

**START OF QUIZ**

**Student ID:**

**22039382,Amal,Cenith**

## Question 1

Topic: Lecture 6

Source: Lecture 6

Based on the Swartz et al (2013) study of personality on social media, give an example of how emotion classification intersects with the identification of personality traits. (1)

## Question 2

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 3

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 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

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 6

Topic: Lecture 6

Source: Lecture 6

How does modeling author personality help in the detection of sentiment (think about how it might help us determine sarcasm or interpret reviews). (2)

## Question 7

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 8

Topic: Lecture 5

Source: Lecture 5

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

## Question 9

Topic: Long

Source: Lecture 5

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**