

Research on DQN Stock Trading Strategy Based on Investor's Compound Sentiment

Ningjing Yang^{1,a*}, Deyi Li^{1,2b*}, Yicheng Gong^{1,2c*}, Guici Chen^{1,2d*}

¹Science College, Wuhan University of Science and Technology, Wuhan 430065, China;

²Hubei Provincial Key Laboratory of Metallurgical Industry Process System Science, Wuhan University of Science and Technology, Wuhan 430081, China

Abstract

Both individual investor and institutional investor sentiment affect investors' stock trading decisions. Most existing investor sentiment analyses ignore institutional investor sentiment and treat individual investor sentiment as overall investor sentiment. Quantifying the two types of investor sentiment and expressing them as comprehensive investor sentiment, then exploring their impact on stock investment strategies are conducive to optimizing investment decisions and their investment returns. Based on individual and institutional investors in the stock market, this paper proposes a stock composite sentiment score to measure overall investor sentiment through the text mining method and VADER sentiment analysis. It constructs a DQN single stock trading model based on the composite investor sentiment score using the reinforcement learning DQN algorithm. Through experimental comparison with buy-and-hold and DQN strategies on real stock data, the results show that investor sentiment can effectively optimize stock trading strategies and improve investment returns; compared with individual investor sentiment, comprehensive investor sentiment is better optimized.

Keywords

stock trading, sentiment analysis, VADER, reinforcement learning

1. Introduction

The efficient market hypothesis holds that: market participants are all rational economic persons, the prices in the financial market have already reflected all market information, and rational economic persons will make reasonable decisions based on stock prices. However, the efficient market hypothesis cannot explain some phenomena, such as the equity premium puzzle and the herding effect. Such phenomena suggest that investors' behavior is often influenced by imitative learning among groups and emotional contagion and that changes in investor sentiment can also affect stock prices and trading volumes. Therefore, studying investor sentiment is also important for studying the financial market.

The measurement of investor sentiment in the financial market is an evolving topic. Initially, researchers usually chose a single indicator to reflect investor sentiment; Fisher^[1] and Schmeling^[2] used the consumer confidence index as an indicator of investor sentiment and found that investor sentiment could predict stock market returns to a certain extent. Chinese scholars mostly use the CCTV watch index to measure investor sentiment. Gao^[3] found that investor sentiment is related to short-term market returns. Subsequent studies have found that a single indicator is not representative enough to reflect investor sentiment. Scholars have tried to construct investor sentiment through multiple basic indicators. Baker^[4] constructed an emotion index through principal component analysis using 6 basic indicators. Liu^[5] also used multiple market variables to constitute an investor sentiment index and showed a positive effect of sentiment on the market reflection of surplus announcements through empirical analysis. The sentiment indicators expressed through the composite of multiple indicators result from the mutual equilibrium of multiple macro variables other than sentiment, which expresses the overall market sentiment and cannot directly represent investors' sentiment.

AHPCAI2022@2nd International Conference on Algorithms, High Performance Computing and Artificial Intelligence
EMAIL: ^{a*}yangningjing@wust.edu.cn (Ningjing Yang), ^{b*}lideyi@wust.edu.cn (Deyi Li), ^{c*}gongyicheng@wust.edu.cn (Yicheng Gong),
^{d*}chengguici@wust.edu.cn (Guici Chen)



© 2022 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

With the development of technology, more and more financial investors are gathering in financial forums and expressing their opinions. Through text analysis, researchers can extract investor sentiment directly from forum posts. Smailovic^[6] and Bollen^[7] showed that sentiment analysis of large Twitter text datasets could effectively predict stock market movements. Li^{[8][9]} constructed a quantitative trader that uses publicly available online news and social media data and company-specific news sentiment data to predict stock price movements. Picasso^[10] combined sentiment and technical analysis indicators analyzed from news articles to build a robust stock price prediction model. These studies extracted investor sentiment directly from social media for financial markets through sentiment analysis and achieved good results. However, these studies only focus on the sentiment of individual investors, ignoring the sentiment of institutional investors.

