http://ansinet.com/itj

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (20): 5771-5774, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.5771.5774


# Printing Fault Classification Based on Hybrid Method

[1,2]Qi ya-Li and [2]Li ye-Li
[1]School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China
[2]The Computer Science and Technology Department, Beijing Institute of Graphic Communication, Beijing102600, China

**Abstract:** For the characteristics of printing malfunction diagnose system, a model to classify printing fault based on Incremental Reduced Support Vector Machine (IRSVM) and C4.5 is discussed. IRSVM is an improved method based on Support Vector Machine (SVM) which has been promising method to classify for its solid mathematical foundation. However it is not favored for large-scale, because the training complexity of SVM is highly dependent on the size of data set. This study uses IRSVM to classify root-classes, then uses C4.5 algorithm for further diagnosis to remedy the defect of IRSVM in classing subclasses. The hybrid method makes fully use of the IRSVM efficiency in multidimensional character space but it also brings the accuracy of C4.5 algorithm into full play. That is suited to class the complicated print faults. Computational results indicate the hybrid method has a good efficiency for adjustable printing fault and its computational times as well as its memory usage are much smaller than those of conventional SVM.

**Key words:** Incremental reduced support vector machines (IRSVM), C4.5, malfunction diagnose system, decision tree

## INTRODUCTION

The fault diagnosis expert system is to identify the reason of malfunction based on acquired information. It can understand the relation between entities and phenomena in a complex system. Diagnosis system usually applies to medical diagnosis, electronic mechanism, software fault diagnosis and so on. In this study, we use fault detection method in printing process. The study on printing malfunction diagnose system can improve the press performance in terms of better reliability, print quality and reduce operational costs.

The key of a fault diagnosis expert system is classifying the faults based on the acquired information. There are many methods to distinguish faults. The complexity of the method is often limited its application. In recent years, Artificial Intelligence (AI) is often used to classify which is based on Empirical Risk Minimization (ERM). But it easily causes over-learning because of the advanced demand for samples and the poor extensive performance of the model (Wang *et al*., 2004). Support Vector Machines (SVM) brought from Statistical Learning Theory which is based on the Structural Risk Minimization (SRM). It is a compromise between empirical risk and confidential scope. However, when the number of class samples gets larger, the efficiency of SVM gets poor (Lee and Mangasarian, 2001; Sabzekar and Naghibzadeh, 2013; Wu *et al*., 2006; Platt, 1999). By the inspiration from Lee *et al*. (2003) we use Incremental Reduced Support Vector Machines( IRSVM) to construct classifier which has a good efficiency for adjustable printing fault. Not only computational times but memory usages are much smaller for IRSVM than those of conventional SVM. But the IRSVM adapts to classify the root-class of printing fault, for subclass we need other method to classify. So in this study we use C4.5 algorithm for further diagnosis to remedy the defect of IRSVM in classing subclasses.

The rest of the study is organized as follows. Section 2 briefly introduces IRSVM. And in section3 describe the decision tree based on C4.5 to distinguish the subclasses. In section4 the experiments results are reported to classify printing fault based on the IRSVM and C4.5. Finally section5 concludes the study and proposes further research topics.

## INCREMENTAL RDEUCED SUPPORT VECTOR MACHINES (IRSVM)

**Reduced support vector machines (RSVM):** RSVM generates a nonlinear kernel-based separating surface that requires as little as 1 to 10% of a large dataset. To

**Corresponding Author:** Qi ya-Li, School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China

generate a nonlinear surface, the entire dataset is used as a constraint in an optimization problem with very few variables corresponding to the 10% of the data kept. The remainder of the data can be thrown away after solving the optimization problem.

