

Daily Sales Forecasting for Grapes by Support Vector Machine*

Qian Wen^{1,2}, Weisong Mu³, Li Sun¹, Su Hua⁴, and Zhijian Zhou^{1,**}

¹ College of Science, Applied Mathematics,
China Agricultural University, Beijing, China, 100083

² Department of Mathematics and Statistics, University of West Florida,
Pensacola, Florida, 32514

³ College of Information and Electrical Engineering,
China Agricultural University, Beijing China 100083

⁴ Department of Statistics, North Dakota State University, Fargo, North Dakota, 58102
qw2@students.uwf.edu, {salsally, zhijianzh}@163.com,
wsmu@cau.edu.cn, su.hua@my.ndsu.edu

Abstract. In this article, the quantity of grapes sold in one fruit shop of an interlocking fruit supermarket is forecasted by the method of support vector machine (SVM) based on deficient data. Since SVMs have a lot advantages such as great generalization performance and guarantying global minimum for given training data, it is believed that support vector regression will perform well for forecasting sales of grapes. In order to improve forecasting precision (FP), this article quantifies the factors affecting the sales forecast of grapes such as weather and weekend or weekday, results are suitable for real situations. In this article, we apply ε -SVR and LS-SVR to forecast sales of three varieties of grapes. Moreover, the artificial neural network (ANN) and decision tree (DT) are used as contrast and numerical experiments show that forecasting systems with SVMs is better than ANN and DT to forecast the daily sales of grapes overall.

Keywords: support vector machine, artificial neural networks, grape sale forecasting, ε -SVR, LS-SVR.

1 Introduction

Grapes are special fruits that usually become ripe in summer, i.e., from July to September, and are very popular among fruit customers. Because people are more and more recognizing the nutritional value of grapes, the sales of grapes have also dramatically increased during summer. Unlike the large consumption of grape products in Europe and America, Chinese consumers prefer table grapes. However,

* This paper is supported by the China Agricultural Research System (CARS-30).

** Corresponding author.

grapes are difficult to store because they are perishable. For grape retailers, there are two causes for the loss of grapes: loss caused by storage because of a lack of refrigeration equipment and loss caused by customers who pick and excise grapes according to their preferences. Therefore it is very important for grape retailers to make the right decision in ordering because an insufficient quantity of grapes will not meet the customer demand and the shop owner will obtain less profit. On the other hand, too many grapes may result in a lack of freshness, thereby allowing the grapes to be sold at a discounted price, bringing a loss of profit for retailers. As such, decision makers need an accurate method that is based on mathematics rather than on their experience, to determine the appropriate order quantity of grapes.

Sales forecasting is one of the major tasks in business administration. Precise forecasting of demands can not only decrease inventory costs but also improve the quality of customer service and gain competitive advantages. Recently, some techniques have been implemented to develop some models of forecasting demand for agricultural products with the aim of controlling inventory costs. However, for some fruits, especially grapes, the various factors involved (e.g. climate changes, holidays, and unfixed preference of consumers) are so complicated and changeable that forecast errors significantly influence inventory costs and profits (Roy and Samanta 2011). In this paper, a new algorithm with higher accuracy based on SVM is developed to forecast the demand for grapes, a method which has rarely been applied in such a field before. Because SVMs have greater generalization performance and can guarantee global minima for given training data, it is believed that support vector regression will perform well in forecasting grapes sales.

The rest of this article is organized as follows. The sales forecasting methods review is given in detail in Section 2, and models of SVM are presented in Section 3. Then the forecasting system framework based on the SVM is explored in Section 4. In Section 5, the proposed model is presented, and numerical examples are used to investigate the forecasting performance of the model. The conclusion, contributions of this article, limitations of the research, and some future research directions are provided in Section 6.

2 Selection of Forecasting Method

During the last few decades, many sales forecasting models such as time series, regression analysis, decision tree, ANN and SVM have been developed in the field of perishable product. However not all these methods are suitable for grapes sales forecasting. Next, we will briefly introduce the traditional forecasting models and the SVM sales forecasting model.

2.1 Traditional Method for Forecasting

The traditional methods for forecasting models are mostly based on statistic methods. These methods range from the moving average and exponential smoothing to linear and nonlinear regression. Nonetheless all these models have deficiencies and cannot solve the problem of this article. ARMA model of time series is a method that uses the law of variation of the past variable to forecast future variation of the variable; however, this

method cannot reflect what factors affect the quantity of sale. The regression analysis method is used to reflect the relationship between the quantity of sale and one or more independent variables, but this method is always based on a large number of data to solve the problem. As such it is not feasible to adopt regression analysis to forecast daily sales of grapes. Recently, ANN has received much attention in solving the problem of demand forecasting because of its competent performance in forecasting and pattern recognition. Many studies have attempted to apply the ANN model to time series forecasting. However, ANN models adopt the steepest descent algorithm to find optimal solutions, but they are unable to make sure that the error function of the neural networks converges to a global optimal solution. Moreover, a critical issue concerning neural networks is the over-fitting problem.