In recent years, deep reinforcement learning has been successfully applied to optimize stock trading strategies and portfolio allocation. Xiong^[11] used the deep deterministic strategy gradient algorithm to conduct stock trading, which significantly improved the trading profit. Carta^[12] proposed a method of repeatedly training DQN agents to reduce strategic risk and maximize investment return. Xu^[13] proposed a deep reinforcement learning automated trading algorithm that combines CNN and experimental results on real stock data, showing that the method significantly outperforms other benchmark methods. Prahlad^[14] proposed an adaptive deep reinforcement learning method to train agents to allocate their portfolios. This method not only uses historical stock price data but also senses the market sentiment of the portfolio. Experiments show that this method has a more robust investment return than the existing baseline. Existing research shows that reinforcement learning algorithm has been successfully applied to simulate stock trading, optimize stock trading strategies.

Financial investors are mainly composed of individual and institutional investors., this paper investigates the impact of investor sentiment on the trading strategy of single stock based on these two types of investors. This paper uses the reinforcement learning DQN algorithm to simulate the trading behavior of a single stock. It compares and analyzes the impact of single individual investor sentiment and composite investor sentiment on the stock trading strategy.

2. DQN stock trading model based on composite investor sentiment

In order to study the influence of investors' composite sentiment on stock trading strategies and investment returns, this paper combines sentiment analysis and reinforcement learning to construct a DQN single stock trading model (named DQN-CE) based on investors' composite emotions.

