

**Professor Kyoung-Sook MOON, PhD**  
**E-mail: ksmoon@gachon.ac.kr**  
**Mathematical Finance, Gachon University**  
**Professor Hongjoong KIM, PhD**  
**E-mail: hongjoong@korea.ac.kr**  
**Mathematics, Korea University**

## **PERFORMANCE OF DEEP LEARNING IN PREDICTION OF STOCK MARKET VOLATILITY**

***Abstract.** Volatility forecasting is an important issue for investment analysis and risk management in finance. Based on the Long Short Term Memory (LSTM) deep learning algorithm, we propose an accurate algorithm for forecasting stock market index and its volatility. The proposed algorithm is tested on the data from 5 stock market indices including S&P500, NASDAQ, German DAX, Korean KOSPI200 and Mexico IPC over a 7-year period from 2010 to 2016. The highest prediction performance is observed with hybrid momentum, the difference between the price and the moving average of the past prices, for the predictions of both market index and volatility. Unlike stock index, the prediction accuracy for the volatility does not show dependency on other financial variables such as open, low, high prices, volume, etc. except the volatility itself.*

***Keywords:** volatility prediction; forecasting stock index; deep learning; long short term memory algorithm.*

**JEL Classification: C53, G17**

### **1. Introduction**

Accurate prediction of stock market volatility, the standard deviation of the underlying asset prices, is an important issue in the areas such as investment analysis of derivative securities, decision making and risk management in finance. Since financial markets are not only uncertain and complex but also globalized, it has become more and more difficult to predict financial parameters such as asset prices, indices and their volatilities.

In the early studies in financial derivatives, it was assumed that the volatility was constant. However, from analyzing the data, it has been generally accepted that the volatility is also a stochastic process and there have been studies to build different financial models for the volatility prediction, such as GARCH model, stochastic volatility models etc., see (Heston, 1993; Satchell and Knight, 2007).

Inspired by the great success of advanced data science in many application areas, there have been reported successful results for the prediction in financial market based on various machine learning algorithms, see (Kara et. al., 2011; Tsai et. al., 2011; Ballings et. al., 2015; Patel et. al., 2015; Oztekin et. al., 2016; Moon et. al., 2018; Rana et.al., 2018). To improve the over or under fitting problems in machine learning algorithms, hybridizations of existing classifiers obtained the promising results as in (Nayak et. al., 2015; Qiu et. al., 2016; Zhong and Enke, 2017; Chen and Hao, 2017). Starting from various financial models for volatility prediction, there also have been studies to combine volatility models such as GARCH model, EGARCH or GJR-GARCH and machine learning algorithms, see (Monfared and Enke, 2014; Dash and Dash, 2016; Peng et. al., 2018; Hurduzeu et. al. 2018).

Recently deep learning or hierarchical learning algorithm is introduced and produces superior results in many applications such as computer vision, bioinformatics, speech recognition etc., see (Goodfellow et. al. 2016; Geron, 2017). In (Hochreiter and Schmidhuber, 1997), an efficient deep learning called Long Short Term Memory (LSTM) was introduced and show superiority to machine learning algorithms, see also (Colah, 2015). In this paper, we apply the LSTM deep learning algorithm to financial market in order to predict the trend or values of stock indices and their volatilities.

Many studies use various indicators to identify the trend of the financial time series and develop machine learning or deep learning algorithms to forecast future trends or values. In order to improve the credibility or accuracy of the prediction, several methods can be combined to produce ensemble methods, multiple hidden layers can be introduced with many neurons in deep learning algorithms, or the quantity of input data for training or validation may be increased. In this study, the performance of the deep learning algorithm in the prediction of stock market volatility is analyzed and then compared to that of the market index.

We apply the Long Short Term Memory (LSTM) deep learning algorithm and consider following four aspects:

- the kind of the financial variable (i.e. volatility vs. stock)
- the way the variable of interest is estimated
- the number of features used in the training
- the kind of the market (i.e. United States or Europe vs Korean or Mexico).

The LSTM with above aspects is tested on the data from 5 stock market indices including *S&P500*, *NASDAQ*, *German DAX*, *Korean KOSPI200* and *Mexico IPC* over a 7-year period from 2010 to 2016. In Section 4, the higher prediction performance for predictions of market index and volatility is obtained with the standard and hybrid momentums. In particular, the increase of the number of features does not improve the accuracy for the volatility prediction, while it does for the index.

The remainder of the paper is organized as follows. The explanation of technical indicators and target variables is discussed in Section 2. Section 3

describes LSTM algorithm with parameters in detail and the empirical results are presented in Section 4. Section 5 concludes the paper and outlines future directions.

## 2. Methodology

The main goal of this work is to predict the trends or values of stock indices or their volatilities accurately based on LSTM algorithm. In this study, a close stock index price  $S_t$  at time  $t$  and its volatility  $\sigma_t = \sqrt{Var(S_t)}$  are the variables of interest. In general, machine learning algorithms have two parts: training and testing. During the training process, the algorithm learns  $S_t$  or  $\sigma_t$  or classifies their up or down trends according to the technical indicators based on the features.

### 2.1 Technical indicators

Let us first describe three different ways to make feature values. Firstly, *Moving Average* with period  $p$  is an average of  $S_t$  over the last  $p$  data points. The average can be computed with the same weights for those  $p$  values, or with different weights. In this study,  $MA(S_t, p)$  represents the *simple moving average* of  $S_t$  with period  $p$  given by

$$MA(S_t, p) = \frac{1}{p} \sum_{i=0}^{p-1} S_{t-i} \quad (1)$$

and *Exponential Moving Average*, denoted  $EMA(S_t, p)$  is the exponentially weighted average defined by

$$EMA(S_t, p) = \sum_{i=0}^{\infty} \alpha(1 - \alpha)^i S_{t-i} \quad (2)$$

where  $\alpha = 2/(1 + p)$ . The *momentum*  $M(S_t, m)$  represents the price difference between two different points with the lag  $m$ ,

$$M(S_t, m) = S_t - S_{t-m}. \quad (3)$$

Similarly,  $MA(\sigma_t, p)$ ,  $EMA(\sigma_t, p)$ , and  $M(\sigma_t, m)$  are defined with volatility  $\sigma_t$  instead of the price  $S_t$ .

### 2.2 Classification and Value Estimation

Let us consider the classification of  $S_t$ . Suppose that we have the partition of an interval  $(-\infty, \infty) = (-\infty, v_1] \cup (v_1, v_2] \cup \dots \cup (v_{K-2}, v_{K-1}] \cup (v_{K-1}, \infty)$ , where  $-\infty < v_1 < v_2 < \dots < v_{K-1} < \infty$ . When the value of  $S_t$  belongs to  $k^{th}$  interval  $(v_{k-1}, v_k]$ , we can set the value  $k$  as the label of  $S_t$ , denoted by  $L(S_t)$ . In this study, the training data are partitioned into  $K$  equal-sized buckets based on the quantities in each interval.

