Sudan Crisis 4

July 11, 2025

Human Rights Monitoring and Protection of Population in Sudan Using Twitter Data

```
[1]: #Topic Modeling Analysis
     # Install Required Packages

     %pip install bertopic[flair] sentence-transformers pandas
```

Requirement already satisfied: sentence-transformers in
/home/hadoop/py311env/lib/python3.11/site-packages (5.0.0)
Requirement already satisfied: pandas in
/home/hadoop/py311env/lib/python3.11/site-packages (2.2.3)
Requirement already satisfied: bertopic[flair] in
/home/hadoop/py311env/lib/python3.11/site-packages (0.17.3)
Requirement already satisfied: hdbscan>=0.8.29 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(0.8.40)
Requirement already satisfied: umap-learn>=0.5.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(0.5.9.post2)
Requirement already satisfied: numpy>=1.20.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(1.26.4)
Requirement already satisfied: plotly>=4.7.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(6.2.0)
Requirement already satisfied: scikit-learn>=1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(1.6.1)
Requirement already satisfied: tqdm>=4.41.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(4.67.1)
Requirement already satisfied: llvmlite>0.36.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(0.44.0)
Requirement already satisfied: flair>=0.7 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(0.15.1)
Requirement already satisfied: torch>=1.4.0 in

```
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(2.7.0)
Requirement already satisfied: transformers>=3.5.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from bertopic[flair])
(4.53.1)
Requirement already satisfied: scipy in
/home/hadoop/py311env/lib/python3.11/site-packages (from sentence-transformers)
(1.15.3)
Requirement already satisfied: huggingface-hub>=0.20.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from sentence-transformers)
(0.32.3)
Requirement already satisfied: Pillow in
/home/hadoop/py311env/lib/python3.11/site-packages (from sentence-transformers)
(11.2.1)
Requirement already satisfied: typing_extensions>=4.5.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from sentence-transformers)
(4.13.2)
Requirement already satisfied: filelock in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers>=3.5.1->bertopic[flair]) (3.18.0)
Requirement already satisfied: packaging>=20.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers>=3.5.1->bertopic[flair]) (25.0)
Requirement already satisfied: pyyaml>=5.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers>=3.5.1->bertopic[flair]) (6.0.2)
Requirement already satisfied: regex!=2019.12.17 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers>=3.5.1->bertopic[flair]) (2024.11.6)
Requirement already satisfied: requests in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers>=3.5.1->bertopic[flair]) (2.32.3)
Requirement already satisfied: tokenizers<0.22,>=0.21 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers>=3.5.1->bertopic[flair]) (0.21.1)
Requirement already satisfied: safetensors>=0.4.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers>=3.5.1->bertopic[flair]) (0.5.3)
Requirement already satisfied: fsspec>=2023.5.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from huggingface-
hub>=0.20.0->sentence-transformers) (2025.3.0)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from huggingface-
hub>=0.20.0->sentence-transformers) (1.1.2)
Requirement already satisfied: python-dateutil>=2.8.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from pandas) (2025.2)
```

```
Requirement already satisfied: tzdata>=2022.7 in
/home/hadoop/py311env/lib/python3.11/site-packages (from pandas) (2025.2)
Requirement already satisfied: boto3>=1.20.27 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (1.39.4)
Requirement already satisfied: conllu<5.0.0,>=4.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (4.5.3)
Requirement already satisfied: deprecated>=1.2.13 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (1.2.18)
Requirement already satisfied: ftfy>=6.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (6.3.1)
Requirement already satisfied: gdown>=4.4.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (5.2.0)
Requirement already satisfied: langdetect>=1.0.9 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (1.0.9)
Requirement already satisfied: lxml>=4.8.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (5.4.0)
Requirement already satisfied: matplotlib>=2.2.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (3.10.3)
Requirement already satisfied: more-itertools>=8.13.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (10.7.0)
Requirement already satisfied: mpld3>=0.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (0.5.10)
Requirement already satisfied: pptree>=3.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (3.1)
Requirement already satisfied: pytorch-revgrad>=0.2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (0.2.0)
Requirement already satisfied: segtok>=1.5.11 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (1.5.11)
Requirement already satisfied: sqlitedict>=2.0.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (2.1.0)
Requirement already satisfied: tabulate>=0.8.10 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (0.9.0)
Requirement already satisfied: transformer-smaller-training-vocab>=0.2.3 in
```

