International Journal of Multimedia and Ubiquitous Engineering
Vol.9, No.6 (2014), pp.67-80
http://dx.doi.org/10.14257/ijmue.2014.9.6.08

Application of an Optimized SVR Model of Machine Learning

Zhikun Xu[1], Yabin Gao[2] and Yingying Jin[3]

College of Finance, Hebei Normal University of Science and Technology, Qinhuangdao, Hebei, P.R.China, 066004

[1] xuvikp@126.com, [2] 522865482@qq.com, [3] 695735043@qq.com

Abstract

Machine learning is the core of artificial intelligence. It is a fundamental way to the computer intelligence. Support vector machine is one of the important methods in the field of machine learning. It has the advantages of global optimization and strong generalization ability. It has been successfully applied to face recognition, fault diagnosis, financial forecasting and other fields. In this paper, a novel SVR model is proposed to forecast GDP. In the model, The neighborhood rough set (NRS) is used to reduce the index set and the chaotic genetic algorithm (CGA) is adopted to parameters searching in SVR model. Then the novel model NRS-CGA-SVR is established to predict GDP of Anhui province. The results show that the proposed model is better than the other models presented in this paper on forecasting GDP.

Keywords: *Prediction, Machine learning, SVR, GA, Rough set, Chaotic sequence*

1. Introduction

Machine learning is an active field in the development of artificial intelligence. The purpose of the study of machine learning is to let the computer have the ability to acquire knowledge from the real world like human beings. At the same time, establish computational theory of learning and construct all sorts of learning system which are applied in various fields.

The gross domestic product (GDP) is one of the most important indicators to measure of national economic development. It is also the comprehensive reflection of the economic situation. GDP accounting has become the main methods of economic management departments of country and regional government to understand the functioning of the economy. It is also the main basis for the formulation of economic development strategy, planning, and a variety of macroeconomic policy. So it is very important to predict GDP in a scientific method.

At present, there are lots of methods to predict GDP. The main methods are regression method [1-3], time series prediction method [4-6], gray theory method [7-9], Markova prediction method [10-12] and neural network method [13-15], *etc.* In front of the four methods are the traditional forecasting methods. They mostly analyze the causal relationship between variables. Some problems which are multicollinearity and serial correlation exist in the practical application. The neural network algorithm is a kind of nonlinear optimization method in essence. It also has some shortcomings such as long convergence time and easy to fall into local minimum point.

Support vector machine (SVM) was proposed by Vapnik in 1995 [16]. It is based on structural risk minimization principle and can better solve over learning problem. Since then, pretty of Experts and scholars are focus on SVM. Bouzerdoum *et al.* [17] introduced a new

ISSN: 1975-0080 IJMUE

hybrid model which combines SARIMA and SVM for short-term power forecasting of a grid-connected photovoltaic plant. The comparative study showed the developed hybrid model performs better than both the SARIMA and the SVM model. Li and Xiao [18] used the SVM model to predict GDP of Jiangmen and the satisfactory results had been achieved. Hu *et al.* [19] presented the SVM model to predict GDP of Jiangxi province and it also shows adaptability and accuracy of this model. Qiao *et al.* [20] proposed a Bayesian evidence framework to infer the LS-SVR model parameters. The results show that the Bayesian framework of an LS-SVR significantly improves the speed and accuracy of the forecast. Hong *et al.* [21] employs GA-SA algorithm to choose the suitable parameter combination for a SVR model and used the SVRGA–SA model to forecast Taiwanese 3G mobile phone demand. Li *et al.* [22] proposed a consensus least squares support vector regression (LS-SVR) method for calibrating the near-infrared (NIR) spectra and the predicted results show that consensus LS-SVR model is more robust and reliable than the conventional partial least squares (PLS) and LS-SVR methods. Zheng *et al.* [23] adapted an expectation-maximization (EM) algorithm based on two 2-level hierarchical-Bayes models to fast train multi-scale SVR. Mandal *et al.* [24] used rough set theory to reduce the length of experimental data as well as generate rules and used artificial bee colony (ABC) algorithm for the computational training of SVM. Peng *et al.* proposed a hybrid model based on rough set theory and SVM. RS theory was employed as an attribute reduction tool to identify the significant environmental parameters of a landslide and an SVM was used to predict landslide susceptibility. Compared with the general SVM, the hybrid model has superior prediction skill and higher reliability.

In this paper, the overall structure of the study is as follows: In Section 2, the standard SVR model and the principle of neighborhood rough set is introduced. The chaotic genetic algorithm is also presented in this section. In Section 3, we establish the NRS-CGA-SVR model to predict GDP of Anhui province. At last, the conclusion is given in Section 5.

