more than any other nation.',  
 'And we won't stop.',  
 'We have lost so much to COVID-19. Time with one another. And worst of all, so much loss of life.',  
 'Let's use this moment to reset. Let's stop looking at COVID-19 as a partisan dividing line and see',  
 'line and see it for what it is: A God-awful disease.',  
 'Let's stop seeing each other as enemies, and start seeing each other for who we really are: Fellow',  
 'really are: Fellow Americans.',  
 'We can't change how divided we've been. But we can change how we move forward-on COVID-19 and other',  
 'COVID-19 and other issues we must face together.',  
 'I recently visited the New York City Police Department days after the funerals of Officer Wilbert',  
 'of Officer Wilbert Mora and his partner, Officer Jason Rivera.',  
 'They were responding to a 9-1-1 call when a man shot and killed them with a stolen gun.',

'Officer Mora was 27 years old. \n\nOfficer Rivera was 22.',  
 'Both Dominican Americans who'd grown up on the same streets they later chose  
 to patrol as police',  
 'to patrol as police officers.',  
 'I spoke with their families and told them that we are forever in debt for  
 their sacrifice, and we',  
 'sacrifice, and we will carry on their mission to restore the trust and safety  
 every community',  
 'every community deserves.',  
 'I've worked on these issues a long time.',  
 'I know what works: Investing in crime preventionand community police officers  
 who'll walk the beat,',  
 'walk the beat, who'll know the neighborhood, and who can restore trust and  
 safety.',  
 'So let's not abandon our streets. Or choose between safety and equal  
 justice.',  
 'Let's come together to protect our communities, restore trust, and hold law  
 enforcement',  
 'law enforcement accountable.',  
 'That's why the Justice Department required body cameras, banned chokeholds,  
 and restricted no-knock',  
 'restricted no-knock warrants for its officers.',  
 'That's why the American Rescue Plan provided \$350 Billion that cities, states,  
 and counties can use',  
 'counties can use to hire more police and invest in proven strategies like  
 community violence',  
 'community violence interruption-trusted messengers breaking the cycle of  
 violence and trauma and',  
 'and trauma and giving young people hope.',  
 'We should all agree: The answer is not to Defund the police. The answer is to  
 FUND the police with',  
 'the police with the resources and training they need to protect our  
 communities.',  
 'I ask Democrats and Republicans alike: Pass my budget and keep our  
 neighborhoods safe.',  
 'And I will keep doing everything in my power to crack down on gun trafficking  
 and ghost guns you',  
 'and ghost guns you can buy online and make at home-they have no serial numbers  
 and can't be traced.',  
 'can't be traced.',  
 'And I ask Congress to pass proven measures to reduce gun violence. Pass  
 universal background',  
 'background checks. Why should anyone on a terrorist list be able to purchase a  
 weapon?',  
 'Ban assault weapons and high-capacity magazines.',  
 'Repeal the liability shield that makes gun manufacturers the only industry in  
 America that can't be',

'that can't be sued.',  
 'These laws don't infringe on the Second Amendment. They save lives.',  
 'The most fundamental right in America is the right to vote - and to have it counted. And it's under',  
 'And it's under assault.',  
 'In state after state, new laws have been passed, not only to suppress the vote, but to subvert',  
 'but to subvert entire elections.',  
 'We cannot let this happen.',  
 'Tonight. I call on the Senate to: Pass the Freedom to Vote Act. Pass the John Lewis Voting Rights',  
 'Lewis Voting Rights Act. And while you're at it, pass the Disclose Act so Americans can know who is',  
 'can know who is funding our elections.',  
 'Tonight, I'd like to honor someone who has dedicated his life to serve this country: Justice',  
 'country: Justice Stephen Breyer-an Army veteran, Constitutional scholar, and retiring Justice of',  
 'retiring Justice of the United States Supreme Court. Justice Breyer, thank you for your service.',  
 'One of the most serious constitutional responsibilities a President has is nominating someone to',  
 'someone to serve on the United States Supreme Court.',  
 'And I did that 4 days ago, when I nominated Circuit Court of Appeals Judge Ketanji Brown Jackson.',  
 'Brown Jackson. One of our nation's top legal minds, who will continue Justice Breyer's legacy of',  
 'Breyer's legacy of excellence.',  
 'A former top litigator in private practice. A former federal public defender. And from a family of',  
 'from a family of public school educators and police officers. A consensus builder. Since she's been',  
 'Since she's been nominated, she's received a broad range of support-from the Fraternal Order of',  
 'Fraternal Order of Police to former judges appointed by Democrats and Republicans.',  
 'And if we are to advance liberty and justice, we need to secure the Border and fix the immigration',  
 'fix the immigration system.',  
 'We can do both. At our border, we've installed new technology like cutting-edge scanners to better',  
 'scanners to better detect drug smuggling.',  
 'We've set up joint patrols with Mexico and Guatemala to catch more human traffickers.',  
 'We're putting in place dedicated immigration judges so families fleeing persecution and violence',  
 'and violence can have their cases heard faster.',

