

ANIMAL DISEASE DIAGNOSES EXPERT SYSTEM BASED ON SVM

Long Wan, Wenxing Bao

Department of Computer Science and Engineering, The North University for Ethnic, Yinchuan Ningxia, P. R. China 750021

**Corresponding author, Address: Department of Computer Science and Engineering, The North University for Ethnic, Yinchuan 750021, Ningxia, P. R. China, Email: oneloong@gmail.com*

Abstract: Livestock breeding farms usually distribute in remote areas, with relatively poor condition of disease diagnosis. Generally, it is hard to carry out disease diagnosis rapidly and accurately. But the farms can diagnose animal disease quickly and accurately by the animal disease diagnoses expert system. It could ensure a sound development of the stockbreeding industry. This paper proved the practicality of support vector machine (SVM) which is used in the animal disease diagnoses expert system in theory by studying the disease diagnosis expert system based on SVM. And the experiments proved that SVM can make the disease diagnose accurately.

Keywords: SVM, disease diagnosis, expert system

1. INTRODUCTION

Diagnosing animal disease quickly and accurately has the economic effectiveness. Livestock breeding farms usually is relatively poor condition of disease diagnosis. Generally, it is hard to carry out rapidly and accurately disease diagnosis. It requires veterinary experts a lot of experience and theoretical knowledge to diagnose the disease. The expert system for animal diseases diagnosing can meet the farms for the urgent needs of veterinary experts, since there are very few experts at the farms. SVM (Vapnik et al., 1998) is a new machine learning, which developed on the basis of VC

dimension of statistical theory and structural risk minimization (SRM). SVM has many advantages in the problems in dealing with the pattern recognition of limited samples, nonlinear and high dimensional. To get the best generalization ability, SVM can find the best compromise between the model complexity and the model study ability by limited samples information. In nature, animal disease diagnosis is a problem of pattern classification and recognition. It can judge the disease based on symptoms of animal suffered. In this paper, we designed the model of animal disease diagnoses expert system which was based on SVM and was used to diagnose the cow diseases. It shows that the method is practical and effective. And this practice provides a new approach for animal disease diagnosis.

2. ABOUT EXPERT SYSTEM

Expert System is an intelligent computer program. The intelligence mainly expressed that can mimic human experts' thinking to solve complex problems in the field (Li et al., 1996). Diagnosis expert system for animal diseases belongs to diagnosis expert system in modern medicine. Its mission is to quickly and accurately judge the type of disease from observed symptoms of the disease.

A typical disease diagnosis expert system is composed of knowledge base, reasoning engine, explanation unit, control strategies and human-machine interface, as shown in Fig 1.

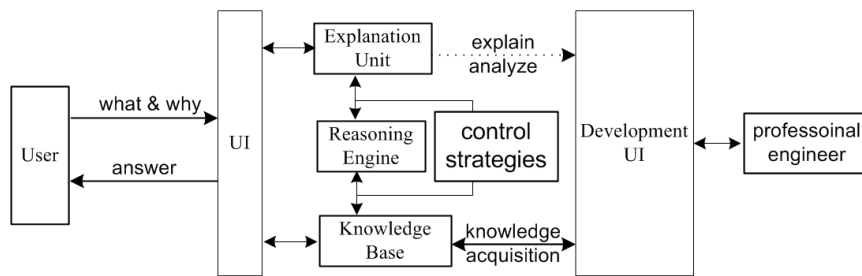


Fig 1. System structure of animal disease diagnosis

Knowledge base: it is the core of the system, which stores professional knowledge of disease diagnosis, the principle knowledge and the data of disease cases, etc. **Control strategy:** it decides that system should be taken what way to find the reasoning rules. **Reasoning engine:** it obtains the inputting symptoms of the disease and analyzing, reasoning them and getting the result of disease diagnosis. **Explanation unit:** in order to make users to understand the reasoning process, it is responsible for answering various

questions use raised. It gives an understandable and adequately explanation based on reasoning rules and reasoning conclusions. Human-machine interface: it is an interface of exchange information between the system and users.

Traditional rule-based expert system has its unique advantages: the description of rules is correspond with the process of knowing about problems and is more easily understood. However, over-reliance on rules to reason, the system has inevitable shortcomings: the rules of reasoning must be exact matches. It does not have the ability of self-learning and predictive, can not be self-improvement in the process of reasoning, and even lack effective treatment measures when there are some unavailable information in known information (Li et al., 1996). Some intelligence algorithm, such as neural networks and genetic algorithms, originate from the theory which is based on samples tended to be infinite, but the size of samples can not have a lot in practice. Hence the expert systems, taking these algorithms, can not get good performance in practice.

