

START OF QUIZ

Student ID:

37157856, Wang, Lusha

Question 1

Topic: Lecture 7

Source: Lecture 7

What is the intuition behind active learning? (1)

Question 2

Topic: Lecture 6

Source: Lecture 6

If you were building your own parallel corpus, what kind of information would you prioritize?
What questions would you ask your stakeholders? (1)

Question 3

Topic: Lecture 5

Source: Lecture 5

Why do MLLMs tend to eventually see a decrease in quality on HRLs? (1)

Question 4

Topic: Lecture 5

Source: Lecture 5

How is the embedding space of MLLMs different from an aligned monolingual space? How is it similar? (1)

Question 5

Topic: Lecture 6

Source: Lecture 6

What step of annotation projection do you think would benefit most from a subword model?
(1)

Question 6

Topic: Lecture 7

Source: Lecture 7

We discussed active learning with respect to classification, but what about regression tasks? What similarities / differences might make active learning suitable or unsuitable to regression? (2)

Question 7

Topic: Lecture 8

Source: Lecture 8

Imagine that we have *no* annotated data for a particular task. How might we address this problem with in-context learning and active learning? (2)

Question 8

Topic: Lecture 8

Source: Lecture 8

Imagine that we are using QbU, but we notice that the quality of our model is not improving with each iteration (or is even decreasing slightly). Where would you look to find out where things are going wrong? (2)

Question 9

Topic: Long

Source: Lecture 5

Imagine you're working on adapting a multilingual LLM for a government that wants it to operate fluently in 10 national languages, including both high- and low-resource languages, and avoid colonial-language bias. Describe a fine-tuning and evaluation pipeline that could help adapt the model fairly across languages. What ethical and linguistic challenges might arise, and how would you mitigate them? How would you include community feedback in the loop? (3)

END OF QUIZ