2. FORECASTING STOCK MARKET DYNAMICS USING BIDIRECTIONAL LONG SHORT-TERM MEMORY[1]

Daehyeon PARK[2]
Doojin RYU[2,*]

Abstract

This study forecasts stock market dynamics using machine learning techniques. Specifically, we use long short-term memory (LSTM) and bidirectional LSTM (Bi-LSTM) networks to predict the spot index return and implied volatility series in the Korean market. The Bi-LSTM model exhibits better out-of-sample forecasting performance than the LSTM and classic autoregressive models do, reflecting the fact that the Bi-LSTM model learns data patterns more accurately through a bidirectional process. The Bi-LSTM model with the longest time lag (i.e., 22 days) exhibits the best performance in predicting returns and volatility over the entire sample period. In contrast, during the global financial crisis and COVID-19 pandemic periods, when the stock market dynamics are unstable, Bi-LSTM models with shorter time lags (i.e., five or ten days) predict volatility more accurately.

Keywords: Bidirectional long short-term memory; Forecasting; Machine learning; Implied volatility; Stock return

JEL Classification: C14, C45, G17

1. Introduction

Forecasting asset price dynamics is essential for portfolio selection, risk management, hedging, derivative pricing, and investment strategy design. Accordingly, economics and empirical finance studies continually attempt to improve forecasts of stock market returns and volatilities and explain their dynamics (Chun, Cho, and Ryu, 2019, 2020; Han, Kutan, and Ryu, 2015; Kim and Ryu, 2015a, 2015b; Kim and Ryu, 2020; Lee and Ryu, 2018; Lee, Lee, and Ryu, 2019; Yang, Kim, Kim, and Ryu, 2018). In their classic study, Box and Jenkins (1970) suggest estimating a univariate time-series model using past values of a time series, as in the autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) methods. The Box-Jenkins methods have been modified in various forms,

[1] *Acknowledgement: The authors appreciate helpful comments and suggestions from Corina Saman (Editor-in-Chief).*

[2] *College of Economics, Sungkyunkwan University, Seoul, Republic of Korea.*

[*] *Corresponding author, E-mail: sharpjin@skku.edu.*

such as models reflecting seasonality or fractional integration. In particular, the autoregressive conditional heteroskedasticity (ARCH) and generalized ARCH (GARCH) models are considered appropriate for explaining volatility clustering in financial markets (Park, Ryu, and Song, 2017; Ryu and Shim, 2017; Shim, Kim, Kim, and Ryu, 2015; Shim, Kim, and Ryu, 2017; Song, Park, and Ryu, 2018). However, the Box-Jenkins methods are simple linear models that have limitations in predicting complicated and nonlinear structures. They also require the strict stationarity of the time-series data for reliable estimation.

A deep learning (or machine learning) algorithm can be a strong alternative candidate for forecasting financial market dynamics more effectively. Chakraborty, Mehrotra, Mohan, and Ranka (1992) suggest using neural networks to forecast nonlinear and trend-dependent time series data. They claim that neural networks yield better forecasting performance relative to classical time-series models, such as Box-Jenkins and GARCH-type models. With the rapid development of computing technology, neural network algorithms are being used to forecast various financial market dynamics (Kim, Cho, and Ryu, 2020; Park and Ryu, 2021). Neural networks can be transformed into various forms to improve forecasting performance. In particular, the recurrent neural network (RNN) algorithm has remarkably improved the ability to process sequential data (Nelson, Pereira, and Oliveira, 2017). The RNN algorithm has a structure in which past information is transmitted recurrently. However, It also has a gradient vanishing problem, as further learning is no longer performed at a certain stage when the time lag becomes longer. This problem, in turn, leads to the long-term dependency problem, whereby past learning results disappear.

Hochreiter and Schmidhuber (1997) propose a long short-term memory (LSTM) structure that can learn long-term dependence to overcome the gradient vanishing problem. The LSTM network, an efficient machine-learning method, uses the gated memory technique within a circulatory neural network to adjust the weight of information from the previous cell during its learning process. Through this process, the LSTM model overcomes the problem of the long-term dependency of data. The LSTM model is a state-of-the-art mechanism in the field of financial forecasting and is used to forecast future stock prices and volatilities (Baek and Kim, 2018; Cao, Li, and Li, 2019; Chen, Zhou, and Dai, 2015; Fischer and Krauss, 2018; Sezer, Gudelek, and Ozbayoglu, 2020). This model is well adapted for financial data forecasting, as it dominates in predicting prices and trends (Kim, Cho, and Ryu, 2021a, 2021b; Kim and Won, 2018; Zhou, Han, Xu, Jiang, and Zhang, 2019). Further development of the LSTM network is the bidirectional LSTM (Bi-LSTM) model, which learns sequential data bidirectionally. The Bi-LSTM model is known to detect data patterns better than a unidirectional LSTM model can (Althelaya, El-Alfy, and Mohammed, 2018).

