# *Capturing Dynamics of Post Earnings Announcement Drift using Genetic Algorithm Optimized Supervised Learnings* -- More Results

Joseph Ye

## Experiment Setup

Following our meeting on 31st July I would like to present more analysis and results for the work on forecasting a company's Post Earnings Announcement Drift (PEAD) and forming portfolios whose actual portfolio returns can be ranked from high to low by ranking stocks according to the forecasted stock returns.

While the overall set up of the models are the same as detailed in the draft paper circulated to you previously, in these new results I've modified and enhanced the overall experimental setups as described below:

- Increase the total number of companies used for model training and target forecasting from 386 to 1106;
- Add more input features to the learning process such as the short interest ratios etc;
- Focus on forecasting returns of stocks whose earnings were released during the same earning release quarter or released on the same day instead of over multiple quarters;
- Experiment training and forecasting within only one industrial sector;
- Experiment more combinations of forecast start date and end date;
- Experiment training and forecasting with only Small Cap, Mid Cap or Large Cap companies;

All of these experiments are conducted using XGBoost trained using Genetic Algorithm. I have not experimented with MLP or Support Vector Machines as I already proved in the previous round of tests that their results were inferior to those of XGBoost.

## Data

In this round of new experiments, I have chosen to source stocks from the Russell 1000 index instead of S&P500 between Q4 1996 and Q4 2018. As new companies were added to or removed from the index over time the total number of companies involved in the training and forecasting experiments is 1106. By including companies that ceased to exist my experiments should not be subject to survival bias. I have sourced the following data for each of these 1106 companies between Q4 1996 and Q4 2018 as long as a company was in operation and data was available on Bloomberg:

### Input data - Financial Reports data

The following set of financial report metrics have been chosen to represent the key information that could affect how the stock market as a whole evaluates and understands how well a company's trading did in the previous quarter and the company's going performance.

| | |
|---|---|
| Cash | Operating Margin |
| Cash from Operating Activities | PB Ratios |
| Cost of Revenue | PC Ratios |
| Current Ratio | PS Ratios |
| Dividend Payout Ratio | Quick Ratio |
| Dividend Yield | Return On Assets |
| Free Cash Flow | Return On Common Equity |
| Gross Profit | Revenue |
| Income from Continued Operations | Short Term Debt |
| Inventory Turnover | Total Asset |
| Net Debt to EBIT | Total Asset |
| Net Income | Total Debt to Total Assets |
| Operating Expenses | Total Debt to Total Equity |
| Operating Income | Total Inventory |
| | Total Liabilities |

Instead of using these financial report metrics as they are, I have created quarterly change and yearly change of each metric and use them as the model's input features instead. If any feature data is unavailable for a particular quarter, the input for that feature is set to 0.

Each company's data is standardized as well as processed for outliers (using Winsorization) at the company level before they are joined together with other companies' data to form a matrix of input features which together represent the model input.

### Input data - Momentum Data

The following momentum data are used as part of the input features. Momentum data are sourced on the forecast start date. For example if forecast is set to start from the market close price prior to the earning reports release, momentum data will be sourced on the date prior to earnings release.

- RSI 9
- RSI 30
- 5 Day Moving Average
- 50 Day Moving Average

- 200 Day Moving Average

## Input data – Short interest ratios

Short interest ratio is released for most companies twice a month and is calculated by dividing the number of shares short in a stock by the stock's average daily trading volume. The short interest ratio is a good gauge on how heavily shorted a stock may be versus its trading volume. The most recent short interest ratio for each company prior to its earnings release is sourced as an input feature to the model for that company.

## Input data – Earnings Per Share (EPS) surprises

Earnings Surprise represents how much a company's actual reported Earnings Per Share (EPS) is ore (or less) than the average of a selected group of stock analysts' estimates on that quarter's EPS. I've engineered the following three features related to Earnings Surprise:

- Current quarter's Earnings Surprise (reported EPS market estimated EPS);
- Difference between current quarter's Earnings Surprise and that of the previous quarter;
- Difference between current quarter's Earnings Surprise and the average Earnings surprise of the preceding three quarters;

## Target data – Cumulative Abnormal Returns

Since we are studying the Post Earning Announcement Drifts (PEAD) which is represented by the Cumulative Abnormal Return of a stock post earnings release, the majority of experiments are conducted on stock price forecast from the last publicly tradable price available prior to earnings release over a certain period of time.

I am using the following rules to determine the *last publicly tradable stock price* that is available to my experiments:

1. If earning is released after market is closed, it's the closing price on the release date;

2. If earning is released before market opens, it's the closing price on the day before release date;
3. If earning is released in the middle of release date, it's the opening price on the release date;

To calculate Cumulative Abnormal Return, I have chosen the S&P500 index to proxy the overall stock market and the daily abnormal return is calculated as $Return_{stock} - Return_{S\&P500}$.

## Forecast Results

Following on from our last meeting I am focusing our discussions on how well the trained model is able to help form portfolios whose actual returns can be ranked from high to low. Therefore all the results shown in this section are returns of portfolios that are formed out of stocks which have ***already been ranked*** according to the model-predicted risk-adjusted return of the stocks. We will go through results in a number of experiment scenarios before we draw certain conclusions.

