Application of Deep Reinforcement Learning Algorithm in Smart Finance

Chunhui CHEN[1] and Yichun ZHOU
Guangdong University of Foreign Studies, China

Abstract. Finance is not only the lifeblood of an economy, but also the lever to adjust the macro-economy. A modern economy is a market economy and essentially a developed financial economy. Based on the analysis of the problems faced by traditional finance and the overview of smart finance, this study puts forward the application of deep learning combined with reinforcement learning in smart finance to solve the problems existing in financial activities for the first time, and verifies through experiments. The model has better data and information processing ability compared with the traditional financial analysis mode. It provides higher quality decision-making information and bring more benefits. Taking a bond rating report as an example, it usually takes about 2 hours for manual in-depth analysis and carding, while it only takes about 2 minutes to interpret and refine the report by using the deep reinforcement learning model. The model has a certain reference value to solve the problems of traditional finance.

Keywords. Deep learning, Reinforcement learning, Deep reinforcement learning, Smart finance, Value quantity, Behavior decision-making

1. Introduction

With the popularity of the concept of wisdom, new concepts such as smart city, smart transportation, smart medical care, and smart finance are gradually emerging in peoples' lives. Driven by AI and big data the two core technologies, smart finance has become the inevitable direction of financial business transformation and high-quality development under the new situation. From the perspective of the demand side, the demand for financial services presents a trend of refinement and ecology. A single financial product and service has become increasingly difficult to meet the actual financial needs of customers. To this end, the financial business is moving towards great integration and collaboration. Its development mode needs to further transform from online and mobile to personalized and intelligent urgently. Through intelligent technology, customers' needs can be quickly perceived. And adaptive financial services can reach customers actively, quickly, and accurately. From the supply side, relying on intelligent technology, the financial business can be analyzed from descriptive and statistical analysis to diagnostic, predictive and decision-making analysis. It can dig out deeper financial business value. Intelligent products can upgrade and improve the existing wired and digital products. It can help to innovate, and improve the supply of financial products, improve the quality and efficiency of financial supply-side reform,

---

[1] Corresponding Author, Chunhui CHEN, Network and Information Center, Experimental Teaching Center, Guangdong University of Foreign Studies, Guangzhou, China. Email: 308469691@qq.com.

optimize the rational allocation of financial assets, reduce the cost of financial business, and enhance the ability of financial services to the real economy. Big data, artificial intelligence and other technologies have been widely used in the financial risk credit, investment research and many other fields of intelligent construction. With the increasing complexity of financial business, the construction of smart finance also puts forward higher requirements for the application and innovation of big data, AI, and other technologies. It is necessary to build a platform-based infrastructure for integrating data resources, computing resources and algorithm resources, to provide more powerful support for the construction of smart finance in data processing, algorithm modeling, and prediction. Aiming at smart finance, this study first analyzes the problems faced by traditional finance and puts forward the concept of smart finance. Then it introduces the deep reinforcement learning model. Finally, it puts forward the application of deep reinforcement learning in smart finance to make smart finance full of the characteristics of transparency, convenience, flexibility, timeliness, efficiency, and security. This study has a certain reference value for the application of artificial intelligence in the financial field.

2. The problems and causes of traditional finance

In the financial market, high-quality information plays an important role in the success of investors. With the modernization of financial transactions and information systems, massive financial data emerge in the financial market. The traditional financial model mainly uses the statistical analysis method to analyze the financial data. The traditional statistical method performs well when the data meet some specific conditions. But the financial price will be affected by various factors, the financial data often presents a non-linear relationship and has high volatility [1]. Therefore, in the actual application, the commonly used statistical analysis methods cannot process the massive financial data, and it is difficult to predict the movement trends of the financial market effectively. It also leads to the financial crisis in the traditional financial market. Therefore, activities of the traditional financial market are often faced with high risk and high cost. At present, there is more and more integration of computer technology to deal with business transactions. Through literature review, most of the methods used in financial analysis and forecasting are fundamental analysis and technical analysis. The premise of using these two methods is market inefficiency. In the economic environment characterized by highly integrated transaction processing information systems, the traditional statistical methods for fundamental analysis and technical analysis have become inefficient and ineffective. [2] The financial industry has the following performances: a) financial institutions are unable to grasp the static and dynamic project information of economic development in time; b) financial institutions are unable to grasp the static and dynamic development information of various industries in time; c) financial institutions are unable to grasp the static and dynamic business information of large and medium-sized enterprises in time; d) all financial institutions can not always be aware of the decisions made by investors for non fully rational people.