Lee and Mangasarian (2001) defined the problem of classifying m points in the n-dimensional real space $R^n$, represented by the $m\times n$ matrix A, according to membership of each point $A_i$ in the classes +1 or -1 as specified by a given $m\times m$ diagonal matrix D with ones or minus ones along its diagonal. A column vector of ones of arbitrary dimension will be denoted by e. For $A\in R^{mn}$ and $B\in R^{n\times l}$, the kernel K(A, B) maps $R^{m\times n}\times R^{n\times l}$ into $R^{m\times l}$. In particular, if x and y are column vectors in $R^n$ then, K (x, y) is a real number, K(x, A) is a row vector in $R^m$ and K(A, A) is an $m\times m$ matrix. The base of the natural logarithm will be denoted by ε. The difference of two sets A and B is defined as $A/B = \{x|x\in A \text{ and } x\notin B\}$.

RSVM achieved by making use of a rectangular $m\times\bar{m}$ kernel K (A, $\bar{A}$) that greatly reduces the size of the quadratic program to be solved and simplifies the characterization of the nonlinear separating surface. Here, the m rows of A represent the original m data points while the $\bar{m}$ rows of $\bar{A}$ represent a greatly reduced $\bar{m}$ data points.

The RSVM is derived from the Generalized Support Vector Machine (GSVM) (Mangasarian, 2000) and smooth support vector machine(Lee and Mangasarian, 2001). It solves the following unconstrained minimization problem for an arbitrary rectangular kernel K(A, $\bar{A}$‘):

$$\min_{(\bar{u},\gamma)\in R^{\bar{m}+1}} \frac{v}{2}\|p(e-D(K(A,A')\bar{A}\bar{u}-e\gamma),\alpha\|_2^2+\frac{1}{2(\bar{u}\bar{u}+\gamma^2)} \tag{1}$$

where, p is a function with a smoothing parameter α and:

$$p(x,\alpha)=x+\frac{1}{\alpha}\log(1+\varepsilon^{-ax}),\ \alpha>0 \tag{2}$$

The positive tuning parameter v here controls the tradeoff between the classification error and the suppression of $(\bar{u}, \gamma)$. A solution of this minimization program for $\bar{u}$ and γ leads to the nonlinear separating surface:

$$K(x,\bar{A})\bar{D}\bar{u}=\gamma \tag{3}$$

The nonlinear classifier generated by RSVM is a linear combination of a set of kernel functions:

$$\{1,K(-,\bar{A}'_1),K(.,\bar{A}'2),\ldots K(.,\bar{A}'_{\bar{m}})\tag{4}$$

The size of this small random subset is pre-specified by users. The key point of IRSVM improved RSVM is using an incremental approach that begins with an extremely small reduced set and then sequentially expands the reduced set according to an information criterion.

## IRSVM

Kantardzic (2003) and Lee *et al.* (2003) propose a process that sequentially adding a kernel function into set (4) only when the function is un-similar to the current set and carrying sufficient extra information over the current set. Using the distance r which between K(A, A’$_i$) and K(A, $\bar{A}$‘) present the scale of un-similar. It start with a very small reduced set $\bar{A}$, typically a size of 2, then add a new data point $A_i$ into the reduced set only when the extra information carried in the vector K(A, A’$_i$) with respect to the column space of K (A, $\bar{A}$‘)is greater than a certain positive threshold. This can be achieved by solving a least squares problem. For $\bar{K} = K(A, \bar{A})\in R^{m\times m}$. The least squares problem which need to solve is:

$$\min_{\beta\in R^{\bar{m}}}\left\|\bar{K}\beta-K(A,\bar{A}'_i)\right\|_2^2 \tag{5}$$

