Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-019-01398-9

ORIGINAL RESEARCH

BiLSTM model based on multivariate time series data in multiple field for forecasting trading area

Jinah Kim[1] · Nammee Moon[2]

Received: 28 December 2018 / Accepted: 8 July 2019

Abstract
An artificial neural network-based model is widely used for analyzing and predicting multivariate time series data. However, the study on the analysis and prediction of multivariate time series data in multiple fields has limitations in that it does not take the features of the fields into account. In this paper, we propose a Bi-directional Long Short-Term Memory model based on multivariate time-series data in multiple fields that considers the fields' features. This model differs from the existing model in that data input into the input layer is divided into fields to learn the features of those fields. In addition, we tried to learn the trend of the time series data at the same time by simultaneously learning the value of the data and its variation. We applied the model proposed in this paper on trading area forecasts to collect purchasing data and SNS data on "restaurants" in the trading area to progress the learning. Experiments and performance evaluation were performed based on the Root Mean Square Error, which was based on whether the learning was done by each field and whether there was variation in the input value. Experimental results show that the proposed model performed better than other models.

Keywords Bi-directional LSTM · LSTM · RNN · Artificial neural network · Multivariate time series analysis

1 Introduction

Analysis and prediction studies on time series data are being conducted in various fields such as economics, voice recognition, transportation, and medical care. In general, several models based on statistical techniques such as Auto-Regressive (AR), Auto-Regressive Moving Average (ARMA), and Auto-Regressive Integrated Moving Average (ARIMA) have been utilized to characterize and model data patterns. However, most time series data has trends or seasonality, and they also have nonlinear characteristics with irregular fluctuations due to various dependent variables that are influenced by the passage of time. Therefore, these have started to become models based on multivariate time series data that use multiple input values that affect the value being predicted and its value. However, since the analysis of mutual relations between data is required, there is a limit to the utilization of statistical techniques in multivariate time series data. Recently, the development of neural networks such as Convolutional Neural Network (CNN), Random Forest (RF), Support Vector Machine (SVM), Naïve Bayes, etc. have enabled learning, analysis, and prediction using modeling based on this (Zeng et al. 2018; Aydadenta 2018; Lee and Moon 2018; Zouina and Outtaj 2017; Meng et al. 2018; Song et al. 2018). Indeed, these underlying models have shown better performance than existing statistical based models (Siami-Namini and Namin 2018).

In this paper, we propose a Bi-directional Long Short-Term Memory (BiLSTM) model based on multivariate time series data in multiple fields. In this case, a field refers to data that occurs independently of any other and affects the fields as if the electrical energy data and the weather data are related to one another. We describe the process of selecting and classifying multivariate time series data according to field and proceed to learn each field. In addition, predictions are made considering the variation that plays the most important role in classifying and forecasting time series data. In other words, both the value of each time-series data set and the variation in the value are used as candidate features

✉ Nammee Moon
nammee.moon@gmail.com

Jinah Kim
jina9406@gmail.com

1 Department of Computer Engineering, Hoseo University, Asan-si, South Korea

2 Division of Computer Information Engineering, Hoseo University, Asan-si, South Korea

Published online: 13 July 2019

and put into the input layer. In this paper, we apply the proposed model to predict the trading area.

This paper is composed as follows. Chapter 2 explains the related works to this paper. Chapter 3 describes the BiLSTM model based on multivariate time series data in multiple fields presented in this paper in detail. Chapter 4 explains the experimental results for the proposed model and compares the performance evaluation for objective evaluation. Finally, Chapter 5 presents the conclusions and future research direction.

2 Related works

2.1 Artificial neural network model for the prediction of the time series data basis

Among various artificial neural network techniques, the Recurrent Neural Network (RNN)-based technique, which is appropriate for time series data, is required here. RNN is characterized by having a recurrent structure because it inputs previous information and connects it. It is suitable for processing time series data in which the temporal order relation is important because learning is performed by reflecting on past data. Thus, it is mainly used in recognizing cursive letters and voice recognition. It is widely adopted in fields with time series data such as stock predictions, temperature, and precipitation (Yin et al. 2017; Miao et al. 2015; Castrejon et al. 2017). However, RNN no longer continues learning if the process is prolonged. There is a limitation that a long-term dependency problem exists where long past learning results are wiped from certain levels.

Therefore, LSTM is one RNN-based model that can improve on this point. Each node of the hidden layer in the RNN is further configured with memory blocks composed of an input gate, forget gate, and output gate. The input gate prints out 0 or 1 and determines what input value it will receive. If this is 0, it does not affect the results. If this is 1, data is sent to affect the results. The forget gate determines whether to keep the output value by calculating using the previous output value and the input value. Finally, the output gate determines how much will be output. Thus, this structure determines how much of the previous condition values are added, forgotten, and outputted. Since the data flow is controlled by each gate, there is no long-term dependency problem and learning is conducted properly. Thus, as an alternative to RNN, time-series data prediction is actively carried out using an LSTM-based model. You et al. (2018) proposed a neural network model of various types of stacked ensemble based on CNN and LSTM for voice detection. Song and Kim (2018) proposed a novel deep neural network model for detecting human activities in untrimmed videos and for detecting human activities from the sequence of extracted feature vectors, they use BiLSTM, a Bi-directional Recurrent Neural Network (BRNN) model.