2.2 SVM for Forecasting

The SVM has recently been proposed as a new kind of learning network based on the statistical learning theories: the Huber robust regression theory and the Wolfe dual programming theory. SVM achieve good performances in terms of higher accuracy, better generalization and the global optimal solution (S.R.Gunn 1998; Vapnik 2000; Doumpos 2004). Originally, SVMs were developed for pattern recognition and classification problems (Cortes and Vapnik, 1995). Tang Hao (2007) used SVM in mechanical failure diagnosis and proposed the combination of principal component analysis (PCA) and SVM to improve the diagnosis rate dramatically. Wu Jiang (2007) applied SVM in computer-aided detection of cancer diseases and provided a reference of diagnosing cancer. In recent years, SVM has been used in regression problems, which making forecasting by SVM possible. Many scholars use SVM to forecast in various subjects. For example, Wu Qi (2008) forecasted the car sales by SVM based on the Gaussian loss function. Du Xiaofang (2011) used SVM, combined with fuzzy theory, to forecast the demand of perishable farm products.

Therefore, SVM is indeed an effective forecasting method as it needs only a small amount of data to forecast sales. This method applies particularly to the sales forecasting of grape, which lacks historical data. At present, there is no such research that applies the SVM method in forecasting the sales of grapes, thus we use SVMs to forecast the sales of grape in this paper.

3 SVMs Forecasting Model

It is well known that SVMs were developed by Cortes & Vapnik (1995) for binary classification and that they can also be applied in regression problems by introducing an alternative loss function. One character of SVMs is it is an algorithm that can only deal with linear problem. When the system is non-linear, the input vector x , is mapped into a high-dimensional feature space z , via a non-linear mapping, and then conducting linear regression in this space. The inner product of this mapping is called kernel function. In this article, the kernel function used is Radial basis function (RBF):

$$K(x, x_j) = \exp\left\{-\frac{\|x_j - x\|^2}{\delta^2}\right\}.$$

The standard support vector regression model given by following equation:

$$y = f(x) = (w \cdot \phi(x)) + b \tag{3.1}$$

where $\phi(x)$ is in the high-dimensional feature space, which is non-linearly mapped from the input space x . The coefficients w and b are estimated by minimizing risk function $R(C)$:

$$\text{Minimizing } R(C) = \frac{1}{2}\|w\|^2 + C \frac{1}{n} \sum_{i=1}^n L_\epsilon(d_i, y_i)$$

where constant $C > 0$ is penalty factor and $L_\epsilon(d_i, y_i)$ is loss function.

3.1 ϵ -SVR Model

In regression, the quality of estimation is measured by the loss function. There are four possible loss functions that can be used: quadratic loss function, Laplacian loss function, Huber’s loss function and ϵ -insensitive loss function. The ϵ -SVM model selects ϵ -insensitive loss function as its error measurement.

$$L_\epsilon(d, y) = \begin{cases} 0 & |d - y| < \epsilon \\ |d - y| - \epsilon & \text{otherwise} \end{cases}$$

Based on ϵ -insensitive loss function, the decision function of ϵ -SVR model (S.R.Gunn, 1988) is

$$f(x) = \sum (\bar{\alpha}_i - \bar{\alpha}_i^*) K(x, x_i) + \bar{b}$$

$$\bar{b} = \begin{cases} y_j - \sum_{i=1}^n (\bar{\alpha}_i - \bar{\alpha}_i^*) K(x_i, x_j) + \epsilon & \bar{\alpha}_i \in (0, C) \\ y_j - \sum_{i=1}^n (\bar{\alpha}_i - \bar{\alpha}_i^*) K(x_i, x_j) - \epsilon & \bar{\alpha}_i^* \in (0, C) \end{cases}$$

Where $K(x_i, x_j)$ is RBF kernel function.

3.2 LS-SVR Model

The LS-SVR model selects quadratic loss function as its loss function. The formula of quadratic loss function is: $L_q(d_i, y_i) = \sum_{i=1}^n (d_i - y_i)^2$.

