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 (19): 5017-5021, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.5017.5021

Soft Decision-Making SVM Using in Fault Detection and Diagnosis

Boyang Li
School of Management Science and Engineering,
Dongbei University of Finance and Economics, Dalian, Liaoning 116025, China

Abstract: One of the central problems of many automation systems is Fault Detection and Diagnosis (FDD). The data from real systems are commonly high-dimension and hard-separated. Various mathematical techniques have been applied on these data. One of the problems in FDD, which so far has been still hard to solve, is how to deal with the data nearly reached the "critical mass". Because the system is very unstable when it nearly reaches the critical condition, it may become very hard to collect data and make decision. An improved Support Vector Machine (SVM) classifier with a soft decision-making boundary is proposed in this paper. The boundary is constructed based on belief degrees of data and reflects the data distribution. A membership function of the critical condition is introduced to extract the critical state data. In order to deal with these critical state data, this paper introduces and discusses two different experimental strategies.

Key words: Support vector machine (SVM), Fault Detection and Diagnosis (FDD), soft decision-making

INTRODUCTION

In most automation systems, FDD is an important problem. Fault detection is recognizing if a problem or an abnormal state has occurred or not. A fault might be led by several "root causes", which are fundamental, underlying problems in the automation systems (Delpha *et al*., 2012). A "root cause" is reflected in one or more observations, but might not be directly observed. Because of the complexities of these systems, it becomes hard to explore the relationships between the fault and observations.

Data from automation systems differ from other classification data in several ways. The most significant distinction lies in the "critical mass" problem. When the system nearly reaches the "critical mass", it becomes very unstable. It is hard to design an effective classifier to deal with these data. And, if these data are used in the training process, the performances of the learning machine may be affected (Dandare and Dudul, 2012).

Because of the existence of this critical state region between the "normal" class and the "fault" class, most hard boundary classifiers are not suitable for FDD in automation systems. A more robust and flexible separation boundary needs to be introduced to improve the classifiers. SVM has general well performances for unseen data classification (FDD data belongs to this kind of data), so we proposed a soft decision-making boundary to improve SVM-based classifiers (Wu *et al*., 2013). In addition, a membership function is designed to measure if a sample belongs to the "critical mass" region or not. For the samples in this region, we have two strategies: Multi-label strategy and Take-off strategy. If the first strategy is used, all samples in the "critical mass" region are given 2 labels, that means they belong to the both of two classes. If we use Take-off strategy, all samples are deleted from the training data. In other words, the data are cleaned, that means no "critical mass" region in the training data.

The remainder of this paper is organized as follows: Section 2 provides the introduction of the proposed SVM classifier model with soft decision-making boundary and membership function of the "critical condition". Section 3 introduces the parameter estimation of the functions. Section 4 discusses a typical FDD problem "Misfire Detection" and presents some simulation results. Finally, concluding remarks are given in the last section.

SOFT DECISION-MAKING SVM

Most of data used in FDD are the typical nonlinear and non-separable real world data. In order to solve this kind of data, we should attempt to minimize the separation error and maximize the margin simultaneously in SVM. The margin is defined as the distance between bounding planes for different classes.

The basic idea of SVM is to map the input vectors into a high-dimension feature space by a mapping function first and then solve the problem in the high-dimension space. The coefficients of the decision-making function are calculated by solving a quadratic programming (QP) problem in its dual form.

Conventionally, SVMs just use a sign function to make decision. This decision function just extracts the signs of decision values as the class labels.

Decision values: Assume that we have a training data set from an automation system denoted as $\{x_i, y_i\}$, where $x_i \epsilon R$ means the i-th sample and y_i is its class label (+1 or -1, Normal and Fault) that denotes the functionality of this sample. The training data set can be divided into these two classes N and F with labels +1 and -1 respectively.

Support vectors (SVs) are the most important samples in the training data. SVs from Class N are the samples N_i in the half space $\{x \epsilon Rn|\omega^T x \leq b+1\}$ and SVs from Class F are those samples Fi in the half space $\{x \epsilon R^n|\omega^T \geq b-1\}$, where ω and b are weights and bias. In SVMs, only these samples are the relevant solutions in the calculation of separating boundary.

