

Time-GAN을 활용한 트레이딩 알고리즘 선택

이재윤, 이주홍*, 최범기**, 송재원***

인하대학교, *인하대학교, **(주)큐헷지, ***주)밸류파인더스

chris0258@naver.com, *juhong@inha.ac.kr, **bgchoi666@gmail.com, ***jwsong@valuefinders.co.kr

Trading Algorithm Selection Using Time-GAN

Jae Yoon Lee, Ju-Hong* Lee, Bum Gi Choi**, Jae Won Song***

Inha Univ., *Inha Univ., **QHedge Co. Ltd., ***ValueFinders Co. Ltd.

요약

본 논문은 주식시장에서 Time-GAN 모델을 이용하여 여러 트레이딩 알고리즘들의 각각의 미래의 수익률 시계열을 생성한 후, 생성된 수익률 시계열의 분산과 평균을 이용하여 각 모델의 미래 수익률 예측 정확도와 수익률의 높고 낮음을 판단하고, Time-GAN에 의한 모델 예측도가 높고 수익률이 좋은 트레이딩 알고리즘을 선택한다. 선택된 트레이딩 알고리즘의 매수, 매도 시그널을 이용하여 안정적으로 최대의 수익률을 내는 투자 모델을 제안하였다.

1. 서론

현대 사회에서 금융시장을 분석하는 것은 그 사회의 경제를 분석하는 것 만큼이나 중요하기에 금융시장을 분석하는 것은 없이 논의할 수 없을 만큼 중요한 역할을 한다. 금융시장을 분석하는 방법은 크게 기본적 분석과 기술적 분석이 있다. 기본적 분석은 가치를 분석하는 방법이고, 기술적 분석은 데이터를 이용해 분석을 하는 방법이다. 기술적 분석을 이용하면 매수 또는 매도 시그널을 이용해 거래를 하는 방법이 있다.[1] 본 논문에서는 기술적 분석을 이용하는 트레이딩 알고리즘들을 이용하여 적은 변동성으로 수익을 내는 방법을 제안한다.

최근에 과거 데이터의 기술적 분석 방법을 기계학습, 신경망, 유전 알고리즘들을 이용하여 고도화한 트레이딩 알고리즘에 관한 연구들이 이루어졌다[2, 3, 4]. 다만 시간에 따라 통계적 특성이 변하는 비정상성 때문에 기술적 분석만으로는 완벽히 분석되지 않는다는 점이 있다. 즉 특정 구간에서는 잘 작동될 수 있지만, 모든 구간에서는 원하는 결과가 나오기 어렵다는 단점이 있다.[5] 이러한 문제점을 해결하기 위하여 Time-GAN (Time series Generative Adversarial Networks) 을 이용하여 각 트레이

딩 알고리즘들이 어떤 성능을 보일지 예측하고 그 결과를 평가하여 여러 트레이딩 알고리즘 중에서 어떤 것이 해당 시점에서 가장 적합할지를 판단하여 투자에 사용할 트레이딩 알고리즘을 선택하는 투자 모델을 제안한다.[6] 이 모델은 2015년부터 2020년까지 5년간 삼성전자 데이터에 대하여 실험하여 좋은 성능을 보여주었다.

이후 사용되는 표기법은 다음과 같다.

- $\bar{x}_t$: t 시점에서의 시가, 고가, 저가, 종가, 거래량을 가진 벡터
- $(X_t)_k$: t-1 에서 t 시점까지 k 번째 알고리즘을 사용했을때의 수익률
- T : 전체 데이터 크기
- S : 주기 고정 길이
- K : 트레이딩 알고리즘의 총 갯수
- $(X_t)_k'$: k번째 Time-GAN 으로 생성한 t 시점 데이터

