Industrial Engineering
& Management Systems
Vol 11, No 1, Mar 2012, pp.26-29
ISSN 1598-7248 | EISSN 2234-6473 |

http://dx.doi.org/10.7232/iems.2012.11.1.026

Support Vector Machine Based on Type-2 Fuzzy Training Samples

Ming-hu Ha*
College of Mathematics and Computer Sciences, Hebei University,
College of Economy and Management, Hebei University of Engineering
Jia-ying Huang
College of Mathematics and Computer Sciences, Hebei University
Yang Yang
College of Mathematics and Computer Sciences, Hebei University
Chao Wang
College of Physics Science and Technology, Hebei University

(Received: November 17, 2011 / Revised: February 18, 2012 / Accepted: February 19, 2012)

ABSTRACT

In order to deal with the classification problems of type-2 fuzzy training samples on generalized credibility space. Firstly the type-2 fuzzy training samples are reduced to ordinary fuzzy samples by the mean reduction method. Secondly the definition of strong fuzzy linear separable data for type-2 fuzzy samples on generalized credibility space is introduced. Further, by utilizing fuzzy chance-constrained programming and classic support vector machine, a support vector machine based on type-2 fuzzy training samples and established on generalized credibility space is given. An example shows the efficiency of the support vector machine.

Keywords: Type-2 Fuzzy Training Samples, Mean Reduction Method, Fuzzy Chance-Constrained Programming, Support Vector Machine

* Corresponding Author, E-mail: minghuha@yahoo.com.cn

1. INTRODUCTION

The fuzzy set was proposed by Zadeh (1965) in 1965, it is an interesting extension of classical set. In classical sets the characteristic function can take only one of two values from the {0, 1}, in fuzzy sets the membership function can take any value from the interval [0, 1]. The membership function plays an important role in fuzzy set theory and its applications. But in some actual problems, it is often very difficult to determine the value of membership function (or membership degree for short). To solve this problem, the concept of type-2 fuzzy set was introduced by Zadeh (1975) in 1975. A type-2 fuzzy set is characterized by a special membership function, the value of which for each element of this set is a fuzzy set in [0, 1]. Compared with ordinary fuzzy set, type-2 fuzzy set could describe the objective phenomenon more accurately because of secondary membership function. It has been applied more efficiently to pattern recognition and other machine learning fields successfully (Mitchell, 2005).

Support vector machine built on statistical learning theory is a kind of effective machine learning method, has advantaged superiority in dealing with small samples classification problems, and has become the standard tool in machine learning field now (Cristianini and Shawe-Taylor, 2000). The traditional support vector machine is based on the real valued random samples and established on the probability space, and it is difficult to handle the small samples classification problems with non-real random samples on non-probability space. Naturally, it is a very interesting and valuable research direction to construct the support vector machine based on non-real valued random samples and established on non-probability space or probability space. On this basis, Lin and Wang (2002) proposed the support vector machine based on a non-real valued random samples-fuzzy samples, constructed the fuzzy support vector machine, which determines membership degree by different weight for each sample. Then Ha *et al*. (2009, 2011) improved the fuzzy support vector machine in determining membership degree, and they constituted a new fuzzy support vector machine and an intuitionistic fuzzy support vector machine. Considering that above support vector machines only make the samples fuzzified and the training samples are not fuzzy input directly, Ji *et al*. (2010) constructed a kind of support vector machine for classification based on fuzzy training data, and discussed its application in the diagnosis of coronary. Nevertheless the support vector machine is limited to the ordinary fuzzy number samples-the triangular fuzzy samples, and it is difficult to deal with the classification problems of type-2 fuzzy number training samples which is a generalization of the ordinary fuzzy number samples. In order to handle the problem, this paper firstly discusses the support vector machine based on type-2 fuzzy training samples and established on generalized credibility space.

