Mathematical Biosciences and Engineering

MBE, 20(5): 9041–9061.
DOI: 10.3934/mbe.2023397
Received: 30 December 2022
Revised: 15 February 2023
Accepted: 27 February 2023
Published: 13 March 2023

http://www.aimspress.com/journal/MBE

Research article

Forecasting of garlic price based on DA-RNN using attention weight of temporal fusion transformers

Eunjae Choi[1], Yoosang Park[2], Jongsun Choi[2,*], Jaeyoung Choi[2] and Libor Mesicek[3]

[1] Department of IT Logistics & Distribution, Soongsil University, Seoul 06978, Republic of Korea
[2] School of Computer Science and Engineering, Soongsil University, Seoul 06978, Republic of Korea
[3] Department of Economics and Management, Jan Evangelista Purkyne University, Pasteurova 3544/1, Usti nad Labem 40096, Czech Republic

* **Correspondence:** Email: jongsun.choi@ssu.ac.kr; Tel: +82028287178.

Abstract: Garlic is a major condiment vegetable grown in South Korea. The price of garlic has a great impact on Korean society and the economy, which requires price stabilization through preemptive supply and demand management. Therefore, the government attempts to keep the price adjusted according to the predicted production cost. However, classic statistical models or well-known deep learning models have lower forecast accuracy when the number of input factors increases. The aforementioned issue could make analysis approaches and their implementation difficult, and the government would confront failure in proper supply and demand management. To solve this problem, we propose a new hybrid deep-learning approach that employs well-known attention models. Recent attention models have achieved outstanding performance in time-series dataset forecasting. However, when input datasets contain dozens or hundreds of variables, the forecasting performance cannot be guaranteed because the prediction accuracy decreases. In this study, a novel approach utilizing attention weights for forecasting prices is introduced. Experience shows that forecasting accuracy can be improved using the proposed model, which deals with different variables related to garlic prices, such as atmospheric conditions, logistics processes, and environmental circumstances. The proposed approach and its model contribute to forecasting outputs for different research domains by using a variety of attention weight models.

Keywords: forecasting garlic price; dual-stage attention-based RNN; temporal fusion transformers; attention mechanism; deep learning

1. Introduction

Garlic is a popular condiment in Korean cuisine and ranks second in total vegetable production according to the Korean Statistical Information Service (KOSIS), accounting for 8% (KRW 903 billion) of agricultural and forestry production. Garlic has a higher product value and more cultivated land compared to peppers, which rank first in KOSIS's "Management Opening of Pepper and Garlic Production Farms." Furthermore, garlic is one of the five major vegetables designated by the Korean government as being sensitive to supply and demand, alongside cabbage, radish, onion, and hot pepper. Hence, there is a requirement for a garlic price forecasting model that can accurately capture both the volatile non-linear component and the linear component supported by statistical models. The forecast of garlic price is crucial in regulating supply and demand, and it can significantly influence production volume and cultivation area [1]. Specifically, the price volatility of garlic is influenced by climate change, logistics processes, and demand from the food industry, as garlic is primarily cultivated in open fields. Crop production is subject to a range of climatic factors, including temperature, precipitation, and sunlight levels [2–4]. Recent atypical climate patterns have had an impact on both garlic harvests and the forecasting of garlic prices [5–7]. Climate change has led to the occurrence of extreme weather phenomena, such as high temperatures, droughts, and flooding. In 2021, several Western European countries, including Germany, Belgium, and Austria, experienced substantial crop damage due to record-breaking heavy rainfall and flooding of up to 150 mm/day [8]. Moreover, in North America, a heatwave with temperatures of up to 49.6℃ in western Canada and 54.4℃ in California led to significant agricultural damage [9]. This has a notable impact on crop price volatility, as crop production is influenced by diverse climatic factors, including temperature, precipitation, and sunlight levels [2,3]. Garlic, in particular, is highly sensitive to climate [4], making it challenging to predict garlic prices under abnormal climate conditions [6,7].

Studies employing various statistical or deep learning models are being conducted to predict and analyze such conditions. Recent research has demonstrated that deep learning models generally outperform statistical models in the realm of time series forecasting involving multiple variables, including predicting stock prices, crop prices, and production costs. However, garlic prices are not only affected by climate but also by various variables such as demand trends and imports. As a result, deep learning models commonly employ dozens of dependent and independent variables in garlic price forecasting. However, a conventional deep learning model alone cannot accurately determine the detailed impact of major variables, leading to less accurate predictions. To address this issue, attention mechanism models in the field of time series have been implemented [10–14]. These models are capable of improving the performance of garlic price forecasting by accounting for variable correlations and interactions.

Despite their advantages, attention mechanism-based models for analyzing and forecasting garlic prices have certain limitations. For example, the dual-stage attention-based RNN (DA-RNN) and temporal fusion transformers (TFT) models both rely on attention mechanisms, but employ different operational processes, leading to potentially different interpretation of results. Thus, it is important to account for the unique characteristics of each model.

In this study, we propose a novel approach called the ATT-Hybrid DA-RNN model using the DA-RNN and TFT, which combines the attention weights of the two models. The attention-based deep learning model showed excellent performance in predictions using multiple variables. In the ATT-Hybrid DA-RNN, the characteristics of TFT using future temporal impacts, such as holidays and DA-

RNN, reflecting the characteristics of time and features in the two attention stages, can be evenly utilized in the proposed model. In addition, various pieces of information affecting prices, such as climate, trading volume, and consumption trends, are entered into the model. The proposed model utilizes the characteristics of both the DA-RNN and TFT to use the effects of various types of information for prediction.

The remainder of this paper is organized as follows. Related works are presented in Section 2, and the methods and details of our proposed model are explained in Section 3. Section 4 presents the dataset and the experimental design. The experimental results are described in Section 5. Section 6 provides conclusions and the future research directions.

2. Related works

Classical statistical models have been widely used in agricultural product prediction studies; however, recent studies employing deep learning models have also been conducted in various ways. In addition, in a study using a deep learning prediction model, it was proven that the performance of the deep learning model was superior to that of the classical statistical model.

Deep learning models have been utilized to agricultural forecasting. For example, Jin et al. [15] used an LSTM model with seasonal and trend decomposition using loess (STL) to predict the prices of Chinese cabbage and radish, achieving accuracies of 92.06 and 88.74%, respectively. In another study, Dahikar and Rode [16] employed an ANN model to forecast agricultural crop yield, using various parameters related to soil and atmosphere. Several studies have applied deep learning models to agricultural forecasting. For instance, Subhasree and Priya [17] used LSTM to predict vegetable prices, while Kumar [18] forecasted the prices of agro-products using LSTM. In both cases, the LSTM model outperformed other methods such as ARIMA and regression, achieving high accuracy scores. Subhasree and Priya also suggested optimal parameters for each crop. Similarly, Kumar found that LSTM had the highest accuracy, with an R-squared score of 0.94 compared to 0.45 for seasonal ARIMA and -5.06 for multiple linear regression.

