START OF QUIZ
Student ID:
94325370,Qiao,Grace

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 8
Source: Lecture 8

What is the purpose of an antecedent in anaphoric resolution? (1)

Question 3

We've seen co-occurrence matrices weighted by TF-IDF- would it make sense to weight them by PMI? Briefly explain. (1)

4

Question 4

Topic: Lecture 5
Source: Lecture 5

What is PMI measuring? That is, what does it mean for two words to have high PMI? (1)

Question 5

Topic: Lecture 8
Source: Lecture 8

Describe a Discourse Unit. (1)

Question 6

Topic: Lecture 6
Source: Lecture 6

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

Question 7

Topic: Lecture 7
Source: Lecture 7

The TextTiling algorithm we looked at just looked at raw word overlap (possibly with stopword removal). Describe a way that we could improve the algorithm to maximize coverage. (2)

Question 8

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 9

Topic: Coding
Source: Coding

Imagine we were trying to find a word that is the best prototype of its synonyms. Write a short function that grabs the lemmas of each synset in wordnet, and calculates which lemma is the best prototype (ie, which lemma is the closest to the centroid of the synset) by using the word embeddings. Ignore words that do not have embeddings in gensim. (3)

END OF QUIZ