

**Biostatistics 546, Spring 2017**  
**Machine Learning for Biomedical Big Data**  
**Homework 2**

**Instructions:**

*Answer the following questions, in full sentences. Solutions should be word-processed. For problems that include coding, pasting output from your R session is not acceptable, unless otherwise indicated; your goal should be to perform statistical learning and the R output is relevant only if it is incorporated as part of the analysis. Append your R code separately, with comments for your future reference. Please upload your solutions as a .pdf file.*

NOTE: While you can (and are encouraged to) work together, your solution to the homework, including the code and the writeup, should be *your own work*.

The first three problems below are from Introduction to Statistical Learning, by James et al (2012).

1. Chapter 6, Problem 1
2. Chapter 6, Problem 3
3. Chapter 6, Problem 8; set your random seed to 1 for part (a) and to 2 for part (b).
4. Chapter 6, Problem 9
5. (*Optional*) Repeat the previous problem on the `prostate` data from `ElemStatLearn` package.