

신경회로망을 이용한 KOSPI 예측 기반의 ETF 매매

황희수

한라대학교 전기전자공학과 교수

ETF Trading Based on Daily KOSPI Forecasting Using Neural Networks

Heesoo Hwang

Department of Electrical and Electronic Engineering, Halla University Professor

요 약 신경회로망은 적합한 수학적 모델에 대한 가정 없이 데이터로부터 유용한 정보를 추출해서 예측에 필요한 입력력 관계를 정의할 수 있어서 주가 예측에 널리 사용되어 왔다. 본 논문에서는 신경회로망 모델을 사용하여 일별 KOrea composite Stock Price Index (KOSPI) 증가를 예측한다. 예측된 증가를 기반으로 KOSPI에 연동해 변동하는 Exchange Traded Funds (ETFs)의 거래를 위한 알파 매매를 제안한다. 본 논문에 제안된 방법으로 KOSPI 예측 신경회로망 모델들을 구현하고 예측 정확도를 평가한다. 구현된 신경회로망 모델(NN1)의 학습 오차(MAPE)는 0.427, 평가 오차는 0.627이다. 평가용 데이터를 사용해 알파 매매를 시뮬레이션하면 수익률은 7.16 ~ 15.29 %를 보인다. 이는 125 거래일 데이터로 거둔 수익률로 제안된 알파 매매가 효과적임을 보인다.

주제어 : 융합, 주가 예측, 주가 모델링, 신경회로망, ETF 매매, 알파 매매, 시계열 예측

Abstract The application of neural networks to stock forecasting has received a great deal of attention because no assumption about a suitable mathematical model has to be made prior to forecasting and they are capable of extracting useful information from data, which is required to describe nonlinear input-output relations of stock forecasting. The paper builds neural network models to forecast daily KOrea composite Stock Price Index (KOSPI), and their performance is demonstrated. MAPEs of NN1 model show 0.427 and 0.627 in its learning and test, respectively. Based on the predicted KOSPI price, the paper proposes an alpha trading for trades in Exchange Traded Funds (ETFs) that fluctuate with the KOSPI200. The alpha trading is tested with data from 125 trade days, and its trade return of 7.16 ~ 15.29 % suggests that the proposed alpha trading is effective.

Key Words : Convergence, Stock Price Forecasting, Stock Price Modeling, Neural Network, ETF, Trading, Alpha Trading, Time Series Forecasting

1. 서론

주식시장에서 종합주가지수나 종목별 지수를 예측하는 것은 잡음과 변동성으로 인해 매우 어려운 일로 인식되고 있지만 그럼에도 주가 예측 모델을 만들기 위한 다양한 방법들이 시도되어 왔다[1-11]. 주가의 움직임을 정

확하게 예측하기 위한 방법은 여전히 해결되지 않은 문제로 퍼지 모델[1], SVM (Support Vector Machine)[2,11], 전문가시스템[3], 인공신경망[4-8]과 같은 방법이 금융 시계열을 예측하는데 널리 사용되었고 예측의 정확도를 높여 왔다. 최근에는 주가 예측에 딥러닝을 적용하는 연구로 일별 주가 지수 예측[5], Recurrent Neural Network

*Corresponding Author : Heesoo Hwang (hshwang@halla.ac.kr)

Received November 2, 2018

Accepted January 20, 2019

Revised December 12, 2018

Published January 28, 2019

(RNN)을 사용한 구글 주가 예측[6] 및 Convolutional Neural Network (CNN)으로 주가에 미치는 이벤트 예측[7]이 있다. 최근에는 종합주가지수를 예측하고 이를 기반으로 매매 전략을 탐색하는 연구가 활발하다. 이를 위해 전통적인 기업 지표를 이용하는 방법[9], 신경회로망으로 매매 신호를 찾는 방법[10]과 기계학습으로 Exchange Traded Funds (ETFs)의 방향을 예측하는 연구[11]가 있다.

