# LSTM-based Sales Forecasting Model

**Jun-Ki Hong***
Department of Computer Engineering, Pai Chai University
Daejeon, South Korea
[e-mail: jkhong@pcu.ac.kr]
*Corresponding author: Jun-Ki Hong

## *Abstract*

In this study, prediction of product sales as they relate to changes in temperature is proposed. This model uses long short-term memory (LSTM), which has shown excellent performance for time series predictions. For verification of the proposed sales prediction model, the sales of short pants, flip-flop sandals, and winter outerwear are predicted based on changes in temperature and time series sales data for clothing products collected from 2015 to 2019 (a total of 1,865 days). The sales predictions using the proposed model show increases in the sale of shorts and flip-flops as the temperature rises (a pattern similar to actual sales), while the sale of winter outerwear increases as the temperature decreases.

# 1. Introduction

According to the Internet Usage Survey of Korea 2019 conducted by the Ministry of Science and ICT of Korea in April 2019, the most popular products purchased online by consumers were apparel, shoes, sporting goods, and accessories (87.8%). The survey found that consumers purchase clothing-related products online, rather than visiting stores, because they are not restricted by time and/or location [1].

In particular, sales of apparel varied rapidly in response to changes in the weather [2] and changes in temperature, unlike home appliances and daily necessities. Therefore, management of apparel inventory is so critical that it can determine the survival of a fashion apparel company. Thus, prediction of sales tied to changes in the weather and temperature is important for efficient stock management and profit maximization.

Research was conducted into efficient stock management of apparel [3-6] and accurate forecasts of sales volume [7-11]. Furthermore, studies were carried out into sales predictions that apply various machine learning techniques [12-15]. However, since fashion companies are reluctant to disclose actual sales numbers, there have been no studies on predictions made with actual sales data.

Therefore, in this study, real sales data from January 1, 2015, to December 31, 2019, for apparel products sold by Company A (2.5 million accounts in October 2020) were used to analyze sales related to temperature changes and to make sales predictions by using long short-term memory (LSTM). To verify the proposed sales prediction model, data on shorts, flip-flop sandals, and winter outerwear, which are immediately impacted by changes in temperature, were used to analyze and predict sales.

# 2. Methodology

The algorithm used to predict variations in sales of products based on temperature changes is explained. **Fig. 1** shows the flowchart of the sales prediction algorithm.
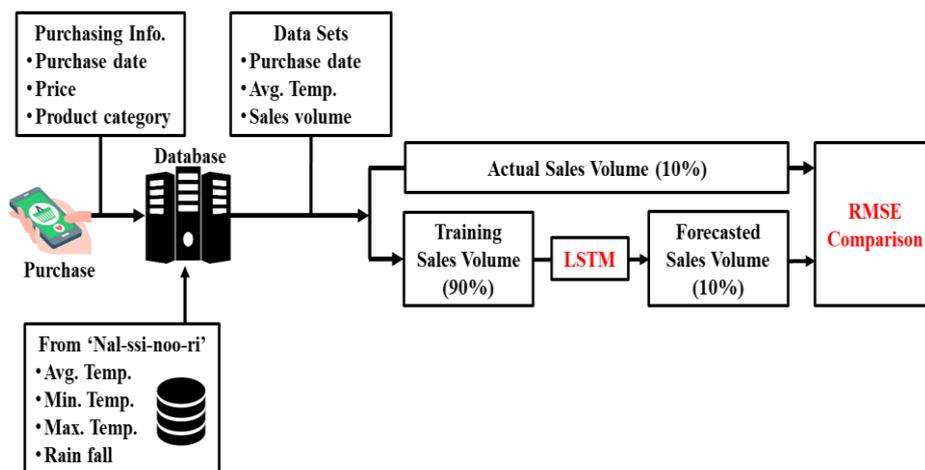


**Fig. 1.** Flow chart of the proposed sales volume prediction algorithm

Subsections 2.1 and 2.2, respectively, detail the big data collection method and the LSTM concept used for the proposed algorithm.

## 2.1 Apparel Sales Data Collection

This subsection discusses the data collected from Company A. As seen in **Fig. 1**, the date the consumer purchased the product, the product category, the number of items purchased, and the price are stored in the database. Also, stored are the average daily temperature, the lowest and highest temperatures, and daily precipitation recorded by the Korea Meteorological Administration.

**Table 1.** Example data sets for short-sleeved shirts

| Purchase Date | Avg. Temp. | Min. Temp. | Max Temp. | Rainfall | Avg. Price | Sales |
|---|---|---|---|---|---|---|
| 20150101 | 4 | 0.3 | 7.9 | 0 | 14864 | 177 |
| 20150102 | 1.5 | -1.3 | 6.2 | 0 | 14587 | 207 |
| 20150103 | 2.3 | -1.1 | 7.3 | 0 | 15319 | 155 |
| 20150104 | -0.1 | -3.7 | 5.6 | 0 | 15275 | 122 |

**Table 1** shows examples of the data stored in the DB for sales predictions of short-sleeved shirts, where Sales is the total volume for each date.

