

# Data Mining Applied to the Predictive Model of Triage System in Emergency Department

Wen-Tsann Lin, Yung-Tsan Jou, Yih-Chuan Wu, and Yuan-Du Hsiao

**Abstract**—The Emergency Department of a medical center in Taiwan cooperated to conduct the research. A predictive model of triage system is constructed from the contract procedure, selection of parameters to sample screening. 2,000 pieces of data needed for the patients is chosen randomly by the computer. After three categorizations of data mining (Multi-group Discriminant Analysis, Multinomial Logistic Regression, Back-propagation Neural Networks), it is found that Back-propagation Neural Networks can best distinguish the patients' extent of emergency, and the accuracy rate can reach to as high as 95.1%. The Back-propagation Neural Networks that has the highest accuracy rate is simulated into the triage acuity expert system in this research. Data mining applied to the predictive model of the triage acuity expert system can be updated regularly for both the improvement of the system and for education training, and will not be affected by subjective factors.

**Keywords**—Back-propagation Neural Networks, Data Mining, Emergency Department, Triage System.

## I. INTRODUCTION

EMERGENCY Department is the frontline for hospital to face patients in emergencies. The members in Emergency Department include the doctors, nursing staff, technicians, social workers, emergency medical technicians, administrative staff, fellow workers, and volunteers. This department runs 24 hours a day, and no matter first aids, observation and surgical operations can all be conducted here. It is like a small hospital in the hospital. Xin-Kai Zhou [1] points out in his research that Emergency Department is the hospital's first line of the medical care service, and is also the place where people go to for timely medical service when they face emergencies. Hence, the medical care service quality and the appropriate allocation of resources are very important to both the public and the hospital.

The so-called Triage System is based on the Appraisal Standards of Emergency Departments from Department of Health, DOH, Executive Yuan, Taiwan, (2009) [2], and there are 4 levels of the triage system for emergency patients: Level 1 (should be treated immediately); Level 2 (should be treated within 20 minutes); Level 3 (should be treated within an hour),

Wen-Tsann Lin is with the Dept. of Industrial Engineering and Management, National Chin-Yi University of Technology, Taichung City 41170, Taiwan.

Yung-Tsan Jou is with the Dept. of Industrial and System Engineering, Chung Yuan Christian University, Chung-Li, Taoyuan 32023, Taiwan.

Yih-Chuan Wu is with the Dept. of Industrial and System Engineering, Chung Yuan Christian University, Chung-Li, Taoyuan 32023, Taiwan (e-mail: jason\_wu1102@yahoo.com.tw, phone:886-921-399529).

Yuan-Du Hsiao is with Department of Business Administration, Chungyu Institute of Technology, Keelung 20103, Taiwan.

and Level 4 (treatment can be held). The meaning of the "consistency" referred in this research means: "After the triage done by the nursing staff and the diagnosis and treatment from the doctors, the percentage of those in the same level is generally presented in accuracy rate".

Currently, the decision making of triage in domestic hospitals is mostly made by senior triage nurses (referred as nursing staff below). A research result shows that the accuracy rate of the consistency between the triage system and the doctors' diagnose and treatment reached to 87% (Wollaston et al.) [3]. The research conducted on the nursing staff and the pediatricians by Russo et al. [4] also indicated that the accuracy rate of the triage consistency reaches to 84%. It is obvious that the consistency of decision making in triage still leaves some room to improve, and on the other hand, the statistics also means that every day, nearly 2,423 patients are over triaged or under triaged because of the triage system and this causes waste of medical care resources and infringement of the patients' rights (J.Y. Zhan) [5]. Moreover, the process of decision making is easily affected by the complexity of work, conflicts, nursing experiences, education and professional knowledge (Brillman, Doezeema, & Tandberg) [6]. Hence, the more emergency patients are, the more complexity the diseases becomes. When facing the mission of triage, how to give consideration to the consistency and robustness of the decision making is the task to concern in this research. Data mining is used for the application to the predictive model of triage system in this research, in attempt to enhance the triage consistency and robustness through an expert system.

## II. LITERATURE REVIEW

The United States started to introduce Triage system back in 1960, in an attempt to reduce the costly economic cost caused by non-urgent patients in emergency departments (Weinerman and Edwards). In 1964, New York Hospital in the States started to employ nursing staff to execute Triage, and senior nursing staff were considered to be competent for the job, and the nursing staff started to be responsible for this ever since (Estrada) [7].

With limited resources, the establishment of Triage system is very helpful to the patients in real urgent situations. However, if the nursing staff and the doctors have different cognitions on patients' urgent extent, patients then could receive different medical care. VanBoxel [8] also mentions in his research that the whether the screening of triage is consistent with the doctors' diagnoses and treatment will significantly influence

patients' satisfaction and waiting time for medical care services.