본 연구는 신경회로망으로 일별 KOSPI 증가를 예측하고 이를 기반으로 KOSPI200 현물 및 선물에 연동된 ETF를 매매하는 전략을 제안한다. 내일 KOSPI 지수 증가가 일정 비율 이상 상승하거나 하락하는 것으로 예측되면 각각 KOSPI 레버리지 ETF와 KOSPI 인버스 ETF를 매수하고 다음 날 마감 전에 매도한다. 증가 예측에는 하나 또는 둘 이상의 은닉층을 갖는 신경회로망을 사용한다. ETF의 매수를 결정하는 예측 증가의 상승 또는 하락 비율(알과 값)은 신경회로망 모델의 오차와 시뮬레이션을 통한 ETF 매매 수익률로부터 최적하게 찾아진다. 제안된 신경회로망 및 ETF 알과 매매는 검증용 데이터를 통해 타당성이 분석된다. 2장은 일별 주가 예측 모델을, 3장은 주가 연동 ETF 매매를, 4장은 주가지수 예측과 ETF 매매의 시뮬레이션 및 성능 평가를, 5장은 결론을 기술한다.

2. 일별 주가 예측 신경망 모델

2.1 신경회로망 모델

1일 앞선 주가지수의 예측을 위해 본 논문에서는 Fig.1의 신경회로망을 사용한다. Fig.1의 신경회로망에서 중간층 및 출력층 뉴런의 전달함수로 logistic sigmoid 함수를 사용하며, 신경회로망의 데이터 흐름은 두 단계로 진행된다.

[단계 1] 전처리로 신경회로망 학습에 사용할 입·출력 데이터를 [-1, 1] 범위로 정규화 한다.

[단계 2] 정규화 된 입력 데이터는 하나 또는 두 개의 은닉층(hidden layer)을 거쳐 출력층에서 주가지수의 증가를 출력한다.

정규화된 입력(In)과 출력(KOSPI 증가) 데이터로 학습이 이루어진다. 학습에는 scaled conjugate gradient descent 역전파 방식을[15] 사용한다.

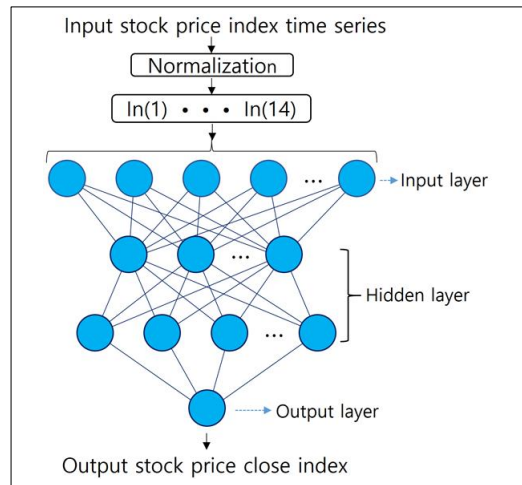


Fig. 1. Architecture of the neural network

2.2 입력 변수

주식 시장은 매우 복잡하고 변동성이 크기 때문에 다양한 기계학습을 적용한 예측 모델이 개발되고 있는데, 이들 모델은 재무 지표, 경제 요인들이나 과거의 주가 움직임 데이터를 사용하였다[1-4,9,10,11]. 또한 동일 모델이라도 입력 변수를 달리 선택하면 그 결과가 달라진다[14]. 주가는 경제 및 정치 등의 많은 요인에 의해 영향을 받는데, 이들 요인 모두를 예측 모델에 반영하는 것은 불가능하다.

