

심층 강화학습을 이용한 A3C 기반의 자동 주식거래시스템 구현

고수민¹, 김성준², 여의주³, 최민경⁴, 이규영⁵

¹동덕여자대학교 데이터사이언스전공

²한국방송통신대학교 컴퓨터학과

³서울과학기술대학교 스마트ICT융합공학과

⁴서울여자대학교 경제학과

⁵한국과학기술원 정보보호대학원

kosumin060404@gmail.com, rlatjdwns09@knou.ac.kr, 23100026@seoultech.ac.kr,

choimin06@swu.ac.kr, leeahn1223@kaist.ac.kr

The Implementation of an Automated Stock Trading System based on A3C using Deep Reinforcement Learning

Sumin Ko¹, Sungjun Kim², EuiJu Yeo³, Minkyung Choi⁴, Kyuyoung Lee⁵

¹Dept. of Data Science, Dongduk Women's University

²Dept. of Computer Science, Korea National Open University

³Dept. of Smart ICT Convergence Engineering, Seoul National University of
Science and Technology

⁴Dept. of Economics, Seoul Women's University

⁵Graduate School of Information Security, KAIST

요 약

최근 금융업에서는 기계학습의 활용범위가 빠르게 확대되고 있다. 본 논문에서는 심층강화학습(DRL)을 활용한 A3C 알고리즘 기반 자동 주식거래시스템을 구현하고, 이를 적용한 결과 저점과 고점에서 우수한 의사결정 판단을 수행하여 높은 수익을 달성하는 것을 실험을 통해 입증하였다.

1. 서론

금융업이 갖는 수리적 또는 규칙(rule) 기반의 의사결정 속성으로 인해 인공지능 도입이 증가하고 있다. 인공지능의 학습능력이 향상될수록 예측 분석의 정밀도가 높아져, 변화하는 주식 시장에서의 핵심 경쟁력으로 작용할 것으로 기대된다[1].

심층 강화학습은 딥러닝과 강화학습을 결합한 머신러닝 기반 방법론이며, 두 학습 방법의 장점을 결합하여 더욱 효과적인 의사결정을 가능하게 한다. 본 연구에서는 DRL 모델로 주가 데이터를 학습하여 주식 시장의 변동성에 적응할 수 있도록 하였다.

2. 관련 연구

기존의 주가 예측 방법으로는 SVM[2]과 ARIMA 등의 기술이 있다. 이러한 주가 예측 방법은 특정 조건에서는 유용하지만, 변동성이 큰 환경에서는 성능에 한계를 보인다[3]. 반면에 강화학습은 시장 데이터를 실시간으로 학습하고 미리 정의한 규칙 대신 환경에 따라 최적의 행동을 선택한다. 구글 딥마인드의

AlphaGo, Tesla의 자동차 자율주행 시스템과 같은 성공 사례를 통해, DRL의 잠재력이 주식 거래에서도 충분히 발휘될 수 있음을 예상할 수 있다.

A3C(Asynchronous Advantage Actor-Critic) 알고리즘은 Actor-Critic 구조를 통해 정책 기반과 가치 기반 방법론을 결합한 방식이다. Actor는 정책 네트워크를 통해 행동을 선택하며, Critic은 상태-행동 쌍의 가치를 평가하고 TD 오차(Temporal Difference Error)를 기반으로 정책을 개선한다.

$$\nabla_{\theta} \log \pi_{\theta}(a_t | s_t) \cdot A(s_t, a_t) \quad (1)$$

$$\delta_t = r_t + \gamma V_{\theta'}(s_{t+1}) - V_{\theta'}(s_t) \quad (2)$$

$$L(\theta) = E[\nabla_{\theta} \log \pi_{\theta}(a_t | s_t) Q_c(s, a)] \quad (3)$$

식(1)은 Actor 업데이트 수식 (정책 경사), 식(2)는 Critic 업데이트 수식 (TD 오차) 그리고 식(3)은 Actor-Critic 손실함수를 뜻한다.

3. 제안 모델

A3C 알고리즘은 게임이나 로봇 제어 등에서 주로 사용되지만, 이를 주가 예측에 적용하려면 주식 시장

의 시계열적 특성을 반영한 최적화가 필요하다.

식(1)~(3)에서 상태 s_t 를 주가, 거래량, 그리고 기술적 지표(이동평균선(SMA), 거래량 이동평균선(VMA))와 같은 다양한 시계열 데이터로 구성했다. 특히 주가 변동성과 과거 365일, 180일, 90일, 30일 동안의 최고점, 최저점 정보를 추가하여, 에이전트가 주가의 과거 추세를 학습하고 미래 변동성을 더 정확히 예측할 수 있도록 했다.

