PSO_LSSVM Prediction Model and Its MATLAB Implementation

Lili Chen¹, Lei Duan², Yanjun Shi³,⁴ and Chenwei Du³

¹ College of Geography and Environmental Engineering, Lanzhou City University, Lanzhou 730000, China
² School of Earth Science and Engineering, Sun Yan-Sen University, Guangzhou 510275, China
³ School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China
⁴ Correspondence should be addressed to Y.S. (email: 123379728@qq.com)

Abstract. The particle swarm optimization (PSO) is added to a least squares support vector machine (LSSVM) prediction model, to achieve an objective optimization of parameters, thus globally optimizing the prediction model and improving the prediction accuracy of the model. According to the experimental results, the estimated values of the model are highly consistent with the actual values, which prove the validity of the prediction via the PSO_LSSVM model.

1. Introduction

Support Vector Machine (called SVM for short) was first stated by Cortes and Vapnik in 1995. It uses the kernel function and follows the Mercer’s theorem to extract features from the original space and map the samples in the original space to a vector in the high-dimensional feature space, to solve the problem of linear indivisibility in the original space. It can solve practical problems such as small samples, non-linearity, high-dimensional problems and local minimum and has a great capacity of generalization. However, it has the constrained quadratic programming problem, with a high complexity of calculation.

The Least Squares Support Vector Machines (LSSVM) proposed by Suykens J.A.K is mainly used to solve problems of pattern classification and function estimation. It uses the least squares linear system as the loss function instead of the traditional quadratic programming method. Therefore, the LSSVM has an obviously faster operating rate, so it is widely applied in function estimation and approximation.

Particle swarm optimization (PSO) algorithm is a new stochastic optimization algorithm based on swarm intelligence. It was first stated by Kennedy and Eberhart, with the basic idea originating from artificial life and evolutionary computation. According to the PSO, the optimal solution search in the complex space is completed through the collaboration and competition among individuals. The PSO algorithm is simple, easy-to-implement and fast in convergence and has a great capacity of global optimization. At present, it has been widely applied in functional optimization, pattern classification, neural network training and other fields.

In this paper, first of all, the LSSVM model and the corresponding MATLAB program are introduced. Secondly, the PSO method is used in the LSSVM prediction model parameter optimization, to build a PSO_LSSVM prediction model. Meanwhile, the corresponding MATLAB program is given. Lastly, the model is applied to predicting the extreme drought event of Xinjiang. The results show that the model can significantly improve the accuracy of the prediction.
2. LSSVM Model

With the given data set \((x_i, y_i)(i=1,2,...,n)\), the LSSVM uses the function \(y = \omega^T \varphi(x_i) + b\) to estimate the input and output. Where \(x_i\) is the \(i^{th}\) input of the \(m\) dimension; \(y_i\) is the \(i^{th}\) first real-valued output; \(n\) is the number of samples; \(\varphi\) is the kernel space mapping function; \(\omega\) is the weight vector; and \(b\) is the amount of deviation.

Find the optimal hyperplane, that is, the minimum value of \(J = \frac{1}{2} \| \omega \|^2 + \frac{1}{2} \sum_{i=1}^{n} e_i^2\). Select the appropriate initial value of the tunable parameter (marginal parameter) and the appropriate kernel function \(K\). The function estimation problem in the original space is described as the solution of the following problem:

\[
\min J (\omega, e) = \frac{1}{2} \| \omega \|^2 + \frac{1}{2} \gamma \sum_{i=1}^{n} e_i^2
\]

(1)

\[
y_i = \omega^T \varphi(x_i) + b + e_i, i = 1,...,n
\]

(2)

Where, \(e_i \in R\) is the error variable.

To get the extreme values of the objective function (1) under Condition (2), the Lagrangian function is defined as:

\[
L (\omega, b, \alpha) = J (\omega, e) - \sum_{i=1}^{n} \alpha_i [\omega^T \varphi(x_i) + b + e_i - y_i]
\]

(3)

Where, \(\alpha_i \in R\) is the Lagrangian coefficient.