Attention mechanisms are being increasingly used in agricultural forecasting due to their ability to improve long-term time-series predictions. Yin et al. [11] proposed an LSTM model based on an attention mechanism and STL, a technique for decomposing time-series data. The resulting STL-ATTLSTM model achieved higher accuracy compared to general deep learning models LSTM and STL-LSTM, with mean absolute performance error (MAPE) of 7% and root mean square error (RMSE) of 380. Similarly, Alwis et al. [19] proposed DA-LSTM, another LSTM model based on the attention mechanism, which outperformed other models such as XGBR and SVR. Gu et al. [20] proposed a dual-input attention LSTM (DIA-LSTM) which achieved even higher accuracy when various variables were used, with a MAPE 1.41 to 4.26% lower than that of benchmark models such as LSTM, GCN-LSTM, and DA-RNN.

Hybrid attention approaches have demonstrated the effectiveness of deep learning models in various application domains, including natural language and image processing. For instance, Gao et al. [21] proposed a deep-learning classification model based on hybrid attention for text-relation classification. In a text relation classification study, the accuracy of the model was improved by combining instance-level attention and feature-level attention. Shen et al. [22] proposed a deep-learning classification model based on hybrid attention in the field of natural language processing. In a natural language processing study, the accuracy of the model was improved by hybridizing soft and hard attention. Islam et al. [23] proposed a HAM-Net with a hybrid attention mechanism that includes temporal soft, semi-soft, and hard attention. The proposed model showed higher accuracy than the state-of-the-art model in the image field. In the papers by Gao et al. [21] and Islam et al. [23], a unique feature is the use of different aspects of attention within a single model. In our study, we applied a hybrid-attention deep learning model to the field of time series forecasting, which is a departure from previous studies. By combining the attention mechanisms from two models, our approach may provide a different perspective compared to previous research.

3. Materials and methods

3.1. Attention mechanism

Bahdanau et al. proposed an attention mechanism in 2014 [24]. They used the attention mechanism to improve the recurrent neural networks (RNN) encoder-decoder proposed by Cho et al. and Sutskever et al. The encoder-decoder model reads a continuous sequence and converts it into another sequence. The encoder summarizes the input data and creates a context vector. The decoder reads the context vector to output results [25,26]. The context vector provides a list of designated factors that have significantly influenced the prediction, and visualizing it can help interpret the results. The end-to-end training of the model makes it well-suited for learning time-series data, such as climate and price. The encoder and decoder components of the model typically consist of a recurrent neural network (RNN), long short-term memory (LSTM), or gated recurrent unit (GRU) cell. Both the DA-RNN and TFT models used in this study were encoder-decoder structures, enabling them to learn temporal dependencies and generate accurate forecasts.

There are two issues with the RNN Encoder-Decoder.

• If the sequence of input values is prolonged, the forecasting accuracy decreases; that is, a vanishing gradient occurs.

• Information loss may occur during the process of compressing the input information into one context vector, resulting in a decrease in forecasting performance.

In this study, the attention mechanism helped the model in several ways. Firstly, it maintained high accuracy even with prolonged input sequences for garlic price forecasting by focusing on highly relevant time steps. Secondly, despite the inclusion of diverse information such as price, import amounts, and climate data, there was no significant loss of information as the attention mechanism prioritized factors significantly related to forecasting. Finally, the attention weights generated by the mechanism enabled users to identify the most influential factor in the forecasting process, thereby enhancing the interpretability of the results. Overall, an attention-mechanism-based deep learning model was shown to improve both the accuracy and interpretability of garlic price forecasting.

In this experiment, the attention mechanism works as follows.

1) After the garlic-related dataset is inserted, the hidden states of the encoder are reviewed to calculate the degree of similarity to the hidden state (s_t) of the decoder in time step t to be predicted. This is known as the attention score (α^t).

2) The probability distribution value of the attention weight is calculated by applying the softmax function to the attention score calculated in Step 1. This is called attention distribution (e^t).

3) The process of multiplying each hidden state and the attention weight of the encoder and adding them is performed. This results in the calculation of the context vector, also known as the

attention value (a^t).

4) The vector ($[a_t; s_t]$) is calculated by concatenating the attention value calculated in Step 3 with the hidden state at time step t of the decoder.

5) After multiplying the calculated v_t in Step 4 by the weight matrix (W_A), the bias (b_A) is added to apply the hyperbolic tangent function. Through this process, a new vector ($\tilde{s}_t$) is calculated and used again as the input of the output layer to calculate the forecasting vector. The final forecast price for garlic is then derived.

$$\alpha^t = softmax(s_t^T h_1 + \cdots + s_t^T h_n) \tag{1}$$

$$a_t = \sum_{k=1}^{n} \alpha_k^t h_k \tag{2}$$

$$\widetilde{s_t} = tanh(W_A[a_t; s_t] + b_A) \tag{3}$$

$$\hat{y}_t = softmax\left(W_y \tilde{s}_t + b_y\right) \tag{4}$$

3.2. Dual-stage attention-based RNN (DA-RNN)

DA-RNN is an attention-based deep learning model used to solve the problem of poor performance and inability to identify the main factor in target forecasting among input factors when non-linear autoregressive exogenous (NARX) model has long target data and input data [27]. This model shows better forecasting performance by sequentially applying both input attention and temporal attention mechanisms to the encoder and decoder, respectively. In this study, the operation process of the DA-RNN was as follows:

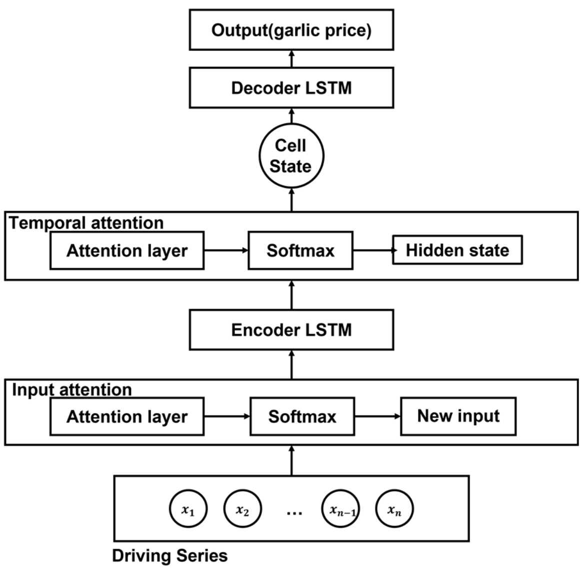

Figure 1. Dual-stage Attention based RNN.

DA-RNN is an RNN-based encoder-decoder network. The encoder component of the architecture takes in a sequence of time-stamped values and encodes them into a fixed-length representation, which summarizes the input information. This representation is then passed to the decoder component, which uses the information to generate a predicted future sequence of time-stamped values. In this study, the input attention mechanism calculates the attention score to identify the driving series with the greatest association with the target series, garlic price, at each time step by referring to the hidden state of the previous encoder. Temporal attention refers to the hidden state of the previous decoder and calculates an attention score to determine which encoder hidden state is most associated with garlic prices during the entire time point. The context vector generated through the previous process is used as the input of the garlic price prediction model decoder. The output of the last decoder is the predicted garlic price [27].

3.3. Temporal fusion transformers (TFTs)

TFT differs from DA-RNN by incorporating two types of inputs. TFT forecasts using variables such as observed input factors (past inputs in Figure 2), such as price or amount of imports, and known future inputs (in Figure 2), which are known in advance at future times, on weekends and holidays. In addition, fixed static metadata (covariates), such as location or type, are used in the model [28]. In this study, the operation process of the TFT was as follows.