'We're securing commitments and supporting partners in South and Central America to host more',  
 'to host more refugees and secure their own borders.',  
 'We can do all this while keeping lit the torch of liberty that has led generations of immigrants to',  
 'of immigrants to this land-my forefathers and so many of yours.',  
 'Provide a pathway to citizenship for Dreamers, those on temporary status, farm workers, and',  
 'farm workers, and essential workers.',  
 'Revise our laws so businesses have the workers they need and families don't wait decades to',  
 'wait decades to reunite.',  
 'It's not only the right thing to do-it's the economically smart thing to do.',  
 'That's why immigration reform is supported by everyone from labor unions to religious leaders to',  
 'leaders to the U.S. Chamber of Commerce.',  
 'Let's get it done once and for all.',  
 'Advancing liberty and justice also requires protecting the rights of women.',  
 'The constitutional right affirmed in Roe v. Wade-standing precedent for half a century-is under',  
 'a century-is under attack as never before.',  
 'If we want to go forward-not backward-we must protect access to health care. Preserve a woman's',  
 'Preserve a woman's right to choose. And let's continue to advance maternal health care in America.',  
 'And for our LGBTQ+ Americans, let's finally get the bipartisan Equality Act to my desk. The',  
 'Act to my desk. The onslaught of state laws targeting transgender Americans and their families is',  
 'their families is wrong.',  
 'As I said last year, especially to our younger transgender Americans, I will always have your back',  
 'have your back as your President, so you can be yourself and reach your God-given potential.',  
 'While it often appears that we never agree, that isn't true. I signed 80 bipartisan bills into law',  
 'bills into law last year. From preventing government shutdowns to protecting Asian-Americans from',  
 'from still-too-common hate crimes to reforming military justice.',  
 'And soon, we'll strengthen the Violence Against Women Act that I first wrote three decades ago. It',  
 'decades ago. It is important for us to show the nation that we can come together and do big things.',  
 'and do big things.',  
 'So tonight I'm offering a Unity Agenda for the Nation. Four big things we can do together.',  
 'First, beat the opioid epidemic.',



'There is so much we can do. Increase funding for prevention, treatment, harm reduction, and',  
 'harm reduction, and recovery.',  
 'Get rid of outdated rules that stop doctors from prescribing treatments. And stop the flow of',  
 'stop the flow of illicit drugs by working with state and local law enforcement to go after',  
 'to go after traffickers.',  
 'If you're suffering from addiction, know you are not alone. I believe in recovery, and I celebrate',  
 'and I celebrate the 23 million Americans in recovery.',  
 'Second, let's take on mental health. Especially among our children, whose lives and education have',  
 'and education have been turned upside down.',  
 'The American Rescue Plan gave schools money to hire teachers and help students make up for lost',  
 'make up for lost learning.',  
 'I urge every parent to make sure your school does just that. And we can all play a part-sign up to',  
 'a part-sign up to be a tutor or a mentor.',  
 'Children were also struggling before the pandemic. Bullying, violence, trauma, and the harms of',  
 'and the harms of social media.',  
 'As Frances Haugen, who is here with us tonight, has shown, we must hold social media platforms',  
 'media platforms accountable for the national experiment they're conducting on our children for',  
 'on our children for profit.',  
 'It's time to strengthen privacy protections, ban targeted advertising to children, demand tech',  
 'demand tech companies stop collecting personal data on our children.',  
 'And let's get all Americans the mental health services they need. More people they can turn to for',  
 'can turn to for help, and full parity between physical and mental health care.',  
 'Third, support our veterans. \n\nVeterans are the best of us.',  
 'I've always believed that we have a sacred obligation to equip all those we send to war and care',  
 'to war and care for them and their families when they come home.',  
 'My administration is providing assistance with job training and housing, and now helping',  
 'and now helping lower-income veterans get VA care debt-free.',  
 'Our troops in Iraq and Afghanistan faced many dangers.',  
 'One was stationed at bases and breathing in toxic smoke from "burn pits" that incinerated wastes of',  
 'wastes of war-medical and hazard material, jet fuel, and more.',  
 'When they came home, many of the world's fittest and best trained warriors

