START OF QUIZ
Student ID: 91905067,Nong-bualang,Bon

Question 1

Topic: Lecture 4
Source: Lecture 4

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

Question 2

Topic: Lecture 4
Source: Lecture 4

What makes dynamic programming methods, such as the Viterbi algorithm, more efficient for sequence prediction tasks compared to brute-force approaches? (1)

Question 3

Desribe the concept of cluster homogeneity, and how it relates to precision. (1)

Question 4

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 5

When is dynamic programming more efficient than brute force programming? (ie, what assumptions do we make about a problem when we use dynamic programming?) (1)

Question 6

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 7

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 8

Imagine we were using k-means to cluster misspellings around their correct spellings. How many clusters would we need, and what would be a good distance function? Explain. (2)

Question 9

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