START OF QUIZ
Student ID:
34157719,Philip,Reshmi

Question 1

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 2

Topic: Lecture 1
Source: Lecture 1

Vowels are often used as a proxy for syllables in words (it's not a perfect correspondence, but it's not bad). Write a function that counts the vowels in a word, without using a loop, using only the tools we went over in Lecture 1 (list comprehension counts as a loop). (2)

Question 3

Topic: Lecture 4
Source: Lecture 4

We discussed two alternative methods for noise reduction: removing all words above a certain frequency, or only removing those from a curated lexicon. Name an advantage to both. (1)

Question 4

Topic: Lecture 1
Source: Lecture 1

How would you reverse a string and keep the result? (1)

Question 5

Topic: Lecture 2
Source: Lecture 2

If we have a new corpus, how might we automatically determine (without ML): A. The language it's written in. B. Whether it is annotated C. If it is multilingual D. genre? Briefly explain your reasoning. (2)

Question 6

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 7

Topic: Lecture 3
Source: Lecture 3

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

Question 8

Topic: Lecture 3
Source: Lecture 3

Why do we not care about the extra space required to create a reverse index? (2 reasons) (1)

Question 9

Topic: Coding
Source: Coding

Write a function that determines whether the last consonant of the root of a verb has been doubled. For example: win -> winning. The function can take any string. What kind of error testing should you perform? (Hint - consonant doubling only occurs in certain environments). (3)

END OF QUIZ