Journal of information and
communication convergence engineering
http://jicce.org

J. Inf. Commun. Converg. Eng. 17(1): 8-13, Mar. 2019

Regular paper

Predicting Crop Production for Agricultural Consultation Service

Soong-Hee Lee* **and Jae-Yong Bae,** ***Member, KIICE***

Department of Information and Communications Engineering, Inje University, Gimhae 50834, Korea

Abstract

Smart Farming has been regarded as an important application in information and communications technology (ICT) fields. Selecting crops for cultivation at the pre-production stage is critical for agricultural producers' final profits because over-production and under-production may result in uncountable losses, and it is necessary to predict crop production to prevent these losses. The ITU-T Recommendation for Smart Farming (Y.4450/Y.2238) defines plan/production consultation service at the pre-production stage; this type of service must trace crop production in a predictive way. Several research papers present that machine learning technology can be applied to predict crop production after related data are learned, but these technologies have little to do with standardized ICT services. This paper clarifies the relationship between agricultural consultation services and predicting crop production. A prediction scheme is proposed, and the results confirm the usability and superiority of machine learning for predicting crop production.

Index Terms: Agricultural Consultation Service, Machine Learning, Prediction of Crop Production

I. INTRODUCTION

Smart farming is regarded as an important information and communications technology (ICT) practical application. ITU-T SG13 has developed a standard document, the Recommendation for Smart Farming (Y.4450/Y.2238) [1]. Service models are in development for pre-production stage work, which is critical for agricultural producers' profits given that over-production and under-production from poor crop choices can result in uncountable losses [2]. To prevent these losses, predicting crop production is required. However, human predictions are not effective with increasing amounts of agricultural data. Instead, machine learning has been raised as a promising option for this goal.

There have been numerous studies on machine learning-based applications in the agricultural field [3-9]. Mishra et al. introduced and compared a number of machine learning mechanisms for agricultural crop production: artificial neural network, information fuzzy network, decision tree, regression analysis, clustering, Bayesian belief network, time series analysis, and Markov chain model [3]. Priya et al. proposed the random forest algorithm for predicting crop yields based on existing data [4], and Rajasekaran et al. proposed an algorithm called ZeroR for predictive analysis [5]. Ghadge et al. proposed a system to predict crop yield based on BPN machine learning [6], Manjula et al. proposed a data mining model for predicting crop yield [7], and Balakrishnan et al. proposed a crop production-ensemble machine learning model for prediction [8]. Sanchez et al. also showed the predictive ability of machine learning methods for massive crop yield prediction [9].

These previous studies show that machine learning technologies can be used for predicting crop production, which is essential for agricultural producers during pre-production.

Received 03 October 2018, **Revised** 08 January 2019, **Accepted** 08 January 2019
***Corresponding Author** Soong-Hee Lee (E-mail: icshlee@inje.ac.kr, Tel: +82-55-320-3557)
Department of Information and Communications Engineering, Inje University, Gimhae 50834, Korea.

Open Access https://doi.org/10.6109/jicce.2019.17.1.8 print ISSN: 2234-8255 online ISSN: 2234-8883

However, there are few studies on applying machine learning technologies to predict crop production as a part of standardized consultation service at the pre-production stage.

In this paper, we first present a basic framework for an agricultural consultation service for agricultural producers at the pre-production stage. We also introduce a scheme for applying machine learning technology to predict crop production based on simple climate data. We then present our results and discuss.

II. FRAMEWORK OF AGRICULTURAL CONSULTATION SERVICE

Smart farming plays a large role in consultation for agricultural producers because it affects their financial profit. In this chapter we discuss issues around smart farming as part of agricultural consultation service.

A. Consultation on Smart Farming

1) Importance of Consultation in Smart Farming

People who make the urban-to-rural transition often suffer from unexpected events and financial difficulties because they lack experience; they need help from experienced experts. Separately, in agriculture, pre-production is critical for producers' final profits because over- and under-production can result in uncountable losses. Predicting crop production is necessary to prevent these losses, and this prediction can help with selecting crops during pre-production. The ITU-T presented a reference architecture for this consultation service, Y.4450/Y.2238.

