

# 거시지표와 딥러닝 알고리즘을 이용한 자동화된 주식 매매 연구

홍성혁  
백석대학교 ICT학부, 부교수

## A Research on stock price prediction based on Deep Learning and Economic Indicators

Sunghyuck Hong  
Baekseok University, Division of ICT, Associate Professor

요 약 거시경제는 한 나라 경제 전체의 움직임을 보여주기 때문에 주식을 분석할 때 선행되어 분석되는 지표 중 하나이다. 실업률, 이자율, 물가, 국민소득, 환율, 통화량, 국제수지 등 국가차원의 경제 상황 전반은 주식시장에 직접적인 영향을 미치고, 경제 지표는 개별 주가와 상관계수가 있기 때문에 주식을 예측하기 위해 많은 증권사 애널리스트들이 관심을 갖게 지켜보고, 개별 주가에 영향을 고려하여 매수와 매도를 판단하는 주요한 근거자료가 되고 있다. 주가에 영향을 미치는 경제 지표를 선행지표로 분석하고, 주가예측을 딥러닝 기반의 예측을 통하여 예측 후 실제 주가를 비교하여 차이가 발생하면 거시지표에 대한 가중치를 조절하여 지속적인 반복학습을 통하여 주식의 매수와 매도를 판단한다면, 주식은 더 이상 도박과 같은 투기가 아닌 건전한 투자가 될 수 있다. 따라서 본 연구는 거시지표와 인공지능의 딥러닝 알고리즘방식을 이용하여 자동화된 주식매매가 가능하도록 연구를 수행하였다.

주제어 : 주가분석, 빅데이터, 텍스트마이닝, 인공지능, 예측 시스템

Abstract Macroeconomics are one of the indicators that are preceded and analyzed when analyzing stocks because it shows the movement of a country's economy as a whole. The overall economic situation at the national level, such as national income, inflation, unemployment, exchange rates, currency, interest rates, and balance of payments, has a great affect on the stock market, and economic indicators are actually correlated with stock prices. It is the main source of data for analysts to watch with interest and to determine buy and sell considering the impact on individual stock prices. Therefore, economic indicators that impact on the stock price are analyzed as leading indicators, and the stock price prediction is predicted through deep learning-based prediction, after that the actual stock price is compared. If you decide to buy or sell stocks by analysis of stock prediction, then stocks can be investments, not gambling. Therefore, this research was conducted to enable automated stock trading by using macro-indicators and deep learning algorithms in artificial intelligence.

Key Words : Stock analysis, Big data, Text mining, AI, Prediction

\*This research was supported by 2020 Baekseok University Research Fund.

\*Corresponding Author : Sunghyuck Hong(shong@bu.ac.kr)

Received October 5, 2020

Revised October 12, 2020

Accepted November 20, 2020

Published November 28, 2020

## 1. 서론

거시경제란 국민경제 전체의 전반적 흐름이며, 경제 성장이나 경기변동을 알려주는 지표로 국가경영을 하기 위한 중요한 지표 중 하나이다. 다양한 거시지표가 있지만 주요지표로는 경제성장률, 실업률, 물가상승률, 빈부 격차, 정부 재정적자, 기준금리, 환율, 신용등급 등이 있다 [1]. 대외 거시지표가 주가에 영향을 미치는 지표로는 IMF 외환위기 이후에는 미국정책금리, 미국주가의 변화가 외국자본의 유입으로 더욱더 민감하게 반응하고 있다 [2-4]. 국내 주식 시장을 예측하기 위해서는 국내 거시지표와 함께 미국 거시지표도 살펴보고 각 지표별 상관관계를 과거 데이터를 통하여 분석하면 가까운 미래를 예측할 것으로 기대한다. 따라서 본 연구를 대내외 과거 거시 지표를 분석하고, 딥러닝을 통하여 예측한 후 실제 주가와와의 상관관계를 밝혀내고 머신러닝을 통하여 개별 주가를 예측하고 실제 주가와 비교하여 차이점을 분석한 후 거시지표가 주가에 미치는 영향을 연구하여 좀 더 정밀하게 주식을 예측하는 시스템을 제안하여 예측 모델을 제공하는 것이 본 연구의 목적이다.

