Capturing dynamics of post-earnings-announcement drift using a genetic algorithm-optimized XGBoost

Zhengxin Joseph Ye [*], Björn W. Schuller

GLAM, Department of Computing, Imperial College London, UK

1. Introduction

The stock market is characterised by nonlinearities, discontinuities, and multi-polynomial components because it continuously interacts with many factors such as an individual company's news, political events, macro economic conditions, and general supply and demand, etc. (Göçken, Özçalıcı, Boru, & Dosdoğru, 2016). The non-stationary nature of the stock market is supported by a widely accepted, but still hotly contested economic theory *Efficient Market Hypothesis* (Fama, 1970Fama, 1991) which states that asset prices fully reflect all available information and the market only moves by reacting to new information. Such a theory implies that the stock market behaves like a martingale and knowledge of all past prices is not informative regarding the expectation of future prices.

Ball and Brown (Ball & Brown, 1968) were the first to note that after earnings are announced, estimated cumulative abnormal returns continue to drift up for firms that are perceived to have reported good financial results for the preceding quarter and drift down for firms whose results have turned out worse than the market had expected. The discovery of Post Earnings Announcement Drift (PEAD), which is a violation of a semi-strong Efficient Market Hypothesis, seems to suggest that while stock markets are generally efficient, there may be information leakages around the announcement dates, coupled with post-earnings drift, resulting in price movement anomalies. It also seems to suggest that past stock price information or other past economic or financial information can potentially be used to predict price movement following a significant economic event such as an earnings announcement.

Research on Post Earnings Announcement Drift proliferated in the late 1980s and 1990s. Fama and French (Fama & French, 1993) show

* Corresponding author.
E-mail addresses: z.ye18@imperial.ac.uk (Z.J. Ye), bjoern.schuller@imperial.ac.uk (B.W. Schuller).

that average stock returns co-vary with three factors, namely, the market risk factor, the book-to-market factor, and the size factor. Bhushan suggests that the existence of sophisticated and unsophisticated investors, transaction costs, and economies of scale in managing money can explain the market's delayed response to earnings (Bhushan, 1994). We notice a lot of previous research would pool companies with negative or positive earnings surprises when measuring its effect on abnormal returns and regress the absolute value of earnings surprise as well as other factors against the absolute value of abnormal return (Qiu, 2014). However, we have found that stock markets do not just react symmetrically to negative and positive earnings surprises and there are a lot more factors in play that drive the near term risk adjusted returns of a stock following an earnings release.

Rather than trying to analyse the link between PEAD and economic and accounting factors as commonly seen in the literature, we manage to leap straight to the more important goal of predicting the direction of PEAD by using machine learning models. In this process we have overcome several constraints commonly seen in the previous research: we are including a much wider range of factors including both fundamental and technical/momentum factors; we achieve a higher level of generality without having to pre-group companies by the value of their earnings surprises or other attributes prior to the analysis or prediction (*subsample analysis*) which is common in the literature (Baker, Ni, Saadi, & Zhu, 2016). Additionally we have chosen 1 106 stocks that are or once existed as components of the Russell 1 000 index (which tracks approximately the 1 000 largest public companies in the US) during the chosen time period between 1997 and 2018. Our selection includes companies that either went bankrupt or dropped out of Russell 1 000, significantly reducing survivorship bias in our training data. This test population is larger than most earlier studies of similar nature. For example, Beyaz and colleagues only chose 140 stocks from S&P500 when they attempted to forecast stock prices both six months and a year out based on fundamental analysis and technical analysis (Beyaz, Tekiner, Zeng, & Keane, 2018), and Bradbury used a sample of only 172 firms to research the relationships among voluntary semi-annual earnings disclosures, earnings volatility, unexpected earnings, and firm size (Bradbury, 1992). Our results generalise better with the universe of stocks on the US markets.

Recognising the highly nonlinear nature of stock price movements, we have chosen to run our experiments using XGBoost which is a state-of-the-art supervised learning model. We divide the training data into in-sample and out-of-sample periods of varying lengths and use part of the in-sample data set to optimise the model's hyperparameters before training it. Our earlier experiments show that grid search (Liashchynskyi & Liashchynskyi, 2019) as a traditional way of finding an optimal parameter set is inexhaustive and can be very slow. Instead we have chosen to use the highly adaptable Genetic Algorithm (GA) (Deng, Yoshiyama, Mitsubuchi, & Sakurai, 2015) to optimise our models. We recognise hyperparameter optimisation is a delicate step and searching with a limited set of parameters will result in a non-optimal model which will not able to fit the essential structure of the training data set. To avoid this potential problem, we have chosen to use a broad value range and a small granular step for each of the hyperparameters. A 5-fold cross validation (CV) is employed within each GA iteration during the optimisation.

Our machine learning-based approach is in direct contrast to most earlier financial research work in the literature as typified by (Kim & Kim, 2003), which sought to devise different portfolios *a priori* by different factor characteristics and tried to analyse and make sense of the link between portfolio returns and the corresponding economic factors that segregated the portfolios. Instead, our model directly learns the intrinsic link between the input feature space and stock price returns. We find that stocks that belong to different industrial sectors can have their PEAD movements driven by different primary factors and such factors can also change from quarter to quarter. Despite such differences and changes in the driving factors, a GA optimised XGBoost model is able to pick up the underlying signals embedded in our engineered features and forecast the 30-day post earnings drift direction with reasonable accuracies. We also study the possibility of grouping stocks into portfolios according to their predicted levels of Cumulative Abnormal Return (CAR). We have found that ranking the out-of-sample stocks by their predicted 30-day abnormal returns could help form portfolios which consistently offer higher positive returns and lower negative returns, a result that could potentially form the basis of further usage in market neutral long-short trading strategies. In the end, we also look at the challenges of applying predictive models in real life markets due to ever changing market prices and asymmetrical level of information access by certain market participants. We share a tactic that can turn a model's forecasts into actionable signals.

2. Related work

Since the discovery of Post Earnings Announcement Drift as a stock market anomaly by Ball and Brown (Ball & Brown, 1968) who documented the return predictability for up to two months after the annual earnings announcements, extensive research has been carried out in the literature though with varying results. For example, Foster, Olsen, and Shevlin (Foster, Olsen, & Shevlin, 1984) found that systematic post-announcement drifts in security returns are only observed for a subset of earnings expectations models when testing drifts in the [+1, +60] trading day period. In recent years, the literature has become less limited to the specific study of PEAD and instead put more focus on the direct predictions of stock price movement using stocks' fundamental and/or technical information, again with varying rates of success. Malkiel studied the impact of price/earnings (P/E) ratios and dividend yields on stock prices using the Campbell-Shiller model. He conceded his work demonstrating that exploitable arbitrage did not exist for investors to earn excess risk-adjusted returns and he could not find a market timing strategy capable of producing returns exceeding a buy-and-hold on a broad market index (Malkiel, 2004). Olson and Mossman on the other hand not only showed that an artificial neural network (ANN) outperforms traditional regression based methods when forecasting 12-month returns by examining 61 financial ratios for 2352 Canadian stocks, but, more importantly, shows that by using fundamental metrics sourced from earning reports, they were able to achieve excessive risk-adjusted returns (Olson & Mossman, 2003).

Other authors went beyond metrics from earnings reports and attempted stock forecast using both fundamental and technical analysis. Sheta et al. explored the use of ANNs, Support Vector Machines (SVMs), and Multiple Linear Regression for prediction of S&P500 market index. They selected 27 technical indicators as well as macro economic indicators and reported that SVM contributed to better predictions than the other models tested (Sheta, Ahmed, & Faris, 2015). Hafezi et al. considered both fundamental and technical analyses in a novel model called Bat-neural Network Multi-agent System when forecasting stock returns. The resulted mean absolute percentage error showed that the new model performed better than a typical Neural Network coupled with a GA (Hafezi, Shahrabi, & Hadavandi, 2015). Alternative data are becoming popular, too. Solberg and Karlsen investigated the possibility to predict the direction of stock prices using scripts of earnings conference calls. By analysing 29 330 different earnings call scripts between 2014 and 2017 using four different machine learning algorithms they managed to achieve a classification error rate of 43.8% using logistic regression and beat the S&P500 benchmark using both logistic regression and gradient boosting. Their results showed that earnings calls contain predictive power for the next day's stock price direction post earnings release (Solberg & Karlsen, 2018).

Researchers also studied how machine learning would directly benefit financial trading. Through a series of applications involving hundreds of predictors and stocks, Huck looked at how to apply some of the state-of-the-art machine learning techniques to manage a long-short portfolio. In that process he also explored a series of practical questions with regard to the predictor data and was able to show that the techniques he examined generated useful trading signals for portfolios with short holding periods (Huck, 2019). Sant'Anna and Caldeira applied Lasso regression for index tracking and long-short investing strategies. They used stocks from three benchmarks, S&P100, Russell 1000, and the Ibovespa Index from Brazil from 2010 to 2017 to assess the quality of Lasso-based tracking portfolios. By using co-integration as a benchmark method to solve the same problems, they demonstrated that the Lasso regression based approach was able to form portfolios that produced similar returns compared to using co-integration, but incurred significantly less transaction costs (Sant'anna, Caldeira, & Filomena, 2020).

As a model that has only recently burst on the scene, there is limited study of XGBoost in financial applications. Chatzis et al. (Chatzis, Siakoulis, Petropoulos, Stavroulakis, & Vlachogiannakis, 2018) evaluate the possibility of a market crash over a 1-day and 20-day horizon across the global markets. By using a vast set of data from global stock markets, bond markets and FX markets, the paper explores a large set of supervised learning models including Logistic Regression, Decision Trees, Random Forest, Support Vector Machines, Deep Neural Networks, and XGBoost. The paper draws conclusions by declaring the superiority of XGBoost over others by examining the forecast results on stock returns through a list of statistical measurement metrics. Li and Zhang (Jidong & Ran, 2018) use XGBoost to dynamically predict the value of a set of seven factors that contribute a stock selection process. Dynamically generated factors are then used to select a portfolio of different stocks whose return is measured over a multiple year period. Portfolios of dynamically selected stocks are shown to perform better than benchmark portfolios.

The only machine learning-based study on PEAD known to the authors was carried out by Schnaubelt et al. (Schnaubelt & Seifert, 2020) which used Random Forecast, another decision tree-based supervised learning model. While presenting a similar approach in selecting portfolios through ranking stocks by their model-predicted risk adjusted returns, the paper didn't delve much on benchmarking the model's general classification prowess in such an application context nor showing great details in model hyperparameter optimization which is a key step in model setup. On the other hand, we justify the suitability of taking the machine learning approach in PEAD studies by showcasing a detailed system/model setup as well as granular forecasting results at both the single stock and portfolio levels, accompanied by analysis on sector-specific driving factors. We hope to use this paper to define a benchmark approach in carrying out post earnings stock movement studies using machine learning.

