Automatic Stock Profit Maximization via Deep Reinforcement Learning and Correlation Analysis

Division of Computer Science, Sookmyung Women's University, Seoul, Republic of Korea
E-mail: minjoo@sookmyung.ac.kr, soohyun.park@sookmyung.ac.kr

Abstract—**In this paper, we propose a novel Deep Deterministic Policy Gradient (DDPG)-based stock profit maximization approach with enhanced predictability and strategic decision-making by integrating correlation analysis between stocks. The proposed algorithm is designed to consider user budgets as well as evaluated to utilize real-world stock market data. Lastly, the performance of the proposed method is evaluated and verified that our proposed method outperforms the others.**

I. Introduction

Background and Motivation. With the advancement of Artificial Intelligence (AI) technologies, algorithms based on AI are being applied across various fields. Among these, the stock market is a critical area where the application of AI is essential due to its inherent complexities. Research on decision-making for stock price predictions using deep reinforcement learning (DRL) algorithms has gained significant attention. Particularly, the Deep Deterministic Policy Gradient (DDPG) algorithm offers a promising approach by enabling agents to learn optimal trading strategies through continuous interaction with the market environment. Many studies have effectively maximized returns by designing DRL models that account for market dynamics, achieving notable gains compared to average stock performance. However, challenges remain in improving performance regarding volatility prediction and correlation analysis among stocks. If it becomes possible to predict price fluctuations or expected price changes of stocks with similar trends, this could lead to maximized returns or minimized losses. Nevertheless, there is a lack of research considering the relationships between stocks within the DRL-based price prediction framework. This gap underscores the need for further exploration into the correlation analysis of stock relationships.

For this reason, this paper proposes a DDPG-based autonomous stock revenue maximization system that utilizes correlation analysis to examine stock relationships. The correlation analysis will provide insights that can inform strategic trading decisions. Through those points, the motivation of this research is to improve the leverage DDPG to identify and exploit the correlation between the returns of two stocks, thereby improving the expected user values by controlling the optimal stock decision under the correlation between the returns of two stocks.

Contribution. By integrating this analysis with considerations of user capital, the proposed DDPG-based algorithm offers a promising approach by enabling agents to learn optimal trading strategies through continuous interaction with the market environment. As a result, it will autonomously determine whether to buy or sell stock. To validate the effectiveness of our proposed approach, real-world stock market data from 2020 to 2023 is used for performance evaluation.

Organization. The rest of this paper is organized as follows. Sec. II introduces related research trends regarding DRL-based stock market control. Sec. III presents the main concept of the proposed DDPG-based autonomous stock trading control with correlation analysis between stocks. Sec. IV evaluates the performance of the proposed idea. Lastly, Sec. V concludes the paper and discusses future research directions.

II. Related Work

A. Deep Reinforcement Learning

Reinforcement Learning (RL) is a methodology that enables agents to get optimal sequential actions, named optimal policy, through interactions with their environment. As authors described in [1], several RL algorithms, such as reward-based learning and policy optimization methods, can be applied to numerous problem-solving scenarios.

DRL, which combines the concept of RL with deep learning, allows the agent to learn the optimal policy in complex state spaces. Through the deep learning process of DRL, we can learn the relation between the complex state and action spaces [2]–[4]. DRL can be distinguished according to whether the action space is discrete or continuous. The representative RL algorithm of the discrete case is Deep Q-Network (DQN). It has achieved human-level performance in Atari games [5], demonstrating the potential of DRL and validating the effectiveness of Convolutional Neural Networks (CNNs) in addressing challenges posed by high-dimensional state spaces. When the action space is continuous, DDPG can be used based on actor-critic architecture. DDPG algorithm enables simultaneous learning of both policy and value functions, facilitating efficient exploration and exploitation in complex environments [6]. Through the characteristics of DDPG, it is suitable for stock price prediction applications such as determining optimal buy and sell points in financial markets.

B. Prediction of Stock Price

In the field of stock price prediction, there has been considerable research focused on reinforcement learning (RL) approaches that utilize a continuous action space. Notably, several studies have employed algorithms such as DDPG and

Advantage Actor-Critic (A2C) to effectively model trading strategies. In [7], the A2C algorithm was utilized for portfolio allocation in the LQ45 index on the Indonesian Stock Exchange. Compared to the Mean Variance Portfolio and Equal Weight Portfolio methods, the A2C approach demonstrated superior performance with a smaller number of portfolios (five). In [8], the authors integrated the prediction and allocation steps of portfolio management to enable optimal decision-making. Utilizing the TD3 algorithm, they proved that Deep Learning Reinforcement Learning is more effective than other types of machine learning algorithms.

