

TimeGAN을 활용한 트레이딩 알고리즘 선택

(Trading Algorithm Selection Using Time-Series Generative Adversarial Networks)

이재윤*, 이주홍**, 최범기**, 송재원***

(Jae Yoon Lee, Ju Hong Lee, Bum Ghi Choi, Jae Won Song)

요약

주식 시장에서 안정적으로 높은 수익을 얻기 위하여 많은 트레이딩 알고리즘에 대한 연구들이 이루어졌다. 트레이딩 알고리즘들이 미국 주식시장의 거래량에서 차지하는 비율은 80 퍼센트가 넘는 정도로 많이 사용된다. 많은 연구에도 불구하고 항상 좋은 성능을 나타내는 트레이딩 알고리즘은 존재하지 않는다. 즉, 과거에 좋은 성능을 보이는 알고리즘이 미래에도 좋은 성능을 보인다는 보장이 없다. 그 이유는 주가에 영향을 주는 요인은 매우 많고, 미래의 불확실성도 존재하기 때문이다. 따라서 본 논문에서는 알고리즘들의 수익률에 대한 과거 기록을 바탕으로 미래의 수익률을 잘 예측하고 수익률도 높을 것으로 추정되는 알고리즘을 선택하는 TimeGAN을 활용한 모델을 제안한다. LSTM기법은 미래 시계열 데이터의 예측이 결정론적임에 반하여 TimeGAN은 확률적이다. TimeGAN의 확률적인 예측의 이점은 미래에 대한 불확실성을 반영하여 줄 수 있다는 점이다. 실험 결과로써, 본 논문에서 제안한 방법은 적은 변동성으로 높은 수익률을 달성하고, 여러 다수의 비교 알고리즘에 비해 우수한 결과를 보인다.

■ 중심어 : 트레이딩 알고리즘 ; TimeGAN ; 주식 시장 ; 알고리즘 선택 ; 금융

Abstract

A lot of research is being going until this day in order to obtain stable profit in the stock market. Trading algorithms are widely used, accounting for over 80% of the trading volume of the US stock market. Despite a lot of research, there is no trading algorithm that always shows good performance. In other words, there is no guarantee that an algorithm that performed well in the past will perform well in the future. The reason is that there are many factors that affect the stock price and there are uncertainties about the future. Therefore, in this paper, we propose a model using TimeGAN that predicts future returns well and selects algorithms that are expected to have high returns based on past records of the returns of algorithms. We use TimeGAN because it is probabilistic, whereas LSTM method predicts future time series data is deterministic. The advantage of TimeGAN probabilistic prediction is that it can reflect uncertainty about the future. As an experimental result, the method proposed in this paper achieves a high return with little volatility and shows superior results compared to many comparison algorithms.

■ keywords : Trading Algorithm ; TimeGAN ; Stock Market ; Algorithm Selection ; Finance

I. 서론

주식 시장은 현대 경제에서 필수적인 금융 시장의 중심 역할을 한다. 주식 시장은 많은 자본이 투입되어 살아있는 생물체처럼 동적으로 움직이기 때문에, 종종 투자자들을 놀라게 할 정도로 변동성이 매우 심하다. 주식 시장은 변동성이 심한 만큼 투자자가 수익을 낼 수도 있지만, 손실을 얻을 수도 있다. 따라서

이러한 시장의 특성을 반영하여 좀 더 과학적인 방법으로 투자하는 트레이딩 알고리즘들이 개발되었다[8-12]. 트레이딩 알고리즘이란 컴퓨터 프로그램이 과거 데이터를 분석하고 매수 또는 매도 시그널을 발생시키고 거래하는 방법이다. 트레이딩 알고리즘을 이용한 거래를 Algorithmic Trading 이라고도 하는데, 이는 미국 금융시장의 전체 거래량에서 80퍼센트가 넘는 비율을 차지하고 있을 만큼 현재 미국 금융시장에서 많은 부분을

* 학생회원, 인하대학교 전기컴퓨터공학과

** 정회원, 인하대학교 컴퓨터공학과 교수

*** 정회원, 밸류파인더스

이 논문은 2021년도 한국연구재단의 지원을 받아 수행된 기초연구사업임 (2019R1F1A1062094, 2021R1F1A1050120, NRF-2020R1F1A1069361)

접수일자 : 2021년 12월 6일

게재확정일 : 2022년 2월 8일

수정일자 : 1차 2022년 1월 10일, 2차 2022년 2월 7일

교신저자 : 이주홍 e-mail : juhong@inha.ac.kr

차지하고 있다[14]. 트레이딩 알고리즘에 대한 많은 연구에도 불구하고, 현재 모든 기간에 걸쳐 항상 좋은 성능만을 보이는 트레이딩 알고리즘은 존재하지 않는다[1]. 이는 아무리 복잡하고 정교한 알고리즘이라 하더라도 주가에 영향을 미치는 요인들은 너무 많고, 시장에는 항상 불확실성이 존재하기 때문에, 트레이딩 알고리즘이 시장의 모든 요인들과 불확실성에 모두 다 대처하는 것은 불가능하기 때문이다.

트레이딩 알고리즘들의 특징을 보면 크게 3가지가 있다.

첫 번째, 금융 관련 과거 데이터는 시간에 따라 통계적 특성이 변하는 비정상성(non-stationary)이라는 특징을 가진다. 트레이딩 알고리즘이 사용하는 데이터의 비정상성 특징 때문에 트레이딩 알고리즘이 항상 좋은 성능을 보이는 것이 매우 어렵다. 즉 트레이딩 알고리즘이 어떤 구간에서는 좋은 성능을 보이다가도 다른 구간에서는 나쁜 성능을 보일 수 있다.

