Design of Algorithm for Forecasting Agricultural Output Quality Based on Genetic Algorithm-Back Propagation Neural Network

Wang Yuehui\textsuperscript{1,a}, Wen Bin\textsuperscript{2,b}, Liao Xin\textsuperscript{3,c}, Chen Lan\textsuperscript{4,d}

\textsuperscript{1}College of Communication Engineering, Chengdu University of Information Technology, Chengdu China
\textsuperscript{2}College of Communication Engineering, Chengdu University of Information Technology, Chengdu China
\textsuperscript{3}College of Communication Engineering, Chengdu University of Information Technology, Chengdu China
\textsuperscript{4}College of Communication Engineering, Chengdu University of Information Technology, Chengdu China
\textsuperscript{a} wangyuehuii@163.com, \textsuperscript{b}573991760@qq.com, \textsuperscript{c}3190402020@stu.cuit.edu.cn, \textsuperscript{d}2641787848@qq.com

Abstract. Nowadays, agricultural technology is developing rapidly, and the concept of smart agriculture is gradually popularized. Economic development has made consumers pay more attention to the quality of agricultural output, so the prediction of the quality of agricultural output has also become important. This research takes the sugar content level of apples as an example of the quality of agricultural output, combined with the meteorological conditions of each crop period in the apple producing area as an influencing factor, and uses the back propagation neural network optimized by genetic algorithm to establish the quality of agricultural output Algorithm model. The experimental results show that the accuracy of GA-BP predicting apple sugar content grade, MAE and RMSE are 90\%, 0.1 and 0.316 respectively, which is more ideal than the traditional method.

1. Introduction

In recent years, with the continuous development of the agricultural level, agricultural planting technology and management methods have been gradually completed. The output of fruits and vegetables has been increasing, and people's living standards have improved. These making people's demand for fruits and vegetables has been changed. The quality and price of fruits and vegetables are the main concerns of consumers. Yield and quality are the main concerns of producers. All in all, the quality of fruits and vegetables becomes the key. Therefore, the quality prediction of agricultural output can better analyze whether different climatic environments are suitable for planting specific crops, and it can predict the changes in the quality of agricultural output caused by changes in the climate and environment each year due to meteorological disasters. At present, the commonly used methods of agricultural output quality evaluation include analytic hierarchy process, gray correlation analysis method, membership function analysis method, DTOPSIS method, entropy weight method, standard deviation weight method, etc.[1-7].
By analyzing the relationship between meteorological data and agricultural output quality indicators, it is found that the input-output data set for agricultural output quality prediction is suitable for the training and prediction of BP neural network, and at the same time, to solve the shortcomings of BP neural network, genetic algorithm is used for optimization. This research applies the machine learning method to the practical problem of agricultural output quality prediction, which reduces the difficulty of model building compared with the traditional method, and also achieves the desired results.

2. Materials and Methods

This research selects apples as an example of agricultural output for analysis and research. The apple quality data in this article comes from apple orchards in 20 villages and towns in Fujian Province, China from 2010 to 2020, and the meteorological data are taken from the daily data from the regional automatic stations of the towns and towns in the corresponding years.

2.1. Feature selection

Feature engineering is required before machine learning. Selecting good input and output features can improve the accuracy of model prediction.

After analyzing the purpose of the study, it is known that quality is the output feature in machine learning of this model, and the sugar content ratio of apples is used as the quality in this research [8-10]. Referring to the technical manual of agricultural meteorological services issued by the Meteorological Bureau of Linfen City, Shanxi Province, the apple sugar content detection values were classified into sweetness classes, and the final classification results were used as output features, and the specific classifications are shown in Table 1.

| Sugar content ratio | grade | Quality category |
|--------------------|-------|------------------|
| <11%               | 001   | Poor             |
| 11%-13%            | 010   | medium           |
| >13%               | 100   | Better           |

The selection of input features is also based on the content provided in the manual. The meteorological conditions that affect the sugar content of apples in different crop periods are summarized, and the meteorological conditions of 32 different crop periods shown in Table 2 are obtained as input features. A linear correlation analysis was performed between these 32 input features and the sugar content ratio, and |r| was greater than 0.3, so these 32 features could be used as input features.

| Crop period          | Weather condition                      |
|----------------------|----------------------------------------|
| Dormant period       | average temperature                    |
|                      | Minimum temperature                    |
|                      | Maximum temperature                    |
|                      | Average relative humidity              |
| Bud period           | average temperature                    |
|                      | Minimum temperature                    |
|                      | Maximum temperature                    |
|                      | Average relative humidity              |
| Flowering period     | average temperature                    |
|                      | Minimum temperature                    |
|                      | Maximum temperature                    |
|                      | Average relative humidity              |
|                      | Average daily temperature >10°C        |
|                      | accumulated temperature                |
| Young fruit stage    | maximum wind speed                     |
|                      | average temperature                    |
2.2. Back Propagation Neural Network
The machine learning model in this paper selects the most widely used back propagation neural network. Back propagation neural network is a common method for teaching artificial neural networks how to perform a given task, and it is a supervised learning method. The instructor adjusted the output to correspond to expectations. Back propagation neural network requires the activation function used by artificial neurons (or "nodes") to be differentiable. [11]

Propagation and weight correction are the core steps of back propagation neural network training. Propagation includes forward and reverse directions: the forward direction is to pass the input data through each layer and then the output; the reverse direction is to propagate the error between the final result obtained in the forward direction and the actual result back to each layer. Weight correction is a key step after backpropagation. The purpose is to adjust the weight so that each layer corrects the error.