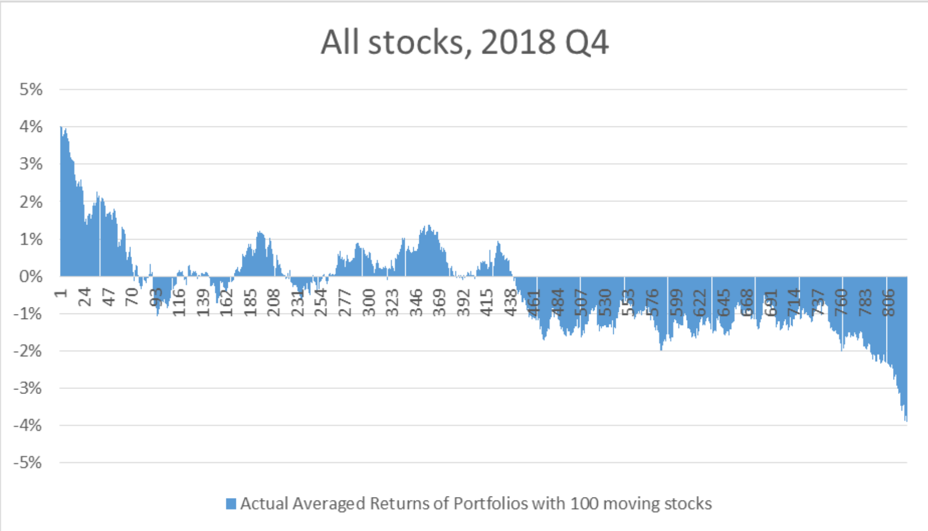
Unless explicitly stated otherwise, data from all the thirty years' quarters, excluding data points from those quarter(s) or date(s) chosen for forecasting, are used for model training and cross validation.

Unless explicitly stated otherwise, the starting price of a PEAD forecast is the last publicly tradable stock price prior to earnings release which is chosen using aforementioned rules.

## Scenarios 1: Thirty day post-release Cumulative Abnormal Returns forecast, Q4 2018

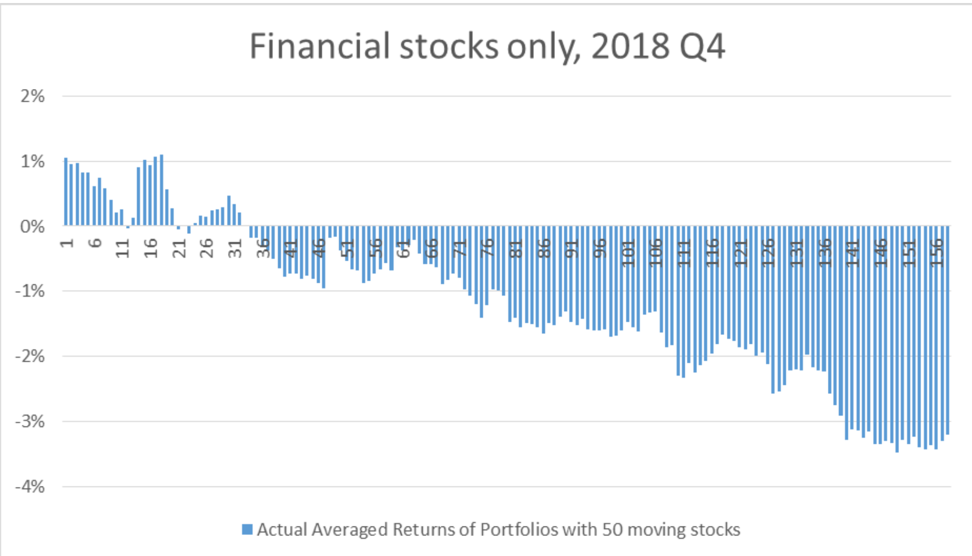
### Test 1.1:

- **30 day** PEAD forecast for **2018 Q4** releases.
- **All sectors** totalling 924 stocks are included in the forecasts.
- Portfolio includes 100 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **57%** classification success rate.