This study examines whether the Bi-LSTM and LSTM models can effectively forecast asset market dynamics in the Korean stock market, a representative emerging market. We forecast daily data on the Korea Composite Stock Price Index 200 (KOSPI 200) and the volatility of the KOSPI 200 (VKOSPI) using Bi-LSTM, LSTM, and autoregressive (AR) models. We consider three time lags: one week (i.e., five trading days), two weeks (i.e., ten trading days), and one month (i.e., 22 trading days). To investigate whether the Bi-LSTM and LSTM models perform well even during highly volatile or crisis periods, such as the global financial crisis and COVID-19 pandemic periods, we analyze and compare their forecasting performances for relevant subsample periods. Our findings are as follows. First, we find that the Bi-LSTM model exhibits the best performance in terms of forecasting both stock market returns and volatilities. Second, the Bi-LSTM model also maintains satisfactory prediction performance in subperiods of high volatility. Third, although the Bi-LSTM model with a long time lag (i.e., 22 trading days) exhibits the best forecasting performance for the entire sample, Bi-LSTM models with shorter time lags (i.e., five or ten trading days) perform better during subsample periods of high volatility.

The remainder of this paper is organized as follows. Section II explains our machine learning forecasting methods, the LSTM and Bi-LSTM models. Section III describes the sample data. Section IV shows the forecasting results of the LSTM and Bi-LSTM models and evaluates their performances. Finally, Section V summarizes and concludes the study.

2. Method

Financial market data can be predicted using several machine learning methods, such as the support vector machine and random forest methods. However, RNN-based models are more suitable than other machine learning methods are for dealing with financial time series data. In particular, the LSTM and Bi-LSTM models have excellent predictive power in financial markets. Thus, we focus on using the LSTM and Bi-LSTM models to predict stock market returns and volatilities. The LSTM network is a kind of RNN, that is, a neural network with a structure in which units are recurrently connected. Because this structure maintains past information during the learning process, it can model and capture time-varying (i.e., dynamic) features. Thus, the RNN structure is more appropriate for processing sequential data. However, because basic RNNs cannot continuously remember past data because of the gradient vanishing problem, their performance decreases as the time lag increases (i.e., they have long-term dependencies). The LSTM model is a type of RNN that solves the long-term dependency problem of vanilla RNN using gated memory. Figure 1 shows the structure of an LSTM cell. In the LSTM cell, information from past cells is transferred to the next cell as a cell state (C). The hidden state (h) from the previous cell updates the cell state using three gates. In the LSTM structure, the three gate types are input (i), forget (f), and output gates (o). Equation (1) shows the processes followed by each gate. x, W, b, and σ are input data, the weight matrix, bias, and the sigmoid function, respectively.

$$
\begin{aligned}
i_t &= \sigma\left([x_t, h_{t-1}] \cdot W^i + b_i\right); \\
f_t &= \sigma([x_t, h_{t-1}] \cdot W^f + b_f); \\
o_t &= \sigma([x_t, h_{t-1}] \cdot W^o + b_o).
\end{aligned} \tag{1}
$$

The input and forget gates update the cell state using Equation (2). A candidate for the next cell state ($\tilde{C}$) is derived from input data and the hidden state. The next cell state is then derived based on this candidate and the past cell state. Through this process, unnecessary information contained in the previous cell state is dropped, and the information is updated based on the new input.

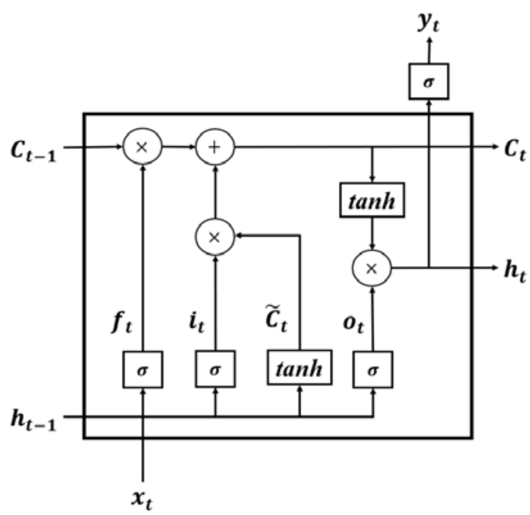
$$
\begin{aligned}
\widetilde{C_t} &= tanh\left([x_t, h_{t-1}] \cdot W^g + b_g\right); \\
C_t &= f_t \cdot C_{t-1} + i_t \cdot \widetilde{C_t}.
\end{aligned} \tag{2}
$$

