

The Application Based on Decision Tree SVM for Multi-class Classification

Hou Huifang^{1, a}, Han Ping^{1, b} and Cao Dan^{1, c}

¹College of Information Science and Engineering, Henan University of Technology, Zhengzhou Henan 450001, China

^ahouhf72@163.com, ^bmuziq@sina.com, ^chautcd@126.com

Keywords: Support vector machines (SVM), decision tree, classification, multi-class.

Abstract. Wastewater treatment process is a multi-class classification problem. While the traditional Support Vector Machines only deal with the binary classification. The traditional SVM has difficulty in solving the multi-class problems like wastewater treatment process. This paper presents a method by introducing the idea of decision binary tree to support vector machine for solving the assessment problem of wastewater water quality. This method distributes classifier to each nodes of decision binary tree. Simulated experiments show the method is effective.

Introduction

Wastewater treatment process is a biochemical process which contains a number of complex processes, involving many of the mechanical equipments, electrical equipments, automation instruments and so on. At the same time the wastewater treatment process is in the unstable state, due to the sharp fluctuations in inflow feed, composition and concentration. Therefore how to monitor the wastewater treatment process, namely how to monitor automatically the operation state of the process by the measurable information of each processing stage, is the key of operation and management wastewater factory. There are some kinds of wastewater, so monitor the operational status of the wastewater treatment process is essentially a multi-class classification problem.

Support Vector Machines (SVM) has good classification accuracy and is designed originally as a binary classifier. That is SVM cannot be directly applied to wastewater treatment process which is a multi-class classification problem. How to extend effectively binary classifier to multi-class classification is still an ongoing research issue [1].

The paper applies a new technique in multi-class SVM to wastewater treatment process, which has the good application prospects.

Multi-class Support vector machine

From the Pattern Recognition viewpoint, Support Vector Machines (SVM), proposed by Vapnik, is a powerful method for machine learning developed on the basis of empirical risk minimization principle and statistical learning theory (SLT) [2, 3]. It solves effectively the problems of small-size limited samples, nonlinearity, and high dimensional pattern recognition, and has been successfully applied to various areas including fault diagnosis. SVM can obtain the global optimization value with higher accuracy by using the standard multiplier. The basic idea of SVM is briefly stated as follows: a hyperplane is used as a decision plane, which can correctly divide two-class samples and maximize the classification margin. That is, the optimal hyperplane problem will be converted into a convex quadratic programming (QP) problem.

However, it is difficult to apply SVM to the most real applications which are as to the multi-class classification, because the binary classifier has a structural limit. Currently, The main methods of multi-class classifiers include: the conventional methods 'one-against-all', 'one-against-one' and decision tree [4, 5].

Among these three methods, decision tree algorithm has good comprehensibility, accuracy rate is higher, training cost is smaller, learning cost is normal, the samples trained repeatedly are less. That

is decision tree which wins popularity due to simple structure and fast training is the better method in multi-class SVM.

The basic idea of decision tree is briefly stated as follows: Divide recognition task into series binary classification. The root node includes the entire sample data. Each inner nodes of the decision tree is a binary SVM classifier. Each leaf node is the class. The inner nodes can be divided continuously, while the leaf nodes can't be divided. The majority number separable classes should be separated at the upper nodes in the decision tree. M classes problem has M-1 classifiers [6, 7]. There are 2M-1 nodes and M leaf nodes in the multi-class binary tree.

Wastewater treatment process based on multi-class SVM decision tree

There are 200 samples in the whole data set. Each sample dimension is 3, there are 3 categories, normal state, normal state above the average value, abnormal state caused by storm. All samples are divided into two parts, 50 percent of samples for training, 50 percent samples for testing. The detail process is as the following:

Construct decision-binary tree.

There are three levels of state: Class 1: normal state; Class 2: normal state over average value; Class 3: abnormal state caused by storm. So two classifiers are constructed. SVM1 is used to discriminate Class 1, SVM2 is used to discriminate Class 2 and Class 3. Decision-binary tree is showed in Fig. 1.

