Journal of Computational Science & Engineering 58 (2022) 20-28

Available at http://www.asocse.org ISSN 1710-4068

Multi-index quantitative stock selection strategy based on machine learning

Shupei Du[1], Li Wu[1], Jieyu Ding [1,2*]

[1]School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China

[2]Center for Computational Mechanics and Engineering Simulation, Qingdao University, Qingdao 266071, China

Abstract: **With the gradual development and improvement of China's financial market, investors are more and more keen to predict and analyze stock prices. Compared with traditional fundamental technical analysis, modern analytical trading tools such as quantitative trading and machine learning based on big data show unique advantages. This paper selects CSI A-Shares excluding all ST shares in all trading days in 2021 for research, uses stock volume price technical indicators-opening price, closing price, MACD, MA, turnover, volume ratio, etc., to screen the daily stock pool, and then uses machine learning algorithms such as XGBoost to predict the rise and fall of the selected stocks, and finally introduces activity indicators such as rising speed, trading volume, capital inflow and so on. Combined with the Topsis comprehensive evaluation method to score the stock, select the three stocks with the highest score to simulate buying, and calculate the rate of return within five trading days after the position. At the end of the paper the model is summarized and prospected.**

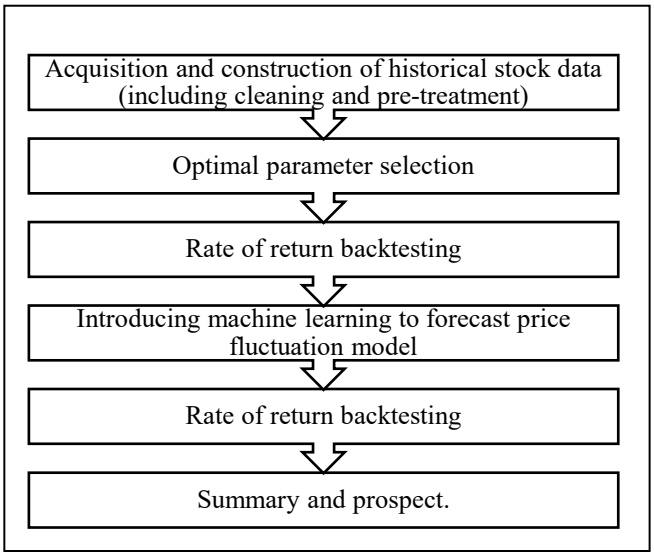
Keywords: Quantitative trading; Multiple indicators; XGBoost; Back-tested returns

Received 03 September 2022, Revised 10 September 2022, Accepted 17. September 2022

*Corresponding Author: Jieyu Ding, djy@qdu.edu.cn

1. Introduction

In today's secondary financial market, the daily trading volume of CSI A-shares has increased significantly, breaking the record since 2015 with 44 consecutive trading days in 2021. The upsurge of this investment is attracting more and more investors to pour into the market. So how to grasp more data and information in this huge secondary market, make use of the intrinsic value of transaction data to obtain higher profits has become a hot topic for investors to discuss. The price of stock has the characteristics of volatility and uncertainty, and will also be affected by macroeconomic factors such as economic policy, panic emotion and so on. The common models for analyzing stock prices include traditional stock forecasting models, such as Time series model, Hidden Markov Mode; and new stock forecasting models, such as machine learning and XGboost and so on, which are based on common ensemble learning algorithms. In this paper, we first compare and analyze the current stock price forecasting models, and then put forward a multi-index quantitative model based on machine learning. Finally, we calculate the return rate of the stock data, which verifies the feasibility of the model.



2. Summary of the current research situation at home and abroad

The traditional stock price forecasting model is mainly based on mathematical statistics and probability theory.

2022@ The American Computational Science Society. All rights reserved.

