

START OF QUIZ

Student ID:

37083607,zeng,zejiao

Question 1

Topic: Lecture 4

Source: Lecture 4

Iterative algorithms often require a stopping condition. Briefly explain why this is necessary, and why perplexity is a metric to use for stopping HMMs. (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 1

Source: Lecture 1

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

Question 4

Topic: Lecture 2

Source: Lecture 2

Discuss the purpose of the linkage criterion in hierarchical clustering (1)

Question 5

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 6

Topic: Lecture 3

Source: Lecture 3

Imagine that we are doing machine translation instead of POS-tagging. What would be the equivalent of emission probabilities and transition probabilities? Explain. (2)

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 3

Source: Lecture 3

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

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