



Journal of Knowledge Information Technology and Systems

ISSN 1975-7700

<http://www.kkits.or.kr>

A Study on Financial Time Series Data Volatility Prediction Method Using AI's LSTM Method

Han Jin Song, Heung Sik Choi*, Sun-Woong Kim, Su-Hun Oh

Department of Trading Systems Graduate School of Business-IT Kookmin University

ABSTRACT

The purpose of this study is to predict the future volatility of financial instruments more effectively by utilizing Long Short Term Memory (LSTM) which is one of the artificial intelligence (AI) techniques. This study focused on the method of predicting the future forecasts based on the historical data through the LSTM technique. The future price of a financial instrument is inversely proportional to the instrument's size of future volatility. Therefore, if future volatility can be predicted for the targeted financial instrument, the fluctuation in the instrument's price in the future can be projected more effectively. This paper predicts the future price volatility by making the historical data of a financial instrument to learn by using AI technique. As a result, the volatility movement predicted by AI was significantly similar to the actual volatility value during the same period. Through this, the results of this study are expected to be used to forecast the price of financial products more effectively. In addition, it is expected that the results of this study, together with other existing methods for volatility prediction, will further contribute to the accuracy of projecting the future volatility, and ultimately contribute to the improvement in investment returns of financial investments products including funds.

© 2019 KKITS All rights reserved

KEYWORDS: Volatility, LSTM, Forecast, Future price, Financial investment

ARTICLE INFO: Received 1 November 2019, Revised 8 November 2019, Accepted 7 December 2019.

*Corresponding author is with the Department of Trading Systems Graduate School of Business-IT, Kookmin University, 77 Jeongneung-ro, Seongbuk-gu, Seoul, 02707,

KOREA

E-mail address: hschoi@kookmin.ac.kr

1. 서론

본 연구는 금융시장에서 투자에 따른 리스크를 측정하는 방법으로 사용되는 변동성을 인공지능 기법을 사용하여 예측하는 것을 목적으로 한다.

일반적으로 특정 투자자산의 가격은 변동성과 반비례하는 성향이 있다. 특정 투자자산의 가격이 상승하면 그 자산의 위험의 척도인 변동성은 축소되고, 그 자산의 가격이 하락하면 반대로 변동성은 확대되는 경향이 있다. 따라서 주식, 채권 등 특정 투자자산의 변동성을 예측할 수 있다면 그 투자대상 자산의 미래 가격의 방향성과 그 변동폭을 용이하게 예측할 수 있게 된다[1].

현재 증권시장에서 투자위험을 대표하는 변동성은 파생상품 옵션의 가격을 기초로 하여 향후 30일간의 변동성을 예상하는 VIX 지수(Volatility Index)를 가장 보편적으로 사용하고 있다[2]. VIX 지수는 시카고 옵션 거래소에 상장된 S&P500 지수옵션의 변동성을 예측하는 지수로 VIX 지수가 최고치에 이르면 투자자들의 불안심리가 최고치에 도달했다고 판단하여 ‘공포지수’라고도 불린다[2].

본 논문은 증권시장에서 거래되는 특정 자산의 변동성 예측에서 기존의 VIX 지수에 추가하여 인공지능을 이용한 변동성 예측 방법을 제시하고 시물레이션을 통하여 그 결과를 확인하였다. 2장에서는 사용하고자 하는 인공지능 기법인 LSTM(Long Short Term Memory)에 대한 이론적 배경과 변동성이 자산가격과 어떤 관계인지를 고찰하고자 하며, 3장에서는 본 연구의 방법과 절차를 제시하였다. 즉, 데이터와 인공지능 학습과정을 설명하고자 하였다. 4장에서는 인공지능을 이용한 각 자산별 변동성 시물레이션 결과를 제시하였다. 학습 대상 기간과 검증 기간으로 구분하여 검증 기간의 변동성 예측 결과 값을 그 기간동안 발생했던 실제 변동성과 비교하여 인공지능을 통한 변동성 예측값의

정확성을 검증하였다. 마지막 5장에서는 본 연구결과와 현실 투자 및 펀드 운용에서의 활용방안에 대해 기술하였다.

