

**START OF QUIZ**

**Student ID:**

**52196086,Ahluwalia,Max**

## Question 1

Topic: Lecture 4

Source: Lecture 4

Why can we use logarithms for the Viterbi algorithm, but not the forward algorithm? (1)

## Question 2

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 3

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 4

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 5

Topic: Lecture 3

Source: Lecture 3

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

## Question 6

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 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 NOT equal? (1)



## Question 8

Topic: Lecture 1

Source: Lecture 1

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

## Question 9

Topic: Long

Source: Lecture 4

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

**END OF QUIZ**