

START OF QUIZ

Student ID:

91877605, Ren, Justin

Question 1

Topic: Lecture 6

Source: Lecture 6

Why do we say that the analogy task is an "intrinsic" evaluation of our word embeddings? (1)

Question 2

Topic: Lecture 7

Source: Lecture 7

Explain salience with respect to entities in a sentence (ie, when identifying Cf). (1)

Question 3

Topic: Lecture 6

Source: Lecture 6

What is the purpose of negative sampling in a Word2Vec model? (1)

Question 4

Topic: Lecture 8

Source: Lecture 8

Describe the recency criterion for anaphor resolution. Why can't we just backtrack from the current word (at least in English)? (2)

Question 5

Topic: Lecture 7

Source: Lecture 7

How is the TextTiling algorithm similar to the Lesk algorithm? How is it different? (2)

Question 6

Topic: Lecture 5

Source: Lecture 5

When we were calculating PMI of a symmetric matrix, why is it not a case of double counting the word in our document? ie., why do the counts of (attorney, fun) and (fun, attorney) not count as two counts each of attorney and fun (such as when we are calculating the total sum of the matrix? (2)

Question 7

Topic: Lecture 5

Source: Lecture 5

What is PMI measuring? That is, what does it mean for two words to have high PMI? (1)

Question 8

Topic: Lecture 8

Source: Lecture 8

What is an anaphor? (1)

Question 9

Topic: Coding

Source: Coding

Write a short function that tries to find a good value for k in truncated SVD. You'll essentially be writing your own version of the evaluate word analogies function. For each of our the analogies, you'll need to do the vector math we were doing in the capital city determination, and return the 1-closest vector - if it's what we're looking for, it's correct. If not, it's wrong. (3)

END OF QUIZ