

recruit ()

June 7, 2024

DP

```
[2]: !pip install ja-ginza
```

```
Requirement already satisfied: ja-ginza in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (5.2.0)
Requirement already satisfied: spacy<4.0.0,>=3.4.4 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from ja-ginza)
(3.7.5)
Requirement already satisfied: sudachipy<0.7.0,>=0.6.2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from ja-ginza)
(0.6.8)
Requirement already satisfied: sudachidict-core>=20210802 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from ja-ginza)
(20240409)
Requirement already satisfied: ginza<5.3.0,>=5.2.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from ja-ginza)
(5.2.0)
Requirement already satisfied: plac>=1.3.3 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
ginza<5.3.0,>=5.2.0->ja-ginza) (1.4.3)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
spacy<4.0.0,>=3.4.4->ja-ginza) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
spacy<4.0.0,>=3.4.4->ja-ginza) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
spacy<4.0.0,>=3.4.4->ja-ginza) (1.0.10)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
spacy<4.0.0,>=3.4.4->ja-ginza) (2.0.8)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
spacy<4.0.0,>=3.4.4->ja-ginza) (3.0.9)
Requirement already satisfied: thinc<8.3.0,>=8.2.2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
```

spacy<4.0.0,>=3.4.4->ja-ginza) (8.2.4)
 Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (1.1.3)
 Requirement already satisfied: srsly<3.0.0,>=2.4.3 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (2.4.8)
 Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (2.0.10)
 Requirement already satisfied: weasel<0.5.0,>=0.1.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (0.4.1)
 Requirement already satisfied: typer<1.0.0,>=0.3.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (0.12.3)
 Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (4.66.2)
 Requirement already satisfied: requests<3.0.0,>=2.13.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (2.31.0)
 Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (2.7.0)
 Requirement already satisfied: jinja2 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (3.1.3)
 Requirement already satisfied: setuptools in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (69.5.1)
 Requirement already satisfied: packaging>=20.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (23.2)
 Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (3.4.0)
 Requirement already satisfied: numpy>=1.19.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 spacy<4.0.0,>=3.4.4->ja-ginza) (1.26.4)
 Requirement already satisfied: language-data>=1.2 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 langcodes<4.0.0,>=3.2.0->spacy<4.0.0,>=3.4.4->ja-ginza) (1.2.0)
 Requirement already satisfied: annotated-types>=0.4.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<4.0.0,>=3.4.4->ja-ginza) (0.6.0)
 Requirement already satisfied: pydantic-core==2.18.1 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from

pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<4.0.0,>=3.4.4->ja-ginza) (2.18.1)
 Requirement already satisfied: typing-extensions>=4.6.1 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<4.0.0,>=3.4.4->ja-ginza) (4.11.0)
 Requirement already satisfied: charset-normalizer<4,>=2 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 requests<3.0.0,>=2.13.0->spacy<4.0.0,>=3.4.4->ja-ginza) (3.3.2)
 Requirement already satisfied: idna<4,>=2.5 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 requests<3.0.0,>=2.13.0->spacy<4.0.0,>=3.4.4->ja-ginza) (3.7)
 Requirement already satisfied: urllib3<3,>=1.21.1 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 requests<3.0.0,>=2.13.0->spacy<4.0.0,>=3.4.4->ja-ginza) (1.26.18)
 Requirement already satisfied: certifi>=2017.4.17 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 requests<3.0.0,>=2.13.0->spacy<4.0.0,>=3.4.4->ja-ginza) (2024.2.2)
 Requirement already satisfied: blis<0.8.0,>=0.7.8 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 thinc<8.3.0,>=8.2.2->spacy<4.0.0,>=3.4.4->ja-ginza) (0.7.11)
 Requirement already satisfied: confection<1.0.0,>=0.0.1 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 thinc<8.3.0,>=8.2.2->spacy<4.0.0,>=3.4.4->ja-ginza) (0.1.5)
 Requirement already satisfied: click>=8.0.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 typer<1.0.0,>=0.3.0->spacy<4.0.0,>=3.4.4->ja-ginza) (8.1.7)
 Requirement already satisfied: shellingham>=1.3.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 typer<1.0.0,>=0.3.0->spacy<4.0.0,>=3.4.4->ja-ginza) (1.5.4)
 Requirement already satisfied: rich>=10.11.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 typer<1.0.0,>=0.3.0->spacy<4.0.0,>=3.4.4->ja-ginza) (13.7.1)
 Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 weasel<0.5.0,>=0.1.0->spacy<4.0.0,>=3.4.4->ja-ginza) (0.18.1)
 Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 weasel<0.5.0,>=0.1.0->spacy<4.0.0,>=3.4.4->ja-ginza) (7.0.4)
 Requirement already satisfied: MarkupSafe>=2.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 jinja2->spacy<4.0.0,>=3.4.4->ja-ginza) (2.1.5)
 Requirement already satisfied: marisa-trie>=0.7.7 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from language-
 data>=1.2->langcodes<4.0.0,>=3.2.0->spacy<4.0.0,>=3.4.4->ja-ginza) (1.2.0)
 Requirement already satisfied: markdown-it-py>=2.2.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
 rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy<4.0.0,>=3.4.4->ja-ginza) (3.0.0)
 Requirement already satisfied: pygments<3.0.0,>=2.13.0 in
 /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from

```
rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy<4.0.0,>=3.4.4->ja-ginza) (2.17.2)
Requirement already satisfied: wrapt in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from smart-
open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy<4.0.0,>=3.4.4->ja-ginza)
(1.16.0)
Requirement already satisfied: mdurl~=0.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from markdown-it-
py>=2.2.0->rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy<4.0.0,>=3.4.4->ja-ginza)
(0.1.2)
```