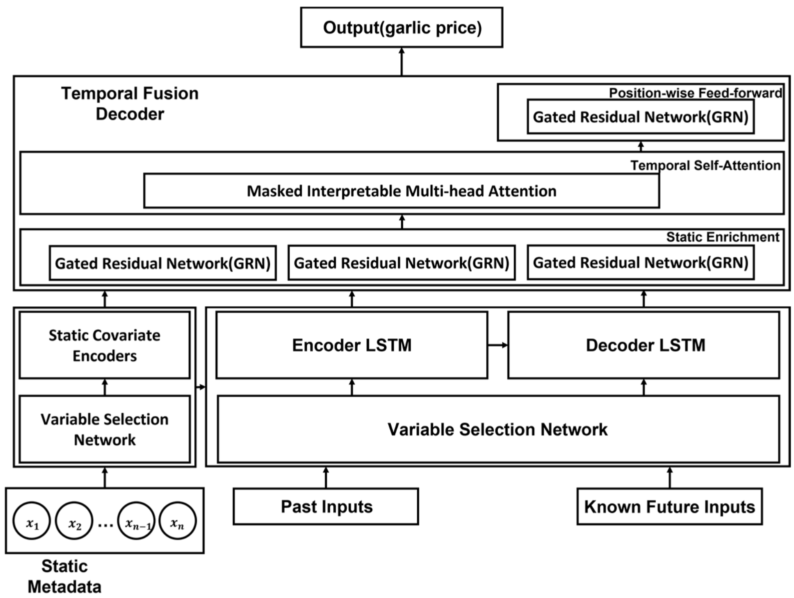

Figure 2. Temporal fusion transformers.

TFT inputs static metadata related to garlic prices, time-varying past inputs, and time-varying a priori known future inputs that can affect garlic sales, such as holidays and weekends. The prediction process of TFT involves entering observed variables from past time points into the LSTM encoder, and known variables from both past and future time points into the LSTM decoder. This is done using the time series data. Because static covariates are fixed variables, sufficient information is passed through the variable selection network, LSTM encoder, and Gated Recurrent Network (GRN) to the entire layer. A GRN uses gates to determine which information to keep and which information to discard, allowing it to better capture long-term dependencies in sequential data. All variables are evaluated at every point through the selection network process and only important variables are entered into the model. Inputs from a series of processes are passed to the Masked Interpretable Multi-Head Attraction Layer. The information produced by the Attention Layer is entered via the Position-Wise Feed-Forward Nonlinear Layer (GRN). The final output value reduces the complexity of the model by reflecting the output information from the LSTM decoder. [28].

3.4. Proposed ATT-hybrid DA-RNN method

The DA-RNN and TFT used in this study calculate the attention weight. However, each model has a different structure and shape of the dataset used, and the results of the calculations are different. In addition, since the performance advantages of the two models vary depending on the dataset, it is necessary to build an advanced performance model by combining the attention weights of the two models. As the results and attention weights of the two models used in the prediction process differed, an appropriate combination was needed to evenly utilize the advantages of both models and address errors that occurred during prediction. In this study, the attention weight of the TFT was extracted as a time-series window corresponding to days 2, 5, 8, and 10, and combined with DA-RNN to reflect the influence of the TFT's attention weight in the calculation process of DA-RNN. The study proceeded in the following order: 1) data preprocessing and variable processing, 2) training TFT and DA-RNN models, 3) model training using attention extracted from TFT combined with DA-RNN (i.e., the proposed model), and 4) forecasting garlic price and visualizing the attention weight using the proposed model. The proposed model followed the steps outlined below:

1) Train DA-RNN and TFT using pre-processed datasets, respectively.
2) Extract an attention of TFT and save it.
3) Build a DA-RNN that uses the attention weight of TFT as an initial attention value to reflect the attention value of TFT (proposed model).
4) Forecast garlic price and visualize the attention weight using the proposed model.

Figures 3 and 4 depict the pseudocode and architectural diagram of the proposed model, respectively.

The attention mechanism is effective for processing multiple variables, as it can apply the main variables that affect forecasting differently based on their attention weights. These weights can also be used to interpret the model's forecasted results. However, using two types of models can make it difficult to interpret the results, and thus a solution is needed.

Therefore, we propose a model that combines the attention of these two models to improve performance and interpretation. The proposed model can also solve the problems presented in the Introduction. The characteristics of the proposed model are as follows: First, our garlic price forecasting model was constructed using deep learning technology, that is suitable for long-term and multiple time-series data. The attention mechanism, encoder-decoder, and LSTM were used in the proposed model. Next, time series data were input for both the DA-RNN and TFT used for attention hybridizing in this study. Therefore, preprocessing that considers the characteristics of the time-series data was necessary. Hence, the following two techniques were used: First, to ensure the stationarity of the time series data, the STL technique is used. Loess, an abbreviation of local regression, is applied to eliminate seasonal trends that cause instability in the time series data. Second, technical indicators, which are a type of technical analysis mainly used for predicting stock prices, were used to improve the performance of the model. Subsequently, the attention weights calculated from two models, DA-RNN and TFT, with an attention mechanism were hybrid-trained.

Algorithm: ATT-Hybrid DA-RNN

```
Input : Pre-processed dataset(P_d)
Output : Predicted garlic price
Procedure ATT-Hybrid DA-RNN(A_TFT, P_d)
input_data := P_d
model := Temporal Fusion Transformers
length := [2, 5, 8, 10] //variable window size
for epoch i = 1, …, N do
        t := Length P_d
        Save TFT_att_tensor(α_0, …, α_t) into T_α
        for length = 2, 5, 8, 10 do
                Slice T_α_1, …, T_α_N as length into ST_α_length
                Save ST_α_length into Attention weight of TFT(A_TFT)
        end
end
feature_attention := load_attention to tensor
proposed_model := ATT-Hybrid DA-RNN
for epoch i = 1, …, N do
        Train proposed_model(feature_attention, input_data)
        output Predicted garlic price
end
end Procedure
```

Figure 3. Pseudo code of the proposed model.

The proposed model considered the attention of both models, making it easier to interpret the results. Furthermore, the results derived by each model can complement each other, thereby improving the prediction results.

The forecasting of garlic prices was conducted in four steps. The first step was the data preparation. After collecting the raw data, the stationarity of the time series data was determined through the augmented Dickey-Fuller test (ADF test). Thereafter, the non-stationary time series data extracted only the residual, which is a stable component, through STL. Garlic price data were also processed to generate technical indicators. The second step was data preprocessing. In this process, the data prepared in Step 1 were cleaned. Data cleaning is the process of examining whether there are any missing ones or outliers. The outliers were removed, and the missing values were interpolated. Next, data with different scales were normalized using min-max scaling. Finally, we split the training and test data for model training. Step 3 in the proposed model was attention hybridization. Training of the ATT-hybrid DA-RNN was conducted in such a way that after training the TFT model, the attention weight was delivered to the DA-RNN. Step 4 involves the evaluation of the model. The accuracy of the forecast values calculated by the proposed model, DA-RNN, and TFT was evaluated using RMSE and MAPE. In addition, the results of the attention weight calculated in the model were visualized. Detailed descriptions of the technical and evaluation indicators are described in detail in Section 4.