In [9], DDPG algorithm was employed using Tucker decomposition along with technical analysis and stock return covariates, providing a comprehensive analysis by incorporating technical indicators. [10] proposed two strategies: the Critic-only trading strategy known as GDQN (Gated Deep Q-learning trading strategy) and the Actor-critic trading strategy known as GDPG (Gated Deterministic Policy Gradient trading strategy). These methods applied the Gated Recurrent Unit (GRU) to extract informative financial features that can represent the unique characteristics of the stock market for adaptive trading decisions. Both strategies were found to be particularly effective during periods of sharp price fluctuations and declining market trends. [11] introduced the Trading Deep Q-Network algorithm (TDQN), which is a modified version of the DQN approach.

III. DDPG Algorithm for Stock Revenue Maximization using Correlation Analysis

In this section, we describe a new DRL-based stock price prediction with an analysis of the relationships between stocks.

A. Correlation Analysis

Before implementing reinforcement learning algorithms for stock price prediction, it is imperative to conduct a correlation analysis among the selected stocks. The rationale for conducting this correlation analysis lies in its ability to facilitate the adjustment of the exploration-exploitation ratio within DRL model. This adjustment allows for a greater emphasis on exploitation when stocks exhibit a higher correlation.

For the correlation analysis, the correlation coefficient is a vital analytical instrument that facilitates our understanding of the relationships between different stocks. By quantifying the degree to which stocks move to one another, we can identify potential co-movements that may impact our predictive model.

Comprehending these relationships is essential for several reasons. It enables us to assess the strength and direction of interactions among stocks, thereby informing our decisions regarding which stocks to incorporate into our reinforcement learning framework. Stocks that exhibit strong correlations may share similar market dynamics, which could significantly influence our modeling approach. To calculate the correlation coefficient, it is necessary to determine the rate of price change for each stock. The rate of price change means daily return, which can be denoted as $R^d(t)$ for date t, and it can be formulated as follows,

$$R^d(t) = \frac{P(t) - P(t-1)}{P(t-1)}, \tag{1}$$

where $P(t)$ and $P(t-1)$ correspond to today's price of one stock and yesterday's price, respectively. Through (1), it is possible to normalize the price data and also get a more meaningful comparison of the correlation performance between the two stocks.

In addition, by utilizing the Pearson correlation coefficient which is widely used for correlation measurement between two values, it is possible to quantify the degree of linear relationship between the daily returns of two stocks. The formulation for two random variable values X and Y can be expressed as follows,

$$C_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}. \tag{2}$$

In our stock market scenario, the variables X and Y represent the daily returns of each stock price calculated by (1). In (2), $Cov(X, Y)$ represents the covariance between the variables X and Y while σ_X and σ_Y denote the standard deviations of X and Y. This formula quantifies the degree to which the two variables move in relation to one another. The Pearson correlation coefficient formulated in (2), denoted as C_{XY}, is a statistical measure that quantifies the strength and direction of the linear relationship between two continuous variables [12]. It ranges from -1 to $+1$. This coefficient is widely used in various fields, including finance, to assess the relationship between different assets, helping traders and analysts understand how movements in one asset may influence another. In the context of stock trading, a strong correlation between stocks can inform investment strategies and risk management decisions.

We can interpret the correlation coefficient value to determine the strength and direction of the relationship. Specifically, if the value is close to 1, there is a strong positive correlation, suggesting that the stocks move in the same direction. On the contrary, if the value is close to -1, that indicates a strong negative correlation, implying that the stocks move in opposite directions. Finally, if the value is near 0, that suggests no significant correlation between the two stocks. This analysis provides valuable insights into the co-movement of these stocks, which is essential for effective portfolio trading strategies.

B. DRL-based Autonomous Stock Trading Control

The model for the DRL-based Autonomous Stock Trading Control System is designed to facilitate intelligent trading decisions within complex financial environments. By employing the DDPG algorithm, this model seeks to optimize trading strategies through the utilization of continuous action spaces, thereby enabling more nuanced decision-making compared to traditional discrete-action methods.