/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (0.4.2)
Requirement already satisfied: wikipedia-api>=0.5.7 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (0.8.1)
Requirement already satisfied: bioc<3.0.0,>=2.0.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
flair>=0.7->bertopic[flair]) (2.1)
Requirement already satisfied: jsonlines>=1.2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
bioc<3.0.0,>=2.0.0->flair>=0.7->bertopic[flair]) (4.0.0)
Requirement already satisfied: intervaltree in
/home/hadoop/py311env/lib/python3.11/site-packages (from
bioc<3.0.0,>=2.0.0->flair>=0.7->bertopic[flair]) (3.1.0)
Requirement already satisfied: docopt in
/home/hadoop/py311env/lib/python3.11/site-packages (from
bioc<3.0.0,>=2.0.0->flair>=0.7->bertopic[flair]) (0.6.2)
Requirement already satisfied: sentencepiece!=0.1.92,>=0.1.91 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers[sentencepiece]<5.0.0,>=4.25.0->flair>=0.7->bertopic[flair]) (0.2.0)
Requirement already satisfied: protobuf in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers[sentencepiece]<5.0.0,>=4.25.0->flair>=0.7->bertopic[flair])
(4.25.8)
Requirement already satisfied: botocore<1.40.0,>=1.39.4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
boto3>=1.20.27->flair>=0.7->bertopic[flair]) (1.39.4)
Requirement already satisfied: jmespath<2.0.0,>=0.7.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
boto3>=1.20.27->flair>=0.7->bertopic[flair]) (1.0.1)
Requirement already satisfied: s3transfer<0.14.0,>=0.13.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
boto3>=1.20.27->flair>=0.7->bertopic[flair]) (0.13.0)
Requirement already satisfied: urllib3!=2.2.0,<3,>=1.25.4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
botocore<1.40.0,>=1.39.4->boto3>=1.20.27->flair>=0.7->bertopic[flair]) (2.4.0)
Requirement already satisfied: six>=1.5 in
/home/hadoop/py311env/lib/python3.11/site-packages (from python-
dateutil>=2.8.2->pandas) (1.17.0)
Requirement already satisfied: wrapt<2,>=1.10 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
deprecated>=1.2.13->flair>=0.7->bertopic[flair]) (1.14.1)
Requirement already satisfied: wcwidth in
/home/hadoop/py311env/lib/python3.11/site-packages (from
ftfy>=6.1.0->flair>=0.7->bertopic[flair]) (0.2.13)
Requirement already satisfied: beautifulsoup4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
gdown>=4.4.0->flair>=0.7->bertopic[flair]) (4.13.4)

Requirement already satisfied: joblib>=1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
hdbscan>=0.8.29->bertopic[flair]) (1.5.1)
Requirement already satisfied: attrs>=19.2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
jsonlines>=1.2.0->bioc<3.0.0,>=2.0.0->flair>=0.7->bertopic[flair]) (25.3.0)
Requirement already satisfied: contourpy>=1.0.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
matplotlib>=2.2.3->flair>=0.7->bertopic[flair]) (1.3.2)
Requirement already satisfied: cycler>=0.10 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
matplotlib>=2.2.3->flair>=0.7->bertopic[flair]) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
matplotlib>=2.2.3->flair>=0.7->bertopic[flair]) (4.58.1)
Requirement already satisfied: kiwisolver>=1.3.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
matplotlib>=2.2.3->flair>=0.7->bertopic[flair]) (1.4.8)
Requirement already satisfied: pyparsing>=2.3.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
matplotlib>=2.2.3->flair>=0.7->bertopic[flair]) (3.2.3)
Requirement already satisfied: jinja2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
mpld3>=0.3->flair>=0.7->bertopic[flair]) (3.1.6)
Requirement already satisfied: narwhals>=1.15.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
plotly>=4.7.0->bertopic[flair]) (1.46.0)
Requirement already satisfied: threadpoolctl>=3.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from scikit-
learn>=1.0->bertopic[flair]) (3.6.0)
Requirement already satisfied: sympy>=1.13.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (1.14.0)
Requirement already satisfied: networkx in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (3.5)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (12.6.77)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (12.6.77)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (12.6.80)
Requirement already satisfied: nvidia-cudnn-cu12==9.5.1.17 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (9.5.1.17)

```
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (12.6.4.1)
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (11.3.0.4)
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (10.3.7.77)
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (11.7.1.2)
Requirement already satisfied: nvidia-cusparse-cu12==12.5.4.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (12.5.4.2)
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (0.6.3)
Requirement already satisfied: nvidia-nccl-cu12==2.26.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (2.26.2)
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (12.6.77)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (12.6.85)
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (1.11.1.6)
Requirement already satisfied: triton==3.3.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
torch>=1.4.0->bertopic[flair]) (3.3.0)
Requirement already satisfied: setuptools>=40.8.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
triton==3.3.0->torch>=1.4.0->bertopic[flair]) (65.5.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
sympy>=1.13.3->torch>=1.4.0->bertopic[flair]) (1.3.0)
Requirement already satisfied: accelerate>=0.26.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
transformers>=3.5.1->bertopic[flair]) (1.8.1)
Requirement already satisfied: psutil in
/home/hadoop/py311env/lib/python3.11/site-packages (from
accelerate>=0.26.0->transformers>=3.5.1->bertopic[flair]) (7.0.0)
Requirement already satisfied: numba>=0.51.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from umap-
learn>=0.5.0->bertopic[flair]) (0.61.2)
```