Since financial data is a data with time series characteristics, it contains many factors such as trend changes, seasonal changes and periodic changes. Robert, an economist, put forward ARCH model in 1980s and predicted the British inflation index at that time[1]. Majumder et al. obtained the highest average prediction accuracy of ARIMA time series algorithm by using a variety of stock prediction algorithms such as ARIMA, linear model and Holt-Winter method[2]. The fluctuation of stock price is a nonlinear movement affected by many factors such as economy, politics and even investor sentiment. The data has the characteristics of strong volatility, chaos and ambiguity[3]. The prediction model based on traditional time series statistics The generalization ability to the sample is weak, and it cannot well predict the new data that appears outside the sample. With the development of machine learning and deep learning, new stock prediction models are gradually improving. Support vector machine (SVM) theory is a new type of learning machine proposed by Vapnik in the 1990s under the framework of statistical learning theory[4]. Compared with the stock price forecasting model based on mathematical statistics, SVM can solve nonlinear and multi-dimensional data problems very well, and its generalization ability is stronger. Kim applied SVM to stock forecasting, and demonstrated that the method is better than the traditional neural network method through experiments[5]. Decision tree is a mature data mining technology, and many integrated learning algorithms derived from decision tree can learn and predict nonlinear data sets. XGBoost is an improved algorithm based on Gradient Boosting Decision Tree (GBDT), which can effectively build boosted trees for classification or regression. Based on the Cart decision tree, Wang Yu et al. used the method of boosting cascaded multiple decision trees to establish a stock prediction model to solve the problem of fit[6]. It shows that the algorithm can process large-scale data, has high prediction accuracy and strong noise tolerance.

3.Construction of multiple index system

Table 1: Symbolic description

Symbol	Significance	Symbol	Significance
N	trading days	w	Weight
MA_{20}	20-day average	P_N	Sum of N-day prices
DIF	Difference	$MACD$	Moving Average Convergence and Divergence
DEA	Difference Exponential Average	S	The total number of hands currently traded
EMA_{12}	12-day exponential moving average	S_t	The current cumulative opening time (minutes)
EMA_{26}	26-day exponential moving average	Va_5	Average volume per minute for the past 5 days
L_b	Quantity ratio	G	Total number of shares currently traded
T_o	Turnover rate	C_g	Number of outstanding shares
C_v	Circulating market value	V_N	Volume
μ	Rate of return	Z_s	Rising speed
NET_2	Net inflow of funds within 2 days	$Score$	Comprehensive evaluation of stock score

Note: if not specified, N=1.

Stock price is affected by many technical indicators and even macro-environmental factors, so it is inherently volatile and uncertain. It is precisely because of its volatility that it brings great returns to short-term investors with both risks and coexistence. Short-term investors are not as stable as long-term investors, they usually consider the cost of time, and need to constantly adjust their positions and candidate stock pools to update their asset allocations. therefore, short-term investors need to spend a lot of time on stock selection, and the labor complexity is extremely high.

The stock price prediction model based on machine learning only predicts the price of one stock, it needs to learn the data of this stock for several months or even several years to make predictions, and its calculation time is large, so it cannot predict all 4,000 stocks of CSI A-Shares. Therefore, this paper first constructs a multi-index system, so as to screen all the stocks at once.

3.1 Data acquisition

This paper uses the quantitative trading open platform JoinQuant to build the stock database, which can dynamically select all the stocks listed by CSI A-shares on the return date. Since ST shares and sub-new shares have high company operating risks and many uncertain factors, they are excluded when constructing the back-testing stock database.

