An Enhanced Multiple SVM for Pattern Classification

Mansun Kim, Sangyong Lee
Computer Engineering Kongju National University
182 Shinkwan-dong, Kongju, Chungchongnam-do Republic of Korea, 314-701
email : {mansun, sylee}@kongju.ac.kr

Abstract - As it has been found that the use of multiple SVMs is more efficient than that of a single SVM in classifying patterns, there are many relevant researches in progress. However, classification based on multiple SVMs has problems such as long learning time, difficulties in linear separation among classes in expanding multiple classes, and class redundancy. In order to solve such problems in multiple SVMs, this study proposes an improved multiple SVM system and tests its performance using different data sets. From the results of the test, the author expects that the improved system minimizes input data by extracting input data, which reduces classification time significantly. It is also expected to improve the structure of multiple SVMs, solving difficulties in linear separation among classes and class redundancy. Moreover, the improved system may enhance efficiency in classification with less time and cost, and can be utilized widely from pattern recognition to data mining and machine learning.

Ⅰ. Introduction

The SVM proposed by Vapnik, a kind of standard feedforward, is entering upon a new turning point of researches concerning classification of patterns.
SVM is characterized as follows: first, it can identify only the elements affecting success of learning through simple and plain algorithm; second, it performs classification very well in actual application; third, it can minimize the chance of misclassification of a set of data with unknown probability distribution since it is based on statistical theories [1].
For all these advantages, it also has a defect of requiring a lot of time and capacity for learning Support Vector. We can save time and capacity by applying approximated algorithm in actual implementations to it but its classification performance decreases. To fix this defect, multi-SVM has been suggested.
A lot of researches on multi-SVM are in progress after its performance is more excellent than when a single SVM is used [2][3][4].

Ⅱ. Related study

This paper intended to point out problems related to the researches and suggest improved multi-SVM to overcome those problems.
The following researches have been in progress

after multi-SVM shows better performance than when a single SVM is used [5][6][7].

Scholkopf suggested multi-SVM with one-against-all method, applied it to vowel sound classification, and it showed 54.9%~79.6% classification rates. On the other hand, Clarkson suggested multi-SVM with one-against-one method, what the one-against-all method has been improved, and, as a result of experiments, it showed improved classification rates by 4~6% than the one-against-all method.

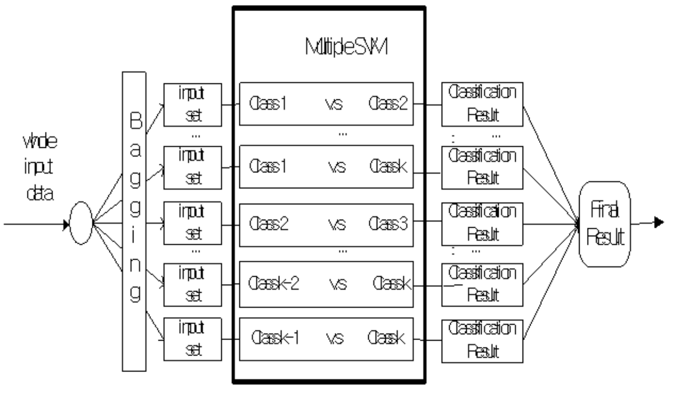
If we point out recent problems related to the researches: first, if the entire input data is used as it is when classifying large quantity of data, it will take too much time to learn when extending all of it into classes. Second, it is difficult to conduct linear classification between other classes or classes can be duplicated when extending into multi-class. Third, it is impossible to guarantee reliability about a final combination method of the multi-SVM, which is currently used.

Ⅲ. Discussion

For these reasons, I suggest improved multi-SVM as follows:

First, when classifying a large quantity of data, it has a problem of requiring too much time for learning if we use the entire input data for extending into classes as it is. Therefore, we should use minimum input data by extracting three of the entire input data [8]. This study used minimum input data by extracting three entire input data randomly.

Multi-SVM has a defect of requiring too much time for learning. To fix this defect, re-construct three input data so that you can draw a part of the entire data and then learn them independently. As shown in [Figure 1], draw the input data different from each other from the entire input data using bagging technique according to the number of SVM.



[Figure 1] A architecture of multi-SVM

Second, when extending multi-classes, it can be difficult to execute linear classification among other classes (in the case of one-against-all method) or classes can be overlapped (in the case of one-against-one method). For example, if it is divided into three classes, in the case of one-against-all method, it is possible to separate one class (class-1) from the other classes (class-2 and class-3) linearly, while it is impossible to separate the other two classes) class-2 and class-3) from each other linearly. The solution to this problem was the one-against-one method, but repetition occurred like class1-class2, class2-class3, and class3-class2.

