

**START OF QUIZ**

**Student ID:**

**61305504,Jia,Hao**

## Question 1

Topic: Topic4

Source: Lecture 4

Why are the Forward and Viterbi algorithms considered to be dynamic programming, and why do we care?

## Question 2

Topic: Topic2

Source: Lecture 2

Why is the Forgy initialization sub-optimal?

### Question 3

Topic: Topic1

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 2, 2, and 4, respectively. What is the value we will put in cell  $(x, y)$ , given that the letters are equal?

## Question 4

Topic: Topic1

Source: Lecture 1

When is Manhattan distance more appropriate than Euclidean distance?

## Question 5

Topic: Topic4

Source: Lecture 4

Briefly describe why soft EM might provide more accurate tagging results than hard EM.

## Question 6

Topic: Topic3

Source: Lecture 3

Describe the noisy channel model, and how it can be used to represent POS-Tagging.

## Question 7

Topic: Topic2

Source: Lecture 2

How do we choose the number of clusters for K-means? What are the consequences if we choose poorly?



## Question 8

Topic: Topic3

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.

## Question 9

Topic: Coding

Source: Lecture 2

Imagine we have three clusters  $[[X, Y], [M, N, P], [A, B, C, D]]$ , and a point  $[R]$ . Write a function that determines which cluster to add  $R$  to, given the min linkage criterion.

**END OF QUIZ**