

딥러닝분석을 이용한 고빈도 매매 최적화 성능평가

Performance Evaluation of High Frequency Trading using a Deep Learning

이우식¹

Woosik Lee¹

요약

현대 금융 시장은 과거의 금융 시장과 달리 전 세계 투자자들이 장소에 구애받지 않고 금융 상품을 거래할 수 있게 되었으며 주문 및 체결도 백만분의 1초 단위에서 일어나는 이전과 비교할 수 없을 정도로 빠르게 움직이는 움직이고 있다. 이에 컴퓨터 알고리즘 트레이딩이 더욱 요구되고 진화하고 있지만, 변화하는 시장 데이터, 많은 종류의 시그널 값을 빠르고 지능적인 방식으로 변화시키는데 제한이 있을 뿐 아니라 최적의 속도와 정교함을 가진 실행 논리를 개발·유지하는 것은 매우 어렵고 복잡하다. 본 연구에서는 최적화 알고리즘에 따른 딥러닝 모형기반의 고빈도 매매성능을 비교 분석한다. 분석 결과, 심층 신경망에 의한 고빈도 매매전략 기반에 따른 총 누적 수익이 기초자산의 수익과 비교하여 높은 수익을 얻을 수 있었으며 장기간 지속하는 특성을 보여주었다. 둘째, 일반적으로 성능이 뛰어나다고 알려진 Adam이 유로/달러 환율 고빈도 데이터에 대한 심층 신경망 학습성능에서는 RMSprop과 AdaGrad에 비해 상대적으로 성능을 떨어지는 것으로 나타났다.

핵심어 : 인공지능, 비즈니스 애널리틱스, 핀테크, 국제 금융, 알고리즘 트레이딩

Abstract

Compared to the past financial markets, large investment banks, hedge funds and institutional investors in modern financial markets use powerful computer programs to execute a large number of trades in less than a millionth of a second anywhere in the world. Even though more sophisticated computer algorithm trading systems should be available to market participants upon request, there are limitations to understand market dynamics and market signals and accurate execution logic with optimal speed. This study compares and analyzes the high-frequency trading performance based on the deep learning model according to the stochastic gradient descent algorithm. The main result shows that high-frequency trading performance based on deep learning was likely to learn trading rules. The total cumulative return based on the high-frequency trading strategy based on the deep neural network was able to obtain a high return compared to the return of the underlying asset and showed long-lasting characteristics. Second, Adam, which is generally known to have excellent performance, showed relatively poor performance compared to RMSprop and AdaGrad in deep neural network learning performance for high-frequency data of euro/dollar exchange rates.

Keyword : Artificial Intelligence, Business Analytics, FinTech, International Finance, Intelligent Algorithm Trading

¹ College of Business and Economics, Gyeongsang National University, Jinju, Korea [Professor]
e-mail: woosiklee@gnu.ac.kr

Received(August 21, 2021), Review Result(1st: September 18, 2021), Accepted(October 8, 2021), Published(October 31, 2021)



© 2021 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

1. 서론

우리나라 환율제도의 첫 도입 당시에는 환율의 급격한 변동에 따른 환 리스크(Foreign Exchange Risk)로부터 기업과 은행을 보호한다는 측면에서 일별 환율 변동을 기준 환율 중심으로 $\pm 0.4\%$ 로 허용하였고, 그 이후 허용 변동폭을 늘렸으며 1997년 12월 16일부터 자유변동 환율제를 채택하였다 [1]. 자유변동 환율제하에서는 급격한 환율 변동이 이루어질 수 있으므로 미래를 예측하기가 힘들다. 실제로 1997년 IMF 외환위기와 2008년 글로벌 금융위기에 환율이 급등하면서 경제적 어려움을 겪었다. 특히 2008년 글로벌 금융위기 당시 우리나라 중·소기업들은 환율 예측을 잘못하여 키코(KIKO)라는 외환 헤지 상품으로 막대한 손실을 보았다. 이러한 손실을 막기 위해서도 환율의 예측은 매우 중요하다 [2].

환율예측은 현재 주어진 정보를 이용하여 미래의 환율에 대해 합당한 추정치를 선택하는 작업이다 [3]. 일반적으로 단변량 시계열 분석(Univariate Time Series Analysis)과 다변량 시계열 분석(Multivariate Time Series Analysis)같은 일반적인 통계 모형으로 금융 시계열을 예측하는 데 적용하였다. 하지만 외환시장을 포함한 금융 시장은 경제, 정치 그리고 투자자들의 투자심리 등 복잡한 요인에 의해 영향을 받아 전통적인 방법들은 금융 시계열의 비선형(Non-Linearity)과 복잡성을 해결하지 못함으로써 예측 정확도가 떨어졌다 [4]. 더불어 현대 글로벌 금융 시장은 과거의 금융시장과 달리 전 세계 투자자들이 장소에 구애받지 않고 여러 금융 상품을 거래할 수 있게 되었으며 주문 및 체결도 백만분의 1초 단위에서 일어나는 이전과는 달리 빠르게 움직이는 움직이고 있다. 이에 컴퓨터 알고리즘을 이용한 초단타 매매 또한 계속해서 늘어나는 추세이다. 영국 금융행위감독청(The Financial Conduct Authority)에 따르면 헤지펀드 등이 초단타 매매를 통해 2018년 한 해 동안 미국에서 약 3조원, 중국에서 약 6천억원, 일본에서 약 3천억원 그리고 한국 주식시장에서도 약 1천억원의 수익을 벌어들였다 [5].

