## Week 05 Assessment

Due Oct 7 at 9amPoints 9Questions 9Available until Oct 28 at 9amTime Limit NoneAllowed Attempts 2

## **Instructions**

Reminder: In all weekly assessments, you will be allowed two attempts (though only one is required), and your average score will be kept. You're welcome to watch the lecture videos between or even during your attempts.

## **Attempt History**

|        | Attempt   | Time       | Score      |  |
|--------|-----------|------------|------------|--|
| KEPT   | Attempt 2 | 1 minute   | 8 out of 9 |  |
| LATEST | Attempt 2 | 1 minute   | 8 out of 9 |  |
|        | Attempt 1 | 25 minutes | 8 out of 9 |  |
|        |           |            |            |  |

Score for this attempt: **8** out of 9 Submitted Oct 6 at 7:43pm This attempt took 1 minute.

| Question 1   | 1 / 1 pts |
|--|-----------|
| Property testing algorithms for a specific property:                           |           |
| Pass inputs that have the property   |           |
| Fail inputs that are not even close (in some predefined sense) to hav property | ing the   |
| May pass or fail inputs that do not have the property, but are close to        | having it |
| Run in sub-linear time   |           |

| 10/6/2019 | 1 | Week 05 Assessment: Introduction to the Challenges and Opportunities of Big Data, the Internet of Things, and Cybersec | urity |
|-----------|---|--|-------|
| Correct!  | I | All of the above   |       |

|          | Question 2                                | 1 / 1 pts |
|----------|---|-----------|
|          | Property testing algorithms:              |           |
|          | Are useful when an exact answer is needed |           |
| Correct! | Are useful when when time is crucial      |           |
|          | Are only useful on graphs                 |           |

|          | Question 3                                     | 1 / 1 pts |
|----------|--|-----------|
| Correct! | Name one advantage of streaming over sampling: |           |
|          | O Low space usage                              |           |
|          | No data element is missed                      |           |
|          | O Low running time                             |           |

| Question 4  | 1 / 1 pts  |
|---|------------|
| Which of the following situations are not amenable to coreset com | ıpression: |
| GPS travel data   |            |
| Hospital patient logs   |            |

Using topic models in multi-aspect summarization increases prediction accuracy because:

It helps to predict the number of aspects

It helps to disambiguate word usage in the context of the corresponding aspects

It makes summaries more fluent

It identifies common words that can be excluded from an output summary

|             | Question 6  | 1 pts |
|-------------|---|-------|
|             | The method for multi-aspect summarization utilizes unlabeled data to le         | earn: |
|             | The likelihood of transition between topics                                     |       |
|             | The likelihood of sentence label given its topic and the words which it contain | าร    |
| ou Answered | The likelihood of a sentence given its topic                                    |       |
|             |   |       |

| orrect Answer | Week 05 Assessment: Introduction to the Challenges and Opportunities of Big Data, the Internet of Things, a  The likelihood of a sentence given its topic and the likelihood of transition between topics | nd Cybersec |
|---------------|---|-------------|
|               | Question 7 1/1  | pts         |
|               | Which of the following is most likely to provide useful information about changes in the state of a hospitalized patient?   |             |
|               | Billing records   |             |
| Correct!      | Pharmacy records  |             |
|               | Demographic information about the patient.  |             |
|               | Question 8 1/1  | pts         |
|               | Consider building a feature vector by binarizing the data in an EMR. The resulting feature vector is likely to be:  |             |
|               | Poorly correlated with the patient's health status.   |             |
| Correct!      | Very sparse.  |             |
|               | Cow dimensional.  |             |
|               | Question 9 1/1  | nte         |

What do linear-time algorithms allow us to do that most sublinear-time

https://canvas.harvard.edu/courses/63118/quizzes/153628

algorithms do not?

| 10/6/2019 | Week 05 Assessment: Introduction to the Challenges and Opportunities of Big Data, the Internet of Things, and Cybersecu |  |
|-----------|---|--|
| Correct!  | Execute the computation faster  |  |
|           | Approximate the desired answer  |  |
|           | Determine the exact answer  |  |
|           | Code the algorithm faster   |  |

Quiz Score: 8 out of 9