

Prediction Algorithm for the Capability of Health Service Support Based on Support Vector Machine

Wen-Ming ZHOU^{1, a}, Yong-Yue FAN^{2, b}, Wen-Xiang XIA^{3, c}
and Chun-Rong ZHANG^{4, d}

^{1, 2, 3}*Simulation Training Center, The Logistics Academy, Taiping Rd, No.23, Beijing 100858, China*

⁴*Department of Material and Oil, the Second Artillery, Beijing 100085, China*

^a*zwmedu@163.com*, ^b*fanyyedu@163.com*, ^c*xiawxedu@163.com* ^d*zhangcr edu@163.com*

Abstract.

The health service support is very important for maintaining and enforcing effective military personal power to consolidate and improve battle effectiveness. Key links and main influencing factors of the health service support are analyzed. The index system of efficiency evaluation is presented. Due to the merits of the effective non-line method established on statistics and characteristics of generalization for support vector machine (SVM), propose a prediction algorithm for the capability of health service support based on it. The logistics command decision and system optimum design of the health service support are strongly supported by the prediction algorithm.

Keywords: Support Vector Machine (SVM), Health Service Support, Pattern Classification, Prediction.

1 Introduction

Health service support is to prevent and cure the wounded and illness of military person, execute health service organized by military health service agency in the army. It is an important part of the logistics support [1]. Its basic task is to maintain and enforce effective military personal power, consolidate and improve battle effectiveness, ensure completing training and combating from constitution and psychology. Health service support affects the process of the ware, even the result.

Currently, study reports about health service support simulation and evaluation is very less, joint theater-level simulation (JTLS) [2] and joint warfare system (JWARS) [3] of American only simply described and simulated health service

support, it couldn't embody the complex and law of it in modern information war. Due to the merits of the effective non-line method established on statistics, characteristics of generalization, good classification precision for support vector machine (SVM), it is widely used in pattern recognition and classification evaluation. Through analyzing key links, main influencing factors and main proceeding of health service support, this paper establishes a prediction algorithm for the capability of health service support. Evaluation examples of ten units are presented to verify its rationality and validity.

2 SVM Modeling

SVM [4] established on statistical learning theory is an effective method for solving non-linear problems. Even through less training samples, it can also get the characteristics of generalization. SVM make sure structural risk minimization (SRM) according to Vapnik and Chervonenkis (VC) dimension theory rather than other machine learning algorithms only thinking empirical risk minimization (ERM).

Basic Theory of SVM

Supposed a training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^l$, where $\mathbf{x}_i \in \mathbf{R}^n$ $y_i \in \{\pm 1\}$ $i = 1 \dots l$. The algorithm based on SVM substantively is a process of optimizing weight parameters. The purpose is to ensure SRM, and maximization of the distance between classification hyper planes. When training set is linear and classifiable, exist (\mathbf{w}, b) , we can get that:

$$\mathbf{w} \cdot \mathbf{x}_i + b \geq 1, \forall \mathbf{x}_i \in A \quad (1)$$

$$\mathbf{w} \cdot \mathbf{x}_i + b \leq -1, \forall \mathbf{x}_i \in B \quad (2)$$

Here A, B are results of classification. Finally get that:

$$f_{\mathbf{w},b} = \text{sign}(\mathbf{w} \cdot \mathbf{x} + b) \quad (3)$$

Translate to optimization problem as follows:

$$\text{Max} \frac{1}{\|\mathbf{w}\|} \text{ or } \text{Min} \frac{1}{2} \|\mathbf{w}\|^2 \quad \text{s.t.} \quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad i = 1, \dots, l \quad (4)$$

Construct Lagrange function (5) to solve it.

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^l \alpha_i [y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1] \quad (5)$$

Give the answer to the optimization problem by the saddle point of Lagrange function. Making decision function is transfer to that:

$$f(x) = \text{sign}\left(\sum_{i=1}^l y_i \alpha_i^* (\mathbf{x} \cdot \mathbf{x}_i) + b^*\right) \quad (6)$$

Where $b^* = y_i - \mathbf{w}^* \cdot \mathbf{x}_i$ $\mathbf{w}^* = \sum_{i=1}^l \alpha_i^* y_i \mathbf{x}_i$ α_i^* is the optimal solution to dual quadratic programming as follows:

$$\begin{aligned} \text{Maximize } F(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad \text{s.t.} \\ \sum_{i=1}^l \alpha_i y_i = 0, \alpha_i \geq 0 \end{aligned} \quad (7)$$

