

Soft measurement of the cell concentration based on SVM and PSO

Hua Meng^{1, a *}, Hui Gao^{2, b} and Liting Han^{3, c}

^{1,2,3} Department of Electrical Engineering, Hebei University of Science and Technology, 26 Yuxiang Street, Shijiazhuang, 050018, P R China

^amenghua0311@126.com, ^bgaohuigogh@126.com, ^c18732182156@163.com

Keywords: Soft sensor ;SVM; cell concentration; PSO; fermentation

Abstract. The process of *Pichia pastoris* fermentation has a long period and less offline data . The cell concentration and some other important variables can not be measured on line. The soft sensor modeling at present is mainly the artificial neural network (ANN). This paper introduces the Support Vector Machine (SVM). Selected fewer off line data and established soft sensor modeling about cell concentration. To address the difficulty of parameters selection in SVM, parameters was optimized by using Particle Swarm Optimization (PSO) , Simulation experiment proves that the support vector machine has better prediction effect and generalization ability and PSO can quickly find out the best parameters of SVM..

Introduction

Microbial fermentation process is a common and complex reaction process, It has many characteristics such as highly nonlinear, time-varying and uncertainty. The fermentation process has plenty of influence factors and complex mechanism, and classical model equation has considerable difficult in describing the important biomass parameters[1] such as the cell concentration , it is the key parameter of the fermentation process, has a great impact on the cost of production and the quality of fermented products. But the current sensor technology can't meet the requirements of online measurement. So building the relationship between the key parameter and the measurable variables by offline data, and prediction cell concentration online, it is the focus of the present study.

Soft measurement method is the most widely used method ,Artificial neural networks(ANN) has been widely applied to the identification and control of nonlinear dynamic systems[2,3]. It is based on the empirical risk minimization principle, Therefore, it has some drawbacks such as over learning and local minimum[4], so the model generalization ability is bad, so as to affect the prediction accuracy. The SVM based on the VC dimension theory and the structure risk minimum principle[5,6]. Choose a group of feature subset from a training set, make the linearly separable problems of feature subset equivalent to divided the whole data set. Under the condition of small samples, obtaining the optimal solution under the existing information and not just the optimal value of the sample tends to infinity. Solve the problem that the general method of study is difficult to solve. Such as the local minimum problem, over-learning, over-dependent on the experience about the choice of the structure and the type, etc. The other machine learning problems are comed down to a quadratic programming problem by SVM, and the optimal solution. In recent years, SVM is used to the soft measurement of bacteria concentration in other fermentation[7]. However, the selection of SVM parameter is difficult, the range of the parameter is very large. According to different objects, the parameter selection range is different. The commonly used methods are the trial and error method, the grid search method, and intelligent heuristic algorithm such as GA[8].

The SVM theory

The SVM regression problem becomes the following minimization problem:

$$\min_{w,b,x} J = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (x_i + x_i^*) \quad (1)$$

$$s.t. \begin{cases} y_i - w \square x_i - b \leq e + \mathbf{x}_i \\ w \square x_i + b - y_i \leq e + \mathbf{x}_i^* \\ \mathbf{x}_i \geq 0 \\ \mathbf{x}_i^* \geq 0 \end{cases} \quad i = 1, 2, \dots, l$$

Introduce the type by the dual principle.

$$\max_{\mathbf{a}, \mathbf{a}^*} \left\{ L_D = -\frac{1}{2} \sum_{i,j=1}^l (\mathbf{a}_i - \mathbf{a}_i^*)(\mathbf{a}_j - \mathbf{a}_j^*) x_i \square x_j - e \sum_{i=1}^l (\mathbf{a}_i + \mathbf{a}_i^*) + \sum_{i=1}^l y_i (\mathbf{a}_i - \mathbf{a}_i^*) \right\} \quad (2)$$

$$\begin{cases} \sum_{i=1}^l (\mathbf{a}_i - \mathbf{a}_i^*) = 0 \\ 0 \leq \mathbf{a}_i \leq C \\ 0 \leq \mathbf{a}_j \leq C \end{cases}$$

For the nonlinear regression problem, it needs to introduce the concept of kernel function. Via a nonlinear function $f(x)$ to map the training data to a higher dimensional linear feature space. And the kernel function is satisfied $K(x_i, x_j) = f(x_i) \square f(x_j)$. Don't need to explicitly know and the feature space. As long as choosing the appropriate kernel function, can determine the inner product in the high dimension space, thus, we avoid dimension disaster caused by high dimensional feature space. So the nonlinear regression problem can make use of dual principle into the following quadratic programming problem:

$$\max_{\mathbf{a}, \mathbf{a}^*} \left\{ L_D = -\frac{1}{2} \sum_{i,j=1}^l (\mathbf{a}_i - \mathbf{a}_i^*)(\mathbf{a}_j - \mathbf{a}_j^*) K(x_i, x_j) - e \sum_{i=1}^l (\mathbf{a}_i - \mathbf{a}_i^*) + \sum_{i=1}^l y_i (\mathbf{a}_i - \mathbf{a}_i^*) \right\} \quad (3)$$