We transform the input data from the low dimension data space into a high dimensional feature space by a nonlinear mapping function $\varphi(x)$. And then we can construct the plane function (1).

$$y_i[\omega T\ \varphi(x_i)+b] \geq 1\text{-}1\xi_i,\ \forall_i \tag{1}$$

Then, the optimal hyper plane problem is changed into an optimization problem (2),

$$\min_{\omega,b,\xi} J(\omega,\xi) = \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{N}\xi_i \tag{2}$$

$$\text{s.t.}\begin{cases} y_i\left[\omega^T\varphi(x_i)+b\right] \geq 1-\xi_i, \\ \xi_i \geq 0, i = 1,\ldots,N. \end{cases}$$

in which parameter C is used to control the degree of tolerance, which is the only changeable parameter in SVM.

Introducing Lagrange multipliers $\alpha = (\alpha_1, ..., \alpha_N)$, this optimization problem is rewritten as a QP problem in its dual form:

$$\max_{\alpha} Q(\alpha) = \frac{1}{2}\sum_{i,j=1}^{N} y_i y_j K(x_i, x_j)\alpha_i\alpha_j + \sum_{j=1}^{N}\alpha_j \tag{3}$$

$$\text{s.t.}\begin{cases} \sum_{i=1}^{N}\alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C, \forall i \end{cases}$$

where K (xi, yi) = φ(xi)Tφ (x_i) is the kernel function. The i-th sample is a support vector, if its corresponding $\alpha \neq 0$. The decision value function is shown as following:

$$v = \sum_{i=1}^{N}\alpha_i y_i K(x, x_i) + b \tag{4}$$

where x_i are the samples from the training data and x is the input data. Only SVs have nonzero Lagrange multipliers, so we can only use them to determine the separation boundary.

Soft decision-making: Conventional SVMs only use a sign function as the decision-making function. However, the hard separation boundary is commonly unsuitable for real-world data, especially data from automation systems. In these systems, noises and "critical conditions" result in the difficulty to describe the distribution of data. They lead the separation boundary between classes to become a gray zone.

We found the data used for FDD are commonly imbalanced. Because the samples in Class Normal is much more than the samples in Class Fault, the sign decision-making function may misclassifies same samples near the separation boundary. Thus, we designed an offset parameter δ to modify the midline of decision-making function a more suitable position.

Figure 1 shows the shape of soft decision-making boundary, comparison with the conventional hard separation boundary.

It is clear that the conventional separation boundary has a threshold 0. As shown in the figure, the offset para meter δ modifies the boundary excursion caused by the data imbalance (Li *et al*., 2008).

Based on the observation, analysis and experiments, we designed the soft decision-making functions by using arc-tangent functions with reflection of decision values. Let $f_N(v)$ and f_F (v) denote the belief degrees of Class Normal and Class Fault. The decision-making curves of Class N and Class N are defined as Eq. 5 and 6 respec tively:

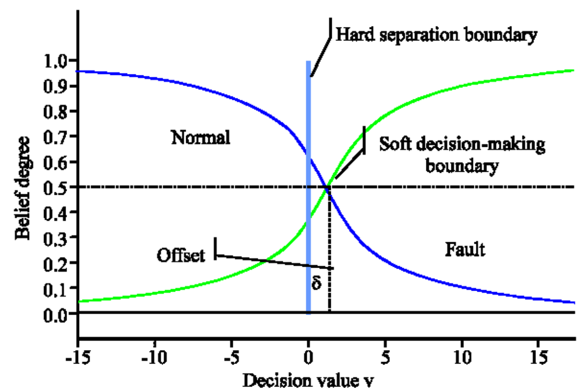

Fig. 1: Soft decision-making boundary

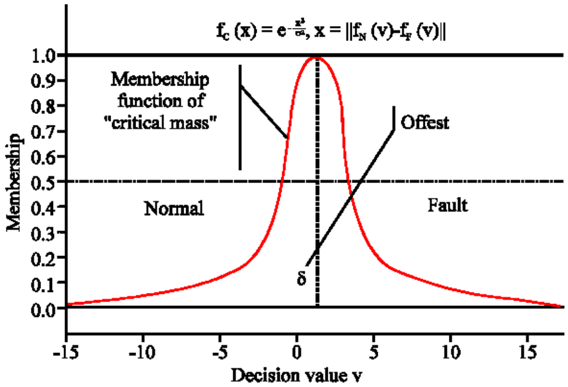

Fig. 2: Membership function of "Critical mass"

