

XGBoost 기반 포지션 진입 제어와 PPO 강화학습을 결합한 나스닥 100 선물 지수 트레이딩 알고리즘 제안

김파란하늘*, 한재웅*, 이연서, 현장훈†

국립한밭대학교

lights040901@gmail.ac.kr*, hanjaewoong1233@gmail.com*, dustj0406@naver.com,
jhhyeon@hanbat.ac.kr†

A Proposed Nasdaq-100 Futures Index Trading Algorithm Combining XGBoost Based Position Entry Control and PPO Reinforcement Learning

Paranhaneul Kim*, Jaewoong Han*, Yeonseo Lee, Janghun Hyeon†

Hanbat National University

요약

본 논문은 금융 시계열 예측 모델의 레깅(Lagging) 현상과 기존 PPO 강화학습 모델의 학습 과정에서 무분별한 포지션 진입으로 인한 불필요한 손실이 누적되는 문제를 완화하기 위해 PPO와 XGBoost를 결합한 포지션 진입 제어 알고리즘을 제안한다. 제안 알고리즘은 XGBoost가 추정한 포지션 진입 성공률을 보상함수에 반영하여 PPO의 학습을 보조하고, 성공 확률이 50% 미만일 경우 PPO가 진입을 선택하더라도 포지션 진입을 억제한다. 제안 알고리즘의 성능을 평가하기 위해 E-mini NA SDAQ 100 선물 데이터를 활용하여 트레이딩 시뮬레이션을 수행하였으며, 실험 결과, 단일 PPO 대비 15.79%, XGBoost 대비 11.3% 수익률을 향상시켜 제안 알고리즘의 우수성을 증명하였다.

I. 서론

시스템 트레이딩은 과거의 주가, 거래량, 기술적 지표 등을 활용하여 매매 규칙을 설계하고 자동으로 매매를 진행하는 투자 방법이다. 최근에는 AI 기술의 발전과 함께 금융 시계열 데이터를 활용한 시스템 트레이딩 연구가 활발히 연구되고 있으며, 대표적으로 XGBoost와 같은 머신러닝 모델 및 LSTM 기반 딥러닝 모델 등 다양한 알고리즘이 제안되었다[1, 2]. 그러나 금융 시계열은 금리, 환율 등의 다양한 경제 상황에 의해 비정상성(non-stationarity)인 특징을 가지고, 모델이 시장의 변동성을 학습하지 못할 경우 과거 데이터의 추세를 따라가는 레깅(Lagging)현상으로 인해 예측 결과가 실제 매매에서는 재현되지 않는 문제가 발생할 수 있다. 이러한 한계를 보완하기 위해 최근에는 강화학습(Reinforcement Learning)을 이용한 트레이딩 연구가 진행되고 있으며[3], 대표적으로 클리핑(Clipping) 기반 목적 함수를 이용해 모델의 학습 안전성을 확보하는 PPO(Proximal Policy Optimization)기반 알고리즘이 널리 활용되고 있는 상황이다[4]. 그러나, 이러한 강화학습 기반 알고리즘은 보상 함수 설계에 따라 투자 성능 편차가 크게 발생하며, 학습 과정에서 모델의 무분별한 포지션 진입으로 인해 불필요한 손실이 누적되는 등에 의해 투자 성능이 떨어지는 문제가 존재한다[5]. 본 논문에서는 이러한 한계를 완화하기 위해 PPO와 XGBoost를 결합한 포지션 진입 제어 알고리즘을 제안한다. 제안 알고리즘은 기술적 지표로 구성하였다. 단기·장기 추세를 반영하기 위해 이동평균(MA), 지수이동평균(EMA)을 활용하였고, 변동성 정보를 파악하기 위해 볼린저 밴드(Bollinger Band)기반 지표를 사용하였다.

II. 본론

2.1. 데이터 수집 및 전처리

본 연구에서 사용한 데이터는 E-mini NASDAQ 100으로, 2019년 ~ 2025년까지 총 7년간의 데이터를 사용하였다. 이를 각 1년씩 총 7개의 기간으로 분할한 뒤, 3년을 Train data set, 1년을 Validation set, 1년을 Test set으로 구성하였으며 시장의 최신 데이터를 반영하기 위해 Sliding Window를 사용하여 실험을 진행하였다.

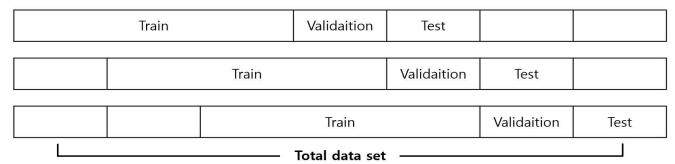


Fig 1. Sliding Window를 이용한 데이터 분할

모델 입력으로 사용된 특징(Feature)은 가격 및 거래량으로부터 산출한 기술적 지표로 구성하였다. 단기·장기 추세를 반영하기 위해 이동평균(MA), 지수이동평균(EMA)을 활용하였고, 변동성 정보를 파악하기 위해 볼린저 밴드(Bollinger Band)기반 지표를 사용하였다.