were never the same.',  
 'Headaches. Numbness. Dizziness. \n\nA cancer that would put them in a flag-  
 draped coffin. \n\nI know.',  
 'I know. \n\nOne of those soldiers was my son Major Beau Biden.',  
 'We don't know for sure if a burn pit was the cause of his brain cancer, or the  
 diseases of so many',  
 'diseases of so many of our troops.',  
 'But I'm committed to finding out everything we can.',  
 'Committed to military families like Danielle Robinson from Ohio.',  
 'The widow of Sergeant First Class Heath Robinson.',  
 'He was born a soldier. Army National Guard. Combat medic in Kosovo and Iraq.',  
 'Stationed near Baghdad, just yards from burn pits the size of football  
 fields.',  
 'Heath's widow Danielle is here with us tonight. They loved going to Ohio State  
 football games. He',  
 'football games. He loved building Legos with their daughter.',  
 'But cancer from prolonged exposure to burn pits ravaged Heath's lungs and  
 body.',  
 'Danielle says Heath was a fighter to the very end.',  
 'He didn't know how to stop fighting, and neither did she.',  
 'Through her pain she found purpose to demand we do better. \n\nTonight,  
 Danielle-we are.',  
 'The VA is pioneering new ways of linking toxic exposures to diseases, already  
 helping more veterans',  
 'more veterans get benefits.',  
 'And tonight, I'm announcing we're expanding eligibility to veterans suffering  
 from nine respiratory',  
 'nine respiratory cancers.',  
 'I'm also calling on Congress: pass a law to make sure veterans devastated by  
 toxic exposures in',  
 'toxic exposures in Iraq and Afghanistan finally get the benefits and  
 comprehensive health care they',  
 'health care they deserve.',  
 'And fourth, let's end cancer as we know it.',  
 'This is personal to me and Jill, to Kamala, and to so many of you.',  
 'Cancer is the #2 cause of death in America-second only to heart disease.',  
 'Last month, I announced our plan to supercharge',  
 'the Cancer Moonshot that President Obama asked me to lead six years ago.',  
 'Our goal is to cut the cancer death rate by at least 50% over the next 25  
 years, turn more cancers',  
 'turn more cancers from death sentences into treatable diseases.',  
 'More support for patients and families.',  
 'To get there, I call on Congress to fund ARPA-H, the Advanced Research  
 Projects Agency for Health.',  
 'It's based on DARPA-the Defense Department project that led to the Internet,  
 GPS, and so much more.',  
 'and so much more.',

'ARPA-H will have a singular purpose-to drive breakthroughs in cancer, Alzheimer's, diabetes, and',  
 'diabetes, and more.',  
 'A unity agenda for the nation. \n\nWe can do this.',  
 'My fellow Americans-tonight , we have gathered in a sacred space-the citadel of our democracy.',  
 'In this Capitol, generation after generation, Americans have debated great questions amid great',  
 'amid great strife, and have done great things.',  
 'We have fought for freedom, expanded liberty, defeated totalitarianism and terror.',  
 'And built the strongest, freest, and most prosperous nation the world has ever known.',  
 'Now is the hour. \n\nOur moment of responsibility.',  
 'Our test of resolve and conscience, of history itself.',  
 'It is in this moment that our character is formed. Our purpose is found. Our future is forged.',  
 'Well I know this nation. \n\nWe will meet the test.',  
 'To protect freedom and liberty, to expand fairness and opportunity. \n\nWe will save democracy.',  
 'As hard as these times have been, I am more optimistic about America today than I have been my',  
 'than I have been my whole life.',  
 'Because I see the future that is within our grasp.',  
 'Because I know there is simply nothing beyond our capacity.',  
 'We are the only nation on Earth that has always turned every crisis we have faced into an',  
 'have faced into an opportunity.',  
 'The only nation that can be defined by a single word: possibilities.',  
 'So on this night, in our 245th year as a nation, I have come to report on the State of the Union.',  
 'And my report is this: the State of the Union is strong-because you, the American people, are',  
 'people, are strong.',  
 'We are stronger today than we were a year ago.',  
 'And we will be stronger a year from now than we are today.',  
 'Now is our moment to meet and overcome the challenges of our time. \n\nAnd we will, as one people.',  
 'One America. \n\nThe United States of America. \n\nMay God bless you all. May God protect our troops.']

## 1.7 Watsonx embedding model

### 1.7.1 Model description

In this section, you will use IBM `slate-125m-english-rtrvr` model as an example embedding model.

The `slate.125m.english.rtrvr` model is a [standard sentence](#) transformers model based on bi-encoders.

The model produces an embedding for a given input, e.g., query, passage, document, etc. At a high level, the model is trained to maximize the cosine similarity between two input pieces of text, e.g., text A (query text) and text B (passage text), which results in the sentence embeddings  $q$  and  $p$ . These sentence embeddings can be compared using cosine similarity, which measures the distance between sentences by calculating the distance between their embeddings.

The embedding model, `slate.125m.english` formerly known as WatBERT, has the same architecture as a RoBERTa base transformer model and has ~125 million parameters and an embedding dimension of 768.

Model name	API model_id	Maximum input tokens	Number of dimensions	More information
slate-125m-english-rtrvr	ibm/slate-125m-english-rtrvr	512	768	<a href="#">model card</a>

### 1.7.2 Build model

The following code shows how to build the `slate-125m-english-rtrvr` model from IBM watsonx.ai API.

