

A Comparative Study of Fuzzy Neural Network and Multivariate Statistics in Discriminant Analysis

Song Weicai

College of Computer Science and Technology, Jiangxi University of Traditional Chinese Medicine,
Nanchang, China

songweicai2000@sina.com

Keywords: Fuzzy neural network, Bayes discriminant, discriminant function, Multivariate statistics.

Abstract: Objective: Find out the effect of Fuzzy Neural Network (FNN) and Bayes method in discriminant analysis. Methods: Fuzzy Neural Network (ANN) and Bayes Method were used to discriminate and diagnose 250 cases of abdominal pain in nine diseases. Results: Fuzzy neural network (FNN) is superior to Bayes method in discriminating accuracy.

1. Introduction

Discriminant analysis, also known as "resolution method", is a multivariate statistical analysis method to determine the type attribution of a research object according to its various eigenvalues under the condition of classification determination. At present, the commonly used discriminant methods in research are mainly artificial neural network discriminant method and multivariate statistical discriminant method. Artificial Neural Network (ANN) is an algorithmic model for distributed parallel information processing by imitating the behavioral characteristics of ANN. It can process information by adjusting the interconnection between a large number of nodes in the network. Multivariate statistical discriminant method establishes one or more discriminant functions according to certain discriminant criteria, determines the undetermined coefficients in the discriminant function with a large amount of data of the object of study, and calculates the discriminant index, so as to determine what kind a sample belongs to. In medicine, we often encounter problems that need to be distinguished. For example, clinicians need to diagnose what disease the patient suffers from according to a series of symptoms, signs and examination results of the patient, and forensic doctors need to judge whether the deceased is suicide or homicide. In order to compare the discriminant effect of the two methods in medical diagnosis, this paper makes a discriminant diagnosis based on the clinical symptoms of some patients with abdominal pain provided by the provincial Chinese medicine hospital in order to achieve some research results.

2. Materials

A total of 250 cases of abdominal pain in Jiangxi Hospital of Traditional Chinese Medicine from January 2010 to December 2016 were collected, there were 12 cases of large intestine agglomeration, 31 cases of liver-depressed hypochondriac pain, 13 cases of damp-heat jaundice, 22 cases of damp-heat diarrhea, 14 cases of stomachache due to food stagnation, 15 cases of large intestine, 22 cases of diarrhea due to deficiency of cold, 79 cases of spleen-heart pain and 42 cases of Qi stagnation abdominal pain. The diagnosis results were confirmed by pathology. therefore the diagnosis results were confirmed by pathological examination.

3. Methods

3.1. Fuzzy Neural Network Discrimination

In order to adapt to various clinical symptoms of TCM diagnosis, this paper introduces fuzzy

logic and constructs a fuzzy neural network system (FNN). It can fuzz some non-quantitative information of patients[1-5].

Firstly, 11 test results were extracted as input vectors of the first level, including large intestine agglomeration, liver depression and hypochondria, damp-heat jaundice, damp-heat diarrhea, stomachache due to food stagnation, large intestine carbuncle, deficiency-cold diarrhea, spleen-heart pain, qi stagnation and abdominal pain.

Secondly, a fuzzy neural network model is established, which consists of three layers of forward BP neural network. The first layer is data preprocessing layer, which extracts 11 characteristic values of patients and fuzzifies them. The second layer is hidden layer, the third layer is output layer, and the output layer is 9 abdominal pain diseases of digestive tract system.

Finally, the training function of the network is trainlm, the learning function is learngdm, the error performance function is mse, the transfer function of each layer is logsig, and the training times are set to 1000. Through the training, validation and testing of the fuzzy neural network, the diagnostic results of 174 training groups, 38 validation groups and 38 test groups are obtained as follows.

In order to understand the effect of fuzzy neural network in TCM clinical diagnosis of abdominal pain, the diagnostic results of training group, validation group and test group are compared, shown in Table 1: (Notes: number of diagnostics(ND); actual number (AN); percentage of brackets for diagnostic accuracy).

Table 1 training, validation, test results.

Disease Type	Training Group		Verification Group		Test Group	
	ND	AN	ND	AN	ND	AN
large intestine agglomeration	9(100%)	9	1(100%)	1	2(100%)	2
liver-depressed hypochondriac pain	21(100%)	21	2(100%)	2	8(100%)	8
damp-heat jaundice	9(100%)	9	1(100%)	1	2(66.7%)	3
damp-heat diarrhea	13(100%)	13	4(100%)	4	5(100%)	5
stomachache	11(100%)	11	1(100%)	1	1(50%)	2
large intestine	13(100%)	13	1(100%)	1	1(100%)	1
diarrhea due to deficiency of cold	13(100%)	13	7(100%)	7	2(100%)	2
spleen-heart pain	53(100%)	53	15(100%)	15	13(84.6%)	11
Qi stagnation abdominal pain	32(100%)	32	6(100%)	6	4(100%)	4

It can be seen from Table 1 that fuzzy neural network (FNN) has good recognition ability. The accuracy rate of disease diagnosis in training group and validation group is 100%. Of all 250 sample cases, only 2 cases have been misdiagnosed, that is, one of the third and fifth diseases has been misdiagnosed as the eighth disease, the misdiagnosis rate is 0.8%, and the accuracy rate of discriminant diagnosis of fuzzy neural network is 99.2%.

