ISSN: 1813-4890

Research on Quantitative Investment Strategies Based on Deep Reinforcement Learning

Zheng Li [1, a], Xiaoyu Pan[1], Sijia Chen [1], Ling Jiang [2,b,*]

[1]School of Management Science and Engineering, Anhui University of Finance and Economics, Bengbu, Anhui, China

[2]School of Economics, Anhui University of Finance and Economics, Bengbu, Anhui, China

[a]3497299918@qq.com, [b]382104110@qq.com

Abstract

Financial services play an important role in the allocation of resources in the real economy, quantitative investment uses mathematics and computer technology to formulate strategies, can effectively use financial data, compared with traditional investment can get more efficient investment strategies. At present, investment transactions in the domestic financial market have gradually shifted from subjective trading based on traditional technical analysis and other methods to programmatic quantitative investment trading, and machine learning models have shined in this field. However, the current quantitative investment models need to improve their predictive ability to obtain high accuracy under dynamic complex change scenarios. In order to improve the prediction ability of quantitative investment model, this paper constructs a dynamic complex network for Q function training, and proposes a quantitative investment model based on Deep Q-Network (DQN). After learning from a single stock price training set, and then using a one-year data set as a test set for prediction verification, the experimental results show that DQN prediction fit is high, which shows the effectiveness of the policy model applied to financial quantitative investment. This project promotes the integration of artificial intelligence technology and finance, and has important theoretical and practical value.

Keywords

Trading strategy; Deep reinforcement learning; Deep Q-Network; Quantitative investing.

1. Introduction

Financial services play an important role in the allocation of resources in the real economy, and financial activities affect the economic development of countries around the world and occupy an important position in the modern social economy. In the information age, data is the most valuable asset. Financial data such as financial transactions, investment and wealth management, and risk control continue to accumulate with the development of the financial industry. Traditional investment methods are difficult to effectively analyze a large amount of financial data and rich types to assist investors' decision-making.

At present, domestic quantitative trading has only entered the initial stage, the volume of quantitative strategies in the market is small compared with foreign markets, and there are still a large number of investors and even investment institutions in the mainstream market. With the development of financial technology, the addition of computer technology has promoted the rapid development of quantitative investment. The performance of the trading system has been improved, and the trading based on programmatic quantitative strategy has won the favor of more investment institutions and scholars at home and abroad. The research of domestic and foreign scholars on quantitative investment mainly focuses on price prediction and quantitative stock selection, and most of its models are based on traditional measurement methods, classical machine learning algorithms, deep learning, etc., less application of reinforcement learning, and the number of research reports and papers based on deep reinforcement learning model construction is small. As the world's second largest economy, China has active financial transactions, and the demand for high-precision quantitative investment models is extremely high, requiring more efficient and rapid execution of investment actions and building a financial trading network.

Quantitative investment can help investors rationally, efficiently and quickly carry out investment activities, and it is also the frontier research direction of financial investment in recent years. At present, most quantitative investment models are based on machine learning and deep learning, and there is still a lack of research that combines deep reinforcement learning with quantitative investment. To this end, this project will summarize and learn the current research status and existing achievements of domestic and foreign scholars in quantitative investment through experimental methods, mathematical methods, literature research methods and other methods, and establish a quantitative investment model based on Deep Q-network. Quantitative investment uses mathematics and computer technology to formulate strategies and can effectively use financial data. Reinforcement learning realizes the cycle of "perception-cognition-self-decision-learning", showing the characteristics of self-learning. As a highly data-based, mathematical, and programmatic industry, quantitative investment is a good application scenario for reinforcement learning.

This project constructs dynamic change network based on stock related data, conducts Q function training, combines Q-learning and neural network to obtain deep Q network (DQN), and uses dynamic change neural network to replace Q-learning for dynamic price prediction. It can adapt to the complex situation and large spatial state to make high-precision prediction results and propose appropriate investment strategies to help investors respond to changes in the trading market more quickly, improve the accuracy and efficiency of investment strategies, reduce investment risks, optimize investment portfolio, promote the development of quantitative investment, and promote the integration of artificial intelligence technology in the financial field.

2. Literature Review

Quantitative investment refers to taking stock price, daily turnover, daily transaction amount and a large number of investment-related data as samples, establishing appropriate mathematical models and formulas by quantitative means, using computer technology to write efficient programs, studying and analyzing the future returns and risks of financial products, and issuing buying and selling orders to realize trading investment according to the judged market trend.