Scholars continually have discussions on the decision making concerning triage since 1970. Currently, the researches on triage system mainly focus on the followings: (1) Questionnaire surveys on factors that influence variations of triage; (2) Empirical Researches on the consistency of triage; (3) Discussions on information technology application for variations of triage and the improvement of the system.

In recent years, because of the demands of information application, the Database and Data Warehousing have evolved into Data Mining nowadays. Data Mining means "the process of digging out non-obvious, un-known and potentially useful information". Reinschmidt, Gottschalk, Kim, & Zwietering [9] believe that "Data Mining is to extract effective, useful, unknown but understandable information from the database as the basis for enterprises to do decision making". It is thus clear that making use of the information properly can provide abundant information for management and decision making to enterprises.

Through many scholars' researches, currently, the technology of data mining is successfully applied in industrial, business, medical care fields and so on. In medical care, Huang and Chen [10] applied many aided machines for data mining to learn how to judge glaucoma, and they combined Linear Discriminant Analysis, LDA, with the Principle Component Analysis, PCA, and the accuracy rate reaches to 91.5%; while the accuracy rate obtained from the combination of Artificial Neural Networks, ANN, and the PCA, is 99.1%. Abdelfattah et al. [11] also combined Multi-group Discriminant Analysis, MDA, with the serum examination and radiotherapy and produced a set of linear functions (to produce new variables), and through ROC (Receiving Operation Characteristic) curve to analyze the explanation capability of these new variables, and to predict whether Hepatitis C will turn into liver cirrhosis with the accuracy rate of 93%. We can see the application of data mining technology has been practiced for years.

Turban, Aronson, & Liang [12] believe that "the expert system is to express experts' knowledge and induction process through computer, and through a virtual consultant, people can solve related problems of a specific type". Related literature is collected first in this research to have a general understanding of the current situation of triage system in emergency departments, and on the other hand, to obtain patients' basic statistics and data from the emergency database of the hospital in this case. Then the research is conducted by the support of the administrative supervisors on professional knowledge in hospital management and nursing in an attempt to properly utilize resources in database for information digging, and to combine theories and practices of data mining technology to effectively improve the consistency of triage in emergency medicine.

### III. METHODS

The "Predictive Model of triage system" and "Predictive Model of Data Mining" are categorized and organized by means of the visit process and the current situation of the triage

in this medical center. On one hand, the parameters with diagnostic meanings clinically are found through this predictive model of triage system. On the other hand, the concrete steps need to be taken can be defined through predicative model of data mining. After the confirmation and approval from the emergency department, the samples needed retrieved from the patients database of this emergency department are sorted and organized for necessary processing of information.

#### *A. Predictive Model of Data Mining*

There are many tools to choose from in data mining. Discriminant Analysis, Logistic Regression, and Artificial Neural Network are adopted in this research. Multi-group Discriminant Analysis and the Discriminant Analysis share the same purpose. Because the Discriminant Analysis is often used to compare with other categorization technologies, Multi-group Discriminant Analysis is also adopted in this research as the comparison basis (Sharda & Delen) [13].

Logistic Regression can divide the research results into two categories through a game theory (Such as: Success-Failure). Since categorical dependent variables are included in the parameters of this research, Multi-group Discriminant Analysis is chosen as the tool for triage.

Since this is a multi-group triage in this research, the structure of date can be more complex, Back-propagation Neural Networks that can use non-linear transfer function is adopted as the triage tool, in an attempt to obtain predictive information with better accuracy rate for triage (Heidar, Nicolaos, & Mahesh) [14]. The predictive model for date mining is shown in Fig. 1. To enhance the credibility of the research, increase the accuracy rate of triage, the data is divided into the training group and testing group before analysis and 10-fold cross-validation is conducted.

