



CBOE Volatility Index Forecasting under COVID-19: An Integrated BiLSTM-ARIMA-GARCH Model

Min Hyung Park¹, Dongyan Nan^{2,3}, Yerin Kim¹ and Jang Hyun Kim^{1,2,3,*}

¹Department of Applied Artificial Intelligence, Sungkyunkwan University, Seoul, 03063, Korea

²Department of Human-Artificial Intelligence Interaction, Sungkyunkwan University, Seoul, 03063, Korea

³Department of Interaction Science, Sungkyunkwan University, Seoul, 03063, Korea

*Corresponding Author: Jang Hyun Kim. Email: alohakim@skku.edu

Received: 12 June 2022; Accepted: 29 January 2023; Published: 26 May 2023

Abstract: After the outbreak of COVID-19, the global economy entered a deep freeze. This observation is supported by the Volatility Index (VIX), which reflects the market risk expected by investors. In the current study, we predicted the VIX using variables obtained from the sentiment analysis of data on Twitter posts related to the keyword “COVID-19,” using a model integrating the bidirectional long-term memory (BiLSTM), autoregressive integrated moving average (ARIMA) algorithm, and generalized autoregressive conditional heteroskedasticity (GARCH) model. The Linguistic Inquiry and Word Count (LIWC) program and Valence Aware Dictionary for Sentiment Reasoning (VADER) model were utilized as sentiment analysis methods. The results revealed that during COVID-19, the proposed integrated model, which trained both the Twitter sentiment values and historical VIX values, presented better results in forecasting the VIX in time-series regression and direction prediction than those of the other existing models.

Keywords: Forecasting VIX; sentiment analysis; COVID-19; ARIMA; GARCH; bidirectional LSTM

1 Introduction

To meet the increased necessity for measuring market fluctuations, the Chicago Board Options Exchange (CBOE) developed the Volatility Index (VIX) in 1993 [1]. Using the real-time prices of S&P 500 index options, the VIX represents the expected volatility of the financial market for the following 30 days [2]. Known as the “fear gauge of investors,” the VIX reflects the level of perceived risk by investors [3]. It also has an inverse relationship with stock prices [4]. Therefore, it is often highlighted whenever there is a strong geopolitical influence on the global economy. The outbreak of COVID-19 is of interest here as it brought about a severe recession in the global economy [5].

Trends in the VIX have been forecasted in previous research; there have been suggestions for exploiting the arbitrage opportunities in VIX options trading and providing useful references for risk management in volatility derivative markets [1,6,7]. For forecasting, several VIX index values with



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

fixed timesteps and S&P 500 (SPX) option prices are used as explanatory variables for multivariate analyses and multiple regressions. These analyses are commonly used as tools for prediction in the general financial market [1,6,7].

In the finance sector, sentiment analysis on social media text has been used in some studies to predict stock prices [8,9]. Given that the VIX and stock prices have a statistically significant relationship [4], sentiment analysis can be employed for VIX forecasting. Accordingly, this study used sentiment features as variables for VIX prediction.

Moreover, the outbreak of the COVID-19 pandemic has left an unprecedented impact on people globally, leading to a high frequency of social media posts with keywords related to COVID-19, clearly describing the general sentiments of people. These posts have been used in multiple studies to examine the correlation between public sentiment and financial market prices, such as stock or bitcoin prices [10–13].

According to this change in the financial market, sentiment features were extracted from Twitter posts related to COVID-19 and were utilized for VIX forecasting. Between December 1, 2019, and August 5, 2020, 1,000 daily Twitter posts were collected; these data were collected from December 2019, when the first case of COVID-19 was reported. After preprocessing, Linguistic Inquiry and Word Count (LIWC) and Valence Aware Dictionary and Sentiment Reasoner (VADER) were employed for sentiment analysis. LIWC has been used in existing research for the sentiment analysis of Twitter data [14–16] and for predicting stock prices [9,17]. Similarly, VADER has been used in previous research for forecasting cryptocurrency prices and stock prices using social data [17–19].

In the current study, two different analysis methods were implemented: time-series regression prediction and direction prediction. For the time-series prediction, several time-series neural networks—such as the bidirectional long short-term memory (BiLSTM), bidirectional gated recurrent unit (BiGRU), long short-term memory (LSTM), gated recurrent unit (GRU), and attention-BiLSTM—were used as the base model in the experiments. These models were utilized to forecast time-series data in finance areas [20–24]. Furthermore, the time-series forecasting model, the BiLSTM, was combined with multivariate time-series forecasting models such as autoregressive integrated moving average (ARIMA) and generalized autoregressive conditional heteroskedasticity (GARCH). ARIMA has been applied to the prediction of financial time-series data [25–27] and GARCH can be combined with ARIMA to improve forecasting performance [27]. Inspired by the studies that unified linear and non-linear models for forecasting financial data [25,26], we considered integrating BiLSTM, ARIMA, and GARCH. The integrated BiLSTM-ARIMA-GARCH model outperformed the single BiLSTM model trained with sentiment features, a single statistical model, and a model that combined only statistical models. The integrated BiLSTM-ARIMA-GARCH model was also modified to implement binary classification to predict the direction of VIX movement.

The contributions of the current study are as follows:

First, this study shows sentiment reflected in social media texts to be an effective feature to predict the financial volatility index in the early stage of a pandemic.

Second, our suggested model is considered to be efficient enough to implement daily prediction. Though it consists of multiple models including neural networks, it requires low computations for both training and inference and takes less than an hour without a GPU.

