International Journal of Business and Management Study - IJBMS
Volume 3 : Issue 1 [ISSN 2372-3955]
Publication Date : 18 April, 2016

Applying Backpropagation Neural Network to Predict the Trading Volume of Cherry Tomatoes in Taiwan

Srayut Tongnoy, Wen-Ming Hung, Deng-Neng Chen

Abstract— **The agricultural issue is always the critical factor to achieving the sustainability of a country. Due to the growing of plants is affected by weather, season, and lots of external influences, the harvest of crops is unstable and might cause the imbalance between supply and demand in the market. Therefore, the precise prediction of demand in the crops transaction market is important. It is helpful to the farmers to make the cultivating plan and also beneficial to the development of agriculture. In our research, we apply back propagation neutral network (BPNN) to develop a time-series prediction model. The model is used to predict the trading volume of cherry tomatoes in the fruit transaction market in Taiwan. The trading volume indicates the demand in the market, for that reason, the farmers and government can make cultivating plan effectively by the prediction results. We collected the transaction information of cherry tomatoes from 2011 to 2014. The transaction information is used to train the BPNN model and evaluate the accuracy of prediction. The analysis results show we models have higher than 80% accuracy rate. It implicates that BPNN can be used to predict the trading volume of crops in the market.**

***Keywords*—neural network, time-series analysis, trading volume prediction, cherry tomato.**

I. Introduction

Agricultural products are crucially important to the human's life and economy worldwide. Due to the growing of plants is affected by weather, season, and lots of external influences, the harvest of crops is unstable and might cause the imbalance between supply and demand in the market. Therefore, the precise prediction of demand in the crops transaction market is important. The prediction of agricultural demand is an complicated problem. That is because of the supply, demand, and price of agricultural products in the market are fluctuating and affected by lots of factors. However, the price fluctuations are a matter of concern among consumers, farmers and policy makers. Therefore, how to make a precise prediction is extremely important for efficient monitoring and planning the agricultural market. Several attempts have been conducted in the past to develop price forecasting models for various commodities [1].

A variety of data mining techniques have been employed in the past to study price fluctuations of crops. For improving the accuracy of prediction, data mining techniques have been rising analysis field [2]. There are lots of data mining and machine learning technologies were used in agricultural research, such as K-Means, K-Nearest Neighbor (KNN), Support Vector Machine (SVM) and Artificial Neural Network (ANN). The results show that the information analysis technologies and algorithms are useful in agricultural prediction research [8]. There are some price prediction systems that are designed based on multilayer perceptron model (MLP) neural network. The system is trained by backpropagation technique. The experimental results show the systems re effective [3].

Artificial Neural Network (ANN) model is applied to predict the environmental influents of potato production. A back-propagation (BP) learning algorithm was chosen to conduct the experiment. The predicting variables used in ANN model were chemical fertilizers, FYM, biocides, seeds, water for irrigation, machinery, diesel fuel and farm size while the six Life Cycle Assessment (LCA) were selected as output parameters. The results obtained, the developed model gave satisfactory predictions in the studied region and appears to be an appropriate tool for prediction of environmental indices of cherry tomato production [4]. In the other, Moving Average (MA) and Artificial Neural Network (ANN) techniques were used to reduce the fluctuation of Agricultural price based on 134 agricultural products during 2009 to 2013. This developed model can be applied to forecast agricultural product price accurately [5].

The model of back-propagation neural network (BPNN) have been addressed in vegetable price prediction [9]. Because of the rapidly changing dimensions for vegetable price and unstable which makes impact in daily life. Data mining classification techniques can be used to develop an innovative model to predict the market. Price prediction is highly useful for farmers to plan their crop cultivation activities. It can be obtained more prices in their crop cultivation and knowing the future price in the advance market. A prediction model is designed based on neural network. There are many kinds of prediction method on basis of Neural Network, among them the application of BP Neural Network algorithm is the most popular one.

Srayut Tongnoy
Department of Tropical Agriculture and International Cooperation
National Pingtung University of Science and Technology, Taiwan

Wen-Ming Hung
Department of Management Information Systems
National Pingtung University of Science and Technology, Taiwan

Deng-Neng Chen
Department of Management Information Systems
National Pingtung University of Science and Technology, Taiwan

There are many studies have been compared different methods of price prediction of crops. Several attempts have been made in the past to develop price prediction models for various commodities [5]. In our research, we apply propagation neutral network (BPNN) to develop time-series models to predict the trading volume of cherry tomatoes in the fruit transaction market in Taiwan. Cherry tomato is one of the crops that are always exposed at risk. In most cases, the farmers are uncertain about their future production and income. The price fluctuates in each year due to the variations in production and imbalance between supply and demand in the market. Thus, modelling and forecasting the annual, seasonal, and monthly trading volume and price is important in practice. We have collected the transaction data from 2011 to 2014, and three time-series models have been developed, that are year time-series model, season time-series model, and month time-series model. The transaction data is used to train and evaluate the time-series BPNN models. The remainder sections of this paper are arranged as followings. The BPNN model is introduced in section two. Our research design and modelling method are described in section three. The data analysis and prediction results are shown in section four, and conclusions are discussed in section five.