0.1 1.

```
[3]: !pip install spacy
      !pip install sklearn
      !pip install python-dp
```

```
Requirement already satisfied: spacy in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (3.7.5)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(1.0.10)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(2.0.8)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(3.0.9)
Requirement already satisfied: thinc<8.3.0,>=8.2.2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(8.2.4)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(2.4.8)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(0.4.1)
```

Requirement already satisfied: typer<1.0.0,>=0.3.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(0.12.3)

Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(4.66.2)

Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(2.31.0)

Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(2.7.0)

Requirement already satisfied: jinja2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(3.1.3)

Requirement already satisfied: setuptools in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(69.5.1)

Requirement already satisfied: packaging>=20.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(23.2)

Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(3.4.0)

Requirement already satisfied: numpy>=1.19.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from spacy)
(1.26.4)

Requirement already satisfied: language-data>=1.2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
langcodes<4.0.0,>=3.2.0->spacy) (1.2.0)

Requirement already satisfied: annotated-types>=0.4.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (0.6.0)

Requirement already satisfied: pydantic-core==2.18.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (2.18.1)

Requirement already satisfied: typing-extensions>=4.6.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (4.11.0)

Requirement already satisfied: charset-normalizer<4,>=2 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
requests<3.0.0,>=2.13.0->spacy) (3.3.2)

Requirement already satisfied: idna<4,>=2.5 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
requests<3.0.0,>=2.13.0->spacy) (3.7)

Requirement already satisfied: urllib3<3,>=1.21.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
requests<3.0.0,>=2.13.0->spacy) (1.26.18)

```

Requirement already satisfied: certifi>=2017.4.17 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
requests<3.0.0,>=2.13.0->spacy) (2024.2.2)
Requirement already satisfied: blis<0.8.0,>=0.7.8 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
thinc<8.3.0,>=8.2.2->spacy) (0.7.11)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
thinc<8.3.0,>=8.2.2->spacy) (0.1.5)
Requirement already satisfied: click>=8.0.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
typer<1.0.0,>=0.3.0->spacy) (8.1.7)
Requirement already satisfied: shellingham>=1.3.0 in
/home/jun/.pyenv/versions/3.11.8/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/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
typer<1.0.0,>=0.3.0->spacy) (13.7.1)
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
weasel<0.5.0,>=0.1.0->spacy) (0.18.1)
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
weasel<0.5.0,>=0.1.0->spacy) (7.0.4)
Requirement already satisfied: MarkupSafe>=2.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
jinja2->spacy) (2.1.5)
Requirement already satisfied: marisa-trie>=0.7.7 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from language-
data>=1.2->langcodes<4.0.0,>=3.2.0->spacy) (1.2.0)
Requirement already satisfied: markdown-it-py>=2.2.0 in
/home/jun/.pyenv/versions/3.11.8/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/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (2.17.2)
Requirement already satisfied: wrapt in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from smart-
open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy) (1.16.0)
Requirement already satisfied: mdurl~=0.1 in
/home/jun/.pyenv/versions/3.11.8/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)
Collecting sklearn
  Using cached sklearn-0.0.post12.tar.gz (2.6 kB)
  Installing build dependencies ... done
  Getting requirements to build wheel ... error
  error: subprocess-exited-with-error

```

```

× Getting requirements to build wheel did not run
successfully.
  exit code: 1
  > [15 lines of output]
    The 'sklearn' PyPI package is deprecated, use 'scikit-learn'
    rather than 'sklearn' for pip commands.

    Here is how to fix this error in the main use cases:
    - use 'pip install scikit-learn' rather than 'pip install
sklearn'
    - replace 'sklearn' by 'scikit-learn' in your pip requirements
files
      (requirements.txt, setup.py, setup.cfg, Pipfile, etc ...)
    - if the 'sklearn' package is used by one of your dependencies,
      it would be great if you take some time to track which package
uses
      'sklearn' instead of 'scikit-learn' and report it to their
issue tracker
    - as a last resort, set the environment variable
      SKLEARN_ALLOW_DEPRECATED_SKLEARN_PACKAGE_INSTALL=True to avoid
this error

    More information is available at
    https://github.com/scikit-learn/sklearn-pypi-package
    [end of output]

```

note: This error originates from a subprocess, and is likely not a problem with pip.

error: subprocess-exited-with-error

```

× Getting requirements to build wheel did not run
successfully.
  exit code: 1
  > See above for output.

