

START OF QUIZ

Student ID:

74488446,Zhang,Yue Yun

Question 1

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 2

Topic: Lecture 6

Source: Lecture 6

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

Question 3

Topic: Lecture 5

Source: Lecture 5

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

Question 4

Topic: Lecture 8

Source: Lecture 8

What is an anaphor? (1)

Question 5

Topic: Lecture 7

Source: Lecture 7

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

Question 6

Topic: Lecture 8

Source: Lecture 8

Describe a Discourse Unit. (1)

Question 7

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 8

Topic: Lecture 6

Source: Lecture 6

Explain why extrinsic evaluation can be a much more desirable method of evaluating the quality of word vectors than intrinsic evaluation (we didn't have this in the slides, (2)

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