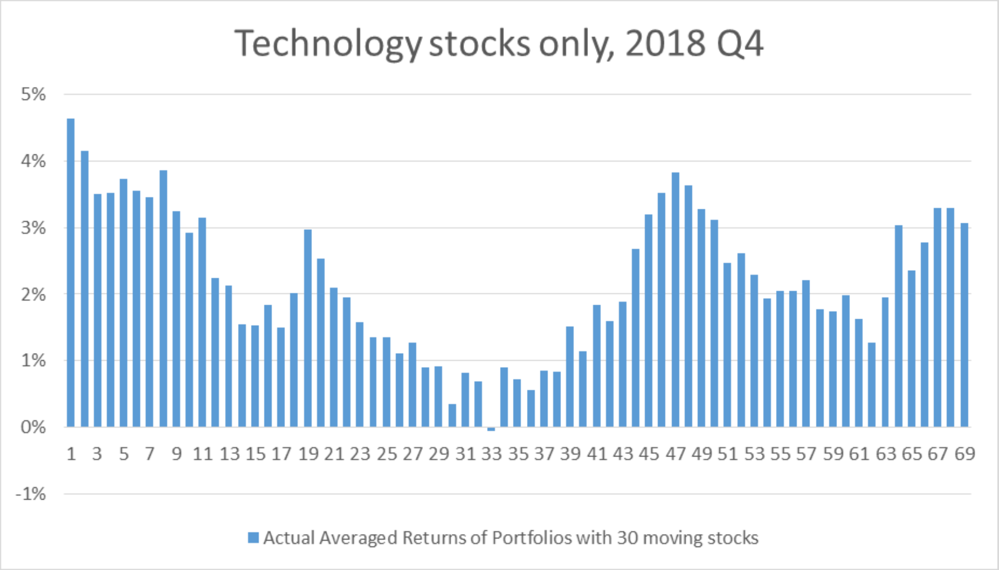

### Test 1.2:

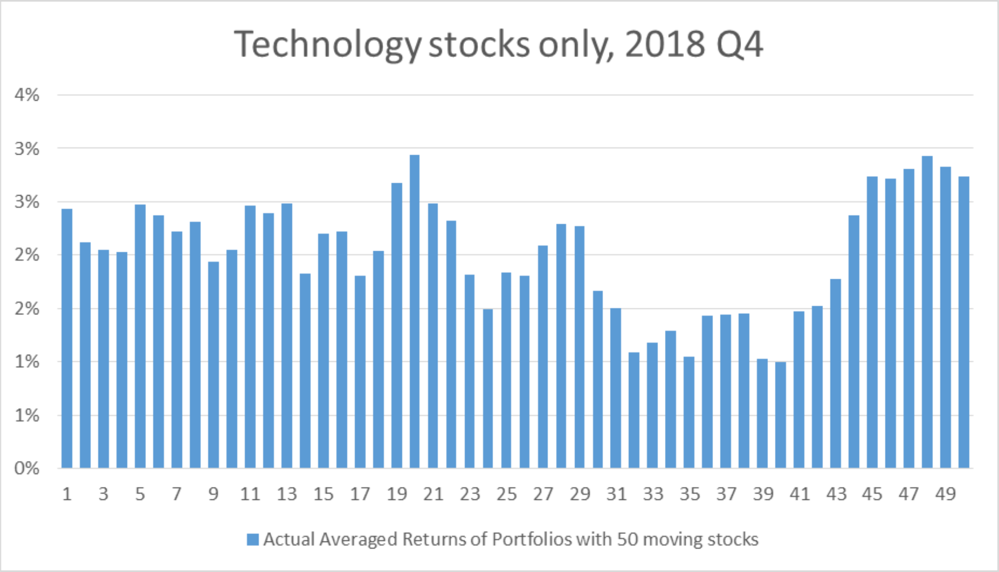
- **30 day** PEAD forecast for **2018 Q4** releases.
- Only the **Financial** sector totalling 207 stocks are included in the forecasts.
- 7923 data points are used for model training.
- Portfolio includes 50 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **63%** classification success rate.
- The actual averaged 30 day return over the same period was -1.13%



### Test 1.3:

- **30 day** PEAD forecast for **2018 Q4** releases.
- Only the **Technology** sector totalling 99 stocks are included in the forecasts.
- 3658 data points are used for model training.
- Two portfolios includes 30 and 50 moving stocks respectively; Chart shows actual portfolio risk adjusted returns.
- **56%** classification success rate.

