Engineering Applications of Artificial Intelligence 82 (2019) 1–12

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A new metric for individual stock trend prediction☆

Guang Liu, Xiaojie Wang *

Center for Intelligence of Science and Technology (CIST), School of Computer Science, Beijing University of Posts and Telecommunications, Beijing, China

ARTICLE INFO

Keywords:
Stock trend prediction
Model evaluation
Classification
Machine learning
Artificial intelligence

ABSTRACT

Individual stock trend prediction is extremely valuable for investment management. Previous studies mainly focused on proposing effective approaches to make profits. However, there is an ineffectiveness in model evaluation due to the inconsistency between model's performance and profitability. We name this inconsistency profit bias. In order to address the profit bias in model evaluation, this paper proposes a new effective metric, Mean Profit Rate (MPR). The effectiveness of metric is measured based on the correlation between the metric value and profit of the model. Experiments on five stock daily index data among four countries show that MPR outperforms the classification metrics in correlating to profit. In view of these findings, we suggest that MPR is a more effective metric than the classification metrics in stock trend prediction.

1. Introduction

The individual stock trend prediction is challenging due to the high volatility, irregularity, and noisy signal in the environment of the stock markets. In recent years, it draws a lot of attention from researchers in various areas, especially Artificial intelligence. Some studies treat it as a regression problem (Bollen et al., 2011; Schumaker and Chen, 2009), aiming to predict the future value of stock price or profits. While other studies treat it as a classification problem (Hsieh et al., 2011; Huang et al., 2008), aiming to predict the trend of stock price movement. In most cases, the classification approaches achieve higher profits than the regression ones (Leung et al., 2000). The effectiveness of different classification approaches in individual stock prediction has been widely explored (Choudhry and Garg, 2008; Lin et al., 2013; Wang and Choi, 0000; O'Connor and Madden, 2006; Lee, 2009a).

Evaluation metric plays a key role in model improvement despite of which specifical model is used to solve the problem. Therefore, it is essential to understand and account for the evaluation process. At present, the evaluation of a model for individual stock trend prediction consists of two stages (Kaastra and Boyd, 1996a; Ding et al., 2015; Li and Tsang, 1999; Zhai et al., 2007; Atsalakis and Valavanis, 2009a; Schumaker and Chen, 2009). The first stage is the model selection. At this stage, the optimal model is chosen by the performance under a given metric, especially the classification metric. Traditionally, Accuracy (Qian and Rasheed, 2007) (or Hit ratio (Huang et al., 2005)) and F-measure (F1) (Patel et al., 2015a), which are utilized to evaluate models both in binary and multiple classifications, are used to select the optimal model. Recently, the Matthews Correlation Coefficient (MCC) (Ding et al., 2015) is used to evaluate the effect of models in binary stock trend classification. The second stage estimates the profitability of the optimal model. The profitability of the selected model is often estimated by simulated trading which has many variables (Kaastra and Boyd, 1996a; Leung et al., 2000; O'Connor and Madden, 2006). It is obvious that the consistency between classification performance and profitability of the model is fairly important, i.e. the model with the highest metric value should achieve the biggest profit.

However, there are usually inconsistencies (Brownstone, 1996; Chang et al., 2009; Schumaker and Chen, 2009; Teixeira and De Oliveira, 2010a), as shown in Fig. 1. We call the inconsistency profit bias. Traditionally, model-oriented and data-oriented methods are used to address the profit bias. In model-oriented methods, researchers aim to alleviate profit bias by incorporating profit information into model's learning targets. In Saad et al. (1998), they assign different weights to samples in the loss calculation to avoid false alarm. However, this also makes the model difficult to converge and affects the performance of other classes. In Moody and Saffell (2001) and Deng et al. (2017a), they use the direct reinforcement learning model to learn the trading signals directly. However, this method is less flexible in strategy settings. In the data-oriented methods, researchers aim to reduce the probability of profit bias by changing the original data distribution. In Kaastra and Boyd (1996b), they suggest removing some small changes from the dataset to avoid profit bias. In Luo and Chen (2013), they use the statistical-based method to segment and label the datasets. All data-oriented methods change the actual distribution of the original stock data. However, these above methods fail to fully address the profit bias.

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.engappai.2019.03.019.

* Corresponding author.
E-mail addresses: peter_liu@bupt.edu.cn (G. Liu), xjwang@bupt.edu.cn (X. Wang).

https://doi.org/10.1016/j.engappai.2019.03.019
Received 22 April 2018; Received in revised form 8 November 2018; Accepted 19 March 2019
Available online 27 March 2019

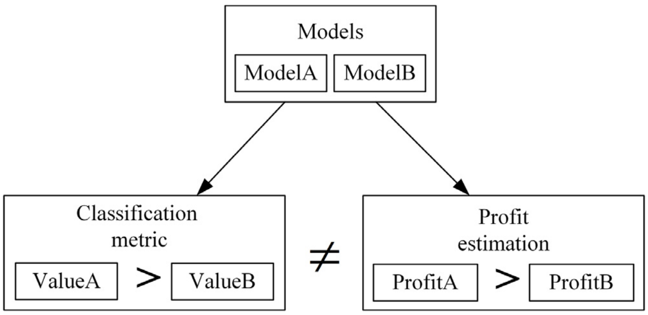

Fig. 1. The inconsistency in evaluation stock trend models.

The key factor to trigger the profit bias is the inconsistency between the metrics used in the two-stage evaluation methods.

A simple example is given to demonstrate the profit bias. Suppose two models (A and B) are used to predict the trend of a stock in three consecutive trading days. These trading days have profit rate 1.00%, 2.00%, and −4.00%, respectively. The results of each model are used as the trade signals for a day trader which invests \$100 on each trade. Model A gives results of up, up and up. It suggests holding a long position each day. The profit of model A is −1. Model B gives a result of down, down, down. It suggests holding a short position for three days. The profit of Model B is 1. Model A's accuracy is 0.667 (2 out of 3), while model B is 0.333. On one hand, model A has a higher accuracy, but a lower level of profit. On the other hand, model B with the lower accuracy but achieves more profits. So, there is an inconsistency between the model's performance and the model's profitability. This is just a very simple example of profit bias, a more detailed example of the problem will be discussed in the next section.

The present study is aimed at overcoming profit bias by improving the effectiveness in the evaluation of individual stock trend prediction. Unlike previous works, we believe the present models can learn the profit information within stock data, and the profit bias is due to the ineffectiveness in selecting the optimal model. Therefore, this paper proposes a new metric, Mean Profit Rate (MPR), to evaluate models without profit bias. Experiments on five stock index data among four countries show that our metric can effectively select models due to its consistency with model's profitability. The findings of the present study may be of some help in improving the effectiveness during the evaluation of individual stock trend prediction.

The rest of this paper is arranged as follows. Section 1 describes the background of stock market prediction, including the classification metrics and simulated trading strategies. Section 2 shows the theoretical proof of the proportional relation between MPR and profitability of model after a careful analysis on a more detailed example of profit bias. Section 3 gives the settings of data, models, and evaluations in our experiments. The experiment results are given and discussed in Section 4. Section 5 presents the conclusions.

2. Related works

This section firstly describes the stock market prediction. Then, the classification metrics used in model evaluation is described. Finally, the definition of simulated trading, which is used to estimate the profitability of a model, is summarized.

2.1. Stock market prediction

Researches on predicting stock markets mainly focused on proposing effective stock trend prediction methods. According to the emphasis of a research, previous researches can be divided into three categories: feature-oriented methods, model-oriented methods, and integration-oriented methods. Feature-oriented methods mainly use statistical-based methods such as principal component analysis (PCA) (Tsai and Hsiao, 2010) and information gain (IG) (Lee, 2009b) to select effective features. The performance of a given model will be improved if the low relevant features in the input are removed. Most of the features are real values of a certain range, and are weakly expressive for trend information. In order to obtain trend expressive features, some studies have introduced fuzzy logic to process feature data (Atsalakis and Valavanis, 2009b; Chang and Liu, 2008). They transform each feature value into a probability distribution over multiple categories, thereby improving the feature's expressive for classification. The model-oriented approach focuses on improving the fitting ability of the model. The support vector machines (SVM) and neural networks (NN) have proven to be very effective for stock market predictions (Kara et al., 2011a). The SVM training algorithm has a cubic time complexity, and both NN and SVM are easily over-fitting due to the excessive parameter size. The Extreme Learning Machine (ELM), which can speed up training and improve generalization performance through randomly generated hidden layer units, is used to predict the stock market (Li et al., 2016). In recent years, deep learning models have many successful applications in computer vision and Natural Language Processing (NLP) tasks due to their powerful feature extraction capabilities. Therefore, deep learning has also been applied to many scenarios in the stock market prediction (Ding et al., 2015; Akita et al., 2016; Deng et al., 2017b). These above studies mainly use a single model. However, an ensemble of multiple models usually achieves better prediction performance than a single model. Therefore, ensemble models, such as Random Forest (RF) (Patel et al., 2015b), is applied for stock market prediction. The integrated-oriented methods often combine multiple artificial intelligence or statistical-based techniques to predict the stock market. Some studies combine the text classification (De Fortuny et al., 2014; Shynkevich et al., 2015) or sentiment analysis (Bollen et al., 2011) in NLP with a classification model to predict stock markets. Some studies have proposed feature selection methods based on genetic algorithm (GA) combined with NN model to select useful features (Kim and Han, 2000; Wu et al., 2007). Some studies have used the technical indicators used in technical analysis as the representation of stock data, combined with the classification model for prediction (Mizuno et al., 1998; Zhai et al., 2007; Teixeira and De Oliveira, 2010b). Some studies use turning points to segment the target sequence based on the piecewise linear regression, and then use NN to predict these points (Oh and Kim, 2002; Luo and Chen, 2013). It is worth noticing that many studies have used two-stage way to predict individual stock trend.