두 번째는 트레이딩 알고리즘에 의한 수익의 변동성이 실제 주가의 변동성보다 작은 경향이 있다.

세 번째는 트레이딩 알고리즘은 대체로 거래횟수가 많다. 매수 시에는 거래 수수료, 매도 시에는 거래 수수료와 거래 세금이 있다. 이러한 비용들은 작은 것처럼 보일 수 있으나 거래가 많아지면 누적되어 무시할 수 없을 만큼 그 값이 커진다.

수익을 내기 위해서는 시장을 분석해야 하는데 시장을 분석하는 방법은 크게 두 가지로 나뉜다. 기업의 내재적 가치를 분석하는 기본적 분석과 과거의 데이터를 활용해 분석하는 기술적 분석으로 나뉜다[1, 2]. 기본적 분석은 재무제표, 경영, 경쟁우위성 등을 분석하여 기업의 가치를 판단하고 이 판단을 토대로 투자를 할 때 사용된다. 기본적 분석은 시장에 대한 정보를 분석할 능력이 없는 사람에게는 어렵다. 하지만 기술적 분석은 외부 정보 없이 과거의 데이터만을 이용하여 분석이 가능하기 때문에 복잡한 시장 분석이 없이도 가능하다는 장점이 있다.

본 논문에서는 어떤 완벽한 하나의 알고리즘을 개발하는 것보다는, 이미 개발되어 사용되는 많은 알고리즘들 중에서 미래의 짧은 기간 동안 높은 확률로 좋은 성능을 보일 것으로 추정되는 알고리즘을 기간별로 따로 선택하여 미래의 긴 기간 동안 좋은 성능을 보일 확률이 높아지는 것이 가능해지도록 하는 트레이딩 방법으로써 알고리즘을 선택하는 기법을 제안하고자 한다. 선택의 대상이 되는 트레이딩 알고리즘들을 베이스 알고리즘이라고 부르는데, 본 논문에서는 베이스 알고리즘으로 기술적 분석 알고리즘들을 사용하였다.

본 논문에서 제안하는 알고리즘 선택 기법의 목표는 충분히 긴 기간 동안 높은 수익률과 함께 낮은 변동성을 보이는 것이다. 다만 변동성과 수익률은 상충관계(trade-off)를 가진다는 단점이 있다. 그림 1의 초록색으로 나타나는 알고리즘의 결과와 같이 수익률이 높으면 변동성도 그에 따라 높고, 보라색으로 나타나는 알고리즘의 결과와 같이 변동성을 낮게 하려면 수익률

도 낮아지는 것이 일반적이다. 즉 수익률이 높으면서 동시에 변동성도 낮은 알고리즘은 거의 존재하지 않는다. 검은색으로 나타나는 트레이딩 알고리즘의 결과는 수익률은 초록색보다는 낮지만 보라색보다는 높으며, 변동성에 있어서는 초록색보다는 낮고 보라색보다는 높다. 즉 검은색으로 나타나는 알고리즘은 초록색과 보라색의 두 가지 극단적인 알고리즘들보다 변동성과 수익률의 균형을 더 잘 맞추고 있다. 이처럼 변동성을 낮추면서도 수익률도 높은 트레이딩 기법을 개발하는 것이 본 논문의 목표이다.

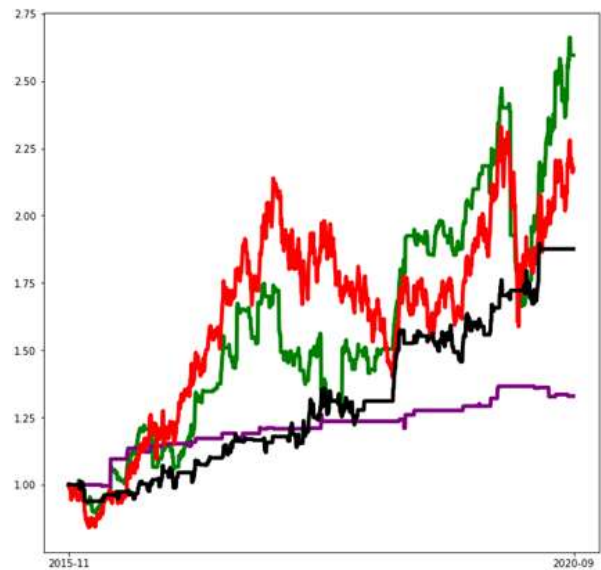


그림 1. 빨간색은 실제 주가, 보라색은 수익률은 높으나 변동성도 높은 알고리즘의 결과, 초록색은 수익률이 낮으며 변동성도 낮은 알고리즘의 결과, 검은색은 제안한 알고리즘의 결과

수익률과 변동성의 상충관계를 가지고 트레이딩 알고리즘의 성능을 측정하는 평가지표로서 샤프지수(sharpe ratio)를 사용한다. (1)은 샤프지수의 식을 나타내며, 매 기간 수익률의 평균을 수익률로서 간주하고 수익률의 표준편차를 변동성으로 간주하여, 변동성 대비 수익률의 비율을 보여주는 지표이다.

$$sharpe \ ratio = \frac{mean(r)}{stddev(r)} \tag{1}$$

본 논문의 구성은 2장에서는 관련 연구를 소개하고, 3장에서는 연구 동기와 기존 한계를 극복할 아이디어를 제시하고 4장에서는 그 아이디어를 사용할 수 있는 모델과 TimeGAN을 사용하는 이유에 대해 설명한다. 5장에서는 제안한 투자 방법의 결과를 평가할 평가지표와 제안 투자모델의 평가 결과를 보여준다. 6장에서는 실험결과와 향후 과제에 대하여 서술한다.

