START OF QUIZ
Student ID:
31181845,Yuan,Su

Question 1

Topic: Lecture 4
Source: Lecture 4

How is it that EM can arrive at a good solution, even if we have a random initialization of parameters? (1)

Question 2

Topic: Lecture 3
Source: Lecture 3

If we have the sentence “You keep using that word - I do not think it means what you think it means”, what is the probability of the bigram “you think”, assuming that the sentence is the entire corpus? (1)

Question 3

Topic: Lecture 2
Source: Lecture 2

Describe the purpose of linkage in hierarchical clustering. (1)

Question 4

Topic: Lecture 4
Source: Lecture 4

How does semi-supervised learning differ from unsupervised and fully-supervised learning? (1)

Question 5

Topic: Lecture 2
Source: Lecture 2

Describe the intuition behind K-means++ (ie, why do we use it, and what is it trying to accomplish?) (1)

Question 6

Topic: Lecture 3
Source: Lecture 3

Imagine you were trying to pitch a new version of Scrabble to Hasbro that included "digraphs" (ie, combinations of two consecutive letters, like "th"). Do you think that you could score them as a simple combination of the single letter scores (ie, "th" is worth "t" + "h"), or would you need to do some more complex scoring calculations? Explain. (2)

Question 7

Topic: Lecture 1
Source: Lecture 1

What is the primary concern of a semantic vector space (ie, a vector space representing meaning), and how does it relate to our use of cosine similarity to measure word similarity? Can you think of any sorts of words for which it might be very difficult to satisfy this concern? (2)

Question 8

Topic: Lecture 1
Source: Lecture 1

Let's consider a variant of the string alignment problem where instead of aligning characters, we're aligning sequences of characters (maybe we're doing machine translation...). What would need to be modified to handle a situation where we likely have a much higher vocabulary, and there's a lot less copying going on? What assumptions would we be making about the data? Would any of these assumptions make Levensthein distance inappropriate? (2)

Question 9

Topic: Long
Source: Lecture 2

Imagine you are tasked with clustering social media posts to identify trends or topics. You have access to a large amount of unstructured text data. What kind of features do you think would be helpful, how would you preprocess the data, and how would you verify that the clustering is a good one? (3)

END OF QUIZ