Research Application of Support Vector Machine in Fault Diagnosis of Certain Type Engine

1, 2 Donghua FENG
1 School of Computer Science and Technology, Wuhan University of Technology, Hubei Wuhan, 430070, China
2 Dept. of Computer and Information Engineering, Nanyang Institute of Technology, Henan Nanyang, 473004, China
Tel.: 13733116699
E-mail: fengdonghuaedu@163.com

Received: 2 December 2013   /Accepted: 28 January 2014   /Published: 28 February 2014

Abstract: For the engine fault diagnosis in real problems, the number of samples available are limited, and the progress of research on the theory of the most limited to assume that the data samples, so that the network training data examples, in engineering applications has been slow, in this paper, the application of support vector machine in fault diagnosis of engine, the segmentation of the training sample set, in order to achieve the optimal analysis of the machine, the reasoning ability best. First introduced the two classification method of support vector machine and multi classification method based on two classification methods of the study, and applied to the fault diagnosis of engine, and then the simulation test for this method, and compared with the existing methods, the results show the effectiveness of the classification method, the results of the analysis also can use the tree diagram or table form, simple and intuitive; but also can save the contribution to some extent in time.

Keywords: Support vector machine, Multi class classification method, Engine, Fault diagnosis, Fault analysis.

1. Introduction

Analysis of mechanical failure is a pattern recognition problem of statistical significance, also is the machine learning problem based on data. Data based machine learning is a very important aspect of modern intelligent technology, mainly studies how to from some observation data (sample) obtained at present can not be obtained by analyzing the principles of law, to analysis the object using these rules, to predict future data or can not be observed data [1]. This requires that the machine (computer) with certain reasoning ability, only with the reasoning ability or acceptable certain learning machine has the practical value in engineering.

But in reality, the number of samples available are limited, and the progress of research on the theory of the most limited to assume that the data samples, so that the network training data examples, in engineering applications has been slow, in this paper, the application of support vector machine in fault diagnosis of engine, by dividing the training samples set in order to achieve the optimal machine, analysis, reasoning ability best.
2. Study Examples and Two Classification Methods of Support Vector Machine

2.1. Two Classification Method of Support Vector Machine

Support vector machine classification includes two linear separable, nonlinear and linear inseparability can be divided into three cases [2]. Fault identification problem can be transformed into the problem of engine classification of data, which belongs to the nonlinear classification can be divided into two. The main idea to support vector machines for data classification is: the input vectors are mapped into a high-dimensional feature space by a nonlinear mapping prior selection, structure to meet the requirements of the linear optimal classification hyper plane to partition the training sample set in this space, and make the training sample set point from the optimal hyper plane as far as possible even if the hyper plane, on both sides of the blank area (margin), the former can make the experience risk minimum, the latter can guarantee confidence range generalization bounds of the minimum, by both can make the real risk to the minimum [3].

Nonlinear classification algorithm described below:

For the d dimension vector space in general, a training sample set \((x_i, y_i)\), \(i=1,2,\ldots,n\), \(x \in \mathbb{R}^d\), \(y_i \in \{+1,-1\}\) is the desired output the corresponding. First, the nonlinear mapping function \(\Phi(x)\) the input data from the original space to N dimension feature space, the structure of optimal classification hyperplane in the feature space, equation (1).

\[
\omega \cdot \Phi(x) + b = 0 \tag{1}
\]

The discriminant function normalized condition, all the sample can be separated from the equation (2):

\[
y_i [\omega \cdot \Phi(x_i) + b] - 1 \geq 0 \tag{2}
\]

Formula (2) is to satisfy the first condition of structure risk minimization: ensure the empirical risk minimization. In the formula (2), the equality of the point is called support vector. To ensure the premise established the first condition, again looking for confidence minimum condition generalization bounds, namely the class interval (margin) maximum conditions [4]. After normalization, it can prove, classification interval equal to \(2 / ||\omega||\). So the interval is equivalent to the \(||\omega||\) minimum. Two uniform conditions can be expressed as a constrained optimization problem as the following, namely (2) in the constraints, for function (3).

\[
\Phi(\omega) = \frac{1}{2} ||\omega||^2 = \frac{1}{2} (\omega \cdot \omega) \tag{3}
\]

For get function (3) minimum. This constrained optimization problem can be solved by constructing a Lagrange function, Lagrange function following function (4).

\[
L(\omega, b, a) = \frac{1}{2} (\omega \cdot \omega) - \sum_{i=1}^{n} a_i [y_i (\omega \cdot \Phi(x_i) + b) - 1] \tag{4}
\]

where \(a_i \geq 0\), Lagrange multiplier. The Lagrange function \(L(\omega, b, a)\) on the W and B for the minimum value, the minimum value shall be calculated at the saddle points.