The wide application and continuous development of machine learning algorithms have improved the accuracy of quantitative investment, and more machine learning algorithms have been applied to quantitative investment and achieved good results. Irwin (1986) and other scholars wrote automated trading programs to automate trading in the futures market, which had a certain transformative effect on the trading of the futures market [1]. Kimk (2003) and others applied the Support Vector Machie (SVM) to financial time series forecasting, and compared the SVM-based strategy with the neural network-based strategy in the study, and the experimental results showed that the SVM-based strategy had a more stable rate of return [2]. Varela (2016) builds an investment strategy based on Q-learning, which learns based on feedback in the investment environment, but presents some limitations when the state space is large [3]. Machine learning was booming when quantitative investment was introduced into

China, and many domestic scholars used machine learning algorithms in quantitative investment research.

Zhao Zhiyong (2014) and other scholars built a strategy based on the constrained Boltzmann machine in deep learning, and the results showed that the strategy could perform feature extraction well [4]. Yue Den (2017) and other scholars applied deep reinforcement learning to the representation and learning of financial signals, using deep learning to perceive the dynamic changes of financial markets and complete feature extraction [5]. Chao Chen (2018) and other scholars apply reinforcement learning to the combination of prediction model and investment strategy, propose a framework for stock market forecasting and trading, and use investment strategy learning information to optimize the parameters of the prediction model [6]. Zhizhong Zhou (2021) and other scholars constructed a new quantitative investment strategy for commodity futures based on information fusion and strategy conversion on the basis of cointegration theory and fractal market theory, and tested the effectiveness and robustness of the strategy through empirical evidence [7]. Li Yafeng (2021) uses LSTM neural network to predict the underlying price, and then obtains trading signals on this basis, and then constructs a trading strategy [8].

At the same time, domestic scholars have also studied deep Q networks and proposed a variety of improved algorithms. Zhu Fei introduced the concept of priority in view of the shortcomings of random sampling when the traditional DQN algorithm uses the empirical playback mechanism, which improves the accuracy of sample selection, further shortens and improves the training effect [9]. Yaoyao Zhou designed a competitive network structure and a DQN algorithm based on priority playback to reduce the training time of the Deep Q network algorithm [10]. Wu Jinjin proposed a Deep Double-Q network algorithm based on weight average, which integrates the weighted double estimator into the network to solve the problem of DQN high estimated action value and DDQN low estimated action value [11].

The above scheme proves that deep reinforcement learning is expected to obtain good results in the field of quantitative investment, and the research on algorithms is also constantly improving, based on this, this paper will build a complex network of dynamic changes in stock prices, and establish a Deep Q-Network model for price prediction.

3. The Development of Quantitative Investment Model and Empirical analysis

3.1. Dynamic multi-layer complex network construction

Dynamic multilayer network is a dynamic complex system in which the components of a network interact with each other from part to whole and interact on the timeline. If there is a network G composed of multiple relationships, a dynamic multilayer network $G_\tau = (\mathrm{t_{min}}, \mathrm{t_{max}})$ is formed in the time period $\mathrm{S} = (\mathrm{t_{min}}, \mathrm{t_{max}})$, and the entire observation period $(\mathrm{t_{min}}, \mathrm{t_{max}})$ of the dynamic multilayer network is divided into T time windows, and the size of each time window is $\tau = (\mathrm{t_{min}}, \mathrm{t_{max}})$ /T, and T equally spaced $\{(\mathrm{t_{min}}, \mathrm{t_{min}} + \tau)$, $(\mathrm{t_{min}} + \tau, \mathrm{t_{min}} + 2\tau)$,...,$(\mathrm{t_{min}} + (\mathrm{T} - 1)\tau, \mathrm{t_{max}})\}$, non-overlapping and continuous time windows can be obtained, then the dynamic multilayer network $G_1, G_2, \ldots, G_T$ is divided into T discrete and ordered multilayer networks. So there is$G_T = (\Phi_t, C_t)(t = 1,2, \ldots, T)$, where $\Phi_t$represents the network set of the multilayer network at time t, specifically: Φ_t={G_α^T: α ∈{1, 2, 3, ..., L}}, where G_α^T=(V_α^T, V_β^T), L represents the number of layers of the multilayer network; $C_t = \{E_{\alpha\beta}^T \in V_\alpha^T * V_\beta^T;\ \alpha, \beta \in \{1, 2, 3, \ldots, \mathrm{L}\}\}$, where $E_{\alpha\beta}^T$ represents the edges between the layer nodesα，β at the t moment.