Combine the above loss function and RBF kernel function with the equation (3.1), we get the LS-SVR decision function (S.R.Gunn, 1988):

$$y = f(x) = \sum_{i=1}^n \bar{\alpha}_i K(x_i, x) + \bar{b}$$

where $\bar{b} = d_i - \frac{\alpha}{C} - \sum_{j=1}^n \bar{\alpha}_j K(x_j, x_i)$ and $K(x_i, x_j)$ is RBF kernel function.

4 Forecasting System Based on SVM

4.1 Selection of SVM Toolbox

For the moment, there are some toolboxes we can utilize such as LS-SVM toolbox and LIBSVM toolbox of MATLAB. Firstly, LIBSVM is a library for Support Vector Machines (SVMs) and developed by Chih-Jen Lin. This package has been actively developed by researchers since the year 2000 to help users to apply SVM easily. Also there has an easier edition at present for the user who does not know anything about SVM. This edition makes everything automatic--from data scaling to parameter selection (Chih-Chung Chang, 2011). Those are the reasons why we select LIBSVM tool in this paper. Secondly, The LS-SVM toolbox is mainly used with the commercial Matlab package. The Matlab toolbox is compiled and tested for different computer architectures including Linux and Windows. LS-SVM lab's interface for Matlab consists of a basic version for beginners (K. Pelckmans,2003).

4.2 Forecasting Framework

The forecasting system framework is showed in the figure 1. At first, we need deal with data in two ways: One is data normalization processing (in order to avoid data overflow) including smooth the historical sale data (in order to eliminate singular values and noise). Another is to process dynamic information, such as weather data, week data, etc., which is corresponding to the historical sale data, as mentioned in Section 5.2. After that the set of date will separate into two parts, one is called training set and the other one is called testing set.

Subsequently, the training set inputted to SVM model is trained and learnt for adjusting the parameters to the optimal values. The future request is forecasted by the system after the machine completes learning. In addition, we obtain the best parameters C (penalty factor) and γ (a parameter of kernel function) by grid-search on C and γ with cross-validation. At last, forecasting is performed and the values

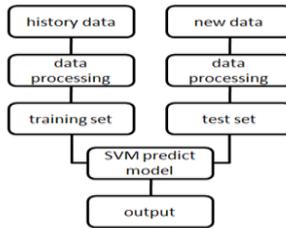


Fig. 1. Framework of forecasting

are obtained after test set is inputted to the trained SVM model. All these process will be done by the SVM toolboxes. What must be mentioned is we perform single-point forecast every time, in other word, there is only one value output in every iteration process. When the sale of a certain day is forecasted, we will put the real value of that day into the training set to renew the history data.

5 The Sale Forecasting of Grape

5.1 About Data

The data we used is obtained from a fruit supermarket called “Fu Man Jia”. The data cover the time from the beginning of July 2011 up to the end of September in year 2012 since grape ripe on the large scale during this period. There are three kinds of grape sold in this market and they are XiaoMiFeng, JuFeng and MeiGuiXiang. We use all of them to test the efficiency of SVM forecasting model in this paper. The data of weather is collected from Website.

5.2 Index of Variables

Forecasting for sale of grape is a complicated procedure that involves multiply variables, and could be treated as regression function $y = f(x) = (w \cdot \phi(x)) + b$. The output value of the regression function is sale quantity y , and the input variable x contain many relevant factors, which control the sale, such as historical sales, weather information, holidays information, etc. The objective of this model is to find a mapping that has a high generalization performance from factors x to sale quantity y . According to historical sales, weather data, holiday's data, etc., we form 8 styles of training samples. Input variable of grape sale forecasting model are shown as:

$$X = (S_{d-1}, S_{d-7}, W_d, W_{d-1}, P_d, P_{d-1}, T_d, T_{d-1})$$

where

S_{d-1}	Sales quantity at the day before the forecasting day
S_{d-7}	Sales quantity at the day 7 days before the forecasting day
W_d	Type of date at the forecasting day (workday or weekend)
W_{d-1}	Type of date at the day before forecasting day (workday or weekend)
P_d	Sale price of grape at the forecasting day
P_{d-1}	Sale price of grape at the day before forecasting day
T_d	The weather condition of the forecasting day
T_{d-1}	The weather condition of the day before forecasting day

Since the data's type of weather condition and holiday are not numerical value, we quantify them as follow:

(1) Quantified value of weather condition

Table 1. Quantified value of weather condition

Weather	value
sunny	1
cloudy	0.9
overcast	0.8
Light rain	0.7
moderate rain	0.5
Showery rain	0.4
downpour	0.2

(2) Quantified value of type of working day date

$$W_d = \begin{cases} 0 & \text{monday, tuesday, wednesday, thursday, friday} \\ 1 & \text{saturday, sunday} \end{cases}$$

5.3 Criteria of Forecasting System

In order to verify the validity of the prediction performance of SVM method, we use the day absolute error as statistical metrics as we only forecast one day's sale every time. Definition of criteria is illustrated in the following expression:

$$\text{DAE} = \left| \frac{d(i) - f(x_i)}{d(i)} \right| \times 100\%$$

Where $d(i)$ and $f(x_i)$ represent actual sales and the forecasting values respectively.

