START OF QUIZ
Student ID:
37469715,Sharma,Prakul

Question 1

Topic: Lecture 4
Source: Lecture 4

What are the assumptions we are making when we are implementing when we are creating a multi-task learner? Why wouldn't we just use the extra labels as input features to a single task learner? Wouldn't that be simpler? (2)

Question 2

Topic: Lecture 1
Source: Lecture 1

Describe the Pollyanna principle, and how it complicates sentiment analysis. (1)

Question 3

Topic: Lecture 3
Source: Lecture 3

Identify the aspect of the following sentence. Paul Giamatti's performance was a highlight of The Holdovers. (1)

Question 4

Topic: Lecture 2
Source: Lecture 2

One of the goals of embeddings is that similar words are close to each other, and unrelated words are far apart. If we are using embeddings in our sentiment analysis toolkit, explain why we can't just "flip the polarity" of words modified by a negator (ie, [0.1, 0.3, 0.5] -> [-0.1, -0.3, -0.5]) (2)

Question 5

Topic: Lecture 1
Source: Lecture 1

Why is sentiment so tightly bound with domain? (1)

Question 6

Topic: Lecture 4
Source: Lecture 4

What benefit does a CNN have over a standard RNN that makes it particularly suited to sentiment analysis? (1)

Question 7

Topic: Lecture 2
Source: Lecture 2

In class, we talked about how repeated use of words is not cumulative (ie, using good 5 times is not 5 times as positive as using it once). Briefly explain why this is the case. (1)

Question 8

Topic: Lecture 3
Source: Lecture 3

We mentioned in class that "but clauses" are intensifiers. Do you think all (or at least most) concessions work the same way (some other concession words are "although", "nevertheless", "nonetheless", "even though", "considering that")? Briefly explain why or why not. (2)

Question 9

Topic: Long
Source: Lecture 3

In class, we've discussed ways of obtaining and expanding polarity lexicons, but we didn't talk about how to identify PPIs / NPIs. Write out pseudocode (ie, codish-looking stuff) that dives through a corpus of sentiment annotated documents across multiple domains, and identifies "potential potential items". If you make any assumptions about the data, be sure to list them. Don't actually write the code - this should be a designed algorithm, not a runnable piece of code. (3)

END OF QUIZ