First, import the necessary dependencies. - `WatsonxEmbeddings` is a class/dependence that can be used to form an embedding model object. - `EmbedTextParamsMetaNames` is a dependence that controls the embedding parameters.

```
[11]: from ibm_watsonx_ai.metanames import EmbedTextParamsMetaNames
      from langchain_ibm import WatsonxEmbeddings

[12]: embed_params = {
      EmbedTextParamsMetaNames.TRUNCATE_INPUT_TOKENS: 3,
      EmbedTextParamsMetaNames.RETURN_OPTIONS: {"input_text": True},
      }

      watsonx_embedding = WatsonxEmbeddings(
          model_id="ibm/slate-125m-english-rtrvr",
          url="https://us-south.ml.cloud.ibm.com",
          project_id="skills-network",
          params=embed_params,
      )
```

### 1.7.3 Query embeddings

Now, create an embedding based on a single sentence, which can be treated as a query.

Use the `embed_query` method.

```
[13]: query = "How are you?"

      query_result = watsonx_embedding.embed_query(query)
```

Let's see the length/dimension of this embedding.

```
[14]: len(query_result)
```

```
[14]: 768
```

It has a dimension of 768, which aligns with the model description.

Next, take a look at the first five results from the embeddings.

```
[15]: query_result[:5]
```

```
[15]: [-0.06722454, -0.023729993, 0.017487843, -0.013195328, -0.039584607]
```

Embeddings des nachfolgenden Modells: [0.027106120, 0.0113318320, -0.00195238508, -0.0369513258, 0.017764887] Wenn man sich auf bestimmte Embeddings festgelegt hat, ist man daran gebunden!

```
[16]: query_result
```

```
[16]: [-0.06722454,  
      -0.023729993,  
      0.017487843,  
      -0.013195328,  
      -0.039584607,  
      0.035013206,  
      0.026268983,  
      -0.016346082,  
      0.0017328745,  
      0.0571103,  
      -0.0721077,  
      -0.017297767,  
      -0.04961109,  
      -0.017423011,  
      -0.028412413,  
      -0.008028798,  
      0.017073669,  
      -0.0052801697,  
      0.02541991,  
      -0.033137497,  
      0.009124655,  
      0.0023437312,  
      0.002702213,  
      -0.023014594,  
      -0.012872512,  
      0.011373165,  
      -0.0090638185,  
      -0.03357957,  
      0.0021336142,  
      0.0397427,  
      -0.0070087253,
```

0.017540198,  
0.013501195,  
-0.020656323,  
0.031942666,  
0.010870434,  
-0.008687372,  
0.023959273,  
0.0007889645,  
-0.007972021,  
-0.028409995,  
-0.009587482,  
-0.027156807,  
-0.022130553,  
-0.002896199,  
-0.016902613,  
0.07734626,  
0.010192929,  
0.044747256,  
0.016257776,  
0.066523395,  
-0.00929883,  
0.05224961,  
0.009617802,  
0.047355816,  
-0.038022686,  
0.0165344,  
0.018556952,  
-0.0104811415,  
0.0061104964,  
-0.01024595,  
0.051244803,  
-0.055859342,  
-0.008510792,  
0.028069686,  
0.011818036,  
-0.040075816,  
0.029689146,  
-0.087874584,  
0.024186028,  
-0.06456673,  
-0.0076045888,  
0.01654361,  
-0.058433313,  
-0.034822233,  
0.02772002,  
-0.03268237,  
-0.02774234,

-0.024907194,  
-0.006553365,  
0.023166403,  
-0.0041369274,  
0.015799565,  
0.001904313,  
0.01980826,  
-0.0018761983,  
-0.018335281,  
-0.027995942,  
-0.034934834,  
-0.013229686,  
0.032818478,  
-0.035363052,  
0.028499516,  
-0.05107064,  
0.015736636,  
0.009241721,  
0.023862502,  
0.043013494,  
0.030826705,  
0.024911856,  
0.061978586,  
-0.021643378,  
0.03621004,  
-0.031883728,  
0.028464902,  
0.052406836,  
-0.03415915,  
0.026003491,  
-0.012133338,  
-0.0033279215,  
0.0723308,  
-0.07607222,  
-0.04539286,  
0.026521046,  
0.04766782,  
0.03823079,  
0.008952164,  
-0.05222688,  
0.041690867,  
-0.04328276,  
-0.07080701,  
-0.042033132,  
-0.076321214,  
-0.06832438,  
0.05429533,

0.029673679,  
0.041617557,  
0.035298593,  
0.021892173,  
-0.01035526,  
-0.007343944,  
0.016643193,  
0.020906948,  
-0.018093234,  
-0.0057902075,  
-0.020870501,  
0.015880644,  
0.0421515,  
-0.019949656,  
-0.0115661295,  
0.017166004,  
-0.0039207703,  
0.021390766,  
-0.00051412696,  
-0.028978363,  
0.028017879,  
0.02085358,  
-0.078443736,  
-0.03266977,  
-0.0041859746,  
-0.037841912,  
-0.05245778,  
0.021955198,  
0.060466062,  
-0.027981356,  
-0.025444549,  
0.020927113,  
0.04582684,  
0.0068703396,  
-0.015889412,  
0.030123811,  
0.007785729,  
-0.040102743,  
0.0071524004,  
0.032464936,  
0.0124876965,  
-0.034773845,  
-0.018295875,  
-0.0012087027,  
-0.07167369,  
0.024570715,  
-0.009726627,



