

START OF QUIZ

Student ID:

**89702757,MacFar-
lane,Jarrett**

Question 1

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 2

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 3

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 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 2

Source: Lecture 2

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

Question 6

Topic: Lecture 1

Source: Lecture 1

Explain why edit distance (given our formulation) will always choose a substitution, if it can. (1)

Question 7

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 8

Topic: Lecture 2

Source: Lecture 2

Are both K-means and agglomerative clustering iterative? Explain, and for each that is, explain when the algorithm ends. (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