2. Machine Learning

Machine learning is the extension of human cognitive behavior and is also an important branch of artificial intelligence. Machine learning uses some basic methods to understand the objective world and acquires all kinds of knowledge and skills from the study of human behavior. With the help of computer technology, it establishes all sorts of learning models, so as to give the learning ability to computer system. Figure 1 shows the basic idea about the machine learning system. Information may be collected from multiples sources.

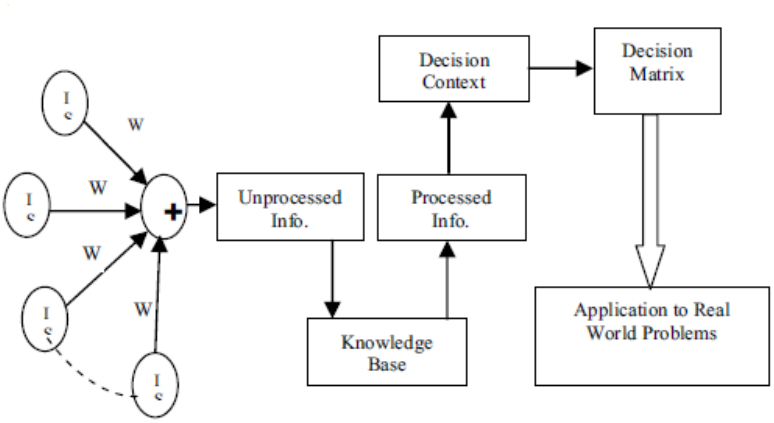

Figure 1. Leaning System

One of the common theoretical bases is statistical from the existing machine learning method, including neural network. Vapnik [16] began to study machine learning problems in the finite sample from 1960s and formed a perfect theory system, named Statistical Leaning Theory in 1990s. It achieved good results on the learnability, correctness, over learning and less learning and local optimum.

3. The optimized SVR model

3.1. The principle of Standard SVR Model

Set a sample set $S=\{(x_1,y_1),(x_2,y_2),\cdots,(x_m,y_m)\}$. Through a nonlinear mapping ϕ, the data from sample set S are mapped into a high dimensional space F. And the following function is used for linear regression in F space.

$$f(x)=w\cdot\phi(x)+b \quad , \quad \phi: R^n \to F \ , \quad w\in F \tag{1}$$

Where b is threshold value. w is regression coefficient vector. The influence factors of w are the sum of empirical risk and the flatness of the high-dimensional space $\|w\|^2$, that is

$$R(w)=\frac{1}{2}\|w\|^2+C\sum_{i=1}^{m}\varepsilon(f(x_i)-y_i) \tag{2}$$

$$\varepsilon(f(x_i)-y_i)=\begin{cases}0, & |f(x_i)-y_i|<\varepsilon\\ |f(x_i)-y_i|, & |f(x_i)-y_i|\geq\varepsilon\end{cases} \tag{3}$$

Where m is the number of samples. $\varepsilon(f(x_i)-y_i)$ is loss function. C is penalty factor.

In order to control the complexity of function, linear regression function should be brought to flat as far as possible. The relaxation factors ξ and ξ^* are also introduced considering the regression error that beyond the precision. The relaxation factors are used to deal with the points which do not satisfy the Eq. (3). Based on structural risk minimization criterion of statistical learning theory, the variable of w and b are determined through support vector regression method by minimizing the objective function $R(w)$:

$$R(w,\xi_i,\xi_i^*)=\frac{1}{2}\|w\|^2+C\sum_{i=1}^{m}(\xi_i+\xi_i^*)$$

$$s.t.\begin{cases}y_i-w\cdot x_i-b\leq\varepsilon+\xi_i\\ w\cdot x_i+b-y_i\leq\varepsilon+\xi_i^*\\ \xi_i,\xi_i^*\geq 0\end{cases} \tag{4}$$

In the Eq.(4), $\frac{1}{2}\|w\|^2$ makes the regression function more flat and have better generalization ability. The second part of Eq.(4) is used to reduce the error. The penalty factor C is a constant number and $C>0$. It is used to control the degree of punishment of samples beyond the error ε. Then establish Lagrange's equation:

$$L(w,\xi_i,\xi_i^*)=\frac{1}{2}\|w\|^2+C\sum_{i=1}^{m}(\xi_i+\xi_i^*)-\sum_{i=1}^{m}\alpha_i[(\varepsilon+\xi_i)+y_i+w\cdot\phi(x_i)+b] \\ -\sum_{i=1}^{m}\alpha_i^*[(\varepsilon+\xi_i^*)+y_i+w\cdot\phi(x_i)-b]-\sum_{i=1}^{m}(\lambda_i\xi_i+\lambda_i^*\xi_i^*) \tag{5}$$

In order to make the Eq. (5) be minimum, function $L(w,\xi_i,\xi_i^*)$ of all parameters of partial derivative is equal to zero. Then the following dual optimization problem can be get.

