Int. Journal of Math. Analysis, Vol. 7, 2013, no. 38, 1869 - 1875
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ijma.2013.35113

Agricultural Product Forecasting Using Machine Learning Approach

Thoranin Sujjaviriyasup

Department of Industrial Engineering, Faculty of Engineering
Khon Kaen University, Thailand 40002
Thoraninbki@gmail.com

Komkrit Pitiruek

Department of Production Technology, Faculty of Technology
Khon Kaen University, Thailand 40002
komkri@kku.ac.th

Abstract

In this study we develop the machine learning models for forecasting agricultural products. The main concept of building the models is because machine learning is flexible and convenient to implement and it can be potential applications for a naïve user. The proposed model of Support Vector Machine (SVM) is able to forecast nonlinear or linear forecasting function upon kernel function. Many experiments were performed on the development of SVM and the most precision model by using statistical criteria was also selected. Real data of Thailand's Pacific white shrimp export and Thailand's produced chicken were used to validate candidate models. Autoregressive Integrate Moving Average (ARIMA) is also selected as a benchmarking to compare other developed models. For Pacific white shrimp export case, comparing to ARIMA, the error reduction from MAE, RMSE, and MAPE is 25.76%, 18.11%, and 19.05%, respectively. Moreover, the error reduction from MAE, RMSE, and MAPE is 21.78%, 18.76%, and 18.11%, respectively, for the case of produced chicken.

Keywords: Agricultural products, ARIMA, Support Vector Machine

1. Introduction

In the present, an international trading on agricultural products is very competitive. The international trading becomes complicate as a result of internal impacts such as geography, climate, plague, and so forth. Moreover, external impacts such as strategic planning of competitor are equally important. According to those impacts, an agricultural country that they play both producer and exporter roles have to improve their performance in order to obtain sustainable achievement under constraints. Unfortunately, harvesting and demand are sometimes unpredictable so it is difficult to manage and make any plan for the product. A forecasting model plays an important role in order to capture the future information [8] in planning. An accuracy of the forecasting model is useful to reduce any risk in strategic planning. A forecasting is not only a model [2, 4] that employs statistics to formulate the model [5, 6, 7] such as ARIMA and Holt-Winters but it is also a model that employs machine learning algorithm [3, 7] to formulate the model such as Artificial Neural Networks (ANNs) or Support Vector Machine (SVM). An advantage of machine learning model is flexible and convenient for using. Moreover, we can create applications for agriculture by using forecasting model to predict the future information. In this research, we develop and apply the SVM in forecasting the exports of agricultural product. The proposed model also compare to traditional model using real data sets of Thailand's Pacific white shrimp and Thailand's produced chicken in order to validate performance of the forecasting models.

2. Numerical data

For Pacific white shrimp, the export data [1] from January 2007 to December 2012 are used and represented in Figure 1.

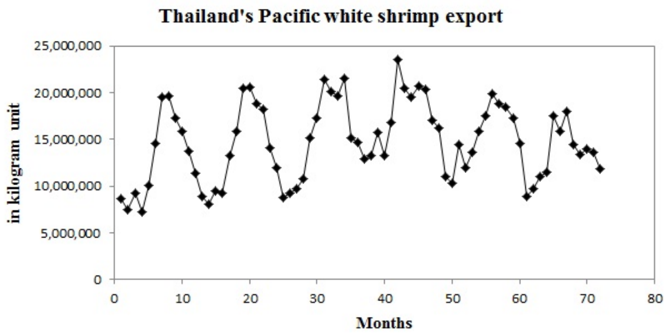

Figure 1: Thailand's Pacific white shrimp export data

For produced chicken, the export data [1] from January 2007 to December 2012 is used and displayed in Figure 2.

