JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 40, 781-797 (2024)
DOI: 10.6688/JISE.202407_40(4).0006

Constructing the Financial Asset Allocation Method Using Deep Reinforcement Learning Algorithm for Financial Transactions

WAN-DONG GAO AND YU-MIN FEI[+]
Economics Management College
Weifang University of Science and Technology
Shouguang, Shandong, 262700 P.R. China
E-mail: feiym162@nenu.edu.cn[+]

In the realm of financial transactions, the allocation of household assets often lacks proper guidance, resulting in suboptimal utilization and limited income for residents. This study aims to address this issue by introducing rational asset allocation strategies, improving efficiency, and increasing household income. Specifically, the study focuses on enhancing Deep Reinforcement Learning (DRL) algorithms, particularly the Deep Q-Network algorithm. The current network is optimized by employing a dual network structure, leading to improved performance. Moreover, the proposed model incorporates Grubbs' improved algorithm to effectively denoise stock data samples, enabling iterative training of diverse stock agents. The results demonstrate the following findings: (1) The application of the Grubbs' improved algorithm to the cumulative returns of different stocks significantly surpasses the results without noise reduction; (2) The Sharpe ratio of the stock agents using the Grubbs improvement algorithm is noticeably higher than that of the unimproved DRL algorithm. The Sharpe ratio of stock agents under the unimproved algorithm consistently remains below 1.0; (3) The maximum drawdown of the stock agents using the Grubbs' improved algorithm is significantly lower than that of the unimproved DRL algorithm. The unimproved algorithm exhibits relatively higher maximum drawdown rates. These findings indicate that Grubbs' improved algorithm possesses notable advantages in noise reduction and enhancing the cumulative returns of assets.

Keywords: deep learning, reinforcement learning, financial assets, asset allocation, Grubbs

1. INTRODUCTION

The ongoing reform of China's financial system has drawn increasing attention from market participants regarding the allocation of financial assets. Participants in the market are striving for high efficiency, high profits, and effective allocation of financial assets. Residential property has also become an indispensable component of financial products in the financial market. However, the current allocation of household financial assets fails to achieve the objective of obtaining investment returns. It is characterized by a continuous decline in the number of financial assets, limited product holdings, high portfolio risk, and a lack of reasonable guidance in asset allocation. Therefore, the study of asset planning is of significant importance as it aims to reasonably allocate household financial assets and enhance residents' financial literacy [1].

With the continuous development and reform of China's financial market, household financial asset allocation has become an important issue. In the past, most households re-

Received April 20, 2023; revised May 25 & June 10 & June 27 & July 6, 2023; accepted July 13, 2023.
Communicated by Mu-Yen Chen.
[+] Corresponding author.

lied on bank deposits and real estate for their financial asset allocation. However, this traditional way of asset allocation can no longer meet the needs of market participants for efficient profits and diversified investment. The allocation of household financial assets is mainly manifested in several aspects. First, the number of household financial assets continues to decline. This may be due to a family's unfamiliarity with the financial markets or a lack of investment awareness. Second, household financial assets have limited product holdings. Households tend to hold only a few financial products and lack diversified and risk-diversifying investments. In addition, the portfolio risk of household financial assets is high, and there is a lack of reasonable asset allocation guidance. Many families lack professional knowledge and financial skills in investment decisions and are prone to high-risk investments.

Based on the above research background, the remaining research structure is as follows. Section 2 analyzes the application status of Deep Learning (DL) algorithms in financial asset allocation in the context of financial transactions, including an overview of relevant research literature and a summary of research status. Section 3 focuses on the assets and financial situation of households with moderate economic conditions and analyzes the data. In addition, Deep Reinforcement Learning (DRL) algorithms are introduced. This study realizes noise reduction optimization and improves the stability and performance of the algorithmic trading system by integrating Grubbs into the DRL algorithm. Meanwhile, a stock denoising model based on the improved algorithm of Grubbs is constructed. Section 4 conducts experiments on the designed model and tests its performance. Additionally, the experimental results are presented and analyzed. Finally, the experimental results are discussed in many aspects. Section 5 summarizes the findings. The research goal of this study is to enhance the noise reduction ability of stocks in the financial trading market, optimize the trading market, achieve a more reasonable asset allocation for residents, and improve returns. Besides, the shortcomings and the future outlook methods are pointed out.

This study innovatively improves the DRL algorithm and explores the optimal investment strategy. It is important for improving household asset management decisions, reducing the cost of participation in financial markets, and increasing household wealth accumulation. This research provides theoretical and technical support for the study of quantitative financial trading.

2. LITERATURE REVIEW

In financial asset allocation research, Huang and Lu [2] conducted an analysis on the interaction and impact of risk assets and the household financial asset allocation structure. They used data from the China Household Finance Survey and asset investment behavior theory. Probit and Tobit's models were employed to empirically analyze the influencing factors on Chinese households' financial asset allocation structure [2]. Ge *et al.* [3] utilized panel data from Chinese households spanning from 2014 to 2018 to examine the impact of political background on Chinese households' asset allocation behavior [3]. Zhang and Lin [4] explored the effect of digital finance development on household income, consumption, and financial asset holdings using extreme value theory. They also employed a binary expansion test to detect the nonlinear dependence between digital finance and household economic variables [4]. Gabrielle *et al.* [5] investigated the association between different types of financial assets and the sense of purpose among the elderly population. The study analyzed the relationship between the sense of purpose and financial assets to enhance the investment enthusiasm of older individuals [5]. Zou and Deng [6] focused on studying the influence of financial literacy and housing value on household financial market participation. They examined the proportion of asset allocation among different households and the structural differences between urban and rural areas, providing relevant suggestions [6]. Baker and Kueng [7] obtained micro-data on household financial transactions from various sources, analyzed their availability, and discussed the potential future applications of this flexible data in corporate research, real-time policy analysis, and macro statistics [7]. Li [8] analyzed the impact of financial literacy on household financial services allocation and examined the investment behavior of low-risk and high-risk financial assets among households [8]. Currently, research on household financial asset allocation primarily focuses on household income distribution, member structure, age, and the impact of financial literacy on asset allocation. However, with the rapid development of the financial investment market and innovative technologies, machine learning (ML), DL, and Reinforcement Learning (RL) have expanded into numerous domains.