Third, it is challenging to accurately forecast the steep rise and fall periods resulting from the outbreak of the pandemic. We present a way to improve the prediction of unexpected patterns,

including the steep rise and fall, resulting from the outbreak of the pandemic: adding statistical models to neural networks that capture features of direction and volatility from recent trends.

The rest of the paper is organized as follows. This section is followed by the Method section, which elaborates on the data collection, preprocessing, models, and analytical details in the study. Then, the Results section reports the outcomes of the predictive and comparative analyses. Finally, the Conclusion section summarizes the study and provides suggestions for future research.

2 Method

2.1 Data Collection and Preprocessing

We collected data on Twitter posts and CBOE VIX data. These Twitter posts were English posts hashtagged with keywords that are related to COVID-19, such as “COVID,” “COVID19,” “COVID-19,” “pandemic,” “corona,” “corona-virus,” and “covid-death.” The VIX data were collected from Google Finance during the same period the Twitter data were collected. The collected data included 171 days of VIX data, based on the business days when the financial market was open.

Regarding the preprocessing of Twitter posts, the posts with websites were removed from the data as those were considered advertisements. For accurately calculating sentiment scores with LIWC and VADER, which are lexicon-based sentiment analysis methods, additional preprocessing was performed on the word level. The words in each post that were not pronouns, nouns, verbs, adjectives, and adverbs were removed. Then, the remaining words were lemmatized before analysis. Repetitive characters such as “o” in “Things will get better soooooo” were included in some of the posts, which were replaced with the corresponding single character.

VIX time series data, which are the training data of the BiLSTM, were normalized. The sentiment features of the previous four days were used to predict the VIX of the following day using the BiLSTM, based on the trials with different timestep values.

2.2 Sentiment Analysis Features—LIWC and VADER

Using Twitter social data, sentiment analysis scores were generated from the posts and used as features in the VIX index prediction. LIWC and VADER were used for the analysis. LIWC is a text analysis program that shows the computed scores of more than 80 sentiments and other content features using the dictionary; the words here are classified categorically. The main categories include linguistic, emotional, grammatical, and psychological categories; the scores of each word in these categories (e.g., “positive emotion,” “negative emotion,” “anxiety,” “anger,” “sad,” and “social”) are provided [28–30]. VADER is a lexicon and rule-based analytical model that assigns sentiment scores, such as “positive,” “negative,” “neutral,” and “compound,” for text data such as social media data. This method is renowned for showing more accurate information on texts from various domains [31].

2.3 Prediction Model

2.3.1 LSTM and Bidirectional LSTM

Among neural networks, the recurrent neural network (RNN) is widely used as a sequential model to forecast time series data since the model provides the corresponding vector with sequenced input vectors. However, the RNN model suffers from the vanishing and exploding gradients problem. To avoid these issues, the LSTM model was devised and used; this model can train long-term sequence data with deeper neural models without encountering such a problem.

The LSTM model consists of an input gate (i_t); forget gate (f_t); output gate (o_t); and memory cell (c_t); the structure is described in Fig. 1. The three gates employed the output of the previously hidden state vector (h_{t-1}) and input vector (x_t) for computation. The forget gate (f_t) was trained to limit the preservation of the former information from the previous state (h_{t-1}). In the memory cell (c_t), the use of memory information was regulated by the forget gate. The input gate (i_t) supported adding the features to enter the memory cell (c_t). The output gate (o_t) controlled the effect of information from the memory cell (c_t) to yield the output of the current LSTM unit. With the elaborately designed gates and cells, the model can train long-term and short-term sequence data better without input and output weight conflicts [24]. The equations for the forward pass of the LSTM units are defined in Eqs. (1)–(5):

$$f_t = \sigma (U_f x_t + V_f h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma (U_i x_t + V_i h_{t-1} + b_i) \quad (2)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tanh (U_c x_t + V_c h_{t-1} + b_c) \quad (3)$$

$$o_t = \sigma (U_o x_t + V_o h_{t-1} + b_o) \quad (4)$$

$$h_t = o_t \otimes \tanh (c_t) \quad (5)$$

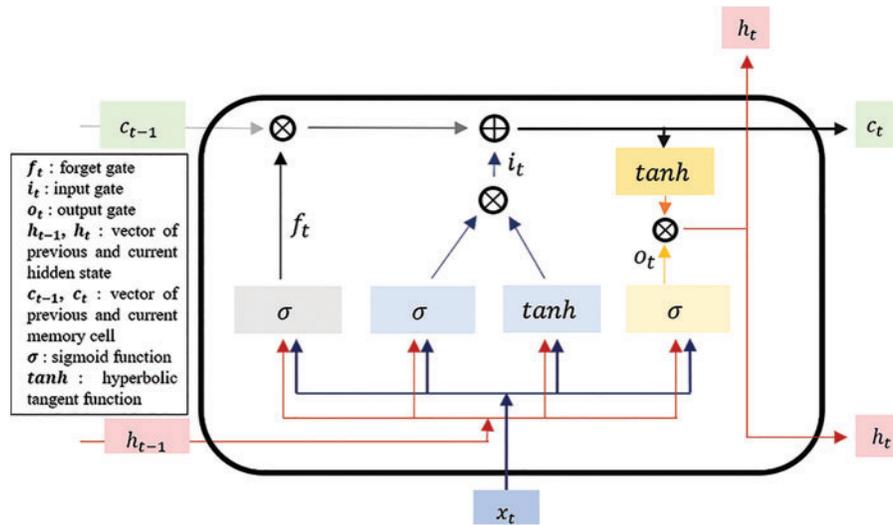


Figure 1: LSTM memory unit

where, U and V are the weight matrices multiplied by the input vector (x_t) and output vector of the previous hidden state (h_{t-1}). The operator σ , \tanh , and \otimes represent the sigmoid function, hyperbolic tangent function, and the Hadamard product, which is also known as the pointwise operation, respectively.