-0.038413875,  
-0.019947456,  
-0.033164926,  
-0.07943935,  
0.010327819,  
-0.040965326,  
0.01806271,  
0.040272914,  
0.099162795,  
-0.014621564,  
-0.025262315,  
0.01583893,  
-0.020925328,  
0.0671015,  
-0.045359034,  
-0.005166108,  
0.01785917,  
0.017828135,  
-0.036870833,  
0.018707165,  
0.025993181,  
0.0048187757,  
0.037456173,  
-0.0108644515,  
0.030485839,  
-0.021875918,  
0.0009316794,  
-0.021685513,  
-0.030767525,  
0.03251512,  
-0.04346948,  
-0.024791716,  
0.01850806,  
-0.025016801,  
-0.03589958,  
0.037892673,  
0.043621857,  
-0.059322376,  
-0.004409154,  
-0.003259454,  
-0.015706513,  
-0.28357172,  
-0.01765032,  
0.016408194,  
-0.132328,  
-0.0054026702,  
0.009985612,

0.008647153,  
-0.017288662,  
0.011719118,  
-0.009293158,  
0.005329986,  
-0.014740604,  
-0.032484908,  
0.029534059,  
-0.04015033,  
-0.041663893,  
0.026231203,  
0.018610755,  
-0.029331755,  
0.01970585,  
0.0232227,  
-0.026472427,  
-0.062451843,  
0.04024927,  
-0.009222635,  
-0.017783197,  
0.028875386,  
-0.04387314,  
-0.009172956,  
-0.0001155504,  
0.06445172,  
0.01852699,  
-0.0006696567,  
-0.0072787013,  
-0.013408191,  
0.01192863,  
-0.05409385,  
-0.018798893,  
-0.013029467,  
-0.2408881,  
-0.0006449026,  
-0.023697007,  
0.045447573,  
0.022037933,  
0.020504631,  
0.024958521,  
-0.0034134309,  
0.017238991,  
0.0021879466,  
-0.009537015,  
0.031147776,  
-0.0470248,  
-0.0036238723,

0.017840993,  
0.0034489844,  
0.03137266,  
0.029044384,  
-0.04737336,  
0.010220993,  
0.06280561,  
0.06503777,  
0.009405424,  
-0.039099466,  
0.031956807,  
-0.013036608,  
-0.013754215,  
0.003991459,  
0.028008541,  
-0.0031983512,  
0.034079324,  
-0.019345928,  
-0.022447912,  
-0.04008743,  
0.01969634,  
-0.022922145,  
-3.4059885e-05,  
0.026170112,  
-0.020740965,  
0.0028476717,  
0.029027516,  
-0.033788297,  
0.025670784,  
0.054839414,  
0.0065655974,  
-0.016915193,  
-0.035949495,  
0.010433692,  
-0.009507914,  
-0.043013297,  
0.028793942,  
-0.0011624522,  
-0.023622556,  
-0.033449072,  
0.02394655,  
-0.00822601,  
-0.01473033,  
0.029443318,  
0.0072977357,  
-0.037048135,  
-0.046621494,

-0.015566333,  
0.026013657,  
-0.017892016,  
0.006063957,  
0.0023254964,  
0.020134445,  
0.0037855709,  
-0.020259904,  
0.019973245,  
0.062036473,  
0.04664174,  
-0.02034029,  
0.021675821,  
0.07472238,  
0.025873106,  
0.042626016,  
-0.007866034,  
-0.070485845,  
-0.007841872,  
-0.0017894959,  
0.0060411086,  
0.025519196,  
-0.026332626,  
0.021141779,  
-0.019905431,  
-0.036675386,  
0.034737747,  
0.02220566,  
0.05096013,  
0.0021026682,  
-0.028781628,  
-0.030223235,  
-0.008009587,  
-0.008597355,  
-0.08181257,  
0.022011764,  
0.03758298,  
0.06818832,  
-0.022083169,  
0.009199552,  
0.012032968,  
-0.036478195,  
-0.06414284,  
0.013720073,  
0.027904207,  
-0.035400555,  
-0.02965272,

-0.013544973,  
0.03268437,  
0.018884385,  
-0.03589316,  
0.0400757,  
-0.004288737,  
0.0065071275,  
-0.035135183,  
0.019213362,  
-0.0051839794,  
-0.02296399,  
-0.020977207,  
0.025904847,  
0.022376189,  
-0.027651103,  
-0.014076218,  
0.03092889,  
-0.079350606,  
-0.012231418,  
0.002005919,  
0.00331962,  
-0.047333047,  
0.041038707,  
-0.027747346,  
-0.062386192,  
0.0344266,  
-0.0006287379,  
-0.021185417,  
0.01605974,  
0.019775243,  
-0.016278237,  
0.10593479,  
0.024622362,  
0.047323518,  
0.057055544,  
-0.032622688,  
0.0837405,  
0.013158537,  
-0.017449645,  
-0.009003906,  
-0.0025256437,  
0.042810705,  
-0.0435028,  
-0.013352283,  
-0.012529934,  
0.08813047,  
-0.08433992,