The next hidden state is derived based on the output gate and the next cell state. Finally, the output vector (y_t) is derived based on the hidden state.

$$
\begin{aligned}
h_t &= \tanh(C_t) \cdot o_t; \\
y_t &= \sigma(h_t).
\end{aligned} \tag{3}
$$

Figure 1

LSTM Cell Structure



LSTM models outperform RNN models, especially in learning and predicting long sequential data. However, LSTM models tend to repeat past data patterns as predicted data because their learning processes heavily depend on historical data patterns. The Bi-LSTM model overcomes this limitation of LSTM models and is known to outperform them (Thireou and Reczko, 2007). The Bi-LSTM model learns the state of each LSTM cell in both the forward and backward directions. Figure 2 shows the structure of the Bi-LSTM model. The boxes labeled *LSTM Cell* in Figure 2 represent the LSTM cell structure shown in Figure 1. The Bi-LSTM model learns the forward (C^F, h^F) and backward states (C^B, h^B) and then synthesizes each hidden state to derive an output vector. We use time lags of one week (lag=5), two weeks (lag=10), and one month (lag=22) to forecast KOSPI returns and the VKOSPI using LSTM and Bi-LSTM models. We also use an AR model as a benchmark. Equation (4) shows the AR(p) model used in this study. y_t is a time-series data point, and β_0 is a constant term. y_t is regressed on past data points in the series (y_{t-i}). p is the same time lag as in the LSTM and Bi-LSTM models, and β_i is the coefficient of each time lag i. ϵ_t is an error term that follows a white noise process with a mean of zero and unit variance ($e_t \sim WN(0,1)$).

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta_i \cdot y_{t-i} + \epsilon_t. \tag{4}$$

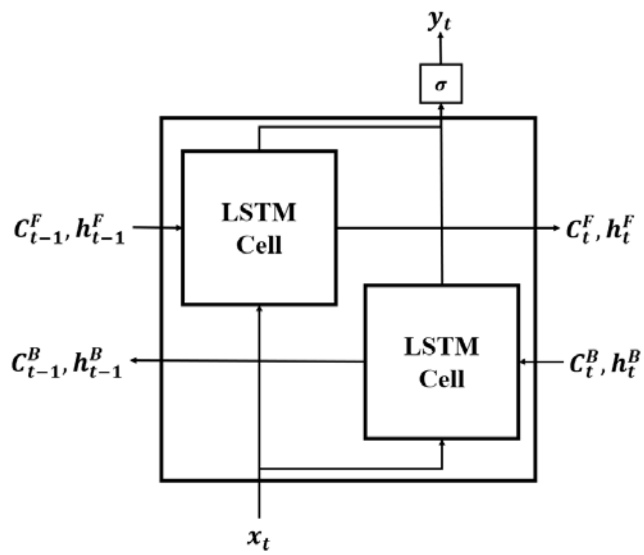
We evaluate forecasting performance using the root mean squared error (RMSE) and mean absolute error (MAE), which are defined as in Equations (5) and (6), respectively. We define them based on the difference between a predicted value ($\hat{x}_{pt}$) and the corresponding real value (x_{rt}). The forecasting period is indicated by T, and RMSE and MAE are defined as the squared and absolute means of the difference, respectively.

$$RMSE = \sqrt{\frac{1}{T}\sum\left(\hat{x}_{pt} - x_{rt}\right)^2}. \tag{5}$$

$$MAE = \frac{1}{T}\sum\left|\hat{x}_{pt} - x_{rt}\right|. \tag{6}$$

Figure 2

Bidirectional LSTM Structure



3. Sample Data

We analyze a long period of daily time series data for the KOSPI 200 and VKOSPI provided by the Korea Exchange. We focus on the dynamics of the Korean financial market for several reasons. First, the Korean market is an important and influential market that affects emerging markets globally (Shim, Kim, Kim, and Ryu, 2016). Second, the Korean financial market is a highly liquid market in which various types of investors actively participate (Yu and Ryu, 2021). Third, because the KOSPI 200 index derivatives markets are world-class markets in terms of their trading volumes, precisely forecasting their underlying return and volatility dynamics is an important issue to global investors.

The entire sample period spans from January 1, 2005, to April 26, 2021. We define two subsamples with high volatility because asset market dynamics differ in highly volatile periods. The first subsample period is the global financial crisis from January 1, 2007, to December 31, 2019. Before and after the 2008 financial crisis, global stock market volatility increased, which affected the Korean stock market. The second subsample covers the COVID-19 pandemic from January 1, 2020 to April 26, 2021. COVID-19, which began spreading at the end of 2019 and became a worldwide pandemic, had a major impact on the global economy. In particular, in February 2020, global stock prices plunged, increasing global stock market volatility. The COVID-19 pandemic also affected the Korean stock market. We analyze these two subsamples to examine whether the Bi-LSTM and LSTM models perform well even during these highly volatile periods.

