

**START OF QUIZ**  
**Student ID: 97190441,Ra-**  
**majayam,Jayathilaga**

## Question 1

Topic: Lecture 7

Source: Lecture 7

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

## Question 2

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 3

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 4

Topic: Lecture 8

Source: Lecture 8

What is the reasoning behind substituting TF-IDF with Okapi BM25? (1)

## Question 5

Topic: Lecture 8

Source: Lecture 8

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

## Question 6

Topic: Lecture 5

Source: Lecture 5

Why can we represent a rank- $m$  matrix as the sum of  $m$  rank-1 matrices \*or\* the product of an  $n \times m$  matrix and an  $m \times n$  matrix (ie, what is matrix multiplication doing that we can take advantage of)? Explain. (2)

## Question 7

Topic: Lecture 7

Source: Lecture 7

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



## Question 8

Topic: Lecture 6

Source: Lecture 6

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

## Question 9

Topic: Coding

Source: Coding

Write a function that returns the most likely  $n$  documents given a term-document matrix, a smoothing parameter, and a query. (3)

**END OF QUIZ**