http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (23): 7754-7757, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.7754.7757

A New Fuzzy Combination Method Based on Parametric Triangle Norm

Jia Pengtao and Liang Shuhui
School of Computer Science, Xi'an University of Science and Technology,
Xi'an, 710054, Shaanxi, China

Abstract: Triangular norms can improve the generalization capability of pattern classification problems, so they are introduced to improve the performance of ensemble learning. This study pays attention to the different impacts of base classifiers and put forward a new fuzzy combination method for ensemble learning. It is based on parametric triangle norms. Firstly, a new combination model was constructed. The base learners in it are set with different weights. Secondly, a set of fuzzy combination rules were generated according to the parametric triangle norms. Thirdly, genetic algorithm was used as the parameters estimation module of the new fuzzy rules. Finally, experiments were conducted on seven different datasets from the University of California, Irvine machine learning repository (UCI). The experimental results show that the fuzzy rules generated by the new combination method have better performance than the base learners and the fixed combination rules. And when set proper weights to base learners, the new fuzzy rules obtain better performance than the ones set a same weight to base learners.

Key words: Fuzzy combination rule, parametric triangle norms, different weights, genetic algorithm

INTRODUCTION

Ensemble learning is a very popular research topic of machine learning. Ensemble methods try to construct a set of hypotheses and combine them to use (Zhou, 2009). Typically, an ensemble has much stronger generalization ability than the base learners in it and it can lead to improved accuracy compared to a single classifier or regression mode (Tsoumakas *et al*., 2008).

A good ensemble lies in the diversity of base learners and the method for combining the results of base learners. There have already some fixed rules can be used to combine the results of base learners, such as product rule, mean rule, median rule, max rule, min rule, the simple voting method (majority voting) (Duda *et al*., 2001) or the weighted voting method (Ko *et al*., 2008), the stacking (stacked generalization) architecture (Chitra and Uma, 2010) and combining the nearest neighbor classifiers through multiple feature subsets (Bay, 1998). But most of them do not consider the different influence of the base learners. So, new rules should be found to concern the different effect of base learners and combine the results of base classifiers effectively.

As we all know, fuzzy theory is a widely used mathematical tool for combining the outputs of individual classifiers. Moreover, triangular norms are important family operations in fuzzy logic and they can improve the generalization capability of pattern classification problems. In reference (Farahbod and Eftekhari, 2012), triangular norms were used in the process of constructing the fuzzy rule based classification systems and altered the overall accuracy. So, in this study, they are introduced to construct the new fuzzy combination method of ensemble learning to improve the performance.

This study presents a new fuzzy combination method based on parametric triangular norms. This method can generate a set of new fuzzy combination rules for ensemble learning system. In experiments, five different triangular norms were applied separately for generating fuzzy rules. There are seven different datasets from the University of California, Irvine machine learning repository (UCI) were used for testing. Experimental results show that the new fuzzy combination rules based on triangular norms have better performance than the used base learners and fixed combination rules. When give appropriate weights to base learners, the new rules can combine them better and obtain higher accuracy than the operators with same weight.

FUZZY COMBINATION RULE BASED ON PARAMETRIC TRIANGLE NORMS

Triangular norms: According to reference (Mizumoto, 1989), a t-norm is a function T: $[0,1]\times[0,1]\rightarrow[0,1]$ which satisfies the following properties:

Corresponding Author: Jia Pengtao, School of Computer Science, Xi'an University of Science and Technology, Xi'an, 710054, Shaanxi, China

- Commutativity: T(x, y) = T(y, x)
- Monotonicity: T(x1, y1) = T(x2, y2) if x1 = x2 and y1 = y2
- Associativity: T(x, T(y, z)) = T(T(x, y), z)
- Boundary conditions: T(0, y) = 0, T(1, y) = y

T-norms can be constructed based on the following theorem:

Theorem 1: Let f: [0,1]→[0, +8] be a strictly decreasing function, so for all x, y in [0, 1], f(1) = 0 and f(x) + f(y) is in the range of f, equal to $f(0^+)$ or +8. Then the function T: [0,1]×[0,1]→[0,1] defined as $T(x, y) = f^{-1}(f(x)+f(y))$ is a t-norm.

The function f(x) in Theorem 1 is called additive generators of T(x, y).

According to the associativity: $t(x_1, x_1,..., x_n) = t(t_{n-1}(x_1, x_1,..., x_{n-1})\ x_n$, t-norms can be extended from binary operators to multivariate ones, so as to meet the needs of practical applications.

