START OF QUIZ
Student ID:
26566455,Lai,Minsi

Question 1

Topic: Lecture 3
Source: Lecture 3

Imagine that we have a parallel corpus (ie, a corpus containing sentences in two languages), and we want to extract a bilingual lexicon. What are some simple steps we could do to identify words that could be translations of each other? (2)

Question 2

Topic: Lecture 1
Source: Lecture 1

In class, we talked about how .isdigit() is insufficient for determining whether we can convert a string to a float. Write a short function "isfloat" that determines whether a provided string is a valid floating point number. (2)

Question 3

Topic: Lecture 3
Source: Lecture 3

What is the Big O time complexity of finding the elements in a set that intersect with an iterable (ie, string, list, etc)? Briefly explain. (1)

Question 4

Topic: Lecture 4
Source: Lecture 4

How would we sort a dictionary alphabetically by the reverse of its keys (assuming the keys are strings)? Write a short piece of code, and briefly explain your logic. (1)

Question 5

Topic: Lecture 2
Source: Lecture 2

Do you think it's possible for a language not to follow a Zipfian curve? What consequences might that have on communication (if, let's say, if the curve were linear)? (2)

Question 6

Topic: Lecture 4
Source: Lecture 4

In class, we removed stopwords by using a lexicon. Can you think of another way that we could remove all closed class words? (1)

Question 7

Topic: Lecture 1
Source: Lecture 1

Why is strip() such a useful function? (1)

Question 8

Topic: Lecture 2
Source: Lecture 2

As we expand the size of our corpus, we increase the number of Hapaxes. Do you think this is also true of stopwords? Briefly explain. (1)

Question 9

Topic: Coding
Source: Lecture 2

Imagine we have a large corpus in an unknown language. We don't have any ML tools to analyze the data. How might we determine the stopwords in our corpus? How might we test our theory of stopwords? (I'll make it easy on you - the tokens are space separated, and we have some way of separating sentences.) (3)

END OF QUIZ