Table 1
The literature of stock market prediction.

Literature	Model type	Metric	Strategy
Mizuno et al. (1998)	C[a]	ACC	L/S[c],M[e]
Li and Tsang (1999)	C	ACC	L,M
Zhai et al. (2007)	C	ACC	L/S,M
Atsalakis and Valavanis (2009a)	C	ACC	L[d],M
Schumaker and Chen (2009)	L[b]	ACC	L/S,M
Hagenau et al. (2013)	C	ACC	L/S,M
Li et al. (2014)	L	ACC	L/S,M
Ding et al. (2015)	C	ACC,MCC	L/S,A[f]
Li et al. (2015)	L	ACC	L/S,M
Patel et al. (2015a)	C	ACC,F1	–
Deng et al. (2017b)	C	ACC	–
Hu et al. (2018)	C	ACC	L/S,M

[a]C in model type: classification.
[b]L in model type: level estimation.
[c]L/S in strategy: long and short.
[d]L in strategy: long only.
[e]M: multiplicative profit.
[f]A: addictive profit.

2.2. Classification metrics

The metric used for choosing the optimal model is the critical component in model selection. Table 1 lists the classification metrics, such as Accuracy (ACC), Matthews Correlation Coefficient (MCC) and F-measure (F1), in model selection. In this study, we use the Accuracy, MCC, and F1 as the represent of classification metrics. These metrics usually are calculated basing on the confusion matrix which has four components: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). Accuracy is the rate of correct prediction and is calculated by the formula in Eq. (1).

$$acc = \frac{TP + TN}{N}. \tag{1}$$

Here, N is the total number of samples in dataset. This metric is easy to understand and calculate. Therefore, Accuracy is often used in the evaluation of stock trend prediction models. The stock data usually is unbalanced data, which means that the number of one kind of trend is significantly more than the others. As in the Eq. (2), MCC is calculated to measure the quality of binary classification.

$$mcc = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \tag{2}$$

F-measure, usually F1, is another metric for classification. It considers both the recall and precision of the classification performance. The calculation of F1 is displayed in Eq. (3)

$$f1 = \frac{2 \times P \times R}{P + R}, \tag{3}$$

$$P = \frac{TP}{TP + FP}, \tag{4}$$

$$R = \frac{TP}{TP + FN}. \tag{5}$$

Although these metrics have been successfully applied to evaluate models in individual stock trend prediction, there is an inefficiency due to the weak correlation between the outcome of classification metrics and model's profitability. The latter is often estimated by simulated trading. In order to clarify the correlations between these metrics and the simulated trading, the definition of simulated trading is summarized in the next section.

2.3. Simulated trading

Simulated trading is the key component in profit estimation. It is a group of algorithms for estimating the model's profitability in real stock market. Table 1 also lists the studies containing simulated trading. Simulated trading can be adapted to data with different frequencies, such as daily and minutely data. In order to simplify the highly complicated trading actions in real life, the algorithms just concern major factors about trading in the stock market. The model usually predicts the trend in the next time step. Therefore, the strategy in simulation is usually designed to hold the securities in one time step. The profitability of model is the profit earning per share. So, the simplified version of operations with the shares number ignored is used. For example, that we buy \$1000 stock at price \$11 means we will have \$1000 value of stock. In the simulation, we simulate the behaviors of a day trader or intraday trader in the stock market. As defined in Eq. (6), a well-defined simulated trading algorithm usually includes trading method (TM), profit method (PM) and the trading cost (TC),

$$profit = \sum f_{tm}(T_i, \widetilde{T_i}) \times f_{pm}(r_i, tc). \tag{6}$$

Here, the function f_{tm} generates the trading signal, which is a value in $[-1, 1]$. The function f_{pm} is the profit function, which calculates the profit for each trade, the variable tc is the transaction cost. Advanced operations such as loss stop and gain stop (Teixeira and De Oliveira, 2010a), which have similar effects to profits as TC, are not mentioned in this work. T_i is the real trend in price list, and $\widetilde{T_i}$ is the predicted trend. The variable r_i is the profit rate at time i. The Long (L) and Short (S) operation are trading methods. The L operation means to buy the stock shares or call option. It will gain profits if the price rises and vice versa (Hsieh et al., 2011). The S operation means to sell the stock shares or buy the put option. It will gain profits if the price declines and vice versa (Wen et al., 2010). PM contains two profit calculation methods. One is the addictive profit (A) method and the other is the multiplicative profit (M) method (Breiman et al., 1984). The A method accumulates the profits on every trade (Hsieh et al., 2011), while the M method multiplies one plus profit rate by the previous investment on each trade (O'Connor and Madden, 2006). TC includes many complicated components such as commissions, taxes and bid–ask spreads (Lee et al., 2014; Wu et al., 2006a). Therefore, TC is usually estimated by a certain rate of stock price (for example, 0.2%) for simplification (Wen et al., 2010).

In simulation, a stock agent will execute a strategy based on model's prediction. The agent has fixed the initial capital and adequate credits to carry out investment operations. We use the form of PM–TM–TC to represent the strategy. The strategy L-A-0 means the long only, additive profit method and no transaction cost strategy. The trade method L/S means the strategy can use L or S operation on each trade. For example, the commonly used L/S-M-0 strategy means both long and short operation are supported, the multiplicative profit method and zero transaction cost. The diversity of trade strategy can lead to various possible profit estimations. One model may make a profit under one strategy while suffer loss under other strategy. In summary, the simulated trading is an algorithm to estimate the profitability (earning per share), and it is bonded with certain strategy.

3. Proposed metric

This section gives the theoretical proof of the proportional relation between MPR and model's profitability. Before the theoretical derivation, we will give the framework to evaluate the metrics consistency with the model's profitability and analyze a more detailed example to explain the profit bias.

3.1. Framework

Fig. 2 illustrates the framework to evaluate the consistency between a metric and the model's profitability. It has four mainly steps (Kaastra and Boyd, 1996a; Lin et al., 2013; Wang and Choi, 0000). The first step is to extract features from stock data. Some techniques, such as Technical Indicator (Teixeira and De Oliveira, 2010a), are applied to extract useful signals from the noisy stock data. The second step is model selection. The optimal model is selected by the outcome of the metric. The third step is profit estimation, which estimates the profitability of the optimal model by simulated trading. The last step is to evaluate the consistency between the metric and the model's profitability.

3.2. Profit bias

A detailed example is given to explain the profit bias. Suppose the trend prediction is proceeded on six consecutive stock trading days $\langle D1, D2, D3, D4, D5, D6\rangle$. The daily closing prices are shown in the second line of Table 2, the profit rate r_t of close prices at tth trading day can be calculated by Eq. (7).

$$r_t = |\frac{C_{t+1}}{C_t} - 1|. \tag{7}$$

Here, C_t is the close price at day t. The values of profit rate from $D1$ to $D5$ are shown in the 3rd line of Table 2. The trends of the close price have two possible values: 1 or −1. 1 means the trend is upward, −1 means the trend is downward. The trends of first five days are shown in the 4th line of Table 2. Assume two models, A and B, give the trend predictions as the 5th and 6th line in Table 2.

F1 and Accuracy are shown in Table 3 which can be calculated by the trend prediction results of model A and B in Table 2. The scores of F1 and Accuracy of model A are better than those in model B. Because of this, model A is the optimal model according to the

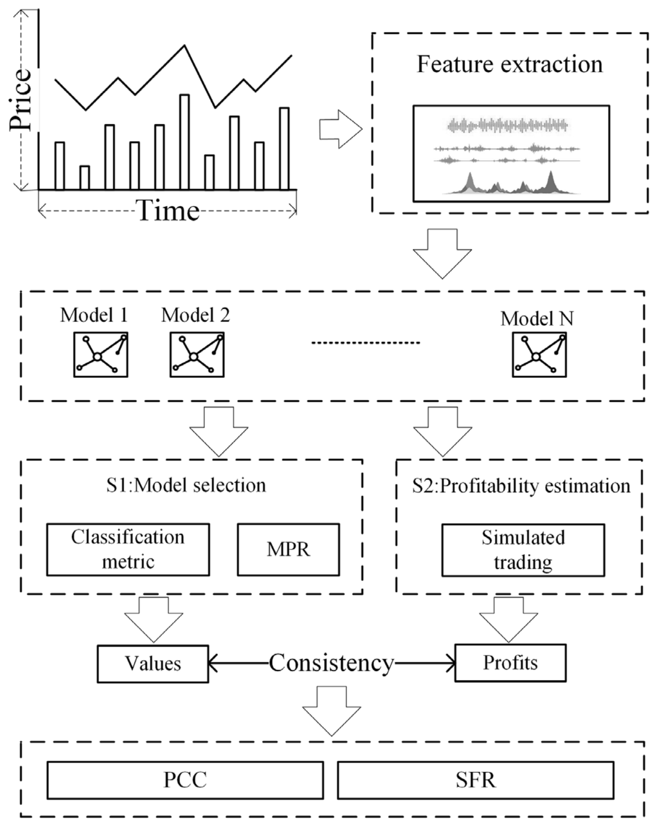

Fig. 2. The framework of stock price trend prediction.

Table 2
The information of stock price and model's prediction.

