

Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network

Mei-Ling Huang*, Yung-Yan Hsu

Department of Industrial Engineering & Management, National Chin-Yi University of Technology, Taichung, Chinese Taipei
Email: huangml@ncut.edu.tw

Received 1 June 2012; revised 5 July 2012; accepted 15 July 2012

ABSTRACT

Fetal distress is one of the main factors to cesarean section in obstetrics and gynecology. If the fetus lack of oxygen in uterus, threat to the fetal health and fetal death could happen. Cardiotocography (CTG) is the most widely used technique to monitor the fetal health and fetal heart rate (FHR) is an important index to identify occurs of fetal distress. This study is to propose discriminant analysis (DA), decision tree (DT), and artificial neural network (ANN) to evaluate fetal distress. The results show that the accuracies of DA, DT and ANN are 82.1%, 86.36% and 97.78%, respectively.

Keywords: Fetal Distress; Cardiotocography (CTG); Discriminant Analysis; Decision Tree; Artificial Neural Network

1. INTRODUCTION

Data mining is the key core of knowledge discovery in database. Through retrieving hidden, previously unknown, credible, innovative, or effective knowledge from large dataset, data mining is powerful on realization, analysis and prediction of the unknown phenomena [1].

With the advances in medical technology, clinical difficult exists for doctors to diagnose diseases when facing the tremendous amounts of disease inspection reports generating from the latest medical examination equipment. Data mining has successfully assisted medical diagnosis in the past a few years to prevent misdiagnosis.

Caesarean section is one of the common surgical procedures in the obstetrics and gynecology. In Taiwan, the proportion of caesarean section has increased year by year. The reasonable explanation is the parents' belief on fortunate birth timing. Regardless of parents' belief, caesarean section must be taken when abnormal fetal position, fetal distress, or other emergency situation occurs. Fetal distress denotes the possibility of fetal hypoxia

and/or metabolic acidosis [2]. It is usually a complication of the labor [3]. Adequate fetal oxygenation is essential for a healthy neonate [4], and efficient and effective diagnosis on fetal distress is an important issue.

The cardiotocography (CTG) and fetal scalp blood sampling (FBS) are commonly used on monitoring fetal distress [1]. CTG provides fetal development and health information, especially the maturation status of autonomous nervous system. The change of fetal heart rate (FHR) represents the response to external stimuli. In general, average of FHR, change of FHR, acceleration and deceleration of FHR, and fetal movement are essential parameters on medical diagnosis of fetal distress [5].

This study aims to predict the occurrence of fetal distress by applying data mining technique to establish classification models.

Discriminant analysis (DA) is to predict group membership for newly sampled data based upon a group of observations, whose memberships are already identified. The discriminant function is to minimize their group misclassifications [6]. Polat *et al.* [7] used generalized discriminant analysis and least square support vector machine on diagnosis of diabetes disease. Multivariate discriminant analysis was applied to build a predictive model and perform tissue-type classifications [8].

Decision tree (DT) is an approach to build a classification model and a tree-shaped structure was produced using inductive reasoning. Decision tree is highly suitable for conducting medical predictions and data analysis explanations [9]. One example is the application of contingent valuation and decision tree to investigate public preferences on water quality improvements by Atkins *et al.* [10]. And classification and regression trees (CART) algorithm is examined to classify hyper-spectral data of experimental corn plots into different categories of water stress, presence of weeds and nitrogen application rates by Waheed *et al.* [11].

The basic concept of artificial neural network (ANN) is that computer can imitate human nervous system and function like human thinking to solve problems. Kohonen [12] defined ANN are massively parallel intercon-

*Corresponding author.

nected networks of simple elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as biological nervous systems do. Lin *et al.* [13] developed regression and artificial neural network models to predict the outcome of surgical treatment in older patients with hip fracture. Chang and Chen [9] combining decision tree with artificial neural network to construct the best predictive classification model on diagnosis of dermatology.

Due to fetal distress causing severe sequela on fetus, the purpose of this study is to establish classification models and compare accuracies among models to help obstetricians on diagnosis of fetal distress.

The rest of sections of this paper are organized as follows. Section 2 introduces the database and methods. Section 3 shows the analytic results of proposed models. In Section 4, we discuss the results and compare the results with other research and explain the limitations of our study.