Table 1 shows the summary statistics of KOSPI 200 returns and the VKOSPI for each sample period. The KOSPI 200 return (r_t={ln (p_t)-ln (p_(t-1))}·100.) at time t is defined as the log return of the KOSPI 200 price (p_t). The VKOSPI is model-free options-implied volatility derived from KOSPI 200 spot and options prices. In Table 1, the columns labeled Mean, Max., Min., and Std.D. show the mean, maximum, minimum, and standard deviation of each variable. The columns labeled ADF and P.-P. show augmented Dickey-Fuller (ADF) and Phillips-Perron statistics, respectively. Both of these statistics are used for the unit root test. KOSPI 200 returns are stationary in all sample periods, but the VKOSPI is stationary only in the whole sample. The standard deviation of KOSPI 200 returns and the mean of the VKOSPI are both higher in the subsample periods than in the whole sample, indicating that the stock market is unstable in those periods.

Table 1

Summary Statistics of KOSPI200 Return and VKOSPI

		Mean	Max.	Min.	Std.D.	ADF	P.-P.
KOSPI200 Return	Whole Sample	0.033	11.54	0.033	1.293	-63.5^{***}	-63.5^{***}
	Subsample1: Financial Crisis	0.024	11.54	0.024	1.916	-27.3^{***}	-27.3^{***}
	Subsample2	0.119	8.755	0.119	1.732	-19.2^{***}	-19.2^{***}
VKOSPI	Whole Sample	20.20	89.3	20.20	9.297	-2.34^{**}	-2.34^{**}
	Subsample1: Financial Crisis	30.65	89.3	30.65	12.83	-1.00	-1.00
	Subsample2	26.21	69.24	26.21	9.026	-0.67	-0.67

Note: The ADF column and P.-P. column show augmented Dickey-Fuller statistics and Phillips-Perron statistics, and ** *and* *** *indicate significance at the 0.05 and 0.01 level.*

4. Results

In this section, we examine out-of-sample forecasts using three models: the Bi-LSTM, LSTM, and AR models. We train each model with 90% of the sample data (January 1, 2005 – September 6, 2019), and we test each model's forecasting performance using 10% of the sample data (September 7, 2019 – April 26, 2021). In the LSTM and Bi-LSTM models, the number of hidden units and the maximum epoch are set to 100 and 50, respectively. We utilize the adaptive moment estimation optimizer. Table 2 shows the three models' forecasting performances, which are evaluated based on the RMSE and MAE values

Table 2 shows that the Bi-LSTM model exhibits better forecasting performance than the other two models do in all cases. In particular, the Bi-LSTM model with a one-month lag (lag=22) performs the best, indicating that the Bi-LSTM model performs better with a larger sample of past data. The LSTM model outperforms the AR model in forecasting the VKOSPI but not in forecasting KOSPI 200 returns. These results reflect that unidirectional LSTM is less capable of learning patterns than bidirectional LSTM is. Figure 3 shows the forecasting results for KOSPI 200 returns.

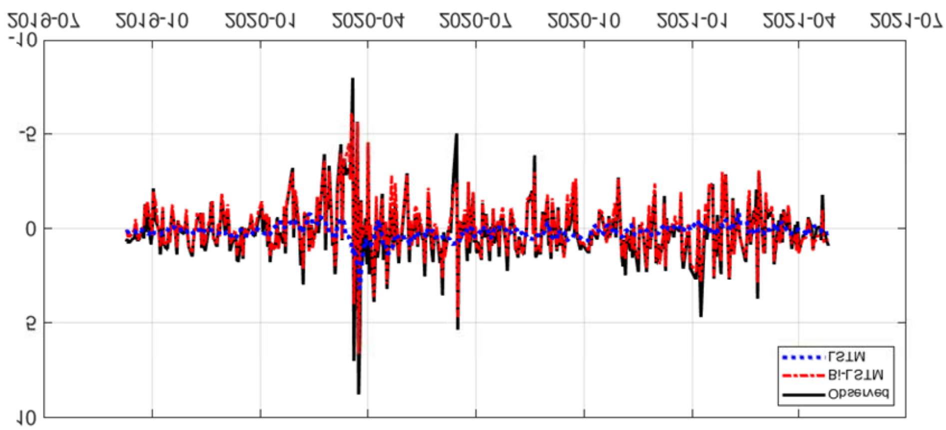
Table 2

Forecasting Performance (Whole sample)