2016년 알파고가 이세돌과의 바둑대전에서 우승한 이후 금융권에서는 딥러닝을 포함한 인공지능 경망에 관심이 증가하고 있으며, 이에 따라 인공지능경망 기반의 알고리즘 트레이딩 연구에 큰 관심을 보이고 있다. 알고리즘 기술의 발달과 함께 지도 학습(Supervised Learning), 비지도 학습(Unsupervised Learning) 그리고 강화학습(Reinforcement Learning)을 포함한 기계학습(Machine Learning)이 금융 시계열 예측에 활용되고 있고, 복잡하고 비선형적인 특성을 기존의 통계 모형에 비해 상대적으로 잘 파악해내는 것으로 나타났다 [6]. 하지만 아직 충분히 검증된 상태가 아니고, 인공지능경망(Artificial Neural Network)과 알고리즘트레이딩이 연계된 연구들은 많이 이뤄지지 않고 있는 상태이다. 기존 연구 중, Chinco와 2인 [7]연구에서는 라소 회귀(Lasso Regression)분석을 통해 1분 단위에서 주가 예측을 할 때 수익률에 대한 패턴 정보가 불규칙적으로 발생한다는 것을 보였

고, 박석진외 1인 [8]연구에서는 서포트벡터머신(Support Vector Machine)과 랜덤포레스트(Random Forest)모형을 이용하여 가격의 상승과 하락 방향에 대한 예측력을 분석하였다. 본 연구에서는 최적화 알고리즘에 따른 딥러닝모형기반의 고빈도 매매성능을 비교 분석하고자 한다. 본 논문의 구성은 다음과 같다. 제 1절의 서론에 이어 제 2절에서는 주요 방법론인 AdaGrad, RMSprop 그리고 Adam 최적화 알고리즘에 대한 설명을 소개하였으며, 제 3절에서는 실증분석 및 결과를 확인한다. 마지막으로 제 4절에서는 결론과 시사점을 제시한다.

2. 이론적 배경

2.1 인공신경망

2016년 알파고(AlphaGo)가 이세돌과의 바둑대전에서 우승한 이후 게임, 로봇, 자율주행, 주식거래 그리고 국방 등 다양한 분야에서 인공지능을 사용하고 있다. 인공지능은 인간이나 생물의 뇌신경망에서 영감을 얻어 설계된 알고리즘으로 [그림 1]과 같이 다수의 유닛(Unit)의 가중치(Weight) 결합으로 구성된 네트워크(Network)를 통해 입력 정보와 출력 정보 간의 관계를 근사한다. 즉 인공신경망(Artificial Neural Network) 모델은 입력층, 은닉층 그리고 출력층으로 구성되어 있고 각 층마다 처리단위인 유닛을 가지고 있는 자료처리 시스템으로 다차원의 입력값 $X = (x_1, \dots, x_N)$ 을 받아서 수식(1)에 의해 y 를 출력한다. $W = (w_1, \dots, w_N)$ 에서 w_i 는 입력의 i 번째 요소 x_i 에 대한 결합 가중치이고 함수 $a(\cdot)$ 는 비선형(Non Linear Function) 혹은 활성화함수(Activation Function)라고 한다.

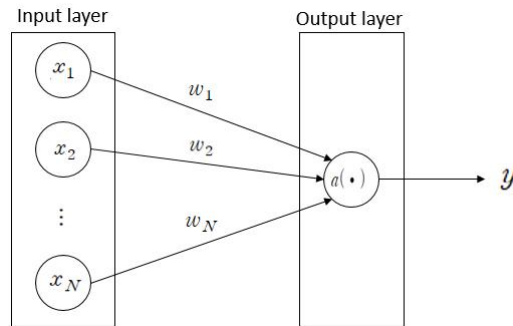
$$y = a\left(\sum_{i=0}^N w_i x_i\right) = a(W^T X), x_0 = 1, i = 1, \dots, N \quad (1)$$

입력 자료 $X_i = (x_{i1}, x_{i2}, \dots, x_{ij})$ 와 이에 대한 목표값 $y = (y_{i1}, \dots, y_{ik})$ 이 쌍으로 주어졌을 때, 입력 자료에 대한 인공신경망 모형의 출력값이 목표값에 가까워지도록 결합 가중치를 계속 갱신하면서 학습을 진행한다. 즉 인공신경망모형을 통한 추정값 \hat{y}_i 과 목표값 y_i 을 비교하여 차이를 줄여나가는 방법으로 결합 가중치를 조절하면서 실제 값에 수렴하게 되고 이것을 학습 알고리즘이라고 한다. 이때 손실함수(Loss Function)로 주로 수식(2)와 수식(3)에 의한 평균 제곱 오차(Mean Squared Error)함수와 교차 엔트로피 오차(Cross-Entropy)함수를 사용한다.

