

START OF QUIZ
Student ID: 31003346,De
Jager,Rachelle

Question 1

Topic: Lecture 4

Source: Lecture 4

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

Question 2

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 3

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 4

Topic: Lecture 1

Source: Lecture 1

Why is the `.split()` method useful when working with sentences or phrases? (1)

Question 5

Topic: Lecture 2

Source: Lecture 2

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

Question 6

Topic: Lecture 2

Source: Lecture 2

Is it possible for a corpus of a low-resource language to follow Zipf's law? What factors might influence the degree to which the law applies in such languages? (2)

Question 7

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 8

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 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