2) Composition of agricultural consultation service

The ITU-T's reference architecture of an agricultural consultation service is shown in Fig. 1 [2]. The Environment Monitoring function gathers the measured data for temperature, rainfall, humidity, pH, etc.; the measured data may include cultivation resource information such as numbers or amounts of available agricultural machines, labor force, etc. The Data Accumulation function gathers all related information such as past cultivation records and final profits per crop, and the Knowhow Base Management function gathers expertise information from experts and skilled user communities. All of this gathered information is transferred to the Data Analysis function, which analyzes the information and produces meaningful results that will help the consultation process. The analyzed results are transferred to the Plan Consulting function, which interacts with the service users to help them make decisions.

The Data Analysis function, the main part of the agricultural consultation service, could include a scheme for predicting crop production based on the agricultural data from other functions and the Agricultural Information Repository. This scheme demands an illative method for more accurate prediction. Machine learning technology is adopted to predict crop production in this paper, building on previous studies that revealed its usefulness in the agricultural field.

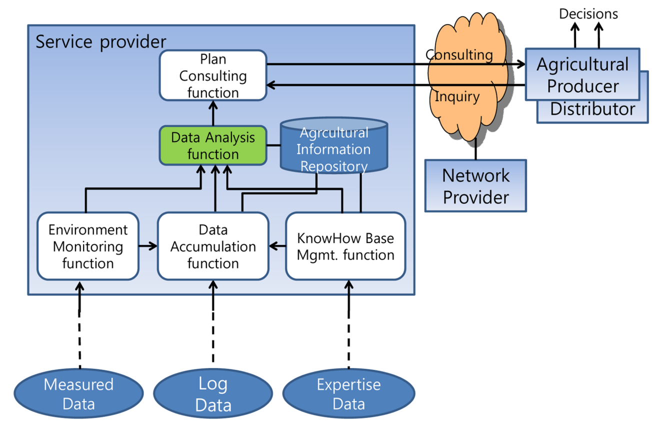


Fig. 1. A reference architecture for an agricultural consultation service (ITU-T Y.smpp [2]).

B. Predicting Crop Production in Agricultural Consultation Service

1) Selecting Crops for Seeding

With predicted crop production before actual cultivation, agricultural producers can select crops for seeding that will prevent over-production or under-production. In addition, the agricultural producers can reduce financial losses by selectively deciding the cultivation area per each crop. Hence, the agricultural consultation service must provide the predicted crop production to service users who will select crops and decide the cultivation area per each crop.

2) Predicting Agricultural Producers' Profits

The final profit of agricultural producers depends on the market conditions and crop production at harvest time. The crop production at harvest time needs to be predicted and merged with the market data from the financial experts to be applied to predict the final profit of agricultural producers.

3) Predicting Crop Production with an Agricultural Consultation Service

Crop production can be predicted from environment-related data such as on climate or soil and expertise data such as market data and past and current crop production, as shown in Fig. 2. The predicted crop production is applied to the financial speculation for predicting the agricultural producers' profits. Predicting crop production accurately and efficiently will make an agricultural consultation service more useful to service users. Therefore, a scheme for pre-

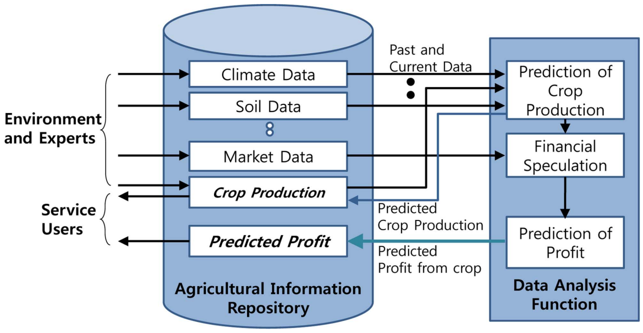


Fig. 2. Predicting crop production with an agricultural consultation service.

dicting crop production from climate data is proposed in the next chapter.

III. PROPOSED PREDICTION SCHEME

A. Input Data

To predict agricultural production, it is first necessary to acquire agricultural production data. For convenience in implementation, rice is chosen in this paper as the crop for predicting production. In general, it is perceived that rice yield depends greatly on rainfall and temperature because it originated from a subtropical region. This means that it is necessary to acquire large amounts of data about rainfall and temperature throughout the country for the prediction. Fortunately, the Korean government organization responsible for statistics, KOSTAT, has released a wide variety of statistical data on the KOSIS website [10], including rainfall and temperature. Total yearly rainfall and average yearly temperature per district in Korea were gathered for predicting rice production, as shown in Table 1.

B. Regression Model

For reasonable prediction results with machine learning, it is necessary to apply an appropriate regression model. With all of the currently available regression methods to choose from[11]~[13], it is difficult to decide which model to apply for prediction. The simple linear regression model is not appropriate for the prediction scheme in this paper because rice production is not directly proportional to rainfall or temperature data.

Palmer indicated that polynomial regression, a special case of linear regression, is useful because even if polynomials do not represent the true model, they take a variety of forms, and they may be close enough for a variety of purposes including the prediction in this paper [14]. Therefore, the following model, with this polynomial regression approach, is applied:

Table 1. Input data for machine learning for the proposed scheme (from KOSIS website)

District	Parameter	'97	'98	'99	…	'16	'17
	Product (kg/10a)	431	404	428	…	513	484
Seoul	Rainfall (mm)	1,210.2	2,349.1	1,733.1	…	991.7	1,233.2
	Temp. (°C)	12.9	13.8	13.2	…	13.6	13
	Product (kg/10a)	491	417	463	…	529	520
Busan	Rainfall (mm)	1,598.1	2,028.8	2,396.7	…	1,760.2	1,014.4
	Temp. (°C)	15.2	15.8	15	…	15.7	15.2
:	:	:	:	:	…	:	:
	Product (kg/10a)	444	430	418	…	419	396
Jeju	Rainfall (mm)	1,287.7	1,836.5	2,748.3	…	1,810.5	1,053.7
	Temp. (°C)	16.6	17.3	16.3	…	17	16.9

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 \quad (1)$$

where x_1 and x_2 denote climate parameters (i.e., yearly rainfall, average temperature), β_k denotes the coefficient for each term, and y denotes the amount of rice produced. With iterative learning, the coefficients β_k will be updated until the end of the learning. After the learning process ends, the predicted amount of rice production is attained from the final value of y with the updated coefficient β_k.

C. Machine Learning Process

For predicting with the above regression model, we need a method of optimizing the coefficients of the polynomial, comparing the result with the expected value. Machine learning is one of these methods. Hence, TensorFlow is adopted as the machine learning tool because it showed sufficient reliability and efficiency [15].

The machine learning process is initiated with *tf.Variable* (*tf.random_uniform*( )) and *tf.placeholder*( ) functions. After applying all input data with *tf.summary.merge_all*( ) function, cost, i.e., the distance between the result and target value, is obtained with *tf.reduce_mean*(*tf.square*(*Result-Y*)). The learning process is iterated until the desired iteration numbers, and the polynomial with the updated coefficients is used for the prediction from the current data, as shown in the flow chart of the proposed prediction scheme in Fig. 3.

For the proposed prediction, actual data of total rainfall, average temperature, and rice production from 1997 to 2016 in Table 1 are applied as input data for the machine learning; Fig. 4 shows the intermediate results during the process. Red dots denote actual rice production data applied to the learning process, while the curved plane denotes the prediction according to the applied regression polynomial with updated coefficients in the figure. One thousand iterations of learning produces immature results, as shown in Fig. 4(a). However, 10,000 iterations of learning produces a closer prediction plane, as shown in Fig. 4(b). Finally, 100,000 iterations of learning produces a prediction plane tightly aligned with actual data, as shown in Fig. 4(c). Therefore, the predicted rice productions after 100,000 iterations of machine learning from input data from 1997 to 2016 are finally selected for predicting the rice production in 2017. The final prediction from the proposed scheme will be compared with actual 2017 rice production.