$$\min\{\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)[\phi(x_i)\cdot\phi(x_j)]+\sum_{i=1}^{m}\alpha_i(\varepsilon-y_i)+\sum_{i=1}^{m}\alpha_i^*(\varepsilon+y_i)\} \tag{6}$$

$$s.t.\begin{cases}\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)=0\\ \alpha_i,\alpha_i^*\in[0,C]\end{cases}$$

Then, the support vector regression problems can be summed up in quadratic programming problem. By solving the quadratic programming problem Eq. (6), we can get the function w which is described by the training sample points.

$$w=\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)\phi(x_i) \tag{7}$$

Where a_i and α_i^* is the solution of $R(w,\xi_i,\xi_i^*)$. Then the regression function can be got. That is

$$f(x)=\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)k(x,x_i)+b \tag{8}$$

Where $k(x,x_i)=\phi(x)\cdot\phi(x_i)$ is the kernel function. Different SVR models can be got by selecting different kernel functions. In this paper, the radial basis function(RBF) is chosen in SVR model. The RBF is shown in bellow

$$K(x,x_i)=\exp(-\gamma|x-x_i|^2) \tag{9}$$

3.2. The Principle of Neighborhood Rough Set

Rough set is proposed in 1982 by professor Pawlak [26]. It uses the method of rough approximation from mass, desultorily, strong interference of data to dig out the potential and essential information. Therefore, the rough set is widely used in attribute reduction, decision rule extraction, classification prediction, etc. But classical rough set is only suitable for dealing with nominal variable. It needs to do discretization for numeric data through various methods, and discretization process inevitably brings effective information loss [27].

Lin [28] first put forward the concept of neighborhood model. The model adopted the neighborhood of space points to grain domain space. It considered the neighborhood as basic information particle and used it to describe the other concepts of the space. Neighborhood rough set model form a neighborhood δ based on each point in real number space. The δ neighborhood group constitutes the basic information particle which describes any concepts in the space. The basic description is as follows:

Set $IS=<U,A,V,f>$ is the information system. $U=\{x_1,x_2,\cdots,x_n\}$ is said to be finite nonempty set. It is called discourse domain. The discourse domain consists of the sample

space set. A is a set of attributes. Here it refers to the prediction index set. V is numerical range. The information function is

$$f : U \times \vec{A} V$$

It describes the relationship of mapping between samples and their attribute values. If $A = C \bigcup D$, where C is condition attribute set. D is decision attribute set. And $C \bigcap D = \varnothing$. Then $IS = < U, A, V, f >$ is a decision table.

Let $x_i \in U$, $B \subseteq C$, the neighborhood of x_i in attribute set B can be defined as

$$\delta_B(x_i) = \{x_i \,|\, x_i \in U, d_B(x_1, x_2) \le \delta\} \tag{10}$$

Where d is the distance function. For $\forall x_1, x_2, x_3 \in U$, d satisfies the following relations.

$$\begin{cases} d(x_1, x_2) \ge 0 \\ d(x_1, x_2) = 0, if \text{ and only } if \ x_1 = x_2 \\ d(x_1, x_2) = d(x_2, x_1) \\ d(x_1, x_3) \le d(x_1, x_2) + d(x_2, x_3) \end{cases} \tag{11}$$

When the sample set has N attributes, the distance function can be defined as

$$d_p(x_1, x_2) = (\sum_{i=1}^{N} | f(x_1, a_i) - f(x_2, a_i) |^p)^{\frac{1}{p}} \tag{12}$$

Where $f(x_i, a_i)$ is the value of sample x_i in attribute a_i. It is worth noting that the distance function d is used to numeric attribute set. But the neighborhood model is easy to extend the numeric attribute set to numeric data and the data which has symbols in distance calculation. For symbol type attribute a_i, set

$$\begin{cases} f(x_1, a_i) - f(x_2, a_i) = 0, when \ x_1 \ and \ x_2 \ have \ same \ value \ on \ a_i \\ f(x_1, a_i) - f(x_2, a_i) = 1, otherwise \end{cases} \tag{13}$$

Then the lower approximation and upper approximation of neighborhood rough set are defined as

$$\begin{cases} \underline{N} X = \{x_i \,|\, \delta(x_i) \subseteq X, x_i \in U\} \\ \overline{N} X = \{x_i \,|\, \delta(x_i) \bigcap X \ne \varnothing, x_i \in U\} \end{cases} \tag{14}$$

The approximate boundary of X is $BN(X) = \overline{N} X - \underline{N} X$.