	D1	D2	D3	D4	D5	D6
Close price	100	101	99.99	100.99	111.09	109
Profit rate	1.00%	1.00%	1.00%	10.00%	1.88%	–
Trends	1	−1	1	1	−1	–
$Prediction_A$	1	−1	1	−1	−1	–
$Prediction_B$	−1	1	−1	1	−1	–

Table 3
Classification performance and profit of each model.

Model	ACC	F1	Profit
A	0.8	0.8	−55
B	0.4	0.4	47

existing model selection method. However, one may ask whether model A achieves higher profit than model B. The L/S-M-0 simulated trading with 1000 initial capital, allowing long and short operations with multiplicative profit method and zero transaction cost, is carried out using the predictions of each model to estimate the profitability of each model. It turns out that model A has a higher level of accuracy but a lower level of profitability.

This findings indicate that the major cause for profit bias lies in the ignorance of profit information in classification metrics. In the above example, model A accurately predicts the trends of 4 days, D1, D2, D3 and D5. The sum of the profit rates of these correctly predicted days is less than the profit rate of the single day D4. Although model A has made correct predictions on most days, it cannot make a profit in simulated trading. These findings indicate three features of an effective metric in model selection. First, the metric should be able to pick up the profitable model under a given strategy. Second, the metric should be consistent with model's profitability under the given strategy. Third, the metric should be strategy free, which means the metric is able to indicate the profitability of model no matter what strategy is used.

In summary, classification metrics suffer from profit bias due to the ignorance of profit information. They are ineffective in the evaluation of individual stock trend prediction, since the model with optimal metric value cannot guarantee the optimal profitability. Therefore, the metric consistency with the model's profitability is needed.

3.3. Mean profit rate

This section firstly defines Mean Profit Rate (MPR), which is consistent with model's profitability. Then, theoretical proof of the proportion relation between MPR and model's profitability is given.

Definition of MPR. MPR is the expectation of profit rate on individual trend prediction. Assume r_i is the profit rate sequence of a stock for n time steps. The real trend of the stock at ith time step is $T_i \in [-1, 1]$. And $\widetilde{T}_i$ is the model's prediction at time step i. MPR is calculated by the formula in Eq. (8) below

$$MPR = \frac{1}{n}\sum_{i=1}^{n} I(T_i, \widetilde{T}_i) \times r_i \,. \tag{8}$$

Here $I(\cdot)$ is an indicator function. It returns 1 when the T_i equals $\widetilde{T}_i$, and returns -1 when T_i differs from $\widetilde{T}_i$. MPR is supposed to be consistent with model's profitability, so it can overcome the profit bias in model selection. The consistency can be proved if MPR is proportional to model's profitability. Next, we define the simulated trading and then give the theoretical proof of the theorems about the proportion relations.

Formulation of simulated trading. The simulated trading estimates the profit of a given model with fixed initial capital. Its outcome is equivalent to model's profitability under a given strategy. In the simulated trading introduced in Section 2.3, TM, PM and TC are three major components to set up a strategy. Although long-only or short-only can be used in simulated trading, they cannot fully estimate the profitability of a model. Therefore, it is reasonable to use the L/M trading method in simulated trading under most circumstances. The simulated trading can be calculated by

$$profit = \begin{cases} C \times \sum_{i=1}^{n}(I(T_i, \widetilde{T}_i) \times r_i - tc), & PM = A \\ C \times \{\prod_{i=1}^{n}(1 + I(T_i, \widetilde{T}_i) \times r_i - tc) - 1\}, & PM = M \end{cases} \tag{9}$$

Here, $profit$ is the estimated profit. C is the initial capital. tc is the transaction cost rate.

Theoretical proof. In this section ,we will give the theoretical proof of the proportional relation between MPR and model's profitability. The profitability is equivalent to how much profit can be made when using a given strategy. Eq. (9) shows the two major series of strategies (A and M profit method) in simulated trading.

Theorem 1. *Fix the transaction cost tc and initial capital C, MPR is proportional to the profit with addictive profit method. That is,*

$$MPR \propto profit_A \,, \tag{10}$$

where $profit_A$ is the profit estimated by simulated trading with addictive profit method.

Proof. The quantity $profit_A$ can be formulated as Eq. (9) when $PM = A$. That is,

$$profit_A = C \times \sum_{i=1}^{n}(I(T_i, \widetilde{T}_i) \times r_i - tc) \tag{11}$$

$$= C \times n \times \frac{1}{n}\sum_{i=1}^{n}(I(T_i, \widetilde{T}_i) \times r_i - C \times n \times tc) \tag{12}$$

$$= C \times n \times MPR - C \times n \times tc. \quad (13)$$

This shows that MPR is proportional to the profit under the additive profit method.

Theorem 2. *Fix the transaction cost tc and initial capital C, MPR is proportional to the profit with multiplicative profit method. That is,*

$$MPR \propto profit_M, \quad (14)$$

where $profit_M$ is the profit estimated by simulated trading with multiplicative profit method.

Proof. The quantity $profit_M$ can be formulated as Eq. (9) when $PM = M$. That is,

$$profit_M + C = C \times \prod_{i=1}^{n}(1 + I(T_i, \widetilde{T}_i) \times r_i - tc). \quad (15)$$

Take the logarithm of both sides of Eq. (15), one obtains

$$\log(profit_M + C) = \log C + \sum_{i=1}^{n} \log(1 + I(T_i, \widetilde{T}_i) \times r_i - tc). \quad (16)$$

For simplicity, R_i is used to replace the term $I(T_i, \widetilde{T}_i) \times r_i - tc$ in Eq. (16), this gives

$$\log(profit_M + C) = \log C + \sum_{i=1}^{n} \log(1 + R_i). \quad (17)$$

Using Taylor expansion, $\log(1 + x) = x + o(x^2)$. Them Eq. (17) can be rewritten as follows,

$$\log(profit_M + C) = \log C + \sum_{i=1}^{n}(R_i + o(R_i^2)) \quad (18)$$

$$= \log C + C \times n \times \frac{1}{n} \sum_{i=1}^{n} I(T_i, \widetilde{T}_i) \times r_i$$

$$- C \times n \times tc + n \times o(R_i^2) \quad (19)$$

$$= \log C + C \times n \times MPR - C \times n \times tc + n \times o(R_i^2). \quad (20)$$

Here, $o(R_i^2)$ is a very small value, so it can be ignored. The log function is a monotonically increasing function. Hence it can be concluded that MPR is proportional to $profit_M$.

In summary, MPR is proportional to the profitability estimated by simulated trading. This explains why MPR is consistent with model's profitability. Although MPR can be served as an indicator to model's profitability in simulated trading. MPR is strategy free while simulated trading is bonded with certain trade strategy.

4. Experimental settings

In this section, we first describe the data and feature used in our experiments. Then, we introduce the details of the models. Finally, the evaluation methods for metric in stock price prediction are given.

4.1. Data and feature

Data. Five daily index data in the stock markets around four different countries are used in our experiments. They consist of Dow Jones Industrial Average (DJIA), Standard and Poor's 500 (S&P500), Hang Seng Index (HSI), Nikkei 225(N225) and Shanghai Stock Exchange composite index (SSE). We download the data from Yahoo Finance API between 01/22/2007 and 12/30/2017. Each dataset has approximately 2700 data points. In order to test the effectiveness of metrics under different time spans, we split the datasets using the slice windows method (Zhai et al., 2007), as shown in Fig. 3. Each dataset is splitted into 50 subsets with window size 30. Each subset contains 330 data points, 300 for training and 30 for testing.

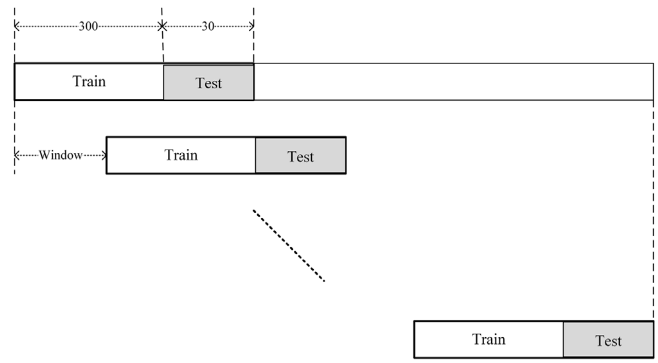

Fig. 3. Slice windows of dataset.

Table 4
Technical indicators for stock trend prediction.

Technical indicator	Explanation
Ten days moving average (MA)	Moving averages are used to emphasize the direction of a trend and smooth out price and volume fluctuations that can confuse interpretation
Ten days relative strength index (RSI)	RSI compares the magnitude of recent gains to recent losses in an attempt to determine overbought and oversold conditions of an asset
Ten days stochastic line (K,D)	The stochastic line K and line D are used to determine the signals of over-purchasing, over-selling, or deviation
Moving average convergence and divergence (MACD)	MACD shows the difference between a fast and slow exponential moving average (EMA) of closing prices. fast:short-period average. slow: long period average
Ten days Commodity Channel Index (CCI)	CCI is a versatile indicator that can be used to identify a new trend or warn of extreme conditions.
Ten days Momentum (MOM) and Rate of Change (ROC)	The MOM and ROC indicators show trend by remaining positive while an uptrend is sustained, or negative while a downtrend is sustained.
Ten days Williams %R(%R)	%R is a technical analysis oscillator showing the current closing price in relation to the high and low of the past 10 days
Ten days Advance/Decline line (A/D)	The A/D indicates changes in the value of the advance-decline index over a certain time period.