$$Mean\ Squared\ Error = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|^2 \quad (2)$$

$$Cross - Entropy Error = - \sum_{i=1}^N \sum_{k=1}^K (y_{ik}) \ln(\hat{y}_{ik}) \tag{3}$$

기존의 인공지능망 모형 경우, 입력된 데이터에 대한 특징을 크게 향상시키기 위해 은닉층에 유닛수(Hidden Unit)를 증가시키거나, 여러 개의 은닉층을 추가하여 사용하였다 [9]. 하지만 여러 은닉층을 추가할수록 기울기 소실 문제가 발생하였고, 이를 해결하기 위해 2006년 제프리 힌튼교수가 딥러닝(Deep Learning)을 제시하였다. 딥러닝은 다수의 계층으로 구성된 심층 인공지능망(Deep Neural Networks)을 기반으로 입력과 출력 사이의 복잡한 비선형 관계를 나타낼 수 있을 뿐 아니라 자료로부터 높은 수준으로 추상화된 정보를 학습한다.



[그림 1] 인공지능망의 구조

[Fig. 1] Structure of Artificial Neural Network

2.2 연구모형

본 연구에서는 심층 신경망을 이용한 고빈도 매매 최적화 성능평가를 위해 여러 가지 초매개변수(Hyperparameter)중에서 최적화 알고리즘인 경사하강법(Gradient Descent)에 중점을 두고 자료를 분석하였다. 경사하강법은 인공지능망을 학습을 위해 손실 함수(Loss Function) 또는 비용 함수(Cost Function)를 최소화되도록 가중치(Wight)를 갱신하기 위해 이용하는 알고리즘이다. 하지만 경사하강법의 단점은 한 번 가중치를 업데이트하기 위해 전체 데이터를 사용하는 것이고 이로 인해 비효율적이고 학습시간이 오래 걸린다. 이를 위해 확률적 경사하강법(Stochastic Gradient Descent)을 이용하는데 즉 자료를 샘플링하여 적은 자료로 학습을 하고, 이 과정을 계속 반복하면서 확률적으로 전체자료를 학습하는 것과 같은 효과가 나타나게 하는 것이다. 그러나 경사하강법과 확률적 경사하강법은 학습 도중 극소점(Local Minimum Point)에 빠져 실질적인 극소값을 찾을 수 없는 상태가 발생할 수 있다. 이에 본 연구에서는 Adaptive Gradient(AdaGrad), Root Mean Square

Propagation(RMSProp) 그리고 Adaptive Moment Estimation(Adam) 경사하강법 알고리즘을 통해 심층 신경망을 학습한다 [10].

2.2.1 Adaptive Gradient(AdaGrad) 최적화 알고리즘

실제 인공신경망 학습 시, 자료의 특성에 따른 효과적인 학습률(Learning Rate) 선택은 중요하다. 학습률이 너무 작으면 최적점을 찾기 위해 수많은 학습 횟수가 요구되고, 학습률이 너무 크면 학습을 통한 추정값과 실제 목표값의 오차를 줄이는 방향으로 가지 못하고 발산해서 학습이 제대로 이루어지지 않는다. 이런 문제를 학습률 규제(Learning Rate Decay)를 통해 해결하는데, AdaGrad 최적화 알고리즘은 가중치가 조정됨에 따라 학습률도 자동으로 갱신되도록 설계되었다. AdaGrad 최적화알고리즘의 업데이트 방법은 다음과 같다 [10].

$$G = G + \nabla_w J(w) \odot \nabla_w J(w) \quad (4)$$

$$w = w - \frac{\eta}{\sqrt{G + \epsilon}} \odot \nabla_w J(w) \quad (5)$$

여기서 \odot 는 행렬의 원소별 제곱을 의미하고, ϵ 는 분모가 0이 될 때 오류를 제한하기 위한 작은 값이다 [10].

2.2.2 Root Mean Square Propagation(RMSProp) 최적화 알고리즘

RMSProp 최적화 알고리즘은 AdaGrad 최적화 알고리즘을 개선한 알고리즘으로 AdaGrad 최적화 알고리즘에서 G 는 현재까지의 변화량의 총합으로 시간이 지나면서 학습률은 감소하지만, RMSProp 최적화 알고리즘은 과거와 현재의 변화량의 지수 평균으로 인해 학습률이 급감을 방지할 수 있다 [10]. 즉 최근에 학습한 자료가 가중치 조정에 좀 더 많은 영향을 미치도록 설계되었다.

