

START OF QUIZ

Student ID:

90660002,Benipal,Jaskirat

Question 1

Topic: Lecture 5

Source: Lecture 5

In class, we said that “fake” fake reviews are often too prototypical when they are generated by hand. Given the tools you’re familiar with, how do you think we could generate fake reviews automatically? Do you think they would suffer from the same problem? (2)

Question 2

Topic: Lecture 7

Source: Lecture 7

Imagine that we had a strange representation of the date: “Year 23 in the 21st century on the 3rd day of March, at 11 minutes past 17”. Using `strptime`, what is the format that we would need to provide to recognize this time? (1)

Question 3

Topic: Lecture 5

Source: Lecture 5

When is ordinal classification more suitable for sentiment analysis than binary classification (2 factors)? (1)

Question 4

Topic: Lecture 8

Source: Lecture 8

Suggest one way that normalization of non-standard social data can help sentiment analysis, and one that can hurt it. (1)

Question 5

Topic: Lecture 7

Source: Lecture 7

Can you think of any disadvantages to representing data in a choropleth? When might it be more advantageous to use a different visualization method? (2)

Question 6

Topic: Lecture 6

Source: Lecture 6

Which of the following Tweets is most likely to be sarcastic? Give a brief explanation of why. A. That sounds like a really great idea! #Awesome! B. That sounds like a reeeeeeeally great idea! C. That sounds like a really great idea! (_) D. That sounds like a really great idea! :+1: (2)

Question 7

Topic: Lecture 8

Source: Lecture 8

What is code-switching, and why is it a problem for NLP? (1)

Question 8

Topic: Lecture 6

Source: Lecture 6

In class, we looked at 2 different ways of identifying personality traits - a self-applied questionnaire, and a data-driven prediction model. Give a brief description of which setup you think would be more reliable, and why. Are there any conditions that might change your answer? (1)

Question 9

Topic: Coding

Source: Lecture 6

Imagine a detective approaches you as a data analyst and says that they have been receiving letters purporting to be from a serial killer. The detective is worried that some of the letters might be copycats. What are some tests (at least 3) that you can run to try to determine if the letters were written by the same person? (3)

END OF QUIZ