START OF QUIZ
Student ID:
38595138,Christilaw,Tim

Question 1

Topic: Lecture 1
Source: Lecture 1

When would you choose to preserve the original case of text during data processing, rather than converting everything to lowercase? (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 3
Source: Lecture 3

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

Question 4

Topic: Lecture 2
Source: Lecture 2

What role does linguistic annotation provide for corpora, specifically for computational linguistics? (1)

Question 5

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 6

Topic: Lecture 4
Source: Lecture 4

Attributive adverbs are a type of adverb that provides ”flavour” to speech verbs (example: ”she said quickly”; ”he spoke loudly”). They are often frowned upon in formal writing, because they can be replaced with other verbs: ”blurted” or ”shouted”, in the example. Write a quick function that finds them in the Brown corpus, and reports how many sentences in 1000 have them. (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

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 9

Topic: Long
Source: Lecture 2

Imagine you are working with a corpus in a language you don't know, and you need to identify the stopwords in it. You cannot use machine learning but can perform basic statistical analysis. How would you approach identifying stopwords? What metrics would help you confirm that you've identified them correctly? (3)

END OF QUIZ