Meanwhile, since RNN-based neural network models are only processed in time order, the results tend to reflect the previous state. BRNN is a method of adding hidden layers in the reverse direction to both previous data and data to be learned in the future. It has two separate hidden layers that are not connected in the forward or backward directions. When all input values are transferred to the two hidden layers, both output values of both hidden layers are calculated and thus the final output is calculated. The forward direction is the same as for general RNN. However, input values are input in the forward and backward hidden layers in the opposite direction and output layer values are calculated after all inputs have been applied to hidden layers in both directions. When updating the weights in the reverse direction, Back Propagation Through Time (BPTT) is used similarly to general RNN. However, the difference is that the error value is calculated first for all times in the output layer and then transmitted in the opposite direction to the forward hidden layer and the reverse hidden layer. Data learned in this manner is better than the existing unidirectional model because it uses both previous data and future data. Thus, many BRNN or BiLSTM-based studies are carried out. Ma et al. (2017) proposed a novel model, named Dipole, to address the challenges of modeling the Electronic Health Record (EHR) data and interpreting the prediction results by employing BRNN. Eapen et al. (2019) proposed a novel deep learning model that combined CNN with a BiLSTM artificial neural network model based on multivariate time series data for stock market prediction.

2.2 Artificial neural network model based on time series data in multiple field

The most important thing in the artificial neural network model that uses time series data is the characteristics of the input data. The data should be deeply understood and well reflected in the model. Previously, analyses and prediction have been conducted using only single time-series data. However, as data becomes more diverse and voluminous, multivariate time series data, including data on other related dependent variables, has become important in recent years to reflect the true data characteristics.

The study of multivariate time series data can be classified into two types. The first deals with multivariate time series data for individual fields that occur at the same time. For example, when analyzing and forecasting stock price data, we mainly use data that consists of opening price, closing price, high and low price, and the volume of stock traded. However, in the case of multivariate time series data for one field, it is difficult to fully reflect the features of the time series data because time series data are nonlinear and include trends and other irregularities; therefore, it changes to a prediction model that deals with multivariate time series data in multiple fields affecting each other. Lang et al. (2018) proposed a short-term load forecasting method based on the multivariate time series prediction scheme and the KNNRW model using the electricity load and temperature data. Toubeau et al. (2018) presented a new approach to generate short-term multivariate predictive scenarios using the electric power market and weather data.

These models mainly combine all the data of several fields at a time to learn. It is difficult to derive a feature of the effect of each field and to reflect it. Therefore, learning for this is separately required for each field. And then it is needed to predict by combining all the fields afterwards. Zhuge et al. (2017) proposed an LSTM model for forecasting opening stock prices and combined the results of classification and analysis of the naïve Bayesian-based emotions on forum data. In this model, emotions and stock price data are learned separately, and then they are combined and learned.

In other words, it is required to neural network model based on multivariate time-series data in multiple fields that can find dependent variables that can affect prediction results and performs learning through being combined with other data. Also, learning for each field should be performed separately in order to extract features between fields. Previous studies have entered the artificial neural network model for several series at a time, so they are limited by the fact that the classification of fields is ambiguous and does not reflect their features. This study proposes a model that reflects the features of each field in multiple fields through bidirectional learning. In addition, we have added the variation of data used in the existing statistical techniques to the input value. Table 1 shows the results of the comparative analysis of related works and the proposed model.

3 BiLSTM model based on multivariate time series data in multiple fields

3.1 Structure of the proposed model

Figure 1 shows the BiLSTM model based on multiple field multivariate time series data proposed in this paper. It is a many-to-one type that collects data and predicts one time-series data through learning by inputting preprocessed multivariate time series data according to each data feature. S_1 represents the main data to be predicted, and the other data are related to $D_{S_1,1}$ and are time series data used for learning. Here, the data are divided into two series S_1 and S_2. Here, S_1 is composed of related time series data for $D_{S_1,1}$ fields such as the existing multivariate time series data-based model. S_2 consists of related time series data for the other fields that affect $D_{S_1,1}$. S_2 learns the influence of S_2 on $D_{S_1,1}$ and has an auxiliary role for forecasting $D_{S_1,1}$. Therefore, S_2 must include $D_{S_1,1}$. Since the data structure and update cycle may be different for each S, it is necessary to catch the binding criterion and the preprocessing process is important. Next, learning is performed separately for each S and the obtained result is combined again to finally predict $D_{S_1,1}$. Here, each learning process is based on BiLSTM.

In addition, both the value data D of each time-series data and the variation (DD) for each D are summed for learning. Like existing statistical-based prediction models that predict the current state using the past state and error values, it reflects the learning model to update the weight and bias better.

