START OF QUIZ
Student ID:
66605874,Li,Mingcong

Question 1

Topic: Lecture 8
Source: Lecture 8

What is an anaphor? (1)

Question 2

Topic: Lecture 6
Source: Lecture 6

Think back to week 1 of this block when we were doing word sense disambiguation. Do you think there would be benefits or disadvantages to disambiguating all words before running word2vec? Explain. (2)

Question 3

Topic: Lecture 5
Source: Lecture 5

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 4

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 5

Topic: Lecture 7
Source: Lecture 7

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

Question 6

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 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 8
Source: Lecture 8

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