Research on forecast model of Chinese coffee export trade market based on SVM

Lijin Chen

Fuzhou University of International Studies and Trade, Fuzhou, China

Corresponding author and e-mail: Lijin Chen, chenlijin@fzfu.edu.cn

Abstract. Aiming at the problem of the low share of China's coffee exports in the international market, this paper studies the relationship between China's coffee export amount and five major factors influencing on this. The results show that the coffee market should not only pay attention to internal factors-planting area, tea culture and export volume, but also pay close attention to international trends. It is necessary to analyze the requirements of the international market and keep up with the trend. At the same time, the results of using SVM and BP neural network to establish a mathematical model show that SVM has many unique advantages in solving small sample and nonlinear problems with acceptable stability. The prediction effect prove the SVM model has good generalization ability.

1. Introduction

In recent years, coffee culture has prevailed in Asia, and Asia has become the area with the fastest growth in global coffee production, especially in China. According to statistics from China Industry Information Network, from 1998 to 2013, China’s coffee industry showed an overall upward trend. The coffee planting area increased from 244,500 acres in 1998 to 1.4 million acres in 2013, with an average annual growth rate of 12%. In 2018, Yunnan has been regarded as one of the two major regions for coffee cultivation in China. The production area accounts for 98% of the China’s area, and the output value accounts for 98.2% of the China’s total output. Among them, Pu’er City has a unique industrial development, with coffee planting area reaching 77%. 10,000 acres, with an output of 50,000 tons, and an output value of 1.7 billion yuan. The coffee planting area, output quantity and output value rank first in the province [1]. According to the forecast of China's coffee consumption market in the next two years by China Investment Consultants, the average annual compound growth rate of China's coffee in the future is about 17.56%, and the size of China's coffee consumption market will reach 408.2 billion yuan in 2022. However, from a global perspective, coffee production and sales are distributed asymmetrically. Coffee planting and raw material sales are still concentrated in developing countries, refined processing and terminal sales are mainly concentrated in developed countries, and China is still mainly exporting coffee beans. China's imports are mainly coffee products. According to Ding Li and Hou Yuanyuan’s analysis of China’s coffee industry situation and development suggestions, the international market share of China’s coffee exports has gradually increased from 0.29% in 2006 to 0.78% in 2015. Under the strategic promotion of the Belt and Road Initiative and the Free Trade Zone, the competitiveness of China's coffee exports in the
international market has steadily increased, but there is still a certain gap with Brazil, which is the strongest export country [2].

When Ding Li and Zhou Xiangyang analyzed the competitiveness of China's coffee industry, they used the diamond model, a macro analysis tool to analyze national and regional competitiveness, and reached relevant conclusions. In terms of production factors in China’s coffee industry, the continuous increase in labor costs has resulted in higher coffee planting costs compared to other major producing countries, and the immaturity of domestic processing technology has led to a shorter coffee industry chain and weak competitiveness, which is very important for exporting coffee. It is disadvantageous. In the longer and later periods, the export price of coffee of the same quality is higher than that of other major producing countries, which indirectly leads to a smaller share of China's coffee export market [3]. Ouyang Huan and others have also made a comprehensive qualitative analysis on the non-traditional factors affecting the development of China's coffee industry [4]. Therefore, in order to solve the problem of the failure of a quantitative analysis model for China's coffee export, this paper proposes the use of support vector machines (SVM) to establish a prediction model for China's coffee export trade market to achieve effective control and guidance of China's coffee market.

2. Principles and algorithms

The support vector machine itself is a classifier. Suppose there are some data points in the two-dimensional plane, and they belong to two different categories. If a straight line can be drawn in this plane to separate the two types of points, these data points are called linearly separable, as shown in Figure 1. This dividing line is also called the "hyperplane", and the three points on the extreme edge are also called supporting points. If these data points are in an n-dimensional space, the dividing hyperplane is obviously n-1 dimensional. At this time, the three points on the edge are called vectors, and the equation of this hyperplane can be expressed as:

\[ W_1X + b = 0 \]  

Where, \( W \) is the weight; \( B \) is the bias term.

![Figure 1. Schematic diagram of linear classification in two-dimensional space.](image)

However, in most cases, the data is non-linear and it is difficult to get the segmentation line. Therefore, we need to use kernel function to solve the problem, mainly using the following three categories:

- **Linear kernel:**
  \[ K(x, xi) = x \times xi \]  

- **RBF:**
  \[ K(x, xi) = \exp\left(-\frac{(x-xi)^2}{\delta^2}\right) \]  

- **Poly:**
\[ K(x, x_i) = ((x \ast x_i) + 1)^d \]  

(4)

Sigmoid:

\[ K(x, x_i) = \tanh(\eta < x, x_i > + \theta) \]  

(5)

Their main function is to map the linearly indivisible data in the space to a high-dimensional feature space, so that the data in the feature space is separable, and the inner product of the feature space can be performed only in the input space.