$$G = \gamma G + (1 - \gamma) \nabla_w J(w) \odot \nabla_w J(w) \quad (6)$$

$$w = w - \frac{\eta}{\sqrt{G + \epsilon}} \odot \nabla_w J(w) \quad (7)$$

2.2.3 Adaptive Moment Estimation(Adam) 최적화 알고리즘

Adam 최적화 알고리즘은 가중치 수정값에 관성(Momentum)을 추가한 Momentum알고리즘과 RMSProp 최적화 알고리즘의 특성을 모두 추가한 알고리즘으로, 최적화에 의한 업데이트 경로를 변경하는 운동량(Momentum) 개념을 기반으로 과거의 업데이트 정보를 부분적으로 반영하고, 새롭

계 계산된 경사 방향과의 조합으로 최종적인 업데이트 크기를 계산한다. 즉 운동량처럼 현재까지 계산해온 기울기의 지수 평균을 저장하고, 아래와 같이 RMSProp 최적화 알고리즘처럼 기울기의 제곱 값의 지수 평균(Exponential Average)을 저장한다 [10].

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla_w J(w) \tag{8}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) (\nabla_w J(w))^2 \tag{9}$$

위 식에서 m_t 과 v_t 는 각각 1차 운동량과 2차 운동량의 추정값을 나타내고, β_1 와 β_2 는 감쇠 비율로 운동량 측정에 사용되는 고정 변수로서 일반적으로 β_1 는 0.9, β_2 는 0.999가 사용된다. 이러한 갱신 방식만을 이용할 경우, 운동량의 초기 설정값이 0으로 초반 업데이트 크기가 지나치게 작아진다. 이를 방지하기 위해 Adam 최적화 알고리즘에서는 아래와 같이 1차와 2차 운동량 추정치의 편향 보정한다 [10].

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^n} \tag{10}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^n} \tag{11}$$

여기서 \hat{m}_t 과 \hat{v}_t 은 보정된 운동량 추정치이다. 따라서 RMSProp 최적화 알고리즘의 특성에 추가적으로 운동량 개념을 적용한 Adam 최적화 알고리즘은 식은 수식(12)와 같다 [10].

$$w = w - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \tag{12}$$

3. 실증분석

3.1 자료의 구성

본 연구에서 사용할 표본은 캐글(Kaggle)에서 제공하는 일중 유로/달러 환율(EURUSD)이고 실험을 위해 2000년 5월 30일부터 2020년 6월 30일까지 1분 종가데이터를 수집했다. 유로/달러 환율 일중 종가에 대한 기초통계량은 [표 1]에서 살펴볼 수 있다. 이 환율은 음의 왜도를 보이는데 이는 부정적인 극단 현상이 나타날 가능성이 정규 분포(Normal Distribution)보다 높음을 의미하고, 음의 첨도라는 것은 꼬리가 정규 분포보다 얇다는 것을 의미한다 [11].

$$\text{환율변화율} = \ln(\text{현물환율}(t) / \text{현물환율}(t-1))$$

[표 1] 기초통계량

[Table 1] Descriptive Statistics

	유로/달러 환율(EURUSD)
평균	1.2242
중앙값	1.2309
최대값	1.6038
최소값	0.8233
표준편차	0.1526
왜도	-0.2347
첨도	-0.1296

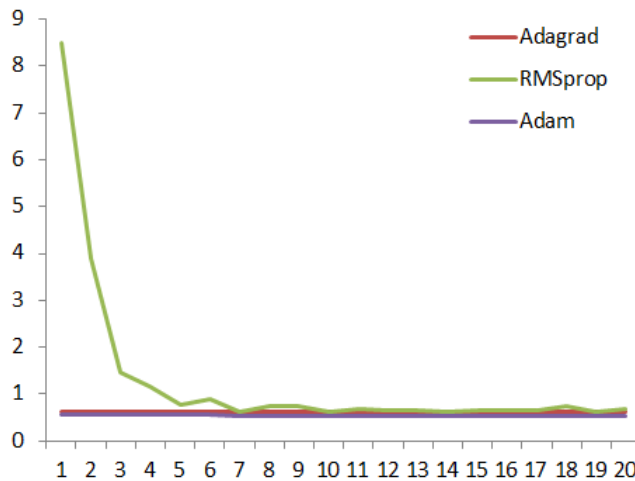
3.2 모형의 추정 및 분석

심층 신경망에 의한 고빈도 매매 최적화 성능평가분석에 다음과 같은 방법을 사용하였다. 전체 자료의 80%는 심층 신경망을 통해 학습시키고, 나머지 20%는 훈련된 모형에 입력변수로 투입시킴으로써 예측되는 값과 실제 값과의 비교와 검증을 통하여 최종적으로 최적의 모형을 확인하였다.

