START OF QUIZ
Student ID:
49919301,Keigan,Jonathan

Question 1

Topic: Lecture 8
Source: Lecture 8

How does silver data differ from synthetic data? (1)

Question 2

Topic: Lecture 6
Source: Lecture 6

What is the intuition behind annotation projection? What assumptions does it make, and how much do you think they matter? (1)

Question 3

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 4

Topic: Lecture 5
Source: Lecture 5

Describe the difference between parallel and pseudo-parallel. (1)

Question 5

Topic: Lecture 7
Source: Lecture 7

Describe situations where margin based selection is more and less appropriate, and why. (1)

Question 6

Topic: Lecture 7
Source: Lecture 7

You've been using AL with multiple annotators. As a sanity check, you have several instances labeled by multiple annotators, but find that the annotations are inconsistent. How can you remedy the problem and select good examples, without knowing the language you are having annotated? (2)

Question 7

Topic: Lecture 5
Source: Lecture 5

You're working with MT5, and you find it's not doing very well on your target language, even after fine-tuning. What do you do? Would your answer change if the model were mBert, instead? (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