Stock Price Predicting Using SVM Optimized by Particle Swarm Optimization Based on Uncertain Knowledge

Jin Xin, Kang Yuhong, Zhang Keyi

1, First Author School of Management, Harbin Institute of Technology, P.R.China, jinxinhit@yahoo.com.cn
2 School of Management, Harbin Institute of Technology, P.R.China, kangyuhong62@yahoo.com.cn
*3, Corresponding Author School of Management, Harbin Institute of Technology, P.R.China, ikeyi007@gmail.com

Abstract

Stock prices have the characteristics of nonlinearity, randomicity and uncertainty, so it is difficult to accurately depict the change rules of stock prices using traditional linear forecasting methods, which lead to low stock price prediction accuracy. In order to improve the stock price prediction precision, this paper proposed a stock price predicting model using SVM optimized by particle swarm optimization based on uncertain knowledge (PSO-UK). We used the great optimization ability of PSO-UK to optimize the parameters of SVM, enhanced the learning ability of SVM, and used the SVM to predict stock price. We compared the accuracies of PSO-UK-SVM and PSO-SVM using SSE Composite Index. Experimental results showed that PSO-UK-SVM model performed better degree of fitting and accuracy. The model proposed in this paper has some guiding significance for investors.

Keywords: Uncertain Knowledge, Particle Swarm Optimization, SVM, Stock Price Forecasting

1. Introduction

The characteristics of the stock market decided the positive correlation between its benefits and risks, usually high return accompanied by high risk. The focus question that individuals and institutional investors pay the most attention to is to predict and judge the stock price trend effectively and achieve maximum investment profit under the circumstances of minimizing risks.

In recent years, stock price predicting has aroused wide concern among the scholars. The traditional forecasting methods mainly have multiple linear regression, time series analysis, markov forecasting and so on, whose premise suppositions are based on linear change of stock price to modeling. Although the process is simple and easy to operate, the prediction accuracy is not high. Stock price is influenced by politics, economy, policy, investors’ psychological factors, so it presents obviously nonlinear characteristics. In recent years, with the development of a series of nonlinear intelligence technology, the artificial neural network (ANN) and support vector machine are gradually applied in the stock price forecasting field. Artificial neural network has parallel processing capabilities with great volume, high error tolerance and noise filtering, for the nonlinear relationship with strong approximation ability, but the shortcoming of which, including locally optimal solutions and over-fitting limits the practical application of artificial neural network. Support vector machine (SVM), developed by Vapnik etc, which seeks to minimize an upper bound of the generalization error instead of the empirical error as in conventional neural networks, has been widely applied to time series forecasting for its good nonlinear mapping ability, making up for the lack of ANN.

In SVM, the inappropriate choices of penalty parameter and the kernel function parameters such as the gamma parameter for the radial basis function (RBF) kernel can lead to over-fitting or under-fitting. Particle swarm optimization (PSO), which was proposed by Kennedy and Eberhart, was presented to optimize the SVM parameters. PSO algorithm has fast convergence rate and is conveniently realizable, but is easily run into local optimization and produces early-maturing problem in the process of evolution. Many scholars studied parameter selection strategy for PSO algorithm, PSO algorithm and other algorithms hybrid strategy, These studies, to a certain degree, improved PSO algorithm overall search ability, but can’t essentially...
overcome early-maturing problem. This paper will use particle swarm optimization based on uncertain knowledge to optimize the parameters of SVM, and use the optimized SVM model to predict SSE Composite Index, compared with PSO-SVM model, verifying the superiority of the improved method.

2. particle swarm optimization

The particle swarm optimization (PSO) was originally designed by Eberhart and Kennedy[8] together in 1995, the basic idea came from the inspiration of their early study for birds group behavior modeling and simulation results. The basic idea of this algorithm was to random a group of initial particles, which had no volume and quality, defined each particle as a feasible solution for the optimization problem, good or bad particles were decided by fitness function.

Suppose a group of particles in which the number of particle is M in D-dimensional space, flying at a certain speed. The status attributes of particle i at time t are as follows:

- **Position**: \( x_i^t = (x_{i1}^t, x_{i2}^t, \ldots, x_{id}^t)^T \), \( x_{id}^t \in [L_d, U_d] \), \( L_d, U_d \) are respectively the lower limit and upper limit of the search space;
- **Speed**: \( v_i^t = (v_{i1}^t, v_{i2}^t, \ldots, v_{id}^t)^T \), \( v_{id}^t \in [v_{min,d}, v_{max,d}] \), \( v_{min,d}, v_{max,d} \) are respectively the minimum and maximum speed;
- **Personal best particle position**: \( p_i^t = (p_{i1}^t, p_{i2}^t, \ldots, p_{id}^t)^T \);
- **Global optimal position**: \( p_g^t = (p_{g1}^t, p_{g2}^t, \ldots, p_{gd}^t)^T \).