For a neighborhood decision system $NDT = < U, C \bigcup D, V, f >$, U is divided into N equivalence class by D : $X_1, X_2, \cdots, X_N$, $\forall B \subseteq C$. The lower approximation, upper approximation and approximate boundary of decision attribute set D about B can be defined as

$$\underline{N_B} D = \bigcup_{i=1}^{N} \underline{N_B} X_i \tag{15}$$

$$\overline{N_B} D = \bigcup_{i=1}^{N} \overline{N_B} X_i \tag{16}$$

$$BN(D) = \overline{N_B}D - \underline{N_B}D \tag{17}$$

The lower approximation of decision attribute set D is also called decision positive domain, that is $POS_B(D)$. The size of decision positive domain reflects the degree of separation of classification problem in certain attribute space. If the positive domain is greater, The overlapping region is less. Then the dependence of the decision attribute D to condition attribute B can be defined as

$$\gamma_B(D) = \frac{|POS_B(D)|}{|U|} \tag{18}$$

Eq.(18) shows the radio of the samples which can be completely contains by one kind of decision account for the entire samples based on the description of condition attribute B in sample set.

3.3. The Chaotic Genetic Algorithm

Genetic algorithm is the search algorithm that combines the rule of survival of the fittest in biological process and random information exchange within the group based on the theory of natural selection and genetic theory. It can achieve parallel global search and has lots of advantages, such as simple, rapid and good robustness, etc. It provides a general framework for solving complex optimization problem. The problem does not depend on the specific field and the algorithm has very strong robustness in the types of the problem.

The chaos has the feature of randomness, ergodicity and regularity. It can't repeat traversal of all state in a certain scope according to its own laws. It usually uses logistic iterative equation to describe chaotic phenomena.

$$x^{i+1} = ux^i(1 - x^i) \tag{19}$$

Where $0 < u \le 4$. The chaotic sequence $X = \{x^1, x^2, \cdots, x^n\}$ can be got by Eq. (19). **Fig2.** shows the chaotic dynamics of the system where $x_0 = 0.001$, $i = 300$

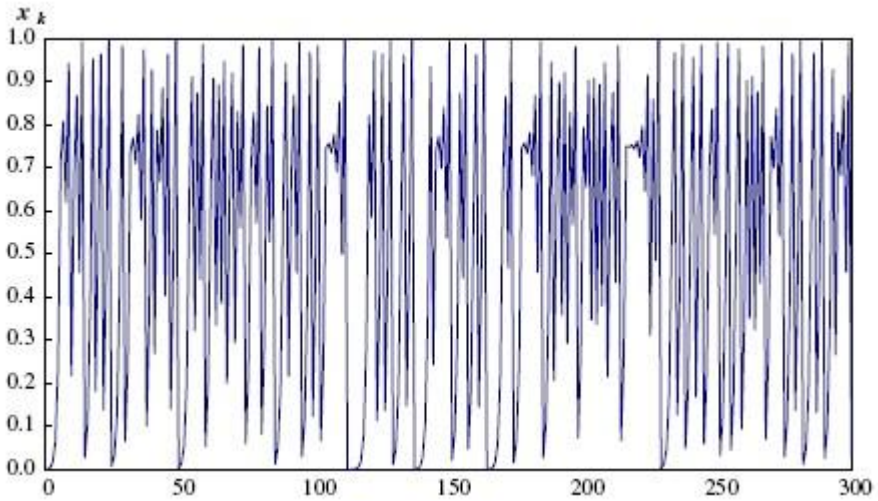

Figure 2. Diversity of Logistic Function

Chaos genetic algorithm is combined with the periodicity of chaotic phenomena and the inversion of the genetic algorithm, using the chaos mapping formula to transform the original population and gets the chaos population according to their own multiple iterations. Then use the new population to do genetic operations. Due to the use of the characteristics of the chaos, the chaos genetic algorithm has better performance on speed and accuracy of searching. Figure 3 shows the flowchart of chaotic genetic algorithm and the main steps of chaotic genetic algorithm are as follows.

Step 1: Chaotic population initialization. There are three parameters $A=(C,\gamma,\varepsilon)$ in a SVR model in this paper. So the individual x^i has three components. x^i is given l different initial values. Set the ith iteration be X^i. Then adopt Eq.(20) to make x^i locate in the interval $(0,1)$.

$$x^i = \frac{X^i - Min}{Max - Min} \tag{20}$$

Where (Min, Max) is the biggest interval of parameters. Then adopt the Eq.(19) to generate the next iteration chaotic variable x^{i+1}. Transform x^{i+1} to X^{i+1} by Eq.(21).

$$X^{i+1} = Min + x^{i+1}(Max - Min) \tag{21}$$

After the transformation, encode individuals into binary format..

Step 2: Fitness evaluation. The fitness function is used to evaluate each individual. In this paper, the reciprocal of mean absolute percentage error(MAPE) is used as the fitness function. The function is shown in Eq.(22).

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - f_i}{y_i}\right| \times 100\% \tag{22}$$

where y_i represents actual values, f_i represents forecasting values, and N is the forecasting periods.

Step 3: Selection. The purpose of selection is to choose a certain number of excellent individuals from current population. Here, the roulette wheel selection principle [29] is used to selection operation.