```

note: This error originates from a subprocess, and is likely not a problem with pip.
Requirement already satisfied: python-dp in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (1.1.4)

0.2 2.

```

[4]: import pandas as pd

data_path = './data/reviews_with_sentiment.csv'

df = pd.read_csv(data_path)

```

```
df
```

```
[4]:
```

	review	sentiment
0		neutral
1		neutral
2		positive
3		positive
4		positive
...
5548		negative
5549		negative
5550		neutral
5551		negative
5552	...	neutral

[5553 rows x 2 columns]

```
[5]: import pandas as pd
import numpy as np
from sklearn.feature_extraction.text import CountVectorizer
import spacy
import itertools
from typing import List, Tuple
from pydp.algorithms.laplacian import Count
import math

# spaCy
nlp = spacy.load('ja_ginza')

#
data_path = './data/reviews_with_sentiment.csv'
df = pd.read_csv(data_path)

#
POS = ['ADJ', 'ADV', 'INTJ', 'PROPN', 'NOUN', 'VERB']
MAX_TERMS_IN_DOC = 5
NGRAM = 1
MAX_DF = 1.0
MIN_DF = 0.0
NUM_VOCAB = 10000
TOP_K = 20

def flatten(*lists) -> list:
    res = []
    for l in list(itertools.chain.from_iterable(lists)):
```



```

        for e in l:
            res.append(e)
    return res

def remove_duplicates(l: List[Tuple[str, float]]) -> List[Tuple[str, float]]:
    d = {}
    for e in l:
        d[e[0]] = e[1]
    return list(d.items())

def preprocess_for_private_counts(tf: np.ndarray) -> List[np.ndarray]:
    repeated_words = []
    for i, term in enumerate(tf):
        repeated_words.append(np.repeat(i, term))
    return repeated_words

def cal_private_count(epsilon: float, max_partition_contributed: float,
    ↪max_contributions_per_partition: float, repeated_words: List[np.ndarray]) ->
    ↪List[int]:
    private_counts = []
    for repeated_word in repeated_words:
        counter = Count(epsilon, max_partition_contributed,
    ↪max_contributions_per_partition)
        count = counter.quick_result(repeated_word)
        private_counts.append(count)
    return private_counts

def top_k_words_and_counts(k: int, tf: np.ndarray, vocab: dict) ->
    ↪List[Tuple[str, int]]:
    indices_topk = np.argsort(tf)[::-1][:k]
    reverse_vocab = {vocab[key]: key for key in vocab.keys()}
    words = [reverse_vocab[i] for i in indices_topk]
    counts = [tf[i] for i in indices_topk]
    return list(zip(words, counts))

#
df["doc"] = [nlp(review) for review in df["review"]]

```

```

/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-
packages/torch/cuda/__init__.py:118: UserWarning: CUDA initialization: CUDA
unknown error - this may be due to an incorrectly set up environment, e.g.
changing env variable CUDA_VISIBLE_DEVICES after program start. Setting the
available devices to be zero. (Triggered internally at
../c10/cuda/CUDAFunctions.cpp:108.)
    return torch._C._cuda_getDeviceCount() > 0

```

0.3 3. Bag-of-Words

```
[6]: # Bag-of-Words
bows = {}
cvs = {}

for sentiment in df["sentiment"].unique():
    tokens = []
    for doc in df[df["sentiment"] == sentiment]["doc"]:
        similarities = [(token.similarity(doc), token.lemma_) for token in doc
        ↪if token.pos_ in POS]
        similarities = remove_duplicates(similarities)
        similarities = sorted(similarities, key=lambda sim: sim[1],
        ↪reverse=True)[:MAX_TERMS_IN_DOC]
        tokens.append([similarity[1] for similarity in similarities])

    cv = CountVectorizer(ngram_range=(1, NGRAM), max_df=MAX_DF, min_df=MIN_DF,
    ↪max_features=NUM_VOCAB)
    bows[sentiment] = cv.fit_transform(flatten(tokens)).toarray()
    cvs[sentiment] = cv
```

/tmp/ipykernel_942954/2647833454.py:8: UserWarning: [W008] Evaluating
Token.similarity based on empty vectors.

```
similarities = [(token.similarity(doc), token.lemma_) for token in doc if
token.pos_ in POS]
```

```
[7]: !pip install numpy
```

Requirement already satisfied: numpy in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (1.26.4)

