DATS 7860 Statistical and Machine Learning for Big Data

Homework 7 (Chapter 9. Support Vector Machine)

There are 3 problems and they are all from Exercises 9.7 of Chapter 9 of the textbook (ISL):

1. Problem 5 (Page 369).

Some specific notes for this problem:

- Part (i): please summarize and compare the performance of different methods (logistic regression with the two original predictors, logistic regression with non-linear functions of the two predictors, support vector classifier with the two original predictors, and SVM with a non-linear kernel) in your comments.

- Please submit your complete R codes for this problem.

2. Problem 7 (Page 371).

Some specific notes for this problem:

- Part (b): please try at least 5 different values of the parameter cost across a broad range (e.g. 0.01, 0.1, 1, 10, 100).

- Part (c): for the polynomial basis kernel, please try at least 5 different values of the parameter cost and 4 different values of the parameter degree; for the radial basis kernel, please try at least 5 different values of the parameter cost and 4 different values of the parameter gamma.

- Part (d): please also briefly summarize and the compare the performance of different methods (support vector classifier, SVMs with polynomial and radial basis kernels).

- Please submit your complete R codes for this problem.

3. Problem 8 (Page 371).

- Please submit your complete R codes for this problem.