One may use the momentum  $M(S_t, m)$  instead of  $S_t$  for the classification. That is, when the value of  $M(S_t, m)$  belongs to  $k^{th}$  interval  $(v_{k-1}, v_k]$ , we can set the value  $k$  as the label, denoted by  $L^M(S_t, m)$ . Note that the classification based on

the momentum can be regarded as the estimation of the *trend* of the movement. For instance, when  $K = 2$ , the volatilities with the label 1 have *decreasing* momentums while those with the label 2 have *increasing* momentums. The volatilities can be partitioned into 3 groups with  $K = 3$  (i.e. the momentums *decrease, do not change, or increase*) or into 4 groups with  $K = 4$  (i.e. the momentums *decrease much, decrease little, increase little, or increase much*.)

Note that the *future value* can be predicted from the momentums in two ways. First,  $S_t$  can be obtained from  $S_{t-m}$  by adding the momentum,

$$S_t = S_{t-m} + M(S_t, m). \quad (4)$$

Alternatively, if the label based on the momentum is known, for instance  $L^M(S_t, m) = k$ , then

$$S_t \approx S_{t-m} + \mu_k \quad (5)$$

can be used as an approximate value of  $S_t$ , where  $\mu_k$  denotes the mean of the momentums belonging to the  $k^{th}$  bin.

### 2.3 Hybrid Momentum

Even though the momentum  $M(S_t, m)$  guides the trend of  $S_t$ , its label is oscillatory due to the noise in the financial time series. Thus, we introduce a *hybrid momentum*  $HM(S_t, m, p)$  defined by

$$HM(S_t, m, p) = S_t - MA(S_{t-m}, p) \quad (6)$$

and note that  $HM(S_t, m, 1) = M(S_t, m)$ .

The hybrid momentum can be used to efficiently measure the trends in financial data. That is, given the partition of  $R$  above, when the value of  $HM(S_t, m, p)$  belongs to the  $k^{th}$  interval, the value  $k$  can be set as the label, denoted by  $L^H(S_t, m, p)$ . Then the classification based on the hybrid momentum can be used for the estimation of the trend. For example, the upward or downward trend in volatility can be predicted with  $K = 2$ , and steep or gradual change in each direction can be considered with  $K = 4$ .

Similarly to (4) or (5) based on the standard momentum,  $S_t$  can be obtained from the hybrid momentum  $MA(S_{t-m_1}, m_2)$  by

$$S_t = MA(S_{t-m}, p) + HM(S_t, m, p) \quad (7)$$

or

$$S_t \approx MA(S_{t-m}, p) + \mu_k^h \quad (8)$$

where  $\mu_k^h$  represents the mean of the hybrid momentums  $HM(S_t, m, p)$  corresponding to the label  $L^H(S_t, m, p) = k$ . Table 1 summarizes the indicators used in this study.

**Table 1. The summary of indicators used in the algorithm**

| Indicators      | Definitions  |
|-----------------|--|
| $MA(S_t, p)$    | $\frac{1}{p} \sum_{i=0}^{p-1} S_{t-i}$   |
| $EMA(S_t, p)$   | $\sum_i^{\infty} \alpha(1 - \alpha)^i S_{t-i}, \quad \alpha = \frac{2}{1 + p}$ |
| $M(S_t, m)$     | $S_t - S_{t-m}$  |
| $HM(S_t, m, p)$ | $S_t - MA(S_{t-m}, p)$   |

Table 2 describes the statistics including the count, mean, standard deviation, minimum, first-, second- and third-quartiles, and maximum of variables and indicators for the volatility  $\sigma_t = \sqrt{Var(S_t)}$ . MA10, MA20, MA50 and EMA10, EMA20, EMA50 represent standard and exponential moving averages,  $MA(\sigma_t, p)$  and  $EMA(\sigma_t, p)$ , respectively, for  $p = 10, 20, 50$ . Mom and HMom in the last two columns of Table 2 are momentums  $M(\sigma_t, m)$  and hybrid momentums  $HM(\sigma_t, m, p)$  with  $p = 10, m = 5$ , respectively.

**Table 2. Statistics (count, mean, standard deviation, minimum, first-, second-, and third-quartiles, and maximum) of input variables, indicators (MA10, MA20, MA50 and EMA10, EMA20, EMA50 are  $MA(\sigma_t, p)$  and  $EMA(\sigma_t, p)$  for  $p = 10, 20, 50$ ) and momentums (Mom and HMom are the momentums  $M(\sigma_t, 5)$  and the hybrid momentums  $HM(\sigma_t, 5, 10)$ ) for stock volatilities  $\sigma_t = \sqrt{Var(S_t)}$  for S&P500, NASDAQ, DAX, KOSPI200 and IPC.**

