

START OF QUIZ

Student ID:

**91905067,Nong-
bualang,Bon**

Question 1

Topic: Lecture 8

Source: Lecture 8

Do you think that we want recommender systems to make recommender systems to make recommendations more in the manner of active learning or self-training? Explain. (1)

Question 2

Topic: Lecture 7

Source: Lecture 7

Why is entropy a good measure to use when using QbC? (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 8

Source: Lecture 8

Explain the role of clustering when performing self-training? (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 6

Source: Lecture 6

L1 interference is a phenomenon whereby L2 language learners make use of properties of their L1 when speaking an L2. Phonetically, this can present as an accent, but it can also impact syntax. Describe this process as if humans were doing projection. (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 5

Source: Lecture 5

Imagine we have a multilingual encoder-model like mBERT, and a multilingual decoder-only model. Do you think we could train the encoder on one set of languages, and then the decoder on a larger set, and better understand the new languages? What kind of adaptations would need to be done? Do you think it would improve zero-shot learning on languages not included in either? (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