

START OF QUIZ

Student ID:

44918563,Piche,Cole

Question 1

Topic: Lecture 6

Source: Lecture 6

Why can't we just run an HMM over documents to discover the latent states like we do for POS-tagging? (1)

Question 2

Topic: Lecture 7

Source: Lecture 7

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

Question 3

Topic: Lecture 7

Source: Lecture 7

What is the benefit (in terms of efficiency) of placing the most discriminative search terms first in a boolean search? (1)

Question 4

Topic: Lecture 8

Source: Lecture 8

What is the intuition behind MAP? (1)

Question 5

Topic: Lecture 5

Source: Lecture 5

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

Question 6

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 7

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 8

Topic: Lecture 6

Source: Lecture 6

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

Question 9

Topic: Long

Source: Lecture 6

Imagine that we have a Beta distribution for each document, and a Theta distribution for each topic. We are at the Maximization state of EM write a short function that calculates the probability of a document, given these distributions. Pay special attention to edge cases and special considerations... (3)

END OF QUIZ