2. PRELIMINARIES

2.1 The Type-2 Fuzzy Set

Definition 2.1(Zadeh, 1965): A fuzzy set $\tilde{A}$ in X is characterized by a mapping $\mu_{\tilde{A}}: X \rightarrow [0,1]$, $x \mapsto \mu_{\tilde{A}}(x)$, $\mu_{\tilde{A}}$ is called the membership function of $\tilde{A}$ and $\mu_{\tilde{A}}(x)$ is called the membership degree of x in $\tilde{A}$. The symbol $\tilde{A} = \int \frac{\tilde{A}(x)}{x}$ represents $\tilde{A}$. A collection of all fuzzy sets of X can be denoted by $\tilde{\mathcal{F}}(X)$.

Definition 2.2: Suppose $a, b, c \in R$. $\tilde{a}$ is a fuzzy set in R, and its membership function is as follows

$$\mu_{\tilde{a}}(x) = \begin{cases} \frac{x-a}{b-a} & a < x < b \\ 1 & x = b. \\ \frac{x-c}{b-c} & b < x < c \end{cases}$$

Then $\tilde{a}$ is called triangular fuzzy number, and is represented by $\tilde{a} = (a, b, c)$ (refer to Ha *et al*., 2010)

Definition 2.3(Zadeh, 1975): A type-2 fuzzy set in X, denoted $\bar{\tilde{A}}$, is characterized by a secondary membership function $\mu_{\bar{\tilde{A}}}(x, u)$, where $x \in X, u \in J_x \subseteq [0,1]$, and $\mu_{\bar{\tilde{A}}}(x) = \int_{u \in J_x} f(u)/u$. $\bar{\tilde{A}}$ can also be expressed as

$$\bar{\tilde{A}} = \int_{x \in X} [\int_{u \in J_x} f_x(u)/u]/x, \ J_x \subseteq [0,1].$$

Definition 2.4 (Qin *et al.*, 2011): A type-2 fuzzy set $\bar{\tilde{a}}$ is called triangular if its secondary membership function $\mu_{\bar{\tilde{A}}}(x,u)$ is

$$\left(\frac{x-r_1}{r_2-r_1} - \theta_l \min\{\frac{x-r_1}{r_2-r_1}, \frac{r_2-x}{r_2-r_1}\}, \frac{x-r_1}{r_2-r_1}, \frac{x-r_1}{r_2-r_1} + \theta_r \min\{\frac{x-r_1}{r_2-r_1}, \frac{r_2-x}{r_2-r_1}\}\right)$$

for $x \in [r_1, r_2]$, and

$$\left(\frac{r_3-x}{r_3-r_2} - \theta_l \min\{\frac{r_3-x}{r_3-r_2}, \frac{x-r_2}{r_3-r_2}\}, \frac{r_3-x}{r_3-r_2}, \frac{r_3-x}{r_3-r_2} + \theta_r \min\{\frac{r_3-x}{r_3-r_2}, \frac{x-r_2}{r_3-r_2}\}\right)$$

for $x \in [r_2, r_3]$, where $\theta_1, \theta_2 \in [0,1]$ are two parameters characterizing the degree of uncertainty that $\bar{\tilde{a}}$ takes the value x. For simplicity, we denote the type-2 triangular fuzzy number $\bar{\tilde{a}}$ with the above condition by $(r_1, r_2, r_3; \theta_l, \theta_r)$.

2.2 The Mean Reduction Method

Definition 2.5(Qin *et al.*, 2011): Let X be a nonempty set, $P(X)$ be the class of all subsets of X, Pos be a possibility measure in $P(X)$. Suppose $\tilde{a}$ is a fuzzy number, then the possibility measure of fuzzy event $\tilde{a} \le b$ is defined by $Pos(\{\tilde{a} \le b\}) = \sup\{\mu_{\tilde{a}}(x), x \in R, x \le b\}$. Similarly, $Pos(\{\tilde{a} < b\}) = \sup\{\mu_{\tilde{a}}(x), x \in R, x < b\}$, $Pos(\{\tilde{a} = b\}) = \mu_{\tilde{a}}(b)$.