0.4 4.

```
[8]: #
vocab = {}
term_frequencies = {}

for sentiment in df["sentiment"].unique():
    bow = bows[sentiment]
    cv = cvs[sentiment]

    vocab = cv.vocabulary_
    term_frequency = np.sum(bow, axis=0)
    vocab[sentiment] = vocab
    term_frequencies[sentiment] = term_frequency

    indices_topk = np.argsort(term_frequency)[::-1][:TOP_K]
    bow_topk = np.take(bow, indices_topk, axis=1)
```

```

reverse_vocab = {vocab[k]: k for k in vocab.keys()}
words = [reverse_vocab[i] for i in indices_topk]

print(sentiment, ":")
for w, c in zip(words, term_frequency[indices_topk]):
    print(w, ":", c)

```

neutral :

```

: 216
: 93
: 67
: 60
: 59
: 50
: 47
: 45
: 43
: 43
: 34
: 32
: 30
: 27
: 27
: 25
: 25
: 23
: 22
: 21

```

positive :

```

: 463
: 315
: 206
: 198
: 193
: 181
: 175
: 144
: 141
: 140
: 124
: 106
: 102
: 99
: 96
: 90
: 90
: 87

```

```

: 84
: 80
negative :
: 99
: 97
: 48
: 42
: 39
: 36
: 34
: 33
: 27
: 24
: 22
: 19
: 16
: 16
: 16
: 16
: 15
: 15
: 15
: 14

```

0.5 5.

```

[9]: #
    epsilons = [0.01, 0.05, 0.1, 0.3, 0.7, 1.0, 2.0, 3.0, 7.0, 10.0]
    MAX_DUPLICATED_TERMS = 1

    #
    results = {sentiment: {"no DP": None, **{eps: None for eps in epsilons}} for
    ↪ sentiment in df["sentiment"].unique()}

    for eps in epsilons:
        print(" : ", eps)
        for sentiment in df["sentiment"].unique():
            repeated_words =
            ↪ preprocess_for_private_counts(term_frequencies[sentiment])
            private_counts = cal_private_count(eps, MAX_TERMS_IN_DOC,
            ↪ MAX_DUPLICATED_TERMS, repeated_words)

            words_and_counts = top_k_words_and_counts(TOP_K, private_counts,
            ↪ vocabs[sentiment])
            results[sentiment][eps] = words_and_counts
            print(sentiment, ":")
            print(words_and_counts)

```

```

: 0.01
neutral :
[(' ', 4047), (' ', 3106), (' ', 3092), (' ', 3017), (' ', 2689), (' ',
2586), (' ', 2534), (' ', 2462), (' ', 2415), (' ', 2396), (' ', 2336),
(' ', 2282), (' ', 2202), (' ', 2006), (' ', 1950), (' ', 1882),
(' ', 1836), (' ', 1794), (' ', 1768), (' ', 1733)]
positive :
[(' ', 4306), (' ', 3461), (' ', 3053), (' ', 2745), (' ', 2569),
(' ', 2507), (' ', 2391), (' ', 2354), (' ', 2346), (' ', 2282),
(' ', 2255), (' ', 2141), (' ', 2101), (' ', 2056), (' ', 2052),
(' ', 2048), (' ', 2038), (' ', 2006), (' ', 1998), (' ', 1970)]
negative :
[(' ', 2921), (' ', 2699), (' ', 2426), (' ', 2324), (' ', 2266), (' ',
2106), (' ', 2017), (' ', 2007), (' ', 1981), (' ', 1853), (' ', 1844),
(' ', 1834), (' ', 1826), (' ', 1809), (' ', 1773), (' ', 1761), (' ',
1738), (' ', 1715), (' ', 1669), (' ', 1655)]
: 0.05
neutral :
[(' ', 619), (' ', 597), (' ', 583), (' ', 546), (' ', 508), (' ', 495),
(' ', 494), (' ', 492), (' ', 478), (' ', 458), (' ', 457), (' ', 444),
(' ', 421), (' ', 399), (' ', 390), (' ', 389), (' ', 389), (' ',
388), (' ', 382), (' ', 372)]
positive :
[(' ', 1090), (' ', 870), (' ', 655), (' ', 498), (' ', 469), (' ', 463),
(' ', 461), (' ', 450), (' ', 447), (' ', 439), (' ', 435), (' ',
430), (' ', 406), (' ', 403), (' ', 402), (' ', 398), (' ', 398), (' ',
396), (' ', 391), (' ', 389)]
negative :
[(' ', 520), (' ', 496), (' ', 480), (' ', 462), (' ', 457), (' ',
450), (' ', 428), (' ', 392), (' ', 385), (' ', 383), (' ', 356), (' ',
353), (' ', 347), (' ', 342), (' ', 320), (' ', 316), (' ', 312), (' ',
308), (' ', 297), (' ', 296)]
: 0.1
neutral :
[(' ', 374), (' ', 291), (' ', 289), (' ', 262), (' ', 260), (' ',
243), (' ', 230), (' ', 211), (' ', 197), (' ', 194), (' ', 188),
(' ', 188), (' ', 185), (' ', 183), (' ', 183), (' ', 182), (' ',
175), (' ', 169), (' ', 165), (' ', 164)]
positive :
[(' ', 499), (' ', 386), (' ', 354), (' ', 284), (' ', 268), (' ', 261),
(' ', 256), (' ', 255), (' ', 249), (' ', 246), (' ', 233), (' ', 223),
(' ', 221), (' ', 220), (' ', 218), (' ', 213), (' ', 211), (' ', 206),
(' ', 204), (' ', 203)]
negative :
[(' ', 253), (' ', 232), (' ', 229), (' ', 219), (' ', 218), (' ', 213),
(' ', 197), (' ', 196), (' ', 194), (' ', 183), (' ', 182), (' ',
182), (' ', 179), (' ', 178), (' ', 176), (' ', 173), (' ', 168), (' ',
166), (' ', 164), (' ', 163)]