Use slack variable set and penalty factor to solve line non-classification problem. Map training samples x to a certain higher dimensional characteristic-space $\phi(x)$, and excute line classification in the space. Thus line non-classification and non-line problems can be solved.

$$\text{Max}F(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j [\phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)] \quad (8)$$

Supposed $\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$ function (8) transfer to that:

$$\text{Max}F(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \mathbf{K}(\mathbf{x}_i, \mathbf{x}_j) \quad (9)$$

Making decision function is transfer to that:

$$f(x) = \text{sign}\left(\sum_{i=1}^l y_i \alpha_i^* \mathbf{K}(\mathbf{x} \cdot \mathbf{x}_i) + b^*\right) \quad (10)$$

Where $\mathbf{K}(\mathbf{x} \cdot \mathbf{x}_i)$ is kernel function. Use it to substitute for vector product in higher dimensional characteristic-space. Three kind of kernel function is as follows: line kernel function $\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = \mathbf{x}^T \mathbf{x}_i$; polynomial kernel function

$\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = (\gamma \mathbf{x}^T \mathbf{x}_i + r)^p, \gamma > 0$; radial basis function

$\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = \exp(-\gamma \|\mathbf{x} - \mathbf{x}_i\|^2), \gamma > 0$

Svm training focus on solving the problem of quadratic programming. This paper use sequential minimal Optimization (SMO) put forward by Platt [5].

Sort Making Decision Model of SVM

Use dichotomy SVM to construct multi-element sort method as follows: one against one, it need $N(N-1)/2!$ designed classifiers; one against rest, it also need n classifiers. This paper combine SVM with binary tree making decision method [6], and it need $n-1$ classifiers as figure 1.

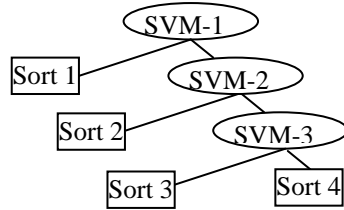


Fig.1. Multielement sorter

3 Factor Analysis of Health Service Support

Health service support is a multi-hierarchy, multi-factor, multi-index complex system. Its support element and feature index are a little much, and it easy bring dimension disaster. Considering the principle of systematicness, objectivity, completeness, measurability and independence during establishing index system, this paper analysis the process of health service support [1,7], adopt similarity compress and singularity compress of information theory to optimize the index [8]. The optimized index system is as figure 2.

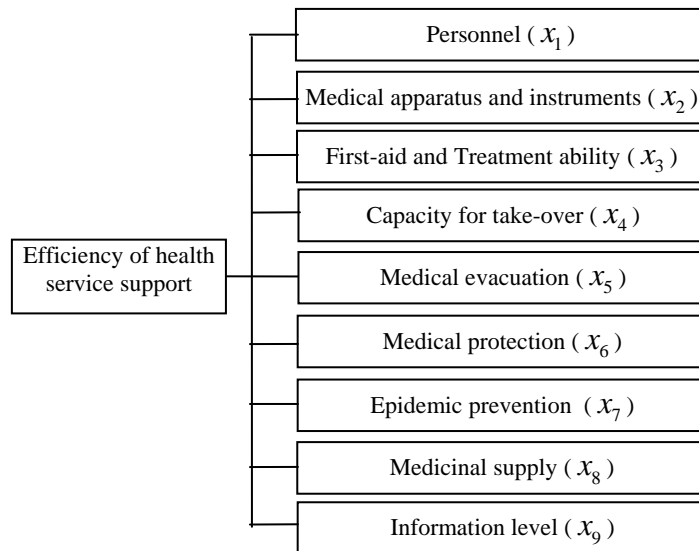


Fig.2. Efficiency index of health service support

4 The Method for Computing Index Value

According to designed index system, adopt three kind of method to compute index value as follows: operations research method, mathematics definition and examination

Operations Research Method

The factor of staff always play key role in all support. Staff with higher degree and stronger capability will be a multiplier to efficiency of health service support. Personnel expressed as x_1 , refer to [9], through three indexes of the second level and twenty indexes of the third level, adopt fuzzy comprehensive evaluation to calculate the value of index x_1 .