2. MATERIAL AND METHODS

2.1. Data

The Cardiotocography data set used in this study is from UCI Machine Learning Repository [14]. It contains the

Fetal Heart Rate, measurements from Cardiotocography, and the diagnosis group classified by gynecologist. There are 21 attributes, including 11 continuous, 9 discrete and 1 nominal scales. The number of observations is 2126 cases in total, including 1655 normal cases, 295 suspects, and 176 fetal distress cases. **Tables 1** and **2** display the details for the dataset.

2.2. Methods

The methods used in this study including multiple group discriminant analysis, decision tree, and back propagation artificial neural network. Combinations of different number of neurons and different transfer functions were executed for artificial neural network. The dataset contains no missing values, and all the continuous attributes were used as the input for three analytic models. While attributes with high correlation were deleted from the original input dataset, the final input dataset includes attribute number 8, 9, 10, 11, 12, 13, 17, 18, 19 and 20 in **Table 1**. **Figure 1** shows the study scheme for this study.

2.3. Discriminant Analysis (DA)

Discriminant analysis comes to understand the difference between/among groups with respect to the combination

Table 1. Description of fetal distress dataset.

No.	Attributes	Explanation	Scale
1	LB	FHR baseline (beats per minute)	Discrete
2	AC	Accelerations per second	Discrete
3	FM	Fetal movements per second	Discrete
4	UC	Uterine contractions per second	Discrete
5	DL	Light decelerations per second	Discrete
6	DS	Severe decelerations per second	Discrete
7	DP	Prolonged decelerations per second	Discrete
8	ASTV	Percentage of time with abnormal short term variability	Continuous
9	MSTV	Mean value of short term variability	Continuous
10	ALTV	Percentage of time with abnormal long term variability	Continuous
11	MLTV	Lean value of long term variability	Continuous
12	Width	Width of FHR histogram	Continuous
13	Min	Minimum of FHR histogram	Continuous
14	Max	Maximum of FHR histogram	Continuous
15	Nmax	Histogram peaks	Discrete
16	Nzeros	Histogram zeros	Discrete
17	Mode	Histogram mode	Continuous
18	Mean	Histogram mean	Continuous
19	Medium	Histogram median	Continuous
20	Variance	Histogram variance	Continuous
21	Tendency	Histogram tendency	Nominal

Table 2. ANOVA for data attributes.

Attributes	Normal	Suspect	Pathologic	p-value
LB	131.98 ± 9.45	141.68 ± 7.88	131.6875 ± 9.4062	<0.001
AC	0.0040 ± 0.0040	0.0003 ± 0.0007	0.0004 ± 0.001	<0.001
FM	0.00080 ± 0.0410	0.0083 ± 0.0416	0.0257 ± 0.0855	<0.001
UC	0.0048 ± 0.0027	0.0024 ± 0.0026	0.0038 ± 0.0036	<0.001
DL	0.0019 ± 0.0029	0.0005 ± 0.0015	0.0037 ± 0.0041	<0.001
DS	5.38E-07 ± 2.188E-05	0 ± 0	3.82E-05 ± 0.0002	<0.001
DP	5.12E-05 ± 0.0003	8.65E-05 ± 0.0004	0.0013 ± 0.0013	<0.001
ASTV	42.4659 ± 15.4927	61.9017 ± 11.7413	64.5398 ± 14.4111	<0.001
MSTV	1.4306 ± 0.8187	0.6389 ± 0.6599	1.5756 ± 1.1960	<0.001
ALTV	5.0447 ± 11.5226	29.03051 ± 20.2309	22.8466 ± 33.8990	<0.001
MLTV	8.7055 ± 5.8204	8.0261 ± 3.7476	3.5886 ± 4.0802	<0.001
Width	73.4 ± 36.3373	49.1593 ± 39.4459	78.3466 ± 48.9322	<0.001
Min	91.0858 ± 27.24531	113.2949 ± 30.9589	83.9830 ± 34.0797	<0.001
Max	164.4858 ± 17.7593	162.4542 ± 16.3088	162.3295 ± 21.6127	0.085
Nmax	4.1631 ± 2.8590	3.3119 ± 3.0947	4.4432 ± 3.3058	<0.001
Nzeros	0.3353 ± 0.6904	0.2441 ± 0.7992	0.3466 ± 0.6736	0.111
Mode	138.2586 ± 13.4132	146.5559 ± 10.9648	114.608 ± 26.0725	<0.001
Mean	135.1021 ± 13.0117	144.7593 ± 10.5010	112.9773 ± 22.7181	<0.001
Median	138.4598 ± 12.5937	147.0712 ± 10.4735	119.5625 ± 19.2831	<0.001
Variance	17.4798 ± 22.2891	7.2136 ± 18.2193	50.7330 ± 60.4095	<0.001
Tendency	0.3420 ± 0.5891	0.4339 ± 0.5893	-0.0739 ± 0.6910	<0.001