DRL can learn control policies directly from high-dimensional data. Jiang and Lynch [9] proposed a Deep Neural Network (DNN)-based controller that utilized a model-free RL algorithm to train the controller to achieve hover stabilization of a quadrotor Unmanned Aerial Vehicle [9]. Fujimoto *et al.* [10] examined the problem of function approximation errors leading to high estimation and suboptimal strategies. They proposed a novel mechanism to minimize the impact of these errors on optimization outcomes [10]. Wu *et al.* [11] delved into recent research on underestimating evaluation results in a dual-delay deep deterministic evaluation algorithm. They provided theoretical justifications for this phenomenon [11]. Dong *et al.* [12] introduced a dynamic hyperparameter optimization method that adaptively optimized hyperparameters for a given sequence using an action prediction network based on continuous deep Q-learning [12]. However, the application of DL algorithms in researching the relationship among family risk, asset structure, and allocation choices within the realm of financial transactions has been limited. Trenta *et al.* [13] proposed an autoregressive integrated moving average and exponential smoothing model for analyzing financial data in financial markets [13].

3. RESEARCH METHODOLOGY

3.1 Family Asset Structure

Presently, households possess a diverse array of assets, including stocks, bonds, real estate, and cash. These assets can be systematically classified into different categories, namely fixed assets, current assets, tangible assets, intangible assets, and long short-term investment assets. Table 1 elucidates the distinction between tangible and intangible household assets [14].

Household financial assets encompass various components, including household deposits, various bonds, stock holdings, and insurance policies. These assets can be classified into relatively safe and risky types based on the level of risks [15]. In the case of middle-level residents, there is generally limited asset planning for risk-based assets, such as stocks, funds, and financial bonds with less emphasis on financial wealth management products.

Besides, relatively safe financial assets, such as deposits, government bonds, and cash, receive more attention and planning.

Table 1. Classification of household assets.

Classification	Definition	Content	Example
Intangible household assets	Assets without physical form	Family reputation and social relations	bloodline, reputation, and social reputation
Tangible household assets	Assets retained in the form of specific material products	Family-owned assets and family-owned human resources	Household physical assets and household financial assets

3.2 Current State of Household Financial Asset Allocation

The allocation of assets among Chinese households displays a significant imbalance, with bank savings dominating the landscape. Allocation to financial risk assets remains relatively low. In addition to cash and deposits, financial products are considered risky assets [16]. Most households possess a limited understanding of financial matters and lack knowledge about other wealth management products [17]. Consequently, these households tend to adopt a conservative approach, choosing low-risk bank deposits as their primary asset allocation strategy. According to relevant data, the specific distribution of household asset allocation in China in 2022 is shown in Fig. 1.

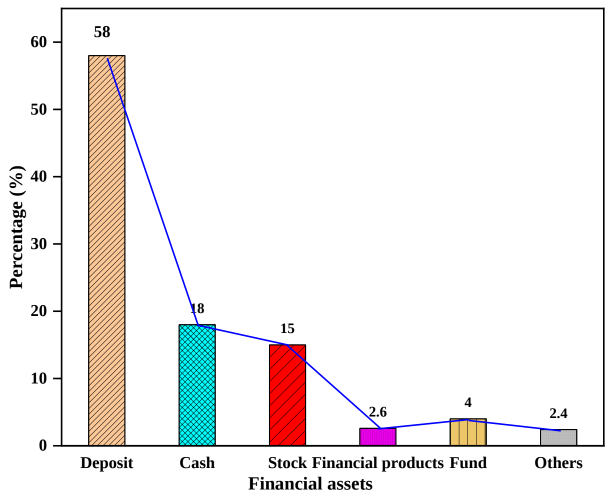

Fig. 1. Financial asset allocation of Chinese households in 2022.

In Fig. 1, a significant portion of households exhibits a preference for investing in financial assets with minimal or no risk, as indicated by the highest proportion allocated to the lowest risk category. Among these allocations, bank deposits emerge as the dominant choice, constituting 58% of the total. Cash holdings closely trail behind at 18%, while equity asset investments account for 15% of the portfolio. Various funds represent 4% of the allocation, with wealth management products comprising 2.6%. Collectively, low-risk financial assets substantially dominate the allocation, constituting approximately 80% of the overall proportion. In regions characterized by higher risk levels, stocks occupy a leading position. However, their proportion remains smaller compared to bank deposits. Financial products represent only 2.6% of the allocation. The distribution of financial assets distinctly illustrates an imbalanced scenario.

The allocation of financial assets among Chinese households is characterized by its unevenness and relatively homogenous structure. This can be attributed to the low level of financial literacy among residents, their limited familiarity with professional financial management, and a lack of comprehension regarding diverse financial products. Furthermore, the financial market confronts challenges stemming from an imbalanced main structure and an imperfect system [18]. The enhancement of financial literacy assumes pivotal importance in reshaping households' perspectives on financial management, particularly among families harboring divergent investment inclinations. Moreover, optimizing the financial system represents a critical avenue for progress.

3.3 Construction of the DRL Algorithm Model

RL involves mapping the state of the environment to a behavioral policy and engaging in a continuous learning process. It utilizes interaction data between the agent and the environment to achieve an optimal policy as an ML approach [19].

