

Economics 187: Project 2
Spring 2022, UCLA
Instructor: Dr. Rojas

Due Date: Apr 19, 2022 by 5:00PM (PST)

For this project you will need to use two different datasets of your choice. One will be used for classification and the other for regularization. For the classification dataset make sure you have more than two classes. For your regularization dataset, since the methods focus on variable selection, please make sure to have as many predictors as possible (e.g., 10s or 100s).

1. (50%) Classification: Fit the following models to your data:

- Logistic Regression
- LDA
- QDA
- kNN
- k-Means

Based on your fits, identify whether a linear or non-linear model is more appropriate. Make sure to discuss your results (including plots and tables), and to use CV and/or bootstrap to evaluate your models' performance.

2. (50%) Regularization: Fit the following models to your data:

- LASSO
- Ridge
- Elastic Net
- PCA

Based on your fits, identify the best model taking into consideration the bias-variance tradeoff. Make sure to discuss your results (including plots and tables), and to use CV and/or bootstrap to evaluate your models' performance.