START OF QUIZ
Student ID:
37105848,Yuen,Shing Hei Robin

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

Topic: Lecture 5
Source: Lecture 5

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 3

Topic: Lecture 6
Source: Lecture 6

When running a window-based approach to vector embeddings (such as CBOW or skip-gram), when would it make sense to keep stopwords, and when would it make sense to remove them? (1)

Question 4

Topic: Lecture 7
Source: Lecture 7

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

Question 5

Topic: Lecture 7
Source: Lecture 7

Why are we interested in backward-facing centers (Cb)? Why not just consider the entities in the current sentence? (1)

Question 6

Topic: Lecture 8
Source: Lecture 8

Describe a Discourse Unit. (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 5
Source: Lecture 5

In class, we talked about how a "typical" dimensionality for embeddings is in the range of 100-500. What might be some consequences if we estimated too low or too high? (2)

Question 9

Topic: Long
Source: Lecture 7

Identify the shifts in the following discourse (show your work): Jonathan Harker was a solicitor from England. He was sent to Transylvania to meet with the mysterious Count Dracula. Dracula wanted to buy property in London. That's where all the wealthiest nobles lived. Dracula had other plans, too, but Harker didn't know that. (3)

END OF QUIZ