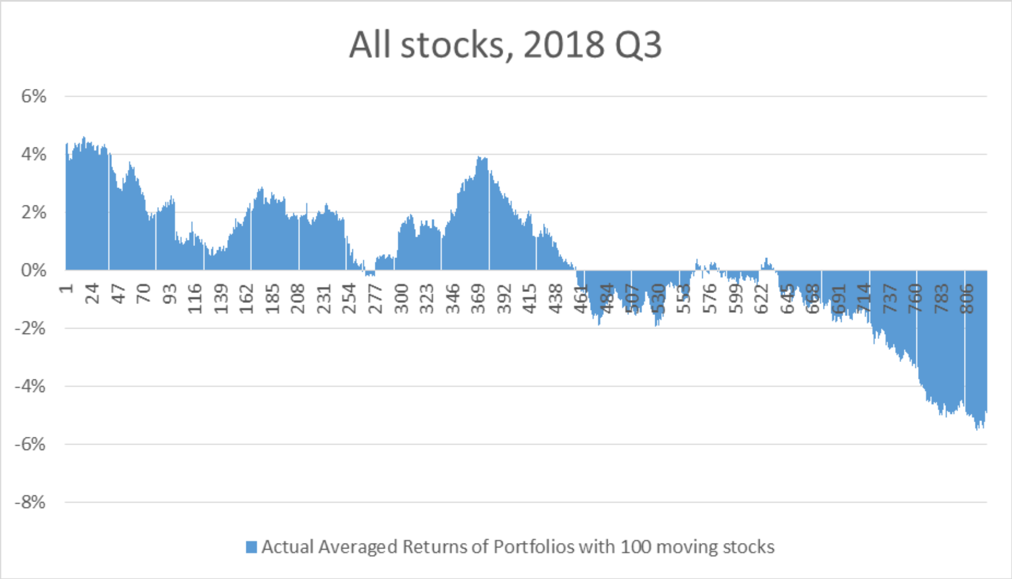


## Scenarios 2: Thirty day post-release Cumulative Abnormal Returns forecast, Q3 2018

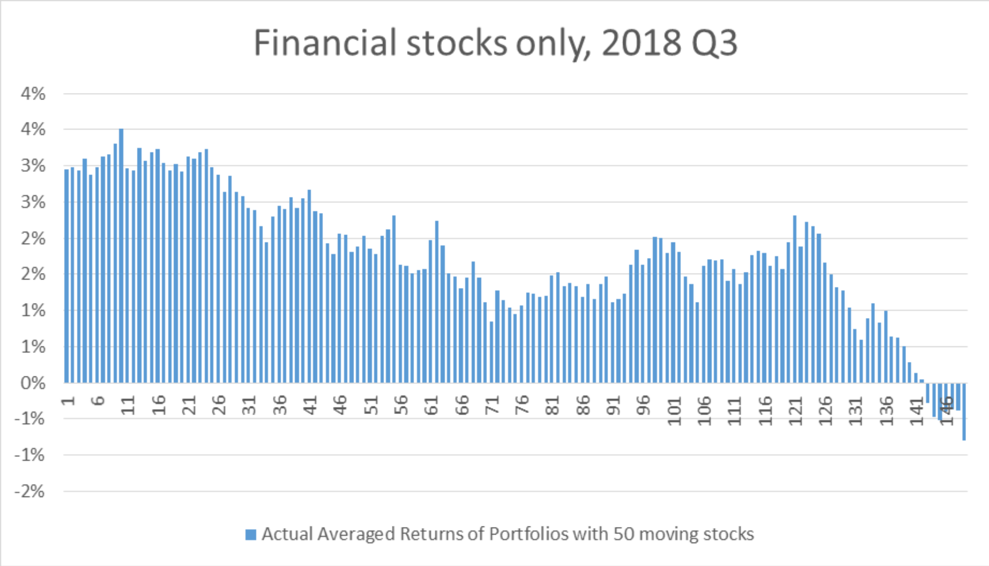
### Test 2.1:

- **30 day** PEAD forecast for **2018 Q3** releases.
- **All sectors** totalling 899 stocks are included in the forecasts.
- 35671 data points are used for model training.
- Portfolio includes 100 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **59%** classification success rate.



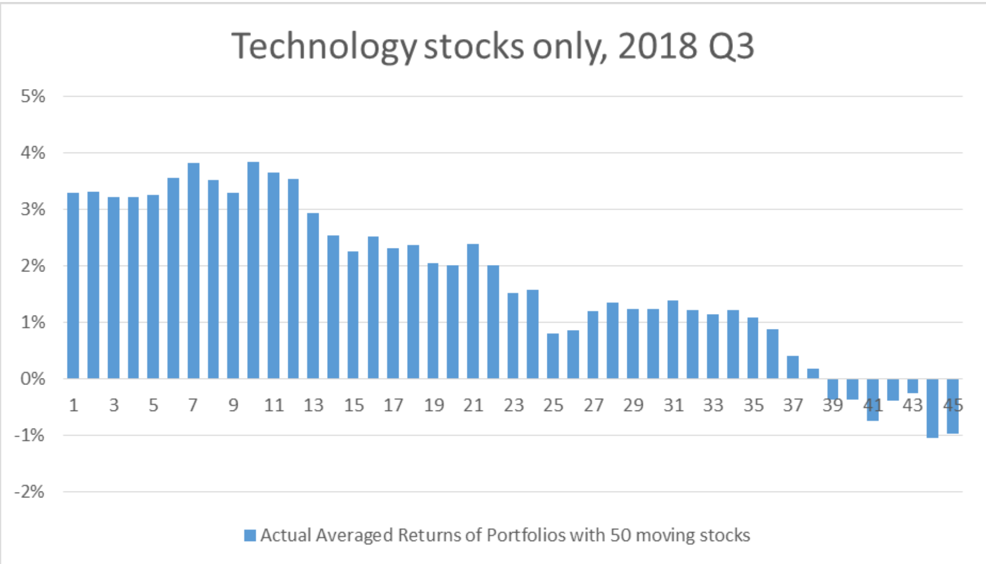

### Test 2.2:

- **30 day** PEAD forecast for **2018 Q3** releases.
- Only the **Financial** sector totalling 207 stocks are included in the forecasts.
- 7923 data points are used for model training.
- Portfolio includes 50 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **62%** classification success rate.




## Test 2.3:

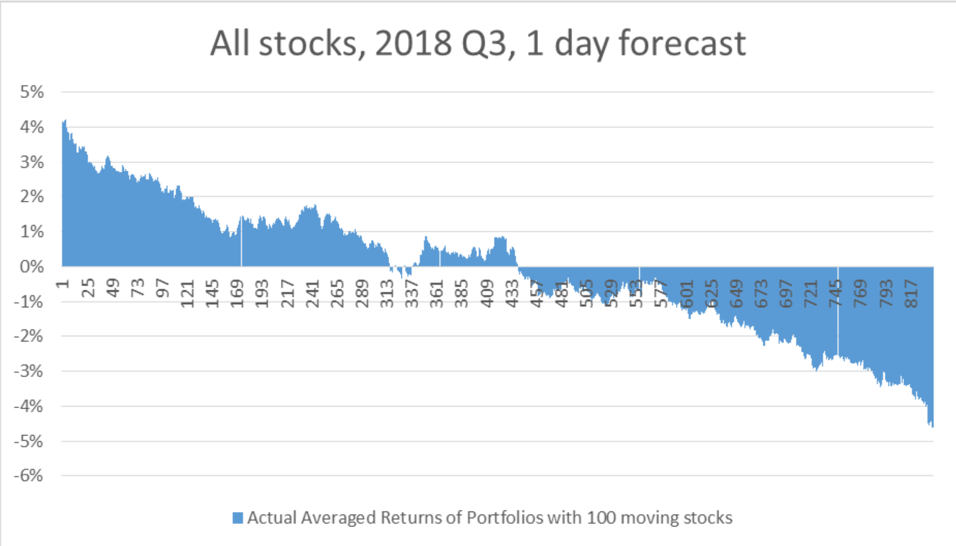
- **30 day** PEAD forecast for **2018 Q3** releases.
- Only the **Technology** sector totalling 94 stocks are included in the forecasts.
- 3564 data points are used for model training.
- Portfolio includes 50 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **61%** classification success rate.



