CSCanada
Advances in Natural Science
Vol. 8, No. 1, 2015, pp. 21-26
DOI: 10.3968/6404

ISSN 1715-7862 [PRINT]
ISSN 1715-7870 [ONLINE]
www.cscanada.net
www.cscanada.org

Using Genetic Algorithm to Optimize Parameters of Support Vector Machine and Its Application in Material Fatigue Life Prediction

ZHANG Lanlan[a],*; LEI Juyang[a]; ZHOU Qilin[a]; WANG Yudong[a]

[a]School of Mechanical Engineering, Shanghai University of Engineering Science, Shanghai, China.
*Corresponding author.

Supported by Graduate Research and Innovation Funded Project of the Shanghai University of Engineering Science (E109031104028).

Received 14 October 2014; accepted 19 January 2015
Published online 26 March 2015

Abstract

Support vector machine is a new kind of learning method based on solid theoretical foundation, but this method has the characteristic of sensitivity to parameter. According to this characteristic, this paper use genetic algorithm to optimize the parameters of SVM and cross validation is introduced to reduce the dependence of the parameters on the training samples. Through the analysis of fatigue data for the relevant literature, take the parameters of the best generalization ability as the final parameters and apply the obtained model (GA-SVR) in material fatigue life prediction. Compared with the conventional SVR model and PSO-SVR model, the mean square error and the square of correlation coefficient are used to verify the reliability and accuracy of the three models. The results show that, the GA-SVR model can predict the fatigue life of materials with high accuracy.

Key words: Support vector machine; Genetic algorithm; Parameter optimization; Material fatigue life prediction; Application

Zhang, L. L., Lei, J. Y., Zhou, Q. L., & Wang, Y. D. (2015). Using Genetic Algorithm to Optimize Parameters of Support Vector Machine and Its Application in Material Fatigue Life Prediction. *Advances in Natural Science*, *8*(1), 21-26. Available from: http://www.cscanada.net/index.php/ans/article/view/6404
DOI: http://dx.doi.org/10.3968/6404

INTRODUCTION

Support vector machine (SVM) is a small sample learning method, the core content was proposed by Vapnik and his collaborators (1992). Support vector machine is based on the VC dimension theory and structure of statistical learning theory based on the minimization principle, is by far the most successful statistical learning theory algorithm (Deng et al., 2009; Chen et al., 2011). It can find the best compromise between complexity and learning ability of the model based on limited sample information in order to obtain good generalization performance (Gu et al., 2014). Support vector machine is a quadratic optimization problem; its final decision function is determined by only a small number of support vectors. Because of its excellent learning performance, has become the hotspot of machine learning. And it has been successfully applied to pattern recognition, machine learning, text classification, bioinformatics and information security and other fields.

With the development of artificial intelligence technology, intelligent prediction method of fatigue life has attracted more and more attention, has been a lot of applications in the life study. Mathew et al. (2008) using the neural network can predict the low cycle fatigue life of 316L stainless steel within twice the range of the experimental values of different test temperature and nitrogen content; Lotfia et al. (2013) using a neural network to predict the fatigue limit of the iron-based powder metallurgy parts; Li et al. (2001) using BP neural network establish the relationship between external loads, damage parameter and changes of material properties estimate the fatigue life of the structure; Abdalla et al. (2011) using artificial neural network RBF model, taking the maximum tensile strain and pressure ratio as input, and put forward the model of fatigue life of steel reinforcing bars. But in the forecasting process, there will be over-learning problems, need to collect a lot of material fatigue reliable experimental data adequately trained predictive model, which limits the wide application of this method to a large extent. Compared with neural networks, support vector machine has a rigorous theoretical basis and mathematical foundations of the small sample size problem also has a strong generalization ability, weak dependence on the number of samples. SVM predictor of performance is sensitive to the selection of parameters, most of the practical application of empirically determined parameters or using test algorithms, resulting in inaccurate due to parameter selection leaving the final prediction accuracy is lower than the target accuracy. Therefore, the author builds support vector machine prediction model based on the genetic algorithm to optimize the parameters by genetic algorithm to automatically search for and identification of parameters. The application of support vector machine to predict fatigue life of materials, can reduce the fatigue test cost, shorten the test time, make the research results as soon as possible to apply in the practical engineering.

