

START OF QUIZ

Student ID:

49171606,Bai,Desmond

Question 1

Topic: Lecture 8

Source: Lecture 8

What do we mean by interpolation? (1)

Question 2

Topic: Lecture 7

Source: Lecture 7

Define $P @ R$. (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 8

Source: Lecture 8

What are some assumptions that we make when we are interpolating between a document and a corpus? When should we trust the corpus more, and when should we trust the document more? (2)

Question 5

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 6

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 7

Topic: Lecture 5

Source: Lecture 5

Why do we need methods like t-SNE? (1)

Question 8

Topic: Lecture 6

Source: Lecture 6

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 7

Imagine that we have 2 information retrieval systems, and we are evaluating on the same test set, which has 10 relevant documents. The first system returns them in positions [1, 5, 7, 15, 25, 50, 60, 70, 71, 90]. The second returns the documents at positions [2, 3, 6, 8, 10, 62, 80, 83, 91, 95]. Make an argument for each system being better, and provide support for both. Explain which system you would rather use, and why. If there are any other considerations, list them. (3)

END OF QUIZ