본 논문에서는 주가의 움직임 자체에만 주목한다. 즉, 과거 주가의 움직임과 유사한 상황이 되면 주가는 동일한 방식으로 움직인다고 가정한다. 이는 캔들스틱(candlestick) 차트는 시간과 분야에 독립적이어서 과거의 주가 변동 패턴이 그 때와 유사한 상황이 되면 현재나 미래에 반복될 수 있다는 연구[12,13]를 적용한 것이다. 연구결과에 의하면 의미가 있는 캔들스틱 패턴의 패턴의 99.8 %는 단일 및 이중 캔들스틱이기[12]

Table 1. Input variables of the neural network model

KOSPI	Input variables	
	2 days ahead	1 day ahead
High price - Close price	ln(1)	ln(8)
Low price - Close price	ln(2)	ln(9)
Close price	ln(3)	ln(10)
3 day moving average	ln(4)	ln(11)
5 day moving average	ln(5)	ln(12)
10 day moving average	ln(6)	ln(13)
20 day moving average	ln(7)	ln(14)

때문에 본 논문에서는 Table 1에 보인대로 예측 1일 전과 2일 전의 14 개의 변수를 예측 모델의 입력으로 사용한다. 증가와 함께 고가와 저가에서 증가를 뺀 값과 3일, 5일, 10일 및 20일 증가 이동평균을 변수로 사용한다.

3. 주가 연동 ETF 매매

본 논문에서는 매매 대상 ETF로 Tiger레버리지 ETF와 Tiger인버스 ETF를 고려한다. Tiger레버리지는 KOSPI200 지수를 기초지수로 하여 1좌당 순자산가치의 일간변동률을 기초지수의 일간변동률의 2배수로 연동하는 ETF이고 Tiger인버스는 1좌당 순자산가치의 일간변동률을 기초지수인 F-KOSPI200 지수 일간변동률의 음의 1배수로 연동하는 ETF이다. 이는 상승장에서 Tiger레버리지 ETF의 매매로 변동율의 2배에 해당하는 수익을, 하락장에서 Tiger인버스 ETF의 매매로 변동율의 1배에 해당하는 수익을 낼 수 있다.

2015년 10월 2일부터 2018년 4월13일까지 KOSPI, KOSPI200 및 F-KOSPI200 지수 데이터를 사용해 이들의 상관관계를 알아보면 Table 2와 같다. KOSPI와 KOSPI200 그리고 KOSPI와 F-KOSPI200 지수의 상관관계가 매우 높고 이들 지수의 전일대비 변동률 (%)간의 상관관계도 높음을 알 수 있다. 따라서 본 논문에서는 KOSPI200과 F-KOSPI200 지수 대신에 KOSPI 지수를 예측하고 이를 토대로 Tiger레버리지 및 Tiger인버스 ETF를 매매한다.

Table 2. The correlation coefficients of KOSPI, KOSPI200 and F-KOSPI200

KOSPI Index	Coefficient
KOSPI and KOSPI200	0.99324
KOSPI and F-KOSPI200	0.99324
KOSPI (%) and KOSPI200 (%)	0.98398
KOSPI (%) and F-KOSPI200 (%)	0.95025

KOSPI 지수 예측에 의한 ETF 매매는 KOSPI와 KOSPI200과 같은 종합주가지수가 장마감 시간 근방(5분 이내)에서 크게 변하지 않는다는 가정에 기초한다. 장마감 5분 전에, 현재 KOSPI 시가를 증가로 간주해서 내일 KOSPI 증가를 예측한다. 예측된 증가에 의한 변동율이 ETF 매매 수수료율 보다 크면 ETF를 매매한다. 즉 예측

된 증가 변동율이 Tiger레버리지 ETF 매매 수수료율보다 크면 Tiger레버리지 ETF를, 음의 Tiger인버스 ETF 매매 수수료율 보다 작으면 Tiger인버스 ETF를 매수한다. 그러나 KOSPI 지수 예측은 정확하지 않으며 오차를 포함하고 있기 때문에 상승 예측에도 실제 하락하거나 하락 예측에도 상승하여 ETF 매매 손실이 발생할 수 있다. 이런 예측 오차로 인한 매매 손실을 줄이기 위해 다음과 같은 알파 매매를 제안한다. 상승 예측에서 그 값이 식(1)을 만족하면 레버리지 ETF를 매수하고, 하락 예측에서 그 값이 식(2)를 만족하면 인버스 ETF를 매수한다. 여기서 α (%)는 예측 오차로 인한 매매 손실을 줄이기 위한 구간으로 식(3)을 만족해야 한다.