Requirement already satisfied: pynndescent>=0.5 in
/home/hadoop/py311env/lib/python3.11/site-packages (from umap-
learn>=0.5.0->bertopic[flair]) (0.5.13)
Requirement already satisfied: soupsieve>1.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
beautifulsoup4->gdown>=4.4.0->flair>=0.7->bertopic[flair]) (2.7)
Requirement already satisfied: sortedcontainers<3.0,>=2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
intervaltree->bioc<3.0.0,>=2.0.0->flair>=0.7->bertopic[flair]) (2.4.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
jinja2->mpld3>=0.3->flair>=0.7->bertopic[flair]) (3.0.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
requests->transformers>=3.5.1->bertopic[flair]) (3.4.2)
Requirement already satisfied: idna<4,>=2.5 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
requests->transformers>=3.5.1->bertopic[flair]) (3.10)
Requirement already satisfied: certifi>=2017.4.17 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
requests->transformers>=3.5.1->bertopic[flair]) (2025.4.26)
Requirement already satisfied: PySocks!=1.5.7,>=1.5.6 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
requests[socks]->gdown>=4.4.0->flair>=0.7->bertopic[flair]) (1.7.1)
Note: you may need to restart the kernel to use updated packages.

```
[2]: %pip install kafka-python pandas geopandas matplotlib seaborn nltk spacy torch↵
     ↪transformers pyspark findspark scikit-learn shapely geopy tensorflow==2.15.0↵
     ↪tensorflow-hub==0.15.0 tensorflow-text==2.15.0 keras==2.15.0 numpy
```

Requirement already satisfied: kafka-python in
/home/hadoop/py311env/lib/python3.11/site-packages (2.2.10)
Requirement already satisfied: pandas in
/home/hadoop/py311env/lib/python3.11/site-packages (2.2.3)
Requirement already satisfied: geopandas in
/home/hadoop/py311env/lib/python3.11/site-packages (1.0.1)
Requirement already satisfied: matplotlib in
/home/hadoop/py311env/lib/python3.11/site-packages (3.10.3)
Requirement already satisfied: seaborn in
/home/hadoop/py311env/lib/python3.11/site-packages (0.13.2)
Requirement already satisfied: nltk in
/home/hadoop/py311env/lib/python3.11/site-packages (3.9.1)
Requirement already satisfied: spacy in
/home/hadoop/py311env/lib/python3.11/site-packages (3.8.7)
Requirement already satisfied: torch in
/home/hadoop/py311env/lib/python3.11/site-packages (2.7.0)
Requirement already satisfied: transformers in
/home/hadoop/py311env/lib/python3.11/site-packages (4.53.1)

```
Requirement already satisfied: pyspark in
/home/hadoop/py311env/lib/python3.11/site-packages (4.0.0)
Requirement already satisfied: findspark in
/home/hadoop/py311env/lib/python3.11/site-packages (2.0.1)
Requirement already satisfied: scikit-learn in
/home/hadoop/py311env/lib/python3.11/site-packages (1.6.1)
Requirement already satisfied: shapely in
/home/hadoop/py311env/lib/python3.11/site-packages (2.1.1)
Requirement already satisfied: geopy in
/home/hadoop/py311env/lib/python3.11/site-packages (2.4.1)
Requirement already satisfied: tensorflow==2.15.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (2.15.0)
Requirement already satisfied: tensorflow-hub==0.15.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (0.15.0)
Requirement already satisfied: tensorflow-text==2.15.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (2.15.0)
Requirement already satisfied: keras==2.15.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (2.15.0)
Requirement already satisfied: numpy in
/home/hadoop/py311env/lib/python3.11/site-packages (1.26.4)
Requirement already satisfied: absl-py>=1.0.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(2.3.0)
Requirement already satisfied: astunparse>=1.6.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(1.6.3)
Requirement already satisfied: flatbuffers>=23.5.26 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(25.2.10)
Requirement already satisfied: gast!=0.5.0,!=0.5.1,!=0.5.2,>=0.2.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(0.6.0)
Requirement already satisfied: google-pasta>=0.1.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(0.2.0)
Requirement already satisfied: h5py>=2.9.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(3.13.0)
Requirement already satisfied: libclang>=13.0.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(18.1.1)
Requirement already satisfied: ml-dtypes~=0.2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(0.2.0)
Requirement already satisfied: opt-einsum>=2.3.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(3.4.0)
Requirement already satisfied: packaging in
```

