START OF QUIZ
Student ID: 78076577,Strafforello,Francesco

Question 1

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 2

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

Question 3

Topic: Lecture 7
Source: Lecture 7

Explain the underlying assumption of the TextTiling algorithm. (1)

Question 4

Topic: Lecture 7
Source: Lecture 7

The TextTiling algorithm we looked at just looked at raw word overlap (possibly with stop-word removal). Describe a way that we could improve the algorithm to maximize coverage. (2)

Question 5

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 6
Source: Lecture 6

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

Question 7

Which is likely to have the highest PMI? A rare word and a frequent word that appear together frequently, or two frequent words that appear together frequently? (1)

Question 8

Topic: Lecture 5
Source: Lecture 5

Generally speaking, why are we not interested in negative PMI? (1)

Question 9

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