

CAIS 380: Intro Machine Learning

Spring 2024

In-class Activity 05: Linear Model Regularization Methods

Activity is DUE as indicated on the course schedule.

This is a group activity! Work with 1-2 other people on it and submit as a group.

Goals:

- Code:
 - Subset selection
 - Shrinkage methods
 - PCR and PLS

Part 1

Start at chapter 6.5 in your textbook. Work through section 6.5.3 in the lab.

All of your work should be recorded in a Markdown or Jupyter file that you collaborate on with your partner(s) [it's probably a good idea to set up a GitHub repo to share code!]. All group members should work together and participate on every section of the lab; do not break it up and each individual work on a section.

It is okay if you do not finish the entire lab. You will be graded on showing an honest effort and working collaboratively with your group.

Part 2

Find a large dataset that is of interest to you. Perform EDA on the dataset to understand the variables in it. Use your exploration to come up with 3 research questions you might ask and answer with your dataset.

Submission

At the end of class, submit the PDF report of your work (showing code and output) on Gradescope. Be sure to tag your group members.