/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(25.0)
Requirement already satisfied:
protobuf!=4.21.0,!=4.21.1,!=4.21.2,!=4.21.3,!=4.21.4,!=4.21.5,<5.0.0dev,>=3.20.3
in /home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(4.25.8)
Requirement already satisfied: setuptools in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(65.5.0)
Requirement already satisfied: six>=1.12.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(1.17.0)
Requirement already satisfied: termcolor>=1.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(3.1.0)
Requirement already satisfied: typing-extensions>=3.6.6 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(4.13.2)
Requirement already satisfied: wrapt<1.15,>=1.11.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(1.14.1)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(0.37.1)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(1.71.0)
Requirement already satisfied: tensorboard<2.16,>=2.15 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(2.15.2)
Requirement already satisfied: tensorflow-estimator<2.16,>=2.15.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from tensorflow==2.15.0)
(2.15.0)
Requirement already satisfied: google-auth<3,>=1.6.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
tensorboard<2.16,>=2.15->tensorflow==2.15.0) (2.40.3)
Requirement already satisfied: google-auth-oauthlib<2,>=0.5 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
tensorboard<2.16,>=2.15->tensorflow==2.15.0) (1.2.2)
Requirement already satisfied: markdown>=2.6.8 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
tensorboard<2.16,>=2.15->tensorflow==2.15.0) (3.8)
Requirement already satisfied: requests<3,>=2.21.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
tensorboard<2.16,>=2.15->tensorflow==2.15.0) (2.32.3)
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
tensorboard<2.16,>=2.15->tensorflow==2.15.0) (0.7.2)

Requirement already satisfied: werkzeug>=1.0.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
tensorboard<2.16,>=2.15->tensorflow==2.15.0) (3.1.3)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from google-
auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (5.5.2)
Requirement already satisfied: pyasn1-modules>=0.2.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from google-
auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (0.4.2)
Requirement already satisfied: rsa<5,>=3.1.4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from google-
auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (4.9.1)
Requirement already satisfied: requests-oauthlib>=0.7.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from google-auth-
oauthlib<2,>=0.5->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (2.0.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (3.4.2)
Requirement already satisfied: idna<4,>=2.5 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (2025.4.26)
Requirement already satisfied: pyasn1>=0.1.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from rsa<5,>=3.1.4->google-
auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (0.6.1)
Requirement already satisfied: python-dateutil>=2.8.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in
/home/hadoop/py311env/lib/python3.11/site-packages (from pandas) (2025.2)
Requirement already satisfied: pyogrio>=0.7.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from geopandas) (0.11.0)
Requirement already satisfied: pyproj>=3.3.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from geopandas) (3.7.1)
Requirement already satisfied: contourpy>=1.0.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from matplotlib) (1.3.2)
Requirement already satisfied: cycler>=0.10 in
/home/hadoop/py311env/lib/python3.11/site-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from matplotlib) (4.58.1)
Requirement already satisfied: kiwisolver>=1.3.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from matplotlib) (1.4.8)

Requirement already satisfied: pillow>=8 in
/home/hadoop/py311env/lib/python3.11/site-packages (from matplotlib) (11.2.1)
Requirement already satisfied: pyparsing>=2.3.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from matplotlib) (3.2.3)
Requirement already satisfied: click in
/home/hadoop/py311env/lib/python3.11/site-packages (from nltk) (8.2.1)
Requirement already satisfied: joblib in
/home/hadoop/py311env/lib/python3.11/site-packages (from nltk) (1.5.1)
Requirement already satisfied: regex>=2021.8.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from nltk) (2024.11.6)
Requirement already satisfied: tqdm in
/home/hadoop/py311env/lib/python3.11/site-packages (from nltk) (4.67.1)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (1.0.13)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (2.0.11)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (3.0.10)
Requirement already satisfied: thinc<8.4.0,>=8.3.4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (8.3.4)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (2.5.1)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (0.4.1)
Requirement already satisfied: typer<1.0.0,>=0.3.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (0.16.0)
Requirement already satisfied: pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (2.11.5)
Requirement already satisfied: jinja2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (3.1.6)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from spacy) (3.5.0)
Requirement already satisfied: language-data>=1.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
langcodes<4.0.0,>=3.2.0->spacy) (1.3.0)
Requirement already satisfied: annotated-types>=0.6.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy) (0.7.0)
Requirement already satisfied: pydantic-core==2.33.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from

pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy) (2.33.2)
Requirement already satisfied: typing-inspection>=0.4.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy) (0.4.1)
Requirement already satisfied: blis<1.3.0,>=1.2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
thinc<8.4.0,>=8.3.4->spacy) (1.2.1)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
thinc<8.4.0,>=8.3.4->spacy) (0.1.5)
Requirement already satisfied: shellingham>=1.3.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
typer<1.0.0,>=0.3.0->spacy) (1.5.4)
Requirement already satisfied: rich>=10.11.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
typer<1.0.0,>=0.3.0->spacy) (14.0.0)
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
weasel<0.5.0,>=0.1.0->spacy) (0.21.1)
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
weasel<0.5.0,>=0.1.0->spacy) (7.1.0)
Requirement already satisfied: filelock in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (3.18.0)
Requirement already satisfied: sympy>=1.13.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (1.14.0)
Requirement already satisfied: networkx in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (3.5)
Requirement already satisfied: fsspec in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (2025.3.0)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (12.6.77)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (12.6.77)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (12.6.80)
Requirement already satisfied: nvidia-cudnn-cu12==9.5.1.17 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (9.5.1.17)
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (12.6.4.1)
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (11.3.0.4)
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (10.3.7.77)
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (11.7.1.2)
Requirement already satisfied: nvidia-cusparse-cu12==12.5.4.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (12.5.4.2)

```
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (0.6.3)
Requirement already satisfied: nvidia-nccl-cu12==2.26.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (2.26.2)
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (12.6.77)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (12.6.85)
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (1.11.1.6)
Requirement already satisfied: triton==3.3.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch) (3.3.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.30.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from transformers) (0.32.3)
Requirement already satisfied: pyyaml>=5.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from transformers) (6.0.2)
Requirement already satisfied: tokenizers<0.22,>=0.21 in
/home/hadoop/py311env/lib/python3.11/site-packages (from transformers) (0.21.1)
Requirement already satisfied: safetensors>=0.4.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from transformers) (0.5.3)
Requirement already satisfied: hf-xet<2.0.0,>=1.1.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from huggingface-
hub<1.0,>=0.30.0->transformers) (1.1.2)
Requirement already satisfied: py4j==0.10.9.9 in
/home/hadoop/py311env/lib/python3.11/site-packages (from pyspark) (0.10.9.9)
Requirement already satisfied: scipy>=1.6.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from scikit-learn) (1.15.3)
Requirement already satisfied: threadpoolctl>=3.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from scikit-learn) (3.6.0)
Requirement already satisfied: geographiclib<3,>=1.52 in
/home/hadoop/py311env/lib/python3.11/site-packages (from geopy) (2.0)
Requirement already satisfied: wheel<1.0,>=0.23.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
astunparse>=1.6.0->tensorflow==2.15.0) (0.45.1)
Requirement already satisfied: marisa-trie>=1.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from language-
data>=1.2->langcodes<4.0.0,>=3.2.0->spacy) (1.2.1)
Requirement already satisfied: oauthlib>=3.0.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from requests-
oauthlib>=0.7.0->google-auth-
oauthlib<2,>=0.5->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (3.3.1)
Requirement already satisfied: markdown-it-py>=2.2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (3.0.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (2.19.1)
Requirement already satisfied: mdurl~=0.1 in
```

/home/hadoop/py311env/lib/python3.11/site-packages (from markdown-it-py>=2.2.0->rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (0.1.2)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from sympy>=1.13.3->torch)
(1.3.0)
Requirement already satisfied: MarkupSafe>=2.1.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
werkzeug>=1.0.1->tensorboard<2.16,>=2.15->tensorflow==2.15.0) (3.0.2)
Note: you may need to restart the kernel to use updated packages.

[3]: `%pip install seqeval torchcrf`

Requirement already satisfied: seqeval in
/home/hadoop/py311env/lib/python3.11/site-packages (1.2.2)
Requirement already satisfied: torchcrf in
/home/hadoop/py311env/lib/python3.11/site-packages (1.1.0)
Requirement already satisfied: numpy>=1.14.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from seqeval) (1.26.4)
Requirement already satisfied: scikit-learn>=0.21.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from seqeval) (1.6.1)
Requirement already satisfied: torch>=1.0.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torchcrf) (2.7.0)
Requirement already satisfied: scipy>=1.6.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from scikit-
learn>=0.21.3->seqeval) (1.15.3)
Requirement already satisfied: joblib>=1.2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from scikit-
learn>=0.21.3->seqeval) (1.5.1)
Requirement already satisfied: threadpoolctl>=3.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from scikit-
learn>=0.21.3->seqeval) (3.6.0)
Requirement already satisfied: filelock in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(3.18.0)
Requirement already satisfied: typing-extensions>=4.10.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(4.13.2)
Requirement already satisfied: sympy>=1.13.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(1.14.0)
Requirement already satisfied: networkx in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(3.5)
Requirement already satisfied: jinja2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(3.1.6)
Requirement already satisfied: fsspec in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)

(2025.3.0)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(12.6.77)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(12.6.77)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.6.80 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(12.6.80)
Requirement already satisfied: nvidia-cudnn-cu12==9.5.1.17 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(9.5.1.17)
Requirement already satisfied: nvidia-cublas-cu12==12.6.4.1 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(12.6.4.1)
Requirement already satisfied: nvidia-cufft-cu12==11.3.0.4 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(11.3.0.4)
Requirement already satisfied: nvidia-curand-cu12==10.3.7.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(10.3.7.77)
Requirement already satisfied: nvidia-cusolver-cu12==11.7.1.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(11.7.1.2)
Requirement already satisfied: nvidia-cusparse-cu12==12.5.4.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(12.5.4.2)
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.3 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(0.6.3)
Requirement already satisfied: nvidia-nccl-cu12==2.26.2 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(2.26.2)
Requirement already satisfied: nvidia-nvtx-cu12==12.6.77 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(12.6.77)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.6.85 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(12.6.85)
Requirement already satisfied: nvidia-cufile-cu12==1.11.1.6 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(1.11.1.6)
Requirement already satisfied: triton==3.3.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from torch>=1.0.0->torchcrf)
(3.3.0)
Requirement already satisfied: setuptools>=40.8.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from

```
triton==3.3.0->torch>=1.0.0->torchcrf) (65.5.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
sympy>=1.13.3->torch>=1.0.0->torchcrf) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/home/hadoop/py311env/lib/python3.11/site-packages (from
jinja2->torch>=1.0.0->torchcrf) (3.0.2)
Note: you may need to restart the kernel to use updated packages.
```