In the BiLSTM, another layer of LSTM units exists, as shown in Fig. 2. One layer is for training the forward information (\vec{h}_t) of the input data and the other is for training the backward information (\overleftarrow{h}_t). The output of the hidden state at the timestep t (y_t) is calculated by multiplying both these pieces

of information with the weight matrices (W). The mathematical expressions of the BiLSTM are written as follows:

$$\vec{h}_t = U_{\vec{h}}x_t + V_{\vec{h}}\vec{h}_{t-1} + \vec{b}_h \tag{6}$$

$$\overleftarrow{h}_t = U_{\overleftarrow{h}}x_t + V_{\overleftarrow{h}}\overleftarrow{h}_{t+1} + \vec{b}_h \tag{7}$$

$$y_t = W_{\vec{h}_y}\vec{h}_t + W_{\overleftarrow{h}_y}\overleftarrow{h}_t + b_y \tag{8}$$

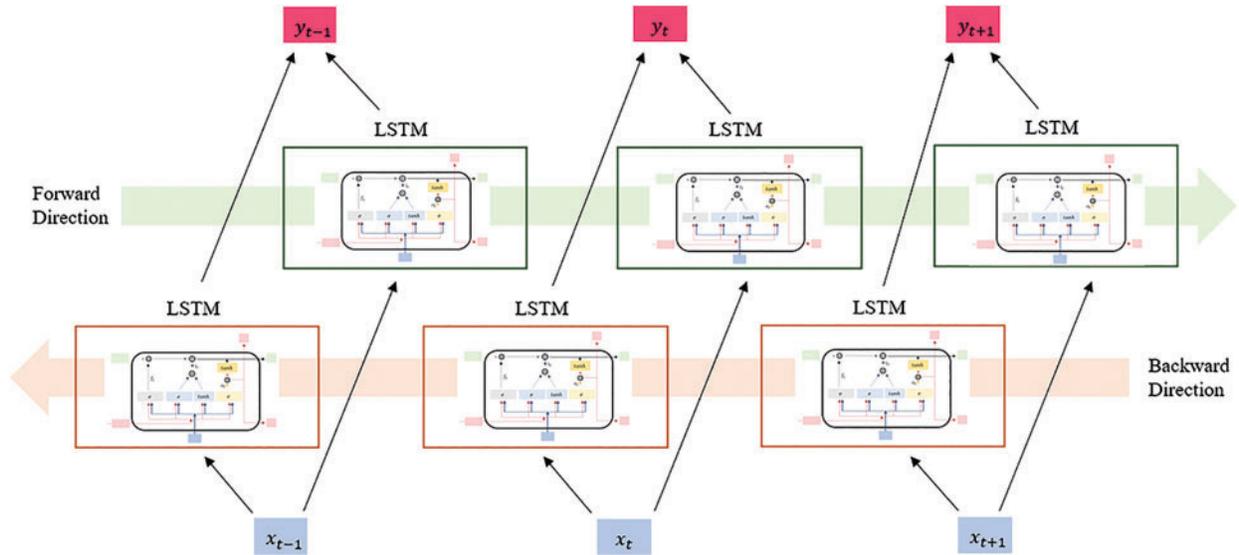


Figure 2: BiLSTM structure

Considering the several models—BiLSTM, BiGRU, LSTM, GRU, and Attention-BiLSTM—used as the base neural network in the experiments, the BiLSTM was chosen to be integrated with the linear time-series prediction models, ARIMA and GARCH, because it demonstrated a better performance than the other models. The model consists of 32 nodes of two BiLSTM layers and 16 nodes of two dense layers. The number of nodes (m) for each layer was obtained from the trials with the nodes among $m \in \{16, 32, 64, 128, 256, 512\}$. Dropout layers with a rate of 0.2 were added after each layer. The BiLSTM was trained with a learning rate of 10^{-5} and 500 epochs using the early-stopping method. The ReduceLROnPlateau class was employed for the learning rate scheduler with a patience of three (3) and a factor of 0.1. The minimum learning rate was set to 10^{-8} . The dataset was split into training and validation sets with a ratio of 8:2 through random choice without any shuffling. The timestep was empirically set to four (4) for the input sentiment features.

2.3.2 ARIMA and GARCH

The ARIMA model is a traditional statistics model and has been applied to time-series forecasting in financial fields [25–27]. ARIMA is a generalized form of the autoregressive moving average (ARMA) model, which combines the autoregressive (AR) and moving average (MA) models, which perform data differencing. The AR(p) model, an autoregressive model of order p , is given in Eq. (9):

$$Y_t = \sum_{k=1}^p \delta_k y_{t-k} + \omega_t \tag{9}$$

where, $\omega_t \sim N(0, \sigma^2)$, and $|\delta_k| < 1$ to ensure that the time series data are stationary. The model parameters (δ) of AR(p) can be estimated to predict the value of the present period (Y_t) with data on the past periods (y_{t-k}). These data can be explained using the linear combination of the past values of the variables, which can be considered autoregression. When δ is equal to zero, the data are considered white noise. They are also considered stationary, meaning that the data of each state are independent of the values of the other time states of data. When $|\delta_k| \geq 1$, the data are non-stationary.