$$\frac{KOSPI_{(t+1)} - KOSPI_{(t)}}{KOSPI_{(t)}} \times 100 \geq \alpha \quad (1)$$

여기서 $KOSPI_{(t+1)}$ 는 예측 값이고 $KOSPI_{(t)}$ 는 예측 전일 증가이다.

$$\frac{KOSPI_{(t+1)} - KOSPI_{(t)}}{KOSPI_{(t)}} \times 100 \leq -\alpha \quad (2)$$

$$\alpha > \max(F_{Tiger-Leverage}, F_{Tiger-Inverse}) \quad (3)$$

$F_{Tiger-Leverage}$ 와 $F_{Tiger-Inverse}$ 는 각각 Tiger레버리지 ETF와 Tiger인버스 ETF의 매매 수수료율이다.

직전일 대비 예측된 증가의 변화율로 알파 매매 구간을 표시하면 Fig. 2와 같다. 실전에서는 장마감 전에 ETF의 매매를 위해서 $KOSPI_{(t)}$ 로 예측 시점(예, 장마감 5분전)의 KOSPI 시가를 사용한다. α 는 예측 모델의 오차와 알파 매매의 수익률을 고려해서 결정된다.

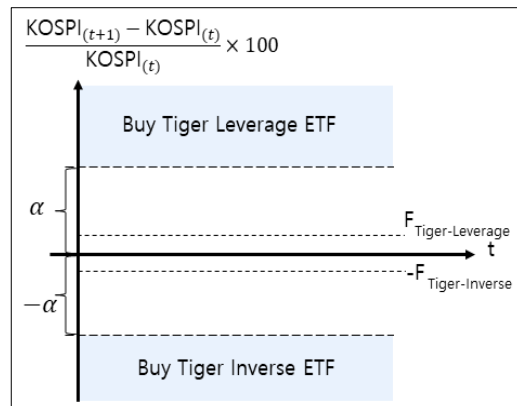


Fig. 2. Trading interval for buying leverage and inverse ETFs

4. 주가 예측에 의한 ETF 매매

신경회로망 모델은 입력으로 Table 1에 보인 1일 및 2일 전의 주가지수 데이터를 사용하여 다음 날의 주가지수 증가를 예측한다. 2015년 11월 20일부터 2018년 5월 2일까지 총 600 거래일의 KOSPI 데이터에서 신경회로망 모델의 학습에 450 거래일 데이터를, 학습된 모델의 평가에 나머지 150 거래일 데이터를 사용한다. 신경회로망 모델의 최대 학습 횟수는 100이다. 본 논문에서 제안한 신경회로망으로 구현된 주가지수 예측 모델(NN1, NN2, NN3)의 매개변수는 반복 학습을 통해 얻은 최적화된 값으로 Table 3과 같다.

Table 3. Neural network parameters for the KOSPI prediction models

Model	i	n_{h1}	n_{h2}
NN1	14	5	
NN2	27	3	5
NN3	74	2	7

i : The number of iteration in the learning
 n_{h1} : The number of neurons in the 1st hidden layer
 n_{h2} : The number of neurons in the 2nd hidden layer
 NN: Neural Network

Table 4는 본 논문의 예측 모델(NN1, NN2, NN3)의 Mean Absolute Percent Error (MAPE)를 비교한 것이다. 세 모델의 오차 차이가 학습에서는 최대 0.063, 평가에서는 최대 0.020으로 매우 작음을 볼 수 있다. 기존의 KOSPI 예측 연구 결과를 보면 퍼지 모델 예측[1]은 학습 오차 1.0390, 평가 오차 1.0630이고 심층 신경회로망 모델 [5]은 학습 오차가 0.6412과 0.6896이고 평가 오차는 0.6748과 0.6672이다. 기존 연구와 사용된 데이터의 기간이 일치하지 않지만 수년 동안의 KOSPI 데이터를 사용하였다는 점에서 간접 비교를 하면 본 논문의 예측 결과가 우수함을 알 수 있다.