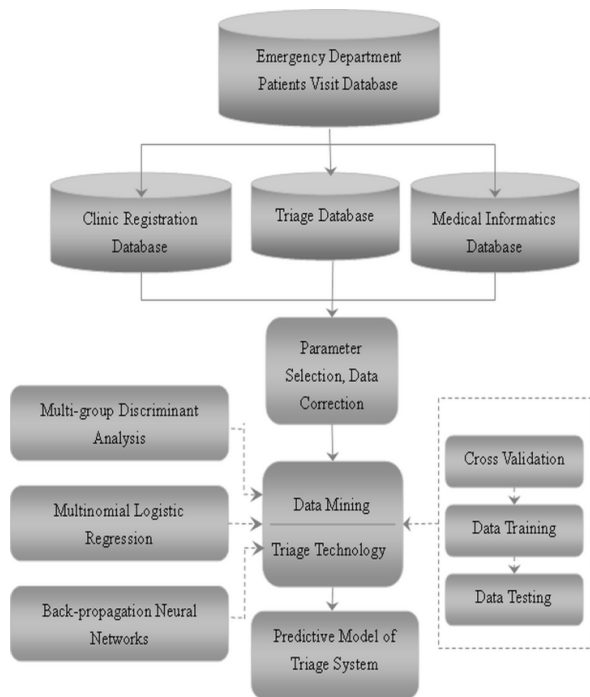


Fig. 1 Predictive Model of Data Mining

**B. Parameter Measurement**

As for the parameters, the norms in are based on the literature of Handyside [15] and the Triage Screening Scale for Adults from Department of Health, Executive Yuan, Taiwan, (2009). The measurement is done by the patients’ complaint, medical history, general appearances, vital signs, symptoms and signs, and the results of Physical Assessment.

The 6 parameters of Vital Sign are all Ratio Scales. Temperature  $\geq 41^{\circ}\text{C}$  or  $\leq 32^{\circ}\text{C}$  belongs to Level 1; temperature between  $39^{\circ}\text{C}$  to  $40^{\circ}\text{C}$  or  $32^{\circ}\text{C}$  to  $35^{\circ}\text{C}$  belongs to Level 2; temperature  $< 39^{\circ}\text{C}$  belongs to Level 3; while others belong to Level 4. Systolic pressure  $> 220\text{mmHg}$  or  $\leq 80\text{mmHg}$  belongs to Level 1; systolic pressure between  $180\text{mmHg}$  to  $220\text{mmHg}$  belongs to Level 2. There is no specific definition of Triage Level 3 and Level 4 when it comes to the systolic pressure.

Since the chief complaint, medical history, general appearance, symptoms and signs and the results of physical assessment cannot be quantified; besides, complication usually occurs on patients, the Triage Screening Scale for Adults from DOH, Executive Yuan, Taiwan, (2009) is used for assessment in this research. 4 points are given when the combined complaint is assessed as Level 1 in this scale; 3 points are given to Level 2; 2 points are given to Level 3; and 1 point is given to Level 4. If there is no point from the norm in the Triage Screening Scale for Adults, then 0 point is given. Finally, the combination of complaints is regarded as the Ordinal Scale to conduct assessment.

**C. Data Retrieve and Organization**

108,480 pieces of patient data of the emergency department was retrieved (including 245 pieces of missing data). When processing the data, the “Medical Record Number” is the primary key while the “Visit Time” is the secondary key, and the 6 parameters in “Parameter Measurements” are the basis for screening. After organization, the distribution of patient triage is shown as Table I.

The number of Triage Level 1 patients accounts for 8.01%; that of Level 2 accounts for 41.30%; that of Level 3 accounts for 50.16%; that of Level 4 accounts for 0.53% of the total patient numbers. Since the triage results will be influenced by triage technology, selection of parameters and data retrieve, to increase the accuracy rate of triage, 2,000 pieces of data are divided into training (1,800 pieces) and testing (200 pieces), and 10-fold cross-validation.

TABLE I  
GENERAL DESCRIPTION OF PATIENT TRIAGE

	Level 1	Level 2	Level 3	Level 4	Total
Triage Variation	15	354	121	11	501
Triage Consistency	8,622	44,325	54,268	519	107,734
Total triage person-time	8,688	44,801	54,411	580	108,480
%	8.01%	41.30%	50.16%	0.53%	100%

**IV. RESULTS**

Three methods of data mining (Multi-group Discriminant Analysis, Multinomial Logistic Regression, Back-propagation Neural Networks) are actually utilized in the predictive model of triage. Through the empirical research of the three methods, the target value (doctor’s diagnosis) and the predictive value (system simulation) can be matched one by one by means of simulation. Not only the applicability of data mining to triage but also the performance of the three methods in data mining can be understood this way. Detailed information on the three methods of data mining and the ranking on their performance are described as below.

**A. Multi-Group Discriminant Analysis (MDA)**

In this research, the operation process of Multi-group Discriminant Analysis is divided into 4 stages: (1) Examination of the prominence the hypothesis; (2) Identification of variables that have differentiation; (3) Formation of new variables (indexes); (4) Decision of the attribution of new observation point. In principle, just follow the above stages; the multi-group discriminant analysis can be done. But the data is divided into training and testing groups in this research, and 10-fold cross-validation is adopted, so multi-group discriminant analysis has to be done in the same steps 10 times (the training and testing of the data is included each time). The software used for Multi-group Discriminant Analysis is SPSS.