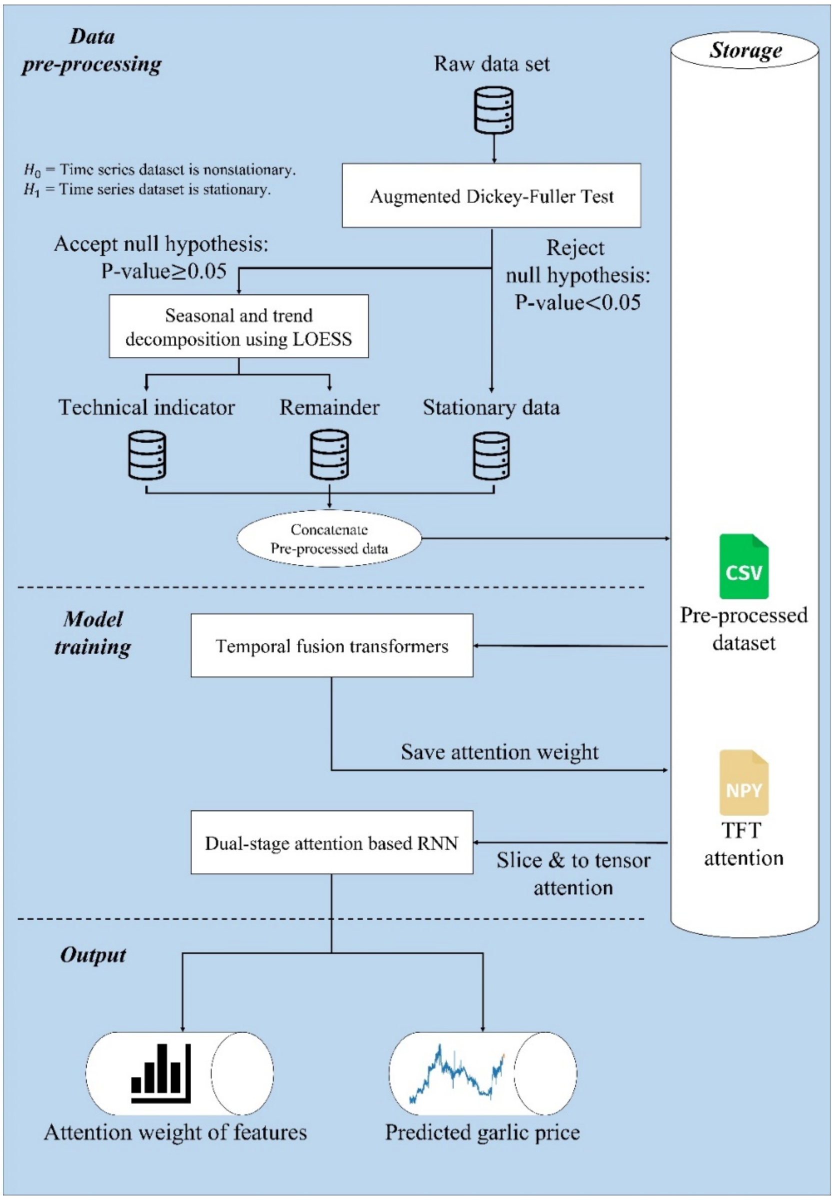

Figure 4. Structure of the proposed model.

4. Experiment

4.1. Dataset

4.1.1. Tables and data

The data used in the garlic price forecasting model in this study can be classified into climate, price, quantity, oil price, search trend, and technical indicators. The schema for each variable in the dataset is given in Table 1. The dataset contains 2738 rows and 36 features collected from January 1, 2014, to June 30, 2021, including the price of garlic, which was the target variable. For the climate data, a dataset from the main garlic production sites prepared by the National Institute of Horticultural and Oriental Medicine, a national research institute in Korea, was used. The main garlic production areas, namely Muan, Jeju, Seosan, and Changnyeong, as described in Section 4.1.2, provided the climate data. The model also employed the price and quantity data of garlic in the wholesale market, as described in Section 4.1.3, and oil data associated with the logistics process.

The oil dataset contains price data for crude oil, gasoline, and diesel, and the details are described in Section 4.1.4. Delivery, HMR, and online shopping trend data related to supply and demand were also used, as described in Section 4.1.5. Additionally, three major technical indicators were processed to analyze trends and momentum, and the details are provided in Section 4.3. Finally, the data were divided into training and testing sets with a ratio of 0.8:0.2, respectively.

Table 1. Input variables and sources.

Main category	Parent subcategory	Child subcategory	Source
Climate	Main product area: Muan, Jeju, Seosan, Changnyeong	Minimum temperature Maximum temperature Humidity Precipitation Solar insolation amount Sunshine duration Average wind speed Maximum wind speed	Korea Meteorological Administration / National Institute of Horticultural and Herbal Science
Price	Product price (Garak market)	Onion (Korean) Garlic (Korean) Garlic (Imported)	NongNet
Volume	Carry-in (Garak market)	Onion (Korean) Garlic (Korean) Garlic (Imported)	NongNet
Oil price	Crude	Price	Opinet
	Diesel	Competitive price Entire market price	
	Gasoline	Competitive price Entire market price	
Search trend	Garlic	Garlic	Google Trends
	Home Meal Replacement	HMR, Meal kit, Cooking box	
	Delivery	Delivery, Application 1[1], Application 2[2]	
	Online shopping	Online shopping, Shipping	
Technical indicator	Price data residuals and indicators	EMA MACD RSI	-

[1, 2] The top 2 delivery service application in South Korea (1: BAEMIN, 2: Yogiyo)

4.1.2. Climate data

In this study, climate data from Muan, Jeju, Seosan, and Changnyeong, the locations of the main garlic production selected by the National Institute of Horticultural and Herbal Science, were used to apply the climate impact to the price model. To verify the impact of the climate data, a model was built for each main production area. Muan, Jeju, Seosan, and Changnyeong are located in Jeollanam-do, Jeju-do, Chungcheongnam-do, and Gyeongsangnam-do, respectively, which are considered as the main garlic-producing areas in various regions of Korea. Given that climate data consist of temperature, humidity, precipitation, amount of sunlight, and wind speed, various climate trends in each main production area can be input into the model. By individually constructing models according to datasets for each main production area, we attempted to compare them to verify the effectiveness of the climate. In particular, the temperature was given as input by dividing the minimum and maximum temperatures so that extreme values could be properly considered.

4.1.3. Price and volume data

Price and volume data from the Garak Market, provided by NongNet, were used. The Garak Market is Korea's first and largest public wholesale market. In addition, considering the volume of agricultural and marine products traded per unit area in this market, it is the densest wholesale market worldwide. The Garak market has an area of 543,451 m^2, and the transaction volume is approximately 2.3 million tons per year and 7500 tons per day. This accounts for 49% of the total agricultural and marine product consumption in Seoul, and 40% of the total nationwide transaction volume.

According to a survey by the 'Korea Agro-Fisheries & Food Trade Corporation', Japan's Ota wholesale market, which is a similar area, has an annual trading volume of approximately 800,000 tons. Italy's Roman wholesale market, which is approximately three times larger, also has an annual trading volume of approximately 800,000 tons, while Spain's Madrid wholesale market has an annual trading volume of approximately 1.6 million tons [29]. An analysis of the wholesale market import volume from 2014 to 2021 by the Agricultural Products Distribution Information System (NongNet) revealed that the Garak market accounts for 30% of the transaction volume of garlic by market. Moreover, peeled garlic makes up a significant portion of the garlic trading volume, accounting for 46% of the garlic variety trading volume [30]. With the expansion of the online market and the recent spread of COVID-19, the preference for peeled garlic has increased since it is easy to use and can be purchased in small packages [31].