$$f_N(v) = \frac{\text{Arctan}(-v\cdot s+\delta\cdot s)}{\pi} + 0.5 \quad (5)$$

$$f_F(v) = \frac{\text{Arctan}(v\cdot s-\delta\cdot s)}{\pi} + 0.5 \quad (6)$$

where, v is the decision value. S is the scale parameter used to describe the shape of curves. In this method, the decision-making function is:

$$y(x) = \begin{cases} -1 \text{ if } (V_r < f_N(v), v < \delta) \\ \quad \text{or } (V_r > f_F(v), v > \delta) \\ +1 \text{ if } (V_r < f_F(v), v > \delta) \\ \quad \text{or } (V_r > f_N(v), v < \delta) \end{cases} \quad (7)$$

where, V_r is a random value between 0 and 1.

Membership function of critical: Soft decision-making function can be deal with the imbalance data and improve the performance for noisy data. However, if we can extract the data in the "critical mass" region, then we can design some strategies for FDD in automation systems.

When an automation system nearly reaches the "critical mass", the samples are almost randomly separated to normal and fault state. Thus, in this region, belief degrees of two classes should be approximately equal. So the function of the "critical" mass region is designed as following:

$$f_C(x) = e^{-\frac{x^2}{\sigma^2}}, x = f_N(v) - f_F(v) \quad (8)$$

σ is the width parameter. The "critical mass" decision-making function is:

$$y'(x) = \begin{cases} 0 \text{ if } (V'_r \le f_C(v)) \\ y(x) \text{ if } (V'_r > f_C(v)) \end{cases} \quad (9)$$

where v'_r is a random value between 0 and 1. If y' (x) equals to 0, then the sample is detected as the data in the "critical mass" region.

PARAMETER ESTIMATION

In Eq. 5, 6 and 8, the offset parameter δ, the scale parameter and the width parameter σ need to be estimated. We tried to use curve-fitting approach to estimate these parameters and make soft decision-making boundary fit the distribution of belief degree in the decision value domain.

Distribution of belief degree: Let us consider an FFD example. Our method uses a soft decision-making boundary to replace the conventional hard one. This soft boundary should describe the distribution of data accurately.

Fist, the decision values for training samples are calculated by Eq. 4. The decision values denote the distances between the training samples and separation boundary. Secondly, we divide the whole decision value domain into some intervals. Assume that the number of samples from Class N and Class F in the i-th interval are M^i_N and M^i_F. Let the number of SVs equals to the number of critical state samples, which are denoted as M^i_C.

Define the percentage of samples from a class in a certain interval as the belief degree, then the belief degrees of Class N, Class F and "critical mass" region are written as follows:

$$g^i_N(v_i) = \frac{M^i_N}{M^i_N + M^i_F} \quad (10)$$

$$g^i_F(v_i) = \frac{M^i_F}{M^i_N + M^i_F} \quad (11)$$

$$g^i_C(v_i) = \frac{M^i_C}{M^i_N + M^i_F - M^i_C} \quad (12)$$

v_i is the center decision value of the i-th interval, where i = 1, 2,.. , k.

Parameter estimation: s, δ and σ are estimated to minimize the summation of the squares of the residuals r_N, r_F and r_C,

$$r_N = \sum_{i=1}^{N}\left(g^i_N(v_i) - f_N(v_i)\right)^2 \quad (13)$$

$$r_F = \sum_{i=1}^{N}\left(g^i_F(v_i) - f_F(v_i)\right)^2 \quad (14)$$

$$r_C = \sum_{i=1}^{N} \left(g_C^i(v_i) - f_C(v_i)\right)^2 \qquad (15)$$

The least squares curve-fitting algorithm, Newton's method combined with the gradient descent, is used to solve this regression problem (Avriel, 2003). The value of s, δ and σ can be easily estimated. The soft decision-making function and the membership function of "critical mass" are constructed automatically.

MISFIRE DETECTION

"Misfire" is a typical FDD problem. In this research, we used the data from Ford Motor Company in our experiments (Robert *et al*., 2008).

Background: In a typical combustion engine, fuel is ignited within a cylinder in the engine. This process forces air in the cylinder to expand and moving a piston. The piston in turn pushes a portion of a crankshaft. Then, the shaft rotates. If the fuel is not ignited, for some reason, no force is transmitted to the crankshaft. We call these occurrences "misfires". Misfire problem relates to combustion failures within the engine and adversely affects engine efficiency.