Features. The technical indicators can help reduce the noise in the stock data. In this work, ten technical indicators, as in Kara et al. (2011b) and Patel et al. (2015a), are used as the representation of the stock data on each dataset. They are describe in Table 4. The min–max normalization in Eq. (21) is applied to each feature to normalize each data into the range $[0, 1]$:

$$X = \frac{X - X_{min}}{X_{max} - X_{min}}. \quad (21)$$

Here, X is the feature list, X_{max} is the maximum in X, and X_{min} is the minimum in X.

4.2. Classification models

Classification models predict the movement direction, while level estimation models predict the range of profit or price. In some cases, the classification ones can achieve more profit than level estimate ones (Leung et al., 2000). As a result, eight classification models are compared in this study, some of them achieves the state-of-the-art performance in stock trend prediction.

- K-nearest neighbor (KNN) (Teixeira and De Oliveira, 2010b). KNN is a machine learning algorithm that is considered very simple to implement (Aha et al., 1991). The stock prediction problem can be mapped into a similarity based classification (Teixeira and De Oliveira, 2010a). It assigns new data to a class according to the most common class amongst k-nearest neighbors (Altman, 1992).

- Support Vector Machine (SVM) (Patel et al., 2015b). SVM is firstly introduced by Vapnik (Cortes and Vapnik, 1995). It often achieves the state-of-the-art in stock trend prediction. It projects the input data into a higher dimensional space by the kernel function and separates different classes of data using a hyper-plane. The trade-off between margin and misclassification errors is controlled by the regularization parameter c. SVM with polynomial kernel (SVM_{poly}) uses kernel function $K(x_i, x_j) = (x_i \cdot x_j + 1)^d$, and SVM with radial basis kernel functions (SVM_{rbf}) uses $K(x_i, x_j) = exp(-\gamma \left\| x_i - x_j \right\|^2)$. Here d is the degree of the polynomial function and γ is the constant of the radial basis function.
- Decision Tree (DT) (Wu et al., 2006b). DT is widely used in data mining (Breiman et al., 1984) and also in stock trend prediction (Huang et al., 2005). A decision tree is a connected acyclic graph with the leaves represented by decisions. The process of determining decisions starts from the root. Then, the transition into either the right or the left node takes place depending on the node value. At the final stage, the leaf at the end is taken.
- Random Forest (RF) (Ballings et al., 2015). RF is firstly proposed by Ho in 1995 (Ho, 1995). It belongs to the category of ensemble learning algorithms. It uses decision tree as the base learner of the ensemble. The idea of ensemble learning is that a single classifier is not sufficient for determining class of test data. After the creation of n trees, when testing data is used, the decision on which majority of trees come up with is considered as the final output. This also avoids problem of over-fitting.
- Kernelized Extreme Learning Machine (K-ELM) (Li et al., 2016). The key idea of ELM is to randomly initialize weights of the hidden layers. K-ELM introduces kernel function into ELM to improve the prediction performance. Due to the reduction of free parameters, K-ELM has faster training speed and better performance than NN model.
- Fuzzy Deep Neural Network (FDNN) (Deng et al., 2017b). FDNN uses fuzzy-neural layers and fully connected layers to learn the fuzzy representation and neural representation separately. Then, these two representations are fused by a two layer fully connected layer. The fused representations is fed to a softmax activation to get the trend predict results.
- Genetic Algorithm and Neural Network (GA–NN) (Inthachot et al., 2016). The Genetic Algorithm is integrated with Neural Network to predict the stock price index trend. The fitness is measured by the accuracy of NN.

These models are used in our experiments. We set the ranges of the values for the parameters of each model as in Table 5. The degree (d) and regularization constant C are key parameters for SVM_{poly}. The parameter gamma (γ) and C are important parameters for SVM_{rbf}. The number of neighbors (k) is the key parameter for KNN. One of the important parameters in DT is the max depth of the tree. Comparing to DT, number of the trees is an additional parameter in RF. K-ELM uses RBF kernel, so it has to set up γ and the number of hidden units n. For FDNN, the key parameter is the number of hidden units in the fuzzy layers (num_{fuzzy}) and neural representation layers (num_{neural}). In GA–NN, we test different settings of crossover probability ($p_{crossover}$) and mutation probability ($p_{mutation}$).

4.3. Evaluation

Pearson Correlation Coefficient (PCC) and Selection Failure Rate (SFR) are used to evaluate the consistency between the metric and profit.

Table 5
The setting of parameters for each model.

Model	Parameter1	Parameter2
KNN	$k = 2n + 1, n \in [1, 10]$	–
SVM_{poly}	$C \in [0.5, 1, 5, 10, 100, 1000]$	$\gamma = 0.5 * n, n \in [1, 20]$
SVM_{rbf}	$C \in [0.5, 1, 5, 10, 100, 1000]$	$d \in [1, 4]$
DT	$depth_{max} \in [1, 5]$	–
RF	$depth_{max} \in [1, 5]$	$num_{tree} \in [3, 5, 7, 9]$
K-ELM	$\gamma = 0.01 * 2^k, k \in [0, 15]$	$n \in [10, 20, 30, 40, 50]$
FDNN	$num_{fuzzy} \in [32, 64, 128]$	$num_{neural} \in [32, 64, 128]$
GA–NN	$p_{crossover} = 0.1 * k, k \in [1, 5]$	$p_{mutation} = k * 0.01, k \in [1, 20]$

PCC is often used to evaluate the correlation between metric values and profit. As shown in Eq. (22), it is the criterion for the linear correlation between the two variables X and Y.

$$PCC_{X,Y} = \frac{cov(X, Y)}{\sigma_X \sigma_Y}. \tag{22}$$

Here, $cov(X, Y)$ is the covariance of X and Y, σ_X and σ_Y are the variance of X and Y, respectively. The value is a real number between -1 and 1. The value 1 means total positive correlation, 0 means no linear correlation, while -1 means total negative correlation.

SFR is the failure rate of the metric used in model selection. Assume model A and model B make predictions on the same dataset. We using metric X to evaluate the models. We get the pair $\langle x_A, p_A \rangle$ for model A, where x_A is the metric value of model A under metric X and p_A is the profit of model A. Similarly, we get $\langle xB, pB \rangle$ for model B. When we select the optimal model using metric X, there are two possible conditions. As shown in Eqs. (23) and (24);

$$condition1 : [x_A > x_B \ and \ p_A > p_B] \ or \ [x_A \leqslant x_B \ and \ p_A \leqslant p_B], \tag{23}$$

$$condition2 : [x_A > x_B \ and \ p_A \leqslant p_B] \ or \ [x_A \leqslant x_B \ and \ p_A > p_B]. \tag{24}$$

It suggests that condition2 is a failure in model selection. If we have to select the optimal model from N models, there are $C_N^2 = N \times (N - 1)/2$ result pairs. Then we obtain

$$SFR_X = \frac{2 \times (number \ of \ condition2)}{N \times (N - 1)}. \tag{25}$$

Here, SFR_X is the SFR when metric X is used for choosing the optimal model.

5. Results and analysis

In this section, we evaluate the effectiveness of Mean Profit Rate (MPR) in stock trend prediction compared to classification metrics. First, MPR's ability in finding profitable models is compared with those of the classification metrics. Second, the consistency between MPR and model's profitability is measured by Pearson Correlation Coefficient (PCC). Third, we use Selection Failure Rate (SFR) to evaluate our MPR in the model selection. Then, the strategy free characteristic of MPR is examined by multiple strategies with different settings. Finally, a demonstration of evaluation multiple models by MPR is given.

5.1. Profitable model selection

In this experiment, we use MPR and classification metrics to select the optimal model for each subset of five datasets. The profit of the selected model is estimated by simulated trading (L/S-A-0 and L/S-M-0). The metric which can get higher profits in simulated trading is considered to be more effective. Firstly, every model is trained using different parameters (Table 5) on the training set of each subset. Secondly, the models are evaluated by each metric separately on the testing set, and the model with the highest level of metric value is chosen as the optimal model. Finally, the profitability of the optimal model selected by each metric is estimated by the simulated trading profit on the testing set.

Table 6
The average profit of optimal model selected by each metric.

Model	L/S-A-0				L/S-M-0			
	ACC	MCC	F1	MPR	ACC	MCC	F1	MPR
KNN	526.48	523.59	499.89	**700.85**	529.34	515.81	503.28	**711.77**
SVM_{poly}	621.75	500.20	458.36	**718.78**	633.68	512.09	468.65	**738.11**
SVM_{rbf}	804.71	767.80	734.57	**970.72**	835.62	796.86	765.08	**1016.89**
DT	670.87	553.87	418.89	**720.61**	699.50	561.21	418.90	**753.06**
RF	806.91	684.74	541.95	**982.18**	841.80	691.37	543.68	**1030.78**
K-ELM	862.21	781.41	671.80	**1146.46**	895.37	807.59	684.36	**1154.11**
FDNN	971.83	890.19	750.3	**1173.28**	1028.88	933.36	773.90	**1260.68**
GA–NN	925.84	898.57	830.86	**1046.84**	976.62	949.17	880.23	**1114.38**

Bold numbers indicate the best results.

Table 7
The average profit of optimal model selected by each metric.

Model	L/S-A-0				L/S-M-0			
	ACC	MCC	F1	MPR	ACC	MCC	F1	MPR
KNN	738.22	734.48	691.14	**839.59**	747.85	743.89	698.38	**856.85**
SVM_{poly}	674.31	641.20	531.99	**769.74**	691.87	650.12	536.18	**794.57**
SVM_{rbf}	798.71	850.48	700.70	**1031.22**	826.46	876.34	722.55	**1079.30**
DT	603.10	619.55	511.67	**799.54**	613.67	624.73	517.71	**818.89**
RF	908.15	900.52	815.34	**1162.94**	934.97	929.66	844.20	**1221.46**
K-ELM	1104.56	1007.53	822.44	**1383.76**	1164.68	1059.54	841.85	**1483.10**
FDNN	960.29	1013.31	763.89	**1128.04**	1001.10	1061.10	783.24	**1185.82**
GA–NN	925.42	857.71	603.18	**1090.66**	959.88	878.66	617.22	**1141.36**

Bold numbers indicate the best results.

