START OF QUIZ
Student ID:
76324177,Chiu,Hayden

Question 1

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 2

Topic: Lecture 7
Source: Lecture 7

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 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 5
Source: Lecture 5

Which is likely to have the lowest PMI? A rare word and a frequent word that appear together frequently, or two frequent words that appear together frequently? (1)

Question 5

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 6

Topic: Lecture 8
Source: Lecture 8

What is an anaphor? (1)

Question 7

Topic: Lecture 7
Source: Lecture 7

Explain salience with respect to entities in a sentence (ie, when identifying Cf). (1)

Question 8

Topic: Lecture 6
Source: Lecture 6

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