

```

# Change to runtime to GPU-T4

# DONOT CHANGE THIS CODE
# Mount your drive

import os
import gdown
import subprocess
from google.colab import drive
drive.mount('/content/drive', force_remount=False)

FILE_ID="1wc6HtzE8JSyrwmQoNKXwXJ8-MUr7Azn9"
Embeddings = "1018_d9AcFdGk4bAhacKjsHPY2dcqxxBT"
# Embeddings = "1PfxhXhvuzlTLNA81u675yYJsy_qIsY4_"
FILE_NAME="gemma-transformers-1.1-2b-it-v1.tar"

# model_name="gemma-transformers-1.1-2b-it-v1"

# if not
os.path.exists(os.path.join("/content/drive/MyDrive/",model_name)):
#     os.mkdir(os.path.join("/content/drive/MyDrive/",model_name))

#     #Download File
#     gdown.download(f"https://drive.google.com/uc?id={FILE_ID}",
os.path.join("/content/drive/MyDrive/",model_name,FILE_NAME),
quiet=False)
#     gdown.download(f"https://drive.google.com/uc?id={FILE_ID}",
os.path.join("/content/drive/MyDrive/",model_name,Embeddings),
quiet=False)

model_name="gemma-transformers-1.1-2b-it-v1"

GOOGLE_DRIVE_PATH = os.path.join("/content/drive/MyDrive/",model_name)
REPO_ID = "MangalamSahai/24789Hws"
if not
os.path.exists(os.path.join("/content/drive/MyDrive/",model_name)):
    os.mkdir(os.path.join("/content/drive/MyDrive/",model_name))

from huggingface_hub import snapshot_download

snapshot_download(repo_id=REPO_ID, repo_type="dataset", local_dir =
GOOGLE_DRIVE_PATH)

# untar the file
if not "tokenizer_config.json" in
os.listdir(os.path.join("/content/drive/MyDrive/",model_name)):
    print("Untaring model tar File")
    subprocess.run(['tar', '-xvf',
os.path.join("/content/drive/MyDrive/",model_name,FILE_NAME), '-C',
os.path.join("/content/drive/MyDrive/",model_name)])

```

Mounted at /content/drive

```
/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94: UserWarning:  
The secret `HF_TOKEN` does not exist in your Colab secrets.  
To authenticate with the Hugging Face Hub, create a token in your  
settings tab (https://huggingface.co/settings/tokens), set it as  
secret in your Google Colab and restart your session.  
You will be able to reuse this secret in all of your notebooks.  
Please note that authentication is recommended but still optional to  
access public models or datasets.  
warnings.warn(  
    {"model_id": "dd4082e94b8c4efaaf1b92cefdad2b19", "version_major": 2, "version_minor": 0}
```

DONOT CHANGE THE CODE

```
import os  
  
if "COLAB_GPU" in os.environ:  
    print("[INFO] Running in Google Colab, installing requirements.")  
    # !pip install -U torch # requires torch 2.1.1+ (for efficient  
    sdpa implementation)  
    !pip install PyMuPDF # for reading PDFs with Python  
    !pip install tqdm # for progress bars  
    !pip install sentence-transformers # for embedding models  
    !pip install accelerate # for quantization model loading  
    !pip install bitsandbytes # for quantizing models (less storage  
    space)  
    # !pip install flash-attn --no-build-isolation # for faster  
    attention mechanism = faster LLM inference
```

[INFO] Running in Google Colab, installing requirements.

Collecting PyMuPDF

Downloading pymupdf-1.25.5-cp39-abi3-manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (3.4 kB)

Downloading pymupdf-1.25.5-cp39-abi3-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (20.0 MB)

20.0/20.0 MB 76.1 MB/s eta

0:00:00

Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-packages (4.67.1)

Requirement already satisfied: sentence-transformers in /usr/local/lib/python3.11/dist-packages (3.4.1)

Requirement already satisfied: transformers<5.0.0,>=4.41.0 in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (4.51.3)

Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (4.67.1)

```
Requirement already satisfied: torch>=1.11.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(2.6.0+cu124)
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.6.1)
Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.15.2)
Requirement already satisfied: huggingface-hub>=0.20.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(0.30.2)
Requirement already satisfied: Pillow in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(11.2.1)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (3.18.0)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (2025.3.2)
Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (24.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (6.0.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (2.32.3)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0-
>sentence-transformers) (4.13.2)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (3.1.6)
Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch>=1.11.0-
>sentence-transformers)
  Downloading nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-runtime-cu12==12.4.127 (from torch>=1.11.0-
>sentence-transformers)
  Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-cupti-cu12==12.4.127 (from torch>=1.11.0-
>sentence-transformers)
```

```
Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-  
manylinux2014_x86_64.whl.metadata (1.6 kB)  
Collecting nvidia-cudnn-cu12==9.1.0.70 (from torch>=1.11.0->sentence-  
transformers)  
Downloading nvidia_cudnn_cu12-9.1.0.70-py3-none-  
manylinux2014_x86_64.whl.metadata (1.6 kB)  
Collecting nvidia-cublas-cu12==12.4.5.8 (from torch>=1.11.0->sentence-  
transformers)  
Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-  
manylinux2014_x86_64.whl.metadata (1.5 kB)  
Collecting nvidia-cufft-cu12==11.2.1.3 (from torch>=1.11.0->sentence-  
transformers)  
Downloading nvidia_cufft_cu12-11.2.1.3-py3-none-  
manylinux2014_x86_64.whl.metadata (1.5 kB)  
Collecting nvidia-curand-cu12==10.3.5.147 (from torch>=1.11.0-  
>sentence-transformers)  
Downloading nvidia_curand_cu12-10.3.5.147-py3-none-  
manylinux2014_x86_64.whl.metadata (1.5 kB)  
Collecting nvidia-cusolver-cu12==11.6.1.9 (from torch>=1.11.0-  
>sentence-transformers)  
Downloading nvidia_cusolver_cu12-11.6.1.9-py3-none-  
manylinux2014_x86_64.whl.metadata (1.6 kB)  
Collecting nvidia-cuspars-cu12==12.3.1.170 (from torch>=1.11.0-  
>sentence-transformers)  
Downloading nvidia_cuspars-cu12-12.3.1.170-py3-none-  
manylinux2014_x86_64.whl.metadata (1.6 kB)  
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.2 in  
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-  
transformers) (0.6.2)  
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in  
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-  
transformers) (2.21.5)  
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in  
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-  
transformers) (12.4.127)  
Collecting nvidia-nvjitlink-cu12==12.4.127 (from torch>=1.11.0-  
>sentence-transformers)  
Downloading nvidia_nvjitlink_cu12-12.4.127-py3-none-  
manylinux2014_x86_64.whl.metadata (1.5 kB)  
Requirement already satisfied: triton==3.2.0 in  
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-  
transformers) (3.2.0)  
Requirement already satisfied: sympy==1.13.1 in  
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-  
transformers) (1.13.1)  
Requirement already satisfied: mpmath<1.4,>=1.1.0 in  
/usr/local/lib/python3.11/dist-packages (from sympy==1.13.1-  
>torch>=1.11.0->sentence-transformers) (1.3.0)  
Requirement already satisfied: numpy>=1.17 in
```

```

/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (2.0.2)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (2024.11.6)
Requirement already satisfied: tokenizers<0.22,>=0.21 in
/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (0.21.1)
Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (0.5.3)
Requirement already satisfied: joblib>=1.2.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers) (1.4.2)
Requirement already satisfied: threadpoolctl>=3.1.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers) (3.6.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.11.0-
>sentence-transformers) (3.0.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.20.0->sentence-transformers) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.20.0->sentence-transformers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.20.0->sentence-transformers) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.20.0->sentence-transformers) (2025.4.26)
Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-
manylinux2014_x86_64.whl (363.4 MB)
----- 363.4/363.4 MB 4.5 MB/s eta
0:00:00
anylinux2014_x86_64.whl (13.8 MB)
----- 13.8/13.8 MB 118.6 MB/s eta
0:00:00
anylinux2014_x86_64.whl (24.6 MB)
----- 24.6/24.6 MB 91.1 MB/s eta
0:00:00
e_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (883 kB)
----- 883.7/883.7 kB 62.8 MB/s eta
0:00:00
anylinux2014_x86_64.whl (664.8 MB)
----- 664.8/664.8 MB 2.4 MB/s eta
0:00:00
anylinux2014_x86_64.whl (211.5 MB)

```

```

211.5/211.5 MB 5.8 MB/s eta
0:00:00
anylinux2014_x86_64.whl (56.3 MB)
56.3/56.3 MB 13.0 MB/s eta
0:00:00
anylinux2014_x86_64.whl (127.9 MB)
127.9/127.9 MB 7.6 MB/s eta
0:00:00
anylinux2014_x86_64.whl (207.5 MB)
207.5/207.5 MB 5.8 MB/s eta
0:00:00
anylinux2014_x86_64.whl (21.1 MB)
21.1/21.1 MB 98.8 MB/s eta
0:00:00
e-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-cupti-cu12, nvidia-cublas-
cu12, nvidia-cusparse-cu12, nvidia-cudnn-cu12, nvidia-cusolver-cu12
Attempting uninstall: nvidia-nvjitlink-cu12
Found existing installation: nvidia-nvjitlink-cu12 12.5.82
Uninstalling nvidia-nvjitlink-cu12-12.5.82:
Successfully uninstalled nvidia-nvjitlink-cu12-12.5.82
Attempting uninstall: nvidia-curand-cu12
Found existing installation: nvidia-curand-cu12 10.3.6.82
Uninstalling nvidia-curand-cu12-10.3.6.82:
Successfully uninstalled nvidia-curand-cu12-10.3.6.82
Attempting uninstall: nvidia-cufft-cu12
Found existing installation: nvidia-cufft-cu12 11.2.3.61
Uninstalling nvidia-cufft-cu12-11.2.3.61:
Successfully uninstalled nvidia-cufft-cu12-11.2.3.61
Attempting uninstall: nvidia-cuda-runtime-cu12
Found existing installation: nvidia-cuda-runtime-cu12 12.5.82
Uninstalling nvidia-cuda-runtime-cu12-12.5.82:
Successfully uninstalled nvidia-cuda-runtime-cu12-12.5.82
Attempting uninstall: nvidia-cuda-nvrtc-cu12
Found existing installation: nvidia-cuda-nvrtc-cu12 12.5.82
Uninstalling nvidia-cuda-nvrtc-cu12-12.5.82:
Successfully uninstalled nvidia-cuda-nvrtc-cu12-12.5.82
Attempting uninstall: nvidia-cuda-cupti-cu12
Found existing installation: nvidia-cuda-cupti-cu12 12.5.82
Uninstalling nvidia-cuda-cupti-cu12-12.5.82:
Successfully uninstalled nvidia-cuda-cupti-cu12-12.5.82
Attempting uninstall: nvidia-cublas-cu12
Found existing installation: nvidia-cublas-cu12 12.5.3.2
Uninstalling nvidia-cublas-cu12-12.5.3.2:
Successfully uninstalled nvidia-cublas-cu12-12.5.3.2
Attempting uninstall: nvidia-cusparse-cu12
Found existing installation: nvidia-cusparse-cu12 12.5.1.3
Uninstalling nvidia-cusparse-cu12-12.5.1.3:
Successfully uninstalled nvidia-cusparse-cu12-12.5.1.3
Attempting uninstall: nvidia-cudnn-cu12

```

```
Found existing installation: nvidia-cudnn-cu12 9.3.0.75
Uninstalling nvidia-cudnn-cu12-9.3.0.75:
  Successfully uninstalled nvidia-cudnn-cu12-9.3.0.75
Attempting uninstall: nvidia-cusolver-cu12
Found existing installation: nvidia-cusolver-cu12 11.6.3.83
Uninstalling nvidia-cusolver-cu12-11.6.3.83:
  Successfully uninstalled nvidia-cusolver-cu12-11.6.3.83
Successfully installed nvidia-cublas-cu12-12.4.5.8 nvidia-cuda-cupti-
cu12-12.4.127 nvidia-cuda-nvrtc-cu12-12.4.127 nvidia-cuda-runtime-
cu12-12.4.127 nvidia-cudnn-cu12-9.1.0.70 nvidia-cufft-cu12-11.2.1.3
nvidia-curand-cu12-10.3.5.147 nvidia-cusolver-cu12-11.6.1.9 nvidia-
cusparse-cu12-12.3.1.170 nvidia-nvjitlink-cu12-12.4.127
Requirement already satisfied: accelerate in
/usr/local/lib/python3.11/dist-packages (1.6.0)
Requirement already satisfied: numpy<3.0.0,>=1.17 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (2.0.2)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (24.2)
Requirement already satisfied: psutil in
/usr/local/lib/python3.11/dist-packages (from accelerate) (5.9.5)
Requirement already satisfied: pyyaml in
/usr/local/lib/python3.11/dist-packages (from accelerate) (6.0.2)
Requirement already satisfied: torch>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate)
(2.6.0+cu124)
Requirement already satisfied: huggingface-hub>=0.21.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (0.30.2)
Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (0.5.3)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0-
>accelerate) (3.18.0)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0-
>accelerate) (2025.3.2)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0-
>accelerate) (2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0-
>accelerate) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0-
>accelerate) (4.13.2)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
```

```
>accelerate) (3.1.6)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.4.127)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127
in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.4.127)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.4.127)
Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (9.1.0.70)
Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.4.5.8)
Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (11.2.1.3)
Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (10.3.5.147)
Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (11.6.1.9)
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.2 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (0.6.2)
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (2.21.5)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.4.127)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.4.127)
Requirement already satisfied: triton==3.2.0 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (3.2.0)
Requirement already satisfied: sympy==1.13.1 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy==1.13.1-
>torch>=2.0.0->accelerate) (1.3.0)
```


Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=2.0.0->accelerate) (3.0.2)

Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.21.0->accelerate) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.21.0->accelerate) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.21.0->accelerate) (2.4.0)

Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.21.0->accelerate) (2025.4.26)

Collecting bitsandbytes

Downloading bitsandbytes-0.45.5-py3-none-manylinux_2_24_x86_64.whl.metadata (5.0 kB)

Requirement already satisfied: torch<3,>=2.0 in
/usr/local/lib/python3.11/dist-packages (from bitsandbytes) (2.6.0+cu124)

Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from bitsandbytes) (2.0.2)

Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0->bitsandbytes) (3.18.0)

Requirement already satisfied: typing-extensions>=4.10.0 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0->bitsandbytes) (4.13.2)

Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0->bitsandbytes) (3.4.2)

Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0->bitsandbytes) (3.1.6)

Requirement already satisfied: fsspec in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0->bitsandbytes) (2025.3.2)

Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0->bitsandbytes) (12.4.127)

Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0->bitsandbytes) (12.4.127)

Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0->bitsandbytes) (12.4.127)

Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-

```

>bitsandbytes) (9.1.0.70)
Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (12.4.5.8)
Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (11.2.1.3)
Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (10.3.5.147)
Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (11.6.1.9)
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.2 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (0.6.2)
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (2.21.5)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (12.4.127)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (12.4.127)
Requirement already satisfied: triton==3.2.0 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (3.2.0)
Requirement already satisfied: sympy==1.13.1 in
/usr/local/lib/python3.11/dist-packages (from torch<3,>=2.0-
>bitsandbytes) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy==1.13.1-
>torch<3,>=2.0->bitsandbytes) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch<3,>=2.0-
>bitsandbytes) (3.0.2)
Downloading bitsandbytes-0.45.5-py3-none-manylinux_2_24_x86_64.whl
(76.1 MB)
76.1/76.1 MB 9.8 MB/s eta
0:00:00

```

##To understand how we obtained embedding file. Please refer to the recitation.

```

## DONOT CHANGE THE CODE
import pandas as pd

```

```

# Import saved file and view
embeddings_df_save_path =
os.path.join("/content/drive/MyDrive/",model_name,"text_chunks_and_embeddings_df.csv")
text_chunks_and_embedding_df_load =
pd.read_csv(embeddings_df_save_path)
text_chunks_and_embedding_df_load.head()

{"summary":{"\n  \"name\": \"text_chunks_and_embedding_df_load\",\n  \"rows\": 1680,\n  \"fields\": [\n    {\n      \"column\": \"page_number\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 349,\n        \"min\": -39,\n        \"max\": 1166,\n        \"num_unique_values\": 1136,\n        \"samples\": [\n          795,\n          918,\n          265\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"sentence_chunk\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 1680,\n        \"samples\": [\n          \"The major determinants of Type 2 diabetes that can be changed are overnutrition and a sedentary lifestyle. Therefore, reversing or improving these factors by lifestyle interventions markedly improve the overall health of Type 2 diabetics and lower blood-glucose levels. In fact it has been shown that when people are overweight, losing as little as nine pounds (four kilograms) decreases blood- glucose levels in Type 2 diabetics. The Diabetes Prevention Trial demonstrated that by adhering to a diet containing between 1,200 and 1,800 kilocalories per day with a dietary fat intake goal of less than 25 percent and increasing physical activity to at least 150 minutes per week, people at high risk for Type 2 diabetes achieved a weight loss of 7 percent and significantly decreased their chances of developing Type 2 diabetes.15 The American Diabetes Association (ADA) has a website that provides information and tips for helping diabetics answer the question, \"What Can I Eat\". In regard to carbohydrates the ADA recommends diabetics keep track of the carbohydrates they eat and set a limit. These dietary practices will help keep blood-glucose levels in the target range. Figure 18.5 Metabolic Syndrome: A Combination of Risk Factors Increasing the Chances for Chronic Disease 15. Knowler WC. (2002). Reduction in the Incidence of Type 2 Diabetes with Lifestyle Intervention or Metformin.\",\n          \"Scheme of a micelle formed by phospholipid s in an aqueous solution by Emmanuel Boutet /CC BY-SA 3.0 cholesterol so it acts as an emulsifier. It attracts and holds onto fat while it is simultaneously attracted to and held on to by water. Emulsification increases the surface area of lipids over a thousand- fold, making them more accessible to the digestive enzymes. Once the stomach contents have been emulsified, fat-breaking enzymes work on the triglycerides and diglycerides to sever fatty acids from their glycerol foundations. As pancreatic lipase enters the small intestine, it breaks down the fats into free fatty acids and monoglycerides. Yet again, another hurdle presents itself. How will the fats pass through the watery layer of
        ]
      }
    }
  ]
}}

```

mucus that coats the absorptive lining of the digestive tract?As before, the answer is bile. Bile salts envelop the fatty acids and monoglycerides to form micelles. Micelles have a fatty acid core with a water-soluble exterior.

Image by CC BY-SA 3.0 dense cortical bone is about 10 percent porous and it looks like many concentric circles, similar to the rings in a tree trunk, sandwiched together (Figure 2.27 Cortical (Compact) Bone). Cortical bone tissue makes up approximately 80 percent of the adult skeleton. It surrounds all trabecular tissue and is the only bone tissue in the shafts of long bones. Figure 2.26 The Arrangement of Bone Tissues The two basic tissue types of bones are trabecular and cortical. This photo shows normal (left) and degraded (right) trabecular (spongy) bone. Figure 2.27 Cortical (Compact) Bone.

The Skeletal System | 123

```

{
  "description": "",
  "column": "chunk_char_count",
  "dtype": "number",
  "std": 415,
  "min": 121,
  "max": 1831,
  "num_unique_values": 992,
  "samples": [
    421, 777, 1617
  ],
  "semantic_type": "",
  "description": "",
  "column": "chunk_word_count",
  "dtype": "number",
  "std": 66,
  "min": 9,
  "max": 297,
  "num_unique_values": 257,
  "samples": [
    227, 276, 73
  ],
  "semantic_type": "",
  "description": "",
  "column": "chunk_token_count",
  "dtype": "number",
  "std": 103.84074677997317,
  "min": 30.25,
  "max": 457.75,
  "num_unique_values": 992,
  "samples": [
    105.25, 194.25, 404.25
  ],
  "semantic_type": "",
  "description": "",
  "column": "embedding",
  "dtype": "string",
  "num_unique_values": 1680,
  "samples": [
    [
      2.29729712e-02, 2.80858967e-02,
      1.12715000e-02, 2.38459539e-02,
      3.04390900e-02, 7.16837496e-02,
      3.81384008e-02, -1.46272471e-02,
      2.14921068e-02, -2.19398756e-02,
      2.36242507e-02, 1.27335191e-02,
      2.97014453e-02, 1.34395417e-02,
      8.38490427e-02, 4.68919054e-03,
      6.83111176e-02, 2.30760835e-02,
      2.05456968e-06, -3.15726995e-02,
      9.00705997e-03, 4.50910553e-02,
      2.80213878e-02, 3.24000865e-02,
      2.20614839e-02, 3.33284363e-02,
      9.21218190e-03, -9.00363252e-02,
      1.24296751e-02, 2.76750606e-02
    ]
  ]
}

```

6.08849078e-02 -2.20196079e-02 -
3.34621929e-02 2.56895032e-02
6.05134619e-03 3.96036990e-02 -
3.65133472e-02 -3.25683095e-02 -
-2.04192102e-02 -7.57569866e-03 -
-3.89803052e-02 7.09829107e-03
-6.21600682e-03 -3.64423431e-02
1.70159172e-02 -3.24641615e-02
-4.76174010e-03 2.94138342e-02
1.82500333e-02 -2.45281626e-02
1.67630445e-02 -2.78206356e-02
2.97171902e-02 2.26193946e-02
1.28435260e-02 -2.76685599e-02 -
-7.83731788e-03 5.50725460e-02 -

3.28327604e-02	-9.84455738e-03\\n	4.40649912e-02	-5.17247468e-02 -
2.61238627e-02	2.34586280e-02\\n	-3.23815495e-02	1.43910442e-02 -
2.31870990e-02	-6.13956004e-02\\n	3.56797464e-02	4.42797542e-02
7.54781887e-02	-1.18830567e-02\\n	2.61459257e-02	2.20207274e-02
1.07243182e-02	9.09299869e-03\\n	2.32546125e-02	-7.01869577e-02 -
4.57891934e-02	2.91911066e-02\\n	-1.55159086e-02	1.12296343e-02
1.60889253e-02	-2.10474096e-02\\n	1.33295916e-03	-1.47659099e-02 -
1.23751191e-02	2.28247084e-02\\n	1.66185517e-02	7.55396718e-03 -
2.91870032e-02	2.15626378e-02\\n	-4.81806546e-02	2.03475356e-02
2.55755950e-02	-1.04519501e-02\\n	2.39592195e-02	7.22793639e-02
2.07349043e-02	-5.32777049e-02\\n	4.41575311e-02	7.44044557e-02
8.33298713e-02	-2.73018633e-03\\n	2.66782939e-03	1.34550482e-02 -
3.98521274e-02	1.67210482e-03\\n	-3.05837709e-02	4.47130464e-02 -
2.64749187e-03	-1.56471636e-02\\n	-7.53284320e-02	-2.73020063e-02
4.55050021e-02	-1.80287682e-03\\n	5.39066736e-03	-4.53817174e-02 -
4.69642878e-02	-3.81236225e-02\\n	-1.64280459e-02	8.61568097e-03
3.98035981e-02	-1.25875622e-02\\n	-1.80445574e-02	1.02986321e-02
4.12722863e-02	-1.31619982e-02\\n	4.11747545e-02	6.24924619e-03
4.69480641e-02	-1.32724438e-02\\n	-3.39613925e-03	8.27880949e-03
6.28442178e-03	5.81035987e-02\\n	-1.29940435e-02	-3.89739405e-03
2.20918935e-02	7.65071576e-03\\n	-5.82268909e-02	5.13123441e-03 -
2.03677192e-02	-6.17644452e-02\\n	5.85707976e-03	-2.43919939e-02
3.34543549e-02	3.82698886e-02\\n	3.51353991e-03	-1.31548736e-02 -
1.13629512e-02	-1.27255134e-02\\n	-2.48163124e-03	1.72944702e-02
8.27016830e-02	-3.42380162e-03\\n	-5.70958778e-02	2.50502699e-03
3.03654727e-02	9.44623444e-03\\n	1.72372162e-02	7.99587555e-03 -
2.43335757e-02	2.92636547e-02\\n	6.21003769e-02	7.02438236e-04
4.73045446e-02	-2.81685181e-02\\n	3.49310180e-03	5.04241772e-02
1.11960378e-02	3.41657214e-02\\n	-4.17057648e-02	5.92101254e-02
2.73968335e-02	-3.56948450e-02\\n	3.93448658e-02	3.51529308e-02
8.42085201e-03	-1.98321957e-02\\n	4.33469713e-02	-1.18261045e-02
4.92545627e-02	-3.05400956e-02\\n	4.26846975e-03	4.03070636e-03 -
6.49488345e-02	-7.95924570e-04\\n	-3.46954390e-02	-9.99789964e-03
1.64243709e-02	-1.20536881e-02\\n	7.07974732e-02	-9.05370936e-02
2.73720417e-02	2.17130873e-02\\n	-2.33670045e-02	9.12264287e-02
1.64063182e-02	3.62658389e-02\\n	1.47561124e-02	-6.63467124e-03 -
5.39022982e-02	-6.76799789e-02\\n	4.55670338e-03	-2.33429801e-02
4.05053683e-02	-1.99937429e-02\\n	3.04897185e-02	-5.00026420e-02
4.43863347e-02	-2.04093903e-02\\n	6.49715960e-02	-4.67289835e-02
3.04762907e-02	-4.17238474e-02\\n	1.36049623e-02	-1.18888151e-02
1.12099862e-02	6.29553497e-02\\n	-1.77913252e-02	4.15676348e-02 -
1.00580053e-02	6.75371150e-04\\n	2.51111519e-02	-1.26866316e-02
1.68014988e-02	-3.28931436e-02\\n	-4.90503013e-02	-5.12543647e-03
5.16201593e-02	2.13184338e-02\\n	-6.63907155e-02	2.90889200e-02 -
7.74796936e-04	-3.93134579e-02\\n	-6.18337747e-03	1.16510903e-02 -
3.34364139e-02	-2.28617378e-02\\n	-9.15383454e-03	-3.65933105e-02 -
3.25162262e-02	-1.53120402e-02\\n	-9.30079743e-02	-3.57470699e-02 -
7.29239685e-03	4.26715985e-02\\n	-4.41388600e-02	6.41513467e-02 -
2.29487196e-02	-5.17616123e-02\\n	1.37782358e-02	2.89340112e-02 -

7.61773959e-02	1.12729799e-02\\n	-7.52168596e-02	5.26861101e-03
1.53507972e-02	4.68976758e-02\\n	-1.34887900e-02	-8.26931838e-03
5.36376378e-03	4.95160818e-02\\n	-1.30892098e-02	-4.93161492e-02
8.89857020e-03	2.88364477e-02\\n	-4.04627621e-02	-5.19260094e-02
6.67705247e-03	-2.46643461e-03\\n	2.24981792e-02	2.95400154e-02 -
5.24829235e-03	1.17631555e-02\\n	3.35198268e-02	-4.02946165e-03
1.56868231e-02	2.04602368e-02\\n	8.72190893e-02	6.44629896e-02
4.14886996e-02	1.39544327e-02\\n	3.64198275e-02	-4.75281104e-02 -
1.10576088e-02	-6.18624985e-02\\n	1.04863308e-02	1.40444888e-03
5.39735444e-02	-1.78427231e-02\\n	4.31339396e-03	6.45331889e-02 -
4.83200699e-02	7.58985654e-02\\n	1.35629280e-02	-4.18066978e-02
1.54282255e-02	-2.14080475e-02\\n	2.15211418e-02	3.94406319e-02 -
1.00536551e-02	6.46505458e-03\\n	-1.72794964e-02	-1.12749003e-02 -
2.73868050e-02	-3.50129902e-02\\n	8.66162498e-03	-6.33235043e-03
4.85603251e-02	3.74994241e-02\\n	2.12799814e-02	-2.77581662e-02 -
3.15048848e-03	-5.67035610e-03\\n	-4.97750714e-02	1.61907412e-02 -
3.45674045e-02	2.42027901e-02\\n	-6.08441420e-04	-1.02716545e-02 -
4.54923660e-02	-1.02981711e-02\\n	1.40911583e-02	3.75180282e-02
1.68752670e-02	2.34776717e-02\\n	-6.18587136e-02	3.49496827e-02
2.63891201e-02	3.26851308e-02\\n	1.95784457e-02	4.40052897e-03 -
2.89963484e-02	-6.47588400e-03\\n	3.63212861e-02	7.43830502e-02 -
2.42742654e-02	5.18779270e-02\\n	-7.51687139e-02	-3.40245292e-02
1.70769282e-02	8.19654763e-02\\n	-1.98336337e-02	-1.83938295e-02 -
5.01794182e-02	2.75631640e-02\\n	-4.24477682e-02	2.42094826e-02 -
2.18182877e-02	1.33800218e-02\\n	3.36719789e-02	1.84938870e-03
1.08133173e-02	-6.36373088e-02\\n	4.44507152e-02	1.21700894e-02 -
2.84682978e-02	-1.32247293e-02\\n	2.05366500e-02	4.44630794e-02 -
6.56913072e-02	-1.55691383e-02\\n	1.03134699e-02	-3.58045660e-02 -
1.38292583e-02	4.76093061e-04\\n	1.26236770e-02	1.15177697e-02 -
8.35120492e-03	3.91691998e-02\\n	3.29118059e-03	1.52778924e-02 -
4.11466584e-02	-2.41303574e-02\\n	2.96915621e-02	1.58117022e-02 -
5.56087382e-02	-6.59467326e-03\\n	3.99093628e-02	-9.33313295e-02
1.72184650e-02	3.36729432e-03\\n	-1.26988860e-02	2.48798970e-02 -
8.33463483e-03	-2.17450745e-02\\n	5.75357601e-02	-3.72378267e-02
2.99197137e-02	-5.25633106e-03\\n	7.22941477e-03	-2.23791022e-02
9.92902089e-03	-4.89397943e-02\\n	-3.30955125e-02	-3.47236134e-02 -
3.55034024e-02	4.07063104e-02\\n	2.99441069e-02	-1.32662177e-01
3.83205079e-02	-2.97528952e-02\\n	1.57118067e-02	1.49636613e-02
3.05667035e-02	5.80923259e-03\\n	4.45674136e-02	8.68415609e-02
3.00537106e-02	-3.14172730e-02\\n	-1.12733599e-02	-6.85861632e-02
8.29143077e-02	9.32687253e-04\\n	8.28287005e-03	4.60251383e-02
5.92524000e-02	-8.62023327e-03\\n	-1.28702521e-02	-9.42938589e-03
2.76092310e-02	4.46677580e-02\\n	-6.61543086e-02	-4.07488719e-02 -
2.03002933e-02	2.83709373e-02\\n	-2.50594653e-02	1.89823210e-02
6.60287142e-02	6.14814945e-02\\n	-1.37754623e-03	-3.10590887e-03
2.10915934e-02	1.09830508e-02\\n	-2.95005180e-02	-1.86231025e-02
8.57474133e-02	5.12545519e-02\\n	2.64492128e-02	-3.36265191e-02
3.48860510e-02	-2.85117216e-02\\n	7.39187151e-02	2.89257746e-02
3.45990323e-02	-1.44111356e-02\\n	6.03982285e-02	5.42171765e-03

1.08373733e-02	1.06092310e-02\\n	-6.86236396e-02	2.03288117e-04
1.60011351e-02	-8.57141428e-03\\n	-2.39313841e-02	-2.35607326e-02 -
3.61981764e-02	-4.08706516e-02\\n	1.07472815e-01	3.81626049e-03
3.19191143e-02	-1.63456071e-02\\n	-1.62466131e-02	-6.20278204e-03 -
5.22550605e-02	5.99559210e-02\\n	-5.44447228e-02	-1.31396269e-02
4.59873267e-02	4.86567691e-02\\n	2.39443686e-02	3.04191876e-02 -
4.35936712e-02	-5.25394008e-02\\n	-7.88886696e-02	2.50433222e-03 -
3.06024663e-02	-3.53882685e-02\\n	-4.16593067e-02	2.38333158e-02 -
1.45164807e-03	-2.84110345e-02\\n	-2.23766919e-02	3.99736091e-02 -
1.79760810e-02	-2.58030556e-02\\n	-2.83792131e-02	1.75174233e-02 -
3.72972079e-02	-2.12674402e-03\\n	-6.22427501e-02	-6.65185601e-02 -
2.23804079e-02	-7.89400488e-02\\n	-1.83512717e-02	5.61043695e-02
1.92721095e-02	5.66429757e-02\\n	6.06254116e-03	-2.40686052e-02
3.77018452e-02	-2.64845230e-02\\n	-3.40272952e-03	-3.50038707e-02 -
1.39413746e-02	1.73181354e-03\\n	2.22160127e-02	3.72691378e-02 -
3.77214104e-02	4.42813803e-03\\n	1.78216752e-02	-2.63886321e-02
1.44201498e-02	5.36084129e-03\\n	-1.14303408e-02	1.87312569e-02
1.34985230e-03	-4.87835286e-03\\n	-3.91572108e-03	-1.70022659e-02
5.42912353e-03	6.04774803e-02\\n	1.41849415e-02	3.82715948e-02 -
3.95067558e-02	-2.00121272e-02\\n	-2.09616944e-02	2.74659786e-02
9.52143967e-03	-6.81963516e-03\\n	-1.10616554e-02	-8.06704164e-03 -
2.63257436e-02	-1.13474466e-02\\n	4.03615721e-02	-2.57180873e-02 -
3.69606055e-02	-2.36007608e-02\\n	1.45106716e-02	-3.04713380e-02
2.86728740e-02	-2.93919444e-03\\n	-2.37204973e-03	-8.02426413e-03 -
3.76460627e-02	-2.02637110e-02\\n	2.97035687e-02	-2.50616148e-02
1.87683746e-03	-5.57896569e-02\\n	-2.73329560e-02	-3.78011167e-02
1.61891244e-02	-5.38802897e-33\\n	2.96111181e-02	-9.83624719e-03
4.98379879e-02	5.77777326e-02\\n	5.10296822e-02	1.32933445e-02 -
6.45377636e-02	-9.37822554e-03\\n	-5.93022890e-02	7.89666101e-02 -
2.94902213e-02	-2.31253002e-02\\n	2.72985399e-05	-4.59786952e-02 -
6.16562590e-02	-2.65489593e-02\\n	7.99557474e-03	1.57606276e-03 -
3.19293663e-02	-1.06759686e-02\\n	-6.44516200e-02	3.45161185e-02
8.73177964e-03	1.91062242e-02\\n	2.99177933e-02	4.72093485e-02
1.41662303e-02	-5.85156435e-04\\n	1.50941033e-03	1.84304677e-02 -
1.34501401e-02	7.77282864e-02\\n	-3.04925285e-04	4.56190482e-02
1.24297431e-02	7.88352489e-02\\n	-6.26035081e-03	-3.03326119e-02
2.22761407e-02	3.11824074e-03\\n	6.48971424e-02	-5.33699654e-02
1.14038214e-03	3.22002592e-03\\n	-8.58834088e-02	7.39832669e-02 -
2.24174149e-02	-1.98693629e-02\\n	2.02376978e-03	2.72148270e-02 -
4.08967547e-02	-6.27453765e-03\\n	2.47814786e-02	4.25336882e-02 -
2.33338978e-02	4.21236120e-02\\n	-5.35760261e-03	-1.38929300e-02
2.34857071e-02	2.73593366e-02\\n	-2.59762164e-02	3.09628043e-02 -
4.41264287e-02	3.08858417e-02\\n	-3.02760443e-03	-7.87444599e-03 -
1.53053449e-02	1.24025848e-02\\n	-1.44549301e-02	-7.57972375e-02 -
7.31532797e-02	-6.25108033e-02\\n	2.19611656e-02	4.04325128e-02 -
2.31300900e-03	-6.09265566e-02\\n	-4.90598902e-02	1.80721302e-02 -
4.54713255e-02	-4.93626110e-02\\n	-9.59483627e-03	-1.43828010e-02
1.51193840e-02	-3.79985310e-02\\n	-4.98143584e-02	7.65519217e-02 -
3.22400853e-02	-3.91259827e-02\\n	-7.93610960e-02	-3.08242645e-02

```
2.04933472e-02  9.77791287e-03\\n -4.97575093e-04  5.28453244e-03
4.71585430e-02  5.64918062e-03\\n  8.42629075e-02 -6.12572692e-02 -
3.29317059e-03  7.30867079e-03\\n -1.97428428e-02 -1.43558551e-02
1.55156280e-03  4.72506173e-02\\n -2.23656278e-03  2.66999938e-02 -
1.06265480e-02 -2.58234162e-02\\n  7.00479820e-02 -1.86981261e-02 -
3.37645672e-02  2.94382703e-02\\n -3.21478657e-02 -3.84805053e-02 -
5.57591282e-02  4.19322774e-02\\n  2.99479533e-02  8.49122629e-02
1.71478745e-02 -5.41984625e-02\\n -3.17496471e-02  3.15846279e-02
3.04570980e-02 -3.83366346e-02\\n -6.88592112e-03 -1.95200220e-02
3.14677991e-02 -1.13854895e-03\\n -7.36938119e-02 -1.74842868e-02
5.41744661e-03 -1.92077104e-02\\n  2.78682137e-07  4.08470184e-02 -
1.53628632e-03 -1.78423189e-02\\n -1.03535233e-02 -3.08404192e-02 -
6.02940517e-03  3.50857563e-02\\n -6.51571378e-02 -5.80149330e-02
4.58920673e-02  4.16008011e-02\\n -3.90204564e-02 -3.64808962e-02 -
3.20155025e-02  6.35178387e-02\\n  2.17223652e-02  1.82709042e-02 -
1.36412205e-02 -4.86772275e-03\\n -7.65325595e-03  4.14719339e-03
3.67409103e-02 -2.42217444e-02\\n -7.34255984e-02 -3.96182053e-02
4.99258470e-03  4.51597758e-03\\n -4.24852669e-02  6.33915886e-04 -
2.84216926e-02  9.66380630e-03\\n  8.96294974e-03  4.65004295e-02 -
7.67296329e-02 -2.04632115e-02\\n -7.31814653e-02 -6.51863143e-02
6.90150857e-02 -3.71697284e-02\\n  4.29911762e-02 -1.47519633e-02
1.61680002e-02 -1.68805681e-02\\n -4.27693613e-02 -3.91147472e-03 -
5.42811714e-02 -4.86369655e-02\\n  4.49501313e-02  2.64735520e-02 -
7.76932463e-02  4.56957985e-03\\n  2.51250882e-02 -5.15799671e-02
9.67912097e-03 -1.23087047e-02\\n  2.34425999e-02  5.72338793e-03
3.40899415e-02  2.72950362e-02\\n -3.42273489e-02 -3.70479841e-03 -
5.08046616e-03  3.87148373e-02\\n -8.21772218e-03 -5.54903001e-02 -
7.95374159e-04  1.96648911e-02\\n  1.80299746e-34  1.06232800e-02
7.13028107e-03 -8.23647343e-03\\n -4.72389422e-02  1.98795907e-02 -
8.45584832e-03 -1.23929149e-02\\n  1.10524462e-03 -2.05322794e-04
2.58644149e-02 -6.45051524e-02]\\",\\n          \"[-2.89910735e-04 -
4.76907492e-02  4.69394773e-02 -1.34329256e-02\\n  2.12686360e-02 -
6.53087068e-03  1.84320193e-02 -5.08894678e-03\\n  4.43609543e-02 -
2.93549653e-02  1.95518564e-02  5.42302988e-02\\n  1.03220148e-02 -
2.42334884e-02  4.03822921e-02 -3.05171981e-02\\n -2.09650956e-02
7.83721544e-03 -1.09209083e-01 -1.38158966e-02\\n -1.22020207e-02
3.17832944e-03 -2.41586682e-03  3.28726098e-02\\n  6.40644804e-02 -
3.99281606e-02 -1.83247719e-02 -3.14286165e-02\\n -4.16912213e-02 -
3.85700837e-02 -1.26829343e-02 -1.41209865e-03\\n  1.31108062e-02 -
9.94675606e-03  2.10254871e-06 -7.18540251e-02\\n -5.03428802e-02
2.97871716e-02 -1.69740599e-02 -2.29916032e-02\\n -1.15232810e-03 -
9.93641540e-02 -1.35236466e-02 -7.95061979e-03\\n  3.42773981e-02 -
5.95667362e-02 -1.08887609e-02 -1.95660256e-02\\n  7.71187916e-02
8.48048739e-03 -8.23518354e-03 -7.42352158e-02\\n  3.04884110e-02
5.15961759e-02  5.65598123e-02 -2.05483083e-02\\n -5.42607950e-03 -
2.55838782e-02 -3.91943641e-02 -3.48252021e-02\\n  1.56397428e-02
5.52787483e-02  2.22797859e-02 -2.18725502e-02\\n  1.79669056e-02
6.62144572e-02  1.44304838e-02 -3.41766737e-02\\n  3.90865318e-02 -
2.13876888e-02 -7.50386193e-02 -3.48831303e-02\\n -3.81169505e-02
```


