START OF QUIZ
Student ID:
52196086,Ahluwalia,Max

Question 1

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 2

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

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 7
Source: Lecture 7

Why do you think that pronouns must be high salience items from previous sentences? (1)

Question 5

Topic: Lecture 8
Source: Lecture 8

Why is it easy to create negative examples for lexical coherence tests? (1)

Question 6

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

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

Question 8

Topic: Lecture 8
Source: Lecture 8

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

Question 9

Topic: Long
Source: Lecture 6

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