D. Results

After 100,000 iterations of learning for the collected data shown in Table 1, predictions are obtained with the coefficients of the regression polynomial in the output screen of the program regarding the proposed scheme, as shown in Fig. 5.

The rice productions for other districts are also predicted using the already obtained coefficients of the regression polynomial of (1) (β_0 = 1.17109466, β_1 = 0.2731905, β_2 = -0.03467211, β_3 = -0.7962535, β_4 = 0.63746864, β_5 = -0.68467444). The predicted rice production is compared with actual rice production as shown in Fig. 6.

The results in Fig. 6 show predictions that were very close to actual rice production with only a large gap in Jeju District.

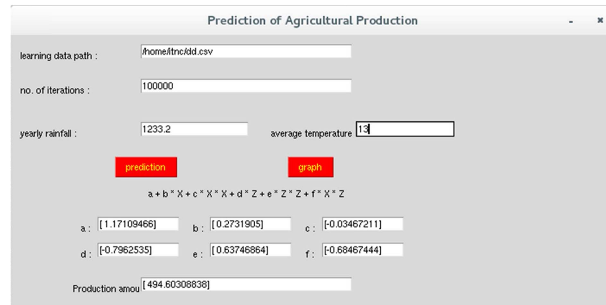


Fig. 5. Predicted rice production in Seoul at 2017.

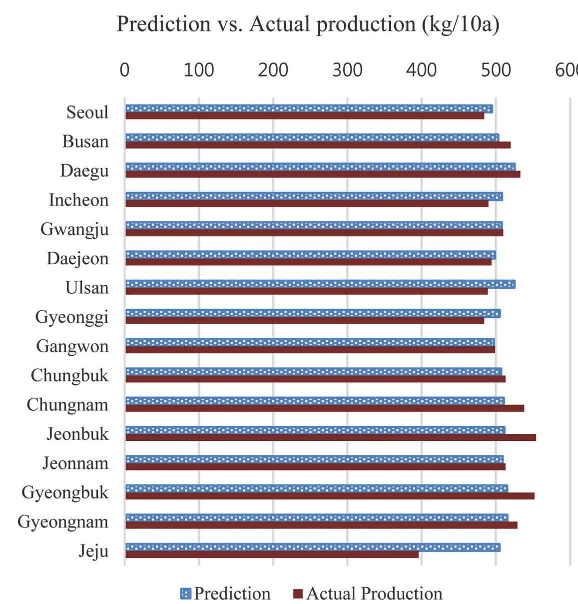


Fig. 6. Comparison between predicted rice production and actual rice production in 2017 per each district.

