CSc 59929
Introduction to Machine Learning
Assignments for Week 11
Due November 14 by 10:59 AM

- Compete the programming exercise described in the following charts.

Programming Exercise

1. Download the housing dataset referenced in the textbook on page 279, https://archive.ics.uci.edu/ml/datasets/Housing

2. Pick a number of features that you think may be correlated and plot pairs of them to confirm that they are correlated.

3. Pick one feature that you think can be predicted by the other features in the dataset. The feature to be predicted needs to have numerical values.

The City College of New York
CSc 59929 – Introduction to Machine Learning
Fall 2017 – Erik K. Grimmelmann, Ph.D.

2

Programming Exercise

4. Separate the data into training, validation, and test sets.

5. Apply various **Scikit Learn regression methods** to the data.

6. Quantify how good a job each method did in predicting the value of the dependent variable (using your test data set). Discuss why you believe each method performed as it did.

7. Repeat steps 3. to 6. for a second feature.

The City College of New York
CSc 59929 – Introduction to Machine Learning
Fall 2017 – Erik K. Grimmelmann, Ph.D.

3

Programming Exercise

8. Upload your .ipynb file to Blackboard. Include your discussions (Step 6.) in either a Word or pdf file or (preferably) in a cell (or cells) in your notebook with your code.

The City College of New York
CSc 59929 – Introduction to Machine Learning
Fall 2017 – Erik K. Grimmelmann, Ph.D.

4