Research on Fault Diagnosis and Prediction Method about Driving Motor for Seeding Metering of Plot Seeder Based on SVMs Classification and Regression

Yuntao Hou\textsuperscript{1,2}, Zequan Wu\textsuperscript{1,2}, Xiaohua Cai\textsuperscript{1,2} and Zhongge Dong\textsuperscript{1,2}

\textsuperscript{1}Heilongjiang Academy of Agricultural Machinery Engineering Sciences, Heilongjiang Academy of Agriculture Sciences, Harbin, Heilongjiang,150081, China

\textsuperscript{2}Heilongjiang WOER Agricultural Equipment Technology co. LTD, Harbin, Heilongjiang,150081, China

*Corresponding author’s e-mail:736126975@qq.com

Abstract. To solve the problem of fault diagnosis and prediction about driving motor (brushless DC motor) for seeding metering of plot seeder, the method based on SVMs classification(SVC) and regression(SVR) is proposed in this paper: firstly, the three-phase current signals of brushless DC motor(BLDCM) are analyzed to extract characteristic values of various samples; then the SVC model based on clustering combine SVMs dynamic pruned binary tree is constructed, and the SVR models about all and each class of training samples are all constructed, too. The first input sample is firstly predicted according to the SVR model about all training samples, then the predicted value is identified that belongs to a certain class of faults by the SVC model, and the next operation will be carried out according to the classification result. This method is verified in the fault diagnosis and prediction experiments of BLDCM, the diagnosis accuracy is 96.37% and the Root-Mean-Square Error is 0.217. It might be seen that the method presented in the paper is of validity.

1. Introduction
In order to meet the operational requirements of plot breeding experiments, the way of motor-drive is adopted by seeding metering of plot seeder [1,2], and because of simple structure and high precision, etc, BLDCM is widely used. Therefore, monitoring the running state of BLDCM is of great significance to ensure the accuracy of plot breeding experiments.

2. The fault diagnosis and prediction method based on SVMs classification and regression

2.1 Support vector machines
The basic theory of support vector machines(SVMs) is to use a few support vectors to represent the whole sample set, then the kernel function of the original space is used to replace the point product operation of high-dimensional feature space \(\mathbf{w}\) and \(\Phi(x)\).

SVMs has unique advantages in solving classification problems [3,4]. In a given set of samples \(\{(x_i, y_i) | i = 1, 2, ..., n\}\), \(x_i \in \mathbb{R}^n\) is the input value and \(y_i \in \{1, -1\}\) is the predicted value. The kernel function \(k(x_i, x)\) and the penalty parameter \(C\) are introduced, and the corresponding optimal classification function can be expressed as
The complete classification of n-class can be realized by constructing classification decision tree. However, the performance of the entire classification model will be greatly influenced by different binary tree structures, and such effect may result in "error accumulation". Therefore, the most easily separate class should be separated first in the process of constructing a binary tree [5,6].

Based on the above analysis, clustering combine SVMs dynamic pruned binary tree algorithm is taken as the SVC algorithm in this paper. The class similarity is proposed which considers not only the class distance but also the class distribution. Here are some formulas used in the algorithm.

The center of class i in the feature space is defined as

\[ m_i = \frac{1}{l_i} \sum \phi(x_i) \]  

The distance between class i and j in the feature space

\[ D_{ij} = \|m_i - m_j\| = \sqrt{\frac{1}{l_i} \sum \sum \Phi(x_i, x_j) - \frac{2}{l_i l_j} \sum \sum \sum \Phi(x_i, x_j) + \frac{1}{l_j} \sum \sum \sum \Phi(x_i, x_j)} \]  

In the equation (3), \( \Phi(x_i, x_j) = \Phi(x_i) \bullet \Phi(x_j) \).

The class similarity is defined as

\[ similar(i, j) = R_i^2 + R_j^2 / \|m_i - m_j\|^2 \]  

The procedure of clustering combine SVMs dynamic pruned binary tree algorithm is as follows:

Step1: the class similarity of samples computed by the equation (5).

Step2: two classes with the biggest similarity(i, j) are united as a new class and then the SVMs are trained; the class-center of the new class and the distances between the new class and the other classes are calculated by the equation (3) and (4); the similarity between the new class and the other classes are calculated by the equation (5).

Step3: the above two steps are repeated until all the classes are clustered two classes as the root nodes, thus the uppermost two nodes of binary tree as root nodes are reproduced. According to this principle, the lower nodes are reproduced in order. Then the SVMs classifiers between the binary treenodes on the same level are constructed in turn by the equation (1).

Step4: identify the class of input sample according to the SVMs binary tree constructed in step3 and record the characteristics of input sample.
Step5: the samples with small similarity to the input sample are removed and the rest samples are taken as the root node to reconstruct the SVMs binary tree.

