International Journal of Advanced Smart Convergence Vol.10 No.4 98-103 (2021)
http://dx.doi.org/10.7236/IJASC.2021.10.4.98

IJASC 21-4-12

Proposal of An Artificial Intelligence Farm Income Prediction Algorithm based on Time Series Analysis

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

Recently, as the need for food resources has increased both domestically and internationally, support for the agricultural sector for stable food supply and demand is expanding in Korea. However, according to recent media articles, the biggest problem in rural communities is the unstable profit structure. In addition, in order to confirm the profit structure, profit forecast data must be clearly prepared, but there is a lack of auxiliary data for farmers or future returnees to predict farm income. Therefore, in this paper we analyzed data over the past 15 years through time series analysis and proposes an artificial intelligence farm income prediction algorithm that can predict farm household income in the future. If the proposed algorithm is used, it is expected that it can be used as auxiliary data to predict farm profits.

Keywords: *Time Series Analysis, Predicting Income Algorithm, Data Analysis, Big Data, Artificial Intelligence*

1. Introduction

Recently, as internal and external interest in the future food shortage has increased, support projects for stable food supply and demand are expanding in Korea. However, according to a recent media article, it was confirmed that the biggest problem faced by farmers in the current rural society is the unstable profit structure. In addition, there are a lot of difficulties in planning the future profit structure because there is currently a lack of auxiliary data that can predict the profit structure in Korea.

Therefore, this paper analyzes data over the past 15 years through time series analysis and proposes an artificial intelligence farm income prediction algorithm that can predict farm household income in the future. If the proposed algorithm is used, it is expected that it can be used as auxiliary data to predict farm profits.

2. Time Series Analysis Method

The time series data analysis method is a method of predicting future data by grasping the characteristics of data existing at regular time intervals [1]. There are various algorithms for time series analysis, but in this

Manuscript Received: October. 12, 2021 / Revised: October. 15, 2021 / Accepted: October. 19, 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

paper, time series analysis is conducted through the least squares method.

The least squares method refers to a method of minimizing the sum of square errors and effectively selecting coefficients. That is, it is a method of obtaining β capable of minimizing the value of Equation 1.

Equation (1) shows the least squares method arithmetic expression.

$$\sum_i (y_i - f(x_i, \beta))^2 \tag{1}$$

x_i and y_i each represent a variable, β means a given constant, and $f(x_i, \beta)$ represents a linear linear function.

3. Analysis Data

In order to implement the forecasting system proposed in this paper, we use the average farm income data from 2003 to 2020 provided by the National Statistical Office.

This system, which operates based on artificial intelligence, aims to learn new data every year.

Table 1 shows the average farm income used for data analysis.

Table 1. Average farm income (2003~2019)

Year	Average Income(unit: 1,000 won)
2003	26,878
2004	29,001
2005	30,503
2006	32,303
2007	31,967
2008	30,523
2009	30,814
2010	32,121
2011	30,148
2012	31,031
2013	34,524
2014	34,950
2015	37,215
2016	37,197
2017	38,239
2018	42,066
2019	41,182
2020	45,029

4. Implement of AI Farm Income Prediction Algorithm

The prediction system proposed in this paper is an algorithm that creates a future prediction model using past data (2003~2017), and checks the prediction accuracy by comparing the results confirmed through the prediction model with the actual data (2018~2020).

The prediction algorithm development environment is as follows. The OS is Windows 64bit, the language used is Python 3.7, and the development tool used is Pycharm.

Figure 1 shows a part of the revenue prediction algorithm code.

```
predict_vals=[]
upper_limit=[]
lower_limit=[]
for l in future:
    predict_val=b0+b1*(n+l)
    s2=np.sum(np.square(y-b0-b1*time))/(n-2)
    x=np.array([[1],[n+l]])
    variance_factor=np.sqrt(1+x.T.dot(X_tX_inv.dot(x)))
    limit=t_val*np.sqrt(s2)*variance_factor[0][0]
    predict_vals.append(predict_val)
    upper_limit.append(predict_val+limit)
    lower_limit.append(predict_val-limit)

start_data='2018-01-01'

pred_data={
    'pred':predict_vals,
    'ul':upper_limit,
    'll':lower_limit
}

pred_df=pd.DataFrame(pred_data)
pred_df.index=pd.date_range(start_data, periods=future[-1], freq='Y')

fitted_val=b0+b1*time
print(fitted_val)
vals=pd.concat([fitted_val, pred_df['pred']])
print(vals)

fig=plt.figure(figsize=(8,8))
fig.set_facecolor('white')

marker_config=dict()
```

Figure 1. Income prediction algorithm code

Figure 2 shows the revenue forecast graph. The dotted line in the upper right corner indicates the predicted data, and the shaded area indicates the area of the minimum and maximum values.