In this paper, we apply this model to trading area forecasting. For this, we used data of purchase related to the trading

Table 1 Comparative table of related works and the proposed model

Model	Bidirectional learning	Multivariate time series data		Features extraction of field		Use variation
		One field	Multiple fields	Each field	Combining fields	
SAnD-Multi (Song et al. 2018)		O			O	
Dipole (Ma et al. 2017)	O	O			O	
The Multiple Pipelines Models with CNN and BiLSTM (Eapen et al. 2019)	O	O			O	
WKNNRW (Lang et al. 2018)			O		O	
The Probabilistic BLSTM + copula (Toubeau et al. 2018)	O		O		O	
The LSTM Model on Behavior and Opinion data space (Zhuge et al. 2017)			O	O	O	O
The proposed model	O		O	O	O	O

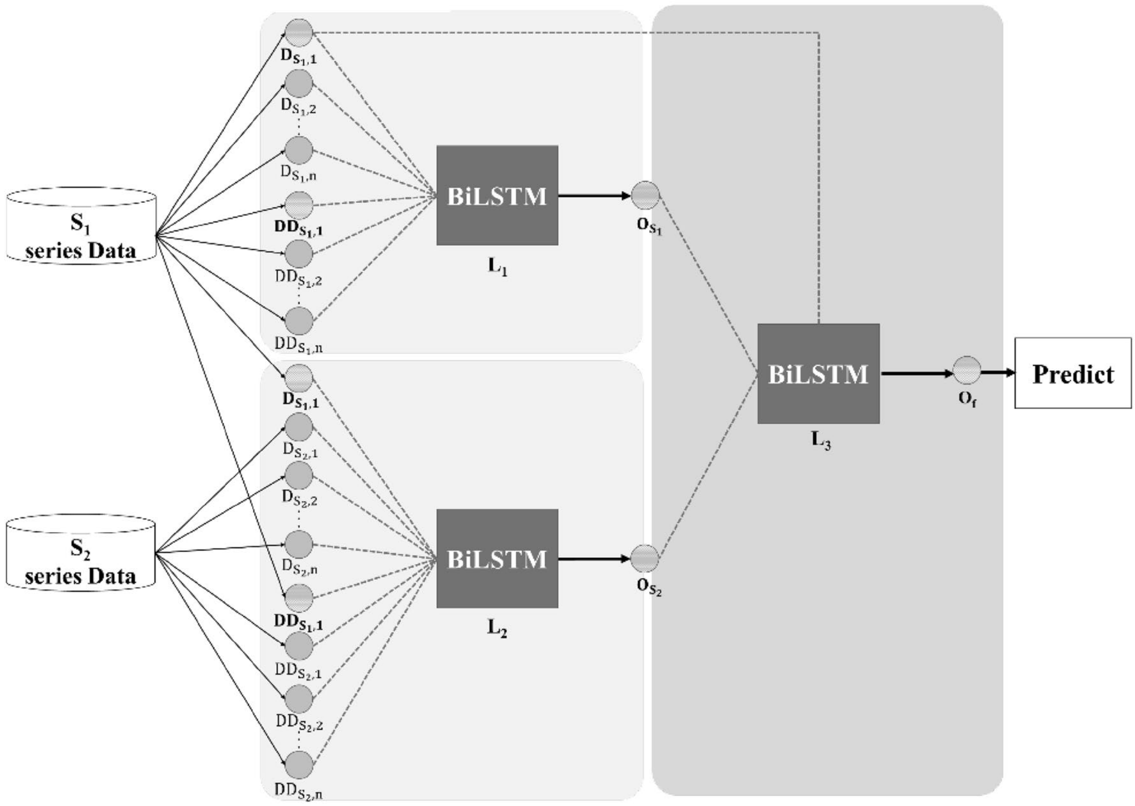


Fig. 1 Overview of the proposed learning model

area and SNS data that reflects consumption trends. The data of purchase is the total amount of purchase, frequency of purchase, etc. and the SNS data is about the SNS Buzz amount related to the purchase. Since the proposed model assumes that S_2 affects S_1, we need data that is suitable for S_2. Since the SNS data is focused on restaurants that has the largest amount of data related to purchasing, it is based on the collection of SNS data for restaurants. As shown in Table 2, the restaurant field includes 11 sectors.

3.2 Process of the proposed model

3.2.1 Structure of multivariate time series data in multiple fields

In this study, the data were composed of trading area-related data of purchase and SNS data for trading area forecasting. Prior to data collection, we must select and collect data that is highly influential for trading area forecasting.

First, in the case of data of purchase related to trading area, it is necessary to check what factors affect the trading area. Table 3 presents an analysis of factors remarkably influencing analysis of trading area mentioned in earlier research on existing trading area. This table shows the necessity of information on sales, transaction frequency, income level, moving population, and resident population that are common for analyzing trading areas.

