A New Framework of Quantitative analysis Based on WGAN

Xingru Jiang[1], Kaiwen Jiang[2]*

[1]Software Engineering Department, School of software, Beihang university, Beijing, China, 100083
[2]Tianjin Key Lab of Intelligent Computing & Novel software Technology, Tianjin University of Technology, Tianjin, China, 300191

Abstract. This paper follows the logic of financial investment strategies based on WGAN, one of AI algorithms. The trend prediction module and the distribution characteristics of price time series are on the basis of the WGAN. Multiple factors extraction and analysis are on the basis of natural language processing technology. Buy-sell decision module is based on DQN algorithm which is one of reinforcement learning algorithms. And a designed risk control network is used as a protector for capital of investors. A multiple feature combination is proposed to describe the stock market. In the end, four Sector ETFs were selected to make simulation experiments.

1. Introduction

Tendency of Stock Market or Sector ETF will be predict by using the DRL algorithms and proper decision of investors could be provided by AI [1]. The useful results is showed already by the application of AI in the realm of quantitative investment, but there are still problems , which are in the following:

1) Although the large amount of market raw data used in model training conforms to the characteristics of the diversity of financial features, it relies excessively on the feature extraction ability of the network and does not effectively utilize prior knowledge of the market characteristics of human behavior.

2) The strategic model does not provide risk control for trading behavior, and simple trading operations can easily induce serious risks.

This paper proposes a new framework of quantitative analysis based on WGAN that analysis and in-depth research multiple features from the three levels of trend prediction, feature selection, and operational decision-making.

2. Principle of our algorithm

2.1 Wasserstein GAN Algorithm

WGAN [2] is an improved algorithm of GAN, in which Wasserstein metric W (P, Q) is used to evaluate the similarity between two distributions. As the formula (1).

$$W\left(P_{data}, P_f\right) = \sup_{\|f\|_L \le 1} E_{x\sim P_{data}} f(x) - E_{x\sim Pf} f(x) \tag{1}$$

Wasserstein GAN has achieved the following:

1) the problem of unstable GAN training was solved ;2) the problem of collapse mode was solved ;3) there is a value to indicate the training process finally.

2.2 DQN Algorithm

DQN[3] is a type of DRL[4] (Deep Reinforcement Learning), a combination of deep learning and Q learning. Which action can obtain the Q value, it is calculated by a deep neural network. DQN uses a deep learning neural network to predict the Q value and acquire the optimal action path by constantly updating the neural network.

2.3 Analysis of Non-stationary Financial Time Series

Time series[5] refers to the value series of various exchange data generated in chronological order within a certain period of exchange market. The time trend of non-stationary financial time series[6] has two types of trends: deterministic and stochastic. The modeling of financial time series analysis can generally be divided into three steps: data processing, model selection, and model evaluation.

2.4 Quantitative Analysis of KDJ

KDJ[7], also known as a random index, consists of three lines, the K line, the D line, and the J line. Several data derived from the highest, lowest, and closing prices. As shown in the figure1.

According to the intersection principle of fast and slow-moving averages, an upward breakthrough of the K line through the D line is a signal for buying, which means that the market is an obvious upward trend. The K line falling below the D line is a sell signal, meaning that the market is a significant decline.

*Corresponding author: jiang512@email.tjut.edu.cn

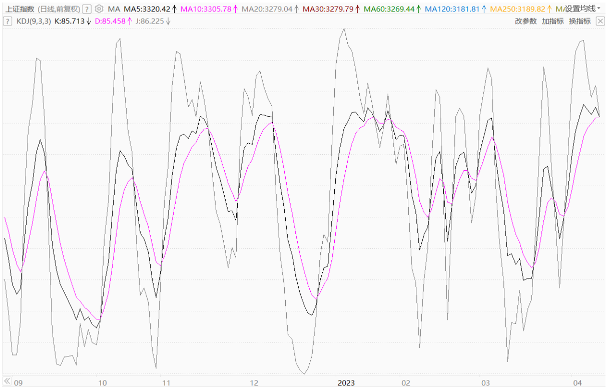

Fig1. KDJ illustration

2.5 Quantamental

Because this paper is a quantitative analysis of sector funds, only macro analysis[8] and sector analysis are considered. There are three important financial indicators that need to be looked at: money supply, interest rates, and exchange rates. Besides these indicators, there are other factors that need to be considered, such as wars, political factors, political stability and so on.