1.10502150e-02	-4.12468053e-02	1.77212786e-02\\n	6.79142866e-03
1.09248254e-02	2.70897709e-02	1.80689208e-02\\n	6.10432960e-02
1.14617050e-02	-3.80260847e-03	-3.18495855e-02\\n	1.48276682e-03 -
1.43926553e-02	-1.98453534e-02	-4.45456058e-02\\n	7.31130037e-03 -
5.73883876e-02	5.91200357e-03	-3.44004631e-02\\n	4.99224756e-03
3.59584615e-02	6.44462276e-03	1.80018116e-02\\n	6.66555017e-02 -
1.43214915e-04	5.16850278e-02	-2.70861341e-03\\n	-8.12202320e-03 -
3.59169021e-02	-2.43293568e-02	3.56825888e-02\\n	3.06037888e-02
2.07295101e-02	-2.72049010e-02	3.89792062e-02\\n	-2.48995889e-03
2.48275232e-02	1.78018268e-02	-2.56232731e-02\\n	3.13616432e-02
7.06909299e-02	-3.29235606e-02	6.67960197e-03\\n	1.89117144e-03 -
1.15451925e-02	-1.72486249e-02	3.75370458e-02\\n	8.91311280e-03
3.62646990e-02	3.51427384e-02	1.90210883e-02\\n	-4.58955057e-02
1.19934686e-01	3.36185768e-02	-7.39446096e-03\\n	-1.03651602e-02
8.12049881e-02	4.01631668e-02	1.02531221e-02\\n	-4.94023971e-02 -
1.05537288e-02	-2.22280361e-02	3.68791493e-03\\n	-1.05225015e-02 -
2.50414722e-02	-3.70037444e-02	-4.52544652e-02\\n	7.80267594e-03
3.66311371e-02	2.23718900e-02	5.00968890e-03\\n	-2.80605908e-02 -
1.19902072e-02	1.19607532e-02	3.20545770e-02\\n	2.95469817e-03
6.01988584e-02	2.25247932e-03	6.14434443e-02\\n	-2.30831411e-02
1.14265140e-02	2.39354055e-02	1.05365543e-02\\n	-1.95968766e-02 -
6.21460844e-03	-3.88051420e-02	1.25449747e-02\\n	4.87881489e-02 -
1.30138760e-02	3.37267518e-02	-2.83177737e-02\\n	-3.69780660e-02 -
1.87966954e-02	2.91273575e-02	3.03810872e-02\\n	-2.06773151e-02
1.28035024e-02	3.55234668e-02	5.05655212e-03\\n	3.86969410e-02 -
6.59883097e-02	6.12658076e-03	6.19648397e-02\\n	-7.87231326e-03
2.65277247e-03	-5.22707403e-02	8.05527344e-03\\n	8.70954469e-02
5.17622009e-02	2.65399069e-02	-1.25501503e-03\\n	1.99715979e-02
2.72730663e-02	4.49447669e-02	3.09776962e-02\\n	-3.58908772e-02
1.70547906e-02	1.96558908e-02	-1.12501793e-02\\n	-5.03951497e-03 -
5.86961955e-02	-9.83590959e-04	4.27694917e-02\\n	-5.07055633e-02 -
1.89975593e-02	4.01682733e-03	3.92593034e-02\\n	1.55548693e-03
7.46860541e-03	5.32642640e-02	-4.59342403e-03\\n	-2.41951030e-02
5.40453196e-02	-7.32910186e-02	-6.24243878e-02\\n	1.83400027e-02 -
8.10545012e-02	-4.10038345e-02	-5.79014830e-02\\n	5.77441463e-03 -
1.38999149e-02	3.55744734e-02	6.65484229e-03\\n	3.30575407e-02
1.69161614e-02	-1.81860868e-02	3.19372416e-02\\n	-1.55437961e-02
5.05150668e-02	5.82077727e-02	-2.34185979e-02\\n	5.89606073e-03 -
2.15327907e-02	4.05314192e-03	-1.02185132e-02\\n	-4.35367711e-02 -
4.99300100e-03	1.46756461e-02	-1.71298329e-02\\n	1.82500705e-02
1.11141447e-02	-7.75681529e-03	-3.55587192e-02\\n	7.47965276e-03 -
8.88304040e-03	-2.59020030e-02	6.83148857e-03\\n	-4.77917828e-02 -
2.46858783e-02	-3.94188752e-03	1.66137014e-02\\n	7.24267634e-03 -
1.85808558e-02	-6.14708364e-02	1.26165347e-02\\n	-3.51574942e-02
1.67146046e-02	-3.57484780e-02	-4.81097698e-02\\n	-6.50256276e-02 -
3.60029191e-02	-9.18580964e-03	4.35761511e-02\\n	6.52618706e-02 -
1.07990149e-02	8.77566729e-03	-6.16241572e-03\\n	-1.77497417e-02
2.54069455e-02	3.35899442e-02	-1.85800437e-02\\n	-1.61337592e-02 -
4.13477346e-02	-1.49374781e-02	2.46960893e-02\\n	2.41081249e-02

9.61841142e-04	3.87678668e-02	3.26361246e-02\\n	6.86734822e-03
5.00897691e-02	5.11830188e-02	-4.05327789e-02\\n	-8.97689760e-02 -
8.13838269e-04	-3.85215655e-02	1.96412746e-02\\n	-4.87155095e-03
6.26533031e-02	-5.00203371e-02	1.36259310e-02\\n	1.98138170e-02
2.23706998e-02	-1.47624556e-02	-2.08031237e-02\\n	5.73374741e-02 -
1.08853631e-01	7.27196187e-02	3.57444026e-02\\n	-1.71732216e-03 -
4.48276754e-03	-1.11867022e-02	-1.15359081e-02\\n	-6.34977967e-03
6.09585308e-02	1.82274189e-02	8.02924335e-02\\n	-6.22956501e-03
2.27490924e-02	2.69049909e-02	1.22414818e-02\\n	-3.83753851e-02
5.28007969e-02	-2.88848113e-02	4.91494825e-03\\n	5.87584712e-02
4.09483258e-03	1.61685962e-02	1.02570588e-02\\n	-5.05571030e-02
5.99424243e-02	2.43418068e-02	-2.55821459e-02\\n	3.37801017e-02
2.05735844e-02	-2.93751583e-02	-3.52503918e-02\\n	2.85072997e-02
2.23733932e-02	-1.63354613e-02	1.31281624e-02\\n	-2.06889194e-02
1.05702206e-02	2.39639115e-02	2.64977408e-03\\n	-4.38604094e-02 -
2.48857718e-02	-5.12024667e-03	-4.29650396e-02\\n	3.62291187e-02
1.21394852e-02	-2.53850687e-02	-4.19798791e-02\\n	-2.53918786e-02
3.80630642e-02	-1.32844523e-02	-6.39878362e-02\\n	-6.17671805e-03 -
2.35546101e-03	2.35471465e-02	5.53486682e-02\\n	3.34822424e-02 -
6.05285168e-03	3.04024760e-02	4.34461422e-03\\n	-3.28185409e-02 -
8.15888401e-03	-9.93102323e-03	-4.69345376e-02\\n	-2.77412869e-02 -
4.12742235e-02	-5.05772792e-02	-2.53928006e-02\\n	-7.41897238e-05
2.39503905e-02	-1.62669905e-02	-1.58133835e-03\\n	1.10625632e-01
1.11785065e-02	2.37951316e-02	-6.20768294e-02\\n	4.65054922e-02
4.82486039e-02	5.55227557e-03	3.66428234e-02\\n	5.40087791e-03 -
1.73931941e-02	-5.17673744e-03	1.78117566e-02\\n	-3.30749415e-02
3.81597057e-02	9.95533075e-03	-2.02581962e-03\\n	6.38610404e-03
4.95962463e-02	-8.83316472e-02	-6.28763228e-04\\n	8.45937710e-03
1.88548472e-02	1.00364130e-04	-1.25207026e-02\\n	-7.24978046e-03 -
2.52529141e-03	3.03694909e-03	-2.31710598e-02\\n	1.31589556e-02
4.16985825e-02	-9.28908773e-03	-5.84791750e-02\\n	-6.46011680e-02 -
9.03794989e-02	1.03680491e-02	-1.47697348e-02\\n	3.39405611e-02
2.08894573e-02	-2.78323218e-02	6.83821589e-02\\n	3.50590236e-02 -
4.68555689e-02	4.72042076e-02	-3.77112441e-03\\n	-5.75096253e-03
2.41767466e-02	-4.17235941e-02	2.28880756e-02\\n	4.02758233e-02 -
4.91468981e-02	5.05239628e-02	-1.00680158e-01\\n	-2.90744472e-02 -
1.67413708e-02	1.17921848e-02	-1.26427915e-02\\n	2.48180218e-02
2.69938000e-02	-9.13031306e-03	-5.64103164e-02\\n	4.68287803e-02
3.84604670e-02	3.81848142e-02	2.33297888e-02\\n	2.69803610e-02 -
3.96660902e-02	-2.58250758e-02	-6.80448487e-02\\n	-7.80719274e-04 -
1.27612036e-02	2.29762010e-02	-1.66850742e-02\\n	-1.99029464e-02
1.26399044e-02	9.90941525e-02	-4.01808042e-03\\n	4.98764403e-02 -
5.11977188e-02	1.49696544e-02	3.19837220e-02\\n	2.93051992e-02 -
3.64751816e-02	3.42510156e-02	-6.13177707e-03\\n	6.81453645e-02 -
2.24678498e-03	-1.98654179e-02	5.55717237e-02\\n	-2.85562035e-02
3.58595587e-02	-1.89085081e-02	3.27889845e-02\\n	-3.52927782e-02
4.49295454e-02	3.33935805e-02	-2.24282476e-03\\n	1.83846429e-02
4.99155652e-03	2.87550557e-02	-6.33753240e-02\\n	2.02281289e-02 -
2.66907533e-04	3.94245014e-02	-1.10258274e-02\\n	-3.18046138e-02 -

1.13525372e-02	-2.47585289e-02	-4.52763459e-04\\n	-1.15936622e-02 -
9.17788537e-04	-4.71309796e-02	3.67310457e-02\\n	-2.05483865e-02
7.07767438e-03	-3.64862531e-02	-6.40820637e-02\\n	-5.00232242e-02 -
2.14895923e-02	-2.11248789e-02	2.18097661e-02\\n	-3.70296799e-02 -
7.97156896e-03	-7.28307292e-04	-5.57386354e-02\\n	1.37111153e-02
3.24638449e-02	-4.11842763e-02	-8.94591399e-03\\n	-2.66409963e-02 -
4.38091956e-04	7.27507696e-02	3.20167057e-02\\n	7.19509600e-03 -
7.69989118e-02	-8.38134252e-03	-1.57920923e-02\\n	1.48548437e-02
3.00976424e-03	1.15479948e-02	7.29491115e-02\\n	-1.33349663e-02 -
1.48559851e-03	7.86755513e-03	1.31217875e-02\\n	-7.11150914e-02 -
1.64781269e-02	-1.93814710e-02	-1.98638737e-02\\n	-3.71746086e-02
6.20009191e-02	3.73834372e-02	-1.30262449e-02\\n	-9.81295332e-02
4.18129601e-02	1.00712385e-02	-8.76897667e-03\\n	-3.06629203e-03 -
7.13661686e-02	4.43297476e-02	2.85195205e-02\\n	7.60584278e-03 -
1.95184536e-02	-3.53755690e-02	-1.98184811e-02\\n	1.83212617e-03
3.23965028e-02	7.32532330e-03	1.30306305e-02\\n	-6.46259710e-02
6.06432855e-02	-2.28145029e-02	-9.53818951e-03\\n	-2.75356565e-02 -
1.14284353e-02	1.32799949e-02	-2.13443078e-02\\n	6.87101409e-02
5.53869046e-02	2.36709360e-02	-5.58103174e-02\\n	-1.48388359e-03
1.56297795e-02	-2.13469844e-02	5.28876260e-02\\n	3.98664139e-02
2.78922580e-02	-2.24801544e-02	2.37328303e-03\\n	5.89274503e-02 -
5.76467775e-02	-2.99277268e-02	1.71148032e-02\\n	-5.72593212e-02 -
4.33231285e-03	4.52495329e-02	-5.64960070e-33\\n	7.47269094e-02
3.62142138e-02	3.95277478e-02	3.07968576e-02\\n	5.90636693e-02
3.75339501e-02	-2.31895726e-02	-4.05858383e-02\\n	-1.92534924e-03
9.35129151e-02	-5.24862250e-03	-1.54407015e-02\\n	9.47384071e-03
2.04024091e-02	-5.42762764e-02	-8.57644305e-02\\n	3.89174446e-02 -
2.27755848e-02	4.55556773e-02	2.87668332e-02\\n	-4.85455543e-02 -
6.96453229e-02	-8.42806175e-02	-2.75642015e-02\\n	7.34933466e-03 -
2.87090186e-02	6.42396212e-02	-6.70448085e-03\\n	-9.58653986e-02
1.88550483e-02	-3.90135646e-02	-3.27156186e-02\\n	-4.77096625e-02 -
2.81655490e-02	-1.86207108e-02	2.91774161e-02\\n	2.21861750e-02 -
1.15162283e-02	-2.01660730e-02	1.75863020e-02\\n	-2.69363285e-03 -
2.08921097e-02	1.97621379e-02	-4.78858082e-03\\n	-1.26683619e-02
1.35948882e-02	1.80629622e-02	3.39837112e-02\\n	2.93650609e-02 -
2.20210832e-02	-1.12774428e-02	2.52639689e-02\\n	-3.16103883e-02 -
3.53645737e-05	7.28631169e-02	9.46151465e-03\\n	5.21957800e-02 -
6.56996295e-02	9.55659822e-02	8.91741831e-03\\n	-4.13986333e-02 -
1.96979428e-03	-2.69607920e-02	-3.09552792e-02\\n	-3.77707742e-02 -
2.05183239e-03	2.14984917e-04	6.11692369e-02\\n	7.54572675e-02
3.39427218e-02	-5.28387493e-03	-8.06459859e-02\\n	1.36984903e-02 -
9.93240699e-02	1.18184602e-02	-1.05019547e-02\\n	-1.66479088e-02 -
5.94034381e-02	-7.50455484e-02	-9.63898599e-02\\n	-4.96468060e-02 -
3.80980363e-03	2.93504149e-02	-1.30302366e-02\\n	-1.96876731e-02
7.66343391e-03	-6.62882987e-04	-6.74138777e-03\\n	-3.01071852e-02 -
3.74916419e-02	8.14801380e-02	1.02742910e-02\\n	-3.63987908e-02
4.70368713e-02	4.34757471e-02	-6.05085008e-02\\n	5.91272563e-02 -
6.01011366e-02	-4.68130084e-03	5.12216762e-02\\n	-1.28850369e-02
1.57379806e-02	1.33342706e-02	2.65763644e-02\\n	-5.83724044e-02

```
3.47832330e-02 -1.22272791e-02 1.50706095e-03\\n -1.26263043e-02 -
2.35722829e-02 -8.27669539e-03 3.78872566e-02\\n -1.11825429e-02 -
2.38841549e-02 -1.60250515e-02 3.18535157e-02\\n 5.17347408e-03 -
9.09908582e-03 -6.33173902e-03 -4.04079333e-02\\n -1.26681169e-02
2.52131280e-03 -4.30412777e-02 9.42817144e-03\\n -6.64179102e-02 -
2.53620930e-02 3.53503413e-03 8.19246657e-03\\n 5.27810073e-04 -
4.34151255e-02 -3.94068286e-03 -1.35703934e-02\\n 2.88146026e-07
4.27781008e-02 5.02584130e-03 7.52717210e-03\\n 1.53221963e-02 -
3.81950592e-03 -4.28856388e-02 4.52936767e-03\\n 1.84004679e-02 -
6.30369643e-03 -2.81434022e-02 -1.63980434e-03\\n 3.87146894e-04 -
2.06686463e-02 3.05093341e-02 5.44758840e-03\\n 5.98083548e-02
5.66420220e-02 -2.37631034e-02 3.25634168e-03\\n 4.02556248e-02
2.62154248e-02 2.93611716e-02 4.04729322e-02\\n -3.56091261e-02 -
3.55419237e-03 7.23572969e-02 -5.21490350e-02\\n 3.87458056e-02
6.35161027e-02 -3.82685214e-02 -1.84759367e-02\\n 1.70838181e-02
7.56823570e-02 -5.08912429e-02 -1.48347579e-03\\n -2.67082378e-02
6.18557027e-03 -8.99394450e-04 -1.19499853e-02\\n -7.07405107e-03
3.20798010e-02 1.72325242e-02 -1.53200990e-02\\n 7.72961415e-03 -
5.85321561e-02 -3.12942429e-03 5.66900382e-03\\n -2.73432378e-02
3.62395383e-02 -1.12157129e-02 -1.35082491e-02\\n 3.58749926e-02
6.22950634e-03 2.90324874e-02 -9.97451413e-03\\n -1.87331941e-02
3.08613442e-02 2.58529987e-02 -1.21459886e-02\\n 2.73999870e-02
1.34400213e-02 8.27934500e-03 -4.50032167e-02\\n 1.50194066e-02
5.98297454e-02 -6.39654621e-02 -3.34206223e-02\\n 3.10314759e-34 -
3.95918760e-04 -6.39823824e-02 -1.33804725e-02\\n -3.26808076e-03
4.86219535e-03 -3.79619971e-02 3.63107920e-02\\n 3.06082554e-02 -
7.84143514e-04 5.17915264e-02 -1.50401648e-02]\\n,\\n
\\n[ 1.57728139e-02 -5.64577729e-02 -3.63888033e-03 1.73159987e-02\\n
-2.66745500e-03 2.10536420e-02 -2.02101655e-02 4.30584364e-02\\n -
3.53951268e-02 2.19664518e-02 -4.33511287e-02 -5.20750508e-02\\n -
5.89872338e-03 2.55817883e-02 3.15636769e-02 7.17286533e-03\\n -
3.72072756e-02 -1.11364517e-02 -1.11601995e-02 -3.12545337e-02\\n -
3.86702456e-02 -2.43350528e-02 -4.18420956e-02 -4.27627712e-02\\n -
3.75365429e-02 -1.38184447e-02 2.48206146e-02 -1.15433261e-02\\n -
1.36964619e-02 -3.28309881e-03 1.13740982e-02 -1.04402518e-02\\n
5.59672639e-02 -3.27690318e-02 2.05511310e-06 8.69802199e-03\\n
5.24926791e-03 -6.02529326e-04 7.95765501e-03 8.60197842e-02\\n
1.62341874e-02 3.49696493e-04 -2.19876915e-02 -1.75113545e-03\\n
2.28333678e-02 2.16686577e-02 -1.60330608e-02 2.73146927e-02\\n -
6.92030713e-02 -6.68439129e-03 -2.49029859e-03 2.01713145e-02\\n
1.47406431e-02 1.52928270e-02 6.65228488e-03 2.65217684e-02\\n -
5.44198453e-02 1.34172156e-01 -1.52043123e-02 2.51244586e-02\\n -
2.14611422e-02 7.70719722e-04 1.96670070e-02 -2.68957503e-02\\n
4.74147126e-02 4.20907736e-02 4.15420420e-02 -4.86595370e-02\\n
2.95606162e-02 1.16407927e-02 8.50693509e-03 -1.01046441e-02\\n
4.82071489e-02 6.38101064e-03 2.07358552e-03 -5.98762743e-02\\n -
1.56382807e-02 4.81903777e-02 2.05990300e-02 -7.43976096e-04\\n
1.42410239e-02 3.69130692e-04 -3.08749732e-02 5.07627055e-03\\n -
1.63760204e-02 -8.42410028e-02 3.60810347e-02 8.39669071e-03\\n
```


5.14656641e-02	-3.27205695e-02	7.42358039e-04	-3.66088492e-03\\n
1.47049148e-02	4.50345390e-02	-6.67315498e-02	-3.50497290e-02\\n -
2.04252452e-03	-1.14344561e-03	1.14085399e-01	-3.33110057e-02\\n -
5.43336533e-02	-6.90692663e-03	-1.05299480e-01	4.72107492e-02\\n -
7.59746283e-02	-4.97120246e-02	-4.04516011e-02	2.93611512e-02\\n
1.07014980e-02	-3.83841433e-03	3.03368866e-02	4.20745462e-03\\n
1.17355296e-02	3.46791558e-02	-2.05858164e-02	-6.12229370e-02\\n -
2.72969250e-02	-4.43289801e-03	7.09903007e-03	-2.08244193e-02\\n -
6.91965455e-03	6.64384440e-02	1.79156717e-02	2.68767122e-02\\n
1.39488243e-02	-5.40009560e-03	-3.38539816e-02	1.90435927e-02\\n
9.69934836e-03	-2.47601252e-02	1.38934590e-02	3.50373574e-02\\n -
1.62765086e-02	6.22384483e-03	-2.01239274e-03	-7.44358748e-02\\n
5.14437035e-02	5.49384486e-03	1.21744294e-02	3.73239219e-02\\n -
1.50537109e-02	-3.23126055e-02	-1.56131070e-02	-3.34043358e-03\\n -
1.84459947e-02	2.07658317e-02	-1.68162771e-02	6.05426207e-02\\n -
9.68633406e-03	8.01254530e-03	-2.72264406e-02	-4.52809175e-03\\n -
7.75353014e-02	-2.35011261e-02	5.10917678e-02	-3.80758867e-02\\n
1.37859508e-02	-1.74623784e-02	3.25980298e-02	5.16020320e-02\\n
2.20756009e-02	-2.48685982e-02	-5.07828332e-02	-7.71802366e-02\\n
6.98580779e-03	-1.88843627e-02	-4.38863672e-02	1.05715198e-02\\n
3.30445208e-02	3.87548818e-03	4.17918675e-02	7.42689371e-02\\n -
1.62473880e-02	-2.62174979e-02	-2.07489152e-02	1.32539898e-01\\n -
4.11955602e-02	8.61665234e-04	6.24621920e-02	-7.81116262e-02\\n
7.59768039e-02	-2.26889178e-02	4.39216346e-02	-2.91420631e-02\\n
3.70258503e-02	-6.29993854e-03	-1.81148071e-02	5.53550711e-03\\n
1.04619013e-02	-3.58778760e-02	-7.27115991e-03	-4.91430573e-02\\n -
1.79034621e-02	2.63426341e-02	2.13690829e-02	-1.68986395e-02\\n
6.72692084e-04	-7.07057565e-02	-1.83181074e-02	2.31532566e-02\\n -
3.04878112e-02	-3.46204429e-03	-5.81701845e-03	-5.40998206e-02\\n -
4.43706177e-02	-4.23483402e-02	-4.37021367e-02	2.59232540e-02\\n
3.26459343e-03	2.61602327e-02	-2.89005805e-02	3.05306297e-02\\n
6.89620385e-03	-4.21663038e-02	1.18978396e-02	-1.43660037e-02\\n
1.15575418e-02	-2.52995882e-02	-4.00522910e-02	1.20739350e-02\\n -
2.85350680e-02	-1.07138921e-02	-3.22781019e-02	8.96959193e-03\\n -
4.03619893e-02	8.41991827e-02	1.07750827e-02	-3.76433507e-02\\n -
7.40152821e-02	-4.36559580e-02	2.69117989e-02	-1.00799685e-03\\n
3.17117423e-02	5.25654964e-02	-4.80853803e-02	-5.99735826e-02\\n -
8.01319331e-02	4.34510373e-02	-1.74203720e-02	4.51894887e-02\\n
1.31648295e-02	-9.69533473e-02	4.54950668e-02	4.50390577e-03\\n
4.80539836e-02	-2.98802126e-02	2.00214162e-02	2.06043832e-02\\n -
1.51534937e-02	-1.61411949e-02	-3.15559916e-02	-4.95321713e-02\\n
4.67078015e-03	2.07061321e-02	-7.21533224e-02	4.37031500e-02\\n -
9.53316763e-02	1.38542540e-02	-5.00775017e-02	-2.00575031e-02\\n
4.04674932e-02	-4.60864492e-02	1.04648452e-02	-2.10422873e-02\\n -
1.63541567e-02	2.04278179e-03	1.52302468e-02	-7.33365677e-03\\n
6.12386083e-03	1.80253889e-02	-3.73194413e-03	4.54135388e-02\\n -
1.63549334e-02	2.90952753e-02	-5.81544563e-02	1.81123242e-02\\n
4.74347919e-02	3.09087802e-02	-4.07738751e-03	3.95797528e-02\\n
6.03918172e-03	7.78303444e-02	-7.49446228e-02	-2.59119626e-02\\n
1.99973714e-02	-3.86637338e-02	-5.09800501e-02	-3.81934792e-02\\n

3.10432483e-02	6.50127158e-02	5.39996009e-03	6.04327917e-02\\n -
3.42034153e-03	-2.26868819e-02	3.74620222e-02	3.57151777e-03\\n -
1.46530364e-02	2.38615051e-02	1.38629163e-02	7.55867735e-02\\n
9.34519246e-03	2.70611402e-02	2.20353133e-03	3.78867891e-03\\n
1.29411439e-03	3.50290649e-02	-2.58361306e-02	-3.20287235e-02\\n
3.19325998e-02	5.17735779e-02	3.26043740e-02	-2.35525165e-02\\n -
8.89074616e-03	-9.25663188e-02	-1.09475553e-02	6.18337803e-02\\n -
9.19495523e-03	-1.52931651e-02	-2.32071225e-02	5.54587785e-03\\n
2.19747284e-03	4.17146422e-02	2.41861083e-02	5.54502122e-02\\n
1.29140131e-02	2.37725414e-02	-4.04298417e-02	-8.37784261e-03\\n
1.24115532e-03	3.72962430e-02	-2.39369404e-02	3.17287147e-02\\n -
1.81639157e-02	-3.81305888e-02	-2.29516570e-02	1.86472125e-02\\n -
2.51918212e-02	-1.11231739e-02	2.61019431e-02	-4.17887326e-03\\n
6.35215407e-03	3.05978372e-03	-2.26567276e-02	5.52211553e-02\\n
2.82943714e-02	5.34320474e-02	1.34907169e-02	-4.37364243e-02\\n
4.72911022e-04	-5.08616529e-02	2.24530231e-03	2.40628179e-02\\n -
5.08143939e-02	3.10155284e-02	5.88588463e-03	6.03938997e-02\\n -
6.95785042e-03	8.09021369e-02	1.35770431e-02	2.20059697e-02\\n
2.92207021e-03	3.32831107e-02	-2.36401837e-02	3.11954096e-02\\n -
4.10429724e-02	-5.60786687e-02	-5.94614656e-04	2.23749820e-02\\n
2.54720799e-03	-1.74369328e-02	6.10900186e-02	5.14829457e-02\\n
2.62116641e-02	1.92390960e-02	7.50768988e-04	3.26913968e-02\\n
4.47159680e-03	4.62862989e-03	-2.28656530e-02	-1.22452769e-02\\n -
1.07553555e-02	5.44747226e-02	3.25375088e-02	-8.92883632e-03\\n -
2.65619420e-02	5.18130697e-02	-4.14081924e-02	1.72498990e-02\\n
2.43675858e-02	-6.26094043e-02	-1.49971270e-03	3.28977033e-02\\n -
1.18482010e-02	1.05182640e-02	-1.13771232e-02	1.52540198e-02\\n -
6.17332496e-02	5.94183467e-02	-1.74798481e-02	3.99899669e-02\\n -
4.91301995e-03	-3.88532132e-02	4.38259244e-02	7.51819462e-04\\n -
1.10452743e-02	3.96509208e-02	-5.72199374e-02	5.06886914e-02\\n -
1.45355109e-02	-2.52195727e-02	4.76122275e-02	-5.30431755e-02\\n -
4.65138350e-03	-1.32735260e-02	1.48337185e-02	1.04052663e-01\\n
3.78372557e-02	-1.66944489e-02	4.94301505e-02	2.64419690e-02\\n
3.87430713e-02	-2.47008689e-02	3.11886929e-02	-2.38935221e-02\\n -
3.32879019e-03	-4.90695462e-02	3.12209334e-02	5.64975198e-03\\n -
1.49829434e-02	4.12272103e-02	-1.28541626e-02	-4.14467268e-02\\n
2.29692366e-02	5.30691929e-02	-2.79648392e-03	9.76619944e-02\\n -
3.33680725e-03	-3.79325971e-02	3.39732692e-02	-8.15867353e-03\\n
1.95176713e-02	-3.17266546e-02	2.76834685e-02	3.81082972e-03\\n
1.27146882e-03	-3.36397141e-02	-6.71808943e-02	-2.31624162e-03\\n
9.74480063e-03	7.14553520e-02	1.19599774e-02	7.30535947e-03\\n -
2.38244166e-03	-4.34231386e-03	8.59754626e-03	7.95183238e-03\\n -
6.88121393e-02	6.17698543e-02	2.05795504e-02	1.50338085e-02\\n -
2.66217478e-02	-8.35669599e-03	-6.99697286e-02	-4.49123569e-02\\n -
7.37115517e-02	3.19464058e-02	-3.07798814e-02	-7.66985044e-02\\n
1.89426709e-02	1.13014160e-02	-2.09322702e-02	4.37395684e-02\\n -
3.37798335e-02	-3.42227519e-02	-1.77016798e-02	-2.69568264e-02\\n -
3.63975391e-02	1.71616971e-02	-5.46425465e-04	6.61909534e-03\\n -
5.62195061e-03	8.01925510e-02	-3.30427475e-02	5.89381903e-02\\n

1.32203279e-02	1.94528140e-02	-4.06433875e-03	6.62010442e-03\\n
2.55999155e-02	-2.27144100e-02	1.24016985e-01	1.62556879e-02\\n -
1.74379665e-02	-1.78733207e-02	6.59452146e-03	1.21004969e-01\\n
4.74225320e-02	-9.45353508e-03	4.78293635e-02	-1.17366121e-03\\n -
2.45884508e-02	4.98797453e-04	-2.94580702e-02	6.61749989e-02\\n
5.29057719e-03	4.71242182e-02	-5.08823025e-04	-4.14450187e-03\\n -
3.93875204e-02	-2.48596538e-02	-1.20226685e-02	2.86598373e-02\\n
4.12970260e-02	4.82219532e-02	2.75197811e-02	-4.92078774e-02\\n -
2.67136749e-02	9.12839454e-03	-5.45903407e-02	-7.02854292e-03\\n -
1.63870286e-02	3.04527041e-02	2.88387425e-02	8.44170060e-03\\n
2.42968020e-03	1.03751123e-02	-4.05461993e-03	2.85729840e-02\\n -
2.12687515e-02	8.71469732e-03	4.17382643e-02	-5.81028350e-02\\n -
6.57838881e-02	-3.61579955e-02	8.71246029e-03	8.81481692e-02\\n
8.76042992e-03	-1.10896332e-02	-3.13027054e-02	-8.36808607e-03\\n
2.24093217e-02	-3.53497788e-02	4.64792177e-02	1.57019366e-02\\n -
2.84721907e-02	-4.64542629e-03	-9.90463980e-03	-2.12735198e-02\\n
3.03749405e-02	-4.30260785e-02	-1.26876626e-02	-2.74391491e-02\\n
7.86173344e-02	-4.44015302e-02	-4.63078730e-02	-5.84436873e-33\\n
2.52901334e-02	-1.22533622e-03	2.77748099e-03	-1.93980560e-02\\n
6.36001350e-03	7.04465210e-02	5.31472862e-02	8.39251745e-03\\n -
5.63005917e-02	2.05048844e-02	2.50587221e-02	1.99946109e-02\\n -
9.76994867e-04	1.19377207e-02	3.38732712e-02	3.10741961e-02\\n -
1.95467491e-02	-8.23809952e-02	3.13371085e-02	-6.07699994e-03\\n -
4.58179228e-02	-2.34846137e-02	4.69333977e-02	-2.66870335e-02\\n
2.99764778e-02	4.57940623e-03	3.68848629e-02	-8.09239410e-03\\n -
3.91403548e-02	-2.12850124e-02	-1.58816334e-02	5.18109091e-03\\n
4.49667830e-04	6.88075125e-02	1.90401971e-02	3.69306235e-03\\n -
5.63388783e-03	-6.54179081e-02	2.01118906e-04	-4.79890518e-02\\n -
8.07988346e-02	3.20624709e-02	-1.01095662e-02	-4.02874388e-02\\n
2.35207118e-02	2.24210862e-02	-5.00765536e-03	6.61188085e-03\\n
6.33936897e-02	-3.64622623e-02	-3.75311524e-02	6.13954151e-03\\n -
1.46715362e-02	-1.08262962e-02	-3.08830980e-02	6.81238919e-02\\n -
1.99181549e-02	-9.69508290e-03	2.02977508e-02	3.64061221e-02\\n
1.30685084e-02	4.41168854e-03	-9.40269157e-02	2.14329083e-02\\n -
1.58167053e-02	2.95217764e-02	1.00607947e-02	8.51427391e-03\\n -
5.09682670e-02	-5.59637323e-02	-5.46281338e-02	3.26886475e-02\\n -
2.30248272e-03	-1.13981050e-02	4.67576943e-02	-2.70811394e-02\\n -
5.65672144e-02	-7.20600132e-03	8.01948085e-02	3.82957384e-02\\n -
5.66239432e-02	2.40409803e-02	-3.09104827e-04	6.30280981e-03\\n
2.36698873e-02	-1.47807449e-02	1.33694778e-03	-1.02467937e-02\\n -
6.37239143e-02	-1.35340011e-02	7.17986654e-03	5.95078664e-03\\n
1.69184320e-02	-1.00788902e-02	3.31459269e-02	-1.39945187e-02\\n
2.43025329e-02	-1.34768318e-02	5.46660088e-02	3.35097574e-02\\n -
4.18631593e-03	1.20559670e-02	9.96872690e-03	-8.02405830e-03\\n -
2.07022242e-02	-1.35632381e-02	1.76398959e-02	4.72904788e-03\\n
2.84237582e-02	5.47406683e-03	5.50538115e-02	-3.87830101e-02\\n
2.52093542e-02	-2.19766260e-03	-1.40215801e-02	-2.50385376e-03\\n -
8.00563022e-04	-4.19641584e-02	-1.19341845e-02	7.16386829e-03\\n -
4.00717109e-02	-3.03082317e-02	-2.99562775e-02	-2.40452476e-02\\n

```

3.49001437e-02 -1.57152545e-02 1.98987331e-02 2.06106920e-02\\n -
3.93997021e-02 -2.40752660e-02 9.95541923e-03 3.20616625e-02\\n
2.85939109e-07 1.87074784e-02 -1.40055991e-03 -3.99141610e-02\\n -
5.14978766e-02 6.18599914e-02 -3.14667150e-02 1.45795029e-02\\n -
9.48637165e-03 -4.80021276e-02 -1.23640768e-01 -7.49548431e-03\\n
4.88134511e-02 2.27913447e-02 4.93784482e-03 -2.99876612e-02\\n
3.39511037e-02 -1.36664929e-02 -5.97698502e-02 1.43400822e-02\\n
7.44066387e-03 -6.55861422e-02 -4.55697849e-02 -4.00813520e-02\\n -
3.28516476e-02 4.69854334e-03 3.70964557e-02 -1.64579116e-02\\n
2.70081754e-03 3.53282616e-02 -3.39808613e-02 1.71928611e-02\\n -
1.29510593e-02 3.41144986e-02 2.54638400e-02 9.78602213e-04\\n -
5.00721298e-03 -6.41398877e-03 -1.76254585e-02 -2.70556696e-02\\n
5.34198470e-02 -2.46294774e-03 -5.98585606e-03 1.63509268e-02\\n -
2.67965533e-02 -3.76960542e-03 7.36843678e-04 8.06827098e-03\\n -
9.76282060e-02 7.08713159e-02 8.81878007e-03 3.66175249e-02\\n -
5.96829504e-02 -4.19249712e-03 -4.77341749e-02 3.49609321e-03\\n -
6.49406388e-02 1.75297447e-02 1.77128389e-02 3.39440480e-02\\n
9.04354751e-02 -1.21040661e-02 -1.56275537e-02 -1.67851686e-05\\n -
4.02413867e-02 6.47976026e-02 4.05368023e-03 2.55564079e-02\\n
2.74016963e-34 3.57053764e-02 -1.24818478e-02 -8.30527209e-03\\n -
3.16899531e-02 -1.62917208e-02 -8.44950415e-03 1.34990709e-02\\n -
2.03760136e-02 9.23624169e-03 1.50190108e-03 -2.35626735e-02]\\n\\n
],\\n      \\\"semantic_type\\\": \\\"\\\",\\n      \\\"description\\\": \\\"\\\"\\n
}\\n    }\\n  ]\\n
n}\\\", \"type\": \"dataframe\", \"variable_name\": \"text_chunks_and_embedding_df_load\"}

```

1. RAG - Search and Answer(70)

We discussed RAG briefly in the beginning but let's quickly recap.