1. SUPPORT VECTOR MACHINE THEORY

1.1 SVM Regression

In general, the requirement of regression problems is: given a new model x, according to the training set T to infer its corresponding output y is. The basic idea of the SVM regression is through nonlinear mapping to turn the sample space to a high-dimensional or infinite dimensional feature space. And seek an optimal hyper plane in the feature space.

Given a training set $T=\{(x_i, y_1),(x_2, y_2),\cdots,(y_l, y_l)\}$, among them $x_i \in R_l$, $y\mathrm{i} \in R, i=1,\cdots l$. According to the structural risk minimization principle, SVM regression function is obtained by for minimizing the following objective function:

$$\min\left\{\frac{1}{2}\|W\|^2 + C\sum_{i-1}^{l}\left(\xi_i+\xi_i^*\right)\right\}. \tag{1}$$

$$\text{Subject to: } \begin{array}{l} y_i-(w\cdot x_i)-b\le\varepsilon+\xi_i \\ (w\cdot x_i)+b-y_i\le\varepsilon+\xi_i^* \\ \xi_i,\xi_i^*\ge 0, i=1,\ldots l \end{array}. \tag{2}$$

Here the penalty factor C,C>0, shows the degree of punishment beyond the error of the sample; ξ_i, ξ^*_i are the relaxation factor; ε is the insensitive loss function. Though Lagrange function method and applies KKT conditions, you can get the dual optimization problem of the Equations (1) and (2):

$$\max\left[-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)K(x_i,x_j)- \varepsilon\sum_{i=1}^{l}(\alpha_i^*+\alpha_i)+\sum_{i=1}^{l}y_i(\alpha_i-\alpha_i^*)\right]. \tag{3}$$

$$\text{Subject to: } \begin{array}{l} \sum_{i=1}^{l}(\alpha_i^*-\alpha_i)=0 \\ \alpha_i,\alpha_i^*\in[0,C],\ i=1,\ldots l \end{array}. \tag{4}$$

Solving the dual optimization problem can obtain the regression problem decision function:

$$f(x)=\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)K(x_i,x_j)+b. \tag{5}$$

1.2 Kernel function and Its Characteristics

The introduction of kernel function is an approach to solve support vector machine nonlinear problem. The kernel function of input data samples is mapped to high dimensional feature space, regression in high dimensional space. As long as the selection of the proper kernel function need not determine the specific forms of nonlinear transform of ϕ, with K kernel function satisfying the Mercer conditions (x_i, x_j) to replace the corresponding inner product ($\phi(x_i)\cdot\phi(x_j)$), still can get the decision function of the original space. The introduction of kernel function can establish the nonlinear model in the feature space under the condition of us don't need to know the concrete expression of the nonlinear mapping. The results obtained with different kernel function are different. The most commonly used Gaussian kernel function is applied in this paper:

$$K(x_i,x_j)=\exp\left(-\gamma\|x_i-x_j\|^2\right),\gamma>0. \tag{6}$$

The kernel function has the following characteristics:

(a) by introducing the kernel function, avoid the curse of dimensionality problem, can greatly reduce the amount of calculation.

(b) do not need to determine the specific expressions of nonlinear transformation.

(c) the kernel function and parameters will lead to different mapping from the input space to the feature space is different, have different properties of the feature space.

1.3 The Influence of Parameters on the Model

As a new machine learning method, support vector machine has some room for improvement, the parameter selection is one of the problems urgent to perfect. Due to the lack of theoretical guidance, the traditional parameter selection is through repeated test, artificial selection out satisfactory solution. This method needs the guiding of people's experience, and its selection needs higher cost of time, so the traditional parameter selection method and can not adapt to the development of support vector machine.

Because there are many kinds of kernel function, an important step in SVM design is the selection of kernel function and kernel parameters. Research has pointed out the different kernel function has little effect on the performance of SVM, so here kernel function is Gauss function the most commonly used. Reasonable selection of design parameter values is vital to ensure the performance of support vector machines.

