START OF QUIZ
Student ID:
68022458,Chan,Douglas

Question 1

Topic: Lecture 6
Source: Lecture 6

What are two significant shortcoming of the Word2Vec model? (1)

Question 2

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 3

Topic: Lecture 8
Source: Lecture 8

What is the purpose of an antecedent in anaphoric resolution? (1)

Question 4

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 5

Topic: Lecture 6
Source: Lecture 6

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

Question 6

Topic: Lecture 7
Source: Lecture 7

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

Question 7

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 8

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