In the process of finding solutions of the conditional extreme values, the first is to use the above formula to seek the partial derivatives of \(\omega, b, e, \alpha_i\) and make them zero, to get an equation. The second step is to use the equation to find solutions of \(b, \alpha_i\) and the support vector=\(x_i\). At the last step, the LSSVM regression function is constructed as:

\[
y(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b
\]

(4)

In Formula (4), where \(y(x)\) is the object of prediction; \(x_i\) is the support vector obtained via training; \(x\) is the prediction sample; \(\alpha_i\) and \(b\) are the Lagrangian coefficient and the deviation amount solved through training; \(K(x, x_i)\) is the kernel function.

Different LCSVM models can be built by selecting different kernel functions. At present, the commonly-used kernel functions include the RBF kernel function (5), polynomial kernel function (6) and Sigmoid kernel function (7).

\[
K(x, x_i) = \exp(-\sigma \| x - x_i \|^2)
\]

(5)

\[
K(x, x_i) = (x_i + x)^d, d = 1,2,...
\]

(6)

\[
K(x, x_i) = \tanh(\beta(x_i) + C)
\]

(7)

The SVM algorithm was designed originally for solving the binary classification problem. However, in practical applications, the multiple-class classification is a common problem, so it is necessary to building a suitable multiple-class classifier. Common methods of building SVM multi-class classifiers include the direct one and indirect one. The direct one is to make multi-class classifications by solving the optimization problem corresponding to the parameters of multiple-class classification planes. This method seems simple, but is not easy-to-implement in practice, due to its high computational complexity. Therefore, the indirect method is frequently used in practice. The
indirect method is to build multiple-class classifiers by combining various two-class classifiers. The common coding schemes include the One versus All Coding (OneVsAll), One Versus One Coding (OneVsOne), Error Correcting Output Coding (ECOC) and Minimum Output Coding (MOC).

The LSSVM prediction model is build through the following process. First of all, normalize the experimental data, and divide the experimental data into training data and prediction data. Secondly, on the basis of the training data, the specified kernel function and the multi-class coding scheme, the crossover algorithm is used to seek the least squares support vector machine parameters, with given initial parameters. Thirdly, use the least squares support vector machine parameters to construct the prediction model. Lastly, use the test set to test the model (Figure 1).

![Flow of LSSVM Modeling](image)

**Figure 1.** Flow of LSSVM Modeling

Owing to its open source codes, easy-to-use features and other advantages, MATLAB’s LSSVM toolbox has been widely used. In this toolbox, the selection of the kernel function and the determination of related parameters are a quite important link. Due to various advantages of the Gaussian radial basis kernel function, many scholars select the Gaussian radial basis function as a kernel function in the LMASVM toolbox of MATLAB. The calculation process involves 2 parameters: $\gamma$ and $\sigma$. $\gamma$ is a regularization parameter, which is used to control the complexity of the model and amount of deviation, while $\sigma$ is the kernel parameter, responsible for adjusting the smoothness of the kernel function. These two parameters, to a great extent, determine the learning ability and generalization ability of the model.

### 3. PSO_LSSVM Model

The LSSVM model is widely applied. However, due to the difficulty in identifying its parameters and its poor robustness and low prediction accuracy, scholars have put forward many methods to optimize and improve the LSSVM model. An intelligent optimization algorithm called the Particle Swarm Optimization (PSO) is one of these methods. Owing to its fast convergence speed, great global optimization ability and high prediction accuracy, especially the penalty factor and nuclear parameters of the intelligent optimization LSSVM model, the PSO_LSSVM model has earned the favor of many
scholars.

![PSO_LSSVM Parameter Optimization Flowchart](image)

**Figure 2. PSO_LSSVM Parameter Optimization Flowchart**

The PSO algorithm arises from research on the bird predation behavior. In the predation process, the easiest and most effective way for each bird to find food is to search for the area around the bird closest to the food. The first step of this algorithm is to initialize a group of random particles and search for the optimal solution through iterations. In each iteration, the particle updates itself by tracking 2 “extreme values”: the local optimum \( p_{best} \) and global optimum \( g_{best} \). The local optimum is the optimal position that a particle once went through, while the global optimum is the best one of the optimal positions that all particles in a swarm have had. The particles update their velocity and position based on the above 2 optimums. When these 2 optimums are found, the particles update their velocity and position according to Equation (8) and Equation (9).

\[
\begin{align*}
\dot{v} &= \omega \times v + C_1 \times \text{Rand}() \times (p_{best} - x) + C_2 \times \text{Rand}() \times (g_{best} - x) \\
\dot{x} &= x + v
\end{align*}
\]

Where, \( \omega \) is the coefficient of elasticity; \( V \) is the velocity of the particle; \( X \) is the current position of the particle; \( \text{Rand}() \) is a random number between 0 and 1; \( C_1 \) and \( C_2 \) are the learning factors, which are 1.5, in most cases.