## Scenarios 3: One day post-release Cumulative Abnormal Returns forecast

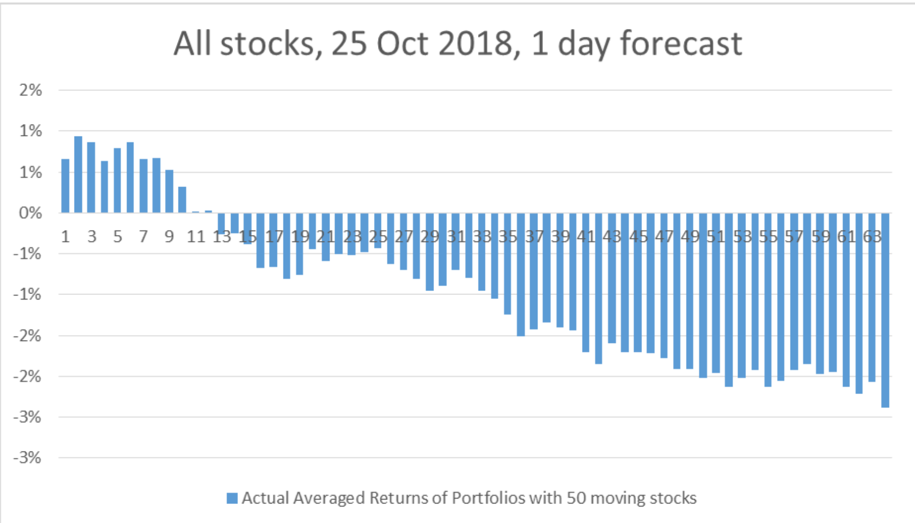
### Test 3.1:

- **1 day** PEAD forecast for **Q3 2018** releases.
- **All sectors** totalling 938 stocks are included in the forecasts.
- 46528 data points are used for model training.
- Portfolio includes 100 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **62%** classification success rate.



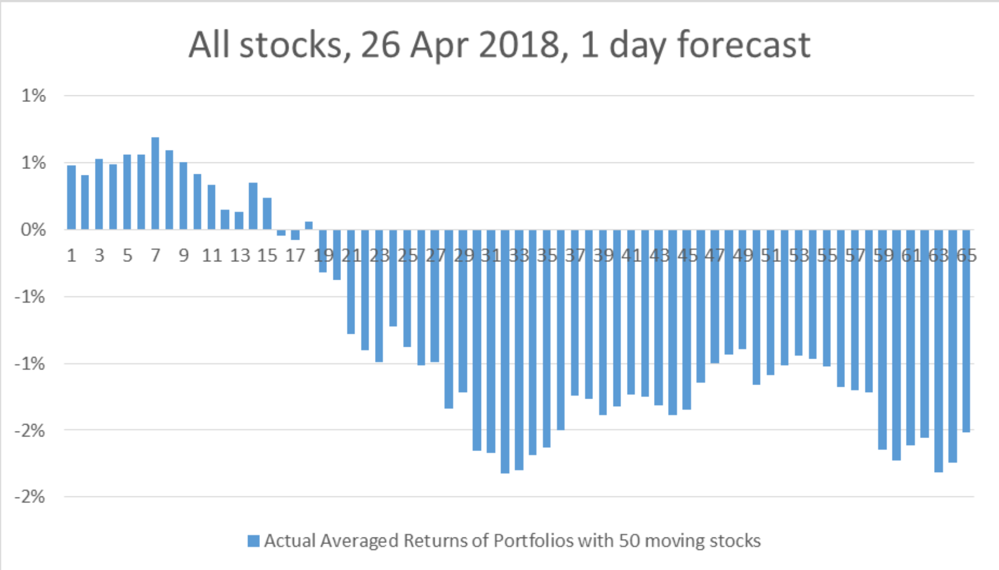

### Test 3.2:

- **1 day** PEAD forecast for stocks with earnings release on **25 Oct 2018**.
- **All sectors** totalling 113 stocks are included in the forecasts.
- 48278 data points are used for model training.
- Portfolio includes 50 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **61%** classification success rate.