II. 관련 연구

시장예측의 결과를 트레이딩 알고리즘에 적용하는 방법이 연구 되었다[7]. 시장예측 결과를 도출하기 위한 방법으로 K-means와 LSTM을 사용한 연구[4], CNN Sliding Window를 사용한 연구[5], 유전 알고리즘의 최적화 기법을 사용하는 연구[6] 등이 있다.

과거 데이터의 통계적 분석과 기술적 지표를 사용하는 여러 종류의 트레이딩 알고리즘들이 있다. 사용되는 기술적 지표로는 이동 평균(Moving Average), MACD, 최고가와 최저가, Stochastic, Bollinger Bands 등이 있다[2].

머신러닝 모델을 사용하는 트레이딩 알고리즘들도 많이 연구되고 있다[8, 9, 10, 11, 12]. 신경망 기반의 트레이딩 알고리즘[8], Hybrid Neural Networks를 이용하는 트레이딩 알고리즘[9], CNN을 이용하여 패턴을 감지해 시그널을 발생시키는 트레이딩 알고리즘[10], 강화학습을 이용하는 트레이딩 알고리즘 등이 있다[11]. 주가 정보 외에 경제 및 시장 정보 등 환경 변수를 입력으로 받아들이고 강화학습으로 트레이딩을 학습하는 알고리즘에 대한 연구가 있다[12]. 이처럼 최근까지도 여러 가지 다양한 방법을 이용한 트레이딩 알고리즘들이 연구되었지만, 어떤 트레이딩 알고리즘이든 모든 구간에서 좋은 성능을 내는 트레이딩 알고리즘은 없다[1].

III. 제안 아이디어

1. 연구 동기

트레이딩의 목적은 높은 수익률을 얻는 것이다. 그러나 하나의 알고리즘에 의한 트레이딩은 그 특징들 때문에 높은 수익률을 얻기 위해서는 높은 변동성을 피할 수 없다. 그러나 높은 변동성은 투자자에게 심리적인 타격을 심하게 가하기 때문에, 투자자들은 낮은 변동성을 선호한다. 낮은 변동성을 얻기 위해서는 수익률이 낮아질 수밖에 없다. 즉, 변동성과 수익률은 상충관계를 가진다.

여러 트레이딩 알고리즘의 성능을 그림 2와 같이 한 개의 그래프에서 종합하여 관찰하면, 전체 구간을 작은 구간으로 나누었을 때 각 구간 별로 제일 높은 성능을 나타내는 알고리즘은 항상 다른 것을 확인할 수 있다. 즉, 그림 2에서 1,200일 동안의 주가 (삼성전자, 2015년 11월~2020년 9월, 빨간색)의 변화에 대해서, 서로 다른 4개의 트레이딩 알고리즘의 결과를 보면, 각 구간별로 성능이 좋은 알고리즘이 다르다.

따라서 본 논문에서는 단일 트레이딩에 의한 트레이딩이 아닌 여러 알고리즘들을 베이스 알고리즘으로 두고, 전체 구간을 작은 구간들로 나누고 각 구간별로 성능이 좋은, 즉 높은 신뢰도로 높은 수익률이 예측되는, 알고리즘을 선택해가는 것이다.

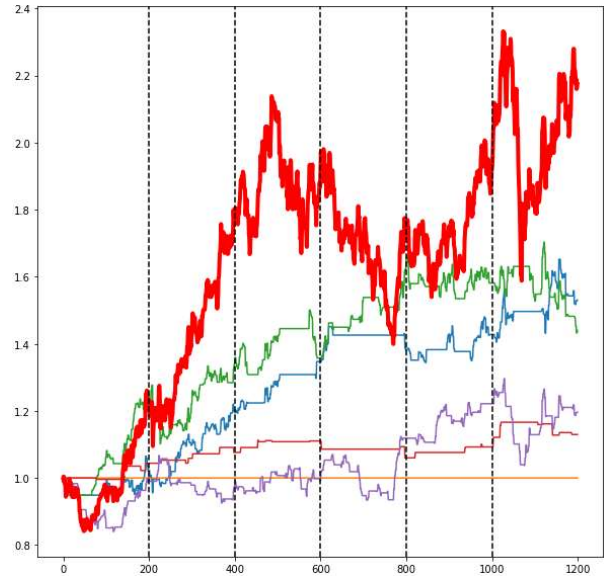


그림 2. 주가와 4개의 트레이딩 알고리즘 구간별로 제일 좋은 성능 (초록/파랑/파랑/초록/보라/파랑)

2. 제안 방법

본 논문에서 제안하는 방법은 다음과 같다. 적은 변동성으로 높은 수익률을 만족하기 위해 다수의 트레이딩 알고리즘들을 사용하고, 각각의 트레이딩 알고리즘의 수익률을 미래의 불확실성과 함께 예측하여, 예측의 정확도가 높으면서, 높은 수익률을 보일 것으로 예측되는 알고리즘을 선택하여 거래한다. 이때 거래량의 빈도를 조절하여 거래비용을 줄이는 기법을 적용한다.