Therefore, domestic peeled garlic price data, which is the subject of the study, and the price data of onions and imported garlic used as additional related variables were collected. As this study builds a daily price forecasting model, day-to-day data were used.

4.1.4. Oil price data

Oil prices reflect the logistics costs and prices. If oil prices soar, logistics costs increase, and hence, the prices of various items, including garlic, increase, and vice versa. Oil prices were included in the dataset and diesel and gasoline were classified into separate variables and used as independent variables in the model. To collect the data, we obtained information from Opinet, which is operated by the Korea National Oil Corporation and provides data on the KRX market of the Korea Exchange. Diesel and gasoline price data are divided into the total transaction price and competitive transaction price, which is the average value of the competitive transaction price and negotiated transaction price. The competitive transaction price is divided into the price of the volume of competitive transactions sold by a large number of unspecified sellers and buyers through the e-commerce system.

4.1.5. Search trend data

Consumption trends can be observed in the search trend data. To apply the consumption trend of the rapidly growing delivery market and the meal kit market post COVID-19, trend data of Korea's representative delivery applications and meal kits, HMR, and cooking boxes were used. In addition, the consumption interest in garlic itself was reflected in the garlic trend data. The data were provided by Google Trends.

In particular, among the data belonging to the delivery category, BAEMIN, which corresponds to Application 1, and Yogiyo, which corresponds to Application 2, accounted for about 69 and 20% of the market share, respectively. In addition, since 2019, the frequency of use of delivery and online shopping has increased owing to COVID-19; therefore, data from delivery, online shopping, and shipping have been used to reflect this trend.

4.2. Pre-processing

4.2.1. Augmented Dickey-Fuller test (ADF test)

Time-series data were recorded over time. These time-series data (y_t) can be divided into three elements, assuming additive decomposition. Seasonality, trend, and remainder are included. Each component is defined as follows:

• Seasonality (S_t): Seasonality indicates periodic fluctuations. Seasonality may appear based on various frequencies, such as semi-annual, quarterly, monthly, or weekly.

• Trend (T_t): Trends indicate that data increase or decrease linearly. Time series data may show a trend over time, and it is interpreted that the direction changes as the trend changes.

• Remainder (R_t): The remainder is the component from which seasonality and trends have been removed from time series data.

Datasets with seasonality or trends are referred to as nonstationary time-series data. When non-stationary time-series data are input to the deep learning model, the stability of the derived results decreases, and the forecasting accuracy decreases. The augmented Dickey-Fuller test (ADF test) can validate these non-stationary data. The hypothesis of the ADF test is as follows:

• H_0: The data does not satisfy stationary conditions. (= There is a unit root in the data)

• H_1: The data satisfies stationarity. (= There is no unit root in the data)

That is, if the p-value is less than 0.05, as a result of the ADF test, the null hypothesis is rejected and it can be determined that the stationary hypothesis is satisfied. The time-series data used in this study consisted of climate, import volume, price, and trend. Because it is not known whether the data satisfies stationarity with the actual measured data, non-stationary time series data are determined through the ADF test.

4.2.2. Seasonal and trend decomposition using local regression (STL decomposition)

The seasonality and trend that undermine the stationarity of the time series were removed from the data determined through the ADF test. This process was performed via STL decomposition. STL is a technique for decomposing time series data into seasonal, trend, and remaining components, as described in Section 4.2.1 [32]. Unlike techniques such as X11 [33] and SEATS [34], which can only use monthly or quarterly data, STL decomposition has the advantage of being applicable to all types of data. It also has the advantage of robustness against outliers, because it obtains a robust weight from the outer loop. y_t represents the original data at time t, and S_t, T_t and R_t represent the seasonality , trend and residual components, respectively. The STL decomposition can be expressed using the following equation:

$$y_t = S_t + T_t + R_t \tag{5}$$

The STL decomposition algorithm was performed using the outer and inner loops. The outer loop is a process of minimizing the influence of outliers, and the inner loop is the process of calculating and removing seasonal and trend components. First, in the inner loop, seasonal performance trends were eliminated in six steps.

1). De-trending
2). Cycle-subseries smoothing
3). Low-pass filtering of smoothed cycle-subseries
4). De-trending of smoothed cycle-subseries
5). De-seasonalizing
6). Trend smoothing

In the outer loop, the remaining components and robustness weights were calculated using the seasonal and trend components calculated in the initial inner loop. Subsequently, the inner loop is repeated. In this study, STL decomposition was used for the ADF test results from the data determined as a non-stationary time series. Each dataset was separated by seasonality, trends, and residuals using STL decomposition. Only the remaining data were used to secure the normality of the data.

4.3. Technical indicator

Technical indicators are primarily used in stock market analyses. This is a way to focus on the rise and fall in stock prices. Thus, price data that move in this direction, such as stock prices, can also be used. The technical indicators used in this study were divided into the following categories. Trend indicator: This was used to determine the direction of the indicator. The trend indicators used in this study were the exponential moving average (EMA) and moving average convergence divergence (MACD).

Momentum indicator: This is used to identify the inflection point at which the trend of the indicator changes. The momentum indicator used in this study is the relative strength index (RSI).

4.3.1. Exponential moving average (EMA)

The moving average is the average value of the data for a certain period. It is an indicator designed to examine how data change over time. EMA is designed to increase the influence of recent data by reflecting weights on the moving average and lowering the influence of past data. In this study, setting the power line to 20 days in the technical analysis of stock prices is applied to the EMA calculation period. The calculation formula of the EMA is shown in Eq (6). W is the weight, t is the calculation period, and p_t is the current price of garlic.

$$EMA_{(t)} = (1 - W) \cdot EMA_{(t-1)} + p_{(t)} \tag{6}$$

4.3.2. Moving average convergence divergence (MACD)

MACD is an indicator that reflects the convergence and divergence of the long- and short-term moving averages. Changes in the trend or flow of garlic prices can be observed through MACD. In general, the 12-day EMA and 26-day EMA are used as short-term index movement averages (EMA_s) and long-term index movement averages (EMA_i), respectively. Therefore, we used the 12-day garlic price EMA and 26-day garlic price EMA. The calculation formula for MACD is given by Eq (7).

$$MACD = EMA_s - EMA_i \tag{7}$$

4.3.3. Relative strength index (RSI)

RSI is an indicator of the inflection point of the trend. The strength of the current rise and fall is expressed as a percentage. Based on the stock market, if the index is below 30 lines, it is judged as the lower limit, and if the index is above 70 lines, it is judged as the upper limit. If U is the increase in garlic price, D is the decrease in garlic price, AU is the average of U for k days, and AD is the average of D for k days, then the formula for the RSI is given by Eq (8).

$$RSI = \frac{AU}{AU+AD} \tag{8}$$

4.4. Measurement criteria

The root mean square error (RMSE) and mean absolute performance error (MAPE) were used as performance indices to compare the forecasting accuracy of the models. The RMSE is an index that shows the average error between the forecast and actual values, and makes intuitive evaluation feasible. To calculate the RMSE, the difference between the actual (y_i) and forecast ($\hat{y}_i$) values of n data points is squared and added together. i is an index variable that runs over the number of data points in the sample, n. It represents the ith data point in the sample. This value was again divided by n. The RMSE value is calculated when the square root of the result is obtained. The closer the RMSE value is to 0, the smaller the error. The RMSE formula is expressed by Eq (9).

