Assignment 8

Learning Objectives:

1. Practice tuning and troubleshooting skills

Description:
The data set you will be working with for this assignment is again the same data set you used for the midterm, but now it has been divided into 5 train/test folds using the stratified remove folds filer. The original dataset, which you used for Assignment 7, is also included. You might remember we discussed it in an earlier lecture.

Step-by-Step Guide:

1. Complete the week 10 and 11 assigned readings about linear and statistical models, and review the lecture slides related to doing optimization in the Experimenter.

2. Use the experimenter to find the optimal setting for the Exponent parameter of the polynomial kernel on the SMO algorithm. Try 1.0, 2.0, and 3.0 as potential settings for that parameter. Evaluate the results in terms of Kappa. Fill in the following table (except for the Test set performance) from the results you see in the Experimenter:

Fold	Train Set Performance For Exponent = 1.0	Train Set Performance For Exponent = 2.0	Train Set Performance For Exponent = 3.0	Optimal Setting	Test Set Performance
1					
2					
3					
4					
5					

Also take a screen shot of the Experimenter when you do the analysis so that we can verify that you are able to interpret the output correctly.

3. Now, using the Explorer and the optimal setting you observed on each fold, compute the test set performance by training a model using the optimal setting over the training data on that fold and then applying that model to the test data for that fold. Fill in these values under Test Set Performance.

4. Now using the Explorer and the whole set of data rather than the separate folds, determine what setting you would use to build a model over the whole set. What is that setting, and how would you estimate what performance you would get for that model over a new set of data?

5. Looking at the comparison across settings and thinking about what the different settings mean in terms of what the algorithm is doing with the data when it builds its model, what conclusions can you draw?

Deliverables: Turn in the table you filled in, the screen shot, and the write up for 3 – 5.

Submission Mode:

Submit the assignment to blackboard as usual.