The moving average model, MA(q), explains Y_t with the mean of the time series (μ) and white noise error term (ω) of each period. It is a linear regression form between the predicted error of the term and the estimated model parameters. MA(q) is defined in Eqs. (10) and (11):

$$Y_t = \sum_{k=1}^q \psi_k \omega_{t-k} + \omega_t \quad (10)$$

$$\omega_{t-k} = Y_{t-k} - \hat{Y}_{t-k} \quad (11)$$

The terms p and q are defined through the Akaike Information Criterion (AIC) value, which is known to increase with lower values. The equation for the AIC is stated in Eq. (12):

$$AIC = 2pr - 2 \ln(\widehat{ML}) \quad (12)$$

where, pr represents the values of the parameters of the model through the estimation, and \widehat{ML} denotes the maximum value from the likelihood function. Based on the AIC, an ARIMA model with an order of (1, 0, 2) was employed for the prediction.

The mathematical expression of ARIMA (1, 0, 2) can be restated in Eq. (13):

$$Y_t = \delta_1 y_{t-1} + \psi_1 \omega_{t-1} + \psi_2 \omega_{t-2} + \omega_t \quad (13)$$

The GARCH model captures the feature of the variance of the time series data. Owing to the significance of the risk, GARCH is commonly used in studies on the financial market. The model explains the volatility at the time t (σ_t^2) with the squared residual returns (y_{t-k}^2) and the squared past volatility (σ_{t-k}^2). The GARCH model is a generalized form of the autoregressive conditional heteroskedasticity (ARCH) model, and it can be expressed in Eq. (14):

$$\sigma_t^2 = \eta + \sum_{k=1}^q \alpha_k y_{t-k}^2 + \sum_{k=1}^p \beta_k \sigma_{t-k}^2 \quad (14)$$

where, $\eta > 0, \alpha_k, \beta_k \geq 0, \sum_{k=1}^q \alpha_k + \sum_{k=1}^p \beta_k > 1, \eta \sim N(0, \sigma^2)$ makes the output value of the model always positive; η, α , and β are the model parameters generally estimated through the maximum likelihood estimation method.

According to the related studies [1,32–34], GARCH (1, 1) is dominantly used to the forecast VIX index or compared with other forecasting models on VIX prediction. Therefore, GARCH (1, 1) is considered to be the most suitable for the forecasting VIX index among the standard GARCH models with other parameter settings. GARCH (1, 1) is defined in Eq. (15):

$$\sigma_t^2 = \eta + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (15)$$

2.3.3 Integrated Bidirectional LSTM-ARIMA-GARCH Model

The BiLSTM trained with sentiment features and ARIMA were unified to capture the linear and non-linear patterns of the data to predict the target. Although the models are different, the cases of integrating ARIMA and non-linear models are shown in previous studies in the finance sector [25,26]. Based on the idea from a study in the financial domain [27], we also combined GARCH with ARIMA to help the model capture volatility features from recent target trends. The unified model

consists of three base models; the new artificial neural network (ANN) model was added to it. The additional network took the outputs of the BiLSTM, ARIMA, and GARCH models as new input features (Fig. 3). Then, the output of the ANN was used as the final prediction value. The ANN, which was combined later, had two dense layers (32, 16 nodes): two dropout layers after each layer and an output layer. The dense layers, except for the output layer, used the rectified linear unit (ReLU) as the activation function and He initialization. The other hyperparameters were applied in the same manner as those in the BiLSTM.

Our model requires low computation to implement daily prediction. The number of parameters of the neural networks in our model is approximately 0.28M, which is 226× smaller than that of the base model of Transformer [35], which is well-known for using sequence data. In terms of using a light model, the optimization is generally completed quickly. Although the maximum number of training epochs was 500, the integrated model stopped early after 60 epochs, and took less than an hour, even when not using a GPU. Based on the details above, the integrated BiLSTM-ARIMA-GARCH could be regarded as a significantly efficient model to perform daily forecasting.

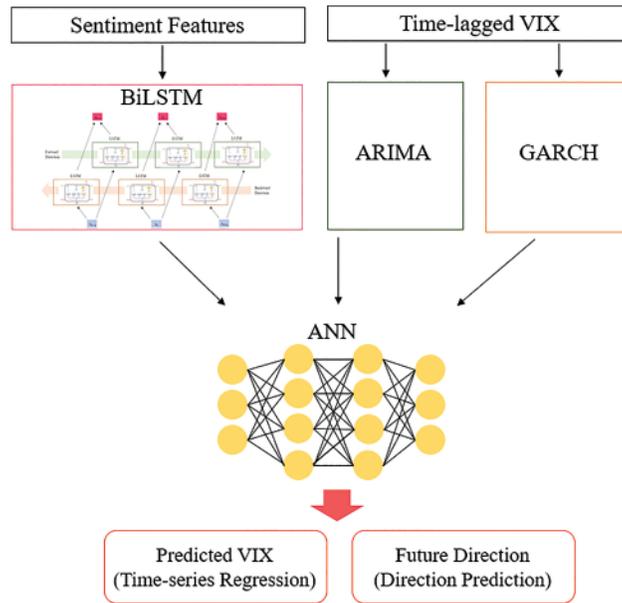


Figure 3: Structure of integrated bidirectional LSTM-ARIMA-GARCH

2.4 Evaluation Metrics

As shown in recent studies on VIX forecasting [1,21,36], the root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) were employed to evaluate the ability of the model for time-series regression prediction. The three metrics are defined in Eqs. (16)–(18):

$$RMSE = \sqrt{\frac{1}{M} \sum_{t=1}^M y_t - \hat{y}_t} \tag{16}$$

$$MAE = \frac{1}{M} \sum_{t=1}^M |y_t - \hat{y}_t| \tag{17}$$

$$MAPE = \frac{1}{M} \sum_{t=1}^M \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (18)$$

where, M represents the number of the samples, y_t represents the true index value, and \hat{y}_t represents the predicted value from the model.