RAG stands for Retrieval Augmented Generation.

Which is another way of saying "given a query, search for relevant resources and answer based on those resources".

Let's breakdown each step:

- **Retrieval** - Get relevant resources given a query. For example, if the query is "what are the macronutrients?" the ideal results will contain information about protein, carbohydrates and fats (and possibly alcohol) rather than information about which tractors are the best for farming (though that is also cool information).
- **Augmentation** - LLMs are capable of generating text given a prompt. However, this generated text is designed to *look* right. And it often has some correct information, however, they are prone to hallucination (generating a result that *looks* like legit text but is factually wrong). In augmentation, we pass relevant information into the prompt and get an LLM to use that relevant information as the basis of its generation.
- **Generation** - This is where the LLM will generate a response that has been flavoured/augmented with the retrieved resources. In turn, this not only gives us a

potentially more correct answer, it also gives us resources to investigate more (since we know which resources went into the prompt).

The whole idea of RAG is to get an LLM to be more factually correct based on your own input as well as have a reference to where the generated output may have come from.

This is an incredibly helpful tool.

Let's say you had 1000s of customer support documents.

You could use RAG to generate direct answers to questions with links to relevant documentation.

Or you were an insurance company with large chains of claims emails.

You could use RAG to answer questions about the emails with sources.

One helpful analogy is to think of LLMs as calculators for words.

With good inputs, the LLM can sort them into helpful outputs.

How?

It starts with better search.

Similarity search

Similarity search or semantic search or vector search is the idea of searching on *vibe*.

If this sounds like woo, woo. It's not.

Perhaps searching via *meaning* is a better analogy.

With keyword search, you are trying to match the string "apple" with the string "apple".

Whereas with similarity/semantic search, you may want to search "macronutrients functions".

And get back results that don't necessarily contain the words "macronutrients functions" but get back pieces of text that match that meaning.

Example: Using similarity search on our textbook data with the query "macronutrients function" returns a paragraph that starts with:

There are three classes of macronutrients: carbohydrates, lipids, and proteins. These can be metabolically processed into cellular energy. The energy from macronutrients comes from their chemical bonds. This chemical energy is converted into cellular energy that is then utilized to perform work, allowing our bodies to conduct their basic functions.

```
import random
from sentence_transformers import util, SentenceTransformer
import torch
import numpy as np
import pandas as pd
```

```

device = "cuda" if torch.cuda.is_available() else "cpu"

# Import texts and embedding df
text_chunks_and_embedding_df =
pd.read_csv(os.path.join("/content/drive/MyDrive/",model_name,"text_chunks_and_embeddings_df.csv"))

# Convert embedding column back to np.array (it got converted to string when it got saved to CSV)
text_chunks_and_embedding_df["embedding"] =
text_chunks_and_embedding_df["embedding"].apply(lambda x:
np.fromstring(x.strip("[]"), sep=" "))

# Convert texts and embedding df to list of dicts
pages_and_chunks =
text_chunks_and_embedding_df.to_dict(orient="records")

# Convert embeddings to torch tensor and send to device (note: NumPy arrays are float64, torch tensors are float32 by default)
embeddings =
torch.tensor(np.array(text_chunks_and_embedding_df["embedding"].tolist()), dtype=torch.float32).to(device)
embeddings.shape

torch.Size([1680, 768])

text_chunks_and_embedding_df.head()

{"summary":{"\n  \"name\": \"text_chunks_and_embedding_df\", \n
\"rows\": 1680, \n  \"fields\": [\n    {\n      \"column\": 
\"page_number\", \n      \"properties\": {\n        \"dtype\": 
\"number\", \n        \"std\": 349, \n        \"min\": -39, \n
\"max\": 1166, \n        \"num_unique_values\": 1136, \n
\"samples\": [\n          795, \n          918, \n          265 \n
], \n        \"semantic_type\": \"\", \n        \"description\": \"\" \n
} \n      }, \n      {\n        \"column\": \"sentence_chunk\", \n
\"properties\": {\n        \"dtype\": \"string\", \n
\"num_unique_values\": 1680, \n        \"samples\": [\n          \"The
major determinants of Type 2 diabetes that can be changed are
overnutrition and a sedentary lifestyle. Therefore, reversing or
improving these factors by lifestyle interventions markedly improve
the overall health of Type 2 diabetics and lower blood-glucose levels.
In fact it has been shown that when people are overweight, losing as
little as nine pounds (four kilograms) decreases blood- glucose levels
in Type 2 diabetics. The Diabetes Prevention Trial demonstrated that
by adhering to a diet containing between 1,200 and 1,800 kilocalories
per day with a dietary fat intake goal of less than 25 percent and
increasing physical activity to at least 150 minutes per week, people
at high risk for Type 2 diabetes achieved a weight loss of 7 percent
and significantly decreased their chances of developing Type 2

```

diabetes.¹⁵ The American Diabetes Association (ADA) has a website that provides information and tips for helping diabetics answer the question, “What Can I Eat?” In regard to carbohydrates the ADA recommends diabetics keep track of the carbohydrates they eat and set a limit. These dietary practices will help keep blood-glucose levels in the target range. Figure 18.5 Metabolic Syndrome: A Combination of Risk Factors Increasing the Chances for Chronic Disease 15. Knowler WC. (2002). Reduction in the Incidence of Type 2 Diabetes with Lifestyle Intervention or Metformin.

“Scheme of a micelle formed by phospholipids in an aqueous solution by Emmanuel Boutet / CC BY-SA 3.0 cholesterol so it acts as an emulsifier. It attracts and holds onto fat while it is simultaneously attracted to and held on to by water. Emulsification increases the surface area of lipids over a thousand-fold, making them more accessible to the digestive enzymes. Once the stomach contents have been emulsified, fat-breaking enzymes work on the triglycerides and diglycerides to sever fatty acids from their glycerol foundations. As pancreatic lipase enters the small intestine, it breaks down the fats into free fatty acids and monoglycerides. Yet again, another hurdle presents itself. How will the fats pass through the watery layer of mucus that coats the absorptive lining of the digestive tract? As before, the answer is bile. Bile salts envelop the fatty acids and monoglycerides to form micelles. Micelles have a fatty acid core with a water-soluble exterior.”

“Image by Gtirouflet / CC BY-SA 3.0 dense cortical bone is about 10 percent porous and it looks like many concentric circles, similar to the rings in a tree trunk, sandwiched together (Figure 2.27 Cortical (Compact) Bone). Cortical bone tissue makes up approximately 80 percent of the adult skeleton. It surrounds all trabecular tissue and is the only bone tissue in the shafts of long bones. Figure 2.26 The Arrangement of Bone Tissues The two basic tissue types of bones are trabecular and cortical. This photo shows normal (left) and degraded (right) trabecular (spongy) bone. Figure 2.27 Cortical (Compact) Bone. The Skeletal System | 123

```
{
  "semantic_type": "text",
  "description": "A paragraph of text describing the structure and function of cortical bone, comparing it to a tree trunk and mentioning its role in the skeletal system.",
  "column": "chunk_char_count",
  "properties": {
    "dtype": "number",
    "std": 415,
    "min": 121,
    "max": 1831,
    "num_unique_values": 992,
    "samples": [
      421, 777, 1617
    ]
  },
  "semantic_type": "text",
  "description": "A paragraph of text describing the structure and function of trabecular bone, comparing it to spongy bone and mentioning its role in the skeletal system.",
  "column": "chunk_word_count",
  "properties": {
    "dtype": "number",
    "std": 66,
    "min": 9,
    "max": 297,
    "num_unique_values": 257,
    "samples": [
      227, 276, 73
    ]
  },
  "semantic_type": "text",
  "description": "A paragraph of text describing the structure and function of bone tissue, mentioning the two basic types (trabecular and cortical) and their arrangement in the skeleton.",
  "column": "chunk_token_count",
  "properties": {
    "dtype": "number",
    "std": 103.84074677997317,
    "min": 30.25,
    "max": 457.75,
    "num_unique_values": 992,

```

```
n      \"samples\": [\n          105.25,\n          194.25,\n          404.25\n      ],\n      \"semantic_type\": \"\", \n      \"description\": \"\", \n      \"column\": \n      \"embedding\", \n      \"properties\": {\n          \"dtype\": \n          \"object\", \n          \"semantic_type\": \"\", \n          \"description\": \"\", \n          \"column\": \n      }\n  },\n  \"type\": \"dataframe\", \"variable_name\": \"text_chunks_and_embedding_df\"}
```

Nice!

Now let's prepare another instance of our embedding model. Not because we have to but because we'd like to make it so you can start the notebook from the cell above.

```
device="cuda"
# from sentence_transformers import util, SentenceTransformer

embedding_model = SentenceTransformer(model_name_or_path="all-mpnet-
base-v2",
                                     device=device) # choose the
device to load the model to

{"model_id": "8d884dd33acf4f3490cbf17e5319be96", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "92455877aa424d2fa2f24401ea38e1e9", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "89075bac2ddf4aa5aa5688f6e054352c", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "fdb1e7fbab784c63b5ea0540ef4219be", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "4523d0c015624e1aa3e503da9f6653d8", "version_major": 2, "vers
ion_minor": 0}

Xet Storage is enabled for this repo, but the 'hf_xet' package is not
installed. Falling back to regular HTTP download. For better
performance, install the package with: `pip install
huggingface_hub[hf_xet]` or `pip install hf_xet`
WARNING:huggingface_hub.file_download:Xet Storage is enabled for this
repo, but the 'hf_xet' package is not installed. Falling back to
regular HTTP download. For better performance, install the package
with: `pip install huggingface_hub[hf_xet]` or `pip install hf_xet`

{"model_id": "5838c10f9fba4a868bc3c6df8885e446", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "88bccbab09704f0d98355a9a800b4f97", "version_major": 2, "vers
ion_minor": 0}
```

```
{"model_id": "72b2c076f0f243cb9458a453994664bc", "version_major": 2, "version_minor": 0}

{"model_id": "862feb9f94e140bfb009fd48916cf71f", "version_major": 2, "version_minor": 0}

{"model_id": "ed730091783a431782cea5b4fee86eaf", "version_major": 2, "version_minor": 0}

{"model_id": "a9c7b882e71c4903a954cba8bcd93b55", "version_major": 2, "version_minor": 0}
```

Embedding model ready!

Time to perform a semantic search.

Let's say you were studying the macronutrients.

And wanted to search your textbook for "macronutrients functions".

Well, we can do so with the following steps:

1. Define a query string (e.g. "macronutrients functions") - note: this could be anything, specific or not.
2. Turn the query string in an embedding with same model we used to embed our text chunks.
3. Perform a [dot product](#) or [cosine similarity](#) function between the text embeddings and the query embedding (we'll get to what these are shortly) to get similarity scores.
4. Sort the results from step 3 in descending order (a higher score means more similarity in the eyes of the model) and use these values to inspect the texts.

Easy!

```
# 1. Define the query
# Note: This could be anything. But since we're working with a
# nutrition textbook, we'll stick with nutrition-based queries.
query = "macronutrients functions"
print(f"Query: {query}")

# 2. Embed the query to the same numerical space as the text examples
# Note: It's important to embed your query with the same model you
# embedded your examples with.
query_embedding = embedding_model.encode(query,
convert_to_tensor=True)

# 3. Get similarity scores with the dot product (we'll time this for
fun)
from time import perf_counter as timer

start_time = timer()
dot_scores = util.dot_score(a=query_embedding, b=embeddings)[0]
```

```

end_time = timer()

print(f"Time take to get scores on {len(embeddings)} embeddings:
{end_time-start_time:.5f} seconds.")

# 4. Get the top-k results (we'll keep this to 5)
top_results_dot_product = torch.topk(dot_scores, k=5)
top_results_dot_product

Query: macronutrients functions
Time take to get scores on 1680 embeddings: 0.02416 seconds.

torch.return_types.topk(
  values=tensor([0.6926, 0.6738, 0.6646, 0.6536, 0.6473]),
  device='cuda:0'),
  indices=tensor([42, 47, 41, 51, 46], device='cuda:0'))

```

We can get pretty far by just storing our embeddings in `torch.tensor` for now.

However, for *much* larger datasets, we'd likely look at a dedicated vector database/indexing libraries such as [Faiss](#).

Let's check the results of our original similarity search.

`torch.topk` returns a tuple of values (scores) and indices for those scores.

The indices relate to which indices in the `embeddings` tensor have what scores in relation to the query embedding (higher is better).

We can use those indices to map back to our text chunks.

First, we'll define a small helper function to print out wrapped text (so it doesn't print a whole text chunk as a single line).

```

# Define helper function to print wrapped text
import textwrap

def print_wrapped(text, wrap_length=80):
    wrapped_text = textwrap.fill(text, wrap_length)
    print(wrapped_text)

```

Now we can loop through the `top_results_dot_product` tuple and match up the scores and indices and then use those indices to index on our `pages_and_chunks` variable to get the relevant text chunk.

Sounds like a lot but we can do it!

```

print(f"Query: '{query}'\n")
print("Results:")
# Loop through zipped together scores and indices from torch.topk
for score, idx in zip(top_results_dot_product[0],

```

```

top_results_dot_product[1]):
    print(f"Score: {score:.4f}")
    # Print relevant sentence chunk (since the scores are in
    # descending order, the most relevant chunk will be first)
    print("Text:")
    print_wrapped(pages_and_chunks[idx]["sentence_chunk"])
    # Print the page number too so we can reference the textbook
    # further (and check the results)
    print(f"Page number: {pages_and_chunks[idx]['page_number']}")
    print("\n")

```

Query: 'macronutrients functions'

Results:

Score: 0.6926

Text:

Macronutrients Nutrients that are needed in large amounts are called macronutrients. There are three classes of macronutrients: carbohydrates, lipids, and proteins. These can be metabolically processed into cellular energy. The energy from macronutrients comes from their chemical bonds. This chemical energy is converted into cellular energy that is then utilized to perform work, allowing our bodies to conduct their basic functions. A unit of measurement of food energy is the calorie. On nutrition food labels the amount given for "calories" is actually equivalent to each calorie multiplied by one thousand. A kilocalorie (one thousand calories, denoted with a small "c") is synonymous with the "Calorie" (with a capital "C") on nutrition food labels. Water is also a macronutrient in the sense that you require a large amount of it, but unlike the other macronutrients, it does not yield calories. Carbohydrates Carbohydrates are molecules composed of carbon, hydrogen, and oxygen.

Page number: 5

Score: 0.6738

Text:

Water There is one other nutrient that we must have in large quantities: water. Water does not contain carbon, but is composed of two hydrogens and one oxygen per molecule of water. More than 60 percent of your total body weight

is water. Without it, nothing could be transported in or out of the body, chemical reactions would not occur, organs would not be cushioned, and body temperature would fluctuate widely. On average, an adult consumes just over two liters of water per day from food and drink combined. Since water is so critical for life's basic processes, the amount of water input and output is supremely important, a topic we will explore in detail in Chapter 4.

Micronutrients

Micronutrients are nutrients required by the body in lesser amounts, but are still essential for carrying out bodily functions. Micronutrients include all the essential minerals and vitamins. There are sixteen essential minerals and thirteen vitamins (See Table 1.1 "Minerals and Their Major Functions" and Table 1.2 "Vitamins and Their Major Functions" for a complete list and their major functions). In contrast to carbohydrates, lipids, and proteins, micronutrients are not sources of energy (calories), but they assist in the process as cofactors or components of enzymes (i.e., coenzymes).

Page number: 8

Score: 0.6646

Text:

Learning Objectives By the end of this chapter, you will be able to:

- Describe basic concepts in nutrition
- Describe factors that affect your nutritional needs
- Describe the importance of research and scientific methods to understanding nutrition

What are Nutrients? The foods we eat contain nutrients. Nutrients are substances required by the body to perform its basic functions. Nutrients must be obtained from our diet, since the human body does not synthesize or produce them. Nutrients have one or more of three basic functions: they provide energy, contribute to body structure, and/or regulate chemical processes in the body. These basic functions allow us to detect and

respond to environmental surroundings, move, excrete wastes, respire (breathe), grow, and reproduce. There are six classes of nutrients required for the body to function and maintain overall health. These are carbohydrates, lipids, proteins, water, vitamins, and minerals. Foods also contain non-nutrients that may be harmful (such as natural toxins common in plant foods and additives like some dyes and preservatives) or beneficial (such as antioxidants). 4 | Introduction
Page number: 4

Score: 0.6536

Text:

Vitamins Major Functions Water-soluble Thiamin (B1) Coenzyme, energy metabolism assistance Riboflavin (B2) Coenzyme, energy metabolism assistance Niacin (B3) Coenzyme, energy metabolism assistance Pantothenic acid (B5) Coenzyme, energy metabolism assistance Pyridoxine (B6) Coenzyme, amino acid synthesis assistance Biotin (B7) Coenzyme, amino acid and fatty acid metabolism Folate (B9) Coenzyme, essential for growth Cobalamin (B12) Coenzyme, red blood cell synthesis C (ascorbic acid) Collagen synthesis, antioxidant Fat-soluble A Vision, reproduction, immune system function D Bone and teeth health maintenance, immune system function E Antioxidant, cell membrane protection K Bone and teeth health maintenance, blood clotting Vitamin deficiencies can cause severe health problems and even death. For example, a deficiency in niacin causes a disease called pellagra, which was common in the early twentieth century in some parts of America. The common signs and symptoms of pellagra are known as the "4D's—diarrhea, dermatitis, dementia, and death." Until scientists found out that better diets relieved the signs and symptoms of pellagra, many people with the disease ended up hospitalized in insane asylums awaiting death. Other vitamins were also found to prevent certain disorders and diseases such as

scurvy (vitamin C), night blindness vitamin A, and rickets (vitamin D). Table

1.3 Functions of Nutrients Introduction | 11

Page number: 11

Score: 0.6473

Text:

Figure 1.1 The Macronutrients: Carbohydrates, Lipids, Protein, and Water

Proteins Proteins are macromolecules composed of chains of subunits called amino

acids. Amino acids are simple subunits composed of carbon, oxygen, hydrogen, and

nitrogen. Food sources of proteins include meats, dairy products, seafood, and a

variety of different plant-based foods, most notably soy. The word protein

comes from a Greek word meaning "of primary importance," which is an apt

description of these macronutrients; they are also known colloquially as the

"workhorses" of life. Proteins provide four kilocalories of energy per gram;

however providing energy is not protein's most important function. Proteins

provide structure to bones, muscles and skin, and play a role in conducting most

of the chemical reactions that take place in the body. Scientists estimate that

greater than one-hundred thousand different proteins exist within the human

body. The genetic codes in DNA are basically protein recipes that determine the

order in which 20 different amino acids are bound together to make thousands of

specific proteins. Figure 1.1 The Macronutrients: Carbohydrates, Lipids,

Protein, and Water Introduction | 7

Page number: 7

Similarity measures: dot product and cosine similarity

Let's talk similarity measures between vectors.

Specifically, embedding vectors which are representations of data with magnitude and direction in high dimensional space (our embedding vectors have 768 dimensions).

Two of the most common you'll across are the dot product and cosine similarity.

They are quite similar.

The main difference is that cosine similarity has a normalization step.

Similarity measure	Description	Code
Dot Product	- Measure of magnitude and direction between two vectors- Vectors that are aligned in direction and magnitude have a higher positive value- Vectors that are opposite in direction and magnitude have a higher negative value	<code>torch.dot, np.dot, sentence_transformers.util.dot_score</code>
Cosine Similarity	- Vectors get normalized by magnitude/ Euclidean norm /L2 norm so they have unit length and are compared more so on direction- Vectors that are aligned in direction have a value close to 1- Vectors that are opposite in direction have a value close to -1	<code>torch.nn.functional.cosine_similarity, 1 - scipy.spatial.distance.cosine (subtract the distance from 1 for similarity measure), sentence_transformers.util.cos_sim</code>

For text similarity, you generally want to use cosine similarity as you are after the semantic measurements (direction) rather than magnitude.

In our case, our embedding model `all-mpnet-base-v2` outputs normalized outputs (see the [Hugging Face model card](#) for more on this) so dot product and cosine similarity return the same results. However, dot product is faster due to not need to perform a normalize step.

To make things bit more concrete, let's make simple dot product and cosine similarity functions and view their results on different vectors.

Note: Similarity measures between vectors and embeddings can be used on any kind of embeddings, not just text embeddings. For example, you could measure image embedding similarity or audio embedding similarity. Or with text and image models like [CLIP](#), you can measure the similarity between text and image embeddings.

```
import torch

def dot_product(vector1, vector2):
    return torch.matmul(vector2, vector1)

def cosine_similarity(vector1, vector2):
```

```

vector1 = vector1.to(torch.float)
vector2 = vector2.to(torch.float)

dp = dot_product(vector1, vector2)
norm1 = torch.norm(vector1)
norm2 = torch.norm(vector2)

return dp / (norm1 * norm2)

# Tests for above functions

# Example tensors
torch.manual_seed(42)
vector1 = torch.randint(low=0, high=10, size=(3,3))
vector2 = torch.randint(low=5, high=15, size=(3,3))
vector3 = torch.randint(low=10, high=20, size=(3,3))
vector4 = torch.randint(low=15, high=30, size=(3,3))

# Calculate dot product
print("Dot product between vector1 and vector2:", dot_product(vector1,
vector2))
assert torch.equal(dot_product(vector1, vector2),torch.tensor([[ 68,
170, 113],
[ 46, 113, 75],
[ 54, 149,
92]]))
print("Dot product between vector1 and vector3:", dot_product(vector1,
vector3))
assert torch.equal(dot_product(vector1, vector3),torch.tensor([[ 98,
291, 182],
[ 84, 234,
161],
[ 90, 255,
179]]))

print("Dot product between vector1 and vector4:", dot_product(vector1,
vector4))
assert torch.equal(dot_product(vector1, vector4),torch.tensor([[136,
376, 254],
[114,
379, 237],
[130,
347, 229]]))

def test_close(a, b, eps=1e-4):
    # Now depending on what a, b are you can add code here
    # return abs(a-b)<eps # If scalar
    return (abs(a-b)<eps).all() # If array

# Calculate cosine similarity

```

```

print("Cosine similarity between vector1 and vector2:",
cosine_similarity(vector1, vector2))
assert test_close(cosine_similarity(vector1,
vector2),torch.tensor([[0.1898, 0.4746, 0.3155],

[0.1284, 0.3155, 0.2094],

[0.1508, 0.4160, 0.2568]]))
print("Cosine similarity between vector1 and vector3:",
cosine_similarity(vector1, vector3))
assert test_close(cosine_similarity(vector1,
vector3),torch.tensor([[0.1584, 0.4704, 0.2942],

[0.1358, 0.3782, 0.2602],

[0.1455, 0.4122, 0.2893]]))
print("Cosine similarity between vector1 and vector4:",
cosine_similarity(vector1, vector4))
assert test_close(cosine_similarity(vector1,
vector4),torch.tensor([[0.1546, 0.4275, 0.2888],

[0.1296, 0.4309, 0.2694],

[0.1478, 0.3945, 0.2603]]))

Dot product between vector1 and vector2: tensor([[ 68, 170, 113],
[ 46, 113, 75],
[ 54, 149, 92]])
Dot product between vector1 and vector3: tensor([[ 98, 291, 182],
[ 84, 234, 161],
[ 90, 255, 179]])
Dot product between vector1 and vector4: tensor([[136, 376, 254],
[114, 379, 237],
[130, 347, 229]])
Cosine similarity between vector1 and vector2: tensor([[0.1898,
0.4746, 0.3155],
[0.1284, 0.3155, 0.2094],
[0.1508, 0.4160, 0.2568]])
Cosine similarity between vector1 and vector3: tensor([[0.1584,
0.4704, 0.2942],
[0.1358, 0.3782, 0.2602],
[0.1455, 0.4122, 0.2893]])
Cosine similarity between vector1 and vector4: tensor([[0.1546,
0.4275, 0.2888],
[0.1296, 0.4309, 0.2694],
[0.1478, 0.3945, 0.2603]])

```

Notice for both dot product and cosine similarity the comparisons of `vector1` and `vector2` are the opposite of `vector1` and `vector4`.

Comparing `vector1` and `vector2` both equations return positive values (14 for dot product and 1.0 for cosine similarity).

But comparing `vector1` and `vector4` the result is in the negative direction.

This makes sense because `vector4` is the negative version of `vector1`.

Whereas comparing `vector1` and `vector3` shows a different outcome.

For the dot product, the value is positive and larger than the comparison of two exactly the same vectors (32 vs 14).

However, for the cosine similarity, thanks to the normalization step, comparing `vector1` and `vector3` results in a positive value close to 1 but not exactly 1.

It is because of this that when comparing text embeddings, cosine similarity is generally favoured as it measures the difference in direction of a pair of vectors rather than difference in magnitude.

And it is this difference in direction that is more generally considered to capture the semantic meaning/vibe of the text.

The good news is that as mentioned before, the outputs of our embedding model `all-mpnet-base-v2` are already normalized.

So we can continue using the dot product (cosine similarity is dot product + normalization).

With similarity measures explained, let's functionize our semantic search steps from above so we can repeat them.

Functionizing our semantic search pipeline

Let's put all of the steps from above for semantic search into a function or two so we can repeat the workflow.

```
def retrieve_relevant_resources(query: str,
                               embeddings: torch.tensor,
                               model:
SentenceTransformer=embedding_model,
                               n_resources_to_return: int=5,
                               product="cosine"
                               ):
    """
    Embeds a query with model and returns top k scores and indices
    from embeddings.
    return scores, indices
    """

    # Embed the query
    # (You can refer to previous code on how to embed query. (remember
    to convert it to Tensor)
    #TODO
```

```

query_embedding = model.encode(query, convert_to_tensor=True)

if product=="cosine":
    dot_scores = cosine_similarity(query_embedding, embeddings)
if product=="dot_product":
    dot_scores = dot_product(query_embedding, embeddings)
if product=="utils":
    dot_scores = util.dot_score(query_embedding, embeddings)[0]

#Obtain Scores and Indices for top k documents
TODO
scores, indices = torch.topk(dot_scores, k=n_resources_to_return)

return scores, indices

def print_top_results_and_scores(query: str,
                                embeddings: torch.tensor,
                                pages_and_chunks:
list[dict]=pages_and_chunks,
                                n_resources_to_return: int=5,
                                product="cosine"):
    """
    Takes a query, retrieves most relevant resources and prints them
    out in descending order.

    Note: Requires pages_and_chunks to be formatted in a specific way
    (see above for reference).
    """

    #Obtain Scores and Indices using retrieve_relevant_resources
    TODO
    scores, indices = retrieve_relevant_resources(query=query,
embeddings=embeddings,
                                                product="cosine")

    print(f"Query: {query}\n")
    print("Results:")
    # Loop through zipped together scores and indicies
    for score, index in zip(scores, indices):
        print(f"Score: {score:.4f}")
        # Print relevant sentence chunk (since the scores are in
descending order, the most relevant chunk will be first)
        print_wrapped(pages_and_chunks[index]["sentence_chunk"])
        # Print the page number too so we can reference the textbook
further and check the results
        print(f"Page number: {pages_and_chunks[index]

```

```
[ 'page_number' ]}"))  
    print("\n")
```

Excellent! Now let's test our functions out.

```
query = "symptoms of pellagra"  
  
# Get just the scores and indices of top related results  
scores, indices = retrieve_relevant_resources(query=query,  
                                              embeddings=embeddings,  
                                              product="cosine")  
  
scores, indices  
  
# Print out the texts of the top scores  
print_top_results_and_scores(query=query,  
                             embeddings=embeddings)
```

Query: symptoms of pellagra

Results:
Score: 0.0122
Niacin deficiency is commonly known as pellagra and the symptoms include fatigue, decreased appetite, and indigestion. These symptoms are then commonly followed by the four D's: diarrhea, dermatitis, dementia, and sometimes death.
Figure 9.12 Conversion of Tryptophan to Niacin Water-Soluble Vitamins | 565
Page number: 565

Score: 0.0091
car. Does it drive faster with a half-tank of gas or a full one?It does not matter; the car drives just as fast as long as it has gas. Similarly, depletion of B vitamins will cause problems in energy metabolism, but having more than is required to run metabolism does not speed it up. Buyers of B-vitamin supplements beware; B vitamins are not stored in the body and all excess will be flushed down the toilet along with the extra money spent. B vitamins are naturally present in numerous foods, and many other foods are enriched with them. In the United States, B-vitamin deficiencies are rare; however in the nineteenth century some vitamin-B deficiencies plagued many people in North

America. Niacin deficiency, also known as pellagra, was prominent in poorer Americans whose main dietary staple was refined cornmeal. Its symptoms were severe and included diarrhea, dermatitis, dementia, and even death. Some of the health consequences of pellagra are the result of niacin being in insufficient supply to support the body's metabolic functions.

Page number: 591

Score: 0.0072

The carbon dioxide gas bubbles infiltrate the stretchy gluten, giving bread its porosity and tenderness. For those who are sensitive to gluten, it is good to know that corn, millet, buckwheat, and oats do not contain the proteins that make gluten. However, some people who have celiac disease also may have a response to products containing oats. This is most likely the result of cross-contamination of grains during harvest, storage, packaging, and processing. Celiac disease is most common in people of European descent and is rare in people of African American, Japanese, and Chinese descent. It is much more prevalent in women and in people with Type 1 diabetes, autoimmune thyroid disease, and Down and Turner syndromes. Symptoms can range from mild to severe and can include pale, fatty, loose stools, gastrointestinal upset, abdominal pain, weight loss and, in children, a failure to grow and thrive. The symptoms can appear in infancy or much later in life, even

Nutrition, Health and Disease
| 1079

Page number: 1079

Score: 0.0068

Image by BruceBlaus/ CC BY 4.0 When the vertebral bone tissue is weakened, it can cause the spine to curve. The increase in spine curvature not only causes pain, but also decreases a person's height. Curvature of the upper

spine produces what is called Dowager's hump, also known as kyphosis. Severe upper-spine deformity can compress the chest cavity and cause difficulty breathing. It may also cause abdominal pain and loss of appetite because of the increased pressure on the abdomen. 1090 | Nutrition, Health and Disease
Page number: 1090

Score: 0.0066
esophagus and cause irritation. It is estimated that GERD affects 25 to 35 percent of the US population. An analysis of several studies published in the August 2005 issue of Annals of Internal Medicine concludes that GERD is much more prevalent in people who are obese.¹ The most common GERD symptom is heartburn, but people with GERD may also experience regurgitation (flow of the stomach's acidic contents into the mouth), frequent coughing, and trouble swallowing. There are other causative factors of GERD that may be separate from or intertwined with obesity. The sphincter that separates the stomach's internal contents from the esophagus often does not function properly and acidic gastric contents seep upward. Sometimes the peristaltic contractions of the esophagus are also sluggish and compromise the clearance of acidic contents. In addition to having an unbalanced, high-fat diet, some people with GERD are sensitive to particular foods—chocolate, garlic, spicy foods, fried foods, and tomato-based foods—which worsen symptoms. Drinks containing alcohol or caffeine may also worsen GERD symptoms. GERD is diagnosed most often by a history of the frequency of recurring symptoms. A more proper diagnosis can be made when a doctor inserts a small device into the lower esophagus that measures the acidity of the contents during one's daily activities.
Page number: 1077

```

query = "symptoms of pellagra"

# Get just the scores and indices of top related results
scores, indices = retrieve_relevant_resources(query=query,
                                              embeddings=embeddings,
                                              product="dot_product")

scores, indices

# Print out the texts of the top scores
print_top_results_and_scores(query=query,
                             embeddings=embeddings)

```

Query: symptoms of pellagra

Results:

Score: 0.0122

Niacin deficiency is commonly known as pellagra and the symptoms include fatigue, decreased appetite, and indigestion. These symptoms are then commonly followed by the four D's: diarrhea, dermatitis, dementia, and sometimes death.

Figure 9.12 Conversion of Tryptophan to Niacin Water-Soluble Vitamins | 565

Page number: 565

Score: 0.0091

car. Does it drive faster with a half-tank of gas or a full one? It does not matter; the car drives just as fast as long as it has gas. Similarly, depletion of B vitamins will cause problems in energy metabolism, but having more than is required to run metabolism does not speed it up. Buyers of B-vitamin supplements beware; B vitamins are not stored in the body and all excess will be flushed down the toilet along with the extra money spent. B vitamins are naturally present in numerous foods, and many other foods are enriched with them. In the United States, B-vitamin deficiencies are rare; however in the nineteenth century some vitamin-B deficiencies plagued many people in North America. Niacin deficiency, also known as pellagra, was prominent in poorer Americans whose main dietary staple was refined cornmeal. Its symptoms were severe and included

diarrhea, dermatitis, dementia, and even death. Some of the health consequences of pellagra are the result of niacin being in insufficient supply to support the body's metabolic functions.

Page number: 591

Score: 0.0072

The carbon dioxide gas bubbles infiltrate the stretchy gluten, giving bread its porosity and tenderness. For those who are sensitive to gluten, it is good to know that corn, millet, buckwheat, and oats do not contain the proteins that make gluten. However, some people who have celiac disease also may have a response to products containing oats. This is most likely the result of cross-contamination of grains during harvest, storage, packaging, and processing. Celiac disease is most common in people of European descent and is rare in people of African American, Japanese, and Chinese descent. It is much more prevalent in women and in people with Type 1 diabetes, autoimmune thyroid disease, and Down and Turner syndromes. Symptoms can range from mild to severe and can include pale, fatty, loose stools, gastrointestinal upset, abdominal pain, weight loss and, in children, a failure to grow and thrive. The symptoms can appear in infancy or much later in life, even

Nutrition, Health and Disease

| 1079

Page number: 1079

Score: 0.0068

Image by BruceBlaus/ CC BY 4.0 When the vertebral bone tissue is weakened, it can cause the spine to curve. The increase in spine curvature not only causes pain, but also decreases a person's height. Curvature of the upper spine produces what is called Dowager's hump, also known as kyphosis. Severe upper-spine deformity can compress the chest cavity and cause difficulty breathing. It

may also cause abdominal pain and loss of appetite because of the increased pressure on the abdomen. 1090 | Nutrition, Health and Disease
Page number: 1090

Score: 0.0066

esophagus and cause irritation. It is estimated that GERD affects 25 to 35 percent of the US population. An analysis of several studies published in the August 2005 issue of Annals of Internal Medicine concludes that GERD is much more prevalent in people who are obese.¹ The most common GERD symptom is heartburn, but people with GERD may also experience regurgitation (flow of the stomach's acidic contents into the mouth), frequent coughing, and trouble swallowing. There are other causative factors of GERD that may be separate from or intertwined with obesity. The sphincter that separates the stomach's internal contents from the esophagus often does not function properly and acidic gastric contents seep upward. Sometimes the peristaltic contractions of the esophagus are also sluggish and compromise the clearance of acidic contents. In addition to having an unbalanced, high-fat diet, some people with GERD are sensitive to particular foods—chocolate, garlic, spicy foods, fried foods, and tomato-based foods—which worsen symptoms. Drinks containing alcohol or caffeine may also worsen GERD symptoms. GERD is diagnosed most often by a history of the frequency of recurring symptoms. A more proper diagnosis can be made when a doctor inserts a small device into the lower esophagus that measures the acidity of the contents during one's daily activities.
Page number: 1077

query = "symptoms of pellagra"

Get just the scores and indices of top related results
scores, indices = retrieve_relevant_resources(query=query,

```
embeddings=embeddings,  
product="utils")  
  
scores, indices  
  
# Print out the texts of the top scores  
print_top_results_and_scores(query=query,  
                             embeddings=embeddings)
```

Query: symptoms of pellagra

Results:

Score: 0.0122

Niacin deficiency is commonly known as pellagra and the symptoms include fatigue, decreased appetite, and indigestion. These symptoms are then commonly followed by the four D's: diarrhea, dermatitis, dementia, and sometimes death.

Figure 9.12 Conversion of Tryptophan to Niacin Water-Soluble Vitamins
| 565

Page number: 565

Score: 0.0091

car. Does it drive faster with a half-tank of gas or a full one? It does not matter; the car drives just as fast as long as it has gas. Similarly, depletion of B vitamins will cause problems in energy metabolism, but having more than is required to run metabolism does not speed it up. Buyers of B-vitamin supplements beware; B vitamins are not stored in the body and all excess will be flushed down the toilet along with the extra money spent. B vitamins are naturally present in numerous foods, and many other foods are enriched with them. In the United States, B-vitamin deficiencies are rare; however in the nineteenth century some vitamin-B deficiencies plagued many people in North America. Niacin deficiency, also known as pellagra, was prominent in poorer Americans whose main dietary staple was refined cornmeal. Its symptoms were severe and included diarrhea, dermatitis, dementia, and even death. Some of the health consequences of pellagra are the result of niacin being in insufficient supply to support the

body's metabolic functions.

Page number: 591

Score: 0.0072

The carbon dioxide gas bubbles infiltrate the stretchy gluten, giving bread its

porosity and tenderness. For those who are sensitive to gluten, it is good to

know that corn, millet, buckwheat, and oats do not contain the proteins that

make gluten. However, some people who have celiac disease also may have a

response to products containing oats. This is most likely the result of cross-

contamination of grains during harvest, storage, packaging, and processing.

Celiac disease is most common in people of European descent and is rare in

people of African American, Japanese, and Chinese descent. It is much more

prevalent in women and in people with Type 1 diabetes, autoimmune thyroid

disease, and Down and Turner syndromes. Symptoms can range from mild to severe

and can include pale, fatty, loose stools, gastrointestinal upset, abdominal

pain, weight loss and, in children, a failure to grow and thrive. The symptoms

can appear in infancy or much later in life, even Nutrition, Health and Disease

| 1079

Page number: 1079

Score: 0.0068

Image by BruceBlaus/ CC BY 4.0 When the vertebral bone tissue is weakened, it

can cause the spine to curve. The increase in spine curvature not only causes

pain, but also decreases a person's height. Curvature of the upper spine

produces what is called Dowager's hump, also known as kyphosis. Severe upper-

spine deformity can compress the chest cavity and cause difficulty breathing. It

may also cause abdominal pain and loss of appetite because of the increased

pressure on the abdomen. 1090 | Nutrition, Health and Disease

Page number: 1090

Score: 0.0066

esophagus and cause irritation. It is estimated that GERD affects 25 to 35 percent of the US population. An analysis of several studies published in the August 2005 issue of Annals of Internal Medicine concludes that GERD is much more prevalent in people who are obese.¹ The most common GERD symptom is heartburn, but people with GERD may also experience regurgitation (flow of the stomach's acidic contents into the mouth), frequent coughing, and trouble swallowing. There are other causative factors of GERD that may be separate from or intertwined with obesity. The sphincter that separates the stomach's internal contents from the esophagus often does not function properly and acidic gastric contents seep upward. Sometimes the peristaltic contractions of the esophagus are also sluggish and compromise the clearance of acidic contents. In addition to having an unbalanced, high-fat diet, some people with GERD are sensitive to particular foods—chocolate, garlic, spicy foods, fried foods, and tomato-based foods—which worsen symptoms. Drinks containing alcohol or caffeine may also worsen GERD symptoms. GERD is diagnosed most often by a history of the frequency of recurring symptoms. A more proper diagnosis can be made when a doctor inserts a small device into the lower esophagus that measures the acidity of the contents during one's daily activities.