of valuable attributes which best explains the data. Identifying the best set of attributes with significantly statistical differences between/among groups is the first step of discriminant analysis. The combination of attributes is used to form the discriminant function to separate groups of observations into predefined groups. It can be further applied to predict the associated group for new observations.

The idea underlying this classification method is to develop a linear combination, F , of n variables as $F = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$ with values for $\beta_1, \beta_2, \dots, \beta_n$ chosen to maximize the difference between groups and to minimize the difference within groups. This widely used nonparametric statistical method is based on a projection of the multidimensional feature vector x onto an optimal chosen vector w ($y = w^T x$) to maximize the ratio of between-class scatter to within-class scatter. Detailed descriptions of the statistical procedures involved in the assessment and use of a DA model can be found in the literature [15].

2.4. Decision Tree (DT)

Decision tree consists of root, decision nodes, branch nodes, and leaf nodes. It begins with the root, and each branch is a new decision node or a new leaf node of the

tree. Each decision node represents a decision with respect to the attribute of classification observation. And each leaf node reflects one possible classification result.

The attribute with maximum information is chosen to develop the next node based upon the information gain from information theory. And the specific value for this attribute is determined to set up the branch. Continue this step to develop the decision tree until satisfying the predetermined stopping criteria. The following steps describe the details of building the decision tree:

Step 1: Split the whole dataset into training and testing sets.

Step 2: Use the training set as the input for tree root.

Step 3: Grow the tree by selecting one best attribute base upon the information gain theory for each node.

Step 4: Prune the grown tree by using the testing set until each node has only one node.

Step 5: repeat step 1 to 4 until all the nodes are leaf nodes.

2.5. Artificial Neural Network (ANN)

The artificial neural network is derived from imitating the picture of human neural system. ANN consists of many nonlinear neurons and the linkages communicating the neurons in the system. The neurons perform numer-