-0.014972823,  
0.11374519,  
-0.03298366,  
0.009955626,  
-0.031106567,  
0.024170713,  
0.03699106,  
0.0028586984,  
-0.0026328673,  
0.037273373,  
0.031983905,  
0.040924147,  
0.006574421,  
-0.048946306,  
-0.016951323,  
0.017962558,  
0.001709784,  
0.2177823,  
-0.012243276,  
-0.05004646,  
-0.0020600175,  
-0.053124685,  
-0.0031153879,  
0.039001524,  
-0.03362239,  
0.023183882,  
-0.02117483,  
0.09611747,  
0.0095118815,  
-0.05250059,  
-0.01817278,  
0.03729259,  
-0.057749152,  
-0.013294511,  
-0.06821871,  
0.016727619,  
-0.034375787,  
0.021188265,  
0.022696516,  
0.037687916,  
-0.0039334456,  
-5.932546e-05,  
-0.023479592,  
0.028875569,  
0.015842682,  
0.013222266,  
0.045494396,

0.029369412,  
-0.016189879,  
-0.0132519165,  
0.010400355,  
-0.0095192855,  
-0.030444242,  
-0.01872283,  
-0.019748425,  
-0.029347025,  
0.017578196,  
0.05879046,  
0.031586163,  
-0.026485048,  
-0.056373414,  
0.007970381,  
-0.0067538866,  
-0.035401486,  
-0.01612609,  
0.07254415,  
0.037870564,  
-0.047262013,  
-0.021215025,  
-0.014108702,  
0.043487906,  
0.05258437,  
0.006874368,  
-0.026191575,  
0.044459417,  
-0.01797606,  
-0.032685425,  
0.033352982,  
0.049406167,  
-0.012074338,  
-0.00035741003,  
-0.06456156,  
-0.044894427,  
-0.0181325,  
-0.007603064,  
-0.08239568,  
0.0011747673,  
0.027255842,  
0.01614236,  
0.0025581708,  
0.023169067,  
0.008250359,  
0.0047441036,  
-0.022313116,

-0.047295287,  
-0.006546512,  
0.02982424,  
-0.004806572,  
0.038411133,  
0.019149443,  
0.031204265,  
-0.02160092,  
-0.047022402,  
0.007123535,  
0.01284781,  
-0.00096812384,  
0.054738738,  
-0.049797673,  
0.027554417,  
0.027115365,  
0.011576999,  
-0.0395702,  
0.010787675,  
-0.012926269,  
0.026221,  
0.005134827,  
0.004114412,  
-0.005465977,  
-0.01978783,  
-0.00021866143,  
-0.02575867,  
-0.041323543,  
-0.029870367,  
0.009251497,  
-0.0067668837,  
0.016422248,  
-0.015582096,  
-0.0032229403,  
0.021139067,  
0.002345128,  
0.041876413,  
-0.03170716,  
-0.020851465,  
0.00444637,  
-0.008263837,  
-0.021481758,  
-0.0840727,  
-0.024300944,  
0.008947599,  
0.08880224,  
-0.024432592,



0.023595653,  
0.0508692,  
-0.018014988,  
-0.031196216,  
0.009017828,  
-7.596886e-05,  
-0.013442756,  
0.030178037,  
0.0033020454,  
0.023084454,  
0.05374295,  
-0.037115037,  
0.04399153,  
0.0058427374,  
-0.004270568,  
-0.006344175,  
0.004707057,  
0.037882797,  
-0.016168524,  
-0.03673947,  
-0.003819329,  
-0.021599503,  
0.031486075,  
-0.034679566,  
-0.013894746,  
0.0013890695,  
0.008054422,  
-0.033693112,  
0.005422246,  
0.024470221,  
-0.014058833,  
0.014632633,  
0.0036097178,  
0.0064672716,  
0.0020671377,  
0.023427159,  
-0.0853438,  
0.014969556,  
-0.00970658,  
0.021453416,  
-0.022904865,  
-0.023744116,  
-0.060693994,  
-0.036001194,  
0.04096359,  
0.017321426,  
2.0463765e-05,

-0.027393986,  
-0.04250415,  
-0.03714639,  
0.02418469,  
-0.03634261,  
0.03509443,  
0.012959197,  
-0.069309,  
-0.05688327,  
-0.022972703,  
0.04825098,  
-0.0070261746,  
0.02460965,  
0.03518499,  
0.04548682,  
-0.02426543,  
-0.052501865,  
-0.008089482,  
-0.026447006,  
0.008933223,  
-0.008084308,  
-0.013452941,  
-0.025808847,  
0.037585612,  
-0.004602825,  
0.011124059,  
0.0058883755,  
0.025275715,  
-0.003333772,  
-0.022864768,  
0.043936156,  
0.020140165,  
-0.0006580436,  
0.021422973,  
0.019339869,  
0.012094279,  
0.03681956,  
0.011538912,  
-0.02468816,  
-0.016712403,  
-0.058771417,  
-0.0070213126,  
-0.030203639,  
-0.06518787,  
0.005810488,  
0.04143981,  
0.0025484671,