Deep reinforcement learning is a new data analysis technology. It abandons the premise that the data needs to meet specific conditions. It can analyze and forecast even when the market is effective. Through deep reinforcement learning, financial analysts can analyze business processes and related internal controls effectively. [3] They can

also analyze and evaluate all transactions recorded and determine deviations automatically and quantitatively. [4]

3. Overview of smart finance integrated with deep reinforcement learning

3.1 The concept of smart finance integrated with deep reinforcement learning

Deep learning [5] is the general name of the deep neural network. It is composed of an input layer, hidden layer, and output layer [6]. The output layer is a single-layer structure. The hidden layer is a single-layer or multi-layer structure. Deep learning refers to self-learning according to training data without programming every problem to be solved. The goal of the deep learning model is to model the data, to analyze the deep correlation within the data and to establish a knowledge framework [7]. The model is used in prediction, classification, and feature extraction. There are three main learning models of deep learning: deep confidence network [8], convolutional neural network and recurrent neural network. [9] Reinforcement learning is a branch of machine learning. Reinforcement learning aims to achieve a goal through multi-step appropriate decision-making under a series of scenarios. Deep reinforcement learning [10] combines the perceptual analysis ability of deep learning with the decision-making ability of reinforcement learning. The most representative researcher is Google's DeepMind [11] team. This team has published two articles on deep reinforcement learning in nature, namely *Deep reinforcement learning algorithm based on video games and the AlphaGo program*. Compared with traditional finance, smart finance is a faster, more efficient, and safer financing. Integrated with deep reinforcement learning, smart finance is to use deep learning model to deal with many complex data. [12] It can analyze the internal correlation of data and establish the corresponding knowledge architecture. Through reinforcement learning, we can make decisions under different scenarios and make more accurate judgments on the consequences of decisions. [13] And we will achieve convenient, safe, high-speed, and efficient financial activities. The operation process of smart finance integrated with deep reinforcement learning is shown in Figure 1.
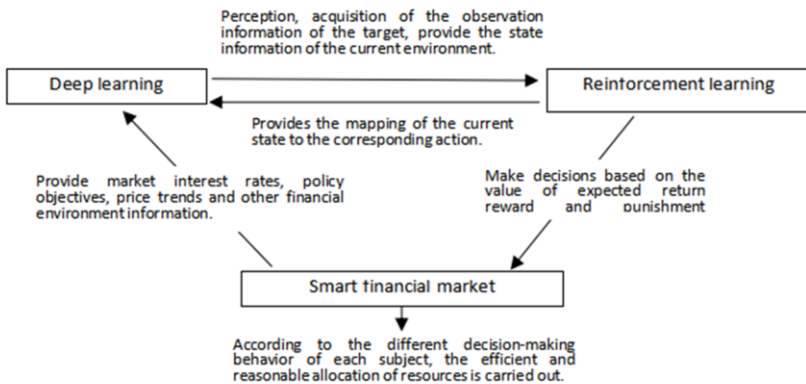

Figure 1. Operation process of smart finance integrating deep reinforcement learning

3.2 Basic algorithm of deep reinforcement learning

In a certain state, the value obtained by performing an action is Q, $\chi_t \in R^{m \times n}$ is the observation of time t ($t=t_1, t_2, \cdots, t_n$), symbol $a_t \in \Lambda$ indicates the action performed under observation χ_t. Λ is the set of all possible actions, the symbol r_i indicates reward (or punishment) for performing the action a_t under observation χ_t. In addition:

$$R_t = \Sigma_{t'=t}^{T} \gamma^{(t'-t)} \cdot r_{t'} \tag{1}$$

$\gamma \in (0,1)$ is discounted factor, R_t is the sum of all cumulative consequences from time t to t_n.

The symbol Q (s, a) is the state action-value function, the state s at time t is

$$s_t = (\chi_1, a_1, \cdots \chi_{t-1}, a_{t-1}, \chi_t) \tag{2}$$

Next, the state action-value function is optimized according to the iterative formula.

$$\begin{cases} Q_{k+1}(s_t, a_t) = Q_k(s_t, a_t) + a_k \cdot \delta_k \\ \delta_k = r_{t+1} + \gamma \cdot \underset{a' \in \Lambda}{maxQ_k(s_t, a')} - Q_k(s_t, a_t) \end{cases} \tag{3}$$

a_k is the learning rate, s_t and a_t are the state and action corresponding to time t, δ_k and γ are the time difference and the discounted factor, a' express the action that state action value function can execute under the kth iteration and s_{t+1}.

