START OF QUIZ
Student ID:
38419826,Zeng,Lingsong

Question 1

Topic: Lecture 2
Source: Lecture 2

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

Question 2

Topic: Lecture 4
Source: Lecture 4

What is the main difference between the Viterbi algorithm and the Forward algorithm, and why does it allow us to find the optimal path through a sequence? (1)

Question 3

Topic: Lecture 1
Source: Lecture 1

What intuition about substitutions allows the DP version of Levenstein distance to work as it does? Briefly explain. (1)

Question 4

Topic: Lecture 1
Source: Lecture 1

Suppose we are filling the table for the Levenshtein distance algorithm. We are in cell (x, y). The values of cell (x-1, y-1), (x-1, y), and (x, y-1) are 4, 3, and 3, respectively. What is the value we will put in cell (x, y), given that the letters are NOT equal? (1)

Question 5

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 6

Topic: Lecture 2
Source: Lecture 2

What kinds of data might be difficult to cluster using k-means? Is it a shortcoming of the algorithm, or does it just need very careful feature engineering and distance calculations? (2)

Question 7

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 8

Topic: Lecture 3
Source: Lecture 3

Imagine that we have a trigram model that encounters a trigram where none of the tokens are in the vocabulary. How do you think that might impact our probability calculation for the sentence? How might we go about finding a solution? (2)

Question 9

Topic: Long
Source: Lecture 3

In class, we built a collocation matrix for a bigram language model. Modify the function so that it can handle a trigram language model and implements "add-alpha" smoothing, instead of "add-one" smoothing. (3)

END OF QUIZ