DL, also known as deep structured learning, aims to understand the inherent laws and representation levels within sample data by extracting features layer by layer and exploring the underlying feature distribution. DL enables tasks, such as data classification and regression, by automatically learning data features, providing machines with the ability to analyze and learn independently, similar to human cognition, and recognize data, such as text and images [20]. Neural networks of this nature enable RL to address decision-making problems that are previously intractable, particularly in scenarios with high-dimensional state and action spaces. DRL can be categorized into three types based on the role played by DNN in RL: DRL methods based on value function, DL methods based on policy, and DL methods based on Actor-Critic approaches.

DL is combined with the Q-learning algorithm to form the Deep Q-Network (DQN) algorithm. The primary objective of Q-learning is to approximate the optimal state-action value function directly. To facilitate training, the algorithm utilizes labeled sample data and employs the Bellman equation to calculate the target value $r + \gamma \max_{a'} Q(s^{\wedge'}, a^{\wedge'}, \theta^{-})$, which serves as the sample label. The DQN algorithm continually optimizes the loss function to update the parameters. The loss function quantifies the discrepancy between the estimated values and the target values. The specific calculation of the loss function is depicted in Eq. (1).

$$L_i(\theta_i) = E[r + \gamma \max_{a'} Q(s', a', \theta_i^{-}) - Q(s, a, \theta_i)]^2 \tag{1}$$

In Eq. (1), s' represents the next state. a' indicates the next action to be performed. θ^{-} indicates the target network parameters. θ indicates the parameters of the evaluation network. The mean square deviation between two Q values is minimized to train the network.

DQN algorithm uses Q-learning to calculate the target value approximating DQN, which is marked as Y_i^{DQN}. The target value is calculated according to Eq. (2),

$$Y_i^{DQN} = r + \gamma \max_{a'} Q(s', a', \theta_i^-). \tag{2}$$

The two network parameters θ^- and θ are updated by the gradient descent method. The partial derivative of the loss function and the gradient are obtained, as shown in Eq. (3).

$$\nabla_{\theta_i} L(\theta_i) = E_{s,a,r,s'}[(Y_i^{DQN} - Q(s,a,\theta_i))\nabla_{\theta_i} Q(s,a,\theta_i)] \tag{3}$$

The random gradient descent method is used to update the target network, as shown in Eq. (4),

$$\theta^- \leftarrow \tau\theta + (1 - \tau)\theta^-. \tag{4}$$

In Eq. (4), τ is the learning rate.

During the training process of the DQN algorithm, the gradient descent algorithm for the neural network and the Q-learning algorithm are executed in parallel. Labeled samples obtained through Q-learning are utilized for training the neural network, enabling it to approximate the optimal action-value function and facilitate the learning of the optimal strategy through RL.

Fig. 2 presents the model structure of the DQN algorithm.

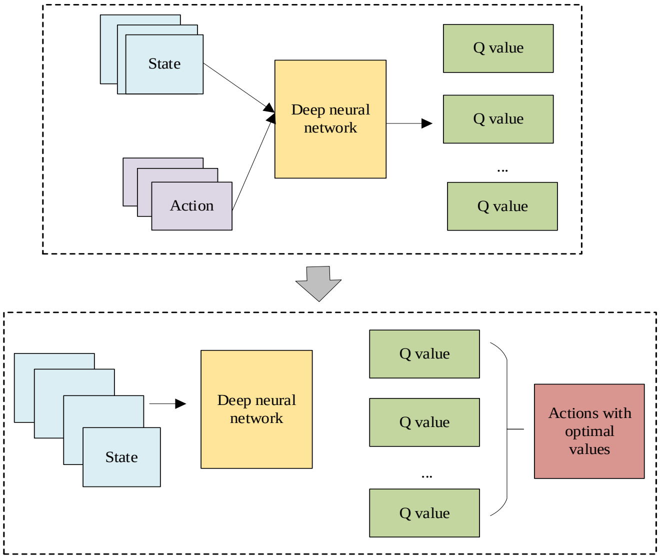

Fig. 2. Schematic representation of the DQN model structure.

Fig. 2 illustrates two models of the DQN algorithm. The upper model involves inputting all states and actions into the DNN and obtaining the Q-values for all actions through network calculations. However, the convergence of the neural network becomes challenging when the state dimension is high, and the action dimension is low. Each action generates distinct inputs, significantly increasing the computational workload. The lower model presents an improvement over the previous model. Only the states are input, and the Q-values for all actions are obtained through DNN calculations. Based on the principles of Q-learning, the action with the optimal value is selected as the next action to be executed.

The DQN algorithm utilizes a dual network structure to optimize the current network and introduces a target value network denoted by θ^{-}. The target value network mitigates the correlation between the target and predicted Q-values using its parameters. It copies the parameters of the current value network. As training progresses, the parameters of the current value network are copied every N steps to acquire the target Q-value. The dual network structure maintains relatively stable target Q-values, reduces the correlation between the current and target value networks, and enhances the algorithm's stability.

Based on the aforementioned content, this study employs the Double DQN algorithm as an enhanced version of DQN to achieve better optimization of financial asset allocation. The Double DQN algorithm addresses the overestimation bias in Q-value estimation. It consists of two neural networks: one for action selection and the other for estimating the Q-value of that action. Double DQN can more accurately estimate the Q-value of the optimal action by separating the networks for action selection and Q-value estimation, thereby improving performance and stability.

DRL is employed to design stock trading systems by leveraging the strategy optimization approach. The agent, consisting of a DNN, is utilized to construct an exploration trading strategy, as illustrated in Fig. 3.