In the saddle point, \(\omega, b\) and \(a_i\) solutions. \(\omega^*, b^*\) and \(a^*_i\) must meet the following conditions:

\[
\frac{\partial L(\omega, b, a)}{\partial b} = 0 \tag{5}
\]

\[
\frac{\partial L(\omega, b, a)}{\partial \omega} = 0 \tag{6}
\]

By equation (5) and (6), it can get the optimal classification face the following characteristics:

1. The optimal hyper plane, coefficient \(a^*_i\) a constraint:

\[
\sum_{i=1}^{n} y_i a^*_i = 0, \quad a^*_i \geq 0, i=1,2,\ldots,n \tag{7}
\]

2. Optimal hyperplane (vector \(\omega^*\)) is a linear combination of vector in the training set:

\[
\omega^* = \sum_{i=1}^{n} y_i a^*_i \Phi(x_i), \tag{8}
\]

where coefficient of non support vectors corresponding to \(a^*_i = 0\), is ignored in the calculation, only the support vectors, coefficient of the corresponding \(a^*_i > 0\), to take effect in the calculation. Therefore, only the support vectors to determine the optimal hyper plane, while the other learning samples will not affect the optimal hyper plane. Get equation (7) and (8) get into (4), as is shown by equation (9) [5].

\[
L(\omega, b, a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} y_i y_j a_i a_j (\Phi^T(x_i) \cdot \Phi(x_j)) \tag{9}
\]
SVM does not directly solve the high-dimensional feature space plot points \( \Phi^T(x_i) \cdot \Phi(x_j) \), instead of the original space by kernel function; the function is to satisfy the symmetric function Mercer:

\[
K(x, x_j) = \Phi^T(x) \cdot \Phi(x_j)
\]  

(10)

According to (8), the optimal hyperplane is determined by the Lagrange multiplier \( a_i^* \) and support vector decision, therefore, to construct the optimal hyperplanes, is called \( a_i^* \), which can be determined by the maximum value (9) obtained. The support vector \( a_i^* \) and the corresponding parameters of \( b^* \) press, to obtain equation (11).

\[
b^* = \frac{1}{2} \left[ \omega^* \cdot x^* (1) + \omega^* \cdot x^*(-1) \right],
\]  

(11)

where \( x^* (1) \) that belong to the first class (which corresponds to the \( y=1 \)) of an arbitrary support vector, \( x^*(-1) \) that belong to the second class (corresponding to \( y=-1 \)) of an arbitrary support vector.

After the training is completed, to obtain the optimal data classification hyper plane, constitute the support vectors and the corresponding parameters of the classifier [6]. The test samples by equation (12) type classification prediction.

\[
f(x) = \text{sgn}(\omega^* \cdot \Phi(x_j) + b^*) = \text{sgn}\left[ \sum_{SV} y_i a_i^* K(x_i, x) + b^* \right],
\]  

(12)

where \( \text{sgn} \) is the sign function.

Choosing different kernel function, it can construct different SVM. Common kernel function with polynomial kernel function, Gauss kernel function, Sigmoid kernel function, we use the Gauss kernel function, as is shown by equation (13).

\[
K(x_i, x_j) = \exp\left[ -\frac{|x_i - x_j|^2}{2\sigma^2} \right]
\]  

(13)

2.2. An Example of Single Fault Diagnosis of Engine

The fuel system and cooling system of a certain type of engine failure is more local, therefore, the fault of the two parts as an example, the fault pattern recognition using support vector machine. The main failure mode of the two parts: a cylinder not injection, cooling water temperature is too high, the temperature sensor circuit breaker, torque sensor, air flow sensor [7]. The parameters of support vector machine input for 6, including: torque, temperature, solar term door opening, air flow, speed, pulse width. At the time of diagnosis, the input parameters in the [0, 1]. For the two classifications, only asked to judge the engine components have no fault, so only the training of a SVM can be. In Table 1 in a sample of 1 - 99 SVM training, inspection Table 1 in the 1 - 13 samples. Here, we can use the normal state as a state, the other all the fault status as a state (also can put any kind of failure as a state, and all other states as a state). Take \( C=100 \), \( \sigma^2 = 0.2 \). In the two case of classification, the trained SVM on all test samples were given a correct classification results.