5.4 The Result of Forecasting and Analysis

The result is showed in the following figures and tables. The curves of figure 2, 3 and 4 show the comparison among real data, ε -SVR forecasting value, LS-SVR forecasting value and ANN and DT forecasting values. It is revealed from those figures that SVRs forecasting value are closer to real data than ANN and DT forecasting values. From table 2, 3 and 4, we find that the SVRs have smaller average relative error and maximum relative error. Even though we find the decision tree performs faster than other methods, that's not a decisive advantage with respect to forecasting grapes sale, since we just have small amount of data. As a conclusion, the forecasting systems of SVRs, though does not that satisfy the sale quantity, outperform the ANN and DT method.

All the points that have relative large forecasting error fall into following two categories. One is forecasting value is greater than the sale quantity. In this case, we found that there is no stock of grapes in the store at most of those situations, i.e., the demand quantity is greater than sale quantity. Therefore, with our method, we can not only satisfy the customer demand, but also increase profit of the store owner at those points. Another case is forecasting value is less than sale quantity. There are some special activities that need large amount of grapes may happen at those point. In this case, the customers always place order at least one day advance, our method will not bring loss to the store owner at some points. What is more, those situations rarely happen. Overall, the SVM methods we used are great ways to help store owner to gain higher profit.

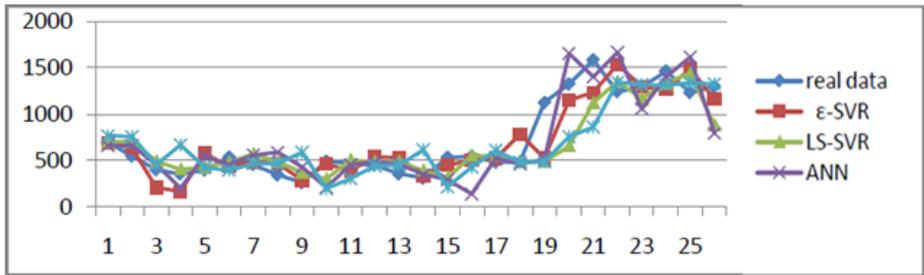


Fig. 2. The result of sale forecasting for XiaoMiFeng

Table 2. Comparison of real data and forecasting result for XiaoMiFeng

method	Average relative error	Maximum relative error	Time spent(s)
ϵ -SVR	0.2124	0.5395	50.32
LS-SVR	0.2229	0.5632	185.81
ANN	0.2890	0.7440	61.12
DR	0.2508	1.2559	5.02

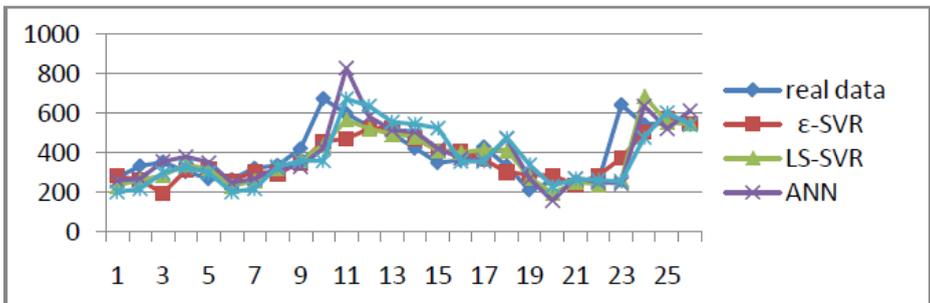
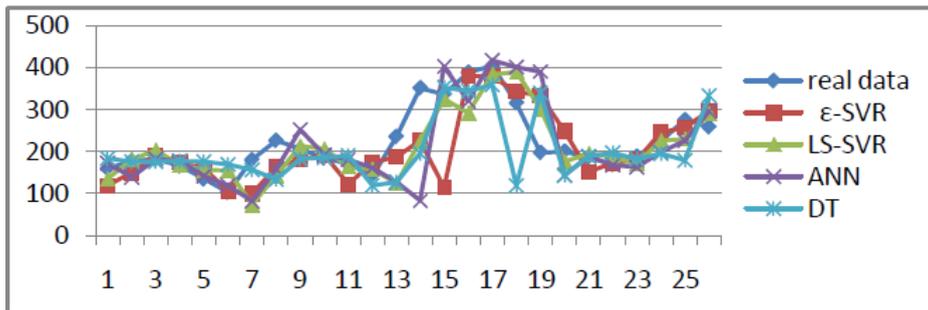


