START OF QUIZ
Student ID:
33836685,LAM,HUI YIN

Question 1

Topic: Lecture 8
Source: Lecture 8

Do you think we could use word embeddings for coreference resolution? What kind of assumptions would we be making, and why do you think it might still be a very difficult task? (2)

Question 2

Topic: Lecture 6
Source: Lecture 6

When running a window-based approach to vector embeddings (such as CBOW or skip-gram), when would it make sense to keep stopwords, and when would it make sense to remove them? (1)

Question 3

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 4

Topic: Lecture 8
Source: Lecture 8

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

Question 5

We took a look at 2 different ways of implementing the TextTiling algorithm - one with vector overlap, and one with BERT. Can you think of how we might modify the algorithm further to strengthen up its weaknesses? (No is not a valid answer.) (2)

Question 6

Topic: Lecture 5
Source: Lecture 5

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

Question 7

Why are we interested in backward-facing centers (Cb)? Why not just consider the entities in the current sentence? (1)

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, but remember that intrisic evaluation is something like the analogy task, that tries to measure the quality of the vectors directly). (2)

Question 9

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