

START OF QUIZ
Student ID: 97190441,Ra-
majayam,Jayathilaga

Question 1

Topic: Lecture 6

Source: Lecture 6

What are two significant shortcoming of the Word2Vec model? (1)

Question 2

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 3

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, (2)

Question 4

Topic: Lecture 8

Source: Lecture 8

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

Question 5

Topic: Lecture 7

Source: Lecture 7

Explain the underlying assumption of the TextTiling algorithm. (1)

Question 6

Topic: Lecture 7

Source: Lecture 7

How is the TextTiling algorithm similar to the Lesk algorithm? How is it different? (2)

Question 7

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 8

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 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