Considering the maximization of short-term returns for short-term investors, this paper should select indicators that can represent the short-term activity of stocks:

2022@ The American Computational Science Society. All rights reserved.

$$MA_{20} = \frac{P_N}{N} \quad (3.1)$$

$$DIF = EMA_{12} - EMA_{26} \quad (3.2)$$

$$EMA_{12} = EMA_{12y} * \frac{11}{13} + P_N * \frac{2}{13} \quad (3.3)$$

$$EMA_{26} = EMA_{26y} * \frac{25}{27} + P_N * \frac{2}{27} \quad (3.4)$$

$$L_b = \frac{S}{S_t * Va_5} \quad (3.5)$$

$$T_o = \frac{G}{C_g} * 100\% \quad (3.6)$$

3.2 Data preprocessing of indicator database

Due to the status of suspension (trading suspension) in the stock trading market, there are no indicators such as volume ratio and turnover rate for individual stocks on the day. We can take the following three options:

(1) Completely use the data of the previous normal trading day to replace the data of the day of trading suspension;

(2) Except that the price (opening price, closing price, etc.) adopts the data of the previous normal trading day, others are set to 0;

(3) Directly delete the trading day data of the current day (this method may cause a mismatch of the trading date dimension when backtesting the overall stock).

Then carry on the positive treatment to the index, in order to use for the later comprehensive evaluation model.

Step1: Forward the four indicators, that is, convert all extremely small indicators into very large indicators, and obtain a positive matrix.

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1j} \\ \vdots & \ddots & \vdots \\ x_{i1} & \cdots & x_{ij} \end{pmatrix} \quad (3.7)$$

Step2: Standardize the positive matrix X to get the normalized matrix Z, where:

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2}} \quad (3.8)$$

If there is a negative number in the matrix Z, then another standardization method is used for X, and the normalized matrix $\tilde{z}$ is obtained, where:

$$\tilde{z}_{ij} = \frac{x_{ij} - \min\{x_{ij}, x_{2j}, ..., x_{nj}\}}{\max\{x_{ij}, x_{2j}, ..., x_{nj}\} - \min\{x_{ij}, x_{2j}, ..., x_{nj}\}} \quad (3.9)$$

The following normalization matrices are unified as Z.

3.3 Setting of Strategy Parameter Values (Grid Search)

Usually, the aggregation characteristics of MACD can reflect the state of both longs and shorts. When DIF crosses DEA from below and is greater than 0, it is a buy signal, indicating that the stock will be idling in a short period of time, and the possibility of profit is high; the closing price of the day needs to be Satisfy greater than MA_{20}: The moving average support theory commonly used by short-term customers; L_b is less than 1, indicating that the stock is in a shrinking stage, with low popularity. 1~1.5: Volume and popularity; potential main accumulation; 1.5~2.5: Moderate heavy volume; More than 2.5: stocks are extremely active, and high yields are also accompanied by high risks.

Since there are still many daily stocks that meet the above index conditions, and the computational time cost of the machine learning model is relatively large, this article gives a better graphical strategy based on the experience of historical K-line charts—that is, it opened higher yesterday, but finally closed negative, satisfied:

$(P_k - P_s) / P_k < w$ (3.10)

$Z_N > 0\ \%$ (3.11)

2022@ The American Computational Science Society. All rights reserved.

Table 2: Indicator Strategy Table

	Strategy 1	Strategy 2	Strategy 3	Strategy 4	Strategy 5	Strategy 6	Strategy 7
$MACD$	>0	>0	>0	>0	>0	Not limited	Not limited
P_N	$>MA_{20}$	$>MA_{20}$	$>MA_{20}$	$>MA_{20}$	Not limited	$>MA_{20}$	Not limited
L_b	>1.1	>1	>1.1	>1.2	>1.2	>1.2	>1.5
w	<1%	<0.5%	<0.8%	<1%	<2%	<2%	<2%
Z	>0	>0	>0	>0	>2	>1	Not limited
T_o	>1	>1	>1.5	>2	>2	>2	Not limited
P_N	<60	<60	<60	<60	<60	<60	<60
C_v	<300	<300	<300	<300	<300	<300	Not limited

3.4 Backtesting return

Define backtest date: January 4, 2021 ~ December 31, 2021

The red line is the strategy yield curve and the blue line is the benchmark line of the Shanghai Composite Index.

Due to the limited space, this paper only shows the return curve of the four strategies and the selection of individual stocks, and the four strategies are all the back-tested data at the same time period.