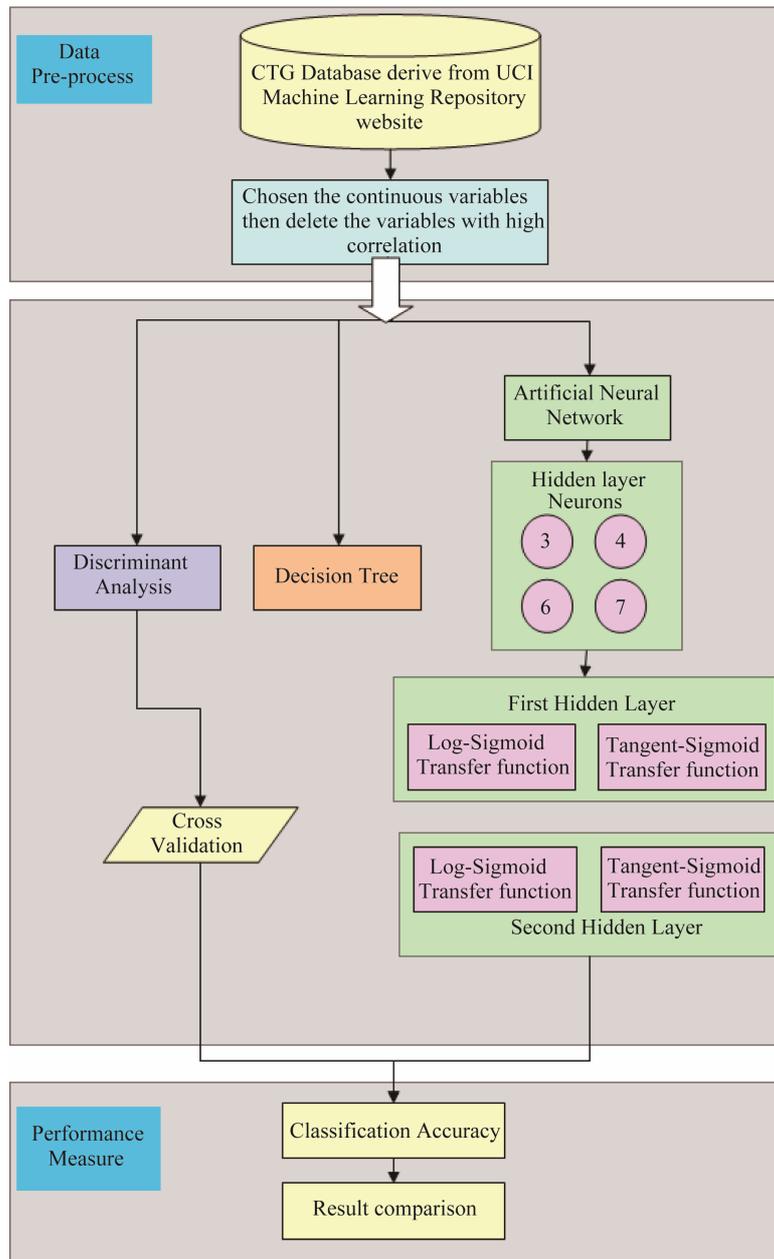


Figure 1. Flow chart of this study.

ous operations simultaneously. And the operation results are treated as the input for other neurons. The relationship between input and the output can be expressed as below:

$$Y_j = f\left(\sum_{i=1}^n W_{ij} X_i - \theta_j\right) \quad (4)$$

where X_i is the input of imitating neurons, Y_j is the output of imitating neurons, f is the transfer function dealing with the transformation, which is the sum of weighted products from input neurons, W_{ij} is the strength of the linkage between neurons, and θ_j is the threshold value.

2.6. Back Propagation Neural Network (BPNN)

Back propagation neural network is a multilayer perceptron (MLP) and the generally used learning algorithm is error back propagation. This combination is called as back propagation neural network, which is one of the most commonly applied neural networks. One weight is determined for each attribute, and the transformation of sum of the weighted attributes are treated as the input of next neuron. An ANN consists of one input layer, one output layer, and one or more hidden layers that extract essential information during learning. The parameters are adjusted iteratively with a supervised learning process

during which the network learns to associate input vectors with the appropriate pattern class. The gradient steepest descent algorithm with back-propagation of error and a momentum term were used during training in this study. **Figure 2** displays a standard BPNN, consisting of the following layers:

1) Input Layer: represent the input of network, and the number of neurons depends on the associated problems.

2) Hidden Layer: represent the interaction of input units. There is no standard to determine the number of neurons, and the most commonly used method is try and error. The number of hidden layer could be zero, one, or more than one.

3) Output Layer: represent the output of the network, and the number of neurons depends on the associated problems.

Matlab7.6 software was used for BPNN model in this study. Two hidden layers with Log-sigmoid transfer function and Tangent-sigmoid transfer function were used in this study. There is no standard guide line for setting the number of neurons for hidden layers. Although the more neurons, the less error, the error could not be further reduced after reaching some certain number of neurons. **Table 3** displays suggestions for determining the number of neurons from the related research, where n_i denotes the number of neurons for input layer; n_j denotes the number of neurons for output layer.

The transfer function is to convert the aggregation of input information to allow the output values locate within a reasonable range. Log-sigmoid transfer function and Tangent-sigmoid transfer function are commonly used for MLP and the output values rang 0 and 1, -1 and 1, respectively. **Figures 3** and **4** show the transformation for Log-sigmoid transfer function and Tangentsigmoid transfer function.

3. RESULTS AND DISCUSSION

3.1. Discriminant Analysis Result

Minitab 14 was used for discriminant analysis. All dataset was first normalized. Attribute "Max", marked as high correlation with other attributes, was deleted before analysis. Therefore, the attributes fed in the discriminant models are attribute number 8, 9, 10, 11, 12, 13, 17, 18, 19 and 20 in **Table 1**. **Table 4** shows the results after cross validation of discriminant analysis. The classification accuracies for normal, suspect and pathologic groups are 0.847, 0.753, and 0.688, respectively. And the overall accuracy is 0.820 Units.