Table 4. KOSPI prediction errors

Model	MAPE	
	Learning	Test
NN1	0.427	0.628
NN2	0.487	0.635
NN3	0.424	0.648

Fig. 3은 평가용 데이터에 대해 KOSPI 실 데이터와 예측 모델 NN1의 출력을 비교한 그림으로 실선은 실 데이터이고 •은 예측된 데이터이다. 추세가 바뀌는 날에 서 상대적으로 오차가 큼을 볼 수 있다.

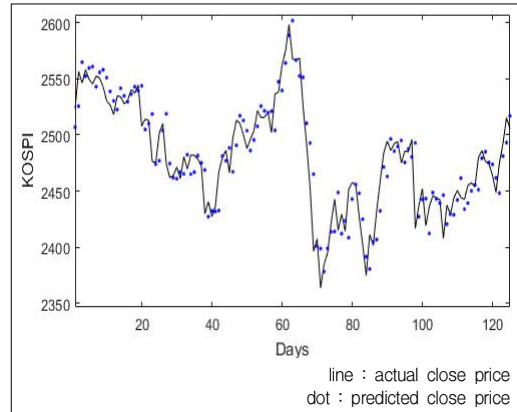


Fig. 3. Output comparison of KOSPI daily close prices in the test

3개의 예측 모델에 대하여 평가용 데이터를 사용하여 ETF의 일별 매매(Day Trading)를 시뮬레이션 하였다. 신경회로망 모델의 다음 날 지수 예측 결과가 식(1)을 만족하면 Tiger레버리지를, 식(2)를 만족하면 Tiger인버스 ETF를 매수한다. 매수된 ETF는 다음 날 증가를 기준으로 매도된다. ETF 매매 수수료는 0.1 %로 계산하였다. Table 5는 3개의 신경회로망 모델에 대해 최적의 결과를 보인 α 값, 매매율 (%), 매매 일 (일), 손실 일 (일)과 매매 성공률(전체 매매에서 수익이 발생한 날의 비율) (%)을 보인다. NN2는 매매 횟수가 4일로 작지만 매매로 손실이 발생하지 않았다. NN1과 NN3은 매매 횟수가 11일로 가장 많고 NN1은 1일의 손실이, NN3은 3일의 손실이 발생하였다.

Table 5. Trade performance of the neural network models

Model	α	Trade rate(%)	Trade days	Loss days	Success rate(%)
NN1	0.5	9.7	11	1	90.9
NN2	0.4	4.0	5	0	100.0
NN3	0.4	8.9	11	3	72.7

Table 6은 예측 모델에 대한 수익률을 표시한 것이다. 이는 평가용 데이터인 125 거래일(약 6.2개월)에서 거둔 수익률이다. NN1이 상승장에서 10.87 %와 하락장에서 4.42 %로 총 15.29 %로 가장 높은 수익률을 보였고 손실이 없는 안전한 매매인 NN2는 7.16 %의 수익률을 보였다. Table 5와 Table 6에 의하면 매매 성공률 90.0 %, 수익률 15.29 %를 보인 NN1이 최적으로 거래임을 보였다.