Panel A. KOSPI200 Return						
	lag=5		*lag*=10		*lag*=22	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Bi-LSTM	1.2593	0.8498	1.1935	0.7901	0.5082	0.3320
LSTM	1.6035	1.1242	1.8475	1.2103	1.6017	1.1083
AR	1.6022	1.1189	1.6008	1.1162	1.6124	1.1329
Panel B. VKOSPI						
	lag=5		*lag*=10		*lag*=22	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Bi-LSTM	1.6229	0.9569	1.5228	0.9101	1.0432	0.6379
LSTM	2.0817	1.2678	2.0781	1.2873	2.2304	1.3161
AR	2.0959	2.1002	2.1002	1.2584	2.1376	1.2809

Figure 3

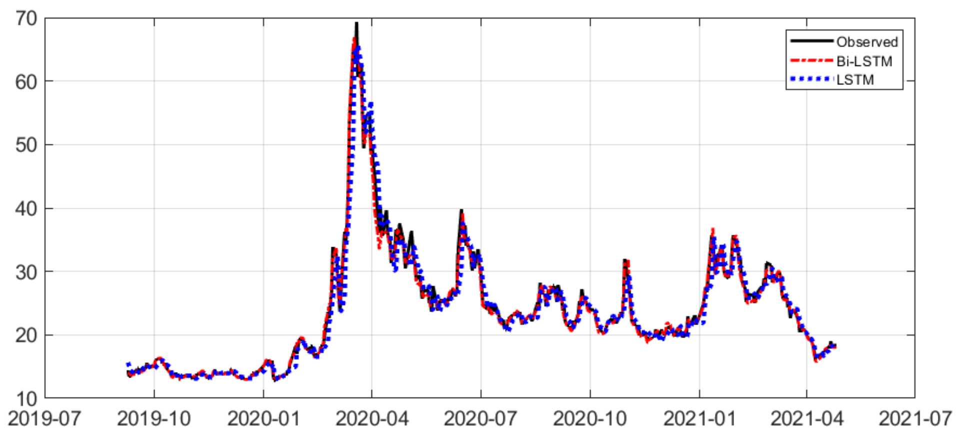
Forecasting of KOSPI200 Returns



Because the LSTM model only learns the patterns of past data, it learns less about data patterns than the Bi-LSTM model does. For this reason, the LSTM model reduces its error by forecasting KOSPI 200 returns with values around zero. This error reduction limits the LSTM model’s ability to improve its forecasting performance. Consequently, the LSTM model performs worse than the AR model does in forecasting returns. Conversely, the Bi-LSTM model predicts the pattern of returns better because it uses a bidirectional process to learn patterns. This difference is also exhibited in the VKOSPI forecasts. Figure 4 shows the VKOSPI forecasting results. In Figure 4, the LSTM model tends to follow actual data by forecasting the previous pattern as it is. Thus, the Bi-LSTM model forecasts the actual data pattern more accurately than the LSTM model does. We conduct the same analysis for two subsample periods in which the stock market is unstable.

Figure 4

Forecasting of VKOSPI



The first subsample period is the global financial crisis from 2007 to 2009. Table 3 shows the forecasting performance using this subsample. In this subsample, the Bi-LSTM model performs the best. As in the whole sample, the Bi-LSTM model performs the best in forecasting KOSPI 200 returns for this subsample when the lag is 22. However, in forecasting the VKOSPI, the Bi-LSTM model performs the best when the lag is 10. We find similar results for the second subsample period, which reflects the COVID-19 pandemic.

Table 3

Forecasting Performance (Subsample1: Global Financial Crisis)

Panel A. KOSPI200 Return						
	lag=5		*lag*=10		*lag*=22	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Bi-LSTM	1.0095	0.6876	0.9349	0.6921	0.7058	0.5696
LSTM	1.3187	1.0416	1.2581	0.9694	1.5929	1.3489
AR	1.2102	0.9239	1.2320	0.9431	1.2329	0.9463

Panel B. VKOSPI						
	lag=5		*lag*=10		*lag*=22	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Bi-LSTM	0.6200	0.4678	0.4960	0.3850	0.5248	0.4137
LSTM	0.8084	0.5682	0.8287	0.6068	0.8282	0.6287
AR	0.8064	0.5966	0.8177	0.6136	0.8348	0.6270

Table 4

Forecasting Performance (Subsample2: COVID-19 Pandemic)