## 2. 주가 예측 시스템

주가는 기업의 미래 가치이고, 그 가치는 재무재표 분석을 통하여 알 수 있기 때문에, 기존의 인공지능을 이용한 주가 예측 시스템은 분석 대상 기업의 재무제표를 분석하여 주가를 예측하고 있다. 하지만, 기업의 재무재표는 큰 틀에서 거시지표가 반영된 것이므로 거시지표 분석을 선행할 경우, 기업의 재무지표의 예측이 용이하고, 매수와 매도의 최적의 시기를 잡을 수 있다. Table 1은 전통적인 기업 분석을 위한 재무제표이다.

Table 1. Major corporate financial statements affecting the stock price

Return of Equity	Debt ratio
Price Book-value Ratio	Net profit growth rate

최근 연구에 따르면 기업의 재무제표가 주가에 미치는 영향을 최대 19%이기 때문에 기존의 재무제표 분석만으로 기업의 미래 가치를 판단하는 시대는 지났으며, 거시 지표와 같은 선행지표도 함께 분석하고 예측하는 것이

필요하다. Table 2는 미국주식에 영향을 미치는 주요 미국내 거시 지표이다.

Table 2. Major corporate financial statements affecting the stock price in the U.S.A

DFE	UNRATE
CPIAUCSL	CSCICP
PCU)	HSN
CCSA	FGI
BAMLH	RAW

Table 2의 지표외에도 다양한 지표들이 존재하지만 본 연구에서는 주요 10개 지표와 개별 주가와와의 상관관계를 살펴보고, 특정 거시 지표가 선행인지 후행인지를 판단하여 주가 예측을 좀 더 정확하게 예측한다. Table 2에서 DFE는 중앙은행금리이며, Untrate은 실업률, CPIAUCSL은 소비자지수, CSCICP는 소비자물가, PCU는 자동차판매지수, CCSA는 실업수당청구건수, HSN은 주택판매지수, FGI는 공포와 탐욕지수, BAMLH는 하일드 채권스프레드이며, 마지막으로 RAW는 원자재 지수이다. 주택판매지수와 자동차 판매지수는 거시 지표와의 관계는 밀접한 상관관계가 있으며, 미국내 주택 구입 평균 금액은 377,200 달러(2020년 기준)이고, 미국내 자동차 평균 구매 금액은 35,000 달러(2020년 기준)로 소비자 경제 측면에서 살펴보면 높은 비중을 차지하고 있다. 따라서 주택과 자동차 지표는 경기를 예측하는 선행 지표로 여겨진다. 자동차의 경우에는 수만 가지의 부품으로 이루어져 있기 때문에 자동차 판매대수는 경기가 회복되는 국면에서는 동행지표 또는 후행지표의 역할을 한다. 신차판매나 리스가 줄어들면 회사에서 실직을 했거나 비즈니스가 잘 안 되는 의미로 받아들일 수 있다. 또한 소비자 심리 지수는 지수가 오르면 소비 욕구가 넘친다는 의미이고, 낮아지면 소비가 위축된다는 의미이며 경제와 밀접한 관계가 있다. 2008년과 같은 미국발 금융위기때에는 소비자 심리지수가 먼저 하락하고 증시가 저점으로 내려가는 것을 알 수 있다. 소비자 심리지수는 직접 소비자에게 설문 조사를 하기 때문에 정확성이 떨어지고 변동성도 크다는 단점은 있다 [5].

Fig. 1은 2004년부터 2020년까지 실업률(청색)과 신규 주택 판매 지수(적색)를 비교한 그래프로 정확하게 정반대로 진행되는 것을 알 수 있다.