For the direction prediction, the metrics precision, recall, and F1-score were utilized to measure the classification performance. Each metric can be calculated by Eqs. (19)–(21):

$$Precision = \frac{TP}{TP + FP} \quad (19)$$

$$Recall = \frac{TP}{TP + FN} \quad (20)$$

$$F1 = \frac{2}{precision^{-1} + recall^{-1}} \quad (21)$$

3 Results

3.1 Defining ARIMA and GARCH Model

Before using the 'ARIMA-GARCH' model, the augmented Dickey–Fuller test (ADF) was implemented with the VIX data to check if the data were stationary. An ARIMA, with the order of $(p, 0, q)$, was utilized for the prediction based on the ADF statistics value, p-value, and critical values, implying that time-series data were applied with no differencing.

Considering the autocorrelation function (ACF) (Fig. 4) and partial autocorrelation function (PACF) plots (Fig. 5), the values of p and q are limited as $p \in \{0, 1, 3, 7, 8\}$ and $q \in \{0, 1, 2\}$, respectively.

Through the AIC comparison among the models with the specified order of the terms, $(1, 0, 2)$ was employed for the ARIMA with the lowest AIC of 798.726. For the GARCH model, GARCH (1, 1) was adopted based on a recent study on the VIX [36].

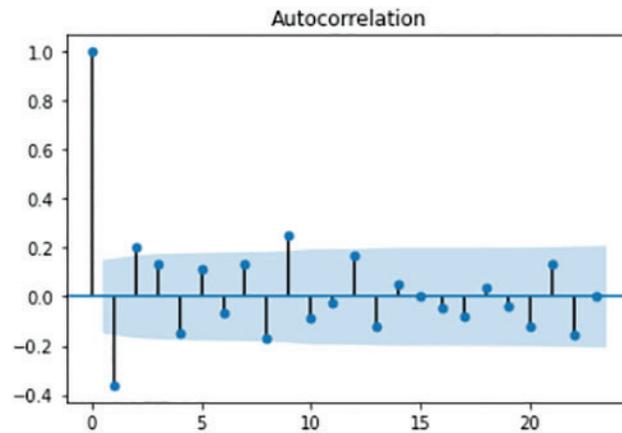


Figure 4: ACF

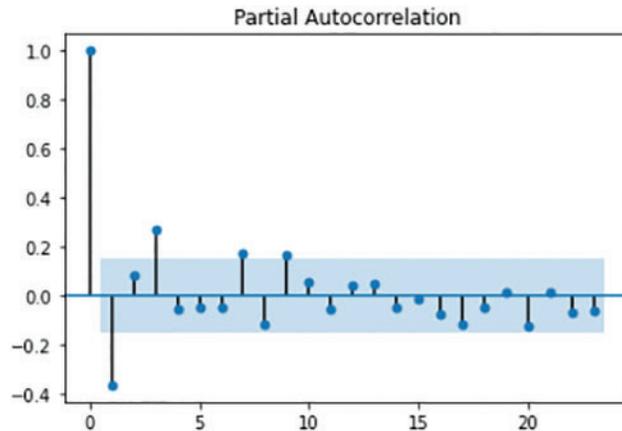


Figure 5: PACF

3.2 Time-Series Regression Prediction

In this section, the predictions of the single BiLSTM were compared with those of the integrated model. The single BiLSTM with the test data returned MAPE values ranging from 12 to 14, indicating a low forecasting error. However, when predicting the VIX of the whole period, the model was found to underfit the data for the overall period (Fig. 6).

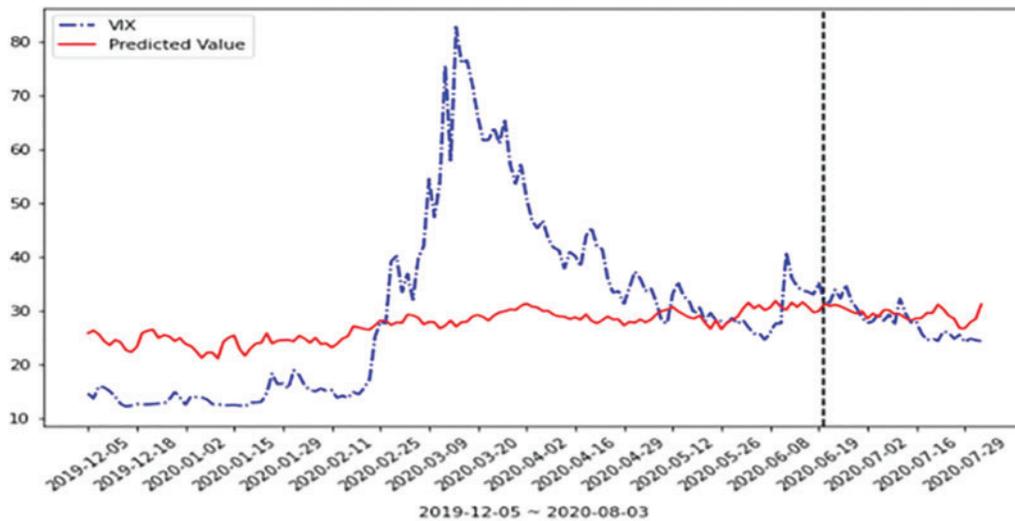


Figure 6: Predictions of single BiLSTM

Note: Dotted line indicates the boundary between the training and test data.

