UDC 004.021

MULTI-CLASS CLASSIFICATION SVM METHODS

ZHAO DI

Belarusian State University of Informatics and Radioelectronics, Republic of Belarus

Received March 9, 2022

Abstract. Support Vector Machine (SVM) was first proposed by Cortes and Vapnik in 1995. It is developed from the linear separable classification problem. The basic principle of SVM is to use a hyperplane to divide data into two categories, but in many classification problems, the sample is not linear, and the classification problem is multi-classification. We need to solve the multi-classification problem on a dichotomy basis.

Keywords: Support Vector Machine (SVM), multi-classification.

Introduction

Multi-class classification SVM has been applied in many fields, and there are two main categories of SVM ideas. One is to calculate all the classification decision functions simultaneously to solve the multi-classification problem at a time. However, the optimization problem solving process of this method is very complex, very computational and difficult to achieve, so it has not been widely used. The other is the decomposition of multi-classification problems into multiple binary classification problems, which are also divided into three categories: one-vs-one (OVO), one-vs-rest (OVR) and so on [1].

OVO multi-class classification SVM algorithm

The OVO multi-class classification SVM algorithm designs an SVM between any two classes of samples. When the unknown sample is classified, and classified by the voting method, the category with the most votes is the category of the unknown sample. Suppose samples have k categories, then we need to design binary SVM. When each sample was tested, each classifier judged it and output the category, and recorded the number of times judged as each category, and the category with the largest number of discriminants is the final category of the sample, the flowchart of this classification method shown as Figure 1.

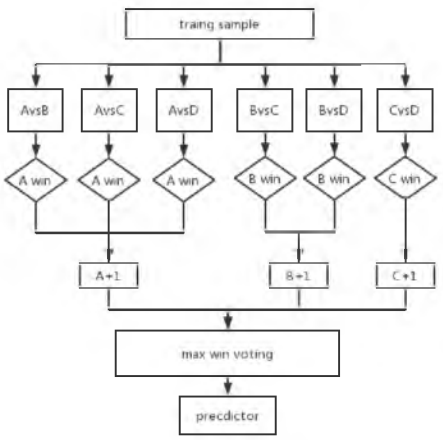

Figure 1. The flowchart of One-vs-One (OVO) multi-class classification SVM method

Pros of One-vs-One (OVO) multi-class classification SVM method:

– relative speed is faster when training individual models;

– when new samples are added without retraining all SVM, only retraining and increasing the sample-related classifier.

Cons of One-vs-One (OVO) multi-class classification SVM method:

– he number of binary classifiers required to construct and test regarding *k* into quadratic functions grows, and the total training time and test time are relatively slow.

Directed Acyclic Graph(DAG) multi-class classification SVM algorithm

From the OVO method, the Directed Acyclic Graph multi-class classification appears [2]. The training process is similar as OVO method, but the specific test process is shown in Figure 2.

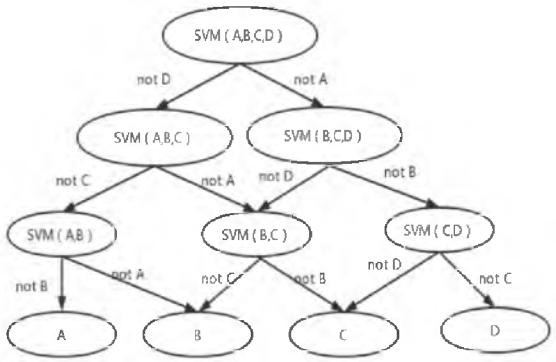

Figure 2. The flowchart of DAG multi-class classification SVM method

Pros of DAG multi-class classification SVM method:

– without having to traverse through all the dichotomized SVM, the decision speed is faster than the OVO voting method;

– it optimizes the training and decision time, is easier to calculate, learning effect is better;

– there is no inseparable region, and the sample number asymmetry problem is solved.