In the next step, changes in sales based on temperature variations are predicted utilizing the purchase date, category, average temperature, and sales data. In this study, data for shorts, flip-flops, and winter jackets are used for the analysis and predictions.

## 2.2 LSTM

LSTM was introduced by Hochreiter and Schmidhuber [16]. Effective applications based on LSTM have been reported in many fields, such as natural language translation [17], encrypted speech retrieval [18], speech recognition [19], image captioning [20-21] and physical activity prediction [22].

In the proposed sales forecasting model, LSTM is used to predict sales volumes related to temperature changes. Representative deep learning models can be largely categorized into the convolution neural network (CNN) and the recurrent neural network (RNN). The CNN performs well when analyzing characteristics that do not depend on the order, such as image recognition, whereas the RNN model performs well when analyzing time series characteristics. However, the RNN has the problem of vanishing gradients when learning long-term patterns; the gradient gradually decreases during backpropagation, significantly degrading the learning ability, and restricting consideration of long-term dependency in the data.

However, LSTM is a derived model of an RNN that resolves the vanishing gradient problem; the learning rate is not updated when carrying out long-term pattern learning by determining the storage/deletion of the previous weight through the cell state layer.
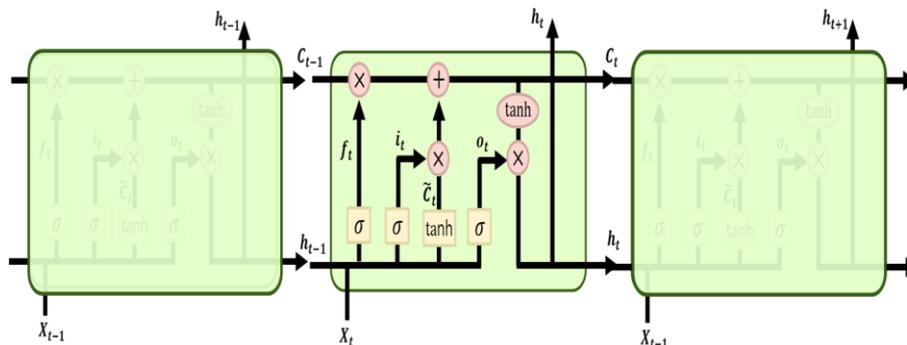


**Fig. 2.** Structure of LSTM

As shown in **Fig. 2**, LSTM is a model based on the RNN, and is composed of cells with multiple gates attached; $x_t$ and $h_t$, respectively, represent the input and hidden state at time $t$. Furthermore, LSTM nodes are connected continuously. In each time step, the hidden state and the cell state from the previous time step are received, as is the input value from the current time step. The hidden state and the cell state are updated through gates, and are transmitted to the next time step. Moreover, $i$, $f$, and $o$ denote input gate, forget gate, and output gate, respectively.

The roles of these cells include delete, store, update, and export functions. Each cell determines what to store, when to export the information, when to write it, and when to delete it based on the connected gate values. With this cell structure, LSTM can overcome the representative issues of the RNN, including long-term dependency, vanishing gradients, and divergence.

The cell state is updated through multiple gates pertaining to the repeating module. The roles of each gate are as follows. First, forget gate $f_t$ determines by using a sigmoid function what information to throw away. Input gate $i_t$ determines the information to store in the cell state by using sigmoid and tanh functions. Output gate $o_t$ determines what information to export based on the updated cell state. Finally, $C_t$ denotes the cell state updated from the previous cell state, $C_{t-1}$, and the trained data, $\tilde{C}_t$.

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big) \qquad (1)$$

$$i_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_i\big) \qquad (2)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \qquad (3)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \qquad (4)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \qquad (5)$$

$$h_t = \sigma * \tanh(C_t) \qquad (6)$$

In the equations above, $x_t$ represents the sales volume at time series $t$. Hence, time series sales values are the input data used in the neural network. The length of the $N$ input sales volumes is denoted $[x_1, x_2, \ldots, x_N]$. Then, the series $[h_1, h_2, \ldots, h_N]$ is calculated from (1) to (6). As shown in **Fig. 2**, this process inputs $x_t$ and $h_{t-1}$ to obtain $h_t$, and this is repeated sequentially from $t = 1$ to $t = N$. In this study, the last $h$ value, $h_N$, is the input for the final calculation of (7), and the obtained result is used to predict sales in the next step. Thus, output, $y$ represents the predicted sales volume obtained in (7) with the estimated value of $x_{N+1}$:

$$y = W_{hy}h_N + b_y \qquad (7)$$

where $W$ and $b$ represent the weight and bias, respectively. However, sales volume and temperature data are used as input to predict temperature. The results of temperature prediction using LSTM can be found in Section 3.1.

Furthermore, sales predictions from the proposed model were evaluated in relation to actual sales through the root mean square error (RMSE), as shown in (8):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(S_{A,i}-S_{P,i})^2}{n}} \qquad (8)$$

where, $S_{A,i}$ and $S_{P,i}$ represent the actual and the predicted sales volumes, respectively, while $n$ denotes the total number of days predicted.