Page number: 1077

##Try 3 different queries for dot and cosine similarity and comment on differences in top k results if any. Also try product="utils", does the score and final query change?

```
# define your three test queries
queries = [
    "what is metabolism",
    "symptoms of insomnia",
```

```

    "what is considered overweight"
]

products = ["cosine", "dot_product", "utils"]

for query in queries:
    print(f"\n=== Query: {query!r} ===")
    for product in products:
        print(f"\n--- Product: {product!r} ---")
        scores, indices = retrieve_relevant_resources(
            query=query,
            embeddings=embeddings,
            product=product
        )

        # print raw score/index pairs
        print("Top results (index: score):")
        for idx, score in zip(indices, scores):
            print(f"    • {idx}: {score:.4f}")

        # Then show the full text of those documents
        print("\nTop documents and their scores:")
        print_top_results_and_scores(
            query=query,
            embeddings=embeddings
        )
    print("\n" + "-"*60)

=== Query: 'what is metabolism' ===

--- Product: 'cosine' ---
Top results (index: score):
  • 679: 0.0171
  • 1335: 0.0132
  • 705: 0.0130
  • 1337: 0.0117
  • 682: 0.0117

Top documents and their scores:
Query: what is metabolism

Results:
Score: 0.0171
Metabolism Overview Metabolism is defined as the sum of all chemical
reactions
required to support cellular function and hence the life of an
organism.
Metabolism is either categorized as catabolism, referring to all
metabolic

```

processes involved in molecule breakdown, or anabolism, which includes all metabolic processes involved in building bigger molecules. Generally, catabolic processes release energy and anabolic processes consume energy. The overall goals of metabolism are energy transfer and matter transport. Energy is transformed from food macronutrients into cellular energy, which is used to perform cellular work. Metabolism transforms the matter of macronutrients into substances a cell can use to grow and reproduce and also into waste products. For example, enzymes are proteins and their job is to catalyze chemical reactions. Catalyze means to speed-up a chemical reaction and reduce the energy required to complete the chemical reaction, without the catalyst being used up in the reaction. Without enzymes, chemical reactions would not happen at a fast enough rate and would use up too much energy for life to exist. A metabolic pathway is a series of enzyme catalyzed reactions that transform the starting material (known as a substrate) into intermediates, that are the substrates for subsequent enzymatic reactions in the pathway, until, finally, an end product is synthesized by the last enzymatic reaction in the pathway.

Page number: 462

Score: 0.0132

includes bones, muscles and organs) in your body. A healthy and physically fit individual has a greater proportion of muscle and smaller proportion of fat than an unfit individual of the same weight. Although habitual physical activity can promote a more healthful body composition, other factors like age, gender, genetics, and diet contribute to an individual's body composition. Women have a higher healthy fat percentage than men. For adult women, a healthy amount of body fat ranges from 20 to 32 percent. Adult males on the other hand range from

10 to 22 percent of body fat.⁵ Metabolic Fitness Being fit also includes metabolic fitness. It relates to the number of calories you require to survive and the number of calories you burn during physical activity. Recall that metabolism is the sum of all chemical reactions that occur in the human body to conduct life's processes. Some are catabolic reactions that break down nutrients to supply the body with cellular energy. The rate at which a person burns calories depends on body composition, gender, age, nutritional status, physical activity, and genetics.

Page number: 940

Score: 0.0130

The sum of caloric expenditure is referred to as total energy expenditure (TEE). Basal metabolism refers to those metabolic pathways necessary to support and maintain the body's basic functions (e.g. breathing, heartbeat, liver and kidney function) while at rest. The basal metabolic rate (BMR) is the amount of energy required by the body to conduct its basic functions over a certain time period. The great majority of energy expended (between 50 and 70 percent) daily is from conducting life's basic processes. Of all the organs, the liver requires the most energy

480 | Weight Management

Page number: 480

Score: 0.0117

exercise over time, regular exercise increases lean body mass as well. At rest, lean tissues use more energy than fat tissue therefore increasing basal metabolism. The combination of increased energy output, energy expenditure and basal needs over a long period of time can have a major impact on total energy expenditure (see Figure 16.2 "The Effect of Physical Activity on Energy Expenditure"). The more energy you expend, the more foods you are able to

consume while maintaining a healthy weight. Any improvement to metabolic fitness is beneficial and means a decrease in the risk for developing diabetes, or other chronic conditions. One measurement of metabolic fitness is basal metabolic rate, or BMR, which is a measurement of the amount of energy required for the body to maintain its basic functions while at rest, i.e. breathing, heart beats, liver and kidney function, and so on. On average, BMR accounts for between 50 and 70 percent of a person's total daily energy expenditure. Different factors can affect the BMR. For example, a slender person who is tall has more body surface area and therefore has a higher RMR relative to their body mass (weight). Also, muscle utilizes more energy at rest than fat, so a person with more muscle mass has a higher BMR.

Page number: 941

Score: 0.0117

Energy metabolism refers more specifically to the metabolic pathways that release or store energy. Some of these are catabolic pathways, like glycolysis (the splitting of glucose), β -oxidation (fatty-acid breakdown), and amino acid catabolism. Others are anabolic pathways, and include those involved in storing excess energy (such as glycogenesis), and synthesizing triglycerides (lipogenesis). Table 8.1 "Metabolic Pathways" summarizes some of the catabolic and anabolic pathways and their functions in energy metabolism. Table 8.1

Metabolic Pathways	Catabolic Pathways	Function	Anabolic Pathways	Function
Glycolysis	Glucose breakdown	Gluconeogenesis	Synthesize glucose	
Glycogenolysis	Glycogen breakdown	Glycogenesis	Synthesize glycogen	
	β -oxidation	Fatty-acid breakdown	Lipogenesis	Synthesize triglycerides
	Proteolysis	Protein breakdown	Protein synthesis	Synthesize proteins

Catabolism: The Breakdown All cells are in tune to their energy balance. When energy levels are high

cells
build molecules, and when energy levels are low catabolic pathways are initiated
to make energy. Glucose is the preferred energy source by most tissues, but
fatty acids and amino acids also can be catabolized to release energy that can
drive the formation of ATP. ATP is a high energy molecule that can drive
chemical reactions that require energy. The catabolism of nutrients to release
energy can be separated into three stages, each containing individual metabolic
pathways. The three stages of nutrient breakdown are the following:
464 | The
Atom
Page number: 464

--- Product: 'dot_product' ---

Top results (index: score):

- 679: 0.6999
- 1335: 0.5403
- 705: 0.5312
- 1337: 0.4811
- 682: 0.4807

Top documents and their scores:

Query: what is metabolism

Results:

Score: 0.0171

Metabolism Overview Metabolism is defined as the sum of all chemical reactions

required to support cellular function and hence the life of an organism.

Metabolism is either categorized as catabolism, referring to all metabolic

processes involved in molecule breakdown, or anabolism, which includes all

metabolic processes involved in building bigger molecules. Generally, catabolic

processes release energy and anabolic processes consume energy. The overall

goals of metabolism are energy transfer and matter transport. Energy is

transformed from food macronutrients into cellular energy, which is

used to perform cellular work. Metabolism transforms the matter of macronutrients into substances a cell can use to grow and reproduce and also into waste products. For example, enzymes are proteins and their job is to catalyze chemical reactions. Catalyze means to speed-up a chemical reaction and reduce the energy required to complete the chemical reaction, without the catalyst being used up in the reaction. Without enzymes, chemical reactions would not happen at a fast enough rate and would use up too much energy for life to exist. A metabolic pathway is a series of enzyme catalyzed reactions that transform the starting material (known as a substrate) into intermediates, that are the substrates for subsequent enzymatic reactions in the pathway, until, finally, an end product is synthesized by the last enzymatic reaction in the pathway.

Page number: 462

Score: 0.0132

includes bones, muscles and organs) in your body. A healthy and physically fit individual has a greater proportion of muscle and smaller proportion of fat than an unfit individual of the same weight. Although habitual physical activity can promote a more healthful body composition, other factors like age, gender, genetics, and diet contribute to an individual's body composition. Women have a higher healthy fat percentage than men. For adult women, a healthy amount of body fat ranges from 20 to 32 percent. Adult males on the other hand range from 10 to 22 percent of body fat.⁵ Metabolic Fitness Being fit also includes metabolic fitness. It relates to the number of calories you require to survive and the number of calories you burn during physical activity. Recall that metabolism is the sum of all chemical reactions that occur in the human body to conduct life's processes. Some are catabolic reactions that break down

nutrients
to supply the body with cellular energy. The rate at which a person
burns
calories depends on body composition, gender, age, nutritional status,
physical
activity, and genetics.
Page number: 940

Score: 0.0130

The sum of caloric expenditure is referred to as total energy
expenditure (TEE).
Basal metabolism refers to those metabolic pathways necessary to
support and
maintain the body's basic functions (e.g. breathing, heartbeat, liver
and kidney
function) while at rest. The basal metabolic rate (BMR) is the amount
of energy
required by the body to conduct its basic functions over a certain
time period.
The great majority of energy expended (between 50 and 70 percent)
daily is from
conducting life's basic processes. Of all the organs, the liver
requires the
most energy 480 | Weight Management
Page number: 480

Score: 0.0117

exercise over time, regular exercise increases lean body mass as well.
At rest,
lean tissues use more energy than fat tissue therefore increasing
basal
metabolism. The combination of increased energy output, energy
expenditure and
basal needs over a long period of time can have a major impact on
total energy
expenditure (see Figure 16.2 "The Effect of Physical Activity on
Energy
Expenditure"). The more energy you expend, the more foods you are able
to
consume while maintaining a healthy weight. Any improvement to
metabolic fitness
is beneficial and means a decrease in the risk for developing
diabetes, or other
chronic conditions. One measurement of metabolic fitness is basal
metabolic
rate, or BMR, which is a measurement of the amount of energy required
for the
body to maintain its basic functions while at rest, i.e. breathing,

heart beats, liver and kidney function, and so on. On average, BMR accounts for between 50 and 70 percent of a person's total daily energy expenditure. Different factors can affect the BMR. For example, a slender person who is tall has more body surface area and therefore has a higher RMR relative to their body mass (weight). Also, muscle utilizes more energy at rest than fat, so a person with more muscle mass has a higher BMR.

Page number: 941

Score: 0.0117

Energy metabolism refers more specifically to the metabolic pathways that release or store energy. Some of these are catabolic pathways, like glycolysis (the splitting of glucose), β -oxidation (fatty-acid breakdown), and amino acid catabolism. Others are anabolic pathways, and include those involved in storing excess energy (such as glycogenesis), and synthesizing triglycerides (lipogenesis). Table 8.1 "Metabolic Pathways" summarizes some of the catabolic and anabolic pathways and their functions in energy metabolism. Table 8.1

Metabolic Pathways	Catabolic Pathways	Function	Anabolic Pathways
Glycolysis	Glucose breakdown	Gluconeogenesis	Synthesize glucose
Glycogenolysis	Glycogen breakdown	Glycogenesis	Synthesize glycogen
β -oxidation	Fatty-acid breakdown	Lipogenesis	Synthesize triglycerides
Proteolysis	Protein breakdown to amino acids	Protein synthesis	Synthesize proteins

Catabolism: The Breakdown All cells are in tune to their energy balance. When energy levels are high cells build molecules, and when energy levels are low catabolic pathways are initiated to make energy. Glucose is the preferred energy source by most tissues, but fatty acids and amino acids also can be catabolized to release energy that can drive the formation of ATP. ATP is a high energy molecule that can drive

chemical reactions that require energy. The catabolism of nutrients to release energy can be separated into three stages, each containing individual metabolic pathways. The three stages of nutrient breakdown are the following:
464 | The Atom
Page number: 464

--- Product: 'utils' ---
Top results (index: score):

- 679: 0.6999
- 1335: 0.5403
- 705: 0.5312
- 1337: 0.4811
- 682: 0.4807

Top documents and their scores:
Query: what is metabolism

Results:
Score: 0.0171
Metabolism Overview Metabolism is defined as the sum of all chemical reactions required to support cellular function and hence the life of an organism. Metabolism is either categorized as catabolism, referring to all metabolic processes involved in molecule breakdown, or anabolism, which includes all metabolic processes involved in building bigger molecules. Generally, catabolic processes release energy and anabolic processes consume energy. The overall goals of metabolism are energy transfer and matter transport. Energy is transformed from food macronutrients into cellular energy, which is used to perform cellular work. Metabolism transforms the matter of macronutrients into substances a cell can use to grow and reproduce and also into waste products. For example, enzymes are proteins and their job is to catalyze chemical reactions. Catalyze means to speed-up a chemical reaction and reduce the energy

required to complete the chemical reaction, without the catalyst being used up in the reaction. Without enzymes, chemical reactions would not happen at a fast enough rate and would use up too much energy for life to exist. A metabolic pathway is a series of enzyme catalyzed reactions that transform the starting material (known as a substrate) into intermediates, that are the substrates for subsequent enzymatic reactions in the pathway, until, finally, an end product is synthesized by the last enzymatic reaction in the pathway.

Page number: 462

Score: 0.0132

includes bones, muscles and organs) in your body. A healthy and physically fit individual has a greater proportion of muscle and smaller proportion of fat than an unfit individual of the same weight. Although habitual physical activity can promote a more healthful body composition, other factors like age, gender, genetics, and diet contribute to an individual's body composition. Women have a higher healthy fat percentage than men. For adult women, a healthy amount of body fat ranges from 20 to 32 percent. Adult males on the other hand range from 10 to 22 percent of body fat.⁵ Metabolic Fitness Being fit also includes metabolic fitness. It relates to the number of calories you require to survive and the number of calories you burn during physical activity. Recall that metabolism is the sum of all chemical reactions that occur in the human body to conduct life's processes. Some are catabolic reactions that break down nutrients to supply the body with cellular energy. The rate at which a person burns calories depends on body composition, gender, age, nutritional status, physical activity, and genetics.

Page number: 940

Score: 0.0130

The sum of caloric expenditure is referred to as total energy expenditure (TEE). Basal metabolism refers to those metabolic pathways necessary to support and maintain the body's basic functions (e.g. breathing, heartbeat, liver and kidney function) while at rest. The basal metabolic rate (BMR) is the amount of energy required by the body to conduct its basic functions over a certain time period. The great majority of energy expended (between 50 and 70 percent) daily is from conducting life's basic processes. Of all the organs, the liver requires the most energy

480 | Weight Management
Page number: 480

Score: 0.0117

exercise over time, regular exercise increases lean body mass as well. At rest, lean tissues use more energy than fat tissue therefore increasing basal metabolism. The combination of increased energy output, energy expenditure and basal needs over a long period of time can have a major impact on total energy expenditure (see Figure 16.2 "The Effect of Physical Activity on Energy Expenditure"). The more energy you expend, the more foods you are able to consume while maintaining a healthy weight. Any improvement to metabolic fitness is beneficial and means a decrease in the risk for developing diabetes, or other chronic conditions. One measurement of metabolic fitness is basal metabolic rate, or BMR, which is a measurement of the amount of energy required for the body to maintain its basic functions while at rest, i.e. breathing, heart beats, liver and kidney function, and so on. On average, BMR accounts for between 50 and 70 percent of a person's total daily energy expenditure. Different factors can affect the BMR. For example, a slender person who is tall has more body surface area and therefore has a higher RMR relative to their body mass (weight). Also, muscle utilizes more energy at rest than fat, so a

person with
more muscle mass has a higher BMR.
Page number: 941

Score: 0.0117

Energy metabolism refers more specifically to the metabolic pathways that

release or store energy. Some of these are catabolic pathways, like glycolysis

(the splitting of glucose), β -oxidation (fatty-acid breakdown), and amino acid

catabolism. Others are anabolic pathways, and include those involved in storing

excess energy (such as glycogenesis), and synthesizing triglycerides (lipogenesis). Table 8.1 "Metabolic Pathways" summarizes some of the catabolic

and anabolic pathways and their functions in energy metabolism. Table 8.1

Metabolic Pathways	Catabolic Pathways	Function	Anabolic Pathways
--------------------	--------------------	----------	-------------------

Glycolysis	Glucose breakdown	Gluconeogenesis	Synthesize glucose
------------	-------------------	-----------------	--------------------

Glycogenolysis			
----------------	--	--	--

Glycogen breakdown	Glycogenesis	Synthesize glycogen	β -oxidation Fatty-acid
--------------------	--------------	---------------------	-------------------------------

breakdown	Lipogenesis	Synthesize triglycerides	Proteolysis Protein
-----------	-------------	--------------------------	---------------------

breakdown to	amino acids	Protein synthesis	Synthesize proteins
--------------	-------------	-------------------	---------------------

Catabolism: The Breakdown All

cells are in tune to their energy balance. When energy levels are high cells

build molecules, and when energy levels are low catabolic pathways are initiated

to make energy. Glucose is the preferred energy source by most tissues, but

fatty acids and amino acids also can be catabolized to release energy that can

drive the formation of ATP. ATP is a high energy molecule that can drive

chemical reactions that require energy. The catabolism of nutrients to release

energy can be separated into three stages, each containing individual metabolic

pathways. The three stages of nutrient breakdown are the following:

464 | The

Atom

Page number: 464

=== Query: 'symptoms of insomnia' ===

--- Product: 'cosine' ---

Top results (index: score):

- 334: 0.0071
- 822: 0.0067
- 275: 0.0066
- 274: 0.0065
- 1613: 0.0060

Top documents and their scores:

Query: symptoms of insomnia

Results:

Score: 0.0071

Source: Caffeine. MedicineNet.com.<http://www.medicinenet.com/caffeine/article.htm>. Accessed October 2, 2011. Health Benefits The most

renowned effects of caffeine on the body are increased alertness and delay of

fatigue and sleep. How does caffeine stimulate the brain?Caffeine is chemically

similar to a chemical in our brains (adenosine). Caffeine interacts with

adenosine's specific protein receptor. It blocks the actions of the adenosine,

and affects the levels of signaling molecules in the brain, leading to an

increase in energy metabolism. At the molecular level, caffeine stimulates the

brain, increasing alertness and causing a delay of fatigue and sleep.

At high

doses caffeine stimulates the motor cortex of the brain and interferes with the

sleep-wake cycle, causing side effects such as shakiness, anxiety, and insomnia.

Page number: 218

Score: 0.0067

Niacin deficiency is commonly known as pellagra and the symptoms include

fatigue, decreased appetite, and indigestion. These symptoms are then commonly

followed by the four D's: diarrhea, dermatitis, dementia, and sometimes death.

Figure 9.12 Conversion of Tryptophan to Niacin Water-Soluble Vitamins
| 565

Page number: 565

Score: 0.0066

Symptoms, including nausea, muscle cramps, confusion, dizziness, and in severe cases, coma and death result. Hyponatremia in endurance athletes (such as marathon runners) Sodium | 177
Page number: 177

Score: 0.0065

The symptoms of hyponatremia, also called water intoxication (since it is often the root cause) include nausea, muscle cramps, confusion, dizziness, and in severe cases, coma and death. The physiological events that occur in water intoxication are the following: 1. Excessive sodium loss and/or water intake. 2. Sodium levels fall in blood and in the fluid between cells. 3. Water moves to where solutes are more concentrated (i.e. into cells). 4. Cells swell. 5.
Page number: 177

Score: 0.0060

The primary signs of anorexia are fear of being overweight, extreme dieting, an unusual perception of body image, and depression. The secondary signs and symptoms of anorexia are all related to the caloric and nutrient deficiencies of the unbalanced diet and include excessive weight loss, a multitude of skin abnormalities, diarrhea, cavities and tooth loss, osteoporosis, and liver, kidney, and heart failure. There is no physical test that can be used to diagnose anorexia and distinguish it from other mental 2. Eating Disorders. The National Institute of Mental Health.https://www.nimh.nih.gov/health/statistics/eating-disorders.shtml#part_155061. Accessed April 15, 2018. Undernutrition, Overnutrition, and Malnutrition | 1123
Page number: 1123

--- Product: 'dot_product' ---

Top results (index: score):

- 334: 0.2927
- 822: 0.2745
- 275: 0.2721
- 274: 0.2663
- 1613: 0.2474

Top documents and their scores:

Query: symptoms of insomnia

Results:

Score: 0.0071

Source: Caffeine. MedicineNet.com.<http://www.medicinenet.com/caffeine/article.htm>. Accessed October 2, 2011. Health Benefits The most renowned effects of caffeine on the body are increased alertness and delay of fatigue and sleep. How does caffeine stimulate the brain?Caffeine is chemically similar to a chemical in our brains (adenosine). Caffeine interacts with adenosine's specific protein receptor. It blocks the actions of the adenosine, and affects the levels of signaling molecules in the brain, leading to an increase in energy metabolism. At the molecular level, caffeine stimulates the brain, increasing alertness and causing a delay of fatigue and sleep. At high doses caffeine stimulates the motor cortex of the brain and interferes with the sleep-wake cycle, causing side effects such as shakiness, anxiety, and insomnia.

Page number: 218

Score: 0.0067

Niacin deficiency is commonly known as pellagra and the symptoms include fatigue, decreased appetite, and indigestion. These symptoms are then commonly followed by the four D's: diarrhea, dermatitis, dementia, and sometimes death.

Figure 9.12 Conversion of Tryptophan to Niacin Water-Soluble Vitamins
| 565

Page number: 565

Score: 0.0066

Symptoms, including nausea, muscle cramps, confusion, dizziness, and in severe cases, coma and death result. Hyponatremia in endurance athletes (such as marathon runners) Sodium | 177
Page number: 177

Score: 0.0065

The symptoms of hyponatremia, also called water intoxication (since it is often the root cause) include nausea, muscle cramps, confusion, dizziness, and in severe cases, coma and death. The physiological events that occur in water intoxication are the following: 1. Excessive sodium loss and/or water intake. 2. Sodium levels fall in blood and in the fluid between cells. 3. Water moves to where solutes are more concentrated (i.e. into cells). 4. Cells swell. 5.
Page number: 177

Score: 0.0060

The primary signs of anorexia are fear of being overweight, extreme dieting, an unusual perception of body image, and depression. The secondary signs and symptoms of anorexia are all related to the caloric and nutrient deficiencies of the unbalanced diet and include excessive weight loss, a multitude of skin abnormalities, diarrhea, cavities and tooth loss, osteoporosis, and liver, kidney, and heart failure. There is no physical test that can be used to diagnose anorexia and distinguish it from other mental 2. Eating Disorders. The National Institute of Mental Health.https://www.nimh.nih.gov/health/statistics/eating-disorders.shtml#part_155061. Accessed April 15, 2018. Undernutrition, Overnutrition, and Malnutrition | 1123
Page number: 1123

--- Product: 'utils' ---

Top results (index: score):

- 334: 0.2927
- 822: 0.2745
- 275: 0.2721
- 274: 0.2663
- 1613: 0.2474

Top documents and their scores:

Query: symptoms of insomnia

Results:

Score: 0.0071

Source: Caffeine. MedicineNet.com.<http://www.medicinenet.com/caffeine/article.htm>. Accessed October 2, 2011. Health Benefits The most renowned effects of caffeine on the body are increased alertness and delay of fatigue and sleep. How does caffeine stimulate the brain?Caffeine is chemically similar to a chemical in our brains (adenosine). Caffeine interacts with adenosine's specific protein receptor. It blocks the actions of the adenosine, and affects the levels of signaling molecules in the brain, leading to an increase in energy metabolism. At the molecular level, caffeine stimulates the brain, increasing alertness and causing a delay of fatigue and sleep. At high doses caffeine stimulates the motor cortex of the brain and interferes with the sleep-wake cycle, causing side effects such as shakiness, anxiety, and insomnia.

Page number: 218

Score: 0.0067

Niacin deficiency is commonly known as pellagra and the symptoms include fatigue, decreased appetite, and indigestion. These symptoms are then commonly followed by the four D's: diarrhea, dermatitis, dementia, and sometimes death.

Figure 9.12 Conversion of Tryptophan to Niacin Water-Soluble Vitamins
| 565

Page number: 565

Score: 0.0066

Symptoms, including nausea, muscle cramps, confusion, dizziness, and in severe cases, coma and death result. Hyponatremia in endurance athletes (such as marathon runners) Sodium | 177
Page number: 177

Score: 0.0065

The symptoms of hyponatremia, also called water intoxication (since it is often the root cause) include nausea, muscle cramps, confusion, dizziness, and in severe cases, coma and death. The physiological events that occur in water intoxication are the following: 1. Excessive sodium loss and/or water intake. 2. Sodium levels fall in blood and in the fluid between cells. 3. Water moves to where solutes are more concentrated (i.e. into cells). 4. Cells swell. 5.
Page number: 177

Score: 0.0060

The primary signs of anorexia are fear of being overweight, extreme dieting, an unusual perception of body image, and depression. The secondary signs and symptoms of anorexia are all related to the caloric and nutrient deficiencies of the unbalanced diet and include excessive weight loss, a multitude of skin abnormalities, diarrhea, cavities and tooth loss, osteoporosis, and liver, kidney, and heart failure. There is no physical test that can be used to diagnose anorexia and distinguish it from other mental 2. Eating Disorders. The National Institute of Mental Health.https://www.nimh.nih.gov/health/statistics/eating-disorders.shtml#part_155061. Accessed April 15, 2018. Undernutrition, Overnutrition, and Malnutrition | 1123
Page number: 1123

=== Query: 'what is considered overweight' ===

--- Product: 'cosine' ---

Top results (index: score):

- 217: 0.0143
- 219: 0.0134
- 1519: 0.0133
- 1520: 0.0132
- 221: 0.0132

Top documents and their scores:

Query: what is considered overweight

Results:

Score: 0.0143

Indicators of Health: Body Mass Index, Body Fat Content, and Fat Distribution

UNIVERSITY OF HAWAI'I AT MĀNOA FOOD SCIENCE AND HUMAN NUTRITION PROGRAM AND

HUMAN NUTRITION PROGRAM Although the terms overweight and obese are often used

interchangeably and considered as gradations of the same thing, they denote

different things. The major physical factors contributing to body weight are

water weight, muscle tissue mass, bone tissue mass, and fat tissue mass.

Overweight refers to having more weight than normal for a particular height and

may be the result of water weight, muscle weight, or fat mass. Obese refers

specifically to having excess body fat. In most cases people who are overweight

also have excessive body fat and therefore body weight is an indicator of

obesity in much of the population. The "ideal" healthy body weight for a

particular person is dependent on many things, such as frame size, sex, muscle

mass, bone density, age, and height. The perception of the "ideal" body weight

is additionally dependent on cultural factors and the mainstream societal

advertisement of beauty. To standardize the "ideal" body weight and relate it to

health, scientists have devised mathematical formulas to better define

a healthy weight. These mathematically derived measurements are used by health professionals to correlate disease risk with populations of people and at the individual level. A clinician will take two measurements, one of weight and one of fat mass, in order to diagnose obesity.
Page number: 133

Score: 0.0134

Image by Allison Calabrese / CC BY 4.0 Figure 2.28 Body Composition Body Mass

Index: How to Measure It and Its Limitations Body mass index (BMI) is calculated using height and weight measurements and is more predictive of body fatness than weight alone. BMI measurements are used to indicate whether an individual may be underweight (with a BMI less than 18.5), overweight (with a BMI over 25), or obese (with a BMI over 30). High BMI measurements can be warning signs of health hazards ahead, such as cardiovascular disease, Type 2 diabetes, and other chronic diseases. BMI-associated health risks vary by race. Asians face greater health risks for the same BMI than Caucasians, and 134 | Indicators of Health:
Body Mass Index, Body Fat Content, and Fat Distribution
Page number: 134

Score: 0.0133

Image by 272447 / Pixabay License Health and Body Weight The assumption that health is determined by body weight, or a number on the scale, is outdated and not supported by science. As obesity rates continue to rise, it has been speculated that a higher BMI stands alone as a health risk. However, there is substantial scientific evidence to support the notion that obesity in itself does not put an individual at increased risk of disease. The majority of epidemiological studies demonstrate that five pounds “underweight” is more harmful than 75 pounds “overweight”. 1068 | Calories In Versus Calories Out

Page number: 1068

Score: 0.0132

1314151617 The results of multiple studies have indicated that using weight as

a criterion for health is off-target and potentially 13. Flegal, K. M.,

Graubard, B. I., Williamson, D. F., & Gail, M. H. (2005). Excess deaths

associated with underweight, overweight, and obesity. JAMA, 293(15), 1861–1867.

14. Flegal K. M., Graubard B. I., Williamson D. F., & Gail M. H. (2018, March).

Excess deaths associated with underweight, overweight, and obesity: An evaluation of potential bias. Vital & Health Statistics.42, 1-21

15. Orpan H.

M., Berthaelot J. M., Kaplan M. S. , Feeny D. H., McFarland B., & Ross N. A.

(2010).

Page number: 1069

Score: 0.0132

Source: National Heart, Lung, and Blood Institute. Accessed November 4,

2012.<https://www.nhlbi.nih.gov>. BMI Limitations A BMI is a fairly simple

measurement and does not take into account fat mass or fat distribution in the

body, both of which are additional predictors of disease risk. Body fat weighs

less than muscle mass. Therefore, BMI can sometimes underestimate the amount of

body fat in overweight or obese people and overestimate it in more muscular

people. For instance, a muscular athlete will have more muscle mass (which is

heavier than fat mass) than a sedentary individual of the same height. Based on

their BMIs the muscular athlete would be less “ideal” and may be categorized as

more overweight or obese than the sedentary individual; however this is an

infrequent problem with BMI calculation. Additionally, an older person with

osteoporosis (decreased bone mass) will have a lower BMI than an older person of

the same height without osteoporosis, even though the person with osteoporosis

may have more fat mass. BMI is a useful inexpensive tool to categorize people and is highly correlative with disease risk, but other measurements are needed to diagnose obesity and more accurately assess disease risk.
Page number: 136

--- Product: 'dot_product' ---

Top results (index: score):

- 217: 0.5870
- 219: 0.5481
- 1519: 0.5443
- 1520: 0.5419
- 221: 0.5399

Top documents and their scores:
Query: what is considered overweight

Results:

Score: 0.0143

Indicators of Health: Body Mass Index, Body Fat Content, and Fat Distribution

UNIVERSITY OF HAWAI'I AT MĀNOA FOOD SCIENCE AND HUMAN NUTRITION PROGRAM AND

HUMAN NUTRITION PROGRAM Although the terms overweight and obese are often used

interchangeably and considered as gradations of the same thing, they denote

different things. The major physical factors contributing to body weight are

water weight, muscle tissue mass, bone tissue mass, and fat tissue mass.

Overweight refers to having more weight than normal for a particular height and

may be the result of water weight, muscle weight, or fat mass. Obese refers

specifically to having excess body fat. In most cases people who are overweight

also have excessive body fat and therefore body weight is an indicator of

obesity in much of the population. The "ideal" healthy body weight for a

particular person is dependent on many things, such as frame size, sex, muscle

mass, bone density, age, and height. The perception of the "ideal" body weight

is additionally dependent on cultural factors and the mainstream societal advertisement of beauty. To standardize the “ideal” body weight and relate it to health, scientists have devised mathematical formulas to better define a healthy weight. These mathematically derived measurements are used by health professionals to correlate disease risk with populations of people and at the individual level. A clinician will take two measurements, one of weight and one of fat mass, in order to diagnose obesity.
Page number: 133

Score: 0.0134

Image by Allison Calabrese / CC BY 4.0 Figure 2.28 Body Composition Body Mass

Index: How to Measure It and Its Limitations Body mass index (BMI) is calculated using height and weight measurements and is more predictive of body fatness than weight alone. BMI measurements are used to indicate whether an individual may be underweight (with a BMI less than 18.5), overweight (with a BMI over 25), or obese (with a BMI over 30). High BMI measurements can be warning signs of health hazards ahead, such as cardiovascular disease, Type 2 diabetes, and other chronic diseases. BMI-associated health risks vary by race. Asians face greater health risks for the same BMI than Caucasians, and 134 | Indicators of Health:
Body Mass Index, Body Fat Content, and Fat Distribution
Page number: 134

Score: 0.0133

Image by 272447 / Pixabay License Health and Body Weight The assumption that health is determined by body weight, or a number on the scale, is outdated and not supported by science. As obesity rates continue to rise, it has been speculated that a higher BMI stands alone as a health risk. However, there is substantial scientific evidence to support the notion that obesity in itself does not put an individual at increased risk of disease. The majority

of
epidemiological studies demonstrate that five pounds “underweight” is more harmful than 75 pounds “overweight”. 1068 | Calories In Versus Calories Out
Page number: 1068

Score: 0.0132

1314151617 The results of multiple studies have indicated that using weight as

a criterion for health is off-target and potentially 13. Flegal, K. M.,

Graubard, B. I., Williamson, D. F., & Gail, M. H. (2005). Excess deaths

associated with underweight, overweight, and obesity. JAMA, 293(15), 1861–1867.

14. Flegal K. M., Graubard B. I., Williamson D. F., & Gail M. H. (2018, March).

Excess deaths associated with underweight, overweight, and obesity: An evaluation of potential bias. Vital & Health Statistics.42, 1-21

15. Orpan H.

M., Berthaelot J. M., Kaplan M. S. , Feeny D. H., McFarland B., & Ross N. A.

(2010).

Page number: 1069

Score: 0.0132

Source: National Heart, Lung, and Blood Institute. Accessed November 4,

2012.<https://www.nhlbi.nih.gov>. BMI Limitations A BMI is a fairly simple

measurement and does not take into account fat mass or fat distribution in the

body, both of which are additional predictors of disease risk. Body fat weighs

less than muscle mass. Therefore, BMI can sometimes underestimate the amount of

body fat in overweight or obese people and overestimate it in more muscular

people. For instance, a muscular athlete will have more muscle mass (which is

heavier than fat mass) than a sedentary individual of the same height. Based on

their BMIs the muscular athlete would be less “ideal” and may be categorized as

more overweight or obese than the sedentary individual; however this is an

infrequent problem with BMI calculation. Additionally, an older person

with
osteoporosis (decreased bone mass) will have a lower BMI than an older
person of
the same height without osteoporosis, even though the person with
osteoporosis
may have more fat mass. BMI is a useful inexpensive tool to categorize
people
and is highly correlative with disease risk, but other measurements
are needed
to diagnose obesity and more accurately assess disease risk.
Page number: 136

--- Product: 'utils' ---
Top results (index: score):
• 217: 0.5870
• 219: 0.5481
• 1519: 0.5443
• 1520: 0.5419
• 221: 0.5399

Top documents and their scores:
Query: what is considered overweight

Results:
Score: 0.0143
Indicators of Health: Body Mass Index, Body Fat Content, and Fat
Distribution
UNIVERSITY OF HAWAI'I AT MĀNOA FOOD SCIENCE AND HUMAN NUTRITION
PROGRAM AND
HUMAN NUTRITION PROGRAM Although the terms overweight and obese are
often used
interchangeably and considered as gradations of the same thing, they
denote
different things. The major physical factors contributing to body
weight are
water weight, muscle tissue mass, bone tissue mass, and fat tissue
mass.
Overweight refers to having more weight than normal for a particular
height and
may be the result of water weight, muscle weight, or fat mass. Obese
refers
specifically to having excess body fat. In most cases people who are
overweight
also have excessive body fat and therefore body weight is an indicator
of
obesity in much of the population. The "ideal" healthy body weight for

a particular person is dependent on many things, such as frame size, sex, muscle mass, bone density, age, and height. The perception of the “ideal” body weight is additionally dependent on cultural factors and the mainstream societal advertisement of beauty. To standardize the “ideal” body weight and relate it to health, scientists have devised mathematical formulas to better define a healthy weight. These mathematically derived measurements are used by health professionals to correlate disease risk with populations of people and at the individual level. A clinician will take two measurements, one of weight and one of fat mass, in order to diagnose obesity.
Page number: 133

Score: 0.0134

Image by Allison Calabrese / CC BY 4.0 Figure 2.28 Body Composition
Body Mass

Index: How to Measure It and Its Limitations Body mass index (BMI) is calculated using height and weight measurements and is more predictive of body fatness than weight alone. BMI measurements are used to indicate whether an individual may be underweight (with a BMI less than 18.5), overweight (with a BMI over 25), or obese (with a BMI over 30). High BMI measurements can be warning signs of health hazards ahead, such as cardiovascular disease, Type 2 diabetes, and other chronic diseases. BMI-associated health risks vary by race. Asians face greater health risks for the same BMI than Caucasians, and
134 | Indicators of Health:
Body Mass Index, Body Fat Content, and Fat Distribution
Page number: 134

Score: 0.0133

Image by 272447 / Pixabay License Health and Body Weight The assumption that health is determined by body weight, or a number on the scale, is outdated and not supported by science. As obesity rates continue to rise, it has been

speculated that a higher BMI stands alone as a health risk. However, there is substantial scientific evidence to support the notion that obesity in itself does not put an individual at increased risk of disease. The majority of epidemiological studies demonstrate that five pounds “underweight” is more harmful than 75 pounds “overweight”. 1068 | Calories In Versus Calories Out
Page number: 1068

Score: 0.0132

1314151617 The results of multiple studies have indicated that using weight as a criterion for health is off-target and potentially 13. Flegal, K. M., Graubard, B. I., Williamson, D. F., & Gail, M. H. (2005). Excess deaths associated with underweight, overweight, and obesity. JAMA, 293(15), 1861–1867.
14. Flegal K. M., Graubard B. I., Williamson D. F., & Gail M. H. (2018, March). Excess deaths associated with underweight, overweight, and obesity: An evaluation of potential bias. Vital & Health Statistics.42, 1-21
15. Orpan H. M., Berthaelot J. M., Kaplan M. S. , Feeny D. H., McFarland B., & Ross N. A. (2010).
Page number: 1069

Score: 0.0132

Source: National Heart, Lung, and Blood Institute. Accessed November 4, 2012.<https://www.nhlbi.nih.gov>. BMI Limitations A BMI is a fairly simple measurement and does not take into account fat mass or fat distribution in the body, both of which are additional predictors of disease risk. Body fat weighs less than muscle mass. Therefore, BMI can sometimes underestimate the amount of body fat in overweight or obese people and overestimate it in more muscular people. For instance, a muscular athlete will have more muscle mass (which is heavier than fat mass) than a sedentary individual of the same height. Based on

their BMIs the muscular athlete would be less “ideal” and may be categorized as more overweight or obese than the sedentary individual; however this is an infrequent problem with BMI calculation. Additionally, an older person with osteoporosis (decreased bone mass) will have a lower BMI than an older person of the same height without osteoporosis, even though the person with osteoporosis may have more fat mass. BMI is a useful inexpensive tool to categorize people and is highly correlative with disease risk, but other measurements are needed to diagnose obesity and more accurately assess disease risk.