Obtain the Lagrange multiplier and by the above formula. The samples which make and not zero are the support vectors. The nonlinear regression function can be obtained by the Lagrange multiplier known:

$$f(x) = \sum_{x_i \in SV} (\mathbf{a}_i - \mathbf{a}_i^*) K(x_i, x) + b \quad (4)$$

The PSO theory

PSO algorithm comes from the study of the birds feeding behavior which is one of the swarm intelligence optimization algorithm[9]. Unlike GA, PSO has no evolution operators such crossover and mutation. There are only few parameters to adjust in PSO. Each particle in PSO represents a potential solution of the problem, and every particle has a fitness value which is the merits of the particles. At the same time, Particles has two features including the speed and position. By comparing the fitness of the new particles and the fitness of the individual and the group extreme value, the particles update the location information and calculate a fitness by the following formula.

$$V_{id}^{k+1} = wV_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k) \quad (5)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (6)$$

Where w is the inertia weight, V_{id} is the particle velocity, X_{id} is the current particle(solution), P_{id} and P_{gd} are pbest and gbest, r_1 and r_2 are the random numbers between (0,1) and c_1 and c_2 are the learning factors.

The PSO implementation of the algorithm has been refined over the years and many variants created. In this paper, the standard PSO has been used. It contemplates some improvements in the implementation[10] and the PSO parameters are set to the values: $c_1 = 1.5$, $c_2 = 1.7$.

Simulation

In this article the simulation data for the modeling is a full reflection of the fermentation process. It includes dissolved oxygen concentration (%), stirring power (W), PH value and CO₂ concentration (mol/L), temperature (K), and cell concentration (g/L). The only inputs to the full simulation were the time in the moment, temperature, pH, flow rate of the gas, dissolved oxygen, speed of agitator, and feeding carbon source and nitrogen source rate. The cell concentration in this moment was as the output variable. Collected a total of 10 batches of data. Selected 7 batches of them as the available data, selected 6 group as the training data, the remaining batch as test data. The parameters of SVM include the kernel function g , penalty coefficient C , insensitive loss function coefficient e . Trial and error method requires a lot of effort and not easy to find the optimal parameters. The grid search method need to traverse the entire search space, although can find the optimal parameters, it is time-consuming. choose the insensitive loss function coefficient e is 0.01. The other two parameters of SVM was optimized by PSO. In the process of optimization, the population size is 20, the evolution algebra is 100. The simulation results are as follows in *Fig. 1*. At the same time use the SVM to predict the concentration of bacteria, In this paper, 5-fold cross-validation was used. Compared the predicted values and real values. The simulation results are as follows in *Fig. 2*.

Conclusions

It can be seen that PSO can well find out the best parameters of SVM And only about 30 generations of evolution. besides, the prediction ability of SVM using the best parameters by PSO is good can be seen from simulation diagram and calculating the MSE (mean square error) of the prediction error (SVM is 0.1238). the method of SVM-PSO is suitable for soft measurement of the bacteria concentration in fermentation.