Panel A. KOSPI200 Return						
	lag=5		*lag*=10		*lag*=22	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Bi-LSTM	0.9342	0.7033	0.7029	0.5821	0.6573	0.4976
LSTM	1.2816	1.0814	1.3303	1.1343	0.9768	0.8304
AR	0.7723	0.6150	0.7322	0.5970	0.6903	0.5349
Panel B. VKOSPI						
	lag=5		*lag*=10		*lag*=22	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Bi-LSTM	0.6938	0.4956	0.7211	0.5472	0.9485	0.7280
LSTM	1.0464	0.7872	1.2578	0.9833	2.0086	1.7587
AR	0.9437	0.7757	0.8751	0.7392	0.8539	0.7031

Table 4 shows the results for this subsample. In this subsample, the Bi-LSTM model exhibits the best forecasting performance in all cases. In forecasting KOSPI 200 returns, the Bi-LSTM model with a lag of 22 performs the best, as in the case of the whole sample and first subsample. However, when forecasting the VKOSPI for the second subsample, the Bi-LSTM model with a lag of five performs the best. The results for the two subsamples suggest that the Bi-LSTM model's volatility forecasting performance can be improved by using fewer past data points during high volatility periods. This result seems to be driven by the many volatility fluctuations during high volatility periods. Thus, a shorter time lag leads to more efficient learning about fluctuations in data patterns. With other time series models, the forecasting performance sharply declines when an unexpected shock occurs. However, the Bi-LSTM model learns the pattern after such a shock and, thus, has better forecasting power. These results suggest that the Bi-LSTM model is a suitable forecasting method for future financial market data.

5. Conclusion

This study confirms that machine learning approaches perform well in predicting stock market return and volatility processes. We forecast daily KOSPI 200 returns and the VKOSPI using two representative machine learning methods, the LSTM and Bi-LSTM models. We adopt time lags of one week (i.e., five trading days), two weeks (i.e., ten trading days), and one month (i.e., 22 trading days). Our out-of-sample forecasting analysis shows that the Bi-LSTM model with the longest time lag is the best predictor of both returns and volatilities for the whole sample. Even in highly volatile periods, such as the global financial crisis and COVID-19 pandemic periods, the Bi-LSTM model exhibits the best prediction performance. This result is because the Bi-LSTM model learns data patterns more accurately through a bidirectional learning process. However, in the subsamples with high volatilities, Bi-LSTM models with shorter time lags perform better in volatility forecasting than the model with the longest time lag does, suggesting that bidirectional learning with fewer past observations can improve the Bi-LSTM model's volatility forecasting performance during high volatility periods.

Reference

Ahn, H.-J., Kang, J. and D. Ryu, 2008. Informed trading in the index option market: The case of KOSPI 200 options. *Journal of Futures Markets*, 28(12), pp. 1118-1146.

Althelaya, K.A., El-Alfy, E.S.M. and Mohammed, S., 2018. Evaluation of bidirectional LSTM for short-and long-term stock market prediction. In 2018 9th international conference on information and communication systems (ICICS), pp.151-156.

Baek, Y. and Kim, H.Y., 2018, ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Systems with Applications*, 113, pp. 457–480.

Box, G. and Jenkins, G., 1970, *Time Series Analysis: Forecasting and Control*. Holden-Day.

Cao, J., Li, Z. and Li, J., 2019, Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, 519, pp. 127–139.

Chakraborty, K., Mehrotra, K., Mohan, C.K. and Ranka, S., 1992, Forecasting the behavior of multivariate time series using neural networks. *Neural networks*, 5(6), pp. 961-970.

Chen, K., Zhou, Y. and Dai, F., 2015, A LSTM-based method for stock returns prediction: A case study of China stock market. *In 2015 IEEE international conference on big data*, pp. 2823–2824.

Chun, D., Cho, H. and Ryu, D., 2018, Macroeconomic structural changes in a leading emerging market: The effects of the Asian financial crisis. *Romanian Journal of Economic Forecasting*, 21(2), pp. 22-42.

Chun, D., Cho, H. and Ryu, D., 2019, Forecasting the KOSPI200 spot volatility using various volatility measures. *Physica A: Statistical Mechanics and its Applications*, 514, pp. 156-166.

Chun, D., Cho, H. and Ryu, D. 2020, Economic indicators and stock market volatility in an emerging economy. *Economic Systems*, 44(2), 100788.

Chung, K.H., Park, S.G. and Ryu, D., 2016, Trade duration, informed trading, and option moneyness. *International Review of Economics and Finance*, 44, pp. 395-411.

Fischer, T. and Krauss, C., 2018, Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), pp. 654–669.

Han, H., Kutan, A.M. and Ryu, D., 2015, Effects of the US stock market return and volatility on the VKOSPI. *Economics: The Open-Access, Open-Assessment E-Journal*, 9(2015-35), pp. 1-34.

