

START OF QUIZ

Student ID:

61305504,Jia,Hao

Question 1

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 2

Topic: Lecture 8

Source: Lecture 8

What is an anaphor? (1)

Question 3

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 4

Topic: Lecture 6

Source: Lecture 6

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

Question 5

Topic: Lecture 5

Source: Lecture 5

What is the primary assumption of the vector space model for semantics, regardless of how it's implemented? (1)

Question 6

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 7

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 to disambiguating all words before running word2vec? Explain. (2)

Question 8

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 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