The core of the PSO algorithm is the intelligent parameter optimization. To be specific, the LSSVM and the PSO algorithm are based to select the global optimal LSSVM parameter. After that, the optimized least squares support vector machine parameter is used to build a LSSVM model with good performance. Lastly, the model is tested using the test set. The PSO parameter optimization process is shown in Figure 2.

4. **Case Analysis and Conclusions**

In this paper, the circulation index of a meteorological observatory in Xinjiang in the period from 1962 to 1997 was used as a predictor to predict the drought level of the meteorological station. The first 36 sets of data (1962-1997) were used as training data, and the last 15 sets of data (1998-2012) were used...
as the prediction data. And the LSSVM (hereinafter referred to as Scheme I) and the PSO_LSSVM (hereinafter referred to as Scheme II) were used to predict the drought levels of the meteorological observatory respectively in spring, summer, autumn and winter. Moreover, the average relative error and prediction accuracy were based and used to analyze and evaluate the prediction results.

Table 1. LSSVM and PSO_LSSVM Model Prediction Results

|             | Spring |       | Summer |       | Autumn |       | Winter |       |
|-------------|--------|-------|--------|-------|--------|-------|--------|-------|
| Scheme I    | γ      | σ     | γ      | σ     | γ      | σ     | γ      | σ     |
|             | 14.9   | 16.4  | 51.6   | 1.0E-02| 1744.4 | 1.5E-05| 20.0   | 1.0E-02|
| Scheme II   | γ      | σ     | γ      | σ     | γ      | σ     | γ      | σ     |
|             | 8.9    | 10.8  | 114.6  | 1.0E-02| 20.0   | 4.9E-02| 20.0   | 1.0E-02|

As can be seen from Table 1, Scheme II is superior to Scheme I in the back-substitution process, in terms of the prediction error and the prediction accuracy. However, Scheme II has no obvious advantage, with respect to the prediction error. Moreover, compared with Scheme I, Scheme II has a significantly higher accuracy in predicting the drought level of 3 seasons including spring, summer and winter. Overall, introducing the PSO algorithm in LSSVM can significantly improve the prediction results of LSSVM.