Fig1. Strategy 1 Backtesting Chart

Fig3. Strategy 3 Backtesting Chart

Fig2. Strategy 2 Backtesting Chart

Fig4. Strategy 4 Backtesting Chart

It can be seen from the 2021 backtest curve that the annualized returns of strategies 1-4 are 25.1%, 75.5%, 58.4%, 0.4% respectively. Among them, strategies 2 and 3 show good annualized returns, and the final return of strategy 1 is 25.1%, which has a large retracement in the time period from September 2021 to October 2021, which is a point where the model can be further improved. Strategy 4 enhances the limitation of volume ratio and turnover rate, and expands the search scope of weight, which makes the range of stocks selected by the strategy too large and imprecise, so it does not achieve a good annualized rate of return.

For example, the results of backtesting part of strategy 2:

4. The decision tree-based ensemble learning model predicts ups and downs and comprehensive evaluation (individual stock data + score + income curve)

4.1 Principle of XGBOOST

Kearns and Valiant (1989) proved the equivalence of weak learners and strong learners. The method of combining multiple weak learners to obtain a strong learner is called ensemble learning, which can further improve

2022@ The American Computational Science Society. All rights reserved.

the accuracy of the model. this method is called Boosting. This paper uses the most representative XGBoost in the Boosting algorithm as a stock price prediction mode.

Assume that the probability that each learner h_i is different from the real output data f is ε, that is:

$P(h_i(x) \neq f(x)) = \varepsilon$ (4.1)

If there are n learners, assuming that more than half of the learners make accurate judgments, the ensemble learning conclusion can be considered to be true, that is:

$H(x) = \text{sign}(\sum_{i=1}^{n} h_i(x))$ (4.2)

When the output errors of each learner are independent, the error rate of ensemble learning is:

$P(H(x) \neq f(x)) = \sum_{k=0}^{n/2} C_n^k (1-\varepsilon)^k \varepsilon^{n-k} \leq e^{-\frac{1}{2}n(1-2\varepsilon)^2}$ (4.3)

In Boosting, k trees are constructed, and the overall prediction function of the algorithm model can be expressed in the following form:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x), f_k \in F$$
$$F = \{f(x) = w_q(x)\} \quad (4.4)$$
$$q: x \to T, w \in R^T$$

Table 3: Display of results of backtesting part of strategy 2

Date	Yield curve					
2021/1/7	0.984475	ALLMED	CANNY	Tonghua Dongbao	CSSC	
2021/1/14	1.0138124	DAZHENG PROPERTY				
2021/1/21	0.9449745	Hongchang Electronics				
2021/1/28	1.0900281	Sanquan Foods				
2021/2/4	1.1007467	Limin Shares	GUIBAO TECH	TRIANGLE		
2021/2/18	1.1312374	HALEAD				
2021/2/25	1.1690081	KXME	TIANFU	XINGMINITS	Western Construction	Ningxia Building Materials…
2021/3/4	1.1241572	HON-FLEX	Mingtai Aluminum	OK	Zhiguang	MOSO …
2021/3/8	1.0736263	ZYNP	JINGGONG STEEL BUILDING			
2021/3/15	1.0603402	WEPON	Jetsen	CENTRAL CHINA MEDIA	EASTERN PIONEER	Fuan …

Where F is the set of all weak learners, w is the weight, T represents the number of sub-nodes of the decision tree, Q is the structure of each tree, and the objective function is composed of two parts: the loss function and the complexity:

$$Obj = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$
$$= \sum_{i=1}^{n} L\left(y_i, \hat{y}_t^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) + \text{cons tan } t \quad (4.5)$$

Among them, 0bj is the objective function, $\sum_{i=1}^{n} L(y_i, \hat{y}_i)$ is the loss function, that is, the residual between the predicted value of the model and the true value, and $\sum_{k=1}^{K} \Omega(f_k)$ is the penalty term for the complexity of each learner. In order to prevent the tree model from being too large, the penalty term is composed of the following. The formula is composed: T is the number of leaves, γ, λ are the punishment strength, plus $L2$ regularization:

$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \parallel w \parallel^2$ (4.6)

In 2016, Academician Chen Tianqi proposed the XGBoost model, which is a Boosting algorithm model that improves the GBDT. It has the advantages of high accuracy and fast computing speed. When XGBoost expands the objective function, it uses the second-order Taylor formula. And in order to prevent the occurrence of overfitting, it is necessary to control the complexity of the model, and XGBoost adds regular terms to the objective function. Its objective function is[7]:

$\text{0bj} = \sum_{i=1}^{n} L\left(y_i, \hat{y}_t^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) + constant$ (4.7)

2022@ The American Computational Science Society. All rights reserved.

$\Omega(f_i) = \gamma T + \frac{1}{2}\lambda \parallel w \parallel^2$ (4.8)

Carry out a second-order Taylor expansion of the objective function:

$Obj' = \sum_{i=1}^{n} \left[l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_i(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \Omega(f_t) + \text{cons tan } t$ (4.9)

Where $\sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)}\right)$ is a fixed value, which can be omitted from the optimization loss function together with the constant, and the above formula becomes:

$Obj^{(t)} = \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$ (4.10)

The kth CART decision tree maps the input sample data to the leaf nodes, that is, $f_k(x)$, and the output leaf node number is $q(x)$, and the value is $w_{q(x)}, f_k(x) = w_{q(x)}$,

$$\begin{aligned} Obj^t &= \sum_{i=1}^{n} \left[g_i w_{q(x_j)} + \frac{1}{2} h_i w_{q(x_j)}^2 \right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \\ &= \sum_{j=1}^{T} \left[\left(\sum_{i \in l_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in l_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \\ &= \sum_{j=1}^{T} \left[G_i w_j + \frac{1}{2} \left(H_i + \lambda w_j^2 \right] + \gamma T \end{aligned} \quad (4.11)$$

Among them, $\sum_{i \in l_j} g_i = G_i, \sum_{i \in I_j} h_i = H_i$, the above formula is a univariate quadratic function about w_j, and the optimal value of w_j obtained by derivation of it is:

$w_j^* = -\frac{G_j}{H_j + \lambda}$ (4.12)

$Obj^* = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T$ (4.13)

4.2 Feature Engineering

"Data and features determine the upper limit of machine learning". No matter how powerful the model is, if there is no good feature extraction work, then the model will not learn any valuable information. Feature engineering refers to selecting data features that are beneficial to machine learning from the original data in an engineered way to improve the training effect of the model. The model for predicting the rise and fall of stocks on the next day requires a large number of trading characteristic data of individual stocks over multiple months. In view of the high computational cost of model learning, this paper applies machine learning algorithms to the stocks that have been screened by multiple indicators, and adds the closing price, opening price, highest price, minimum price, trading volume and capital inflow for model learning.

4.3 A model for predicting the rise and fall of stocks based on an improved binary classification

In general, the two-category prediction problem of stock rise and fall will mark the next day's rise as 1 and the next day's fall as -1, but the result of such a simple binary classification is relatively inaccurate. Therefore, this paper further expands the label on the basis of binary classification, marking the next day $Z \geq 3\%$ as 2; $0 < Z < 3\%$ as 1; on the contrary, when $-3\% < Z \leq 0$ as -1, $Z \leq -3\%$ are marked as -2, thus turning it into a four-classification problem.

The following is an example of the stocks screened at one time on April 8, 2021 in Strategy 1:

Table 4: 5-day yield table

2021/4/8	rate of return（%）
APG	15.88
Zuojiang Technology	8.05
ZHESHANG DEVELOPMENT	4.74
CSB	4.68
JIUZHITANG	0.75
ELECSPN	-0.19
BRIGHT DAIRY	-0.32
HISOAR	-0.99
BIOLIGHT	-1.4
Microgate	-4.73
JTL	-7.94
TESMI	-8.33
RUITAI TECHNOLOGY	-12.83

2022@ The American Computational Science Society. All rights reserved.