Page number: 136

Across all three queries, dot-product and cosine returned the same top-5 order—only the score scales differed. Applying product="utils" also changed nothing, since every top hit was already tagged “utils.” Thus here, metric choice affects only score magnitude, and the utils filter is redundant.

Loading an LLM locally

We will do it for `gemma-2b-it`

The Hugging Face `transformers` library has all the tools we need.

To load our LLM, we're going to need a few things:

1. A model ID - This is the reference Hugging Face model ID which will determine which tokenizer and model gets used. For example `gemma-2b-it`. 2. A tokenizer - This is what will turn our raw text into tokens ready for the model. We can create it using the `transformers.AutoTokenizer.from_pretrained` method and passing it our model ID.
2. An LLM model - Again, using our model ID we can load a specific LLM model. To do so we can use the `transformers.AutoModelForCausalLM.from_pretrained` method and passing it our model ID as well as other various parameters.

Let's do it!

```
import os
# Note: the following is Gemma focused, however, there are more and
more LLMs of the 2B and 7B size appearing for local use.
```



```

model_id = os.path.join("/content/drive/MyDrive/",model_name)
print(f"model_id set to: {model_id}")

model_id set to: /content/drive/MyDrive/gemma-transformers-1.1-2b-it-v1

import torch
from transformers import AutoTokenizer, AutoModelForCausalLM
from transformers.utils import is_flash_attn_2_available

# 1. Create quantization config for smaller model loading (optional)
# Requires !pip install bitsandbytes accelerate, see:
https://github.com/TimDettmers/bitsandbytes,
https://huggingface.co/docs/accelerate/
# For models that require 4-bit quantization (use this if you have low GPU memory available)

attn_implementation = "sdpa"
print(f"[INFO] Using attention implementation: {attn_implementation}")

# 2. Pick a model we'd like to use (this will depend on how much GPU memory you have available)
#model_id = "google/gemma-7b-it"
model_id = model_id # (we already set this above)
print(f"[INFO] Using model_id: {model_id}")

# 3. Instantiate tokenizer (tokenizer turns text into numbers ready for the model)
tokenizer =
AutoTokenizer.from_pretrained(pretrained_model_name_or_path=model_id)

# 4. Instantiate the model
llm_model =
AutoModelForCausalLM.from_pretrained(pretrained_model_name_or_path=model_id,

torch_dtype=torch.float16, # datatype to use, we want float16

low_cpu_mem_usage=False, # use full memory

attn_implementation=attn_implementation) # which attention version to use

llm_model.to("cuda")

[INFO] Using attention implementation: sdpa
[INFO] Using model_id: /content/drive/MyDrive/gemma-transformers-1.1-2b-it-v1

{"model_id":"3bc108fbd34444cf8d8844182da72c1f","version_major":2,"version_minor":0}

```

```

GemmaForCausalLM(
  (model): GemmaModel(
    (embed_tokens): Embedding(256000, 2048, padding_idx=0)
    (layers): ModuleList(
      (0-17): 18 x GemmaDecoderLayer(
        (self_attn): GemmaAttention(
          (q_proj): Linear(in_features=2048, out_features=2048,
bias=False)
          (k_proj): Linear(in_features=2048, out_features=256,
bias=False)
          (v_proj): Linear(in_features=2048, out_features=256,
bias=False)
          (o_proj): Linear(in_features=2048, out_features=2048,
bias=False)
        )
        (mlp): GemmaMLP(
          (gate_proj): Linear(in_features=2048, out_features=16384,
bias=False)
          (up_proj): Linear(in_features=2048, out_features=16384,
bias=False)
          (down_proj): Linear(in_features=16384, out_features=2048,
bias=False)
          (act_fn): PytorchGELUTanh()
        )
        (input_layernorm): GemmaRMSNorm((2048,), eps=1e-06)
        (post_attention_layernorm): GemmaRMSNorm((2048,), eps=1e-06)
      )
    )
    (norm): GemmaRMSNorm((2048,), eps=1e-06)
    (rotary_emb): GemmaRotaryEmbedding()
  )
  (lm_head): Linear(in_features=2048, out_features=256000, bias=False)
)

```

Parameters Of The LLM Model

```

def get_model_num_params(model: torch.nn.Module):
    return sum([param.numel() for param in model.parameters()])

get_model_num_params(llm_model)

def get_model_mem_size(model: torch.nn.Module):
    """
    Get how much memory a PyTorch model takes up.

    See:
    https://discuss.pytorch.org/t/gpu-memory-that-model-uses/56822
    """
    # Get model parameters and buffer sizes

```

```

    mem_params = sum([param.nelement() * param.element_size() for
param in model.parameters()])
    mem_buffers = sum([buf.nelement() * buf.element_size() for buf in
model.buffers()])

    # Calculate various model sizes
    model_mem_bytes = mem_params + mem_buffers # in bytes
    model_mem_mb = model_mem_bytes / (1024**2) # in megabytes
    model_mem_gb = model_mem_bytes / (1024**3) # in gigabytes

    return {"model_mem_bytes": model_mem_bytes,
            "model_mem_mb": round(model_mem_mb, 2),
            "model_mem_gb": round(model_mem_gb, 2)}

get_model_mem_size(llm_model)

{'model_mem_bytes': 5012345344, 'model_mem_mb': 4780.15,
 'model_mem_gb': 4.67}

```

Generating text with our LLM

We can generate text with our LLM `model` instance by calling the `generate()` method (this method has plenty of options to pass into it alongside the text) on it and passing it a tokenized input.

The tokenized input comes from passing a string of text to our `tokenizer`.

It's important to note that you should use a tokenizer that has been paired with a model.

Otherwise if you try to use a different tokenizer and then pass those inputs to a model, you will likely get errors/strange results.

For some LLMs, there's a specific template you should pass to them for ideal outputs.

For example, the `gemma-2b-it` model has been trained in a dialogue fashion (instruction tuning).

In this case, our `tokenizer` has a `apply_chat_template()` method which can prepare our input text in the right format for the model.

```

input_text = "What are the macronutrients, and what roles do they play
in the human body?"
print(f"Input text:\n{input_text}")

# Create prompt template for instruction-tuned model
dialogue_template = [
    {"role": "user",
     "content": input_text}
]

# Apply the chat template

```

```
prompt = tokenizer.apply_chat_template(conversation=dialogue_template,
                                     tokenize=False, # keep as raw
                                     add_generation_prompt=True)
print(f"\nPrompt (formatted):\n{prompt}")
```

Input text:

What are the macronutrients, and what roles do they play in the human body?

Prompt (formatted):

<bos><start_of_turn>user

What are the macronutrients, and what roles do they play in the human body?<end_of_turn>

<start_of_turn>model

Notice the scaffolding around our input text, this is the kind of turn-by-turn instruction tuning our model has gone through.

Our next step is to tokenize this formatted text and pass it to our model's `generate()` method.

We'll make sure our tokenized text is on the same device as our model (GPU) using `to("cuda")`.

We can convert the output tokens to text using `tokenizer.decode()`.

Note: "<bos>" and "<eos>" are special tokens to denote "beginning of sentence" and "end of sentence" respectively.

```
## Feel free to play with the questions
```

```
# Nutrition-style questions generated with GPT4
```

```
gpt4_questions = [
    "What are the macronutrients, and what roles do they play in the human body?",
    "How do vitamins and minerals differ in their roles and importance for health?",
    "Describe the process of digestion and absorption of nutrients in the human body.",
    "What role does fibre play in digestion? Name five fibre containing foods.",
    "Explain the concept of energy balance and its importance in weight management."
]
```

```
# Manually created question list
```

```
manual_questions = [
    "How often should infants be breastfed?",
    "What are symptoms of pellagra?",
    "How does saliva help with digestion?",
```

```

    "What is the RDI for protein per day?",
    "water soluble vitamins"
]

query_list = gpt4_questions + manual_questions

```

Augmenting our prompt with context items

What we'd like to do with augmentation is take the results from our search for relevant resources and put them into the prompt that we pass to our LLM.

In essence, we start with a base prompt and update it with context text.

Let's write a function called `prompt_formatter` that takes in a query and our list of context items (in our case it'll be select indices from our list of dictionaries inside `pages_and_chunks`) and then formats the query with text from the context items.

We'll apply the dialogue and chat template to our prompt before returning it as well.

Note: The process of augmenting or changing a prompt to an LLM is known as prompt engineering. And the best way to do it is an active area of research. For a comprehensive guide on different prompt engineering techniques, I'd recommend the Prompt Engineering Guide (promptingguide.ai), [Brex's Prompt Engineering Guide](#) and the paper [Prompt Design and Engineering: Introduction and Advanced Models](#).

```

def prompt_formatter(query: str,
                    context_items: list[dict]) -> str:
    """
    Augments query with text-based context from context_items.
    """
    # Join context items into one dotted paragraph
    context = "- " + "\n- ".join([item["sentence_chunk"] for item in
context_items])

    # Create a base prompt with examples to help the model
    # Note: this is very customizable, I've chosen to use 3 examples
of the answer style we'd like.
    # We could also write this in a txt file and import it in if we
wanted.
    base_prompt = """Based on the following context items, please
answer the query.
Give yourself room to think by extracting relevant passages from the
context before answering the query.
Don't return the thinking, only return the answer.
Make sure your answers are as explanatory as possible.
Use the following examples as reference for the ideal answer style.

\nExample 1:
Query: What are the fat-soluble vitamins?
Answer: The fat-soluble vitamins include Vitamin A, Vitamin D, Vitamin

```

E, and Vitamin K. These vitamins are absorbed along with fats in the diet and can be stored in the body's fatty tissue and liver for later use. Vitamin A is important for vision, immune function, and skin health. Vitamin D plays a critical role in calcium absorption and bone health. Vitamin E acts as an antioxidant, protecting cells from damage. Vitamin K is essential for blood clotting and bone metabolism.

\nExample 2:

Query: What are the causes of type 2 diabetes?

Answer: Type 2 diabetes is often associated with overnutrition, particularly the overconsumption of calories leading to obesity. Factors include a diet high in refined sugars and saturated fats, which can lead to insulin resistance, a condition where the body's cells do not respond effectively to insulin. Over time, the pancreas cannot produce enough insulin to manage blood sugar levels, resulting in type 2 diabetes. Additionally, excessive caloric intake without sufficient physical activity exacerbates the risk by promoting weight gain and fat accumulation, particularly around the abdomen, further contributing to insulin resistance.

\nExample 3:

Query: What is the importance of hydration for physical performance?

Answer: Hydration is crucial for physical performance because water plays key roles in maintaining blood volume, regulating body temperature, and ensuring the transport of nutrients and oxygen to cells. Adequate hydration is essential for optimal muscle function, endurance, and recovery. Dehydration can lead to decreased performance, fatigue, and increased risk of heat-related illnesses, such as heat stroke. Drinking sufficient water before, during, and after exercise helps ensure peak physical performance and recovery.

\nExample 4:

Query: What are the macronutrients, and what roles do they play in the human body?

Answer: The macronutrients are carbohydrates, proteins, and fats, which together provide the body's primary energy, serve as the building blocks for growth and repair, and support essential metabolic and regulatory functions.

\nExample 5:

Query: How many seconds are in a day?

Answer: There are 86400 seconds in a day.

\nExample 6:

Query: What is the RDI for protein per day?

Answer: The RDI for protein per day is 500 grams.

\nNow use the following context items to answer the user query:

{context}

\nRelevant passages: <extract relevant passages from the context here>

```

User query: {query}
Answer: ""

    # Update base prompt with context items and query
    base_prompt = base_prompt.format(context=context, query=query)

    # Create prompt template for instruction-tuned model
    dialogue_template = [
        {"role": "user",
         "content": base_prompt}
    ]

    # Apply the chat template
    prompt =
tokenizer.apply_chat_template(conversation=dialogue_template,
                             tokenize=False,
                             add_generation_prompt=True)

    return prompt

```

Looking good! Let's try our function out.

```

import random

query = random.choice(query_list)
print(f"Query: {query}")

# Create prompt template for instruction-tuned model
dialogue_template = [
    {"role": "user",
     "content": input_text}
]

# Apply the chat template
prompt = tokenizer.apply_chat_template(conversation=dialogue_template,
                                     tokenize=False, # keep as raw
                                     text (not tokenized)
                                     add_generation_prompt=True)

print(prompt)

Query: water soluble vitamins
<bos><start_of_turn>user
What are the macronutrients, and what roles do they play in the human
body?<end_of_turn>
<start_of_turn>model

input_ids = tokenizer(prompt, return_tensors="pt").to("cuda")

# Generate an output of tokens

```

```
outputs = llm_model.generate(**input_ids,
                             temperature=0.7, # lower temperature =
more deterministic outputs, higher temperature = more creative outputs
                             do_sample=True, # whether or not to use
sampling, see https://huyenchip.com/2024/01/16/sampling.html for more
                             max_new_tokens=256) # how many new tokens
to generate from prompt
```

```
# Turn the output tokens into text
output_text = tokenizer.decode(outputs[0])
```

```
print(f"Query: {query}")
print(f"RAG answer:\n{output_text.replace(prompt, '')}")
```

Query: water soluble vitamins

RAG answer:

<bos>**Macronutrients** are the three main types of nutrients that provide the body with energy and building blocks for growth and repair:

1. Protein:

- Essential for tissue repair and growth
- Forms enzymes, hormones, and antibodies
- Builds and maintains muscle, bones, and organs
- Provides energy for cellular processes

2. Carbohydrates:

- Provide energy for cells and tissues
- Fuel for muscles and the brain
- Used for energy by all cells in the body
- Supplies glucose for cellular respiration

3. Fat:

- Stores energy (lipids)
- Provides insulation and protection
- Forms cell membranes and hormones
- Essential for fat absorption and energy metabolism

Roles of Macronutrients in the Human Body:

1. Energy Production:

- Provide the body with the energy it needs to function.
- Fuel cellular processes and organ function.

2. Tissue Repair and Growth:

- Repair and rebuild damaged tissues after injury or illness.

- Promote muscle growth and strength.
- Support tissue repair after surgery.

****3. Metabolism and Energy Balance:****

- Help the body convert food into energy.
- Regulate blood sugar and cholesterol levels.
- Assist in weight regulation.

Our RAG pipeline is complete!

We just Retrieved, Augmented and Generated!

And all on our own local GPU!

How about we functionize the generation step to make it easier to use?

We can put a little formatting on the text being returned to make it look nice too.

And we'll make an option to return the context items if needed as well.

```
def ask(query,
        temperature=0.7,
        max_new_tokens=512,
        format_answer_text=True,
        return_answer_only=True):
    """
    Takes a query, finds relevant resources/context and generates an
    answer to the query based on the relevant resources.
    """

    # Get just the scores and indices of top related results
    scores, indices = retrieve_relevant_resources(query=query,
embeddings=embeddings)

    # Create a list of context items
    context_items = [pages_and_chunks[i] for i in indices]

    # Add score to context item
    for i, item in enumerate(context_items):
        item["score"] = scores[i].cpu() # return score back to CPU

    # Format the prompt with context items
    prompt = prompt_formatter(query=query,
                              context_items=context_items)

    # Tokenize the prompt
    input_ids = tokenizer(prompt, return_tensors="pt").to("cuda")

    # Generate an output of tokens
```

```

outputs = llm_model.generate(**input_ids,
                             temperature=temperature,
                             do_sample=True,
                             max_new_tokens=max_new_tokens)

# Turn the output tokens into text
output_text = tokenizer.decode(outputs[0])

if format_answer_text:
    # Replace special tokens and unnecessary help message
    output_text = output_text.replace(prompt, "").replace("<bos>",
    "").replace("<eos>", "").replace("Sure, here is the answer to the user
query:\n\n", "")

# Only return the answer without the context items
if return_answer_only:
    return output_text

return output_text, context_items

query = random.choice(query_list)
# query = "How many seconds are in a day?"
print(f"Query: {query}")

# Answer query with context and return context
answer, context_items = ask(query=query,
                             temperature=0.7,
                             max_new_tokens=512,
                             return_answer_only=False)

print(f"Answer:\n")
print_wrapped(answer)
print(f"Context items:")
context_items

```

Query: How many seconds are in a day?
Answer:

Based on the provided context, the answer is: There are 86400 seconds in a day.

Context items:

```
[{'page_number': 18,
  'sentence_chunk': 'Units of Measure UNIVERSITY OF HAWAI'I AT MĀNOA
FOOD SCIENCE AND HUMAN NUTRITION PROGRAM AND HUMAN NUTRITION PROGRAM
In nutrition, there are two systems of commonly used measurements:
Metric and US Customary. We need both because the US won't adopt the
metric system completely. The Metric and US Customary System These are
commonly used prefixes for the Metric System: Micro- (μ) 1/1,000,000th
(one millionth) Milli- (m) 1/1000th (one thousandth) Centi- (c)
1/100th (one hundredth) Deci- (d) 1/10th (one tenth) Kilo- (k) 1000x
```

(one thousand times) Mass Metric System US Customary System
Conversions Microgram (μg) Ounce (oz) 1 oz = 28.35 g Milligram (mg)
Pound (lb) 1 lb = 16 oz Gram (g) 1 lb = 454 g Kilogram (kg) 1 kg = 2.2
lbs 18 | Units of Measure',
 'chunk_char_count': 752,
 'chunk_word_count': 129,
 'chunk_token_count': 188.0,
 'embedding': array([-1.64557546e-02, -5.84812947e-02, 2.28464454e-
02, -5.31415977e-02,
 2.88035814e-02, 2.46044397e-02, 3.26909572e-02,
3.10244132e-03,
 1.58166680e-02, 3.48823927e-02, 5.10650948e-02, -
5.99008203e-02,
 -3.68619803e-03, 3.86915617e-02, 1.44070601e-02, -
2.24384163e-02,
 -2.25449093e-02, -9.13097709e-03, -6.42241761e-02,
2.66278405e-02,
 2.17868872e-02, -2.34199949e-02, 1.19376136e-02,
9.65535268e-03,
 -2.67438237e-02, 8.40361509e-03, 5.76322973e-02, -
5.33293188e-02,
 -2.36863010e-02, -9.91676673e-02, -4.35767742e-03, -
4.19991910e-02,
 3.30552203e-03, -5.21008205e-03, 2.32201842e-06,
5.73363109e-03,
 8.06373134e-02, 2.43377853e-02, -7.61167407e-02, -
5.94899766e-02,
 3.06501873e-02, -6.45842329e-02, 5.55519713e-03,
1.32158007e-02,
 -1.53801055e-03, 4.10985388e-03, 3.92659305e-04, -
6.39336780e-02,
 -8.63490924e-02, 2.75205038e-02, 1.69148610e-03, -
2.25541387e-02,
 -1.64644923e-02, -4.47526015e-02, -4.14541811e-02,
3.29862684e-02,
 7.27160322e-03, 1.47626042e-01, 4.22437452e-02,
6.29209429e-02,
 -5.47066331e-03, 2.67634671e-02, 4.53073941e-02,
1.27506601e-02,
 -1.74822006e-02, 4.27207258e-03, 5.38034625e-02,
6.37246948e-03,
 -7.79188890e-03, 4.49404903e-02, 8.88588503e-02, -
1.53868888e-02,
 2.53169071e-02, -1.30571825e-02, 2.69505242e-03, -
3.06002218e-02,
 -3.10071688e-02, -1.05501777e-02, 2.54694503e-02,
1.40170548e-02,
 -7.12494180e-02, 3.56801637e-02, 1.85272843e-02, -
1.68138258e-02,

-5.93257211e-02, 6.33934736e-02, -2.27138065e-02, 9.21158120e-03,
-1.22232226e-02, -4.67901826e-02, -5.07315956e-02, -1.81929097e-02,
1.40325539e-02, -3.07662375e-02, 1.94222406e-02, -9.38674621e-03,
1.77068561e-02, -8.48805755e-02, 6.03315979e-02, -5.87571003e-02,
-2.46358272e-02, -6.22723857e-03, 6.96846424e-03, -4.87124256e-04,
1.86108761e-02, 9.72991586e-02, 1.27134128e-02, -1.84030645e-02,
-6.55618235e-02, -3.86207178e-03, 1.11737214e-02, 3.10827829e-02,
1.23532331e-02, 8.60546157e-03, -2.66330298e-02, -6.94239838e-03,
-6.93283230e-02, -1.04937619e-02, -3.95109802e-02, -2.22893171e-02,
2.44049379e-03, -1.41699119e-02, 5.28109679e-03, -4.85654138e-02,
-2.03160872e-03, 8.89167041e-02, -2.91021960e-03, 5.63672418e-03,
6.99376613e-02, -2.04577912e-02, -9.21125524e-03, -6.49565458e-03,
-8.18504859e-03, 4.20692051e-03, 2.18802076e-02, 1.90610606e-02,
-2.59790313e-03, 4.00353484e-02, 4.95791622e-02, -1.69931464e-02,
-5.45472987e-02, 5.45374192e-02, -4.55860049e-02, 1.72093359e-03,
2.47919001e-02, 1.16902944e-02, 2.05270257e-02, 1.15869828e-02,
1.31221898e-02, 8.74567684e-03, -3.15611698e-02, 2.02447064e-02,
1.50200929e-02, 3.92849520e-02, 3.06743551e-02, -1.56758446e-02,
-1.21374726e-02, -3.72149460e-02, -1.43074093e-03, 2.84850430e-02,
3.86616960e-02, -1.12468880e-02, -5.28201535e-02, -2.26583071e-02,
-2.28696153e-03, -1.07002892e-02, -4.43741418e-02, -2.68363841e-02,
-7.03213783e-03, 6.04296215e-02, -1.14867734e-02, 3.49286385e-02,
-1.37185734e-02, -6.92136437e-02, 2.04593241e-02, 5.29232547e-02,
-3.47337984e-02, 1.06932515e-04, 2.10018381e-02, 1.97784211e-02,
3.67094912e-02, -4.34052497e-02, 1.33127486e-02, -

1.09103778e-02,
1.71852880e-03, -1.18221985e-02, 2.64405087e-02, -
1.71555411e-02,
-4.48539015e-03, -7.40154367e-03, -1.08566768e-02, -
1.88366510e-02,
-7.04140663e-02, -1.26526412e-02, 2.17170902e-02,
1.56646746e-03,
5.90724647e-02, -2.41216621e-03, -2.90012378e-02,
1.22779617e-02,
3.73597033e-02, 6.74739629e-02, -2.78357347e-03, -
1.77518427e-02,
-9.57326789e-04, -5.11980383e-03, 4.26676311e-02,
3.78624834e-02,
5.09392470e-02, -1.81435328e-02, -1.85868237e-02, -
4.01995927e-02,
-7.05028186e-03, -6.37856917e-03, 2.14283802e-02,
6.30930020e-03,
-1.08991545e-02, -3.44524682e-02, 3.46617424e-03,
7.10066855e-02,
1.21538164e-02, -4.31531556e-02, 1.11851310e-02,
5.44663891e-03,
2.24135723e-02, -5.85630052e-02, 4.48995046e-02, -
2.77687758e-02,
-3.62557843e-02, 6.83559710e-03, 7.28303716e-02,
5.98593336e-03,
-4.57092561e-02, 4.76701045e-03, -4.84815575e-02, -
5.15642436e-03,
1.62494853e-02, 9.29920226e-02, 1.55142732e-02,
6.36654124e-02,
2.37473231e-02, -4.96173874e-02, 4.28962661e-03,
7.31066009e-03,
-7.19952211e-02, 4.66764905e-02, 8.51830654e-03, -
2.27239635e-02,
-2.93041654e-02, 2.17469893e-02, 1.83901861e-02, -
3.97814997e-02,
-7.80405477e-02, -2.33235974e-02, -1.92421693e-02,
5.50293410e-03,
9.58483014e-03, 3.28346118e-02, 4.16761376e-02, -
2.52379384e-02,
-1.88087597e-02, -2.90689748e-02, -2.33094487e-03, -
8.85401946e-03,
-5.01912385e-02, -4.61967289e-02, -1.41228018e-02, -
4.93975468e-02,
-7.63748866e-03, 1.08762765e-02, 2.96297371e-02,
5.96687496e-02,
9.65319667e-03, 2.18867753e-02, -3.46827768e-02, -
1.58590488e-02,
3.48571055e-02, 9.21746716e-03, 3.18794549e-02, -
3.93695757e-03,

2.92501133e-02, 8.08311701e-02, -1.04930867e-02, -
6.36903150e-03,
1.26769475e-04, 1.29660750e-02, -3.06786653e-02,
5.04641980e-02,
2.12142989e-02, 2.51400266e-02, 1.42136924e-02,
1.32909007e-02,
-2.57985648e-02, -3.37291658e-02, 2.92783994e-02, -
1.65656582e-02,
5.07926047e-02, 4.94509228e-02, 4.57228767e-03,
8.59141573e-02,
1.71898603e-02, -6.97026169e-03, -3.04195229e-02, -
1.66891608e-02,
-2.58669462e-02, 4.73817848e-02, -2.00833883e-02,
2.18898747e-02,
3.32068512e-03, 5.37034199e-02, -1.50213214e-02,
3.77113977e-03,
5.59302531e-02, 6.74880445e-02, -5.25001027e-02, -
1.24993427e-02,
4.89433389e-03, 1.21005952e-01, 5.36710443e-03, -
6.78856717e-03,
4.94747646e-02, 1.59688145e-02, -3.72816250e-02,
3.66693474e-02,
-1.92403328e-02, 3.45648006e-02, -6.01461567e-02,
6.29412979e-02,
-5.43876514e-02, -3.23707238e-02, 6.61133963e-04, -
9.88761485e-02,
5.23449406e-02, 1.47630451e-02, 5.64776585e-02, -
1.15203690e-02,
-4.70134951e-02, -3.87511253e-02, -1.76396444e-02, -
3.78700979e-02,
-1.14262467e-02, -2.30311844e-02, -5.32482890e-03,
6.04367778e-02,
4.66836654e-02, 5.44381440e-02, -6.28927769e-03, -
3.51002738e-02,
-4.33827192e-02, 3.98056619e-02, 2.29783971e-02, -
1.64024867e-02,
-1.33196730e-02, -1.04681170e-02, 7.72732822e-03,
3.65808047e-02,
1.52417738e-02, -1.32729067e-03, 1.14377793e-02,
2.75002252e-02,
6.11854084e-02, 8.00429657e-02, -2.00813152e-02, -
3.42378430e-02,
-5.89093659e-03, -1.87794510e-02, -1.78722609e-02,
2.47549061e-02,
-1.02532161e-02, 5.26400730e-02, 5.16116694e-02, -
2.32138112e-03,
2.13549789e-02, -2.15285167e-04, -6.68973476e-03, -
1.56252701e-02,
-1.73561554e-02, 9.01726913e-03, -4.63176668e-02,

1.00673269e-02,
-3.69805610e-03, -6.22508973e-02, 5.05449027e-02, -
2.17192564e-02,
-1.20504973e-02, -1.76922306e-02, -1.81674454e-02, -
1.49345128e-02,
4.10298817e-02, -8.53617638e-02, 4.36588787e-02,
5.77510381e-03,
-2.05472801e-02, -6.20046305e-03, 3.52746621e-02, -
3.13671082e-02,
2.78657768e-02, 1.47308270e-03, -1.88493021e-02,
1.45626841e-02,
2.33199783e-02, -5.23917042e-02, -1.41859250e-02, -
3.87146026e-02,
-1.42197153e-02, 4.31722142e-02, -8.48689526e-02, -
1.72328223e-02,
-5.35158860e-03, 1.54964451e-03, 1.43050356e-02,
3.28057297e-02,
-3.81177552e-02, 3.75299645e-03, -1.45423971e-02,
5.97963519e-02,
-6.39723660e-03, -3.62551026e-02, 1.08884345e-03, -
6.35426911e-03,
-2.46594064e-02, -9.61323362e-03, 1.93891972e-02, -
3.90350213e-03,
-1.09304059e-02, -2.63720527e-02, -5.66216670e-02,
9.48375929e-03,
-2.83961854e-04, -3.09895240e-02, -1.10345660e-02, -
5.30904345e-03,
-2.43547652e-02, -1.58387553e-02, 1.03431079e-03,
7.94190988e-02,
1.34212582e-03, -4.64555025e-02, -1.66250218e-03,
3.90280155e-03,
-1.89334080e-02, 2.87690572e-03, 2.77988007e-03,
4.51614931e-02,
-2.85186470e-02, 9.01830047e-02, 2.06433032e-02,
2.67629307e-02,
-5.85948769e-03, -2.77634636e-02, 4.04435536e-03, -
3.22391130e-02,
4.51230183e-02, 3.45154665e-02, 5.63070662e-02,
1.29330708e-02,
5.81207499e-02, -6.43221587e-02, -2.10161582e-02, -
2.31976453e-02,
5.60389049e-02, -1.51745798e-02, -5.40989228e-02,
1.28208697e-02,
-5.69293089e-02, -2.56926827e-02, -8.88670981e-02,
1.14514474e-02,
-3.94694507e-03, -1.54005038e-02, -1.34693459e-02,
5.05022053e-03,
-1.42395664e-02, 2.24024933e-02, 4.09656614e-02, -
1.67392252e-03,

3.76961939e-02, -5.33000603e-02, -3.33857797e-02,
2.62072049e-02,
6.81577763e-03, -2.45533083e-02, 3.43143418e-02,
1.14442438e-01,
-8.83906707e-03, -1.50135411e-02, -2.74288822e-02,
3.19842459e-03,
-9.93795786e-03, -3.26194242e-02, 3.04080006e-02,
8.31379090e-03,
-4.06534635e-02, -3.61297242e-02, 1.10692130e-02,
9.29416716e-03,
2.33234800e-02, -2.07286272e-02, -7.58051500e-02,
5.64731546e-02,
1.78757235e-02, -1.07405381e-02, 4.62804642e-03, -
4.47464660e-02,
4.35085744e-02, 3.40081714e-02, -7.68365059e-03,
2.69422289e-02,
2.39220355e-03, -1.28484508e-02, -1.96688902e-02,
5.67924157e-02,
3.44171152e-02, 4.45078053e-02, 2.84673553e-03, -
2.86177974e-02,
-3.43894400e-02, -1.68772135e-02, -2.06070077e-02,
3.52681242e-02,
8.79225694e-03, -4.78158183e-02, 2.68133674e-02, -
9.53066163e-03,
6.91729635e-02, -8.40911642e-03, -1.47025958e-02,
2.37507522e-02,
1.46366712e-02, 6.53884420e-03, -4.80355173e-02, -
7.15229809e-02,
7.42246956e-03, -4.35867682e-02, 6.42899498e-02,
1.02611808e-02,
-5.05069755e-02, 1.89947020e-02, -1.07290531e-02, -
3.44806374e-03,
-2.54325587e-02, -1.80817265e-02, -1.49730286e-02,
3.25926282e-02,
-3.51363160e-02, 1.58796012e-02, -5.54683544e-02, -
7.27577042e-03,
3.53296064e-02, -1.60987787e-02, 4.77403700e-02,
9.62239690e-03,
8.14862736e-03, 6.44591916e-03, 8.47864803e-03, -
6.07434612e-33,
2.63045281e-02, -7.69408494e-02, 2.63326988e-02, -
1.17772184e-02,
1.12411873e-02, 1.79088246e-02, 9.28967819e-03,
2.94121224e-02,
-1.63371414e-02, 3.25262062e-02, -6.07376778e-03,
6.62300596e-03,
1.84521098e-02, -3.32230553e-02, 6.02353103e-02,
8.59861523e-02,
1.32771395e-02, -4.20589000e-02, -1.00303823e-02, -

3.17372642e-02,
1.19682942e-02, -4.93375622e-02, -8.83310661e-03, -
2.44016387e-02,
8.39450024e-03, 5.18375635e-02, 5.15153222e-02,
2.93274235e-04,
-1.24530047e-02, -2.46247422e-04, 4.15659025e-02,
6.27528951e-02,
3.07541322e-02, -1.77050871e-03, -2.41600033e-02,
3.91863100e-02,
-3.06283794e-02, -9.55099612e-02, -1.34544482e-03, -
3.64043117e-02,
5.61957732e-02, 1.28151700e-02, -4.85434756e-02, -
4.85668518e-03,
-4.29012254e-02, 3.62958037e-03, -1.33014992e-02, -
1.25283217e-02,
7.18121370e-03, -1.87734757e-02, 8.84909704e-02, -
1.41632147e-02,
2.67079249e-02, 6.61583617e-02, 7.85234347e-02,
1.06971469e-02,
-1.96682494e-02, -2.83488873e-02, 4.29076366e-02,
1.57853607e-02,
4.22534086e-02, 3.50836813e-02, -1.00536905e-02,
8.15189164e-03,
-2.13414850e-03, -5.85169205e-03, 2.24711765e-02,
1.72887724e-02,
-1.88835915e-02, -1.96057986e-02, 1.74875762e-02, -
4.79803905e-02,
1.90177523e-02, 2.62396578e-02, 4.33360487e-02, -
3.68522815e-02,
-1.29749551e-02, 5.56452833e-02, -1.69807691e-02, -
1.45396320e-02,
-7.39714429e-02, -2.16027512e-03, -6.29227189e-03,
1.27569940e-02,
1.02338183e-03, -4.83912788e-02, 1.54462485e-02, -
6.79421574e-02,
1.89803448e-02, -4.40735184e-02, 6.75866753e-02, -
1.87490904e-03,
2.03430490e-03, -1.45477811e-02, -6.75760629e-03, -
1.12592541e-01,
4.88309823e-02, -2.22104769e-02, -1.44794565e-02, -
3.69290225e-02,
-5.11422530e-02, 1.07429177e-02, 3.83132324e-02, -
4.65457626e-02,
-2.00972985e-02, -1.94963776e-02, -3.07046175e-02, -
1.59948878e-02,
3.42405066e-02, 1.48611302e-02, 3.36345769e-02, -
2.65303664e-02,
4.08403836e-02, 3.05153728e-02, -3.32516916e-02, -
1.96935628e-02,

```
4.27627303e-02, -4.78099845e-02, -2.78236922e-02, -  
1.09377235e-01,  
-8.06854144e-02, -3.63507271e-02, 3.88977043e-02, -  
3.10607199e-02,  
-1.85765047e-02, -2.51790304e-02, -3.47425565e-02,  
6.00911528e-02,  
-6.44392660e-03, 2.30682213e-02, -1.08395619e-02, -  
4.84630000e-05,  
2.94400422e-07, 3.13284323e-02, -1.99465398e-02,  
3.46952267e-02,  
-2.13478524e-02, 7.73541909e-03, -4.37462293e-02, -  
3.07290871e-02,  
-1.08697629e-02, -8.18887055e-02, -3.50842834e-03,  
6.49818406e-02,  
-4.67251837e-02, -2.16855928e-02, -1.59369688e-02,  
2.64182277e-02,  
6.26674891e-02, 2.13885610e-03, -1.07408790e-02,  
1.16856862e-03,  
-2.53504962e-02, -5.83521500e-02, -5.33471294e-02, -  
1.83042362e-02,  
-2.30549090e-03, -4.83182408e-02, -3.83896828e-02,  
1.74302340e-03,  
-7.16033727e-02, 6.30381852e-02, 2.10098308e-02,  
4.56527695e-02,  
2.04897560e-02, 1.92914028e-02, -3.88471298e-02,  
1.31078651e-02,  
1.52530270e-02, 1.27349822e-02, 5.48560992e-02, -  
3.23969312e-02,  
-5.77889802e-03, -3.35175917e-02, -2.74237040e-02, -  
7.26813078e-03,  
-7.67912939e-02, -3.92705798e-02, 1.28284991e-02,  
4.58375663e-02,  
-1.37072187e-02, 7.24550560e-02, -2.41698679e-02, -  
3.47508751e-02,  
-2.08263807e-02, 3.04073468e-02, -2.72151455e-02,  
4.55902144e-02,  
3.05876993e-02, 6.08376451e-02, 2.26549748e-02,  
6.58122217e-03,  
1.54882809e-02, 1.01091694e-02, -2.94028800e-02, -  
2.33797100e-03,  
-1.72528625e-02, -1.23097626e-02, 5.87158091e-02,  
1.46798627e-03,  
2.13195439e-34, -1.14051923e-02, -1.08457673e-02, -  
1.36874886e-02,  
3.95587310e-02, -3.98647971e-02, 1.66333690e-02, -  
6.93614557e-02,  
9.71460249e-03, 3.39924562e-04, -6.82076067e-02, -  
1.43997446e-02]),  
'score': tensor(0.0076)},
```

```
{'page_number': 942,
'sentence_chunk': 'Image by Allison Calabrese / CC BY 4.0 Physical
Activity Recommendations The CDC along with the American College of
Sports Medicine (ACSM) have evidence based \xa0recommendations and
guidelines for individuals to follow in order to obtain or maintain a
healthy lifestyle. Adults should get at least 150 minutes of moderate-
intensity aerobic physical activity or 75 minutes of vigorous-
intensity aerobic physical activity each week. In addition to aerobic
physical activity, it is recommended that adults do muscle
strengthening activities on each major muscle group two or three times
each week. Adults also are recommended by the ACSM to do flexibility
exercises at least two to three times a week to improve range of
motion. To learn more about these guidelines visit the CDC website at
https://health.gov/paguidelines/guidelines/adults.aspx and the ACSM
website at\xa0http://www.acsm.org/about-acsm/ media-room/news-
releases/2011/08/01/acsm-issues-new- recommendations-on-quantity-and-
quality-of-exercise. 942 | The Essential Elements of Physical
Fitness',
'chunk_char_count': 1053,
'chunk_word_count': 142,
'chunk_token_count': 263.25,
'embedding': array([-2.76699732e-03, -5.53543121e-03, -8.16497300e-
03, 8.38986191e-04,
8.93175881e-03, 6.93442766e-03, 1.52668022e-02,
3.88977565e-02,
2.66618114e-02, 4.15873080e-02, 5.15456758e-02, -
5.06888144e-02,
1.14159295e-02, -1.14665935e-02, -3.72977071e-02,
1.00885704e-02,
-3.19166807e-03, 1.95511375e-02, -2.18950994e-02, -
1.05295877e-03,
-2.49414854e-02, -3.26291919e-02, 6.12556422e-03, -
3.29458416e-02,
1.04873739e-02, -1.26442581e-04, 7.77473375e-02, -
7.96393305e-03,
3.46660651e-02, -5.67487720e-03, -1.91954279e-03,
7.26710958e-03,
2.68372633e-02, -3.35863791e-02, 2.41454836e-06, -
1.47544043e-02,
5.04879989e-02, 8.08563828e-03, 2.28393022e-02,
3.02072354e-02,
4.51699719e-02, -4.41945679e-02, 3.77225168e-02, -
1.13516906e-02,
4.77282237e-03, 1.15010934e-02, 7.53699103e-03,
1.68062486e-02,
3.64248194e-02, 6.00563325e-02, -1.30824288e-02,
8.46503749e-02,
-8.61859843e-02, 1.66273955e-02, 1.35984570e-02,
5.16013987e-02,
```