## Test 3.3:

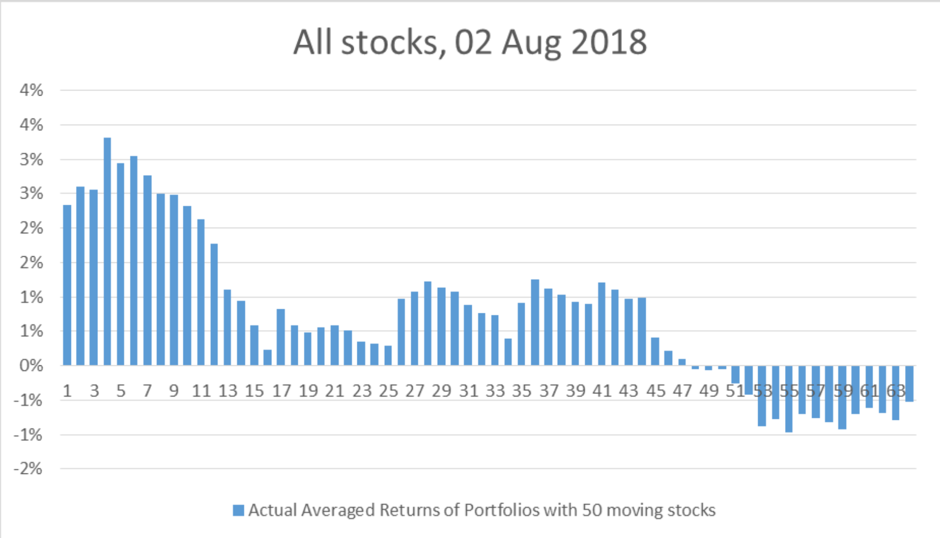
- **1 day** PEAD forecast for stocks with earnings release on **26 April 2018**.
- **All sectors** totalling 116 stocks are included in the forecasts.
- 48275 data points are used for model training.
- Portfolio includes 50 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **57%** classification success rate.



## Scenarios 4: Thirty day post-release Cumulative Abnormal Returns forecast – releases on a single day

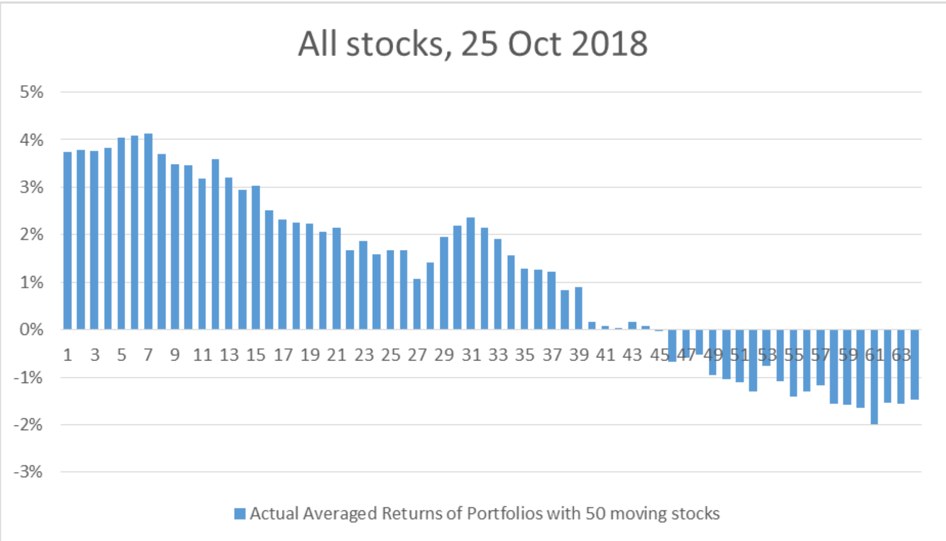
### Test 4.1:

- **30 day** PEAD forecast for stocks with earnings release on **02 Aug 2018**.
- **All sectors** totalling 96 stocks are included in the forecasts.
- 48195 data points are used for model training.
- Portfolio includes 50 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **65%** classification success rate.




### Test 4.2:

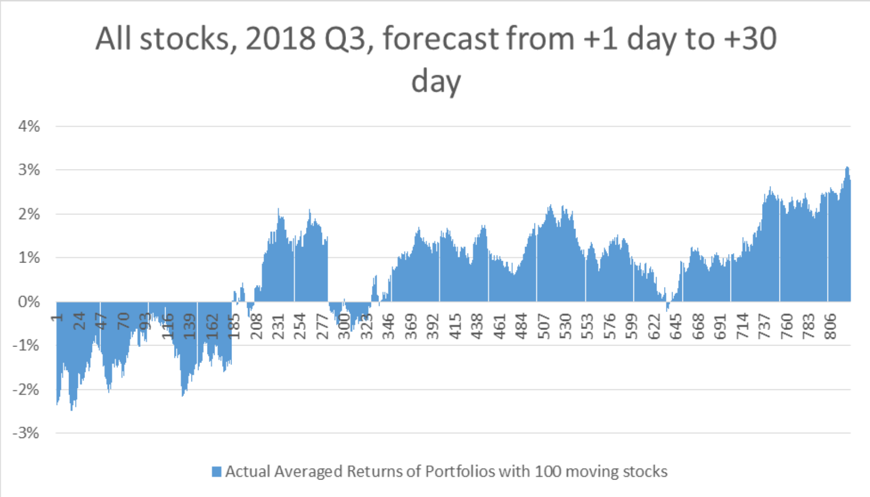
- **30 day** PEAD forecast for stocks with earnings release on **20 Oct 2018**.
- **All sectors** totalling 113 stocks are included in the forecasts.
- 48178 data points are used for model training.
- Portfolio includes 50 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **63%** classification success rate.