## 2.3 Simulation Environment

In this study, an LSTM model with three hidden layers is constructed where each layer has 400 hidden units. An Nvidia RTX 2080 Ti 11 GB graphics processing unit (GPU) was used for the sales prediction simulation, and **Table 2** shows the LSTM model configuration for the simulation.

**Table 2.** Simulation parameters for the Adam optimizer

| Parameter | Value |
|---|---|
| No. of hidden units | 400 |
| Max. epochs | 2000 |
| Initial learning rate | 0.005 |

For optimization of the LSTM model in this study, the Adaptive Moment Estimation (Adam) algorithm was used. The Adam optimizer is advantageous for efficiently finding, in a short period of time, the optimal sales prediction based on temperature by flexibly adjusting the learning rate [23].

## 3. Simulation Results

### 3.1 Average Temperature Prediction

This subsection explains the temperature predictions for 180 days—the result from training using 1,645 days of temperature data to verify the model prior to making sales predictions. This was approximately 90% of the total 1,825 days of average temperatures collected by the Korea Meteorological Administration from January 1, 2015, to May 30, 2019.
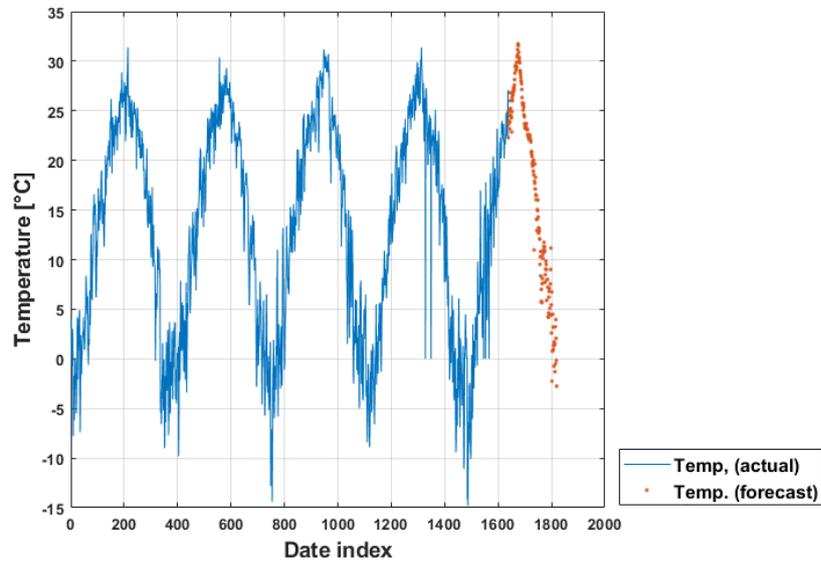
(a)
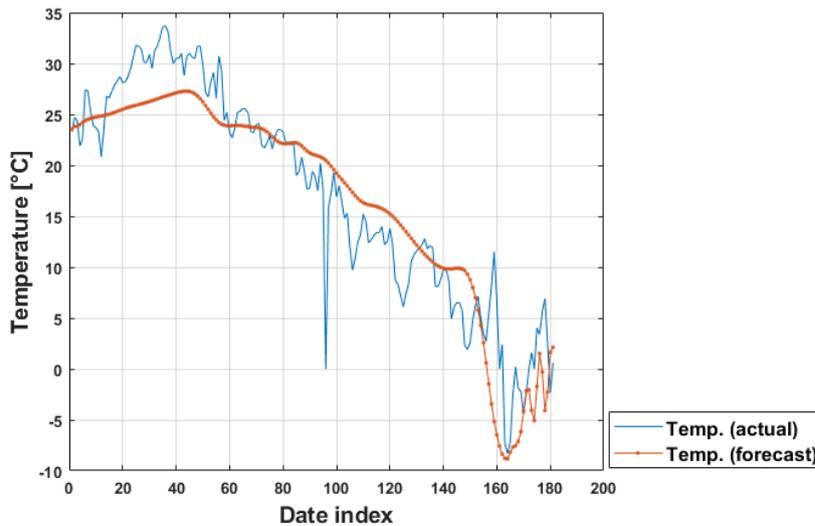


(b)

**Fig. 3.**  (a) Temperature predictions using the LSTM model, and (b) a comparison between actual and predicted temperatures

In **Fig. 3(a)**, the blue plot is 1,645 days of actual temperature data out of the total 1,825 days, and the orange plot shows the temperature predictions for the subsequent 180 days using LSTM. As seen in the orange plot, the temperature cycle that repeats every year is accurately predicted using LSTM. Furthermore, **Fig. 3(b)** shows a comparison between the actual temperature during those 180 days in the orange plot in **Fig. 3(a)** and the temperature predictions using LSTM. The RMSE value between the actual and predicted values was 3.24. As shown in **Figs. 3(a)** and **3(b)**, the proposed model accurately predicted future values using the collected data. As such, future temperatures were accurately predicted using LSTM.