Cons of DAG multi-class classification SVM method:

– this method needs to construct more dichotomy SVM and slower training for classification problems with more samples.

OVR multi-class classification SVM algorithm

OVR multi-class classification is a strategy to compare the original and simple combined multi-classified SVM. Assuming a *k* classification problem, when training the classifier, the sample space is divided into categories and non, and the sample space is constructed into *k* SVM classifiers. In the test stage, the test samples were input into *k* SVM classifiers to compare the classification results of each classifier, and the classifier corresponding to the maximum classification function value was selected to determine the category of the sample, the flowchart of this classification method shown as Figure 3.

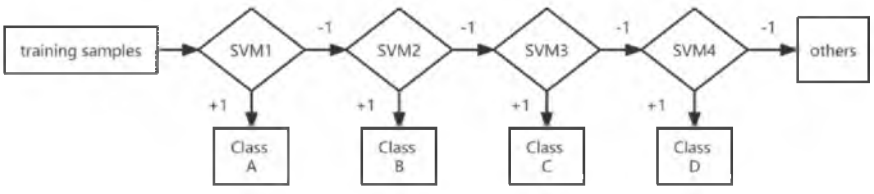

Figure 3. The flowchart of One-vs-Rest (OVR) multi-class classification SVM method

Pros and cons of OVR multi-class classification SVM method:

– this method just need k classifiers and its classification speed is relatively fast.

Cons of OVR multi-class classification SVM method:

– the training of each classifier takes all the samples as the training samples, so that when solving the quadratic planning problem, the training speed will slow down sharply as the number of training samples increases;

– because the data of negative sample is much greater than that of positive sample, sample asymmetry will appear, and this situation tends to be serious with the increase of training data;

– retraining of all models is required when new samples are added.

Decision Tree (DT) multi-class classification SVM algorithm

Following the OVR method, we derived decision tree multi-class classification SVM algorithm [2]. This method first divides all categories into two categories and further divides the subclasses into two secondary subclasses so that all nodes contain only a single category, which are also leaves of the binary tree species. This classification also decomposes the original classification problem into a series of two-class classification problems, where the classification function between the two subclasses adopts SVM. As shown in Figure 4.

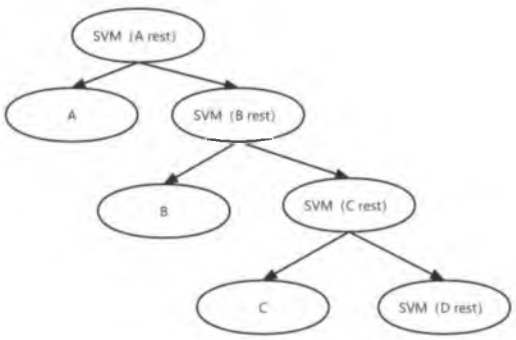

Figure 4. The flowchart of DT multi-class classification SVM method

Pros and cons of DT multi-class classification SVM method:

– the decision tree is easy to understand and interpret, visually analyze and extract rules;

– both nominal and numerical data;

– more suitable for samples with missing attributes;

– it is able to handle irrelevant features;

– faster operation for testing data sets;

– feasible and effective results for large data sources in a relatively short time.

Cons of OVR multi-class classification SVM method:

– easy to overfitting;

– easy to ignore the correlation of attributes in the dataset;

– different judgment criteria will bring different attribute selection tendencies in the decision tree. Retraining of all models is required when new samples are added.

Conclusion

For multi-classification problems, we can combine multiple binary classifiers into multi-classifiers to handle the multi-classification problem. The basic methods include OVO method and OVR method, which have different advantages. In our application, we should choose our own methods according to the characteristics of the data.

References

1. Fleury A., Vacher M., Nury N. // IEEE Trans. On ITB. 2010. Vol. 14. P. 274–283.
2. Sun Z., Fox G. // Proc. of PDPTA. 2012. P. 16–19.