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{9}$$

The MAPE is an indicator of the percentage of errors and can be evaluated relative to each other. To calculate the MAPE, the difference between the actual (y_i) and forecast ($\hat{y}_i$) values is calculated and divided by y_i. The absolute values of the divided values were summed by the number of data, n. Finally, 100 % of n is multiplied. The MAPE formula is expressed by Eq (10).

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{10}$$

4.5. Hyper parameter

DA-RNN, TFT, and the proposed model use time-series data; therefore, the user must set the time-step value of the data that is input to the model. In addition, the user must set the hidden state size, dropout, and ratio of the training and test data. In this experiment, we examined the effect of period by setting the time-step value to {2, 5, 8, 10}. The hidden state size of the encoder and decoder was set to 128, which showed the highest performance in previous studies on DA-RNN and TFT. Dropout was set to 0.2, and the ratio of training and test data was set to 0.8 and 0.2, respectively.

Finally, the epoch was set to 100; however, through early stopping, learning was stopped where val_loss was the lowest, and the best model was saved. The patience of early cessation was 4.

Table 2. Hyper parameters of the proposed model.

Location	Time-step	Encoder hidden states	Decoder hidden states	Batch size	Dropout	Loss function
Muan Jeju Seosan Changnyeong	2, 5, 8, 10	128	128	16	0.2	Smooth L1

5. Results and discussion

In this study, we compared the performance of the benchmark and proposed models. For comparison, data from Muan, one of the main garlic production areas, were used, and the accuracy is shown in Table 3. For comparison on the same basis, a time-step value of 5 was applied to all three models. Consequently, the proposed model showed the lowest error with an RMSE of 350.19 and MAPE of 4.3 %. On the other hand, DA-RNN showed the second lowest error with an RMSE of 439.22 and MAPE of 6.9%. TFT showed the highest error with an RMSE 662.3 and MAPE of 7.1%. As mentioned in Section 4.4, RMSE and MAPE are indicators related to the degree of error, and hence, a lower is better.

Table 3. Comparison of forecasting accuracy between benchmark models and the proposed model.

Location	Metric	DA-RNN	TFT	Proposed model
Muan	RMSE	439.22	662.3	**350.19**
	MAPE	6.9 %	7.1 %	**4.3 %**

Table 4. Forecasting accuracy according to different time-steps.

Location	Metric	Time-steps			
		2	5	8	10
Muan	RMSE	402.96	**350.19**	365.69	384.06
	MAPE	5.12 %	**4.25 %**	5.07 %	5.19 %
Jeju	RMSE	413.78	**363.9**	395.9	444.42
	MAPE	5.72 %	**4.76 %**	5.72 %	6.8 %
Seosan	RMSE	402.72	**349.57**	365.53	374.23
	MAPE	5.23 %	**4.54 %**	5.26 %	5.64 %
Changnyeong	RMSE	401.62	**386.3**	382.18	387.14
	MAPE	5.46 %	**4.97 %**	5.75 %	5.48 %

Next, to compare the performance of each time step of the proposed model, models of 2, 5, 8, and 10 time steps were constructed. In addition, the proposed model was constructed by dividing it into four main production areas, and the results are shown in Table 4. In all the main production models, the time-step value was highest when the time-step value was 5. When classified as the main producing area, the model trained with Muan's climate data showed a high accuracy on average.

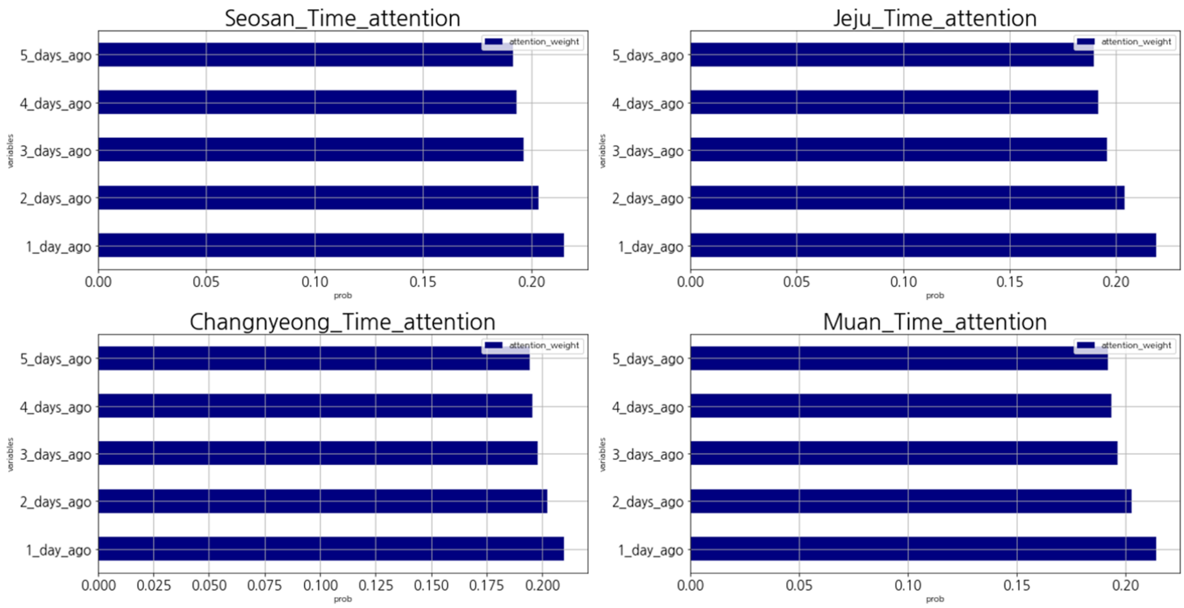

Figure 5. Time attention of the proposed model.

Finally, the result factors of the attention weight from the proposed model are as follows: Figure 5 displays the time attention graph for Muan, Jeju, Seosan, and Changnyeong. Time attention represents the importance of attention by the time step. Figure 6 presents a feature attention graph for Muan, Jeju, Seosan, and Changnyeong. Feature attention represents the importance of the attention to each variable. Figure 5 indicates that the time attention weight of the day before is the highest in all regions. Since the deviation of the attention weight of each time-step is not large, it can be seen that the data of all periods are appropriately reflected. Figure 6 reveals that the feature attention of gasoline prices is the highest in all regions. As climate-related variables also rank at the top, it can be observed that climate variables had a lot of influence on predictions. With regard to technical indicators, it can be confirmed that the influence is meaningful because they occupy the median. Moreover, since the attention scores differ for each region, it can be seen that it is meaningful to build a new model for each region. It is expected that if decision-making can be more effective if variables with high feature attention scores are considered. Lim. H's study [10] showed that the prediction performance would deteriorate if the variables with high attention scores were removed from the model. Therefore, the attention scores also play a significant role in time series prediction models. Through the attention scores of four experimental models, this study provides the following insights. Firstly, the influence of raw materials that directly affect logistics and production, such as gasoline or diesel, is significant. Secondly, the quantity of garlic imports directly affects the domestic garlic price. Additionally, indicators related to onions, a substitute, also had some impact, as did technical indicators of garlic price. However, it was found that delivery or search term indicators have little effect on garlic price prediction. Thus, constructing a dataset focusing on the variables that showed significant attention scores in this study is expected to yield more meaningful results for subsequent research.