[4]: `!jupyter nbextension enable --py widgetsnbextension`

```
usage: jupyter [-h] [--version] [--config-dir] [--data-dir] [--runtime-dir]
               [--paths] [--json] [--debug]
               [subcommand]

Jupyter: Interactive Computing

positional arguments:
  subcommand     the subcommand to launch

options:
  -h, --help     show this help message and exit
  --version      show the versions of core jupyter packages and exit
  --config-dir   show Jupyter config dir
  --data-dir     show Jupyter data dir
  --runtime-dir  show Jupyter runtime dir
  --paths        show all Jupyter paths. Add --json for machine-readable
                 format.
  --json         output paths as machine-readable json
  --debug        output debug information about paths

Available subcommands: console dejavu events execute kernel kernelspec lab
labextension labhub migrate nbconvert notebook run server troubleshoot trust

Jupyter command `jupyter-nbextension` not found.
```

[5]:
```python
import os
os.environ["JAVA_HOME"] = "/usr/lib/jvm/java-11-openjdk-amd64"
os.environ["SPARK_HOME"] = "/opt/spark/spark-3.5.6-bin-hadoop3"
os.environ["HADOOP_HOME"] = "/usr/local/hadoop"
os.environ["TF_USE_LEGACY_KERAS"] = "1"
```

[6]:
```python
import findspark
findspark.init()
import pyspark
```

```python
[7]: #Import necessary libraries
     import pandas as pd    # Pandas is an open source, BSD-licensed library
      ↪providing high-performance, easy-to-use data structures and data analysis
      ↪tools for the Python programming language. Pandas Series is a
      ↪one-dimensional labeled array capable of holding any data type with axis
      ↪labels or index
     import geopandas as gpd
     import time
     import matplotlib.pyplot as plt
     import numpy as np
     import logging
     import hashlib
     from geopy.geocoders import Nominatim
     from geopy.extra.rate_limiter import RateLimiter
     from pyspark.sql import functions as func
     from pyspark.sql.types import IntegerType, StringType, FloatType,
      ↪TimestampType, BinaryType
     from pyspark.sql.functions import col, count, when, mean, stddev, abs, sum,
      ↪row_number
     from pyspark.ml.feature import VectorAssembler, StandardScaler, StringIndexer,
      ↪OneHotEncoder
     from pyspark.ml import Pipeline
     from pyspark.ml.classification import RandomForestClassifier
     from pyspark.ml.tuning import CrossValidator, ParamGridBuilder
     from pyspark.ml.evaluation import MulticlassClassificationEvaluator,
      ↪BinaryClassificationEvaluator
     from sklearn.metrics import precision_recall_curve
     from pyspark.sql import Window
     from transformers import pipeline, AutoTokenizer,
      ↪AutoModelForSequenceClassification
     from shapely.geometry import Point
     from tqdm import tqdm
     import torch
     from bertopic import BERTopic
```

2025-07-11 17:45:37.059981: E
external/local_xla/xla/stream_executor/cuda/cuda_dnn.cc:9261] Unable to register
cuDNN factory: Attempting to register factory for plugin cuDNN when one has
already been registered
2025-07-11 17:45:37.060044: E
external/local_xla/xla/stream_executor/cuda/cuda_fft.cc:607] Unable to register
cuFFT factory: Attempting to register factory for plugin cuFFT when one has
already been registered
2025-07-11 17:45:37.061035: E
external/local_xla/xla/stream_executor/cuda/cuda_blas.cc:1515] Unable to
register cuBLAS factory: Attempting to register factory for plugin cuBLAS when
one has already been registered

```
2025-07-11 17:45:37.068052: I tensorflow/core/platform/cpu_feature_guard.cc:182]
This TensorFlow binary is optimized to use available CPU instructions in
performance-critical operations.
To enable the following instructions: AVX2 FMA, in other operations, rebuild
TensorFlow with the appropriate compiler flags.
2025-07-11 17:45:38.152305: W
tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not
find TensorRT
```

[8]:
```python
# Start SparkSession
from pyspark.sql import SparkSession
spark=SparkSession.builder.appName("Sudan Crisis Analysis").config("spark.
 ↪driver.memory", "4g").getOrCreate()
spark
```

```
25/07/11 17:45:53 WARN Utils: Your hostname, julius-virtual-machine resolves to
a loopback address: 127.0.1.1; using 192.168.145.130 instead (on interface
ens33)
25/07/11 17:45:53 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another
address
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use
setLogLevel(newLevel).
25/07/11 17:45:54 WARN NativeCodeLoader: Unable to load native-hadoop library
for your platform… using builtin-java classes where applicable
```

[8]: <pyspark.sql.session.SparkSession at 0x764508f12f90>

[9]:
```python
# Load CSV data from Event Detection Step
df = pd.read_csv("event_detection_results.csv")
```