0.006904106,  
0.0077010174,  
0.027719207,  
0.036596626,  
-0.015744856,  
-0.009164998,  
-0.022093322,  
-0.025826506,  
0.03720435,  
0.057949778,  
0.01040966,  
0.00042051356,  
-0.022143358,  
0.027365444,  
0.027921537,  
-0.045516405,  
-0.049947355,  
0.02981352,  
0.006173978,  
-0.029749896,  
0.0154333245,  
0.030925909,  
0.014590267,  
-0.016110657,  
-0.011238466,  
-0.0072797085,  
-0.025379356,  
0.01513153,  
-0.035147086,  
0.035199925,  
-0.024340596,  
-1.9044415e-05,  
-0.072061375,  
-0.024322703,  
0.013835519,  
0.068575166,  
0.04664718,  
-0.00018955515,  
-0.03986745,  
0.0070600156,  
0.010219543,  
0.06681889,  
0.00025776066,  
0.024534578,  
0.04103243,  
0.044350315,  
-0.0085192025,

-0.036313064,  
-0.014887768,  
0.0039340192,  
0.018266233,  
0.010035789,  
-0.025854273,  
0.004913163,  
-0.061988845,  
-0.012922391,  
0.027375467,  
0.06313827,  
0.025960613,  
-0.03840754,  
-0.013701233,  
-0.032558024,  
-0.02316364,  
-0.014939711,  
0.03540767,  
-0.00606778,  
0.011539444,  
-0.02555829,  
-0.055061564,  
-0.018397791,  
-0.009948863,  
-0.014422891,  
-0.0018951237,  
-0.043594606,  
0.0191271,  
-0.0038726486,  
-0.011145249,  
0.0043897233,  
-0.020771591,  
-0.044244777,  
-0.051051624,  
0.02264657,  
0.016734788,  
0.004856589,  
-0.06837033,  
-0.016216269,  
-0.033660695,  
0.041094307,  
-0.029574584,  
0.017772948,  
-0.03189445,  
-0.018271524,  
0.016013889,  
0.008760028,

```
0.003604511,  
0.026603373,  
-0.014957636,  
-0.0071086627,  
-0.019532664,  
0.050095722,  
-0.038911674,  
-6.538462e-05,  
-0.029691767,  
0.010813674,  
0.022713948,  
0.013229566,  
0.0034845502,  
-0.016643835,  
0.020702068,  
0.0324832,  
0.0014380844,  
-0.010250552,  
0.056960806,  
-0.030362198,  
-0.014216848,  
-0.036642734,  
-0.027183393,  
0.014595255,  
-0.025707299,  
-0.026838781,  
0.0043701734,  
0.042175166,  
0.0060166903,  
0.019694697,  
0.015346919,  
0.004055402]
```

#### 1.7.4 Document embeddings

After creating the query embeddings, you will be guided on how to create embeddings from documents, which are a list of text chunks.

Use `embed_documents`. The parameter `chunks` should be a list of text. Here, `chunks` is a list of documents you get from before after splitting the whole document.

```
[17]: doc_result = watsonx_embedding.embed_documents(chunks)
```

As each piece of text is embedded into a vector, so the length of the `doc_result` should be the same as the length of `chunks`.

```
[18]: len(doc_result)
```

```
[18]: 570
```

Now, take a look at the first five results from the embeddings of the first piece of text.

```
[19]: doc_result[0][:5]
```

```
[19]: [-0.009708214, -0.0016381955, -0.01340118, -0.005656508, -0.031278424]
```

Check the embedding dimension to see if it is also 768.

```
[20]: len(doc_result[0])
```

```
[20]: 768
```

## 1.8 Hugging Face embedding model

### 1.8.1 Model description

In this section, you will use the `all-mpnet-base-v2` from HuggingFace as an example embedding model.

It is a sentence-transformers model. It maps sentences and paragraphs to a 768-dimensional dense vector space and can be used for tasks like clustering or semantic search. It used the pre-trained Microsoft/money-base model and fine-tuned it on a 1B sentence pairs dataset. For more information, please refer to [here](#).