The predicted average temperatures are used in the next subsection to investigate the changes in sales of shorts, flip-flops, and winter outerwear.
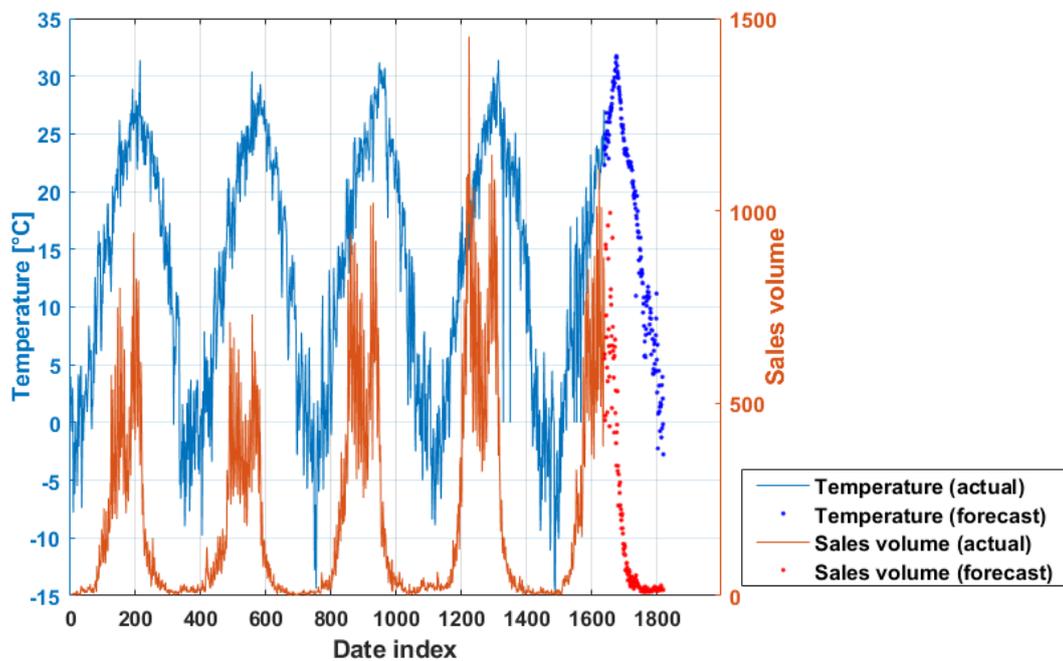
## 3.2 Predicting Sales of Shorts

In this subsection, the sales of shorts are predicted using the proposed model based on actual data. The sales volumes for shorts were analyzed and predicted because they are sensitive to temperature variations, so more accurate analyses and predictions of sales are possible, compared to other products.



**Fig. 4.** Sample images of the actual items for sale

**Fig. 4** shows examples of actual products sold by Company A. Sales were analyzed using data for all shorts sold over the previous five years, and predictions for the sales volumes were based on temperature changes.
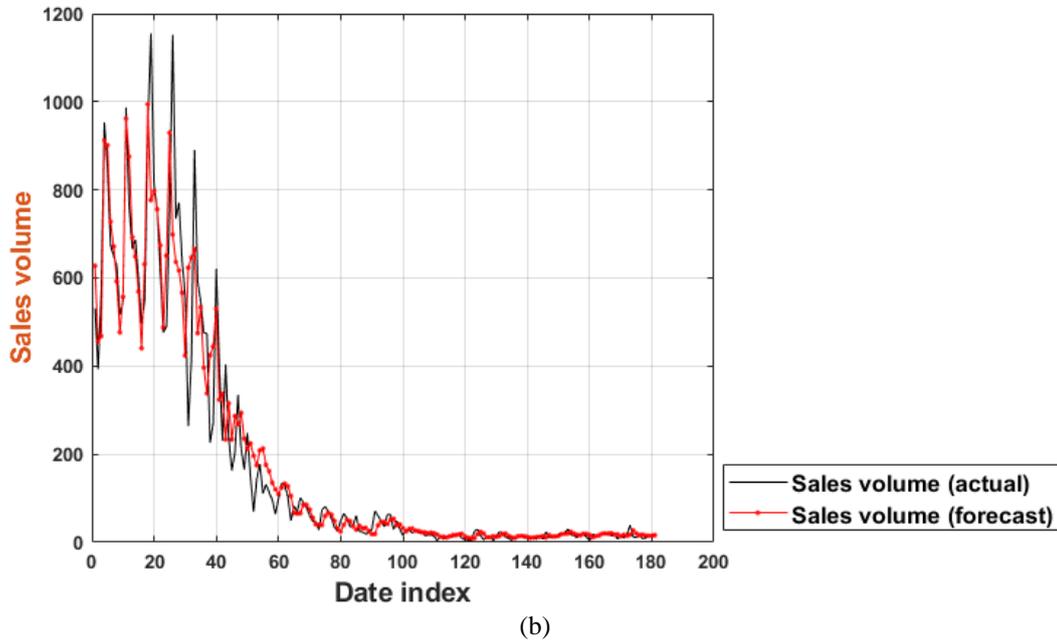


(a)

(b)

**Fig. 5.** (a) Analysis of sales of shorts, and predictions according to temperature variations
(b) comparison between actual and predicted sales

**Fig. 5** shows predictions for sales of shorts over 180 days after training the proposed model using 1,605 days of data out of the total 1,825 days. The blue plot in **Fig. 5(a)** shows the same temperature graph as **Fig. 3(a)**, and the orange and red plots, respectively, represent actual sales and predicted sales using LSTM.