[10]:
```python
# Identify Text Column
print("Columns in DataFrame:", df.columns)
text_col = None
for col in ['text', 'Text', 'tweet', 'Tweet', 'content']:
    if col in df.columns:
        text_col = col
        break
if text_col is None:
    raise ValueError("No text column found! Please check your CSV and set the
 ↪correct column name.")
```

```
Columns in DataFrame: Index(['_id', 'userId', 'userName', 'alias',
'profilePicture', 'userCreatedAt',
       'bio', 'followers', 'following', 'lists', 'statusesCount', 'verified',
       'userLocation', 'lang', 'type', 'text', 'createdAt', 'mentions',
       'totalRetweets', 'favorites', 'replies', 'links', 'firstLink', 'images',
       'impressions', 'isImage', 'firstImage', 'location', 'url',
```

```
           'noteTweetText', 'inReplyTo', 'inReplyToId', 'inReplyToUserId',
           'reportName', 'Latitude', 'Longitude', 'displacement', 'armed conflict',
           'protest suppression', 'sexual violence', 'famine',
           'infrastructure destruction', 'looting', 'Predicted_Events'],
          dtype='object')
```

[11]:
```python
# Prepare Documents for Topic Modeling
docs = df[text_col].astype(str).tolist()
```

[12]:
```python
# Initialize BERTopic with Multilingual Embeddings
topic_model = BERTopic(
    embedding_model="paraphrase-multilingual-MiniLM-L12-v2",
    language="multilingual"
)
```

[13]:
```python
# Fit the Model and Assign Topics
topics, probabilities = topic_model.fit_transform(docs)
```

```
modules.json:    0%|            | 0.00/229 [00:00<?, ?B/s]

config_sentence_transformers.json:    0%|            | 0.00/122 [00:00<?, ?B/s]

README.md: 0.00B [00:00, ?B/s]

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

config.json:    0%|            | 0.00/645 [00:00<?, ?B/s]

model.safetensors:    0%|            | 0.00/471M [00:00<?, ?B/s]

tokenizer_config.json:    0%|            | 0.00/480 [00:00<?, ?B/s]

tokenizer.json:    0%|            | 0.00/9.08M [00:00<?, ?B/s]

special_tokens_map.json:    0%|            | 0.00/239 [00:00<?, ?B/s]

config.json:    0%|            | 0.00/190 [00:00<?, ?B/s]
```

[15]:
```python
# Add Topic Assignments to DataFrame
df['Topic'] = topics
df['Topic_Probability'] = probabilities  # No .max(axis=1) needed
```

