

# 강화학습 기반 주가 예측 모델 기술 동향

김민주, 박수현  
숙명여자대학교

minjoo@sookmyung.ac.kr, soohyun.park@sookmyung.ac.kr

## Technological Trends in Stock Price Prediction Models Based on Reinforcement Learning

Minjoo Kim, Soohyun Park  
Sookmyung Women's University

### 요약

강화학습은 높은 변동성을 지닌 금융 시장에서의 의사 결정 과정을 개선할 잠재성을 지니고 있다. 본 논문은 강화학습을 활용한 주가 예측 모델의 개발 동향과 효과를 분석하고, 이를 통해 금융 시장에서의 의사 결정 과정을 개선할 수 있는 잠재력과 앞으로의 연구 방향에 대해 탐구한다. 더불어, 강화학습 활용 기술의 도전 과제와 한계점에 대하여 논의한다.

### I. 서론

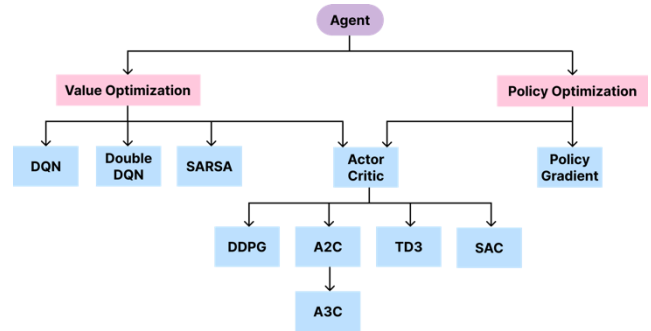
주가 예측은 투자 수익률을 높이고, 투자 위험을 줄이며, 효율적인 자산 운용을 가능하게 한다. 기술의 발전과 함께 주가 예측을 위한 다양한 기술들이 연구되고 있으며, 그 중에서도 강화학습에 대한 탐구가 주목받고 있다. 강화학습은 기계학습의 한 분야로, 주어진 환경에서 최적의 결정을 내리는 방법을 학습한다. 이 과정을 통해 특정 목표를 달성하기 위한 최적의 전략을 찾아간다. 전통적인 주가 예측 모델들이 과거의 데이터 패턴을 기반으로 예측하는데 집중하는 반면 [1], 강화학습을 활용한 접근 방식은 시장의 변화와 불확실성 속에서도 지속적으로 학습하고 적응하는 능력을 갖추게 된다 [2].

본 논문에서는 강화학습을 이용한 주가 예측 모델의 개발 동향을 분석하고, 이를 통해 얻은 인사이트와 향후 연구 방향을 제시하고자 한다. 주가 예측에 강화학습을 사용하면 변동이 크고 예측이 어려운 금융 시장에서의 의사 결정 과정을 개선하고, 잠재적으로 높은 수익을 창출할 수 있는 새로운 기회를 제공할 수 있다. 또한, 강화학습 기반 주가 예측 모델이 직면한 주요 도전 과제와 한계점을 탐구하고, 이를 극복하기 위한 기존 모델들이 효율성과 정확성을 개선할 수 있는 방안을 모색한다.

### II. 본론

과거부터 주식 매매 시스템을 개선하기 위한 연구에서 강화학습이 활용되어 왔다. 특히 다중 에이전트 Q-Learning 알고리즘으로 주식 매매 시스템을 개발하여 수익률을 높이는 효과를 보였다 [3].

최근에는 Deep Q-Network (DQN)과 Advantage Actor-Critic (A2C) 같은 심층 강화학습 기법이 주가 예측 분야에서 성과를 보이고 있다. <그림 1>은 본 논문에서 언급된 강화학습 알고리즘을 분류한 내용이다. [4]에서는 model-free 심층 강화학습 알고리즘들을 가치 최적화와 정책 최적화로 나누어 분류하였다. 이 중 본 논문에서 언급된 강화학습 알고리즘들을 표기하였다. [5]는 Asynchronous Advantage Actor-Critic (A3C)