As seen in the orange plot of **Fig. 5(a)**, the actual sales increased in direct proportion to increases in temperature. Also, the dotted red plot shows that the predictions increased as the temperature increased, and the sales prediction pattern is very similar to the pattern in the actual sales data.

The black plot of **Fig. 5(b)** shows the actual sales of shorts for those 180 days, and the red plot shows the sales predictions using the same LSTM used to obtain the red plot in **Fig. 5(a)**. The sales predicted using LSTM are similar to actual sales, and **Table 3** in Subsection 3.4 shows a set of 10 RMSE values calculated. As such, the proposed sales prediction model is able to accurately predict the sales of shorts based on temperature variations.

## 3.3 Predicting Sales of Flip-flops

Flip-flop sales were analyzed and predicted because, like sales of shorts, they change immediately based on temperature.

**Fig. 6.** Images of actual products corresponding to the collected flip-flop data

**Fig. 6** shows examples of actual flip-flops sold by Company A. These sales were analyzed and predicted according to the temperature variations using total sales data for all flip-flop products sold over five years.
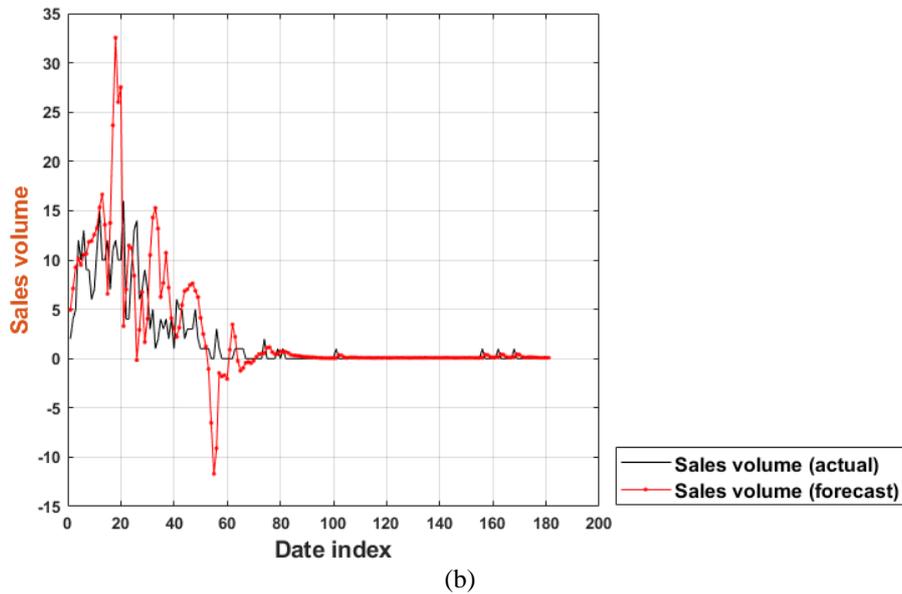


(a)

(b)
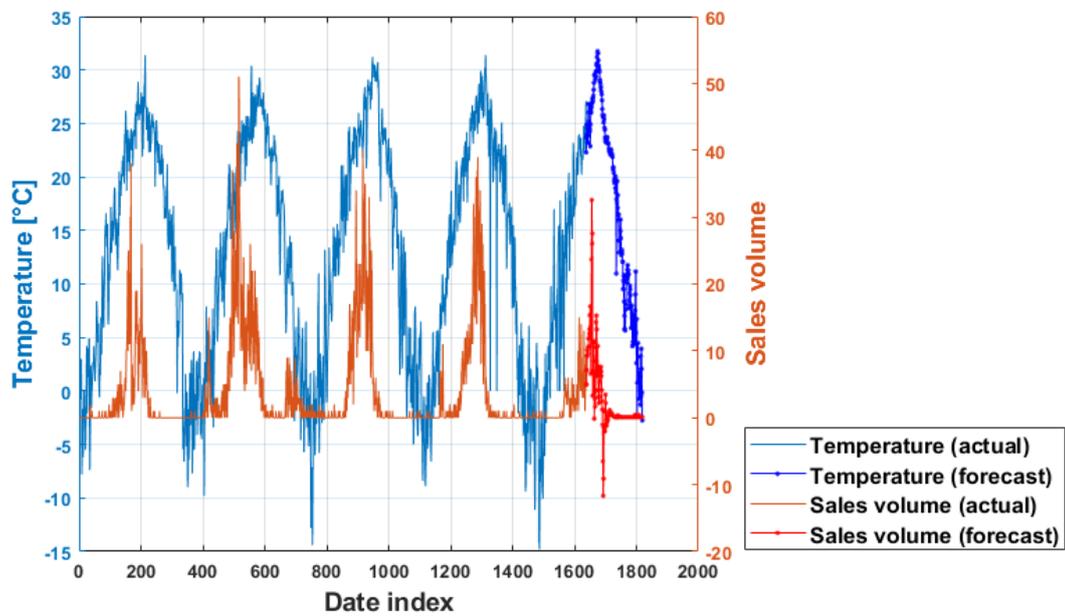
**Fig. 7.**  (a) Flip-flop sales analysis, and predictions according to temperature variations
(b) comparison between actual and predicted sales