3. Model features generation

We have chosen to use 1106 US companies in the Russell-1000 index in total. The data time frame is between the first financial quarter of 1997 (1997 Q1) and the fourth financial quarter of 2018 (2018 Q4). The model output is the 30 day Cumulative Abnormal Return post earnings release of each individual stock and the input space consists of the following set of unadjusted data which we have sourced from Bloomberg:

- Financial statements data
- Earnings Surprise data
- Momentum indicator data
- Short interest data

In total, we have sourced 97 901 quarterly financial statements from our chosen companies over the test time frame. The final population of valid data points used for training and testing whose input features include both financial statement metrics and other economic metrics stands close to 50,000, depending on the test cases. There are a number of reasons for the reduced population: (a) there are no Earnings data, Short interest data or other input feature data on Bloomberg for a good number of historical financial quarters within the test time frame; (b) we have discarded companies in certain historical quarters when the earnings reports suffered badly from missing data; (c) We have been very careful with whether an earnings report was released before market opened, after market closed or during trading hours as such a difference is significant as we would need to alter the forecast starting point accordingly. Bloomberg is missing the release time of day for some financial quarters in earlier years, and we have discarded those quarters.

Table 1
Earnings report metrics chosen as input features.

Cash	Operating Margin
Cash from Operating Activities	Price to Book Ratios
Cost of Revenue	Price to Cashflow Ratios
Current Ratio	Price to Sales 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

3.1. Financial statements data

As shown in Table 1, twenty-nine metrics from earnings reports have been chosen to create training data.

Based on the reported value of these metrics, we have engineered new features as quarterly change and yearly change of each of these financial metrics.

3.2. Earnings Surprise data

Earnings Surprise represents how much a company's actual reported Earnings Per Share (EPS) is more (or less) than the average of a selected group of stock analysts' estimates on the same quarter's EPS. We are not calculating Earnings Surprise as a %change between the reported EPS and market estimated EPS because (a) %change is very volatile when a EPS level is close to zero and a small change can lead to a misleadingly large %change, and (b) we would like to avoid the change-of-signs problem.

We have subsequently engineered the following three features related to Earnings Surprise:

- Current quarter's Earnings Surprise (reported EPS minus 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;

3.3. Momentum indicators

We have chosen the following technical/momentum indicator values calculated on the same day an individual company's quarterly earnings data was released:

- 9-day Relative Strength Index (RSI)
- 30-day Relative Strength Index
- 5-day Moving Average/50-day Moving Average
- 5-day Moving Average/200-day Moving Average
- 50-day Moving Average/ 200-day Moving Average

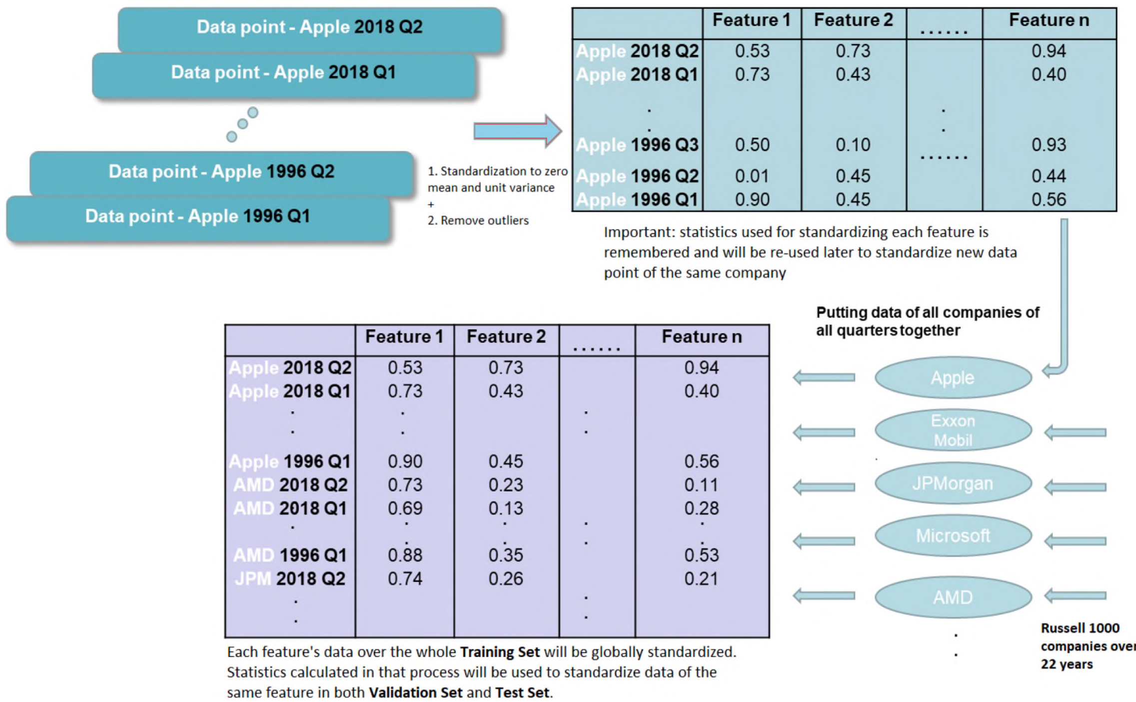

Fig. 1. Steps of Data Pre-processing.

We believe all these indicators should in a way measure how a stock's recent short term movements compare to its historical movements further back in time. The inclusion of momentum indicators is motivated by the intention to allow the prediction process of future stock movements to take into account a stock's recent movement trend as information leakage does happen prior to financial reportings. We have engineered the three ratios of short term moving averages to near or long term moving averages as proxies to the *golden cross* indicators.

3.4. Short interest data

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.

4. Data pre-processing

With totally 1106 companies involved over 21 years, there is a lot of data representing input features for each company at each quarter. In order for them to be understood by the model, we put them into a matrix-like data structure $A \in M_{m\times n}(\mathbb{R})$, where each of the m rows represents a n dimensional training data point, indexed by the pairing of a company name and a historical quarter, and each column holds data of the same feature from all the data points.

Before we put the data of all the companies and of all the quarters into a matrix, we pre-process each company's data to deal with outliers and to standardise data of every company. Firstly, we employ Winsorisation (Duan & Dunlap, 1998) to reduce the number of outliers present in the input features. This is carried out on the feature data of each individual company. Secondly, we standardise a selective group of features of each company. Every company's standardised features will then be stacked back into a full training data set. The pre-processing process is illustrated in Fig. 1.

5. Models and methods

5.1. Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is a scalable machine learning system for tree boosting invented by Tianqi Chen (Chen & Guestrin, 2016), which has gained much prominence in recent years. It distinguishes itself from other existing tree boosting methods (Tyree, Weinberger, Agrawal, & Paykin, 2011; Ye, Chow, Chen, & Zheng, 2009) by having cache-aware and sparsity-aware learnings. The former technology gives the system twice the speed against running a non-cache-aware but otherwise identical greedy tree splitting algorithm, and the latter gives an amazing 50 times speed boosting against a naive implementation handling an Allstate-10k dataset (Chen & Guestrin, 2016). More importantly, XGBoost has achieved algorithmic optimisations by introducing a regularised learning objective within a tree structure which helps achieve smart tree splitting and branch pruning.

For a data set in matrix form $A \in M_{m\times n}(\mathbb{R})$ with m data points and n features, a tree ensemble model uses K base leaner functions to predict the output:

$$\widetilde{y_i} = \phi\left(x_i\right) = \sum_{k=1}^{K} f_k\left(x_i\right), f_k \in \mathbb{F}, \tag{1}$$

where $\mathbb{F}$ is the space of regression trees. Each hypothesis f_k corresponds to an independent tree structure q with leaf scores ω. XGBoost utilises regression trees each of which contains a score on each of its leaves. These scores help form the decision rules in the trees to classify each set of inputs into leaves and calculate the final predicted output by summing up the scores in the related leaves. Unlike other standard gradient boosting models such as AdaBoost and GBM which do not intrinsically perform regularisation, XGBoost minimises a *regularised* loss function in order to learn the set of functions:

$$L\left(\phi\right) = \sum_i \ell\left(\widehat{y}_i, y_i\right) + \sum_k \Omega\left(f_k\right). \tag{2}$$

Here, ℓ is a differentiable convex loss function for the model output and the regularisation term is defined as (though not limited to) $\Omega(f) = \gamma T + \frac{1}{2}\lambda\left|\left|\omega\right|\right|^2$, which reduces the chance of overfitting. As in a typical gradient tree boosting model, a new base learner regression tree f_i which most minimises the loss function in Eq. (2) is greedily and iteratively added to the final loss function. Let $\widehat{y}_{i,t}$ be the model output of the i-th instance at the t-th iteration the loss function can be re-written as

$$L_t\left(\phi\right) = \sum_{i,k}\ell\left(y_i, \widehat{y}_{i,t-1} + f_t\left(x_i\right)\right) + \sum_k \Omega\left(f_{k,t}\right). \tag{3}$$

By taking the Taylor expansion on this loss function up to the second order and removing the constant terms as a result of the expansion the loss function can be simplified to:

$$L_t\left(\phi\right) = \sum_{j=1}^{T}\left[G_j\omega_j + \frac{1}{2}\left(H_j + \lambda\right)\omega_j^2\right] + \lambda T, \tag{4}$$

where

$$G_j = \sum_{i\in I_j} g_i$$

$$H_j = \sum_{i\in I_j} h_i$$

$$I_j = \{i|q(x_i) = j\}$$

$$g_i = \partial_{\widehat{y}_{i,t-1}}\ell\left(y_i, \widehat{y}_{i,t-1}\right) h_i = \partial_{\widehat{y}_{i,t-1}}{}^2\ell\left(y_i, \widehat{y}_{i,t-1}\right).$$

Here, T is the number of leaves in the tree. With ω_j being independent with respect to others, Tianqi (Chen & Guestrin, 2016) has proven that the best ω_j for a given tree structure $q(x)$ should be

$$\omega_j^* = -\frac{G_j}{H_j + \lambda}, \tag{5}$$

which in turn makes the objective function come to its final form:

$$L_j^* = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j + \lambda} + \gamma T. \tag{6}$$

$$L_{split} = \frac{1}{2}\left[\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \lambda} + \frac{\left(\sum_{i\in I} g_i\right)^2}{\sum_{i\in I} h_i + \lambda}\right] - \gamma \tag{7}$$

Ideally, the model would enumerate all possible tree structures with a quality score and pick the best one to be added iteratively. In reality, this is intractable and optimisation has to be executed one tree level at a time. This is made available by the final form of the loss function, as the model uses it as a scoring function to decide on the optimal leaf splitting point. Assume that I_L and I_R are the instance sets of left and right nodes after the split. Letting $I = I_L \cup I_R$, the scoring function for leaf splitting is Eq. (7).