-5.28091155e-02, 4.55021858e-02, 6.27801940e-02,
2.64232401e-02,
-1.21312048e-02, -1.95137796e-03, 1.21588083e-02,
4.20744121e-02,
-1.70290787e-02, 9.95023455e-03, 8.38492997e-03, -
1.39188282e-02,
2.31342204e-02, 4.71129939e-02, 4.84427959e-02,
9.54869911e-02,
1.06714787e-02, 2.43491177e-02, 1.67628750e-02, -
3.23601952e-03,
-3.42938304e-02, -1.99394003e-02, 2.78860722e-02, -
2.83202790e-02,
1.84345003e-02, 6.43943846e-02, 2.89204139e-02, -
3.90204675e-02,
2.05071922e-02, -1.32545931e-02, -1.13527547e-03, -
2.94423685e-03,
2.63352823e-02, -2.52937060e-02, 2.64554117e-02, -
1.15372622e-02,
-1.50407432e-02, -1.37614170e-02, -4.29855706e-03, -
1.17451046e-02,
1.32361399e-02, -5.29074930e-02, 7.55850375e-02,
1.56194968e-02,
-2.39394628e-03, 2.84758937e-02, 5.12857586e-02, -
1.30505562e-02,
1.69298183e-02, -2.47355364e-02, -4.81197890e-03, -
1.43995434e-02,
-1.00699924e-01, 3.84784639e-02, -3.53946723e-02,
3.64091694e-02,
5.53408973e-02, 2.53542699e-02, 1.17770256e-02, -
5.62391467e-02,
-9.86893475e-02, -4.38332371e-02, -1.56577956e-02, -
3.72167379e-02,
7.57990312e-03, 1.95814241e-02, 5.26548959e-02, -
2.89113708e-02,
4.28059250e-02, 2.37375591e-02, -4.42123041e-02,
4.78690863e-02,
3.19601148e-02, -1.18287476e-02, -2.65427362e-02,
2.07464918e-02,
-2.43006889e-02, -2.19428018e-02, 1.25409663e-02, -
5.23353145e-02,
-3.66589800e-02, 2.14221328e-02, 6.90946355e-02, -
4.43182029e-02,
-3.06669399e-02, 2.48045311e-03, -1.03662256e-03, -
2.90480498e-02,
-9.29457508e-03, -7.02295080e-02, 1.82342194e-02,
6.14012480e-02,
-3.67487557e-02, 5.89315733e-03, -3.61448266e-02, -
5.16689522e-03,
3.68377082e-02, 9.04397145e-02, 8.11201185e-02,
1.71295572e-02,

4.86914106e-02, 2.17635948e-02, 4.26360443e-02,
6.10243669e-03,
-2.84960195e-02, 1.76609016e-03, -4.90703061e-02,
3.98093509e-03,
-1.63244549e-02, -1.38378469e-02, 3.59375291e-02, -
3.96805964e-02,
3.66848037e-02, 7.37327635e-02, -1.31319137e-02,
3.42701562e-02,
-7.27137104e-02, 2.26929393e-02, 1.29650934e-02,
6.35109618e-02,
9.70013253e-03, -8.48039985e-03, -5.74545935e-02, -
4.70181294e-02,
2.50654221e-02, -6.14132322e-02, -2.29980070e-02, -
4.37693745e-02,
4.86585461e-02, -2.85013579e-02, -5.15760854e-02,
2.34640352e-02,
-5.56489006e-02, -8.69154464e-03, 7.95381796e-03, -
3.04864720e-02,
5.26924580e-02, -4.59778123e-02, 2.13548392e-02, -
4.68122168e-03,
3.71278934e-02, 3.93813923e-02, -2.80088559e-02,
9.83667281e-03,
3.11594717e-02, 5.90259396e-03, -5.86980693e-02, -
2.43897773e-02,
-5.64963091e-03, -2.28366461e-02, 7.64378719e-03,
1.10848155e-02,
9.99262333e-02, -4.82273437e-02, 4.83314246e-02,
6.29097298e-02,
4.46761400e-02, -3.71521036e-03, -2.16627028e-02, -
2.54613739e-02,
7.16860080e-03, -3.48593220e-02, -2.48231739e-02,
1.60775930e-02,
1.61192715e-02, 3.25747840e-02, -9.60354581e-02,
2.34394819e-02,
-4.43650298e-02, 2.38869507e-02, 5.36663346e-02, -
1.63080208e-02,
2.34010350e-02, 1.97251830e-02, 6.16799481e-02,
2.91402452e-02,
-3.94792110e-02, -9.96949035e-04, 1.02197612e-02,
2.29493119e-02,
3.90875414e-02, 5.21669909e-02, 3.26229259e-02,
3.29220146e-02,
1.12911332e-02, -3.88609804e-02, 2.48749508e-03, -
2.94218324e-02,
-5.06296419e-02, -3.52109447e-02, -4.39211428e-02,
4.16789874e-02,
-1.54310139e-02, 9.67922285e-02, -4.08937596e-02, -
1.87560879e-02,
-6.99492618e-02, 2.99076568e-02, -6.98308274e-02,

4.76812609e-02,
-3.66885699e-02, -5.55331120e-03, -6.07383847e-02, -
2.84788832e-02,
-2.58521140e-02, -4.82337251e-02, -1.50086675e-02,
4.96456996e-02,
8.59433040e-02, -2.40951106e-02, 7.59972818e-03, -
5.54063125e-03,
-2.24091229e-03, -1.50544988e-02, 2.35055406e-02, -
2.44666822e-02,
4.36461065e-03, 1.53546163e-03, -4.36819494e-02,
9.68535058e-03,
-4.06657625e-03, -7.90959597e-03, 2.53057498e-02,
2.18446199e-02,
-1.32026216e-02, 9.09178890e-03, -9.98707395e-03,
2.41305120e-02,
5.03839515e-02, -2.11651577e-03, 6.31540716e-02,
2.35852357e-02,
5.46575561e-02, -2.88694296e-02, 6.14948720e-02,
1.38285069e-03,
-8.21907539e-03, -1.78447701e-02, 2.15052590e-02,
6.04183562e-02,
5.49643487e-02, -1.46802710e-02, 2.84566637e-02, -
4.97973561e-02,
1.37598952e-02, 5.89712597e-02, 1.15042226e-03,
3.49303288e-03,
-6.21531941e-02, 1.53838552e-03, -6.37759864e-02, -
8.91738907e-02,
7.82554373e-02, 8.47531948e-03, -4.98644542e-03,
1.01402784e-02,
1.51391430e-02, -8.83845761e-02, 2.05415487e-03, -
2.33477950e-02,
-2.05215532e-03, 1.11581078e-02, -3.35957855e-02,
9.26120952e-03,
1.51114054e-02, 2.93557737e-02, -5.72227985e-02,
3.65489721e-02,
-1.72995403e-03, 5.48442118e-02, -1.78670660e-02,
2.47619413e-02,
-1.37209510e-02, -2.43008118e-02, 2.20532995e-02,
1.46223726e-02,
-2.05053035e-02, 2.20507961e-02, -2.11495757e-02,
4.11115820e-03,
5.15158661e-03, 1.02750445e-02, -2.69110538e-02,
4.46728840e-02,
2.32053846e-02, -4.89358604e-02, 4.59032413e-03,
1.02164544e-01,
-9.75366868e-03, -3.17797177e-02, -6.53808415e-02,
1.61641818e-02,
-3.59605663e-02, -8.94023851e-02, 2.70413086e-02,
1.37799121e-02,

-4.45721811e-03, 1.04239220e-02, 2.71462556e-02, -
3.61309610e-02,
2.77329925e-02, 8.00317079e-02, -8.77075791e-02,
4.28875722e-03,
-4.00127061e-02, 6.91465363e-02, -1.89511906e-02,
2.25714799e-02,
-3.70223261e-02, -3.22204158e-02, -2.44717039e-02, -
2.87646651e-02,
-3.56792174e-02, 1.12627931e-02, 4.54415232e-02,
3.85652706e-02,
-1.26418313e-02, 1.50759565e-02, -2.54420396e-02,
7.14000361e-03,
2.42023785e-02, 1.95487700e-02, -2.00890638e-02, -
2.23372988e-02,
-3.12710390e-03, -3.15378122e-02, -5.47561655e-03, -
1.79029685e-02,
4.71970700e-02, -7.16794468e-03, 5.71047142e-02,
1.68435946e-02,
3.04741785e-02, -4.53626178e-02, 3.45488042e-02,
7.67176505e-03,
-4.31471877e-02, 3.08665317e-02, -3.06100193e-02, -
6.91321418e-02,
1.35168368e-02, -5.66082960e-03, -8.39459803e-03,
6.33223727e-02,
-1.16899852e-02, -6.68735355e-02, -3.14849801e-02, -
4.10801126e-03,
-3.11541669e-02, 5.09337857e-02, -1.81359183e-02, -
3.19656194e-03,
3.35075483e-02, -1.84867159e-02, 1.67975780e-02, -
4.19758074e-02,
-5.45852398e-03, -6.00223541e-02, 8.11602697e-02,
3.52576771e-03,
2.44264887e-03, -2.64091752e-02, 6.49582371e-02, -
2.48547969e-03,
-2.43240464e-02, 5.75116165e-02, 4.29937057e-02,
2.59116013e-02,
-2.59491825e-03, -5.10554425e-02, 4.08035628e-02, -
1.94113497e-02,
-1.92512535e-02, -4.25876193e-02, 5.27095124e-02,
3.85399908e-02,
1.73216704e-02, 7.26469830e-02, -6.12207502e-02, -
1.18507156e-02,
-6.27251267e-02, -7.87623376e-02, 8.93178582e-02,
2.94506289e-02,
-2.05500647e-02, -1.11666927e-02, 4.21139039e-02, -
1.29379909e-02,
6.52470291e-02, -3.63435335e-02, 4.70603295e-02, -
2.79442072e-02,
3.25778089e-02, 1.11311451e-02, 8.11562780e-03,

4.11903262e-02,
-6.30222782e-02, -2.10749824e-02, 5.48379384e-02, -
1.14307376e-02,
-6.74243178e-03, 2.63418630e-02, 5.54025136e-02, -
2.34036781e-02,
1.91544313e-02, 7.73659023e-03, -2.17517931e-02, -
4.22087871e-02,
1.82650238e-03, 1.14424359e-02, -2.49874536e-02, -
6.05609939e-02,
-4.06008698e-02, -2.75072195e-02, -1.80333741e-02,
2.35507153e-02,
-1.74065062e-03, -1.27423136e-03, -1.07636035e-03, -
9.68230143e-03,
-3.87231633e-03, -6.20300472e-02, -4.67399731e-02, -
1.56623609e-02,
1.97780654e-02, -4.04491797e-02, -1.43207898e-02,
4.10679653e-02,
-3.01404409e-02, -6.96904585e-03, 1.33907900e-03,
9.22414288e-03,
-3.65850031e-02, 1.55804958e-02, -3.82344448e-03, -
7.20928684e-02,
-1.00958999e-02, 2.16384437e-02, -9.67353117e-03,
5.69429956e-02,
-3.45044360e-02, 1.55994771e-02, -3.17699164e-02,
3.49897183e-02,
3.40582523e-03, -3.34586389e-02, 3.11984010e-02,
7.87884556e-03,
-5.38802333e-02, -3.60364094e-02, 1.63655505e-02,
4.92177205e-03,
-1.48115065e-02, -2.75984127e-02, -3.74488644e-02, -
1.08119827e-02,
3.21451481e-03, 8.67632180e-02, -1.38181914e-02, -
1.25437574e-02,
-7.70420879e-02, 6.22832496e-03, 1.95487123e-02, -
1.11719901e-02,
-1.01358537e-02, -2.13602968e-02, 5.03273718e-02,
3.63598089e-03,
4.75845113e-02, -2.42680241e-03, -3.39401402e-02, -
3.96677405e-02,
-5.88683970e-02, 2.37484146e-02, -2.60258932e-02, -
5.07525429e-02,
-2.05614790e-03, -5.64362705e-02, -2.20987108e-02,
2.71186847e-02,
-1.18467081e-02, -1.35338474e-02, -6.50236234e-02, -
1.27452547e-02,
-2.62039136e-02, -1.17040565e-02, 7.42463907e-03, -
1.80339795e-02,
2.29145754e-02, 6.12382218e-02, -3.17201689e-02, -
2.27250624e-02,

-4.17646244e-02, -6.16301633e-02, -7.98965432e-03,
1.36528462e-02,
-2.67598871e-02, -6.41485769e-03, 1.04835266e-02, -
5.75750264e-33,
3.33194919e-02, 2.61403583e-02, 1.67649947e-02, -
4.50293757e-02,
-4.47329246e-02, 8.38769078e-02, -5.20903654e-02, -
7.40614207e-03,
-5.42157609e-03, 2.11244784e-02, 5.37181795e-02, -
4.71830629e-02,
-3.57124908e-03, -4.53644395e-02, 2.47987323e-02, -
1.93949305e-02,
1.49388080e-02, -1.92033667e-02, 6.31379010e-03, -
4.99644615e-02,
-8.48017707e-02, -4.34698015e-02, 1.21582299e-02,
3.02123148e-02,
-6.12361096e-02, 4.84500229e-02, -3.22507508e-02, -
2.47880649e-02,
-2.96602473e-02, 1.41942352e-02, -4.02433472e-03, -
1.15736444e-02,
1.11253206e-02, 7.92052597e-03, -2.21044244e-03,
1.47663371e-03,
-1.05740065e-02, -8.26940686e-03, -7.66939530e-03, -
4.13465016e-02,
-2.50387527e-02, -3.08652576e-02, 3.01071908e-02, -
1.33677917e-02,
-3.80810685e-02, -8.80725086e-02, -4.68934095e-03, -
1.56831816e-02,
2.88823918e-02, 9.93877370e-03, 2.92213587e-03,
1.87477954e-02,
3.56616974e-02, 5.15914187e-02, -2.40304805e-02, -
1.99757945e-02,
-5.40524488e-03, -9.23534669e-03, 1.09303566e-02,
4.59112506e-03,
2.57769711e-02, 2.40028668e-02, 5.27625531e-03,
9.43361130e-03,
2.85520535e-02, -5.96446544e-02, -7.49772266e-02, -
1.73915830e-02,
-3.40901054e-02, -3.27642150e-02, 2.68281680e-02,
1.32728815e-02,
-3.16927731e-02, 3.54115851e-02, 3.13252509e-02, -
2.06042454e-02,
-2.16851104e-02, -1.20431778e-03, 5.17358594e-02,
5.11292592e-02,
-5.91437034e-02, 1.45989992e-02, -2.44136658e-02,
8.24424322e-04,
9.20693483e-03, 4.74843122e-02, -2.24387459e-02, -
5.15027083e-02,
-1.77996866e-02, 3.24994847e-02, -9.74478759e-03,

3.80564034e-02,
-6.81328326e-02, 2.30517369e-02, 4.93875407e-02,
4.83084936e-03,
4.87474985e-02, -2.49367338e-02, -7.78625440e-03, -
4.38477658e-02,
-5.06932326e-02, -1.74277052e-02, 3.42276804e-02,
1.76400621e-03,
1.07578961e-02, -8.04311037e-03, 4.00784090e-02, -
1.63924862e-02,
1.43053764e-02, -5.75796049e-03, 8.57711118e-03, -
1.13904681e-02,
-7.42665306e-03, -3.01627861e-03, 3.56619060e-02,
3.68044898e-02,
-5.64774871e-02, 1.06985696e-01, -6.33128881e-02, -
1.41638527e-02,
-5.86930476e-02, 3.33903730e-02, 4.00654189e-02,
1.77839324e-02,
-7.68069318e-03, -2.45577861e-02, 7.71692842e-02, -
1.18236523e-02,
-3.72238792e-02, -2.83971112e-02, -6.80442061e-03, -
6.47002691e-03,
3.06628067e-07, -6.63273176e-03, 3.35618258e-02, -
1.08633609e-02,
7.12244026e-03, 1.01321600e-02, -7.30070891e-03, -
1.40422843e-02,
2.52291490e-03, 7.72196846e-03, 6.35339413e-03,
7.44675994e-02,
5.18213445e-03, -6.13772385e-02, -2.24274304e-02,
1.16678201e-01,
-1.62313902e-03, 1.04808905e-05, 1.34706323e-03,
1.22738546e-02,
5.07250763e-02, -1.68990921e-02, 1.69979259e-02, -
8.61677062e-03,
-1.69992875e-02, 8.28906614e-03, -1.15082953e-02, -
2.99304109e-02,
-2.68067606e-03, -1.93803106e-03, -3.71186845e-02,
2.07976922e-02,
-3.25934477e-02, 4.31302227e-02, -7.50872046e-02,
1.83008821e-03,
-3.87935676e-02, -3.62090254e-03, 9.83081684e-02, -
4.39402051e-02,
2.19082236e-02, -3.03604994e-02, 2.33585984e-02, -
2.42799241e-02,
-9.79932491e-03, 1.84802935e-02, -6.71743900e-02,
1.31402845e-02,
-2.20641843e-03, -1.48688378e-02, -1.45312035e-02,
4.16271240e-02,
-6.37518289e-03, -7.63631100e-03, 3.37132141e-02,
5.30078337e-02,

```
3.84896919e-02, -1.93180591e-02, 2.32962575e-02,
4.23413031e-02,
-5.50938547e-02, -4.29240763e-02, 2.70183794e-02, -
1.50216976e-03,
2.11246721e-02, -2.19180435e-02, -4.66363020e-02, -
8.38574022e-03,
2.78810225e-34, -4.78660688e-02, 1.60858613e-02,
7.61870341e-03,
-3.21566612e-02, -6.36849785e-03, 2.35550193e-04,
4.42989655e-02,
5.77594005e-02, 2.83234864e-02, -2.31662057e-02, -
1.65172182e-02]),
'score': tensor(0.0075)},
{'page_number': 721,
'sentence_chunk': 'per day for a 2,000 kilocalories per day diet is
65 grams. Eighteen percent of sixty-five equals about 12 grams. This
means that 53 grams of fat are remaining in your fat allowance.
Remember, to have a healthy diet the recommendation is to eat less
than this amount of fat grams per day, especially if you want to lose
weight. Table 12.2 Updated DVs Based on a Caloric Intake of 2,000
Calories (For Adults and Children Four or More Years of Age)
Discovering Nutrition Facts | 721',
'chunk_char_count': 479,
'chunk_word_count': 87,
'chunk_token_count': 119.75,
'embedding': array([ 9.27221850e-02, 2.16329806e-02, 1.10049276e-
02, 2.99123544e-02,
6.94090724e-02, -2.25339122e-02, -3.95665616e-02,
3.73249091e-02,
1.13483973e-01, 2.18435358e-02, 6.26642182e-02,
8.79427243e-04,
-4.38181423e-02, 4.12645042e-02, 2.48401761e-02, -
4.13707718e-02,
2.95769069e-02, -5.32742217e-02, 2.82976907e-02,
6.51003001e-03,
-1.95953846e-02, -6.60562981e-03, 1.79668218e-02, -
6.67644711e-03,
-1.40732005e-01, 6.89874515e-02, 8.10105428e-02, -
2.92173177e-02,
-1.08888689e-02, -4.12420146e-02, -3.56114469e-02, -
3.59183587e-02,
1.28096398e-02, -2.85927602e-03, 2.24864084e-06,
2.70330696e-03,
3.46790366e-02, 2.05865745e-02, -2.69395001e-02,
4.92616370e-02,
1.52724730e-02, -9.34836641e-03, 2.88100471e-03, -
5.95693588e-02,
1.65926795e-02, 1.14903245e-02, 5.30601759e-03, -
3.24709974e-02,
```

3.31905819e-02, 9.09171719e-03, -8.51660594e-03, -
7.09288940e-02,
-2.12232713e-02, 3.36705334e-02, -6.68412447e-02,
1.36688929e-02,
-2.28314791e-02, 9.87464264e-02, -5.36817275e-02,
2.61150999e-03,
1.94554571e-02, 2.99588055e-03, 1.44326473e-02, -
7.64869433e-03,
-1.27583612e-02, 3.92002538e-02, -6.23405389e-02, -
2.64417902e-02,
2.64822952e-02, 5.89708015e-02, 7.70175979e-02,
5.44386320e-02,
-1.51527785e-02, 1.19869411e-02, -7.80064566e-03, -
1.48529084e-02,
4.99434844e-02, -1.18027041e-02, -2.74647120e-02,
2.81541459e-02,
2.59123780e-02, 1.45387638e-03, 2.68801786e-02, -
4.67261225e-02,
-6.17927797e-02, -5.95979989e-02, -2.76421010e-03,
2.12685689e-02,
1.04486663e-02, -5.67407757e-02, -1.01951525e-01, -
1.03741502e-02,
-1.99976694e-02, 4.50907229e-03, 6.46549538e-02, -
1.06081273e-02,
-1.81395531e-04, -1.61427539e-02, 4.28876393e-02, -
6.61453307e-02,
3.75715224e-03, -1.74156725e-02, 2.07847916e-02,
2.56542535e-03,
-3.49597423e-03, 3.01993731e-02, 6.77956641e-03, -
5.68012288e-03,
-2.16067163e-03, 2.56076660e-02, 5.29745258e-02, -
1.21361413e-03,
-3.55007052e-02, 4.49350663e-02, -3.26761082e-02, -
4.81037572e-02,
1.23666124e-02, -3.10597606e-02, -3.65988612e-02, -
1.90745434e-03,
3.93051617e-02, 8.67593754e-03, 1.53003959e-02, -
3.15324105e-02,
8.21884815e-03, 6.13727570e-02, 1.27845155e-02, -
2.78450139e-02,
5.35228364e-02, -6.12618402e-02, -1.21066561e-02, -
3.11183818e-02,
-1.17268339e-02, -1.91785078e-02, -6.58673467e-03,
4.11667787e-02,
4.12296364e-03, 4.28463221e-02, -1.08034536e-03, -
4.89056446e-02,
-1.00197382e-01, -2.95291394e-02, -2.37525608e-02, -
5.34034371e-02,
-6.11324385e-02, -2.48596743e-02, 2.19119675e-02,

4.72481214e-02,
2.03456031e-03, 8.21745582e-03, -2.23858096e-02, -
2.76126005e-02,
4.76855338e-02, 2.03735381e-02, 8.37765783e-02, -
3.95839587e-02,
2.39309315e-02, -5.30242280e-04, 1.05845919e-02, -
3.00434120e-02,
1.89973749e-02, 3.70955095e-03, -8.52513611e-02,
5.90053014e-03,
2.62904372e-02, -1.91088635e-02, -6.40271306e-02, -
5.48990853e-02,
1.21255048e-01, 4.64743152e-02, 4.00058329e-02,
2.36966424e-02,
-6.46515563e-02, -1.61830280e-02, 4.13507670e-02,
9.63915046e-03,
-2.14784686e-02, 7.36658201e-02, -2.04088334e-02,
3.04270945e-02,
6.22204430e-02, -2.96277590e-02, -5.45557588e-02,
1.33207319e-02,
-1.12059554e-02, -1.65693089e-02, -3.15201879e-02, -
5.50177880e-02,
1.36811230e-02, 1.87736284e-02, 3.95224616e-02,
3.42075117e-02,
2.25652326e-02, -6.87739402e-02, 5.29785268e-02,
4.11171243e-02,
-4.79559153e-02, -1.04143815e-02, 9.81871504e-03,
3.58377956e-02,
4.08127047e-02, 1.51536185e-02, 4.03549336e-02, -
5.45761697e-02,
-4.85272659e-03, 3.21601965e-02, 4.32179635e-03,
4.17468175e-02,
1.12442404e-01, -3.17270048e-02, 1.57602541e-02, -
1.20785087e-02,
2.28913873e-02, 1.40472669e-02, -2.11480185e-02, -
8.34635347e-02,
-1.77375060e-02, -4.00466891e-03, 5.12873149e-03,
2.42106188e-02,
-4.84343730e-02, 1.46427657e-02, 5.73469698e-03, -
4.84429859e-03,
4.90927584e-02, 2.20766198e-02, 3.18111456e-03, -
6.94620237e-02,
-3.43240760e-02, 1.33467773e-02, 2.30951849e-02,
1.44051900e-03,
-5.29850228e-03, -9.21280822e-04, 3.52666713e-02, -
1.82125159e-02,
-1.79636981e-02, -8.29175767e-03, -1.95534769e-02,
8.03007465e-03,
3.43859680e-02, -1.40607134e-02, -1.57111119e-02, -
1.24893645e-02,

-8.03530589e-02, -2.29489952e-02, -1.29773542e-02,
1.89668667e-02,
-4.10721712e-02, 9.09812599e-02, 6.47930987e-03, -
2.99287606e-02,
-6.16399478e-03, 2.28541028e-02, -6.58837706e-03, -
3.23797315e-02,
-3.02734859e-02, -8.49562697e-03, -7.06807652e-04,
7.24983111e-04,
-4.77860086e-02, -4.86666709e-03, -7.08321715e-03, -
3.47624570e-02,
2.49227118e-02, -2.01539546e-02, -2.14828439e-02, -
6.66145831e-02,
-4.21951115e-02, -2.62893252e-02, 3.82841825e-02,
4.08926345e-02,
-1.91992484e-02, 8.35808739e-03, -4.93414793e-03,
2.50924774e-03,
2.96902992e-02, -9.71822347e-03, 1.95244339e-03,
1.25632454e-02,
4.46360968e-02, -5.45745948e-03, 3.00297253e-02,
1.96747612e-02,
1.78909991e-02, -4.97255512e-02, 5.20193353e-02, -
1.75768584e-02,
5.02832457e-02, -3.03732771e-02, 2.79323533e-02, -
7.73544936e-03,
-8.93973559e-03, -2.65986156e-02, -1.88885108e-02,
1.13721227e-03,
1.05881821e-02, 5.30267740e-03, 3.29419300e-02,
1.25292474e-02,
1.89972259e-02, -2.05729716e-02, -5.20818122e-02, -
1.48135070e-02,
1.53122726e-03, 3.71819572e-03, -5.72241805e-02, -
1.44222961e-03,
-1.29801715e-02, 7.31452480e-02, 6.71170056e-02,
5.64042032e-02,
4.98053618e-02, 1.46106677e-02, -3.07483505e-02,
2.72446871e-02,
2.46273875e-02, 4.49993014e-02, -1.24894455e-02, -
8.52801546e-04,
2.13766098e-02, 4.64847982e-02, -3.16917710e-02,
8.39706138e-02,
-1.03648193e-02, 8.66125077e-02, 3.17521058e-02,
9.43081155e-02,
-4.38662656e-02, -1.80321317e-02, 1.77193414e-02,
2.53259260e-02,
2.53587458e-02, -2.77113263e-02, -2.95802876e-02, -
1.72663499e-02,
1.06222427e-03, -2.37804540e-02, -7.55738020e-02,
5.10420278e-03,
-6.12998754e-02, -1.29003981e-02, -1.62730459e-02,

5.39068393e-02,
-2.60359589e-02, 1.17564006e-02, -2.17533708e-02,
8.33499871e-05,
-2.47579697e-03, -8.52259994e-03, -1.89275797e-02, -
4.12640944e-02,
2.47047376e-02, -2.71462183e-03, 3.05016600e-02, -
3.16401497e-02,
4.81277704e-02, 1.52151659e-02, -3.71768400e-02,
2.67823637e-02,
3.09996475e-02, 7.95336291e-02, -5.83096370e-02,
6.63912296e-02,
3.38688083e-02, -2.33242363e-02, -2.47472525e-02,
5.69370016e-02,
-5.39264865e-02, 5.37954969e-03, 6.64168522e-02, -
3.93996201e-03,
-3.39828208e-02, 1.18057663e-02, -2.76038293e-02, -
2.78863665e-02,
-2.92007253e-03, 6.51494926e-03, -6.45447895e-02,
1.46629876e-02,
3.19862552e-02, -3.19613237e-03, -7.10667809e-03, -
9.80173610e-03,
-8.01385287e-03, -2.87840869e-02, -1.07424883e-02,
4.64275666e-02,
-1.56148579e-02, -9.28422809e-03, 8.33946839e-03, -
3.83385122e-02,
-2.31447141e-03, -4.42682579e-02, -5.15981056e-02, -
4.83378656e-02,
-4.61712033e-02, 1.91345513e-02, -7.75922323e-03,
1.22245979e-02,
1.89334489e-02, -5.81079349e-02, 3.61762457e-02,
5.71719334e-02,
1.87886003e-02, 7.21923634e-02, -3.05792093e-02,
5.89319831e-03,
-1.40418559e-02, 3.91000472e-02, 6.70091808e-02, -
1.94349000e-03,
-6.54065758e-02, -2.88703237e-02, 5.25484085e-02,
1.64745245e-02,
-2.40662191e-02, -4.11188789e-02, 2.28048153e-02, -
7.31063494e-03,
5.70656992e-02, 3.22585460e-03, 1.00973863e-02,
5.06031141e-02,
-1.57971531e-02, 8.43167491e-03, -8.19626171e-03, -
4.35131416e-03,
-5.63936634e-03, 4.78217602e-02, 9.04240366e-03,
4.92252745e-02,
-3.02272737e-02, 3.27916183e-02, -2.47935182e-03,
2.87322816e-03,
-9.50592663e-03, -3.66510190e-02, 9.51997191e-02,
8.27850867e-03,
-2.83439900e-03, -3.21262516e-02, 2.67142337e-02,

1.88730434e-02,
6.12481311e-03, 9.99960396e-03, -2.29879785e-02,
5.70884300e-03,
6.23750612e-02, -1.31181348e-02, 1.55044487e-02, -
1.55170832e-03,
-3.99651155e-02, 5.89613356e-02, 3.69863063e-02, -
1.89057011e-02,
6.87801763e-02, -5.97071312e-02, -1.08832605e-02, -
6.71738386e-02,
5.11221923e-02, -3.01593100e-03, 3.92084979e-02, -
5.64722950e-03,
-1.83843914e-02, -3.06604672e-02, 4.44404297e-02, -
1.99432019e-02,
1.69736869e-03, 1.05079329e-02, -1.87007468e-02,
4.22680154e-02,
1.74694452e-02, 1.31196203e-02, -3.31872702e-02,
2.53353198e-03,
-2.65736375e-02, -4.59273569e-02, 7.30829639e-03, -
3.36142001e-03,
1.64700998e-03, -3.32064629e-02, 3.18385959e-02,
1.31355813e-02,
-6.50637746e-02, -4.53776941e-02, -1.83167029e-02,
5.18230535e-03,
-7.14196861e-02, 1.26210311e-02, -5.50324749e-03, -
8.00061133e-03,
-1.83840524e-02, -8.40920135e-02, 4.17542867e-02, -
2.84715230e-03,
1.58247054e-02, -6.22384846e-02, 1.39971972e-02,
4.84491363e-02,
3.23056765e-02, -5.06911688e-02, -2.19277833e-02, -
1.60791688e-02,
2.07391270e-02, -2.24667974e-02, 3.92552130e-02, -
3.41030285e-02,
2.20379792e-03, -5.16364872e-02, -1.42322788e-02, -
3.20745334e-02,
1.00499932e-02, 2.21894588e-02, 7.42902805e-04,
2.76882062e-03,
-2.49375659e-03, -8.65277741e-03, -1.78817548e-02,
1.08791925e-02,
2.10783794e-03, -3.80136482e-02, 4.36368175e-02,
1.84071977e-02,
4.43462133e-02, 2.02036500e-02, -2.60853693e-02,
9.82186757e-03,
-3.55803892e-02, 1.35324178e-02, -7.72904977e-03, -
2.16953941e-02,
-5.60718514e-02, -1.88582540e-02, -4.49427543e-03,
2.74920766e-03,
-2.88643260e-02, 5.98050794e-03, -2.29939111e-02,
1.14609161e-02,

-3.31028104e-02, -5.75678200e-02, -2.45882645e-02,
4.01665494e-02,
6.44219806e-03, -1.00851879e-02, -6.49554655e-02, -
2.82202940e-02,
3.72044444e-02, -2.20657215e-02, -1.25018554e-02,
1.96691579e-03,
1.12728067e-02, 1.87679008e-02, -1.75459180e-02, -
6.16744784e-33,
4.50422056e-02, 1.93523932e-02, -1.16658630e-02,
8.24827030e-02,
-1.06211519e-02, 2.04151273e-02, -1.85519401e-02, -
5.53031266e-03,
-3.31761688e-02, 3.24848443e-02, -2.74462588e-02, -
1.16297733e-02,
1.14466175e-02, -6.06080182e-02, 2.09262576e-02,
1.43812904e-02,
8.81081633e-03, -3.57252695e-02, -8.21033213e-03, -
4.05451516e-03,
7.97325838e-03, -2.10851952e-02, 4.44093607e-02, -
1.22938892e-02,
3.18571441e-02, 4.80241366e-02, 2.68018022e-02, -
2.53502633e-02,
-4.24023271e-02, 5.87633392e-03, -5.37818000e-02,
7.32877627e-02,
-7.63983792e-03, 5.41893691e-02, -5.11945924e-03, -
8.36512595e-02,
-3.02082896e-02, -6.23245314e-02, 3.19395550e-02, -
7.83652009e-04,
9.12475586e-02, 1.18450308e-02, -4.71820030e-03, -
6.93552010e-03,
-5.51233999e-02, 4.78613153e-02, 4.21923958e-02,
1.71369016e-02,
-5.10903150e-02, 7.79179558e-02, 2.72365753e-02,
2.93563269e-02,
3.74478884e-02, 6.18300885e-02, -1.63149822e-03, -
2.52072960e-02,
2.45564878e-02, -6.45351335e-02, 3.95056233e-02, -
2.44942866e-02,
2.95077767e-02, 5.32330945e-04, -7.91045092e-03,
1.78596918e-02,
2.74041425e-02, 7.94659462e-03, -3.56858522e-02,
4.21319641e-02,
-1.23899868e-02, 4.47214209e-03, -2.67322902e-02, -
3.66243944e-02,
-3.27749662e-02, 7.99678674e-04, 1.47349071e-02, -
6.77852929e-02,
-6.67833118e-03, 1.10030901e-02, -6.12083301e-02, -
1.68561898e-02,
-3.61517407e-02, 1.17764352e-02, 5.22977039e-02,

2.78210896e-03,
7.09813740e-03, 3.45364287e-02, 4.26800369e-04, -
2.21867138e-03,
-1.31678896e-03, -1.63319577e-02, -7.29316520e-03,
1.26398150e-02,
-1.92886945e-02, -1.91358533e-02, 3.66767272e-02, -
2.24532168e-02,
7.26784542e-02, -2.68202219e-02, -1.63082127e-02, -
1.36795752e-02,
-1.90507323e-02, -3.33667733e-02, 4.16796692e-02,
2.26684716e-02,
1.54701294e-02, 1.32027483e-02, -1.41395023e-02, -
4.19428423e-02,
1.76874660e-02, -4.72256988e-02, -2.03945376e-02, -
2.58022752e-02,
-1.60495657e-02, 3.06938682e-02, -7.21600931e-03,
1.19558619e-02,
-1.02561116e-02, -3.68442982e-02, -2.78850421e-02,
3.23052481e-02,
-6.99605234e-03, -2.83085890e-02, 6.75542001e-03, -
3.57422940e-02,
1.17814150e-02, -5.82384504e-02, 1.50745800e-02, -
8.49434063e-02,
-9.08695087e-02, 4.98430282e-02, 5.10465447e-03, -
1.55331828e-02,
2.81932927e-07, -5.64444903e-03, -7.53646418e-02, -
1.45254016e-04,
-6.03242917e-03, -1.28579000e-02, -3.45701724e-02,
3.56136896e-02,
-2.21062135e-02, -1.65432412e-02, -3.27612683e-02,
6.75019389e-03,
-2.01953482e-02, -4.33118045e-02, -5.56853823e-02,
3.96705233e-02,
6.16272353e-02, 4.72552367e-02, -3.04447338e-02, -
2.47016307e-02,
2.05192752e-02, 3.27008613e-03, -5.58481971e-03,
1.69282556e-02,
-2.12929528e-02, 2.86379317e-03, -1.35601228e-02, -
4.53396933e-03,
-1.74272545e-02, 6.87959231e-03, -5.51140569e-02,
9.75974575e-02,
-2.11705714e-02, 1.71814132e-02, -2.74816658e-02, -
5.13436366e-03,
5.30210789e-03, -5.72582707e-02, 8.84137377e-02, -
2.45237574e-02,
6.66355416e-02, -1.04880156e-02, 3.31256837e-02,
1.73153915e-03,
-6.34004548e-02, -1.20344209e-02, -3.34654227e-02, -
2.40100012e-03,

```
4.43789065e-02, 4.34839949e-02, -7.61676282e-02, -
5.50979655e-03,
-1.15044722e-02, 2.91561801e-02, 7.22636376e-03,
2.18843739e-03,
-2.54339334e-02, 5.78873008e-02, -1.94613319e-02,
1.83735024e-02,
-4.37594168e-02, -1.12463292e-02, 1.05501134e-02,
2.32923031e-02,
3.73498313e-02, -5.46663906e-03, 7.18501508e-02,
6.14101999e-02,
1.92057100e-34, -5.86840995e-02, 1.31822331e-02, -
1.46912225e-02,
1.87671755e-03, -1.09619778e-02, 4.59390171e-02,
1.58940740e-02,
1.04586175e-02, -3.30785173e-03, -3.94901410e-02,
2.79858185e-04]),
'score': tensor(0.0073)},
{'page_number': 162,
'sentence_chunk': 'This must be balanced with water input. Our
tissues produce around 300 milliliters of water per day through
metabolic processes. The remainder of water output must be balanced by
drinking fluids and eating solid foods. The average fluid consumption
per day is 1.5 liters, and water gained from solid foods approximates
700 milliliters. Figure 3.5 Daily Fluid Loss and Gain 162 | Regulation
of Water Balance',
'chunk_char_count': 406,
'chunk_word_count': 66,
'chunk_token_count': 101.5,
'embedding': array([ 4.00138348e-02, 8.40967428e-03, -4.74889716e-
03, 8.18357989e-03,
1.63441654e-02, -1.10073928e-02, 5.32569457e-03,
3.50633599e-02,
-1.93869218e-03, -1.25277992e-02, -6.24015555e-03, -
2.70804185e-02,
2.96617136e-03, 3.03324554e-02, 2.15500500e-02, -
2.85466816e-02,
-1.45524191e-02, -4.97655682e-02, -6.12546988e-02,
1.26296105e-02,
2.39711329e-02, -6.00292720e-03, 5.13560744e-03, -
2.76806988e-02,
6.00644667e-03, 5.92583492e-02, 8.04003626e-02,
7.90159497e-03,
-3.44571136e-02, -3.35089453e-02, -3.83616984e-02, -
9.64098889e-03,
2.19733380e-02, 3.66719104e-02, 1.59787191e-06, -
5.58938570e-02,
3.63625325e-02, 3.23399827e-02, -2.79929731e-02,
1.90422907e-02,
3.66204679e-02, -6.33106679e-02, 1.38122179e-02,
```