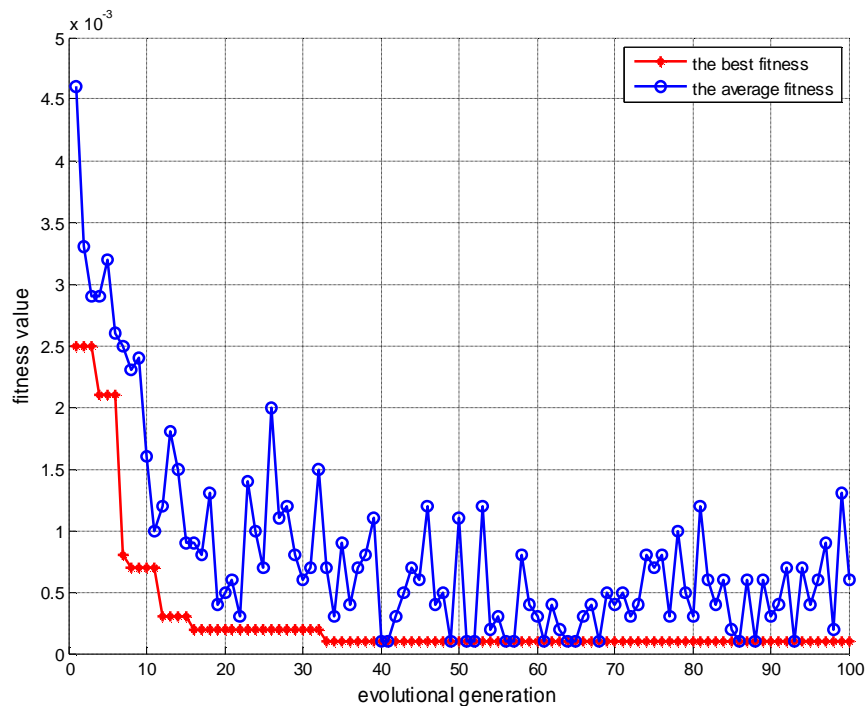


Fig.1. The process of using PSO to optimize the parameters of SVM

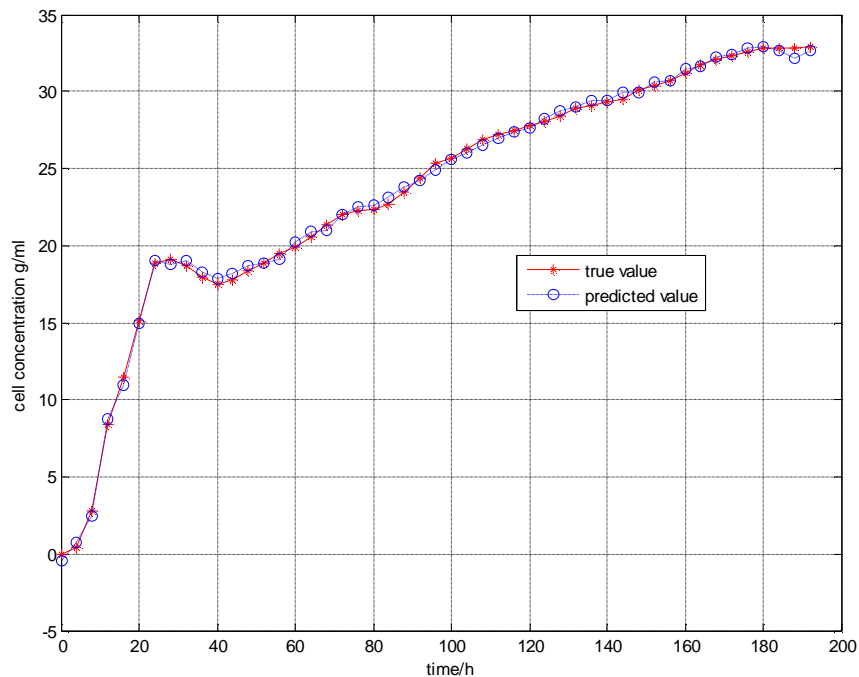


Fig.2. Estimation result of biomass based on SVM

References

- [1] H.F. Sang, F.L Wang, D.K. He. On-line estimate of biomass concentration in fermentation process[J]. Journal of Northeastern University: Natural Science, 2006, 27(6): p602-605. In Chinese
- [2] S.L. Daniel: Dynamic hybrid neural network model of an industrial fed-batch fermentation process to produce foreign protein. Computers and Chemical Engineering. Vol. 31(2007), p163-170.
- [3] M.H.Qiang: Based on RBF neural network in the fermentation process of bacteria concentration of soft measurement.Automation & Instrumentation. Vol. 160(2012),p178-181.
- [4] X.D. Yan, W. Yang: Soft sensor for Ammonia concentration at the Ammonia converter outlet based on an improved Group Search Optimization and BP Neural Network. Chinese Journal of Chemical Engineering. Vol. 20(2012), p. 1184-1190.
- [5] C. Cortes, V. Vapnik. Support-vector networks[J]. Machine Learning, 1995, 20(1): p 273-297.
- [6] V. Vapnik. The nature of statistical learning theory[M]. First Edition. New York: Springer-Verlag, 1995.79-87.
- [7] G.H. Liu. D. W. Zhou Xu H X, Mei C I. Model optimization of a fermentation soft sensor[J]. Expert Systems with Application, 2010, 37(4); 2708-2713.
- [8] E, Pourbasheer, S, Riahi, M.R. Ganjali, Parviz Norouzi. Application of genetic algorithm-support vector machine (GA-SVM) for prediction of BK-channels activity[J]. European Journal of Medicinal Chemistry, 2009,44(12); 5023-5028.
- [9] Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of the IEEE international conference on neural networks IV. Piscataway,NJ,USA: IEEE Press;1995.p.1942–1948.
- [10]Olsson AE. Particle swarm optimization: theory, techniques and applications. New York, NY: Nova Science Pub Inc; 2011.