The strategy of selecting the optimal executable action in a certain state is to maximize the expected value, which has the following formula.

$$\lim_{k \to \infty} Q^k = Q^*(s, a) = \mathbb{E}_{s' \sim \xi}(r + \gamma \cdot \underset{a'}{maxQ^*(s', a')} | s, a) \tag{4}$$

ξ represents the environment, including all states. s' is the state after executing the action a under the state s, and a' is the possible action.

In practical application, the function approximation strategy is usually used to estimate the action-value function.

$$Q (s,a,\theta) \approx Q^*(s, a) \tag{5}$$

θ is the weight parameter to be optimized, assuming e_t plays back the experience at any time t it stores the experience at any time t.

$$e_t = (s_t, a_t, r_t, s_{t+1}) \tag{6}$$

Suppose that all actions end at the time t_n. Then the experience playback corresponds to the set D= $[e_{t_1}, e_{t_2}, \cdots, e_{t_n}]$.

The state action value function is modified to:
Q (s, a, θ) ⟶ Q(φ(s), a; θ_k)
And there are:

$$s_{t+1} = (s_t, a_t, \chi_t) \tag{7}$$

Experience playback is amended as follows:

$$\begin{cases} D \to \overline{D} = [\overline{e_{t_1}}, \overline{e_{t_2}}, \cdots, \overline{e_{t_n}}] \\ \overline{e_t} = [\varphi(s_t), a_t, r_t, \varphi(s_{t+1})] \end{cases} \tag{8}$$

$$L_k(\theta_k) = \mathbb{E}_{\varphi(s),\alpha \sim \rho(\cdot)}[(y_k - Q(\varphi(s), a; \theta_k))^2] \tag{9}$$

$\rho(\cdot)$ represents the behavior distribution, i.e., $\rho(\varphi(s), a)$ is the probability distribution of state $\varphi(s)$ and the probability distribution of behavior a, y_k represents the target Q (output) corresponding to the kth iteration, and has:

$$y_k = \mathbb{E}_{\varphi(s)' \sim \xi}[r + \gamma \cdot \max_{a' \in \Lambda} Q_k(\varphi(s)', a'; \theta_{k-1}) | \varphi(s), a] \tag{10}$$

The parameter θ_0 is known. After getting the target output y_1, by optimizing the objective function, we get the result θ_1. Finally, the convergence of parameters is realized.

$$\lim_{k \to \infty} \theta_k = \theta_* \tag{11}$$

The parameters are updated by the gradient descent method.

$$\nabla_{\theta_k} L_k(\theta_k) = \mathbb{E}_{\varphi(s),\alpha \sim \rho(\cdot); s' \sim \xi}[r + \gamma \cdot \max_{a'} Q_k(\varphi(s)', a'; \theta_{k-1}) - Q(\varphi(s), a, \theta_k)) \cdot \nabla_{\theta_k} Q(\varphi(s), a; \theta_k)] \tag{12}$$

Suppose that the income obtained by financial activities (such as loans, portfolio, etc.) is Q, the action of the actor is a, and the current financial environment is s. s and a are dynamic changes, and there are the following optimal income calculation functions:

$$Q=Q\,(s,\,a) \tag{13}$$

4. Application of deep reinforcement learning model in smart Finance

The core technology layer of applying deep reinforcement learning model to the construction of smart finance is the big data platform and AI platform. The big data platform provides the AI platform with data and computing resources needed for algorithm modeling. The AI platform can feed back the processing results of complex data to the big data platform and enrich the types of data services provided by the big data platform. The two operate in coordination to enable the construction of financial business scenarios and meet and respond to the demands of intelligent applications in different business fields quickly. The following are some application scenarios of the deep reinforcement learning model in smart finance:

4.1 Loan issuance

Issuing loans is the main source of income for banks. To reduce financial risks, banks need to score and screen the credit of loan applicants before making decisions on granting loans. In this process, banks need to pay costs, such as contract formulation, screening technology, customer potential default loss cost, etc. Profit is the goal of bank loans. If the bank loans can not get more than the cost of income, the bank will refuse to lend. Because the decision-making of the bank brings different incomes in different environments, we can use the above optimized income function to calculate the income of the bank in different states.