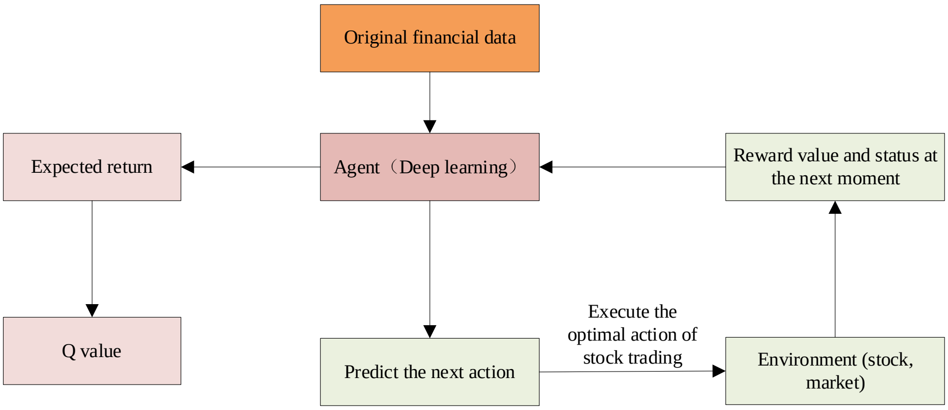

Fig. 3. Stock trading strategies using DRL.

In Fig. 3, the initialization parameter settings for the original quantitative transaction financial data are input into the agent. Then, the agent predicts the next action and executes the optimal stock trading action. During the execution of the stock trading action, the neural network takes inputs from the reward function and the state returned by the stock market at the next time step. This process ultimately yields the expected return and Q-value.

The financial stock market is a complex environment characterized by significant noise. This noise can interfere with the information contained in the trading market data. Representing the raw data of stocks directly as feature vectors may not accurately reflect the information of the financial market. Therefore, this study aims to optimize the model further by incorporating an improved Grubbs algorithm based on the Grubbs criterion. This algorithm adjusts the Grubbs test statistics, considering factors, such as sample size and data distribution, to enhance the accuracy of outlier detection. It demonstrates superior robustness and stability across different sample sizes and data distribution scenarios. Moreover, this algorithm effectively mitigates random errors, reducing the disparity between data noise and actual measurement data and enabling a more objective reflection of the financial market.

For n samples, they are $x_1, x_2, \ldots, x_m, x_n$. The mean $\bar{x}$ is expressed in Eq. (5),

$$\bar{x} = \sum_{m=1}^{n} x_m / n. \tag{5}$$

The standard deviation S is shown in Eq. (6),

$$S = \sqrt{\frac{\sum_{m=1}^{n}(x_m - \bar{x})^2}{n-1}} \tag{6}$$

In Eqs. (5) and (6), the expression of probability event G is set as shown in Eq. (7),

$$G = |x_m - \bar{x}|/S. \tag{7}$$

$G > G(\alpha, n)$ is regarded as a small probability event, and x_m at this time is regarded as an abnormal value. α represents the significance level confidence probability, most of which use $\alpha = 0.01$ and $\alpha = 0.05$. When $G > G(\alpha, n)$, the data is abnormal data and can be discarded and not retained. When $G < G(\alpha, n)$, the data is valid and can be retained.

3.4 Construction of Stock Noise Reduction Model based on Grubbs' Improved Algorithm

3.4.1 Model construction

The denoising of financial stock information is achieved through the application of the improved Grubbs algorithm. Let $X = \{x_1, x_2, \ldots, x_n\}$ represent the set of noisy stock data samples, where X_m denotes the raw stock data sample. The mean $\bar{x}$ and standard deviation, S, of the sample set are calculated using Grubbs' improved algorithm. Subsequently, the value of G is determined using Eq. (3). Data points with abnormal noise are identified and filtered out by comparing the calculated G value with the threshold $G(\alpha, n)$, effectively removing stock noise. Fig. 4 provides a visual representation of the denoising and optimization process of raw financial data using the Grubbs' improved algorithm.

In Fig. 4, the raw data is represented by X_1 to X_n. Among them, X_3, X_5, and X_7 indicate abnormal values, which are highlighted in orange. The application of Grubbs' test enhances the accuracy of the calculations. After removing the outliers, the resulting dataset can serve as a representative sample for subsequent analysis and research. The sample data provides an objective depiction of the economic significance of stocks and accurately reflects the intricate and volatile nature of the financial market by eliminating the outliers.

Two algorithms, namely Grubbs' improved algorithm and the conventional Q-Learning algorithm in DRL, are employed to denoise the original stock data. The difference bet-

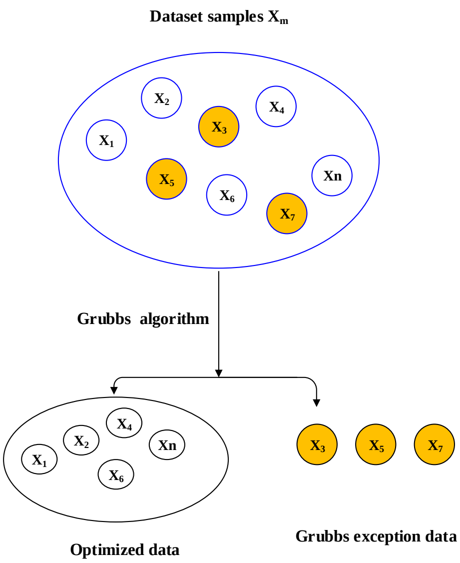

Fig. 4. Denoising and optimization of raw data.

ween these two algorithms lies in their approaches. Grubbs' improved algorithm primarily utilizes statistical methods to identify abnormal data by calculating outliers within the samples. The Q-Learning algorithm aims to find the optimal strategy by learning the value function (Q-values). It is suitable for reinforcement learning tasks involving discrete states and actions within a Markov Decision Process environment. Additionally, this study will analyze the denoising results, Sharpe ratio, and maximum drawdown rate of the proposed stock dataset using the two aforementioned algorithms. The results will be compared to validate the effectiveness of the proposed model.