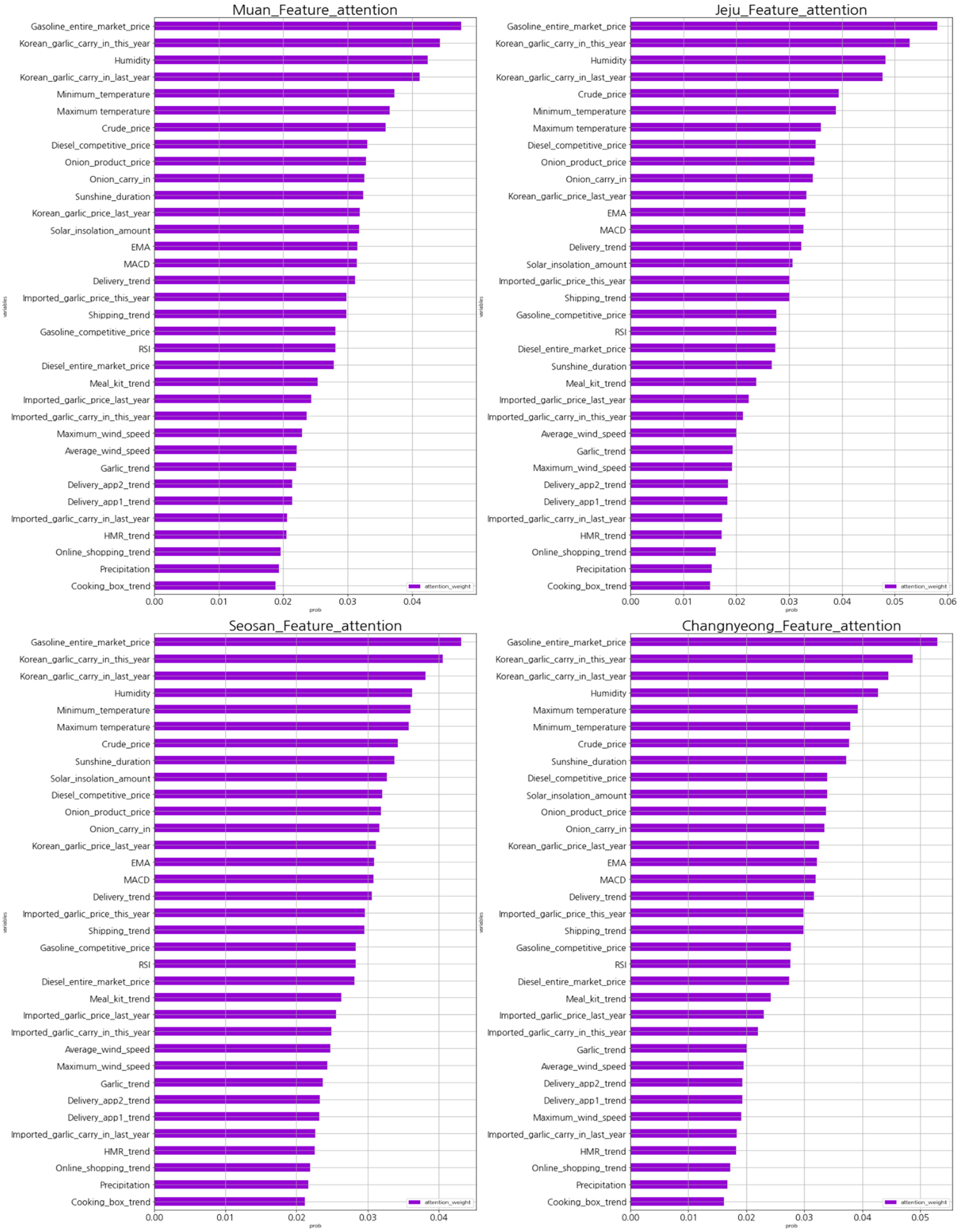

Figure 6. Feature attention of the proposed model.

Table 5. Time attention weight of the proposed model.

Location	Time attention weight				
	1 day ago	2 days ago	3 days ago	4 days ago	5 days ago
Muan	0.214468	0.203157	0.196682	0.193546	0.192148
Jeju	0.219068	0.204119	0.195688	0.191516	0.189610
Seosan	0.215249	0.203399	0.196648	0.193193	0.191511
Changnyeong	0.209997	0.202367	0.197882	0.195490	0.194264

6. Conclusions

Garlic is an essential agricultural product in the Korean diet, but its production varies greatly depending on the climate, making it difficult to predict prices and ensure stability of supply and demand. As a result, the government has identified garlic as one of the top five vegetables for price management. To effectively manage supply and demand based on predicted prices, a deep-learning model capable of processing multiple variables over an extended period of time is required.

In this study, we propose a forecasting model that combines the attention weights of two existing models that show excellent performance in predicting garlic prices. We aimed to improve the accuracy of garlic price forecasting by leveraging the strengths of both models. For the experiment, we utilized climate data from the primary garlic production area, garlic import data from the Garak market, and price data. We also incorporated trend indicators to reflect consumption trends and other indicators, such as oil prices and technical indicators, to capture additional factors that may impact garlic prices. To evaluate the proposed model's performance, we compared it with two existing high-performance models, DA-RNN and TFT, which are time series forecasting models that utilize the attention mechanism. By comparing our model's performance against these benchmark models, we can determine whether our model is an improvement in the field of garlic price prediction.

The training process of the proposed model was as follows. The proposed model was trained by hybridizing the TFT's attention to the DA-RNN. The forecasting accuracy and attention weight of the three models, including the proposed model, were derived, and the forecasting accuracy was compared using the evaluation indicators RMSE and MAPE. DA-RNN showed an RMSE of 439.22, MAPE of 6.9%, TFT RMSE of 662.3, MAPE of 7.1%, proposed model RMSE of 350.19, and MAPE of 4.25%. The model proposed in this study showed better prediction accuracy in garlic price forecasting experiments than benchmark models. Through the visualization of the attention weight, the major variables that affect prices were derived. In four experimental models, gasoline prices and garlic import quantities from the current and previous year, as well as humidity, were identified as key factors. This is expected to help the Korean government manage the supply and demand of garlic and stabilize prices in the future, and has contributed to research on indicators from various perspectives not only for garlic producers but also for consumers.

