

NLP Assignment-3

Name : P.Rahul

Roll No : 2203A51439

Batch : 12

1. Implement Word Embeddings on following text

words = ['king', 'queen', 'man', 'woman', 'paris', 'france', 'london', 'england']. [CO1]

Importing necessary libraries

```
+ Code + Text
[5] !pip install gensim spacy matplotlib scikit-learn
!python -m spacy download en_core_web_sm

Requirement already satisfied: gensim in /usr/local/lib/python3.10/dist-packages (4.3.3)
Requirement already satisfied: spacy in /usr/local/lib/python3.10/dist-packages (3.7.6)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (3.7.1)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.10/dist-packages (1.3.2)
Requirement already satisfied: numpy<2.0,>=1.18.5 in /usr/local/lib/python3.10/dist-packages (from gensim) (1.26.4)
Requirement already satisfied: scipy<1.14.0,>=1.7.0 in /usr/local/lib/python3.10/dist-packages (from gensim) (1.13.1)
Requirement already satisfied: smart-open>=1.8.1 in /usr/local/lib/python3.10/dist-packages (from gensim) (7.0.4)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.0.0,>=0.28.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.0.10)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.8)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.0.9)
Requirement already satisfied: thinc<8.3.0,>=8.2.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (8.2.5)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.4.8)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (0.4.1)
Requirement already satisfied: typer<5.0.0,>=0.3.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (0.12.5)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (4.66.5)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.32.3)
Requirement already satisfied: pydantic<1.8.1,1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.9.1)
Requirement already satisfied: Jinja2 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.1.4)
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from spacy) (71.0.4)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (24.1)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.4.0)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.3.0)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (4.53.1)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.4.7)
Requirement already satisfied: pillow<9.0.0,>=8.3.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (9.4.0)

3s completed at 11:33 AM
```

Implementing **word embeddings** on the given text

```
+ Code + Text
[6] import gensim.downloader as api
import numpy as np

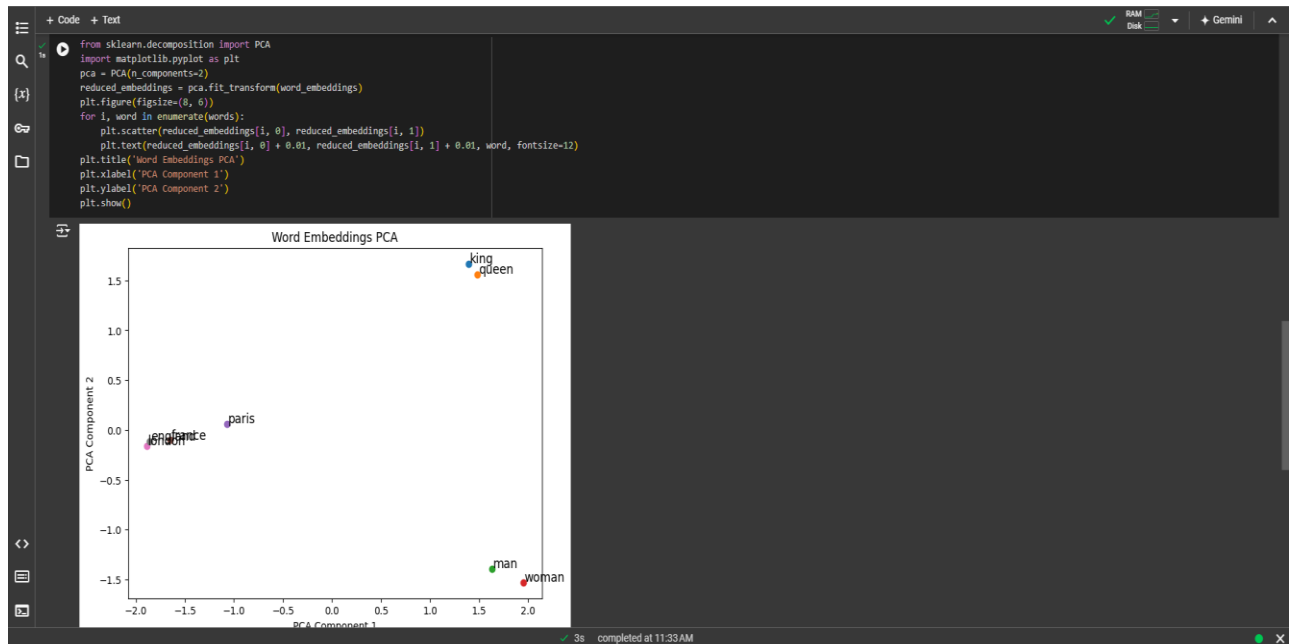
model = api.load('word2vec-google-news-300')
words = ['king', 'queen', 'man', 'woman', 'paris', 'france', 'london', 'england']
word_embeddings = np.array([model[word] for word in words])
word_embeddings

array([[ 0.12597656,  0.02978516,  0.00860596, ..., -0.0859375 ,
         0.09130859,  0.25195312],
       [ 0.00524902, -0.14355469, -0.06933594, ..., -0.046875 ,
         0.16308594,  0.15429688],
       [ 0.32617188,  0.13085938,  0.03466797, ..., -0.30273438,
        -0.08007812,  0.02099609],
       ...,
       [-0.20605469, -0.16699219,  0.19238281, ..., -0.04248047,
        -0.07958984, -0.10058594],
       [-0.4609375 ,  0.0279541 , -0.26757812, ..., -0.16992188,
        -0.02258301,  0.04150391],
       [-0.3671875 , -0.03491211,  0.11083984, ...,  0.07373047,
        -0.22851562,  0.05859375]], dtype=float32)
```