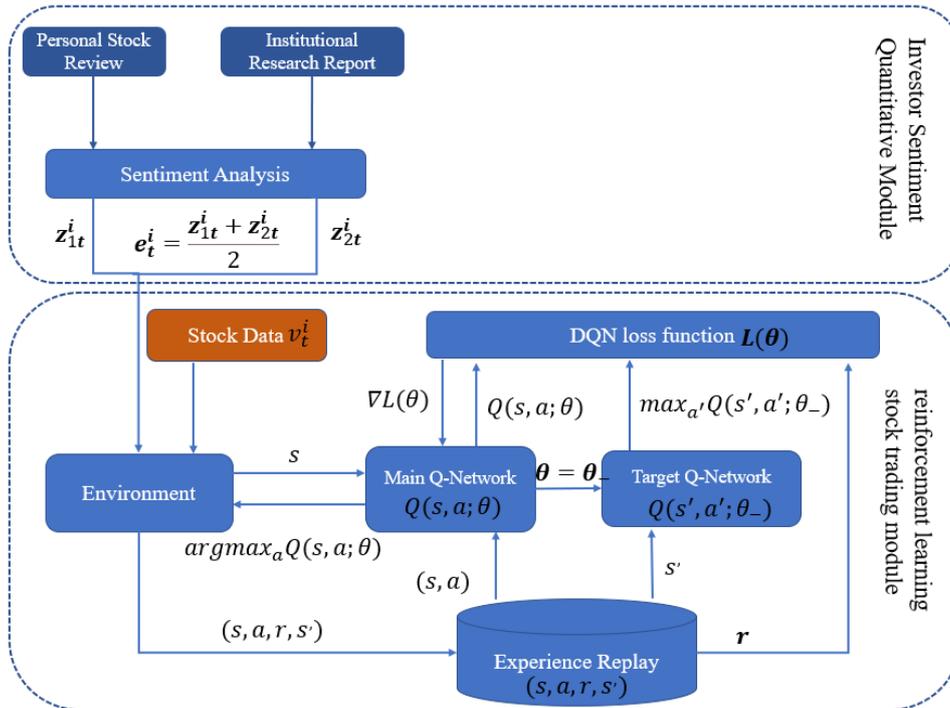


Fig.1 Framework of DQN stock trading model based on composite investor sentiment

The model shown in Fig.1 can be divided into two modules, the investor sentiment quantification module and the reinforcement learning stock trading module. The investor sentiment quantification module is used to measure the overall sentiment of investors. In this module, VADER sentiment analysis is used for the relevant texts of individual and institutional investors to calculate the individual investor sentiment score z_{1t}^i and institutional investor sentiment score z_{2t}^i . The composite investor sentiment score sequence $\{e_t^i\}$ is constructed based on the two types of investor sentiment scores. The sequence $\{e_t^i\}$ will input the reinforcement learning stock trading module with stock price data. The reinforcement learning stock trading module simulates stock trading behavior and learns stock trading strategies.

2.1 Investor sentiment quantification module

VADER is based on a vast dictionary containing sentiment intensity scores for thousands of words, punctuation marks, and web terms. The score of each word in the text is evaluated by querying the intensity score in the dictionary. The sentiment score of the complete text can be obtained by weighting. It usually uses a compound sentiment score to indicate the sentiment tendency of the whole text. A compound sentiment score of -1 is the most negative, and 1 is the most positive.

As shown in the quantitative investor sentiment module in Fig.1, the composite investor sentiment score e_t^i is expressed as the mean of the individual investor sentiment score and the institutional investor sentiment score.

$$e_t^i = \frac{z_{1t}^i + z_{2t}^i}{2}, e_t^i \in [-1,1] \quad (1)$$

In equation (1), z_{1t}^i and z_{2t}^i respectively represent the sentiment scores of individual investors and institutional investors of a single stock i on day t . The composite sentiment score e_t^i represents the overall sentiment of all investors towards stock i , which is more representative than the sentiment of a specific type of investor. The daily composite sentiment score of single stock constitutes a sequence $\{e_t^i\}$.

The text of individual investor sentiment analysis for the single stock is obtained from the respective stock forum on the Eastern Wealth website. All posts were grouped by date to calculate the sentiment score, and the sentiment scores of posts are weighted according to the number of reads and comments on the posts. As equation (2) calculates the individual investor sentiment score z_{1t}^i for day t .

$$z_{1t}^i = \frac{\sum_{j=1}^{n_{1t}^i} [\log(a_{tj}^i + b_{tj}^i) * z_{tj}^i]}{\sum_{j=1}^{n_{1t}^i} \log(a_{tj}^i + b_{tj}^i)} \quad (2)$$

In equation (2), n_{1t}^i denotes the number of posts about stock i on day t , a_{tj}^i, b_{tj}^i denotes the number of reads and comments on the post, respectively, and z_{tj}^i denotes the sentiment score of the stock comment. The individual investor sentiment score z_{1t}^i represents the overall sentiment tendency of individual investors towards stock i on day t .

The text for institutional investor sentiment analysis is derived from individual stock research reports on stock reporting websites. The complete research report is too lengthy to facilitate analysis. Therefore, only the report titles are used for sentiment analysis to calculate the institutional sentiment score z_{2t}^i , as shown in equation (3).

$$z_{2t}^i = \frac{\sum_{j=1}^{n_{2t}^i} z_{tj}^i}{n_{2t}^i} \quad (3)$$

In equation (3), n_{2t}^i denotes the number of stock research reports issued for single stock i on day t , and z_{tj}^i denotes the sentiment score of the research report. The institutional investor sentiment score indicates the sentiment tendency of all institutional investors towards stock i on day t .

2.2 Reinforcement learning based stock trading module

In the reinforcement learning stock trading module in Fig.1, the DQN network receives information about the stock and makes trading actions according to the trading rules. The environment and trading rules of the real stock market are too complex for agent to get a glimpse of the whole picture. In order to facilitate the agent to learn stock trading strategies quickly, we propose three assumptions to simplify the stock trading environment.

Assumption (1): The amount of funds and stocks owned by the agent is limited, which is not enough to affect the stock trading environment.

Assumption (2): An agent can choose to trade once a day, choosing to buy, hold or sell the current stock, and the stock price is based on the closing price of the day.