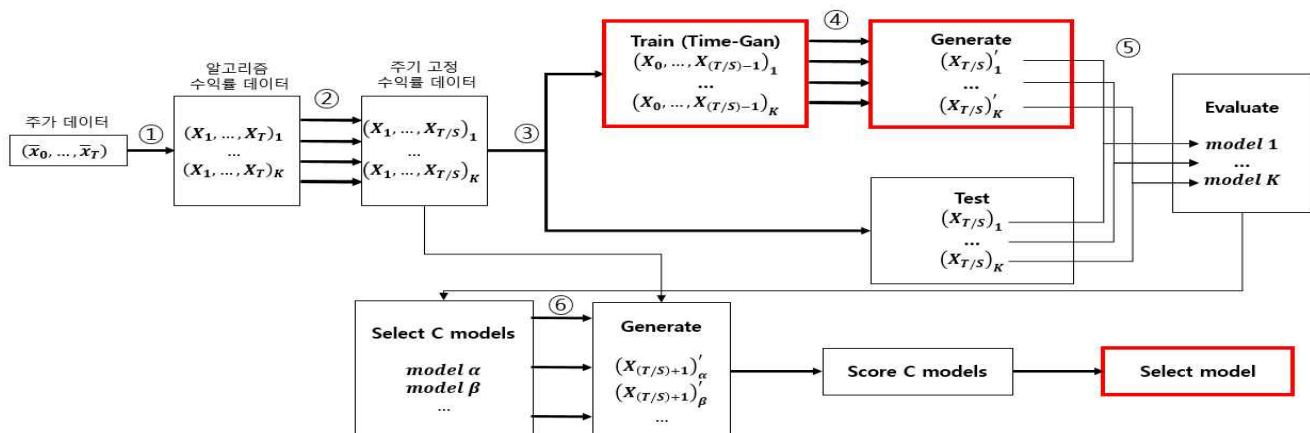


그림 1. 제안된 투자 모델 개요도

II. 본론

본 논문에서 제안하는 투자 모델은 그림 1 과 같은 구성을 가진다. 전처리, 학습, 선택 3 단계로 분류되어 있다. 전처리 단계는 그림 1에서 ①과 ②로서 Time-GAN의 훈련 시계열 데이터를 준비하는 단계이다. ①은 주가 시계열 데이터를 K개의 트레이딩 알고리즘으로 각각 거래를 했을 때의 수익률 시계열들로 변환한다. ②는 각 수익률 데이터를 압축해주는 과정이다. 예를 들면 압축주기가 5일이면 일별 데이터를 주별 데이터로 변환한다. 압축을 하면 학습할 데이터를 줄임으로써 학습시간을 줄일 수 있고, 압축하지 않았을 때 트레이딩 알고리즘들이 거래를 하지 않는 구간이 많이 발생하여 학습의 과적합이 되는 것을 방지하기 위함이다. 압축하게 되면 선택된 트레이딩 알고리즘이 교체되지 않고 압축한 주기 S 동안 동일한 트레이딩 알고리즘을 사용하게 된다. ③를 수행하면, K개의 트레이딩 알고리즘들의 수익률 시계열 (T/S)의 압축된 수익률 시계열로 변환된다. 그림1에서는 훈련하기 위한 하나의 데이터만 표시하였으나 실제로는 과거의 여러 구간에 대해서 훈련 데이터를 생성한다. 학습 단계는 ③, ④, ⑤로 이루어져 있다. ③은 모든 K개의 압축된 수익률 시계열에서 마지막 수익률 데이터인 $X_{(T/S)}$ 를 제외하고 $X_{(T/S)}$ 이전의 시계열을 이용하여 각 트레이딩 알고리즘별 수익률 시계열에 대해서 각각 다른 Time-GAN을 이용해 학습시킨다. ④는 학습된 Time-GAN을 이용해 $(X_{(T/S)})'$ 을 생성한다. ⑤는 생성된 결과의 평균과 분산을 이용하여 실제 값을 비교하여 Time-GAN 예측 성능을 평가한다. ⑥은 선택단계로서 Time-GAN의 예측 성능이 좋은 상위 C개의 알고리즘을 선별한다. 그러면 선별된 알고리즘에 대한 Time-GAN의 예측 결과를 신뢰할 수 있다. 이후에 선별된 알고리즘에 대해서 다음 실제 거래구간에서 각 Time-GAN이 대응되는 알고리즘의 수익률을 예측하게 한다. 알고리즘들의 수익률 예측 결과 중에서 변동성이 작고 수익률이 높은 알고리즘을 최종적으로 선택한다. 만약 i번째 트레이딩 알고리즘이 선택되었다면 알고리즘은 다음 실제 거래 기간인 S일 동안은 i번째 트레이딩 알고리즘이 보내는 매수, 매도 시그널을 따라 투자를 하는 것이다.