Then the best position of particle i can be computed according the following formulation:

\[
\begin{align*}
    v_i^{t+1} &= v_i^t + c_1 r_1 (p_i^t - x_i^t) + c_2 r_2 (p_g^t - x_i^t) \\
    x_i^{t+1} &= x_i^t + v_i^{t+1}
\end{align*}
\] (1)

\[
X_{id}^{t+1}=X_{id}^{t}+V_{id}^{t+1}
\] (2)

In Eq.(1) and Eq.(2), \( r_1 \) and \( r_2 \) are two independently uniformly distributed random variables with range \((0,1)\); \( c_1 \) and \( c_2 \) are called learning factors.

Eq. (1) mainly has three parts: the first part represents the trust of particles’ own motion state, the inheritance of the previous speed; the second part represents the thinking of particles itself, a "cognitive" section, strengthens the learning process; the third part represents the shared information and mutual cooperation between particles. All parts above are considered to be the original particle swarm algorithm.

In order to improve convergence performance of particle swarm algorithm, Shi and Eberhart[13] introduced a self-adapting inertia \( \omega \) to coordinate global and local optimization ability of PSO algorithm, and modify the speed equation:

\[
\begin{align*}
    v_i^{t+1} &= \omega v_i^t + c_1 r_1 (p_i^t - x_i^t) + c_2 r_2 (p_g^t - x_i^t) \\
    x_i^{t+1} &= x_i^t + v_i^{t+1}
\end{align*}
\] (3)

The linear decrease formula for \( \omega \) is as follows:

\[
\omega = \omega_{start} - \frac{\omega_{start} - \omega_{end}}{t_{max}} \times t
\] (4)
In Eq.(3) and Eq.(4), \( t_{\text{max}} \) is the maximum iteration, \( t \) is the current iteration, \( \omega_{\text{start}} \) and \( \omega_{\text{end}} \) are respectively the initial self-adapting inertia and terminate self-adapting inertia. This algorithm is called the standard PSO algorithm (SPSO).

3. Particle swarm optimization based on uncertain knowledge

During the motor process of each particle in PSO algorithm, each particle consistently learned and accessed the information from their own and group process to adjust flight direction and speed, but all the knowledge was acquired regarding the best individual as object, limited to individual knowledge and group knowledge. For example, in Eq.(1), when \( r_1 \) and \( r_2 \) are smaller, the knowledge particles acquired from individual and group is little. At this time, the particle whose cognitive abilities has uncertain characteristics, will mainly rely on inertial to fly. Considering this kind of uncertainty, Mei Cong-li\(^{[9]}\) proposed an improved PSO algorithm based on the PSO algorithm, designed the knowledge structure of the particles into three parts including individual cognition, social cognitive and uncertainty knowledge, which was called particle swarm algorithm based on uncertain knowledge (PSO-UK), evolution formulas are as follows:

\[
\begin{align*}
    v_{id}^{t+1} &= \omega_{id}^{t+1} + c_1 r_1 (p_{gd}^{t} - x_{id}^{t}) + c_2 r_2 (p_{ig}^{t} - x_{id}^{t}) + c_3 (c_1 - x_{id}^{t}) \\
    x_{id}^{t+1} &= x_{id}^{t} + v_{id}^{t+1}
\end{align*}
\]

Where
\[
\begin{align*}
    l_1 &= r_1 / (r_1 + r_2 + r_3) \\
    l_2 &= r_2 / (r_1 + r_2 + r_3) \\
    l_3 &= r_3 / (r_1 + r_2 + r_3)
\end{align*}
\]

\( r_1, r_2, r_3, r_4 \) are four independently uniformly distributed random variables with range \((0,1)\); \( c_1 \) and \( c_2 \) are called learning factors; \( C_D = \text{sgn}(r_4 - 0.5) \text{limit} + p_c \) is the boundary of particle uncertainty knowledge, \( \text{limit} \) is the distance from search space boundary to centre, \( p_c \) is the centre of the search space. This paper will use the PSO-UK algorithm to optimize the parameters of the support vector machine (SVM).