Then the multivariate parametric t-norm can be expressed as:

$$\begin{aligned}T(x_1,x_2,\cdots,x_n,p) &= f^{-1}(x,p)\\ &= f^{-1}(f(x_1,p)+f(x_2,p)+\cdots+f(x_n,p))\end{aligned} \tag{1}$$

where, p is the parameter.

Construct the new fuzzy combination method: Many existing fuzzy rules were binary operators with same weight. But the actual complex systems usually have more than two factors and different weights. So we construct a weighted multimember fuzzy combination method.

Assume the base classifier set is $T = \{C_1, C_2,\ldots,C_N\}$, the new combination method can be described as follows:

Step 1: Assume that the actual values are $y_0(t)$ (t = 1, 2,..., n), n is the number of samples in dataset; the values obtained by C_1, C_2, …, C_N are $y_1(t)$, $y_2(t)$, …, $y_N(t)$, respectively; the final values are y(t). The combination model can be described as:

$$y(t) = F(\alpha_1 y_1(t), \alpha_2 y_2(t), \cdots, \alpha_N y_N(t)) \tag{2}$$

where, F is the combination method, α_i is the weight of $y_i(t)$ and $\alpha_i \in [0, N]$ and:

$$\sum_{i=1}^{N} \alpha_i = N$$

Step 2: Based on the multivariate t-norm in Eq. 1, the combination method F is constructed as:

$$\begin{aligned}F &= f^{-1}(x)=\\ &f^{-1}(\alpha_1 y_1(t)+\alpha_2 y_2(t)+\cdots+\alpha_N y_N(t))\end{aligned} \tag{3}$$

When $f^{-1}(x)$ changes, it can generate different fuzzy combination rules. According to this variability, the fuzzy method gets good generalization ability.

Step 3: p, α_1, α_2,..., α_T are estimated by using genetic algorithm.

Genetic Algorithm (GA) is an adaptive global optimization search algorithm. It is formed by simulating the genetic and evolutionary process of organisms in the natural environment. Given its global optimization ability, GA is used as the parameters estimation module of the new rule.

Due to the objectivity and inevitability of the prediction error, there are errors between the predictive values y(t) and the actual ones $y_0(t)$. Set:

$$E = \sum_{t=1}^{n} (y(t) - y_0(t))^2 \tag{4}$$

Minimizing E is used as the evaluation of the objective function in genetic algorithm.

For comparison, the combination method with same weight can be described as follows:

$$y(t) = F(y_1(t), y_2(t), \cdots, y_N(t))$$

$$F = f^{-1}(y_1(t) + y_2(t) + \cdots + y_N(t))$$

Table 1 shows the used t-norms, the corresponding binary model with same weight and different weights.

In order to simplify the descriptions, the new Fuzzy Combination Method based on T-Norms is called FCMTN shortly. When different t-norms are used, there are different alias for the fuzzy rules. For example, when Aczel-Alsina t-norm is used, the rule with different weights is called FCMTN-d-AA, the rule with same weight is called FCMTN-s-AA. Other rules can be named in the similar way, as shown in the first column of Table 3.

EXPERIMENTS AND RESULTS

To test the effect of the new fuzzy combination method effectively, experiments are conducted with the help of the Pattern Recognition toolbox (PRtool) of MATLAB.

Datasets: The experiments considering seven datasets from the UCI repository of machine learning database (Newman *et al.*, 1998). For training and testing, every dataset is divided into three subsets: U_1, U_2 and U_3 randomly by using the GENDAT function.

Table 1: The used t-norms and the corresponding binary models

T-Norm	Model with same weight	Model with different weights
Aczel-Alsina	$e^{-(\lvert\log x\rvert^p+\lvert\log y\rvert^p)^{1/p}}$	$e^{-(\alpha\lvert\log x\rvert^p+\beta\lvert\log y\rvert^p)^{1/p}}$
Hamacher	$\frac{xy}{p+(1-p)(x+y-xy)}$	$\frac{p}{p-1+\left[\frac{p+(1-p)x}{x}\right]^{\alpha}\left[\frac{p+(1-p)y}{y}\right]^{\beta}}$
Schweizer-Sklar	$\sqrt[p]{\max(0,x^p+y^p-1)}$	$\sqrt[p]{\max(0,\alpha x^p+\beta y^p-1)}$
Sugeno-Weber	$\max(0,\frac{x+y-1+pxy}{p+1})$	$\max(0,\frac{(1+px)^{\alpha}(1+py)^{\beta}-(p+1)}{p(p+1)})$
Yager	$\max(0,1-\sqrt[p]{(1-x)^p+(1-y)^p})$	$\max(0,1-\sqrt[p]{\alpha(1-x)^p+\beta(1-y)^p})$