These scores are then used by a method called the *exact greedy algorithm* to enumerate all the possible splits for continuous features, allowing each level of a tree to be optimised and the overall loss function to be minimised in the process. When deployed on a distributed platform XGBoost employs approximate algorithms instead to alleviate the huge memory consumption demanded by the exact greedy algorithm although this is not needed in our experiments which run on a single machine.

5.2. Hyperparameter optimisation

Model optimisation is one of the most important steps in ensuring the model output can meaningfully capture the underlying dynamics of the dependent variable. In search of optimal hyperparameter sets, we initially experimented a more straightforward approach of grid search but found it less effective in its performance and inexhuastive in the search results. GA as an adaptable and easily extensible heuristic optimisation method is chosen instead to carry out this task. Table 2 gives the list of model hyperparameters we have put through GA. Before we start the optimisation process, we first split the population of data into training data and test data. Selection of the out-of-sample test data varies and depends on the nature of a test which will be explained in subsequent sections. It is the training data that we use to optimise the model. We use 5-fold cross validation to calculate the fitness value on a particular set of hyperparameters examined by the GA. To do that we split the training data into five equal groups, use four groups to train the model and calculate the fitness value using the last group (validation). This process is repeated five times iteratively on each of the five groups and the final fitness value is the averaged fitness of the five iterations.

Table 2
XGBoost hyperparameters optimised by GA + CV.

Hyper Parameters
Gamma
Max depth
Sub sample
Learning Rate
Minimum child weight
Column sample by tree

To optimise the model, each hyperparameter is randomly initialised according to its own valid range of values. This initialisation is repeated 40 times so that we have 40 sets of randomly initialised hyperparameters to start the GA process with. Each set is called a *population*, and each hyperparameter within a set is called a *chromosome*. All of the 40 populations are considered to be part of the current *generation*. The GA process carries out a 5-fold cross-validation on a model using each of the 40 populations of parameters and when finished, keeps the 20 populations that have produced the smallest fitness values in the cross validation step. These 20 sets or *populations* of hyperparameters are considered to have performed better in forecasting post-announcement drifts with the current model than the 20 discarded ones. The 20 better populations are then used to *cross-breed* into 20 new populations and in this process *mutation* is allowed to happen to the cross-bred populations, i. e., chromosomes in the 20 newly created populations are allowed to randomly change value following a predefined level of probability. At the end of this process, we have produced a new and potentially better set of 40 populations of hyperparameters and we call them the *new generation*. The new generation are then fed through a second iteration of the GA process until eventually the minimum fitness value produced by the cross-validation step no longer changes its value within tolerance and at this point we have arrived at the optimal set of hyperparameters which produces the smallest fitness value when being used in the current model. Fig. 2 shows how GA and Cross Validation work together to produce the set of hyperparameters of each model which result in the highest prediction accuracy (smallest fitness value) on the validation set.

6. Results and analysis

All of our experiments centre around the 30 day post-earnings Cumulative Abnormal Return (CAR) as a measure of risk adjusted stock price return. An abnormal return is how much the actual return of a security is above its expected rate of return.

$$AR_i(t) = r_i(t) - E(r_i(t)) \tag{8}$$

where $AR_i(t)$ is the one-day abnormal return for company i on day t, $r_i(t)$ is the actual one-day stock return and $E(r_i(t))$ is the expected one-day return of stock i. As explored by Kim (Kim & Kim, 2003), there is a

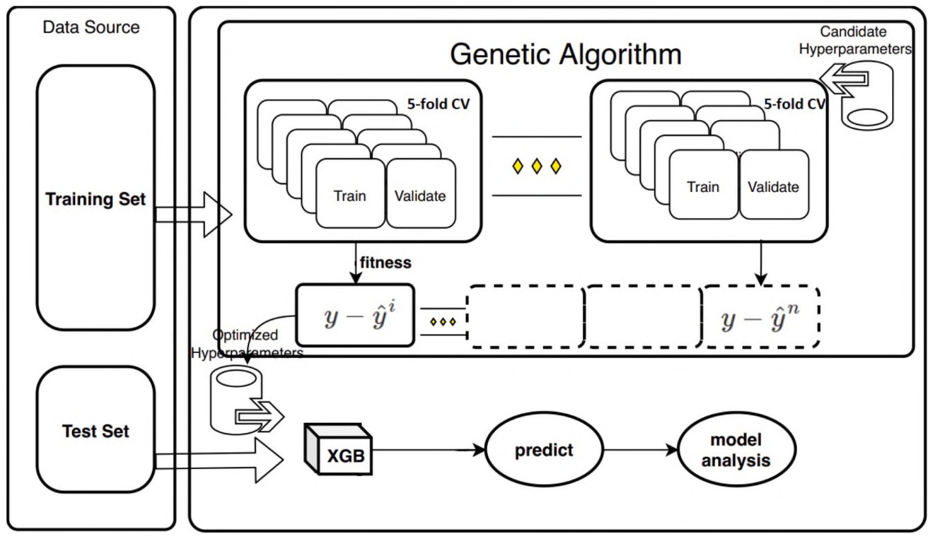

Fig. 2. Hyperparameter Optimisation using GA + CV.

variety of ways of evaluating the expected return including using quantitative models such as the one-factor CAPM model (Sharpe, 1964) and the Fama French three-factor model (Fama & French, 1993). In our experiments, we choose to use the CAPM model to calculate each stock's expected return which is defined as:

$$E(r_i) = r_f + \beta_i (E(r_m) - r_f) \tag{9}$$

where r_f is the risk free rate and $E(r_m)$ is the expected return of the market. We source historical 10Y U.S. Treasury Yields and return of the S&P500 index SPX from Bloomberg at the time of each earning event and use them to estimate r_f and $E(r_m)$ respectively. The $beta_i$ of each company, which measures a company's systematic risk in comparison to the market as a whole, is also sourced from Bloomberg on each historic earnings reporting date.

Consequently, our model output for stock i at time t (one of the many earnings release dates), which is the cumulative abnormal return from T_1 to T_n, is defined below. Throughout our experiments we've fixed n to 30 although other values of n could be chosen to conduct more analysis. We assume a company's beta and the 10Y Treasury yield remain static during the 30-day holding period over which period of time the CAR is calculated.

$$CAR_i\left(T_1, T_n\right) = \sum_{t=T_1}^{T_n}\left(AR_i\left(t\right)\right) = \sum_{t=T_1}^{T_n}\left(r_i\left(t\right) - E\left(r_i\left(t\right)\right)\right) \tag{10}$$

6.1. *Single Stock Forecast*

In this experiment, we have chosen stocks that filed for earnings with SEC in the four quarters in each financial year from 2015 to 2018 as our out-of-sample test population. That means, we first run a forecast on movement direction of all the stocks that reported earnings in 2015 while using all the data prior to 2015 as training data. Once done, we move on to repeat the same exercise on stocks that reported in 2016, etc. It should be noted that a company that filed in each of the 4 quarters of a financial year is considered as four independent data points since the only input data consumed by the XGBoost + GA model at any quarter are the near term financial signals as well as financial statement data of this stock in that particular quarter.

Separately, the same test as described above is also repeated on stocks belonging to a particular industrial sector. Bloomberg categorises US-listed companies into nine sectors: *Industrial, Basic Materials, Consumer Cyclical, Consumer Non-Cyclical, Financial, Technology, Communications, Energy, and Utilities*. Our chosen companies and their data are divided up into 9 groups by industrial sector so that we can run the same tests per industrial group.

Each test whether on all the stocks or stocks belonging to a particular industrial sector are run 100 times. In each run, the same set of training data are used to train the XGBoost model whose performance is verified using the same set of out-of-sample test data. This generates 100 sets of results for each test from which the averaged classification accuracy on the drift direction (up or down) are calculated. Separately, a Multilayer perceptron (MLP) network, as one of the most commonly seen supervised learning models, has been chosen as our benchmark network and the same training and test process is repeated on the MLP network. The MLP's hyperparameters including the *learning rate, number of hidden neurons, number of hidden layers*, and *number of epoches* have all been tuned in the same Genetic Algorithm optimizer in the same way the XGBoost model is optimized.

Fig. 3 presents the averaged classification success rate of 100 runs from each experiment using GA optimized XGBoost and MLP. The success rate is based on the accuracy of a model correctly predicting the

Table 3
Averaged classification success rates of 100 runs from 2015 to 2018 using GA optimized XGBoost and MLP. The success rate is based on the accuracy of a model correctly predicting the movement direction (up or down) of a stock's 30-day PEAD following an earning event.

	XGB			
(accuracy%)	2018	2017	2016	2015
All	57.58	58.71	61.17	59.63
Industrial	63.09	58.62	59.12	62.40
Basic Materials	56.88	60.09	55.13	61.07
Consumer Cyclical	57.02	60.47	58.47	59.22
Consumer Non-Cyclical	57.63	62.29	60.38	59.82
Financial	55.39	53.38	60.15	56.46
Technology	55.93	54.86	58.16	61.08
Communications	53.35	53.93	57.50	57.52
Energy	53.73	51.16	52.56	55.55
Utilities	49.95	49.98	46.82	46.31

	MLP			
(accuracy%)	2018	2017	2016	2015
All	55.99	57.18	60.36	58.45
Industrial	61.14	57.20	60.03	60.12
Basic Materials	54.91	56.76	54.21	57.09
Consumer Cyclical	58.31	58.83	55.74	60.67
Consumer Non-Cyclical	58.41	60.67	58.72	60.62
Financial	53.52	53.30	59.14	51.82
Technology	58.54	56.35	60.78	59.13
Communications	51.78	47.71	54.30	50.56
Energy	48.45	47.00	46.96	51.37
Utilities	49.15	42.06	49.70	47.12

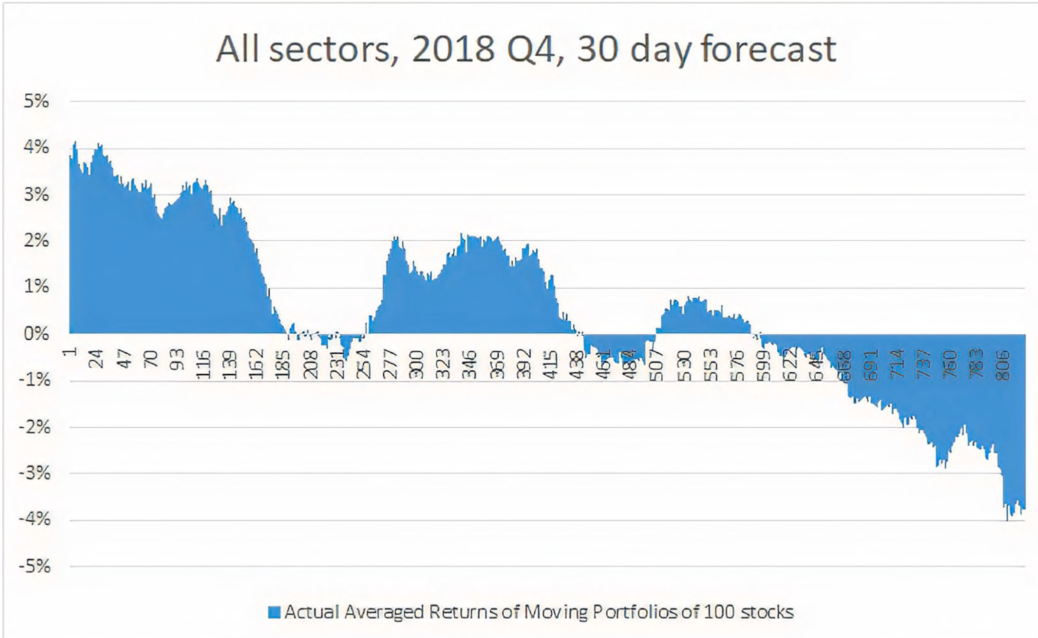

Fig. 3. 2018 Q4 test result. Actual returns of moving portfolios consisting of 100 stocks that reported earnings in this quarter. All stocks have been pre-ranked by the model-predicted stock returns from high to low.