Table 6. Trade profit of the neural network models

Model	Profit (%) in bull market	Profit (%) in bear market	Total trade profit (%)
NN1	10.87	4.42	15.29
NN2	6.09	1.07	7.16
NN3	1.42	5.76	7.18

Fig. 4는 평가용으로 125 거래일 데이터를 사용해서 NN1 모델의 예측 결과가 식(1)을 만족하면 Tiger레버리지 ETF를 매수하고 다음 날 증가로 매도하였을 때의 매매당 수익률을 보인다. Fig. 5는 동일 데이터를 사용해서 NN1 모델의 예측 결과가 식(2)를 만족하면 Tiger인버스 ETF를 매수하고 다음 날 증가로 매도하였을 때의 매매당 수익률을 보인다. Fig. 5에서 첫 번째 Tiger인버스 ETF의 매매는 예측과 달리 손실이 발생하였음을 보인다. Fig. 6은 NN1의 예측에 의한 ETF 매매로 거둔 누적 수익률을 보인다. 첫 매매는 손실을 보이고 나머지 10 번의 매매에서 이득으로 총 15.29 %의 수익률을 보인다.

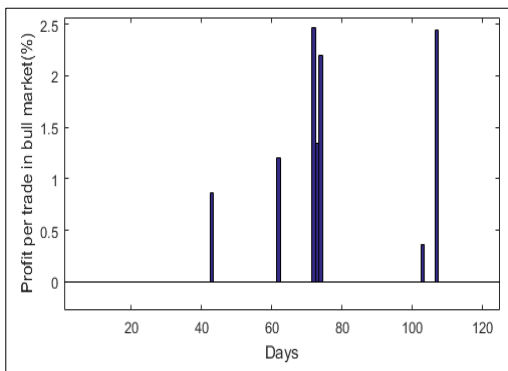


Fig. 4. Profit rate per trade in bull market

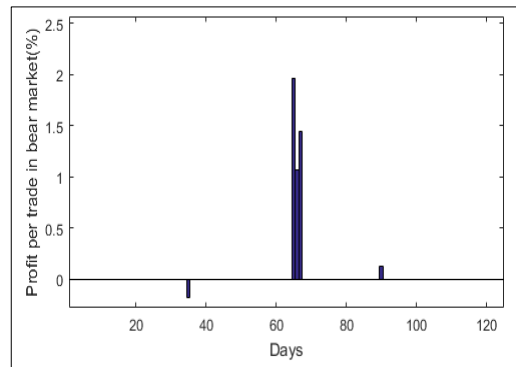


Fig. 5. Profit rate per trade in bear market

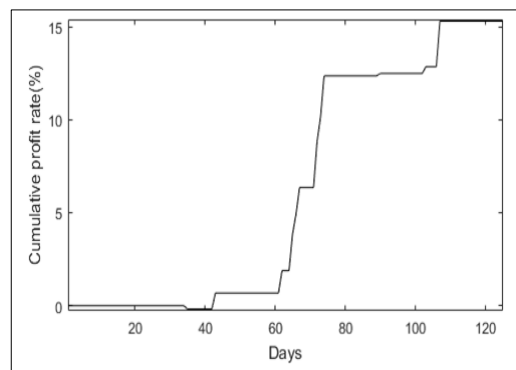


Fig. 6. Cumulative profit rate in the test

5. 결론

본 논문에서는 신경회로망 모델을 사용한 주가지수 예측과 이를 기반으로 지수 연동 ETF를 매매하는 방법에 대해 기술하였다. 14개의 입력 변수를 사용하고 은닉층이 하나 또는 둘인 신경회로망으로 평가용 데이터에 대해 KOSPI 주가의 MAPE가 0.628 ~ 0.648 범위로 효과적인 예측이 가능함을 보였다. 필연적인 예측 오차로 인한 매매 손실을 줄이기 위해 KOSPI 지수 연동 ETF의 알파 매매를 제안하였다. 시뮬레이션 결과 세 개의 예측 모델에서 최소 수익률 7.16 %, 최대 수익률 15.29 %를 보였다. NN1 모델을 사용해 장마감전에 다음 날 증가가 0.5 % 이상 상승하거나 -0.5 % 이하로 하락하는 것으로 예측되면 관련 ETF를 매수한다. 제시된 매매 수익률은 약 6.2개월 동안의 매매 결과로 제안된 알파 매매로 10 % 이상의 년 수익률을 거둘 수 있을 것으로 판단된다.