2.2. 제안 알고리즘

본 연구에서는 PPO 강화학습 모델과 XGBoost를 사용하여 알고리즘을 설계하였으며 이는 Fig 2에서 확인할 수 있다. 강화학습의 거래자에 해당하는 에이전트(Agent)는 매 시점 Sell, Buy, Hold 중 하나를 선택하며, Sell을 선택한 경우 익절/손절 폭을 사전에 미리 정의한 후보중 하나로 선택한다. 상태(State)는 모델의 입력으로 사용된 특징을 기반으로 구성하였으며, 에이전트가 현재 거래 상황을 인지할 수 있도록 포지션 보유 여부, 선택된 익절, 손절 폭, 현재 수익률을 포함하도록 구성하였다. 모델의 보상 함수는 매매의 실현 수익을 최대화하면서 수수료 등 거래 비용을 최소화

* 주저자: 김파란하늘, 한재웅 † 교신저자: 현장훈

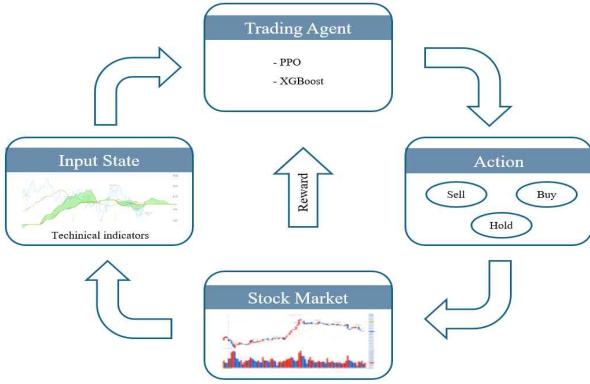


Fig 2. 제안 알고리즘의 프레임워크

하도록 설계하였다. 특히, XGBoost가 예측한 포지션 진입 성공률을 보상함수에 추가하여 강화학습 모델의 학습을 보조하였으며, 성공률이 높게 예측되는 포지션 진입일수록 추가 보상을 부여하여 PPO가 유리한 구간의 진입을 선호하도록 하였다. 또한, XGBoost가 예측한 성공률이 50% 미만일 경우, PPO가 포지션 진입을 선택하더라도 실제 진입을 차단하여 무분별한 거래를 억제하였다. 최종 보상함수는 식 (1)로 정의된다.

$$R = R_{pnL} - R_{fee} + \alpha(P_{win} - P_0) \quad (1)$$

R_{pnL} 은 실현 손익, R_{fee} 는 진입 수수료, P_{win} 은 XGBoost가 예측한 진입 성공률, P_0 는 확률임계값(50%)을 의미한다. XGBoost의 학습 라벨은 포지션 진입 후 일정 기간 내에 이익이 손절보다 먼저 발생하면 Take Profit(TP) 그 외의 경우는 Stop Loss(SL)로 라벨링 하였으며 동일 봉에서 이익과 손절이 동시에 발생하는 경우에는 손절을 우선 처리하였다.

2.3 트레이딩 시뮬레이션

본 연구에서는 나스닥이 장기적으로 우상향하는 특성을 고려하여 Long 포지션만을 사용하여 거래를 진행하였다. PPO와 XGBoost에 의해 매수 신호가 발생하면 다음 봉 시가로 매수를 진행하며, 매도 신호가 발생하면 다음 봉 시가로 매도를 진행한다. 포지션 진입 이후에는 모델이 설정한 이점, 손절 조건 중 하나가 충족될 때까지 포지션을 유지하며, 포지션 청산 시에는 다음 진입 신호가 발생하기 전까지 거래를 진행하지 않는다. 또한, 손절 발생 이후에는 일정 기간동안 포지션 진입을 제한하여 하락 추세인 시장에서 과도한 거래를 억제하였으며, 실제 시장과 유사한 결과를 도출하기 위해 거래 수수료는 1틱으로 부과하였다. 제안 알고리즘의 우수성을 검증하기 위해 비교 모델로 단일 PPO 알고리즘, K-Nearest Neighbor, XGBoost, Decision Tree, Gradient Boosting 총 5개의 모델을 사용하였고, 모델의 성능을 비교하기 위해 수익률, 승률, 최대 손실률(MDD), Payoff ratio(평균수익/평균 손실), Profit factor(총수익/총손실), 거래 횟수를 계산하였다.