## Scenarios 5: Cumulative Abnormal Return forecasts starting after releases

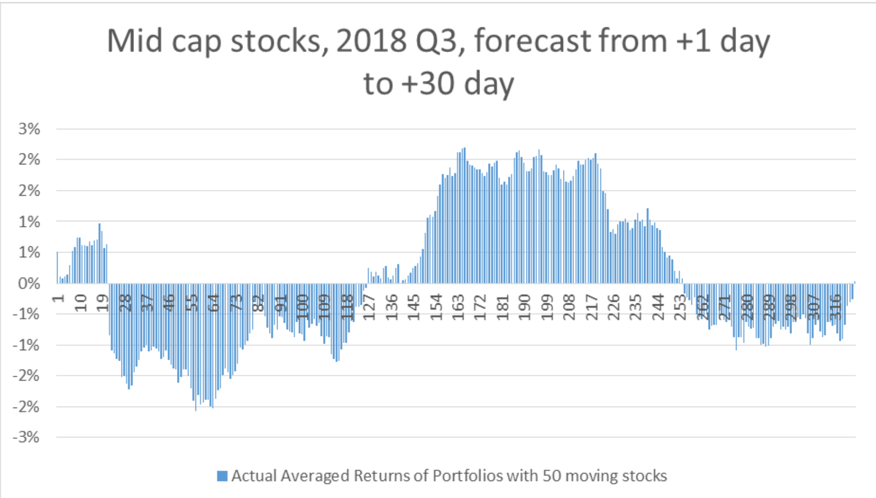
### Test 5.1:

- **29 day** PEAD forecast for **Q3 2018** releases.
- Forecasting from **1D** after release to 30D after release.
- **All sectors** totalling 899 stocks are included in the forecasts.
- 46465 data points are used for model training.
- Portfolio includes 100 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **50%** classification success rate.



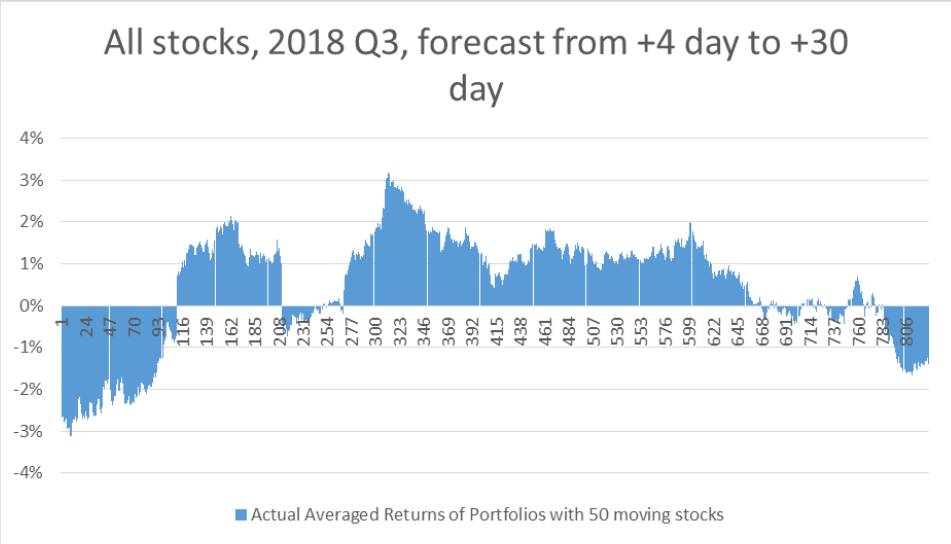

### Test 5.2:

- **29 day** PEAD forecast for **Q3 2018** releases.
- Forecasting from **1D** after release to 30D after release.
- **Mid Cap** stocks only totalling 423 stocks are included in the forecasts.
- 22534 data points are used for model training.
- Portfolio includes 100 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **51%** classification success rate.



### Test 5.3:

- **26 day** PEAD forecast for **Q3 2018** releases.
- Forecasting from **4D** after release to 30D after release.
- **All sectors** totalling 899 stocks are included in the forecasts.
- 46465 data points are used for model training.
- Portfolio includes 100 moving stocks; Chart shows actual portfolio risk adjusted returns.
- **50%** classification success rate.



## Conclusions

As a brief summary of results of the various experiments I would like to draw the following brief conclusions:

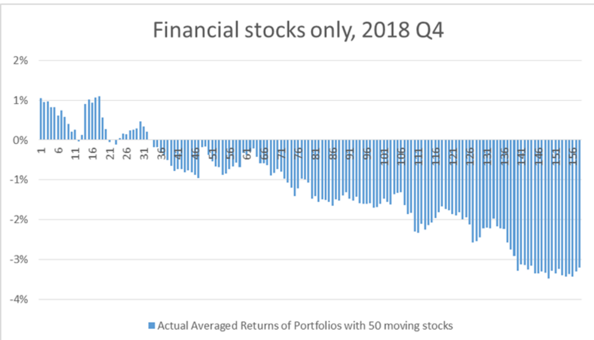
1. In the majority of experiments where the starting price of a forecasted stock price return is the last publicly tradable stock price prior to earnings release, the model is consistently showing a good level of success rate in distinguishing stocks whose actual returns will turn out positive from stocks with negative actual returns. This observation is first captured by the classification success rate of the forecasted stock returns against actual returns and in many scenarios the success rates are consistently around 60%.