movement direction (up or down) of a stock's 30-day PEAD following an earning event. Our results clearly suggest that the XGBoost-GA model has strong prediction power overall and performs better than the benchmark MLP model (also optimized by GA). Our model is able to pick up the patterns in the input data space when there is clear driver in it. It is particularly interesting to see the model performs with a varying degree of success going from sector to sector. Given that the same set of input features is used across the board, there is clear evidence that our data is more impactful to some sectors than others. There are probably two reasons that can explain this observation. First, there are other data that are not included in our feature space that does affect stock movements following earnings release. Second, stocks from different sectors are subject to different drivers, i. e., investment personnel/computer trading algorithms look for signals in different financial metrics for different sectors. Even when the same driving features are examined, the implicit feature weighting must be different for different sectors.

The first reason is true, as there are impactful data that have yet to be included in our research, such as management's guidance, recent revisions of analysts' price forecast, other text information carried in financial reports, and meeting minutes with analysts, etc. It is entirely possible that certain stocks, or stocks from certain sectors are more susceptible to those data and their absence reduces the model's prediction accuracy on such stocks' reaction to their earning events.

We have taken a closer look at the second possible cause and we find indeed stocks from different sectors are driven by different factors. With the 100 tests carried out on each individual sector in each financial year from 2015 to 2018, we have counted the appearance of those factors that appear most often as the top five driving forces as determined by XGBoost. The results are recorded in the Appendix section of this paper and present some very interesting findings. First, most of the time, it is the three EPS related metrics that feature heavily on the top five spots of most influential factors. This finding is consistent with market practice and Earning Per Share surprise/disappointment is indeed one of the most important factors that investors examine. We need to point out that two of these features are engineered by us, which represent how the current quarter's reported EPS compares to those at the preceding quarters. The fact that these two factors also dominate shows that investors look for more complex movements in financial metrics than simply the reported numbers. Second, over the years, we consistently see important albeit less strong features appearing on the top five list for some of the sectors. For instance, the quarterly change in Return On Assets, Price-to-Sales Ratio, and Dividend Payout Ratio are consistently making up the top five spots driving PEAD of stocks from the *Industrial*, *Financial*, and *Basic Materials* sectors respectively. Separately, we carried out specific forecasting experiments using the three EPS features only but did not obtain good results which shows that the model cannot be driven purely by a handful of key features and other, less strong, but also impactful features must not be omitted. Third, in the years when our model produces better prediction results for a particular sector, we are frequently seeing features that are more consistently dominant. This is represented by higher occurrence counts observed for the dominant features. This can be observed in the results for the *Industrial*, *Consumer Cyclical*, and *Consumer Non-Cyclical* sectors.

The opposite is also true. With *Energy* and *Utilities* being the most difficult sectors to predict, the model is returning an inconsistent set of top drivers from the 4 yearly tests among which the occurrence counts are also comparatively lower. Without strong and consistent drivers among our feature data for such sectors, the prediction result is unsurprisingly poorer.

The results of our experiments in this section help us conclude that Post Earnings-Announcement Drift is not merely a market anomaly, but a characteristics of the markets whose direction can be materially predicted. The strength of signal may vary in time and from sector to sector, but machine learning models – especially an XGBoost well optimised by GA – are able to pick up on them. However, the fact that the model performs well with stocks from certain sectors but not on others suggest that there may be limitations in our input feature space or the way the special features have been engineered. This is may be more true with sectors such as *Energy* and *Utilities*. We acknowledge that, when a company makes an earnings announcement, information come out in many different forms such as in financial metric numbers, textual information embedded in the documents filed with SEC, earnings calls with a selected group of equity analysts, let alone information leakage prior to announcement or even insider trading. Trying to capture and take advantage of more forms of drivers on Post Earnings Announcement Drifts is a future research topic.

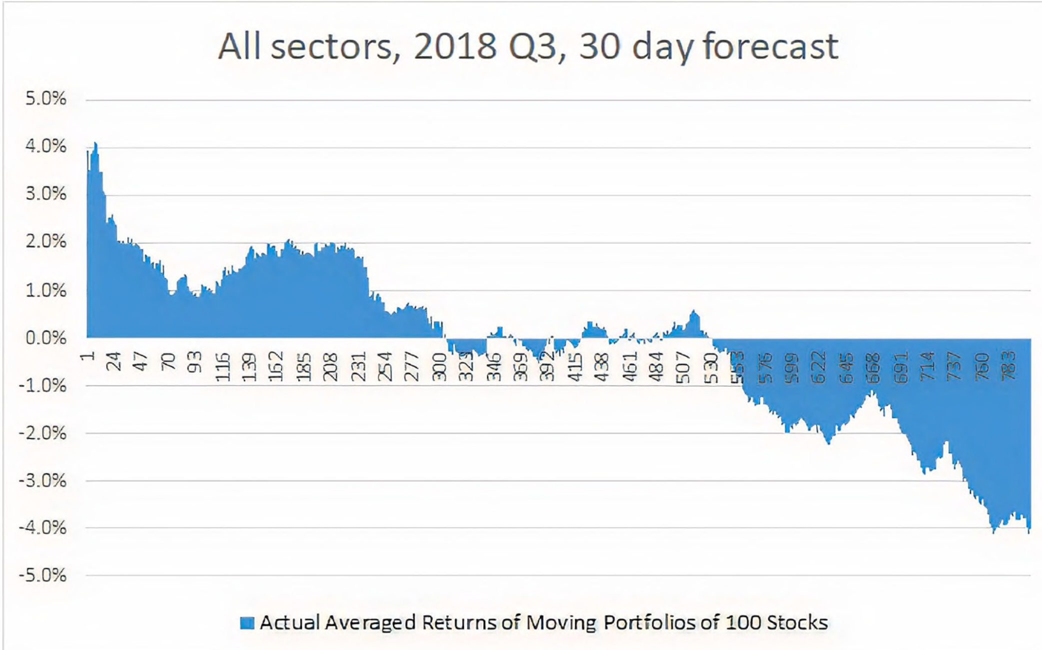

Fig. 4. 2018 Q3 test result. Actual returns of moving portfolios consisting of 100 stocks. All stocks have been pre-ranked by the model-predicted stock returns from high to low.

6.2. *PEAD analysis on a portfolio of stocks*

We believe it is meaningful to evaluate the dynamics of post-earnings drift in the context of portfolios. In the subsequent series of tests, we use the model to forecast the actual level of 30-day post earning cumulative abnormal returns (regression) instead of only the movement direction. In this case, we rank out-of-sample stocks by each stock's *model-predicted returns* from high to low, group stocks from the same quantile into small portfolios and empirically examine the *actual returns* of the portfolios.

6.2.1. *Stocks that reported earnings in the same financial quarter*

We examine stocks that reported earnings in the same financial quarter in the years between 2015 and 2018. In each test we use all the data prior to the test quarter for training and carry out stock return prediction on stocks in the test quarter. Once the stocks have been ranked by their model-predicted post earnings returns, we are consistently observing that portfolios which include stocks from top quantiles of the ranked list are producing higher positive returns, whereas portfolios which include stocks from bottom quantiles are producing lower negative returns. Theoretically, a long-short market neutral strategy (Solberg & Karlsen, 2017) could be formed through longing the top

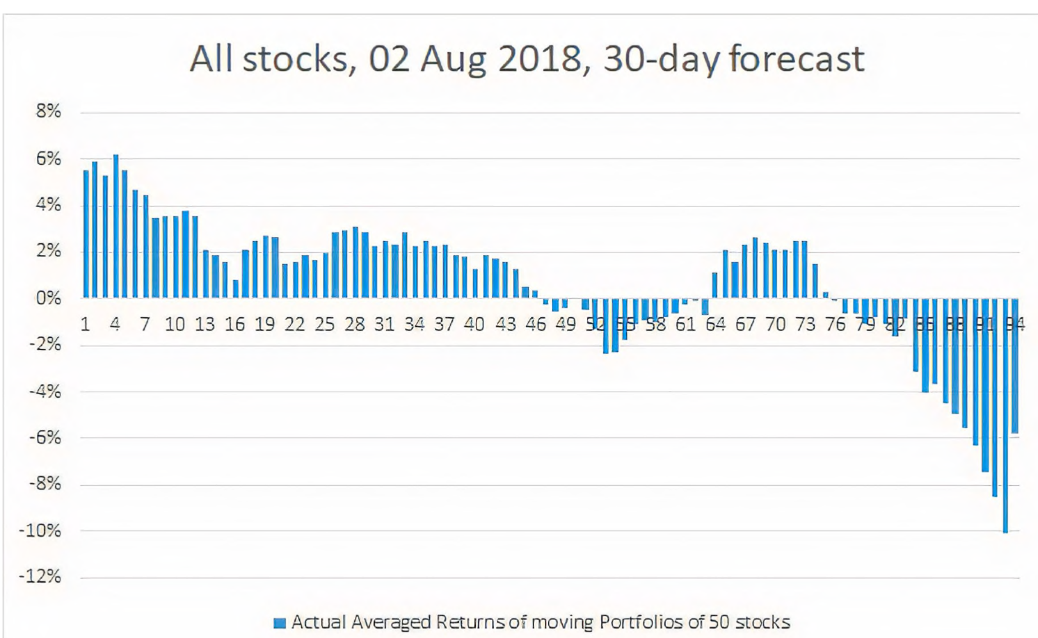

Fig. 5. 02 Aug 2018. Actual returns of moving portfolios consisting of 50 stocks from all sectors that reported earnings on this date. All stocks have been pre-ranked by the model-predicted stock returns from high to low. Portfolio returns are empirically calculated.

