

START OF QUIZ

Student ID:

37083607,zeng,zejiao

Question 1

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 2

Topic: Lecture 5

Source: Lecture 5

What is the primary assumption of the vector space model for semantics, regardless of how it's implemented? (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 6

Source: Lecture 6

We took a look at how vectors can be added / subtracted in vector space. Why does this work? (hint: think back to the general properties of word embeddings that we've wanted from the very start) (1)

Question 6

Topic: Lecture 7

Source: Lecture 7

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

Question 7

Topic: Lecture 7

Source: Lecture 7

Why do you think that pronouns must be high salience items from previous sentences? (1)

Question 8

Topic: Lecture 8

Source: Lecture 8

Why is it easy to create negative examples for lexical coherence tests? (1)

Question 9

Topic: Long

Source: Lecture 8

Draw the RST Tree of the following paragraph (you can use the parenthetical notation we used in class): Although Henry was a professor of archeology, he didn't really like teaching. He preferred to go on adventures, searching for lost artifacts. Sometimes, this meant he had to fight Nazis, and one time, aliens. He time-traveled once, too. But he saved the world a few times, so the University was ok with it. (3)

END OF QUIZ