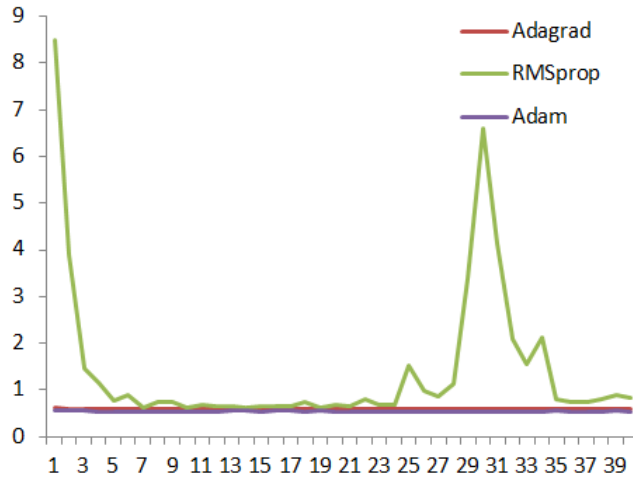
AdaGrad, RMSprop 그리고 Adam 최적화 알고리즘으로 이루어진 각각의 심층 신경망의 학습을 위해 유로/달러 환율의 일중 증가, 단순이동평균, 추세 추종 그리고 변동성 등 기술적 분석 지표에 기반한 시차를 입력층의 입력데이터로 사용하고, 매수와 매도 신호를 목표값으로 지정하였다. 이때 활성화함수에서 발생할 수 있는 기울기 소실문제를 해결하기 위해 활성화함수로 렐루(ReLU)함수를 이용하였다. 심층 신경망의 최적화 함수와 손실함수는 각각 AdaGrad, RMSprop, Adam 그리고 평균 제곱오차(Mean Squared Error)로 설정하였다.

인공신경망을 구성할 때, 은닉층의 수, 은닉층에 존재하는 유닛 수, 활성화 함수, 신경망의 초기화 기법 그리고 최적화 등이 적절한 조합을 통해 최적화되지 못하면 높은 성능을 이끌어 낼 수 없다. 이러한 인공신경망의 초매개변수들의 모든 가능한 조합(Combination)을 통한 최적의 조합을 찾는 것은 매우 어려운 문제이며 많은 계산량이 요구된다. 여러 초매개변수 중 학습과 가장 직접적으로 관련된 최적화 알고리즘은 모형의 손실 값을 최소화하기 위해 가중치 갱신이라는 가장 중요한 역할을 한다 [12]. 동일한 구조와 동일한 초매개변수를 갖는 인공신경망이라도 최적화 알고리즘에 따라 성능이 다르게 나타날 수 있으므로 적합한 최적화 알고리즘을 찾는 것은 매우 중요하다. 이에 본 논문에서는 AdaGrad, RMSprop 그리고 Adam 최적화 알고리즘이 심층 신경망의 학습에 미치는 영향을 측정하는 연구를 수행했다. 즉 AdaGrad, RMSprop, 그리고 Adam 최적화 알고리즘을 사용하여 유로/달러 환율의 일중 증가와 단순이동평균, 추세 추종 그리고 변동성 등 기술적 분석 지표에 기반한 시차를 입력값으로 그리고 매수와 매도 신호에 따른 매매전략 기반의 총 누적 성과

를 계산한다. 그 결과 [그림 3]에서와 같이 심층 신경망에 의한 고빈도 매매전략 기반에 따른 총 누적 수익이 기초자산의 수익과 비교하여 높은 수익을 얻을 수 있었으며 장기간 지속하는 특성을 보여주었다. 테스트 기간 동안 심층 신경망 기반의 고빈도 매매 전략의 모든 성과는 100%이상을 보였고, 특히 최근 5개의 시차를 포함하고 학습진행이 20인 RMSprop 최적화 알고리즘의 경우, 다른 조건에서보다 더 높은 전략의 성과를 보였다 [그림 3(c)]. 더불어 RMSprop 최적화 알고리즘의 경우, [그림 4]에서와 같이 리샘플링(Resampling)을 통해 5분과 10분의 고빈도 데이터의 심층 신경망에 의한 고빈도 매매전략 기반에 따른 총 누적 수익도 기초자산의 수익과 비교하여 높은 수익을 얻을 수 있었으며 장기간 지속하는 특성을 확인할 수 있어, 심층 신경망에 의한 고빈도 매매전략 전략을 활용하여 고빈도 데이터의 움직임을 예측하는데 딥러닝 및 RMSprop의 실효성을 확인할 수 있었다. 둘째, [그림 2]를 통해 학습진행 수를 증가시킬 때, 테스트에서 손실이 어떻게 변화하는지 분석하고자 한다. 심층 신경망이 학습함에 따라 손실 혹은 비용이 감소하고 있음을 알 수 있고, 특히 Adagrad와 Adam의 경우, 테스트 에러가 일정한 수준까지 떨어진 다음 다시 높아지지 않는 것을 통해 모형이 학습에 과적합하지 않다는 것을 알 수 있다. 또한 Adagrad와 Adam의 경우, 학습진행이 20 이후에는 거의 직선에 가깝게 일정하다 [그림 2(a)]. 이는 학습진행의 수를 더 늘리는 것이 유용하지 않음을 의미한다. 이를 확인하기 위해 [그림 2(b)]와 같이 학습진행의 수를 20에서 40으로 2배로 늘려 정확도를 확인해보았더니 역시 큰 변화는 없었다. 즉 학습에 더 많은 시간을 투자한다고 해서 테스트에서 정확도의 결과가 개선되는 것이 아니다. 하지만 심층 신경망 기반의 알고리즘 트레이딩 전략을 전개함으로써 시장 방향의 예측 정확도가 50%보다 크므로 확률보행 구성에 비해 통계적 우위가 있다는 점은 의미가 있다.