```

```

: 0.3
neutral :
[(' ', 204), (' ', 125), (' ', 118), (' ', 113), (' ', 98), (' ', 81),
(' ', 77), (' ', 72), (' ', 72), (' ', 70), (' ', 69), (' ', 69),
(' ', 66), (' ', 66), (' ', 66), (' ', 65), (' ', 65), (' ', 64),
(' ', 62), (' ', 61)]
positive :
[(' ', 452), (' ', 354), (' ', 220), (' ', 197), (' ', 193), (' ', 170),
(' ', 161), (' ', 157), (' ', 148), (' ', 138), (' ', 137), (' ', 137),
(' ', 129), (' ', 118), (' ', 117), (' ', 110), (' ', 108), (' ', 106),
(' ', 105), (' ', 103)]
negative :
[(' ', 133), (' ', 117), (' ', 116), (' ', 110), (' ', 98), (' ', 96),
(' ', 76), (' ', 75), (' ', 74), (' ', 72), (' ', 69), (' ', 66), (' ',
66), (' ', 65), (' ', 65), (' ', 64), (' ', 63), (' ', 63), (' ', 60),
(' ', 59)]
: 0.7
neutral :
[(' ', 222), (' ', 98), (' ', 74), (' ', 66), (' ', 56), (' ', 53), (' ',
52), (' ', 48), (' ', 43), (' ', 41), (' ', 40), (' ', 38), (' ',
37), (' ', 37), (' ', 36), (' ', 35), (' ', 35), (' ', 33), ('no',
33), (' ', 33)]
positive :
[(' ', 465), (' ', 327), (' ', 205), (' ', 201), (' ', 195), (' ', 184),
(' ', 173), (' ', 149), (' ', 146), (' ', 145), (' ', 129), (' ', 109),
(' ', 102), (' ', 100), (' ', 97), (' ', 93), (' ', 93), (' ', 89),
(' ', 87), (' ', 83)]
negative :
[(' ', 90), (' ', 87), (' ', 45), (' ', 42), (' ', 42), (' ', 41), (' ',
39), (' ', 38), (' ', 37), (' ', 37), (' ', 35), (' ', 34), (' ', 33),
(' ', 31), (' ', 31), (' ', 30), (' ', 30), (' ', 30), (' ', 29),
(' ', 29)]
: 1.0
neutral :
[(' ', 216), (' ', 96), (' ', 68), (' ', 63), (' ', 59), (' ', 58), (' ',
58), (' ', 48), (' ', 47), (' ', 45), (' ', 45), (' ', 43), (' ', 40),
(' ', 39), (' ', 38), (' ', 34), (' ', 32), (' ', 31), (' ', 29), (' ',
28)]
positive :
[(' ', 465), (' ', 318), (' ', 203), (' ', 198), (' ', 189), (' ', 183),
(' ', 171), (' ', 158), (' ', 145), (' ', 139), (' ', 126), (' ', 114),
(' ', 104), (' ', 96), (' ', 95), (' ', 91), (' ', 90), (' ', 86), (' ',
83), (' ', 82)]
negative :
[(' ', 98), (' ', 88), (' ', 45), (' ', 40), (' ', 40), (' ', 39), (' ',
38), (' ', 37), (' ', 33), (' ', 30), (' ', 29), (' ', 27), (' ', 26),
(' ', 25), (' ', 25), (' ', 24), (' ', 23), (' ', 22), (' ', 22), (' ',
22)]