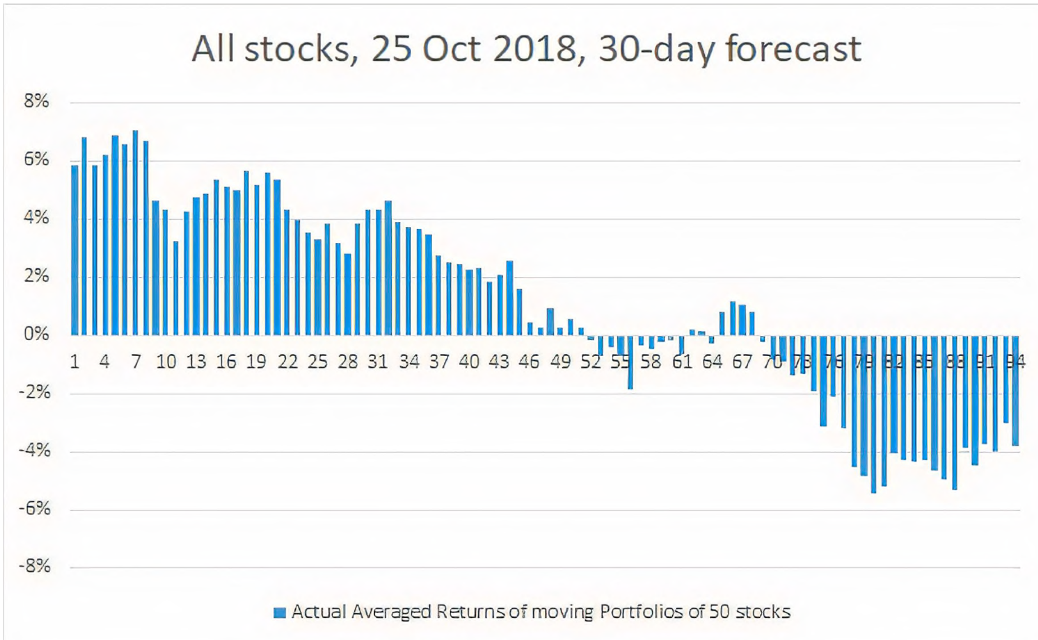

Fig. 6. 25 Oct 2018. Actual returns of moving portfolios consisting of 50 stocks from all sectors that reported earnings on this date. All stocks have been pre-ranked by the model-predicted stock returns from high to low. Portfolio returns are empirically calculated.

quantile portfolios and shorting the bottom ones.

We've selected the Q3 2018 and Q4 2018 earning seasons to demonstrate our findings. One of the reasons for choosing these two quarters is that the US stock market went through two polar opposite phases of development in these two quarters with the S&P500 shedding 20% in the last quarter of 2018 (around the time most Q3 2018 earnings were reported in the US) amid fear of Fed rate rises and US-China trade war escalation among other things but gaining a major rebound in the first quarter of 2019 (when most US companies reported Q4 2018 earnings). Our intention is to evaluate if our model can successfully capture those very different PEAD dynamics given very different macro conditions and different company specific accounts.

Each point on Figs. 3 and 4 corresponds to the empirically calculated return of a portfolio consisting of 100 stocks when we move down the list of out-of-sample stocks which have now been ranked by their predicted 30-day risk-adjusted returns following earnings releases. For instance, the first point is the actual 30-day Cumulative Abnormal Return of a portfolio consisting of the 1st to the 100th stocks and the second point is the actual return of a portfolio including the 2nd to the 101st stocks, etc. We consistently produce similar figures with a downward slope when we run the same tests over many times. We conclude that our model has captured an unseen *collective* trend of movement in such groups of stocks as triggered by their earnings release and other relevant economic factors.

6.2.2. *Stocks that reported earnings on the same date*

If we were to construct market neutral portfolios, practically speaking it would only make sense if we could execute the buying and short-selling of model-chosen stocks within a short time frame, such as within a day or ideally less. Therefore, we run the same portfolio test on stocks which filed for earnings with SEC on the same date. Two dates in 2018 with busy earnings release activities were chosen for demonstration. Figs. 5 and 6 are plots of portfolio returns on 02 Aug 2018 and 25 Oct 2018 respectively. The stocks have been ranked by their model-predicted 30-day post earning CAR before being grouped into quantile portfolios. The same as shown in Figs. 3 and 4, the combination of the XGBoost-GA model and our engineered input features is producing the kind of results which can be used to rank stocks and construct portfolios, which would produce higher positive returns or lower negative returns.

Table 4
Here we present the averaged 30-day Cumulative Abnormal Return of full portfolios under test, as well as the return of portfolios consisting of only the top/bottom quantile group of 100 stocks. Returns from such quantile groups are consistently higher/lower than the whole portfolio's average. All returns listed here are empirically calculated. Stocks have been ranked by their model-predicted returns so as to produce the quantile groups.

Out-of-sample Time Frame	Industries	Forecast Holding Period	Average Return of all out-of-sample stocks	Return of Top Quantile Portfolio	Return of Bottom Quantile Portfolio
Q4 2018	All	30 days	0.45%	3.94%	−3.91%
Q3 2018	All	30 days	1.09%	4.00%	−4.01%
02-Aug-18	All	30 days	1.94%	5.81%	−10.01%
25-Oct-18	All	30 days	1.15%	5.94%	−3.85%

Table 4 gives the stats on how top quantile portfolios and bottom quantile portfolios are performing compared against the average return of all the out-of-sample stocks. In some cases, returns from portfolios consisting of top/bottom quantile stocks are considerably higher/lower than the out-of-sample stock population's average. Such patterns of portfolio returns could have theoretically made them good candidates for a long-short strategy capitalising on the events of earnings release. This, however, is made difficult in reality, as we will discuss more about the timeliness of the signals.

6.3. *Trading on Earning Event Signals*

In the aforementioned experiments we have chosen the last publicly available tradable stock price before an earnings release as the starting point of a 30 day holding period. This is an intuitive choice and commonly seen in the literature. For instance, Erlien (Erlien, 2011) uses the end point of her training window as the beginning of a calculation window for cumulative abnormal returns. Similarly, when examining how numerous factors drives the revision of analysts' consensus forecast on a company's EPS, Ahmed and Irfan (Ahmed & Safdar, 2018) collect the final consensus available prior to earnings announcement to start the forecast period.

However in reality a company's stock price moves on receipt of the first trickle of news. Information is never symmetrical, and some parties always possess greater material knowledge than others. They can and will act on such material information driving the stock price away from the last tradable price before the wider market gains access to the same level of information. Also, the incorporation of earnings information into the latest price is hugely accelerated by the presence of algorithmic trading systems as verified by Frino et al. who studied a unique dataset obtained from the Australian Securities Exchange (Frino, Prodromou, Wang, Westerholm, & Zheng, 2017). Correct forecasting of stock movements upon financial events is not practically useful unless they can be acted upon.

With this in mind, we have attempted to forecast cumulative abnormal returns from 1 day *after* the announcement of news to 30 day after, i. e., CAR from t_1 to t_{30}. Our results show that the forecasting is inferior with accuracy of around 50% and sometimes less and cannot be relied upon. This is not at all surprising, because, as per the efficient market hypothesis any granular earnings information embedded in the financial statements and management's guidance, coupled with the market's own interpretations, will have been mostly consumed by the markets and reflected in the latest stock prices not too long after the announcement. A similar observation was already seen by Allen and Karjalainen (Allen & Karjalainen, 1999) that introducing a one-day delay to trading signals removes most of the forecasting ability when they used GAs to find technical trading rules.

Since we are not able to accurately forecast the direction of CAR from $t_{t+\Delta t}$ to t_{t+T} (with Δt being non-negligible) using newly released financial statements data and a stock's momentum signals prior to announcement, we have devised a tactic to infer a stock's forward movement direction from a delayed-starting position, such as 1 day after an earning event. It is important to point out that since we intend to trade on a stock's movement from t_1 to t_{30}, our standpoint is now 1 day after the earnings announcement, and we are already in possession of the knowledge of how a stock has moved from t_0 to t_1. Standing near the market close 1 day after announcement, we have devised the steps below to infer a stock's movement from t_1 to t_{30}:

1. Stock Exclusion: Exclude all the stocks whose actual movement from t_0 to t_1 are within the interval of [−0.05%, 0.05%] (obtained through empirical analysis) so as to eliminate stocks with weak immediate response to their earnings announcements;
2. Train the model using the remaining stocks in the training set and include in the model input space each stock's known movement direction from t_0 to t_1. Use the trained model to forecast PEAD direction from t_0 to t_{30} on stocks in the test set. The overall up/down classification accuracy has increased to around 70% due to additional inputs to the model as well as having more responsive stocks in the population;
3. Filtering: Select stocks whose real movement from t_0 to t_1 is in opposite direction compared with the model-predicted movement from t_0 to t_{30}, i. e., select stocks which have in reality gone up (down) in the first day despite being forecasted to go down (up) from t_0 to t_{30};
4. **For all the remaining stocks**, deduce a stock's movement direction from t_1 to t_{30} to be the same as the predicted movement direction from t_0 to t_{30}.

We test this tactic using data from 2016, 2017, and 2018, respectively. Data from preceding years are used for training. As noted in an earlier section, a company that filed in each of the four quarters of a financial year is considered as four independent data points. Table 5 gives the results of applying this tactic on the three test years. After stock exclusion and filtering the number of eligible stocks have come down to lower hundreds. With the remaining stocks we observe that we are consistently achieving close to 60% classification accuracy in inferring a stock's movement direction (up or down) from t_1 to t_{30} through empirical study. The crucial thing is that, since this tactic is meant to be exercised by a trader at or near the close of market 1 day after an earnings announcement, this is a signal that can genuinely be acted upon. We also expect the overall accuracy of inferring a stock price's drift direction from $t_{t+\Delta t}$ to t_{t+T} to increase once we are able to further increase the prediction accuracy by the model on the direction of post-earnings announcement drift. We believe this is very likely, as there are other sources of impactful information that have yet to be included in the feature space, such as management's guidance, equity analyst's price revisions, other text data carried in financial reports, and meeting minutes with analysts, etc. This is another potential future research direction.

Table 5
Accuracy of inferring a stock's movement direction (up or down) from t_1 to t_{30} using (a) predicted movement direction from t_0 to t_{30} by the model and (b) the known stock movement from t_0 to t_1. The accuracy is empirically calculated by comparing the inferred movement direction against the actual movement direction of each out-of-sample stock under test, from t_1 to t_{30}.

Year Tested	No. of initial out-of-sample points	No. of points after stock exclusion	No. of points after filtering	Classification accuracy on the remaining points through inferring
2016	3635	3231	254	58.57%
2017	3715	3377	151	60.86%
2018	3728	3387	261	57.53%

7. Conclusion

Post-earnings announcement drift is a well known and well studied stock market anomaly when a stock's risk adjusted price can continue in the direction of an earnings surprise in the near to mid term following an earnings release. Past research was, however, often limited in using simpler regression based methods to explain this phenomenon, and was often confined to using a limited set of explaining factors. Even fewer research was carried out on how to potentially take advantage of this known anomaly and conduct *actionable forecasting* on stock price movements following such a significant economic event. Attempting to fill this gap in the literature, our experiment is including a much bigger set of carefully selected input factors of various types with some being specifically engineered, sourcing the data over a longer historical time frame and attempting to forecast both the level (regression) and the direction (classification) of Cumulative Abnormal Returns (CAR) with a machine learning approach. We have adopted the state-of-the-art XGBoost and put it through a rigorous optimisation process. We not only looked at specific success rate of forecasting drift direction, but also examined if there is a *collective* trend of movement enjoyed by a group of stocks following their individual earnings release.