Table 2: Brief descriptions of datasets

NO.	Name	Attributes	Dataset size	Classes	U_1	U_2	U_3
1	Haberman	3	306	2	154	76	76
2	Iris	4	150	3	75	39	36
3	Liver	7	345	2	173	86	86
4	Pima	8	768	2	384	192	192
5	Statlog-heart	13	270	2	135	68	67
6	Wdbc	32	569	2	285	142	142
7	Wine	13	178	3	90	45	43

GENDAT function is a member of P Rtool and mainly used to generate datasets randomly.

U_1 is used to train the single classifiers and U_2 is used to test the single classifiers trained by U_1 and train the new ensemble rule and U_3 is used to test the single classifiers, the new ensemble rule and other combination rules. Table 2 shows these datasets and their descriptions.

Experimental procedure: The experiment was carried out according to the following steps:

Step 1: Normalize the values of datasets in interval [0, 1] for comparison according to the following formula. Because there is discrepancy in the sequential values of the datasets:

$$\text{norm}(x_i)=\frac{x_i-\min(X)}{\max(X)-\min(X)}$$

Step 2: Set N in Eq. 2 with three and choose three classifiers as base classifiers. Binary Decision Tree classifier, k-Nearest Neighbor classifier and Naive Bayes classifier are chosen and referred to as BDT, KNN and NB separately, where k is set with three

Step 3: U_1 is used to train the base classifiers, U_2 is used to test the base classifiers and to train the new fuzzy ensemble rules based on t-norms

Step 4: Using genetic algorithm to evaluate the parameters of the new fuzzy rule in U_2

Set the following parameters of GA for parameter optimization: the initial population is 20; use binary coding with eight numbers; select operation by using the uniform distribution random model; do crossover operation by the disperse cross; mutate operation by using gauss function.

Step 5: U_3 is used to test the single classifiers, the new fuzzy rule based on t-norms, the combination rule with same weight and some fixed combination rules. In order to validate our rule, we select product rule, mean rule, median rule, maximum rule, minimum rule (Kittler *et al*., 1998) and majority voting rule as comparison rules

Step 6: To compare all the results and analysis the effect of the new fuzzy rules

In order to avoid the randomness and locality of the results, we repeat the experiment for ten times and consider the mean-value of the ten experimental results as the final results for comparison.

RESULTS ANALYSIS AND DISCUSSION

The effect of FCMTN is investigated mainly focuses on the classification error rates. The results are shown in Table 3.

According to Table 3, the fuzzy combination rules based on t-norms have better performance than the base classifiers and fixed combination rules in many cases. For example, on the Haberman dataset, all the new fuzzy rules obtain lower error rates than base classifiers and fixed combination rules. And most of the new fuzzy rules perform better on Liver, Pima and Wdbc dataset.

When consider the different impacts of base classifiers, the new FCMTN can generate a set of fuzzy combination rules with improved performance. These fuzzy rules perform better than the ones that give same weight to base classifiers. Take the fuzzy rules based on Hamacher t-norm and Sugeno-Weber t-norm for instance, FCMTN-d-H obtain lower error rates than FCMTN-s-H on all the datasets except Wine and FCMTN-d-SW obtain lower error rates than FCMTN-s-SW on all the datasets. Moreover, FCMTN-d-AA improves the performance on three datasets, the results obtained by it on other datasets stay the same with the ones obtained by FCMTN-s-AA. FCMTN-d-SS has the similar performance with FCMTN-d-AA.