Support vector machine (SVM) is sensitive to parameters(Wang et al., 2007), the performance of support vector machine (SVM) regression algorithm is good or not mainly depends on the selection of three parameters, the three parameters are not sensitive parameters, error penalty factor and kernel function parameters. They balance between training error and generalization, specific performance of the impact on the model is:

(a) Penalty factor C characterizing the degree of punishment error, C greater the higher the fitting, training error and test error are reduced, the training becomes time consuming. But C is too large, there will be over-fitting phenomenon, and will increase the prediction error.

(b) ε, the greater the lower fitting accuracy, less the number of support vectors, the complexity of the accuracy of the model is reduced; its generalization ability is stronger.

(c) When γ is small, the radial basis function fitting performance is poor, when γ is larger, has a good fitting performance, but will cause poor generalization ability.

2. GENETIC ALGORITHM (GA) TO OPTIMIZE PARAMETERS

2.1 Genetic Algorithm Theory

Genetic algorithm (Zuo, 2011) is a procedure by simulating biological evolution of the law of search optimization method; it is an object of all the individuals in the population, the use of randomization techniques for an efficient encoding parameter space search. Among them, the selection, crossover and mutation constitute the genetic manipulation of genetic algorithms. Genetic algorithm has become a tool to solve complex problems.

The genetic algorithm provides a general framework to solve the optimization problems of complex systems, it does not depend on specific fields of problems, types of the problem have very strong robustness. Different from the traditional search algorithm, search ability of genetic algorithm to solve the model does not depend on the specifics. Genetic algorithm randomly generates a set of initial solutions, become group, begin the search process. Each individual in the population is a solution to the problem, called chromosomes. These chromosomes in the subsequent iteration evolving, called genetic. Genetic algorithm is mainly through the cross, mutation, selection operation to generate the next generation of chromosomes, called offspring. Chromosome is good or bad for the fitness measure. To adapt to the high probability of individual be inherited to the next generation in the group; lower fitness, is genetic to the next generation of small probability. According to the fitness value selecting a certain number of individuals from the previous generation and the offspring, as the next generation of group, and then continue to evolve, so after some generations, the algorithm converges to the best chromosome, it is likely that the optimal solution of the problem or is suboptimal solution. The use of adapted to the fine degree this concept is used to measure the individuals of the population in the optimization calculations may have reached the optimal solution in genetic algorithms. Measure of individual fitness function is called fitness function, the definition of general and specific issues related to.

The basic steps of genetic algorithm to optimize SVR parameters are as follows:

Step 1 determines C, ε and γ may be the range and generate the initial population to be encoded;

Step 2 calculate its mean square error function to determine its fitness;

Step 3 determine the best individual in the population of the fitness function value meets the stop condition, the stop condition is generally reaches a specified error or the maximum genetic algebra;

Step 4 failure to reach a stop condition, the application selection, crossover and mutation to produce new populations, go to step 2 to continue the iterative optimization;

Step 5 repeat steps 2-4 until the end of the training;

Step 6 output optimization results and training samples for training to get the regression model. Specific processes genetic algorithm to optimize the parameters of support vector machine shown in Figure 1.

2.2 The principle of Setting Parameters of Genetic Algorithm

Genetic algorithm parameters settings of the control principle is as follows:

(a) the population size is too small will cause the premature convergence phenomenon is too big, make the fitness evaluation times increase sharply, the convergence speed significantly reduced. In this paper, the population size is 20, the evolution algebra is taken as 200.

(b) the crossover probability is too low may lead to a search block, too high and may lead to destruction of the original model. In this paper, the crossover probability is 0.7.

(c) mutation probability is too small may make some gene bit premature loss of information cannot be recovered, and the mutation probability high genetic search will become a random search.

Optimization of support vector machine in each interval parameter settings are as follows: $C \in [0,100]$, $\gamma \in [0,1000]$, $\varepsilon \in [0.001,1]$.