4. Support vector machine (SVM) optimized by PSO-UK algorithm

Because of the good nonlinear prediction function of the radial basis function (RBF) \( \exp(-\|x_i - x_j\|^2 / 2\sigma^2) \), the SVM model in this paper will take it as kernel function. In SVM model, the parameters \( \sigma \) and \( \varepsilon \) are user-determined parameters, inappropriate choice of the parameters can lead to over-fitting or under-fitting. Compared with PSO algorithm, PSO-UK algorithm has better capability of the global optimal solution searching. In this paper, we will use PSO-UK algorithm to choose the parameters \( \sigma \) and \( \varepsilon \) of SVM model, the procedures of SVM parameters optimized by PSO-UK algorithm are as follows:

1. Initialize the particle set parameters. Initialize the positions and velocities of each particle \( \{c, \sigma, \varepsilon\} \) randomly, set the maximum iterating times, learning factors etc.
2. Take individual extremum point of every particle as current position, evaluate each initialized
particle’s fitness value. The fitness function is defined as 

\[ F = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{F_i - F^*}{F_i} \right| \]

if particle’s fitness value is better than current individual extremum point, than take \( p^i_d \) as the position of the particle and renew individual extremum point. If the best individual extremum point among all the particles is better than current global extremum point, than take \( p^g_d \) as global extremum point.

(3) Particle state update. The positions and velocities of all the particles are updated according to the Eq.(5) and Eq.(6). If \( v_i \geq v_{\text{max}} \), choose \( v_{\text{max}} \); if \( v_i \leq v_{\text{min}} \), choose \( v_{\text{min}} \).

(4) Stop condition checking. The evolutionary process proceeds until stopping criteria (maximum iterations predefined or the error accuracy of the fitness function) are met. Otherwise, go to step(2).

Figure 1[14] is the specific flow:

5. Experimental analysis

In this paper, we chose the opening price of SSE Composite Index, from April 4, 2011 to April 25, 2012 as experimental data sets with 100 data as the training set and 10 data as the test set[15,16]. Mean absolute percentage error was used to evaluate the effectiveness of the proposed forecasting model. The parameters of SVM optimized by PSO-UK are \( c=93.9465 \), \( \sigma=0.0039 \), \( \varepsilon=0.0039 \). In order to compare the forecasting results, we used PSO-UK-SVM model to predict the same data. Comparison of the forecasting results between PSO-SVM and PSO-UK-SVM is shown in Fig.2 and Tab.1.
Stock Price Predicting Using SVM Optimized by Particle Swarm Optimization Based on Uncertain Knowledge
Jin Xin, Kang Yuhong, Zhang Keyi

Figure 2. Predicting results by PSO-UK and PSO

Table 1. Predicting results by PSO-UK and PSO

<table>
<thead>
<tr>
<th>Date</th>
<th>4.12</th>
<th>4.13</th>
<th>4.16</th>
<th>4.17</th>
<th>4.18</th>
<th>MAPE/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>2310.48</td>
<td>2351.51</td>
<td>2346.31</td>
<td>2355.37</td>
<td>2341.14</td>
<td></td>
</tr>
<tr>
<td>Forecasting</td>
<td>2304.51</td>
<td>2284.56</td>
<td>2300.96</td>
<td>2306.76</td>
<td>2348.64</td>
<td></td>
</tr>
<tr>
<td>PSO-UK Error/%</td>
<td>0.259%</td>
<td>2.847%</td>
<td>1.933%</td>
<td>2.064%</td>
<td>0.320%</td>
<td>1.192%</td>
</tr>
<tr>
<td>Forecasting</td>
<td>2303.74</td>
<td>2283.78</td>
<td>2300.96</td>
<td>2306.76</td>
<td>2348.20</td>
<td></td>
</tr>
<tr>
<td>PSO Error/%</td>
<td>0.292%</td>
<td>2.880%</td>
<td>1.989%</td>
<td>2.096%</td>
<td>0.302%</td>
<td>1.214%</td>
</tr>
</tbody>
</table>

According to the experimental result, we used the SVM model optimized by PSO to predict SSE Composite Index from April 12, 2012 to April 25, 2012, the mean absolute percentage error is 1.214%. The mean absolute percentage error, which comes from the model optimized by PSO-UK predicting SSE Composite Index, is 1.192%. The result is better than that of SVM model optimized by PSO.

6. Conclusion

Stock price forecasting has been a hot spot in recent years, the model developed in this paper can predict the stock opening price accurately, has more practical value for the operation of the shareholder. The model used PSO-UK algorithm to optimize the parameters of SVM, taking SSE Composite Index from April 12, 2012 to April 25, 2012, 10 trading days as the test data. The simulation results showed that the SVM model optimized by PSO-UK algorithm can accurately predict the stock opening price, judge accurately the change trend of stock price, has a broad prospect of application in the stock market. In the future, choosing more reasonable stock price analysis indexes as the input vectors, looking for superior parameters optimization algorithm for SVM are further researches.
References