Information level expressed as x_0 , include level of command, control and communication information system, level of specialist knowledge database, information management level of medical warehouse etc. Its value is also calculated through fuzzy comprehensive evaluation [10].

Mathematics Definition

Index of Medical apparatus and instruments expressed as x_2 include advanced technology level x_{21} , percentage of complement (x_{22}), undamaged rate (x_{23}) etc.

Definition1. The express x_{22} is defined as the rate of practical number of medical apparatus and instruments (n_v) to the allocated number (n_w). It is as follows:

$$x_{22} = \begin{cases} \frac{n_v}{n_w}, n_v \leq n_w \\ 1, n_v > n_w \end{cases} \quad (11)$$

Definition2. The express x_{23} is defined as the rate of the undamaged number (n_1) to the allocated number (n_2). It is as follows:

$$x_{23} = \begin{cases} \frac{n_1}{n_2}, n_1 \leq n_2 \\ 1, n_1 > n_2 \end{cases} \quad (12)$$

The value x_2 of can be calculated by analytic hierarchy process.

Definition3. Medical evacuation is a fuzzy index (x_5). It is defined as a synthesis by transportation facilities (m_1), transportation manner (m_2), transportation

environment (m_3), transportation distance (m_4), transportation object (m_5) etc. m_i ($i = 1, 2, \dots, 5$) is confirmed by specialists.

$$x_5 = \sum_{i=1}^5 a_i m_i \quad (13)$$

Where a_i is weight coefficient, and $a_1 + a_2 + \dots + a_5 = 1$.

Examination

According to examination and inspection method of different unit, the first aid and treatment ability (x_3), capacity for take-over (x_4), medical protection (x_6), epidemic prevention (x_7) and medical supply (x_8) are all calculated their values by examination and inspection. Where x_3 include organization factor of medical staff and medical supply; x_4 mainly consider the factors of number of ward bed and hand frame; x_6 mainly consider the protection capability for nuclear weapon, biological weapon, chemical weapon and new concept weapon, including factors of medical reconnaissance, protection means and mastery of all kinds of weapon mechanics; x_7 mainly consider the epidemic reconnaissance, vaccine inoculation and epidemic situation management; x_8 mainly consider medical material acquisition, storage and management, carrying and moving load and so on.

5 Classification Pattern

Set up classification prediction pattern of health service support with Delphi method, classify and evaluate and synthesize the nine indexes respectively. Full mark is set according to the thousand mark examination system. According to performance, it is sorted as four grades: excellent (A), good (B), moderate(C) and inferior (D), as can be seen in table 1.

Table1. Index classification pattern of health service support

Comprehensive index	Four grade pattern			
	D	C	B	A
Personnel	[0 300]	[300 350]	[350 400]	[400 450]
Medical apparatus and	[0 35]	[35 40]	[40 45]	[45 50]
First-aid and treatment ability	[0 35]	[35 40]	[40 45]	[45 50]
Capacity for take-over	[0 35]	[35 40]	[40 45]	[45 50]
Medical evacuation	[0 35]	[35 40]	[40 45]	[45 50]
Medical protection	[0 85]	[85 90]	[90 95]	[95 100]

Epidemic prevention	[0 85]	[85 90]	[90 95]	[95 100]
Medicinal supply	[0 85]	[85 90]	[90 95]	[95 100]
Information level	[0 35]	[35 40]	[40 45]	[45 50]

6 Prediction Example

To set up a stable SVM prediction model, examination data of units are collected and processed. Sixty two samples are acquired. Fifty two of the samples are as training samples to set up the prediction model, the rest ten are as test samples to verify it. Due to the limit to paper length, only part training and test samples can be seen in table 2 and table 3. Visual training data is given for comparing the difference between respective sample-values as figure 3.

