Lesson 05

CSC357 Advanced Topics—Machine Learning

20 January 2020

1. The goal of the exercise in Chapter 2 was prediction. What is the goal of the exercise in Chapter 3?

2. Will we need a URL to fetch the MNIST data set? Explain.

3. What does the MNIST data set contain? What are its dimensions?

4. Some new mathematical terms appear in Chapter 3.

    - What is a gradient?
    - What does stochastic mean?

5. For what purpose might we use scikit-learn's GridSearchCV class?

6. Show how to calculate the precision statistic. What is the signficance of the denominator?

7. Show how to calculate the recall statistic. What is the significance of the denominator?

8. What is a confusion matrix? What is the significance of the rows? What is the significance of the columns?

9. Distinguish between the OvO and OvR strategies.

10. What is the ROC curve?

11. Identify some of the algorithms or classes that Aurélien Géron suggests for the solution of the problem that he outlines for us in Chapter 3.

12. The word 'trade-off' appears several times in Chapter 3. In what context?