

START OF QUIZ

Student ID:

94779345, Wang, Sherry

Question 1

Topic: Topic3

Source: Lecture 3

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

Question 2

Topic: Topic4

Source: Lecture 4

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

Question 3

Topic: Topic2

Source: Lecture 2

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

Question 4

Topic: Topic4

Source: Lecture 4

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

Question 5

Topic: Topic1

Source: Lecture 1

Discuss why one might do unsupervised learning instead of supervised learning.

Question 6

Topic: Topic3

Source: Lecture 3

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

Question 7

Topic: Topic2

Source: Lecture 2

Why is the Forgy initialization sub-optimal?

Question 8

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 9

Topic: Coding

Source: Lecture 4

Please answer the following question.

Imagine we have the following randomly-initialized probabilities:

```
      1    2    3    4    5
EMIT: 1 [0.2 0.3 0.3 0.2 0.0]
      2 [0.3 0.1 0.1 0.1 0.4]
      3 [0.0 0.5 0.4 0.0 0.1]
```

```
      1    2    3
<s>[0.3  0.4 0.3]
TRAN: 1 [0.0  0.3 0.7]
      2 [0.4  0.1 0.5]
      3 [0.5  0.3 0.2]
```

And imagine we had 3 sentences in our corpus that were tagged in the following way after the E-step

```
1 3 2:
2 3 1
```

```
1 4 5:
3 2 3
```

```
5 4 2:
2 1 3
```

Is the initial probability for state 2 going to be higher or lower after this iteration?

What about the transition probability from state 3 to state 1?

And finally, the emission probability of word 4 given state 2?

Briefly explain.

END OF QUIZ