START OF QUIZ
Student ID:
49171606,Bai,Desmond

Question 1

Topic: Lecture 3
Source: Lecture 3

How does ”get” differ from a default dictionary (2 ways)? (1)

Question 2

Topic: Lecture 4
Source: Lecture 4

Why does type-to-token ratio decrease as the size of the corpus increases? What does this suggest about long documents? (1)

Question 3

Topic: Lecture 2
Source: Lecture 2

How does Zipf's law relate to Hapax Legomena? (1)

Question 4

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 5

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 6

Topic: Lecture 1
Source: Lecture 1

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

Question 7

Topic: Lecture 3
Source: Lecture 3

Although lexicons are often good starting points, they are often less capable than ML methods. What are some reasons (at least 2) that lexicons are insufficient for state-of-the-art training. Briefly explain. (2)

Question 8

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 9

Topic: Coding
Source: Coding

Imagine that we have an encrypted data set in a language we don't know, but it is written in the Latin script (ie, the script of English, French, etc.). What are some tests that we could run to try to determine the original language? Please list any assumptions you make. Assume that machine learning is not an option. (3)

END OF QUIZ