START OF QUIZ
Student ID:
95153730,Ihn,Jae

Question 1

Topic: Lecture 5
Source: Lecture 5

The Frobenius norm looks very similar to a distance metric we've already observed. Explain which one. (1)

Question 2

Topic: Lecture 6
Source: Lecture 6

Why do we need a ”human in the loop” for topic modeling? (1)

Question 3

Topic: Lecture 8
Source: Lecture 8

Why do we not simply take the probability of a word given its document (maybe with smoothing added in)? (1)

Question 4

Topic: Lecture 7
Source: Lecture 7

Explain why boolean filtering is usually insufficient for retrieval, and why we normally need some way of scoring the documents. (2)

Question 5

Topic: Lecture 8
Source: Lecture 8

What do we mean by interpolation? (1)

Question 6

Topic: Lecture 5
Source: Lecture 5

What advantages do sparse vectors have over dense ones. (1)

Question 7

Topic: Lecture 7
Source: Lecture 7

Explain why the cosine similarity between a document and query vector is roughly equivalent to adding up the TF-IDF scores of each word in the document that occurs in the query. (2)

Question 8

Topic: Lecture 6
Source: Lecture 6

In some ways, we could consider Beta distributions themselves to be an embedding of a document. Explain, and explain how we might be able to leverage that. (2)

Question 9

Topic: Coding
Source: Coding

Write a short function that confirms that the sum of n rank-1 matrices is identical to the product of an nxk matrix and a kxn matrix. (3)

END OF QUIZ