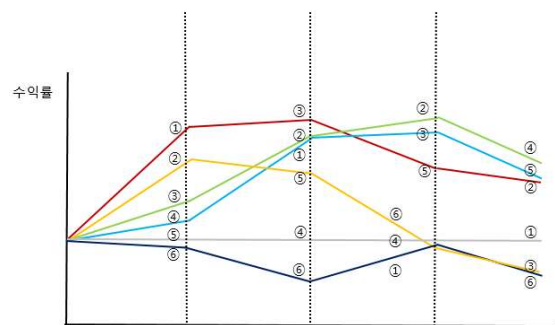


그림 3. 구간 별 알고리즘 수익률 순위

가. 다수의 트레이딩 알고리즘 사용

트레이딩 알고리즘들은 각각 서로 다른 특정 구간에서 좋은 성능을 나타낸다. 이를 활용하여 각각의 구간에서 좋은 성능을 나타내는 알고리즘을 선택할 수 있다면 전체 구간에서 전체 성능을 증가시킬 수 있을 것이다. 그림 3처럼 여러 트레이딩 알고

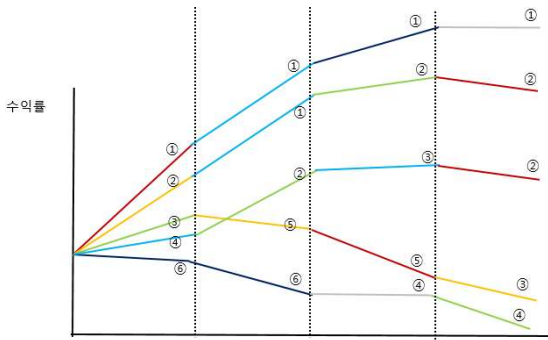


그림 4. 구간 별 다른 알고리즘 사용했을 때

리즘이 존재할 때, 구간별로 수익률이 제일 높은 트레이딩 알고리즘을 사용하면 그림 4의 제일 위처럼 제일 좋은 결과가 나오게 된다.

해당 아이디어를 실제 주가에 적용한 것이 그림 5이다.

그림 2는 200일을 기본 구간 단위로 하여 전체 구간을 6개 구간으로 나누는데 반해, 그림 5는 20일을 기본 구간 단위로 전체 구간을 60개 구간으로 나눈다. 그림 5에서 검은색은 매 구간마다 4개의 알고리즘 중 제일 수익률이 좋은 알고리즘을 한 개를 선택하여 거래를 진행하여 나타난 결과로써 높은 수익률과 낮은 변동성을 달성하고 있다. 반면 구간마다 무작위로 알고리즘을 고를 경우에는 파란색의 결과처럼 낮은 수익률이 나타난다.

나. 거래 제한 및 현금 보유

트레이딩 알고리즘의 문제 중 하나는 많은 거래로 인해 발생하는 큰 거래비용이다. 이 단점을 극복하기 위해서 실제로 발생하는 매수, 매도 시그널 중 일부를 무시하는 방법을 사용한다. 일정한 비율로 시그널을 무시하면 거래의 횟수를 줄일 수 있고 발생하는 수수료와 세금으로 인한 손실을 줄일 수 있다.

또한 다수의(예 90% 이상) 트레이딩 알고리즘의 예측 결과를 통하여 거래를 하지 않는 것이 거래를 하는 것보다 더 좋을 것이라고 판단되면 거래를 하지 않고 현금으로 보유한다. 현금보유는 평가액의 변동이 없으므로 변동성이 줄어드는 효과가 나타난다.

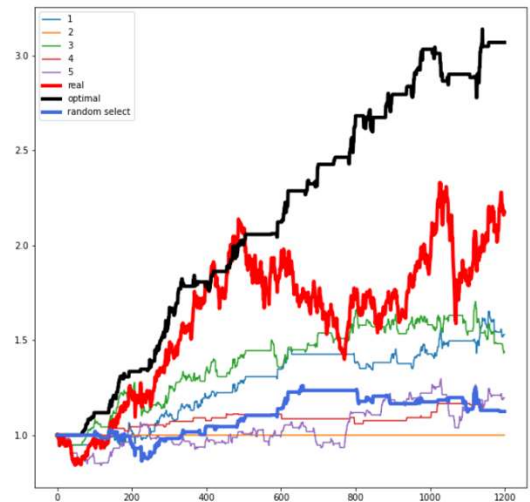


그림 5. 검은색 : 최적의 선택 시
파란색 : 무작위 선택 시

IV. 모델 구성

기존에 존재하는 트레이딩 알고리즘들을 사용한다. 이들의 과거 수익률 데이터로 각각 대응되는 TimeGAN (Time-Series Generative Adversarial Network)[3]을 훈련 시킨다. TimeGAN의 수익률 예측 결과의 정확도(신뢰도)가 높으면서 동시에 예측 수익률이 높은 대응되는 트레이딩 알고리즘을 선택하여 금융 상품을 거래한다.

모델의 전체 구성도는 그림 6 과 같다. 주가 데이터를 각 알고리즘들에 적용하여 각 알고리즘들의 수익률 데이터를 얻는다. 각 알고리즘별로 각각 하나씩 준비된 TimeGAN을 그림 8과 그림9에서 보여준 바와 같이 알고리즘의 수익률 데이터를 학습하여 과거의 각 구간별로 다음 일부 기간을 예측한다. 이 때 각 알고리즘에 대응된 TimeGAN의 수익률 예측의 정확도(신뢰도)를 측정하여 수익률 예측의 정확도가 높은 알고리즘을 선택한다(필터1). 필터1로 선택된 알고리즘들 중에서 그림 9처럼 각 TimeGAN들이 예측한 수익률 중에서 높은 수익률에 대응되는 알고리즘을 선택한다(필터2). 필터 1을 통해 선택된 알고리즘들의 TimeGAN들은 잘 훈련되어 신뢰할 수 있다고 간주되어 그 TimeGAN들이 예측한 수익률은 신뢰도가 높다고 판단한다. 필터 2는 필터1에서 선택된 TimeGAN들 중 예측 수익률이 가장 높은 것의 대응되는 알고리즘을 선택한다.