II. Backpropagation Neural Network

An Artificial Neural Network (ANNs) is a computational model which is composed of many interconnected neurons. The most used kind of ANNs is the multilayer perception, in which neurons are organized in layers. The input layer neurons receive the input signal to feed in the network. The neurons on the output layers are active and the result they provide is considered as the output provided by the network. There are some hidden layers between the input and output layer. Each neuron can receive input signal from the neuron belongs to previous layer and it can send its output to the successive layer. Back-Propagation Neural Network (BPNN) is usually based on the error back propagation to the multi-layer Neural Network [6].

The back propagation is one of the most popular ANNs algorithms. The meanings are shown in layers, and send their signals "forward", and then the errors are propagated backwards. The algorithm can be decomposed in the following four steps *a)* Feed-forward computation *b)* Back propagation to the output layer *c)* Back propagation to the hidden layer *d)* Weight updates The algorithm is stopped when the value of the error function has become sufficiently slight [7]. The back propagation algorithm uses supervised learning. The algorithm can be defined with examples of the inputs and outputs to the network computation, and then the error (difference between actual and expected results) is calculated. The point of the back propagation algorithm is to reduce this error, until the ANN learns the training data. The training begins with random weights, and the target is to fix up them will be closed the error minimal [10].

There has been large number of studying on forecasting of agricultural commodities price. This section presents a very brief review of the related and recent studies. BP neural network based on principal component analysis to the field of yield forecasting [5]. The methodology has the high efficiency and precision through calculation of the sample. The analysis has stronger applicability for grain yield prediction which is influenced by multiple factors affecting grain yield are quite complex, completely including grain production were flexible and randomness. In addition, its application field and scope of application can indicates that the model of principal component analysis method in BP neural network has the characteristics of higher precision and faster convergence speed. A prediction model was applied by the neural network. The main objectives of the pork consumption series and constructs BPNN, GANN and WNN for time series prediction respectively [11]. The method were achieved a more stable and accurate prediction, Neural network techniques were combined for developing and predicts the pork demand based on ensemble model, reveals that the stability and accuracy of the entirety method is better than the ones of each base model alone. The result indicates that method was provided the most stable prediction for livestock product demand prediction. The model of back-propagation neural network (BPNN) was addressed in vegetable price [12]. Due to the rapidly changing dimensions for vegetable price and unstable which makes impact in daily life. Data mining classification techniques can be used to develop an innovative model to predict the market. Price prediction is highly useful for farmers to plan their crop cultivation activities. They could be obtained more price in their crop cultivation and knowing the future price in the advance market. The aim of this paper is to develop Neural Network model that can be used to predict the price of tomato in agricultural commodities market. These methods can help farmers and government policy makers to obtain more efficient monitoring and planning.

III. Research Design

The transaction data of cherry tomato were collected from 2011 to 2014 in fruit transaction market in Taiwan to train our BPNN prediction models. There are three time-series models developed, including year time-series, season time-series and month time-series.

A. *Data Collection and Modelling*

Cherry tomato trading price and volume are affected by several factors such as climate, supply, demand, and festival etc. Therefore, the prediction is more difficult than ordinary commercial products. It is very difficult to collect data based on all these factors. In this paper, four kind of tomato trading prices were collected as input parameters of the BPNN model, they are the highest price, the lowest price, the middle price and average price. In the other, the trading volume is designed as the output parameter in the model.

In the year time-series prediction model, we use the transaction data of 2011 to predict the trading volume of 2012, and use 2012 to predict 2013, and so on. In the other words, the prediction model can predict the trading volume of current year by analyzing the data of previous year.

In the season time-series prediction model, the data were divided into four seasons, spring (3-5), summer (6-8) autumn (9-11), and winter (12-2). In the BPNN model, we apply the data of the season in current year to predict the data of the same season in the next year. For example, we use data of spring 2011 to predict the data of spring 2012, and so on.

In the month time-series prediction model, the transaction data were divided into twelve by monthly. We use the data of the month in current to predict the data of the same month in the following year. For example, we use data of December 2011 to predict the trading data of December 2012, and so on.

B. *Data Processing and Prediction Accuracy*

The parameters of BPNN model is set as followings:

Nodes of input layer: 4
Nodes of hidden layer: 4
Nodes of output layer: 1
Iterations: 1000
Learning rate: 0.1

The input parameters are the four type of trading price and the output one is the trading volume. The trading volume is normalized into ten classification groups from first to tenth, the lowest trading volume to the highest one, respectively.