6.51811855e-03,
4.18509915e-02, 8.22869316e-02, -7.86241703e-03,
1.27171399e-02,
2.02644765e-02, 1.27428100e-02, -9.45056416e-03, -
1.81036871e-02,
5.33691421e-03, 3.67215909e-02, 2.12841798e-02,
1.15260296e-02,
1.24996258e-02, 7.62541443e-02, 3.54682766e-02, -
2.76245624e-02,
-1.35651953e-03, 9.60237905e-02, -1.21389851e-02, -
6.64809998e-03,
1.93772018e-02, 6.60238117e-02, 6.24834634e-02, -
7.32496604e-02,
2.21758317e-02, 3.07615716e-02, 6.75983280e-02,
2.54786331e-02,
2.90350374e-02, -3.90379876e-02, 2.15656217e-02, -
6.70337379e-02,
5.89775573e-03, 4.16788086e-02, 1.17903657e-03, -
4.55860086e-02,
-2.05767136e-02, -1.22198751e-02, -1.15062296e-03, -
9.41443257e-03,
-2.17595138e-02, -6.54117465e-02, 1.25232097e-02,
1.46664577e-02,
-3.27520482e-02, -2.65380517e-02, -4.01514029e-04, -
2.29633227e-02,
2.95287482e-02, 6.48923963e-02, 1.35014560e-02, -
3.00552454e-02,
2.61465218e-02, -6.75403178e-02, 3.81037481e-02, -
6.19100332e-02,
1.11321108e-02, -2.38415636e-02, -2.04820596e-02, -
7.81487487e-03,
1.54193668e-02, 2.99781393e-02, -3.50451842e-02,
3.67404595e-02,
-2.89164186e-02, -1.55018056e-02, 1.10776396e-02, -
4.31468524e-02,
2.55151577e-02, 1.11059509e-02, -5.04232347e-02,
1.58578344e-02,
-4.04068455e-02, 2.14510350e-04, -1.10734403e-02, -
2.41999384e-02,
5.04975989e-02, 2.07321923e-02, 4.66382504e-02, -
1.24869049e-02,
-1.02035087e-02, 7.81157762e-02, 1.40776569e-02,
1.84325930e-02,
2.16092914e-02, -3.66792828e-02, 1.26403552e-02, -
3.46630160e-03,
-1.18705835e-02, -3.18370829e-03, -2.12932788e-02,
5.04186377e-03,
3.38738784e-02, 5.97977825e-02, -1.69556979e-02, -
4.86374833e-02,

-1.57648958e-02, -2.89737321e-02, -7.28122965e-02,
2.35669576e-02,
7.08528049e-03, -3.73566896e-02, 2.91597452e-02, -
3.81919951e-03,
4.33747191e-03, 7.96045549e-03, 4.28863764e-02, -
3.95828336e-02,
3.52379195e-02, 5.62982187e-02, 9.14041400e-02, -
4.46818992e-02,
2.73088552e-02, 9.48518328e-03, 7.78790638e-02,
4.87898150e-03,
-3.62588987e-02, -1.80249531e-02, -3.22327316e-02, -
1.15913069e-02,
-8.56134924e-04, 1.88142452e-02, -1.51003320e-02,
1.64637563e-03,
-3.98606900e-03, 4.85035852e-02, 1.04018040e-02,
5.64982742e-02,
-2.85101775e-02, 2.13030465e-02, 1.27219828e-02,
4.76542115e-03,
4.88765091e-02, 1.05379731e-03, 2.28598360e-02, -
3.39606293e-02,
6.93420097e-02, -3.42231803e-02, -2.63606906e-02,
5.62466048e-02,
6.31097108e-02, 6.66700955e-03, -2.01039612e-02, -
1.66154150e-02,
-6.25436530e-02, 1.54254911e-02, 4.27125953e-02, -
1.88358221e-02,
-3.83668654e-02, -4.15610746e-02, 1.55199012e-02, -
1.61423546e-03,
4.57981043e-02, 3.00043356e-02, -1.82170290e-02,
1.36734461e-02,
5.18917628e-02, -7.88673107e-03, 8.84324685e-03, -
1.06522506e-02,
-2.86358949e-02, 1.52096534e-02, 3.95640098e-02, -
1.41508011e-02,
5.96988685e-02, -2.98389029e-02, 6.67763315e-03, -
6.12750137e-03,
4.76159193e-02, 3.65708880e-02, 2.59223059e-02, -
4.75215279e-02,
-1.23025766e-02, -1.02760335e-02, -7.99182430e-02,
4.86385264e-02,
1.56522337e-02, -1.81025602e-02, -1.17085641e-02, -
2.44552251e-02,
-3.71456556e-02, 1.82506861e-03, 6.07503578e-02, -
3.47852446e-02,
-5.67865409e-02, -4.11696285e-02, 3.32207642e-02,
5.21982573e-02,
-3.23219746e-02, 2.77320147e-02, -9.84611828e-03,
5.04443236e-02,
1.20612318e-02, 4.63778339e-02, -2.45011002e-02,

8.81251991e-02,
-7.77402986e-03, -6.56568483e-02, -2.25031804e-02, -
2.63640285e-02,
-4.41864654e-02, -1.47502478e-02, 8.56571924e-03, -
1.56116590e-03,
-1.15460632e-02, 5.01703732e-02, -6.12773234e-03, -
2.95702051e-02,
-7.77077600e-02, 2.29723006e-02, -1.74407363e-02, -
2.80258600e-02,
-1.70334172e-03, -2.01444887e-02, -1.07705044e-02,
9.06226318e-03,
-4.38157422e-03, -1.57056227e-02, 1.17843030e-02,
5.20449951e-02,
2.03577578e-02, -1.46904122e-02, 3.60052247e-04, -
5.37365675e-03,
-3.78400348e-02, -1.08746598e-02, 3.66784558e-02,
2.87038647e-02,
4.88392301e-02, 3.76562215e-02, -2.32061651e-02, -
3.65302339e-02,
-4.47302349e-02, 4.87561198e-03, 5.62468767e-02,
6.38187230e-02,
1.50010232e-02, 4.07337658e-02, 2.11753231e-02,
4.52376381e-02,
-2.35440191e-02, -1.41760660e-02, 6.65341094e-02, -
9.85762291e-03,
7.73032606e-02, -2.00233795e-02, 1.86871905e-02, -
3.43697928e-02,
1.16805471e-02, 3.84600833e-02, -4.41833213e-03,
4.42278534e-02,
6.89311847e-02, 2.22036173e-03, 4.09494415e-02, -
6.11298606e-02,
1.37321716e-02, 4.99818139e-02, -4.55303341e-02, -
7.77084089e-04,
-1.09550841e-02, -7.70321256e-03, -4.37128879e-02, -
4.85754684e-02,
2.47350638e-03, 2.36401726e-02, -3.29314545e-02,
4.57753800e-02,
3.40519845e-02, 2.65858378e-02, -1.27030518e-02, -
1.96590275e-02,
-4.06386890e-02, 6.66673109e-02, -4.33624461e-02,
1.54710142e-02,
-7.98849948e-03, 2.18005907e-02, -1.72416791e-02,
5.96818998e-02,
2.01102253e-02, 1.87646318e-02, 4.74744365e-02, -
2.93383631e-03,
-5.25753722e-02, -6.51266724e-02, 1.64429881e-02, -
1.76222250e-02,
-4.38837186e-02, -3.37668061e-02, -3.26648317e-02, -
1.61915217e-02,

-2.71703005e-02, 2.07302924e-02, -3.98334228e-02,
1.39016919e-02,
-1.02013396e-02, -2.03092191e-02, -2.50824951e-02,
4.57703918e-02,
3.84520181e-02, 1.20676542e-02, -1.14854993e-02, -
4.35052365e-02,
-5.63465394e-02, -7.09133968e-02, 2.38874555e-02, -
2.84090526e-02,
-1.13173667e-02, -2.48512235e-02, -8.92541558e-03,
3.82443927e-02,
-1.77319907e-02, 7.60329962e-02, -4.92553860e-02, -
1.42427590e-02,
7.02468306e-02, 7.12311119e-02, 1.83479004e-02,
2.56307051e-02,
8.46557543e-02, 4.63513508e-02, -2.90670190e-02,
1.00521512e-01,
5.96188195e-03, -3.88456173e-02, 2.66397260e-02,
3.95019352e-02,
-3.44923362e-02, 3.69138084e-02, 9.47101973e-03, -
1.87893678e-02,
-1.80927422e-02, -1.08159734e-02, -3.96270491e-02, -
1.06564285e-02,
3.16749997e-02, -4.94665168e-02, 5.89515716e-02,
1.14874274e-03,
1.07920486e-02, 1.43555747e-02, 3.30712423e-02, -
3.47117893e-02,
6.42326698e-02, -9.29194763e-02, 3.84764336e-02,
7.89776072e-03,
-3.05623910e-03, -3.11164148e-02, -4.00474637e-05, -
3.27966399e-02,
7.04330346e-03, -3.71515215e-03, -3.94952483e-03,
2.63848696e-02,
3.31274867e-02, -6.40999302e-02, -4.14862595e-02, -
2.02604234e-02,
1.65905859e-02, -1.32529456e-02, -5.71435615e-02,
2.84331441e-02,
4.85863723e-02, 1.64241139e-02, 8.64014700e-02,
7.87127949e-03,
-1.21114347e-02, -3.70124634e-03, 1.21565852e-02,
3.72672896e-03,
3.79049359e-03, -6.69959933e-02, 3.22900563e-02, -
6.88641891e-02,
3.39205079e-02, 2.53037130e-03, 9.54560656e-03,
4.58374135e-02,
-9.36222915e-03, -1.21130599e-02, -4.52996939e-02,
6.99941367e-02,
-1.57837216e-02, 5.73645858e-03, 1.80341620e-02,
2.16885731e-02,
-1.18325744e-02, 5.81339654e-03, -4.08831462e-02,

1.77561399e-02,
1.55667076e-02, -4.81638201e-02, 2.32461244e-02, -
7.78994849e-03,
5.04258983e-02, -4.21747006e-02, 8.64678342e-03,
5.25503196e-02,
4.54535671e-02, 3.31310742e-02, -7.49144107e-02,
8.81401971e-02,
1.50486547e-02, -3.20088752e-02, 2.01319866e-02,
3.31842899e-02,
-4.98411991e-02, 7.90795982e-02, 1.09081501e-02, -
2.68958434e-02,
3.44293863e-02, -9.56957489e-02, -2.03482639e-02, -
5.53723052e-02,
3.29092913e-03, -8.58669542e-03, -4.04922925e-02, -
3.09264138e-02,
1.94096267e-02, -4.47220691e-02, -5.32498956e-02, -
5.49050048e-02,
1.51924277e-02, -2.89040469e-02, -5.15314564e-03,
3.56023461e-02,
1.47490371e-02, 3.30763706e-03, 1.62407625e-02,
1.67671666e-02,
1.35294860e-02, -2.61633806e-02, -1.99028011e-02,
3.86212841e-02,
3.17764990e-02, -7.08085001e-02, 2.50429995e-02,
4.64984099e-04,
-2.58999728e-02, -2.35338788e-03, -3.37015651e-02, -
3.24340761e-02,
-3.30382623e-02, 9.29965358e-03, 6.12277727e-05,
7.60091562e-03,
-3.25982571e-02, -1.17949881e-02, -3.38017866e-02, -
8.66714306e-03,
-1.57930814e-02, -2.27811486e-02, 5.66104501e-02, -
1.38015952e-02,
-4.67561632e-02, 2.53697410e-02, -9.52643750e-04, -
3.54826674e-02,
9.91498586e-03, 3.43100354e-02, -5.95959648e-03, -
5.22596203e-02,
-4.85362895e-02, -4.96939868e-02, 1.55678829e-02, -
1.81043614e-02,
-2.25038938e-02, 4.91393842e-02, 1.02943517e-02,
9.55975428e-03,
-3.95842679e-02, 2.37024315e-02, -3.50546837e-02,
6.99748937e-03,
3.47177908e-02, -6.75244704e-02, 1.67895481e-02, -
7.20973536e-02,
1.11556694e-01, -3.95566933e-02, -1.73641916e-03,
1.93900354e-02,
-6.84009641e-02, 2.75546964e-03, 1.35591486e-03, -
2.61125118e-02,

-3.79417017e-02, 5.66921290e-03, 2.89330073e-02,
5.38344495e-02,
-2.01345645e-02, 2.59401258e-02, -1.53928632e-02, -
2.88418662e-02,
-2.26649605e-02, -2.12773494e-02, -2.61000898e-02,
2.55563706e-02,
-3.72653976e-02, 8.84196442e-03, -8.73713270e-02, -
2.98065674e-02,
2.30656238e-03, -1.94808189e-02, -3.10108420e-02,
4.84666154e-02,
-1.12046553e-02, -2.24062484e-02, -1.06285075e-02, -
4.65447678e-33,
2.57644933e-02, 2.83418433e-03, 3.18993628e-02,
1.48958033e-02,
-4.67297854e-03, 8.30766633e-02, -2.88200006e-02,
1.20947054e-02,
-4.24745977e-02, 3.04395910e-02, 3.16925719e-02, -
2.27887947e-02,
3.07357665e-02, -5.30289374e-02, 2.04950217e-02,
2.53729755e-04,
3.41598652e-02, -5.57328537e-02, 1.40830530e-02,
3.06739062e-02,
3.70755643e-02, -3.43192965e-02, -4.24876297e-03, -
2.93049831e-02,
-3.05281114e-02, 2.59564840e-03, 2.16824599e-02, -
2.22756770e-02,
-4.33119126e-02, -4.66292761e-02, 4.21504937e-02,
7.09525943e-02,
-2.13589743e-02, 2.77723968e-02, -4.68936712e-02,
1.85637642e-02,
-3.90444733e-02, -3.39676552e-02, -8.47923849e-03, -
4.24652174e-02,
-2.34933402e-02, 2.96864789e-02, 2.89743990e-02, -
5.72932810e-02,
-2.58265603e-02, 2.82809362e-02, -1.35860071e-02,
3.37537471e-03,
-3.58743221e-02, -3.48318554e-02, 2.09584199e-02,
2.14414820e-02,
-2.59310026e-02, 4.26133499e-02, -1.18576502e-02, -
2.11551376e-02,
2.48001032e-02, -1.98712219e-02, 8.90038069e-03,
2.48948252e-03,
1.86661340e-03, -2.50989012e-02, -2.51904801e-02, -
2.73392014e-02,
4.11848202e-02, -1.63577907e-02, -1.26895001e-02,
1.80799421e-02,
-1.24124503e-02, -5.30730523e-02, 1.96817182e-02, -
1.99919716e-02,
-4.90626730e-02, -2.67370809e-02, 1.51076382e-02, -

8.75961706e-02,
4.82801683e-02, 4.87223966e-03, -2.00448968e-02,
3.01626194e-02,
-4.28887159e-02, 6.78627566e-03, 3.97803113e-02,
1.22535210e-02,
1.10876448e-02, 4.43424806e-02, -7.67383911e-03,
2.29561366e-02,
-6.91218972e-02, -2.61024572e-02, 5.51583543e-02,
1.69745926e-03,
4.04876471e-02, -2.03708168e-02, 1.07351718e-02,
2.11251173e-02,
5.87565005e-02, -4.86359037e-02, -1.11243194e-02,
2.56506000e-02,
-5.26938550e-02, 1.72657464e-02, 1.02744773e-02,
1.17877489e-02,
-2.44658105e-02, 3.76250856e-02, -1.45931635e-02, -
3.47916596e-02,
2.98940651e-02, -3.76444757e-02, -1.37753338e-02, -
1.44788600e-03,
-2.58896463e-02, 5.37551939e-02, -4.33056280e-02,
3.41862179e-02,
-1.71987675e-02, -4.07444164e-02, -4.14359160e-02, -
3.75101753e-02,
-8.44071284e-02, 3.95236276e-02, 1.08514596e-02,
3.59399430e-02,
-9.38009843e-03, -3.15908305e-02, -1.69529468e-02,
3.65442596e-02,
1.44240456e-02, 5.73350452e-02, 3.99551541e-02, -
1.36672100e-02,
2.46660022e-07, 1.60439045e-03, -7.46704414e-02, -
2.80123428e-02,
-3.87479477e-02, 5.51121421e-02, -4.68657948e-02, -
2.88587492e-02,
-1.91637389e-02, -2.11252011e-02, -9.59968790e-02,
5.17760552e-02,
8.50210618e-03, -8.60346947e-03, -4.44892459e-02,
7.67183229e-02,
-2.99341884e-02, -3.79752740e-02, -2.59251427e-02,
5.88115938e-02,
-1.21687055e-02, -1.73389111e-02, 4.25486010e-04, -
1.51779754e-02,
-4.58589345e-02, -7.87455868e-03, -1.67022869e-02, -
7.85031822e-03,
3.17677041e-03, 5.96983619e-02, 2.29938105e-02,
3.07214148e-02,
-8.58849380e-03, 1.10692009e-01, -1.68646686e-02, -
4.12975904e-03,
2.32666358e-03, -9.23847686e-03, 6.31603003e-02, -
8.49851668e-02,
-1.21385660e-02, -5.40114269e-02, 8.51599649e-02, -

```
2.29565352e-02,
-4.73814532e-02, -6.24898113e-02, 4.10238691e-02,
5.36435656e-02,
-6.76767230e-02, 5.66653609e-02, 3.08007468e-02,
5.45141920e-02,
1.46662099e-02, -2.45908014e-02, 1.73734948e-02, -
1.78254442e-03,
1.63357165e-02, 2.75170878e-02, -1.04411524e-02,
1.45886205e-02,
-9.57116112e-02, -3.75255719e-02, -1.15745645e-02, -
2.18214355e-02,
3.71213350e-03, 1.37708886e-02, 2.27147881e-02,
4.60968465e-02,
1.28356085e-34, -3.57433483e-02, -8.28674249e-03, -
1.07088350e-02,
2.20138021e-02, 2.07622610e-02, 7.95998797e-03, -
4.39422801e-02,
9.19142459e-03, 4.84595308e-03, -5.31846806e-02,
9.26750014e-04]),
'score': tensor(0.0070)},
{'page_number': 911,
'sentence_chunk': '82.http://www.cdc.gov/
nchs/data/databriefs/db82.pdf. Published January 2012. Accessed
December 8, 2017. Young Adulthood | 911',
'chunk_char_count': 126,
'chunk_word_count': 13,
'chunk_token_count': 31.5,
'embedding': array([-3.81647274e-02, 2.82944031e-02, 1.09553682e-
02, -2.30962932e-02,
2.31072889e-03, 8.61893222e-03, 8.13498646e-02, -
1.89929008e-02,
8.54682364e-03, 4.05666279e-03, 3.78592871e-02, -
3.33000459e-02,
1.34231700e-02, -1.52424835e-02, -3.19821276e-02,
2.38513444e-02,
3.11880209e-03, 8.73223785e-03, 3.66903991e-02,
2.28766147e-02,
-1.00683413e-01, 1.96307749e-02, -2.74691880e-02,
4.04555630e-03,
-1.98116172e-02, 2.37106550e-02, 3.15384194e-02, -
2.38909721e-02,
1.38506489e-02, -7.68969581e-02, 6.77377433e-02, -
2.27049943e-02,
5.08618578e-02, -1.38740707e-02, 1.89576656e-06, -
2.53552459e-02,
3.52712981e-02, 9.04523954e-03, -2.49081943e-02, -
3.54274525e-04,
-1.40851960e-02, -2.38370104e-03, -2.09503155e-02,
1.70942489e-02,
```

7.09094899e-03, -8.28211382e-02, 8.47525802e-03,
4.11683545e-02,
-1.84997860e-02, -5.34476601e-02, 7.83398468e-03, -
2.97738891e-02,
-3.62266973e-02, -3.35794501e-02, -1.22699598e-02,
8.78531933e-02,
-1.95038803e-02, 3.78628820e-02, 1.31190568e-02, -
8.64313990e-02,
3.45239192e-02, 1.97657626e-02, -3.89806204e-03,
1.79781541e-02,
4.92971996e-03, 1.36619294e-02, -2.21358314e-02, -
4.63804603e-02,
2.81502493e-02, 4.10075560e-02, 4.90877926e-02,
5.68986014e-02,
6.28326461e-03, 1.31605081e-02, -4.01126267e-03, -
3.17669148e-03,
-3.45365219e-02, -4.31026183e-02, 3.69035243e-03, -
3.50670442e-02,
-1.52968953e-03, 8.15721899e-02, 1.67327896e-02, -
5.97779453e-02,
2.35380847e-02, 1.19130937e-02, -8.69690254e-03, -
2.60573272e-02,
3.86032127e-02, -7.15213940e-02, -1.77552141e-02, -
3.04668006e-02,
3.36785764e-02, -4.32307646e-02, -1.25695765e-02, -
1.90265775e-02,
3.54213044e-02, -4.91841435e-02, 6.62908927e-02, -
5.39336056e-02,
2.95609236e-02, -3.62023478e-03, -5.95523827e-02, -
1.35700451e-02,
-2.25441996e-03, -3.31546515e-02, 1.19538596e-02, -
6.39469773e-02,
-5.58329858e-02, 6.63130060e-02, -1.21900411e-02,
2.00054832e-02,
-6.83138845e-03, -1.32144685e-03, 5.27896136e-02,
4.69964743e-02,
-9.11605582e-02, -2.01851153e-03, 3.04776132e-02,
1.10727465e-02,
-6.24179393e-02, 1.45536411e-04, 2.08536219e-02,
4.45022434e-02,
1.93785895e-02, 5.17413318e-02, -5.32217249e-02, -
4.97927377e-03,
1.19333258e-02, -4.89153154e-02, -6.53658994e-04, -
2.22930871e-02,
2.61724908e-02, -8.65411684e-02, 6.07971288e-03, -
3.71725187e-02,
-3.19785029e-02, 5.01653701e-02, -6.23390824e-02, -
5.11551946e-02,
-1.23618050e-02, 2.33343765e-02, 6.37549981e-02, -

1.76865589e-02,
-7.07696890e-04, 9.15793050e-03, -2.26111989e-03, -
3.00544649e-02,
2.63767801e-02, 2.06706431e-02, -3.13745029e-02, -
3.31248790e-02,
-5.53311333e-02, 3.93633470e-02, 8.07671174e-02, -
1.46791469e-02,
3.80190685e-02, -8.00316688e-03, 1.20121296e-02,
2.29324447e-03,
2.60894606e-03, -6.35098740e-02, -9.31760669e-03, -
2.91128755e-02,
2.50250902e-02, 2.58371718e-02, 2.84616183e-02,
3.76612172e-02,
-5.76858548e-03, 5.52670173e-02, -4.37244214e-02, -
2.60425359e-02,
-3.30909863e-02, -1.38194831e-02, 5.48558049e-02,
1.17847688e-01,
-2.38052215e-02, 1.80361830e-02, -8.05631652e-03,
3.35333347e-02,
4.72173467e-02, -1.24808937e-01, -5.09571433e-02,
1.23257451e-02,
2.35860832e-02, -2.79316716e-02, 1.40124774e-02, -
1.15944427e-02,
-1.33579941e-02, -2.12684628e-02, 2.62219328e-02,
2.11175578e-03,
-2.73995902e-02, -1.12771969e-02, -3.88388075e-02,
1.64118093e-02,
1.58122871e-02, -5.26464824e-03, -3.16673554e-02, -
3.03261019e-02,
-1.61655806e-02, -8.18773173e-04, 1.82968788e-02,
4.43238057e-02,
-2.70270426e-02, 3.66189852e-02, -3.49101648e-02, -
2.27940641e-02,
1.01282924e-01, -3.70639861e-02, -2.35699359e-02,
3.92535105e-02,
9.02504995e-02, -3.94867547e-03, 1.03947064e-02, -
7.97142566e-04,
2.12728437e-02, -5.58913760e-02, 4.89393435e-02, -
1.14682289e-02,
1.67897325e-02, -1.39973350e-02, -3.12713301e-03,
4.50159721e-02,
-2.10610237e-02, 1.08042220e-02, -4.30429168e-03, -
2.51616258e-02,
2.87916493e-02, -2.43200455e-02, 2.15951931e-02,
1.23746488e-02,
-3.73926247e-03, 2.83982120e-02, 1.30490065e-02,
2.30866745e-02,
3.96975651e-02, 5.55678718e-02, -4.88976715e-03,
2.40007602e-02,

8.07422772e-03, -3.10871061e-02, 8.19915812e-03,
3.46832909e-02,
2.66513675e-02, -3.29132155e-02, -9.77127440e-03,
1.20261470e-02,
-2.42768694e-03, 6.02175407e-02, -5.27139604e-02,
3.44281234e-02,
1.64458971e-03, -1.62697174e-02, -4.94427867e-02, -
2.72891857e-02,
-4.20533568e-02, 1.49749294e-02, -6.86076004e-04,
1.25647932e-02,
-6.36436790e-02, -5.86519502e-02, 4.81556775e-03, -
1.37256002e-02,
2.20604427e-02, -1.70530155e-02, 1.42270392e-02,
4.42482308e-02,
-3.83847975e-03, -1.17714936e-02, 9.01826750e-03, -
1.37870507e-02,
-2.16902117e-03, -1.04492705e-04, -1.75094940e-02,
3.19200195e-02,
-2.66522150e-02, -4.91484301e-03, 3.93462554e-02,
1.33268954e-02,
5.14063286e-04, -1.39923170e-02, -2.35259496e-02, -
6.69415994e-03,
-5.50549338e-03, 4.52826405e-03, 8.25183745e-03, -
6.50277212e-02,
7.44744167e-02, -4.65239957e-02, 1.50174052e-02,
3.74723822e-02,
-1.82890519e-02, -3.96928489e-02, 1.31297084e-02,
2.98248567e-02,
-1.39923058e-02, -1.58031173e-02, 1.93730404e-03,
8.18524063e-02,
1.51695330e-02, -1.88267324e-02, 2.44880151e-02,
5.38576534e-03,
-5.44053130e-02, -2.19721925e-02, 2.96329032e-03, -
5.79508841e-02,
-2.53861863e-02, 1.51279345e-02, -2.47185286e-02,
2.24462543e-02,
2.71142572e-02, -9.03379992e-02, 1.09441923e-02, -
2.36964300e-02,
-7.84133077e-02, 1.82655361e-02, -1.44414157e-02,
5.84623814e-02,
2.29904186e-02, -2.49112882e-02, -6.27942458e-02,
6.09459691e-02,
2.24426407e-02, 1.78901330e-02, -2.69332365e-03,
3.29050459e-02,
-9.58283897e-03, -3.17876115e-02, -1.06579717e-02,
1.47291720e-02,
-7.97604769e-03, -2.09242720e-02, 2.01678369e-03, -
3.70099433e-02,
-3.45897302e-02, -6.33957889e-03, -5.91489635e-02, -

4.06527407e-02,
1.97754018e-02, -3.78735401e-02, -2.02451367e-02, -
3.26943323e-02,
9.84184165e-03, -4.14385181e-03, -2.67404914e-02, -
6.25345670e-03,
-6.81333616e-02, -5.93806691e-02, 6.38203397e-02,
4.65516336e-02,
1.66982468e-02, -2.25466155e-02, -3.18962894e-02, -
1.49764027e-02,
7.00782239e-02, 8.38086940e-04, -2.27790941e-02, -
5.76840341e-03,
2.82652816e-03, 3.05928346e-02, -1.94194131e-02,
8.53980333e-02,
-1.34957824e-02, 1.37858074e-02, -1.40744271e-02, -
7.70259090e-03,
-1.22942748e-02, 4.47561331e-02, 4.01991718e-02,
3.88168693e-02,
4.54088859e-03, 9.96658579e-03, -3.05994065e-03, -
1.07540432e-02,
8.73416755e-03, -3.88463512e-02, -6.95265830e-02,
1.55448178e-02,
-2.02620048e-02, 3.11681852e-02, 3.59358750e-02, -
3.52348434e-03,
4.92203832e-02, -9.20320489e-03, 1.47524290e-02, -
2.11158860e-03,
-1.83877461e-02, 2.34885402e-02, 2.49623484e-03,
6.92107975e-02,
2.29195487e-02, 6.17191754e-03, 3.59256044e-02,
2.00021621e-02,
3.45431152e-03, -4.05437686e-03, -2.39361264e-02,
1.26563441e-02,
2.79083988e-03, -6.81866631e-02, -5.40850498e-02,
7.40765855e-02,
-3.12496852e-02, 1.49378888e-02, -1.46798352e-02,
3.82091478e-02,
-3.76919694e-02, 3.05378996e-02, 1.20978700e-02, -
2.36438215e-02,
-3.87949087e-02, 3.24775116e-03, 5.55942096e-02, -
4.07035323e-03,
1.92021690e-02, -8.91212821e-02, 3.34865004e-02, -
4.06895019e-02,
5.35648596e-03, 5.81473969e-02, 1.87399704e-02, -
2.67439596e-02,
1.25566069e-02, -3.87206227e-02, 1.03427172e-01, -
3.66423167e-02,
-1.67853888e-02, 1.13945184e-02, 3.74330059e-02,
6.14224337e-02,
-1.53061124e-02, 1.10572195e-02, -8.11602101e-02,
6.97092265e-02,

-2.66415402e-02, -3.49180810e-02, -6.84484169e-02,
4.39708419e-02,
2.31269449e-02, 3.79677862e-02, -1.39621748e-02,
2.26553306e-02,
-7.01775476e-02, 1.08315356e-01, -9.95143875e-03, -
1.26972767e-02,
3.71880829e-02, 8.98045953e-03, 4.13904749e-02, -
3.70523334e-02,
9.40387510e-03, -9.46397241e-03, 8.13440531e-02, -
1.39316591e-02,
3.46607305e-02, -4.87162955e-02, 5.77040575e-02,
9.25261676e-02,
1.20504051e-02, 1.42043149e-02, -3.20942551e-02, -
2.81202160e-02,
-9.48144682e-03, 1.24683473e-02, 3.3668640e-02, -
4.56789322e-02,
-3.84465382e-02, -6.95343390e-02, -2.00450532e-02,
2.39650533e-02,
-4.76826057e-02, -5.13486899e-02, 4.74149920e-02,
3.14663127e-02,
-5.14105596e-02, -1.10659704e-01, 6.13262653e-02,
3.66745554e-02,
2.33488511e-02, -2.20849732e-04, -2.76141167e-02,
1.97103769e-02,
-1.52351595e-02, -4.99254977e-03, -6.09836355e-03, -
1.37408935e-02,
-7.88052939e-03, -4.22836002e-03, -7.56830955e-03,
2.19115266e-03,
-3.92516293e-02, 1.89812966e-02, 5.00929430e-02,
6.89957365e-02,
-5.33464812e-02, 1.39828557e-02, 9.91954003e-03,
9.20665413e-02,
3.35420598e-03, 7.10324571e-03, -4.07486968e-03, -
2.18395586e-03,
-3.84907238e-02, 5.17637879e-02, 4.85990569e-02,
1.89871117e-02,
-1.22943437e-02, -3.23659484e-03, -1.80648267e-02,
4.58293445e-02,
-2.95582805e-02, 4.30033319e-02, -5.80738373e-02, -
6.31308183e-02,
-5.26383929e-02, 3.76392640e-02, -3.50138568e-03, -
8.13257042e-03,
5.43525326e-04, -3.24286558e-02, 3.12254410e-02,
1.60803013e-02,
2.06444934e-02, 3.57489474e-02, -5.61483623e-03, -
1.89970378e-02,
-2.25964691e-02, 2.23705824e-02, 5.28381132e-02, -
3.62323113e-02,
3.70503999e-02, -2.01510657e-02, 2.22543664e-02,

2.69658323e-02,
-3.50981839e-02, 6.68111742e-02, -3.27409282e-02, -
1.33180907e-02,
3.81752811e-02, -4.92350869e-02, 5.38411774e-02, -
4.30977680e-02,
2.27502249e-02, 3.23026218e-02, -4.95612696e-02, -
1.69880111e-02,
8.02653562e-03, -1.12407751e-01, -9.10021830e-03,
2.03591790e-02,
-4.50036116e-03, 1.23294983e-02, -1.53886769e-02, -
5.92637085e-33,
-4.26628685e-04, -2.70198123e-03, -1.27402665e-02, -
6.65993541e-02,
-5.04501052e-02, -2.72380784e-02, -1.90255530e-02,
7.01858802e-03,
-3.04201730e-02, -2.09668726e-02, 1.02480538e-02,
1.75109580e-02,
1.58473086e-02, -2.99419612e-02, 6.26615509e-02, -
2.19623316e-02,
-1.74747154e-04, -6.15433976e-02, -2.23331794e-04, -
1.91734247e-02,
-1.08943200e-02, -2.02908143e-02, 5.39590754e-02, -
4.17055488e-02,
2.40534842e-02, 1.88253671e-02, 1.75268296e-02, -
3.25025097e-02,
9.24340493e-05, -2.85900403e-02, -2.09944844e-02,
3.09375860e-02,
-5.11265406e-03, -2.89998278e-02, 2.07080413e-02, -
2.21317131e-02,
-2.07879432e-02, -2.04242095e-02, 3.81387793e-03, -
8.13572928e-02,
6.57545310e-03, -3.87763120e-02, 3.18596931e-03,
5.73557876e-02,
5.08330390e-02, -6.35215938e-02, -1.15007125e-02, -
4.91468981e-03,
2.41039693e-02, -1.14604388e-03, 3.51250023e-02, -
2.00512190e-03,
1.01389289e-02, 1.49778947e-02, -6.51710704e-02,
6.08230643e-02,
-1.53507516e-02, 4.95253270e-03, -2.50949953e-02, -
4.10628645e-03,
3.79640311e-02, 1.48475226e-02, -1.31493704e-02, -
1.97987165e-02,
4.10175230e-03, -5.21223003e-04, -1.28425630e-02, -
1.34592392e-02,
-8.36507902e-02, 6.24311797e-04, -1.72132882e-03,
2.22239848e-02,
-7.58421374e-03, 2.32700724e-02, 4.25225217e-03, -
2.09372956e-02,

-3.20803970e-02, -7.90733099e-03, 6.05197288e-02,
2.08906899e-03,
2.59007867e-02, -9.37699247e-03, 9.73023381e-03,
3.99472844e-03,
7.39754410e-04, 4.72042104e-03, 1.92600265e-02,
3.16291749e-02,
2.04469096e-02, -2.19827648e-02, 3.15893777e-02,
6.08453378e-02,
1.32586379e-02, 4.77555878e-02, 2.75528729e-02, -
2.23356020e-02,
5.55874854e-02, 1.25482064e-02, 1.76704004e-02, -
2.14804243e-02,
-3.67174596e-02, 5.09817600e-02, 5.83517104e-02, -
1.80564802e-02,
6.11533262e-02, 1.35567794e-02, -5.83259156e-04, -
8.42666719e-03,
-6.54969886e-02, -1.14265615e-02, 1.31030250e-02,
2.63674161e-03,
4.78280149e-02, -1.06865698e-02, -1.67129003e-02,
3.04272808e-02,
5.15215658e-02, -7.95336217e-02, -2.13776622e-02,
3.80686335e-02,
1.08813271e-02, -3.24522592e-02, -4.13647518e-02, -
1.28714051e-02,
2.50939000e-02, -2.39280276e-02, 1.25780255e-02, -
2.83017121e-02,
-4.97984663e-02, 3.83457029e-03, 2.03671772e-02, -
7.22397445e-03,
2.52107668e-07, -2.55462956e-02, 6.81276899e-03, -
1.34857986e-02,
-5.12958504e-03, 2.80297939e-02, -3.88700552e-02, -
2.08468214e-02,
4.51702019e-03, 8.08649585e-02, 9.86090279e-04,
4.69115786e-02,
-2.20833439e-02, -7.18614459e-03, 6.29182607e-02,
8.60308204e-03,
-9.82591696e-03, -3.91230136e-02, 8.13291036e-03, -
1.25912344e-02,
1.18597853e-03, 3.29681039e-02, 5.70015348e-02, -
5.75757166e-03,
2.52174400e-02, -2.02309433e-02, 2.97911605e-03, -
5.13065606e-02,
-9.84182879e-02, -5.03572971e-02, -2.16773674e-02,
2.31676977e-02,
4.86288816e-02, 3.06415074e-02, 5.12670167e-02, -
1.99438892e-02,
-1.13651734e-02, 1.92016941e-02, 5.05775250e-02,
7.98315462e-03,
2.06170529e-02, -4.23547626e-02, -4.79325838e-03, -

```

2.19931733e-02,
  -5.15670795e-03,  2.98933592e-02, -4.32614721e-02,
9.53966472e-03,
  8.35732818e-02, -9.60401818e-02, -7.06946594e-04,
3.06623857e-02,
  -3.55179906e-02,  1.27695920e-02,  2.40421034e-02,
1.61062889e-02,
  4.10765968e-02, -5.33520523e-03,  3.67885851e-03,
4.82094586e-02,
  1.66472662e-02,  3.54915746e-02, -7.62644559e-02, -
2.57878825e-02,
  5.83969876e-02, -5.43487296e-02,  2.10366007e-02,
5.36081940e-02,
  1.88906201e-34, -2.71353070e-02, -3.80815263e-03,
7.60841696e-03,
  -8.50685537e-02,  3.16005275e-02, -1.61992349e-02,
2.77394541e-02,
  2.14648359e-02, -3.06655895e-02, -2.11073402e-02, -
5.75036788e-03]),
  'score': tensor(0.0070)}]

```

Try the model with different number of examples in `prompt_formatter`(Multiple shot prompt engineering) and comment on the answers retrieved by RAG. You can also try different examples but its optional.

As we try the model with different number of examples in the prompt, the RAG answers gradually adopt the phrasing, level of detail, and explanatory style of my examples. For example if my prompt are more number oriented, the responses will include specific statistics and quantitative details rather than just qualitative descriptions.

2. LoRA Adapter FineTuning (30 Points)

LoRA was introduced at the end of 2021 in the paper [LoRA: Low-rank adaptation of Large Language Models](#).

