

START OF QUIZ

Student ID:

38595138,Christilaw,Tim

Question 1

Topic: Lecture 5

Source: Lecture 5

Why is POS information so important (whether via tagging or embedded information)? (1)

Question 2

Topic: Lecture 7

Source: Lecture 7

Describe elision in terms of edit actions. (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 7

Source: Lecture 7

Describe epenthesis in terms of edit actions. (1)

Question 5

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 6

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 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 6

Source: Lecture 6

Feature engineering can be incorporated into encoder-decoder models through the use of multiple encoders. If you could have any extra annotation for morphological analysis, and were able to pass each through a separate encoder, what types of features would you include? Do you see any potential problems with using this extra annotation? (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Linguistic determinism is the (somewhat outdated, and originally quite racist) idea that the language we speak shapes the way that we see the world.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