START OF QUIZ
Student ID:
10237352,Shan,Zhengyi

Question 1

Topic: Lecture 5
Source: Lecture 5

Why is a majority tagger such a strong baseline for POS tagging? (1)

Question 2

Topic: Lecture 6
Source: Lecture 6

Explain how an MSD differs from a POS tag, and how it's similar. (1)

Question 3

Topic: Lecture 7
Source: Lecture 7

Describe gemination in terms of edit actions. (1)

Question 4

Topic: Lecture 6
Source: Lecture 6

Would F1 score be an appropriate measure for gauging the quality of a morphological analyzer? Explain why or why not. (1)

Question 5

Topic: Lecture 5
Source: Lecture 5

Imagine that we wanted to create POS embeddings in the same way that we create word embeddings. Explain one way that it would be very similar to word embeddings, and one that it would be very different. (1)

Question 6

Topic: Lecture 8
Source: Lecture 8

Imagine we were designing a probe to understand whether a model were gender biased. How might we design such a probe, and if we found the model to exhibit such a bias, what suggestions would you make to neutralize the bias? (2)

Question 7

Topic: Lecture 7
Source: Lecture 7

What benefits might encoding MSDs with a second encoder have over a single encoder approach? Can you think of any disadvantages? (2)

Question 8

Topic: Lecture 8
Source: Lecture 8

Imagine we have a good neural morphological analyzer, and we want to inject the knowledge into a larger NLP DL model (like an LLM, etc.). How might we do so in an efficient way? (2)

Question 9

Topic: Long
Source: Lecture 5

In class, we talked about how POS and morphological information is often latently encoded in word embeddings, but not in character embeddings. Let's think about subword embeddings, since most DL models are going to use subword representations. If a word is split, where do you think this information is encoded, and does it matter? Explain your reasoning. (3)

END OF QUIZ