START OF QUIZ
Student ID:
94779345,Wang,Sherry

Question 1

Topic: Lecture 7
Source: Lecture 7

Can you think of any biases that exist in the datetime library? If you were redesigning the library, what added functionality might you add? (2)

Question 2

Topic: Lecture 6
Source: Lecture 6

What is distant supervision, and why can we apply it to social media? (1)

Question 3

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 4

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 5

Topic: Lecture 8
Source: Lecture 8

What properties of code-switched text are useful for identifying the language of the text? (List at least 2) (1)

Question 6

Topic: Lecture 5
Source: Lecture 5

Why would a tweet history help identify sarcasm in a new tweet? (1)

Question 7

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 8

What is one similarity and one dissimilarity between emojis and emoticons? (1)

Question 9

Topic: Coding
Source: Lecture 8

Imagine you were tasked with building a Sentiment Analyzer for Reddit posts. Reddit is not quite as irregular as Twitter, but it uses a mixture of standard language and internet phenomena. If there were no existing tools for processing Reddit data, how might you go about creating a successful analyzer? Think of the tools you would have to build, and any assumptions you might have to make about them. (3)

END OF QUIZ