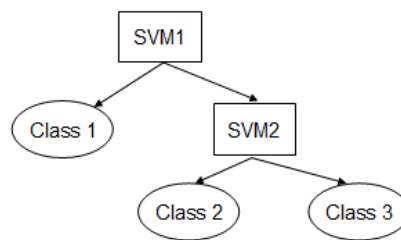


Fig. 1 Multi-class SVM-decision binary tree

Train samples.

Calculate the all kinds of classification function $f(x)$ by testing sample data. The algorithm is as the following:

1. $S = \{ C_1', C_2', C_3' \}$, where C_i' is the class. There are three class, the order of C_1', C_2', C_3' is the order of classification function $f(x)$;
2. $i=1$ to 2, $i++ //3$ class problem needs to construct $3-2=2$ classification function $f(x)$;
3. {
4. Select C_i' ;
5. Change the classification value of training sample C_i' to 1;
6. Set the other classification value of training sample to -1;
7. Apply the binary classification SVM to calculate the classification function of class C_i' .
8. $S = S - C_i'$
9. }

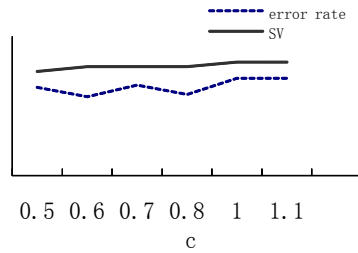
In order to obtain the classifier by using binary classification SVM method, the following steps should be carried out:

- (1) Select support vectors from many training sample data:

Method shows as following:

step1: select suitable kernel function; we use Sigmoid function:

$$K(x, x_i) = \tan(\delta(x \bullet x_i) / 54 + c) \quad (1)$$



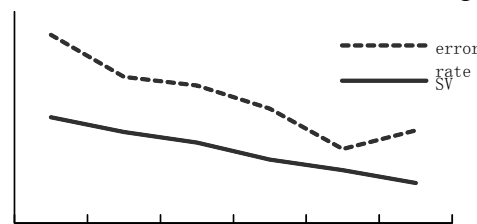
The experiment results show: the number of support vectors increases as soon as c increases.

step2: Determine the value of the penalty parameter C ;

By applying the change of penalty parameter C , we can get the error rate and the number of support vectors.

The experiment results show: When $C=100$, the error rate is minimum and equal to 1.1%.

The error rate is minimum when $C=100$ and $\delta=0.1$. The number of support vectors decreases sharply when C increases, while the error rate decreases first and begins to increase and steady.



step3: Calculate support vector, where the Lagrange multiplier is nonzero and less than C , by substituting sample and corresponding function value into equation.

$$\begin{aligned} \max Q(\alpha) &= \max \left[\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right] \\ &= \Lambda^T I - \frac{1}{2} \Lambda^T Z \Lambda \end{aligned} \quad (2)$$

$$\text{where } \Lambda = (\alpha_1, \alpha_2, \dots, \alpha_n)^T, Z_{ij} = y_i y_j K(x_i, x_j). \quad (3)$$

s.t

$$0 \leq \alpha_i \leq C \quad i=1, \dots, n \quad \text{where } \alpha_i \text{ is the Lagrange multiplier.}$$

$$\sum_{i=1}^n \alpha_i y_i = 0$$

Calculate the Lagrange equation:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i \{y_i (w \cdot x + b) - 1 + \xi_i\} - \sum_{i=1}^l \mu_i \xi_i \quad (4)$$

The support vectors are the sample corresponding the calculate Lagrange function and are input vectors with $\alpha_i > 0$.