Hochreiter, S., Schmidhuber, J., 1997, Long Short-Term Memory. *Neural Computation*, 9(8), pp. 1735-1780.

Kim, H., Cho, H. and Ryu, D., 2020, Corporate default predictions using machine learning: Literature review. *Sustainability*, 12(16), 6325

Kim, H., Cho, H. and Ryu, D., 2021a, Corporate bankruptcy prediction using machine learning methodologies with a focus on sequential data. *Computational Economics*, Forthcoming.

Kim, H., Cho, H. and Ryu, D., 2021b, Forecasting consumer credit recovery failure: Classification approaches. *Journal of Credit Risk*, Forthcoming.

Kim, H.Y. and Won, C.H., 2018, Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, 103, pp. 25–37.
Kim, J.S. and Ryu, D., 2015a, Are the KOSPI 200 implied volatilities useful in value-at-risk models? *Emerging Markets Review*, 22, pp. 43-64.
Kim, J.S. and Ryu, D., 2015b, Return and volatility spillovers and cojump behavior between the U.S. and Korean stock markets. *Emerging Markets Finance and Trade*, 51(S1), pp. S3-S17.
Kim, K. and Ryu, D., 2020, Predictive ability of investor sentiment for the stock market. *Romanian Journal of Economic Forecasting*, 23(4), pp. 33-46.
Kim, K., Ryu, D. and Yu, J., 2021, Do sentiment trades explain investor overconfidence around analyst recommendation revisions? *Research in International Business and Finance*, 56, 101376.
Lee, G. and Ryu, D., 2018, Asymmetry in the stock price response to macroeconomic shocks: Evidence from the Korean market. *Journal of Business Economics and Management*, 19(2), pp. 343-359.
Lee, J. and Ryu, D., 2019, The impacts of public news announcements on intraday implied volatility dynamics. *Journal of Futures Markets*, 39(6), pp. 656-685.
Lee, J., Ryu, D. and H. Yang, 2021, Does vega-neutral options trading contain information? *Journal of Empirical Finance*, 62, pp. 294-314.
Lee, J., Lee, G. and Ryu, D., 2019, The difference in the intraday return-volume relationships of spot and futures: A quantile regression approach. *Economics: The Open-Access, Open-Assessment E-Journal*, 13(2019-26), pp. 1-38.
Lee, J., Kang, J. and Ryu, D., 2015, Common deviation and regime-dependent dynamics in the index derivatives markets. *Pacific-Basin Finance Journal*, 33, pp. 1-22.
Nelson, D.M., Pereira, A.C. and de Oliveira, R.A., 2017, Stock market's price movement prediction with LSTM neural networks. *In 2017 International Joint Conference on Neural Networks*, pp. 1419–1426.
Park S.Y., Ryu, D. and Song, J., 2017, The dynamic conditional relationship between stock market returns and implied volatility. *Physica A: Statistical Mechanics and its Applications*, 482, pp. 638-648.
Park, D. and Ryu, D., 2021, A machine learning-based early warning system for the housing and stock markets. *IEEE Access*, Forthcomming.
Park, S.G. and Ryu, D., 2019, Speed and trading behavior in an order-driven market. *Pacific-Basin Finance Journal*, 53, pp. 145-164.
Ryu, D., 2011, Intraday price formation and bid-ask spread components: A new approach using a cross-market model. *Journal of Futures Markets*, 31(12), pp. 1142-1169.
Ryu, D., 2013, What types of investors generate the two-phase phenomenon? *Physica A: Statistical Mechanics and its Applications*, 392(23), pp. 5939-5946.
Ryu, D. and H. Shim, 2017, Intraday dynamics of asset returns, trading activities, and implied volatilities: A trivariate GARCH framework. *Romanian Journal of Economic Forecasting*, 20(2), pp. 45-61.
Ryu, D. and Yang, H., 2018, The directional information content of options volumes. *Journal of Futures Markets*, 38(12), pp. 1533-1548.
Ryu, D. and Yang, H., 2019, Who has volatility information in the index options market? *Finance Research Letters*, 30, pp. 266-270.