Assumption (3): The number of shares per trade is the number of all shares held by the agent.

After limiting the assumptions about a single stock's trading environment and trading rules, we formalize the stock trading process as a Markov decision process, expressed as (s, a, p, r, γ) , γ represents the discount factor.

Each state s in the state set S is represented as a tuple $(v_t^i, h_t^i, b_t^i, e_t^i)$. v_t^i denotes the closing price of stock i on day t , $v_t^i \in R^+$; h_t^i denotes the number of stock i held by the agent on day t , $h_t^i \in R^+$; b_t^i denotes the agent's remaining cash balance, $b_t^i \in R^+$. e_t^i is the composite sentiment score of stock i calculated by the quantitative investor sentiment module.

A is action space of an agent, $a_t^i \in \{-1, 0, 1\}$, where

$$a_t^i = \begin{cases} -1; & \text{sell} \\ 0; & \text{hold} \\ 1; & \text{buy} \end{cases} \quad (4)$$

The choice of trading behavior is based on greedy rules, and the model initially sets a greedy value ε . The program randomly generates a number that obeys the uniform distribution within the range of (0, 1) and compares it with ε . If the random number is less than the greedy value, the action with the largest Q value of the corresponding state action among the three actions is selected as the trading action; otherwise, any one of $\{-1, 0, 1\}$ is selected as the trading action.

The reward is the immediate reward received by the agent after executing the transaction action. The agent adjusts the selection of subsequent transaction operations according to the received reward value. Moreover, use the accumulated rewards between all trading dates to measure the pros and cons of the trading strategy. Considering the impact of stock price changes and investor sentiment on stock trading strategies, the rewards of the model are defined as two parts, as shown in equation (5).

$$r_t^i = \frac{v_{t+1}^i - v_t^i}{v_t^i} + 0.1a_t^i * e_t^i \quad (5)$$

In equation (5), $\frac{v_{t+1}^i - v_t^i}{v_t^i}$ is the reward part of stock i . v_t^i is the closing price of the stock i on day t , v_{t+1}^i is the closing price on day $t + 1$. $0.1a_t^i * e_t^i$ is the sentiment reward part of the stock i , which is controlled by a factor of 0.1 so that the emotional reward part of the investor and the return reward part have the same size. Agents can also adjust strategies through the reward of investors' sentiment components.

The action value function $Q_\pi(s, a)$ represents the cumulative expected value obtained by selecting action a following the policy π in the state s , which is expressed as equation (6).

$$Q_\pi(s, a) = E_{s' \sim s}[r(s, a, s') + \gamma Q_\pi(s', a')] \quad (6)$$

The goal of Q-learning is to make $Q(s, a)$ maximization, i.e., maximization of the expected value of the future return of a given state-action pair, as shown in equation (7).

$$Q^*(s, a) = E_{s' \sim s}[r(s, a, s') + \gamma \max_{a'} Q(s', a')] \quad (7)$$

As shown in Fig.1, the agent composed of two networks continuously interacts with the environment to generate experience data (s, a, r, s') and store it in the experience replay. When a sufficient amount of experience data is stored in the experience replay, a small batch of data is randomly selected from the experience replay to train the neural network. $Q(s, a; \theta)$ is the output of the main Q -network, which is used to evaluate the value function of the current state action pair, and the parameter θ of $Q(s, a; \theta)$ is updated in real-time; $Q(s, a; \theta_-)$ represents the output of the target Q -network, which is used to calculate the objective function, as shown in equation (8).

$$y = r + \gamma \max_{a'} Q(s', a'; \theta_-) \quad (8)$$

Thus when the agent takes action a in the environment, the loss function $L(\theta)$ can be calculated, as shown in equation (9).

$$L(\theta) = E \left[\left(r + \gamma \max_{a'} Q(s', a'; \theta_-) - Q(s, a; \theta) \right)^2 \right] \quad (9)$$

According to the loss function $L(\theta)$ in equation (9) calculate the gradient of the error function for updating the parameters θ of the main Q -network. The main Q -network copies its parameters θ to the target Q -network θ_- after a certain number of iterations, thus completing the learning process once.