In this paper, based on previous studies, we selected sales, transaction frequency, income level, and ratio of moving population as shown in Table 4. Sales and transaction frequency were preprocessed by constructing the daily statistical data. The income level was the most influential variable for the "restaurant," particularly the sales, according to previous research on trading area analysis (Jeong and Kim 2014). In this paper, we divided the income level into five

Table 2 Details of the sectors related to restaurants

Field	Sector
Restaurant	Korea food, Japanese food, Western food, Chinese food, Bakery, Confectionary, Coffee/beverage, Fast food, Flour-based food, Chicken/Duck, Pub/bar, Foodservice business and others

Table 3 Comparison analysis of the factors that remarkably influenced the analysis of trading areas in earlier research

Name of article	Influential factor
Analysis of trade area for retail industry store using consumer purchase record (Iwasaki et al. 2017)	Parking lots in store, location, number of products, population, number of households
Evaluating trade areas using social media data with a calibrated huff model (Wang et al. 2016)	Action radius of residents, distance between resident and store, visit frequency
Analysis of market district hinterland in Seoul by 'Big Data' analysis (Oh et al. 2017)	Monthly sales according to businesses, moving population, expenditure compared to income, settled population
An analysis of the location factors that affects the sales of campus commercial district (Lee et al. 2014)	Public transportation, moving population, size of trading area
Smart store in Smart City: the development of smart trade area analysis system based on consumer sentiment (Yoo et al. 2018)	Number of local business, > 3-year survival rate, residence area per person, residential environment satisfaction

Table 4 Factors applied in the proposed learning model

Factor	Data source	Subscribe
Sales	South Korea S Card Company (2017–2018)	Statistics on the monthly sales according to businesses
Transaction frequency		Statistics on the monthly transaction frequency for businesses
Income level		Statistics on the sales of the top income layer out of entire sales
Ratio of moving population		Percentage of the moving population compared to the population in the trading area

levels and calculated the sales amount for the highest income class out of the total sales amount. The lower the value of the income level, the higher the income. Finally, the ratio of moving population calculates the ratio of moving population to the total population.

Next is the SNS data for grasping social trends of purchase; this is collected by crawling SNS posts using the Python-based Selenium and BeautifulSoup libraries. In this case, the SNS posting collection criteria is a keyword that includes the name of the trading area and the classifier determines whether the place of the posting is included in the trading area. This classifier extracts national basic district number in the place of the post and classifies them by utilizing the street address API provided by Google API and the Ministry of Public Administration and Security based on high-frequency extracted words from every crawled post. Through this, once all postings have been classified, the buzz amount is obtained by knowing the number of posts related to the trading area. Figure 2 is a flowchart for this process.

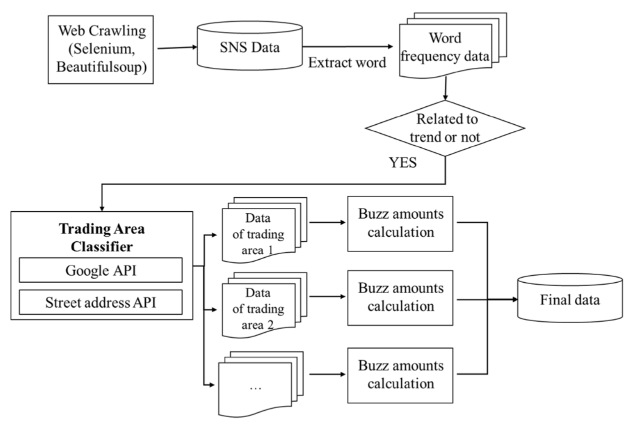


Fig. 2 Process for SNS data collection and preprocessing

3.2.2 Data structure of the input layer

When data collection is complete, preprocessing is performed to place it in the learning model's input layer. A data normalization process is indispensable to proceed with learning. In particular, when learning a large amount of data, if the data value is too large or too small, it greatly affects the result. In this paper, MinMaxScaler is used to normalize the data so that all values are between 0 and 1 through proportional adjustment. However, since the data representing the variation is an element that indicates increases and decreases, scaling is performed between -1 and 1 rather than between 0 and 1.

Next, we proceed with data binding for learning. Since S_1 is the data for the purchase field, there is no need for a separate combination, but this is necessary for S_2 because it is combined with SNS data.

Data of purchase is collected on a daily basis and SNS data is combined on a daily basis since a buzz amount is calculated based on the publication date. Finally, the initial data structure for S_1 and S_2 is shown in Table 5. S_1 is composed of the SA (Sales Amount), TF (Transaction Frequency), IL

Table 5 Example of the integrated data structure

Date	S_1								S_2					
	SA	SA dif	TF	TF dif	IL	IL dif	MP	MP dif	SA	SA dif	TF	TF dif	SB	SB dif
2017 0101	0.41	0.03	0.24	−0.08	0.03	0.08	0.34	−0.05	0.41	0.03	0.24	−0.08	0.37	−0.09
2017 0102	0.21	−0.47	0.20	−0.26	0.03	0.08	0.28	−0.23	0.21	−0.47	0.20	−0.26	0.25	−0.36
2017 0103	0.18	−0.07	0.18	−0.16	0.03	0.08	0.27	−0.06	0.18	−0.07	0.18	−0.16	0.36	0.14
2017 0104	0.23	0.16	0.19	−0.07	0.03	0.08	0.30	0.03	0.23	0.16	0.19	−0.07	0.32	−0.16
2017 0105	0.19	−0.08	0.18	−0.10	0.03	0.08	0.30	−0.03	0.19	−0.08	0.18	−0.10	0.41	0.08

