

START OF QUIZ

Student ID:

19756840,Wang,Rennan

Question 1

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 2

Topic: Lecture 5

Source: Lecture 5

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

Question 3

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 4

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 5

Topic: Lecture 5

Source: Lecture 5

Why do we attach an embedding layer before passing information to the hidden layer(s)? (1)

Question 6

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 7

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 8

Topic: Lecture 7

Source: Lecture 7

Imagine that we want to improve the quality of a morphological inflector/analyser, and we want to do so through multi-task learning (ie, two decoders predicting different tasks, with a composite loss function). Can you think of a task that might benefit morphological processing? Explain. (2)

Question 9

Topic: Long

Source: Lecture 6

Linguistic determinism is the (somewhat outdated, and originally quite racist) idea that the language we speak shapes the way that we see the world. Although it's been shown to only have a small influence on humans (called weak determinism), there haven't really been many studies that look into it in machines. However, given that our methodologies are so heavily influenced by embedding spaces, we can say that modern ML models are entirely shaped by the language they speak (the language here being the constraints of the embedding space). If we had a well-annotated morphological corpus, how might we set up an experiment to test this theory? (3)

END OF QUIZ