START OF QUIZ
Student ID:
47323894,Liao,Jingyi

Question 1

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 2

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 3

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 4

Topic: Lecture 7
Source: Lecture 7

What information about a user/document is required in order to include it in a choropleth (2 items)? (1)

Question 5

Topic: Lecture 7
Source: Lecture 7

How might you modify a standard sentiment analyzer to track change in sentiment over time? (2)

Question 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 reeeeeeally great idea!
C. That sounds like a really great idea! (_)
D. That sounds like a really great idea! :+1: (2)

Question 7

Topic: Lecture 6
Source: Lecture 6

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

Question 8

Topic: Lecture 5
Source: Lecture 5

What is argumentation mining? How is it related to IR? (1)

Question 9

A: Feature vector = (2,1), rating = 2
B: Feature vector = (2,-1), rating = 3
C: Feature vector = (-1,-1), rating = 5
If we are doing SVM-based ranking, give at least one feature vector that can be used as a positive example for our binary SVM classifier, and one feature vector that can be used as a negative example.
Secondly, if the weight vector of our trained SVM classifier is (2,-2), what is Kendall's Tau for the resulting ordinal classification of these 3 documents? (3)

END OF QUIZ