Advanced AI/ML for Finance

**Spring 2020  
Professor Brian Clark**

**Homework #1 – Elastic Net Regressions.**

**Due at the start of class on 2/8/2021.**

**Overview**: Build Ridge and Lasso regression models to **predict** S&P 500 returns using a minimum of 10 features (independent variables) of your choice. The detailed requirements are as follows:

1. Download all necessary data. The only restriction is that the data must be public so I can replicate your results and it must be downloaded directly in your code. You are encouraged to use the `quantmod` package.
2. **Define your outcome variable as the next period return of the S&P 500**. You may use the SPY ETF as a proxy. You are free to choose the return period (days, weeks, months).
3. Your models must include at least 10 features of your choice. A good place to get relevant variables is the St. Louis Fed via the `quantmod` package. You can also engineer variables out of historical return data (e.g., lagged variables, moving averages, volume, etc.).
4. Build two models using R’s `glmnet` package (make sure you calibrate the models by selecting the  values that minimize the out of sample MSE):
   1. Ridge Regression
   2. Lasso Regression
5. **Deliverables**: Please submit a one-page report that includes the following:
   1. Variable definitions
   2. Definition of the test/training data
   3. A table summarizing key model parameters in the final models (e.g., which features are retained in the Lasso model, R2 of your predictions, etc.)
   4. A brief paragraph or two summarizing any insights (e.g., what model worked best? Why?, Do the selected features make sense? Are they directionally correct?).

Please submit the brief report and all code via LMS.