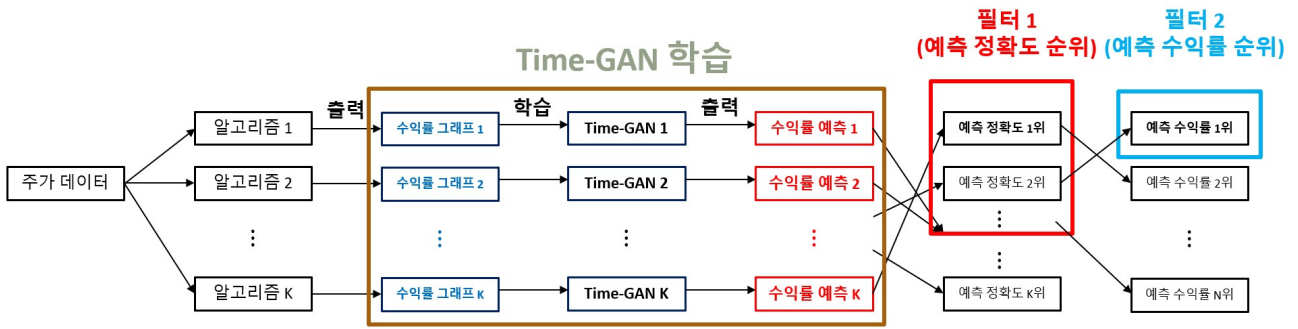


그림 6. 전체 모델 구조

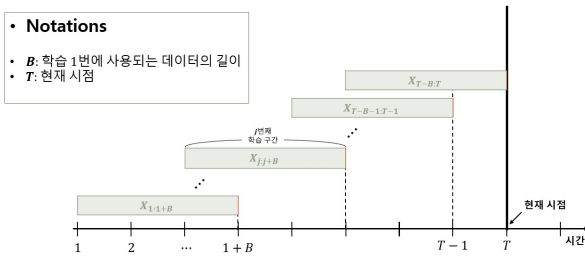


그림 7. TimeGAN의 블록 단위 학습

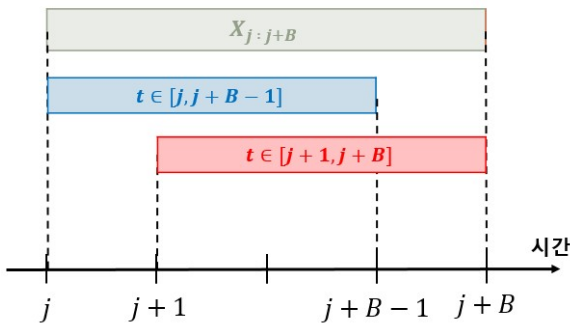


그림 8. TimeGAN의 generator

1. TimeGAN을 이용한 학습 (필터1)

TimeGAN은 시계열 생성에 사용되는 모델이며 그림 7과 같이 블록 단위로 학습을 진행한다. 그림 8은 j 번째 블록이 학습을 하는 과정이다. 한 블록의 크기가 B+1이라고 할때 j 번째 블록은 j~(j+B)의 데이터를 가지고 학습을 한다. 이중 생성자(generator)는 j~(j+B-1)의 데이터를 이용하여 (j+1)~(j+B)의 데이터를 생성한다. 생성된 데이터와 가지고 있는 실제 데이터를 비교하여 supervised loss를 계산한다. 필터 1은 supervised loss를 이용하여 정확도(신뢰도)를 계산한다. supervised loss의 값이 낮을수록 해당 모델이 정확도(신뢰도)가 높다고 간주 되며, 여러 개의 TimeGAN 중에서 정확도가 높은 TimeGAN과 그에 대응되는 알고리즘들이 선택된다.

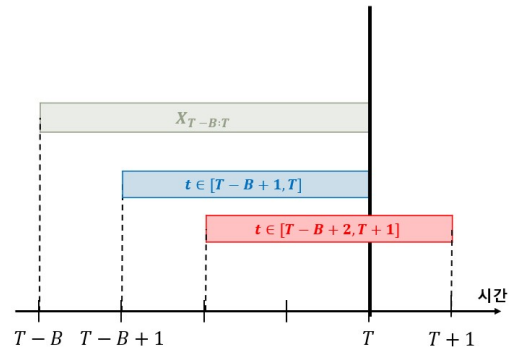


그림 9. TimeGAN을 이용한 예측

2. TimeGAN을 이용한 예측 (필터2)

블록 단위로 학습되는 TimeGAN에서 그림 9와 같이 마지막 블록은 (T-B)~T 데이터를 이용하여 학습을 완료한다. 학습을 완료한 후 (T-B+1)~T의 데이터를 이용하여 생성자(generator)를 이용하여 (T-B+2)~(T+1)의 데이터를 생성한다. T+1의 데이터는 현재 시점 이후의 데이터이므로 예측된 미래 데이터이다. 예측값이 가장 높은 TimeGAN의 대응되는 알고리즘이 최종 선택되는 알고리즘이다.