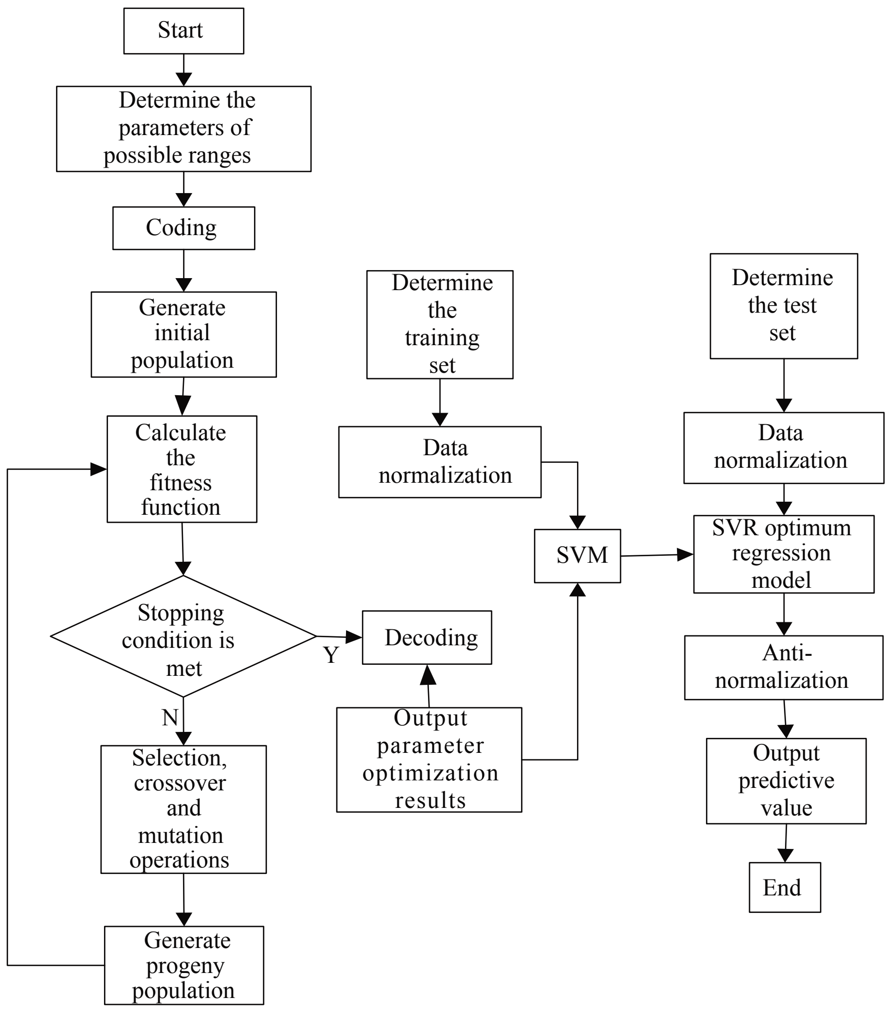

Figure 1
GA - SVR Overall Flow Chart

3. MODELING AND VERIFY THE PREDICTION RESULT

3.1 Experimental Data

In this paper, SVM is applied to material fatigue life prediction; the material is P91 steel base metal and welding consumable. Material samples as shown in Table 1, from the literature (Ji, 2011). The hold time of P91 steel base metal and welding consumables fatigue sample data are used as input and fatigue life are used as output, establishing the fatigue life prediction model. Training with the front 14 sets of data to predict the remaining two groups.

Table 1
The Training Sample Data and Forecast Data (P91 Steel)

Base metal			Welding consumable		
Sample number	**Hold time(t_H/s)**	**Fatigue life($N_f^{t_H}$/time)**	**Sample number**	**Hold time(t_H/s)**	**Fatigue life($N_f^{t_H}$/time)**
M27	85	5,375	H27	85	3805
M23	265	2,708	H21	145	2708
M32	385	1,200	H20	265	1654
M13	445	920	H23	445	903
M33	505	770	H19	565	320
M24	565	680	H25	685	414
M20	685	514	H22	745	412
M25	865	450	H29	925	301
M21	1,345	278	H17	1,165	201
M16	1,465	242	H30	1,225	164
M30	1,525	221	H26	1,345	140
M22	1,645	225	H16	1,465	138

To be continued

Continued

Base metal			Welding consumable		
Sample number	**Hold time(t_H/s)**	**Fatigue life($N^{t_H}_f$/time)**	**Sample number**	**Hold time(t_H/s)**	**Fatigue life($N^{t_H}_f$/time)**
M29	1,705	182	H31	1,585	130
M15	1,765	195	H15	1,765	101
M19	1,045	358	H18	865	362
M17	1,165	315	H24	1045	302