2. 이론적 배경

2.1 LSTM(Long-Short Term Memory)

순환신경망(Recurrent Neural Networks)은 인공신경망 모형 중 하나로 데이터의 순서와 관계없이 학습을 하는 일반 신경망에 시계열 개념을 도입한 것이다.

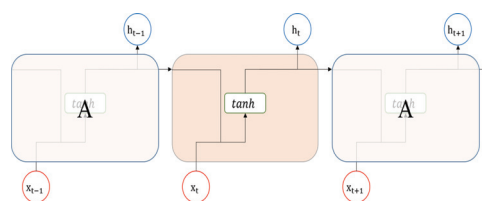


그림 1. RNN 구조
Figure 1. RNN Structure

순환신경망의 구조를 시각적으로 표현한 <그림 1>에서 순환신경망은 시계열 데이터를 입력받아 값을 출력하고 출력 결과가 다음 단계 학습에 영향을 주도록 한다. 하지만 학습 대상인 데이터의 시퀀스(Sequence)가 길어질수록 어느 시점 이후로는 더이상 학습이 진행되지 않는 기울기 소실(Vanishing gradient)의 한계가 존재한다[3-4].

LSTM은 순환신경망의 이러한 장기의존성(Long-Term dependencies) 한계점을 돌파하기 위해 Hochreiter,S., & Schmidhuber,J.(1997)에 의해 소개되었다. LSTM은 셀 상태가 전체 체인을 관통하며 계속 유지되고 순환신경망과 다르게 입력 게이트, 망각 게이트, 출력 게이트를 추가로 이용하여 출력 값을 조정한다[3-5].

$$i_t = \sigma(W_{iz}x_t + W_{ih}h_{t-1} + W_{ic}c_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{fz}x_t + W_{fh}h_{t-1} + W_{fc}c_{t-1} + b_f) \quad (2)$$

$$c_t = f_t c_{t-1} + i_t \text{Tanh}(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (3)$$

$$O_t = \sigma(W_{oz}x_t + W_{oh}h_{t-1} + W_{oc}c_t + b_o) \quad (4)$$

$$h_t = o_t \text{Tanh}(c_t) \quad (5)$$

LSTM의 구조를 나타낸 <그림 2>와 같이 입력 게이트(Input Gate)는 식 1에 의해 입력 데이터를 얼마나 받아들일지 결정하고, 망각 게이트(Forget Gate)는 식 2에 의해 이전의 셀 상태를 얼마나 유지할지 결정하고, 출력 게이트는 식 3에 의해 출력 데이터의 값을 결정한다. 출력 게이트에 해당하는 식 3은 기존 셀 상태에 망각 게이트에서 나온 값을 곱하여 셀 상태값의 일정부분의 값을 망각하고(식 4), 입력 값과 그 이전 단계의 출력 값을 처리한 결과에 입력 게이트의 출력 값을 곱하여 일정량을 유지하며 새로운 셀 상태를 만들어낸다. 또, 식 5를 통해 만들어진 새로운 셀 상태에 출력 게이트의 출력 값을 곱하여 셀의 출력을 결정한다. 위 과정을 통해 LSTM은 기존 순환신경망보다 시퀀스가 긴 시계열 데이터에 더 원활한 학습을 수행할 수 있다[3][6][7].

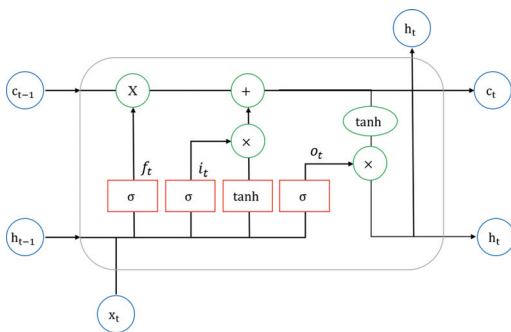


그림 2. LSTM 셀 구조
Figure 2. LSTM Cell Structure

위 그림 <그림 2>에서 c_t 와 h_t 는 각각 셀 상태와

셀의 출력을 나타내고, 시그마와 Tanh는 활성화 함수 Sigmoid와 하이퍼탄젠트(Hyperbolic Tangent)를 의미한다.