First, our results show that when properly configured using a Generic Algorithm, XGBoost produces meaningful prediction accuracy on the direction of PEAD. We demonstrated that our selected input features were genuinely driving PEAD with a classification success rate going up to 63% depending on the test scenarios. In a further breakdown, we observed that stocks from different industrial sectors and at a different time can have their PEADs driven by different primary factors. The strengths of the driving factors are well understood by our model with stocks from certain sectors producing excellent/poor forecast results when the underlying factor dominance is more/less pronounced.

Second, guided by the model's forecast outputs we found that it is possible to build portfolios which consistently offer higher positive returns and lower negative returns and such an observation could potentially form the basis of market neutral long-short trading strategy.

Third, we studied the challenges of applying earning event signals in real trading. Market participants with information advantage can drive the price away before the rest of the markets have an opportunity to act on the signals. Instead of trying to buy in as soon as event data comes out, we have devised a tactic to create opportunities to delay-buy into the market at a later time using the same prediction results by the models as well as public knowledge of market movements immediately following the release of earnings data.

Lastly, future efforts will need to also investigate recent methods of deep learning, which, in our preliminary experiments were inferior to the considered approach. However, their partial or combined usage such as with representation learning (Xie, Gao, Nijkamp, Zhu, & Wu, 2020) or data augmentation appears promising.

CRediT authorship contribution statement

Zhengxin Joseph Ye: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Björn W. Schuller:** Validation, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

We are reporting the most significant driving factors as calculated by

Table 6
Top five driving factors for all stocks in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	90	21%	F1	EPS_EarningsSurprise	4	23%
F2	EPS_EarningsSurprise	62	14%	F2	EPS_Earnings_Surprise_Backward_Diff	24	14%
F3	Total_Liabilities_Q_Change	38	10%	F3	EPS_EarningsSurprise	24	11%
F4	Total_Liabilities_Q_Change	38	10%	F4	Return_On_Common_Equity	26	10%
F5	Operating_Income_Y_Change	26	9%	F5	Return_On_Common_Equity	24	9%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	90	23%	F1	EPS_EarningsSurprise	8	16%
F2	EPS_EarningsSurprise	86	14%	F2	EPS_Earnings_Surprise_Backward_Diff	8	15%
F3	EPS_Earnings_Surprise_Backward_Diff	60	11%	F3	Total_Liabilities_Q_Change	20	10%
F4	Total_Liabilities_Q_Change	38	10%	F4	Return_On_Common_Equity	26	9%
F5	Operating_Income_Y_Change	28	9%	F5	Total_Liabilities_Q_Change	26	8%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	72	21%	F1	EPS_EarningsSurprise	26	21%
F2	EPS_EarningsSurprise	64	15%	F2	EPS_Earnings_Surprise_Backward_Ave_Diff	28	16%
F3	EPS_Earnings_Surprise_Backward_Diff	48	11%	F3	Total_Liabilities_Q_Change	28	10%
F4	Total_Liabilities_Q_Change	46	9%	F4	Return_On_Common_Equity	16	9%
F5	Operating_Income_Y_Change	34	8%	F5	Return_On_Common_Equity	20	8%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	54	17%	F1	EPS_EarningsSurprise	36	17%
F2	EPS_EarningsSurprise	58	17%	F2	EPS_Earnings_Surprise_Backward_Ave_Diff	18	17%
F3	Total_Liabilities_Q_Change	56	11%	F3	EPS_Earnings_Surprise_Backward_Diff	20	12%
F4	Total_Liabilities_Q_Change	34	10%	F4	EPS_Earnings_Surprise_Backward_Diff	32	9%
F5	Operating_Income_Y_Change	50	8%	F5	EPS_Earnings_Surprise_Backward_Diff	12	8%

Table 7
Top five driving factors for stocks in the **Financial** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	95	20%	F1	EPS_Earnings_Surprise_Backward_Diff	5	23%
F2	EPS_Earnings_Surprise_Backward_Diff	62	13%	F2	EPS_EarningsSurprise	11	13%
F3	EPS_EarningsSurprise	28	11%	F3	EPS_EarningsSurprise	28	11%
F4	PS_Ratios	29	9%	F4	EPS_EarningsSurprise	22	9%
F5	PS_Ratios	19	8%	F5	PS_Ratios_Y_Change	14	9%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Diff	52	18%	F1	EPS_Earnings_Surprise_Backward_Ave_Diff	46	19%
F2	EPS_Earnings_Surprise_Backward_Ave_Diff	47	14%	F2	EPS_Earnings_Surprise_Backward_Diff	34	14%
F3	PS_Ratios	31	10%	F3	EPS_EarningsSurprise	27	11%
F4	PS_Ratios	37	10%	F4	PS_Ratios_Y_Change	13	10%
F5	EPS_EarningsSurprise	13	9%	F5	PS_Ratios	11	9%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Diff	50	21%	F1	EPS_Earnings_Surprise_Backward_Ave_Diff	47	18%
F2	EPS_Earnings_Surprise_Backward_Diff	43	15%	F2	EPS_Earnings_Surprise_Backward_Ave_Diff	35	13%
F3	PS_Ratios	33	9%	F3	EPS_EarningsSurprise	27	10%
F4	PS_Ratios	38	9%	F4	EPS_EarningsSurprise	18	9%
F5	EPS_EarningsSurprise	15	9%	F5	PS_Ratios	11	9%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	92	21%	F1	EPS_Earnings_Surprise_Backward_Diff	6	18%
F2	EPS_Earnings_Surprise_Backward_Diff	78	14%	F2	EPS_Earnings_Surprise_Backward_Ave_Diff	8	15%
F3	EPS_EarningsSurprise	49	10%	F3	PS_Ratios_Y_Change	19	9%
F4	PS_Ratios	31	9%	F4	PS_Ratios_Y_Change	24	9%
F5	PS_Ratios_Y_Change	28	8%	F5	PS_Ratios	21	9%

Table 8
Top five driving factors for stocks in the **Industrial** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	94	20%	F1	EPS_Earnings_Surprise_Backward_Diff	6	25%
F2	EPS_Earnings_Surprise_Backward_Diff	67	14%	F2	EPS_EarningsSurprise	20	12%
F3	EPS_EarningsSurprise	64	11%	F3	EPS_Earnings_Surprise_Backward_Diff	19	11%
F4	Free_Cash_Flow_Q_Change	55	9%	F4	Return_On_Assets_Q_Change	9	9%
F5	Return_On_Assets_Q_Change	20	8%	F5	Return_On_Assets_Y_Change	12	8%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	52	22%	F1	EPS_Earnings_Surprise_Backward_Diff	39	20%
F2	EPS_Earnings_Surprise_Backward_Diff	52	16%	F2	EPS_Earnings_Surprise_Backward_Ave_Diff	34	16%
F3	EPS_EarningsSurprise	66	12%	F3	EPS_Earnings_Surprise_Backward_Ave_Diff	9	12%
F4	Return_On_Assets_Q_Change	36	9%	F4	PC_Ratios_Y_Change	16	9%
F5	Return_On_Assets_Q_Change	24	8%	F5	RSI-30D	11	8%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	64	24%	F1	EPS_Earnings_Surprise_Backward_Diff	24	20%
F2	EPS_Earnings_Surprise_Backward_Diff	51	14%	F2	EPS_Earnings_Surprise_Backward_Ave_Diff	22	15%
F3	EPS_EarningsSurprise	46	11%	F3	Return_On_Assets_Q_Change	18	10%
F4	Return_On_Assets_Q_Change	27	9%	F4	EPS_EarningsSurprise	14	9%
F5	Return_On_Assets_Q_Change	23	8%	F5	Free_Cash_Flow_Q_Change	20	8%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	72	19%	F1	EPS_Earnings_Surprise_Backward_Diff	15	17%
F2	EPS_Earnings_Surprise_Backward_Diff	39	14%	F2	EPS_EarningsSurprise	36	13%
F3	EPS_EarningsSurprise	46	12%	F3	EPS_Earnings_Surprise_Backward_Diff	31	11%
F4	Free_Cash_Flow_Q_Change	27	9%	F4	Return_On_Assets_Y_Change	14	9%
F5	Return_On_Assets_Y_Change	19	8%	F5	PC_Ratios_Y_Change	15	9%

Table 9
Top five driving factors for stocks in the **Basic Materials** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	36	13%	F1	Total_Liabilities_Q_Change	11	15%
F2	EPS_Earnings_Surprise_Backward_Ave_Diff	17	12%	F2	Dividend_Payout_Ratio	12	11%
F3	EPS_Earnings_Surprise_Backward_Ave_Diff	10	11%	F3	DMA_50D/200D	9	11%
F4	EPS_EarningsSurprise	9	10%	F4	EPS_EarningsSurprise	9	10%
F5	EPS_EarningsSurprise	8	10%	F5	PC_Ratios	6	10%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	Dividend_Payout_Ratio	31	13%	F1	DMA_50D/200D	9	12%
F2	Dividend_Payout_Ratio	14	11%	F2	EPS_Earnings_Surprise_Backward_Ave_Diff	8	12%
F3	DMA_50D/200D	11	10%	F3	Dividend_Payout_Ratio	10	11%
F4	EPS_Earnings_Surprise_Backward_Ave_Diff	10	10%	F4	DMA_50D/200D	6	10%
F5	PE_Ratios_Q_Change	8	10%	F5	EPS_Earnings_Surprise_Backward_Ave_Diff	6	10%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_EarningsSurprise	15	14%	F1	EPS_EarningsSurprise	15	14%
F2	EPS_Earnings_Surprise_Backward_Diff	13	12%	F2	Cost_Of_Revenue_Q_Change	10	12%
F3	Inventory_Turnover	8	11%	F3	Cost_Of_Revenue_Q_Change	7	11%
F4	EPS_Earnings_Surprise_Backward_Diff	7	10%	F4	EPS_Earnings_Surprise_Backward_Diff	7	10%
F5	EPS_Earnings_Surprise_Backward_Diff	10	10%	F5	EPS_Earnings_Surprise_Backward_Ave_Diff	7	10%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	41	13%	F1	EPS_Earnings_Surprise_Backward_Diff	6	12%
F2	EPS_Earnings_Surprise_Backward_Diff	16	12%	F2	EPS_Earnings_Surprise_Backward_Ave_Diff	13	11%
F3	EPS_Earnings_Surprise_Backward_Diff	10	11%	F3	Dividend_Payout_Ratio	7	11%
F4	EPS_Earnings_Surprise_Backward_Diff	14	10%	F4	Dividend_Payout_Ratio	9	10%
F5	Dividend_Payout_Ratio	11	10%	F5	Total_Liabilities_Q_Change	7	10%