**Fig. 7(a)** shows the flip-flop sales analysis and the predictions based on temperature changes. As seen in **Fig. 7(a)**, actual flip-flop sales increased in direct proportion to the rise in temperature. However, while sales of shorts shown in **Fig. 6(a)** increased gradually based on the temperature, sales of flip-flops increased rapidly as the temperature increased. That is because flip-flops are mostly used during high temperatures, reflecting the characteristic surge of flip-flop sales in the summer.

The analysis of shorts and flip-flop sales shown in **Fig. 7(b)** reveals that the actual and predicted sales patterns are very similar. The black plot in **Fig. 7(b)** shows flip-flop sales for 180 days, and the red plot shows flip-flop sales predictions using the same LSTM used for the red plot of **Fig. 7(a)**.

The flip-flop sales predictions using LSTM are very similar to actual sales but with a larger RMSE value than for predictions for sales of shorts. This larger RMSE value between actual and predicted flip-flop sales is thought to be due to the smaller absolute sales volume of flip-flops, compared to shorts, resulting in greater difficulty in carrying out finer training and prediction.

## 3.4 Winter Outerwear Sales Prediction

In this subsection, the proposed sales prediction model is used to predict and analyze the sales of winter outerwear based on temperature variations. **Fig. 8** shows images of actual winter outerwear included in this study.



**Fig. 8.**  Images of products from the winter outerwear collection

The total sales of various winter outerwear products, including coats, jackets, and lamb skin jackets, were used to analyze and predict sales according to temperature change.
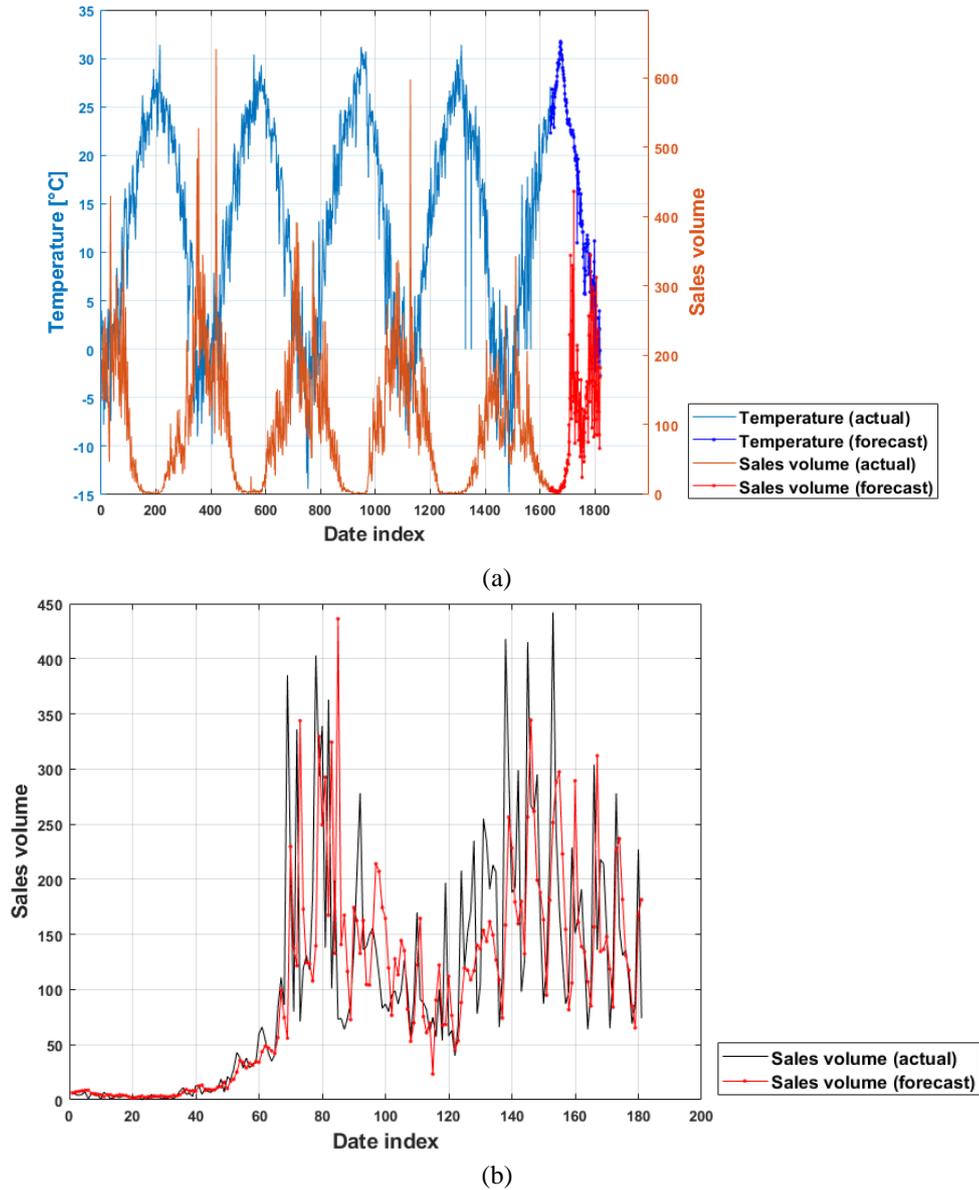