[16]:
```python
# View Topic Info and Top Words
topic_info = topic_model.get_topic_info()
print(topic_info.head())
print("\nTop words for topic 0:\n", topic_model.get_topic(0))
print("\nSample of assigned topics:\n", df[[text_col, 'Topic',
   'Topic_Probability']].head())
```

```
    Topic  Count                                 Name  \
0     -1    814                 -1_islamic_war_sudan_the
1      0    559        0_vikingwarior20_icza1jq9bk_words_no
```

```
2       1       315       1_banan_4akb1fdec6_academic_k1yoroctus
3       2       263   2_waxkagvnhn_invaded_enduring_coordinated
4       3       232             3_tkcabhfyz8_purchase_allowing_566

                                    Representation  \
0  [islamic, war, sudan, the, it, cuts, with, sim…
1  [vikingwarior20, icza1jq9bk, words, no, sudan,…
2  [banan, 4akb1fdec6, academic, k1yoroctus, jour…
3  [waxkagvnhn, invaded, enduring, coordinated, p…
4  [tkcabhfyz8, purchase, allowing, 566, slowed, …

                                  Representative_Docs
0  [RT @whosudan: A #cholera vaccination campaign…
1  [RT @vikingwarior20: No words#Sudan https://t…
2  [RT @BSonblast: Banan hasn't received any dona…
3  [RT @saroyahx: Sudan is being invaded again. N…
4  [RT @blkdesertfairy: still $566 away from hitt…

Top words for topic 0:
 [('vikingwarior20', 0.14354800770847975), ('icza1jq9bk', 0.14332159484063942),
('words', 0.13427903204097877), ('no', 0.07988264786955356), ('sudan',
0.013950284405678831), ('rt', 0.010734224402136637), ('https',
0.01014624694917553), ('co', 0.009416944449052474), ('talkaboutsudan',
0.005171473874933103), ('shadowbanned', 0.003909250467570901)]

Sample of assigned topics:
                                       text  Topic  Topic_Probability
0  RT @BSonblast: Please help this young lady rea…      8           0.569408
1  RT @BSonblast: Please help this young lady rea…      8           0.836380
2  RT @BSonblast: Please help this young lady rea…     67           1.000000
3  RT @blkdesertfairy: still $566 away from hitti…      3           1.000000
4  RT @BSonblast: Please help this young lady rea…      8           0.583514
```

```python
from sentence_transformers import SentenceTransformer
from umap import UMAP
```

```python
# Prepare embeddings
sentence_model = SentenceTransformer("paraphrase-multilingual-MiniLM-L12-v2")
embeddings = sentence_model.encode(docs, show_progress_bar=True)
```

```
Batches:   0%|          | 0/489 [00:00<?, ?it/s]
```

```python
# Visualize Topics
topic_model.visualize_documents(docs, embeddings=embeddings)
```

```python
# Reduce Number of Topics
topic_model.reduce_topics(docs, nr_topics=20)
```

```
print("\nReduced topic info:\n", topic_model.get_topic_info())
```

```
Reduced topic info:
     Topic  Count                                        Name  \
0      -1    814                            -1_the_sudan_in_of
1       0   9498                           0_sudan_the_co_https
2       1   2368                        1_rsf_reported_quick_in
3       2   1117                         2_her_help_please_this
4       3    312              3_thomaserenabh_115_400_paypal
5       4    305             4_eid_send_donationsplease_please
6       5    260             5_xwzbpyidcv_prioritize_this_matter
7       6    212                6_my_surgery_02fzgof0du_goalany
8       7    123                7_water_470_6thone_aztxhamvta
9       8    113             8_alive_kitchen_community_families
10      9    104                9_men_disappearances_nyala_dead
11     10    101                       10_russia_iran_israel_war
12     11     84                          11_only_12600_miss_6202
13     12     59  12_football_adambluebn_tournament_sudanafwerki
14     13     57                       13_mercenary_once_always_
15     14     39              14_teachers_school_students_jezira
16     15     19                15_vetter_eu_wolfram_diplomat
17     16     18                   16_nahud_en_entering_hospital
18     17     15                 17_health_eldaman_appalled_six
19     18     14               18_thief_night_madagascar_senegal

                                      Representation  \
0     [the, sudan, in, of, https, co, to, and, rt, war]
1      [sudan, the, co, https, to, rt, in, and, of, is]
2     [rsf, reported, quick, in, update, bsonblast, …
3     [her, help, please, this, goal, she, donations…
4     [thomaserenabh, 115, 400, paypal, goal, on, 38…
5     [eid, send, donationsplease, please, einass50,…
6     [xwzbpyidcv, prioritize, this, matter, keepeye…
7     [my, surgery, 02fzgof0du, goalany, differenceh…
8     [water, 470, 6thone, aztxhamvta, yqjcpppr8r, t…
9     [alive, kitchen, community, families, kjvaze4r…
10    [men, disappearances, nyala, dead, abductions,…
11    [russia, iran, israel, war, myanmar, iraq, ukr…
12    [only, 12600, miss, 6202, 6398, 8xf2yw3ovc, qy…
13    [football, adambluebn, tournament, sudanafwerk…
14    [mercenary, once, always,    , 1fpl77r4xp, pp…
15    [teachers, school, students, jezira, medani, g…
16    [vetter, eu, wolfram, diplomat, german, m00vd9…
17    [nahud, en, entering, hospital, medical, vktlk…
18    [health, eldaman, appalled, six, infrastructur…
19    [thief, night, madagascar, senegal, ivorycoast…
```

```
                        Representative_Docs
0    [RT @BSonblast: Today's quick update [Jun 16]:…
1    [RT @saroyahx: Sudan is being invaded again. N…
2    [RT @BSonblast: Today's quick update [Jun 13]:…
3    [RT @BSonblast: Banan hasn't received any dona…
4    [RT @ThomaSerenaBH: Still stuck at $115 out of…
5    [RT @einass50: Third Eid Day if you haven't do…
6    [RT @saroyahx: Please prioritize this!! This i…
7    [RT @Banan_Mohamed99: it's been two weeks sinc…
8    [RT @notasjad: We want to get 2 water tanks be…
9    [RT @BSonblast: Literally hundreds of families…
10   [RT @qoga12: #KeepEyesOnSudan Nyala has seen a…
11   [RT @Charles_Lister: 5 months ago #Trump said …
12   [RT @einass50: $6202/$12600 Only 30 days to ra…
13   [RT @AdamblueBN: #Eritrea South #SudanAfwerki'…
14   [RT @Nora_isle: Once a mercenary always a merc…
15   [RT @BSonblast: - Per UNOCHA statement 5 aid w…
16   [RT @PatrickHeinisc1: #German diplomat Wolfram…
17   [RT @missinchident: not only RSF turned a numb…
18   [RT @DrTedros: We are appalled by another dead…
19   [A Thief In The Night (1972) #Apocalypse A Dis…
```

```python
# Save Model and Results
topic_model.save("sudan_bertopic_model")
df.to_csv("event_topic_results.csv", index=False)
print("Saved topic modeling results to event_topic_results.csv")
```

2025-07-11 18:15:37,728 - BERTopic - WARNING: When you use `pickle` to save/load a BERTopic model,please make sure that the environments in which you saveand load the model are **exactly** the same. The version of BERTopic,its dependencies, and python need to remain the same.

Saved topic modeling results to event_topic_results.csv

[ ]: