START OF QUIZ
Student ID:
38522413,Wang,Yuxi

Question 1

Topic: Lecture 5
Source: Lecture 5

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

Question 2

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 3

Topic: Lecture 8
Source: Lecture 8

What do we mean by interpolation? (1)

Question 4

Topic: Lecture 8
Source: Lecture 8

Why don't we use a higher-order language model to perform IR? (1)

Question 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 6

Topic: Lecture 7
Source: Lecture 7

Define P@R. (1)

Question 7

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 8

Why don't we just use k-means to cluster document-vectors (sparse or dense)? (1)

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