(2) Write the optional equation and construct the SVM system

Methods: calculate w^* and b^* and obtain the optimal classification function $f(x)$.

$$w^* = \sum_{SV} \alpha_i^* y_i x_i \quad (5)$$

$$b^* = -\frac{1}{2} w^* \cdot (x_r + x_s) \quad (6)$$

x_r and x_s are any support vectors in the 2-class and their $\alpha_i^* > 0$

$$f(x) = \text{sgn}((w^* \bullet x) + b^*) = \text{sgn}\left(\sum_{SV} (\alpha_i^* y_i (x_i \bullet x) + b^*)\right) \quad (7)$$

We can judge the class of sample by the positive and negative of classification function $f(x)$, Fig. 2 shows the result of classification.

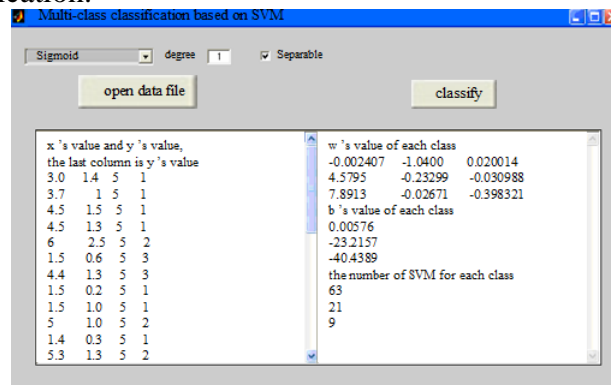


Fig. 2 Result of classification

Recognize the unknown samples.

Put the attribute of object into the classification function $f(x)$ and test its values from the decision tree root. Stop proceeding if the test value equals 1 which arrives at leaf node, which means the class of the example is the class of the leaf. While the test value equals -1, go along the branch and stop when it arrives at a certain leaf node.

The above results are the experiment and results analysis trained by training sample data, the final aim is to recognize the rest unknown sample by trained classification SVM. 50 percent of the rest sample data are experimented by the obtaining classification SVM. The experiment is showed in the Table 1.

Table 1 Recognition results

	Class 1	Class 2	Class 3
Number of test samples	50	20	30
Accuracy rate	87.7%	89.7%	93.6%
False accept rate(FAR)	4.2%	7.1%	4.3%
False reject rate(FRR)	8.1%	3.2%	2.1%
Total correct recognition rate	89.6%		

Conclusions

In this paper, we focus on applying SVM to multi-class classification problems. According to the multi-class character of SVM, this paper proposes a method based on SVM to solve the assessment of wastewater water quality by using decision tree.

Numerical simulations are conducted in order to testify the effectiveness of the proposed method. The simulation results indicate that the proposed method can achieve better performance with higher accuracy rate which is 89.6%. Therefore, it is proved that the proposed method as a multi-class classifier is of great significance on the theory, and also very practical in applications. It also has the high practical value.

References

- [1] C. Hsu, C. J. Lin. A comparison of methods for multi-class support vector machines. IEEE Trans. D Neural Networks, 2002. 13(2): 415-425.
- [2] V. N. Vapnik. An overview of statistical learning theory. IEEE Trans. on Neural Networks, 1999. 10(5): 88-99.
- [3] Ulrich K. Pairwise classification and support vector machines [C]. Advances in Kernel Methods Support vector Learning. Cambridge, MA: MIT Press, 1998: 255-268.

- [4] K. E Bennett, J. A. Blue. A support vector machine approach to decision tree[C]. Proc. of the IEEE World Congress on Computational Intelligence. 1998: 2396-2401.
- [5] Chang C C, Lin C J. LIBSVM: a library for support vector machines [J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): 1-27.
- [6] Platt John C, Probabilistic Output for Support Vector Machine and Comparisons to Regularized Likelihood Methods [M] . Advances in Large Margin Classifier, Cambridge: MIT Press, 2000.
- [7] Huang C, Davis L S, Townshend j R G. An assessment of support vector machines for land cover classification [J]. Int. J. Remote Sensing. 2002, 23(4): 725-749.