3.2. Decision Tree Result

Answer Tree software was used to develop the decision tree in this study. The whole dataset was split into training and testing sets. 80% of the whole dataset was treated

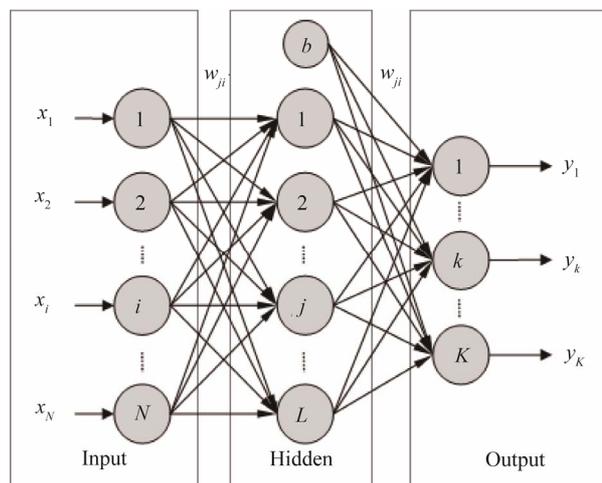


Figure 2. The structure of back propagation neural network.

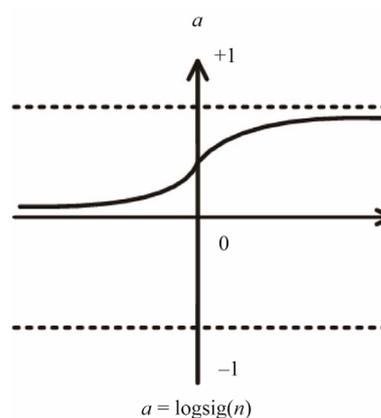


Figure 3. Log-sigmoid transfer function.

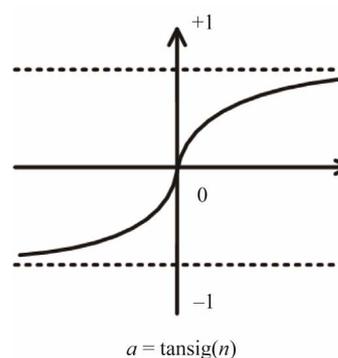


Figure 4. Tangent-sigmoid transfer function.

as training set, and the remaining 20% was used as testing set. Ten-fold cross validation technique was employed to assess how well the tree structure generalized from the data. **Table 5** shows the 10-cross validation results and **Figure 5** illustrates the decision rules and classification results. The classification accuracies for normal, suspect and pathologic groups are 0.932, 0.681, and

Table 3. Calculation of number of neurons for hidden layer.

Author	Formula	Values for this study	Selected number of neurons	
Yeh [16]	$\frac{n_i + n_o}{2}$	6.5	6	7
	$\sqrt{n_i \times n_o}$	3.6	3	4
Lu Tieding <i>et al.</i> [17]	$(\sqrt{0.34n_i n_o + 0.12n_o^2 + 2.54n_i + 0.77n_o + 0.53}) + 0.51$	6.79	6	7
Paola [18]	$\frac{2 + (n_o \times n_i) + [(n_o + n_i) \times 0.5n_o] - 3}{n_o + n_i}$	3.7	3	4
Wang [19]	$\frac{2n_i}{3}$	6.67	6	7

Table 4. Cross-validation results of discriminant analysis.

Predict	Actual		
	Normal	Suspect	Pathologic
Normal	1401	60	8
Suspect	195	222	47
Pathologic	59	13	121
Total	1655	295	176
Accuracy	0.847	0.753	0.688

Total accuracy: 0.820.

Table 5. Cross-validation results of decision tree.

Predict	Actual		
	Normal	Suspect	Pathologic
Normal	1542	92	15
Suspect	106	201	68
Pathologic	10	2	93
Total	1655	295	176
Accuracy	0.932	0.681	0.528

Total accuracy: 0.864.

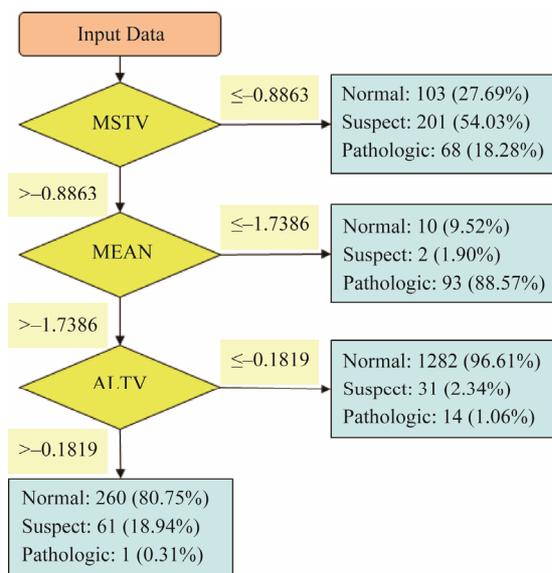


Figure 5. Decision rules and classification results.