3.2. Bayes Discrimination of Multivariate Statistics

Fisher discriminant and Bayes discriminant are commonly used in statistics. Fisher discriminant is for two classifications and Bayes discriminant is for multiple classifications [6-8]. Bayes discriminant is based on the prior probability of the whole, so as to minimize the average loss of misjudgement. Its greatest advantage is that it can be used for multiple sets of discrimination problems.

Bayes discriminant analysis was performed on 250 patients with abdominal pain using SPSS 19.0. (Notes: 12 cases of large intestine agglomeration(LIA), 31 cases of liver-depressed hypochondriac pain(LHP), 13 cases of damp-heat jaundice(DHJ), 22 cases of damp-heat diarrhea(DHD), 14 cases of stomachache due to food stagnation(SFS), 15 cases of large intestine(LI), 22 cases of diarrhea due to deficiency of cold(DDC), 79 cases of spleen-heart pain(SHP) and 42 cases of Qi stagnation abdominal pain(QSA)). Nine Bayes discriminant functions

can be obtained,By judging nine discriminant functions, Bayes discriminant results are obtained, such as Table 2.

Table 2 Bayes discriminant result.

Disease Type	Actual number	Discriminant diagnosis	Misdiagnosis
LIA	12	8 (66. 7%)	Four cases were misdiagnosed as the eighth disease
LHP	31	31 (100%)	0 cases
DHJ	13	11 (84. 6%)	Two cases were misdiagnosed as the eighth disease
DHD	22	21 (95. 4%)	One case was misdiagnosed as the sixth disease
SFS	14	10 (71. 4%)	Four cases were misdiagnosed as the third disease
LI	15	14 (93. 3%)	One case was misdiagnosed as the seventh disease
DDC	22	20 (90. 9%)	Two cases were misdiagnosed as the sixth disease
SHP	79	70 (88. 6%)	Nine cases were misdiagnosed as the first disease
QSA	42	42 (100%)	0 cases

As can be seen from Table 2, 23 of 250 patients with abdominal pain were misdiagnosed, and the accuracy rate was 91.8%. Bayes method of multivariate statistics also had good discriminant ability, but it was inferior to the discriminant ability of fuzzy neural network.

4. Comparative analysis

From Table 1 and Table 2, it can be seen that the first, third, fifth and eighth diseases are most easily misdiagnosed, followed by the sixth and seventh diseases. Why are the first, third, fifth, eighth and sixth and seventh diseases prone to misdiagnosis, while the second and ninth diseases prone to misdiagnosis? The reason is that the clinical symptoms of these diseases have many similarities, which can be explained by the discriminant function diagram of the nine diseases. As shown in Figure 1:

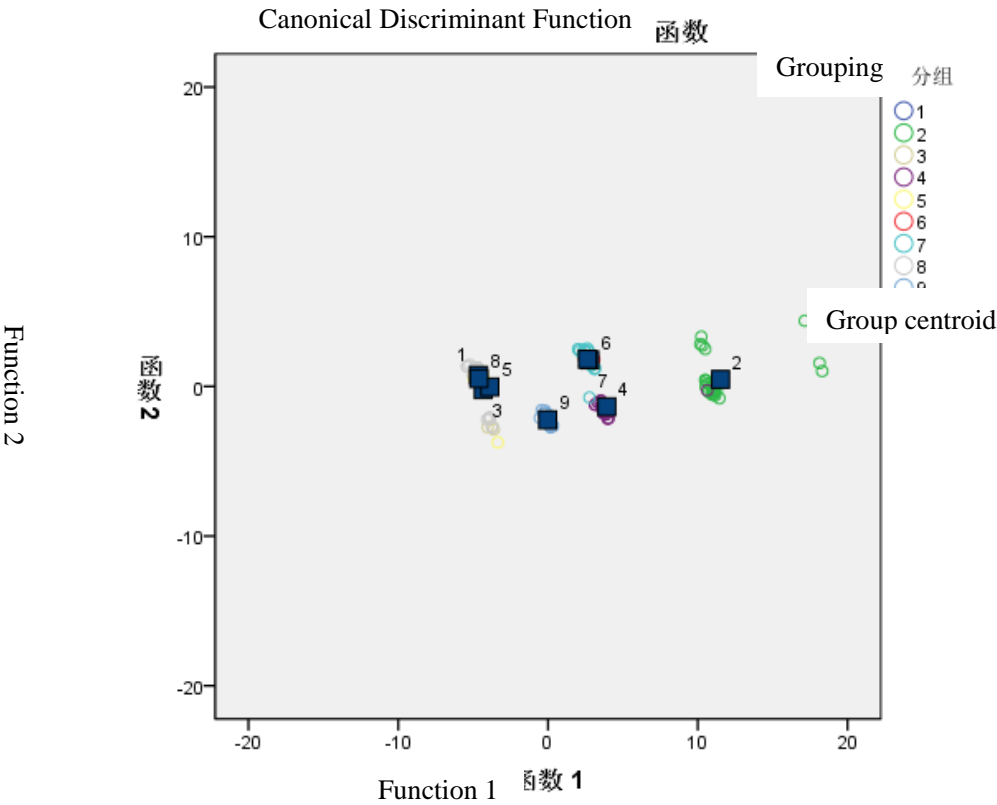


Figure 1 Nine Disease Discriminant Function Diagrams.

As can be seen from Figure 1, the centroid of the first, third, fifth and eighth diseases is very close, the centroid of the sixth and seventh diseases is very close, and the centroid of the second and ninth diseases is far away from the centroid of other groups. Therefore, the first, third, fifth and eighth diseases and the sixth and seventh diseases are easily misdiagnosed in clinic. In addition, from Table 1 and Table 2, it can be concluded that the discriminant accuracy of Fuzzy Neural Network is obviously higher than Bayes discriminant of multivariate statistics, mainly because Fuzzy Neural Network has strong ability of non-linear mapping, self-learning and self-adapting, and fault-tolerance. The use of Bayes discriminant of multivariate statistics must satisfy three assumptions, that is, all variables must obey multivariate positive. The distribution of states, the covariance matrices of each group must be equal, and the mean values of each group of variables have significant differences. The covariance matrices of each variable selected here are somewhat different, and the mean values of some variables are not significant.

5. Conclusion

Through discriminant and comparative analysis of various abdominal pain data in TCM clinic, this paper draws a conclusion that when faced with large sample data and many index variables, the Fuzzy Neural Network (FNN) has obvious pattern discriminant advantages. It can simulate linear and non-linear data, and has strong fault tolerance. However, the discriminant method of Fuzzy Neural Network has its shortcomings. It can not conclude input and output data. Structural analysis. In the discriminant analysis of multivariate statistics, the quality of variable selection will affect the discriminant effect, leaving some variables which have little influence on the discriminant function will interfere with the discriminant results. In addition, multivariate statistical discriminant analysis requires that there should be no linear relationship between variables, otherwise it will seriously affect the discriminant results of multivariate statistics.

Acknowledgements

The authors gratefully thank the financial support from Research on Research Topics of Traditional Chinese Medicine of Jiangxi Health and Family Planning Commission (2017B002).

Reference

- [1] Huang Yongfeng, Cen Kang, etc. Application of fuzzy neural network in segmentation of brain magnetic resonance image [J]. Proceeding of the National Institute of Biomedical Engineering, 2003, 22 (6): 508-512.
- [2] Xu Liping, Shang Dan. Application of Fuzzy Neural Network in CT Diagnosis of Lung Cancer [J]. Journal of Zhengzhou University (Medical Sciences), 2014, 49 (2): 191-194.
- [3] Liu Li, Huo Liqin. Application of Fuzzy BP Neural Network in Early Diagnosis of Neonatal HIE [J]. Biomedical Engineering, 2011, 28 (4): 814-819.
- [4] Song Weicai, Wu Yanxia. Application of Fuzzy Neural Network in Clinical Diagnosis of Abdominal Pain [J]. Science and Technology Square, 2014, (10): 47-52.
- [5] Song Weicai, Wu Yanxia. Study on the Effect of Fuzzy Neural Network in Clinical Diagnosis of Traditional Chinese Medicine [J]. Medical information, 2015, 28(9): 16-17.
- [6] Wen Changping, Bai Yinyong, etc. Bayes discriminant analysis method for natural grassland classification [J]. Journal of Grassland Science of China, 2016, 38(3): 76-79.
- [7] Li Bin, Shi Haibin, etc. Analysis method of groundwater chemical classification based on Bayes discriminant theory [J]. Agricultural Research in Dry Areas, 2015, (4): 246-250.
- [8] Shiyuan, Liu Ruijie. Analysis of EEG Data Based on Bayes Discrimination [J]. value engineering, 2015, (13): 169-170.