III. 트레이딩 시뮬레이션 결과

Table 1은 3년간 평균 투자 성과 지표를 정리한 결과이다. 제안 알고리즘의 수익률은 70.26%로 전체 알고리즘 중 가장 높은 성과를 보였으며, 이는 동일 기간 연평균 성장률(CAGR)인 32.5%보다 37.76% 높은 수익률이다. 또한, 제안 기법의 MDD는 -24.04%로 단일 PPO 알고리즘 대비 26.57% 감소한 모습을 보였다. 이를 통해 제안 알고리즘의 트레이딩이 기존 모델보다 우수함을 알 수 있으며 이는 Fig 3에서 확인할 수 있다. 머신러닝 알고리즘에서는 Gradient Boosting이 66.29%의 수익률로 우수한 성과를 보였으나, MDD가 -42.16%로 안정성 측면에서 한계를 보여주었다.

알고리즘	수익률	승률	MDD	Payoff Ratio	Profit factor	거래 횟수
Proposed	70.26	60.0	-24.04	1.009	1.514	48
PPO	54.47	57.9	-50.61	0.997	1.459	39
KNN	53.95	64.2	-24.67	1.056	2.020	19.3
XGBoost	58.96	61.5	-41.96	1.152	1.858	24
Decision Tree	23.13	51.0	-52.77	1.022	1.317	30.7
Gradient Boosting	66.29	59.1	-42.16	1.136	1.744	29.7

Table 1. 트레이딩 시뮬레이션 결과

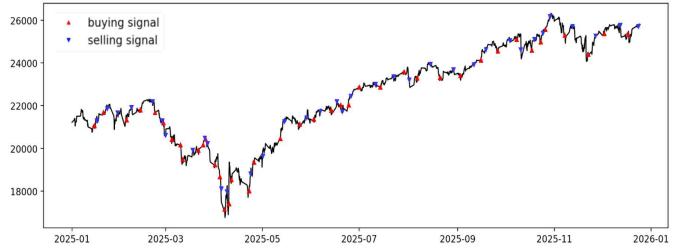


Fig 3. 제안 알고리즘의 트레이딩 시그널

IV. 결론

본 논문에서는 머신러닝 기반 시스템 트레이딩이 실매매에서 레깅 등의 문제로 인한 투자 성능 저하가 발생할 수 있으며, 강화학습 과정에서 모델의 무분별한 포지션 진입으로 인해 불필요한 손실이 누적되는 문제를 완화하기 위해 PPO 알고리즘과 XGBoost를 결합한 포지션 진입 제어 알고리즘을 제안한다. 제안 기법은 XGBoost가 추정한 포지션 진입 성공률을 보상함수에 반영하여 PPO의 학습을 보조하고, 확률 임계값을 통해 무분별한 포지션 진입을 억제함으로써 불필요한 거래로 인한 누적 손실을 완화도록 설계하였다. 트레이딩 시뮬레이션 결과 제안 기법은 3년간 평균 수익률 70.26%를 기록하였으며, MDD -24.04%를 기록하여 단일 PPO 대비 26.57% 감소한 모습을 보였다. 특히, 제안 알고리즘의 연평균 거래 횟수가 타 알고리즘 대비 증가했음에도 불구하고 수익률과 MDD가 더욱 우수한 성능을 보임을 확인하였으며, 이는 제안 알고리즘이 단일 PPO 모델과 머신러닝 모델들에 비해 무분별한 포지션 진입 문제를 보완하고 수익성과 안정성을 동시에 개선할 수 있음을 시사한다.

ACKNOWLEDGMENT

본 연구는 2025년도 산업통상자원부 및 산업기술기획평가원(KEIT) 연구비 지원과(과제번호: RS-2024-00432506) 2025년 과학기술정통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과(2022-0-01068) 및 2025년도 교육부 및 세종특별자치시의 재원으로 세종RISE센터의 지원을 받아 수행된 지역혁신중심 대학지원체계(RISE)의 결과입니다(2025-RIS E-08-004).

참고 문헌

- [1] 주성준. (2024). LSTM 기반의 적대적 생성 네트워크(GAN)를 이용한 코스피200 지수 방향 예측 [석사학위논문, 연세대학교].
- [2] 한예찬, 김재윤. (2021). N-기간 변동성 데이터 라벨링을 활용한 XGBoost 트레이딩 시스템 개발. 한국데이터정보과학회지, 32(5), 1049-1070. 10.7465/jkdi.2021.32.5.1049
- [3] 박명석, 김재윤. (2022-06-22). 강화학습을 이용한 주식 트레이딩 시스템 개발. 한국통신학회 학술대회논문집, 제주.
- [4] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," arXiv preprint arXiv:1707.06347, 2017.
- [5] T. Théate and D. Ernst, "An application of deep reinforcement learning to algorithmic trading," Expert Systems with Applications, vol. 173, p. 114663, 2021.