START OF QUIZ
Student ID:
71936512,Jimenez,Daniel

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

We've seen co-occurrence matrices weighted by TF-IDF; would it make sense to weight them by PMI? Briefly explain. (1)

Question 3

Topic: Lecture 8
Source: Lecture 8

Do you think we could use word embeddings for coreference resolution? What kind of assumptions would we be making, and why do you think it might still be a very difficult task? (2)

Question 4

Topic: Lecture 7
Source: Lecture 7

We took a look at 2 different ways of implementing the TextTiling algorithm - one with vector overlap, and one with BERT. Can you think of how we might modify the algorithm further to strengthen up its weaknesses? (No is not a valid answer.) (2)

Question 5

Topic: Lecture 8
Source: Lecture 8

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

Question 6

Topic: Lecture 7
Source: Lecture 7

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

Question 7

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 8

Topic: Lecture 6
Source: Lecture 6

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

Question 9

Topic: Long
Source: Lecture 5

All of these embeddings we've been looking at have been an effort to translate meaning into math, so that we can use computational algorithms (which are good at math) to process meaning. To what extent do you think that these are a good approximation for how we understand language, and to what extent do you think they are a poor approximation? (3)

END OF QUIZ