<table>
<thead>
<tr>
<th>No</th>
<th>Torque (V)</th>
<th>Temperature (V)</th>
<th>Solar term door opening (V)</th>
<th>Air flow (V)</th>
<th>Speed (t/min)</th>
<th>Pulse width (ms)</th>
<th>Failure mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.371</td>
<td>0.281</td>
<td>0.945</td>
<td>1.404</td>
<td>13179.86</td>
<td>4.56</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>0.41</td>
<td>0.73</td>
<td>1.05</td>
<td>1.624</td>
<td>17974.68</td>
<td>4.156</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0.942</td>
<td>0.251</td>
<td>1.648</td>
<td>2.407</td>
<td>1855.747</td>
<td>7.628</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>3.189</td>
<td>0.398</td>
<td>1.072</td>
<td>1.675</td>
<td>1255.04</td>
<td>5.857</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>0.419</td>
<td>0.572</td>
<td>0.727</td>
<td>3.747</td>
<td>687.424</td>
<td>4.5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>0.483</td>
<td>4.957</td>
<td>0.942</td>
<td>1.396</td>
<td>1053.963</td>
<td>5.224</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>1.26</td>
<td>0.055</td>
<td>1.716</td>
<td>2.341</td>
<td>1414.999</td>
<td>9.376</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0.84</td>
<td>0.154</td>
<td>1.626</td>
<td>2.285</td>
<td>1403.312</td>
<td>8.8</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>0.637</td>
<td>0.389</td>
<td>1.223</td>
<td>1.821</td>
<td>1150.996</td>
<td>5.864</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0.415</td>
<td>0.942</td>
<td>0.857</td>
<td>1.25</td>
<td>1123.007</td>
<td>4.288</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>1.572</td>
<td>0.077</td>
<td>1.938</td>
<td>2.754</td>
<td>1760.563</td>
<td>11.068</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>3.196</td>
<td>0.97</td>
<td>1.38</td>
<td>2.009</td>
<td>1210.38</td>
<td>7.888</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>0.39</td>
<td>0.667</td>
<td>0.721</td>
<td>3.767</td>
<td>620.904</td>
<td>4.5</td>
<td>5</td>
</tr>
</tbody>
</table>

These figures represent the meaning of follows:
0 – Normal;
1 – The cooling water temperature is too high;
2 – One cylinder without injection;
3 – The torque sensor circuit;
4 – The temperature sensor circuit;
5 – Damaged – air flow meter.

According to the calculation results, the fault mode two classifications, 99 training samples with 21 training samples become support vectors, their sample number is: 3, 6, 11, 16, 17, 22, 24, 30, 35, 37, 41, 42, 48, 57, 63, 67, 70, 76, 87, 96, 99. The sample of 3, 11, 16, 17, 24, 30, 35, 42, 57, 63, 67, 70, 76, 96 in the fault free boundary, the remaining samples in fault boundary, the support vector from the classification boundary nearest.

For any test sample, the fault classifier based on \( a_i^* \) \( K(x_i, x) + b^* \) classification, if the discriminant function is greater than 0, belonging to the non fault type, fault or belonging to the class.

Fig. 1 gives all test samples of discriminant function distribution, a sample of 2, 3, 9, 10 of positive discriminant function, belonging to the non fault class, while the rest are fault class samples. As can be seen from Table 2, the classification results are consistent with the actual situation (where N represents the normal, F represents the fault).
3. Multi Class Classification Method
Case Study of Support Vector Machine

3.1. Multi Class Classification Method of Support Vector Machine

The two level classification principles, the multi value classification (k classification, k > 2) is based on two value based classification. There is a certain relationship between K classification and two classification problems [8]. If a problem can be divided into K type, K type between any of the two must be divided; on the contrary, in a K classification problem, if known any of its two can be divided, by the combination of the law, but by two can be divided into the final realization of K classification. Combination law is different form different classification methods [9]. One of the major one on one (one against one), one to many (one against all) and directed acyclic graph (directed acyclic graph) etc. Following is a brief introduction about the above several classification methods.