Step6: compare the similarity between the next input sample (sample2) and the previous input sample (sample1) according to the equation (5). If the similarity is enough big, the SVMs binary tree constructed in step 5 is kept to identify the class of sample 2; otherwise, go to step 4.

End.

This classification algorithm reconstructs the binary tree according to the similarity between classes, and reduces the original n-1 SVMs two-class classifiers to n-1-i. The new binary tree constructed by the remaining n-1-i SVMs two-class classifiers can accelerate the classification speed and ensure the classification accuracy.

Figure 1. Topological structure of clustering combine SVMs dynamic pruned binary tree.

As shown in Figure 1, if fault 4 occurs, the impossible fault 1 and 2 are eliminated immediately, and the new binary tree is constructed to identify the class of current fault 4.

Figure 2. The flow chart of the pre-classification and later regression prediction algorithm.

The flow chart of the pre-classification and later regression prediction algorithm is shown in Figure 2, the SVC models of clustering combine SVMs dynamic pruned binary tree are constructed according to the equation (1), and the SVR models about all and each class of training samples are all constructed according to the equation (2). The working principle of SVMs pre-classification and later regression prediction model is as follows: firstly, the first input sample is predicted by the SVR model about all training samples, then class (normal/fault) of the predictive value is identified. According to the classification result, the next operation will be carried out.

3. The experimental section

The composition of BLDCM is shown in Figure 3. It can be seen from the composition of BLDCM that the inverter circuit fault and the Hall sensor fault can be regarded as the typical faults because of their high fault ratio and great harm [7,8].
Figure 3. BLDCM and its control system composition diagram.

The three-phase winding current signals of BLDCM are shown as periodic signals. The value of current signal will change when faults occur, and the current signal features of different fault types are also different. The range, average value, current change rate, waveform index, peak value index, pulse index and margin index of current waveform are extracted as signal features. The single Hall sensor hard fault (short circuit or open circuit), the single MOSTET soft fault (channel width and length are increased by 50% respectively), the two MOSTET hard faults (two MOSTET on different bridge arms are in circumstances of short circuit or open circuit) and normal mode are selected as 4 working states. The single Hall sensor hard fault has 3 sets of fault modes, the single MOSTET soft fault has 6 sets of fault modes, while the two MOSTET hard faults have 12 sets of fault modes. Counting the normal mode in, there are 22 fault modes in total. The number of samples for each mode is 90, and the total number of samples for all modes is 22×90=1980.

3.1The experiments on performance about classification of the algorithms

OAA-SVMs, OAO-SVMs and clustering combine SVMs dynamic pruned binary tree algorithm are compared to show which is the best for circuit fault diagnosis of BLDCM, the four kernel functions are used, they are

- polynomial kernel function:
  \[ K(x, x_i) = [r(x \cdot x_i) + r]^d \quad r = 2; d = 2 \]  

- radial basis function:
  \[ K(x, x_i) = \exp(-\gamma \|x - x_i\|^2) \quad r = 3 \]  

- sigmoid kernel function:
  \[ K(x, x_i) = \tanh(\gamma (x \cdot x_i) + c) \quad r = 2, c = 2 \]  

- linear kernel function:
  \[ K(x, x_i) = x \cdot x_i \]  

“OAA-SVMs” is the earliest approach for multi-class SVMs. It constructs K binary SVMs to classify K-classes. The ith classifier is trained while labeling all the samples in the ith class as positive and the rest as negative. The classification speed is fast, but its training is the most computationally expensive because each SVM is optimized on all the N training samples.

“OAO-SVMs” use all the binary pairwise combinations of the K-classes. Thus, the model consists of K(K-1)/2 binary SVMs. The overall training speed is faster than “OAA-SVMs”, but the number of classifiers increases greatly along with the increasing of K, so the decision speed is slow.

The 660 sets of samples are used in the experiments, the classification time is measured in seconds and penalty parameter C=10. The classification accuracy and time of the algorithms are shown in Table 1.
4. Conclusion

In order to solve the problem about fault diagnosis and prediction of driving motor (BLDCM) for seeding metering of plot seeder, the pre-classification and later regression prediction algorithm is proposed in this paper, and then the simulation experiments are carried out. The experimental results show that: (1) comparing with “OAO-SVMs” and “OAA-SVMs”, the SVC algorithm (clustering

| Algorithm                  | kernel function       | Accuracy | Time | Accuracy | Time | Accuracy | Time | Accuracy | Time |
|---------------------------|-----------------------|----------|------|----------|------|----------|------|----------|------|
| OAA-SVMs                  | linear                | 0.8238   | 1.05 | 0.8956   | 1.17 | 0.9014   | 1.21 | 0.7570   | 0.87 |
| OAO-SVMs                  | RBF                   | 0.8741   | 1.12 | 0.9480   | 0.93 | 0.9370   | 0.96 | 0.7960   | 1.07 |
| SVMs pruned binary tree   | polynomial            | 0.9185   | 1.31 | 0.9637   | 1.02 | 0.9385   | 1.14 | 0.8647   | 0.95 |
|                           | sigmoid               |          |      |          |      |          |      |          |      |