2.6 Behavioral Finance

Behavioral finance[9] studies people's cognitive and decision-making processes, as well as the important phenomena of these decisions in the financial market. The fear and greed index objectively reflects the fear and greed emotions of the market. It can assist investors in making buying or selling decisions.

3. the new framework of quantitative analysis Based on WGAN

Feature extraction module, index trend prediction module, and position adjustment decision-making module are designed in this paper. Quantitative Investment Strategy of Sector ETF Based on WGAN is proposed in this paper, which bases on financial-related theories such as time series, technical indicators, behavioral science and so on. Wasserstein GAN Generative Adversary Networks, DQN[10] deep learning neural network, technical quantitative analysis, and fear and greed index are used as the basis for establishing a risk prevention, which could help to apply holding strategies for Sector Funds. As shown in the figure2.

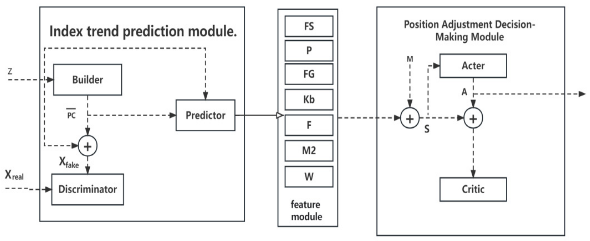

Fig2. The framework of quantitative analysis Based on WGAN

3.1 Index Trend Prediction Module

The index trend prediction module uses the index data of the Shanghai Stock Exchange, Shenzhen Stock Exchange, military ETF, and high-end manufacturing ETF, on each trading day to find out the timing rules contained in the index data. RNN is used as a prototype module, which is the concept of generating rivalries between generators and discriminators, is introduced to create models of the samples probability distribution. The LSTM model with good temporal data processing performance is used to connect the multi-layer perception MLP as the basic structure of the WGAN to achieve the purpose of trend judgment. As shown in the figure4:

3.2 Feature Extraction Module

Multiple factors that affect the financial market are extracted in this paper.

1) Time series prediction factor F: The WAN generator will deduce and forge subsequent time series index data based on the historical index, and calculate the time series prediction factor F (t, T, D) using the formula according to the generator's output Xfake.

2) Financial Time Cycle factor P: Under the background of the Federal Reserve's interest rate hike, US dollar liquidity contracted and US dollars returned to the United States. The US dollar rose against the RMB, and the US dollar appreciated. Once the expectation of RMB depreciation is formed, it will lead to the outflow of hot money and the contraction of domestic liquidity, which is not conducive to the strength of A-shares. The effect of the Federal Reserve's interest rate cut on A-shares is opposite to that of the Federal Reserve's interest rate hike. P=0 is the period of interest rate increase by the Federal Reserve, P=1 is the period of interest rate reduction by the Federal Reserve.

3) Sector Cycle Fs: sector historical data is put in WGAN generator to derive subsequent sector index data.

4) Behavioral Finance FG: If the fear and greed index is below 10, FG=1. It is time suitable for buying. If the fear and greed index has reached around 90, FG=0. It is time for sell. The fear and greed index can be read in the software called Jiuquaner.

5) Technical indicator KDJ: The short-term indicator K line crossing the long-term indicator D, which is the golden cross shape of KDJ. The golden cross is at a low of 50, which is a strong signal of buying, Kb=1.The death cross refers to the K line falls below the D line, and then the J line also falls below the D line, Kb=0.

6) The local currency environment factor M2: M2 refers to the total amount of circulating funds in society. A year-on-year increase in M2 has a positive impact on the stock market, M2=1. while a year-on-year decrease in M2 has a negative impact on the stock market M2=0.

7) Event factor W: In the equity market, after the outbreak of geopolitical conflicts, risk assets will be impacted in the short term, and the stock markets of major economies such as China, Europe, and the United States will significantly decline, W=0.

8)Yield Factor I: The yield statistics of holding sector ETF are the data provided by each exchange market. When I=-10%, the purchase operation is performed.

3.3 Natural Language Processing Technology

Using web crawler technology to capture the information: Firstly, Sina Finance channel as the target page, daily hot news headlines captured from browser. Secondly analyzing the HTML structure of the target web page and locating the tag instance where the desired news title is located. Finally, crawler technology is used to capture the HTML text data of the target webpage and intercept the text information within the class attribute in the tag.