3. Experiments and results analysis

3.1 Experimental setup

The experiments used Python 3.7 to implement the algorithm model, the PyTorch function module to train the deep learning network, and the reinforcement learning DQN agent; the python matplotlib 3.1 library to visualize the data, and the Vader function module in the NLTK library to calculate the sentiment score. The transaction cost is set at 0.05% of the transaction amount. mini-batch size is equal to 512 and the initial learning rate is 0.01.

The experiment data comes from Yahoo Finance, and selects four stocks that have attracted much attention in the A-share market. They are Sany, Gree Electric, China Merchants Bank and UNIS. The data of all stocks are divided into two parts, the data from January 2018 to December 2020 are used as the training dataset for model training (denoted as M_1); the data from January 2021 to January 2022 were used as the test dataset (denoted as M_2) for model back-testing.

To explore the effect of investor sentiment on stock trading strategies and investment returns, a DQN stock trading model based on individual investor sentiment (denoted as DQN-SE) and a DQN stock trading model based on investor sentiment score (denoted as DQN-CE) are constructed, respectively. In order to compare the effectiveness of the models, they are compared with two benchmark strategies, the buy-and-hold strategy (denoted as B&H) and the DQN algorithmic trading strategy (denoted as DQN), and the models are evaluated using the cumulative stock returns and Sharpe ratios.

3.2 Experimental results and analysis

In the DQN-SE model, only the influence of individual investor sentiment on stock trading strategies and returns is considered. The individual investor sentiment score is used as a state component to replace the composite investor sentiment in equation (5), with the return set to $r_t^i = \frac{v_{t+1}^i - v_t^i}{v_t^i} + 0.1a_t^i * z_{1t}^i$. The models DQN-SE, DQN-CE, DQN were back-tested on the test set M_2 after 1000 rounds of training on the training set M_1 .

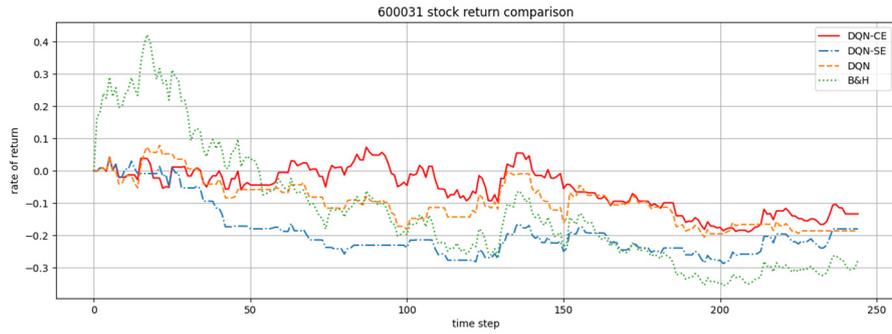


Fig.2 Sany cumulative return

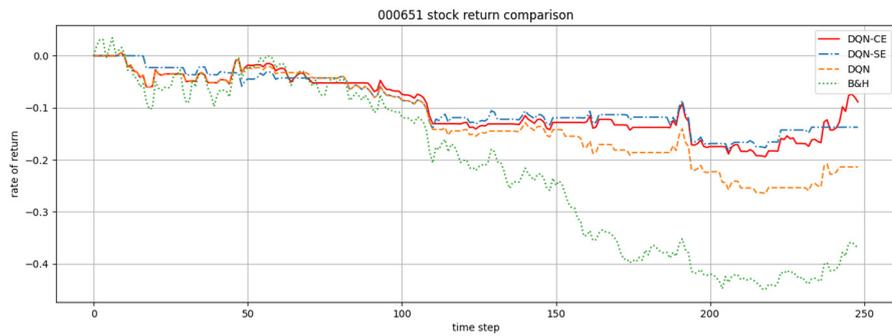


Fig.3 GREE cumulative return

As can be seen in Fig.2 and Fig.3, the stock prices of Sany and GREE showed a clear downward trend, and all four trading strategies showed losses with negative returns. Among them, the return curve of DQN-CE trading strategy is higher than that of other strategies, with the smallest loss and the largest loss of B&H strategy.