Definition 2.6(Qin *et al.*, 2011): Let $\tilde{A}$ be a fuzzy set in X, and $\bar{\tilde{A}}$ be a type-2 fuzzy set in X. $\tilde{A}$ is the reduction of $\bar{\tilde{A}}$ via the mean reduction method, if it satisfies

$$\mu_{\tilde{A}}(x) = \int_0^{\infty} 1 + Pos\{\mu_{\bar{\tilde{A}}} \ge r\} - Pos\{\mu_{\bar{\tilde{A}}} < r\})dr - \int_{-\infty}^{0} 1 + Pos\{\mu_{\bar{\tilde{A}}} \le r\} - Pos\{\mu_{\bar{\tilde{A}}} > r\})dr$$

We denote $\tilde{A} = \varphi(\bar{\tilde{A}})$.

Example 2.1(Qin *et al.*, 2011): Let $\bar{\tilde{a}} = (r_1, r_2, r_3; \theta_l, \theta_r)$ be a type-2 triangular fuzzy number. According to the definition 2.6, we have the reduction $\tilde{a}$ of $\bar{\tilde{a}}$, and the membership function of $\tilde{a}$ is as follows.

$$\mu_{\tilde{a}}(x) = \begin{cases} \frac{(4+\theta_r-\theta_l)(x-r_1)}{4(r_2-r_1)} & r_1 \le x \le \frac{r_1+r_2}{2} \\ \frac{(4-\theta_r+\theta_l)x + (\theta_r-\theta_l)r_2 - 4r_1}{4(r_2-r_1)} & \frac{r_1+r_2}{2} < x \le r_2 \\ \frac{(-4+\theta_r-\theta_l)x + 4r_3 - (\theta_r-\theta_l)r_2}{4(r_3-r_2)} & r_2 < x \le \frac{r_2+r_3}{2} \\ \frac{(4+\theta_r-\theta_l)(r_3-x)}{4(r_3-r_2)} & \frac{r_2+r_3}{2} \le x \le r_3 \end{cases}$$

2.3 The Generalized Credibility Measure

Definition 2.7(Qin *et al.*, 2011): Suppose $\tilde{a}$ is a fuzzy number. The generalized credibility measure $\tilde{C}r$ of the event $\{\tilde{a} \le b\}$ is defined by $\tilde{C}r\{\tilde{a} \le b\} = \frac{1}{2}(\sup_{x\in R}\mu(x) + \sup_{x\ge r}\mu(x) - \sup_{x<r}\mu(x)),\ r \in R.$

Theorem 2.1(Qin *et al.*, 2011): *Let* $\tilde{a}_i$ *be the mean reduction of type-2 fuzzy triangular number* $\bar{\tilde{a}}_i = (r_1^i, r_2^i, r_3^i; \theta_{l,i}, \theta_{r,i}), i = 1, 2, \cdots, n.$ *Suppose* $\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_n$ *are mutually independent, and* $k_i \ge 0, i = 1, 2, \cdots, n$. *Given the generalized credibility level* $\lambda \in (\frac{12-\theta_{r,1}-\theta_{l,1}}{16}, 1]$, *if* $\theta_{r,1} - \theta_{l,1} \le \theta_{r,2} - \theta_{l,2} \le \cdots \le \theta_{r,i} - \theta_{l,i}$, *then* $\tilde{C}r\{\sum_{i=1}^{n} k_i\tilde{a}_i \le t\} \ge \lambda$ *is equivalent to*

$$\sum_{i=1}^{n} k_i \frac{(4(2\lambda-1)+\theta_{r,i}-\theta_{l,i})r_3^i + 8(1-\lambda)r_2^i}{4+\theta_{r,i}-\theta_{l,i}} \le t.$$

Proof: It can be easily got by type-2 fuzzy set theory in the reference.

