

Question for Vector Space Model worked example.

Here is a simplified example of the vector space retrieval model. Consider a very small collection C that consists of the following three documents:

d1: “new york times”

d2: “new york post”

d3: “los angeles times”

- 1) Using the TF-IDF weighting scheme, create vectors to represent each of these documents, so that a query can be processed according to the vector space model.
- 2) Given the query: “new new times”, calculate a TF-IDF vector to represent the query, compute its cosine similarity with each document vector, and output a ranked list of the documents.