Table 8
The average profit of optimal model selected by each metric.

Model	L/S-A-0				L/S-M-0			
	ACC	MCC	F1	MPR	ACC	MCC	F1	MPR
KNN	955.88	922.57	902.31	**1138.25**	1019.46	984.95	966.70	**1215.88**
SVM_{poly}	1121.16	1069.85	803.30	**1222.02**	1198.34	1150.04	860.28	**1305.75**
SVM_{rbf}	1232.49	1091.11	930.84	**1439.90**	1286.75	1131.50	966.90	**1536.37**
DT	738.55	741.96	348.58	**917.81**	744.28	772.95	345.44	**964.12**
RF	1281.22	1211.67	1034.17	**1520.46**	1375.56	1292.08	1095.27	**1638.74**
K-ELM	1529.67	1358.06	1238.29	**1715.89**	1682.84	1486.53	1332.58	**1901.16**
FDNN	1349.80	1308.44	1108.54	**1556.52**	1454.83	1403.54	1172.39	**1696.75**
GA–NN	1263.71	1239.12	1022.16	**1518.74**	1373.91	1344.85	1113.63	**1666.47**

Bold numbers indicate the best results.

Table 9
The average profit of optimal model selected by each metric.

Model	L/S-A-0				L/S-M-0			
	ACC	MCC	F1	MPR	ACC	MCC	F1	MPR
KNN	738.86	843.33	661.52	**987.84**	749.61	880.80	675.15	**1030.54**
SVM_{poly}	985.04	810.77	691.23	**1200.22**	1013.55	827.85	708.69	**1253.46**
SVM_{rbf}	1303.90	1266.77	1175.04	**1575.89**	1369.59	1328.60	1230.06	**1696.05**
DT	962.70	837.82	704.27	**1111.60**	1052.06	927.78	768.95	**1215.86**
RF	1372.99	1312.24	1136.32	**1674.54**	1475.77	1414.80	1206.14	**1823.71**
K-ELM	1535.79	1397.97	1215.47	**1836.00**	1657.30	1500.52	1285.32	**2015.88**
FDNN	1458.45	1404.66	1165.91	**1685.75**	1588.31	1507.71	1244.50	**1865.04**
GA–NN	1211.50	1300.22	1071.93	**1561.61**	1323.84	1417.63	1157.35	**1722.76**

Bold numbers indicate the best results.

Tables 6–10 lists the average profit of the optimal model on different datasets. For example, the 2nd column on the 2nd line is the average profit by L/S-A-0 strategy when using Accuracy to select the optimal KNN model. Table 11 lists the results of the t-test, which is carried out on the results between MPR and the classification metrics under L/S-A-0 strategy. Table 12 lists the t-test results under L/S-M-0 strategy. Three observations can be concluded from these results.

- MPR is a better selector than the classification metrics in model selection. It can be seen from the results that the optimal model selected by MPR has a higher average profit than the optimal model selected by other metrics. On all datasets, except N225 (over 16%), the MPR-selected model can achieve an average gain of more than 20% over the ACC-selected model.
- Under different trading strategies, the profitability of the model selected by MPR is significantly better than that selected by other metrics. The results in Table 11 indicate that the average L/S-A-0 profit of the model selected by MPR is significantly higher than those of the models selected by other indicators. The results in Table 12 show that the average L/S-M-0 profit of the model selected by MPR is significantly higher than those of the models selected by other metrics.
- The K-ELM model outperforms other classification models in stock trend prediction. In four out of five of the datasets, the optimal K-ELM model selected by MPR achieves the highest average profits. In DJIA, the average profit of the MPR optimal model is the second highest, just below the average profit of the FDNN model.

Table 10
The average profit of optimal model selected by each metric.

Model	L/S-A-0				L/S-M-0			
	ACC	MCC	F1	MPR	ACC	MCC	F1	MPR
KNN	713.38	728.43	654.36	**989.78**	736.13	750.35	680.50	**1037.45**
SVM_{poly}	570.90	499.61	327.22	**628.38**	578.05	506.24	323.94	**639.84**
SVM_{rbf}	1160.28	1072.46	920.89	**1341.16**	1223.67	1126.76	971.41	**1428.75**
DT	837.02	819.10	610.47	**928.98**	876.08	857.92	640.48	**977.47**
RF	1081.26	1026.09	891.61	**1443.78**	1141.98	1096.93	938.17	**1551.03**
K-ELM	1292.10	1166.45	1056.66	**1545.77**	1376.30	1233.93	1099.00	**1661.93**
FDNN	1054.45	1142.01	812.74	**1301.15**	1121.44	1217.16	846.10	**1389.86**
GA–NN	1135.32	1060.57	762.05	**1322.42**	1206.40	1131.83	803.91	**1413.81**

Bold numbers indicate the best results.

Table 11
The results of t-test under L/S-A-0.

	DJIA	S&P500	HSI	N225	SSE
ACC($\times 10^{-3}$)	**0.198*****	**0.052*****	**0.004*****	**0.004*****	**0.305*****
MCC($\times 10^{-3}$)	**0.026*****	**0.238*****	**0.014*****	**0.014*****	**0.209*****
F1($\times 10^{-3}$)	**0.039*****	**0.073*****	**0.002*****	**0.001*****	**0.004*****

p-value < 0.01: ***, p-value < 0.05: **, p-value < 0.1: *.

Table 12
The results of t-test under L/S-M-0.

	DJIA	S&P500	HSI	N225	SSE
ACC($\times 10^{-3}$)	**0.097*****	**0.072*****	**0.005*****	**0.005*****	**0.327*****
MCC($\times 10^{-3}$)	**0.014*****	**0.277*****	**0.026*****	**0.019*****	**0.241*****
F1($\times 10^{-3}$)	**0.043*****	**0.100*****	**0.003*****	**0.002*****	**0.006*****

p-value < 0.01: ***, p-value < 0.05: **, p-value < 0.1: *.

From the above experimental results, it can be concluded that MPR is more effective in class selection than the classification indicators. MPR can choose a model with higher profitability.

Next, we analyze the correlation between the metrics (MPR and classification metrics) and the model's profitability.

5.2. Consistency analysis

We analyze the consistency between the metric and its profitability when evaluating a given model. The consistency is measured by Pearson Correlation Coefficient (PCC). A higher level of PCC value between these two indicates a better consistency. First, the results of the optimal model, MPR, classification metrics and profit for each model on each subset are collected in the form $\langle acc, mcc, f1, mpr, profit\rangle$. Then, PCC between metric and its profitability is calculated. In this case, the profit estimated by simulated trading is equal to the model's profitability. We choose the mostly used strategy L/S-M-0 as our strategy in simulated trading.

For example, PCC between ACC and profit is calculated basing on the result pairs $\langle acc, profit\rangle$. We get fifty (50 subsets) PPC results for each metric, and each of them is calculated basing on eight result pairs (eight models) on each subset. PCC between the metric and its profitability will be different due to that the distribution of profit differs on each subset. In order to study the dynamic distribution of PCC, we draw the box chart of PCC results on each data set.

Figs. 4 to 8 show the distribution of PCC between each metric and the model's profitability. In these pictures, we use the metric name to denote the PCC results between the metric and the model's profitability. For instance, we use MPR to denote $PCC_{MPR,profit}$. Firstly, PCC between MPR and model's profitability is close to 1. It can be concluded that MPR strongly correlates with model's profitability. This strong correlation is consistent with the theoretical proof before. Secondly, the classification metrics sometimes have a weak correlation with profit. That PCC is lower than 0.5 indicates a weak correlation. Intuitively, it has a great chance to suffer from profit bias when the metric has a weak correlation with model's profitability. The main reason for this weak correlation is that the test set contains several data points with higher level profit than others. However, this is beyond the range of this paper's main concerns. We will not discuss the details of the conditions which trigger this weak correlation. Finally, we conclude that the order of consistency level is $MPR > ACC > MCC > F1$. The higher PCC indicates the better consistency and means lower possibility to suffer from profit bias. Therefore, MPR is the least possible to suffer from profit bias.

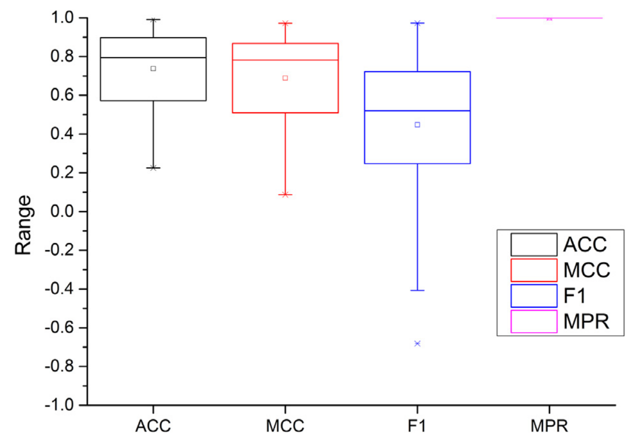

Fig. 4. The box chart of PCC results between metrics and profits on DJIA.