Step 4: Crossover operation. The selected individuals do match operation in the random position according to crossover probability P_c. In this paper, a single point crossover method is used.

Step 5: Annealing chaotic mutation. From crossover operation, we get population X^i. Then Eq.(23) is used to form the crossover chaotic variable space x^i.

$$x^i = \frac{X^i - M\,\mathrm{ax}}{Max - Min} \tag{23}$$

Where (Min, Max) is current solution space. The chaotic mutation variable is in the Eq. (24).

$$x^i = x^i + \delta x^i \tag{24}$$

Where δ is annealing operation. x^i is the ith chaotic variable. Then the chaotic mutation variable is obtained in Eq.(25) which is mapped to the solution interval by the probability of mutation p_m

$$X^i = Min + x^i (Max - Min) \tag{25}$$

Step 6: Stop condition. If the generations reach the maximum number, then stop the algorithm. Otherwise, go back to Step 2.

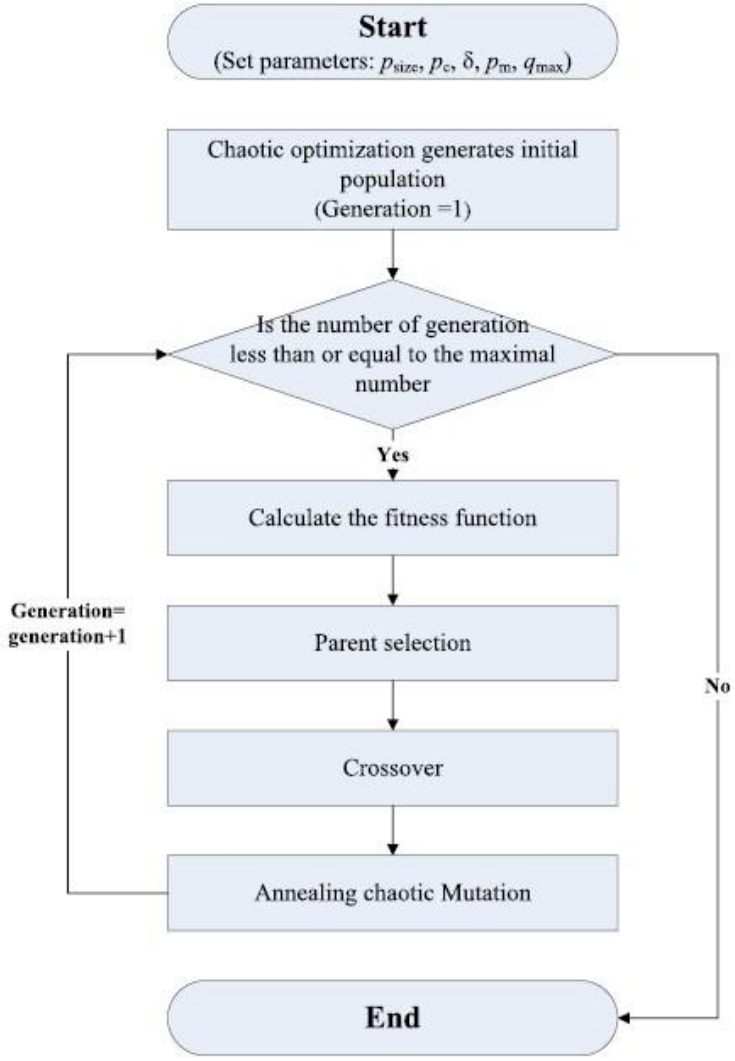

Figure 3. The Flowchart of Chaotic Genetic Algorithm

4. Model Construct and Prediction

The NRS-CGA-SVR model is used to predict GDP of Anhui province. There are many factors that can affect GDP. In this paper, the index set which has influence on GDP is concluded in Table 1. Since the influence of GDP index is up to several dozen. These data often has a great deal of information overlap. They not only affect the SVM generalization ability, but also make the structure of SVR become complex. So the neighborhood RS method is used to reduce the index set so as to eliminate redundant attributes, reduce noise and improve the prediction accuracy. The new index set after attribute reduction is shown in Table 2.