3.2 Evaluation Indicator

In order to get high and low of the prediction precision, the mean square error(MSE) and correlation coefficient square(R^2) are chosen as evaluation indicator of the model. The formula is as follows:

$$\mathrm{MSE} = \frac{1}{l}\sum_{i=1}^{l}\left(f(x_i)-y_i\right)^2 , \tag{7}$$

$$R^2 = \frac{\left(l\sum_{i=1}^{l} f(x_i)y_i - \sum_{i=1}^{l} f(x_i)\sum_{i=1}^{l} y_i\right)^2}{\left(l\sum_{i=1}^{l} f(x_i)^2 - \left(\sum_{i=1}^{l} f(x_i)\right)^2\right)\left(l\sum_{i=1}^{l} y_i^2 - \left(\sum_{i=1}^{l} y_i\right)^2\right)} \tag{8}$$

Here $f(x_i)$shows the prediction value, y_i shows the original sample data.

3.3 Results and Analysis

Using the normalized means to deal with the original sample data and using genetic algorithm to optimize parameters get the best regression model. With this optimization model to predict the training sample and the test sample, and then test the model results. After several parameters optimization training, gained the optimal parameters of base metal are: γ=22.9588, C=29.519, ε=0.001589.The original data and forecasting regression data of base metal as shown in Figure 2.

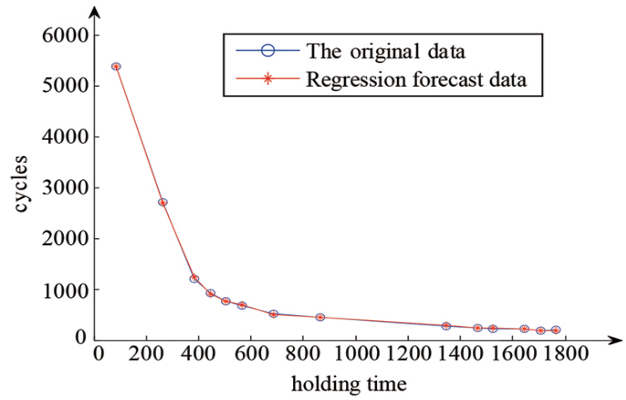

Figure 2
The Contrast Between the Original Data and Regression Data(Base Metal)

Compare with PSO-SVR and SVR model, the contrast of the model evaluation index as shown in Table 2.

Table 2
Base Metal Predictive Model Evaluation Index Contrast

Evaluation index		GA-SVR	PSO-SVR	SVR
Training sample	MSE	7.028×10^{-6}	1.349×10^{-4}	0.0204175
	R^2	99.99%	99.82%	91.12%
Predicting sample	MSE	1.3774×10^{-6}	1.28×10^{-4}	0.126418
	R^2	100%	100%	100%

The optimal parameters of welding consumables are: γ=23.5949, C=99.7016, ε=0.001. The original data and forecasting regression data of welding consumables as shown in Figure 3. The model evaluation index as shown in Table 3.

Table 3
Welding Consumables Predictive Model Evaluation Index Contrast

Evaluation index		GA-SVR	PSO-SVR	SVR
Training sample	MSE	2.39485×10^{-5}	0.000387	0.00220703
	R^2	99.97%	99.5744%	97.92%
Predicting sample	MSE	7.55×10^{-7}	0.0001	0.01083
	R^2	100%	100%	100%

Comparison of three kinds of model predictions, overall, the highest precision is GA-SVR model with the best results, PSO-SVR model is followed, SVR model for the worst. Evaluation of training samples, GA-SVR model prediction means square error is small, square of the correlation coefficient is closer to 1, indicating that GA-SVR model has better approximation ability.