1. A multi classification method

A multi method to establish k two values in the training of the classifier, and k value of two classifiers is to use all the training samples. For example, in the training of i classifier, the fault sample label class i is 1, the rest of the sample label is -1. Classifier training is completed, set up k of two class classifier, and then use the k classifier to discriminate the input sample calculation. In the realization of one to much classification, can be eliminated one by one method: the test sample x shown in Fig. 2 procedures in accordance with the classification, each classifier can identify a fault, using all of the classifier to classify the samples calculated one by one. Discriminant function formula (14):

\[ f^i(x) = \text{sgn}[(\omega^i)^T K(x, x_i) + b^i], \quad i=1, 2, ..., k \quad (14) \]

If the \( f^i(x) = 1 \), judge x belongs to i class, the end of the test. Otherwise the x input to a classifier, until the traversal of all classifier, if all of the \( f^i(x) \) are not equal to 1, then to judge x as fault other outside the k class or normal state.

2. The directed acyclic graph method

Judging flowchart of this approach is a direction, but not cyclic graph, as shown in Fig. 3. A value of two per 2 class classifier, a node of each classifier is in figure. For the k classification, there were k(k-1)/2 nodes. For a given sample x, to determine its genus, can start from the root node, each node (two classes) for inspection, identification and function of each node is the same equation (14). If the discriminant function of the value of a node is -1, then from the left into the next layer, to judge; otherwise, if the discriminant function is +1, then from the right side into the next layer. Until it reaches the bottom, you can classify x.

3. A multi classification method.

The k classification problem, a category k of the training data, the k data of two combinations, Co Construction \( M = C_k^2 = k(k-1)/2 \) training set, respectively, using the SVM two value on the M training set for learning classification method, to generate the M classifier. For the construction of a
class i and class j data training two class classifier problem for finding the following function value, as is shown by equation (15).

\[
\min \frac{1}{2} (\omega^i)^T \omega^i + C \sum_{n=1}^{N} \xi_n (\omega^i)^T (\omega^j)^T + \xi_n \xi_m
\]  

(15)

The constraint conditions are the following equations:

\[
(\omega^i)^T K(x_n) + b^i \geq -1 + \xi_n, \quad \text{if} \quad y_n = i
\]  

(16)

\[
(\omega^j)^T K(x_n) + b^j \leq -1 + \xi_n, \quad \text{if} \quad y_n = j
\]  

(17)

where \( x_n \) is the class i or class j in a training sample, \( y_n = n \) is \( x_n \) (n class =1,2,..., k). Build up M value of two classifiers; it can implement classification of unknown samples.

Judged by the voting method, the decision-making process is shown in Fig. 4.

3.2. Case Study of Engine Fault Classification

In the mechanical fault analysis, the two categories is the most basic requirements, it just runs the state judgment system simple is normal or abnormal. To conduct a comprehensive analysis of the recognition of fault situation of mechanical products, only two of the value of classification is obviously not enough. In order to eliminate failure, to ensure the normal operation of the product, not only to identify the products have no fault, but also to determine the fault type and location specific parts. This is a multi class classification problem.

For one to one for multi class classification method, to diagnose the fault of the 6 types (normal state as a fault condition), you need to train 15 SVM. At the time of diagnosis, with the appendix Table 1 samples as training samples, the data in Table 1 as a test sample. Take C =1, \( \sigma^2 = 0.2 \). As shown in Table 3, the results are as follows: 13 test sample fault categories all correct diagnosis.

<table>
<thead>
<tr>
<th>Sample No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fault type</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Winning number</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Recognition results</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Fig. 5 shows the times w on 3 samples tested in turn using paired SVM when compared to the types of the faults. The 1 samples of second kinds of fault wins most, the sample is second kinds of failures; similarly, 2 samples of zeroth kinds of fault wins most, the sample is zeroth kinds of failures; 6, sample fourth kinds of fault wins most, the sample is fourth kinds of fault. Other test samples have similar results, as shown in Table 3.

4. Conclusions

This paper introduces two classification method of support vector machine and multi classification method based on two classification methods of the study, and applied to the fault identification of the engine; the results show the effectiveness of the
classification method. Taking into account the habits, can help analysts take from top to bottom or from the analysis under the supreme direction; the results of the analysis also can use the tree diagram or table form, simple and intuitive; but also can save the contribution to some extent in time.

Fig. 5. Different fault categories wins.

References