(a) Lag = 5, Epoch = 20



(b) Lag = 5, Epoch = 40

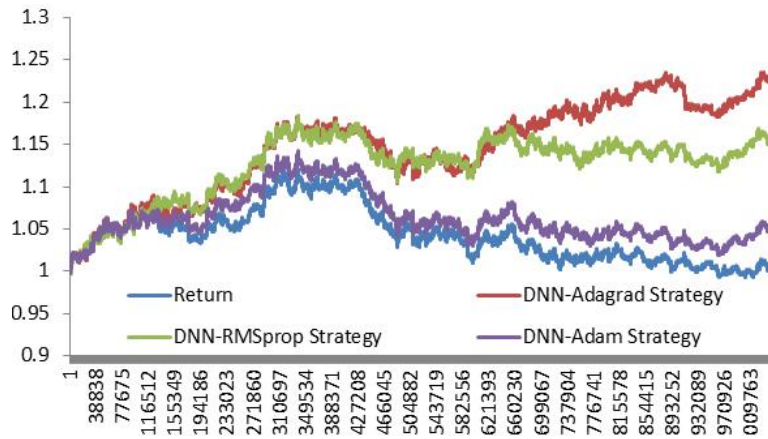
[그림 2] 테스트 손실 값

[Fig. 2] Test Loss Values against Number of Epochs

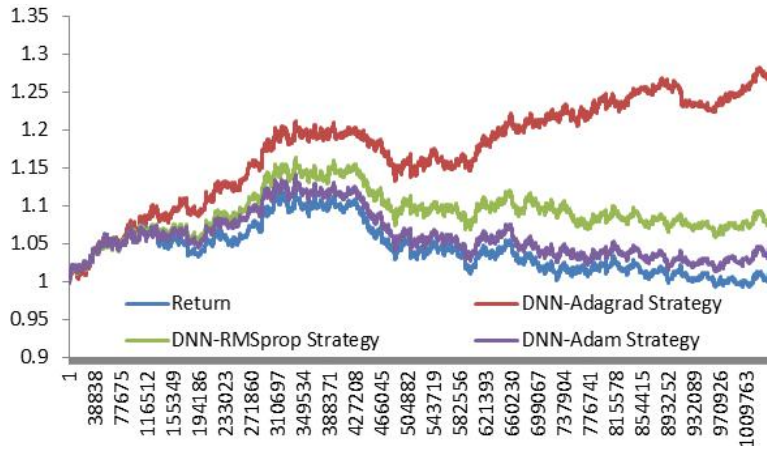
[표 2] 정확도

[Table 2] Accuracy

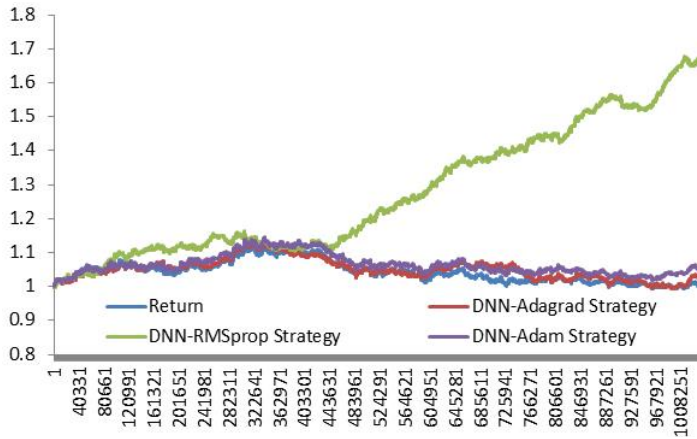
	알고리즘	훈련	평가
학습진행 20	Adagrad	0.6145	0.6454
	RMSprop	0.5437	0.6526
	Adam	0.5881	0.6240
학습진행 40	Adagrad	0.6128	0.6449
	RMSprop	0.4915	0.6369
	Adam	0.5904	0.6267