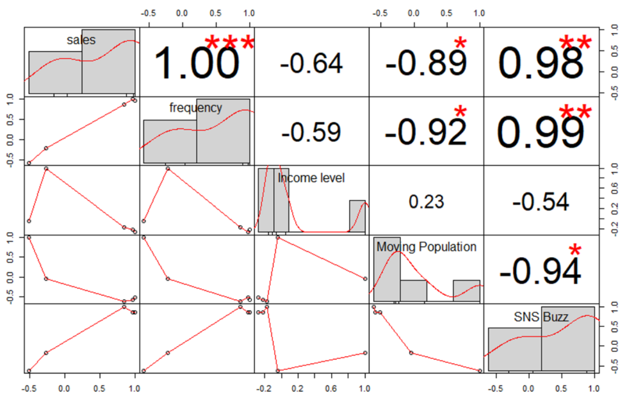


Fig. 3 Results of correlation analysis based on Pearson

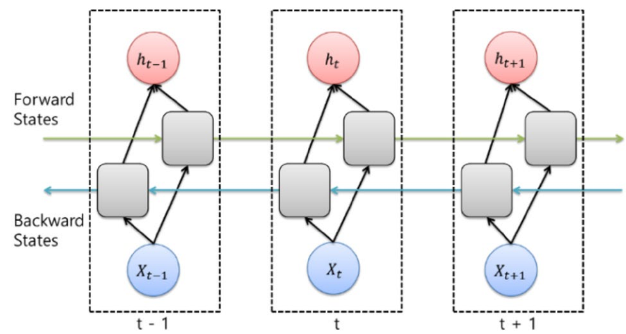


Fig. 4 Structure of the bidirectional LSTM

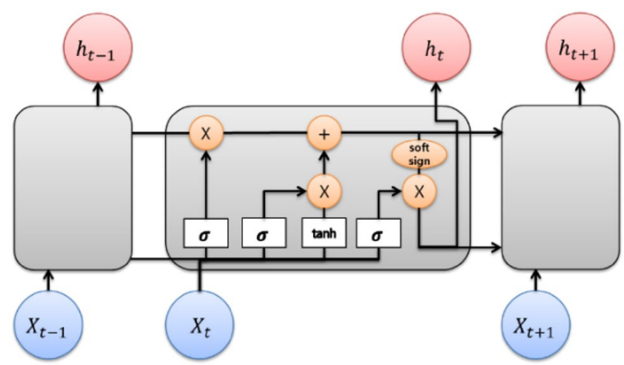


Fig. 5 Structure of the LSTM

(Income Level), MP (ratio of Moving Population), and the variation thereof (SA dif, TF dif, IL dif, MP dif). Figure 3 shows the Pearson-based correlation analysis using R; we can see that there is almost a one-to-one correlation between the sales and frequency and the SNS Buzz. Learning was carried out by including data on the frequency in S_2. Finally, S_2 is composed of the SA (Sales Amount), TF (Transaction Frequency), SB (SNS Buzz), and the variation thereof (SA dif, TF dif, SB dif).

3.2.3 Learning method based on BiLSTM

In Fig. 1, L_1 means learning about S_1, L_2 means learning about S_2, and L_3 means a combination of these. In other words, when applied to trading area forecasting, L_1 is learning about trading-related data of purchase, L_2 is learning about SNS data related to trading area, and L_3 is learning based on the predicted values of L_1 and L_2. In L_3, L_1 and L_2 are received as input values for the final prediction and learned to create output values for actual sales. Each L_1, L_2, and L_3 has a seven-day pattern cycle of purchase, so it predicts the next day with an input value for seven days.

The learning model of L_1, L_2, and L_3 utilizes biLSTM. Figure 4 shows that it is possible to store both the previous information and later information as the current time basis of the time series data and thus good performance can be expected.

In addition, Fig. 5 transmits past learning results through the forget, input, and output gates in the LSTM cell and the information of the forward and backward states are transmitted to the hidden layer. Thus, the activation output $\vec{h}_t$ of the forward hidden layer, the activation output $\overleftarrow{h}_t$ of the backward hidden layer, and y_t of the output layer are calculated in (1), (2), and (3), respectively. Here, σ is the activation function and y_t is generated by updating the output layer repeatedly in the backward and forward layers (Graves et al. 2013). On this basis, learning was performed such that the error between the actual value and the predicted value was reduced by using pre-processed data that had been collected earlier. Here, we used softsign rather than tanh as the activation function, which is commonly used in the existing LSTM model. Softsign is similar to tanh but it has some features that are more robust to the initialization process and shows improved performance when applied to LSTM.