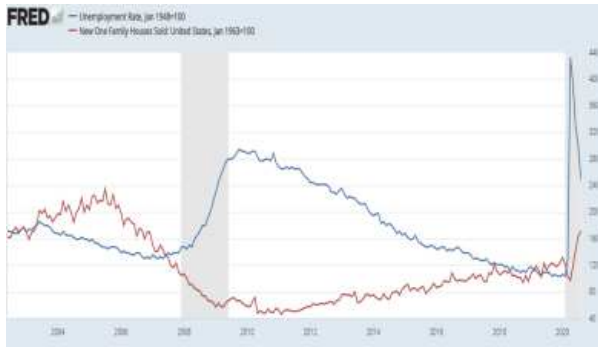


Fig. 1. Unemployed rate and Home Sales Index

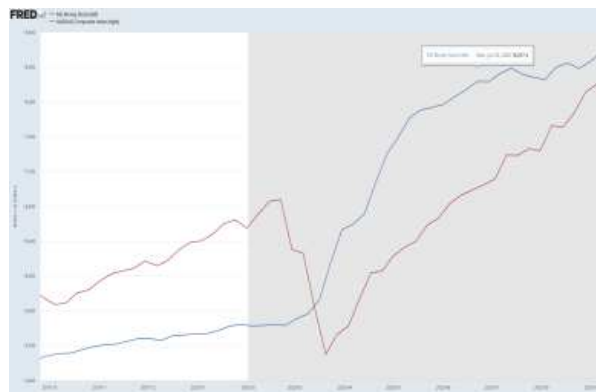


Fig. 2. Money and NASDAQ Index

Fig. 2은 2019년 10월부터 2020년 8월까지 통화량(청색)과 NASDAQ(적색)를 비교한 그래프로 동반 고점으로 진행되는 것을 알 수 있다. 통화량은 시중에 얼마나 통화가 풀렸는지 나타내는 지표로 유통되는 통화와 요구불 예금을 포함한다. 한마디로 시중에 자금을 얼마나 많이 풀었는지 확인하는 유통성 지표이다. 미국 연방준비은행은 주가 하락 할 때마다 돈을 풀어서 경기를 부양하고 있다. 미국 연방준비은행은 양적완화 완화 정책을 중앙은행의 금리를 통해서 실시한다. 금리를 낮추면 시중에 돈이 많이 풀리게 되어 소비와 기업의 매출이 늘어나지만, 금리를 높이면, 자연스럽게 시장에 돈이 회수 된다. 시장에 돈이 줄어들면 기업은 경영이 어려워 직원을 해고하게 된다. 따라서 미국 연방준비은행은 경기 나쁠 때 돈을 유통시키기 위해 기준금리를 거의 제로 금리에 가깝게 유지한다.

주택 판매 지수는 선행지표로 주택이 재산에 차지하는 비중이 크기 때문에 실직을 하거나 은행에서 대출이 불가능할 때 집을 팔아서 부족한 재정을 충당하기 때문에 신규주택 판매 증가는 경기 호황을 의미하고, 기존 주택 판매 증가는 현금을 보유하기 위한 활동으로 경기 위축

을 뜻하게 된다. 주간실업수당 청구 건수는 주가 지수가 지속적으로 하락할 경우 후행지표에 해당한다. 2008년 미국발 금융위기 때 주식이 먼저 하락하고 주간실업수당 청구 건수 증가하였다. 실업수당 청구건수는 불황이 시작될 때는 후행지표이지만, 증시가 좋아질 경우에는 선행지표로서 사용된다.

또한, Fig.3은 공포와 탐욕지수로 CNN 제공하는 지표로 주식 시장에 포함된 낙관의 상대 강도 및 내재된 공포심을 나타내는 지표이다. 0은 극단적인 공포를 나타내며, 100은 극단적인 낙관을 의미하기 때문에 일반적으로 80이상 시에는 매도하는 것이 이득이고, 20미만 시는 매수하는 것이 일반적이다 [6-9].



Fig. 3. Fear and Greedy Index

### 3. 거시지표와 딥러닝을 이용한 주가 예측 모델

미국 주식 거래에서의 자동주가 매매 비중은 꾸준히 상승하여 50%이상 정도로 높은 수준이다. 반면, 빈번한 거래 회사로 공식 분류된 회사의 이익 규모는 2010년 이후 급격히 감소한 후 회복하는 것으로 조사되었지만, 주식거래 수수료 인하가 될 경우 시스템 매매는 더욱 증가할 것으로 추정한다. 시스템 매매는 초단기로 매매를 반복함으로써 이익을 추구한다. 대한민국의 경우 수수료가 거래금액의 0.25%에서 0.3% 이므로 고빈도 매매는 불가능한 수준이다. 하지만, 대한민국도 수수료 인하를 하게 된다면, 자동매매 프로그램으로 인한 거래가 증가될 것으로 기대한다 [10-12].

현재 증권사에서 사용 중인 인공지능 기반 자동트레이딩 시스템은 펀드매니저가 지정해 놓은 60가지의 지표의 조건절들의 조합으로 특정 조건의 수가 많거나 작으면 매매하는 방식으로 진행하기 때문에 딥러닝을 통한 학습 개념이 있는 것은 아니고, 단지 조건을 따지는 연산을 컴

퓨터 프로그램이 대신하여 정확하고 빠르게 처리하고 있는 실정이다 [13-15].

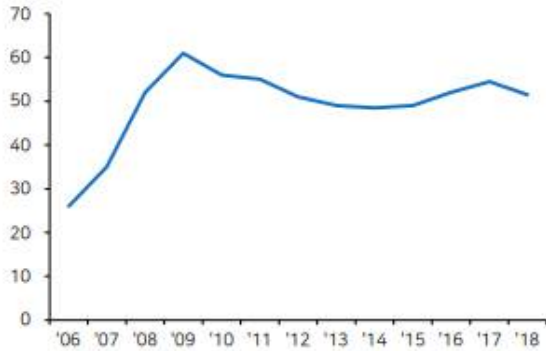


Fig. 4. Ratio of High Frequency Trading in the U.S. Stock Market

### 3.1 주가 예측 모델

Fig. 5는 딥러닝과 거시지표를 이용한 주가 예측 모델 구성도이다. 제안하는 주가 예측 모델은 개별 주가 분석과 거시지표 10개에 대한 예측을 한 후 예측된 주가에 거시지표 예측 값을 가중치를 주고 최종 주가를 예측한 후 실제 주가를 증권사 API를 통하여 가지고와서 예측된 주가와 비교하여 맞으면 예측된 주가의 목표가와 비교하여 매매를 결정하게 한다. 예측된 주가와 실제 주가와 차이가 발생하게 되면, 다시 예측된 지표들( $M^1, M^2, \dots, M^N$ )을 수정하고 학습 시킨 후 다음 시간(시간, 일, 주, 월 단위) 주가 예측 수행한다. Fig. 6는 Table 4를 근거로 작성되었으며, Fig. 6의 순서는 왼쪽부터 Table 2에 있는 지표순서대로 나열되어 있고, 마지막은 예측된 값이다. Table 4는 거시지표와 주가의 가격 변동을 시그모이드 함수를 통해서 정규분포화 시킨 후 LSTM(Long-Short Term Memory)을 이용한 NASDAQ 주가 예측을 하고 실제 주가와 비교한 것이다. NASDAQ 과거 데이터는 야후 파이낸스에서 추출하였고, 거시지표 과거 데이터는 fred.stlouisfed.org 사이트에서 추출하였다. 예측 모델은 2주(10일)을 기준으로 1개를 예측하고 일씩 증가시키면서 10이라는 단위를 늘려가면서 실제 값과 비교하여 학습시키는 것을 반복하도록 훈련 시켰다. 모든 데이터는 시작일을 기준으로 0으로 환산하고, 증가분을 정규분포로 사용하였다. Table 3에서 20-09-01 값들은 LSTM에 의해 예측된 값으로, 실제 NASDAQ 20년 9월1일 값은 11057로 실제 예측된 10180.74 와 약 -11.5% 정도 차이가 난다. 따라서 Table 3에 나타난 거시 지표의 가중치를 계산하여 예측된 NASDAQ 값인 10180 값이 최대

한 11057 값에 근사한 값과 같이 될 수 있도록 지표들의 가중치를 조절하여 실제 가격에 맞추고, 각 지표들의 가중치 값도 반복 학습시켜서 각 주가 종목별 지표의 상관관계를 만들어내면 예측값의 신뢰도를 높을 수 있다.