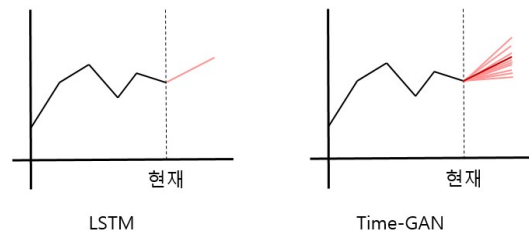


그림 10. 왼쪽 : LSTM , 오른쪽 : TimeGAN

3. TimeGAN을 사용하는 이유

과거 데이터는 불확실성이 제거된 데이터이다. 과거의 어느 시점에는 불확실성이 존재하였지만 이후 시간이 지나면서 1개의 값으로 확정되었기 때문에 과거 당시에 존재했던 불확실성이 사라지는 것이다. 반면에 아직 일어나지 않은 미래는 불확실성이 존재한다. 따라서 과거의 데이터에서 과거 순간에 예측값이 의미를 가지기 위해서는 과거의 불확실성이 반영되어야 한다. 시계열 예측에서 가장 많이 사용되는 모델 중 하나는 LSTM이다. 그런데 LSTM은 그림 10과 같이 단 하나의 값만으로 미래 예측값을 생성한다. 즉 LSTM은 미래에 대한 불확실성을 나타낼 수 없다. 그러나 TimeGAN은 생성자(generator)가 입력된 랜덤 노이즈에 따라서 매우 다양한 값을 생성하므로 훈련하는 과정에서 불확실성을 자동적으로 반영할 수 있으며 실제 예측 시에 노이즈 값의 변이에 따라서 다양한 값을 생성하므로 미래의 불확실성을 생성할 수 있다.

V. 실험

1. 실험 설계

실험에서는 실험 구간 동안 제안한 방법과 트레이딩 알고리즘을 사용하였을 때의 수익률과 변동성이 얼마나 제안한 방법과 차이가 나는지를 비교한다. 수익률과 변동성을 한 개의 수치로 표현하기 위해 샤프지수를 이용하여 결과를 측정한다.

2. 실험 환경

실험은 Intel Xeon Processor 와 VT100 2대의 GPU를 장착한 서버1과, Intel(R) Xeon(R) CPU E5-1620과 RTX 3090 1대의 GPU를 장착한 서버2로 진행하였다. 사용 언어는 python 3, 주요 사용 라이브러리는 tensorflow==2.6.0 (신경망을 사용한 트레이딩 알고리즘 학습에 사용) pytorch==1.7.0 (TimeGAN 학습에 사용)을 사용하였다.

3. 실험 데이터

실험 데이터는 POSCO, 기아, 삼성SDI, 삼성전기, 한국조선해양, 삼성전자, 한국전력, SK텔레콤, 롯데케미칼, 엔씨소프트, LG생활건강, 신한지주 코스피 12개 종목 30분봉 데이터의 시가 고가 저가 종가 거래량을 사용하였다. 훈련기간은 2012년 1월~2017년 10월이고 실험 기간은 2017년 10월~2018년 5월이다.

4. 사용된 트레이딩 알고리즘

사용된 트레이딩 알고리즘은 기술적 지표를 사용한 트레이딩 알고리즘 15개, 머신러닝을 이용한 트레이딩 알고리즘 5개 총 20개의 트레이딩 알고리즘을 사용하였다. 사용된 트레이딩 알고리즘은 다음과 같다. 이동 평균선 정배열(10,30,60), 급등주 포착 알고리즘 (650), MACD(12,26), MACD oscillator(1.01, 0.99), 볼린저밴드 (2,12), Fast stochastic K (0.7, 0.3), Fast stochastic D (10,5), Slow stochastic K (5), Slow stochastic D (5), DMI (12), CCI (20), 이동평균선 밀집 (5,20,60), 20일선 돌파, 60일선 돌파, Swing pattern (1), Logistic regression, Random Forest, Gradient Boosting, 3층의 FC를 쌓아서 만든 ANN을 이용한 트레이딩 알고리즘과, 2층의 LSTM과 FC를 쌓아서 만든 LSTM을 이용한 트레이딩 알고리즘이다. 이외에 현금으로 보유하는 것과 주식 상태로 보유하는 방법까지 총 22개의 선택지가 있는 상태에서 실험을 진행하였다.