Fig.4 UNIS cumulative return

As shown in Fig.4, the overall trend of UNIS stock price is up, and all four trading strategies show profits. The return curve for trading with the DQN-CE strategy is the highest, with a return of over 30%, the DQN-SE strategy also has a return of over 20%, and the DQN and B&H strategies both have a return of around 10%.

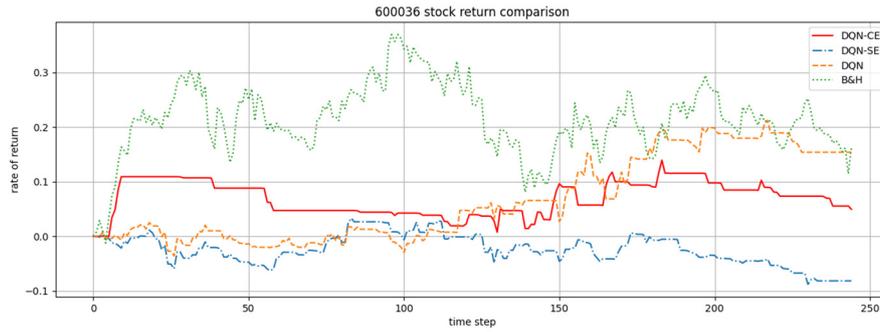


Fig.5 China Merchants Bank cumulative return

As shown in Fig.5, the return curves of the four trading strategies on China Merchants Bank stocks vary widely, with the DQN-CE strategy, DQN strategy, and B&H strategy showing positive returns and the B&H strategy gaining the most; trading with the DQN-SE strategy shows negative returns.

Compared to the B&H strategy, the three trading strategies DQN, DQN-SE, and DQN-CE had more returns on three stocks, Sany, Gree Electric, and UNIS, but the B&H strategy had the highest return curve on the China Merchants Bank stock. China Merchants Bank's stock is a high market capitalization stable stock, and the stock price has been showing a more stable upward trend during the training period, showing a cyclical, more significant fluctuation during the testing period; there is an inevitable lag between individual investors sentiment and the trend of stock price movement, so it isn't easy to profit in the stock market.

Compared to the DQN strategy model without investor sentiment analysis, the DQN-SE and DQN-CE models with investor sentiment analysis have higher return curves on all four stocks, with the DQN-CE model having the highest return curve. The results suggest that investor sentiment has an ameliorating effect on stock investment returns and that the use of composite investor sentiment has a more pronounced effect than individual investor sentiment.

After analyzing the model's investment return performance, the performance of different trading strategies is measured using the Sharpe ratio, which combines investment return and investment risk and describes the excess return that an investor can earn per unit of risk taken.

Tab.1 Performance evaluation of the four models

Approach	Sany Heavy		Gree Electric		China Merchants Bank		UNIS	
	Cumulative Return	Sharpe Ratio	Cumulative Return	Sharpe Ratio	Cumulative Return	Sharpe Ratio	Cumulative Return	Sharpe Ratio
B&H	-27.77%	-0.68	-37.03%	-1.81	16.34%	0.35	11.23%	0.18
DQN	-18.55%	-0.84	-21.38%	-1.82	15.38%	0.75	8.75%	0.39
DQN-SE	-17.95%	-0.94	-13.73%	-1.44	-8.17%	-0.84	25.49%	0.49
DQN-CE	-13.32%	-0.55	-8.87%	-0.83	4.94%	0.19	35.52%	1.06

As can be seen from Tab.1, for the stocks of Sany and GREE, whose share prices continue to fall and are in a loss-making state, the DQN-CE strategy reduces losses while reducing trading risk. For UNIS, whose stock price is in an uptrend, the trading return using the DQN-CE strategy is 35.52%, and the trading risk is also the lowest. As for China Merchants Bank, which has a volatile stock price, using investor sentiment does not optimize the trading strategy and improve investment returns, using the traditional B&H strategy would have generated higher investment returns. Compared to the DQN-SE strategy, the DQN-CE algorithmic strategy performed better on four stocks, with an average improvement in returns of about 8.16%.