### 3.2 텐서플로 예측 소스 코드

Table 3는 파이썬 예측 코드로 일부만 나타내었고, 입력 모델은 10으로 월단위를 10개씩 분석하여 1개를 예측하는 모델을 사용하였다.

Table 3. Part of Tensorflow Source Code by Python

```

From pandas as pd
from keras.models import Sequential
from keras.layers.core import Dropout, Activation, Dense
from keras.layers.recurrent import LSTM
...
model = Sequential()
model.add(LSTM(10, return_sequences=True, input_shape=(10, 1)))
    
```

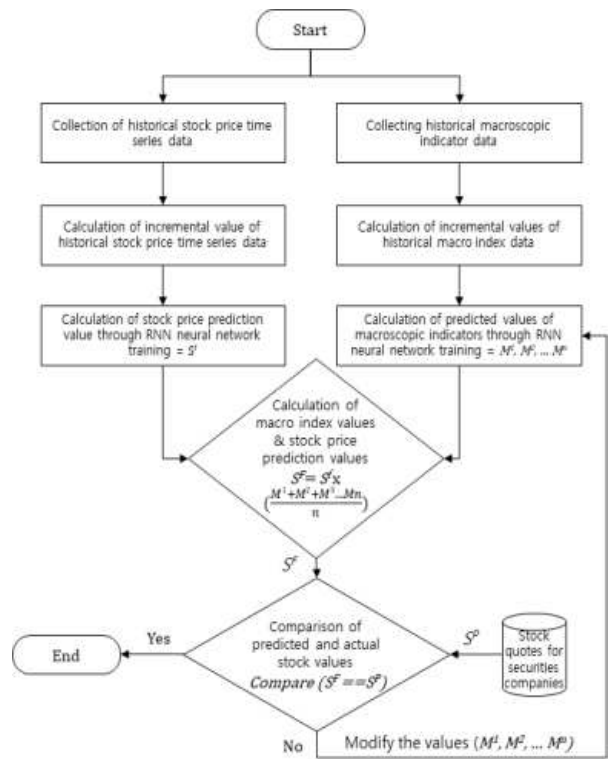


Fig. 5. The Configuration of Proposed Stock Prediction Model

Table 4. 10 Macroeconomic indicators' incremental rates with NASDAQ index

date	DFF	UNRATE	CPIAUCSL	CSCICP	PCU	HSN	CCSA	FGI	BAMLH	RAW	NASDAQ
19-04-01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19-05-01	-0.01	0.00	0.00	0.00	-0.04	-0.10	0.01	0.02	0.13	0.01	-0.02
19-06-01	-0.01	0.03	0.00	0.00	0.01	0.21	0.01	0.02	0.01	0.00	0.00
19-07-01	0.01	0.00	0.00	0.00	-0.01	-0.09	0.00	0.02	-0.08	-0.03	0.05
19-08-01	-0.12	0.00	0.00	0.00	0.00	0.07	0.00	-0.02	0.08	-0.04	-0.04
19-09-01	-0.04	-0.05	0.00	0.00	0.01	0.03	-0.01	0.00	-0.11	0.00	0.02
19-10-01	-0.10	0.03	0.00	0.00	-0.03	-0.03	0.01	-0.04	0.03	-0.01	0.00
19-11-01	-0.15	-0.03	0.00	0.00	-0.01	-0.01	0.00	0.02	-0.03	0.03	0.05
19-12-01	0.00	0.00	0.00	0.00	-0.01	0.05	0.03	-0.02	-0.10	0.00	0.03
20-01-01	0.00	0.03	0.00	0.00	0.03	0.06	0.00	-0.10	0.00	0.03	0.05
20-02-01	0.02	-0.03	0.00	0.00	0.00	-0.07	-0.02	-0.16	0.15	-0.03	0.02
20-03-01	-0.59	0.26	0.00	-0.01	0.03	-0.15	1.05	-0.08	1.44	-0.03	-0.17
20-04-01	-0.92	2.34	-0.01	-0.01	-0.14	-0.07	3.87	0.29	0.01	-0.05	0.07
20-05-01	0.02	-0.10	0.00	-0.01	0.00	0.22	0.29	-0.02	-0.11	-0.01	0.10
20-06-01	0.55	-0.17	0.01	0.00	0.13	0.20	-0.14	0.09	-0.16	0.02	0.08
20-07-01	0.19	-0.08	0.01	0.00	0.21	0.15	-0.13	0.02	-0.07	0.00	0.07
20-08-01	0.03	-0.18	0.00	0.00	0.11	0.05	-0.14	0.00	-0.11	0.03	0.07
<b>20-09-01</b>	<b>0.00</b>	<b>-0.02</b>	<b>0.00</b>	<b>0.00</b>	<b>-0.20</b>	<b>-0.16</b>	<b>0.19</b>	<b>0.00</b>	<b>0.31</b>	<b>-0.05</b>	<b>-0.09</b>