2.2 변동성(Volatility)

변동성이란 자산의 가격, 가치가 시간이 흐름에 따라 변화하는 정도를 의미한다. 이는 곧 그 자산의 가격 변동에 대한 위험을 의미한다. 자산의 변동성이 크면 자산의 가격의 등락폭이 상대적으로 크고, 자산의 변동성이 작으면 상대적으로 가격의 등락폭이 작다. 안정되고 성숙된 시장일수록 변동성이 상대적으로 작고 신흥시장이 상대적으로 변동성이 크다[8].

증권시장에서 일반적인 의미의 주가변동성이란 수익률((금일종가-전일종가)/전일종가)의 일정 기간 표준편차를 계산한 것으로 본 연구에서는 30일간의 수익률로 표준편차를 계산하였다.

표 1. 변동성과 주요 지수 간 상관관계
Table 1. Correlation between Volatility and Major Index

	KOSPI	S&P500	FTSE100	CAC40
Correlation	-0.45177	-0.32263	-0.39186	-0.45368

변동성과 주가 간의 상관관계 분석을 위해 KOSPI(한국), S&P500(미국), FTSE100(영국), CAC40(프랑스) 네 가지의 지수의 2003년 9월 29일부터 2019년 10월 11일까지의 종가 데이터를 이용하여 주가와 변동성과의 상관계수를 구하였다.

<표 1>과 같이 각 지수별 상관계수의 차이는 존재하나 변동성과 주가 사이에는 음의 상관관계가 존재한다는 것을 알 수 있다.

3. 연구 방법 및 절차

3.1 데이터

본 연구의 실험은 <표 2>와 같이 한국, 미국, 일본, 영국, 프랑스, 독일 6개 국가의 주요 지수인 KOSPI, S&P 500, NIKKEI 225, FTSE 100, CAC 40, DAX와 펀드 구성이 오직 채권으로만 구성된 펀드인 iShares Core U.S. Aggregate Bond ETF를 대상으로 하여 진행하였다.

표 2. 데이터 목록
Table 2. Data List

Asset Class	Code	Explanation
Equity	KPP	KOSPI
	SPP	S&P 500
	NIK	NIKKEI 225
	ENG	FTSE 100
	FRC	CAC 40
	GRM	DAX
Fixed Income	BND	iShares Core U.S. Aggregate Bond ETF

표 3. LSTM 입력 변수 설명
Table 3. Explanation of LSTM Factors

Factor	Explanation
Moving Average	A calculation to analyze price points by creating a series of averages of specified time frame of the full time frame.
Momentum	The indicator that shows where the current price is in compare to where the price was in the past.
GARCH(1,1)	A statistical model used in analyzing time series data that describes the variance of the current error term as a function of the actual sizes of the previous time periods' error terms.
Standard deviation	A quantity calculated to indicate the variation or dispersion of a set of values.

본 연구는 총 7개의 종목을 대상으로 2003년 9월 29일부터 2019년 10월 11일 까지의 증가를 기본 데이터로 설정하고 이를 이용하여 각 종목의 증가

30일 이동평균, 30일 모멘텀, GARCH(1,1)모형 결과 값, 표준편차를 LSTM의 입력 변수로 설정하였다. 입력변수는 유사한 연구[9-11]에서 사용되고 있으며 각 입력변수에 대한 설명은 <표 3>에 정리하였다[12-15].

3.2 LSTM 학습 환경

본 연구의 실험 진행을 위해 프로그램 언어는 Python 3.8을 사용하였다. LSTM 학습 진행을 위해 Python 3.8 환경에서 사용 가능한 라이브러리 Tensorflow 2.0과 LSTM 이용을 위한 Keras를 사용하였고, Jupiter Lab에서 진행하였다. LSTM 학습을 위한 데이터 전처리 과정에서는 Visual Basic for Application(Excel)을 이용하였다.