The underfitting of data was the limitation of using only the non-linear model. To resolve this problem, the model was combined with models that could add the linear trends and features of the target data. Applying the multivariate statistical models ARIMA and GARCH, the hybrid model fitted the overall data better, as shown in Fig. 7.

Compared to the results of the existing studies that forecasted the VIX in the COVID-19 era [36], the integrated model showed significant progress in fitting the dynamic data patterns. The model

trained the data pattern mapping the steep rise between February and March. Additionally, it also learned the long-term decline from March to May and the reascent that was maintained until early June. Even with the unseen data, the integrated model showed competitive results in predicting the falling phase of the test data.

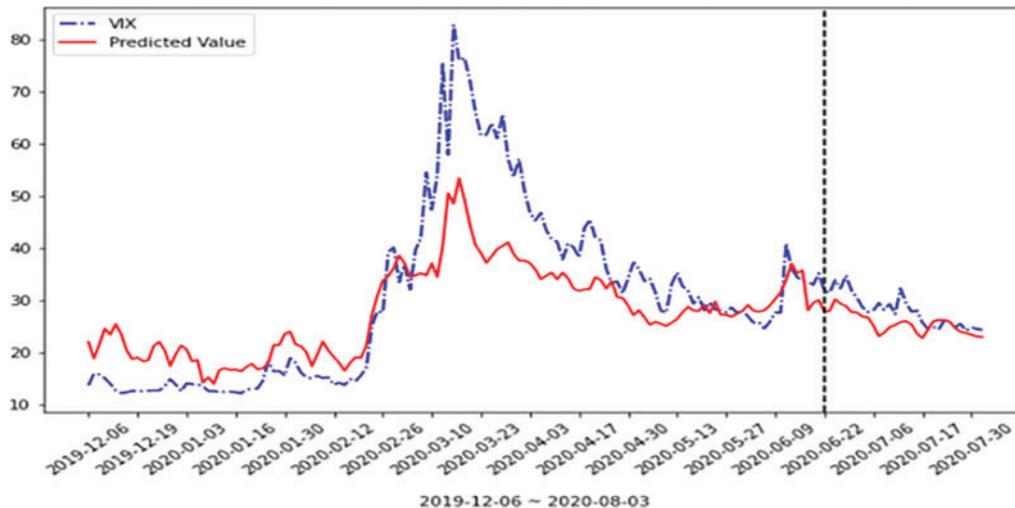


Figure 7: Predictions of the integrated model

We experimented using the hybrid model with other non-linear base models; the observed results are shown in [Table 1](#). We found that the integrated model having the BiLSTM showed the best results among all the models. There was no improvement even after applying the attention mechanism to the BiLSTM.

Table 1: Regression test results of integrated model with different time series base models

	RMSE	MAE	MAPE	Improvement (%)
BiLSTM	3.065	2.494	9.360	–
BiGRU	5.070	4.003	13.989	33.09
LSTM	5.339	4.596	17.548	46.66
GRU	7.772	6.694	25.260	62.95
Attention-BiLSTM	8.040	6.955	26.588	64.80

Note: Improvement values imply MAPE decreases when using our final model.

The integrated model was also compared to the ARIMA-GARCH combined model, which does not use sentiment analysis features for training. The results shown in [Table 2](#) indicated that our integrated BiLSTM-ARIMA-GARCH model was better than the other models. This improved performance can be attributed to the ARIMA and GARCH usually working well with short-term prediction because of the convergence of the predicted values. The base neural network, i.e., the BiLSTM, which learns features with non-linear operations, seems to overcome this weakness in our integrated model by predicting the unfitted data.

Table 2: Comparison of test results between the integrated model and ARIMA based models

	RMSE	MAE	MAPE	Improvement (%)
BiLSTM-ARIMA-GARCH	3.065	2.494	9.360	–
ARIMA	3.014	2.592	9.658	3.09
ARIMA-GARCH	4.824	3.818	12.752	26.60

Note: Improvement values imply MAPE decreases when using our final model.

3.3 Direction Prediction

With the same trained model, we evaluated the model for classifying the VIX future direction of increase or decrease (Table 3) using the metrics of precision, recall, and F1-score. The prediction for the upward trend in the VIX index was slightly better than that for the downward trend.

Table 3: Classification test results of the future direction of the VIX using the integrated model

	Precision	Recall	F1-score
Increase	0.65	0.61	0.63
Decrease	0.42	0.45	0.43

4 Conclusion

This study predicted the global volatility index in the early stage of the pandemic by using sentiments in social media texts. The sentiment information of the texts was extracted through two sentiment analysis methods: (1) LIWC, which is used to extract variable sentiments from text and (2) VADER, which is recognized to accurately analyze sentiments from texts from variable domains. The BiLSTM model, which learned sentimental features, was proven to be effective for the prediction of volatility index in that the integration of the model showed a better performance than when using only a single statistical model (i.e., ARIMA) or combining statistical models (i.e., ARIMA-GARCH).