| Open  | High  | Low   | Volume | Pct   | S     | MA10  | EMA10 | MA20  | EMA20 | MA50  | EMA50 | Mom    |
|-------|-------|-------|--------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| 1698  | 1698  | 1698  | 1698   | 1698  | 1698  | 1698  | 1698  | 1698  | 1698  | 1698  | 1698  | 1698   |
| 0.513 | 0.515 | 0.514 | 0.286  | 0.588 | 0.181 | 0.186 | 0.188 | 0.206 | 0.216 | 0.227 | 0.260 | -0.000 |
| 0.293 | 0.293 | 0.290 | 0.091  | 0.086 | 0.130 | 0.133 | 0.141 | 0.158 | 0.164 | 0.188 | 0.216 | 0.081  |
| 0.000 | 0.000 | 0.000 | 0.000  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | -0.432 |
| 0.237 | 0.239 | 0.241 | 0.233  | 0.550 | 0.094 | 0.099 | 0.095 | 0.103 | 0.110 | 0.112 | 0.123 | -0.041 |
| 0.526 | 0.527 | 0.530 | 0.274  | 0.588 | 0.149 | 0.152 | 0.149 | 0.160 | 0.162 | 0.160 | 0.184 | 0.002  |
| 0.808 | 0.810 | 0.804 | 0.929  | 0.630 | 0.234 | 0.237 | 0.237 | 0.254 | 0.261 | 0.270 | 0.310 | 0.039  |
| 1.000 | 1.000 | 1.000 | 1.000  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.599  |
| Open  | High  | Low   | Volume | Pct   | S     | MA10  | EMA10 | MA20  | EMA20 | MA50  | EMA50 | Mom    |
| 1698  | 1698  | 1698  | 1698   | 1698  | 1698  | 1698  | 1698  | 1698  | 1698  | 1698  | 1698  | 1698   |
| 0.484 | 0.486 | 0.487 | 0.387  | 0.570 | 0.198 | 0.189 | 0.196 | 0.206 | 0.227 | 0.215 | 0.253 | -0.000 |
| 0.297 | 0.298 | 0.295 | 0.089  | 0.091 | 0.132 | 0.136 | 0.142 | 0.153 | 0.161 | 0.194 | 0.204 | 0.086  |
| 0.000 | 0.000 | 0.000 | 0.000  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | -0.382 |
| 0.209 | 0.209 | 0.213 | 0.337  | 0.528 | 0.108 | 0.098 | 0.101 | 0.107 | 0.122 | 0.091 | 0.119 | -0.045 |
| 0.456 | 0.456 | 0.460 | 0.376  | 0.573 | 0.168 | 0.158 | 0.160 | 0.161 | 0.178 | 0.147 | 0.180 | 0.002  |
| 0.779 | 0.784 | 0.780 | 0.419  | 0.618 | 0.252 | 0.235 | 0.245 | 0.254 | 0.275 | 0.262 | 0.313 | 0.043  |
| 1.000 | 1.000 | 1.000 | 1.000  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.573  |
| Open  | High  | Low   | Volume | Pct   | S     | MA10  | EMA10 | MA20  | EMA20 | MA50  | EMA50 | Mom    |
| 1713  | 1713  | 1713  | 1713   | 1713  | 1713  | 1713  | 1713  | 1713  | 1713  | 1713  | 1713  | 1713   |
| 0.466 | 0.468 | 0.466 | 0.217  | 0.564 | 0.281 | 0.283 | 0.278 | 0.286 | 0.277 | 0.286 | 0.300 | -0.000 |
| 0.248 | 0.250 | 0.244 | 0.125  | 0.108 | 0.159 | 0.175 | 0.172 | 0.175 | 0.173 | 0.181 | 0.184 | 0.102  |
| 0.000 | 0.000 | 0.000 | 0.000  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | -0.374 |
| 0.255 | 0.252 | 0.259 | 0.134  | 0.511 | 0.178 | 0.174 | 0.169 | 0.171 | 0.159 | 0.166 | 0.169 | -0.057 |
| 0.451 | 0.451 | 0.454 | 0.183  | 0.368 | 0.249 | 0.245 | 0.239 | 0.249 | 0.241 | 0.241 | 0.250 | -0.003 |
| 0.665 | 0.668 | 0.663 | 0.265  | 0.618 | 0.351 | 0.358 | 0.351 | 0.354 | 0.344 | 0.370 | 0.388 | 0.056  |
| 1.000 | 1.000 | 1.000 | 1.000  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.482  |
| Open  | High  | Low   | Volume | Pct   | S     | MA10  | EMA10 | MA20  | EMA20 | MA50  | EMA50 | Mom    |
| 1170  | 1170  | 1170  | 1170   | 1170  | 1170  | 1170  | 1170  | 1170  | 1170  | 1170  | 1170  | 1170   |
| 0.511 | 0.487 | 0.522 | 0.310  | 0.535 | 0.385 | 0.346 | 0.347 | 0.335 | 0.337 | 0.317 | 0.338 | -0.001 |
| 0.155 | 0.159 | 0.167 | 0.113  | 0.125 | 0.180 | 0.203 | 0.212 | 0.197 | 0.211 | 0.219 | 0.233 | 0.138  |
| 0.000 | 0.000 | 0.000 | 0.000  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | -0.518 |
| 0.417 | 0.391 | 0.421 | 0.237  | 0.472 | 0.238 | 0.204 | 0.203 | 0.187 | 0.190 | 0.160 | 0.172 | -0.075 |
| 0.524 | 0.502 | 0.538 | 0.290  | 0.336 | 0.323 | 0.296 | 0.296 | 0.283 | 0.295 | 0.267 | 0.281 | -0.001 |
| 0.606 | 0.587 | 0.630 | 0.356  | 0.604 | 0.450 | 0.431 | 0.427 | 0.403 | 0.418 | 0.419 | 0.451 | 0.073  |
| 1.000 | 1.000 | 1.000 | 1.000  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.656  |
| Open  | High  | Low   | Volume | Pct   | S     | MA10  | EMA10 | MA20  | EMA20 | MA50  | EMA50 | Mom    |
| 1183  | 1183  | 1183  | 1183   | 1183  | 1183  | 1183  | 1183  | 1183  | 1183  | 1183  | 1183  | 1183   |
| 0.520 | 0.518 | 0.518 | 0.121  | 0.562 | 0.302 | 0.318 | 0.337 | 0.344 | 0.371 | 0.379 | 0.434 | 0.000  |
| 0.206 | 0.206 | 0.211 | 0.058  | 0.106 | 0.162 | 0.163 | 0.172 | 0.189 | 0.189 | 0.229 | 0.225 | 0.126  |
| 0.000 | 0.000 | 0.000 | 0.000  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | -0.675 |
| 0.349 | 0.349 | 0.343 | 0.092  | 0.503 | 0.191 | 0.212 | 0.223 | 0.208 | 0.235 | 0.205 | 0.268 | -0.068 |
| 0.529 | 0.525 | 0.528 | 0.112  | 0.563 | 0.275 | 0.288 | 0.312 | 0.317 | 0.354 | 0.359 | 0.411 | -0.001 |
| 0.679 | 0.675 | 0.685 | 0.139  | 0.626 | 0.384 | 0.401 | 0.421 | 0.446 | 0.468 | 0.525 | 0.562 | 0.069  |
| 1.000 | 1.000 | 1.000 | 1.000  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.700  |

### 3. Predictions

#### 3.1 LSTM algorithm

In this study, the deep learning algorithm based on Long Short Term Memory (LSTM) introduced in (Hochreiter and Schmidhuber, 1997) is used to predict future trends and values of  $\sigma_t = \sqrt{Var(S_t)}$ . LSTM is a special RNN, which can handle *long term dependencies*. Instead of a single neural network, LSTM has four interacting layers. See (Colah, 2015) for details.

Three different numbers of features in the input data are considered to observe the effect of the number of features.

- *One feature*: only target variable  $S_t$  is provided for training.
- *6 features (small number of features)*: target variable  $S_t$  or  $\sigma_t$  with open, high and low values  $O_t, H_t, L_t$ , the volume  $V_t$ , and the daily percentage change  $P_t$  of the stock index are provided.

- 12 features (*many number of features*):the features above with additional 3 standard and exponential moving averages of  $S_t$  , ( $MA(S_t, m_1)$  ,  $MA(S_t, m_2)$ ,  $MA(S_t, m_3)$ ,  $EMA(S_t, m_1)$ ,  $EMA(S_t, m_2)$ ,  $EMA(S_t, m_3)$ ) are provided. For prediction of volatility we use  $\sigma_t$  instead of  $S_t$ .

### 3.2 Prediction of future trends or values

Following three types of target variables are considered in this study, let  $\Phi(v)$  represent the estimation of  $v$  using the deep learning algorithm.