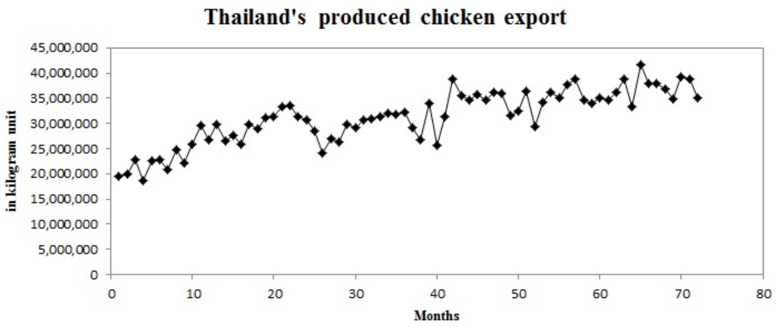

Figure 2: Thailand's produced chicken export data

3. Methodology

ARIMA model is composed of moving average (MA) and Autoregressive (AR) as equation (1)

$$\left(1-\sum_{i=1}^{p}\varphi_i L^i\right)(1-L)^d X_t = \left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t \qquad (1)$$

where L is Lag operator

φ_i is a parameter of autoregressive part

θ_i is a parameter of moving average part

ARIMA model used in this paper that is appropriately selected by AIC (Akaike's Information Criteria), which gives the lowest value. When updated data are realized, the data will be used to forecast for the following information and this procedure will be repeated.

Holt–Winters model or triple exponential smoothing is a mathematical formula that is the combination of the smoothed values of the constant part for time, the sequence of best estimates of the linear trend and the sequence of seasonal correction factor as equation (2).
The additive Holt-Winters model is

$$s_0 = x_0$$
$$s_t = \alpha(x_t - c_{t-L}) + (1-\alpha)(s_{t-1} + b_{t-1})$$
$$b_t = \beta(s_t - s_{t-1}) + (1-\beta)b_{t-1}$$
$$c_t = \gamma(x_t - s_t) + (1-\gamma)c_{t-L}$$
$$F_{t+m} = (s_t + mb_t)c_{t-L+((m-1)(\bmod L))} \qquad (2)$$

Where x_0 is beginning value at time $t = 0$

s_t is the smoothed value of the constant part for time

b_t is the sequence of best estimates of the linear trend

c_t is the sequence of seasonal correction factor

F_{t+m} is the forecasting value of *m* ahead

Holt-Winters method used in this study, the fitted model is decided using Sum Squared Error (SSE). In the model, we use the present data to predict the future data. The updated data will also be used for forecasting the next data repeatedly.

SVM model is machine learning as supervised learning model to find fitted coefficient to set up a line to classify dataset in training process. Not only SVM model can formulate linear model but also SVM model can formulate nonlinear model to classify the nonlinear dataset. In this research, we slot *m* column of historical data and also use *m* - 1 column of historical data to input into the SVM model use kernel function as radial basis and polynomial with degree and coefficient is 1 and 0.2 to predict the future numbers of m^{th} data of Pacific white shrimp and produced chicken, respectively.

4. Cross – validation process

The criteria used in the experiment are mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) in cross-validation check process. All criteria equations are shown in Equation (3) – (5). The ARIMA is selected and used as the benchmarking of univariate time-series forecasting method in order to compare with candidate model.

$$\text{MAE} = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n} \qquad (3)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \qquad (4)$$

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{n} \left| \hat{y}_i - y_i \right| / y_i}{n} \times 100\,\% \tag{5}$$

We separate an export volume of Thailand’s Pacific white shrimp and Thailand’s produced chicken into two datasets as a formulation dataset and a validation dataset. The formulation dataset is 70% of an export data from January 2007 to March 2011. The rest is used to test the validation of forecasting model as shown in Table 1–2.

Table 1: MAE, RMSE, and MAPE of forecasted data of Thailand’s Pacific white shrimp export

Model	Criteria (metric ton unit)		
	MAE	RMSE	MAPE
SVM(5)	**1,504.5250**	**1,978.7900**	**11.22%**
SVM(10)	1696.9475	2,279.8176	12.24%
SVM(20)	1,731.6591	2,499.0625	13.00%
ARIMA	2,026.4584	2,416.3242	13.86%
Holt-Winters(additive)	2,154.2949	2,645.7206	14.90%

Table 2: MAE, RMSE, and MAPE of forecasted data of Thailand’s produced chicken export

Model	Criteria (metric ton unit)		
	MAE	RMSE	MAPE
SVM(10)	**2,171.4905**	**2,600.8397**	**6.06%**
SVM(20)	2,241.2820	2,658.7264	6.10%
Holt-Winters(additive)	2,242.8290	2,806.1192	6.19%
SVM(5)	2,651.3570	3,199.9106	7.24%
ARIMA	2,776.1020	3,201.3756	7.40%

The results of SVM model with actual value in metric ton unit from April 2011 to December 2012. For the Pacific white shrimp and the produced chicken are shown in Figure 3 to 4.