Table 1. Index Set

Variable	Index
X1	The proportion of first industrial population
X2	Total investment in fixed assets
X3	Fiscal expenditure
X4	Loan balance at the end of year
X5	Electricity consumption of whole society
X6	Turnover of freight traffic
X7	Business volume of post and telecommunications
X8	Total retail sales of social consumer goods
X9	CPI
X10	Urbanization rate
X11	Total population
X12	The RMB exchange rate against the dollar
X13	Steel production
X14	Chemical fiber production
X15	Grain production
X16	Pork production
X17	The per capita net income of rural residents
X18	The per capita real income of city residents
X19	Industrial added value
X20	High school students
X21	Total import and export volume
X22	The completed investment in real estate development

Table 2. The New Index Set After Attribute Reduction

Variable	Index
X2	Total investment in fixed assets
X3	Fiscal expenditure
X5	Electricity consumption of whole society
X8	Total retail sales of social consumer goods
X13	Steel production
X17	The per capita net income of rural residents
X18	The per capita real income of city residents
X19	Industrial added value
X21	Total import and export volume

Then we use data from the new index set and GDP data to establish NRS-CGA-SVR model and adopt the model to predict the GDP of Anhui province. The Regression curve of NRS-CGA-SVR model is shown in Figure 4. From Figure 4, we can find intuitively that the regression and forecasting effect of the model are really good. Then NRS-CGA-SVR model is compared with CGA-SVR model, standard SVR model and BP neural network model. The comparison of predicting results are shown in Figure 5, Table 3 and Table 4.

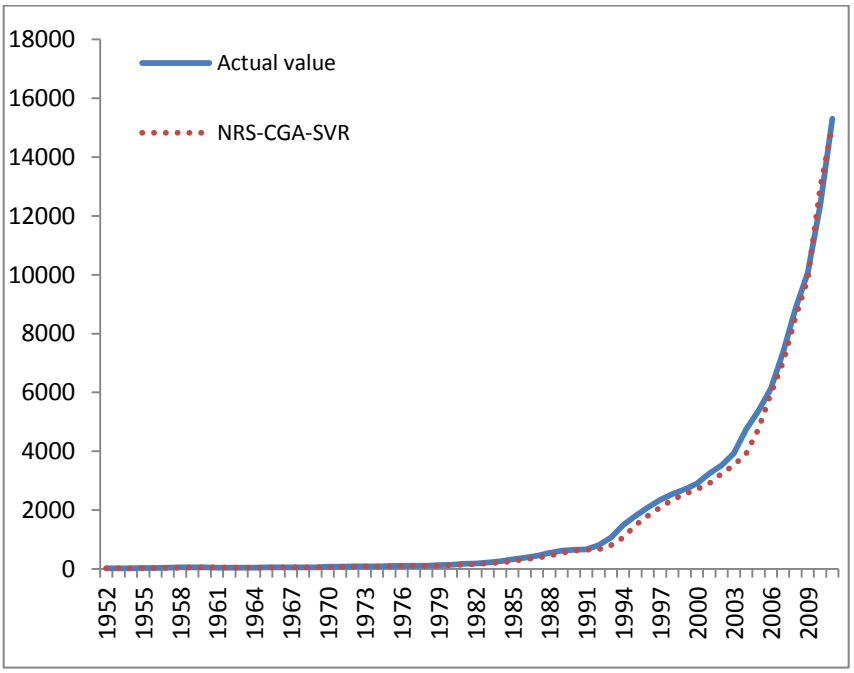

Figure 4. Regression Curve of NRS-CGA-SVR

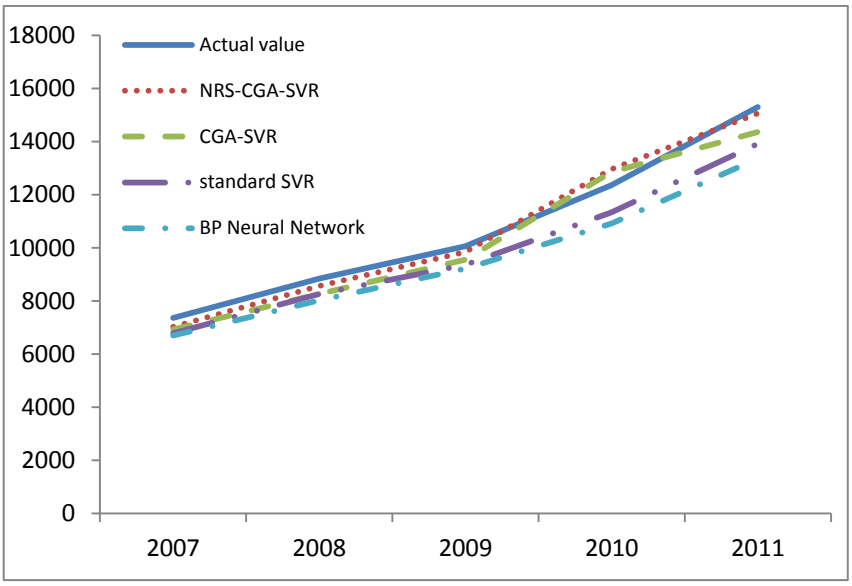

Figure 5. The Prediction Results of Different Models

Table 3. Comparison of Overall Prediction Results

	NRS-CGA-SVR	CGA-SVR	standard SVR	BP Neural Network
MAPE	0.0357	0.0491	0.0824	0.1728
RMSE	14.5482	16.0138	38.2362	55.3411

From Table 3, we can find that the NRS-CGA-SVR model do the best in both MAPE and RMSE. It shows that the model proposed in this paper has good predictive ability. In addition, Table 4 shows the relative error of different models. We can discover that in most of the years from 2007 to 2011, the NRS-CGA-SVR model has better prediction accuracy and the BP neural network model has the worst prediction accuracy of all these models. It demonstrates the superiority of the SVR model. It also indicates that the NRS and CGA methods can enhance the prediction accuracy.

