International Journal of Advanced Smart Convergence Vol.10 No.4 104-109 (2021)
http://dx.doi.org/10.7236/IJASC.2021.10.4.104

IJASC 21-4-13

Proposal of An Artificial Intelligence based Temperature Prediction Algorithm for Efficient Agricultural Activities -Focusing on Gyeonggi-do Farm House-

Eun-Jin Jang[1], Seung-Jung Shin[2]*

[1] *Doctor Candidate, Department of IT Convergence, Hansei University, Korea*
dmswls1061@naver.com
[2*] *Associate professor, Department of ICT Convergence, Hansei University, Korea*
expersin@hansei.ac.kr

Abstract

In the aftermath of the global pandemic that started in 2019, there have been many changes in the import/export and supply/demand process of agricultural products in each country. Amid these changes, the necessity and importance of each country's food self-sufficiency rate is increasing. There are several conditions that must accompany efficient agricultural activities, but among them, temperature is by far one of the most important conditions. For this reason, the need for high-accuracy climate data for stable agricultural activities is increasing, and various studies on climate prediction are being conducted in Korea, but data that can visually confirm climate prediction data for farmers are insufficient. Therefore, in this paper, we propose an artificial intelligence-based temperature prediction algorithm that can predict future temperature information by collecting and analyzing temperature data of farms in Gyeonggi-do in Korea for the last 10 years. If this algorithm is used, it is expected that it can be used as an auxiliary data for agricultural activities.

Keywords: *Temperature Data, AI(Artificial Intelligence),, Gyeonggi-do Farmhouse, Data Analysis, Data Prediction*

1. Introduction

As the flow of international logistics has changed due to the COVID-19 pandemic that occurred in 2019, awareness of food self-sufficiency in each country is increasing. As a result, the importance of agriculture as a primary industrial field is increasing. In Korea, efforts are being made to expand financial support in the agricultural sector to ensure a stable food supply, but the actual resettlement rate and the number of farm households are still low.

Moreover, since Korea belongs to a climate with four distinct seasons and consists of three sides of the sea, climate change is large and climate prediction is difficult. In this situation, if it is possible to predict the temperature in the future for the period of agricultural activities, it can be used as auxiliary data for efficient agricultural activities. Therefore, this paper proposes an artificial intelligence-based temperature prediction

Manuscript Received: October. 14, 2021 / Revised: October. 19, 2021 / Accepted: October. 22, 2021
Corresponding Author: expersin@hansei.ac.kr
Tel: +82-31-450-9585, Fax: +82-31-450-5172
Associate professor, Department of ICT Convergence, Hansei University, Korea

algorithm that can predict future temperature information by collecting and analyzing temperature data of farms in Gyeonggi Province in Korea over the past 10 years.

2. Research Scope

According to data from the National Statistical Office in 2019, 169 farms were confirmed nationwide, of which 31 farms were distributed in Gyeonggi Province. In this paper, agricultural land is used as the Gyeonggi-do area for data collection. And the standard number of farms are the more than 6,000 farmhouse. In this case, the target agricultural land is Pyeongtaek-si, Paju-si, Icheon-si, Anseong-si, Hwaseong-si, Yeoju-si, and Yangpyeong-gun. Paju uses Paju Geumchon data according to the data classification criteria provided.

Table 1 shows the Gyeonggi region where more than 6,000 farm households are distributed as of 2019.

Table 1. Gyeonggi area with more than 6,000 number of farms(2019)

Areas	The number of farms
Pyeongtaek-si	8,437
Paju-si	6,790
Icheon-si	6,976
Anseong-si	8,518
Hwaseong-si	11,726
Yeoju-si	6,676
Yanngpyeong-gun	6,483

For temperature data analysis, the temperature data of the target farmland for the last 10 years (2010-2019) provided by the Korea Meteorological Administration is collected and analyzed. The temperature data is based on data from March to October, which is the period of sowing and harvesting crops. Table 2 shows the average temperature data for the last 10 years in the Pyeongtaek area.