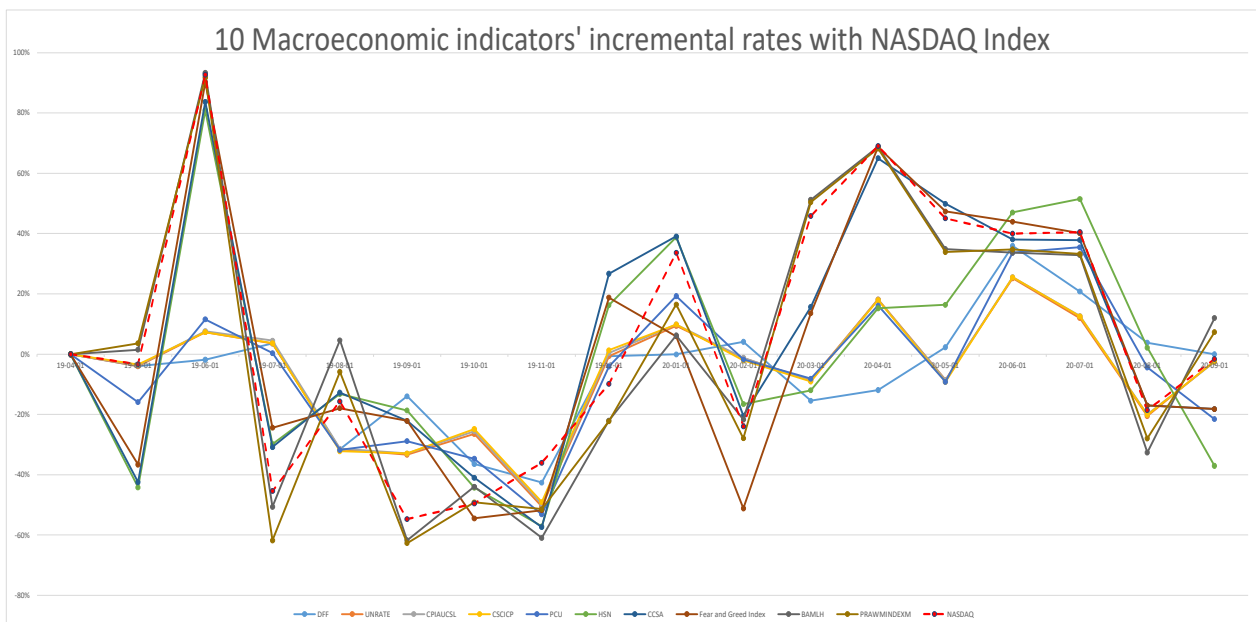


Fig. 6. 10 Macroeconomic indicators' incremental rates with NASDAQ index

#### 4. 결론

인공지능의 예측을 통하여 완벽한 랜덤인 주가를 예측하는 시도는 많이 진행되고 있으며, 기계학습과 반복되는 일정의 패턴을 찾는 것을 알고리즘이 대신하게 되면 기존의 펀드매니저의 일들을 인공지능이 대신하게 될 것으로 기대한다. 또한 사람은 감정에 사로잡혀 하락장에서 상승을 예측하기 어렵기 때문에 손실 매도하는 경우 빈

번히 발생하지만 시스템으로 감정을 배제한 채 매매를 진행하게 된다면, 진정한 시스템 트레이딩으로 인한 손실을 줄일 수 있고, 건전한 주식 투자 문화가 정착될 것으로 기대한다. 향후 연구는 지표별 상관관계를 분석하여 개별 주가에 최적화된 가중치를 계산하고, 뉴스 증시정보를 텍스트 마이닝하여 호재인지 악재인지를 판단하여 가중치를 계산하여 미래의 주식 가격을 예측하는 연구를 계속 진행할 예정이다.