In (5), $\beta\in R^{\bar{m}}$ is a free vector and $\bar{K}\beta\in R^{\bar{m}}$ is a linear combination of the function K(A,$\bar{A}_i$), i = 1,...., $\bar{m}$ that represents the column space of K(A, $\bar{A}$). The optimal solution β* of above problem (5) is equivalent to solving a system of normal equations:

$$\overline{(K'K\beta=K'K(A,A'_i)} \tag{6}$$

If the columns of the rectangular kernel matrix generated by the initial reduced set are linear independent, the least squares problem (5) has a unique solution β*:

$$\beta^*=(\bar{K}'\bar{K})^{-1}\bar{K}'\bar{K}\ (A,\ A'_i \tag{7}$$

So r is the squared root of the optional value of (5):

$$r=\left\|\bar{K}\beta-K(A,\bar{A}'_i)\right\|_2 \tag{8}$$

The r has the information from K (A, $A_i$) and K(A, $\bar{A}$). So we design a threshold which control when the reduced set has been selected completely. The algorithm is showed as following.

**IRSVM algorithm:** Select δ>0 as a given threshold:

- **S1:** Choose an initial reduced set $A0\in^{2\times n}$ which has 2 point from the training data matrix $A\in^{m\times n}$ generate the reduced kernel matrix K(A, $\bar{A}$) and $\bar{A}_{new} = \bar{A}_0$.

- **S2:** Select $_{A}\in A/\bar{A}_0$ and compute the distance $K(A, \bar{A}_j)$ from the kernel vector $K(A, \bar{A}_{new})$ to the column space of $K(A, \bar{A}_{new})$ using equatioon 8
- **S3:** If $r>\delta$, then $\bar{A}_{new} \cup A_j$
- **S4:** Repeat Step 2) until several successive failures happened in S3, then the resulting $(\bar{A}_{new})$ is our final reduced kernel
- **S5:** Apply the Newton-Armijo Algorithm (Lee and Mangasarian, 2001) to solve the objective function:

$$\min_{(\bar{u},\gamma)\in R^{\bar{m}+1}} \frac{\nu}{2}\left\|p(e - D(K(A, \bar{A}')\bar{D}\bar{u} - e\gamma), \alpha\right\|_2^2 + \frac{1}{2}(\bar{u}'\bar{u} + \gamma^2)(\bar{u}, \gamma) \tag{9}$$

the reduced kernel $K(A, \bar{A})$ in front function is which obtained in S4.

- **S6:** The separating surface is given as follows:

$$K(x', \bar{A}')\bar{D}\bar{u} = \gamma \tag{10}$$

where, $(\bar{u}, \gamma)\in R^{\bar{m}+1}$ is the unique solution to(9).

- **S7:** A new point $x\in R_n$ is classified into class +1 or -1 depending on the step function $(K(x', \bar{A}')\bar{D}\bar{u} - \gamma)$ is +1 or zero, respectively

## PRINTING FAULT DETECTION BASED ON IRSVM AND C4.5

**Printing fault detection based on IRSVM and C4.5:** The printing malfunction has some factors such as study, ink, water, block, printing machine and so on. Every factor can be defined as a class. Every class has its inner child factors and the child factor may have grandchild factors. Child and grandchild factor maybe have inter-action, so the printing malfunction system isn't simple arrangement. Using IRSVM only has a good accuracy for the root-classes but poor for subclasses. We need a method to elevate the efficiency when there are many kinds of reason for one output. Though the accuracy of the decision tree based on C4.5 is inferior to more advanced techniques like neural networks or SVM, C4.5 is a fast algorithm what we can use to classify the complexity fault classes. C4.5 algorithm doesn't rely on the distribution of attribution so it is more robust than other statistical learning methods. But because more classes will need high complexity decision tree to describe, we use IRSVM to classify the high-layer faults what distinguish the fault sort and use C4.5 to classify the low-layer faults which detect the concrete reason for the malfunction.

As stated section 2, we use IRSVM to classify the root-class of printing fault. We didn't used randomly selected subset $\bar{m}$ from A but make a criterion which select training point by equal interval. In our test we use different ratio of reduced 1, 5 and 10% to compare with conventional SVM. The class of ink fault use different quantity of data set to test the computational time. The result can be seen at next section.