Suppose from time t to t_n, all states, the action a_t made by the bank, is making a loan (or refusing to make a loan). The amount of value obtained is Q, and the cost of the loan is C. Using the optimized income calculation function above, the target Q is solved. If Q > C, the loan will be granted (or refused). If Q < C, the loan will be refused (or granted). It is assumed that the value Q obtained by investors is obtained when their utility reaches the maximum, and is higher than that obtained by reinvestment.

4.2 Default risk judgment

Here, default is relative to the behavior of the customer, such as the loan applicant. When the customer signs a loan agreement with the bank, the customer needs to pay the corresponding cost for the obtained funds within a certain period. During the contract period, customers need to use the loan funds to obtain income in order to repay the interest and principal. Under different conditions, there are differences in the ability and possibility of customers to obtain income, as well as the solvency and willingness.

Suppose from time t to t_n, all states, the action a_t made by consumers is keeping an appointment or defaulting. The value obtained for performance is Q, and the borrowing cost is C. If Q > C, the probability of customer default is small; if Q < C, the probability of customer default is greatly increased. It is assumed that the value Q obtained by investors is obtained when their utility reaches the maximum, and is higher than that obtained by reinvestment.

4.3 Calculation of portfolio income

A portfolio can effectively avoid unsystematic risks. Its return is the weighted average of the returns of several securities in the portfolio. The yield of various bonds will fluctuate in different environments, so it is relatively complex to calculate the dynamic change of the portfolio yield. Often, all kinds of securities can only be combined according to the yield at a certain point in time. Using the deep reinforcement learning model, we can effectively calculate the portfolio income and maximize the income through portfolio adjustment.

Suppose that from time t to t_n, in each different state, the action a_t made by us is to choose different funds and shares for the portfolio. The value obtained is Q. Using the above optimized income calculation function, we can solve the target Q when we make different behaviors in the dynamic environment. By comparing the value quantity Q, the optimal action is selected to maximize the income. It is assumed that the value Q obtained by investors is obtained when their utility reaches the maximum, and is higher than that obtained by reinvestment.

5. Model validation

The above-mentioned research on the application of deep reinforcement learning in smart finance has achieved certain results. The main extracted contents are shown in Table 1.

Table 1. Summary of main research results

Core mechanism	Problem solving and advantages	Deficiencies	Applicable scenarios
Objective function, objective network and experience playback	Realize the prediction and decision-making under different states, and apply to the income judgment and decision in financial problems	The calculation depends on the state and can't get the specific strategy, so it is not suitable for the too complex state	It is mainly used in portfolio allocation and income calculation

Corresponding to different application goals, data assets also have rich types. Based on financial scenarios such as wealth management, investment research and risk internal control, this study mainly constructs two types of data assets: feature factor library and knowledge map. The dynamic feature factor library is based on the original data. It is combined with the business logic thinking in the financial field to build a feature index. The feature index can be effectively applied to business analysis and modeling prediction. It is also updated with the evolution of the data continuously, such as the basic information of the customer's age and city, the financial information of mortgage value, the purchase of funds in the stock type, and so on. The financial knowledge map is mainly to integrate structured, semi-structured, unstructured, and other different types of data in the financial field through a large-scale semantic network. These data can be shown in the form of graph connection to precipitate them into a structured knowledge system. Through further analysis and excavation, deeper hidden information is obtained. It makes information retrieval and query more intelligent. It also shows the relationship between different subjects more clearly, studies and judges possible risk events. In addition to the application performance, the algorithm is verified from the requirements of equipment and learning time. Table 2 shows the learning time and computer equipment of the algorithm. Table 3 shows the comparison of the time and effect of using statistical methods and the algorithm to analyze the bond investment report.

Table 2. Learning time and equipment of basic deep reinforcement learning algorithm --DNQ algorithm

Learning time	Computer type
8 days	CPU

Table 3. The time and effect of statistical analysis and the deep learning algorithm in analyzing a bond investment research report

	Time required	The results form	Annualized yield
Statistical analysis	About 2 hours	Words and sentences; Weak correlation	4%-6%
Deep reinforcement learning algorithm	About 2 minutes	Atlas; Strong relevance	6%-10%

The proportion of the annual stock of debt securities assets invested by investors in the total annual debt securities investment assets is taken as the weight of different bonds. The average monthly yield of five-year treasury bonds in the current year is taken as the multiplication factor. Then, the annual weighted average long-term bond yield of each region is taken as the yield of investors' bond investment in that year. To sum up, compared with the traditional finance, in the smart finance mode of deep reinforcement learning, the actors can make more efficient decisions and obtain higher income.