$$\vec{h}_t = \sigma\left(W_{x\vec{h}}x_t + W_{\vec{h}\vec{h}}\vec{h}_{t-1} + b_{\vec{h}}\right) \tag{1}$$

$$\overleftarrow{h}_t = \sigma\left(W_{x\overleftarrow{h}}x_t + W_{\overleftarrow{h}\overleftarrow{h}}\overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}\right) \tag{2}$$

$$y_t = W_{\vec{h}y}\vec{h}_t + W_{\vec{h}y}\vec{h}_t + b_y \tag{3}$$

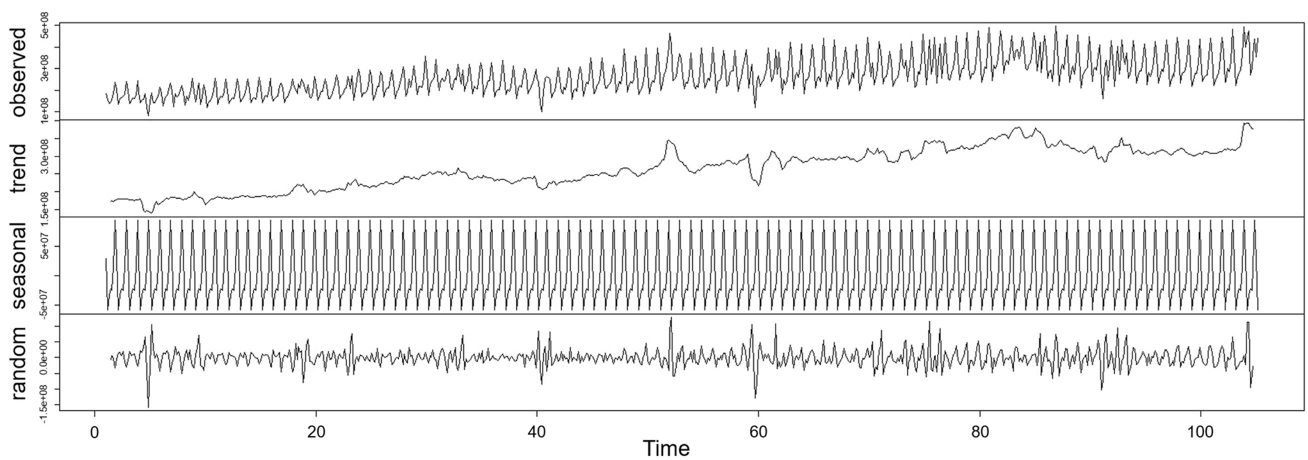
Table 6 Experiment environment

CPU	Intel® Core™ i7-8700 K
GPU	GeForce GTX 1080
RAM	16 GB
Python	3.5.5
TensorFlow	1.12.0

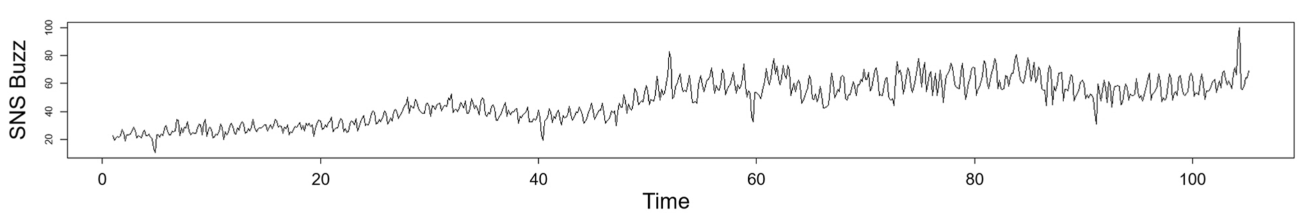
4 Experiments

Recently, various deep learning-based open source libraries such as Tensor-flow, Caffe, and Theano have been provided. Tensor-flow provides an interface to create an artificial network model and learn it in various hardware environments. Table 6 presents the experimental environment in this paper.

The data for the experiment was composed of eight sets: two training and validation sets for region A in Korea, and a test set was applied for region B. The data used for purchasing data was from two years: 2017 and 2018 for S-card companies and the Instagram for SNS data over the same period. Figure 6 (a) shows the time series of data collected for region A, which is a result of the time-series decomposition on sales of region A; (b) is a graph of the SNS Buzz data on region A according to time.



(a) Result of time series decomposition for sales of region A



(b) Graph of SNS Buzz for region A over time

Fig. 6 Time series of data collected for region A

For the performance evaluation, results were compared and analyzed by dividing them into four models according to whether learning was done by field or whether there was variation in the input value. M^1 is a model proposed in this paper that progresses by learning each field and includes variation. M^2 is a model in which learning is carried out by field but that does not include a variation as an input value. M^3 is a model that does not carry out field-by-field learning but includes an input value in the variation, and M^4 is a model that does not include input values in the variation and does not perform field-by-field learning.

For the learning in the four models, the softsign function was used for the activation function and AdamOptimizer was used for the optimization function. In addition, learning was performed by setting the epoch to 1000 and the learning rate to 0.001. The loss function uses the Mean Squared Error (MSE) and performance evaluation uses the Root Mean Square Error (RMSE) between actual and predicted values. The formula for this is (4) and the smaller the value, the better the predictive power.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{n}\left(y_t - \hat{y}_t\right)^2} \tag{4}$$

where y_t is the actual value and $\hat{y}_t$ is the predicted value.