알고리즘으로 동적으로 포트폴리오를 구성하고 시장 전체의 학습이 개별 학습보다 효과적임을 증명하였다. [6]은 DQN 이 Deep Deterministic Policy Gradient (DDPG)보다 안전성이 높다는 결과를 도출하였다. [7]은 기존의 DQN 이 action-value 를 과대평가하고, 거래액이 많을 수록 추정 오차가 커지는 문제를 방지하기 위하여 강화학습 방식의 거래 전략인 Gated Deterministic Policy Gradient Trade Strategy (GDPG)와 Gated Deep Q-learning Trade Strategy (GDQN) 방식을 제시하였다. GDPG 방식이 GDQN 방식보다 더 복잡한 주식 시장 환경을 학습할 수 있으며, 안정적인 수익을 제공하는 것으로 평가되었다. DQN, Policy Gradient (PG), Advantage Actor-Critic (A2C) 알고리즘을 비교 분석한 연구에서는 안전성과 수익성 사이에서 사용자의 선호에 따라 적합한 모델을 선택할 수 있으며 A2C 학습 기법에서 학습 기간을 길게 설정하면 높은 수익성을 기대할 수 있음을 제시하였다 [8]. 이 연구들은 변동성이 큰 주가 데이터를 사용하여 모델을 학습시킬 때 안정성을 개선할 수 있음을 보여주었다 [7, 8]. 또한, DQN, A2C, A3C 모델을 비교한 다른 연구에서는 DQN 과 A2C 가 우수한 수익률을 달성하였으며, A2C 에서 가장 높은 수익성을 보였다 [9]. 이 연구에서는 학습 데이터의 양을 조절하여 수익성과 안전성을 향상시킬 수 있음을 강조하였다. 학습 기간과 변동성에 차이를 두어 더 높은 안전성 또는 수익성을 기대할 수 있다고 제시하였다. Soft Actor-Critic (SAC) 모델을 활용한 연구에서는 시장의 위험 구조와 수익률 간의 최적 균형을 도출하는 결과를 얻었다 [10]. [11]은

그림 1 Model-free 심층 강화학습 알고리즘

기술 지표와 뉴스 기사를 상태 표현에 추가하여 에이전트의 성능을 정성적 분석으로 향상시켰으며, Actor-Critic 기반 알고리즘인 Twin Delayed DDPG (TD3)을 사용하여 포트폴리오 할당 문제를 해결하였다.

강화학습 기반 주가 예측 모델은 여전히 여러 도전 과제에 직면해 있다. 실제 시장 환경을 모델링하는 데 있어서의 복잡성으로 인하여 모델이 현실 세계의 시장을 정확히 반영하지 못할 수 있다. 또한, 강화학습 모델은 과적합되기 쉬워, 새로운 데이터에 대한 일반화 능력이 저하될 수 있다. 특정 기간에 대한 과적합 문제를 해결하기 위하여 다양한 기간 동안 에이전트를 훈련시키는 방법이 연구되었다 [12]. 이 연구에서는 에피소드 랜덤화를 통해 다양한 기간의 데이터를 학습시키는 것과 중간에 전체 기간의 데이터를 학습시키는 것을 결합하는 것이 효과적임을 입증하였다.

### III. 결론

강화학습을 이용한 주가 예측은 금융 시장에서의 의사 결정 과정을 개선할 잠재력을 가지고 있다. 본 논문에서 다양한 연구 사례들을 통해 강화학습이 주가 예측에서 높은 수익성을 제공할 수 있으며, 변동성이 큰 금융 시장에서 효과적으로 적용할 수 있는 능력을 갖추고 있음을 확인하였다. 특히, 심층 강화학습 모델인 GPGD와 GDQN, 그리고 A2C가 주가 예측에서 우수한 성능을 보여주었으며, 기술 지표와 뉴스 기사를 상태 표현에 추가하여 에이전트의 성능을 향상시키는 방법도 유효함을 알 수 있었다.