3.2. The development of Deep Q-Network model

In the Q-learning algorithm of traditional reinforcement learning, because the state space and action space dimensions are discrete and limited, the state-action pairs and corresponding

ISSN: 1813-4890

values can be stored in the Q-Table memory, but when the state and action spaces are continuous And when it is infinite, it will be very difficult to use Q-Table for storage. Even if this problem can be solved by large-capacity storage, it is impossible to quickly search for the corresponding value according to the state-action pair.

The main algorithm model of DQN algorithm is the combination of deep neural network and Q learning algorithm, due to the powerful neural network representation ability, in reinforcement learning, high-dimensional input data is used as a vector of states for processing, and then becomes the input of the neural network model (Agent); Next, the neural network model outputs the value distribution corresponding to each action, and then chooses which action to perform with the value size, and the ultimate goal of reinforcement learning is to obtain the most returns through self-training and self-learning.

DQN combines Q-learning and deep learning neural networks, replacing tables in D-learning with neural networks, and learning to the optimal action path by constantly updating the network. In reinforcement learning, the goal of the agent is to maximize the expected cumulative reward, and this paper designs the reward function based on the node attribute factor on the basis of the adjacency matrix, and constructs a neural network instead of the optimal action-reward value functions Q(S,A), maxQ (representing the largest Q value in each behavior of the next state. Use the greedy strategy to select actions for environment search, and store experience into the experience pool. If the number of samples in the experience pool is greater than BECH, take out the BECH size sample, obtain the training samples according to the target network, update the network parameters by gradient descent, update the target network parameters after reaching a certain number of steps, assign the new state value to the current state, and record the optimal path of each state to reach the target state after the training is completed.

The core of DQN algorithm is as follows:

Objective function: to construct deep learning self-learning function based on Q learning algorithm.

Target network: The target Q value generated based on the deep neural network, and the target Q value of the next state is estimated according to the target Q value.

Empirical playback mechanism: The purpose is to solve the problem of non-static distribution and correlation between data.

Compared with reinforcement learning, deep reinforcement learning has two more important improvements, which are the experience playback pool and the frozen target network. Since the training target value and the predicted value are from the same network, and they have a strong correlation during training, these two measures are proposed to solve it. The experience playback pool is quite a large data sample data collection, each time training, the latest generated data pair into the experience storage pool, and from the experience playback pool randomly take multiple data pairs for training, in this way can reduce the strong correlation of training data, but also can find that DQN algorithm is Off-Policy algorithm, high sampling efficiency, experience pool can be pre-set to a fixed-size queue structure, the training obtained storage in, when the experience pool is full, The next batch of training results is stored, so it is constantly in and out of the experience pool, so that the agent can learn from the past experience and optimize the next action.

In order to verify the accuracy of the model, this paper uses Yahoo Finance's single stock data, learns the stock price data from 2000 to 2010, predicts the stock price in 2011, calculates the prediction fit, and performs empirical analysis.

3.3. Empirical analysis

After model establishment and data set training, GSPC_2011 data was used to predict and verify the accuracy of the model. Based on the data from 2000 to 2010, the model carried out a total

ISSN: 1813-4890

of 38 buying and selling operations, including 19 buying operations and 19 selling operations. Finally, the total profit per share is 249.13, the actual value is 252, the error is within 95%, the model is valid.