From the experimental results, the DQN-CE strategy performs better compared to DQN-SE. The DQN-CE strategy has better performance on stocks with a clear trend of stock price changes, so the DQN-CE strategy is more suitable for stocks with a clear trend of movements. The DQN-CE strategy is less applicable to stocks whose stock prices fluctuate in a range with no obvious trend.

4. Conclusion

This paper combines two types of investor sentiment using composite investor sentiment score to measure investor sentiment in the stock market, simulates stock trading through the reinforcement learning DQN algorithm, and conducts empirical analysis using real data from four stocks. The results show that: (1) for stocks with more obvious stock price change trends, investor sentiment analysis is helpful in optimizing stock trading strategies and reducing investment risks. (2) Compared to single individual investor sentiment, composite investor sentiment is more conducive to optimizing trading strategies and improving investment returns.

5. References

- [1] Fisher Kenneth L, Statman M. Consumer Confidence and Stock Returns[J]. *The Journal of Portfolio Management*,2003,30(1):115-127.
- [2] Schmeling M. Investor Sentiment and Stock Returns: Some International Evidence[J]. *Journal of Empirical Finance*, 2009, 16(3):394-408.
- [3] Gao L. Equity Transfers and Market Reactions[J]. *Journal of Emerging Market Finance*,2008,7(3):293-308.
- [4] Baker M, Wurgler J. Investor Sentiment and the Cross-Section of Stock Returns[J]. *The Journal of Finance*,2006,61(4):1645-1680.
- [5] Liu J, Wen YL, Xu JL. Corporate governance, investor sentiment and surplus response[J]. *Statistics and Decision Making*,2020,36(07):154-158.
- [6] Smailović J, Grčar M, et al. Stream-based Active Learning for Sentiment Analysis in the Financial Domain[J]. *Information Sciences*,2014, 285:181-203.
- [7] Bollen J, Mao H, et al. Twitter Mood Predicts the Stock Market[J]. *Journal of Computational Science*, 2011,2(1):1-8.
- [8] Li Q, Wang T, Gong Q, et al. Media-aware Quantitative Trading Based on Public Web information[J]. *Decision Support Systems*, 2014, 61:93-105.
- [9] Li Q, Wang TJ, Li P, et al. The Effect of News and Public Mood on Stock Movements[J]. *Information Sciences*, 2014, 278:826-840.
- [10] Picasso A, Merello S, et al. Technical Analysis and Sentiment Embeddings for Market Trend Prediction[J]. *Expert Systems with Applications*, 2019, 135:60-70.
- [11] Xiong ZR, Liu XY, Zhong S, et al. Practical Deep Reinforcement Learning Approach for Stock Trading.[J]. *CoRR*,2018,abs/1811.07522.
- [12] Carta S, Ferreira A, Podda AS, et al. Multi-DQN: An Ensemble of Deep Q-learning Agents for Stock Market Forecasting[J]. *Expert Systems with Applications*, 2021,164: 113820.
- [13] Xu J, Zhu YK, Xing CH. Research on financial trading algorithms based on deep reinforcement learning[J]. *Computer Engineering and Applications*,2022,58(07):276-285.
- [14] Koratamaddi P, Wadhvani K, Gupta M, et al. Market Sentiment-aware Deep Reinforcement Learning Approach for Sock Portfolio Allocation[J]. *Engineering Science and Technology an International Journal*, 2021,24(4):848-859.