200 samples in the testing group are regarded 200 new coming patients, and one patient data among the testing group is shown as an example below (as in Table II). Table II shows the patient data (standardized) in the testing group, and it was

applied into 3 discrimination functions, and the discriminant scores are give below.

$$X = Z_1 = 0.118X_1 + 0.136X_3 + 0.022X_4 - 0.095X_5 + 2.565X_6 + 0.004 = 3.524$$

$$Y = Z_2 = 0.426X_1 - 0.361X_3 + 0.587X_4 - 0.710X_5 + 0.167X_6 + 0.007 = 0.432$$

$$Z = Z_3 = 0.378X_1 + 0.649X_3 - 0.700X_4 - 0.509X_5 - 0.024X_6 + 0.008 = 0.504$$

Through the calculation of discrimination functions, the parameter of the patient data can be transformed into coordinates in space  $(X, Y, Z) = (3.524, 0.432, 0.504)$ . By calculating the patient coordinates and the Euclidean distance, it is learnt from the research that the patient is closer to the core of Triage Level 1 ( $0.39 < 2.92 < 10.95 < 23.56$ ). Therefore, the patient was predicted as Level 1 patient, and the result matches the target value. In the same manner, other samples in the testing group show an accuracy rate of 90.50%. The operation of the Multi-group Discriminant Analysis is complete. To increase the accuracy rate, 9 times of cross validation is conducted based on the steps mentioned above. After a 10-fold cross validation, it is confirmed that the 2<sup>nd</sup> test set has better accuracy rate of predication, when the training group has an accuracy rate of 91.30% and the testing group has an accuracy rate of 88.50%, and the total accuracy rate is 91.02%.

### B. Multinomial Logistic Regression

The operation of Multinomial Logistic Regression is similar to that of Multi-group Discrimination Analysis, through the stepwise process; variables with better discrimination capability can be chosen to produce a concise, easily-explained Multinomial Logistic Regression with good compatibility. The software for Multinomial Logistic Regression in this research is SAS 9.1. After adding variables with explanation capability into Multinomial Logistic Regression by Stepwise process, 3 Regression Models are produced (i.e.  $\ln(p_{Y_i} / 1 - p_{Y_i})$ ,  $i = 1$  to 3). For the convenience of follow-up observation and categorization, the regression models are named as  $Y_1$ ,  $Y_2$ , and  $Y_3$ .

$$Y_1 = 1.005X_1 + 0.493X_2 + 1.302X_3 - 1.004X_5 + 25.234X_{6(0)} + 45.451X_{6(1)} + 41.516X_{6(2)} + 2.680X_{6(3)} + 0X_{6(4)} - 21.825 \quad (1)$$

$$Y_2 = 0.807X_1 + 0.508X_2 + 0.839X_3 - 0.136X_5 - 13.685X_{6(0)} + 24.912X_{6(1)} + 26.900X_{6(2)} + 5.815X_{6(3)} + 0X_{6(4)} - 4.846 \quad (2)$$

$$Y_3 = 0.545X_1 + 0.076X_2 + 0.016X_3 - 0.108X_5 - 16.110X_{6(0)} + 3.672X_{6(1)} + 20.358X_{6(2)} + 6.122X_{6(3)} + 0X_{6(4)} - 2.821 \quad (3)$$

$X_1$ : Systolic Pressure  $X_2$ : Diastolic Pressure  $X_3$ : Pulse  $X_5$ : SaO<sub>2</sub>

$X_{6(0)}$ : Combined Complaint = 0

$X_{6(1)}$ : Combined Complaint = 1

$X_{6(2)}$ : Combined Complaint = 2

$X_{6(3)}$ : Combined Complaint = 3

$X_{6(4)}$ : Combined Complaint = 4

when Multinomial Logistic Regression forms a regressions model, little transformation is needed to calculate the “odds-on of the observation point of each triage level”. The formulas for regression model to transform into odds-on model are showed below:

$$\text{Odds-on of Triage Level 1: } p_{\text{Triage1}} = \frac{e^{Y_1}}{1 + e^{Y_1} + e^{Y_2} + e^{Y_3}} \quad (4)$$

$$\text{Odds-on of Triage Level 2: } p_{\text{Triage2}} = \frac{e^{Y_2}}{1 + e^{Y_1} + e^{Y_2} + e^{Y_3}} \quad (5)$$

$$\text{Odds-on of Triage Level 3: } p_{\text{Triage3}} = \frac{e^{Y_3}}{1 + e^{Y_1} + e^{Y_2} + e^{Y_3}} \quad (6)$$

$$\text{Odds-on of Triage Level 4: } p_{\text{Triage4}} = \frac{1}{1 + e^{Y_1} + e^{Y_2} + e^{Y_3}} \quad (7)$$