- $S_t$ : the values of  $S_t$  itself are trained. The result  $\Phi(S_T)$  from deep learning represents the direct prediction of  $S_T$ , and the up-down trend can be estimated by the classification label  $L(\Phi(S_T))$ .
- $M(S_t, m)$  : the momentums of  $S_t$  are trained. Once the return value  $\Phi(M(S_T, m))$  for the momentum is obtained, its classification label  $L^M(\Phi(M(S_T, m)), m)$  predicts the trend of the movement. The value of  $S_T$  can be predicted in two ways, i.e. by  $S_{T-m} + \Phi(M(S_T, m))$  as in (4) or by  $S_{T-m} + \Phi(\mu_k)$  as in (5), where  $\Phi(\mu_k)$  is the mean of the training data corresponding to the label  $L^M(\Phi(M(S_T, m)), m)$ .
- $HM(S_t, m, p)$  : the hybrid momentums of  $S_t$  are trained. Similarly to above, once the return value  $\Phi(HM(S_T, m, p))$  for the hybrid momentum is obtained, its classification label  $L^H(\Phi(HM(S_T, m, p)), m, p)$  predicts the trend of the movement. The value of  $S_T$  can be predicted in two ways, i.e. by  $S_{T-m} + \Phi(HM(S_T, m, p))$  as in (7) or by  $S_{T-m} + \Phi(\mu_k^h)$  as in (8), where  $\Phi(\mu_k^h)$  is the mean of the training data corresponding to the label  $L^H(\Phi(HM(S_T, m, p)), m, p)$ .

Table 3 summarizes the target variables used in this study.

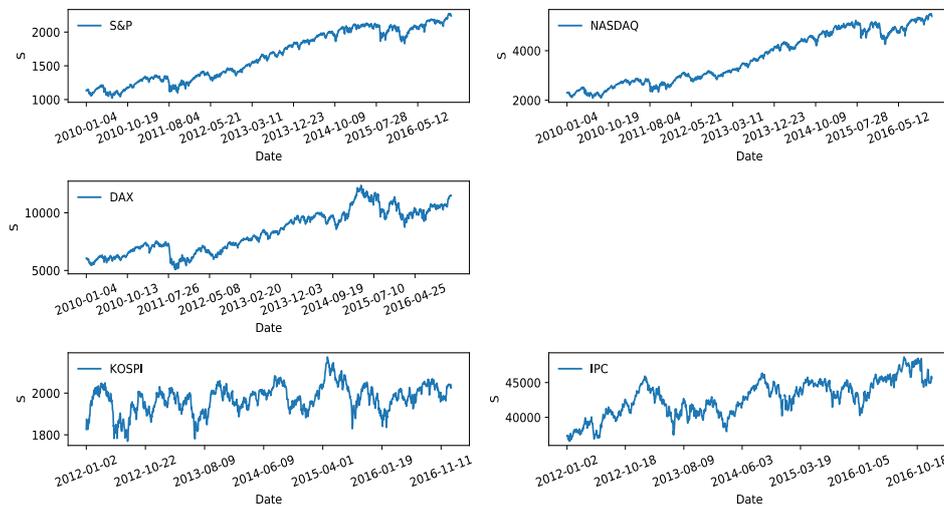
**Table 3. Summary of target variables used for the prediction of trends or values of the volatility.**

| Types of target variable | $S_t$          | $M(S_t, m)$               |                             | $HM(S_t, m, p)$                  |                                 |
|--------------------------|----------------|---------------------------|-----------------------------|----------------------------------|---------------------------------|
| Prediction of trends     | $L(\Phi(S_T))$ | $L^M(\Phi(M(S_T, m)), m)$ |                             | $L^H(\Phi(HM(S_T, m, p)), m, p)$ |                                 |
| Prediction of values     | $\Phi(S_T)$    | $S_{T-m} + \Phi(\mu_k)$   | $S_{T-m} + \Phi(M(S_T, m))$ | $S_{T-m} + \Phi(\mu_k^h)$        | $S_{T-m} + \Phi(HM(S_T, m, p))$ |

## 4. Results

### 4.1 Data Description

The data from 5 stock market indices, S&P500, NASDAQ (United States), DAX (Germany), KOSPI200 (Korea) and IPC (Mexico) for 7 years from April 1, 2010 to December 30, 2016 is used in this study. The daily index values in the form of  $(H_t, L_t, O_t, C_t, V_t)$  of the high, low, open and close values, and the volume, respectively, have been downloaded from the *Yahoo Finance*. Figure 1 shows the trend of the indices of the financial markets. The price index of KOSPI200 and IPC seem to have wider and more continuous fluctuation compared to those of S&P500, NASDAQ and DAX.



**Figure 1. Stock indices**

### 4.2 Prediction of future trends

Given the classification for up or down movement, the *true positive* rate can be used as a measure for the performance of the prediction. It represents the ratio of actual positives that are correctly identified. The parameters used in LSTM algorithm is summarized in Table 4.

**Table 4. Summary of parameters for LSTM algorithm.**

| Parameters                   | Values                             |
|------------------------------|------------------------------------|
| the number of labels ( $K$ ) | 2, 3, 4                            |
| market index                 | S&P500, NASDAQ, DAX, KOSPI200, IPC |

Performance of Deep Learning in Prediction of Stock Market Volatility

|                              |  |
|------------------------------|--|
| the number of input features | 1, 6, 12                                       |
| types of target variable     | $\sigma_t, M(\sigma_t, m), HM(\sigma_t, m, p)$ |

Table 5 represents the average of the true positive rates when the trend of the volatility is estimated by  $L(\Phi(\sigma_t))$ ,  $L^M(\Phi(M(\sigma_t, m)), m)$ , and  $L^H(\Phi(HM(\sigma_t, m, p)), m, p)$  for each case of the number of labels, the index, the number of input features and the type of target variables in Table 4. The last row represents the naive probability that the random classification gives the correct identification. Note that the true positive rates with respect to the values  $S_t$  are quite good but those with respect to the hybrid moments  $HM(\sigma_t, m, p)$  are slightly better regardless of the number of features or indices. More importantly, the rates do not seem to be dependent upon the number of features and the classification with only 1 feature result in better rates than that with 6 or 12 features in some cases.