Table 4. The Relative Error of Different Models

	NRS-CGA-SVR	CGA-SVR	standard SVR	BP Neural Network
2007	-0.045233875	-0.0587745	-0.079020176	-0.089869884
2008	-0.032483127	-0.066318114	-0.066034439	-0.092617306
2009	-0.02099012	-0.050802837	-0.068991804	-0.083194834
2010	0.048824758	0.040749652	-0.082052676	-0.116232733
2011	-0.015777911	-0.061500459	-0.09028455	-0.126884302

5. Conclusion

In order to enhance the prediction accuracy, the NRS-CGA-SVR is proposed in this paper. The neighborhood rough set (NRS) is used to reduce the index set which has influence on GDP. The chaotic genetic algorithm (CGA) is adopted to parameters searching in SVR model.

Forecasting results show that the proposed NRS-CGA-SVR model is superior to other alternatives. By the way, this study still has many improvements. For example, there are other advanced optimization algorithms for parameters determination. The index set may be not comprehensive enough. There are some other methods that can be combined into SVR model.

References

[1] H. Y. Lv, H. W. Li and Y. F. Zhang, "GDP Forecast Based on Linear Regression in the Period of Henan Twelfth Five-Years Plan", Henan Science, vol. 29, no. 10, **(2011),** pp. 1241-1244.
[2] J. F. Wang, "The Comparison of the Multivariant Regression Models of Shanghai's and Tianjin's GDP Increases and the Study of Their Development Foreground", Sci/tech Information Development and Economy, vol. 15, no. 3, **(2005),** pp. 121-123.
[3] Y. M. Wang, X. Z. Chen and N. M. Dong, "The Development of Per Capita GDP in Shanxi Province", Science Technology and Engineering, vol. 12, no. 18, **(2012),** pp. 4575-4578.
[4] G. Y. Gong, "The Application of ARIMA Model in Forecasting GDP of Shenzhen", MATHEMATICS IN PRACTICE AND THEORY, vol. 38, no. 4, **(2008),** pp. 53-57.
[5] Q. Li and C. Yang, "Study on Time Series Analysis Model and Its Application in GDP Forecasting", JOURNAL OF ANHUI AGRICULTURAL SCIENCES, vol. 39, no. 20, **(2011),** pp. 12449-12451.
[6] J. G. Zhang and N. Liang, "Combinational Forecasting Model Based on ARIMA and ANN", JOURNAL OF CHINA THREE GORGES UNIVERSITY(NATURAL SCIENCES), vol. 30, no. 4, pp. 86-88. **(2008)**
[7] Z. Q. Li, L. Zhang and M. Z. Chen, "A GREY FORECASTING MODEL ON GDP OF SHANDONG", JOURNAL OF SHANDONG AGRICULTURAL UNIVERSITY(NATURAL SCIENCE EDITION), vol. 39, no. 2, **(2008),** pp. 263-266.
[8] Q. Li and C. G. Pan, "The Research of High Precision Gray Forecast Model and Its Application in GDP Forecast of 2005", RESEARCH ON FINANCIAL AND ECONOMIC ISSUES, no. 8, . **(2005),** pp. 11-13
[9] L. Y. Wang and X. Li, "The Application and Analysis of the Grey Model in the Gross Domestic Product (GDP) of Jiangxi Province", JOURNAL OF NANCHANG UNIVERSITY(NATURAL SCIENCE), vol. 30, no. 5, **(2006)**, pp. 452-454.
[10] Z. Y. Xiao and X. Wang, "Markov Forecast of Per Capita GDP in Each Part of China and Analysis Variation", JOURNAL OF NORTHWEST SCI-TECH UNIVERSITY OF AGRICULTURE AND FORESTRY(SOCIAL SCIENCE EDITION), vol. 4, no. 4, **(2004),** pp. 70-74.
[11] S. Q. Liang, "MARKOV FORECAST AND CHANGE ANALYSIS ON AVERAGE PER CAPITA GDP OF VARIOUS PREFECTURES AND CITIES IN GANSU PROVINCE", CHINESE JOURNAL OF AGRICULTURAL RESOURCES AND REGIONAL PLANNING, vol. 28, no. 2, **(2007),** pp. 38-4.
[12] R. G. Yan and C. R. Qiu, "The Markov Regimes Switching Model of Quarterly GDP Growth Ratio Series in China", APPLICATION OF STATISTICS AND MANAGEMENT, vol. 27, no. 2, **(2007),** pp. 217-222.
[13] Z. G. Chen, "Application of Genetic Algorithm and BP Neural Network in GDP Forecasts", Computer and Digital Engineering, vol. 37, no. 9, **(2009),** pp. 172-175.
[14] X. F. Li, "Data Preprocessing Study in GDP Forcasting Based on BP NN", Computer and Digital Engineering, vol. 39, no.9, **(2011),** pp. 84-87.
[15] N. Chen and X. H. Tai, "Prediction Research of the Per Capita GDP of Liaoning Province Based on the Improved BP Algorithm", SCIENCE AND TECHNOLOGY MANAGEMENT RESEARCH, vol. 30, no. 13, **(2010),** pp. 247-249.
[16] V. N. Vapnik, "The Nature of Statistic Learning Theory", Springer, USA, **(1995).**
[17] M. Bouzerdouma, A. Mellita, A. M. Pavanc, "A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant", Solar Energy, vol. 98, **(2013),** pp. 226-235.
[18] Y. M. Li and J. H. Xiao, "Application of Support Vector Machine Method in Short Term Prediction of GDP", ACTA SCIENTIARUM NATURALIUM UNIVERSITATIS NEIMONGOL, vol. 35, no. 4, **(2004),** pp. 438-441.
[19] Y. G. Hu, Y. X. Wu, and G. H. Jiang, "Research on applying SVM to the regression of GDP", Control and Automation, vol. 23, no. 33, **(2007),** pp. 17-19.
[20] M. Y. Qiao, X. P. Ma, J. Y. Lan and Y. Wang, "Time-series gas prediction model using LS-SVR within a Bayesian framework", Mining Science and Technology (China), vol. 21, no. 1, **(2011),** pp. 153-157.
[21] W. C. Hong, Y. C. Dong, L. Y. Chen and C. Y. Lai, "Taiwanese 3G mobile phone demand forecasting by SVR with hybrid evolutionary algorithms", Expert Systems with Applications, vol. 37, no. 6, **(2010),** pp. 4452-4462.
[22] Y. K. Li, X. G. Shao and W. S. Cai, "A consensus least squares support vector regression (LS-SVR) for analysis of near-infrared spectra of plant samples, Talanta, vol. 72, no. 1, **(2007),** pp. 217-222.

