START OF QUIZ
Student ID:
50656347,Zhang,Lisa

Question 1

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 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

Topic: Lecture 1
Source: Lecture 1

Unsupervised Learning typically tries to find structure in unlabeled data. Give two reasons why we might want to still annotate a small dataset to use with our algorithm. (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 2
Source: Lecture 2

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 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 1

Do you think that auto-correct has a bias for where in a word an error occurs (ie, the index of the mistake)? If so, how might you approach fixing this problem? If not, explain why the position doesn't matter. As always, list any assumptions you're making. (3)

END OF QUIZ