START OF QUIZ
Student ID: 97190441,Ramajayam,Jayathilaga

Question 1

Topic: Topic2
Source: Lecture 2

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

Question 2

Topic: Topic3
Source: Lecture 3

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

Question 3

Topic: Topic2
Source: Lecture 2

Imagine we were using k-means to cluster misspelling around their correct spellings. How many clusters would we need, and what would be a good distance function? Explain.

Question 4

Topic: Topic4
Source: Lecture 4

What is the main purpose of semi-supervised learning in EM? That is, how does it affect the overall model, and where is the effect the largest?

Question 5

Topic: Topic1
Source: Lecture 1

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

Question 6

Topic: Topic4
Source: Lecture 4

Imagine that we are doing ASR instead of POS tagging. Briefly describe what the emissions and transitions would be.

Question 7

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 NOT equal?

Question 8

Topic: Topic3
Source: Lecture 3

In your own words, explain the Markov assumption, and how it is used for language modeling.

Question 9

Topic: Coding
Source: Lecture 3

In class, we built a collocation matrix for a bigram language model. Modify the function so that it can handle trigram language model and implements "add-alpha" smoothing, instead of "add-one" smoothing.

END OF QUIZ