The four stock features optimized by the Grubbs algorithm are fed into the trading agent. The agent is structured as a shallow neural network consisting of three parts: the input layer, hidden layer, and output layer, with only one hidden layer. The neurons in each layer employ the tanh activation function. The trading agent undergoes tens of thousands of training iterations. Finally, the training results are evaluated based on financial indicators and cumulative returns.

3.4.2 Measurement of financial performance indicators

Financial performance is often evaluated using several common metrics, including the cumulative return on assets, the Sharpe ratio, and the maximum drawdown ratio. The cumulative return on assets measures the overall income generated by all purchased products, taking into account both redeemed and currently held assets. The cumulative return on assets can be calculated according to Eq. (8).

$$P = A - C + B \tag{8}$$

In Eq. (8), P represents the cumulative income of assets. A represents total assets. C represents total costs. B represents redemption assets.

The Sharpe ratio is employed as a metric to evaluate the degree to which the returns of an asset adequately compensate investors for the risk they assume with the goal of maximizing risk-adjusted returns. This ratio is computed using the time series of investment returns, as depicted in Eq. (9).

$$S_T = \frac{Avrg(R_t)}{Std(R_t)} \quad t = 1, \ldots, T \tag{9}$$

In Eq. (9), R_t is the investment return within the trading time interval t; Avrg is the average of R_t; Std is the standard deviation of R_t.

The maximum drawdown is a pivotal risk indicator for financial stocks that provides insights into the maximum potential loss and the resilience of investments. The maximum drawdown ratio can be calculated according to Eq. (10).

$$D = \max \frac{P_i - P_j}{P_i} \tag{10}$$

In Eq. (10), i and j are any two moments in a given period; P_i and P_j are the net asset values of the two moments, respectively.

In this study, an iterative training process is conducted using both the Grubbs' improved algorithm and the unimproved deep reinforcement algorithm for four stock agents. The analysis focuses on evaluating and comparing the cumulative asset return, Sharpe ratio, and maximum drawdown ratio across the different algorithms.

4. EXPERIMENTAL DESIGN AND PERFORMANCE EVALUATION

4.1 Experimental Materials

To compare the performance differences between the improved Grubbs algorithm and the unimproved DRL algorithm in training financial trading agents, the selected evaluation metrics include cumulative return on assets, Sharpe ratio, and maximum drawdown ratio.

4.1.1 Data acquisition

To construct a comprehensive stock dataset covering multiple domains, this study randomly selects four stocks from different industries: computers, energy, chemicals, and finance. These stocks, namely Stock A (Hikvision), Stock B (Minsheng Bank), Stock C (Sinopec), and Stock D (CMB Securities), represent diverse sectors. The dataset of these stocks is divided into a training period and a testing period for model training and evaluation. The training period spans three years from November 2017 to November 2020, while the testing period covers one year from November 2020 to November 2021. Daily stock data, including opening, closing, highest, lowest, and trading volumes, are collected for Hikvision, Minsheng Bank, Sinopec, and CMB Securities during the training period. The testing period data (unseen during training) is used to evaluate the model's performance on new data. Discrete return is defined as the return type, categorizing stock returns into predefined categories instead of continuous numerical values. This simplification enables the application of algorithms and models suitable for handling discrete data for prediction and decision-making.

4.1.2 Data preprocessing

To enhance experimental results, the dataset underwent several preprocessing steps. Firstly, any missing values in the dataset were addressed using interpolation methods for data imputation. Secondly, the Grubbs' improved algorithm was applied to identify and handle outliers, improving data quality. Next, data smoothing techniques were employed to reduce noise and volatility. Furthermore, to eliminate scale and unit differences among different stocks, Min-Max normalization was applied to standardize the data. Finally, additional feature engineering operations were conducted after splitting the dataset into training and testing sets, such as calculating daily returns and moving averages. These preprocessing steps resulted in a reliable dataset that had undergone cleansing, smoothing, standardization, and feature engineering. This processed dataset can be utilized for data analysis, model training, and decision-making purposes.

4.2 Experimental Environment

4.2.1 Hardware environment

The hardware used in this experiment includes Intel Core i7 processor, 16GB memory, and NVIDIA GeForce RTX graphics card.

4.2.2 Software environment

The software used for the experiment includes the TensorFlow 2.0.0 DL framework, the Pandas 1.2.3 data processing library, the Matplotlib 3.3.4 data visualization library, the Stable Baselines 2.10.2 reinforcement learning library, the SciPy 1.6.0 statistical analysis library, and the Scikit-learn 0.24.1 ML library.

4.3 Parameters Setting

In this experiment, the improved Grubbs algorithm is used as the noise reduction method with a noise threshold set to three times the standard deviation. To allow the model to learn and adjust strategies adequately and achieve better performance, both the improved and unimproved models are trained for 2,000 iterations. To balance training speed and model stability, the learning rate for both models is set to 0.003. Furthermore, Z-score normalization is employed in data preprocessing to transform the data into a standard normal distribution with a mean of zero and a standard deviation of one, enabling the model to better learn the data features. The historical trading data is split into training and testing sets based on the time series.

4.4 Performance Evaluation

4.4.1 Analysis of cumulative asset income after noise reduction by Grubbs' improved algorithm

The training of financial trading agents is carried out with a particular emphasis on analyzing the performance of the improved Grubbs algorithm compared to the unimproved deep reinforcement algorithm. The analysis primarily focuses on the cumulative income of assets, as depicted in Fig. 5.