Table 10
Top five driving factors for stocks in the **Cyclical** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	EPS_EarningsSurprise	65	16%	**F1**	EPS_Earnings_Surprise_Backward_Ave_Diff	21	18%
F2	EPS_Earnings_Surprise_Backward_Ave_Diff	41	13%	**F2**	EPS_EarningsSurprise	20	14%
F3	Return_On_Common_Equity	29	12%	**F3**	Return_On_Common_Equity	29	12%
F4	Return_On_Common_Equity	31	10%	**F4**	EPS_Earnings_Surprise_Backward_Diff	26	10%
F5	Net_Income_Y_Change	35	9%	**F5**	Return_On_Common_Equity	14	9%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	70	16%	**F1**	EPS_EarningsSurprise	19	15%
F2	EPS_EarningsSurprise	36	12%	**F2**	EPS_Earnings_Surprise_Backward_Ave_Diff	22	13%
F3	Return_On_Common_Equity	30	11%	**F3**	EPS_Earnings_Surprise_Backward_Diff	21	11%
F4	Return_On_Common_Equity	32	10%	**F4**	EPS_Earnings_Surprise_Backward_Diff	11	10%
F5	Total_Liabilities_Q_Change	11	10%	**F5**	Net_Income_Y_Change	10	9%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_EarningsSurprise	47	16%	**F1**	EPS_Earnings_Surprise_Backward_Diff	25	16%
F2	EPS_Earnings_Surprise_Backward_Diff	35	13%	**F2**	EPS_EarningsSurprise	33	14%
F3	EPS_Earnings_Surprise_Backward_Diff	25	12%	**F3**	EPS_Earnings_Surprise_Backward_Ave_Diff	21	12%
F4	Return_On_Common_Equity	40	10%	**F4**	Net_Income_Y_Change	11	10%
F5	Net_Income_Y_Change	19	9%	**F5**	Return_On_Common_Equity	12	10%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_EarningsSurprise	61	15%	**F1**	EPS_Earnings_Surprise_Backward_Ave_Diff	16	16%
F2	Return_On_Common_Equity	38	12%	**F2**	EPS_EarningsSurprise	20	14%
F3	EPS_Earnings_Surprise_Backward_Diff	29	12%	**F3**	Return_On_Common_Equity	21	12%
F4	Net_Income_Y_Change	27	10%	**F4**	Return_On_Common_Equity	19	11%
F5	Net_Income_Y_Change	32	9%	**F5**	PE_Ratios	12	9%

Table 11
Top five driving factors for stocks in the **Non Cyclical** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	EPS_Earnings_Surprise_Backward_Diff	41	16%	**F1**	EPS_Earnings_Surprise_Backward_Ave_Diff	28	16%
F2	EPS_Earnings_Surprise_Backward_Ave_Diff	27	13%	**F2**	EPS_EarningsSurprise	26	13%
F3	EPS_EarningsSurprise	32	11%	**F3**	EPS_Earnings_Surprise_Backward_Ave_Diff	15	12%
F4	Return_On_Common_Equity	19	10%	**F4**	EPS_EarningsSurprise	10	10%
F5	Return_On_Common_Equity	12	9%	**F5**	Operating_Margin_Y_Change	7	9%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	51	17%	**F1**	EPS_Earnings_Surprise_Backward_Diff	34	18%
F2	EPS_Earnings_Surprise_Backward_Diff	35	15%	**F2**	EPS_Earnings_Surprise_Backward_Ave_Diff	34	15%
F3	EPS_EarningsSurprise	53	12%	**F3**	EPS_Earnings_Surprise_Backward_Diff	17	11%
F4	Operating_Income_Y_Change	22	10%	**F4**	EPS_EarningsSurprise	9	10%
F5	Operating_Income_Y_Change	14	9%	**F5**	Return_On_Common_Equity	13	8%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	52	18%	**F1**	EPS_Earnings_Surprise_Backward_Diff	35	19%
F2	EPS_Earnings_Surprise_Backward_Diff	45	15%	**F2**	EPS_Earnings_Surprise_Backward_Ave_Diff	34	15%
F3	EPS_EarningsSurprise	46	11%	**F3**	EPS_Earnings_Surprise_Backward_Ave_Diff	10	12%
F4	Return_On_Common_Equity	21	9%	**F4**	EPS_EarningsSurprise	7	9%
F5	Return_On_Common_Equity	15	9%	**F5**	Gross_Profit_Y_Change	13	8%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_EarningsSurprise	67	15%	**F1**	EPS_Earnings_Surprise_Backward_Ave_Diff	23	17%
F2	EPS_Earnings_Surprise_Backward_Ave_Diff	37	13%	**F2**	EPS_EarningsSurprise	22	13%
F3	EPS_Earnings_Surprise_Backward_Diff	47	11%	**F3**	EPS_Earnings_Surprise_Backward_Ave_Diff	10	11%
F4	Return_On_Common_Equity	12	9%	**F4**	Total_Liabilities_Q_Change	9	9%
F5	EPS_Earnings_Surprise_Backward_Ave_Diff	8	10%	**F5**	EPS_Earnings_Surprise_Backward_Ave_Diff	8	10%

Table 12

Top five driving factors for stocks in the **Technology** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	57	14%	**F1**	EPS_Earnings_Surprise_Backward_Diff	14	14%
F2	EPS_EarningsSurprise	26	11%	**F2**	EPS_Earnings_Surprise_Backward_Ave_Diff	18	12%
F3	Return_On_Assets_Q_Change	24	11%	**F3**	EPS_Earnings_Surprise_Backward_Diff	16	11%
F4	Return_On_Assets_Q_Change	18	10%	**F4**	EPS_Earnings_Surprise_Backward_Diff	11	10%
F5	Return_On_Assets_Q_Change	12	10%	**F5**	EPS_EarningsSurprise	11	10%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	53	13%	**F1**	EPS_Earnings_Surprise_Backward_Diff	15	16%
F2	EPS_EarningsSurprise	37	11%	**F2**	EPS_Earnings_Surprise_Backward_Ave_Diff	17	12%
F3	Return_On_Assets_Q_Change	13	10%	**F3**	EPS_EarningsSurprise	11	11%
F4	Return_On_Assets_Q_Change	18	10%	**F4**	EPS_EarningsSurprise	12	10%
F5	Dividend_Payout_Ratio_Y_Change	9	10%	**F5**	Return_On_Assets_Q_Change	7	10%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_EarningsSurprise	31	13%	**F1**	EPS_Earnings_Surprise_Backward_Ave_Diff	21	12%
F2	EPS_EarningsSurprise	22	11%	**F2**	EPS_Earnings_Surprise_Backward_Ave_Diff	14	11%
F3	EPS_EarningsSurprise	18	11%	**F3**	EPS_Earnings_Surprise_Backward_Diff	11	11%
F4	Operating_Income_Y_Change	11	10%	**F4**	Short_Term_Debt_Q_Change	10	10%
F5	Operating_Income_Y_Change	15	10%	**F5**	EPS_Earnings_Surprise_Backward_Diff	8	10%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	61	14%	**F1**	EPS_EarningsSurprise	11	14%
F2	EPS_EarningsSurprise	18	11%	**F2**	EPS_Earnings_Surprise_Backward_Diff	16	11%
F3	EPS_EarningsSurprise	19	11%	**F3**	EPS_Earnings_Surprise_Backward_Diff	11	11%
F4	EPS_EarningsSurprise	12	10%	**F4**	EPS_EarningsSurprise	12	10%
F5	EPS_EarningsSurprise	9	10%	**F5**	EPS_EarningsSurprise	9	10%

Table 13

Top five driving factors for stocks in the **Communications** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	Total_Liabilities_Q_Change	27	12%	**F1**	Net_Income_Y_Change	18	12%
F2	Total_Liabilities_Q_Change	26	11%	**F2**	Net_Income_Y_Change	18	11%
F3	Operating_Income_Y_Change	15	11%	**F3**	Total_Liabilities_Q_Change	13	11%
F4	EPS_Earnings_Surprise_Backward_Ave_Diff	11	10%	**F4**	PC_Ratios	6	10%
F5	Net_Income_Y_Change	13	10%	**F5**	Operating_Income_Y_Change	9	10%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	PE_Ratios	37	12%	**F1**	PB_Ratios_Y_Change	8	13%
F2	Net_Income_Y_Change	17	11%	**F2**	PE_Ratios	14	11%
F3	PB_Ratios_Y_Change	12	11%	**F3**	PB_Ratios_Y_Change	12	11%
F4	Net_Income_Y_Change	17	10%	**F4**	Cost_Of_Revenue_Y_Change	9	10%
F5	Total_Liabilities_Q_Change	9	10%	**F5**	EPS_Earnings_Surprise_Backward_Ave_Diff	8	10%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	Total_Liabilities_Q_Change	21	12%	**F1**	PS_Ratios_Y_Change	12	12%
F2	Total_Liabilities_Q_Change	11	11%	**F2**	PS_Ratios_Y_Change	10	11%
F3	Total_Liabilities_Q_Change	12	11%	**F3**	EPS_Earnings_Surprise_Backward_Ave_Diff	10	11%
F4	Total_Liabilities_Q_Change	8	10%	**F4**	PS_Ratios_Y_Change	7	10%
F5	Total_Liabilities_Q_Change	13	10%	**F5**	EPS_Earnings_Surprise_Backward_Ave_Diff	8	10%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	Total_Liabilities_Q_Change	18	12%	**F1**	PE_Ratios	12	12%
F2	Total_Liabilities_Q_Change	14	11%	**F2**	EPS_Earnings_Surprise_Backward_Ave_Diff	11	11%
F3	Operating_Income_Y_Change	10	11%	**F3**	PE_Ratios	8	11%
F4	Total_Liabilities_Q_Change	11	10%	**F4**	EPS_Earnings_Surprise_Backward_Ave_Diff	8	10%
F5	Total_Liabilities_Q_Change	14	10%	**F5**	PE_Ratios	5	10%

Table 14
Top five driving factors for stocks in the **Energy** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	EPS_EarningsSurprise	33	13%	**F1**	DMA_50D/200D	15	12%
F2	PS_Ratios	15	11%	**F2**	EPS_EarningsSurprise	13	11%
F3	EPS_EarningsSurprise	12	10%	**F3**	DMA_50D/200D	8	10%
F4	DMA_50D/200D	9	10%	**F4**	PC_Ratios_Q_Change	8	10%
F5	EPS_EarningsSurprise	6	10%	**F5**	EPS_EarningsSurprise	6	10%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_EarningsSurprise	34	12%	**F1**	DMA_5D/200D	8	11%
F2	EPS_EarningsSurprise	12	11%	**F2**	EPS_EarningsSurprise	12	11%
F3	Current_Ratio	12	11%	**F3**	Return_On_Assets_Y_Change	10	10%
F4	Return_On_Assets_Y_Change	7	10%	**F4**	EPS_EarningsSurprise	6	10%
F5	DMA_50D/200D	9	10%	**F5**	Current_Ratio	8	10%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	DMA_50D/200D	24	11%	**F1**	DMA_5D/200D	13	12%
F2	DMA_50D/200D	15	11%	**F2**	PS_Ratios	14	11%
F3	DMA_50D/200D	13	11%	**F3**	DMA_50D/200D	13	11%
F4	EPS_Earnings_Surprise_Backward_Ave_Diff	18	10%	**F4**	DMA_50D/200D	12	10%
F5	DMA_50D/200D	11	10%	**F5**	PS_Ratios	9	10%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	EPS_Earnings_Surprise_Backward_Ave_Diff	69	12%	**F1**	Gross_Profit_Q_Change	5	12%
F2	DMA_5D/200D	16	11%	**F2**	EPS_Earnings_Surprise_Backward_Ave_Diff	11	11%
F3	PS_Ratios	14	10%	**F3**	EPS_Earnings_Surprise_Backward_Ave_Diff	11	11%
F4	DMA_5D/200D	9	10%	**F4**	PS_Ratios	7	10%
F5	DMA_5D/200D	12	10%	**F5**	PS_Ratios	11	10%

Table 15
Top five driving factors for stocks in the **Utilities** sector in each financial reporting year from 2015 to 2018 as estimated in 100 runs. Each factor's occurrence count and normalized importance score are given.