Furthermore, by integrating the sequence neural network model with the traditional statistics models, the non-linear features from the sentiment data and the linear trends of the target values were utilized simultaneously. Using the three models, the underfitting problem was resolved, and the integrated model fitted the data patterns better for the entire period.

Even though the integrated model consists of multiple models, training and inference are completed quickly enough to support daily forecasting. Since neural networks in the unified model have a small number of parameters, the optimization requires low computations and is completed quickly without using a GPU.

The integrated BiLSTM-ARIMA-GARCH model, which used only social media sentiment data and the historical values of the target, showed lower forecasting errors in regression prediction compared to those shown in a similar study on VIX prediction conducted during the COVID-19 pandemic [36].

Nevertheless, these results are promising, considering that prior studies used larger amounts of data collected over multiple years than the current study [1,37]. In addition, regarding the forecasting

results for the error of the regression, this study reports that the model can be used to predict the long-term movement of the VIX index.

Since the outbreak of COVID-19 is relatively recent, the integrated model still needs more data that might help the model train the dynamic patterns of the VIX during COVID-19. Such improvement will enable the model to make better predictions for extreme patterns in the future. However, this study proved that the sentiment scores of social media data could be an advantageous independent variable for predicting volatility in the finance market. Additionally, the social media posts related to global issues, such as the pandemic, also seemed to reflect the sentiments of people toward the finance market, eventually affecting the changes in the market itself.

However, future studies still need to consider such potential changes. If global issues continue to persist and the public gets used to them, the explanatory power of sentiments based on social media posts related only to keywords based on global issues might decrease from the initial stage. Therefore, using social media texts related to keywords on both global issues and the research domain is expected to show better results with the prediction task. Social media sentiments can contribute to predictions in diverse areas during the pandemic.

It should be noted that contextualized text representations learned from pre-trained language models, such as ELMo and BERT [38,39], can be utilized for predicting the volatility index, instead of extracted sentiment features. These representations are likely to contain more useful information of the texts besides sentiment features. Therefore, using contextual representations to predict the volatility index under supervised learning could be a meaningful approach in future studies.

Funding Statement: This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (NRF-2020R1A2C1014957).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] G. Qiao, J. Yang and W. Li, "VIX forecasting based on GARCH-type model with observable dynamic jumps: A new perspective," *The North American Journal of Economics and Finance*, vol. 53, no. 4, pp. 101186, 2020.
- [2] R. E. Whaley, "Understanding the VIX," *The Journal of Portfolio Management*, vol. 35, no. 3, pp. 98–105, 2009.
- [3] V. Aggarwal, A. Doifode and M. K. Tiwary, "Volatility spillover from institutional equity investments to Indian volatility index," *International Journal of Management Concepts and Philosophy*, vol. 13, no. 3, pp. 173–183, 2020.
- [4] N. Gupta and A. Kumar, "Macroeconomic variables and market expectations: Indian stock market," *Theoretical and Applied Economics*, vol. 27, no. 3, pp. 161–178, 2020.
- [5] World Bank, "The global economic outlook during the COVID-19 pandemic: A changed world," 2020. [Online]. Available: www.worldbank.org/en/news/feature/2020/06/08/the-global-economic-outlook-during-the-covid-19-pandemic-a-changed-world
- [6] A. Thavaneswaran, Y. Liang, S. Das, R. K. Thulasiram and J. Bhanushali, "Intelligent probabilistic forecasts of VIX and its volatility using machine learning methods," in *2022 IEEE Symp. on Computational Intelligence for Financial Engineering and Economics (CIFER)*, Helsinki, Finland, pp. 1–8, 2022.
- [7] J. Osterrieder, D. Kucharczyk, S. Rudolf and D. Wittwer, "Neural networks and arbitrage in the VIX," *Digital Finance*, vol. 2, no. 1, pp. 97–115, 2020.