Fig. 3. The result of sale forecasting for JuFeng

Table 3. Comparison of real data and forecasting result for JuFeng

Method	Average relative error	Maximum relative error	Time spent(s)
ε -SVR	0.1407	0.4495	64.38
LS-SVR	0.1528	0.5938	238.27
ANN	0.1896	0.6223	70.30
DR	0.1467	0.6064	6.57

**Fig. 4.** The result of sale forecasting for Meiguixiang**Table 4.** Comparison of real data and forecasting result for Meiguixiang

Method	Average relative error	Maximum relative error	Time spent (s)
ε -SVR	0.1926	0.6777	48.16
LS-SVR	0.1845	0.5957	196.05
ANN	0.2163	0.9715	50.34
DR	0.1725	0.7113	6.13

6 Conclusions

Forecasting is the foundation of fruit supermarket to make order plan and inventory control, while grapes sale has its own characteristics such as multi-dimension, small sample and nonlinearity. It is difficult for the decision maker to forecast the sale accurately by their experience. In this article, the ε -SVR and LS-SVR are used to forecast daily grapes sale, and the result is acceptable. Thus we provide an advanced intelligent forecasting technique for decision maker.

This article also has significant contribution in applications. For example, the forecasting technique we created can be applied in the management of fruit market more successfully. By applying this technique, the correct quantity of fruits with right quality in the appropriate time will be obtained and the shortages or over-stocking will be avoided properly.

However, there still have some limitations in our model. For example, we do not take into account the substituting fruit of the grape that may affects the grapes sale. Further study will focus on improving the algorithm accuracy while more practical factors are involved, so that more realistic sale forecasting result can be obtained in the future.

References

1. Smola, A.: Regression estimation with support vector learning machines. Master's thesis, Technische University at Munchen (1996)
2. Chakraborty, K., Mehrotra, K., Mohan, C.: Forecasting the behavior of multivariate time series using neural networks. *Neural Networks* 5(6), 961–970 (1992)
3. Chang, C.-C., Lin, C.-J.: LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2, 27:1–27:27 (2011), <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
4. Cortes, C., Vapnik, V.: Support vector networks. *Machine Learning* 20, 273–297 (1995)
5. Doumpos, M.: An Experimental Comparison of Some Efficient Approaches for Training Support Vector Machines. *Operational Research* 4(1), 45–56 (2004)
6. Tang, H., Qu, L.: Fault diagnosis of engine based on support vector machine. *Journal of Xi'an Jiaotong University* 9, 1124–1126 (2007) (in Chinese)
7. Wu, J., Dong, T.: SVM applied to modeling of cancer date. *Science Technology and Engineering* 20(7), 5363–5365 (2007)
8. Pelckmans, K., Suykens, J.A.K.: LS-SVMLab Toolbox User's Guide. Katholieke Universiteit Leuven. ESAT-SCD-SISTA Technical Report, 02-145 (2003)
9. Wu, Q., Yan, H.-S., Yang, H.-B.: A Forecasting Model Based Support Vector Machine and Particle Swarm Optimization. *Power Electronics and Intelligent Transportation System* (2008)
10. Roy, A., Samanta, G.P.: Inventory Model with Two Rates of Production for Deteriorating Items with Permissible Delay in Payment. *International Journal of Systems Science* 42, 1375–1386 (2011)
11. Gunn, S.R.: Support Vector Machines for Classification and Regression. ISIS Technical Report, University of Southampton, Department of Electronics and Computer Science (1998)
12. Vapnik, V.N.: *The Nature of Statistical Learning Theory*, 2nd edn. Springer, New York (2000)
13. Du, X.F., Leung, S.C.H.: Demand forecasting of perishable farm products using support vector machine. *International Journal of Systems Science*, 1–12 (2011)
14. Xu, X.-H., Zhang, H.: Forecasting Demand of Short Life Cycle Products by SVM. In: *International Conference on Management Science & Engineering*, vol. 9, pp. 10–12 (2008)