2018	Highest Occurance	Count	Scores	2018	Second Highest Occurance	Count	Scores
F1	Short_Term_Debt_Q_Change	38	12%	**F1**	Return_On_Assets_Y_Change	12	12%
F2	Return_On_Assets_Y_Change	19	11%	**F2**	Short_Term_Debt_Q_Change	18	11%
F3	Return_On_Assets_Y_Change	10	10%	**F3**	Net_Debt_to_EBIT_Y_Change	6	10%
F4	Short_Term_Debt_Y_Change	8	10%	**F4**	Return_On_Assets_Y_Change	7	10%
F5	Short_Term_Debt_Y_Change	7	10%	**F5**	Return_On_Assets_Y_Change	6	10%
2017	**Highest Occurance**	**Count**	**Scores**	**2017**	**Second Highest Occurance**	**Count**	**Scores**
F1	Return_On_Assets_Y_Change	36	13%	**F1**	Dividend_Yield_Y_Change	28	14%
F2	Dividend_Yield_Y_Change	24	13%	**F2**	Return_On_Assets_Y_Change	18	12%
F3	Short_Term_Debt_Q_Change	20	11%	**F3**	Return_On_Assets_Y_Change	11	11%
F4	Short_Term_Debt_Q_Change	11	10%	**F4**	Dividend_Yield_Y_Change	7	10%
F5	Free_Cash_Flow_Q_Change	11	9%	**F5**	Short_Term_Debt_Q_Change	9	10%
2016	**Highest Occurance**	**Count**	**Scores**	**2016**	**Second Highest Occurance**	**Count**	**Scores**
F1	Short_Term_Debt_Q_Change	20	12%	**F1**	Operating_Margin_Y_Change	18	12%
F2	Return_On_Assets_Y_Change	18	11%	**F2**	Operating_Margin_Y_Change	10	11%
F3	Return_On_Assets_Y_Change	13	10%	**F3**	Operating_Margin_Y_Change	8	11%
F4	Operating_Margin_Y_Change	10	10%	**F4**	Operating_Margin_Y_Change	10	10%
F5	Short_Term_Debt_Q_Change	6	10%	**F5**	Short_Term_Debt_Q_Change	6	10%
2015	**Highest Occurance**	**Count**	**Scores**	**2015**	**Second Highest Occurance**	**Count**	**Scores**
F1	Short_Term_Debt_Q_Change	21	12%	**F1**	Return_On_Assets_Y_Change	13	12%
F2	Short_Term_Debt_Q_Change	19	11%	**F2**	Return_On_Assets_Y_Change	12	11%
F3	Return_On_Assets_Y_Change	13	10%	**F3**	DMA_5D/50D	9	10%
F4	DMA_5D/50D	12	10%	**F4**	Short_Term_Debt_Q_Change	7	10%
F5	PC_Ratios	6	10%	**F5**	PC_Ratios	6	10%

our XGBoost + GA model for all stocks that reported quarterly earnings during the 2015–2018 period. This reporting is also done on stocks from each of the seven industrial sectors. The results are created after running 100 tests on each group of stocks. The occurrence of dominant features has been counted and results are given in Tables 6–15. The features that most frequently appear as the top five driving factors are listed along with their occurrence counts. We use labels *F1* to *F5* represent the top five most impactful features. Features whose name starts with the name of a financial variable and ends with "Q_Change" or "Y_Change" represents the quarterly change or yearly change of the same variable.

During each forecasting exercise, XGBoost calculates and feedbacks the top *n* features which have contributed most in the prediction process. At the end of the 100 runs, we count which feature has appeared most often as the F1 feature (most dominant feature). We do the same counting for the F2 to F5 features. As an example, Table 6 shows that *EPS_Eearnings_Surprise_Backward_Av_Diff* appears 90 times out of 100 as the top dominant feature *F1* with an average normalized importance score of 21% for the 2018 test. Another feature *EPS_EarningsSurprise* also appears as the top dominant factor but this only happens 4 times. On the other hand, this feature appears 62 times as the second most dominant feature *F2* with a score of 14%.

References

Ahmed, A. S., & Safdar, I. (2018). Dissecting stock price momentum using financial statement analysis. *Accounting & Finance, 58*, 3–43.

Allen, F., & Karjalainen, R. (1999). Using genetic algorithms to find technical trading rules. *Journal of Financial Economics, 51*, 245–271.

Baker, H. K., Ni, Y., Saadi, S., & Zhu, H. (2016). Competitive earnings news and post-earnings announcement drift. *International Review of Financial Analysis, 63*, 331–343.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research, 6*, 159–178.

Beyaz, E., Tekiner, F., Zeng, X. J., & Keane, J. (2018). Comparing technical and fundamental indicators in stock price forecasting. In *Proceedings of IEEE 4th international conference on data science and systems* (pp. 1607–1613).

Bhushan, R. (1994). An informational efficiency perspective on the post-eanings announcement drift. *Journal o f Accounting and Economics, 18*, 45–65.

Bradbury, M. E. (1992). Voluntary semiannual earnings disclosures, earnings volatility, unexpected earnings, and firm size. *Journal o f Accounting Research, 30*, 137–145.

Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., & Vlachogiannakis, N. (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Applications, 112*, 353–371.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794).

Deng, S., Yoshiyama, K., Mitsubuchi, T., & Sakurai, A. (2015). Hybrid method of multiple kernel learning and genetic algorithm for forecasting short-term foreign exchange rates. *Computational Economics, 45*, 49–89.

Duan, B., & Dunlap, W. P. (1998). The robustness of trimming and winsorization when the population distribution is skewed. ProQuest Dissertations and Theses.

Erlien, M. (2011). *Earnings announcements and stock returns – A study of efficiency in the Norwegian Capital Market*. Master's thesis. University of Stavanger Norway.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*, 383–417.

Fama, E. F. (1991). Efficient capital markets ii. *The Journal of Finance, 46*, 1575–1617.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 22*, 3–56.

Foster, G., Olsen, C., & Shevlin, T. (1984). Earnings releases, anomalies, and the behavior of security returns. *The Accounting Review, 59*, 574–603.

Frino, A., Prodromou, T., Wang, G. H., Westerholm, P. J., & Zheng, H. (2017). An empirical analysis of algorithmic trading around earnings announcements. *Pacific-Basin Finance Journal, 45*, 34–51.

Göçken, M., Özçalıcı, M., Boru, A., & Dosdoğru, A. T. (2016). Integrating metaheuristics and artificial neural networks for improved stock price prediction. *Expert Systems With Applications, 44*, 320–331.

Hafezi, R., Shahrabi, J., & Hadavandi, E. (2015). A bat-neural network multi-agent system (bnnmas) for stock price prediction: case study of dax stock price. *Applied Soft Computing, 29*, 196–210.

Huck, N. (2019). Large data sets and machine learning: Applications to statistical arbitrage. *European Journal of Operational Research, 278*, 330–342.

Jidong, L., & Ran, Z. (2018). Dynamic weighting multi factor stock selection strategy based on xgboost machine learning algorithm. *Proceedings of IEEE international conference of safety produce informatization (IICSPI)*, 868–872.

Kim, D., & Kim, M. (2003). A multifactor explanation of post-earnings announcement drift. *The Journal of Financial and Quantitative Analysis, 38*, 383–398.

Liashchynskyi, P., & Liashchynskyi, P. (2019). Grid search, random search, genetic algorithm: A big comparison for NAS. *CoRR*. abs/1912.06059.

Malkiel, B. G. (2004). Models of stock market predictability. *Journal of Financial Research, 27*(4), 449–459.

Olson, D., & Mossman, C. (2003). Neural network forecasts of canadian stock returns using accounting ratios. *International Journal of Forecasting, 19*, 453–465.

Qiu, L. (2014). *Earnings announcement and abnormal return of s&p 500 companies*.

Sant'anna, L. R., Caldeira, J. F., & Filomena, T. P. (2020). Lasso-based index tracking and statistical arbitrage long-short strategies. *North American Journal of Economics and Finance 51*, No Pagination.

Schnaubelt, M., & Seifert, O. (2020). *Valuation ratios, surprises, uncertainty or sentiment: How does financial machine learning predict returns from earnings announcements?*

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*, 425–442.

Sheta, A., Ahmed, S., & Faris, H. (2015). A comparison between regression, artificial neural networks and support vector machines for predicting stock market index. *International Journal of Advanced Research in Artificial Intelligence, 4*, 55–63.

Solberg, L. E., & Karlsen, J. (2017). *Pairs trading using machine learning: An empirical study. Master's thesis*. Erasmus University Rotterdam.

Solberg, L. E., & Karlsen, J. (2018). *The predictive power of earnings conference calls: Predicting stock price movement with earnings call transcripts*. Master's thesis. Norwegian School of Economics.

Tyree, S., Weinberger, K., Agrawal, K., & Paykin, J. (2011). Parallel boosted regression trees for web search ranking. In *Proceedings of the 20th international conference on World wide web* (pp. 387–396).

Xie, J., Gao, R., Nijkamp, E., Zhu, S., & Wu, Y. (2020). Representation learning: A statistical perspective. *Annual Review of Statistics and Its Application, 7*.

Ye, J., Chow, J. H., Chen, J., & Zheng, Z. (2009). Stochastic gradient boosted distributed decision trees. In *Proceedings of the 18th ACM conference on Information and knowledge management* (pp. 2061–2064).