Table 3 Error rates of base classifiers and combination rules

Name	Haberman	Iris	Liver	Pima	Statlog-heart	Wdbc	Wine
BDT	0.2761	0.1081	0.4220	0.3375	0.3732	0.0688	0.1071
NB	0.2645	0.0696	0.3569	0.2479	0.1670	0.0575	0.0397
3NN	0.2945	0.0419	0.3362	0.3052	0.3656	0.0803	0.0281
product	0.2604	0.0558	0.3231	0.2463	0.1925	0.0546	0.0279
Mean	0.2671	0.0696	0.3581	0.2479	0.1670	0.0575	0.0281
median	0.2788	0.0805	0.3708	0.2907	0.3537	0.0661	0.0419
max	0.2645	0.0696	0.3569	0.2479	0.1670	0.0575	0.0327
min	0.2788	0.0915	0.3708	0.2907	0.3537	0.0661	0.0677
voting	0.2617	0.0502	0.3244	0.2698	0.2689	0.0518	0.0185
FCMTN-s-AA	0.2395	0.0472	0.1488	0.2380	0.1955	0.0310	0.0582
FCMTN-d-AA	0.2395	0.0222	0.1488	0.2380	0.1955	0.0232	0.0233
FCMTN-s-H	0.2592	0.0695	0.3233	0.2807	0.2925	0.0599	0.0698
FCMTN-d-H	0.2553	0.0389	0.2430	0.2359	0.2134	0.0451	0.0721
FCMTN-s-SS	0.2592	0.0667	0.3105	0.2807	0.2925	0.0599	0.0698
FCMTN-d-SS	0.2592	0.0556	0.2651	0.2807	0.2925	0.0599	0.0651
FCMTN-s-SW	0.2592	0.3000	0.3233	0.2708	0.2925	0.0599	0.3465
FCMTN-d-SW	0.2474	0.0695	0.1454	0.1714	0.1612	0.0289	0.0977
FCMTN-s-Y	0.1698	0.0445	0.0267	0.1099	0.1164	0.0310	0.0372
FCMTN-d-Y	0.1698	0.0472	0.0267	0.1115	0.0881	0.0317	0.0512

CONCLUSION

In practical applications, each base classifier in a ensemble plays a unique role, so proper weights of the base classifiers should be considered. On this basis, this study presents a new fuzzy combination method based on parametric triangular norms (FCMTN). It takes the different effect of base classifiers and the better generalization ability of t-norms into account.

When different t-norms are used, FCMTN can generate different fuzzy combination rules. The experimental results show that the FCMTN is suitable for combination in ensembles and the fuzzy combination rules generated by it can get better performance than the base classifiers and fixed combination rules. The new fuzzy rules that give proper weights to base classifiers obtain better performance than the ones whose base classifiers have same weight, especially when Hamacher t-norm and Sugeno-Weber t-norm are used.

ACKNOWLEDGMENT

This project is supported by Natural Science Basic Research Plan in Shaanxi Province of China (NO. 2012JQ8035) and the Scientific Research Program Funded by Shaanxi Provincial Education Department (NO. 2013JK0870).

REFERENCES

Bay, S.D., 1998. Combining Nearest Neighbor Classifiers through Multiple Feature Subsets. In: Proceeding 15th International Conference on Machine Learning, Shavlik, J. (Ed.). Morgan Kaufmann, San Francisco, pp: 37-45.

Chitra, A. and S. Uma, 2010. An ensemble model of multiple classifiers for time series prediction. Int. J. Comput. Theory Eng., 2: 454-458.

Duda, R.O., P.E. Hart and D.G. Stork, 2001. Pattern Classification. 2nd Edn., John Wiley and Sons, New York.

Farahbod, F. and M. Eftekhari, 2012. Comparison of different t-norm operators in classification problems. Int. J. Fuzzy Logic Syst., 2: 33-39.

Kittler, J., M. Hatef, R.P.W. Duin and J. Matas, 1998. On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell., 20: 226-239.

Ko, A.H.R., R. Sabourin and A.S. Britto Jr., 2008. From dynamic classi?er selection to dynamic ensemble selection. Pattern Recognit., 41: 1718-1731.

Mizumoto, M., 1989. Pictorial representations of fuzzy connectives, part I: Cases of t-norms, t-conorms and averaging operators. Fuzzy Sets Syst., 31: 217-242.

Newman, D.J., S. Hettich, C. Blake and C. Merz, 1998. UCI Repository of Machine Learning Databases. Department of Information and Computer Science, University of California, Berleley, CA.

Tsoumakas, G., I. Partalas and I. Vlahavas, 2008. A taxonomy and short review of ensemble selection. Proceedings of the 18th European Conference on Artificial Intelligence, July 21-25, 2008, Patras, Greece, pp: 1-6.

Zhou, Z.H., 2009. Ensemble Learning. In: Encyclopedia of Biometrics, Li, S.Z. and A. Jain (Eds.). Springer, Berlin, pp: 270-273.