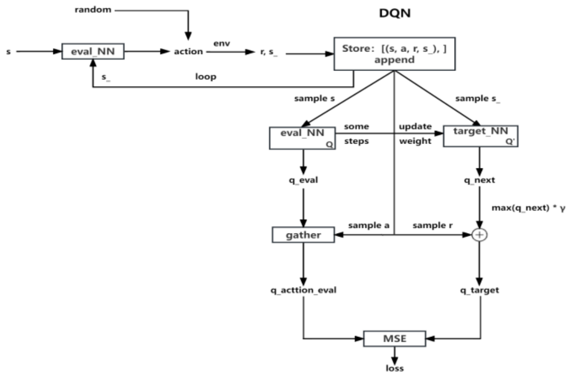

Fig3. DQN Network Structure Diagram

3.4 Decision Module and Network Structure Design

The position control problem of sector ETF can be seen as a Markov decision-making process. The article combines the market environment characteristics F and the amount of funds M as the status S, expressed as S={F, M}, position as the action A. The reward function ρ which based on the income situation. As the formula(2)(3)(4).

$$\rho(s_t, a_t) = \eta_1 \cdot r_{market} + \eta_2 \cdot r_{agent} \quad (2)$$

and

$$r_{market} = \frac{PC_{t+1} - PC_t}{PC_t}\left(a_t M_{Agent}(t) - M_{market}(t)\right) \quad (3)$$

and

$$r_{agent} = a_t \frac{PC_{t+1} - PC_t}{PC_t} M_{Agent}(t) \quad (4)$$

The simulation environment outputs state, reward, and the neural network selects actions based on the state to form a record (state, reward, action, next_state) into the memory. As shown in the figure4.

4. Experiment of A New Framework of Quantitative analysis Based on WGAN

This paper is aimed at help investors with more than 200,000 RMB to invest Sector ETF. Due to four funds correspond to different cycle and identify timing trends through analysis of past data of SSE and four ETF funds based on the principle of multiple characteristics: D={F, FS, Pek, Pec, Pej, Pex, Peg, Peq, FG, Kb, M2, W} to conduct buying and selling. The funds for each sector are divided into 20 portions at 5,000 RMB each, which conduct a buying operation each time. As well as, it is necessary to compare the pre value yield to conduct a selling operation. When the yield factor I=- 10%, conduct a buy operation. As the same time, when four of the five factors Fs, Pe, FG, Kb, and M2 are satisfied, the buying and selling operation have to be performed. Once a significant event occurs, if it is determined that the event is a negative event, W=1 is the selling condition. Considering that the role of artificial intelligence quantitative investment strategies is to complete auxiliary work for investors, completely performing end-to-end selling operations by itself will not happen unless emergency of risk controlling.

Fig4. 512660

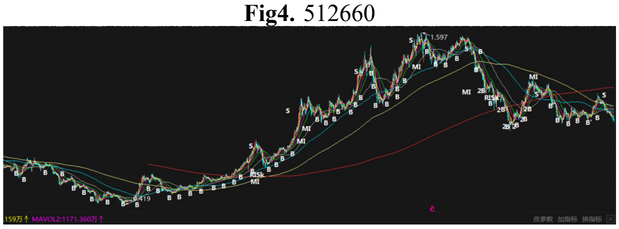

Fig5. 512910

As shown in the figure4. 512660 military ETF trading and risk point tips. In order to verify the effectiveness of the software each trading point was followed, which executed 47 buying operations and 9 selling operations, executed risk selling of COVID-19 and Russia Ukraine war and implemented manual intervention once.

Table1. Results of Simulation Experiment

	Buying Times	Selling Times	Days	Total Profit	Fund Utilization Ratio	Annual Yield
512660 Military ETF	47	9	1900	72%	35%	8.3%
512910high-endmanufacturingETF	30	3	428	24%	76%	18.24%
15994 GEM 50ETF	65	7	1900	56%	78%	14%
512000 brokerage ETF	58	2	1900	40%	100%	13%

As shown in the figure5. 512910 manufacturing ETF trading and risk point tips. Due to the fact that the ETF fund 512910 was not available for trading until January 10, 2022, and the characteristics of newly issued funds, there will be a few months of downward adjustment after trading. 30 buying points were prompted in a t year, as the same time only 3 times the selling point were prompted.

Fig6. 159949

As shown in the figure6. 159949 GEM 50ETF trading and risk point tips. As the stocks held by GEM 50ETF funds are mostly high-tech growth stocks, the fluctuation of GEM 50ETF funds is relatively large, and the speed of rise and fall is relatively fast. The rapid rise and rapid decline of GEM 50ETF is not conducive to the persistence of buying operations, so more manual intervention has been added to the handling of GEM 50ETF buying, selling.