We use two indexes to measure the effectiveness of our prediction models, hit rate and accuracy rate. Hit rate is the percentage of the prediction result hit the real classification group. For example, if our prediction model predict the trading volume is the tenth, the highest trading volume category, and the real trading volume is also classified in the tenth category, that is one hit.

Due to the scale of our trading volume classifications is not only nominal but also continuous interval scale, the hit rate cannot represent the error deviation of the prediction result. We defines accuracy rate to measure the prediction result. The accuracy rate is shown in the following equation (1):

$$\text{Accuracy rate} = \left(1-\frac{|P-x|}{10}\right)*100\% \tag{1}$$

x = real trading volume category
P = predicted trading volume category

IV. Results and Discussion

We used C# to develop our prediction system. The transaction data of cherry tomatoes were used to train and evaluate our model. The analysis results of year time-series, season time-series, and month time-series are shown in table one, two, three, respectively.

TABLE I. PREDICTION RESULTS OF YEAR TIME-SERIES MODEL

Target year	Hit items	Total items	Hit rate	Accuracy rate
2012	181	414	43.72%	91.40%
2013	164	440	37.27%	89.57%
2014	200	479	41.75%	90.42%

The results of year time-series model show that the hit rate are found in 43%, 37%, 40%. The hit rate is not good enough. However, the accuracy rate shows around 90%. That is also said our year time-series prediction model can hit the real category or hit the nearby category. Our model shows it is good in year by year prediction.

TABLE II. PREDICTION RESULTS OF SEASON TIME-SERIES MODEL

	Spring				Summer			
Target year	Hit items	Total items	Hit rate	Accuracy rate	Hit items	Total items	Hit rate	Accuracy rate
2012	28	132	21.21%	86.06%	65	85	76.47%	97.29%
2013	19	146	13.01%	82.26%	58	79	73.42%	96.58%
2014	47	144	32.64%	88.82%	64	115	55.65%	91.74%
	Autumn				Winter			
Target year	Hit items	Total items	Hit rate	Accuracy rate	Hit items	Total items	Hit rate	Accuracy rate
2012	49	79	62.03%	94.81%	35	134	26.12%	88.81%
2013	62	80	77.50%	97.50%	53	138	38.41%	88.26%
2014	45	82	54.88%	95.49%	29	146	19.86%	85.41%

Table 2 shows the hit rate and accuracy rate of season time-series prediction. The accuracy rate of prediction in summer and autumn seasons are higher than spring and winter, it means that these two seasons made tomato market prices fairly regular and stable. While the accuracy rate of prediction in spring and winter seasons, they were lower than the optimistic predictions, it mean both seasons are unable when used the neural network technique in the period.

TABLE III. PREDICTION RESULTS OF MONTH TIME-SERIES MODEL

	Jan.				Feb.			
Target year	Hit items	Total items	Hit rate	Accuracy rate	Hit items	Total items	Hit rate	Accuracy rate
2012	2	34	5.88%	80.29%	7	35	20.00%	87.71%
2013	11	51	21.57%	86.86%	5	34	14.71%	87.35%
2014	10	50	20.00%	86.00%	17	38	44.74%	83.42%
	Mar.				Apr.			
Target year	Hit items	Total items	Hit rate	Accuracy rate	Hit items	Total items	Hit rate	Accuracy rate
2012	16	44	20.45%	90.23%	11	37	29.73%	86.49%
2013	15	54	11.11%	83.70%	2	46	4.35%	80.87%
2014	7	44	9.09%	87.95%	5	49	10.20%	83.27%
	May.				Jun.			
Target year	Hit items	Total items	Hit rate	Accuracy rate	Hit items	Total items	Hit rate	Accuracy rate
2012	1	51	1.96%	79.41%	11	28	39.29%	93.93%
2013	3	46	6.52%	80.87%	8	28	28.57%	92.50%
2014	21	51	41.18%	88.82%	8	46	17.39%	90.22%
	Jul.				Aug.			
Target year	Hit items	Total items	Hit rate	Accuracy rate	Hit items	Total items	Hit rate	Accuracy rate
2012	22	31	70.97%	96.13%	26	26	100.00%	100.00%
2013	15	26	57.69%	95.00%	25	25	100.00%	100.00%
2014	4	40	10.00%	90.75%	25	29	86.21%	98.28%
	Sep.				Oct.			
Target year	Hit items	Total items	Hit rate	Accuracy rate	Hit items	Total items	Hit rate	Accuracy rate
2012	25	25	100.00%	100.00%	15	25	60.00%	95.60%
2013	21	24	87.50%	98.75%	26	26	100.00%	100.00%
2014	23	23	100.00%	100.00%	23	23	100.00%	100.00%
	Nov.				Dec.			
Target year	Hit items	Total items	Hit rate	Accuracy rate	Hit items	Total items	Hit rate	Accuracy rate
2012	7	66	10.61%	91.06%	10	49	20.41%	91.63%
2013	11	30	36.67%	91.00%	16	50	32.00%	91.40%
2014	8	31	25.81%	92.58%	7	50	14.00%	88.00%

Table 3 presents the analysis results of month time-series prediction model. The prediction results are worse in Jan and Feb, and best results in Aug, Sep, and Oct.