Ryu, D. and Yang, H., 2020, Noise traders, mispricing, and price adjustments in derivatives markets. *European Journal of Finance*, 26(6), pp. 480-499.
Ryu, D. and Yang, H., 2021, Intraday option price changes and net buying pressure. *Applied Economics Letters*, forthcoming.
Ryu, D. and Yu, J., 2020, Hybrid bond issuances by insurance firms. *Emerging Markets Review*, 45, 100722.
Ryu, D. and Yu, J., 2021, Informed options trading around holidays. *Journal of Futures Markets*, 41(5), pp. 658-685.
Ryu, D., Ryu, D. and Yang, H., 2021, The impact of net buying pressure on index options prices. *Journal of Futures Markets*, 41(1), pp. 27-45.
Ryu, D., Ryu, D. and Hwang, J.H., 2017, Corporate governance, product-market competition, and stock returns: Evidence from the Korean market. *Asian Business and Management*, 16(1-2), pp. 50-91.
Seo, S.W., Kim, J.S. and Ryu, D., 2019, Effects of the Asian financial crisis on the relation between leverage and employee compensation. *Spanish Journal of Finance and Accounting*, 48(1), pp. 1-20.
Sezer, O.B., Gudelek, M.U. and Ozbayoglu, A.M., 2020, Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, 90, 106181.
Shim, H., Chung, C.Y. and Ryu, D., 2018, Labor income share and imperfectly competitive product market. *B.E. Journal of Macroeconomics*, 18(1), 20160188 (1-16).
Shim, H., Kim, H., Kim, J.Y. and Ryu, D., 2015, Weather and stock market volatility: The case of a leading emerging market. *Applied Economics Letters*, 22(12), pp. 987-992.
Shim, H., Kim, H., Kim, S. and Ryu, D., 2016, Testing the relative purchasing power parity hypothesis: The case of Korea. *Applied Economics*, 48(25), pp. 2383-2395.
Shim, H., Kim, M.H. and Ryu, D., 2017, Effects of intraday weather changes on asset returns and volatilities. *Proceedings of Rijeka Faculty of Economics: Journal of Economics and Business*, 35(2), pp. 301-330.
Sim, M., Ryu, D. and Yang, H., 2016, Tests on the monotonicity properties of KOSPI 200 options prices. *Journal of Futures Markets*, 36(7), pp. 625-646.
Song, J. and Ryu, D., 2016, Credit cycle and balancing the capital gap: Evidence from Korea. *Economic Systems*, 40(4), pp. 595-611.
Song, W., Ryu, D. and Webb, R.I., 2016, Overseas market shocks and VKOSPI dynamics: A Markov-switching approach. *Finance Research Letters*, 16, pp. 275-282.
Song, W., Ryu, D. and Webb, R.I., 2018, Volatility dynamics under an endogenous Markov-switching framework: A cross-market approach. *Quantitative Finance*, 18(9), pp. 1559-1571.
Song, W., Park, S.Y. and Ryu, D., 2018, Dynamic conditional relationships between developed and emerging markets. *Physica A: Statistical Mechanics and its Applications*, 507, pp. 534-543.
Thireou, T. and Reczko, M., 2007, Bidirectional long short-term memory networks for predicting the subcellular localization of eukaryotic proteins. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 4(3), 441-446.
Yang, E., Kim, S., Kim, M.H. and Ryu, D., 2018, Macroeconomic shocks and stock market returns: The case of Korea. *Applied Economics*, 50(7), pp. 757-773.
Yang, H., Kutan, A.M. and Ryu, D., 2018, Option moneyness and price disagreements. *Applied Economics Letters*, 25(3), pp. 192-196.

Yang, H., Kutan, A.M. and Ryu, D., 2019, Volatility information trading in the index options market: An intraday analysis. *International Review of Economics and Finance*, 64, pp.412-426.

Yang, H., Ryu, D. and Ryu, D., 2018, Market reform and efficiency: The case of KOSPI200 options. *Emerging Markets Finance and Trade*, 54(12), pp. 2687-2697.

Yang, H., Ahn, H.-J., Kim, M.H. and Ryu, D., 2017, Information asymmetry and investor trading behavior around bond rating change announcements. *Emerging Markets Review*, 32, pp. 38-51.

Yang, H., Choi, H.-S. and Ryu, D., 2017, Option market characteristics and price monotonicity violations. *Journal of Futures Markets*, 37(5), pp. 473-498.

Yang, H., Lee, J. and Ryu, D., 2018, Market depth, domestic investors and price monotonicity violations. *Applied Economics Letters*, 25(10), pp. 688-692.

Yu, J. and Ryu, D., 2019, Predicting banks' subordinated bond issuances. *Romanian Journal of Economic Forecasting*, 22(4), pp. 87-99.

Yu, J. and Ryu, D., 2021. Effectiveness of the Basel III Framework: Procyclicality in the Banking Sector and Macroeconomic Fluctuations. *Singapore Economic Review*, pp. 1-25.

Zhou, Y.-L., Han, R.-J., Xu, Q., Jiang, Q.-J. and Zhang, W.-K., 2019, Long short-term memory networks for CSI300 volatility prediction with Baidu search volume. *Concurrency and Computation: Practice and Experience*, 31(10), e4721.