0.528, respectively. And the overall accuracy is 0.864. **Table 6** and **Figure 5** illustrate the decision rules and classification results. From **Table 6** and **Figure 5**, rule number 1 predicts more “suspect” when (MSTV ≤ -0.8863); sample will be classified as “Pathologic” if (MSTV > -0.8863) and (MEAN ≤ -1.7386) based on rule number 2; sample will be classified as “Normal” if (MSTV > -0.8863) and (MEAN > -1.7386) based on rule number 3 and 4. Compared with discriminant analysis and artificial neural network, the advantage of decision tree illustrates precise information that helps for immediate diagnosis of fetal distress.

3.3. Back Propagation Neural Network Result

Matlab7.6 software was applied to build the neuron network. Eighty percent, 10%, and the remaining 10% of the whole dataset were randomly selected as training, testing, and validation sets, respectively. Two hidden

Table 6. Decision rules and classification results.

No.	Decision rules	Probability (%)	Number	Type
1	IF (MSTV \leq -0.8863)	27.69	103	Normal
		54.03	201	Suspect
		18.28	68	Pathologic
		9.52	10	Normal
2	IF (MSTV $>$ -0.8863) and (MEAN \leq -1.7386)	1.90	2	Suspect
		88.57	93	Pathologic
		96.61	1282	Normal
3	IF (MSTV $>$ -0.8863) (MEAN $>$ -1.7386) and (ALTV \leq -0.1819)	2.34	31	Suspect
		1.06	14	Pathologic
		80.75	260	Normal
4	IF (MSTV $>$ -0.8863) (MEAN $>$ -1.7386) and (ALTV $>$ -0.1819)	18.94	61	Suspect
		0.31	1	Pathologic

Table 7. Accuracies for ANN with combination of neurons and transfer functions.

The number of neurons	The combination of transfer function for hidden layers							
	First Layer	Second Layer	First Layer	Second Layer	First Layer	Second Layer	First Layer	Second Layer
	Tangent-sigmoid	Tangent-sigmoid	Tangent-sigmoid	Log-sigmoid	Log-Sigmoid	Tangent-sigmoid	Log-sigmoid	Log-Sigmoid
3		0.947		0.934		0.778		0.945
4		0.786		0.923		0.864		0.848
6		0.876		0.963		0.978		0.840
7		0.899		0.961		0.903		0.940

layers with neurons number 3, 4, 6, and 7 were executed, and the two transformation functions were Log-Sigmoid and Tangent-sigmoid functions. **Table 7** shows the classification results. The highest classification accuracy 0.978 was obtained with 6 neurons in hidden layer, and Log-Sigmoid and Tangent-sigmoid functions were used in the first and the second layers, respectively.

The classification models we established were applied on CTG database to predict the occurrence of fetal distress. The accuracies obtain from DA and DT models are 0.820 and 0.864, respectively. The highest accuracy in different combination of transfer function in hidden layers of ANN model is 0.978. The results show that the classification accuracy of ANN model is better than DA and DT models.

Yarnguy and Kanarkard [20] use of radial basis function network (RBF) classifier committees trained on different features for improving classification accuracy in CTG diagnosis. The results showed that the committee correct identification was 0.998, which is higher than our study. The disadvantage of unknown data process via RBF classifier still remains in their study. The reason why our results shows lower classification accuracy might be we only use the continuous attributes in our models, the information implicit in discrete attributes perhaps pos-

sess valuable information.

4. CONCLUSIONS

This study analyzes Cardiocography data set from UCI Machine Learning Repository website with discriminant analysis, decision tree, and artificial neural network techniques to create the classification models to predict fetal distress. Similar study with ANN application of the exceptional accuracy 0.988 was done on breast cancer diagnosis [21]. And the other study with the application of decision tree reached 0.801 accuracy was found on glaucoma diagnosis [22]. Although the highest accuracy is derived from the ANN model in this study, but the processes in ANN model just like a "black box", we couldn't understand how it works inside. The accuracy derived from the DT model isn't higher than ANN model, but it shows the concise and effective decision rules to classify the situation of fetus. Those decision rules could support obstetrician determine whether the fetus is suffering fetal distress effectively and efficiently.