**Table 5. The average of the true positive rates when the trend of the stock volatility  $\sigma_t = \sqrt{Var(S_t)}$  is estimated with the parameter in Table 4.**

| $K$    |             | 2        |          |          | 4        |          |          |
|--------|-------------|----------|----------|----------|----------|----------|----------|
|        |             | $S_t$    | $M$      | $HM$     | $S_t$    | $M$      | $HM$     |
| S&P    | 1 feature   | 0.556522 | 0.860870 | 0.904348 | 0.266667 | 0.655072 | 0.750725 |
|        | 6 features  | 0.588406 | 0.805797 | 0.866667 | 0.260870 | 0.591304 | 0.721739 |
|        | 12 features | 0.548961 | 0.833828 | 0.878338 | 0.284866 | 0.632047 | 0.667656 |
| NASDAQ | 1 feature   | 0.582609 | 0.849275 | 0.898551 | 0.266667 | 0.698551 | 0.684058 |
|        | 6 features  | 0.573913 | 0.820290 | 0.837681 | 0.269565 | 0.437681 | 0.626087 |
|        | 12 features | 0.569733 | 0.747774 | 0.857567 | 0.252226 | 0.554896 | 0.718101 |
| DAX    | 1 feature   | 0.487032 | 0.870317 | 0.847761 | 0.293948 | 0.677233 | 0.650746 |
|        | 6 features  | 0.564841 | 0.841499 | 0.865672 | 0.299712 | 0.628242 | 0.632836 |
|        | 12 features | 0.536873 | 0.861357 | 0.850153 | 0.286136 | 0.489676 | 0.623853 |
| KOSPI  | 1 feature   | 0.523013 | 0.811715 | 0.889362 | 0.238494 | 0.669456 | 0.693617 |
|        | 6 features  | 0.552301 | 0.748954 | 0.851064 | 0.259414 | 0.497908 | 0.668085 |
|        | 12 features | 0.549784 | 0.796537 | 0.867841 | 0.255411 | 0.580087 | 0.726872 |
| IPC    | 1 feature   | 0.493776 | 0.863071 | 0.877119 | 0.261411 | 0.688797 | 0.686441 |
|        | 6 features  | 0.510373 | 0.846473 | 0.838983 | 0.265560 | 0.626556 | 0.686441 |
|        | 12 features | 0.545064 | 0.845494 | 0.802632 | 0.270386 | 0.635193 | 0.662281 |
| Naive  |             | 0.500000 |          |          | 0.250000 |          |          |

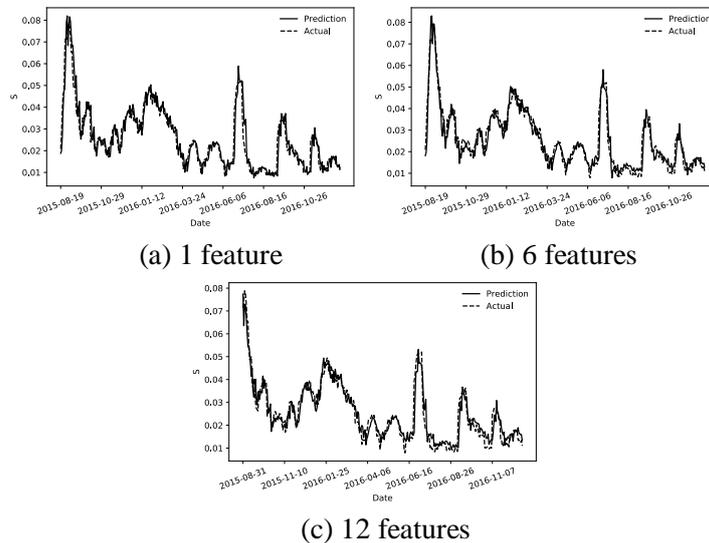
Table 6 shows the average of the true positive rates when the trend of the stock market is estimated, which shows the weakness of the classification with respect to  $S_t$  as follows. Since unlike the volatility, the stock market index such as S&P500, NASDAQ and DAX increased for the past decade as seen in Figure 1, the values of the test data are not observed during the training period (i.e. the range of the test data and that of the training data do not overlap much) so that the corresponding rates are not good compared to the others. Such inadequate training is not observed when the index of KOSPI or IPC is considered. Note that such inadequate training can be avoided by the computation of the momentum or the hybrid momentum even for S&P500, NASDAQ and DAX index. The hybrid momentum produces better prediction accuracies for the trends of both stock market indices and volatilities, but the number of features seems to have positive effects only on the market index, not the volatility.

**Table 6. The average of the true positive rates when the trend of the stock value  $S_t$  is estimated with the parameters in Table 4.**

| K      |             | 2        |          |          | 4        |          |          |
|--------|-------------|----------|----------|----------|----------|----------|----------|
|        |             | $S_t$    | M        | HM       | $S_t$    | M        | HM       |
| S&P    | 1 feature   | 0.488439 | 0.488439 | 0.501441 | 0.268786 | 0.205202 | 0.363112 |
|        | 6 features  | 0.465318 | 0.786127 | 0.749280 | 0.265896 | 0.500000 | 0.475504 |
|        | 12 features | 0.495575 | 0.784661 | 0.867257 | 0.274336 | 0.510324 | 0.640118 |
| NASDAQ | 1 feature   | 0.528902 | 0.528902 | 0.527536 | 0.306358 | 0.225434 | 0.124638 |
|        | 6 features  | 0.528902 | 0.673410 | 0.840580 | 0.306358 | 0.552023 | 0.539130 |
|        | 12 features | 0.536873 | 0.784661 | 0.674556 | 0.312684 | 0.516224 | 0.520710 |
| DAX    | 1 feature   | 0.484241 | 0.742120 | 0.706052 | 0.297994 | 0.197708 | 0.530259 |
|        | 6 features  | 0.484241 | 0.770774 | 0.841499 | 0.295129 | 0.607450 | 0.654179 |
|        | 12 features | 0.494152 | 0.792398 | 0.859649 | 0.304094 | 0.590643 | 0.728070 |
| KOSPI  | 1 feature   | 0.454545 | 0.789256 | 0.851240 | 0.227273 | 0.611570 | 0.574380 |
|        | 6 features  | 0.466942 | 0.793388 | 0.851240 | 0.231405 | 0.595041 | 0.714876 |
|        | 12 features | 0.457265 | 0.773504 | 0.897436 | 0.196581 | 0.568376 | 0.645299 |
| IPC    | 1 feature   | 0.374486 | 0.827160 | 0.754098 | 0.181070 | 0.625514 | 0.532787 |
|        | 6 features  | 0.395062 | 0.818930 | 0.840164 | 0.185185 | 0.662551 | 0.549180 |
|        | 12 features | 0.412766 | 0.825532 | 0.851695 | 0.200000 | 0.608511 | 0.627119 |
| Naive  |             | 0.500000 |          |          | 0.250000 |          |          |