[23] D. N. Zheng, J. X. Wang and Y. N. Zhao, "Training sparse MS-SVR with an expectation-maximization algorithm", Neurocomputing, vol. 69, nos. 13–15, **(2006),** pp. 1659-1664.
[24] S. K. Mandal, F. T. S. Chan and M. K. Tiwari, "Leak detection of pipeline: An integrated approach of rough set theory and artificial bee colony trained SVM", Expert Systems with Applications, vol. 39, no. 3, **(2012),** pp. 3071-3080.
[25] L. Peng, R. Q. Niu, B. Huang, X. L. Wu, Y. N. Zhao and R. Q. Ye, "Landslide susceptibility mapping based on rough set theory and support vector machines: A case of the Three Gorges area, China", Geomorphology, vol. 204, no. 1, **(2014),** pp. 287-301.
[26] Z. Pawlak, "Rough Sets", International Journal of Information and Computer Sciences, vol. 11, no. 5, pp. 341-356, **(1982).**
[27] Q. H. Hu, D. R. Yu, J. F. Liu, and C. X. Wu, "Neighborhood rough set based heterogeneous feature subset selection", Information Sciences, vol. 178, no. 18, pp. 357-359, **(2008).**
[28] T. Y. Lin, "Granular computing on binary relations I: data mining and neighborhood systems", Rough sets in knowledge discovery, vol. 1, **(1998),** pp. 107-121.
[29] J. Holland, "Adaptation in natural and artificial system", Ann Arbor: University of Michigan Press, **(1975).**

Authors

Zhikun Xu, He received his bachelor's degree of Economics in Inner Mongolia Agricultural University, (1993) and master's degree of management in North china Electric Power University, (2010), now he work in hebei normal university of science and technology institute of finance and economics, Qinhuangdao, Hebei. He is a Master's Supervisor. Her major fields of study are practical economics.

Yabin Ga, She received her bachelor's degree of Science in TANGSHAN Normal University Tangshan, Hebei, (2004) and master's degree of Software Engineering in Beijing University of Technology, Beijing, (2012), Now she is a lecturer in Hebei Normal University of Science and Technology, Qinhuangdao, Hebei. Her major fields of study are Education and teaching management, software engineering and education.

Yingying Jin, She received her bachelor's degree of education in Hebei Normal University, Baoding, Hebei (2001) and master's degree of education in Hebei Normal University, Baoding, Hebei. (2005), now she is a lecturer in Hebei Normal University of Science and Technology, Qinhuangdao, Hebei. Her major fields of study are practical economics.