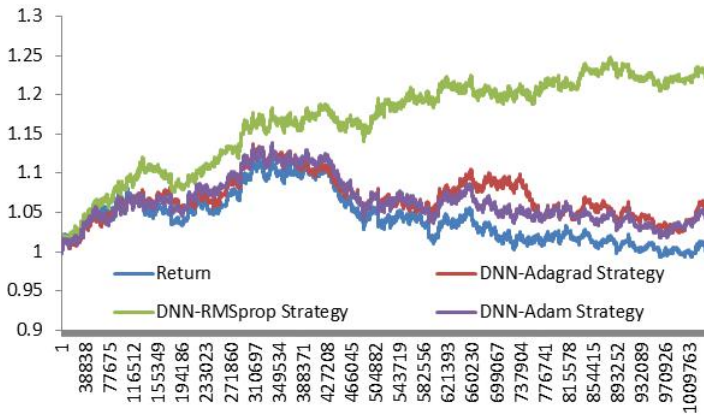
(a) Lag = 1, Epoch = 20



(b) Lag = 1, Epoch = 40



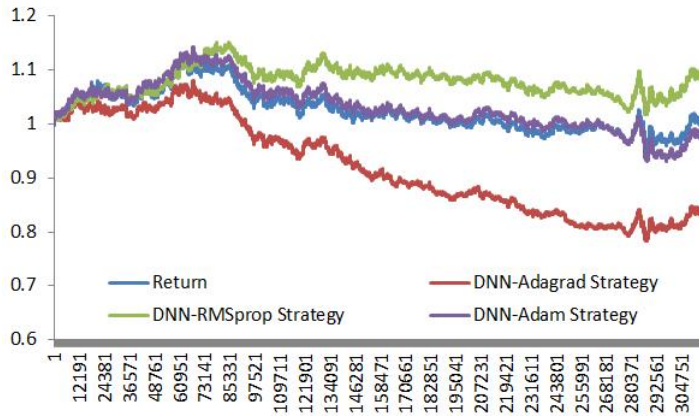
(c) Lag = 5, Epoch = 20



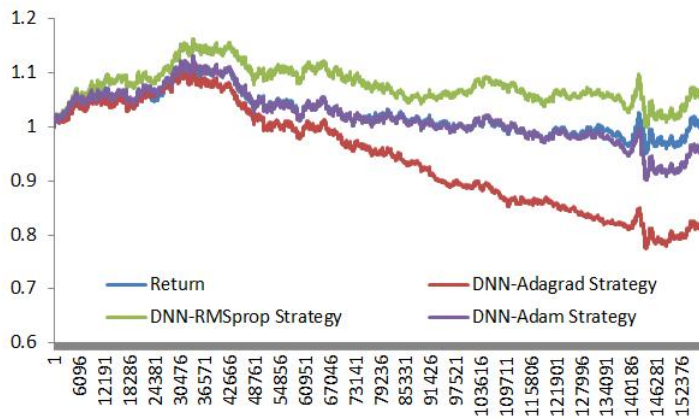
(d) Lag = 5, Epoch = 40

[그림 3] 다양한 최적화에 대한 알고리즘 트레이딩 전략의 총 성과 그래프 - 1분 자료

[Fig. 3] Total Cumulative Returns - One Minute



(a) Lag = 5, Epoch = 20



(b) Lag = 5, Epoch = 40

[그림 4] 다양한 최적화에 대한 알고리즘 트레이딩 전략의 총 성과 그래프 - 5분과 10분 자료

[Fig. 4] Total Cumulative Returns - Five and Ten Minutes

4. 결론

최근 전 세계적으로 딥러닝을 포함한 인공지능경망에 관심이 증가하고 있으며, 이에 따라 인공지능 경망을 활용하는 연구가 큰 폭으로 증가하고 있다. 하지만 인공지능경망을 구성할 때, 은닉층의 수, 은닉층에 존재하는 유닛 수, 활성화 함수, 신경망의 초기화 기법 그리고 최적화 등이 적절한 조합을 통해 최적화되지 못하면 높은 성능을 이끌어 낼 수 없다. 이러한 인공지능경망의 초매개변수들의 모든 가능한 조합(Combination)을 통한 최적의 조합을 찾는 것은 매우 어려운 문제이며 많은 계산량이 요구된다. 여러 초매개변수 중 학습과 가장 직접적으로 관련된 최적화 알고리즘은 모형의 손실값을 최소화 할 수 있도록 가중치를 갱신하는데 가장 중요한 역할을 담당한다 [12]. 동일한 구조

와 동일한 초매개변수를 갖는 인공신경망이라도 최적화 알고리즘에 따라 성능이 다르게 나타날 수 있으므로 적합한 최적화 알고리즘을 찾는 것은 매우 중요하다.

본 연구에서는 심층 신경망을 이용한 고빈도 매매 최적화 성능평가를 위해 AdaGrad, RMSProp 그리고 Adam 최적화 경사하강법 알고리즘 기반의 학습성능을 분석하였다. 분석 결과는 다음과 같다. 첫째, 심층 신경망에 의한 고빈도 매매전략 기반에 따른 총 누적 수익이 유로/달러 환율의 수익과 비교하여 높은 수익을 얻을 수 있었으며 장기간 지속하는 특성을 보여주었다. 이를 통해 고빈도 가격 움직임을 예측하는데 딥러닝의 실효성을 확인할 수 있었다. 둘째, 일반적으로 성능이 뛰어나다고 알려진 Adam과 다르게 유로/달러 환율 고빈도 데이터에 대한 심층 신경망 학습성능에서는 RMSprop과 AdaGrad이 Adam에 비해 상대적으로 성능이 뛰어난 것으로 나타났다. 이를 통해 특히 RMSProp이 직간접적으로 서로 복잡하게 얽혀 있는 수많은 변수들에 의해 불규칙하게 변화하는 고빈도 외환시장 환경에서 보다 잘 작동한다 높은 학습 효과를 확인할 수 있다.