The topological structure of a typical three-layer neural network [12-13] is shown in Figure 1:
2.3. Genetic Algorithm
This article uses genetic algorithm to optimize the back propagation neural network model. Genetic algorithm is an evolutionary algorithm derived from the biological world, and the core idea is as the same as the saying "survival of the fittest" in biology.

The initialization of genetic algorithm is to treat multiple individuals as a population. Firstly the algorithm performs coding(transforms the manifestation into a gene form), and then uses the survival of the fittest in biology and the principle of survival of the fittest to select an approximate solution. According to the fitness formula, the cross-combination mutation genetic operator in each generation. Finally it has a new solution group. Genetic algorithm is the use of biological ideas to obtain an approximate optimal solution to the problem.

2.4. Algorithm for Predicting the Quality of Agricultural Output Based on GA-BP Neural Network
It is well known that back propagation neural network is the most widely used neural network model due to its non-linear nature, high fault tolerance and autonomous training ability. But the problem is also obvious. It is easy to converge to a local minimum. The main idea of using genetic algorithm to optimize back propagation neural network in this research is to use genetic algorithm to obtain the optimal initial weight and threshold of each layer of back propagation neural network, which can eventually reduce the training cost greatly. When the genetic algorithm optimizes the initial weights and thresholds, it is necessary to adjust the individual selection rate through the fitness function. [14] The fitness function set in this article is as follows:

$$\text{fitness} = \frac{1}{\sum_{i=1}^{N} |y_i - x_i| + \Theta}$$ (1)

Among them, $y_i$ represents the quality level of agricultural output predicted by GA-BP neural network, and $x_i$ represents the real quality level of agricultural output, and N represents the total number of samples. The effect of $\Theta$ is to avoid the denominator being 0.

The main steps of the agricultural output quality prediction algorithm based on GA-BP neural network are:

1. Establish a database for predicting the quality of agricultural output based on historical meteorological data and historical output quality grade data.
2. According to the Agricultural Meteorological Service Technical Manual issued by the Meteorological Bureau, determine the meteorological parameters that affect the quality of the output. They are used as the input characteristics of the neural network. Then the input data will be preprocessed, and each layer structure of the neural network will be designed.
3. Select the first weight and threshold of back propagation neural network according to genetic algorithm.
4. Network training.
5. After get the optimal solution which is obtained through multiple training, the test samples are the input for prediction.
6. Compared with the output results of the test samples and the real results. Then evaluate the accuracy of the algorithm.

3. Results & Discussion
In order to verify the correctness of the agricultural output quality prediction algorithm based on GA-BP neural network proposed in this research, the meteorological data and apple sugar content data of apple orchards in 20 villages and towns in Fujian Province, China from 2010 to 2020 are selected for training and verification. Among them, 80% is selected as the training data, and 20% as the test data. The population size, crossover probability, mutation probability, number of evolutionary iterations, and learning rate of the genetic algorithm are set to 100, 0.3, 0.1, 500, 0.01 respectively. The neural network
structure is set to three layers: 32-6-3. A single back propagation neural network is selected to compare the results with the method in this paper.

3.1. Evaluation index
In this research, the Mean Absolute Error (MAE), Root-mean-square error (RMSE) and accuracy are used to evaluate the prediction accuracy of the model, and the formula is as follows:

\[ MAE = \frac{1}{n} \sum_{i=1}^{n} |d_{pre} - d_{real}| \]  
\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (d_{pre} - d_{real})^2} \]  
\[ \text{Accuracy} = \frac{\text{Total number of samples} - \text{Number of error samples}}{\text{Total number of samples}} \]

where \( d_{pre} \) is the predicted value, and \( d_{real} \) is the true value. The lower the MAE and RMSE, the smaller the prediction error, and the higher the accuracy, the better the prediction result.

3.2. Analysis of GA-BP Neural Network Prediction and Simulation Results
The machine learning prediction uses the training data set for network training firstly, then puts the test data into the network for testing, and finally compares and analyzes the results of the back propagation neural network agricultural output quality prediction and the GA-BP agricultural output quality prediction.

Figure 2 shows the comparison of back propagation neural network's agricultural output quality forecast categories and actual categories, and Figure 3 shows the comparison of GA-BP's agricultural output quality forecast categories and actual categories.