```

```

: 2.0
neutral :
[( ' ', 216), ( ' ', 95), ( ' ', 67), ( ' ', 62), ( ' ', 53), ( ' ', 50), ( ' ',
47), ( ' ', 47), ( ' ', 44), ( ' ', 43), ( ' ', 42), ( ' ', 32), ( ' ', 30),
( ' ', 28), ( ' ', 26), ( ' ', 24), ( ' ', 24), ( ' ', 23), ( ' ', 23), ( ' ',
22)]
positive :
[( ' ', 461), ( ' ', 318), ( ' ', 203), ( ' ', 200), ( ' ', 192), ( ' ', 181),
( ' ', 172), ( ' ', 141), ( ' ', 140), ( ' ', 138), ( ' ', 121), ( ' ', 102),
( ' ', 102), ( ' ', 98), ( ' ', 97), ( ' ', 96), ( ' ', 90), ( ' ', 86), ( ' ',
84), ( ' ', 81)]
negative :
[( ' ', 101), ( ' ', 97), ( ' ', 52), ( ' ', 41), ( ' ', 40), ( ' ', 33), ( ' ',
33), ( ' ', 31), ( ' ', 30), ( ' ', 29), ( ' ', 22), ( ' ', 19), ( ' ', 18),
( ' ', 17), ( ' ', 17), ( ' ', 17), ( ' ', 15), ( ' ', 15), ( ' ', 14), ( ' ',
14)]
: 3.0
neutral :
[( ' ', 215), ( ' ', 91), ( ' ', 68), ( ' ', 63), ( ' ', 58), ( ' ', 49), ( ' ',
48), ( ' ', 43), ( ' ', 43), ( ' ', 39), ( ' ', 39), ( ' ', 37), ( ' ', 31),
( ' ', 28), ( ' ', 28), ( ' ', 27), ( ' ', 26), ( ' ', 25), ( ' ', 24), ( ' ',
21)]
positive :
[( ' ', 463), ( ' ', 314), ( ' ', 205), ( ' ', 198), ( ' ', 191), ( ' ', 177),
( ' ', 175), ( ' ', 142), ( ' ', 141), ( ' ', 140), ( ' ', 123), ( ' ', 103),
( ' ', 98), ( ' ', 96), ( ' ', 95), ( ' ', 89), ( ' ', 89), ( ' ', 84), ( ' ',
83), ( ' ', 79)]
negative :
[( ' ', 100), ( ' ', 98), ( ' ', 50), ( ' ', 43), ( ' ', 40), ( ' ', 34), ( ' ',
34), ( ' ', 32), ( ' ', 26), ( ' ', 22), ( ' ', 21), ( ' ', 18), ( ' ', 18),
( ' ', 17), ( ' ', 17), ( ' ', 16), ( ' ', 16), ( ' ', 16), ( ' ', 16), ( ' ',
15)]
: 7.0
neutral :
[( ' ', 216), ( ' ', 95), ( ' ', 67), ( ' ', 60), ( ' ', 59), ( ' ', 49), ( ' ',
47), ( ' ', 46), ( ' ', 45), ( ' ', 44), ( ' ', 32), ( ' ', 31), ( ' ', 30),
( ' ', 28), ( ' ', 27), ( ' ', 25), ( ' ', 24), ( ' ', 23), ( ' ', 22), ( ' ',
21)]
positive :
[( ' ', 463), ( ' ', 316), ( ' ', 205), ( ' ', 198), ( ' ', 192), ( ' ', 182),
( ' ', 175), ( ' ', 146), ( ' ', 143), ( ' ', 140), ( ' ', 124), ( ' ', 106),
( ' ', 102), ( ' ', 97), ( ' ', 95), ( ' ', 90), ( ' ', 89), ( ' ', 88), ( ' ',
87), ( ' ', 81)]
negative :
[( ' ', 100), ( ' ', 98), ( ' ', 49), ( ' ', 43), ( ' ', 38), ( ' ', 37), ( ' ',
34), ( ' ', 32), ( ' ', 28), ( ' ', 25), ( ' ', 23), ( ' ', 19), ( ' ', 18),
( ' ', 17), ( ' ', 17), ( ' ', 16), ( ' ', 16), ( ' ', 15), ( ' ', 15),
( ' ', 15)]

```

```

: 10.0
neutral :
[(' ', 216), (' ', 93), (' ', 67), (' ', 60), (' ', 58), (' ', 50), (' ',
47), (' ', 44), (' ', 43), (' ', 43), (' ', 33), (' ', 32), (' ', 30),
(' ', 28), (' ', 27), (' ', 26), (' ', 25), (' ', 24), (' ', 23), (' ',
21)]
positive :
[(' ', 463), (' ', 315), (' ', 206), (' ', 198), (' ', 195), (' ', 180),
(' ', 175), (' ', 144), (' ', 142), (' ', 140), (' ', 124), (' ', 108),
(' ', 104), (' ', 99), (' ', 96), (' ', 91), (' ', 90), (' ', 87), (' ',
85), (' ', 80)]
negative :
[(' ', 99), (' ', 97), (' ', 48), (' ', 42), (' ', 38), (' ', 36), (' ',
33), (' ', 32), (' ', 28), (' ', 24), (' ', 22), (' ', 19), (' ', 17),
(' ', 16), (' ', 16), (' ', 16), (' ', 15), (' ', 15), (' ', 15),
(' ', 15)]