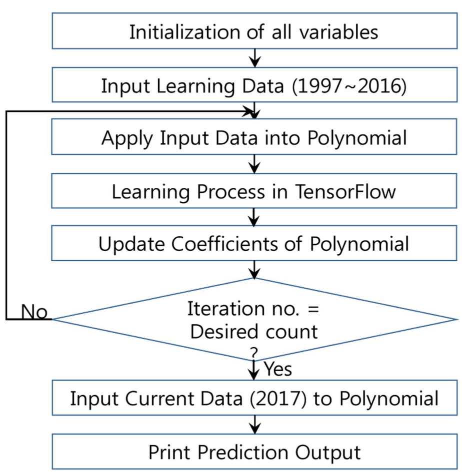
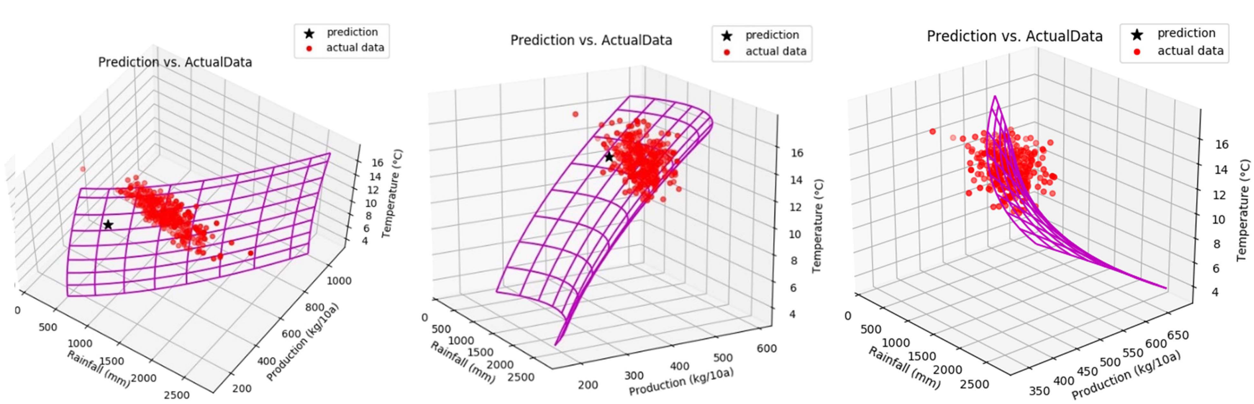


Fig. 3. Flow chart of the proposed prediction scheme.



(a) After 1,000 iterations (b) After 10,000 iterations (c) After 100,000 iterations

Fig. 4. Predicted results after 1000, 10000, and 100000 iterations of learning with input data from 1997 to 2016.

Table 2. Accuracy of predictions for rice production in each district

District	Actual rice production (kg/10a)	Predicted rice production (kg/10a)	Accuracy (%)
Seoul	484	494.6031	97.8093
Busan	520	502.9397	96.7192
Daegu	533	525.0275	98.5042
Incheon	490	507.9302	96.3408
Gwangju	510	507.8251	99.5735
Daejeon	494	498.4509	99.0990
Ulsan	489	524.7574	92.6877
Gyeonggi	484	505.0581	95.6492
Gangwon	499	497.0408	99.6074
Chungbuk	513	506.8620	98.8035
Chungnam	538	510.3421	94.8591
Jeonbuk	554	511.2329	92.2803
Jeonnam	513	508.8654	99.1940
Gyeongbuk	552	514.9648	93.2907
Gyeongnam	529	515.1384	97.3797
Jeju	396	504.8537	72.5117

For a better comparison, the accuracy of the proposed prediction is evaluated with the formula Accuracy (%) = (1 - abs(prediction-actualvalue)) × 100. The results including the evaluated accuracy are shown in Table 2. The table shows accuracy of 92.2803~99.6074%, with an average of 96.7865 excluding the outlier case of 72.5117% on Jeju Island. This case can be explained by the geological feature of volcanic rock, which is porous and holds little water after a rain. The average accuracy of the proposed prediction scheme is 95.2693% including the result for Jeju.

IV. DISCUSSION AND CONCLUSIONS

We first established a relationship between an agricultural consultation service and predicting crop production. A scheme for this prediction applying machine learning was also proposed; the proposed scheme predicted the rice production in 2017 after machine learning with input data from 1997 to 2016. We compared the predicted results with actual rice production in 2017, and with average accuracy over 95%, the results show that an agricultural consultation service can use the proposed prediction scheme at the pre-production stage by providing agricultural producers with predicted crop production. In addition, it was proved that the total yearly rainfall and average yearly temperature for the machine learning were highly effective for the prediction. This means that these two factors play significant roles in rice production.

More works must be carried out to apply the proposed scheme to other cases, including exceptional cases such as that of Jeju. This paper just showed the potential of the machine-learning-based prediction scheme for an agricultural consultation service for the standardization. In a future study, various types of yearly data (soil condition, cultivation methods, etc.) will be added to the input data for the machine learning. The proposed scheme including regression model and input data should be considered for predicting yields of other crops. Finally, the proposed prediction scheme will be applied toward standardizing the service model for an agricultural consultation service.