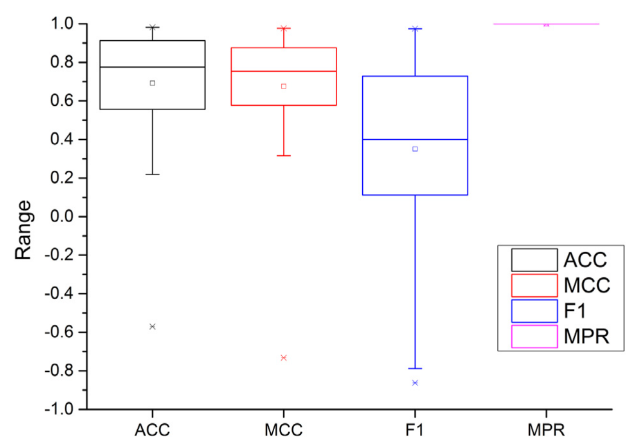

Fig. 5. The box chart of PCC results between metrics and profits on S&P500.

5.3. Effects in model selection

We use Selection Failure Rate (SFR) to evaluate the effects of a metric in model selection. SFR is the rate when the model has a higher level of metric value but a lower level of profit. SFR_{ACC} is calculated basing on the result pairs A $\langle acc1, profit1\rangle$ and B $\langle acc2, profit2\rangle$. Figs. 9 to 13

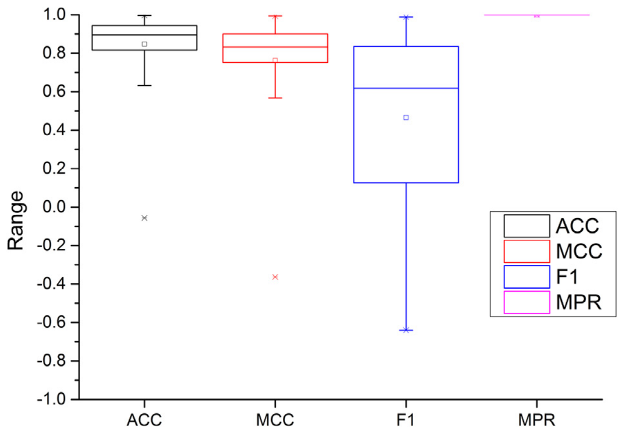

Fig. 6. The box chart of PCC results between metrics and profits on N225.

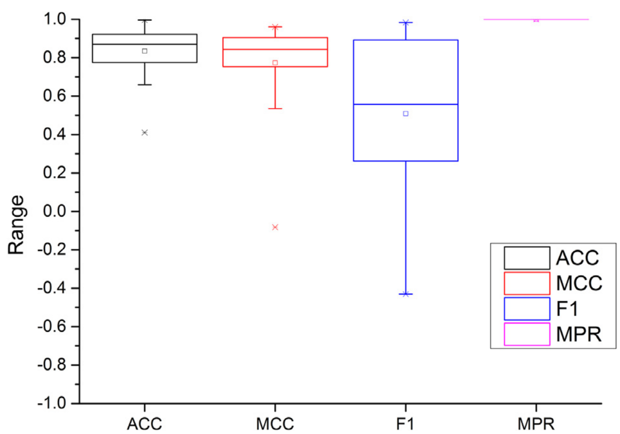

Fig. 7. The box chart of PCC results between metrics and profits on HSI.

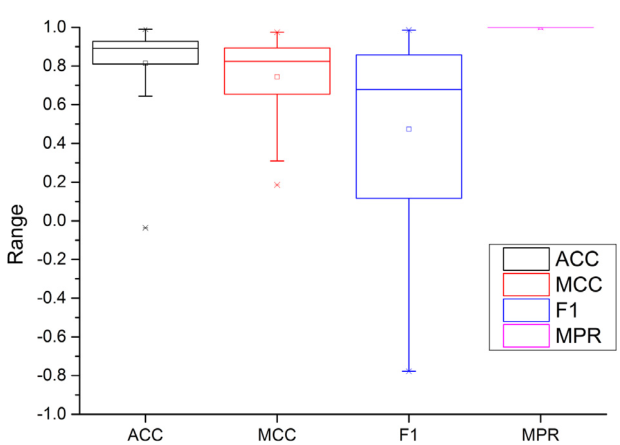

Fig. 8. The box chart of PCC results between metrics and profits on SSE.

show the distributions of SFR between each metric and the model's profitability. We use the name of the metric to denote SFR_{metric}. For instance, we use MPR to denote SFR_{MPR}. Firstly, MPR is unlikely to suffer from profit bias, because SFR_{MPR} distributes near zero. The model has the lowest level of SFR. Hence MPR is unlikely to make incorrect selections in model selection. Secondly, SFR of classification metrics are greater than 0.1 on most subsets. It sounds unreliable when a metric has over 10% of chance making incorrect selections in model

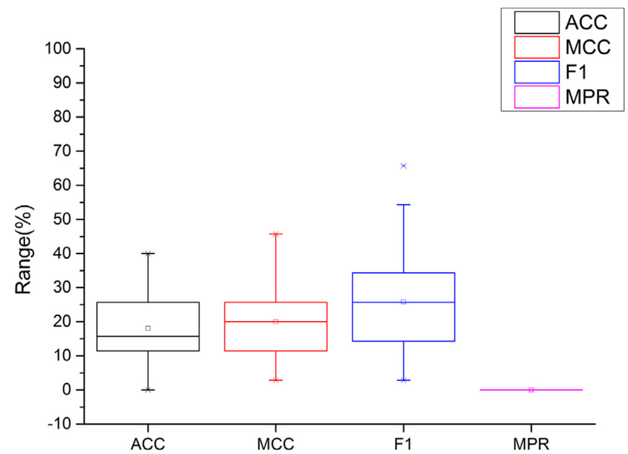

Fig. 9. The box chart of SFR results between metrics and profits on DJIA.

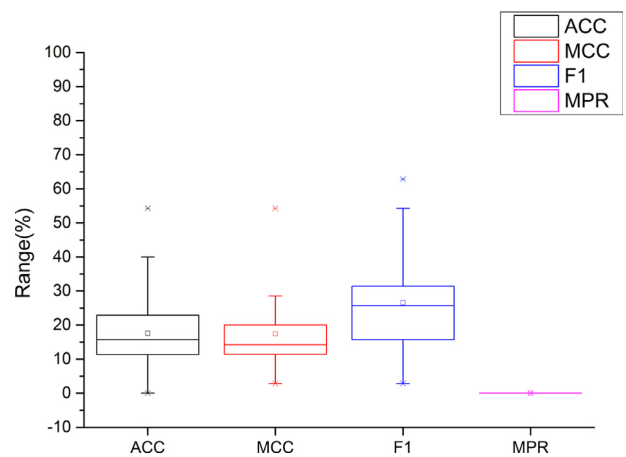

Fig. 10. The box chart of SFR results between metrics and profits on S&P500.

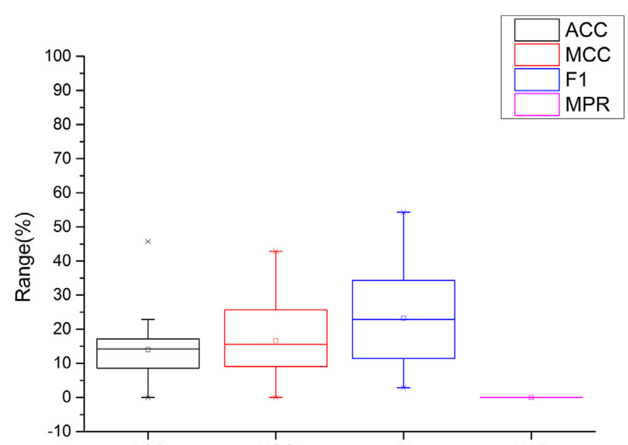

Fig. 11. The box chart of SFR results between metrics and profits on N225.

selection. Thirdly, we get the same effective metric order, $MPR > ACC > MCC > F1$.

In summary, MPR has a lower level of SFR than the classification metrics. Our results indicate that MPR is more effective than the classification metrics in model selection.

Next, we test the consistency between a given metric and the model's profitability under different strategies.

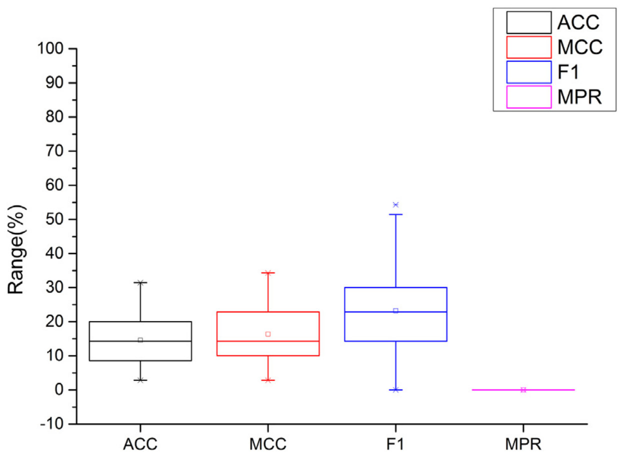

Fig. 12. The box chart of SFR results between metrics and profits on HSI.