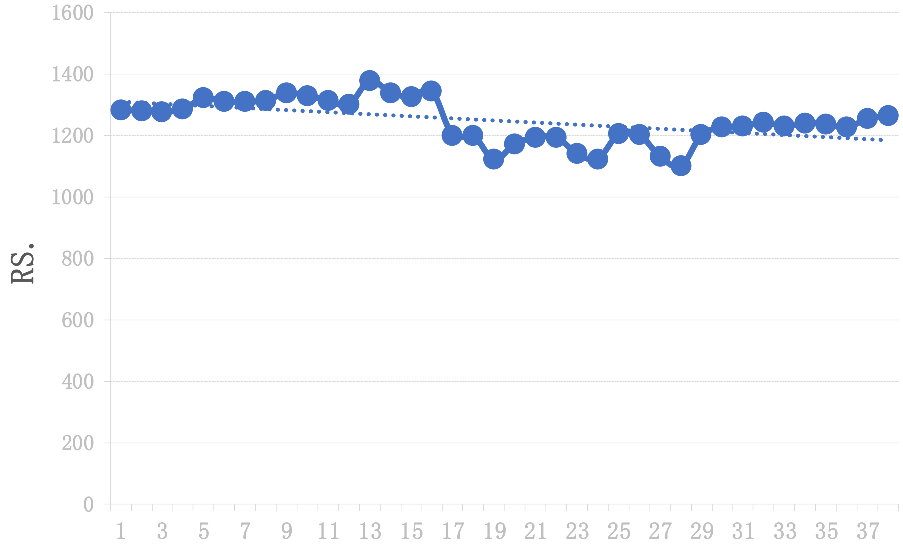

Figure 1. Buy - sell operation balance table

4. Conclusions and Suggestions

With the gradual development of securities trading in China, more and more people are beginning to study the application of deep learning models in quantitative trading, and want to rely on the powerful learning ability of deep neural networks to try to find the law of change behind massive data, and then profit in stock trading.

In this paper, the quantitative investment model based on DQN is obtained by training the year of 2000 to 2010 GSPC. By forecasting the data of 2011 GSPC, the fitting degree is higher than that of the actual value prediction, which shows the effectiveness of this policy model applied to financial quantitative investment. However, the DQN model is a Q learning applied to short-term stock trading, which uses the N-day window of closing price to determine the best decision in a certain event. But the model is not good at making decisions about long-term trends, and it is better at predicting peaks and troughs than other models.

The DQN algorithm solves the problem that Q-TABLE cannot store high-dimensional data by using deep neural network to approximate the function value, and the good nonlinear fitting characteristics of the neural network make the training converge faster and improve the efficiency of training.

In this paper, some ideal conditions are set in the experimental trading environment, such as the transaction does not consider the slip point problem, the transaction price is the closing price of the current time step K line. In the actual market, there will often be a slip point due to the trading system, trading network and other problems, so that the strategy can not be traded at the ideal price. The financial market is complex and the market changes rapidly. The quantitative strategy proposed in this paper needs to be further optimized before it can be used in actual trading. In addition, the current experiment considers trading for a single breed, which can be expanded to pair trading between two breeds and multi-breed portfolio trading strategies as needed.

In addition, in this paper, only the multi-stage growth rate of individual stocks is used as the training label. In addition, there are more individual stock labels that can be used as predictions. The labels can be divided into static labels and dynamic labels. The research conclusions and future prospects of dynamic labels are not included in this article The growth rate adopted can also add dynamically changing labels such as trading volume, opening price, and intraday price to the training indicators; static labels such as the company's market value, whether to pay dividends, and the scale of debts that do not change can also be added to the training in due course in the indicator.

Acknowledgment

This work is supported by Anhui University of Finance and Economics Undergraduate Scientific Research Innovation Fund Project in 2023 (Grant No: XSKY23164).

References

[1] Lukac, Brorsen, Irwin. A comparison of twelve technical trading systems with market efficiency implications [D]. Purdue University, 1986.
[2] Kim K. Financial time series forecasting using support vector machines [J]. Neuro computing, 2003, 55(1): 300-321.
[3] M. Varela, O. Viera, and F. Robledo, A Q-Learning Approach for Investment Decisions [M]. Springer International Publishing, 2016.
[4] Zhiyong Zhao, Feng Wang, Yuanxiang Li. Stock market forecasting based on deep learning[J].Journal of Wuhan University, 201439(1):1-7.
[5] Y. Deng, F. Bao, Y. Kong, Z. Ren, Q. Dai, Deep Direct Reinforcement Learning for Financial Signal Representation and Trading [J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 3, : 653-664.
[6] Chao Chen, Yelin Li, Hui Bu, Junjie Wu, Zhang Xiong. Forecasting on Trading: A Parameter Adaptive Framework Based on Q-Iearning [C]. ICSSSM, 2018.
[7] Zhizhong Zhou, Zuqing Yu. Quantitative investment strategy of commodity futures based on information fusion and strategy conversion [J]. Systems Engineering-Theory Methodology Application,2021,30(02):253-263.
[8] Yafeng Li，Hongbo Wang，Chen Li, etc. LSTM futures investment strategy integrating attention mechanism [J]. Computer system application, 2021,30(8): 22-30.
[9] Fei Zhu, Wen Wu, Quan Liu, Yuchen Fu. A Deep Q network method for maximum confidence upper bound empirical sampling [J]. Computer research and development, 2018,55(08):1694-1705.
[10] Yaoyao Zhou, Ye Li. Competitive Deep Q network learning based on Sort-first experiential playback [J]. Computer application research, 2020,37(02):486-488.
[11] Jinjin Wu, Quan Liu, Song Che, Yan Yan. A deep double Q network method with weighted mean value [J]. Computer research and development,2020,57(03):576-589.