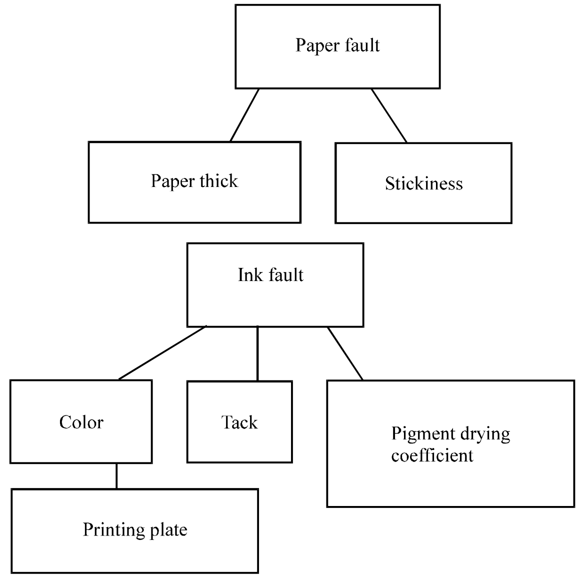


Fig. 1: Classes in printing fault

After classifying by IRSVM classifier which distinguishes root-class, we use decision tree based on C4.5 algorithm to further diagnosis. In (Kantardzic, 2003) describe the algorithm to calculate the information gain rate and based this algorithm we construct the sub-classifier. The decision trees of study fault and ink fault are showed in Fig. 1.

## THE EXPERIMENTS RESULTS

All our experiments run on a personal computer which consists of Pentium-4 2.4GHz processor, 512MB of memory and utilizing the Windows XP operating system. We used LibSVM version 2.36 (Chang and Lin, 2001) to implementation and used Gaussian kernel:

$$K(x, z') = e^{-\mu\|x-z\|_2^2} \qquad x, z \in R^n \tag{11}$$

for the parameter $\mu$, we choose appropriate value $\mu$ = 0.5 based on experiment. As Fig.1 illustrates, study faults have two subclasses and have 16 properties. Ink faults have four subclass divide 21 properties.

For the class of ink fault using different quantity of data set to test the computational time as Table 1 illustrated. In this test we selected $\bar{m}/m$ =5%. Table 2 give

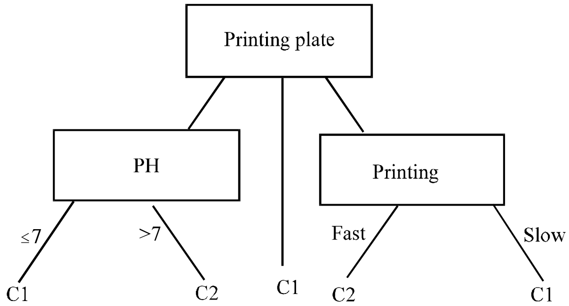


Fig. 2: A decision tree based on C4.5 algorithm

Table 1: Different $\overline{m}$ and different training dataset need the computational time

| Training t Datase (m×n) | IRSVM (1%) Time (sec) | IRSVM (5%) Time (sec) | IRSVM (10%) Time (sec) | SVM Time(sec) |
|---|---|---|---|---|
| Paper fault (466×16) | 14.2 | 10.5 | 9.1 | 25.8 |
| Ink fault (382×21) | 11.4 | 10.6 | 8.4 | 34.2 |

Table 2: Using the trained classifier by table1 to classify

| Test Dataset (m×n) | IRSVM (1%) Correct (%) Time (sec) | IRSVM (5%) Correct (%) Time(sec) | IRSVM (10%) Correct (%) Time (sec) | SVM Correct (%) Time (sec) |
|---|---|---|---|---|
| Print fault | 4.1 | 89.6 | 5.5 | 85.3 |
| (948×16) | 87.2 | 6.1 | 87.2 | 16.9 |
| Ink fault | 3.4 | 81.2 | 3.6 | 86.4 |
| (586×21) | 77.2 | 4.3 | 87.2 | 17.5 |