5. Appendix

5.1. Appendix A

The MATLAB program for the LSSVM model is:

```matlab
[X,Xt]=scaleForSVM(xn_train',xn_test',0,1);
Y=dn_train';
Yt=dn_test';
igam=100;
isig2=0.1;
type='c';
kernel='RBF_kernel';
preprocess='preprocess';
codefct='code_MOC';
[Yc,codebook,old_codebook]=code(Y, codefct);
[gam,sig2]=tunelssvm({X,Yc,type,igam,isig2,kernel,preprocess},[],'gridsearch',{},'crossvalidate',{X,Yc,10,'misclass');
[alpha,b]=trainlssvm({X,Yc,type,gam,sig2,kernel,preprocess},{});
Yd10=simlssvm({X,Yc,type,gam,sig2,kernel,preprocess},[alpha,b],X);
Yd101=code(Yd10,old_codebook,[]);
Result10=1-abs(Yd101-Y);
Percent101=sum(Result10==1)/length(Result10);
Yd11=simlssvm({X,Yc,type,gam,sig2,kernel,preprocess},[alpha,b],Xt);
Yd111=code(Yd11,old_codebook,[]);```
5.2. Appendix B

The MATLAB program for the PSO_LSSVM model is:

c1=1.5; 
c2=1.7; 
maxgen=300; 
sizepop=30; 
popcmax=10^(3); 
popcmin=10^(-1); 
popgmax=10^(3); 
popgmin=10^(-2); 
k =0.5; 
Vcmax =k*popcmax; 
Vcmin=-Vcmax; 
Vgmax=k*popgmax; 
Vgmin=-Vgmax ; 
eps =10^(-7); 
type='function estimation'; 
kernel='RBF_kernel'; 
proprecess='original'; 
for i=1:sizepop 
pop(i,1)=(popcmax-popcmin)*rand(1,1)+popcmin; 
pop(i,2)=(popgmax-popgmin)*rand(1,1)+popgmin; 
V(i,1)=Vcmax*rands(1,1); 
V(i,2)=Vgmax*rands(1,1); 
gam=pop(i,1); 
sig2=pop(i,2); 
model=initlssvm(train_x,train_y,type,gam,sig2,kernel,proprecess); 
model=trainlssvm(model); 
[ptrain,zt,model]=simlssvm(model,train_x); 
trainmse=sum((ptrain-train_y).^2)/length(train_y); 
fitness(i)=trainmse; 
end 
[global_fitness bestindex]=min(fitness); 
local_fitness=fitness; 
global_x=pop(bestindex,:); 
local_x=pop; 
avgfitness_gen=zeros(1,maxgen); 
tic 
for i=1:maxgen 
for j=1:sizepop 
wV=1; 
V(j,:)=wV*V(j,:)+c1*rand*(local_x(j,:)-pop(j,:))+c2*rand*(global_x- pop(j,:)); 
if V(j,1)>Vcmax 
V(j,1)=Vcmax; 
end 
if V(j,1)<Vcmin 
V(j,1)=Vcmin; 
end 
if V(j,2)>Vgmax 
V(j,2)=Vgmax; 
end 
if V(j,2)<gmin 
V(j,2)=gmin; 
end
end

wP = 1;
pop(j,:) = pop(j,:) + wP*V(j,:);
if pop(j,1) > popcmax
    pop(j,1) = popcmax;
end
if pop(j,1) < popcmin
    pop(j,1) = popcmin;
end
if pop(j,2) > popgmax
    pop(j,2) = popgmax;
end
if pop(j,2) < popgmin
    pop(j,2) = popgmin;
end
if rand > 0.5
    k = ceil(2*rand);
    if k == 1
        pop(j,k) = (20-1)*rand+1;
    end
    if k == 2
        pop(j,k) = (popgmax-popgmin)*rand+popgmin;
    end
end
gam = pop(j,1);
sig2 = pop(j,2);
model = initlssvm(train_x,train_y,type,gam,sig2,kernel,proprecess);
model = trainlssvm(model);
[ptrain,zt,model] = simlssvm(model,train_x);
trainmse = sum((ptrain-train_y).^2)/length(train_y);
fitness(j) = trainmse;
end
if fitness(j) < local_fitness(j)
    local_x(j,:) = pop(j,:);
    local_fitness(j) = fitness(j);
end
if fitness(j) == local_fitness(j) and pop(j,1) < local_x(j,1)
    local_x(j,:) = pop(j,:);
    local_fitness(j) = fitness(j);
end
if fitness(j) < global_fitness
    global_x = pop(j,:);
    global_fitness = fitness(j);
end
if abs(fitness(j)-global_fitness) <= eps and pop(j,1) < global_x(1)
    global_x = pop(j,:);
    global_fitness = fitness(j);
end
end
fit_gen(i) = global_fitness;
avgfitness_gen(i) = sum(fitness)/sizepop;
end
toc
6. References

[1] Cortes C, Vapnik V 1995 Support vector networks J. Machine Learning 20 273-297
[2] Lu C, Li W and Liu W 2009 Model of prediction of flow rate of alumina powder based on PSO_LSSVM J. Lanzhou University of Technology 35(3) 80-84
[3] Suykens J A K, Vandewalle J. 1999 Least squares support vector machine classifiers J. Neural Processing Letters 9(3) 293-300
[4] Zhu Y, Wu J, Li Qi and Mao Z 2005 Building model with hybrid kernel function in MISO system J. Control and Decision 20(4) 417-420
[5] Ding X 2007 LSSVM-based online identification for T-S model J. Information and Control 36(4) 451-454
[6] Kennedy J and Eberhart R 1995 Proc of IEEE Int. Conf. on Neural Networks pp 1942-48
[7] Liu R, Mou S, Su H 2006 Soft-sensor modeling based on support vector machine and particle swarm optimization algorithm J. Control Theory & Applications 23(6) 895-99
[8] Li Y, You F, Huang Q and Xu J 2010 Least squares support vector machine model of multivariate predication of stream flow J. Journal of Hydroelectric Engineering 29(3) 28-33