그러나 강화학습 기반 주가 예측 모델은 여전히 여러 도전 과제와 한계에 직면해 있다. 주식 시장의 복잡성을 정확히 예측하기 위해서는 거시경제 지표, 기업 실적, 투자자 심리 등 다양한 요소들을 종합적으로 고려하는 모델링 방법을 연구하여야 한다. 또한 투자자 개인의 심리와 선호를 반영할 수 있는 개인화된 모델링 접근 방식이 필요하다. 이를 통해 각 투자자의 특성에 맞춘 정밀한 예측이 가능해지고, 모델의 효율성과 정확성을 향상시킬 수 있다. 추가적으로, 강화학습을 기반으로, 다른 인공지능 알고리즘을 결합하는 앙상블 기법을 활용한 연구가 필요하다. 이를 통해 단일 모델보다 더 정확하고 신뢰성 있는 모델을 개발할 수 있다. 이와 더불어 실시간 데이터 처리 및 분석 능력을 강화하고, 다양한 데이터로부터 얻은 정보를 효과적으로 통합하여 적용할 수 있는 알고리즘의 개발도 진행되어야 한다. 이는 모델의 실제 투자 환경에서의 적용 가능성을 확대하고, 투자자들에게 신뢰할 수 있는 예측 정보 제공이 가능하다. 이러한 연구는 투자 분야에서의 강화학습 기반 기술 발전을 촉진하고, 금융 산업에서 이점을 제공할 수 있는 기여를 할 것으로 기대된다.

### ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학지원사업의 연구결과로 수행되었음 (2022-0-01087)

### 참 고 문 헌

[1] 박인찬, 권오진 and 김태윤. "시계열 모형을 이용한 주가지수 방향성 예측" 한국데이터정보과학회지 20, no.6 (2009) : 991-998.

[2] Jae Won Lee, "Stock price prediction using reinforcement learning," *ISIE 2001. 2001 IEEE International Symposium on Industrial Electronics Proceedings (Cat. No.01TH8570)*, Pusan, South Korea, 2001, pp. 690-695 vol.1, doi: 10.1109/ISIE.2001.931880.

[3] 김유섭, 이재원, and 이종우, "다중 에이전트 Q-학습 구조에 기반한 주식 매매 시스템의 최적화," 한국정보처리학회논문지 B11.2 (2004): 207-212.

[4] A. K. Shakya, G.Pillai, and S. Chakrabarty, "Reinforcement learning algorithms: A brief survey," *Expert Systems with Applications*, vol. 231. Elsevier BV, p.120495, Nov-2023.

[5] 윤성식, and 이석준, "강화학습을 활용한 주식 포트폴리오 구성 및 트레이딩 시뮬레이션에 관한 연구," *경영과 정보연구* 40.4(2021): 185-204.

[6] N.Kodurupaka, B. H. M, S. T. Pecheti and A. J, "Deep Reinforcement Learning in Stock Trading: Evaluation DDPG and DQN Strategies," *2024 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, 2024, pp. 1-7.

[7] Wu, Xing, et al, "Adaptive stock trading strategies with deep reinforcement learning methods," *Information Science* 538(2020): 142-158.

[8] 황호현, 김용훈, and 이영훈, "강화학습 기반 주식 자동 매매 모델 전략 제안," *대한산업공학회지*, vol. 47, no. 4, pp. 399-405, 2021.

[9] 윤성식 and 이석준, "강화학습을 활용한 주식 포트폴리오 구성 및 트레이딩 시뮬레이션에 관한 연구," *경영과 정보연구*, vol. 40, no. 4, pp. 185-203, 2021.

[10] 이우식, "SAC 강화학습 알고리즘을 활용한 자산배분 전략에 관한 연구," *경영컨설팅연구*, 제 22 권 제 4 호 통권 제 76 호, 2022, 367-373.

[11] T. Kabbani and E. Duman, "Deep Reinforcement Learning Approach for Trading Automation in the Stock Market," in *IEEE Access*, vol. 10, pp. 93564-93574, 2022.

[12] 우세형 and 김도국, "에피소드 랜덤화 및 액션 노이즈를 통한 강화학습 기반의 포트폴리오 최적화 성능 향상," *Journal of KIISE*, vol. 51, no. 4, pp. 370-379, 2024. (10.5626/JOK.2024.51.4.370)