3.3 LSTM 학습 진행 방법

본 연구는 LSTM의 입력 변수로 설정한 다섯 가지의 데이터를 이용하여 30일 후의 변동성(표준편차)을 예측하는 것을 목표로 한다. 전처리 후의 데이터인 2003년 10월 28일부터 2019년 9월 12일까지의 5,799개의 데이터 중 5,445개의 데이터로 학습과 검증을 수행하고 실제 예측은 354개의 데이터를 대상으로 수행한다.

4. 변동성 예측 결과 분석

본 연구에서 학습과 예측 수행 정확도를 평가하는 방법으로는 평균 제곱근 오차인 RMSE(Root Mean Square Error)를 사용한다. RMSE는 수치예측 모델 평가 방법으로 예측 값에서 실제 값을 차감한 값의 제곱 평균을 구한 후에 제곱근하는 방식이다. 이를 통해 모델 간의 잔차를 비교해 볼 수 있다.

RMSE 수치는 평가에 있어 절대적인 기준이 존재하지 않고 수치가 상대적으로 크고 작음을 통해 결과를 확인하는 것으로 RMSE 수치만으로는 예측 결과의 정확도를 판단하는 데 있어 한계점이 존재한다. 따라서 본 연구에서는 예측 결과를 분석함에 있어 RMSE 수치만 확인하는 것이 아니라 <그림 3>등의 그래프를 통해 실제 값과 예측 값 사이에 얼마만큼 오차가 존재하는지 확인하였다.

5,799개의 데이터 중 5,445개의 데이터를 대상으로 학습을 수행하고 354개의 데이터를 대상으로 예측을 수행한 결과는 Table 4와 같다. 단순 RMSE 수치 상으로는 iShares Core U.S. Aggregate Bond ETF가 Training RMSE가 0.00017344, Test RMSE가 0.00001101로 가장 좋은 결과가 나왔지만, <그림 9>를 확인하면 채권의 변동성이 지수의 변동성보다 상대적으로 작기 때문에 RMSE 값도 상대적으로 작게 나온 것으로 볼 수 있다.

표 4. RMSE 결과
Table 4. RMSE Results

Category	KPP	SPP	NIK	ENG	FRC	GRM	BND
Training RMSE	0.00087452	0.00088514	0.00089188	0.00106904	0.00140376	0.00126567	0.00017344
Test RMSE	0.00025656	0.00025656	0.00029719	0.00029701	0.00035366	0.00036733	0.00001101

채권을 제외한 6개 국가의 지수를 대상으로 한 예측 RMSE 값 중에는 KOSPI 예측 RMSE가 Training RMSE가 0.00087452, Test RMSE가 0.00025656으로 가장 작은 값이지만 변동 폭을 고려하여 각 지수별 테스트 결과를 시각적으로 나타낸 <그림 3>, <그림 4>, <그림 5>, <그림 6>, <그림 7>, <그림 8>의 그래프로 분석한 결과 변동성의 범위가 상대적으로 차이가 나기 때문에 이를 고려한다면 KOSPI 예측이 6개 지수 중에서는 실제 데이터와 예측 데이터가 상대적으로 맞지 않는 것을

확인할 수 있다. KOSPI를 제외한 나머지 5개 지수의 경우 RMSE값과 그래프분석을 통해 예측모형 성능이 우수함을 확인하였다.