향후 수익률을 더 높이기 위해 제안된 방법을

KOSPI200선물레버리지 ETF와 KOSPI200선물인버스×2 ETF 매매에 적용하는 연구가 수행될 것이다.

REFERENCES

- [1] H. S. Hwang & J. S. Oh. (2009). Time Series Stock Prices Prediction Based on Fuzzy Model, *Journal of The Korean Institute of Intelligent Systems*, 19(5), 689-694.
- [2] W. Huang, Y. Nakamori & S. Y. Wang. (2005). Forecasting Stock Market Movement Direction with Support Vector Machine, *Computers & Operations Research*, 32(10), 2513-2522.
- [3] R. C. Cavalcante, R. C. Brasileiro, V. L. F. Souza, J. P. Nobrega & A. L. I. Oliveira. (2016). *Computational Intelligence and Financial Markets: A Survey and Future Directions*, *Expert Systems with Applications*, 55, 194-211.
- [4] H. Amin, H. M. Moein & E. Morteza. (2016). Stock market index prediction using artificial neural network, *Journal of Economics, Finance and Administrative Science*. 10.1016/j.jefas.2016.07.002.
- [5] H. S. Hwang. (2018). Daily Stock Price Forecasting Using Deep Neural Network Model, *Journal of the Korea Convergence Society*, 9(6), 39-44.
- [6] L. D. Persio & O. Honcha. (2017). Recurrent Neural Networks Approach to The Financial Forecast of Google Assets, *Int. J. of Mathetics and Computers in Simulation*, 11, 7-13.
- [7] X. Ding X, Y. Zhang, T. Liu & J. Duan. (2015). Deep Learning for Event-driven Stock Prediction, *Proc. of the 24th Int. Joint Conf. on Artificial Intelligence*, 2327-2333.
- [8] Qiu M, Song Y (2016) Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model.
<https://doi.org/10.1371/journal.pone.0155133>
- [9] S. H. Koh. (2016). A Converging Approach on Investment Strategies, Past Financial Information, and Investors' Behavioral Bias in the Korean Stock Market, *Journal of the Korea Convergence Society*, 7(6), 205-212.
- [10] T. Chandima & M. Sidney, M. Musa & H. Cameron. (2007). Predicting stock market index trading signals using neural networks. Pro. of the 14th Annual Global Finance Conference (GFC'07), 171-179.
- [11] L. J. Kyung-Soo & M. Boris. (2017). Forecasting ETFs with Machine Learning Algorithms,
<http://dx.doi.org/10.2139/ssrn.2899520>.
- [12] K. H. Lee & G. S. Jo. (1999). Expert System for Predicting Stock Market Timing Using A Candlestick Chart, *Expert System With Applications*, 16, 357-364.
- [13] L. C. H. Leon, A. Liu & W. S. Chen. (2006). Pattern Discovery of Fuzzy Time Series for Financial Prediction, *IEEE Trans. Knowledge and Data Engineering*, 18(5), 613-625.
- [14] O. Ican & T. B. Celik. (2017). Stock Market Prediction Performance of Neural Networks: A Literature Review, *International Journal of Economics and Finance*, 9(11), 100-108.
- [15] M. F. Moller. (1993). A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning, *Neural Networks*, 6, 525 - 533.

황희수(Heesoo Hwang)

[정회원]



석연구원

- 1988년 2월 : 연세대학교 본대학원 전기과 (공학석사)
- 1993년 2월 : 연세대학교 본대학원 전기과 (공학박사)
- 1993년 4월 ~ 2001년 2월 : 한국고속철도기술개발사업단 선임 및 수

- 2001년 3월 ~ 현재 : 한라대학교 전기전자과 교수
- 관심분야 : 머신러닝, 시계열 예측, 데이터 기반 진단 및 예측
- E-Mail : hshwang@halla.ac.kr