3. FUNDAMENTAL PRINCIPLE OF SVM

At present, machine learning is one of the basic tasks in artificial intelligence field and statistics theory is its main analysis method. Its basic idea was to find the hidden rules through training samples, and using these rules to predict unknown data. V.Vapnik and others proposed a kind of learning machine support vector machine, which is based on limited samples' statistics learning theory. According to SRM theory, SVM can avoid local smallest point and solve over fitting problem effectively, so it has good extending capability and accurate classification. The basic idea of SVM can be illuminated by two kinds of linearly separable cases (training sets). As showed in Fig 2, filled points and hollow points stand for two kinds of samples. If they are linearly separable, the result of machine learning is a hyperplane (a straight line at two dimensions) or called discriminant function, which can divide training samples into categories: positive and negative. Obviously, according to the requirements of experience risk minimize (ERM), there are infinite hyperplanes like that. For training samples, though some hyperplanes have good classification capability, have bad predicting capability, such as hyperplane P_1 in Fig 2. According to the requirements of SRM, the result of learning should be the optimized hyperplane P_0 that it can not only classify two kinds of training samples accurately but also make the biggest classification margin. Making the biggest classification margin is controlling extending capability, which is one of core ideas of SVM.

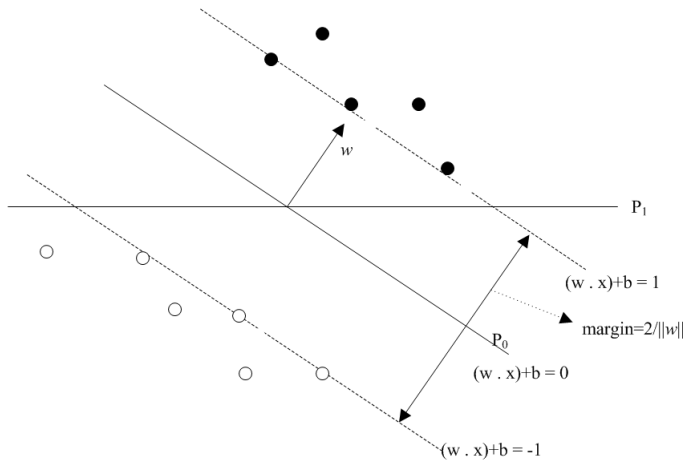


Fig 2. Classification hyperplane on linearly separable

4. ANIMAL DISEASE DIAGNOSIS EXPERT SYSTEM MODEL BASED ON SVM

4.1 Model Principle

The diagnosis methods based on disease symptoms still belong pattern classification problem virtually. It mainly searches the mapping relations between disease symptom and disease category. Animal disease diagnosis expert system based on SVM via training disease samples to obtain the mapping relations, and take disease symptom knowledge to implicitly represent on an optimize classification hyperplane in high dimensional features space. Moreover, the hidden rules among samples can be obtained and we can use them to diagnose diseases.

SVM was proposed for two classification problems primly, but it can be extended to multi-classification problems by adopting the combination method of binary classification. In this model, we choose the one-versus-one (OVO) (Blom et al., 1984, 1988) method, which has good classification capability. In processing of knowledge acquisition, we take the set, S :

$(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$, of disease samples provided by experts and users as the training samples of SVM, and $x_i \in R^n, y_i \in \{\text{set of disease pattern } (m)\}$. The method constructed decision functions among disease categories, which can construct $N = m(m-1)/2$ decision hyperplanes. Each decision function is trained by two categories of corresponding samples. Max-wins-voting strategy is adopted when a disease

sample x is classified. In the process if classification decision functions judge x belong to i category disease and the votes of i category disease add 1, otherwise, the votes of j category disease add 1. After N decision functions make the judgments, the disease category which has the most votes, is the diagnosis result of the unknown sample. For this model, we set a separate SVM for each disease according to the category of disease, and then all classified SVM are integrated, combined to a SVM classification group. The SVM group can reduce operation complexity of single SVM and improve reasoning efficiency and exactness of disease diagnosis.

4.2 Procedure of Animal Disease Diagnosis

The Fig 3 is the model of animal disease diagnosis based on SVM (MADDS)

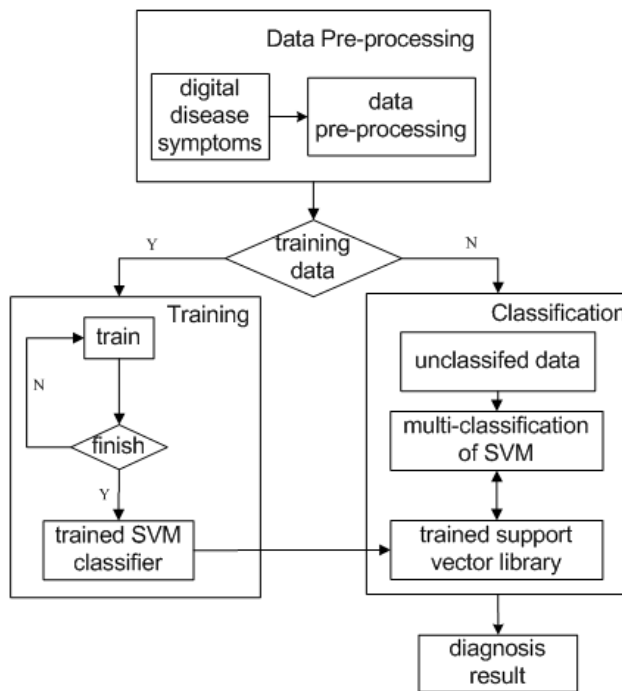


Fig 3. The model of animal disease diagnosis

The model is composed of data pre-processing module, training module and classification of multi-value module. First, the model obtains symptoms of the disease which users have observed. Then data pre-processing module carries out data preprocessing, and works out the input vector form which correspond to SVM. If the vectors are the training data, the model will train

them. And the results of training, a number of supporting vectors, will be saved in the knowledge vector library. The function of the three modules is as follows:

1. Data pre-processing module of MADDs:

In the procedure of disease diagnosis, characteristic extraction is the key step. It usually converts the description of the symptoms of disease into the digital information, which the computer can handle, using standardized criteria. In response to the features of disease diagnosis, we created a relation tree of disease-symptoms about disease and disease symptoms information of disease, as Fig 4 shows. All animal diseases information can be classified, arranged, and formed a unified knowledge statement when using the tree structure as the standard.

In the course of disease diagnosis, the symptoms of the user inputting will be converted into a unified data format by the module. In this procedure, after the model normalizes the data using some of the basic concepts of the data reliability analysis, it can reduce the differences between the sample data and improve the convergence performance of the classifier.

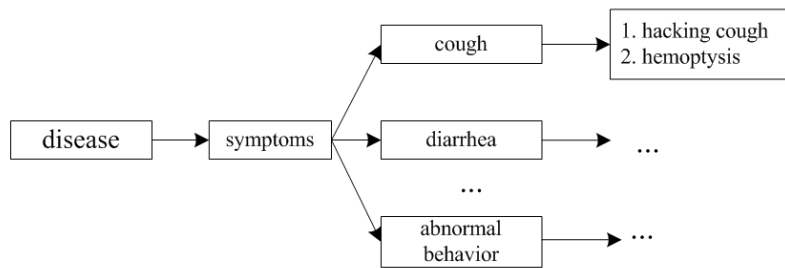


Fig 4. Relation trees of disease-symptoms

2. Training module of MADDs:

Since the diagnosis of animal diseases is a multi-classification problem, we constructed the multi-classification of SVM by the OVO method. The module chooses radial basis function as the kernel function of the classifier and used cross validation to train the data.

3. Multi-valued classification module of MADDs

The module will classify the unclassified data by using the trained support vector library.

5. EXPERIMENTAL STUDY AND RESULTS

In this experiment, we used some data of cow disease to test. According to cow disease symptoms categories, the symptoms were classified as 26

categories, corresponding to 26 components of the support vector machine. Each disease sample data of training and testing should be filled in the corresponding information according to the format of [Table 1](#).

Table 1. Vectors of disease symptom.

	1.psychiatric condition	2.body surface	3.physical features	...	26.behaviors
Sample (preprocessed data)	0.5	0.015625	0.00076293	...	0.015625

In the test, we collected five kinds of cow disease. Each disease had 35 cases. We used 20 cases of them to be train, and other 15 cases to be tested by Lib-SVM ([Chang et al., 2005](#)). The test's results are shown in [Table 2](#).

Table 2. Result of test.

Disease category	Number of training sample	Number of test sample	Number of accuracy	accuracy
cold	20	15	11	73.33%
malignant catarrhal fever	20	15	15	100%
rabies	20	15	13	86.67%
poxviruses	20	15	15	100%
ketosis in cattle	20	15	15	100%

Through the test, it can be found that the more detailed of symptoms was described, then after training sample cases, the more accurate the disease was diagnosed. This required the sample of cases should include as much as possible all the symptoms information of such disease. By the experimental verification, the model can make the diagnosis accurately and rapidly.

6. CONCLUSION

In this paper, we put forward a model of animal diseases diagnosis expert system based on SVM from the perspective of animal disease diagnosis. Because SVM is established on the principle of structural risk minimization, it has a strong ability of generalization. We attempt to make use of its strong generalization ability to resolve the difficulty of the rapid diagnosis processed because of the complexity and diversity of the symptoms of animal diseases. Experimental results showed that the model can be carried out animal diseases diagnosis more accurately, rapidly on the condition of small samples. It shows that animal disease diagnosis expert system based

on support vector machines is a good application in the field of diagnosis of animal diseases.

ACKNOWLEDGEMENTS

This work is supported by the National Key Technology R&D Program of China under Grant No. 2007BAD33B03.

REFERENCES

- Vapnik V. Statistical Learning Theory. New York: Wiley, 1998.
- Blom H A P. A Sophisticated Tracking Algorithm for ATC Surveillance Data. Proc. International Radar Conf. (Paris, France), 1984.
- Blom H A P, Bar-Shalom Y. The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients. IEEE Trans. Automatic Control, 1988, AC-33(8):780-783.
- Li X R. Hybrid Estimation Techniques. Control and Dynamic Systems: Advances in Theory and Applications (C. T. Leondes, ed.), New York: Academic Press, 1996.
- Chang Chih Chung, Lin Chih Jen. LIBSVM: a library for support vector machines, 2005.