Through the odds-on categorization, the training group’s accuracy rate of categorization (1,800 samples) can be obtained. After organization, 405 Triage Level 1 samples were successfully predicted as Level 1; 389 Triage Level 2 samples were successfully predicted as Level 2; 421 Triage Level 3 samples were successfully predicted as Level 3; 431 Triage Level 4 samples were successfully predicted as Level 4. 1,646 pieces of data were accurately triaged in total, with an accuracy rate of 91.40% in the training group.

Like Multi-group Discriminant Analysis, 200 samples in the testing group are regarded 200 new coming patients, and one patient data among the testing group is shown as an example in Table II. After imputing the patient data into 4 odds-on models, the odds-on of each level obtained is listed as below.

$$\text{Odds-on of Triage Level 1: } p_{\text{Triage1}} = 41\%$$

$$\text{Odds-on of Triage Level 2: } p_{\text{Triage2}} = 35\%$$

$$\text{Odds-on of Triage Level 3: } p_{\text{Triage3}} = 12\%$$

$$\text{Odds-on of Triage Level 4: } p_{\text{Triage4}} = 12\%$$

It can be learnt that there is higher odds-on to categorize the patient as Triage Level 1 ( $41\% > 35\% > 12\% = 12\%$ ). Therefore, the patient is triaged as Triage Level 1, and the result matches the target value.

Other samples are tested in the same method, and 71% of accuracy rate in the testing group can be obtained. After a 10-fold cross validation, it is confirmed that the 3<sup>rd</sup> test set has better accuracy rate of predication, when the training group has an accuracy rate of 91.50% and the testing group has an accuracy rate of 72%, and the total accuracy rate is 89.55%.

TABLE II  
PATIENT DATA IN TESTING GROUP

Medical Record Number	Systolic Pressure	Diastolic Pressure	Pulse	Temp.	SaO <sub>2</sub>	Combined Complaint	Target Value
44XX65	0.175	0.635	-0.592	-0.818	0.542	1.423	1

C. Back-Propagation Neural Networks

The basic structure of a Neural Network includes the input layer, the hidden layer and the output layer. Generally, the processing unit of the input layer and the output layer depend on the research topic. However, if the number of layers in the hidden layer or if the numbers of the neuron are not properly set, then the network could be over-complex or bad convergence may occur. The structure of the Neural Network is shown in Fig. 2.

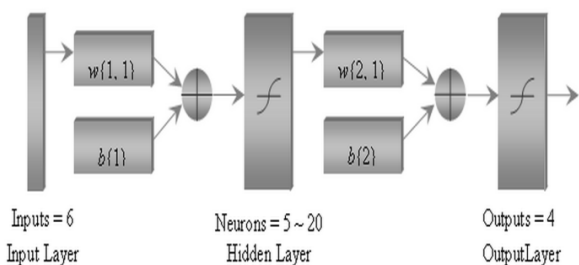


Fig. 2 Structure of Neural Network

The number of neurons in the hidden layer has great impact on the network quality. The more number of neurons there is in the hidden layer, the slower the convergence is, and this is of little help to reduce the error. Comparatively, if there are too few neurons in the hidden layer, the mutual relation between the input and the output cannot be reflected.

The number of neurons in the hidden layer can be based on that of the input layer (6) and the output layer (4) (Y.-C. Ye) [16]:

Simple Questions: Number of neurons in the hidden layer =  $(\text{input neuron number} + \text{output neuron number}) / 2 = 5$

General Questions: Number of neurons in the hidden layer =  $(\text{input neuron number} + \text{output neuron number}) = 10$

Difficult Questions: Number of neurons in the hidden layer =  $(\text{input neuron number} + \text{output neuron number}) \times 2 = 20$

The triage accuracy rate is not greatly influenced in every training process. Therefore, it is confirmed that in the researches on triage, the best neuron number of the hidden layer in the Back-propagation Neural Networks is 14. Secondary, to test the best combination of transfer function, 4 numbers of

neurons with higher triage accuracy rate (12, 13, 14, 19) in this research are retrieved for follow-up analysis.

Since the target value of this research is between 0 and 1, it is expected that the output value of the network does not go beyond the enumerated domain of  $\pm 1$ , and if Purelin is used in the output layer, the output of the network is very possible to go beyond the enumerated domain of  $\pm 1$ . Hence, Tansig and Logsig are used to test the best combination.