In addition, three kinds of model comparison to illustrate the importance of selecting the parameters on the performance of SVR model. If SVR parameter selection is unreasonable, the final forecast accuracy will be lower than the target accuracy. Besides, the prediction accuracy of SVR model on the training sample data is higher than the prediction accuracy on the test data indicating that only by support vector machine model experience

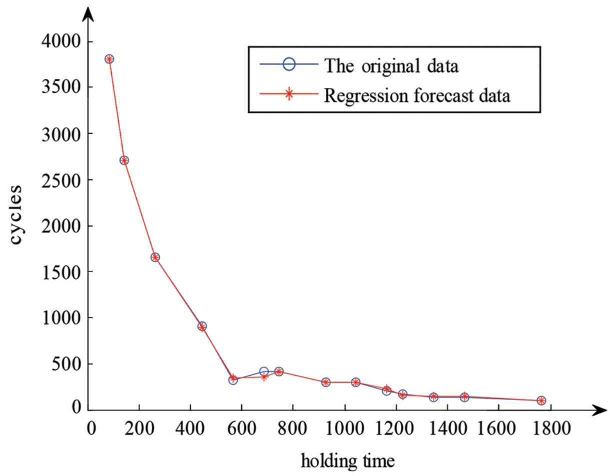

Figure 3
The Contrast Between the Original Data and Regression Data (Welding Consumable)

parameters selection does not have good generalization ability. However, parameter optimization of support vector machine model (GA - SVR and PSO - SVR) won the good performance and achieved a balance between learning ability and generalization ability. Especially GA-SVR full use of the limited sample information is the optimal solution, and finds the best compromise between the learning and generalization performance, the ability to get a good promotion.

CONCLUSION

(a) In the process of SVM prediction, selection and optimization of parameters play a key role. If the parameter selection is unreasonable, the prediction accuracy and running time will be affected. In order to improve the accuracy of parameter selection, we use cross-validation approach to optimizing the parameters. And a set of parameters for the best generalization ability as the final parameters of training samples, validated gained good prediction precision.

(b) Support vector machine parameters optimization based on genetic algorithm has good effect, to illustrate the effectiveness of the proposed method for the prediction of fatigue life of materials. This method is of high precision, high stability prediction accuracy and also can be applied to other forecasting models.

(c) Research on support vector machine has achieved some results, but how to get the effect parameters so that the model has higher reliability is still a problem. In addition，to application in the more fields is also a direction for future research.

REFERENCES

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). *A training algorithm for optimal margin classifiers* (pp.144-152). In D. Haussler (Ed.). Proceeding of the 5th Annual ACM Workshop on Computational Learning Theory. ACM Press.

Chen, Y. Y., & Xiong, Q. F. (2011). *Application of support vector machine method tutorial* (pp.1-83). China Meteorological Press.

Deng, N. Y., & Tian, Y. J. (2009). *Support vector machine-theory, algorithm and expand* (pp.63-111). Beijing, China: Science Press.

Gu, J. Y., Liu, J. F., & Chen, M. (2014). A modified regression prediction algorithm of large sample data based on SVM. *Computer Engineering, 40*(1), 161-166.

Jamal, A. & Abdalla, R. H. (2011). Modeling and simulation of low-cycle fatigue life of steel reinforcing bars using artificial neural network. *Journal of the Franklin Institute, 348,*1393-1403.

Ji, D. M. (2011). Study on the fatigue life of P91 steel creep based on support vector machine. *Pressure Vessel, 28*(15), 15-21.

Li, J., Ping, A., & Wang, D. J. (2001). The intelligent simulation method of the reliability analysis based on the fatigue damage process with non steady state. *Chinese Journal of Mechanical Engineering, 37*(1), 1-5.

Lotfia, B., & Beissb, P. (2013). Application of neural networking for fatigue limit prediction of powder metallurgy steel parts. *Materials & Design, 50,* 440-445.

Mathew, M. D., Kim, D., & Ryu, W.-S. (2008). A Neural network model to predict low cycle fatigue life of nitrogen-alloyed 316L stainless steel. *Material Science and Engineer A 474,* 247-253 .

Wang, J., Jing, Q. B., & Cao, L. L. (2007). Support vector regression facing multi input and multi output system. *Journal of Tsinghua University, 47*(S2), 1737-1741.

Zuo, J. W. (2011). T*he application of MATLAB in mathematical modeling* (pp.37-82). China: Beijing University of Aeronautics and Astronautics Press.