Table 1. The classification accuracy and time of the algorithms

It is concluded that clustering combine SVMs dynamic pruned binary tree algorithm with RBF kernel function did better than "OAA-SVMs” and “OAO-SVMs” in the experiments, and it did not only reduce more classification time, but also have the highest classification accuracy. Though OAO-SVMs algorithm with RBF kernel function or polynomial kernel function even had shorter classification identification time, but as the volume of sample increased, the classification error rate became increasing. Hence, the clustering combine SVMs dynamic pruned binary tree algorithm is feasible to the circuit fault diagnosis of BLDCM.

3.2 The experiments on prediction accuracy of the algorithm models

The analyses of model generally include absolute numerical analysis and relative numerical analysis, Root-Mean-Square Error is absolute numerical analysis, Theil IC and Mean Absolute Percentage Error are relative numerical analysis. The prediction accuracy values of the algorithms are shown in Table 2.

| Model                      | MAPE    | Theil IC | RMSE   |
|----------------------------|---------|----------|--------|
| pre-classification and later regression | 2.540   | 0.0155   | 0.217  |
| unclassified and direct regression       | 2.814   | 0.0175   | 0.351  |

Table 2. The prediction accuracy values of the algorithms

(1) Mean Absolute Percentage Error (MAPE). It is generally believed that if MAPE value is smaller than 10, the prediction accuracy is relatively high. The MAPE value of SVMs pre-classification and later regression prediction algorithm is 2.540, which is smaller than 2.814, indicating that the prediction accuracy of SVMs pre-classification and later regression prediction algorithm is higher.

(2) Theil IC. Theil IC is generally between 0 and 1. The smaller value of Theil IC is, the smaller difference is between the fitting value and the real value, and the prediction accuracy is higher. As the results can be seen from Table 2, Theil IC value of SVMs pre-classification and later regression prediction algorithm is 0.0155, which is smaller than 0.0175 and much smaller than 1, indicating that the pre-classification and later regression prediction algorithm has higher prediction accuracy.

(3) Root-Mean-Square Error (RMSE). As the results can be seen from Table 2, the RMSE value of the pre-classification and later regression prediction algorithm is 0.217, which is smaller than 0.351, indicating that the deviation between predicted value and true value of SVMs pre-classification and later regression algorithm is smaller.
combine SVMs dynamic pruned binary tree) used in this paper has faster classification speed and higher classification accuracy; (2) the three evaluation indexes of SVMs pre-classification and later regression model all meet the criteria and are better than that of SVMs unclassified and direct regression model by absolute numerical and relative numerical analyses.

In conclusion, the SVMs pre-classification and later regression algorithm can solve the problem about fault prediction and diagnosis of BLDCM effectively and has a good application prospect in prediction of small size samples.

Acknowledgments
This work is funded by the National Science and Technology Support Program of China (2014BAD08B10), the Scientific Research Project of Heilongjiang Academy of Agricultural Sciences (2019YYF046) and Heilongjiang WOER Agricultural Equipment Technology Co. LTD.

References
[1] Lian ZG, Wang JG, Yang ZH, et al. (2010) Development of plot-sowing mechanization in China. Transactions of the CSAE, 28 (S.2) : 140-145.
[2] Shang SL, Yang RB, Yin YY, et al. (2010) Current situation and development trend of mechanization of field experiments. Transactions of the CSAE, 26 (S1): 5-8
[3] Vapnik VN. (2000) The Essence of Statistical Learning Theory. Trans. Zhang XG. Tsinghua University Press, Beijing.
[4] Zhang XG. (2000) On Statistical Learning Theory and Support Vector Machine. Journal of Automation. (01): 36-46.
[5] Zhu YP, Dai RW. (2005) Text Classifier Based on SVM Decision Tree. Pattern Recognition and Artificial Intelligence.18 (04): 412-416.
[6] AN Wang, YT Hou, Y Zhao, FY Zhao. (2010) Research on Fault Diagnosis Method of Blast Furnace Based on Clustering Combine SVMs Dynamic Pruned Binary Tree. In: IEEE International Conference on Networking, Sensing and Control. Chicago. pp. 67-70.
[7] Chai YL, Wang W, He WG. (2018) A fault-detection method for brushless DC motor. Small & Special Electrical Machine. 46 (6): 49-52.
[8] Liu K, Wang YH, Xu D. (2017) Fault handling study for hall position sensors for brushless DC motor. Power Electronics. 51 (7): 93-96.