Table III shows the triage accuracy rate of different combinations of transfer function at the neuron number of 12, 13, 14, and 19. On the top of the table is the Accuracy Rate of Training Group; in the middle of the table is the Accuracy Rate of Testing Group; at the bottom of the table is the Total Accuracy Rate. It is indicated from the Accuracy Rate of Training Group that the training group has the highest accuracy rate (96.00%) when Transfer Function Logsig + Logsig and the neuron number is 19, then is that when Transfer Function Tansig + Logsig and the neuron number is 14 and 13, and the accuracy rate of the training group is 95.61% and 95.44% respectively.

TABLE III  
ACCURACY RATE OF NEURONS AND TRANSFER FUNCTION

500 Step iterations		Neuron Number			
Transfer Function		12	13	14	19
Accuracy rate of training group (1800)	TANSIG+TANSIG	93.15%	93.18%	93.58%	93.44%
	TANSIG+LOGSIG	94.56%	95.44%	95.61%	94.83%
	LOGSIG+LOGSIG	94.39%	94.00%	93.89%	96.00%
	LOGSIG+TANSIG	92.89%	92.89%	92.83%	92.50%
Accuracy rate of testing group (200)	TANSIG+TANSIG	89.00%	88.17%	88.17%	88.67%
	TANSIG+LOGSIG	86.00%	88.00%	87.50%	83.50%
	LOGSIG+LOGSIG	88.00%	87.50%	87.50%	81.50%
	LOGSIG+TANSIG	89.00%	87.00%	89.50%	88.50%
Total Accuracy Rate (2000)	TANSIG+TANSIG	92.74%	92.68%	93.04%	92.97%
	TANSIG+LOGSIG	93.70%	94.70%	94.80%	93.70%
	LOGSIG+LOGSIG	93.75%	93.35%	93.25%	94.55%
	LOGSIG+TANSIG	92.50%	92.30%	92.50%	92.10%

Although the training group has the highest accuracy rate when transfer function Logsig + Logsig and the neuron number is 19, it can be learnt from a closer observation of the testing group performance that over learning seems to occur in this network (Accuracy rate of testing group: 81.50%).

Meanwhile, the network performance when transfer function Tansig + Logsig and when the neuron number is 14 and 13, the triage accuracy rates of the training group and the testing group are pretty good. As a result, the Back-propagation Neural Networks has good triage accuracy rate and robustness when

the transfer function of the hidden layer is set at Tansig and the neuron number is 14, and the transfer function of the output layer is set at Logsig.

To increase the accuracy rate of this Back-propagation Neural Networks, the transfer function of the hidden layer is set as Tansig, the number of neurons is set at 14, and the transfer function of the output layer is set as Logsig, and 9 times of cross validation is conducted based on the structure of Neural Networks. After a 10-fold cross validation, it is confirmed that the 4th test set has better accuracy rate of predication, when the training group has an accuracy rate of 95.56% and the testing group has an accuracy rate of 91.00%, and the total accuracy rate is 95.10%.

Histogram is used for the performance appraisal for data mining, and the performance appraisal of training group, testing group and the total accuracy rate of Multi-group Discriminant Analysis, Multinomial Logistic Regression, and Back-propagation Neural Networks are conducted. The accuracy rate of the training group shows the appropriateness of algorithm to triage screening, and it is also the most direct way for assessment. The gap between the accuracy rate of the testing group and that of the training group shows the robustness algorithm is to the application of triage screening. The larger the gap is, the lower the robustness is while the smaller the gap is, the higher robustness is. The total accuracy rate is used to consider the robustness of the testing group. The method to make downward revision of the accuracy rate of the training group is also the main basis for the research to choose the algorithm. The Histogram of Data Mining Accuracy Rate is shown as Fig. 3.

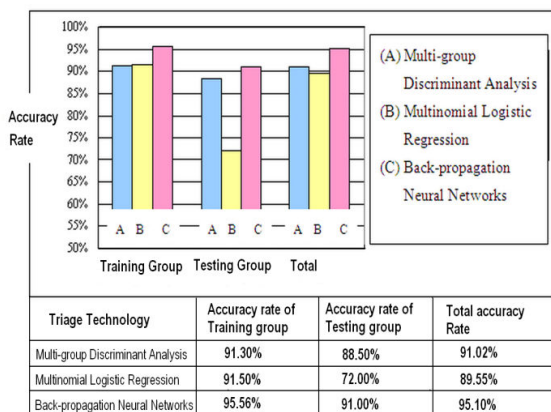


Fig. 3 Histogram of Data Mining Accuracy Rate

#### D. The Expert System

The purpose of this research is to increase the triage consistency in Emergency Medicine with the application of data mining technology. So far, the triage accuracy rate is indeed successfully increased by utilizing Neural Network Technology. However, it is very important to concretely present this theoretical performance and further, let the nursing staff in the front line actually experience it. Therefore, it is aimed to develop the rule of the information flow and decision

control in the Expert System by using IDEF0 ( Integrated Computer-aided Manufacturing, ICAM Definition ), and it is hoped that through the process of IDEF0 and the Syntax Definition between output, input, limitation and resources, the User Interface of the Expert System can be clearly depicted.