3. SUPPORT VECTOR MACHINE BASED ON TYPE-2 FUZZY TRAINING SAMPLES AND APPLICATION EXAMPLE

3.1 Support Vector Machine Based on Type-2 Fuzzy Training Samples

Consider the fuzzy type-2 training samples set $S = \{(\bar{\tilde{X}}_1, y_1), (\bar{\tilde{X}}_2, y_2), \cdots, (\bar{\tilde{X}}_m, y_m)\}$, where $\bar{\tilde{X}}_i = (\bar{\tilde{x}}_{i1}, \bar{\tilde{x}}_{i2}, \cdots, \bar{\tilde{x}}_{in})$, and $\bar{\tilde{x}}_{ij} = (r_1^{ij}, r_2^{ij}, r_3^{ij}; \theta_{r,j}^i, \theta_{l,j}^i), i = 1, 2, \cdots, m, j = 1, 2, \cdots, n$, when $y_i = 1$, then $(\bar{\tilde{X}}_i, y_i)$ is called a positive class; when $y_i = -1$, then $(\bar{\tilde{X}}_i, y_i)$ is called a negative class. The classification based on the type-2 fuzzy training set $S = \{(\bar{\tilde{X}}_1, y_1), (\bar{\tilde{X}}_2, y_2), \cdots, (\bar{\tilde{X}}_m, y_m)\}$ is to find a decision function $g(\bar{\tilde{X}})$, such that the positive class and the negative class can be separated with the low classification error and good generalization performance.

Definition 3.1(Ji *et al.*, 2010): For the type-2 fuzzy training samples set $S = \{(\bar{\tilde{X}}_1, y_1), (\bar{\tilde{X}}_2, y_2), \cdots, (\bar{\tilde{X}}_m, y_m)\}$, if for a given level $\lambda \in \left(\frac{12-\theta_{r,1}-\theta_{l,1}}{16}, 1\right]$, there exist $w \in R^n, b \in R$, such that

$$\bar{C}r\left\{y_i\left(\left(w \cdot \varphi(\bar{\tilde{X}}_i)\right) + b\right) \ge 1\right\} \ge \lambda(i = 1, 2, \cdots, m).$$

Then the type-2 fuzzy training samples set $S = \{(\bar{\tilde{X}}_1, y_1), (\bar{\tilde{X}}_2, y_2), \cdots, (\bar{\tilde{X}}_m, y_m)\}$ is strong type-2 fuzzy linear separable.

The support vector machine for strong linear separable type-2 fuzzy sample set is to solve the fuzzy chance constrained programming:

$$\begin{cases} \min & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\xi_i \\ s.t. & \bar{C}r\left\{y_i\left(\left(w \cdot \varphi(\bar{\tilde{X}}_i)\right) + b\right) \ge 1 - \xi_i\right\} \ge \lambda(i = 1, 2, \cdots, m) \\ & \xi_i \ge 0(i = 1, 2, \cdots, m) \end{cases}$$

To solve the above programming,

$$\bar{C}r\left\{y_i\left(\left(w \cdot \varphi(\bar{\tilde{X}}_i)\right) + b\right) \ge 1 - \xi_i\right\} \ge \lambda(i = 1, 2, \cdots, m)$$

is equivalent to

$$\sum_{j=1}^{n} -y_i w_i \frac{(4(2\lambda-1)+\theta_{r,j}^i-\theta_{l,j}^i)r_3^{ij} + 8(1-\lambda)r_2^{ij}}{4+\theta_{r,j}^i-\theta_{l,j}^i} \le y_i b + \xi_i - 1,$$

$i = 1, 2, \cdots, m,\ j = 1, 2, \cdots, n$ according to Theorem 2.1. Then the fuzzy chance constrained programming is converted into the following classic convex quadratic programming:

$$\begin{cases} \min & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\xi_i \\ s.t. & y_i\left((w \cdot T_i) + b\right) \ge 1 - \xi_i\ (i = 1, 2, \cdots, m). \\ & \xi_i \ge 0\ (i = 1, 2, \cdots, m) \end{cases}$$

We can obtain its dual problem:

$$\begin{cases} \max & L(\alpha) = \sum_{j=1}^{m}\alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} y_i y_j \alpha_i \alpha_j (T_i \cdot T_j) \\ s.t. & 0 \le \alpha_i \le C, \sum_{j=1}^{m} y_i\alpha_i = 0(i = 1, 2, \cdots, m). \end{cases}$$

w^*, b^* are the solution of above dual programming. For a given confidence level $\lambda \in (\frac{12-\theta_r^*-\theta_l^*}{16}, 1]$, if $\tilde{C}r\left\{\left(w^* \cdot \varphi(\bar{\tilde{X}}_i)\right) + b^* \ge 0\right\} \ge \lambda$, then $(\bar{\tilde{X}}_i, y_i)$ belongs to the positive class; if $\tilde{C}r\left\{\left(w^* \cdot \varphi(\bar{\tilde{X}}_i)\right) + b^* \le 0\right\} \ge \lambda$, then $(\bar{\tilde{X}}_i, y_i)$ belongs to the negative class.