인공신경망을 구성할 때, 은닉층의 수, 은닉층에 존재하는 유닛 수, 활성화 함수, 신경망의 초기화 기법 그리고 최적화 등이 적절히 최적화하지 못하면 높은 성능을 이끌어 낼 수 없다. 이러한 인공신경망의 초매개변수들의 모든 가능한 조합(Combination)을 통한 최적의 조합을 찾는 것은 매우 어려운 문제이며 많은 계산량이 요구된다. 여러 초매개변수 중 학습과 가장 직접적으로 관련된 최적화 알고리즘은 모형의 손실 값을 최소화하기 위해 가중치 갱신이라는 가장 중요한 역할을 한다 [2]. 동일한 구조와 동일한 초매개변수를 갖는 인공신경망이라도 최적화 알고리즘에 따라 성능이 다르게 나타날 수 있으므로 적합한 최적화 알고리즘을 찾는 것은 매우 중요하다. 이에 본 논문에서는 AdaGrad, RMSprop 그리고 Adam 최적화 알고리즘에 따른 심층 신경망의 학습에 미치는 영향을 측정하는 연구를 고빈도 자료를 활용하여 수행했다. 하지만 본 논문에도 향후 몇 가지 보완할 점이 필요하다. 보다 정량적인 방식인 그리드 검색 또는 베이지안 최적화와 같은 기법이 필요하고, 심층 신경망 모형에 다양한 정형화된(Structured) 경제지표와 비정형화(Unstructured) 데이터 등을 추가 한 방안과 거래비용 추가 등을 추가한 비교연구가 필요하다.

References

- [1] W. Lee, Y. Lee “Time Series Analysis of Won/Yen Exchange Rate”, Korea Trade Review, vol. 36, no. 3, June 2011, pp. 1-21.
- [2] H. Lee, “Time series models based on relationship between won/dollar and won/yen exchange rate”, Journal of the Korean Data and Information Science Society, vol. 27, no. 6, November 2016, pp. 1547-1555, doi: 10.7465/jkdi.2016.27.6.1547.
- [3] J. A. Han, C. H. An, “A Comparative Study on the Prediction Model of Won/Dollar Exchange Rate Using Time Series Model”, The Society of Convergence Knowledge Transactions, vol. 7, no. 4, December 2019,

- pp. 69-78, doi: 10.22716/sckt.2019.7.4.052.
- [4] S. Sun, Y. Wei, S. Wang, "AdaBoost-LSTM ensemble learning for financial time series forecasting", *Computational Science - ICCS 2018*, June 11-13, 2018, Wuxi, China, pp. 590-597, doi: 10.1007/978-3-319-93713-7_55.
- [5] A. Osipovich, "Ultrafast Trading Costs Stock Investors Nearly \$5 Billion a Year, Study Says", *wsj.com*, <https://www.wsj.com/articles/ultrafast-trading-costs-stock-investors-nearly-5-billion-a-year-study-says-11580126036>, (accessed January 27, 2020).
- [6] L. d. P. Marcos, "Ten Financial Applications of Machine Learning", *SSRN Electronic Journal*, January 2018, pp. 459-470, doi: 10.2139/ssrn.3197726.
- [7] A. Chincio, A. D. Clark-Joseph, M. Ye, "Sparse Signals in the Cross-Section of Returns", *Journal of Finance*, vol. 74, no. 1, February 2019, pp. 449-492, doi: 10.1111/jofi.12733.
- [8] S. J. Park, C. S. Chung, "Forecasting Ability of Machine Learning Algorithms using High-frequency", *Journal of Money & Finance*, vol. 33, no. 4, December 2019, pp. 31-60, doi: 10.21023/JMF.33.4.2.
- [9] H. Choi, Y. Min, "Introduction to deep learning", *Korea Information Processing Society Review*, vol. 22, no. 1, January 2015, pp. 7-21.
- [10] W. Lee, "Performance Evaluation of Portfolio using a Deep Q-Networks", *Journal of Next-generation Convergence Information Services Technology*, vol. 10, no. 4, August 2021, pp. 459-470, doi: 10.29056/jncist.2021.08.10.
- [11] W. Lee, "A deep learning analysis of the KOSPI's directions", *Journal of the Korean Data and Information Science Society*, vol. 28, no. 2, March 2017, pp. 287-295, doi: 10.7465/jkdi.2017.28.2.287.
- [12] G. Gihun, C. Park, H. Im, "Performance Evaluation of Machine Learning Optimizers", *Journal of Institute of Korean Electrical and Electronics Engineers*, vol. 24, no. 3, September 2020, pp. 766-776, doi: 10.7471/ikeee.2020.24.3.766.