REFERENCES

[1] ITU-T SG13 Recommendation ITU-T Y.2238, Overview of Smart Farming based on networks, pp. 1-24, 2015. Available: https://www.itu.int/rec/T-REC-Y.2238/en.

[2] ITU-T SG13 TD, Draft recommendation ITU-T Y.smpp: Service model for the pre-production stage on Smart Farming, pp. 1-13, 2018. Available: https://www.itu.int/md/T17-SG13-180418-TD-WP3-0107.

[3] S. Mishra, D. Mishra, and G. H. Santra, "Applications of machine learning techniques in agricultural crop production: a review paper," *Indian Journal of Science and Technology*, vol. 9, no. 38, pp. 1-14, 2016. DOI: 10.17485/ijst/2016/v9i38/95032.

[4] P. Priya, U. Muthaiah, and M. Balamurugan, "Predicting yield of the crop using machine learning algorithm," *International Journal of Engineering Sciences & Research Technology*, vol. 7, no. 1, pp. 1-7, 2018. DOI: 10.5281/zenodo.1212821.

[5] T. Rajasekaran, P. Jayasheelan, and K. R. Sri Preethaa, "Predictive analysis in agriculture to improve the crop productivity using ZeroR algorithm," *International Journal of Computer Science and Engineering Communications*, vol. 4, no. 2, pp. 1397-1401, 2016.

[6] R. Ghadge, J. Kulkarni, P. More, S. Nene, and Priya, "Prediction of crop yield using machine learning," *International Research Journal of Engineering and Technology*, vo. 5, no. 2, pp. 2237-2239, 2018.

[7] E. Manjula and S. Djodiltachoumy, "A model for prediction of crop yield," *International Journal of Computational Intelligence and Informatics*, vo. 6, no. 4, pp. 298-305, 2017.

[8] N. Balakrishinan and G. Muthykumarasamy, "Crop production-ensemble machine learning model for prediction," *International Journal of Computer Science and Software Engineering*, vo. 5, no. 7, pp. 148-153, 2016.

[9] A. G. Sanchez and J. F. Solis, "Predictive ability of machine learning methods for massive crop yield prediction," *Spanish Journal of Agricultural Research*, vol. 12, no. 2, pp. 314-328, 2014. DOI: 10.5424/sjar/2014122-4439.

[10] Korean Statistical Information Service, Available: http://kosis.kr/.

[11] S. Lee, J. Lee, and M. Hwang, "A study of the transformation of the dependent variable in non-linear regression," pp. 1-22, 2012. Available: http://kostat.go.kr/ attach/journal/12-2-3.pdf.

[12] A. Ruckstuhl, "Introduction to nonlinear regression learning," pp. 1-29, 2010. Available: https://stat.ethz.ch/~stahel/courses/cheming/nlreg10E.pdf.

[13] P. Dalgaard, "Nonlinear regression analysis," pp. 1-75, 2008. Available: http://staff.pubhealth.ku.dk/~pd/V+R/slides/nonlin.pdf.

[14] M. Palmer, "Multiple regression," 2010. Available: http://ordination.okstate.edu/MULTIPLE.htm.

[15] An Open Source Machine Learning Library for Research and Production, Available: https://www.tensorflow.org/.

Soong-Hee Lee
received the B.S., M.S., and Ph.D. degrees from Kyungpook National University, Korea in 1987, 1990 and 1995, respectively. From 1987 to 1997, he was a member of research staff in Electronics and Telecommunications Research Institute. Since 1997, he has been with the Dep't of Information and Communications Engineering, Inje University. His research activities are in the area of next generation network technologies and services. He was designated as an IT Standard Expert on behalf of Korea by Ministry of Information and Communications in 2001. He is an Editor in the ITU-T SG13 and a Member of KIICE.

Jae-Yong Bae
will receive the B.S. degree from Inje University, Korea in 2019. Since 2013, he has been with the Dep't of Information and Communications Engineering, Inje University. His research activities are in the area of machine learning technologies and services. He is a Member of KIICE.