**4.2 Prediction of future values**

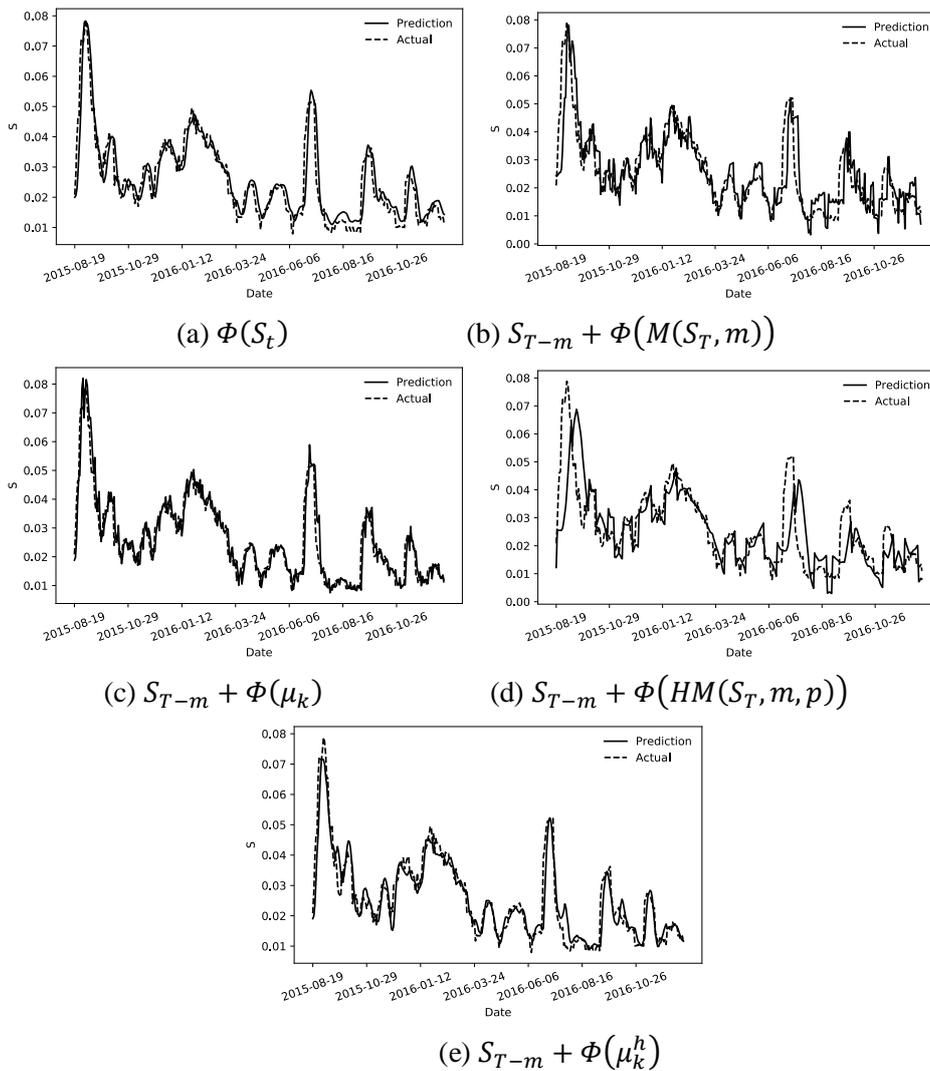
Now let us consider the prediction of future volatility values for each case of the number of labels, the index, the number of input features and the type of target variables in Table 4. The value of the volatility can be estimated by 5 different ways,  $\Phi(S_T)$ ,  $S_{T-m} + \Phi(M(S_T, m))$ ,  $S_{T-m} + \Phi(\mu_k)$ ,  $S_{T-m} + \Phi(HM(S_T, m, p))$ ,  $S_{T-m} + \Phi(\mu_k^h)$ . For instance, Figure 2 shows the S&P500 stock market volatility and its prediction by  $S_{T-m} + \Phi(\mu_k)$  with respect to the number of:



**Figure 2. Actual S&P500 stock market volatility and its prediction by  $S_{T-m} + \Phi(\mu_k)$  when the numbers of features are 1, 6 and 12.**

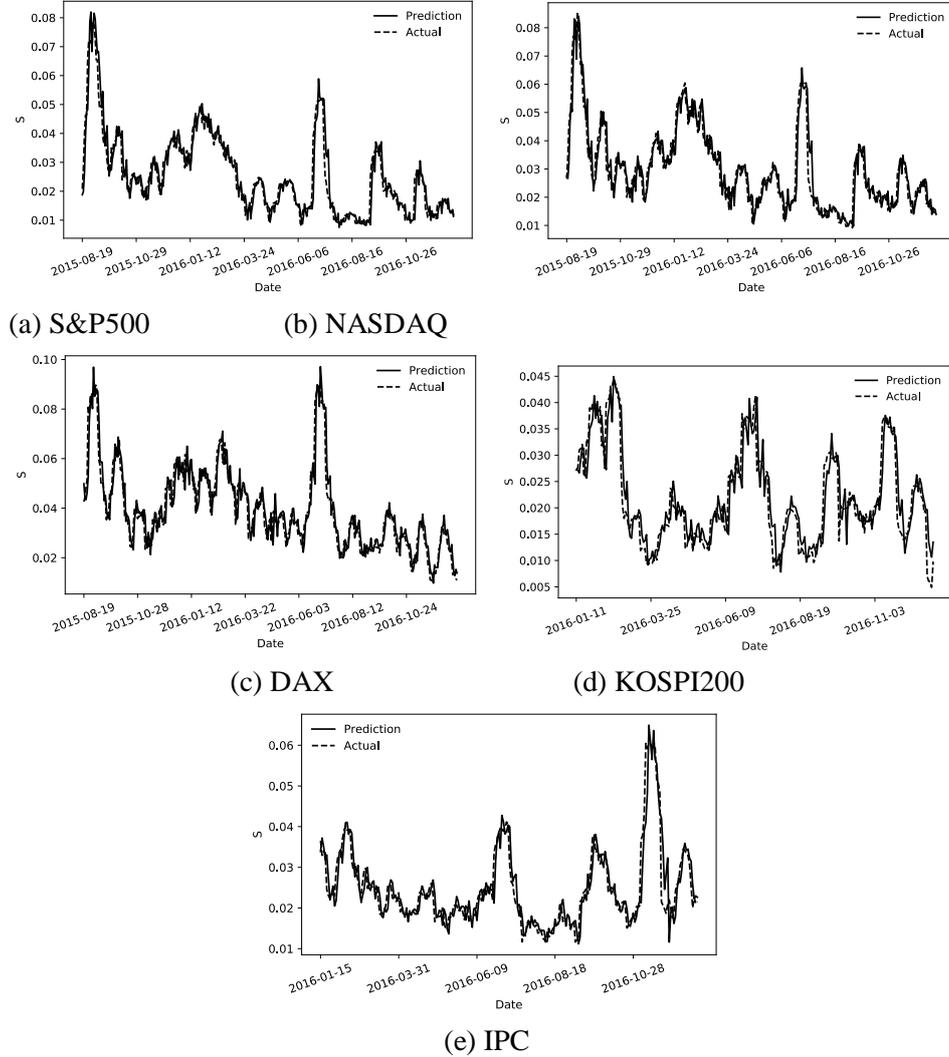
Performance of Deep Learning in Prediction of Stock Market Volatility

When the way the target variable is predicted is changed, the S&P500 stock market volatility prediction results are as follows:



**Figure 3. Actual S&P500 stock market volatility and its prediction with only 1 feature when the way the target variable is predicted is changed.**

The difference between the actual values and predictions is much bigger when different target variables are used in Figure 3 compared to the difference when different number of features is used in Figure 2. Following Figure 4 compares the volatility prediction results for 5 different stock market indices:



**Figure 4. Actual volatility and its prediction by  $S_{T-m} + \Phi(\mu_k)$  with 1 feature from 5 stock market indices: S&P500, NASDAQ, DAX, KOSPI200 and IPC.**

#### 4.2.1 Measures

Following two errors are used to measure the accuracies in the prediction of values.