- **State**: The holding ratio represents the proportion of the portfolio's value that is attributed to the currently held stocks. Profit and loss is another critical metric, indicating the ratio of the current portfolio value relative to the initial capital, thereby reflecting the extent to which the portfolio value has appreciated since its inception. The price fluctuation relative to the average purchase price quantifies the degree of change in the current stock price compared to the average purchase price, calculated as,

$$\frac{\text{Current Price}}{\text{Average Purchase Price}} - 1. \tag{3}$$

The correlation coefficient is calculated on a 60-day basis.

- **Action**: In the context of autonomous stock trading control, the model must determine whether to buy, sell, or refrain from taking any action regarding a particular stock. The actions are denoted as $A_{\text{buy}}, A_{\text{sell}}, A_{\text{hold}} \in A$, with each action corresponding to the values $[-1, 0, 1]$, respectively.
- **Reward**: The reward function is based on the Return on Investment (ROI) for the portfolio. The reward $R(t)$ is calculated when $B(t)$ represents the balance, $P(t)$ denotes the current price of the stock, and $N(t)$ indicates the number of shares held. Therefore, the $R(t)$ can be directly formulated as follows,

$$R(t) = B(t) + P(t) \times N(t). \tag{4}$$

For the implementation of our proposed stock trading model, which employs reinforcement learning and correlation analysis, we have established a systematic approach to adjust the exploration parameter, referred to as ϵ. This adjustment is predicated on the strength of the correlation between predicted stock prices and their actual values. Specifically, when the correlation coefficient exceeds the threshold of $\epsilon = 0.7$, we reduce the ϵ by 0.9. This strategic modification diminishes the likelihood of exploration—where the agent seeks new information—and concurrently enhances the probability of utilizing its learned results, whereby the agent capitalizes on its previously acquired knowledge and experiences.

The reinforcement learning-based decision-making process regarding exploration versus the application of learned knowledge is fundamentally influenced by the relationship between a generated random value and the ϵ. When the random value is less than ϵ, the agent opts to explore, thereby engaging in random decision-making within the predefined action space. This approach allows the agent to investigate potentially beneficial actions that it has not previously considered. Conversely, when the random value exceeds the ϵ threshold, the agent shifts its focus towards leveraging its learned results. During this phase, the agent relies on its existing predictions to identify the action that is anticipated to yield the highest return. This decision-making process is grounded in the agent's learned experiences, allowing it to make informed choices based on prior knowledge rather than engaging in random exploration.

By dynamically adjusting the exploration parameter in response to the correlation between predicted and actual stock prices, our model effectively balances the dual imperatives of exploration and the application of learned knowledge. This balance optimizes trading performance within the complexities of financial markets, ultimately enhancing the agent's ability to make informed and strategic trading decisions.

IV. Performance Evaluation

We analyze the stocks of Samsung Electronics and SK Hynix to examine the correlation between different stock assets within the KOSPI index, which comprises n stocks. Notice that the reasons why the two companies are selected are (i) they are in the same industry sector, i.e., electrical engineering and manufacturing, and also (ii) they are the two largest electronics companies in the Republic of Korea.

The data utilized includes daily stock closing prices from January 2020 to December 2023, sourced from the Korea Exchange (KRX). The dataset is divided into training and testing subsets.

- Training set: January 1, 2015 $\sim$ December 31, 2019
- Test set: January 1, 2020 $\sim$ November 30, 2021

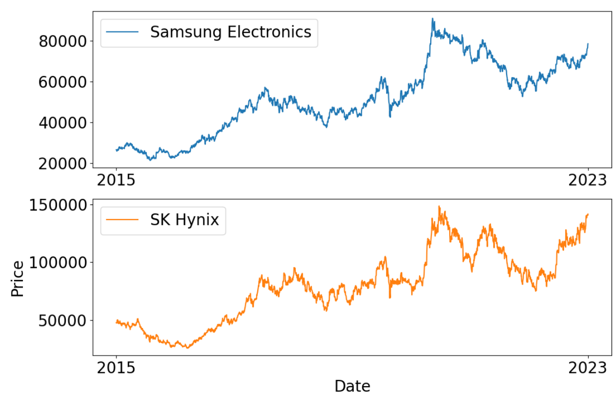

Fig. 1: Price Dynamics of Samsung Electronics and SK Hynix.