- [8] T. H. Nguyen, K. Shirai and J. Velcin, "Sentiment analysis on social media for stock movement prediction," *Expert Systems with Applications*, vol. 42, no. 24, pp. 9603–9611, 2015.
- [9] B. S. Kumar, V. Ravi and R. Miglani, "Predicting Indian stock market using the psycho-linguistic features of financial news," *Annals of Data Science*, vol. 8, no. 3, pp. 517–558, 2019.
- [10] C. Chen, L. Liu and N. Zhao, "Fear sentiment, uncertainty, and bitcoin price dynamics: The case of COVID-19," *Emerging Markets Finance and Trade*, vol. 56, no. 10, pp. 2298–2309, 2020.
- [11] M. Costola, M. Nofer, O. Hinz and L. Pelizzon, "Machine learning sentiment analysis, COVID-19 news and stock market reactions," *SAFE Working Paper*, no. 288, 2020. [Online]. Available: <https://www.econstor.eu/handle/10419/224131>
- [12] Y. Duan, L. Liu and Z. Wang, "COVID-19 sentiment and the Chinese stock market: Evidence from the official news media and Sina Weibo," *Research in International Business and Finance*, vol. 58, no. 4, pp. 101432, 2021.
- [13] H. S. Lee, "Exploring the initial impact of COVID-19 sentiment on US stock market using big data," *Sustainability*, vol. 12, no. 16, pp. 6648, 2020.
- [14] A. Tumasjan, T. O. Sprenger, P. G. Sandner and I. M. Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment," in *Fourth Int. AAAI Conf. on Weblogs and Social Media*, Washington, USA, pp. 178–185, 2010.
- [15] Y. Bae and H. Lee, "Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 12, pp. 2521–2535, 2012.
- [16] S. Volkova, Y. Bachrach and B. Van Durme, "Mining user interests to predict perceived psychodemographic traits on twitter," in *2016 IEEE Second Int. Conf. on Big Data Computing Service and Applications (BigDataService)*, Oxford, UK, pp. 36–43, 2016.
- [17] J. Abraham, D. Higdon, J. Nelson and J. Ibarra, "Cryptocurrency price prediction using tweet volumes and sentiment analysis," *SMU Data Science Review*, vol. 1, no. 3, pp. 1, 2018.
- [18] X. Li, Y. Li, H. Yang, L. Yang and X. Y. Liu, "DP-LSTM: Differential privacy-inspired LSTM for stock prediction using financial news," arXiv preprint arXiv:1912.10806, 2019.
- [19] X. Li, P. Wu and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong," *Information Processing & Management*, vol. 57, no. 5, pp. 102212, 2020.
- [20] K. A. Althelaya, E. S. M. El-Alfy and S. Mohammed, "Stock market forecast using multivariate analysis with bidirectional and stacked (LSTM, GRU)," in *2018 21st Saudi Computer Society National Computer Conf. (NCC)*, Secunderabad, India, pp. 1–7, 2018.
- [21] A. Banerjee, "Forecasting of India VIX as a measure of sentiment," *International Journal of Economics and Financial Issues*, vol. 9, no. 3, pp. 268–276, 2019.
- [22] A. Dutta, S. Kumar and M. Basu, "A gated recurrent unit approach to bitcoin price prediction," *Journal of Risk and Financial Management*, vol. 13, no. 2, pp. 23, 2020.
- [23] Y. Hu, J. Ni and L. Wen, "A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction," *Physica A: Statistical Mechanics and Its Applications*, vol. 557, no. 5, pp. 124907, 2020.
- [24] J. Qiu, B. Wang and C. Zhou, "Forecasting stock prices with long-short term memory neural network based on attention mechanism," *PLoS One*, vol. 15, no. 1, pp. e0227222, 2020.
- [25] Y. Wang and Y. Guo, "Forecasting method of stock market volatility in time series data based on mixed model of ARIMA and XGBoost," *China Communications*, vol. 17, no. 3, pp. 205–221, 2020.
- [26] S. Verma, S. Prakash Sahu and T. Prasad Sahu, "Ensemble approach for stock market forecasting using ARIMA and LSTM model," in *Proc. of Third Int. Conf. on Intelligent Computing, Information and Control Systems*, Singapore, pp. 65–80, 2022.

- [27] M. Zolfaghari and S. Gholami, "A hybrid approach of adaptive wavelet transform, long short-term memory and ARIMA-GARCH family models for the stock index prediction," *Expert Systems with Applications*, vol. 182, no. 4, pp. 115149, 2021.
- [28] J. W. Pennebaker, R. L. Boyd, K. Jordan and K. Blackburn, "The development and psychometric properties of LIWC2015," 2015, Available at: <https://repositories.lib.utexas.edu/handle/2152/31333>
- [29] J. H. Kim, D. Nan, Y. Kim and M. H. Park, "Computing the user experience via big data analysis: A case of uber services," *CMC-Computers, Materials & Continua*, vol. 67, no. 3, pp. 2819–2829, 2021.
- [30] J. H. Kim, H. S. Jung, M. H. Park, S. H. Lee, H. Lee *et al.*, "Exploring cultural differences of public perception of artificial intelligence via big data approach," in *Proc. of Int. Conf. on Human-Computer Interaction*, vol. 1580, pp. 427–432, 2022.
- [31] C. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. of the Int. AAAI Conf. on Web and Social Media*, Ann Arbor, Michigan USA, vol. 8, no. 1, pp. 216–225, 2014.
- [32] P. C. Pati, P. Barai and P. Rajib, "Forecasting stock market volatility and information content of implied volatility index," *Applied Economics*, vol. 50, no. 23, pp. 2552–2568, 2018.
- [33] H. Y. Kim and C. H. Won, "Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models," *Expert Systems with Applications*, vol. 103, no. 1, pp. 25–37, 2018.
- [34] S. Sharma, V. Aggarwal and M. P. Yadav, "Comparison of linear and non-linear GARCH models for forecasting volatility of select emerging countries," *Journal of Advances in Management Research*, vol. 18, no. 4, pp. 526–547, 2021.
- [35] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones *et al.*, "Attention is all you need. Advances in neural information processing systems," in *Proc. of 31st Conf. on Neural Information Processing Systems*, Long Beach, California, USA, pp. 5998–6008, 2017.
- [36] R. Lit, "Forecasting the VIX in the midst of COVID-19," 2020, Available at: https://timeserieslab.com/articles/rlit_vix_v2.pdf
- [37] P. H. Kumar and S. B. Patil, "A hybrid regression and deep learning LSTM based technique for predicting volatility index (VIX) direction of change (trend)," *Indian Journal of Science and Technology*, vol. 11, no. 47, pp. 1–9, 2018.
- [38] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark *et al.*, "Deep contextualized word representations," in *Proc. of the 2018 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, New Orleans, Louisiana, pp. 2227–2237, 2018.
- [39] J. Devlin, M. W. Chang, K. Lee and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018.