(a)



(b)

**Fig. 9.** (a) Winter outerwear sales analysis and predictions based on temperature variations
(b) comparison between actual and predicted sales

The orange plot in **Fig. 9(a)** shows actual winter outerwear sales, which decreased as the temperature increased, and increased as the temperature decreased, unlike the sales of shorts and flip-flops. As seen from the red plot in **Fig. 9(a)**, winter outerwear sales were predicted to increase as the temperature decreased, just like the conventional sales pattern based on temperature variations.

The black plot in **Fig. 9(b)** shows actual winter outerwear sales for 180 days, and the red plot shows the predicted winter outerwear sales, which exhibit a sales pattern very similar to actual sales.

Additionally, the winter outerwear sales pattern for a one-year period showed an increase in sales followed by a decline in the middle, and then another increase in sales. Unlike shorts and flip-flops, winter outerwear takes longer to manufacture and is prepared in advance (before temperatures fall); stock is reordered when supplies are exhausted. This phenomenon is also predicted to match the actual sales pattern, as shown in **Fig. 9(b)**.

**Table 3.** RMSE values for weather, shorts, flip-flops, and winter outerwear

| No. | Weather | Shorts | Flip-flop | Winter Outerwear |
|-----|---------|--------|-----------|------------------|
| 1 | 3.5232 | 64.3354 | 11.3149 | 3.1541 |
| 2 | 3.2534 | 65.6082 | 9.2075 | 2.6718 |
| 3 | 3.8126 | 75.1091 | 3.2924 | 2.3142 |
| 4 | 3.8553 | 60.7491 | 13.3793 | 2.4859 |
| 5 | 2.4467 | 71.0822 | 10.1840 | 2.4340 |
| 6 | 4.1687 | 84.3910 | 5.1584 | 3.0600 |
| 7 | 3.7116 | 68.7939 | 3.6899 | 1.6672 |
| 8 | 2.6527 | 62.3343 | 9.1514 | 2.3988 |
| 9 | 2.7469 | 62.2786 | 12.9468 | 3.5097 |
| 10 | 2.4511 | 69.3698 | 4.3579 | 3.7404 |
| Avg. | 3.2476 | 68.4052 | 8.2682 | 2.7436 |

As seen in **Table 3**, RMSE values for shorts are higher than the other products since the sales volume is up to about 1,200, which is much higher than the other products, as shown in the left *y*-axis of **Fig. 5(a)**. However, the trend in sales volume for shorts due to temperature change is precisely predicted.

# 4. Conclusion

In this study, a model employing LSTM for predictions of product sales according to temperature change is proposed. To verify the proposed model, product sales data from 1,825 days (January 1, 2015, to December 31, 2019) for Company A in Korea was used to analyze and predict the sales of short pants, flip-flop sandals, and winter outerwear. The sales predictions for all three product categories according to temperature changes exhibited patterns very similar to actual sales. Moreover, the LSTM simulation showed that using more sales data in training will result in finer sales predictions. As further study, a sales prediction model will be implemented taking into consideration not only temperature, but various external factors, as well as the selling price.

# Acknowledgement

# References

[1] M. S. No, H. N. Heo, Y. J. Choi, and H. S. Lee, "Survey on the Internet Usage 2019," *Ministry of Science and ICT (MSIT) and Korea Internet and Security Agency (KISA)*, 2020. Article (CrossRef Link)

[2] J. K. Hong, "Analysis of Sales Volume by Products According to Temperature Change Using Big Data Analysis," *The Journal of Bigdata*, vol. 4, no. 2, pp. 85-91, Dec. 2019. Article (CrossRef Link)

[3] S. Thomassey, "Sales Forecasts in Clothing industry: The Key Success Factor of The Supply Chain Management," *International Journal of Production Economics*, vol. 128, no. 2, pp. 470-483, Dec. 2010. Article (CrossRef Link)

[4] W. K. Wong and Z. X. Guo, "A Hybrid Intelligent Model for Medium-term Sales Forecasting in Fashion Retail Supply Chains using Extreme Learning Machine and Harmony Search Algorithm," *International Journal of Production Economics*, vol. 128, no. 2, pp. 614-624, 2010. Article (CrossRef Link)