본 논문에서는 제안한 투자 모델을 성능을 평가하기 위해 2015년 11월 11일부터 2020년 9월 29일까지 영업일 기준 1200일 삼성전자 주가에 대하여 sliding window 방식을 이용해 실험을 진행하였다. 초기 훈련기간은 2000년 1월 4일부터 2015년 11월 10일부터 영업일 기준 3100일의 데이터를 사용하였다. 결과는 그림 2 와 같다. 1200일 동안의 트레이딩 알고리즘들의 수익률과 주가의 수익률을 나타낸 그래프로 빨간색 선은 주가의 증가이고 검은색은 투자 모델을 사용했을때의 결과, 나머지는 14개의 트레이딩 알고리즘들 이다.



그림 2. 삼성전자 2015년-2020년

	샤프 지수	최종 수익률	변동성
주가	0.73	117%	0.0169
투자 모델	1.13	116%	0.0096

표 1. 결과

표 1에서 보는 바와 같이 최종 수익률은 주가가 117% 상승 하는 동안 투자 모델도 116%의 수익을 내었다. 투자 모델의 수익률은 비슷하지만 투자 모델의 변동성은 주가보다 약 2배 정도 작기 때문에 샤프 지수가 높게 나와서 안정적인 투자를 할 수 있는 모델임을 확인할 수 있다.

III. 결론

본 논문은 트레이딩 알고리즘이 항상 좋은 결과를 내기는 힘들다는 한계점을 가진다는 단점을 보완하기 위해 여러 트레이딩 알고리즘을 사용해서 구간별로 적은 변동성으로 좋은 수익률이 예상되는 트레이딩 알고리즘을 찾는 투자 모델을 제안하였다. 제안 방법은 대상 종목의 주가와 비교해 더 높은 샤프지수를 보여주며, 제안된 투자 모델의 성과가 좋음을 보여주었다.

ACKNOWLEDGMENT

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (2019R1F1A1062094, 2017R1D1A1A02018319, NRF-2020R1F1A1069361)

참 고 문 헌

- [1] Nuti, G., Mirghaemi, M., Treleven, P., & Yingsaeree, C. "Algorithmic trading." Computer, 44(11), 61-69. 2011
- [2] Shao, X., Ma, D., Liu, Y., & Yin, Q. "Short-term forecast of stock price of multi-branch LSTM based on K-means." In 2017 4th International Conference on Systems and Informatics (ICSAI) (pp. 1546-1551). IEEE. 2017
- [3] Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. "Stock price prediction using LSTM, RNN and CNN-sliding window model." In 2017 international conference on advances in computing, communications and informatics (icacci) (pp. 1643-1647). IEEE. 2017
- [4] Chung, H., & Shin, K. S. "Genetic algorithm-optimized long short-term memory network for stock market prediction." Sustainability, 10(10), 3765. 2018
- [5] Mills, T. C. "Technical analysis and the London Stock Exchange: Testing trading rules using the FT30." International Journal of Finance & Economics, 2(4), 319-331. 1997
- [6] Yoon, J., Jarrett, D., & Van der Schaar, M. "Time-series generative adversarial networks." NeurIPS, 2019