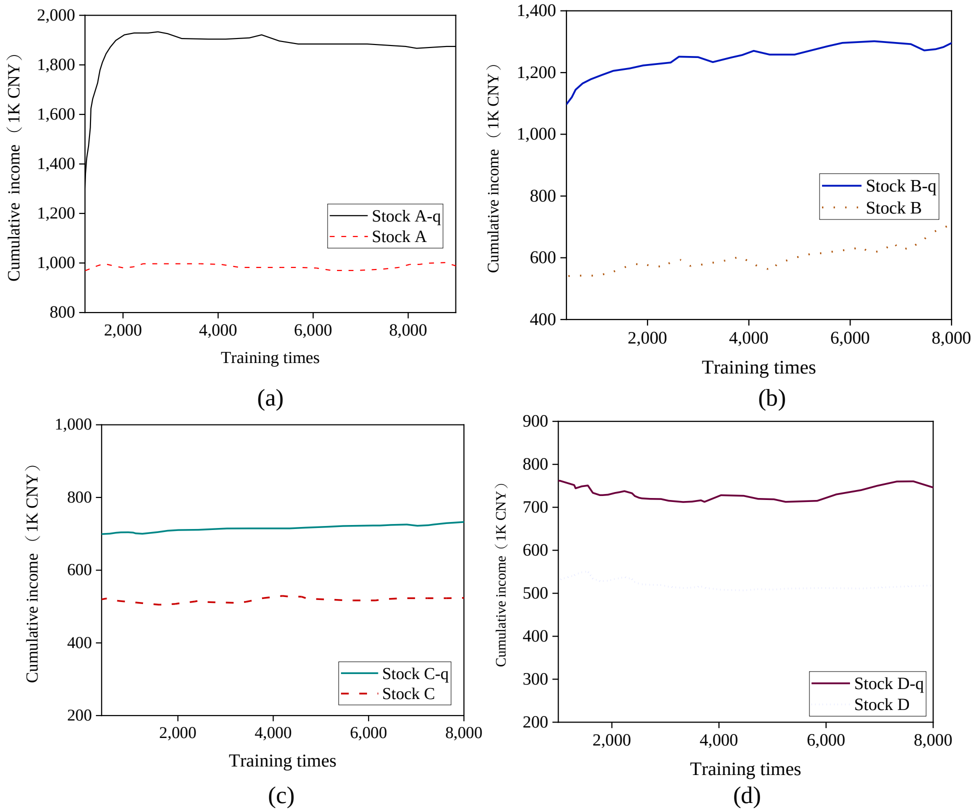

Fig. 5. Cumulative returns of agent training assets under different algorithms; (a) The training result of stock A; (b) The training result of stock B; (c) The training result of stock C; (d) The training result of stock D.

In Figs. 5 (a)-(d) represent the accumulated income results of the agent's assets for Stock A, Stock B, Stock C, and Stock D, respectively. To distinguish the use of the Grubbs' improved algorithm for noise reduction, the results after applying the Grubbs improved algorithm are denoted by "stock name-q." In Fig. 5 (a), after 2,000 training sessions, Stock A reaches a relatively stable state. The cumulative income of assets, such as Stock A-q after noise reduction, approaches RMB 2 million. Conversely, the cumulative income of assets without noise reduction remains relatively low and does not exceed one million yuan. The disparity between the two noise reduction methods is evident. From Fig. 5 (b), Stock B utilizes the improved Grubbs algorithm for noise reduction. After 2,000 training sessions, the accumulated asset income stabilizes and shows a slight increase, surpassing RMB 1.2 million. In addition, the cumulative income of assets without noise reduction does not exceed RMB 800,000. The results for Stock C exhibit a similar pattern to Stock B, albeit with a relatively smaller increase in cumulative income. Stock D gradually stabilizes after 2,000 training sessions. The improved Grubbs algorithm outperforms the unimproved deep reinforcement algorithm in denoising the original stock data, resulting in significantly higher cumulative returns for assets. The improved algorithm can largely eliminate noise interference.

4.4.2 Sharpe ratio analysis of improved grubbs algorithm and traditional reinforcement algorithm after noise reduction

The Grubbs' improved algorithm and the unimproved deep reinforcement algorithm are employed, and the Sharpe ratios are tested after training iterations for the four stock agents, as plotted in Fig. 6.

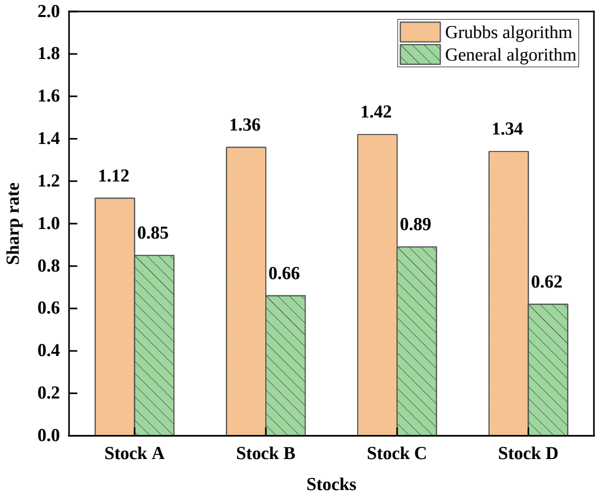

Fig. 6. Sharpe ratio of agent training under different algorithms.