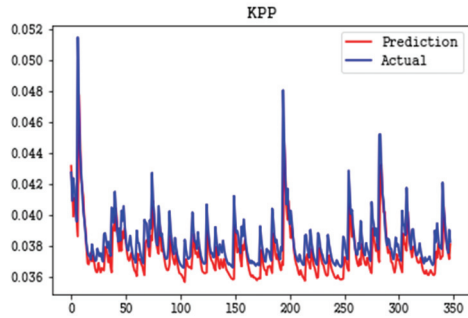


그림 3. KOSPI 테스트 결과
Figure 3. KOSPI Test Result

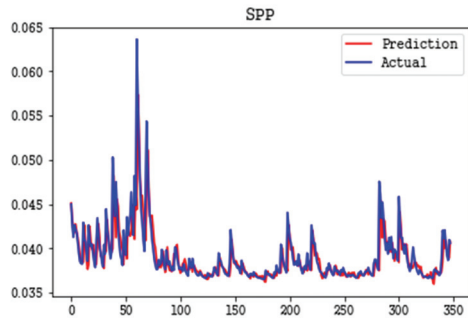


그림 4. S&P 500 테스트 결과
Figure 4. S&P 500 Test Result

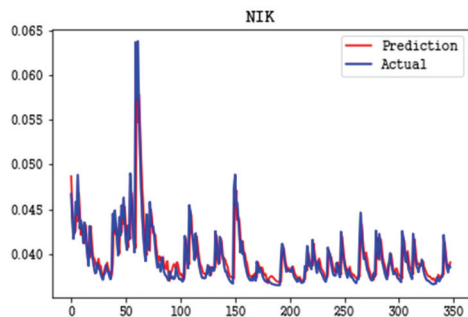


그림 5. NIKKEI 테스트 결과
Figure 5. NIKKEI 225 Test Result

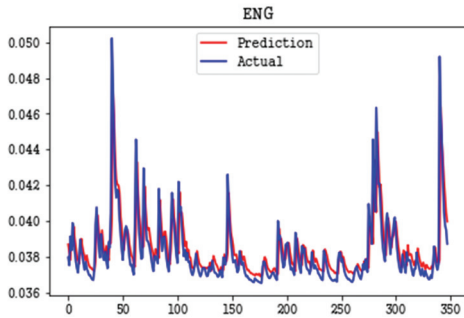


그림 6. FTSE 100 테스트 결과
Figure 6. FTSE 100 Test Result

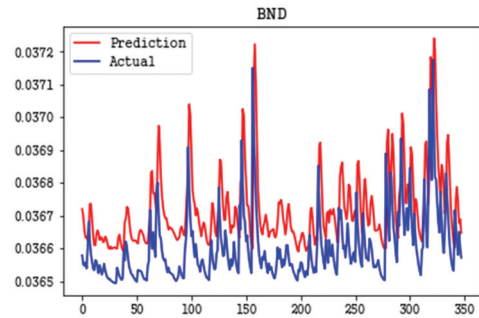


그림 9. iShares Core U.S. Aggregate Bond ETF 테스트 결과
Figure 9. iShares Core U.S. Aggregate Bond ETF Test Result

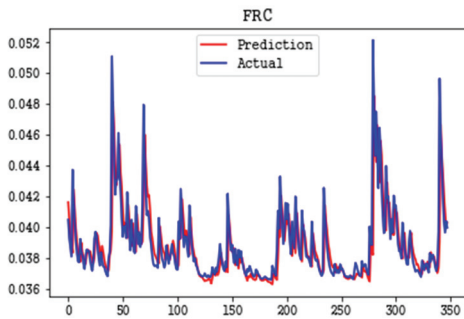


그림 7. CAC 40 테스트 결과
Figure 7. CAC 40 Test Result

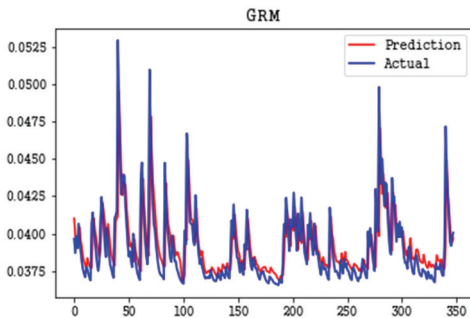


그림 8. DAX 테스트 결과
Figure 8. DAX Test Result

5. 결론

본 연구는 금융시장에서 개별 종목에 대한 투자뿐 아니라 포트폴리오, 즉 펀드를 운용하는데 리스크의 중요한 척도가 되는 변동성을 인공지능을 활용하여 예측하는 것이 목표이다. 현재 증권시장에서 투자위험을 대표하는 변동성은 VIX 지수를 가장 많이 사용하고 있다. 그러나 증권시장에서 변동성을 측정할 때 VIX지수 뿐만 아니라 또 다른 기법을 통한 변동성 측정값을 같이 사용한다면 증권시장의 투자위험을 예측하고 그에 따른 투자전략을 수행하는 데 보다 효과적인 것이다.