The average return of the stocks selected by strategy 1 is -0.2% after 5 days, which has many characteristics that cannot be learned by quantitative indicators. We use XGBOOST to learn all its trading data in the last year and predict the rise and fall of the next day.

Table 5: Probability prediction table of ups and downs

code	Rising probability	Falling probability
HISOAR	0.842799	0.157201
ELECSPN	0.714864	0.285136
BIOLIGHT	0.644983	0.355017
CSB	0.619643	0.380357
JTL	0.627322	0.372678
APG	0.573544	0.426456
JIUZHITANG	0.576076	0.423924
TESMI	0.513867	0.486133
Microgate	0.487937	0.512063
BRIGHT DAIRY	0.441809	0.558191
ZHESHANG DEVELOPMENT	0.443563	0.556437
Zuojiang Technology	0.321241	0.678759
RUITAI TECHNOLOGY	0.199491	0.800509

4.4 Comprehensive evaluation of stock selection model (Topsis)

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} \tag{4.14}$$

Among them:

$$D_i^+ = \sqrt{\sum_{j=1} w_j (z_j^+ - z_{ij})^2} \tag{4.15}$$

$$D_i^- = \sqrt{\sum_{j=1} w_j (z_j^- - z_{ij})^2} \tag{4.16}$$

$$\begin{aligned} z^+ &= (z_1^+, z_2^+, ..., z_6^+) = (\max\{z_{11}, z_{21}, ..., z_{402,1}\} \\ &, \max\{z_{12}, z_{22}, ..., z_{402,2}\}, \\ &..., \max\{z_{16}, z_{26}, ..., z_{402,6}\}) \\ z^- &= (z_1^-, z_2^-, ..., z_6^-) = (\min\{z_{11}, z_{21}, ..., z_{402,1}\} \\ &, \min\{z_{12}, z_{22}, ..., z_{402,2}\}, \\ &..., \min\{z_{16}, z_{26}, ..., z_{402,6}\}) \end{aligned} \tag{4.17}$$

We believe that, to a certain extent, volume ratio, net capital inflow, trading volume, growth rate of these new indicators can reflect the activity of individual stocks.

Therefore, using the Topsis model to score the stocks selected by Strategy 1 above on April 8, 2021, we get:

Table 6: Comprehensive evaluation model indicators and scores

code	lb	net2	volume	zhangsu_%	score
HISOAR	1.082931	-2668.6	50836750	1.401401	0.417552
ELECSPN	0.964063	902.3589	1983055	0.800493	0.491126
BIOLIGHT	1.277277	269.324	2594668	1.168884	0.464347
CSB	0.686087	-15.3092	2143339	1.005025	0.454608
JTL	1.377076	-6092.16	90030791	1.741294	0.396606
APG	1.412199	144.8026	11756637	1.160542	0.475002
JIUZHITANG	0.757759	-34.2755	7633063	0.636132	0.45672
TESMI	1.200938	468.7781	4425199	2.070828	0.496626

2022@ The American Computational Science Society. All rights reserved.

Microgate	1.070835	1392.707	17646657	2.214452	0.578743
BRIGHT DAIRY	0.959213	2952.403	23229356	1.258206	0.625654
ZHESHANG DEVELOPMENT	1.272625	-498.503	13641456	3.11284	0.488837
Zuojiang Technology	1.11995	-210.465	837768	4.239799	0.500022
RUITAI TECHNOLOGY	0.71913	597.4598	5264470	2.588686	0.5221

We define the weight of stocks predicted by XGboost to be 0.8, and the auxiliary weight of Topsis comprehensive evaluation model to be 0.2.

That is, the probability of rising the next day a is the final score of the stock.