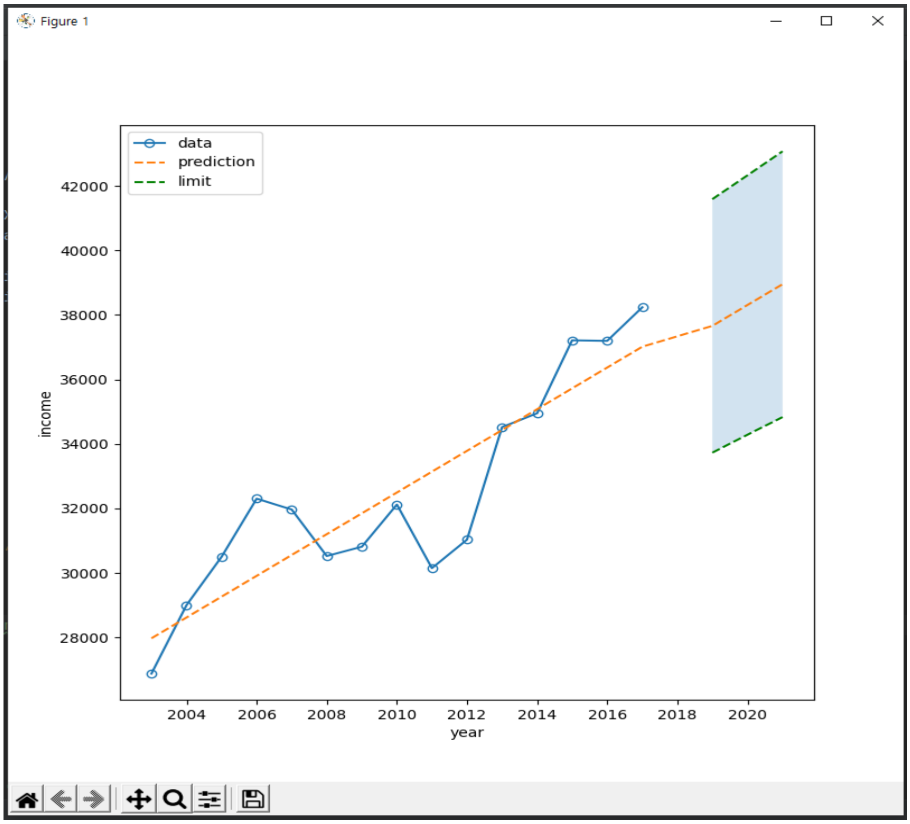

Figure 2. Income forecast graph

Table 3 shows the farmhouse revenue data and actual farmhouse revenue data for the last three years confirmed through the prediction algorithm.

Table 3. Predicted income data and actual income data (2018~2020)

Year	Predicted income (unit: 1,000 won)	Actual income (unit: 1,000 won)
2018	37,662	42,066
2019	38,309	41,182
2020	38,953	45,029

In order to confirm the prediction accuracy of the predicted data through the recommendation algorithm proposed in this paper and the actual data, the Pearson correlation coefficient of the data of the two groups was analyzed, and a similarity of 73% was confirmed.

Figure 3 shows a part of the Pearson correlation coefficient analysis code and the operation result.

```
data1=[37662,38309,38953] #예상수익
data2=[42066,41182,45029] #실제수익

def mean(inp):
    result=0
    len_inp=len(inp)
    for i in inp:
        result+=i
    result=result/len_inp
    return result

def sqrt(inp):
    result=inp/2
    for i in range(30):
        result=(result+(inp/result))/2
    return result

def cor(inp1, inp2):
    mean_inp1=mean(inp1)
    mean_inp2=mean(inp2)
    son=0
    for i in range(len(inp1)):
```

```
C:\Users\User\Desktop\CropProduction\venv\Scr
0.7343425612366594
```

Figure 3. Income forecast graph

When the prediction algorithm training data was set to 2003 to 2019 and the 2020 data was predicted, the result value of 40569 could be confirmed. This suggests that the greater the number of learned data, the higher the prediction accuracy.

Figure 4 shows the expected revenue data for 2020, which is confirmed when the learning period is set to 2003-2019.

```
2020-12-31     40569.294118
2021-12-31     41347.176471
2022-12-31     42125.058824
```

Figure 4. Income forecast data (2020)

5. Conclusions

In this paper, we proposed an artificial intelligence-based algorithm that can predict future farmhouse profits by using past farmhouse profit data. When learning using the range of data currently provided by the National Statistical Office, the data prediction accuracy in the last three years (2018-2020) was 73%, but it was confirmed that increasing the number of learning data sets significantly reduces the error range of predicted data. This suggests that the higher the cumulative number of data collected in the future, the higher the accuracy of the revenue forecast data.

In the current rural community, where unstable profit structure is a problem, if this algorithm is used, it is expected that more accurate farm income information can be predicted based on the accumulated income data.

In the future, additional research should be conducted to expand the scope of collected data and to subdivide the scope of data collection and classification.

References

[1] K. S. Woo, and J. H. Suh "Time Series Analysis of Park Use Behavior Utilizing Big Data -Targeting Olympic Park-," The Journal of the Korean Institute of Landscape Architecture(KILA), Vol. 46, No. 2, pp. 27-36, Apr 2018.
DOI: https://doi.org/10.9715/KILA.2018.46.2.027

[2] H. Choi, "Stock prediction analysis through artificial intelligence using big data," The Journal of The Korea Institute of Information and Communication Engineering, Vol.25, No.10, pp.1435-1440, Oct 2021
DOI: http://doi.org/10.6109/jkiice.2021.25.10.1435

[3] 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

[4] C. W. Hong and K. Hur "Prognosis of the Remaining Useful Life of a Turbofan Engine Using Deep Neural Network and Pearson Correlation Coefficient," The Journal of Korea Society for Naval Science & Technology, Vol.4, No.1, pp.078-083, Mar 2021
DOI: https://doi.org/10.31818/JKNST.2021.03.4.1.78

[5] J. M. Jeon and D. H. Jang "A Study on Factors Influencing Agricultural Income of Peach Farmers," The Journal of Industrial Economics and Business, Vol.33, No.4, pp.1201-1223, Aug 2020
DOI: http://dx.doi.org/10.22558/jieb.2020.8.33.4.1201

[6] 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

[7] H. K. Jeong, "Effects of Utilizing of Weather and Climate Information of Farmer's Income," The Journal of The Korean Society of Climate Change Research, Vol. 9, No. 3, pp. 283-291, Sep 2018.
DOI: https://doi.org/10.15531/KSCCR.2018.9.3.283

[8] 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