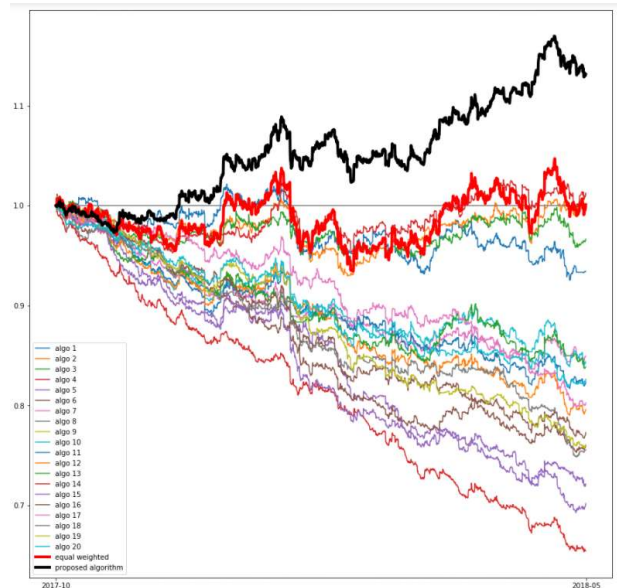


그림 11. 전체 실험 결과

5. 실험 결과

그림 11는 결과 그래프로 빨간색 선은 12개의 종목을 동일 가중치를 주었을 때의 포트폴리오를 나타낸 것이다. 검은색은 제안 방법의 결과이다. 나머지 실선은 트레이딩 알고리즘들의 결과이다. 최종 결과는 표 1로 확인할 수 있다. 표 1은 제안한 방법과 12개 종목에 동일 가중치를 주었을 때의 결과, 20개의 트레이딩 알고리즘의 결과를 수익률, 변동성, 샤프지수로 나타낸 표이다. 제안방법의 샤프지수는 0.48로 다른 트레이딩 알고리즘들의 샤프지수와 12개 종목을 동일 가중치를 주었을 때의 결과보다 높은 성능을 보인다.

표 1. 제안 투자모델, 포트폴리오, 각 알고리즘 결과

	수익률	변동성	샤프지수
제안 방법	0.131424	0.032735	0.488120
equal weighted	0.001002	0.043155	0.024484
Bollinger_Bands	-0.065975	0.030334	-0.266211
RSI	0.01179	0.026717	0.068223
Fast stochastic K	-0.033634	0.026843	-0.145989
Fast stochastic D	-0.345746	0.025784	-2.044690
Slow stochastic K	-0.298267	0.026750	-1.642449
Slow stochastic D	-0.227195	0.024253	-1.316778
MACD	-0.150592	0.027134	-0.738688
n봉 거래량 평균	0.0	0.000000	0
이평선 밀집	0.000895	0.042860	0.024046
DMI	-0.156423	0.027115	-0.770964
n일 돌파	-0.176951	0.027826	-0.861263
cci	-0.204118	0.028099	-1.002002
경매열	-0.161421	0.027664	-0.781986
이격도	0.012456	0.028160	0.069054
Swing pattern	-0.278561	0.027358	-1.478770
DNN	-0.239796	0.025955	-1.308003
LSTM	-0.197494	0.025405	-1.070311
Linear regression	-0.244146	0.023806	-1.458466
random forest	-0.23983	0.023894	-1.423210
gradient boosting	-0.17393	0.023407	-1.009144

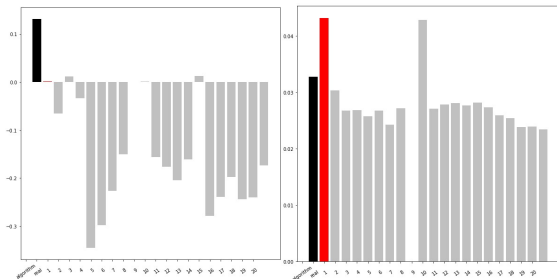


그림 12. 수익률 결과

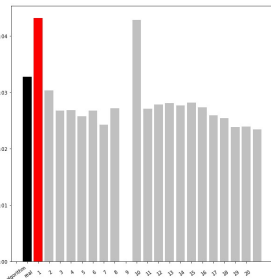


그림 13. 변동성 결과

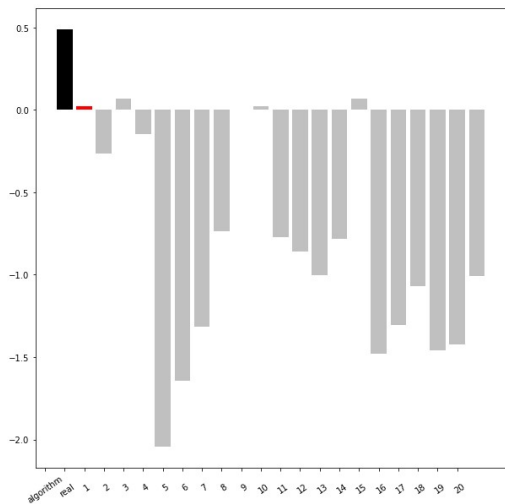


그림 14. 샤프 지수 결과

그림 12, 그림 13, 그림 14 는 각각 수익률, 변동성, 샤프지수를 나타낸 그래프로 검은색은 제안 투자모델의 결과, 빨간색은 12종목이 동일 가중치로 구성된 포트폴리오, 회색은 20개의 트레이딩 알고리즘 각각의 결과이다.

VI. 결 론

본 논문은 적은 변동성으로 높은 수익률을 내는 것을 목표로 하는 트레이딩 시스템을 제안하였다. 짧은 특정 구간에서만 작동하는 트레이딩 알고리즘 특성상 구간을 늘리면 대부분의 트레이딩 알고리즘은 손해를 보는 것에 비해 제안한 트레이딩 시스템은 주가가 떨어지는 구간에서 주가에 비해 비교적 적게 떨어지고 올라가는 부분에서 실제 올라가는 만큼 비슷하게 따라 올라가며 전체적으로 횡보를 하는 구간에서 좋은 수익과 낮은 변동성으로 높은 샤프지수를 달성하였다. 매 시점마다 사용할 트레이딩 알고리즘을 적절히 선택하여 높은 성능을 달성하였다. 또 잦은 거래로 인한 큰 거래비용을 시그널을 무시하는 방법을 통해 극복하였다. 현 제안 방법의 문제점은 다음과 같다. 현재 시점에서 실제로 가장 좋은 성능을 내는 알고리즘이 있더라도, TimeGAN의 훈련이 충분히 되지 않는다면 필터 1에서 걸러지게 되어 제일 좋은 성능을 내는 알고리즘이, 선택되지 않을 수 있다. 향후 과제로서, 모델의 예측 정확도를 평가하는 필터1을 개선하여 좋은 알고리즘이 제외되는 상황을 방지하고, 더 많은 트레이딩 알고리즘을 적용함으로써 알고리즘 선택 폭의 다양성을 늘리고, TimeGAN 학습을 병렬처리함으로써 훈련 시간을 줄이는 과제가 있다.