Fig. 6 depicts the Sharpe ratios of stock agents using the Grubbs' improved and traditional algorithms. The results show that after applying the Grubbs noise reduction technique, the Sharpe ratios for Stock A, Stock B, Stock C, and Stock D are 1.12, 1.36, 1.42, and 1.34, respectively. In contrast, the Sharpe ratios obtained using the traditional algorithm are below 1.0. This indicates that the Grubbs' improved algorithm demonstrates better performance in stock trading. The reason for this phenomenon can be explained by the effectiveness of the Grubbs algorithm in removing outliers and noisy data, reducing noise interference in the data, and improving the model's understanding and predictive capabilities of real market conditions. The Grubbs' improved algorithm provides a more stable income curve and enhances decision-making accuracy by reducing the impact of noise on trading strategies. In conclusion, through the improvement of data processing and noise reduction techniques, the Grubbs' improved algorithm exhibits the potential to outperform traditional algorithms in financial trading. It enhances the performance of trading agents and increases the potential for higher returns.

4.4.3 Analysis of maximum retraction rate after noise reduction: Improved grubbs algorithm vs. traditional reinforcement algorithm

To assess the impact of noise reduction using the improved Grubbs algorithm and the traditional reinforcement algorithm, the experiment conducts an evaluation of the maximum drawdown rates. Four stock agents are trained using both methods, and the results are presented in Fig. 7.

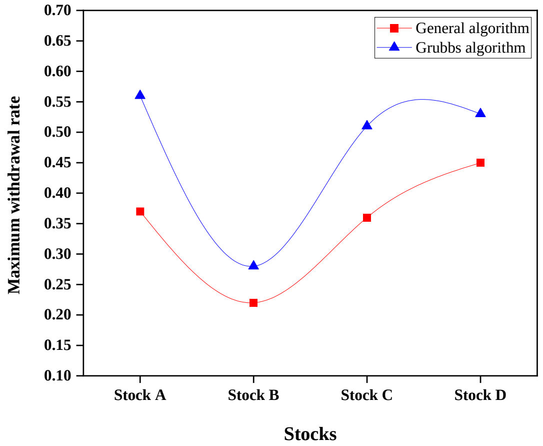

Fig. 7. The maximum drawback rate of agent training under different algorithms.

In Fig. 7, the generic algorithm represents the unimproved deep reinforcement algorithm. By applying the Grubbs noise reduction technique, the maximum drawdown ratios for Stock A, Stock B, Stock C, and Stock D are 0.56, 0.28, 0.51, and 0.53, respectively. These values are significantly higher than those obtained using the unimproved deep reinforcement algorithm. The experimental results demonstrate that the improved Grubbs algorithm has a distinct advantage in terms of maximum drawdown ratios, thus performing better in financial indicators. The reason for this phenomenon lies in the ability of the Grubbs algorithm to effectively reduce the influence of noisy data, remove outliers and noisy data, and allow the model to focus more on real market trends. The Grubbs noise reduction technique lowers the maximum drawdown ratio and enhances the portfolio's protective capability by reducing the interference of noise on trading decisions. Therefore, it is possible to significantly improve the maximum drawdown ratio, protect assets, and enhance the stability of returns by applying the Grubbs noise reduction technique to improve the deep reinforcement algorithm, thereby achieving better performance in financial trading.

4.5 Discussion

The experiments in this study apply the improved Grubbs algorithm to reduce the noise of stock agents and compare the effects compared to traditional reinforcement algorithms. Here are some rich discussions of the experimental results.

(1) Advantages of the improved Grubbs algorithm. Experimental results show that the improved Grubbs algorithm has significant advantages in removing outliers and reducing noise. The algorithm provides a more stable cumulative return curve of assets and improves the accuracy of decision-making by reducing the interference of noisy data on the model. This is critical for financial trading, as noise and outliers can lead to misleading decisions that can affect portfolio performance.
(2) Earnings stability and growth. After using the improved Grubbs algorithm to reduce noise, the cumulative return of equity agents has been significantly improved and shown a

more stable growth trend. This means that the improved algorithm can filter out unwanted noise fluctuations, allowing the model to better capture real market trends. Investors can earn long-term returns more reliably by improving the cumulative return and stability of assets.

(3) Improvement in the Sharpe ratio. The Sharpe ratio is an important measure of the relationship between asset return and risk. The experimental results show that after noise reduction using the improved Grubbs algorithm, the Sharpe ratio of stock agents is significantly improved. This suggests that improved algorithms can reduce uncertainty and volatility in investment strategies, resulting in higher returns for portfolios at the same level of risk. A higher Sharpe ratio reflects more effective investment decisions and better risk control.

(4) Improvement of the maximum drawdown rate. The maximum drawdown rate measures the degree to which asset prices fall from peak to trough. The experimental results show that the maximum drawdown rate of stock agents is significantly reduced by applying the improved Grubbs algorithm for noise reduction. This means that improved algorithms better protect portfolios from price fluctuations and reduce potential losses. A lower maximum drawdown indicates better risk management and asset protection.

Overall, the experimental results highlight the value of the improved Grubbs algorithm in financial trading. Through noise reduction technology, the algorithm provides a more stable and reliable cumulative return on assets and improves key financial indicators, such as the Sharpe ratio and maximum drawdown ratio. This provides better decision support and risk management tools for investors, helping to optimize investment portfolios and achieve higher returns. However, it should be noted that the experimental results may be affected by factors, such as data samples and model parameters. Therefore, further verification and optimization are required in practical applications.

