START OF QUIZ
Student ID:
29404548,Lee,Ai Ka

Question 1

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 2

Topic: Lecture 4
Source: Lecture 4

Imagine that we are doing ASR instead of POS tagging. Briefly describe what the emissions and transitions would be. (2)

Question 3

Topic: Lecture 3
Source: Lecture 3

Describe the noisy channel model, and how it can be used to represent [Machine Translation, ASR, POS-tagging]. (1)

Question 4

Topic: Lecture 4
Source: Lecture 4

Why are the forward and Viterbi algorithms considered to be dynamic programming, and why do we care? (1)

Question 5

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 5, 3, and 1, respectively. What is the value we will put in cell (x, y), given that the letters are equal? (1)

Question 6

Topic: Lecture 3
Source: Lecture 3

Why do we use log-probability intstead of linear probability? (1)

Question 7

Topic: Lecture 2
Source: Lecture 2

How do we choose the number of clusters for K-means? What are the consequences if we choose incorrectly? (2)

Question 8

Topic: Lecture 1
Source: Lecture 1

What is the main difference between Hamming Distance and Edit Distance? (1)

Question 9

Topic: Long
Source: Lecture 2

Imagine that we are creating a bilingual dictionary, and we want to cluster words that are likely translations of each other (this task is known as "Bilingual Lexicon Induction", or BLI). What kind of features might be good features for this task, and how would we convert them to numerical representations? You can assume that we have a large bilingual corpus that is sentence aligned, but no further information. Do you think we could use K-Means for this task? If not, why not? If so, what kind of special considerations would we need to make, if any?

END OF QUIZ