- Mean squared error (MSE):  $MSE = \frac{1}{n} \sum_{i=1}^n (\sigma_{t_i} - \widehat{\sigma}_{t_i})^2$
- Mean absolute percentage error (MAPE):  $MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\sigma_{t_i} - \widehat{\sigma}_{t_i}}{\sigma_{t_i}} \right|$ ,

Performance of Deep Learning in Prediction of Stock Market Volatility

where  $\sigma_{t_i}$  is the value at time  $t_i$  and  $\widehat{\sigma}_{t_i}$  is its prediction. Table 7 represents the mean squared errors (MSE) for each case of the index, the number of input features and the type of target variables in Table 4.

**Table 7. MSE for the prediction of the stock market volatility with respect to the indices, input types and target types.**

|        |             | $\Phi(\sigma_t)$ | $\Phi(\mu_k)$ | $\Phi(M)$ | $\Phi(\mu_k^h)$ | $\Phi(HM)$ |
|--------|-------------|------------------|---------------|-----------|-----------------|------------|
| S&P    | 1 feature   | 0.000026         | 0.000063      | 0.000019  | 0.000113        | 0.000022   |
|        | 6 features  | 0.000029         | 0.000067      | 0.000016  | 0.000117        | 0.000027   |
|        | 12 features | 0.000020         | 0.000049      | 0.000016  | 0.000097        | 0.000022   |
| NASDAQ | 1 feature   | 0.000029         | 0.000075      | 0.000022  | 0.000136        | 0.000030   |
|        | 6 features  | 0.000086         | 0.000102      | 0.000044  | 0.000170        | 0.000064   |
|        | 12 features | 0.000036         | 0.000069      | 0.000036  | 0.000134        | 0.000026   |
| DAX    | 1 feature   | 0.000046         | 0.000090      | 0.000032  | 0.000179        | 0.000044   |
|        | 6 features  | 0.000072         | 0.000101      | 0.000037  | 0.000198        | 0.000055   |
|        | 12 features | 0.000067         | 0.000092      | 0.000040  | 0.000191        | 0.000043   |
| KOSPI  | 1 feature   | 0.000027         | 0.000037      | 0.000015  | 0.000063        | 0.000024   |
|        | 6 features  | 0.000023         | 0.000038      | 0.000024  | 0.000065        | 0.000025   |
|        | 12 features | 0.000031         | 0.000039      | 0.000018  | 0.000063        | 0.000022   |
| IPC    | 1 feature   | 0.000029         | 0.000055      | 0.000016  | 0.000095        | 0.000021   |
|        | 6 features  | 0.000040         | 0.000056      | 0.000022  | 0.000106        | 0.000021   |
|        | 12 features | 0.000035         | 0.000057      | 0.000019  | 0.000105        | 0.000020   |

Table 7 shows that the predictions based on the momentum and the hybrid momentum,  $\sigma_{T-m} + \Phi(\mu_k)$  and  $\sigma_{T-m} + \Phi(HM(\sigma_T, m, p))$ , are better than those based on the value  $\Phi(\sigma_t)$  or the mean averages,  $\sigma_{T-m} + \Phi(\mu_k)$  and  $\sigma_T + \Phi(\mu_k^h)$ . In addition, the accuracy of the prediction is not improved when 6 or 12 features are used compared to the prediction with single feature only as observed in the prediction of the up-down trends. Table 8 represents the mean absolute percentage errors (MAPE) for the stock volatility for each case of parameters in Table 4. The results are similar to those from the MSE errors. On the other hand, the number of features affects the prediction of stock market indices as in Table 9.

**Table 8. MAPE for stock volatility with respect to the indices, input types and target types.**

|        |             | $\Phi(\sigma_t)$ | $\Phi(\mu_k)$ | $\Phi(M)$ | $\Phi(\mu_k^h)$ | $\Phi(HM)$ |
|--------|-------------|------------------|---------------|-----------|-----------------|------------|
| S&P    | 1 feature   | 0.163513         | 0.242556      | 0.117751  | 0.286171        | 0.138758   |
|        | 6 features  | 0.166877         | 0.250906      | 0.136487  | 0.292519        | 0.153767   |
|        | 12 features | 0.169169         | 0.237896      | 0.145978  | 0.287685        | 0.146001   |
| NASDAQ | 1 feature   | 0.144711         | 0.226552      | 0.115783  | 0.278826        | 0.139860   |
|        | 6 features  | 0.249747         | 0.275721      | 0.198779  | 0.335995        | 0.255263   |
|        | 12 features | 0.162544         | 0.232119      | 0.163614  | 0.316072        | 0.129283   |
| DAX    | 1 feature   | 0.131749         | 0.180979      | 0.110242  | 0.236755        | 0.133250   |
|        | 6 features  | 0.175659         | 0.187149      | 0.125327  | 0.244958        | 0.156644   |
|        | 12 features | 0.169671         | 0.200711      | 0.144669  | 0.247046        | 0.136845   |
| KOSPI  | 1 feature   | 0.234685         | 0.267634      | 0.153841  | 0.344652        | 0.214972   |
|        | 6 features  | 0.211396         | 0.274907      | 0.200791  | 0.351643        | 0.193113   |
|        | 12 features | 0.227424         | 0.280337      | 0.171994  | 0.345195        | 0.201980   |
| IPC    | 1 feature   | 0.150047         | 0.202207      | 0.110848  | 0.254583        | 0.136609   |
|        | 6 features  | 0.160959         | 0.203139      | 0.139105  | 0.268954        | 0.129007   |
|        | 12 features | 0.162059         | 0.205530      | 0.119393  | 0.270246        | 0.121897   |