Figure 2  Comparison of back propagation neural network prediction results with real values
From Figures 2 and 3, the number of sample groups with incorrect predictions of the sugar content of apples based on GA-BP is less than that of a single back propagation neural network prediction of the sugar content of apples. Comparing the accuracy evaluation indexes before and after optimization, the MAE and RMSE of back propagation neural network are 0.2 and 0.447 respectively, and the MAE and RMSE of genetic algorithm optimization are 0.1 and 0.316 respectively. The accuracy of the compared test samples is shown in Table 3, which shows that the GA-BP neural network predicts the sugar content level of apples with an accuracy of 90%, which is 10% higher than before optimization. And the accuracy of each individual category was also improved. Therefore, it is concluded that GA-BP has more accurate results for the prediction of the sugar content of apples.

| Sample category | Back propagation neural network accuracy | GA-BP accuracy |
|----------------|----------------------------------------|----------------|
| Poor           | 0.750                                  | 0.917          |
| medium         | 0.875                                  | 0.937          |
| Better         | 0.750                                  | 0.833          |
| total          | 0.800                                  | 0.900          |

4. Conclusions
The result analysis shows that the prediction accuracy of apple sugar content based on GA-BP is higher than that of single back propagation neural network apple sugar content. GA-BP predicts apple sugar content grade accuracy rate, MAE and RMSE respectively 90%, 0.1 and 0.316. The prediction result is relatively ideal. The use of GA-BP neural network technology in the quality prediction of agricultural output can predict in advance the changes in the quality of agricultural output that may be caused by climate change or climate disasters, allowing farmers to make agricultural adjustments and preparations. This research applies the machine learning method to the practical problem of agricultural output quality prediction, which reduces the difficulty of model building compared with the traditional method, and also achieves the desired results. In the future, the prediction of the agricultural output’s quality can be applied to the crop’s fitness area prediction, and the result is more specific than a single fitness area prediction.

Acknowledgment
We would like to express our gratitude for financial support from Key Funded Project of Science and Technology Support Program of Sichuan Provincial Department of Science and Technology (2018NZ0051).
References
[1] Chen X, Yang R P, Zhao Y, et al. Comprehensive Evaluation on Commerciality of Tomato Fruits by Using AHP and TOPSIS[J]. Guizhou Agricultural Sciences, 2008.
[2] Xia X, Wang Q, Yin G, et al. Application of DTOPSIS Method on Comprehensive Evaluation of Tomato Combinations in Solar Greenhouse in Early Spring[J]. Journal of Changjiang Vegetables, 2012.
[3] Yan H O, Lu F K. APPLICATION OF DTOPSIS IN INTEGRATIVE EVALUATION OF TOMATO VARIETIES[J]. Journal of Inner Mongolia Agricultural University(Natural Science Edition), 2011.
[4] Chen X, Yang R P, Zhao Y, et al. Comprehensive Evaluation on Commerciality of Tomato Fruits by Using AHP and TOPSIS[J]. Guizhou Agricultural Sciences, 2008.
[5] Wang Z, Cong P, Zhou J, et al. Method for identification of external quality of wheat grain based on image processing and artificial neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2007.
[6] Pang S Q, Zhao Y, Ge L I. Comprehensive Evaluation on Processing Tomato Breeds with Gray Relationship Analysis[J]. Journal of Shihezi University(Natural Science), 2006.
[7] Lin W H. The artificial neural networks method for synthetic evaluation of rice quality[J]. Journal of Fujian Agricultural University(Natural Science), 2002.
[8] Zhao Y, Liu D, Ren Y, et al. Establishing an Index System of Physiochemical Quality Evaluation in Apples[J]. Northern Horticulture, 2018.
[9] Hong Y X, Ping X Y, Dong M X. Influence of climate change on apple growth and countermeasures in Tianshui City[J]. Agricultural Research in the Arid Areas, 2006.
[10] Zhang M, Zhang K, Jiang H, Zhou Z, et al. Temperature Affects the Quality and Quality Regionalization of Red Fuji Apple in East Gansu[J]. Chinese Journal of Eco-Agriculture, 2020.
[11] Pengbo L I, Yan X, Dongrui X U. Comparison of grain yield spatial distribution forecast between the models of BP neural network and multiple linear regression[J]. Journal of Arid Land Resources and Environment, 2014.
[12] Tian H, Zhang H, Feng W, et al. CG Lightning Prediction Method of Based on BP Neural Network and Atmospheric Electric Field Characteristics[J]. Insulators and Surge Arresters, 2018.
[13] Gu S, Tang C, Deng Z, et al. Lightning Trip Warning Based on GA-BP Neural Network Technology[C]// 2020 4th International Conference on HVDC (HVDC). 2020.
[14] Hao J, Ren H, Jing W. BP Neural Network Stock Price Forecast Based on Adaptive Genetic Algorithm Optimization[J]. Henanence, 2017.