In order to eliminate data level and reduce error of prediction, this paper adopts maximum-minimum normalization method to normalize the samples data as follows:

$$f : x \rightarrow y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (14)$$

Where $x, y \in R^n$, $x_{\min} = \min(x)$, $x_{\max} = \max(x)$

Table2. The value of training samples

Index Sample	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
1	268	37	35	39	36	86	86	88	35
2	282	38	36	36	36	85	87	87	38
3	277	41	37	35	37	86	85	86	38
4	411	34	37	36	35	87	81	88	36
5	311	35	38	37	38	87	80	89	36
6	301	36	33	32	40	86	89	86	37
7	369	42	45	45	49	94	94	92	44
8	322	34	40	38	39	88	89	90	37
9	401	40	39	39	41	89	90	88	40
10	428	47	48	50	49	98	98	97	48
...
52	423	45	46	51	41	87	79	87	45

Table3. The value of test samples

Index \ Sample	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
1	271	38	36	39	37	86	86	87	36
2	278	38	37	37	36	86	88	88	38
3	277	41	36	35	37	87	85	87	37
4	401	33	36	34	35	86	80	88	36
5	312	37	39	36	39	88	81	89	37
6	301	36	34	33	40	85	89	87	36
7	368	42	46	45	48	96	95	93	45
8	323	36	39	39	40	87	89	91	37
9	398	39	40	39	42	90	91	88	39
10	431	48	47	50	49	98	98	96	48

This paper adopts SOM method to train the four sort making decision pattern and get the optimization making decision hyperplane. Adopt radial basis function as kernel function as follows:

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{|\mathbf{x} - \mathbf{x}_i|^2}{\sigma^2}\right) \quad (15)$$

Input test samples value to the trained prediction model, the classification result is as figure .4, among ten test sample, only one is classified error, accuracy is up to ninety percent.

7 Conclusion

According to the result of prediction example, the proposed prediction method for the capability of health service support is valid and applicable. The logistics command decision making and system optimum design of the health service support information system will be strongly supported by it.

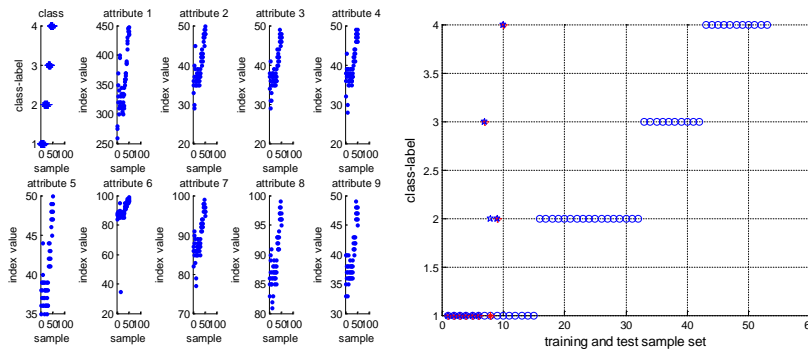


Fig.3. Visual training data of Model

○: training set classification; *: test set classification; ☆: practical classification

Fig.4. Training and prediction set classification

Acknowledgement

This research is financially supported by the Postdoctoral National Science Foundation of China Grant #201150M1560.

References

- [1] HUA H, NAN Z.Q, WANG S.Q. *Logistics support* [M], Beijing: Chinese Encyclopedia Publishing House, 2007.
- [2] U.S. Joint forces command joint warfighting center. JTLS (3.4.3.0) version description document[EB/OL]. [2010-06-08]. http://www.rolands.com/jtls/j_vdds/vdd_3200.pdf

- [3] George F S, Gregory A M. The joint warfare system (JWARS): a modeling and analysis tool for the defense department[C]//*Proc. of the winter simulation conference*. 2001:78-99.
- [4] Haykin S. *Neural Neetworks: a comprehensive foundation*[M]. Ye S W, Shi Z Z, trans. Beijing: China Machine Press, 2004.
- [5] Platt J C. Sequential minimal optimization: a fast algorithm for training support vector machines [M]// Scholkopf B, Burges C J C, Smola A J. *Advances in Kernel Methods-Support Vector learning*. Cambridge: MIT Press 1999.
- [6] FU Y G, SHEN R M. Learning effect evaluation system based on support vector machine [J]. *Computer Engineering*, 2004, 30(8):15-16, 74.
- [7] SUN X.D, HUANG C.L, ZHANG Q. *Science of military logistics* [M], Beijing: The Publishing House of PLA, 2002.
- [8] Chen Y S. Simple in complex and certainty in uncertainty [EB/OL]. <http://survivor99.com/entropy/chen21.htm>.
- [9] REN H Q. *Science of military command decision* [M].Beijing: National Defense University Publishing House, 2007.
- [10] SHI H P, HAN T. Evaluation in support ability of equipment maintenance personnel based on fuzzy comprehensive evaluation [J].*Modern Electronics Technology*, 2008, 246(1):96-98.