**Table 9. MAPE for stock index with respect to the indices, input types and target types.**

|        |             | $\Phi(S_t)$ | $\Phi(\mu_k)$ | $\Phi(M)$ | $\Phi(\mu_k^h)$ | $\Phi(HM)$ |
|--------|-------------|-------------|---------------|-----------|-----------------|------------|
| S&P    | 1 feature   | 0.030996    | 0.017059      | 0.014602  | 0.023147        | 0.020217   |
|        | 6 features  | 0.024696    | 0.012460      | 0.009102  | 0.017692        | 0.010432   |
|        | 12 features | 0.026296    | 0.010857      | 0.008947  | 0.013056        | 0.009810   |
| NASDAQ | 1 feature   | 0.031296    | 0.020007      | 0.018066  | 0.026136        | 0.021438   |
|        | 6 features  | 0.040861    | 0.016977      | 0.011808  | 0.018899        | 0.014817   |
|        | 12 features | 0.026391    | 0.013682      | 0.010911  | 0.021501        | 0.012013   |
| DAX    | 1 feature   | 0.025035    | 0.020352      | 0.022385  | 0.030767        | 0.019756   |
|        | 6 features  | 0.024217    | 0.020089      | 0.013362  | 0.026504        | 0.013078   |
|        | 12 features | 0.042276    | 0.018737      | 0.014445  | 0.022207        | 0.013319   |
| KOSPI  | 1 feature   | 0.013398    | 0.010458      | 0.006633  | 0.009972        | 0.008507   |
|        | 6 features  | 0.008474    | 0.010366      | 0.006727  | 0.009167        | 0.008141   |
|        | 12 features | 0.010701    | 0.010553      | 0.006529  | 0.008468        | 0.007261   |
| IPC    | 1 feature   | 0.013797    | 0.011148      | 0.007269  | 0.014482        | 0.011570   |
|        | 6 features  | 0.014798    | 0.011512      | 0.007918  | 0.012573        | 0.008223   |
|        | 12 features | 0.018094    | 0.011713      | 0.007442  | 0.012534        | 0.007420   |

### 5. Conclusions

The prediction of the stock market index and volatility has been observed in several aspects. The stock market index and volatility share some similarities but also have some distinctions. The usage of standard and hybrid momentums as target variable is superior to the usage of the value of the variable or classification label in estimating the up-down trend or predicting the future value. In particular, the hybrid momentum shows good results for the prediction. The increase of the number of features does not improve the accuracy for the volatility prediction, while it does for the index itself.

### Acknowledgements

*This work was supported by the Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (2017R1D1A1B03035543); and National Research Foundation of Korea [NRF-2018R1D1A1B07050046].*

### REFERENCES

- [1] Ballings, M., den Poel, D.V., Hespeels, N., Gryp, R. (2015), *Evaluating Multiple Classifiers for Stock Price Direction Prediction*; Expert Systems with Applications, 42:7046—7056;
- [2] Chen, Y., Hao, Y. (2017), *A Feature Weighted Support Vector Machine and K-Nearest Neighbor Algorithm for Stock Market Indices Prediction*; Expert Systems with Applications, 80:340—355;
- [3] Dash, R., Dash, P.K. (2016), *An Evolutionary Hybrid Fuzzy Computationally Efficient EGARCH Model for Volatility Prediction*; Applied Soft Computing, 45:40—60;
- [4] Geron, A. (2017), *Hands-on Machine Learning with Scikit-Learn and Tensor Flow: Concepts, Tools, and Techniques to Build Intelligent Systems*; O'Reilly Media, ISBN-13: 978-1491962299;
- [5] Goodfellow, I., Bengio, Y., Courville, A. (2016), *Deep Learning (Adaptive Computation and Machine Learning)*; The MIT Press, ISBN-13:978-0262035613;
- [6] Heston, S. L. (1993), *A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*; The Review of Financial Studies, 6(2): 327—343;
- [7] Hochreiter, S., Schmidhuber, J. (1997), *Long Short-term Memory*; Neural Computation 9(8):1735—1780;
- [8] Hurduzeu, G., Lolea, I.-C., Giurea, A.-M., Popescu, M.F. (2018), *Does Real-time Microeconomic Data Ensure an Accurate Volatility Forecasting? A Two States Approach for The US Equity Market*; Economic Computation and Economic Cybernetics Studies and Research; ASE Publishing, 52(2):37—50;
- [9] Kara, Y., Boyacioglu, M.A., Baykan, O.K. (2011), *Predicting Direction of Stock Price Index Movement Using Artificial Neural Networks and Support Vector Machines: The Sample of the Istanbul Stock Exchange*; Expert Systems with Applications, 38: 5311—5319;
- [10] Monfared, S. A., Enke, D. (2014), *Volatility Forecasting Using a Hybrid GJR-GARCH Neural Network Model*; Procedia Computer Science 36:246—253;
- [11] Moon, K-S., Jun, S., Kim, H. (2018), *Speed up of the Majority Voting Ensemble Method for the Prediction of Stock Price Directions*; Economic Computation and Economic Cybernetics Studies and Research; ASE Publishing; 52(1):215—228;

- [12] **Nayak, R. K., Mishra, D., Rath, A. K. (2015),** *A Naïve SVM-KNN Based Stock Market Trend Reversal Analysis for Indian Benchmark Indices*; Applied Soft Computing, 35:670—680;
- [13] **Colah, C. (2015),** *Understanding LSTM Networks*, available at: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>, last accessed 08.31.2018;
- [14] **Oztekin, A., Kizilaslan, R., Freund, S., Iseri, A. (2016),** *A Data Analytic Approach to Forecasting Daily Stock Returns in an Emerging Market*; European Journal of Operational Research, 253: 697—710;
- [15] **Patel, J., Shah, S., Thakkar, P., Kotecha, K. (2015),** *Predicting Stock and Stock Price Index Movement Using Trend Deterministic Data Prediction and Machine Learning Techniques*; Expert Systems with Applications, 42: 259—268;
- [16] **Peng, Y., Albuquerque, P.H.M., de Sa, J.M.C., Padula, A.J.A., Montenegro, M.R. (2018),** *The Best of Two Worlds: Forecasting High Frequency Volatility for Cryptocurrencies and Traditional Currencies with Support Vector Regression*; Expert Systems with Applications, 97: 177—192;
- [17] **Qui, M., Song, Y., Akagi, F. (2016),** *Application of Artificial Neural Network for the Prediction of Stock Market Returns: The Case of the Japanese Stock Market*; Chaos, Solitons and Fractals, 85:1—7;
- [18] **Rana, S., Dhhan, W., Midi, H. (2018),** *Fixed Parameters Support Vector Regression for Outlier Detection*; Economic Computation and Economic Cybernetics Studies and Research; *ASE Publishing*; 52(2):267—282;
- [19] **Satchell, S., Knight, J. (2007),** *Forecasting Volatility in the Financial Markets (Quantitative Finance)*, Butterworth-Heinemann, ISBN-13:978-0750669429;
- [20] **Tsai, C.F., Lin, Y.C., Yen, D.C., Chen, Y.M. (2011),** *Predicting Stock Returns Byclassifier Ensembles*; Applied Soft Computing, 11: 2452—2459;
- [21] **Zhong, X., Enke, D. (2017),** *A Comprehensive Cluster and Classification Mining Procedure for Daily Stock Market Return Forecasting*; Neurocomputing, 267:152—168.