Table 2. Average temperature of pyeongtaek-si (2010~2019)

Date	Average Temperature	Date	Average Temperature
Mar-10	5	Mar-15	5.7
Apr-10	10	Apr-15	12.8
May-10	17.9	May-15	18.3
Jun-10	23.8	Jun-15	22.7
Jul-10	26.8	Jul-15	25.2
Aug-10	27.8	Aug-15	25.7
Sep-10	22.6	Sep-15	21.6
Oct-10	14.7	Oct-15	15.2
Mar-11	4.2	Mar-16	6.6
Apr-11	10.8	Apr-16	13.9
May-11	18.6	May-16	19.1
Jun-11	22.7	Jun-16	23.2
Jul-11	26	Jul-16	25.8

Aug-11	26.5	Aug-16	27.3
Sep-11	21.8	Sep-16	22.6
Oct-11	14	Oct-16	15.6
Mar-12	4.9	Mar-17	5.7
Apr-12	12	Apr-17	13.5
May-12	19.2	May-17	18.9
Jun-12	23.7	Jun-17	22.5
Jul-12	25.8	Jul-17	26.7
Aug-12	27	Aug-17	25.4
Sep-12	20.3	Sep-17	21
Oct-12	14.2	Oct-17	15.3
Mar-13	4.8	Mar-18	7.9
Apr-13	9.7	Apr-18	12.7
May-13	17.7	May-18	18
Jun-13	23.4	Jun-18	22.5
Jul-13	26	Jul-18	27.2
Aug-13	27.1	Aug-18	28.2
Sep-13	21.1	Sep-18	20.8
Oct-13	15.3	Oct-18	12.5
Mar-14	7.3	Mar-19	6.5
Apr-14	13.2	Apr-19	11.8
May-14	18.3	May-19	18.6
Jun-14	22.8	Jun-19	21.9
Jul-14	25.4	Jul-19	25.4
Aug-14	24.2	Aug-19	26.4
Sep-14	21.3	Sep-19	21.9
Oct-14	14.7	Oct-19	15.8

3. Data Analysis and Verification Method

The method used for data analysis in this paper is the time series analysis method. The time series analysis method is a data analysis method that predicts future data by analyzing it based on a series of past data.

Among various algorithms for time series analysis, this paper analyzes data using the Prophet Library. Prophet Library is an open source library for time series analysis developed by Facebook and predicts future data by analyzing seasonality and trend of data. Prophet Library can implement data through Python programming.

For data validation, the cosine similarity is checked by comparing the predicted data with the actual data for 2020. The cosine similarity is an algorithm that allows you to check the degree of similarity between two groups of data. As the cosine result value is closer to 1, it can be said that the data of the two groups have a strong similarity [1]. Equation (1) shows the formula for cosine similarity.

$$\cos(\theta) = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (A_i)^2}} \quad (1)$$

A_i and B_i represent vectors to be compared, respectively.

4. Implementation and Verification of the Algorithm

The algorithm proposed in this paper predicts future data by using a predictive model learned based on the collected data. The prediction algorithm is implemented in the following development environment. The OS is Window64bit, the development tool is Pycharm, and the development language is Pythoon 3.7. In addition, Prophet Library is used for data analysis.

Figure 1 shows a part of the temperature prediction algorithm code proposed in this paper, and Figure 2 shows a graph of the predicted temperature in the Pyeongtaek area. The trend line on the right in Figure 2 shows the predicted temperature for 2020.

```
model=Prophet()
model.fit(data)

last_1year=list()
for i in range(3, 11):
    last_1year.append(['2020-%02d'%i])
    print(last_1year)
last_1year=pd.DataFrame(last_1year, columns=['ds'])
last_1year['ds']=pd.to_datetime(last_1year['ds'])
print(last_1year)

forecast = model.predict(last_1year)
print("forecast")
print(forecast)
model.plot(forecast)
plt.xlabel('Year')
plt.ylabel('temperature')
plt.show()
```