2. Implement Word Embeddings using PCA (Principal Component Analysis) on following text

words = ['king', 'queen', 'man', 'woman', 'paris', 'france', 'london', 'england']. [CO1]

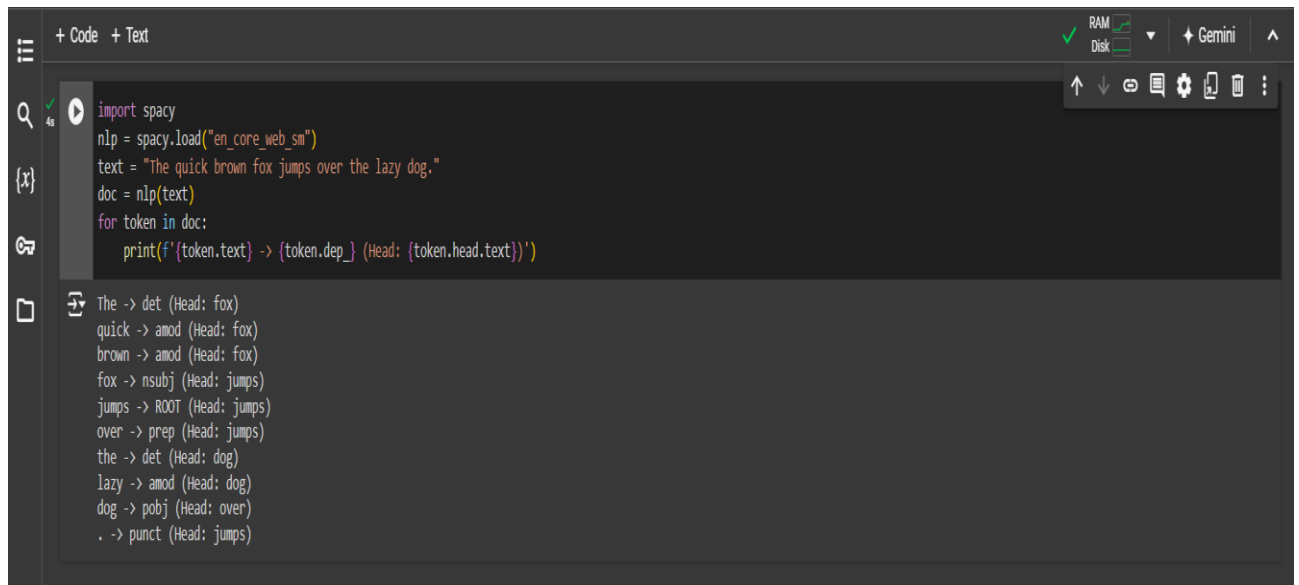
Implementing **Word Embeddings using PCA** on given text



3. Execute Dependency Parsing on any text or on following text [CO1]

The quick brown fox jumps over the lazy dog

Executing **Dependency Parsing** on given text



```
+ Code + Text
import spacy
nlp = spacy.load("en_core_web_sm")
text = "The quick brown fox jumps over the lazy dog."
doc = nlp(text)
for token in doc:
    print(f'{token.text} -> {token.dep_} (Head: {token.head.text})')

The -> det (Head: fox)
quick -> amod (Head: fox)
brown -> amod (Head: fox)
fox -> nsubj (Head: jumps)
jumps -> ROOT (Head: jumps)
over -> prep (Head: jumps)
the -> det (Head: dog)
lazy -> amod (Head: dog)
dog -> pobj (Head: over)
. -> punct (Head: jumps)
```