5 CONCLUSIONS

5.1 Research Contribution

This study aims to address the rational allocation of household financial assets, delve into research on stock trading in the financial market, and improve upon the shortcomings of traditional quantitative trading. The proposed model utilizes a dual network structure based on DRL and the DQN algorithm, integrating Grubbs' improved algorithm to denoise stock data samples. The improved Grubbs algorithm and the unimproved deep reinforcement algorithm are employed for the iterative training of four stock agents. The analysis focuses on cumulative returns, Sharpe ratios, and maximum drawdown ratios for different algorithms. The results reveal the superiority of the improved Grubbs algorithm over the unimproved deep reinforcement algorithm in denoising the original stock data and effectively eliminating noise interference. The application of the improved algorithm for noise reduction leads to higher cumulative asset income compared to the unimproved deep reinforcement algorithm. All four stocks exhibit Sharpe ratios above 1.0, and the maximum drawdown performance indicator shows significant advantages. The improved Grubbs algorithm achieves better returns, thereby improving performance indicators for stocks and significantly enhancing the return performance of stock trading algorithms. The main contribution of this study lies in the enhancement of depth and RL algorithms, the development of an improved DQN algorithm, and the utilization of the Grubbs improved algorithm to construct a stock noise reduction model. These algorithms are employed to identify optimal investment strategies.

5.2 Future Works and Research Limitations

One limitation of this study is the narrow scope of financial indicators and the relatively small sample size. Future research should involve expanding the sample size and conducting a comprehensive study of financial indicators. This study provides valuable insights into stock market risk control research and serves as a reference for optimizing the allocation of household financial assets.

REFERENCES

1. H. F. Gholipour and M. E. Dunkley, "Economic policy uncertainty and household financial assets," *Applied Economics Quarterly*, Vol. 65, 2019, pp. 101-114.
2. Z. Huang and Y. Lu, "Analysis of asset risk and household financial asset allocation structure – Empirical analysis from a nonlinear model," *E3S Web of Conferences, EDP Sciences*, Vol. 235, 2021, p. 01039.
3. Y. Ge, H. Chen, L. Zou, and Z. Zhou, "Political background and household financial asset allocation in China," *Emerging Markets Finance and Trade*, Vol. 57, 2021, pp. 1232-1246.
4. L. Zhang and H. Lin, "The impacts of digital finance development on household income, consumption, and financial asset holding: an extreme value analysis of China's microdata," *Personal and Ubiquitous Computing*, Vol. 1, 2022, pp. 1-21.
5. Pfund, L. Patrick, and N. Gabrielle, "Considering financial assets when promoting sense of purpose in older adulthood," *Journal of Aging & Social Policy*, Vol. 1, 2022, pp. 11-13.
6. J. Zou and X. Deng, "Financial literacy, housing value and household financial market participation: Evidence from urban China," *China Economic Review*, Vol. 55, 2019, pp. 2-66.
7. S. R. Baker and L. Kueng, "Household financial transaction data," *Annual Review of Economics*, Vol. 14, 2022, pp. 47-67.
8. L. Li, "Analysis on the influence of financial literacy on the household dinancial assets allocation," *Frontiers in Economics and Management*, Vol. 2, 2021, pp. 41-48.
9. Z. Jiang and A. F. Lynch, "Quadrotor motion control using deep reinforcement learning," *Journal of Unmanned Vehicle Systems*, Vol. 9, 2021, pp. 234-251.
10. S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Proceedings of International Conference on Machine Learning*, 2018, pp. 1587-1596.
11. D. Wu, X. Dong, J. Shen, and S. C. Hoi, "Reducing estimation bias via triplet-average deep deterministic policy gradien," *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 31, 2020, pp. 4933-4945.
12. X. Dong, J. Shen, W. Wang, L. Shao, H. Ling, and F. Porikli, "Dynamical hyperparameter optimization via deep reinforcement learning in tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 43, 2019, pp. 1515-1529.

13. F. Rundo, F. Trenta, A. L. Stallo, and S. Battiato, "Machine learning for quantitative finance applications: A survey," *Applied Sciences*, Vol. 9, 2019, pp. 55-74.
14. T. Zhu, J. J. Xiao, "Consumer financial education and risky financial asset holding in China," *International Journal of Consumer Studies*, Vol. 46, 2022, pp. 56-74.
15. D. Jia, R. Li, S. Bian, and C. Gan, "Financial planning ability, risk perception and household portfolio choice," *Emerging Markets Finance and Trade*, Vol. 57, 2021, pp. 2153-2175.
16. S. P. M. Broekema and M. M. Krame, "Overconfidence, financial advice seeking and household portfolio under-diversification," *Journal of Risk and Financial Management*, Vol. 14, 2021, pp. 553-553.
17. F. A. Fabozzi, J. Simonian, and F. J. Fabozzi, "Risk parity: The democratization of risk in asset allocation," *The Journal of Portfolio Management*, Vol. 47, 2021, pp. 41-50.
18. X. Lu, J. Guo, and H. Zhou, "Digital financial inclusion development, investment diversification, and household extreme portfolio risk," *Accounting & Finance*, Vol. 61, 2021, pp. 6225-6261.
19. Z. Wang and T. Hong, "Reinforcement learning for building controls: The opportunities and challenges," *Applied Energy*, Vol. 269, 2020, pp. 115036.
20. W. Y. Gao and C. Su, "Analysis on block chain financial transaction under artificial neural network of deep learning," *Journal of Computational and Applied Mathematics*, Vol. 380, 2020, p. 112991.

Wan-Dong Gao was born in Changchun, Jilin, P.R. China, in 1964. He received the bachelor's degree from Changchun Normal University, P.R. China. Now, he is a Professor of Weifang University of Science and Technology. His research mainly focuses on the financial sector. E-mail: gwd.2022-3@wfust.edu.cn

Yu-Min Fei was born in RiZhao, ShanDong, P.R. China, in 1967. He received the bachelor's degree from Jilin Business and Technology College, P.R. China. Now, he is a Professor of Weifang University of Science and Technology. His research mainly focuses on the auditing sector. E-mail:feiym162@nenu.edu.cn