

START OF QUIZ

Student ID:

88179403,Du,Yiyang

Question 1

Topic: Lecture 7

Source: Lecture 7

Describe epenthesis in terms of edit actions. (1)

Question 2

Topic: Lecture 6

Source: Lecture 6

What is the purpose of using normalized edit distance in the evaluation of a morphological analyzer? (1)

Question 3

Topic: Lecture 8

Source: Lecture 8

Nonce words are almost always in open classes of words, but there are some exceptions (such as pronouns, which do allow some flexibility). Do you think that LLMs would react to new words in traditionally closed classes the same way as in open classes? (1)

Question 4

Topic: Lecture 8

Source: Lecture 8

Where do you think pragmatic learning (ie, intent) might fall within the layers of an LLM?
Explain briefly. How might we test for it? (1)

Question 5

Topic: Lecture 6

Source: Lecture 6

Even if we're only interested in lemmas, do you think it's worthwhile to produce MSDs, as well? Why or why not? (1)

Question 6

Topic: Lecture 5

Source: Lecture 5

Imagine that we have some pre-trained multilingual embeddings of really high quality. We train a POS tagger for a very common language, with lots of data, embedding the data with the multilingual embeddings. At inference, we then replace the input with another language. Do you think the tagger would beat a majority baseline? Explain your reasoning, and list any assumptions. (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 5

Source: Lecture 5

In DSCI 563, we discussed EM for POS tagging. Let's make it neural. Assume we have a small set of gold annotated sentences (100). How could we use contextualized embeddings to bootstrap more annotated data (assume that fine-tuning doesn't work)? (2)

Question 9

Topic: Long

Source: Lecture 8

Humans are pretty good at learning the morphology of their language, but they make mistakes. Kids will typically learn irregular morphology before regular morphology, then forget it, and then reacquire it (for example, kids will learn “I went outside”, then transition to “I goed outside”, before eventually coming back to “I went outside”). From your understanding of how DL models work, do you expect something similar to happen within neural language tools? Explain why this might or might not be the case. (3)

END OF QUIZ