START OF QUIZ
Student ID: 97132088,Balakrishnan,Aswin Chidambaram

Question 1

Topic: Lecture 8
Source: Lecture 8

In class, we discussed that internet speech may be emerging as its own language (or at least, as a dialect). What features of an emerging language does it demonstrate? Does it lack anything to make you consider it a language? Finally, do you think that separate social media sites could be considered different dialects? Briefly explain. (2)

Question 2

Topic: Lecture 5
Source: Lecture 5

Describe metadata. Why is it useful? (1)

Question 3

Topic: Lecture 6
Source: Lecture 6

We saw that age and gender are relatively easy to predict from tweet history, but that personality traits are a lot harder. Why do you think that is? (1)

Question 4

Topic: Lecture 7
Source: Lecture 7

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

Question 5

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

Question 6

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 7

Topic: Lecture 8
Source: Lecture 8

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

Question 8

Topic: Lecture 5
Source: Lecture 5

SVM ranking takes advantage of the fact that an ordinal problem can be transformed into a binary "larger than" problem by simple subtraction of feature vectors. It's typically done with a linear SVM. Do you think we could apply a similar trick with a neural model? Why or why not? (2)

Question 9

Topic: Coding
Source: Lecture 7

Imagine that you have gotten a long string of unstructured data, and you need to recognize and normalize dates and times in either YY(YY)/MM/DD, MM/DD/YY(YY), or [X]Day, Month Name Day of Month, Year format (ie. Monday, March 20, 2023 - this was actually part of a Capstone a few years ago). Write pseudocode to first identify a date in running text, and then to normalize it. (3)

END OF QUIZ