All the three time-series prediction models show precise prediction capability. Most of the accuracy rates are higher than 80% and even 90%. That is also said that our time-series BPNN model is developed well and can be applied to predict the trading volume of cherry tomatoes in the market effectively.

V. Conclusions

This study used the cherry tomatoes trading prices to predict the trading volume in the market by BPNN. Three time-series prediction models are developed and evaluated. We have provided an insight that how to use BPNN to develop a prediction model for agricultural production and transaction. Our method is contributive to monitor the transaction market and is helpful to make an effective cultivating plan for the farmers.

In this research, we have proved BPNN can be used to predict the behavior of agricultural trading market. Another machine learning classification methods can be considered to develop the prediction model for agriculture, such as decision tree, support vector machine, genetic algorithm and so on. In this "big data" era, there are huge amount of data sources that might contain valuable knowledge. If we can use the analysis tools well, we can do much things that we cannot image before and will be helpful to people.

Although our research has achieved our objectives, there are still some limitations. First, we only collected the data from 2011 to 2014. By a time-series prediction model, if we can get more data to train our model, the prediction power will be enhanced. Second, we only used the price data to be the input parameters, although we have get good prediction accuracy rate, there are lots of another influential factors might be considered to trained the model, such as weather, economic, and government policy factors.

Acknowledgment

This study is conducted under the "Online and Offline integrated Smart Commerce Platform (2/4) Project" of the Institute for Information Industry which is subsidized by the Ministry of Economy Affairs of the Republic of China.

References

[1] H. Ghosh and Prajneshu, Non-linear time series modelling of volatile onion price data using AR(p)-ARCH(q)-inmean, Calcutta Statistical Association Bulletin ,vol 54, pp. 231-47, 2003.

[2] R. Kalpana, N. Shanthi, and S. Arumuga, A survey on data minig techniqus, Int J Adv Comput Sci Tech, vol 3, pp 426-431, 2014.

[3] C. juncha and R. montean, The system of rubber price forecast using neural technique a case study : rubber research institute of Thailand, NCCIT, pp. 845-850, 2009.

[4] B. Khoshnevisan, S. Rafiee, M. Omid, H. Mousazadeh, and P. Sefeedpari, Prognostication of environmental indices in potato production using artificial neural networks, Journal of cleaner Production, Vol 52, pp 402-409, 2013.

[5] P. Tanyamai and B. pudsadee, Forecasting Agricultural Product Prices using Moving Average and Artificial Neural Network Techniques, NCCIT2014, pp. 456-461, 2014.

[6] D. E. Rumelhar and J. L. McClelland, An interactive activation model of context effects in letter perception, Part II: The contextual enhancement effect and some tests and extensions of the model. Psychological Review, 89, 60 – 94. 1982

[7] J. L. McClelland and J. E. Elman, The TRACE model of speech perception. Cognitive Psychology, 18, 1-86. 1986.

[8] D. Ramesh, N. and V. Vardhan, Data mining techniques and applications to agricultural yield data, IJARCSEE, vol 2, pp. 3477-3480, 2013.

[9] F. Khan and D. SinghInt, Association rule mining in the field of agriculture: a survey, J Sci Eng Res, vol 4, pp. 1-4, 2014.

[10] Z. Wang and B. Li, Application of BP neural network based on principal component analysis in grain yield prediction, IEEE, pp. 2010.

[11] Q. Liu, T. Li, W. Xu, and C. Cheng, Livestock product demand prediction based on neural network ensemble, IJACT, vol 5, pp. 499-505, 2013.

[12] G. M. Nasira and N. Hemageetha, Forecasting model for vegetable price using back propagation neural network, Int J Comput Intell Healthc Informat, vol 2, pp. 110-115, 2012.

About Authors:

Srayut Tongnoy is currently pursuing his doctorate in DTAIC, with a major in Management Information Systems, National Pingtung University of Science and Technology, Taiwan. His research interests include prediction techniques in data mining, e-commerce, etc.

Wen-Ming Hung is currently pursuing his master degree in Management Information Systems, National Pingtung University of Science and Technology, Taiwan. His research interests include artificial intelligence and data mining.

Deng-Neng Chen is an associate professor at the department of Management Information Systems, National Pingtung University of Science and Technology, Taiwan. His research interests include knowledge management, e-commerce, knowledge engineering, and artificial intelligence applications. He has published several articles in various journals and conferences.