```

0.6 6.

```

[10]: #
for sentiment in df["sentiment"].unique():
    words_and_counts = top_k_words_and_counts(TOP_K,
    ↪term_frequencies[sentiment], vocabs[sentiment])
    results[sentiment]["no DP"] = words_and_counts

#
def display_results(results):
    for sentiment, data in results.items():
        print(f"\nSentiment: {sentiment}\n")
        df = pd.DataFrame(columns=["rank", "word", "count"] + [f"={eps}" for
    ↪eps in epsilons])
        for rank, (word, count) in enumerate(data["no DP"], start=1):
            row = {"rank": rank, "word": word, "count": count}
            for eps in epsilons:
                if eps in data and rank <= len(data[eps]):
                    row[f"={eps}"] = data[eps][rank-1][1] if rank-1 <
    ↪len(data[eps]) else None
            df = pd.concat([df, pd.DataFrame([row])], ignore_index=True)
        print(df)

display_results(results)

```

Sentiment: neutral

	rank	word	count	=0.01	=0.05	=0.1	=0.3	=0.7	=1.0	=2.0	=3.0	=7.0 \
0	1		216	4047	619	374	204	222	216	216	215	216
1	2		93	3106	597	291	125	98	96	95	91	95

2	3	67	3092	583	289	118	74	68	67	68	67
3	4	60	3017	546	262	113	66	63	62	63	60
4	5	59	2689	508	260	98	56	59	53	58	59
5	6	50	2586	495	243	81	53	58	50	49	49
6	7	47	2534	494	230	77	52	58	47	48	47
7	8	45	2462	492	211	72	48	48	47	43	46
8	9	43	2415	478	197	72	43	47	44	43	45
9	10	43	2396	458	194	70	41	45	43	39	44
10	11	34	2336	457	188	69	40	45	42	39	32
11	12	32	2282	444	188	69	38	43	32	37	31
12	13	30	2202	421	185	66	37	40	30	31	30
13	14	27	2006	399	183	66	37	39	28	28	28
14	15	27	1950	390	183	66	36	38	26	28	27
15	16	25	1882	389	182	65	35	34	24	27	25
16	17	25	1836	389	175	65	35	32	24	26	24
17	18	23	1794	388	169	64	33	31	23	25	23
18	19	22	1768	382	165	62	33	29	23	24	22
19	20	21	1733	372	164	61	33	28	22	21	21

=10.0

0	216
1	93
2	67
3	60
4	58
5	50
6	47
7	44
8	43
9	43
10	33
11	32
12	30
13	28
14	27
15	26
16	25
17	24
18	23
19	21

Sentiment: positive

	rank	word count	=0.01	=0.05	=0.1	=0.3	=0.7	=1.0	=2.0	=3.0	=7.0	\
0	1	463	4306	1090	499	452	465	465	461	463	463	
1	2	315	3461	870	386	354	327	318	318	314	316	
2	3	206	3053	655	354	220	205	203	203	205	205	
3	4	198	2745	498	284	197	201	198	200	198	198	

4	5	193	2569	469	268	193	195	189	192	191	192
5	6	181	2507	463	261	170	184	183	181	177	182
6	7	175	2391	461	256	161	173	171	172	175	175
7	8	144	2354	450	255	157	149	158	141	142	146
8	9	141	2346	447	249	148	146	145	140	141	143
9	10	140	2282	439	246	138	145	139	138	140	140
10	11	124	2255	435	233	137	129	126	121	123	124
11	12	106	2141	430	223	137	109	114	102	103	106
12	13	102	2101	406	221	129	102	104	102	98	102
13	14	99	2056	403	220	118	100	96	98	96	97
14	15	96	2052	402	218	117	97	95	97	95	95
15	16	90	2048	398	213	110	93	91	96	89	90
16	17	90	2038	398	211	108	93	90	90	89	89
17	18	87	2006	396	206	106	89	86	86	84	88
18	19	84	1998	391	204	105	87	83	84	83	87
19	20	80	1970	389	203	103	83	82	81	79	81

=10.0

0	463
1	315
2	206
3	198
4	195
5	180
6	175
7	144
8	142
9	140
10	124
11	108
12	104
13	99
14	96
15	91
16	90
17	87
18	85
19	80

Sentiment: negative

	rank	word count	=0.01	=0.05	=0.1	=0.3	=0.7	=1.0	=2.0	=3.0	=7.0	\
0	1	99	2921	520	253	133	90	98	101	100	100	
1	2	97	2699	496	232	117	87	88	97	98	98	
2	3	48	2426	480	229	116	45	45	52	50	49	
3	4	42	2324	462	219	110	42	40	41	43	43	
4	5	39	2266	457	218	98	42	40	40	40	38	
5	6	36	2106	450	213	96	41	39	33	34	37	