본 연구는 과거 실제가격과 그 가격이 변동했던 데이터를 바탕으로 AI기법을 활용하여 변동성을 예측하고자 하였으며, 그 결과 인공지능(LSTM 기법)을 이용하여 금융자산의 미래 변동성을 예측하는 것이 상대적으로 유용하다는 결론을 각 자산별 시뮬레이션을 통하여 제시하였다. 즉, LSTM 기법을 활용하여 특정 금융상품의 과거 데이터(변동성 크기와 가격)를 인공지능으로 학습 시켜서 미래 전망치를 예측하였으며, 그 결과 인공지능이 예측한 검증기간의 변동성 움직임이 그 기간의 실제 변동성 값과 매우 유사한 결과를 도출함으로써 그 유용성

을 검증하였다.

본 연구 결과를 기존의 금융자산 투자와 펀드 운용에 활용한다면 일반적인 주식형 펀드는 물론 TDF, FoF 등 자산배분형 펀드의 운용에서 미래 특정기간에 대한 리스크를 보다 합리적인 방법으로 예측하고 이에 맞는 투자전략 수립과 효율적 수행을 가능하게 하여, 궁극적으로는 금융자산을 포함한 펀드 투자의 수익률 개선에 기여할 수 있을 것으로 본다.

References

- [1] C. H. Won, *Predicting stock index volatility using LSTM and GARCH-type models*, Master's thesis, Ajou University, pp. 1-16, 2019.
- [2] J-B. Moon, *A study on VKOSPI volatility trading strategy using VIX index*, Master's thesis, Kookmin University, pp. 16-23, 2016.
- [3] J. Ilonen, J-K. Kamarainen, J. Lampinen, *Differential evolution training algorithm for feed-forward neural networks*, Neural Processing Letters, 17, pp. 93-105, 2003.
- [4] W. Jang, *A study to solve trailing problem in the financial time series forecasting model using LSTM*, Master's thesis, The Catholic University of Korea, pp. 10-19, 2017.
- [5] Y-M. Seo, *Analysis of prediction accuracy of fine dust concentration for Seoul region using LSTM model*, Master's thesis, Dept. of Geoinformation Engineering The Graduate School Sejong University, pp. 12-21, 2019.
- [6] J-H. Lee, *A comparative study on stock price forecasting models using LSTM and bidirectional neural networks*, Master's thesis, Seoul National University of Science and Technology, pp. 5-15, 2019.
- [7] G. Lee, *A study on the lead/lag relationship among VIX, VIX futures, and S&P500*, Master's thesis, Pusan National University, pp. 5-16, 2015.
- [8] C-B. Lee, *An empirical study of the Korean stock return volatility using the GARCH models*, Master's thesis, Pusan National University, pp.5-21, 1999.
- [9] S. H. Park, *The weather effect on stock returns, volatility, and trading behavior*, Master's thesis, Sungkyunkwan University, pp. 5-10, 2015.
- [10] Y. J. Yoon, *A study on relation between bond yields and equity volatility*, Master's thesis, Pusan National University, pp. 11-22, 2014.
- [11] B. G. Park, *The volatility analysis of the Korea stock market by using regime switching model*, Master's thesis, Department of Business Administration Graduate School of Soongsil University, pp. 7-14, 2014.
- [12] C. Kim, *A study on dynamic asset allocation using momentum and development of ETF asset allocation system*, Master's thesis, Kookmin University, pp. 4-6, 2018.
- [13] T. Bollerslev, *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, Vol. 31, No. 3, pp. 307-327, 1986.
- [14] Y-W. Choo, *A study on the comparison of forecasting volatility of stock return between the GARCH model and the stochastic volatility model*, Master's thesis, Yonsei University, pp. 5-10, 2000.
- [15] S-P. Hong, *Performance and fitness of risk forecasting about characteristics of stock*

with GARCH and VaR, Master's thesis, Chonbuk National University, pp. 6-15, 2016.