Figure 1. Temperature prediction algorithm code

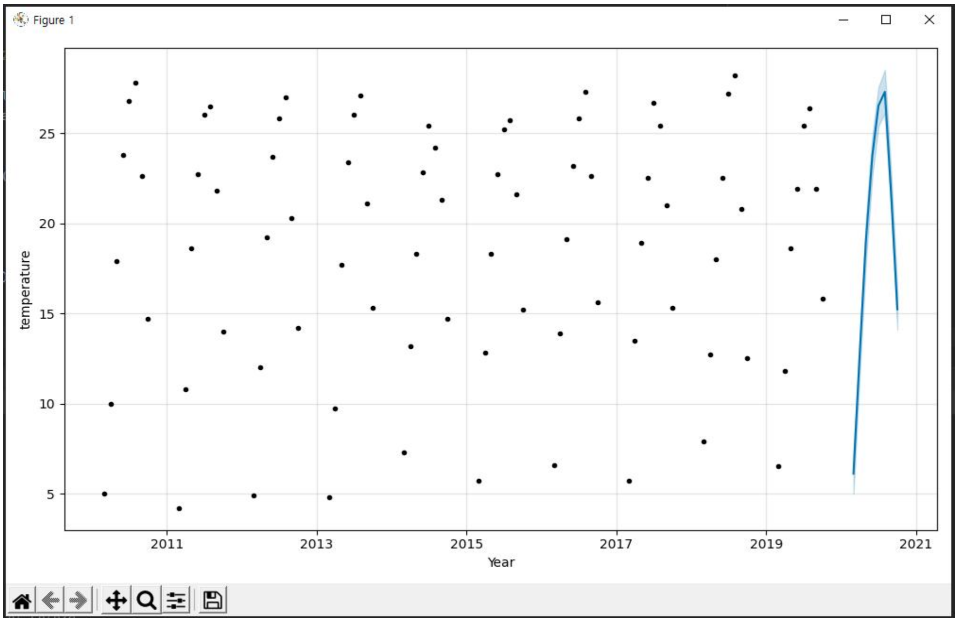

Figure 2. Temperature forecast graph of pyeongtaek-si

Table 3 shows the temperature data predicted through the proposed algorithm in the Pyeongtaek area and the actual observed temperature data in 2020.

Table 3. Predicted temperature data and actual temperature data (Pyeongtaek-si)

	Predicted Temperature	Actual Temperature
Mar-2020	6.1	7.2
Apr-2020	12.8	10.7
May-2020	19.1	17.8
Jun-2020	23.7	23.1
Jul-2020	26.5	24.7
Aug-2020	27.2	26.6
Sep-2020	21.5	20.7
Oct-2020	15.2	13.6

The cosine similarity was checked for the predicted temperature and the actual temperature in the Pyeongtaek area, and a similarity of 0.99 was confirmed.

In addition, the cosine similarity was checked by comparing the predicted data with the actual data for six agricultural fields to be analyzed, and the following similarity can be confirmed. Table 4 shows the cosine similarity for the temperature of target farmhouse.

Table 4. Temperature data cosine similarity by region

Region	Cosine Similarity
Pyeongtaek-si	0.99
Paju-si	0.99
Icheon-si	0.99
Anseong-si	0.99
Hwaseong-si	0.99
Yeoju-si	0.99
Yanngpyeong-gun	0.99

5. Conclusions

The algorithm proposed in this paper is an algorithm that predicts future temperature data by collecting and analyzing temperature data for agricultural land in Gyeonggi Province over the past 10 years (2010-2019). Through this algorithm, the similarity between the predicted data and the actual data was confirmed, and the cosine similarity of 0.99 figures was confirmed in all target regions.

This paper implemented an analysis model for some farms in Gyeonggi Province, but if this algorithm is used, it will be possible to expand and utilize the region as auxiliary data for future temperature analysis as an artificial intelligence analysis model. In addition, it is expected that the similarity of predicted data may also increase as the accumulated data increases every year with the raw data collection cycle as one year. Also, further research to expand the scope of the agricultural land area to be investigated in the future should be conducted.

References

[1] C. W. Yoo, "Estimation of Cosine Similarity Change in Singular Value Decomposition," The Journal of the Institute of Korean Institute of Information Technology(KIIT), Vol. 18, No. 3, pp. 31-37, Mar 2020.
DOI: http://dx.doi.org/10.14801/jkiit.2020.18.3.31

[2] W. S. Lee, K. J. Son, D. H. Jun, and Y. T. Shin, "Big Data Activation Plan for Digital Transformation of Agriculture and Rural," The Journal of The Korea Information Processing Society(KIPS), Vol.9, No.8, pp.235-242, May 2020
DOI: https://doi.org/10.3745/KTSDE.2020.9.8.235

[3] D. M. Seo, "Development of Artificial Intelligence Model for Predicting Citrus Sugar Content based on Meteorological Data," The Journal of The Korea Contents Association, Vol.21, No.6, pp.35-43, Jun 2021
DOI: https://doi.org/10.5392/JKCA.2021.21.06.035

[4] S. W. Gil and K. Y. Lee, "A Behavior-based Authentication Using the Measuring Cosine Similarity," The Journal of The Institute of Internet, Broadcasting and Communication, Vol.20, No.4, pp.17-22, Aug 2020
DOI: https://doi.org/10.7236/JIIBC.2020.20.4.17

[5] E. J. Jang and S. J. Shin, "Proposal of New Data Processing Function to Improve the Security of Self-driving Cars System," The Journal of The Institute of Internet, Broadcasting and Communication, Vol.20, No.4, pp.81-86, Aug 2020
DOI: http://doi.org/10.7236/JIIBC.2020.20.4.81

[6] M. S. Kim, S. H. Jung, J. G. Kim, H. S. Lee, and S. S. Kim "A Study on Solar Radiation Forecasting Based on Long Short-term Memory Considering Hourly Weather Changes," The Journal of Korean Institute of Intelligent System, Vol.31, No.1, pp.088-094, Feb 2021
DOI: http://dx.doi.org/10.5391/JKIIS.2021.31.1.088