START OF QUIZ
Student ID:
37105848,Yuen,Shing Hei Robin

Question 1

Topic: Lecture 8
Source: Lecture 8

What do we mean by interpolation? (1)

Question 2

Define P@R. (1)

Question 3

Topic: Lecture 6
Source: Lecture 6

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

Question 4

Topic: Lecture 8
Source: Lecture 8

In class, I mentioned that high k value for BM25 TF weigthing rewards documents with many, many instances of a term in them. Explain why that's the case. (2)

Question 5

Topic: Lecture 7
Source: Lecture 7

What is the benefit of evaluating boolean queries using set operations instead of loops? (1)

Question 6

Topic: Lecture 5
Source: Lecture 5

Why can we be confident that a low-rank approximation of a matrix contains the most important information in a document? (1)

Question 7

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

Question 8

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

Question 9

Topic: Long
Source: Lecture 8

In class, we considered two different types of information retrieval systems - one that uses Boolean terms to find matches, and one that uses a language model to allow for "natural language" queries. Can you think of a way that we might be able to leverage the strengths of both, while minimizing the disadvantages? Briefly explain how that might work. (2)

END OF QUIZ