6. Conclusion

In this study, we take the construction of smart finance as the research object. We discuss the application of deep reinforcement learning algorithm to the field of smart finance in the upsurge of artificial intelligence. Compared with the traditional statistical analysis method of finance, the application of deep reinforcement learning in smart finance can obtain valuable information from many nonlinear and complex financial data efficiently. It also helps the actors in financial activities to make more efficient decisions.   Through the application of deep reinforcement learning in smart finance, the entire financial system faces lower risks and obtains higher returns. In the practice of business intelligence, financial institutions need to combine the data system, technology system and application system. It also needs to build a smooth vertical interaction architecture among the big data platform, the AI platform, and the application platform to realize the landing application of intelligent products in pricing, risk assessment, trend prediction and other scenarios. The results show that the application of deep reinforcement learning in smart finance can bring higher benefits for financial activities.

7. Limitations and Prospects

In the application of the model, the updating of micro information and macro information needs timely, dynamic, diversified, and multilateral information resources. The data in the database should be updated continuously to meet these requirements. At present, we can adopt the way of classifying database construction to solve the above problems, such as the existing historical database (static database) and real-time database (dynamic database). And we can adopt the data warehouse to complete the task of providing data management. It is gratifying that many data warehouse products have their own perfect data mining tools. It reduces the time and procedures for the development process of the users' strategic decision support system, but also brings the problem of data lack of scalability. Although the deep reinforcement learning algorithm applied to the smart finance has achieved good results in the simulation test platform, it has not migrated the algorithm to the real trading environment such as investment banks, commercial banks, and exchanges. In the future, with the support of big the data platform and the AI platform, we will carry out the development of data assets and the construction of algorithm model. The deep reinforcement learning model will finally realize the landing and application of intelligent scenario products, such as investment research, risk internal control and wealth management in the financial field.

References

[1]   Labatt S. Environmental Finance: A Guide to Environmental Risk Assessment and Financial Products[J]. Transplantation, 2002, 66(8):405-9.
[2]   Tian M W, Wang L, Yan S R, et al. Research on Financial Technology Innovation and Application Based on 5G Network[J]. IEEE Access, 2019, PP (99):1-1.
[3]   B Z L A, C D H, D X L, et al. Information structures in a covering information system[J]. Information Sciences, 2020, 507:449-471.
[4]   Li, Y. (2019). Reinforcement learning applications. arXiv preprint arXiv:1908.06973.
[5]   Qiu Liwei. Application of deep reinforcement learning in video games [D]

[6]   Scott, H, Lee, et al. Chief complaint classification with recurrent neural networks. [J]. Journal of Biomedical Informatics, 2019.

[7]   Li Y, Fang S, Bai X, et al. Parallel Design of Sparse Deep Belief Network with Multi-objective Optimization[J]. Information Sciences, 2020, 533.

[8]   Charpentier, A., Elie, R., & Remlinger, C. (2021). Reinforcement learning in economics and finance. Computational Economics, 1-38.

[9]   Fu T, Zang Y, Zou Q, et al. Using Deep Learning to Identify Molecular Junction Characteristics[J]. Nano Letters, 2020, 20(5):3320-3325.

[10] Tang J, Dong S, Cui C, et al. Sampled-Data Modeling for Wireless Power Transfer Systems[J]. IEEE Transactions on Power Electronics, 2019, PP (99):1-1.

[11] Groman S M, Keistler C, Keip A J, et al. Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes[J]. Neuron, 2019, 103(4):734-746.e3.

[12] Ismail A, Khalil S, Safieddine A, et al. Smart investments by smart money: Evidence from acquirers' projected synergies[J]. Journal of Corporate Finance, 2019, 56:343-363.

[13] Saufi M S R B M, Ahmad Z A B, Leong M S, et al. Gearbox fault diagnosis using a deep learning model with limited data sample[J]. IEEE Transactions on Industrial Informatics, 2020, PP (99):1-1.