# START OF QUIZ
# Student ID:
# 10237352,Shan,Zhengyi

## Question 1

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 2

Topic: Lecture 8
Source: Lecture 8

How does ICL differ from fine-tuning? (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

What are the key differences between BERT and BART, and what concept from DSCI 563 does BART imitate? (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 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 8

Topic: Lecture 6
Source: Lecture 6

In the lab, you likely saw that a certain part took much longer than others, and produced noisy output. How might you adjust your strategy, speed up the methodology, and what simple tools could you use to decrease noise? (2)

## Question 9

Topic: Long
Source: Lecture 6

Imagine that instead of projecting tags, we project embeddings, instead (that is, we attach a high-resource embedding to a low resource word). What advantages might this have over tag projection, and what difficulties might we still encounter? Would it introduce new difficulties? (3)

# END OF QUIZ