Table 3: Flat database of printing fault based on paper

| Printing plate | PH | Printing speed | Class |
|---|---|---|---|
| PS plate | 7 | Fast | C1 |
| PS plate | 9 | Fast | C 2 |
| PS plate | 8.5 | Slow | C 2 |
| PS plate | 9.5 | Slow | C 2 |
| PS plate | 7 | Slow | C 1 |
| Paper plate | 9 | Fast | C 1 |
| Paper plate | 7.8 | Slow | C 1 |
| Paper plate | 6.5 | Fast | C 1 |
| Paper plate | 7.5 | Slow | C 1 |
| GSS plate | 8 | Fast | C 2 |
| GSS plate | 7 | Slow | C 2 |
| GSS plate | 8 | Slow | C 1 |
| GSS plate | 8 | Fast | C 1 |
| GSS plate | 9.6 | Fast | C 1 |

us a result as the training set get larger, the efficiency of conventional SVM get poorer than IRSVM.

After classifying by IRSVM classifier, we use decision tree based on C4.5 algorithm to further diagnosis. As Table 3 illustrated, a subset of printing fault caused by study are described as a flat file database.

According to the information gain rate, we use study plate as the first classifying basis. The first property has the highest information gain rate. The final decision tree presented as Fig. 2.

## SUMMARY

In this study, we used IRSVM to classify the root-class of printing fault. We didn't use randomly selected subset $\overline{m}$ from A but make a criterion which select training point by equal interval. The computational result proved IRSVM can avoid the computational difficulties in generating a nonlinear support vector machine classifier for a massive dataset. The reduced kernel matrix cuts the problem size and does not scarify the prediction accuracy. Computational results show IRSVM only adapt to classify the root-class of printing fault, so we use C4.5 to detection the low-layer faults for remedying the defect of IRSVM. The experiments show the hybrid method has a high efficiency for what caused by adjustable reason such as study or water. But if the fault resulted with press machine, the hybrid method gets poor. It is hoped to give new insight to the total system approach.

## ACKNOWLEGMENTS

This research is supported by Beijing University Youth Excellence Program YETP1467 and it is supported by Beijing Institute of Graphic Communication accented term.

## REFERENCES

Chang, C.C. and C.J. Lin, 2001. LIBSVM: A library for support vector machines. http://www.csie.ntu. edu.tw /~cjlin/libsvm/

Kantardzic, M., 2003. Data Mining: Concepts, Models, Methods and Algorithm. 2nd Edn., Tsinghua University Press, China, ISBN:13-978111802-9138, pp: 122-128.

Lee, Y.J. and O.L. Mangasarian, 2001. SSVM: A smooth support vector machine. Computational optimization and applications. Comput. Optim. Applic., 20: 5-22.

Lee, Y.J., H.Y. Lo and S.Y. Huang, 2003. Incremental reduced support vector machines. Proceedings of the International Conference on Informatics, Cybernetics and Systems, December 14-16, 2003, Kaohsiung, Taiwan.

Mangasarian, O.L., 2000. Advances in Large Margin Classifiers. MIT Press, Cambridge, MA., pp: 135-146.

Platt, J.C., 1999. Fast Training of Support Vector Machines Using Sequential Minimal Optimization Support Vector Learning. MIT Press, Cambridge, MA., USA., ISBN:0-262-19416-3, Pages 185-208.

Sabzekar, M. and M. Naghibzadeh, 2013. Fuzzy c-means improvement using relaxed constraints support vector machines. Applied Soft Comput., 13: 881-890.

Wang, H., X. Zhang and J. Yu, 2004. Fault diagnosis based on support vector machine. J. East China Uni. Sci. Technol., 30: 179-182.

Wu, M., B. Scholkopf and G. Bakir, 2006. A direct method for building sparse kernel learning. J. Mach. Learn. Res., 7: 603-624.