REFERENCES

- [1] G. Nuti, M. Mirghaemi, P. Treleaven, et al. "Algorithmic trading," *Computer* vol. 44, no. 11, pp. 61-69, Nov. 2011.
- [2] V. Drakopoulou, "A review of fundamental and technical stock analysis techniques," *Journal of Stock & Forex Trading*, vol. 5, no. 1, Nov. 2016.
- [3] J. Yoon, D. Jarrett and M. Van der Schaar "Time-series generative adversarial networks," *NeurIPS*, 2019.
- [4] XL. Shao, D. Ma, Y. Liu, et al. "Short-term forecast of stock price of multi-branch LSTM based on K-means," *2017 4th International Conference on Systems and Informatics (ICSAI)*, IEEE, Nov. 2017.
- [5] S. Selvin and R. Vinayakumar, et al., "Stock price prediction using LSTM, RNN and CNN-sliding window model," *2017 international conference on advances in computing, communications and informatics (icacci)*, IEEE, Sep. 2017.
- [6] H Chung and K Shin "Genetic algorithm-optimized long short-term memory network for stock market prediction," *Sustainability*, vol. 10, no. 10, Oct. 2018.
- [7] BX. Yong, MRA. Rahim, et al. "A stock market

trading system using deep neural network," *Asian Simulation Conference, Springer*, pp. 356-364, Singapore, 2017.

- [8] B. Taylor, M. Kim and A. Choi, "Automated stock trading algorithm using neural networks," *Proc. 2nd International Conference on Intelligent Technologies and Engineering Systems (ICITES2013)*, Springer, pp. 849-857, Cham, 2014.
- [9] W. Haotian, "Trading Decision Making Based on Hybrid Neural Network," *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)*, IEEE, Apr. 2021.
- [10] M. Ugur Gudelek, S. Arda Boluk, et al., "A deep learning based stock trading model with 2-D CNN trend detection," *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, IEEE, Dec. 2017.
- [11] Fiorini, Pierre M. and Pierce-Gabriel Fiorini, "A Simple Reinforcement Learning Algorithm for Stock Trading," *2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, IEEE, pp. 824-830, Sep. 2021.
- [12] C. Ma, J. Zhang, J. Liu, et al., "A parallel multi-module deep reinforcement learning algorithm for stock trading," *Neurocomputing* vol. 449, pp. 290-302, Apr. 2021.
- [13] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical review letters*, vol. 88, no. 17 : 174102, 2002.
- [14] Algo Trading Dominates 80% Of Stock Market (2019)<https://seekingalpha.com/article/4230982-algo-trading-dominates-80-percent-of-stock-market> (accessed Dec, 09, 2021.)
- [15] 천성길, 이주홍, 최범기, 송재원 "대규모 외생 변수 및 Deep Neural Network 기반 금융 시장 예측 및 성능 향상," *스마트미디어저널*, 제9권 26-35 쪽. 2020
- [16] 이윤선, 이주홍, 최범기, 송재원 "비정형, 정형 데이터의 이미지 학습을 활용한 시장예측," *스마트미디어저널*, 제10권, 제2호, 16-21쪽, 2021
- [17] 윤동진, 이주홍, 최범기, 송재원 "부분복제 지수 상향 추종을 위한 진화 알고리즘 기반 3 단계 포트폴리오 선택 앙상블 학습," *스마트미디어저널*, 제10권, 39-47쪽, 2021

저자 소개



이재운(학생회원)

2020년 인하대학교 컴퓨터학과 학사 졸업.

2020년 ~ 현재 인하대 전기컴퓨터공학과 석사 과정.

<주관심분야 : 시계열 데이터, 금융IT>



이주홍(정회원)

1983년 서울대학교 전자계산기공학과 (현 컴퓨터공학과) 학사 졸업.

1985년 서울대학교 전자계산기공학과 (현 컴퓨터공학과) 석사 졸업.

2001년 KAIST 컴퓨터공학전공 박사 졸업.

2001년 ~ 현재 인하대 컴퓨터공학과 교수.

2019년 ~ 현재 ㈜큐헛지 대표이사

<주관심분야 : 머신러닝, 금융투자공학>



최범기(정회원)

1986년 서울대학교 수학과 학사 졸업.

1995년 Florida State university computer science 석사 졸업.

2017년 ~ 2020년 인하대학교 컴퓨터공학과 겸임교수.

2019년 ~ 현재 ㈜큐헛지 이사

<주관심분야 : 머신러닝, 금융투자공학>



송재원(정회원)

2005년 성공회대학교 컴퓨터공학과 학사 졸업.

2007년 인하대학교 컴퓨터공학과 석사 졸업.

2013년 인하대학교 컴퓨터공학과 박사 졸업.

2013~2014 대구경북과학기술원 개발팀장.

2014~2017 인천국제공항공사 선임 연구원

2017~2017 가천길대학병원 연구교수

2017~현재 ㈜ 밸류파인더스 대표이사

<주관심분야 : 머신러닝, 데이터마이닝, 금융IT>