The data we used contain the information of time series of 50000 samples (for 10-cylinder internal combustion engine). Each sample consists of four inputs elements and one output label (normal firing or misfire).

The four elements of input vectors represent cylinder identifier (1st), engine crankshaft speed in Revolutions Per Minute (RPM) (2nd), load (3rd) and crankshaft acceleration (4th) respectively. The number of samples from the normal firing class in the training data and the test data are 45093 and 45395. The numbers of samples from the misfire class in the training and test data sets are 4907 and 4605. It is obvious that that this data is extremely imbalanced.

Classification results: For the samples in the "critical mass" region, we tried 2 strategies in our experiments: Multi-label strategy and Take-off strategy.

These 2 strategies are used to test if the data near the "critical mass" are still useful for SVM training. If the Multi-label strategy outperforms Take-off strategy, that means data clearing based on the latter may lead to losing some information.

Let C in SVM equal to 1, 10, 100 and use RBF kernel and Polynomial kernel respectively. The comparison results between the proposed approach and WSVM, Multi-label strategy and Take-off strategy are shown in Table 1 and Table 2. In Table 1, we used Multi-label strategy for soft decision-making SVMs. MultiLable-RBF in Table 2 and Soft-RBF-SVM in Table 1 are the same method. MultiLable-Poly in Table 2 and Soft-Poly-SVM in Table 1 are same.

Table 1: Comparison between Soft decision making SVMs and classical SVMs (%)

C=	1	10	100
Soft-RBF-SVM	95.07	95.34	95.70
Soft-Poly-SVM	92.91	93.75	94.65
RBF-SVM	92.15	91.78	91.98
Poly-SVM	90.80	92.48	93.22

Table 2: Comparison between Multi-label Strategy and Take-off Strategy (%)

C=	1	10	100
MultiLabel-RBF	95.07	95.34	95.70
MultiLabel-Poly	92.91	93.75	94.65
TakeOff-RBF	94.84	94.91	95.45
TakeOff-Poly	92.88	93.56	94.30

DISCUSSIONS

As shown in Table 1, the proposed method presents better classification performances than classical SVMs. The accuracy of Poly kernel SVMs increased by turning C to a higher value. The performances of RBF kernel SVMs are almost flat.

From Table 2, we can find that Multi-label strategy outperformed Take-off strategy. This result verifies the hypothesis we suggested: using Take-off strategy results in some information losing. In addition, this result also proved that the soft decision-making boundary is suitable for dealing with multi-label classification problems.

CONCLUSIONS

In order to deal with "critical mass" problems in FDD, we proposed a soft decision-making method to improve SVM classifiers. The main problems addressed in this research are how to extract samples in critical condition and how to deal with them.

In this study, we defined belief degrees based on the decision values of samples. The distribution of this belief degree in the decision value domain is reflected by the soft decision-making boundary. A "critical mass" membership function is proposed to extract samples in critical.

Through experiments, the proposed method was verified to be useful for FDD and showed better performances than classical methods.

ACKNOWLEDGMENT

The author would like to thank for the support by National Statistical Research Program of China (National Bureau of Statistics of China, No. 2013LZ25).

REFERENCES

Avriel, M., 2003. Nonlinear Programming: Analysis and Methods. Dover Publishing, Mineola, New York.

Delpha, C., H. Chen and D. Diallo, 2012. SVM based diagnosis of inverter fed induction machine drive: A new challenge. Proceedings of the 38th Annual Conference on IEEE Industrial Electronics Society, October 25-28, 2012, Montreal, QC., pp: 3931-3936.

Dandare, S.N. and S.V. Dudul, 2012. Support vector machine based multiple fault detection in an automobile engine using sound signal. J. Electron. Electr. Eng., 3: 59-63.

Li, B., J. Hu and K. Hirasawa, 2008. Support vector machine classifier with WHM offset for unbalanced data. J. Adv. Comput. Intell. Intell. Inform., 12: 94-101.

Robert, B.G., K. Marko and P. Sun, 2008. Neural network-based engine misfire detection systems and methods. US US20080243364, WO2008154055A2. http://www.google.com/patents/WO2008154055A3.

Wu, X., Y. Chang, J. Mao and Z. Du, 2013. Predicting reliability and failures of engine systems by single multiplicative neuron model with iterated nonlinear filters. Reliab. Eng. Syst. Saf., 119: 244-250.