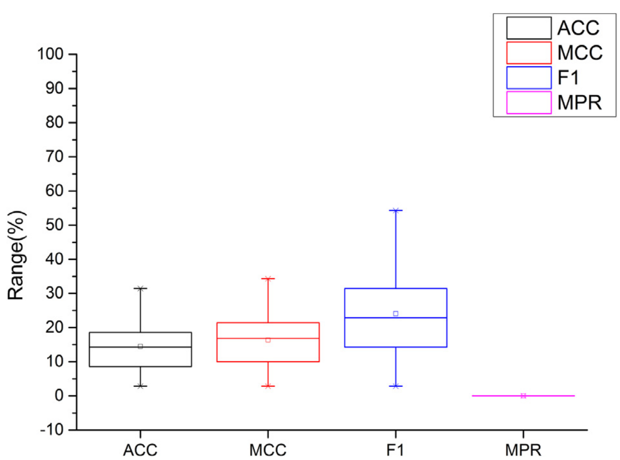

Fig. 13. The box chart of SFR results between metrics and profits on SSE.

Table 13
The parameters in simulated trading.

Trade method	Profit method	Transaction cost (%)
L/S	$[A, M]$	$[0, 0.1, 0.2, 0.3]$

5.4. Strategy free

An effective metric for individual stock trend prediction should be strategy free. It means that the metric is always consistent with the profit estimation when using different strategies to estimate the model's profitability. The definition of strategy indicates that the model's profitability is correlated to Trading Method (TM), Profit Method (PM) and Transaction Cost (TC). We design eight strategies to represent various strategies in real life. We use the combination of settings, listed in Table 13, to form strategies. In order to test whether the metric is strategy free, we design an experiment with three steps. First, eight models are trained with randomly chosen parameters on each training set of five chosen subsets (DJIA-23, S&P500-28, HIS-9, N225-24 and SSE-39). Second, the classification metrics and MPR are calculated on the testing set. Finally, PCC between each metric and the profitability under each strategy is calculated on the testing set.

Table 14 lists the PCC values of eight strategies. PCC of ACC and the profit under L/S-A-0 strategy is listed on the second row and the second column. An effective metric should have better consistency with profit. The higher PCC indicates the metric has better correlation with profit. It will lead to higher consistency.

Table 14
The PCC results under different strategies.

Metric	A-0	A-0.1	A-0.2	A-0.3
ACC	0.70945	0.68393	0.65563	0.62426
MCC	0.68488	0.66116	0.63480	0.60556
F1	0.38452	0.35729	0.63480	0.60556
MPR	**1.00000**	**0.99937**	**0.99728**	**0.99345**
Metric	M-0	M-0.1	M-0.2	M-0.3
ACC	0.67525	0.67521	0.67517	0.67513
MCC	0.65332	0.65329	0.65326	0.65322
F1	0.36586	0.36584	0.36582	0.36580
MPR	**0.99559**	**0.99558**	**0.99558**	**0.99557**

Bold numbers indicate the best results.

Table 15
The evaluation results of models on DJIA-23.

Dataset	Model	ACC	MCC	F1	MPR (%)	L/S-A-0	L/S-M-0
DJIA-23	KNN	0.63	0.25	0.69	0.02	53.54	47.47
	SVM_{poly}	0.53	−0.16	0.70	0.12	361.65	361.96
	SVM_{rbf}	0.57	0.00	0.72	0.14	434.58	437.80
	DT	0.47	−0.07	0.50	−0.07	−200.36	−204.42
	RF	0.67	0.39	0.64	0.17	516.69	523.86
	K-ELM	0.67	0.32	0.71	0.05	141.42	136.17
	FDNN	0.67	0.39	0.64	**0.29**	**855.44**	**886.48**
	GA–NN	**0.70**	**0.39**	**0.77**	0.12	363.82	364.12

Bold numbers indicate the best results.

Our results show two major findings. First, PCC between MPR and profit is the highest among four metrics no matter what strategy is taken. Second, $PCC_{MPR,profit}$ is nearly equal to one while other PCC values are below 0.71. This means MPR has stronger correlation with the profit no matter what strategy is chosen. These results suggest that MPR is strategy free.

5.5. Demonstration of usage

We demonstrate how to use MPR to evaluate models in individual stock tend prediction. Five subsets in the above Section are chosen as the dataset. The optimal models are chosen from the parameters in Table 5. Tables 15 to 19 list the evaluation results of the different models on five subsets. DJIA-23 is the 23rd subset of DJIA dataset, SVM_{rbf} is the SVM model with RBF kernel, and SVM_{poly} is the SVM model with POLY kernel. We use L/S-M-0 to denote the profit estimated under L/S-M-0 strategy and L/S-A-0 to denote the profit estimated under L/S-A-0 strategy.

We have three main conclusions. Firstly, MPR is consistent with the model's profitability. The optimal model selected by MPR has the best profitability. For instance, FDNN has 0.29% MPR on DJIA-23 and over 800 units of profits, which is the highest among the eight models, under both strategies. Secondly, the classification metrics are not consistent with the model's profitability in most cases. For instance, the model with the highest ACC value on DJIA-23, S&P500, and SSE-39 do not achieve the highest profits. Thirdly, the using of multiple classification metrics may not improve the efficiency in model selection. On HSI-9, ACC, MCC, and F1 select K-ELM, FDNN, and GA–NN as the optimal model, respectively. The using of multiple metrics can lead to confusing results, making model selection more inefficient.

6. Conclusion

This paper proposed a new metric, Mean Profit Rate (MPR), to overcome the profit bias in evaluating models for stock trend prediction. MPR is the expectation of profit rate on each trend prediction. We theoretically proved that MPR is proportional to the model's profitability. Therefore, MPR is a good indicator of model's profitability due to the consistency. To evaluate our metric, we compared our MPR with the classification metrics in model selection. The experimental results

Table 16
The evaluation results of models on S&P500-28.

Subset	Model	ACC	MCC	F1	MPR (%)	L/S-A-0	L/S-M-0
S&P500-28	KNN	0.53	0.07	0.50	−0.17	−506.41	−540.91
	SVM_{poly}	0.33	−0.31	0.44	−0.59	−1771.53	−1665.17
	SVM_{rbf}	0.43	0.00	0.60	−0.40	−1199.64	−1174.58
	DT	0.37	−0.25	0.39	−0.70	−2088.03	−1924.85
	RF	0.57	0.14	0.55	0.14	419.31	377.30
	K-ELM	0.63	0.28	0.62	**0.25**	**763.15**	**740.29**
	FDNN	**0.67**	**0.36**	**0.67**	0.22	668.61	639.22
	GA–NN	0.60	0.25	0.62	−0.10	−314.63	−357.78

Bold numbers indicate the best results.

Table 17
The evaluation results of models onHSI-9.

Subset	Model	ACC	MCC	F1	MPR (%)	L/S-A-0	L/S-M-0
HSI-9	KNN	0.50	0.03	0.52	0.20	594.70	534.33
	SVM_{poly}	0.57	0.11	0.63	0.33	976.75	944.69
	SVM_{rbf}	0.50	−0.03	0.57	0.20	611.93	552.47
	DT	0.50	0.09	0.44	0.28	828.93	784.28
	RF	0.43	0.15	0.11	−0.16	−475.48	−535.84
	K-ELM	**0.63**	0.27	0.67	**0.62**	**1845.56**	**1938.45**
	FDNN	**0.63**	**0.39**	0.59	0.37	1118.59	1100.94
	GA–NN	0.60	0.14	**0.68**	0.34	1033.32	1106.63

Bold numbers indicate the best results.

Table 18
The evaluation results of models on N225-24.

Subset	Model	ACC	MCC	F1	MPR (%)	L/S-A-0	L/S-M-0
N225-24	KNN	0.57	0.13	0.55	0.24	723.02	614.84
	SVM_{poly}	0.53	0.11	0.61	0.38	1145.63	1074.66
	SVM_{rbf}	0.53	0.09	0.59	1.00	2991.28	3329.92
	DT	0.63	0.27	0.48	0.96	2890.21	3195.19
	RF	0.57	0.16	0.61	−0.03	−95.89	−219.94
	K-ELM	**0.70**	**0.40**	**0.67**	1.01	3040.37	3393.96
	FDNN	**0.70**	**0.40**	0.64	**1.37**	**4112.87**	**4912.27**
	GA–NN	0.67	0.34	**0.67**	1.24	3726.99	4347.74

Bold numbers indicate the best results.

Table 19
The evaluation results of models on SSE-39.

Subset	Model	ACC	MCC	F1	MPR (%)	L/S-A-0	L/S-M-0
SSE-39	KNN	0.70	0.34	0.77	0.38	1150.32	1200.66
	SVM_{poly}	0.63	0.20	0.72	0.42	1254.61	1318.07
	SVM_{rbf}	0.67	0.23	0.76	0.23	693.49	700.27
	DT	0.50	−0.02	0.57	0.20	598.77	599.60
	RF	0.60	0.29	0.60	0.25	742.90	753.29
	K-ELM	0.63	0.17	0.73	0.23	693.24	700.17
	FDNN	**0.70**	**0.31**	**0.80**	0.40	1199.95	1256.36
	GA–NN	0.67	0.26	0.75	**0.43**	**1289.16**	**1357.27**

Bold numbers indicate the best results.

suggested MPR can serve as a better selector in model selection than the classification metrics. The results of consistency analysis suggested MPR has higher Pearson Correlation Coefficient (PCC) with model's profitability and lower Selection Failure Rate (SFR) than the classification metrics. The results of PCC between metrics and the model's profitability under various strategies indicate MPR is insensitive to the change of strategies. We also demonstrated the effectiveness of MPR in evaluating stock prediction models. The findings of this study may shed some light on effectively evaluating models for individual stock trend prediction.

Acknowledgments

This research was funded in part by the National Social Science Fund of China No. 2016ZDA055, and in part by the Discipline Building Plan in 111 Base No. B08004.