AI의 LSTM기법을 이용한 금융시계열 데이터 변동성 예측방법 연구

송한진¹, 최홍식², 김선웅², 오수훈³

¹국민대학교 비즈니스IT전문대학원 박사과정

²국민대학교 비즈니스IT전문대학원 교수

³국민대학교 비즈니스IT전문대학원 석사과정

요 약

본 연구의 목적은 인공지능 기법 중 하나인 LSTM(Long Short Term Memory)을 사용하여 금융상품들에 대한 미래 변동성을 보다 효과적으로 예측하고자 하는 것이다. 기존의 주식, 채권 등 금융상품에 대한 미래값 예측은 기본적 분석과 기술적 분석을 통하여 이루어져 왔다. 그 중에서 본 연구는 기술적 분석의 한 방법으로 과거 데이터를 기초로 한 미래 전망치를 예측하는 방법에 집중하였다. LSTM 기법을 통해 보다 미래 전망치의 예측률을 높이는 것을 목적으로 하였다. 금융상품의 미래가격 전망치는 그 상품의 미래 변동성의 크기에 반비례한다. 즉, 내일의 주가변동성이 커지면 내일의 주가는 하락할 가능성이 크다는 것이다. 따라서 대상 금융상품에 대한 미래의 변동성을 예측할 수 있다면 그 상품의 미래가격 등락을 보다 효과적으로 전망할 수 있다. 본 논문은 LSTM 기법을 활용하여 어떤 금융상품의 과거데이터(변동성 크기와 가격)를 인공지능으로 학습시켜서 미래의 가격변동성을 예측하였으며, 그 결과 인공지능이 예측한 검증 기간의 변동성 움직임이 동 기간의 실제 변동성 값과 매우 유사한 결과를 도출했다. 이를 통해 본 연구결과를 활용하여 향후 금융상품(주식, 채권 등)에 대한 미래 가격전망을 보다 효과적으로 할 수 있을 것으로 본다. 또한 기존의 다른 변동성 예측방법과 더불어서 본 연구결과를 활용할 경우 미래의 변동성 예측 정확성에 보다 기여할 수 있을 것으로 기대하며, 궁극적으로는 펀드를 포함한 금융투자 상품의 투자수익률 개선에 기여할 것으로 기대한다.



Han Jin Song received the bachelor's degree in the Department of Business administration and the M.S. degree in the Department of Economics from the So-Gang University in 1990 and 1997. He is attending the Ph.D. degree in the Department of Trading-Systems Graduate School of Business-IT, Kookmin University. From 1990 to this day, he is also working in Asset Management industry for 30 years. He has many experiences in the investment and finance business including the fund management. He serves as an adjunct professor in the Department of Economics of the Soong-Sil University from 2017 up to now. He is a regular member of the KKITS.

E-mail address: hanjin.song@shbnpam.com



Heung Sik Choi received the M.S. degree in the Department of Management Science from KAIST and received Ph.D. degree in the Computers and Information Systems from the University of Rochester. He has been a professor in the School of Business and Graduate School of Business Information Technology at Kookmin University since 1995. His current research interests include options volatility trading, system trading and global asset allocation.

E-mail address: hschoi@kookmin.ac.kr

Sun-Woong Kim received the bachelor's degree in the Department of Business Administration from the Seoul National University in 1981. He received the M.S. degree and the Ph.D. degree in the Department of



Management Science from KAIST in 1983 and 1988. He worked as a fund manager in the investment management industry for more than 20 years. He has been a professor in the Graduate School of Business Information Technology at Kookmin University since 2009. His current research interests include strategy trading and asset management.

E-mail address: swkim@kookmin.ac.kr



Su-Hun Oh received the bachelor's degree in the Department of Management Information System from the Kookmin University in 2019. He is attending the Master's degree in the Department of Trading- Systems Graduate School of Business-IT, Kookmin University. He is a regular member of the KKITS.

E-mail address: holdrang@naver.com