매수 및 매도 수량을 조정하는 Action은 0주부터 잔고가 있는 만큼(max)의 매수와 매도, 관망으로 총 $2 \times \max + 1$ 가지의 Action이 존재한다. Action을 추가함으로써 학습 모델이 자산을 헤지(Hedge)할 수 있는 기능을 갖추게 했다. 이로써 세밀한 거래 전략이 가능하며 리스크 관리에 있어 유리해졌다.

마지막으로 실험에 사용한 주가 데이터들에 대해 가장 우수한 성능을 보인 시드(seed) 값을 채택하여 최적화된 모델을 구성하였다. 엡실론 탐험 과정에서 발생하는 랜덤성을 보정하고, 동일 조건에서 동일 결과를 얻을 수 있도록 한 것이다.

4. 실험 및 고찰

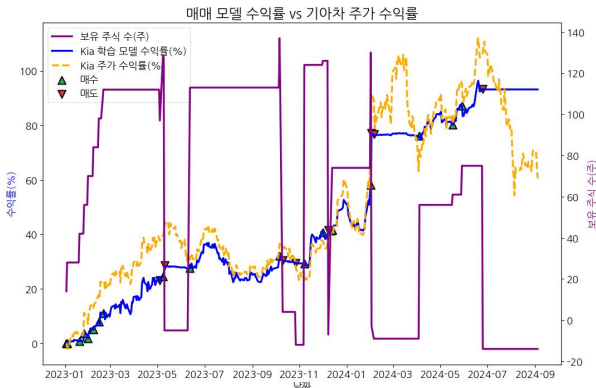
4.1 실험 환경

Train 기간은 2004년부터 2022년까지 19년이고, Test 기간은 2023년 1월부터 2024년 9월까지 1년 9개월간 일일 종가 실시간 데이터를 사용했다. 이에 따라 Time step은 '일' 단위로 설정해 1개 종목(기아자동차)을 A3C 알고리즘으로 모델링하여 실험하였다.

Agent의 초기 현금은 한국투자증권 모의주식 투자 시작금액에 착안해 동일하게 1000만원으로 설정했다. 또한 초기 주식 보유량을 0주로 하고, 거래 수수료 0.00015%와 거래세 0.002%를 적용했다.

4.2 실험 결과

그림 1에서 모델의 수익률이 전반적으로 주가와 유사한 양상을 보이며 변동하고 있음을 확인할 수 있다.



(그림 1) 모델 수익률과 주가 비교 그래프.

자동 매매 초기인 2023년 1월, 소량의 주식을 여러 번에 걸쳐 매수하는 모습을 통해 상승장을 타고 있는

모습을 보인다. 2023년 5월과 6월경 주가의 소폭 급락과 급등하는 구간에 맞춰 각각 매도와 매수 결정을 내린 것은 해당 모델의 Train 데이터 훈련 과정에서 실제 학습이 잘 이루어졌음을 의미한다.

특히 2024년 7월, 가지고 있는 주식을 다량 매도하는 모습은 주가 하락 추세 가능성에 따른 행동 결정이라 볼 수 있다. 실제로 매도 이후에 주가 급락이 확인되며 모델의 우수한 리스크 관리 능력을 확인했다.

4.3 실험 고찰

상기 실험 결과를 통해, A3C 기반의 주식 거래 시스템이 병렬 학습을 통해 빠른 수렴 속도를 가지며, 에이전트의 행동 범위를 확장함으로써 보다 안정적인 수익을 창출할 수 있음을 확인했다. 특히, 강화학습의 누적보상을 추구하는 특성에 기인하여 장기적 안목에서 수익을 극대화하는 모습을 볼 수 있었다.

아울러 테스트 기간 기아자동차(kia) 주식의 누적 수익률은 +60.58%를 보이며 A3C 알고리즘 기반 모델이 최적의 매도/매수 시점을 결정하는 데 효과적임을 입증하였다. 또한 주가 움직임에 대한 예측 정확도가 우수하여 중장기적인 상승세의 수익률을 확보하였음을 확인했다.

5. 결론

본 연구는 강화학습 기술을 주식 거래에 적용한 A3C 모델을 제안하고, 모델 수익률과 주가를 비교 분석하고자 강화학습 기반의 자동 매매 모델을 개발하였으며, 실험을 통해 그 유용성을 입증했다.

※본 논문은 과학기술정보통신부 대학디지털교육역량강화사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.

참고문헌

[1] 서정호(토론). "금융업의 인공지능 활용과 정책과제." KIF 정책분석보고서, vol. 2022, no. 1, pp. 1-114, 2022.

[2] Kim Y, Shin E & Hong T, "Comparison of stock Price Index Prediction Performance Using Neural Networks and Support Vector Machine", Korea Internet e-Commerce Association, vol.4, no.3, pp.221-243, 2004

[3] Pokou, F., Sadefo Kamdem, J., Benhmad, F., "Hybridization of ARIMA with Learning Models for Forecasting of Stock Market Time Series.", Computational Economics, vol.63, pp. 1349-1399, 2023