START OF QUIZ
Student ID:
80041163,Zhao,Ryan

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 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 1, 3, and 5, respectively. What is the value we will put in cell (x, y), given that the letters are NOT equal? (1)

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 3
Source: Lecture 3

If our vocabulary consists of just symbols A and B and our corpus consists of the sequence: A B B A A B and we build a bigram language model by applying add-one smoothing to the maximum likelihood estimate from the corpus, what is the probability P(B|A)? Please show your work. (2)

Question 5

Topic: Lecture 1
Source: Lecture 1

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 2
Source: Lecture 2

When is it more appropriate to use hierarchical clustering than k-means? (1)

Question 7

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 8

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 9

Topic: Long
Source: Lecture 4

Please refer to the ”Long” question from Lecture 4.

END OF QUIZ