[9] Kaelbling L P, Littman M L, Moore A W 1996 Reinforcement learning: A survey J. Journal of Artificial Intelligence 4 237-285
[10] Sutton R S and Barto A G 1998 Reinforcement Learning (Cambridge, MA: The MIT Press)
[11] Yin H 2010 Gas time series analysis method and its early warning application in coal mine D. Jiangsu: China University of Mining and Technology 123-125
[12] Zhao D, Li C and Lan S 2013 Time-delay control algorithm for structural seismic response based on LSSVM J. Journal of Vibration and Shock 32(9) 165-172
[13] Lv Y, Wang Y Lu X and Lu J 2009 GPS elevation fitting support vector machine model J. Gns World of China 3 11-13
[14] Zhang W, Wu, Li X and Hu X 2001 Dam stress prediction model based on least squares support vector machine J. Journal of Water Resources and Architectural Engineering 9(1) 26-29
[15] Xu H and Yang L 2008 Prediction foundation pit deformation based on least square support vector machine regression J. of Nanjing Tech. University (Natural Science Edition) 32(2) 51-58
[16] Jiang S, Gui W, Yang C and Dai X 2009 Fault diagnosis method based on rough set and LS-SVM and its application J. Journal of Central South University (Science and Technology) 40(2) 447 -451
[17] Tang Y, Wang H and Wang W 2006 Prediction of traffic flow at intersection based on Gaussian Radial Basis Function Neural Network J. Agricultural Equipment & Vehicle Engineering 3 41-43
[18] Xu W, Liu G and Liu H 2008 GA-based parameter optimization of Gauss kernel function J. Electric Power Automation Equipment 28(6) 52-55
[19] Ai Q, Lu Y and Liu Q 2013 Recognition of sEMG based on reconstructed feature by Gaussian radial basic function Computer Engineering and Applications 49(12) 182-222
[20] Liu K, Zhang R, Xu H, Min J and Zhu W 2007 Error correction of the area index of subtropical high in the T106 numerical prediction based on support vector machine-Kalman filter model Acta Meteorologica Sinica 65(3) 450-57
[21] Liu K, Zhang R, Hong M, Yu D and Wang Z 2009 Subtropical high forecast model of least squares sport vector machine Journal of Applied Meteorological Science 20(3) 354-59
[22] Liu K, Zhang R, Yu P, Wang Y and Yu D 2007 Western pacific subtropical high forecast model based on wavelet decomposition and least squares support vector machine Journal of Tropical Meteorology 23(5) 491-96
[23] Wang H, Zhang R, Liu W and Wang G 2001 Kriging interpolation method optimized by support vector machine and its application in oceanic data. Transactions of Atmospheric Sciences 34(5) 567-73
[24] Xu S and Long W 2012 Parameters selection for LSSVM based on differential evolution to mid-long term runoff prediction Science Technology and Engineering 12(27) 6955-59
[25] Zhu X 2012 Power short-term load forecasting based on SA-LSSVM Science Technology and Engineering 12(24) 6171-74
[26] Ma J 2009 Application of GA_LSSVM model in the prediction of strength of fly ash concrete Ningxia Engineering Technology 8(4) 325-27
[27] Tang C, Yang C, Gui W and Zhu H 2010 KPCA and LSSVM model-based slag basicity prediction for siliconmanganese smelting process Chinese Journal of Scientific Instrument 31(3) 689-93
[28] Chen G and Guo R 2010 Application of PSO_LSSVM classification model in logging lithology recognition Journal of Xi’an Shiyou University 25(1) 96-99
[29] Xu B 2011 Research costs forecasting model of military engineering machinery based on improved PSO_LSSVM Ordnance Industry Automation 30(10) 43-45
[30] Sun B and Yao H 2012 The short-term wind speed forecast analysis based on PSO_LSSVM prediction model Power System Protection and Control 40(5) 85-89
[31] Long W, Liang X, Long Z and Yan G 2013 Application of PSO_LSSVM grey combination model in groundwater depth prediction Systems Engineering - Theory and Practice 33(1) 243-48
[32] Chen G. 2010 Aircraft maintenance support cost prediction model based on PSO_LSSVM Computer Applications and Software 27(9) 181-183
[33] Li W, Shi L and Liang C 2008 Forecasting model of research octane number based on PSO_LSSVM Control and Instruments in Chemical Industry 35(2) 25-27
[34] Xu L, Zhang F and Cheng J 2007 Research on fault diagnosis method based on PSO neural network Computer Engineering and Design 28(15) 3640-43