

**START OF QUIZ**

**Student ID:**

**78076577, Stra-  
forello, Francesco**

## Question 1

Topic: Lecture 5

Source: Lecture 5

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

## Question 2

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 3

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 4

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 5

Topic: Lecture 7

Source: Lecture 7

Why is datetime functionality necessary? That is, why can't we just use the date and time separately? (1)

## Question 6

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 7

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 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 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**