

START OF QUIZ

Student ID:

76324177,Chiu,Hayden

Question 1

Topic: Lecture 1

Source: Lecture 1

How would you quickly identify the words in a sentence? (1)

Question 2

Topic: Lecture 2

Source: Lecture 2

Why is it important to know when a corpus was constructed, and who constructed it? (1)

Question 3

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 4

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 5

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 6

Topic: Lecture 4

Source: Lecture 4

Do you think that children's (age 3-5) picture books would have a higher or lower ratio of adjectives than university literature? Briefly explain your logic. (1)

Question 7

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 8

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 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