References

Aha, D.W., Kibler, D., Albert, M.K., 1991. Instance-based learning algorithms. Mach. Learn. 6 (1), 37–66.
Akita, R., Yoshihara, A., Matsubara, T., Uehara, K., 2016. Deep learning for stock prediction using numerical and textual information. In: Computer and Information Science (ICIS), 2016 IEEE/ACIS 15th International Conference on. IEEE, pp. 1–6.
Altman, N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. Amer. Statist. 46 (3), 175–185.
Atsalakis, G.S., Valavanis, K.P., 2009a. Forecasting stock market short-term trends using a neuro-fuzzy based methodology. Expert Syst. Appl. 36 (7), 10696–10707.
Atsalakis, G.S., Valavanis, K.P., 2009b. Forecasting stock market short-term trends using a neuro-fuzzy based methodology. Expert Syst. Appl. 36 (7), 10696–10707.
Ballings, M., Van den Poel, D., Hespeels, N., Gryp, R., 2015. Evaluating multiple classifiers for stock price direction prediction. Expert Syst. Appl. 42 (20), 7046–7056.
Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. J. Comput. Sci. 2 (1), 1–8.
Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. Classification and Regression Trees. CRC press.
Brownstone, D., 1996. Using percentage accuracy to measure neural network predictions in stock market movements. Neurocomputing 10 (3), 237–250.
Chang, P.-C., Liu, C.-H., 2008. A tsk type fuzzy rule based system for stock price prediction. Expert Syst. Appl. 34 (1), 135–144.
Chang, P.-C., Liu, C.-H., Lin, J.-L., Fan, C.-Y., Ng, C.S., 2009. A neural network with a case based dynamic window for stock trading prediction. Expert Syst. Appl. 36 (3), 6889–6898.
Choudhry, R., Garg, K., 2008. A hybrid machine learning system for stock market forecasting. World Acad. Sci. Eng. Technol. 39 (3), 315–318.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297.
De Fortuny, E.J., De Smedt, T., Martens, D., Daelemans, W., 2014. Evaluating and understanding text-based stock price prediction models. Inf. Process. Manage. 50 (2), 426–441.
Deng, Y., Bao, F., Kong, Y., Ren, Z., Dai, Q., 2017a. Deep direct reinforcement learning for financial signal representation and trading. IEEE Trans. Neural Netw. Learn. Syst. 28 (3), 653–664.
Deng, Y., Ren, Z., Kong, Y., Bao, F., Dai, Q., 2017b. A hierarchical fused fuzzy deep neural network for data classification. IEEE Trans. Fuzzy Syst. 25 (4), 1006–1012.
Ding, X., Zhang, Y., Liu, T., Duan, J., 2015. Deep learning for event-driven stock prediction. In: Ijcai. pp. 2327–2333.
Hagenau, M., Liebmann, M., Neumann, D., 2013. Automated news reading: Stock price prediction based on financial news using context-capturing features. Decis. Support Syst. 55 (3), 685–697.
Ho, T.K., 1995. Random decision forests. In: Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on, Vol. 1. IEEE, pp. 278–282.
Hsieh, T.-J., Hsiao, H.-F., Yeh, W.-C., 2011. Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm. Appl. Soft Comput. 11 (2), 2510–2525.
Hu, Z., Liu, W., Bian, J., Liu, X., Liu, T.-Y., 2018. Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. In: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. ACM, pp. 261–269.
Huang, W., Nakamori, Y., Wang, S.-Y., 2005. Forecasting stock market movement direction with support vector machine. Comput. Oper. Res. 32 (10), 2513–2522.
Huang, C.-J., Yang, D.-X., Chuang, Y.-T., 2008. Application of wrapper approach and composite classifier to the stock trend prediction. Expert Syst. Appl. 34 (4), 2870–2878.
Inthachot, M., Boonjing, V., Intakosum, S., 2016. Artificial neural network and genetic algorithm hybrid intelligence for predicting thai stock price index trend. Comput. Intell. Neurosci..
Kaastra, I., Boyd, M., 1996a. Designing a neural network for forecasting financial and economic time series. Neurocomputing 10 (3), 215–236.
Kaastra, I., Boyd, M., 1996b. Designing a neural network for forecasting financial and economic time series. Neurocomputing 10 (3), 215–236.
Kara, Y., Boyacioglu, M.A., Baykan, Ö.K., 2011a. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. Expert Syst. Appl. 38 (5), 5311–5319.
Kara, Y., Boyacioglu, M.A., Baykan, Ö.K., 2011b. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. Expert Syst. Appl. 38 (5), 5311–5319.
Kim, K.-j., Han, I., 2000. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. Expert Syst. Appl. 19 (2), 125–132.
Lee, M.-C., 2009a. Using support vector machine with a hybrid feature selection method to the stock trend prediction. Expert Syst. Appl. 36 (8), 10896–10904.
Lee, M.-C., 2009b. Using support vector machine with a hybrid feature selection method to the stock trend prediction. Expert Syst. Appl. 36 (8), 10896–10904.
Lee, H., Surdeanu, M., MacCartney, B., Jurafsky, D., 2014. On the importance of text analysis for stock price prediction. In: LREC. pp. 1170–1175.

Leung, M.T., Daouk, H., Chen, A.-S., 2000. Forecasting stock indices: a comparison of classification and level estimation models. Int. J. Forecast. 16 (2), 173–190.
Li, Q., Jiang, L., Li, P., Chen, H., 2015. Tensor-based learning for predicting stock movements. In: AAAI. pp. 1784–1790.
Li, J., Tsang, E.P., 1999. Improving technical analysis predictions: An application of genetic programming. In: FLAIRS Conference. pp. 108–112.
Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., Chen, Y., 2014. The effect of news and public mood on stock movements. Inform. Sci. 278, 826–840.
Li, X., Xie, H., Wang, R., Cai, Y., Cao, J., Wang, F., Min, H., Deng, X., 2016. Empirical analysis: stock market prediction via extreme learning machine. Neural Comput. Appl. 27 (1), 67–78.
Lin, Y., Guo, H., Hu, J., 2013. An svm-based approach for stock market trend prediction. In: Neural Networks (IJCNN), The 2013 International Joint Conference on. IEEE, pp. 1–7.
Luo, L., Chen, X., 2013. Integrating piecewise linear representation and weighted support vector machine for stock trading signal prediction. Appl. Soft Comput. 13 (2), 806–816.
Mizuno, H., Kosaka, M., Yajima, H., Komoda, N., 1998. Application of neural network to technical analysis of stock market prediction. Stud. Inform. Control 7 (3), 111–120.
Moody, J., Saffell, M., 2001. Learning to trade via direct reinforcement. IEEE Trans. Neural Netw. 12 (4), 875–889.
O'Connor, N., Madden, M.G., 2006. A neural network approach to predicting stock exchange movements using external factors. Knowl.-Based Syst. 19 (5), 371–378.
Oh, K.J., Kim, K.-j., 2002. Analyzing stock market tick data using piecewise nonlinear model. Expert Syst. Appl. 22 (3), 249–255.
Patel, J., Shah, S., Thakkar, P., Kotecha, K., 2015a. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Syst. Appl. 42 (1), 259–268.
Patel, J., Shah, S., Thakkar, P., Kotecha, K., 2015b. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Syst. Appl. 42 (1), 259–268.
Qian, B., Rasheed, K., 2007. Stock market prediction with multiple classifiers. Appl. Intell. 26 (1), 25–33.
Saad, E.W., Prokhorov, D.V., Wunsch, D.C., 1998. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. IEEE Trans. Neural Netw. 9 (6), 1456–1470.
Schumaker, R.P., Chen, H., 2009. Textual analysis of stock market prediction using breaking financial news: The azfin text system. ACM Trans. Inf. Syst. (TOIS) 27 (2), 12.
Shynkevich, Y., McGinnity, T., Coleman, S., Belatreche, A., 2015. Predicting stock price movements based on different categories of news articles. In: Computational Intelligence, 2015 IEEE Symposium Series on. IEEE, pp. 703–710.
Teixeira, L.A., De Oliveira, A.L.I., 2010a. A method for automatic stock trading combining technical analysis and nearest neighbor classification. Expert Syst. Appl. 37 (10), 6885–6890.
Teixeira, L.A., De Oliveira, A.L.I., 2010b. A method for automatic stock trading combining technical analysis and nearest neighbor classification. Expert Syst. Appl. 37 (10), 6885–6890.
Tsai, C.-F., Hsiao, Y.-C., 2010. Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. Decis. Support Syst. 50 (1), 258–269.
Wang, Y., Choi, I.-C., 0000. Market index and stock price direction prediction using machine learning techniques: an empirical study on the kospi and hsi, arXiv preprint arXiv:1309.7119.
Wen, Q., Yang, Z., Song, Y., Jia, P., 2010. Automatic stock decision support system based on box theory and svm algorithm. Expert Syst. Appl. 37 (2), 1015–1022.
Wu, M.-C., Lin, S.-Y., Lin, C.-H., 2006a. An effective application of decision tree to stock trading. Expert Syst. Appl. 31 (2), 270–274.
Wu, M.-C., Lin, S.-Y., Lin, C.-H., 2006b. An effective application of decision tree to stock trading. Expert Syst. Appl. 31 (2), 270–274.
Wu, C.-H., Tzeng, G.-H., Goo, Y.-J., Fang, W.-C., 2007. A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. Expert Syst. Appl. 32 (2), 397–408.
Zhai, Y., Hsu, A., Halgamuge, S.K., 2007. Combining news and technical indicators in daily stock price trends prediction. In: International Symposium on Neural Networks. Springer, pp. 1087–1096.