

START OF QUIZ

Student ID:

34447581,Ong,Claudia

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

Explain why HMMs are a generative model, and how that differs from a discriminative model. (1)

Question 4

Topic: Lecture 3

Source: Lecture 3

Explain the purpose of Laplace smoothing, and how it accomplishes its goal. (1)

Question 5

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 6

Topic: Lecture 2

Source: Lecture 2

Why do outliers cause problems for clustering algorithms like k-means? How can we deal with them? (1)

Question 7

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 8

Topic: Lecture 1

Source: Lecture 1

Explain what modifications would need to be made to our dynamic edit distance algorithm to incorporate weighted edit distance. (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