Fig. 4 shows the Inquiry Screen of the Interface of the Patient's Basic Information in the Expert System. When the patient is sent to the Emergency Room, they will be sent to the Screening station for Triage Screening. Nursing staff can inquire patients' basic information by inputting the patients' names or ID numbers. When the Patient's Name or the ID Number is input, the Patient's Basic Information can be sent to all the interfaces of the Expert System through the Patients Visit Database, including the Gender, Weight, Medical Record, Transfusion Reaction and so on. Wrong information or date can also be deleted, saved, or exited in this interface.

Fig. 4 Interface of the Expert System (Patient's Basic Information)

Fig. 5 Interface of the Expert System (Triage Screening Information)

Fig. 5 shows the Prediction Screen of the Triage Screening Information of this Expert System. When the nursing staff input the Patient's Name or ID Number in the Inquiry Screen, information about the patient's the Medical Record, Allergy to Drugs and Transfusion Reaction can be shown on the

Prediction Screen through the Patients Visit Database. The interface needed to be input on the Triage Screening Information includes the Triage Level decided by the nursing staff, Department for Emergency Visit and Department for Consultation, and the Vital Signs and Combined Chief Complaint for the prediction. Norms defined in the Triage Screening Scale for Adults are listed in the Dropdown Menu for the input of Combined Chief Complaint; undefined ones are described with words. After inputting all the information, the nursing staff can click the Prediction button to start the function of triage prediction.

Because this triage expert system is developed from the data mining technology, this simulated triage expert system can use the nursing staff's decision making experiences and formed artificial intelligence in the Back-propagation Neural Networks so that the expert system can act like a senior nursing staff and make predictions based on the past experiences of decision making, and the system can assist the nursing staff in executing triage screening. Besides, since the operation of the expert system is done through computer programs, the decision making will not be influenced by the complexity of work, conflicts of work, nursing experiences, education and professional knowledge. Therefore, when the nursing staff's triage is different from that predicted by the system, the nursing staff can reassess the patient's triage level carefully.

In addition, as far as the operation of the expert system is concerned, the triage screening screen is referred to when programming, so the expert system interface is very much similar to the current operation screen, and the nursing staff should be used to the operation very soon. On the other hand, the method for information input is also similar to that used currently, but it used to be needed to key in the patient's chief complaint, but now input can be done by clicking, and the operation time can be shortened. The only difference is that one Prediction Button is added on the "Prediction Screen" in the expert system, and it takes less than one second to predict after inputting the patient's information, and result of the suggested triage level can be obtained.

## V. CONCLUSION

Three methods of data mining (Multi-group Discriminant Analysis, Multinomial Logistic Regression, Back-propagation Neural Networks) are actually utilized to increase the consistency of emergency triage screening. Because Back-propagation Neural Networks can solve the problem of the non-linear triage, it can have better performance in triage prediction (95.10%), then is the Multi-group Discriminant (91.02%), and lastly, the Multinomial Logistic Regression (89.55%).

Although the Back-propagation Neural Networks has better triage effect, the nursing staff cannot be told the performance of each parameter (explanation capability). Concerning this, the answers can be obtained from the Discriminant Function of the Multi-group Discriminant Analysis and the Stepwise process of the Multinomial Logistic Regression. As far as the Multi-group Discriminant Analysis is concerned, the parameters are listed from the one with the best performance to the least: Combined

Chief Complaint, SaO<sub>2</sub>, Pulse, Systolic Pressure, Temperature, and Diastolic Pressure. For Multinomial Logistic Regression, the order of the parameter performance goes: Combined Chief Complaint, Pulse, SaO<sub>2</sub>, Diastolic Pressure, Systolic Pressure, and Temperature. The Combined Chief Complaint has better performance in the Multi-group Discriminant Analysis and in the Multinomial Logistic Regression, which conforms to the nursing staff's implementation of triage screening.