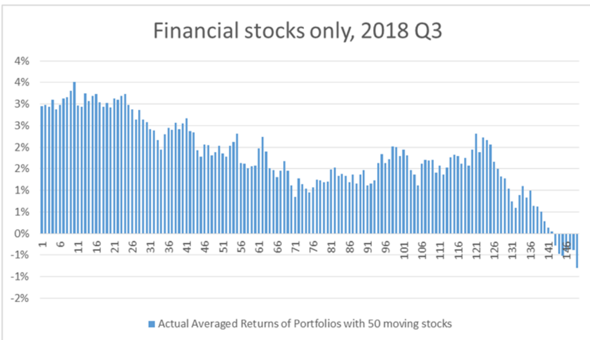
   A more important observation however is that, if stocks are firstly ranked according to their forecasted stock returns from high to low, portfolios that consist of a subset of moving stocks that have been ranked from high to low are generating portfolio returns also from high to low, with portfolios consisting of stocks from the top tiles or bottom tiles showing more pronounced results. For example, once stocks have been ranked according to their forecasted returns, portfolios that consist of stocks in the top tile will very likely show the highest return among other combinations of stocks, whereas portfolios that consist of stocks from the bottom tile will very likely show the lowest negative returns among other combinations of stocks.

   These observations are backed by results in tests from Scenarios 1, 2, 3, and 4, except for test 1.3.

2. The main reason I've chosen to conduct experiments against Q3 and Q4 2018 is that the US stock market went through this period of time with a great deal of volatilities and it would be very interesting to see how the model handles more extreme market conditions. The markets went through the last quarter of 2018 with huge distress spooked by the Fed's rate rises among other things with the S&P500 losing close to 20% in that period of time but then enjoyed a major rebound in the first quarter of 2019 claiming back all the previous losses.

   Interestingly some of these market volatilities have been captured by my results. Tests 1.2 and 2.2 both carry out forecasts on Financial stocks but respectively in 2018 Q3 and Q4. In 2018 Q4 the averaged actual 30-day risk-adjusted returns of Financial stocks was -1.13% whereas in 2018 Q3 the averaged actual 30-day risk-adjusted returns of Financial stocks was 1.46% with 63% of stocks recording a positive returns. These two contrasted performances were accurately captured by the following two results using the forecasted 30-day returns by the model.




3. Of all the 1-day or 30-day forecast experiments starting from the last publicly tradable stock price prior to earnings release, there is one test case that has failed to provide distinguishable results.

Results of test 1.3 indicate that there were no portfolios that could be formed that would generate a negative return, which isn't true. After investigation, I can attribute this failure partly to the model's inability to adapt to such a rapid market rebound and partly to the much smaller set of stocks used for model training.

In fact the risk-adjusted returns of Technology stocks were un-usually high in that period of time (first quarter of 2019), with the 99 chosen Technology stocks achieving an averaged 30-day risk-adjusted actual returns of 2.59%. This is exceptionally high because if the population of stocks used for forecasting is large enough, these stocks' average actual risk-adjusted return should really be close to zero.

4. To demonstrate the operability of using the model results to form tradable portfolios, Scenarios 4 shows that we can achieve the same kind of results as in the previous three test scenarios when we carry out forecast on Cumulative Abnormal Returns on stocks which report their earnings on the same day. Since the crust of our model is about being able to help form portfolios with a fixed holding period whose returns can be viewed from high to low, practically speaking it only makes sense if we are able to form portfolios by buying the model-chosen stocks within a short time frame, such as within a day or ideally less.

   In the two tests conducted here, I've chosen 02 Aug 2018 and 20 Oct 2018 when a larger group of stocks filed their financial reports. Both tests have in fact shown a higher degree of classification success rate (63% and 65% respectively) compared to other tests, possibly due to a consistent market environment for all the stocks involved when the model forecasts their 30-day risk adjusted returns all from the same starting day.

5. Results from Scenario 5 show that the model is not able to produce the same results if stock forecasting is carried out after financial reports have come out. On the surface this seems understandable because the majority of the earnings information will have been consumed by the initial tradings in the market and hence the impacts of my model inputs, majority of which come from the financial reports, would be weaken. The signals are seemingly gone.

   However, I believe the quality of a company's earning results and its forward guidance should continue to impact its price movements in the medium term. However this can sometimes be a big ask as we already see that value trading hasn't worked for a lot of stocks for a long time.

   Another important note I would like to make is that, stock movements post earning releases can be significantly affected by company management's forward guidance on the following quarter(s). Since I've not been able to source this kind of data for my model it is entirely possible that, with the help of such kind of data, carrying out forecast after results have come out would potentially become more accurate.

6. Finally, I attribute the model's ability to distinguish positive returning stocks from negative ones to the set of inputs I've chosen and the XGBoost model which is the core of my model. It is clear that there is a fundamental relationship between the chosen inputs and the realised Post Earnings-Announcement Drifts. When given the input features of any data point (a stock from an earning-releasing quarter), a trained-up XGBoost is having a good success at correctly identifying a positive-returning stock as such, by generating a positive forecasted return for this data point.

Another note is that, although the overall classification success rate stands around 60%, in a lot of the tests I've conducted I have observed greater success rates at the two ends of the distribution of the forecasted returns. This is understandable as the model should be more sensitive to those data points that react more promptly to the inputs. This explains why the portfolios consisting of stocks at either end of the spectrum seem to more reliably produce higher positive or lower negative returns.