START OF QUIZ
Student ID:
19668508,Li,Julian

Question 1

Topic: Lecture 1
Source: Lecture 1

What method would you use to check if a string contains only numeric digits (including decimals) without using any additional libraries? (1)

Question 2

Topic: Lecture 2
Source: Lecture 2

Why is it important to understand the intended audience and time period of a corpus when conducting linguistic analysis? (1)

Question 3

Topic: Lecture 4
Source: Lecture 4

What are two potential drawbacks of removing stopwords from a text before conducting a sentiment analysis? (1)

Question 4

Topic: Lecture 4
Source: Lecture 4

Why does the lexical diversity (type-to-token ratio) typically increase when analyzing smaller sub-corpora rather than larger ones? What does this suggest about the content of smaller texts? (1)

Question 5

Topic: Lecture 2
Source: Lecture 2

As we increase the size of a corpus, the frequency of Hapax Legomena generally increases. Would the frequency of function words like ”the” or ”is” also increase? Why or why not? (1)

Question 6

Topic: Lecture 3
Source: Lecture 3

Lexicons are useful for initial text analysis but often lack the adaptability needed for advanced NLP tasks. Why is this the case? Provide at least 2 reasons with brief explanations. (2)

Question 7

Topic: Lecture 1
Source: Lecture 1

You are given a sentence. Write a function to count how many words in the sentence start with a vowel, without using loops or list comprehensions. (2)

Question 8

Topic: Lecture 3
Source: Lecture 3

Imagine you have a large text corpus in English and Spanish and want to automatically align sentences for machine translation. What are some straightforward methods you could use to identify sentence pairs that are likely translations of each other? (2)

Question 9

Topic: Long
Source: Lecture 1

Write a function that validates if a string matches a phone number format, such as (123) 456-7890. What types of invalid inputs should the function check for? Are there edge cases we would be willing to accept? How would we handle those? Write 3 test cases - 2 that should pass, and one that should fail. (3)

END OF QUIZ