We can expect more stable and better performance when multi-SVM is used than a single SVM. Therefore, this study suggests a multi-SVM model, which excludes the repetition from the one-against-one method, based on the previous research method on multi-SVM. For example, in [Figure 1], we need 12 SVMs if we don't exclude the repetition when K is 4, but total six SVMs including class1-class2, class1-class3, class1-class4, class2-class3, class2-class4, class3-class4, and class3-class4 are generated and the classification is in rapid progress after removing the repetition.

Third, there is a problem with a final combination method of the current multi-SVM. It requires a passing process of combining independently studied several SVMs using an appropriate method after the machine learning is finished.

The combination methods of end results are divided into Abstract Level, Rank Level, and Measurement Level. If we subdivide the methods, there are linear and non-linear methods. The former is to combine several SVMs using linear methods like majority voting and least square estimation (LSE), while the latter is to construct two-level SVM by placing other SVM in upper level while combining the SVMs of lower levels.

The majority voting method of Abstract Level, which is mainly used at present, can get good results only when each SVM has excellent performance. Additionally, the final combination method composed of two-level SVM proposed by Je [9] also has been suggested, but it can't guarantee end result reliability due to complicated structure that adds another SVM to upper level of SVMs to incorporate output of several SVMs.

In order to fix the majority combination method currently used for multi-SVM and solve the problem with the final combination method composed of two-level SVMs, we could get a better classification result, if we use the Measurement Level method using the classification reliability of all classes of each SVM.

Ⅳ. Simulation

Iris data has four attributes and three classes, and consists of fifty patterns for each class. As for the characteristics of the data, the attribute1 and attribute2 don't represent the characteristics of each class since a lot of patterns of each class are overlapped there, while third and fourth attributes represent them well.

Among 150 data, 90 data were randomly extracted and then designated a set of data for learning. In the case of bagging, 60 out of these data (20 for each class) are randomly extracted for learning.

To construct unit SVMs, classes, c, are organized as c(c-1)/2 to eliminate the repetition of Iris data composed of three classes. For example, the problem consisting of 10 classes like number recognition is organized as 45 unit SVMs.

After the machine learning is finished using a Matlab tool, the results of several independently learned unit SVMs are united using the calculated reliabilities through experiments and then final decisions are made.

Ⅴ. Conclusions

In general, when actually implementing SVM, the estimated algorithm is used to save computation time and capacity. As a result, it sometimes doesn't meet theoretical values of SVM's classification performance.

In order to overcome this problem, this paper discussed the multi-SVM composed of several unit SVMs.

It is expected to be able to enhance classification effectiveness by shortening the learning time through bagging and reducing the number of each classification.

For a follow-up study, the boosting method capable of extracting input data at random via probability distribution, which can be compared with a bagging method, is considered and it is necessary to study application of kernel functions and the effect of a few parameters on generalization.

REFERENCES

[1] C. Cortes and V. Vapnik, "Support vector netwo -rks," Machine Learning, vol.20, pp. 273-297, 1995.

[2] L. Hanson, P. Salamon, "Neural Network Ensembles", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol 12, pp. 993-1001, 1990.

[3] M. Haris, V. Ganapathy, "Neural Network Ensembles for financial for trend prediction", IEEE TENCON, Preceedings, vol 3, pp. 157-161, 2000.

[4] G.Giacinto, F.Roli, L. Bruzzone, "Combination of Neural and Statistical Algorithms", Pattern Recognition Letters, vol 21. issues5,2000,pp.385-397.

[5] G. Guo, S. Li and K. Chan, "Support vector machines for face recognition," Image and Vision Computing, Volume 19, Issues 9-10, pp. 631-638, 2001.

[6] Y. Cai and S. Lin, "Support vector machines for predicting rRNA-, RNA-, and DNA-binding proteins from amino acid sequence," Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics, Volume 1648, Issues 1-2, pp. 127-133, 2003.

[7] H. Drucker, B. Shahrary and D. Gibbon, "Support vector machines: relevance feedback and information retrieval," Information Processing and Management, Volume 38, Issue 3, 2002, pp. 305-323.

[8] Hong-Mo Je, Sung-Yang Bang, "Improving SVM Classification by Constructing Ensemble", Journal of KISS, Vol 30, No.3·4, pp. 251~258, 2003.

[9] Hong-Hee Won, Sung-Bae Cho, "Ensemble Classifier with Negatively Correlated Features for Cancer Classification", Journal of KISS, Vol 30, No.11·12, pp. 1124~1134, 2003.