START OF QUIZ
Student ID:
90660002,Benipal,Jaskirat

Question 1

Topic: Lecture 7
Source: Lecture 7

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

Question 2

Topic: Lecture 6
Source: Lecture 6

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

Question 3

Topic: Lecture 7
Source: Lecture 7

Explain salience with respect to entities in a sentence (ie, when identifying Cf). (1)

Question 4

Topic: Lecture 8
Source: Lecture 8

What tools are required to build an entity grid? (not structures - matrices, etc. are interesting, but I'm asking what kind of NLP tools are necessary to fill the grid - there are at least 2. (1)

Question 5

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 6

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 7

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 8

Topic: Lecture 6
Source: Lecture 6

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

Question 9

Topic: Coding
Source: Coding

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