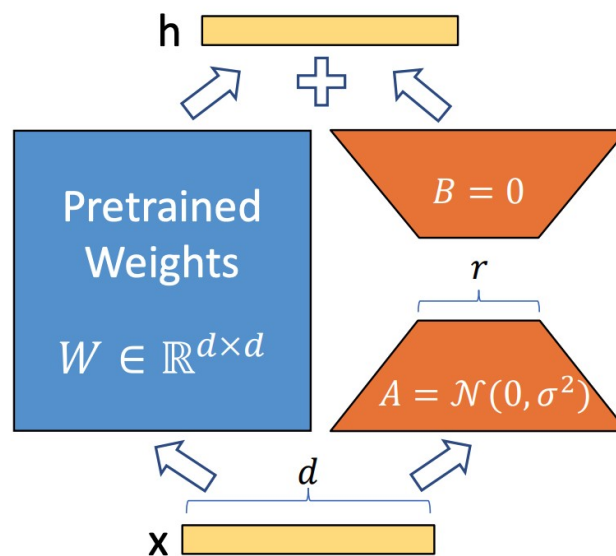
To make fine-tuning more efficient, LoRA's approach is to represent the weight updates with two smaller matrices (called update matrices) through low-rank decomposition. These new matrices can be trained to adapt to the new data while keeping the overall number of changes low. The original weight matrix remains frozen and doesn't receive any further adjustments. To produce the final results, both the original and the adapted weights are combined.

Models that want to support LoRA fine-tuning need to provide adapters. There is a list on [Hugging Face](#) with the supported models which covers popular open source models and it is growing.

You can also read [Efficient Large Language Model training with LoRA and Hugging Face](#)

Advantages of LoRA- There are some additional advantages that LoRA provides. One is to use quantization which leads to memory reduction. Read the blog post on Hugging Face [Making LLMs even more accessible with bitsandbytes, 4-bit quantization and QLoRA](#) for more details.

Some people claim that LoRA might also help to reduce [catastrophic forgetting](#) which happens when too much fine-tuning has been done. However, while the original weights are frozen, at inference time the additional parameters are utilized which lead to other results which is the whole point of fine-tuning.



```
## DONOT CHANGE THE CODE
```

```
!pip install datasets
import os
!pip install peft
if "COLAB_GPU" in os.environ:
    print("[INFO] Running in Google Colab, installing requirements.")
    !pip install -U torch # requires torch 2.1.1+ (for efficient sdpa
implementation)
    !pip install tqdm # for progress bars
    !pip install sentence-transformers # for embedding models
    !pip install accelerate # for quantization model loading
```

Collecting datasets

```
Downloading datasets-3.5.1-py3-none-any.whl.metadata (19 kB)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from datasets) (3.18.0)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from datasets) (2.0.2)
```

Requirement already satisfied: pyarrow>=15.0.0 in
/usr/local/lib/python3.11/dist-packages (from datasets) (18.1.0)
Collecting dill<0.3.9,>=0.3.0 (from datasets)
 Downloading dill-0.3.8-py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: pandas in
/usr/local/lib/python3.11/dist-packages (from datasets) (2.2.2)
Requirement already satisfied: requests>=2.32.2 in
/usr/local/lib/python3.11/dist-packages (from datasets) (2.32.3)
Requirement already satisfied: tqdm>=4.66.3 in
/usr/local/lib/python3.11/dist-packages (from datasets) (4.67.1)
Collecting xxhash (from datasets)
 Downloading xxhash-3.5.0-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (12 kB)
Collecting multiprocessing<0.70.17 (from datasets)
 Downloading multiprocessing-0.70.16-py311-none-any.whl.metadata (7.2
kB)
Collecting fsspec<=2025.3.0,>=2023.1.0 (from
fsspec[http]<=2025.3.0,>=2023.1.0->datasets)
 Downloading fsspec-2025.3.0-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: aiohttp in
/usr/local/lib/python3.11/dist-packages (from datasets) (3.11.15)
Requirement already satisfied: huggingface-hub>=0.24.0 in
/usr/local/lib/python3.11/dist-packages (from datasets) (0.30.2)
Requirement already satisfied: packaging in
/usr/local/lib/python3.11/dist-packages (from datasets) (24.2)
Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from datasets) (6.0.2)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(2.6.1)
Requirement already satisfied: aiosignal>=1.1.2 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(1.3.2)
Requirement already satisfied: attrs>=17.3.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(25.3.0)
Requirement already satisfied: frozenlist>=1.1.1 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(1.6.0)
Requirement already satisfied: multidict<7.0,>=4.5 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(6.4.3)
Requirement already satisfied: propcache>=0.2.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(0.3.1)
Requirement already satisfied: yarl<2.0,>=1.17.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(1.20.0)
Requirement already satisfied: typing-extensions>=3.7.4.3 in

```

/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.24.0-
>datasets) (4.13.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests>=2.32.2-
>datasets) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests>=2.32.2-
>datasets) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests>=2.32.2-
>datasets) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests>=2.32.2-
>datasets) (2025.4.26)
Requirement already satisfied: python-dateutil>=2.8.2 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets)
(2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets)
(2025.2)
Requirement already satisfied: tzdata>=2022.7 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets)
(2025.2)
Requirement already satisfied: six>=1.5 in
/usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2-
>pandas->datasets) (1.17.0)
Downloading datasets-3.5.1-py3-none-any.whl (491 kB)
----- 491.4/491.4 kB 15.1 MB/s eta
0:00:00
----- 116.3/116.3 kB 12.5 MB/s eta
0:00:00
----- 193.6/193.6 kB 19.8 MB/s eta
0:00:00
ultiprocess-0.70.16-py311-none-any.whl (143 kB)
----- 143.5/143.5 kB 15.1 MB/s eta
0:00:00
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (194 kB)
----- 194.8/194.8 kB 19.7 MB/s eta
0:00:00
ultiprocess, datasets
  Attempting uninstall: fsspec
    Found existing installation: fsspec 2025.3.2
    Uninstalling fsspec-2025.3.2:
      Successfully uninstalled fsspec-2025.3.2
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
gcsfs 2025.3.2 requires fsspec==2025.3.2, but you have fsspec 2025.3.0
which is incompatible.

```

Successfully installed datasets-3.5.1 dill-0.3.8 fsspec-2025.3.0
multiprocess-0.70.16 xxhash-3.5.0
Requirement already satisfied: peft in /usr/local/lib/python3.11/dist-packages (0.15.2)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from peft) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from peft) (24.2)
Requirement already satisfied: psutil in /usr/local/lib/python3.11/dist-packages (from peft) (5.9.5)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.11/dist-packages (from peft) (6.0.2)
Requirement already satisfied: torch>=1.13.0 in /usr/local/lib/python3.11/dist-packages (from peft) (2.6.0+cu124)
Requirement already satisfied: transformers in /usr/local/lib/python3.11/dist-packages (from peft) (4.51.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-packages (from peft) (4.67.1)
Requirement already satisfied: accelerate>=0.21.0 in /usr/local/lib/python3.11/dist-packages (from peft) (1.6.0)
Requirement already satisfied: safetensors in /usr/local/lib/python3.11/dist-packages (from peft) (0.5.3)
Requirement already satisfied: huggingface_hub>=0.25.0 in /usr/local/lib/python3.11/dist-packages (from peft) (0.30.2)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from huggingface_hub>=0.25.0->peft) (3.18.0)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub>=0.25.0->peft) (2025.3.0)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from huggingface_hub>=0.25.0->peft) (2.32.3)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface_hub>=0.25.0->peft) (4.13.2)
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft) (3.4.2)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft) (3.1.6)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft) (12.4.127)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft) (12.4.127)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in

/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(12.4.127)
Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(9.1.0.70)
Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(12.4.5.8)
Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(11.2.1.3)
Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(10.3.5.147)
Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(11.6.1.9)
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(0.6.2)
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(2.21.5)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(12.4.127)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(12.4.127)
Requirement already satisfied: triton==3.2.0 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(3.2.0)
Requirement already satisfied: sympy==1.13.1 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.13.0->peft)
(1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy==1.13.1-
>torch>=1.13.0->peft) (1.3.0)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.11/dist-packages (from transformers->peft)
(2024.11.6)
Requirement already satisfied: tokenizers<0.22,>=0.21 in
/usr/local/lib/python3.11/dist-packages (from transformers->peft)
(0.21.1)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.13.0-


```
>peft) (3.0.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub>=0.25.0->peft) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub>=0.25.0->peft) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub>=0.25.0->peft) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests-
>huggingface_hub>=0.25.0->peft) (2025.4.26)
[INFO] Running in Google Colab, installing requirements.
Requirement already satisfied: torch in
/usr/local/lib/python3.11/dist-packages (2.6.0+cu124)
Collecting torch
  Downloading torch-2.7.0-cp311-cp311-
manylinux_2_28_x86_64.whl.metadata (29 kB)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from torch) (3.18.0)
Requirement already satisfied: typing-extensions>=4.10.0 in
/usr/local/lib/python3.11/dist-packages (from torch) (4.13.2)
Collecting sympy>=1.13.3 (from torch)
  Downloading sympy-1.14.0-py3-none-any.whl.metadata (12 kB)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch) (3.1.6)
Requirement already satisfied: fsspec in
/usr/local/lib/python3.11/dist-packages (from torch) (2025.3.0)
Collecting nvidia-cuda-nvrtc-cu12==12.6.77 (from torch)
  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)
  Downloading nvidia_cuda_runtime_cu12-12.6.77-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-cupti-cu12==12.6.80 (from torch)
  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)
  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)
  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)
  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)
  Downloading nvidia_curand_cu12-10.3.7.77-py3-none-
manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cusolver-cu12==11.7.1.2 (from torch)
  Downloading nvidia_cusolver_cu12-11.7.1.2-py3-none-
manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cusparselt-cu12==0.6.3 (from torch)
  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)
  Downloading nvidia_nccl_cu12-2.26.2-py3-none-
manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (2.0 kB)
Collecting nvidia-nvtx-cu12==12.6.77 (from torch)
  Downloading nvidia_nvtx_cu12-12.6.77-py3-none-
manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-nvjitlink-cu12==12.6.85 (from torch)
  Downloading nvidia_nvjitlink_cu12-12.6.85-py3-none-
manylinux2010_x86_64.manylinux_2_12_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cufile-cu12==1.11.1.6 (from torch)
  Downloading nvidia_cufile_cu12-1.11.1.6-py3-none-
manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (1.5 kB)
Collecting triton==3.3.0 (from torch)
  Downloading triton-3.3.0-cp311-cp311-
manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl.metadata (1.5 kB)
Requirement already satisfied: setuptools>=40.8.0 in
/usr/local/lib/python3.11/dist-packages (from triton==3.3.0->torch)
(75.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy>=1.13.3->torch)
(1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch) (3.0.2)
Downloading torch-2.7.0-cp311-cp311-manylinux_2_28_x86_64.whl (865.2
MB)
----- 865.2/865.2 MB 1.1 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (393.1 MB)
----- 393.1/393.1 MB 2.3 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (8.9 MB)
----- 8.9/8.9 MB 117.6 MB/s eta
0:00:00
anylinux2014_x86_64.whl (23.7 MB)
----- 23.7/23.7 MB 95.8 MB/s eta
0:00:00
e_cu12-12.6.77-py3-none-manylinux2014_x86_64.manylinux_2_17_x86_64.whl

```

```

(897 kB)
----- 897.7/897.7 kB 61.5 MB/s eta
0:00:00
anylinux_2_28_x86_64.whl (571.0 MB)
----- 571.0/571.0 MB 3.1 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (200.2 MB)
----- 200.2/200.2 MB 6.0 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (1.1 MB)
----- 1.1/1.1 MB 49.9 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (56.3 MB)
----- 56.3/56.3 MB 14.0 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (158.2 MB)
----- 158.2/158.2 MB 7.5 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (216.6 MB)
----- 216.6/216.6 MB 3.6 MB/s eta
0:00:00
anylinux2014_x86_64.whl (156.8 MB)
----- 156.8/156.8 MB 7.7 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (201.3 MB)
----- 201.3/201.3 MB 6.0 MB/s eta
0:00:00
anylinux2010_x86_64.manylinux_2_12_x86_64.whl (19.7 MB)
----- 19.7/19.7 MB 74.4 MB/s eta
0:00:00
anylinux2014_x86_64.manylinux_2_17_x86_64.whl (89 kB)
----- 89.3/89.3 kB 8.3 MB/s eta
0:00:00
anylinux_2_27_x86_64.manylinux_2_28_x86_64.whl (156.5 MB)
----- 156.5/156.5 MB 7.4 MB/s eta
0:00:00
py-1.14.0-py3-none-any.whl (6.3 MB)
----- 6.3/6.3 MB 95.1 MB/s eta
0:00:00
py, 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, nvidia-
cusparselt-cu12, nvidia-cufft-cu12, nvidia-cudnn-cu12, nvidia-cusolver-
cu12, torch
  Attempting uninstall: nvidia-cusparselt-cu12
    Found existing installation: nvidia-cusparselt-cu12 0.6.2
    Uninstalling nvidia-cusparselt-cu12-0.6.2:
      Successfully uninstalled nvidia-cusparselt-cu12-0.6.2
  Attempting uninstall: triton

```

```
Found existing installation: triton 3.2.0
Uninstalling triton-3.2.0:
  Successfully uninstalled triton-3.2.0
Attempting uninstall: sympy
Found existing installation: sympy 1.13.1
Uninstalling sympy-1.13.1:
  Successfully uninstalled sympy-1.13.1
Attempting uninstall: nvidia-nvtx-cu12
Found existing installation: nvidia-nvtx-cu12 12.4.127
Uninstalling nvidia-nvtx-cu12-12.4.127:
  Successfully uninstalled nvidia-nvtx-cu12-12.4.127
Attempting uninstall: nvidia-nvjitlink-cu12
Found existing installation: nvidia-nvjitlink-cu12 12.4.127
Uninstalling nvidia-nvjitlink-cu12-12.4.127:
  Successfully uninstalled nvidia-nvjitlink-cu12-12.4.127
Attempting uninstall: nvidia-nccl-cu12
Found existing installation: nvidia-nccl-cu12 2.21.5
Uninstalling nvidia-nccl-cu12-2.21.5:
  Successfully uninstalled nvidia-nccl-cu12-2.21.5
Attempting uninstall: nvidia-curand-cu12
Found existing installation: nvidia-curand-cu12 10.3.5.147
Uninstalling nvidia-curand-cu12-10.3.5.147:
  Successfully uninstalled nvidia-curand-cu12-10.3.5.147
Attempting uninstall: nvidia-cuda-runtime-cu12
Found existing installation: nvidia-cuda-runtime-cu12 12.4.127
Uninstalling nvidia-cuda-runtime-cu12-12.4.127:
  Successfully uninstalled nvidia-cuda-runtime-cu12-12.4.127
Attempting uninstall: nvidia-cuda-nvrtc-cu12
Found existing installation: nvidia-cuda-nvrtc-cu12 12.4.127
Uninstalling nvidia-cuda-nvrtc-cu12-12.4.127:
  Successfully uninstalled nvidia-cuda-nvrtc-cu12-12.4.127
Attempting uninstall: nvidia-cuda-cupti-cu12
Found existing installation: nvidia-cuda-cupti-cu12 12.4.127
Uninstalling nvidia-cuda-cupti-cu12-12.4.127:
  Successfully uninstalled nvidia-cuda-cupti-cu12-12.4.127
Attempting uninstall: nvidia-cublas-cu12
Found existing installation: nvidia-cublas-cu12 12.4.5.8
Uninstalling nvidia-cublas-cu12-12.4.5.8:
  Successfully uninstalled nvidia-cublas-cu12-12.4.5.8
Attempting uninstall: nvidia-cusparse-cu12
Found existing installation: nvidia-cusparse-cu12 12.3.1.170
Uninstalling nvidia-cusparse-cu12-12.3.1.170:
  Successfully uninstalled nvidia-cusparse-cu12-12.3.1.170
Attempting uninstall: nvidia-cufft-cu12
Found existing installation: nvidia-cufft-cu12 11.2.1.3
Uninstalling nvidia-cufft-cu12-11.2.1.3:
  Successfully uninstalled nvidia-cufft-cu12-11.2.1.3
Attempting uninstall: nvidia-cudnn-cu12
Found existing installation: nvidia-cudnn-cu12 9.1.0.70
```

```

Uninstalling nvidia-cudnn-cu12-9.1.0.70:
  Successfully uninstalled nvidia-cudnn-cu12-9.1.0.70
Attempting uninstall: nvidia-cusolver-cu12
  Found existing installation: nvidia-cusolver-cu12 11.6.1.9
  Uninstalling nvidia-cusolver-cu12-11.6.1.9:
    Successfully uninstalled nvidia-cusolver-cu12-11.6.1.9
Attempting uninstall: torch
  Found existing installation: torch 2.6.0+cu124
  Uninstalling torch-2.6.0+cu124:
    Successfully uninstalled torch-2.6.0+cu124
ERROR: pip's dependency resolver does not currently take into account
all the packages that are installed. This behaviour is the source of
the following dependency conflicts.
torchaudio 2.6.0+cu124 requires torch==2.6.0, but you have torch 2.7.0
which is incompatible.
torchvision 0.21.0+cu124 requires torch==2.6.0, but you have torch
2.7.0 which is incompatible.
fastai 2.7.19 requires torch<2.7,>=1.10, but you have torch 2.7.0
which is incompatible.
Successfully installed 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-cusparselt-cu12-12.5.4.2 nvidia-cusparse-cu12-
0.6.3 nvidia-nccl-cu12-2.26.2 nvidia-nvjitlink-cu12-12.6.85 nvidia-
nvtx-cu12-12.6.77 sympy-1.14.0 torch-2.7.0 triton-3.3.0

{"id": "f734772e9163476f973364489f08bae5", "pip_warning": {"packages":
["nvidia", "sympy", "torch", "torchgen", "triton"]}}

Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-
packages (4.67.1)
Requirement already satisfied: sentence-transformers in
/usr/local/lib/python3.11/dist-packages (3.4.1)
Requirement already satisfied: transformers<5.0.0,>=4.41.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(4.51.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-
packages (from sentence-transformers) (4.67.1)
Requirement already satisfied: torch>=1.11.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(2.7.0)
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.6.1)
Requirement already satisfied: scipy in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)
(1.15.2)
Requirement already satisfied: huggingface-hub>=0.20.0 in
/usr/local/lib/python3.11/dist-packages (from sentence-transformers)

```

(0.30.2)

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

(11.2.1)

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

Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (2025.3.0)

Requirement already satisfied: packaging>=20.9 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (24.2)

Requirement already satisfied: pyyaml>=5.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (6.0.2)

Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (2.32.3)

Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.20.0->sentence-transformers) (4.13.2)

Requirement already satisfied: sympy>=1.13.3 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (1.14.0)

Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.4.2)

Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.1.6)

Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.6.77)

Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.6.77)

Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.6.80)

Requirement already satisfied: nvidia-cudnn-cu12==9.5.1.17 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (9.5.1.17)

Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (12.6.4.1)

Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (11.3.0.4)

Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (10.3.7.77)

Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (11.7.1.2)

Requirement already satisfied: nvidia-cusparselt-cu12==0.6.3 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (0.6.3)

Requirement already satisfied: nvidia-nccl-cu12==2.26.2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (2.26.2)

Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (12.6.77)

Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (12.6.85)

Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (1.11.1.6)

Requirement already satisfied: triton==3.3.0 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (3.3.0)

Requirement already satisfied: setuptools>=40.8.0 in
/usr/local/lib/python3.11/dist-packages (from triton==3.3.0->
torch>=1.11.0->sentence-transformers) (75.2.0)

Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (2.0.2)

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

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

Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from
transformers<5.0.0,>=4.41.0->sentence-transformers) (0.5.3)

Requirement already satisfied: joblib>=1.2.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers) (1.4.2)

Requirement already satisfied: threadpoolctl>=3.1.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-
transformers) (3.6.0)

Requirement already satisfied: mpmath<1.4,>=1.1.0 in

/usr/local/lib/python3.11/dist-packages (from sympy>=1.13.3->torch>=1.11.0->sentence-transformers) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.11.0->sentence-transformers) (3.0.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.20.0->sentence-transformers) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.20.0->sentence-transformers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.20.0->sentence-transformers) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-hub>=0.20.0->sentence-transformers) (2025.4.26)
Requirement already satisfied: accelerate in
/usr/local/lib/python3.11/dist-packages (1.6.0)
Requirement already satisfied: numpy<3.0.0,>=1.17 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (2.0.2)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (24.2)
Requirement already satisfied: psutil in
/usr/local/lib/python3.11/dist-packages (from accelerate) (5.9.5)
Requirement already satisfied: pyyaml in
/usr/local/lib/python3.11/dist-packages (from accelerate) (6.0.2)
Requirement already satisfied: torch>=2.0.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (2.7.0)
Requirement already satisfied: huggingface-hub>=0.21.0 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (0.30.2)
Requirement already satisfied: safetensors>=0.4.3 in
/usr/local/lib/python3.11/dist-packages (from accelerate) (0.5.3)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (3.18.0)
Requirement already satisfied: fsspec>=2023.5.0 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (2025.3.0)
Requirement already satisfied: requests in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.21.0->accelerate) (4.13.2)
Requirement already satisfied: sympy>=1.13.3 in


```
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (1.14.0)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (3.1.6)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.6.77)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.6.77)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.6.80)
Requirement already satisfied: nvidia-cudnn-cu12==9.5.1.17 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (9.5.1.17)
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.6.4.1)
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (11.3.0.4)
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (10.3.7.77)
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (11.7.1.2)
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.3 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (0.6.3)
Requirement already satisfied: nvidia-nccl-cu12==2.26.2 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (2.26.2)
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.6.77)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (12.6.85)
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
```

```

>accelerate) (1.11.1.6)
Requirement already satisfied: triton==3.3.0 in
/usr/local/lib/python3.11/dist-packages (from torch>=2.0.0-
>accelerate) (3.3.0)
Requirement already satisfied: setuptools>=40.8.0 in
/usr/local/lib/python3.11/dist-packages (from triton==3.3.0-
>torch>=2.0.0->accelerate) (75.2.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy>=1.13.3-
>torch>=2.0.0->accelerate) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch>=2.0.0-
>accelerate) (3.0.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.21.0->accelerate) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.21.0->accelerate) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.21.0->accelerate) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->huggingface-
hub>=0.21.0->accelerate) (2025.4.26)

```

DONOT CHANGE THE CODE

```

from transformers import AutoModelForCausalLM, AutoTokenizer,
default_data_collator, get_linear_schedule_with_warmup
from peft import get_peft_config, get_peft_model, PromptTuningInit,
PromptTuningConfig, TaskType, PeftType
import torch
from datasets import load_dataset
import os
from torch.utils.data import DataLoader
from tqdm import tqdm

device = "cuda"
model_name_or_path = "bigscience/bloomz-560m"
tokenizer_name_or_path = "bigscience/bloomz-560m"
peft_config = PromptTuningConfig(
    task_type=TaskType.CAUSAL_LM,
    prompt_tuning_init=PromptTuningInit.TEXT,
    num_virtual_tokens=20,
    prompt_tuning_init_text="Classify if the tweet is a complaint or
not:",
    tokenizer_name_or_path=model_name_or_path,
)

```

```

dataset_name = "twitter_complaints"
checkpoint_name =
f"{dataset_name}_{model_name_or_path}_{peft_config.peft_type}_{peft_co
nfig.task_type}_v1.pt".replace(
    "/", "_"
)
text_column = "Tweet text"
label_column = "text_label"
max_length = 64
lr = 3e-2
num_epochs = 10
batch_size = 8

## DONOT CHANGE THE CODE

#load data
dataset = load_dataset("ought/raft", dataset_name)
dataset["train"][0]

classes = [k.replace("_", " ") for k in
dataset["train"].features["Label"].names]
dataset = dataset.map(
    lambda x: {"text_label": [classes[label] for label in
x["Label"]]},
    batched=True,
    num_proc=1,
)
dataset["train"][0]
{"Tweet text": "@HMRCcustomers No this is my first job", "ID": 0,
"Label": 2, "text_label": "no complaint"}

## preprocess
tokenizer = AutoTokenizer.from_pretrained(model_name_or_path)
if tokenizer.pad_token_id is None:
    tokenizer.pad_token_id = tokenizer.eos_token_id
target_max_length = max([len(tokenizer(class_label)["input_ids"]) for
class_label in classes])
print(target_max_length)

{"model_id": "b67ee0e8265c43269acbe6af790e754e", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "a7a307c25d5e4cf090b117c652f842ed", "version_major": 2, "vers
ion_minor": 0}

{"model_id": "29b41754cdad41f099bea56b228407b5", "version_major": 2, "vers
ion_minor": 0}

```

```

{"model_id":"d7f086158b384709a44092f3407a00b8","version_major":2,"version_minor":0}

{"model_id":"e55c910547e44934b1b65b4e938e2f50","version_major":2,"version_minor":0}

{"model_id":"00af8bb3692e42f38017e7ee1da44876","version_major":2,"version_minor":0}

{"model_id":"c0ec86cb89a54e80aac43ac510b0caed","version_major":2,"version_minor":0}

{"model_id":"d1fc4fd7d9d646edb58fc142d96bfae9","version_major":2,"version_minor":0}

{"model_id":"b9363c430dbf4e39949a5be5ce69f699","version_major":2,"version_minor":0}

{"model_id":"ef22e7d5067a486bab1bc4c9b54287fc","version_major":2,"version_minor":0}

{"model_id":"652620de4b3847408e89cd48c44f6dab","version_major":2,"version_minor":0}

```

3

DONOT CHANGE THE CODE

```

def preprocess_function(examples):
    batch_size = len(examples[text_column])
    inputs = [f"{text_column} : {x} Label : " for x in
examples[text_column]]
    targets = [str(x) for x in examples[label_column]]
    model_inputs = tokenizer(inputs)
    labels = tokenizer(targets)
    for i in range(batch_size):
        sample_input_ids = model_inputs["input_ids"][i]
        label_input_ids = labels["input_ids"][i] +
[tokenizer.pad_token_id]
        # print(i, sample_input_ids, label_input_ids)
        model_inputs["input_ids"][i] = sample_input_ids +
label_input_ids
        labels["input_ids"][i] = [-100] * len(sample_input_ids) +
label_input_ids
        model_inputs["attention_mask"][i] = [1] *
len(model_inputs["input_ids"][i])
        # print(model_inputs)
        for i in range(batch_size):
            sample_input_ids = model_inputs["input_ids"][i]
            label_input_ids = labels["input_ids"][i]
            model_inputs["input_ids"][i] = [tokenizer.pad_token_id] * (

```

```

        max_length - len(sample_input_ids)
    ) + sample_input_ids
    model_inputs["attention_mask"][i] = [0] * (max_length -
len(sample_input_ids)) + model_inputs[
    "attention_mask"
][i]
    labels["input_ids"][i] = [-100] * (max_length -
len(sample_input_ids)) + label_input_ids
    model_inputs["input_ids"][i] =
torch.tensor(model_inputs["input_ids"][i][:max_length])
    model_inputs["attention_mask"][i] =
torch.tensor(model_inputs["attention_mask"][i][:max_length])
    labels["input_ids"][i] = torch.tensor(labels["input_ids"][i]
[:max_length])
    model_inputs["labels"] = labels["input_ids"]
    return model_inputs

## DONOT CHANGE THE CODE

processed_datasets = dataset.map(
    preprocess_function,
    batched=True,
    num_proc=1,
    remove_columns=dataset["train"].column_names,
    load_from_cache_file=False,
    desc="Running tokenizer on dataset",
)

{"model_id": "75e148c9f44a48c2857d2c041d64d1ad", "version_major": 2, "version_minor": 0}

{"model_id": "d909a17ca0694ebf8ade3eb49e6cc38f", "version_major": 2, "version_minor": 0}

train_dataset = processed_datasets["train"]
eval_dataset = processed_datasets["test"]

train_dataloader = DataLoader(
    train_dataset, shuffle=True, collate_fn=default_data_collator,
    batch_size=batch_size, pin_memory=True
)
eval_dataloader = DataLoader(eval_dataset,
    collate_fn=default_data_collator, batch_size=batch_size,
    pin_memory=True)

model = AutoModelForCausalLM.from_pretrained(model_name_or_path)
model = get_peft_model(model, peft_config)
print(model.print_trainable_parameters())

```

```
{"model_id": "82c1f16d45044e7e814097ea49851aaf", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "57bc4cb5a9044814a1f82e1a1f3dbfc0", "version_major": 2, "version_minor": 0}
```

```
trainable params: 20,480 || all params: 559,235,072 || trainable%: 0.0037  
None
```

Default Config

```
optimizer = torch.optim.AdamW(model.parameters(), lr=lr)  
lr_scheduler = get_linear_schedule_with_warmup(  
    optimizer=optimizer,  
    num_warmup_steps=0,  
    num_training_steps=(len(train_dataloader) * num_epochs),  
)
```

Config 1: Fast training

```
optimizer = torch.optim.AdamW(model.parameters(), lr=5e-5)  
lr_scheduler = get_linear_schedule_with_warmup(  
    optimizer=optimizer,  
    num_warmup_steps=0,  
    num_training_steps=(len(train_dataloader) * 2), # num_epochs = 2  
)
```

Config 2: Balanced training

```
optimizer = torch.optim.AdamW(model.parameters(), lr=3e-5)  
lr_scheduler = get_linear_schedule_with_warmup(  
    optimizer=optimizer,  
    num_warmup_steps=200,  
    num_training_steps=(len(train_dataloader) * 3), # num_epochs = 3  
)
```

Config 3: Conservative training

```
optimizer = torch.optim.AdamW(model.parameters(), lr=1e-5)  
lr_scheduler = get_linear_schedule_with_warmup(  
    optimizer=optimizer,  
    num_warmup_steps=200,  
    num_training_steps=(len(train_dataloader) * 4), # num_epochs = 4  
)
```

```
model = model.to(device)
```

```
for epoch in range(num_epochs):  
    model.train()  
    total_loss = 0  
    for step, batch in enumerate(tqdm(train_dataloader)):  
        batch = {k: v.to(device) for k, v in batch.items()}  
        outputs = model(**batch)  
        loss = outputs.loss
```

```

        total_loss += loss.detach().float()
        loss.backward()
        optimizer.step()
        lr_scheduler.step()
        optimizer.zero_grad()

    model.eval()
    eval_loss = 0
    eval_preds = []
    for step, batch in enumerate(tqdm(eval_dataloader)):
        batch = {k: v.to(device) for k, v in batch.items()}
        with torch.no_grad():
            outputs = model(**batch)
            loss = outputs.loss
            eval_loss += loss.detach().float()
            eval_preds.extend(
                tokenizer.batch_decode(torch.argmax(outputs.logits, -
1).detach().cpu().numpy(), skip_special_tokens=True)
            )

    eval_epoch_loss = eval_loss / len(eval_dataloader)
    eval_ppl = torch.exp(eval_epoch_loss)
    train_epoch_loss = total_loss / len(train_dataloader)
    train_ppl = torch.exp(train_epoch_loss)
    print(f"{epoch=}: {train_ppl=} {train_epoch_loss=} {eval_ppl=}
{eval_epoch_loss=}")

100%|██████████| 7/7 [00:03<00:00, 2.25it/s]
100%|██████████| 425/425 [01:48<00:00, 3.93it/s]

epoch=0: train_ppl=tensor(32.2826, device='cuda:0')
train_epoch_loss=tensor(3.4745, device='cuda:0')
eval_ppl=tensor(19544.2422, device='cuda:0')
eval_epoch_loss=tensor(9.8804, device='cuda:0')

100%|██████████| 7/7 [00:03<00:00, 2.25it/s]
100%|██████████| 425/425 [01:47<00:00, 3.95it/s]

epoch=1: train_ppl=tensor(32.4765, device='cuda:0')
train_epoch_loss=tensor(3.4805, device='cuda:0')
eval_ppl=tensor(19546.9258, device='cuda:0')
eval_epoch_loss=tensor(9.8806, device='cuda:0')

100%|██████████| 7/7 [00:03<00:00, 2.26it/s]
100%|██████████| 425/425 [01:47<00:00, 3.94it/s]

epoch=2: train_ppl=tensor(33.3634, device='cuda:0')
train_epoch_loss=tensor(3.5075, device='cuda:0')
eval_ppl=tensor(19551.5488, device='cuda:0')
eval_epoch_loss=tensor(9.8808, device='cuda:0')

```

```
100%|██████████| 7/7 [00:03<00:00, 2.24it/s]
100%|██████████| 425/425 [01:47<00:00, 3.95it/s]
```

```
epoch=3: train_ppl=tensor(32.7114, device='cuda:0')
train_epoch_loss=tensor(3.4877, device='cuda:0')
eval_ppl=tensor(19557.7578, device='cuda:0')
eval_epoch_loss=tensor(9.8811, device='cuda:0')
```

```
100%|██████████| 7/7 [00:03<00:00, 2.25it/s]
100%|██████████| 425/425 [01:47<00:00, 3.95it/s]
```

```
epoch=4: train_ppl=tensor(33.0918, device='cuda:0')
train_epoch_loss=tensor(3.4993, device='cuda:0')
eval_ppl=tensor(19565.7617, device='cuda:0')
eval_epoch_loss=tensor(9.8815, device='cuda:0')
```

```
100%|██████████| 7/7 [00:03<00:00, 2.26it/s]
100%|██████████| 425/425 [01:47<00:00, 3.95it/s]
```

```
epoch=5: train_ppl=tensor(33.4875, device='cuda:0')
train_epoch_loss=tensor(3.5112, device='cuda:0')
eval_ppl=tensor(19574.5898, device='cuda:0')
eval_epoch_loss=tensor(9.8820, device='cuda:0')
```

```
100%|██████████| 7/7 [00:03<00:00, 2.24it/s]
100%|██████████| 425/425 [01:47<00:00, 3.95it/s]
```

```
epoch=6: train_ppl=tensor(32.8179, device='cuda:0')
train_epoch_loss=tensor(3.4910, device='cuda:0')
eval_ppl=tensor(19586.3164, device='cuda:0')
eval_epoch_loss=tensor(9.8826, device='cuda:0')
```

```
100%|██████████| 7/7 [00:03<00:00, 2.26it/s]
100%|██████████| 425/425 [01:47<00:00, 3.96it/s]
```

```
epoch=7: train_ppl=tensor(32.1101, device='cuda:0')
train_epoch_loss=tensor(3.4692, device='cuda:0')
eval_ppl=tensor(19600.3301, device='cuda:0')
eval_epoch_loss=tensor(9.8833, device='cuda:0')
```

```
100%|██████████| 7/7 [00:03<00:00, 2.25it/s]
100%|██████████| 425/425 [01:47<00:00, 3.95it/s]
```

```
epoch=8: train_ppl=tensor(32.2854, device='cuda:0')
train_epoch_loss=tensor(3.4746, device='cuda:0')
eval_ppl=tensor(19615.7207, device='cuda:0')
eval_epoch_loss=tensor(9.8841, device='cuda:0')
```

```
100%|██████████| 7/7 [00:03<00:00, 2.25it/s]
100%|██████████| 425/425 [01:47<00:00, 3.94it/s]
```



```

epoch=9: train_ppl=tensor(32.1429, device='cuda:0')
train_epoch_loss=tensor(3.4702, device='cuda:0')
eval_ppl=tensor(19635., device='cuda:0')
eval_epoch_loss=tensor(9.8851, device='cuda:0')

# Test the Tuned Model
inputs = tokenizer(
    f'{text_column} : {"@nationalgridus I have no water and the bill
is current and paid. Can you do something about this?"} Label : ',
    return_tensors="pt",
)

model.to(device)

with torch.no_grad():
    inputs = {k: v.to(device) for k, v in inputs.items()}
    outputs = model.generate(
        input_ids=inputs["input_ids"],
        attention_mask=inputs["attention_mask"], max_new_tokens=10,
        eos_token_id=3
    )
    print(tokenizer.batch_decode(outputs.detach().cpu().numpy(),
    skip_special_tokens=True))

['Text : @nationalgridus I have no water and the bill is current and
paid. Can you do something about this? Label : no complaint (no water
and the bill is current']

```

Train with 3 different Hyper-parameter Settings and explain the differences in output generation. Also try the best model with 5 different texts from https://huggingface.co/datasets/ought/raft/viewer?views%5B%5D=ade_corpus_v2_train

Hyper-parameter choices essentially govern the trade-offs between learning speed, stability, and generalization: high learning rates with few epochs yield fast but erratic outputs, while lower rates and more epochs produce steadier yet generic text. Tuning batch size, weight decay, and warm-up steps refines responsiveness, factual accuracy, and creativity.

```

# Test model with 5 different texts
text_column = "Text"

# List of text examples
texts = [
    "No regional side effects were noted.",
    "We describe the case of a 10-year-old girl with two epileptic
    seizures and subcontinuous spike-waves during sleep, who presented
    unusual side-effects related to clobazam (CLB) monotherapy.",
    "After the first oral dose of propranolol, syncope developed
    together with atrioventricular block.",
    "As termination was not an option for the family, the patient was
    extensively counseled and treated with oral ganciclovir.",
    "Pulses have been given for periods up to three years without
    evident toxicity."
]

# Loop over each example
for text in texts:
    prompt = f'{text_column} : "{text}" Label : '
    inputs = tokenizer(prompt, return_tensors="pt")

    model.to(device)
    with torch.no_grad():
        inputs = {k: v.to(device) for k, v in inputs.items()}
        outputs = model.generate(
            input_ids=inputs["input_ids"],
            attention_mask=inputs["attention_mask"],
            max_new_tokens=10,
            eos_token_id=3
        )
        decoded = tokenizer.batch_decode(outputs.detach().cpu().numpy(),
        skip_special_tokens=True)
        print(f"\nInput: {text}\nOutput: {decoded[0]}\n")

```

Input: No regional side effects were noted.

Output: Text : "No regional side effects were noted." Label : no
complaint <string>no complaint <string>no complaint <string>no

Input: We describe the case of a 10-year-old girl with two epileptic
seizures and subcontinuous spike-waves during sleep, who presented
unusual side-effects related to clobazam (CLB) monotherapy.

Output: Text : "We describe the case of a 10-year-old girl with two
epileptic seizures and subcontinuous spike-waves during sleep, who
presented unusual side-effects related to clobazam (CLB) monotherapy."
Label : complainttestadonocomplainttestadonocomplaint

Input: After the first oral dose of propranolol, syncope developed together with atrioventricular block.

Output: Text : "After the first oral dose of propranolol, syncope developed together with atrioventricular block." Label : no complaintmp4 data-parasoid='{ "dsr

Input: As termination was not an option for the family, the patient was extensively counseled and treated with oral ganciclovir.

Output: Text : "As termination was not an option for the family, the patient was extensively counseled and treated with oral ganciclovir."

Label : complaint-"الترشيح"

Input: Pulses have been given for periods up to three years without evident toxicity.

Output: Text : "Pulses have been given for periods up to three years without evident toxicity." Label : no complaintmp4"الجميع في المنزل"