### 1.8.2 Build model

To build the model, you need to import the `HuggingFaceEmbeddings` dependence first.

```
[21]: from langchain_community.embeddings import HuggingFaceEmbeddings
```

Then, you specify the model name.

```
[22]: model_name = "sentence-transformers/all-mpnet-base-v2"
```

Here we create a embedding model object.

```
[23]: huggingface_embedding = HuggingFaceEmbeddings(model_name=model_name)
```

```
/tmp/ipykernel_453/103517808.py:1: LangChainDeprecationWarning: The class
`HuggingFaceEmbeddings` was deprecated in LangChain 0.2.2 and will be removed in
0.3.0. An updated version of the class exists in the langchain-huggingface
package and should be used instead. To use it run `pip install -U langchain-
huggingface` and import as `from langchain_huggingface import
HuggingFaceEmbeddings`.
```

```
huggingface_embedding = HuggingFaceEmbeddings(model_name=model_name)
```

```
modules.json: 0%|          | 0.00/349 [00:00<?, ?B/s]
```

```
config_sentence_transformers.json: 0%|          | 0.00/116 [00:00<?, ?B/s]
```

```
README.md: 0%|          | 0.00/10.4k [00:00<?, ?B/s]
```

```
sentence_bert_config.json: 0%|          | 0.00/53.0 [00:00<?, ?B/s]
```

```
config.json: 0%|          | 0.00/571 [00:00<?, ?B/s]
model.safetensors: 0%|          | 0.00/438M [00:00<?, ?B/s]
tokenizer_config.json: 0%|          | 0.00/363 [00:00<?, ?B/s]
vocab.txt: 0%|          | 0.00/232k [00:00<?, ?B/s]
tokenizer.json: 0%|          | 0.00/466k [00:00<?, ?B/s]
special_tokens_map.json: 0%|          | 0.00/239 [00:00<?, ?B/s]
config.json: 0%|          | 0.00/190 [00:00<?, ?B/s]
```

### 1.8.3 Query embeddings

Let's create the embeddings from the same sentence, but using the Hugging Face embedding model.

```
[24]: query = "How are you?"
```

```
[25]: query_result = huggingface_embedding.embed_query(query)
```

```
[26]: query_result[:5]
```

```
[26]: [0.02710619755089283,
       0.011331832036376,
       -0.0019523875089362264,
       -0.036951325833797455,
       0.01776488684117794]
```

Do you see the differences between embeddings that are created by the watsonx embedding model and the Hugging Face embedding model?

### 1.8.4 Document embeddings

Next, you can do the same for creating embeddings from documents.

```
[27]: doc_result = huggingface_embedding.embed_documents(chunks)
      doc_result[0][:5]
```

```
[27]: [0.04028427228331566,
       0.030196484178304672,
       -0.006393850315362215,
       0.06266555190086365,
       -0.007770839612931013]
```

```
[28]: len(doc_result[0])
```

```
[28]: 768
```

## 2 Exercises

### 2.0.1 Exercise 1 - Using another watsonx embedding model

Watsonx.ai also supports other embedding models, for which you can find more information [here](#). Can you try to use another embedding model to create embeddings for the document?

```
[29]: from ibm_watsonx_ai.metanames import EmbedTextParamsMetaNames
      from langchain_ibm import WatsonxEmbeddings

      embed_params = {
          EmbedTextParamsMetaNames.TRUNCATE_INPUT_TOKENS: 3,
          EmbedTextParamsMetaNames.RETURN_OPTIONS: {"input_text": True},
      }

      watsonx_embedding = WatsonxEmbeddings(
          model_id="ibm/slate-30m-english-rtrvr",
          url="https://us-south.ml.cloud.ibm.com",
          project_id="skills-network",
          params=embed_params,
      )

      doc_result = watsonx_embedding.embed_documents(chunks)

      doc_result[0][:5]
```

```
[29]: [0.020980379, 0.0008661887, 0.018803626, 0.028992167, -0.0032073553]
```

Click here for Solution

```
from ibm_watsonx_ai.metanames import EmbedTextParamsMetaNames
from langchain_ibm import WatsonxEmbeddings

embed_params = {
    EmbedTextParamsMetaNames.TRUNCATE_INPUT_TOKENS: 3,
    EmbedTextParamsMetaNames.RETURN_OPTIONS: {"input_text": True},
}

watsonx_embedding = WatsonxEmbeddings(
    model_id="ibm/slate-30m-english-rtrvr",
    url="https://us-south.ml.cloud.ibm.com",
    project_id="skills-network",
    params=embed_params,
)

doc_result = watsonx_embedding.embed_documents(chunks)

doc_result[0][:5]
```



## 2.1 Authors

### Kang Wang

Kang Wang is a Data Scientist in IBM. He is also a PhD Candidate in the University of Waterloo.

#### 2.1.1 Other Contributors

### Joseph Santarcangelo

Joseph has a Ph.D. in Electrical Engineering, his research focused on using machine learning, signal processing, and computer vision to determine how videos impact human cognition. Joseph has been working for IBM since he completed his PhD.

### Hailey Quach

Hailey is a Data Scientist Intern at IBM. She is also pursuing a BSc in Computer Science, Honors at Concordia University, Montreal.

{## Change Log}

{|Date (YYYY-MM-DD)|Version|Changed By|Change Description||-|-|-||2024-07-22|0.1|Kang Wang|Create the lab|}

Copyright © IBM Corporation. All rights reserved.