In future work, we plan to conduct an attention hybrid study using models with other attention mechanisms. In addition, we can verify and analyze which model attention is effective when used for model training. Then, we can analyze various technical indicators that reflect the time series characteristics of garlic price through attention weight. It is expected that a garlic price prediction model having enhanced performance can be built through future research. Further, it can help the government or individuals in making proper decisions. Finally, decision makers can identify signal variables that affect garlic prices through the attention received from the proposed model.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. J. H. Ha, S. T. Seo, S. W. Kim, Improving forecasting performance for onion and garlic prices, *J. Korean Soc. Rural Plann.*, **25** (2019), 109–117. https://doi.org/10.7851/ksrp.2019.25.4.109
2. A. Ranjana, S. C. Mehta, Weather based forecasting of crop yields, pests and diseases-IASRI models, *J. Ind. Soc. Agril Stat.*, **61** (2007), 255–263.
3. F. Tao, M. Yokozawa, Z. Zhang, Modelling the impacts of weather and climate variability on crop productivity over a large area: a new process-based model development, optimization, and uncertainties analysis, *Agric. For. Meteorol.*, **149** (2009), 831–850. https://doi.org/10.1016/j.agrformet.2008.11.004
4. S. Choi, J. Kim, H. Seo, Estimation of garlic's bulb weight at harvest using a multi-level model based on growth and meteorological data, *Hortic. Sci. Technol.*, **39** (2021), 521–529. https://doi.org/10.7235/HORT.20210047
5. S. Choi, J. Baek, Garlic yields estimation using climate data, *J. Korean Data Inf. Sci. Soc.*, **27** (2016), 969–977. https://doi.org/10.7465/jkdi.2016.27.4.969
6. J. Choi, P. Helmberger, How sensitive are crop yields to price changes and farm programs?, *J. Agric. Appl. Econ.*, **25** (1993), 237–244. https://doi.org/10.1017/S1074070800018794
7. *KREI (Korea Rural Economic Institute)*, A Study on the Development of Agri-food Trade Model Reflecting Climate Change, 2021. Available from: https://repository.krei.re.kr/bitstream/2018.oak/26046/1/R901.pdf
8. Hortidaily, 2021. Available from: https://www.hortidaily.com/article/9340224/heavy-rain-causes-severe-damage-to-open-field-fruit-and-vegetable-crops-in-western-europe.
9. *REUTERS*, 2021. Available from: https://www.reuters.com/world/us/wither-away-die-us-pacific-northwest-heat-wave-bakes-wheat-fruit-crops-2021-07-12/.
10. H. Lim, *Dual Attention-based LSTM Model for Produce Price Prediction*, Master thesis, Sejong university graduate school, Seoul, Korea, 2020.
11. H. Yin, D. Jin, Y. H. Gu, C. J. Park, S. K. Han, S. J. Yoo, STL-ATTLSTM: vegetable price forecasting using STL and attention mechanism-based LSTM, *Agriculture*, **10** (2020), 612. https://doi.org/10.3390/agriculture10120612
12. S. Y. Shih, F. K. Sun, H. Y. Lee,Temporal pattern attention for multivariate time series forecasting, *Mach. Learn.*, **108** (2019), 1421–1441. https://doi.org/10.1007/s10994-019-05815-0
13. S. Huang, D. Wang, X. Wu, A. Tang, DSANet: Dual self-attention network for multivariate time series forecasting, in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, (2019), 2129–2132. https://doi.org/10.1145/3357384.3358132
14. T. Hollis, A. Viscardi, S. E. Yi, A comparison of LSTMs and attention mechanisms for forecasting financial time series, preprint, arXiv:1812.07699.
15. D. Jin, H. Yin, Y. Gu, S. J. Yoo, Forecasting of vegetable prices using STL-LSTM method, in *2019 6th International Conference on Systems and Informatics (ICSAI)*, (2019), 866–871. https://doi.org/10.1109/ICSAI48974.2019.9010181

16. S. S. Dahikar, S. V. Rode, Agricultural crop yield prediction using artificial neural network approach, *Int. J. Innovative Res. Electr., Electron., Instrum. Control Eng.*, **2** (2014), 683–686.
17. M. Subhasree, C. A. Priya, Forecasting vegetable price using time series data, *Int. J. Adv. Res.*, **3** (2016), 535–641.
18. R. K. Kumar, *Short term forecasting of Agro-products pricing using Multivariate time series analysis*, Doctoral dissertation, Dublin, National College of Ireland, 2020.
19. S. D. Alwis, Y. Zhang, M. Na, G. Li, Duo attention with deep learning on tomato yield prediction and factor interpretation, in *Pacific Rim International Conference on Artificial Intelligence*, Springer, Cham, 2019. https://doi.org/10.1007/978-3-030-29894-4_56
20. Y. H. Gu, D. Jin, H. Yin, R. Zheng, X. Piao, S. J. Yoo, Forecasting agricultural commodity prices using dual input attention LSTM, *Agriculture*, **12** (2022), 256. https://doi.org/10.3390/agriculture12020256
21. T. Gao, X. Han, Z. Liu, M. Sun, Hybrid attention-based prototypical networks for noisy few-shot relation classification, in *Proceedings of the AAAI Conference on Artificial Intelligence*, **33** (2019), 6407–6414. https://doi.org/10.1609/aaai.v33i01.33016407
22. T. Shen, T. Zhou, G. Long, J. Jiang, S. Wang, C. Zhang, Reinforced self-attention network: a hybrid of hard and soft attention for sequence modeling, preprint, arXiv:1801.10296. https://doi.org/10.48550/arXiv.1801.10296
23. A. Islam, C. Long, R. Radke, A hybrid attention mechanism for weakly-supervised temporal action localization, in *Proceedings of the AAAI Conference on Artificial Intelligence*, **35** (2021), 1637–1645.
24. D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, preprint, arXiv:1409.0473.
25. K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, et al., Learning phrase representations using RNN encoder-decoder for statistical machine translation, preprint, arXiv:1406.1078. https://doi.org/10.48550/arXiv.1406.1078
26. I. Sutskever, O. Vinyals, Q. V. Le, Sequence to sequence learning with neural networks, *Adv. Neural Inf. Process. Syst.*, **27** (2014), 3104–3112.
27. Y. Qin, D. Song, H. Chen, W. Cheng, G. Jiang, G. Cottrell, A dual-stage attention-based recurrent neural network for time series prediction, preprint, arXiv:1704.02971.
28. B. Lim, S. Ö. Arık, N. Loeff, T. Pfister, Temporal fusion transformers for interpretable multi-horizon time series forecasting, *Int. J. Forecasting*, **37** (2021), 1748–1764. https://doi.org/10.1016/j.ijforecast.2021.03.012
29. *Seoul Agro-Fisheries & Food Corporation*, Wholesale market information (Garak market), 2021. Available from: https://www.garak.co.kr/cntnts/selectContents.do?cntnts_id=CC010200.
30. *NongNet*, Analysis of the amount brought into the wholesale market (garlic), 2021. Available from: https://www.nongnet.or.kr/anss/tkinVolmInfo.do.
31. Trends and Prospects in Supply and Demand of Seasoning Vegetables in Chapter 14, Agricultural Outlook 2021: Changes in agriculture and rural areas after COVID-19, (2021), 425–441.
32. R. B. Cleveland, W. S. Cleveland, J. E. McRae, I. J. Terpenning, STL: A seasonal-trend decomposition procedure based on loess, *J. Off. Stat.*, **6** (1990), 3–33.
33. E. B. Dagum, S. Bianconcini, 2016, *Seasonal Adjustment Methods and Real Time Trend-Cycle Estimation*, Berlin/Heidelberg, Germany: Springer International Publishing, (2016), 29–57. https://doi.org/10.1007/978-3-319-31822-6_2

34. V. Gómez, A. Maravall, Seasonal adjustment and signal extraction in economic time series, *Course Time Ser. Anal.*, **2001** (2001), 202–247. https://doi.org/10.1002/9781118032978.ch8