ANLY-590 Assignment 1

August 2019

1

Regularization. Using the accompanying *Hitters* dataset, we will explore regression models to predict a player's Salary from other variables. You must python.

https://gist.github.com/keeganhines/59974f1ebef97bbaa44fb19143f90bad

1.1

Use LASSO regression to predict Salary from the other *numeric* predictors. Create a visualization of the coefficient trajectories. Comment on which are the final three predictors that remain in the model. Use cross-validation to find the optimal value of the regularization penality. How many predictors are left in that model?

1.2

Repeat with Ridge Regression. Visualize coeffecient trajectories. Use cross-validation to find the optimal vaule of the regularization penalty.

2

Short Answer. Explain in your own words the bias-variance tradeoff. What role does regularization play in this tradeoff? Make reference to your findings in number (1) to describe models of high/low bias and variance.