Final Projects for FIE 453 Fall 2021

October 11, 2021

Here are some suggestions for final projects. You should think of this as a source of inspiration, rather than as a prescription of what to do. Remember that projects will be judged on both execution and level of ambition.

Basic Projects Use CRSP and Compustat data to predict one of the following variables. (You may also think of an interesting variant on your own.)

- Predict returns.

- Predict excess returns: returns minus market returns.

- Identify which stock beat the market in a month.

- Predict which stocks are in the top decile (or some other quantile) of returns in a month.

- Predicting returns is genuinely hard, so a possible alternative is to predict volatility of stuck prices.

- Rather than trying to predict a stock market outcome, predict some operational characteristic of the firm, such as earnings per share, or revenue.

- Predict a financial ratio, such as market-to-book ratio, or price-earnings ration. These measure roughly how heavily the market weighs the future value of the firm versus it's curent value.

Keep in mind that you may need feature engineering to get good results:

- Unlike vanilla linear or logistic regression, many methods are sensitive to how the variables are scaled. Consider the best way to scale the variables appropriately.

- The firms vary wildly in size. Is it better to use levels or ratios, or even a mix of both? What can we do to avoid dividing by zero?

One additional question that you can try to answer is how would you turn this information into an actual trading strategy.

New Data and Models   These projects use additional data that you must either assemble yourself.

- Suppose you turn your results into a portfolio strategy. We know from investment theory that the excess returns on such a strategy can be a reward for bearing risk. Use a factor model such as the Fama-French 3-factor model to evaluate the riskiness of your approach.

- The stock market is frequently taken as a barometer for the health of the US economy. Can we use macroeconomic data to predict stock returns?

- Can you use firm-level data to improve predictions?

These use models we didn't cover in class:

- The "Dissecting Characteristics Nonparametrically" paper uses a group lasso approach to mix the advantages of splines and lasso. Can you implement that approach here?

- The paper "Empirical Asset Pricing via Machine Learning" compares a range of machine learning techniques, including deep learning models. This uses existing stock market predictors. Can you adapt anything from this paper?

- There's a topic called survival analysis, which provide a class of models that have been extended for machine learning techniques. Is there a question you can answer using survival analysis?