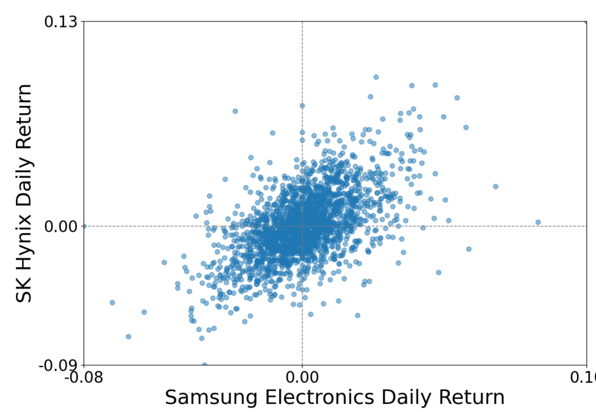

Fig. 2: Comparison of Fluctuation Rate between Samsung Electronics and SK Hynix.

In this study, we aimed to analyze the correlation between the two stock items, i.e., Samsung Electronics and SK Hynix,

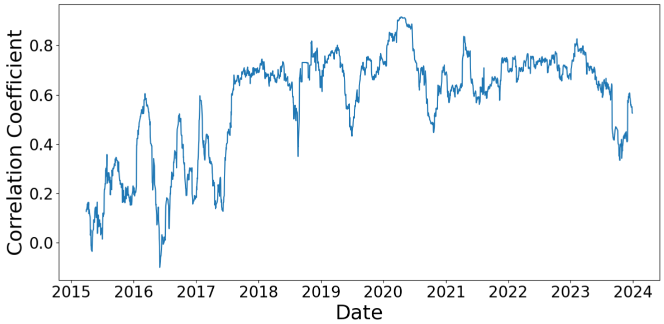

Fig. 3: Correlation Analysis.

and utilize the correlation during the training of a DRL model. To achieve this, we conducted a thorough correlation analysis. Fig. 2 illustrates the price trends of the two stocks over time, while Fig. 3 presents a scatter plot of their respective prices. From these graphs, it is evident that there are a positive correlation between the two stocks.

The results of our correlation analysis are depicted in Fig. 4, which shows the distribution of correlation coefficients. A significant portion of the values exceeds 0.5, indicating a high level of correlation between the two stocks. Specifically, our simulation yielded an average correlation of 0.58, signifying a moderate positive relationship between the predicted values and the actual stock prices. Furthermore, we observed that 70 percent of the correlations exceeded the threshold of 0.5, suggesting that a substantial proportion of the predictions made by our model were positively correlated between the two stocks. According to [13], this indicates a moderate correlation, which can enhance the predictive accuracy of our reinforcement learning model.

In the implementation of the DRL algorithm, the trading commission is set at 0.00015, and the trading tax is established at 0.002. For training purposes, the number of epochs is set to 1000, and the initial capital is defined as 100 million Korean won (KRW).

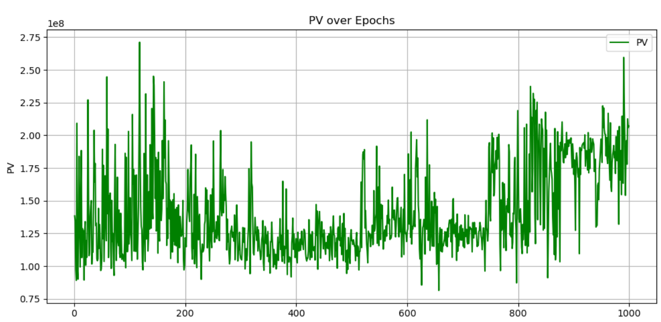

Fig. 4: Reward Convergence in Training.

The test results yielded a portfolio value of $153,181,589$, with a notable return of $54.18\,\%$. As illustrated in Fig. 5, our proposed model demonstrated superior performance in terms of stability.

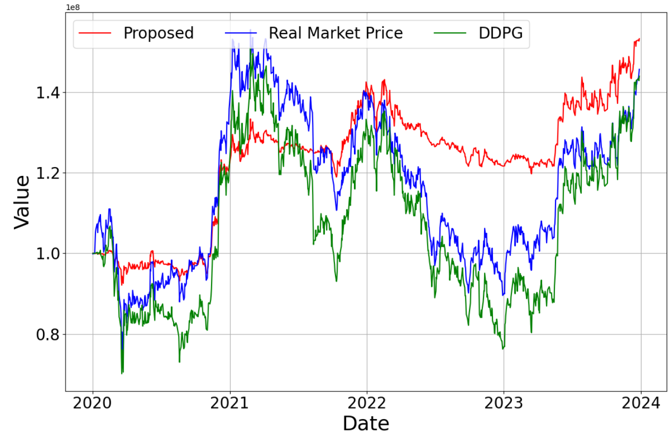

Fig. 5: Time-Series Analysis of Stock Returns.