Compared with the triage consistence implemented by the nursing staff, the accuracy rate of triage screening is usually increased by improving triage system of by implementing educational system when the nursing staff implements the triage task. The Triage Screen System has been practiced for years; besides, the Triage System in Taiwan is governed by DOH. So there is little room but more difficulties to improve the system (Triage Accuracy Rate: 87.60%). For educational training, currently, the most effective way for improvement is to lower the gap between the nursing staff's triage decisions by implementing educational trainings. However, the nursing staff is more or less influenced by subjective factors (the complexity of work, conflicts, nursing experiences, education and professional knowledge), so there are so many uncertainty in the triage decision making process (Triage Accuracy Rate: 89.6%).

The Triage Expert System simulated the triage screening through the Back-propagation Neural Networks Technology in data mining. To make it easier for system analysts to program the triage expert system so that the developed system can meet practical needs more. The system analysts can clearly understand the language definitions of input, output, resources, limitation in the triage expert system, and this can make it easier for the nursing staff to use the system.

Data mining technology is applied to the simulation of triage expert system, and useful information and rules are extracted from the nursing staff's past decision making experiences through the scientific method of data mining. The Triage Expert System developed in this research is just like a senior triage nursing staff. What's different from real nursing staff is that the Expert System can give consideration to the system improvement and educational training through regular update, and is not influenced by emotions.

## ACKNOWLEDGMENT

The authors would like to express their gratitude for the budget support from National Science Council (NSC 100-2410-H-167-010-MY2) which made this research possible.

## REFERENCES

- [1] X. K. Zhou, "Discussion on the features of high emergency resources consumers by data mining technology (unpublished thesis)", National Taiwan University, Taiwan, 2004.
- [2] Department of Health, DOH, Appraisal Standard of Emergency Departments. Website of DOH, Executive Yuan, Taiwan, 2009.
- [3] A. Wollaston, P. Fahey, M. McKay, D. Hegney, P. Miller, & J. Wollaston, "Reliability and validity of the toowoomba adult trauma triage tool: a Queensland, Australia study," *Accident and Emergency Nursing*, vol.12(4), 2004, pp.230-237.

- [4] R. M. Russo, V. J. Gururaj, A. S. Bunye, Y. H. Kim, & S. Ner, "Triage abilities of nurse practitioner vs. pediatrician," *American Journal of Disease of Children*, 129(6), 1975, pp.673-675.
- [5] J. Y. Zhan, "Discussion on the accuracy rate of the nursing staff's emergency triage and its correlation between the decision making capability (unpublished thesis)", National Taipei College of Nursing, Taiwan, 2003.
- [6] J. Brillman, D. Doezema, & D. Tandberg, "Triage : Limitations in predicting the need for emergency care and hospital admissions," *Annals of Emergency Medicine*, vol.27(4), 1996, pp.493-500.
- [7] E. G. Estrada, "Triage Systems," *Nursing Clinics of North America*, vol.16(1), 1981, pp.13-22.
- [8] A. VanBoxel, "How We Do It: Improving the triage process," *Journal of Emergency Nursing*, vol.21(4), 1995, pp.332-334.
- [9] J. Reinschmidt, H. Gottschalk, H. Kim, & D. Zwietering, *Intelligent Miner for Data: Enhance Your Business Intelligence*. USA: IBM International Technical Support Organization. 1999.
- [10] M. L. Huang, & H. Y. Chen, "Development and comparison of automated classifiers for glaucoma diagnosis using stratus optical coherence tomography," *Investigative Ophthalmology & Visual Science*, vol.46(11), 2005, pp.4121-4129.
- [11] M. A. Abdelfattah, A. T. El-Shahat, M. E. Ahmad, A. A. Mosaad, M. O. Mohamed, & E. S. Gamal, "Discrimination function based on hyaluronic acid and its degrading enzymes and degradation products for differentiating cirrhotic from non-cirrhotic liver diseased patients in chronic HCV infection," *Clinica Chimica Acta*, vol.369(1), 2006, pp.66-72.
- [12] E. Turban, J. E. Aronson, & T. P. Liang, *Decision Support and Intelligent Systems (7th ed.)*. Pearson: Prentice Hall, 2005.
- [13] R. Sharda, & D. Delen, "Predicting box-office success of motion pictures with neural networks," *Expert Systems with Applications*, vol.30(2), 2006, pp.243-254.
- [14] A. M. Heidar, B. K. Nicolaos, & B. Mahesh, "Short-term electric power load forecasting using feedforward neural networks," *Expert Systems*, vol.21(3), 2004, pp. 157-166.
- [15] G. Handyside, *Triage in Emergency Practice*. St. Louis, MO: Mosby, 1996.
- [16] Y. C. Ye, *The Application and Practice of Neural Network Models*, Taipei: Scholar Books, 2011.