IRE Assignment 4 - Report

Word2Vec

Query:

integrated analysis

Top 10 results and cosine similarity:

```
[(0.97152615, 'sbr12195.txt'),
 (0.9710187, 'emt10695.txt'),
 (0.969862, 'sbr06195.txt'),
 (0.9689544, 'mat06395.txt'),
 (0.9687294, 'sbr17995.txt'),
 (0.96836835, 'sbr01495.txt'),
 (0.96698225, 'sbr01395.txt'),
 (0.9668887, 'ins20495.txt'),
 (0.966227, 'emt01995.txt'),
 (0.9661464, 'sbr21495.txt')]
```

Expected Result: emt01995.txt

Spacy

Query:

integrated analysis

Top 10 results and cosine similarity:

```
[(0.3569761792529935, 'emt05095.txt'),
 (0.33977543846728053, 'sbr06295.txt'),
 (0.3384800826116411, 'sbr01495.txt'),
 (0.33514201671148763, 'emt04995.txt'),
 (0.3346343367249719, 'eos03795.txt'),
 (0.33456275763460896, 'sbr01395.txt'),
 (0.3337068261071418, 'eos16095.txt'),
 (0.32824074225209005, 'sbr18095.txt'),
 (0.3243281549194393, 'ins20595.txt'),
 (0.32026474294644597, 'inf12795.txt')]
```

Expected Result: emt01995.txt

DPR

Query:

integrated analysis

Top 10 results and cosine similarity:

```
[(0.63890994, 'eos11695.txt'),
 (0.57324654, 'inf11495.txt'),
 (0.5717596, 'inf21695.txt'),
 (0.56599337, 'emt01995.txt'),
 (0.5652106, 'sbr17995.txt'),
 (0.54340976, 'emt13895.txt'),
 (0.5345992, 'sbr17895.txt'),
 (0.53454757, 'eos19995.txt'),
 (0.5314743, 'emt02495.txt'),
 (0.52591914, 'eos07195.txt')]
```

Expected Result: emt01995.txt