Suppose the error between the predicted value by regression through fitting function and the actual output value is \( \varepsilon \), and \( \varepsilon^* \) is the setting error. Ideally all the training sample data to be able to fall in with \( \varepsilon^* \) for radius area. But when the error exceeds \( \varepsilon^* \), beyond the part will be punished. In this way, the problem should be optimized to obtain the parameters of the fitting function that minimizes the loss function, namely:

\[
\begin{align*}
&\min \left[ \frac{1}{2} \|w\|^2 + c \sum_{i=1}^{k} (\delta_i^+ + \delta_i^-) \right] \\
&y_i - w1 \ast x - b \leq \varepsilon^* + \delta_i^+ \\
&-y_i + w1 \ast x + b \leq \varepsilon^* + \delta_i^-
\end{align*}
\]  

(6)

In this equation, \( c \) is the penalty coefficient; \( \delta \) is the relaxation factor. Among them, \( \delta_i^+ \) is the part above \( \varepsilon^* \); \( \delta_i^- \) is the part below the target beyond \( \varepsilon^* \).

Compared with the traditional neural network, SVM has the following advantages: (1) SVM is designed for small sample problems, which can acquire the optimal solution in the limited sample. (2) SVM will eventually be transformed into a quadratic programming problem, which can theoretically obtain a global optimal solution, so as to solve the problem that traditional neural networks cannot avoid local optimality. (3) SVM avoids the problem that traditional nerves need to try to determine the network structure (see Figure 2).

Figure 2. SVM algorithm prediction flowchart.

3. Experimental results and discussion

3.1. Experimental design

This article collects the amount of China’s coffee exports from 2007 to 2018, as well as factors that may affect China’s coffee exports or are indirectly linked to China’s coffee exports, such as China’s coffee export volume, coffee planting area, global green coffee consumption, the average exchange rate of the US dollar and China’s tea exports.

The research model is shown in Figure 3.
Figure 3. Coffee export market forecasting model.

There are the following considerations for selecting the above factors. Evaluate the competitive position of China’s coffee in the international market by the amount of coffee exports; generally, coffee as a kind of goods has a linear relationship with the quantity and unit price. However, due to the wide variety of coffee in the international market, and the difference between the good and the bad, the average price is difficult to grasp and collect, so the direct use of the amount of coffee exports can directly judge the quality of the export situation; the domestic coffee planting area directly affects the annual output and export of coffee; combined with domestic and foreign factors analysis, Global green coffee consumption represents the consumption power of the international market, the stronger the consumption power, the greater the demand; from the perspective of economics, exchange rate changes have a direct effect on international trade [5]. Coffee, tea, and cocoa are the world’s three major beverages. As a major tea exporter in the world, the influence of tea will also affect the export of its commodity substitutes-coffee. Take 2007-2016 as the training set of the SVM prediction model, and the data from 2017-2018 as the verification set. The specific data are shown in Table 1.

Table 1. Coffee export amount and influencing factors from 2007 to 2018.

| Year | Coffee export amount (Million US dollars) | Quantity of coffee exported (tons) | Coffee plantation area (hectare) | Global raw coffee Consumption (tons) | Average us dollar exchange rate | Tea export volume (ton) |
|------|------------------------------------------|-----------------------------------|---------------------------------|------------------------------------|-------------------------------|------------------------|
| 2007 | 42.66                                    | 1.81                              | 20593                           | 738.69                             | 7.604                         | 28.95                  |
| 2008 | 63.43                                    | 2.62                              | 24440                           | 768.26                             | 6.944                         | 29.7                   |
| 2009 | 81.49                                    | 3.27                              | 30271                           | 786.14                             | 6.83                          | 30.3                   |
| 2010 | 102.19                                   | 3.29                              | 43220                           | 806.97                             | 6.769                         | 30.24                  |
| 2011 | 179.74                                   | 3.79                              | 61958                           | 849.16                             | 6.385                         | 32.26                  |
| 2012 | 232.1                                    | 6.14                              | 94133                           | 852.83                             | 6.312                         | 31.34                  |
| 2013 | 216.71                                   | 7.61                              | 97222                           | 854.33                             | 6.278                         | 32.57                  |
| 2014 | 222.42                                   | 7.17                              | 128600                          | 873.82                             | 6.122                         | 30.15                  |
| 2015 | 227.22                                   | 6.64                              | 119624                          | 916.37                             | 6.227                         | 32.49                  |
| 2016 | 472.46                                   | 9.47                              | 118569                          | 923.03                             | 6.642                         | 32.87                  |
| 2017 | 646.36                                   | 10.49                             | 112189                          | 956.76                             | 6.752                         | 35.52                  |
| 2018 | 832.47                                   | 11.55                             | 116970                          | 983.32                             | 6.617                         | 36.5                   |
3.2. Results and discussion

After normalizing the above data, a scatter distribution chart of different factors and coffee export value is established as shown in Figure 4. It can be clearly seen from that there is a strong positive correlation between China’s coffee export value and export volume, which is closely related to the global. There is a certain correlation between green coffee consumption and tea exports, and there is basically no linear relationship with other factors.