Table 7 Results of RMSE

	M^1	M^2	M^3	M^4
Min RMSE	0.0786	0.0797	0.1008	0.0964
MAX RMSE	0.1066	0.1109	0.1595	0.1356
Average RMSE	0.0819	0.0828	0.1546	0.0971

Figure 7 is a graph of the predicted results for the test data and Table 7 shows the results of the comparative analysis of performance evaluation for the four models. We have confirmed that the performance of M^1 proposed in this paper has a lower RMSE and higher performance than the other models. Because the training the test data are not the same area-based data, the overall prediction rate is not high. However, as can be seen from the RMSE values, it is confirmed that the performance is improved when the learning is performed by each field. In addition, when a variation is given to the input value, a slight improvement in performance is confirmed only when learning is performed by fields. Therefore, we confirmed that the proposed M^1 model has high performance. In the case of M^3, the prediction rate is lower than that of M^4, although the data on the variation is added. It is presumed that the input value is too much and the learning is not done properly. Therefore, In the case of the variation

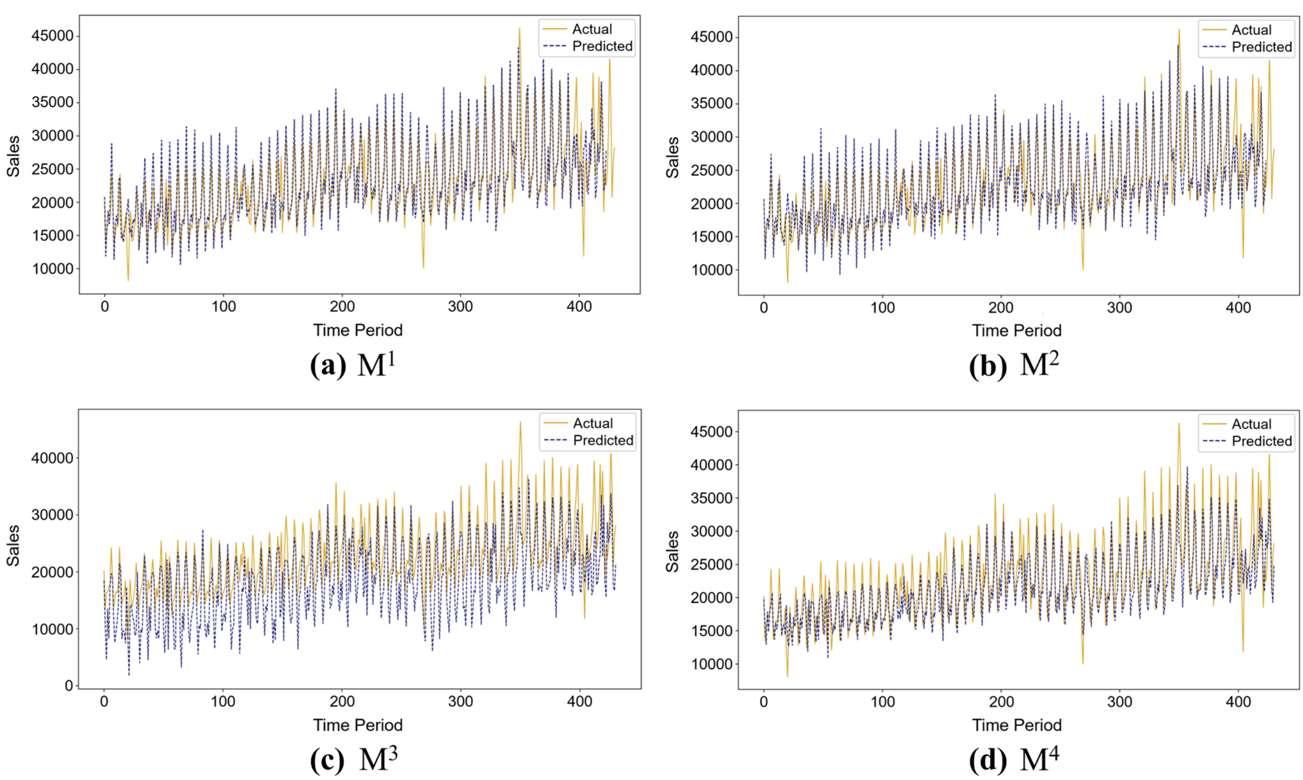


Fig. 7 Graph of the predicted results for the test data

to the input value, it is considered necessary to reconstruct the model.

5 Conclusion

Since most time series data have nonlinear characteristics, artificial neural network-based analysis and prediction methods are more suitable than existing statistical analysis and prediction methods. In addition, we recommend including dependent variables that have an effect rather than single time series data to understand the irregularities in time series data. In this paper, we propose a bidirectional LSTM model based on multivariate time series data. Here, the time series data of other fields that influence the time series data merge. In addition, to reflect the features of the time series data, both the values and the variation in each time series are derived and reflected as input values. Thus, the proposed model is a many-to-one type that takes bi-directional LSTM-based learning with multivariate time series data for two fields as input values and derives the final predicted value.