[5] T. M. Choi and B. Shen, "A System of Systems Framework for Sustainable Fashion Supply Chain Management in the Big Data Era," in *Proc. of 2016 IEEE 14th International Conference on Industrial Informatics(INDIN)*, pp. 902-908, 2016. Article (CrossRef Link)

[6] T. M. Choi, "Incorporating Social Media Observations and Bounded Rationality into Fashion Quick Response Supply Chains in the Big Data Era," *Transportation Research Part E: Logistics and Transportation Review*, vol. 114, pp. 386-397, June 2018. Article (CrossRef Link)

[7] Y. Zhang, C. Zhang, and Y. Liu, "An AHP-Based Scheme for Sales Forecasting in the Fashion Industry," *Analytical Modeling Research in Fashion Business*, pp. 251-267 May 2016. Article (CrossRef Link)

[8] S. Ren, T. Choi, and N. Liu, "Fashion Sales Forecasting with a Panel Data-Based Particle-Filter Model," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 3, pp. 411-421, Mar. 2015. Article (CrossRef Link)

[9] Y. Ni and F. Fan, "A Two-Stage Dynamic Sales Forecasting Model for The Fashion Retail," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1529-1536, Mar. 2011. Article (CrossRef Link)

[10] A. Aksoy, N. Ozturk, and E. Sucky, "A Decision Support System for Demand Forecasting in the Clothing Industry," *International Journal of Clothing Science and Technology*, vol. 24, no. 4, pp. 221-236. July 2012. Article (CrossRef Link)

[11] N. Liu, S. Ren, T. M. Choi, C. L. Hui, and S. F. Ng, "Sales Forecasting for Fashion Retailing Service Industry: A Review," *Mathematical Problems in Engineering*, vol. 2013, pp. 1-9, Oct. 2013. Article (CrossRef Link)

[12] K. F. Au, T. M. Choi, and Y. Yu, "Fashion Retail Forecasting by Evolutionary Neural Networks," *International Journal of Production Economics*, vol. 114, no. 2, pp. 615-630, Aug. 2008. Article (CrossRef Link)

[13] S. Thomassey and M. Happiette, "A Neural Clustering and Classification System for Sales Forecasting of New Apparel Items," *Applied Soft Computing Journal*, vol. 7, no. 4, pp. 1177-1187, Aug. 2007. Article (CrossRef Link)

[14] W. K. Wong and Z. X. Guo, "A Hybrid Intelligent Model for Medium-term Sales Forecasting in Fashion Retail Supply Chains using Extreme Learning Machine and Harmony Search Algorithm," *International Journal of Production Economics*, vol. 128, no. 2, pp. 614-624, Dec. 2010. Article (CrossRef Link)

[15] Z. L. Sun, T. M. Choi, K. F. Au, and Y. Yu, "Sales Forecasting using Extreme Learning Machine with Applications in Fashion Retailing," *Decision Support Systems*, vol. 46, no. 1, pp. 411-419, Dec. 2008. Article (CrossRef Link)

[16] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997. Article (CrossRef Link)

[17] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to Sequence Learning with Neural Networks," *in Proc. of Advanced Neural Information Process Systems*, pp. 3104-3112, 2014. Article (CrossRef Link)

[18] Q. Zhang, Y. Li, and Y. Hu, "An Encrypted Speech Retrieval Scheme Based on Long Short-Term Memory Neural Network and Deep Hashing," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 6, pp. 2612-2633, June 2020. Article (CrossRef Link)

[19] A. Graves and N. Jaitly, "Towards End-to-end Speech Recognition with Recurrent Neural Networks," *in Proc. of the 31st International Conference on Machine Learning(ICML)*, vol. 32, no. 2, pp. 1764-1772, 2014. Article (CrossRef Link)

[20] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and Tell: A Neural Image Caption Generator," in *Proc. of IEEE Conference Computer Vision Pattern Recognition*, pp. 3156-3164, 2015. Article (CrossRef Link)

[21] A. Karpathy and L. Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions," in *Proc. of IEEE Conference Computer Vision Pattern Recognition*, pp. 3128-3137, 2015. Article (CrossRef Link)

[22] J. Kim and K. Chung, "Prediction Model of User Physical Activity using Data Characteristics-based Long Short-Term Memory Recurrent Neural Networks," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 4, pp. 2060-2077, Apr. 2019. Article (CrossRef Link)

[23] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *in Proc. of 3rd International Conference for Learning Representations*, pp.1-15, May 2015. Article (CrossRef Link)

**Jun-Ki Hong** was born in Seoul, South Korea. He received the B.S. in computer systems engineering from Carleton University, Ottawa, Canada, in 2010 and the Ph.D. degree in electronics engineering from Yonsei University, Seoul, South Korea, in 2017. He joined Telecommunications Technology Association, Bundang-gu, South Korea, where he worked on developing the standardizations of V2X from 2016 to 2017. From 2017 to 2020, he was an Assistant Professor with the Department of Electrical and Electronic Engineering, Youngsan University. Since March 2003, he has been with the Department of Computer Engineering, Pai Chai University, Daejeon, South Korea, where he is currently an Assistant Professor. His research interests include big data, AI, IoT and 5G wireless communications.