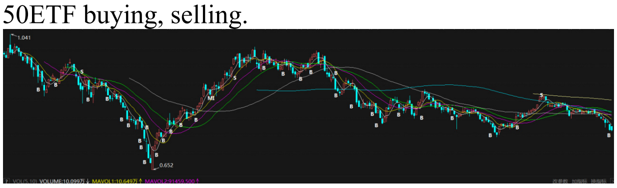

Fig7. 512000

As shown in the figure: 512000 Brokerage ETF trading and risk point tips. Brokerage funds are greatly affected by the Federal Reserve's monetary policy, so manual intervention in the starting point for buying is necessary. From January 1, 2018 to March 17, 2023, a total of 58 buying operations and 2 selling operations were performed, and 3 manual intervention operations were performed.

5. Experiment Results

To the four ETFs, the risk control ability of the strategic model in the many downward period of the market from 2018 to 2022 is excellent. After accounting for capital utilization, the annual return of the four ETFs is greater than 8%, which is twice the bank 3year saving interest rate. As show in table1. In most stable time, this model can also help investors adopt a fund distribution investment strategy to invest funds without excessive interference from external unimportant information, which results in low capital utilization. Therefore, with the help of the framework of quantitative analysis Based on WGAN in this paper, investors can detect changes in market more keenly to obtain more stable excess returns.

References

1. Gomez, E; Brualla-Mora, A. An analysis of qualitative and quantitative morphokinetic parameters automatically annotated using CHLOE (Fairtility), an AI-based tool, finds AI score predictive of blastulation and ploidy[C]. HUMAN REPRODUCTION, 2022, 37: pp.I301-I302
2. Deng, FQ; Wan, Q. Image restoration of motion artifacts in cardiac arteries and vessels based on a generative adversarial network [J]. QUANTITATIVE IMAGING IN MEDICINE AND SURGERY, 2022, 12 (5) :pp.2755-2766
3. Arlery,M (Arlery,M); Rouviere, JL (Rouviere,JL); Widmann,F (Widmann,F). Quantitative characterization of GaN quantum-dot structures in AlN by high-resolution transmission electron microscopy [J].APPLIED PHYSICS LETTERS,1999: 3287-3289
4. Li, Z (Li, Zhen); Gu, JY (Gu, Jiayu); Ding, J (Ding, Jie). Molecular mechanism of ethanol-H-2 co-production fermentation in anaerobic acidogenesis: Challenges and perspectives[J], BIOTECHNOLOGY ADVANCES, 2021. DOI: 10.1016/j.biotechadv. 2020.10767
5. Sahu, SK (Sahu, Santosh Kumar) [1]; Mokhade, A (Mokhade, Anil) [1]; Bokde, ND (Bokde, Neeraj Dhanraj). An Overview of Machine Learning, Deep Learning, and Reinforcement Learning-Based Techniques in Quantitative Finance: Recent Progress and Challenges[J]. APPLIED SCIENCES-BASEL[J],2023(3)
6. Zhang M, Pi D. A New Time Series Representation Model and Corresponding Similarity Measure for Fast and Accurate Similarity Detection[J]. IEEE Access, 2017, 5:24503-24519.
7. Rout A K, Dash P K, Dash R, et al. Forecasting financial time series using a low complexity recurrent neural network and evolutionary learning approach[J]. Journal of King Saud University - Computer and Information Sciences, 2015: S1319157815000944.
8. Gomber P, Koch J A, Siering M. Digital Finance and FinTech: current research and future research directions [J]. ZEITSCHRIFT FUR BETRIEBSWIRTSCHAFT, 2017(5). DOI: 10.3390/en15228658
9. Alexander, S. Price movements in speculative markets: Trends or random walks[J]. Industrial Management Review, 1961(5):7-26.
10. Zheng, LN (Zheng, Leina); Pan, TJ (Pan, Tiejun); Liu, J (Liu, Jun) . Quantitative trading system based on machine learning in Chinese financial market[J], JOURNAL OF INTELLIGENT & FUZZY SYSTEMS, 2020: 38 (2), pp.1423-1433
11. Risk management: Shefrin, Hersh, Statman, Meir. Behavioral Portfolio Theory[J]. Journal of Financial & Quantitative Analysis, 2000, 35(2):127-151.