In this paper, we apply the proposed model to the trading area forecasting and combine it with SNS data to identify purchasing trends and information related to the trading area (sales, transaction frequency, income level, ratio of moving population). The SNS data describes the amount of buzz around the trading area. We combined the two fields of data by day, put in the data for seven days, and proceeded to learn how to forecast the amount of sales for the next day.

To evaluate this model's performance, the results were compared and analyzed by dividing them into four models according to whether learning had been done by each field or whether there was any variation in the input value. We can confirm that the proposed model has a smaller RMSE value and higher prediction rate than the other models.

The proposed model has great significance in multivariate time series analysis and prediction; in addition, it can be applied to other categories. However, this study has limitations because it does not consider the learning speed. Since the amount of data increases and the learning speed slows when combining multivariate time series data of multiple fields, it is necessary to supplement this. In a future study, we will modify the hyperparameter values of this model for each learning to improve its performance by variously experimenting to consider the time series feature more closely. In addition, we will study a forecasting model that combines LSTM with other artificial neural network models to reflect the effects of each field on two or more multiple field multivariate time series data.

Acknowledgements This work has supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIT) (No. NRF-2017R1A2B4008886).

References

Aydadenta H (2018) A clustering approach for feature selection in microarray data classification using random forest. J Inf Process Syst 14(5):1167–1175

Castrejon L, Kundu K, Urtasun R, Fidler S (2017) Annotating object instances with a polygon-rnn. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 5230–5238

Eapen J, Bein D, Verma A (2019) Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction. In: 2019 IEEE 9th annual computing and communication workshop and conference (CCWC), pp 0264–0270

Graves A, Mohamed AR, Hinton G (2013) Speech recognition with deep recurrent neural networks. In: 2013 IEEE international conference on acoustics, speech and signal processing, pp 6645–6649

Iwasaki S, Hashimoto K, Otake K, Namatame T (2017) Analysis of trade area for retail industry store using consumer purchase record. In: International conference on social computing and social media, Springer, Cham, pp 177–189

Jeong DS, Kim HB (2014) An analysis of the distribution and factors influencing sales by business types. GRI Rev 16:101–122

Lang K, Zhang M, Yuan Y, Yue X (2018) Short-term load forecasting based on multivariate time series prediction and weighted neural network with random weights and kernels. Cluster Comput. https://doi.org/10.1007/s10586-017-1685-7

Lee S, Moon N (2018) Location recognition system using random forest. J Ambient Intell Humaniz Comput 9(4):1191–1196

Lee YS, Park HS, Lew SH, Kang JM (2014) An analysis of the location factors that affects the sales of campus commercial district. Seoul Stud 15(1):17–34

Ma F, Chitta R, Zhou J, You Q, Sun T, Gao J (2017) Dipole: Diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks. In: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, ACM, pp 1903–1911

Meng R, Rice SG, Wang J, Sun X (2018) A fusion steganographic algorithm based on faster R-CNN. Comput Mater Cont 55(1):1–16

Miao Y, Gowayyed M, Metze F (2015) EESEN: end-to-end speech recognition using deep RNN models and WFST-based decoding. In: 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), pp 167–174

Oh HR, Bae BS, Moon HJ (2017) Analysis of market district hinterland in Seoul by 'Big Data' analysis. Korean Academic Society of Business Adiministration, pp 344–372

Siami-Namini S, Namin AS (2018) Forecasting economics and financial time series: Arima vs. lstm. arXiv preprint arXiv:1803.06386

Song Y, Kim I (2018) DeepAct: a deep neural network model for activity detection in untrimmed videos. J Inf Process Syst 14(1):150–161

Song H, Rajan D, Thiagarajan JJ, Spanias A (2018) Attend and diagnose: clinical time series analysis using attention models. In: Thirty-second AAAI conference on artificial intelligence

Toubeau JF, Bottieau J, Vallée F, De GZ (2018) Deep learning-based multivariate probabilistic forecasting for short-term scheduling in power markets. IEEE Trans Power Syst 34(2):1203–1215

Wang Y, Jiang W, Liu S, Ye X, Wang T (2016) Evaluating trade areas using social media data with a calibrated huff model. ISPRS Int J Geo Inf 5(7):112

Yin W, Kann K, Yu M, Schutze H (2017) Comparative study of cnn and rnn for natural language processing. arXiv preprint arXiv:1702.01923

Yoo IJ, Seo BG, Park DH (2018) Smart store in Smart City: the development of smart trade area analysis system based on consumer sentiments. J Intell Inf Syst 24(1):25–52

You SD, Liu CH, Chen WK (2018) Comparative study of singing voice detection based on deep neural networks and ensemble learning. Hum Cent Comput Inf Sci 8(1):34

Zeng D, Dai Y, Li F, Sherratt RS, Wang J (2018) Adversarial learning for distant supervised relation extraction. Comput Mater Cont 55(1):121–136

Zhuge Q, Xu L, Zhang G (2017) LSTM neural network with emotional analysis for prediction of stock price. Eng Lett 25(2):64–72

Zouina M, Outtaj B (2017) A novel lightweight URL phishing detection system using SVM and similarity index. Hum Cent Comput Inf Sci 7(1):17