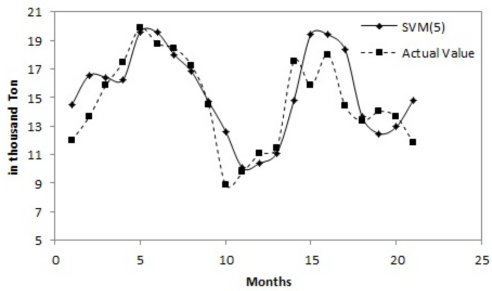

Figure 3: The SVM(5) model for Thailand's Pacific white shrimp dataset

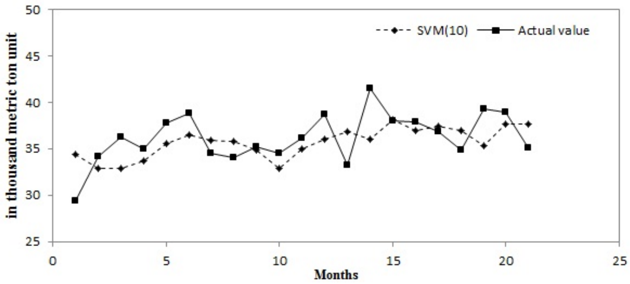

Figure 4: The SVM(10) model for Thailand's produced chicken dataset

5. Conclusion

The results indicate that the developed machine learning model of SVM with the cross-validation check is the most accurate forecasting tool among the others. The machine learning model is flexible, convenient, and can reduce error of forecasting in the case of MAE, RMSE, and MAPE. For the Pacific white shrimp dataset are 25.76% MAE, 18.11% RMSE, and 19.05% MAPE, respectively. For the produced chicken dataset are 21.78% MAE, 18.76% RMSE, and 18.11% MAPE, respectively. In both dataset, the SVM is compared to ARIMA model. Additionally, the machine learning model is appropriate to implement in application and apply to other data sets. Moreover, the SVM model is most appropriate to forecast Thailand's Pacific white shrimp export, and Thailand's produced chicken.

References

[1] http://www.oae.go.th/oae_report/export_import/export.php

[2] J. G. D. Gooijer and R. J. Hyndman, “25 year of time series forecasting”, International Journal of Forecasting 22 (2006) 443 – 473.
[3] M. A. D. Oliveira, “An Application of Neural Networks Trained with Kalman Filter Variants (EKF and UKF) to Heteroscedatic Time Series Forecasting”, Applied Mathematical Sciences, 6 (2012) 3675 – 3686.
[4] M. Paliwal and U. A. Kumar, “Neural networks and statistical techniques: A review of applications”, Expert Systems with Applications 36 (2009) 2 – 17.
[5] M. Shukla and S. Jharkharia, “Applicability of ARIMA Models in Wholesale Vegetable Market: An Investigation”, Proceedings of the 2011 International Conference on Industrial Engineering and Operations Management Kuala Lumpur, Malaysia, January (2011) 22 – 24.
[6] N. H. Miswan and ***et al.***, “On Parameter Estimation for Malaysian Gold Prices Modelling and Forecasting”, Int. Journal of Math. Analysis, 7 (2013) 1059 – 1068.
[7] T. Sujjaviriyasup and K. Pitiruek, “Hybrid ARIMA-Support Vector Machine Model for Agricultural Production Planning”, Applied Mathematical Sciences, 7 (2013) 2833 – 2840.
[8] Y. Acar and E. S. Gardner Jr., “Forecasting method selection in a global supply chain”, International Journal of Forecasting 28 (2012) 842 – 848.

Received: May 5, 2013