START OF QUIZ
Student ID:
49919301,Keigan,Jonathan

Question 1

Topic: Lecture 7
Source: Lecture 7

What is the purpose of an inverted index? (1)

Question 2

Topic: Lecture 5
Source: Lecture 5

Explain the logic behind the IDF part of TF-IDF (ie, why does it give higher weights to more "interesting" words?). (1)

Question 3

Topic: Lecture 6
Source: Lecture 6

In class, we saw a few topics that we were unable to identify. What could be a cause for such pointless topics (ie, how might we ensure that our topics are better? (2 reasons). (1)

Question 4

Topic: Lecture 7
Source: Lecture 7

From a processing perspective, what is one benefit structured data has over unstructured data, and vice versa. (1)

Question 5

Topic: Lecture 8
Source: Lecture 8

What is the intuition behind MAP? (1)

Question 6

Imagine we performed LDA on the classes in this block. What might their Beta distributions look like? (2)

Question 7

Topic: Lecture 5
Source: Lecture 5

We often weight our matrices using something like PMI or TF-IDF. Do you think it would make sense to do this after applying SVD? Why or why not? (2)

Question 8

In class (and in the lab) you saw some examples of using a language model for IR. How do you think we could incorporate an LLM into the IR pipeline? In what ways do you think an n-gram lm might be more appropriate? (2)

Question 9

In class, I mentioned that we might want to provide a weighted ranking of documents in an IR system (for example, we might have a list of relevant documents already sorted for relevancy, and we want our IR system to not only return those documents high in the returned document list, but in the same order). How might we modify the MAP metric to reward the ordering of the documents, as well? Explain (use an example, if you have to). (3)

END OF QUIZ