## REFERENCES

- [1] Tobin, J. (1998). World economy and financial markets. *Japan and the World Economy*, 10(3), 377-379. doi:10.1016/s0922-1425(98)00038-3
- [2] Choi, J., Yoo, S., & Kim, J. (2012). Comparative Analysis of Default Risk of Construction Company during Macroeconomic Fluctuations. *Korean Journal of Construction Engineering and Management*, 13(4), 60-68. doi:10.6106/kjcem.2012.13.4.060
- [3] Choi, J., & Lee, O. (2014). Correlation Analysis Among the Price of Apartments in Seoul, Stock Market and main Economic Indicators. *Journal of Digital Convergence*, 12(2), 45-59. doi:10.14400/jdc.2014.12.2.45
- [4] Tang, J. (2015). The Effect on KOSPI 200 Futures after Launching KOSPI 200 Option. *Proceedings of the 2015 International Conference on Industrial Technology and Management Science*. doi:10.2991/itms-15.2015.347
- [5] Fred Economic Research. (2020). <https://fred.stlouisfed.org/graph/?g=w8ko> Fear and Greed Index. (2020). <https://money.cnn.com/data/fear-and-greed/>
- [6] Hong, S. (2020). A study on stock price prediction system based on text mining method using LSTM and stock market news. *The Society of Digital Policy and Management*, 18(7), 223-228. <https://doi.org/10.14400/JDC.2020.18.7.223>
- [7] Joshi, H., Verma, A., & Mishra, A. (2020). Classification of Social Signals Using Deep LSTM-based Recurrent Neural Networks. *2020 International Conference on Signal Processing and Communications (SPCOM)*. doi:10.1109/spcom50965.2020.9179516
- [8] Heesoo Hwang. (2018). Daily Stock Price Forecasting Using Deep Neural Network Model. *Journal of the Korea Convergence Society*, 9(6), 39-44.
- [9] Yongtaek Lim, Heuseok Lim. (2020). A Comparative Analysis of the Prediction Models for the Direction of Stock Price Using the Online Company Reviews. *Journal of the Korea Convergence Society*, 11(8), 165-171.
- [10] Gyeahyung Jeon, Sun-young Kwon. (2018). A Study on the Influence of Economic Factors on Library Use. *Journal of the Korea Convergence Society*, 9(11), 299-306.
- [11] Yong Jae Shin. (2018). A Study on Effects of 6th Industry types on the Korean Economy. *Journal of the Korea Convergence Society*, 9(10), 325-338.
- [12] Billah, M., Waheed, S., & Hanifa, A. (2016). Stock market prediction using an improved training algorithm of neural network. *2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*. doi:10.1109/icecte.2016.7879611
- [13] Jeon, S., Hong, B., Kim, J., & Lee, H. (2016). Stock Price Prediction based on Stock Big Data and Pattern Graph Analysis. *Proceedings of the International Conference on Internet of Things and Big Data*. doi:10.5220/0005876102230231
- [14] Ghosh, A., Bose, S., Maji, G., Debnath, N., & Sen, S. (n.d.). Stock Price Prediction Using LSTM on Indian Share Market. doi:10.29007/qgcz

홍 성 혁(Sunghyuck Hong)

[중신회원]



- 2007년 8월 : Texas Tech University, Computer Science (공학박사)
- 2012년 3월 ~ 현재 : 백석대학교 ICT 학부 부교수
- 관심분야 : 딥러닝, 블록체인, 사물인터넷 보안, 경량보안프로토콜
- E-Mail : shong@bu.ac.kr