3.2 An Example

We shall apply the support vector machine for two-class classification with type-2 fuzzy training samples to the diagnosis of Coronary. The data is sourced from the reference (Ji *et al.*, 2010), the half of which are healthy

($y_i = 1$), the others are Coronary patients ($y_i = -1$). Let $\theta_{l,1}{}^{i} = \theta_{r,1}{}^{i} = 0.1$, $\theta_{l,2}{}^{i} = \theta_{r,2}{}^{i} = 0.2$, we can obtain type-2 triangular fuzzy numbers $\bar{\bar{x}}_{i1} = (r_1^{i1}, r_2^{i1}, r_3^{i1}; \theta_{l,1}{}^{i}, \theta_{r,1}{}^{i})$ and $\bar{\bar{x}}_{i2} = (r_1^{i2}, r_2^{i2}, r_3^{i2}; \theta_{l,2}{}^{i}, \theta_{r,2}{}^{i})$. The programming was solved by SVM toolbox of matlab, when parameter $C = 1$, $\lambda = 0.8$. The accurate rate of diagnosis is 80%. The result of example illustrates that the proposed support vector machine is effective.

4. CONCLUSIONS

This paper constructs firstly the support vector machine in which the training samples is type-2 fuzzy number input, and the proposed support vector machine is an interesting extension of the traditional support vector machine and the fuzzy support vector machine. We shall focus on the proposed support vector machine's applications in some practical problems.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 61073121 and No. 60773062) and the Natural Science Foundation of Hebei Province (No. F2008000633).

REFERENCES

Cristianini, N. and Shawe-Taylor, J. (2000), *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, Cambridge.

Ha, M. H., Peng, G. B., Zhao, Q. H., and Ma, L. J. (2009), A New fuzzy support vector machine, *Computer Engineering and Applications*, **45**, 151-153.

Ha, M. H., Wang, C., Zhang, Z. M., and Tian, D. Z. (2010), *Uncertainty Statistical Learning Theory*, Science Press, Beijing.

Ha, M. H., Huang, S., Wang, C., and Wang, X. L. (2011), Intuitionistic fuzzy support vector machine, *Journal of Hebei University (Natural Science Edition)*, **3**, 225-229.

Ji, A. B., Pang, J. H., and Qiu, H. J. (2010), Support vector machine for classification based on fuzzy training data, *Expert Systems with Applications*, **37**, 3495-3498.

Liu, B. D. (2002), *Theory and Practice of Uncertain Programming*, Springer-Verlag, Heidelberg.

Lin, C. F. and Wang, S. D. (2002), Fuzzy support vector machines, *IEEE Transaction on Neural Networks*, **13**, 464-471.

Mitchell, H. B. (2005), Pattern recognition using type-II fuzzy sets, *Information Sciences*, **2**, 409-418.

Qin, R., Liu, Y. K., and Liu, Z. Q. (2011), Critical values reduction methods for type-2 fuzzy variables and their applications, *Journal of Computational and Applied Mathematics*, **1**, 1454-1481.

Qin, R., Liu, Y. K., and Liu, Z. Q. (2011), Modeling fuzzy data envelopment analysis by parametric programming method, *Expert Systems with Applications*, **38**, 8648-8663.

Vapnik, V. N. (1995), *The Nature of Statistical Learning Theory*, Springer-Verlag, New York.

Zadeh, L. A. (1965), Fuzzy sets, *Information and Control*, **8**, 338-353.

Zadeh, L. A. (1975), The concept of a linguistic variable and its application to approximate reasoning-I, *Information Sciences*, **8**, 199-249.