The performance evaluation of our proposed *Deep Reinforcement Learning (DRL)-based Autonomous Stock Trading Control* was conducted using two distinct methods to assess the effectiveness of the trading strategies implemented.

- **Trading Based on Initial Investment**: In the first method, the simulation began on the first date of the test set period. An initial investment was made in two selected stocks: Samsung Electronics and SK Hynix. The total investment was allocated equally, with each stock receiving half of the benchmark amount. Specifically, the agent purchased 905 shares of Samsung Electronics and 527 shares of SK Hynix, maximizing the allowable quantity based on the investment constraints. At the conclusion of the test period, the final portfolio value is evaluated. The results indicated that the total amount reached $145,613,000\,\text{KRW}$, yielding a return of $45.61\,\%$. This outcome demonstrates the effectiveness of the trading strategy based on the initial investment and the market conditions during the test period.
- **Analysis without Correlation Coefficient**: The second method involved analyzing the results obtained when the correlation coefficient between the two stocks was not utilized in the trading strategy. All other parameters remained consistent with the baseline reinforcement learning algorithm. In this scenario, the agent executed trades without considering the correlation metrics that typically inform decision-making. The results from this method showed a final portfolio value of $134,215,598\,\text{KRW}$, resulting in a return of $34.22\,\%$. This outcome suggests that, even in the absence of correlation considerations, the trading strategy remained effective, achieving a higher return than the first method.

Overall, the evaluation results illustrate the performance of the DRL-based trading model under different conditions. The first method highlights the importance of initial investment strategies, while the second method demonstrates that the trading strategy can still yield substantial returns without relying on correlation metrics. These findings provide valuable insights into the adaptability and robustness of the proposed trading control system. Furthermore, we also observed that our proposed trading control system presents the superior performance when the system operates for relatively longer time.

V. Conclusions and Future Work

This paper proposes a DDPG-based automatic stock profit maximization method under consideration of correlation analysis between stocks in order to improve expected profit performance from a long-term perspective. Note that the use of correlation analysis is not considered in the literature.

For our future work directions, it is anticipated that analyzing the correlations among a broader range of stocks will lead to even higher returns. This strategy could enhance the model's adaptability and performance in dynamic market conditions, ultimately maximizing profit potential across diverse investment scenarios. By incorporating more comprehensive data, the approach aims to be refined to achieve superior results in stock profit maximization.

Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2024-RS-2024-00436887) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

References

[1] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[2] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 30, no. 1, 2016.

[3] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[4] P. F. Christiano, J. Leike, T. Brown, M. Martic, S. Legg, and D. Amodei, "Deep reinforcement learning from human preferences," *Advances in neural information processing systems*, vol. 30, 2017.

[5] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[6] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[7] B. A. Luthfianti, D. Saepudin, and A. F. Ihsan, "Portfolio allocation of stocks in index lq45 using deep reinforcement learning," in *2022 10th International Conference on Information and Communication Technology (ICoICT)*. IEEE, 2022, pp. 205–210.

[8] T. Kabbani and E. Duman, "Deep reinforcement learning approach for trading automation in the stock market," *IEEE Access*, vol. 10, pp. 93 564–93 574, 2022.

[9] J. Jang and N. Seong, "Deep reinforcement learning for stock portfolio optimization by connecting with modern portfolio theory," *Expert Systems with Applications*, vol. 218, p. 119556, 2023.

[10] X. Wu, H. Chen, J. Wang, L. Troiano, V. Loia, and H. Fujita, "Adaptive stock trading strategies with deep reinforcement learning methods," *Information Sciences*, vol. 538, pp. 142–158, 2020.

[11] T. Théate and D. Ernst, "An application of deep reinforcement learning to algorithmic trading," *Expert Systems with Applications*, vol. 173, p. 114632, 2021.

[12] P. Sedgwick, "Pearson's correlation coefficient," *Bmj*, vol. 345, 2012.

[13] H. Akoglu, "User's guide to correlation coefficients," *Turkish journal of emergency medicine*, vol. 18, no. 3, pp. 91–93, 2018.