The composite scores for the selected stocks for the day are as follows:

Table 7: Composite score

code	Comprehensive score	Ranking
HISOAR	0.75775	1
ELECSPN	0.670117	2
BIOLIGHT	0.608856	3
CSB	0.586636	4
JTL	0.581179	5
APG	0.553836	6
JIUZHITANG	0.552205	7
TESMI	0.510419	8
Microgate	0.506099	9
BRIGHT DAIRY	0.478578	10
ZHESHANG DEVELOPMENT	0.452618	11
Zuojiang Technology	0.356997	12
RUITAI TECHNOLOGY	0.264012	13

We take the top six stocks and get a yield of 1.67%.

Applying the model and backtesting, we get:

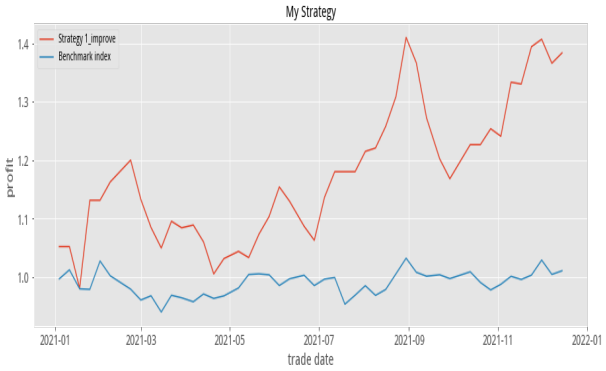


Fig5.Machine Learning revised stock selection Strategy 2

The comprehensive yield is close to 40%, which has increased from 25% to 40% now. It can be seen that the model has identified stocks with negative expectations to a certain extent.

5. Summary and prospect

In this paper, it applies machine learning to predict the short-term active stocks screened by the multi-index system, and combines the comprehensive evaluation method to score the stocks in the candidate stock pool. A considerable rate of return has been achieved during the period from January 2021 to December 2021. The deficiency is that the construction of multiple indicators depends on the experience in the secondary market, and investors can build their own multi-index system according to their own preferences and experience. In the future, we can use clustering method to apply similar index system to stocks with the same trend, and can also analyze

2022@ The American Computational Science Society. All rights reserved.

the historical data of stocks to find out the indicators that have a significant impact on stock price fluctuations. Although the indicators (MSE, RMSE, AUC) for evaluating the learning ability of the model in machine learning can reflect the accuracy of predicting stock prices, in the stock market, stock prices are affected by many uncertain factors, so the accuracy of the model cannot completely guarantee the accurate trend of the stock price. Therefore, in the future, we can conduct in-depth research in feature engineering and introduce some more advanced models of deep learning for model fusion, which may greatly improve the accuracy of stock selection.

Acknowledgments

This study has been supported by the National Natural Science Foundations of China under Grant Nos. 12172186.

Reference

[1] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation[J]. Econometrica: Journal of the econometric society, (1982) 987-1007.
[2] Majumder M M R, Hossain M I, Hasan M K. Indices prediction of Bangladeshi stock by using time series forecasting and performance analysis[C]. 2019 International Conference on Electrical, Computer and Communication Engineering, (2019)1-5.
[3] T Y Huang, F F Chen. Research on Kernel Function Application Based on SVM Stock Price Prediction [J]. Journal of Chongqing University of Technology (Natural Science), 30(2016)89-94
[4] R Hu. Multi-output support vector regression and its application in stock index forecasting [J]. Computer Technology and Development, 10(2007)226-229.
[5] Kim K J. Financial time series forecasting using support-vector machines[J]. Neurocomputing, 55(2003)307-319.
[6] Y Wang, D Y Chen, Y X Tang. Stock prediction based on Cart decision tree and boosting method [J]. Journal of Harbin University of Science and Technology, 24(2019) 98-103.
[7] Y Yang. Research on Liquor Stock Prediction Based on Machine Learning Algorithm Model [D]. Shandong University, (2021).

2022@ The American Computational Science Society. All rights reserved.