![Figure 4](image_url)

Figure 4. (a) Relationship between coffee export quantity and coffee export amount, (b) Relationship between coffee planting area and coffee export amount, (c) Relationship between us dollar exchange rate and coffee export amount, (d) Relationship between quantity of tea exported and coffee export amount, (e) Relationship between global raw coffee consumption and coffee export amount.

The coffee export amount of China in 2017-2018 is predicted by the established SVM mathematical model, and the prediction results are shown in the table below:

| Year | True value | Fitted value | Relative error/% |
|------|------------|--------------|------------------|
| 2017 | 646.36     | 651.11       | 0.736            |
| 2018 | 832.47     | 839.31       | 0.822            |

It can be seen from the prediction results that the relative errors are 0.736% and 0.822%, respectively. There is a small error between the predicted value of the prediction model and the true value, and the fit is good. At the same time, in order to compare the accuracy of the prediction data and the effect of the SVM prediction model, we build a 3-layer BP neural network model with the same sample, and conduct model training on the input and output. The maximum number of iterations is 500. The regression results are shown in Table 3. It can be seen that the average relative error of the SVM regression model is only 2.087%, and the prediction results in recent years are good, and the error is controlled within 1.5%, indicating that the model has strong generalization ability. In contrast, the BP prediction model has large errors in the past three years, and the prediction error in recent years is the core indicator to measure the quality of the model. Compared with the relative error of adjacent years, the fluctuations are strong and unstable, which shows that the SVM model has obvious advantages. The prediction curves of the two models are shown in Figure 5. It can be seen that the SVM prediction curve fits well with the true value.
Table 3. Comparison of coffee export value forecast models in China.

| Year | True value | SVM Fitted value | Relative error | BP Fitted value | Relative error |
|------|------------|-----------------|----------------|----------------|----------------|
| 2007 | 42.66      | 45.59           | 6.89           | 42.47          | -0.45          |
| 2008 | 63.43      | 65.36           | 3.05           | 65.06          | 2.57           |
| 2009 | 81.49      | 84.39           | 3.57           | 83.28          | 2.2            |
| 2010 | 102.19     | 104.39          | 2.16           | 103.07         | 0.86           |
| 2011 | 179.74     | 181.15          | 0.79           | 178.30         | -0.8           |
| 2012 | 232.1      | 231.33          | -0.33          | 230.52         | -0.68          |
| 2013 | 216.71     | 220.74          | 1.86           | 224.55         | 3.62           |
| 2014 | 222.42     | 229.63          | 3.24           | 209.54         | -5.79          |
| 2015 | 227.22     | 226.53          | -0.31          | 217.99         | -4.06          |
| 2016 | 472.46     | 478.55          | 1.29           | 480.73         | 1.75           |
| 2017 | 646.36     | 651.12          | 0.736          | 609.84         | -5.65          |
| 2018 | 832.47     | 839.31          | 0.822          | 852.61         | 2.42           |
| Average relative error | 2.087 | 2.57 |

Figure 5. Curve comparison of the two prediction models.

4. Conclusions
Based on the data of coffee export value from 2007 to 2016 and its five influencing factors, this paper trains the SVM prediction model and predicts the export value from 2016 to 2018, with a relative error of less than 1.5%. At the same time, in order to verify the effect of the model, the SVM model and the BP neural network model are compared, finally it is concluded that SVM has a good solution to the small sample problem, and its training model is more stable and accurate than the BP neural network. Through the results of SVM prediction and analysis, it provides an application reference for forecasting China's coffee export trade market and enhancing China's coffee international competitiveness. For example, the brand effect of Chinese tea is used to promote coffee exports; agricultural enterprises cooperate to introduce advanced planting equipment to expand planting production and enhance quality to expand exports; grasp first-hand information on coffee consumption in overseas countries, and so on. In addition, the accuracy of SVM prediction will not be satisfactory
in some years. It’s estimated that further research is needed in the selection and processing of variable data, or it may be affected by irresistible factors, such as policies and disasters.

References
[1] Yang Lianzhen. Recent trends in global coffee Market [J]. World Tropical Agriculture Information, 2019(12):33-35.
[2] Ding li, hou yuan-yuan. Analysis on the situation and development Suggestions of Chinese coffee industry [J]. Science of tropical agriculture, 2019, 39(03):105-109.
[3] Ding li, zhou xiangyang. Competitiveness analysis and prospect of China's coffee industry [J]. Agricultural outlook, 2017, 13(10):52-57.
[4] Ouyang Huan, Long Cosmos, DONG Yunping, Fu Hongmei. Non-traditional factors affecting the development of China's coffee industry [J]. Agriculture and technology, 2011, 31(02):12-15.
[5] Wang Yuecheng, Zhang Bojiang. An Analysis of the Impact of RMB Exchange rate Fluctuation on China's Import and Export Trade [J]. Modern Finance, 2020(07):26-28.