6	7	34	2017	428	197	76	39	38	33	34	34
7	8	33	2007	392	196	75	38	37	31	32	32
8	9	27	1981	385	194	74	37	33	30	26	28
9	10	24	1853	383	183	72	37	30	29	22	25
10	11	22	1844	356	182	69	35	29	22	21	23
11	12	19	1834	353	182	66	34	27	19	18	19
12	13	16	1826	347	179	66	33	26	18	18	18
13	14	16	1809	342	178	65	31	25	17	17	17
14	15	16	1773	320	176	65	31	25	17	17	17
15	16	16	1761	316	173	64	30	24	17	16	16
16	17	15	1738	312	168	63	30	23	15	16	16
17	18	15	1715	308	166	63	30	22	15	16	15
18	19	15	1669	297	164	60	29	22	14	16	15
19	20	14	1655	296	163	59	29	22	14	15	15

=10.0

0	99
1	97
2	48
3	42
4	38
5	36
6	33
7	32
8	28
9	24
10	22
11	19
12	17
13	16
14	16
15	16
16	15
17	15
18	15
19	15

0.7 7.

```
[11]: !pip install matplotlib
      !pip install seaborn
```

```
Requirement already satisfied: matplotlib in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (3.8.4)
Requirement already satisfied: contourpy>=1.0.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib)
(1.2.1)
Requirement already satisfied: cycler>=0.10 in
```

/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib) (0.12.1)
 Requirement already satisfied: fonttools>=4.22.0 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib) (4.51.0)
 Requirement already satisfied: kiwisolver>=1.3.1 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib) (1.4.5)
 Requirement already satisfied: numpy>=1.21 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib) (1.26.4)
 Requirement already satisfied: packaging>=20.0 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib) (23.2)
 Requirement already satisfied: pillow>=8 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib) (10.3.0)
 Requirement already satisfied: pyparsing>=2.3.1 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib) (3.1.2)
 Requirement already satisfied: python-dateutil>=2.7 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib) (2.9.0.post0)
 Requirement already satisfied: six>=1.5 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
 Requirement already satisfied: seaborn in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (0.13.2)
 Requirement already satisfied: numpy!=1.24.0,>=1.20 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from seaborn) (1.26.4)
 Requirement already satisfied: pandas>=1.2 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from seaborn) (2.2.2)
 Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from seaborn) (3.8.4)
 Requirement already satisfied: contourpy>=1.0.1 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.2.1)
 Requirement already satisfied: cycler>=0.10 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (0.12.1)
 Requirement already satisfied: fonttools>=4.22.0 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (4.51.0)
 Requirement already satisfied: kiwisolver>=1.3.1 in /home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from

```

matplotlib!=3.6.1,>=3.4->seaborn) (1.4.5)
Requirement already satisfied: packaging>=20.0 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (23.2)
Requirement already satisfied: pillow>=8 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (10.3.0)
Requirement already satisfied: pyparsing>=2.3.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from
pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: six>=1.5 in
/home/jun/.pyenv/versions/3.11.8/lib/python3.11/site-packages (from python-
dateutil>=2.7->matplotlib!=3.6.1,>=3.4->seaborn) (1.16.0)

```

```

[12]: #
import matplotlib.pyplot as plt
import matplotlib.font_manager as fm

#
font_path = '/usr/share/fonts/truetype/fonts-japanese-gothic.ttf' #
font_prop = fm.FontProperties(fname=font_path)

def calculate_match_rate(original_top_k, dp_top_k):
    match_count = len(set(original_top_k) & set(dp_top_k))
    return match_count / len(original_top_k) if original_top_k else 0

def plot_match_rate(results, epsilons, sentiment_label):
    top_k_values = [3, 5, 10, 20]
    match_rates = {k: [] for k in top_k_values}

    for eps in epsilons:
        for k in top_k_values:
            original_top_k = [word for word, count in
↪results[sentiment_label]["no DP"][:k]]
            dp_top_k = [word for word, count in results[sentiment_label][eps][:
↪k]]

            match_rate = calculate_match_rate(original_top_k, dp_top_k)
            match_rates[k].append(match_rate)

```

```

plt.figure(figsize=(10, 6))
for k, rates in match_rates.items():
    plt.plot(epsilons, rates, label=f'top{k}')

plt.xlabel(' ', fontproperties=font_prop)
plt.ylabel(' ', fontproperties=font_prop)
plt.title(f'({sentiment_label})', fontproperties=font_prop)
plt.legend(prop=font_prop)
plt.grid(True)
plt.show()

epsilons = [0.01, 0.05, 0.1, 0.3, 0.7, 1.0, 2.0, 3.0, 7.0, 10.0]

plot_match_rate(results, epsilons, 'negative')
plot_match_rate(results, epsilons, 'neutral')
plot_match_rate(results, epsilons, 'positive')

```