This database contains 21 attributes, involving 11 continuous attributes and 10 discrete attributes. Since discrete attribute could not be applied in discriminant model, the continuous attributes were employed in this study. The weakness of the study is the information implicating

by the discrete attributes was not considered. The future study will use the classification techniques which can be fit in discrete attributes, e.g. logistic regression and rough set etc. In addition, feature selection technique can also be considered to use in data preprocessing stage. That is, first, extracting the effective attributes, then use the extracted attributes to establish models. It will reduce the analysis time and drop the stored space significantly and the efficiency of data analysis may be highly improved.

REFERENCES

- [1] Maimon, O. and Rokach, L. (2005) The Data mining and knowledge discovery handbook. Springer, New York. [doi:10.1007/b107408](https://doi.org/10.1007/b107408)
- [2] Chandrharan, E. and Arulkumaran, S. (2007) Prevention of birth asphyxia: Responding appropriately to cardiograph (CTG) traces. *Best Practice & Research Clinical Obstetrics and Gynaecology*, **21**, 609-624. [doi:10.1016/j.bpobgyn.2007.02.008](https://doi.org/10.1016/j.bpobgyn.2007.02.008)
- [3] Lunghi, F., Magenes, G., Pedrinazzi, L. and Signorini, M. G. (2005) Detection of fetal distress through a support vector machine based on fetal heart rate parameters. *Computers in Cardiology*, 25-28 September 2005, 247-250. [doi:10.1109/CIC.2005.1588083](https://doi.org/10.1109/CIC.2005.1588083)
- [4] Vigil-De Gracia, P., Simiti, E. and Lora, Y. (2000) Intrapartum fetal distress and magnesium sulfate. *International Journal of Gynecology & Obstetrics*, **68**, 3-6.
- [5] Romano, M., Bracale, M., Cesarelli, M., et al. (2006) Antepartum cardiotocography: A study of fetal reactivity in frequency domain. *Computers in Biology and Medicine*, **36**, 619-633. [doi:10.1016/j.combiomed.2005.04.004](https://doi.org/10.1016/j.combiomed.2005.04.004)
- [6] Sueyoshi, T. (2001) Extend DEA-discriminant analysis. *European Journal of Operational Research*, **131**, 324-351. [doi:10.1016/S0377-2217\(00\)00054-0](https://doi.org/10.1016/S0377-2217(00)00054-0)
- [7] Polat, K., Günes S. and Arslan, A. (2008) A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine. *Expert System with Applications*, **34**, 482-487. [doi:10.1016/j.eswa.2006.09.012](https://doi.org/10.1016/j.eswa.2006.09.012)
- [8] Piacenti da Sliva, M., Zucchi, O.L.A.D., Ribeiro-Silva, A., et al. (2009) Discriminant analysis of trace elements in normal, benign and malignant breast tissues measured by total reflection X-ray fluorescence. *Spectrochimica Acta Part B*, **64**, 587-592. [doi:10.1016/j.sab.2009.05.026](https://doi.org/10.1016/j.sab.2009.05.026)
- [9] Chang, C.L. and Chen, C.H. (2009) Applying decision tree and neural network to increase quality of dermatologic diagnosis. *Expert Systems with Applications*, **36**, 4035-4041. [doi:10.1016/j.eswa.2008.03.007](https://doi.org/10.1016/j.eswa.2008.03.007)
- [10] Atkins, J.P., Burdon, D. and Allen, J.H. (2007) An application of contingent valuation and decision tree analysis to water quality improvement. *Marine Pollution Bulletin*, **55**, 591-602. [doi:10.1016/j.marpolbul.2007.09.018](https://doi.org/10.1016/j.marpolbul.2007.09.018)
- [11] Waheed, T., Bonnell, R.B., Prasher, S.O. and Paulet, E. (2006) Measuring performance in precision agriculture: CART-A decision tree approach. *Agricultural Water Management*, **84**, 173-185. [doi:10.1016/j.agwat.2005.12.003](https://doi.org/10.1016/j.agwat.2005.12.003)
- [12] Kohonen, T. (1988) An introduction to neural computing. *Neural Networks*, **1**, 3-6. [doi:10.1016/0893-6080\(88\)90020-2](https://doi.org/10.1016/0893-6080(88)90020-2)
- [13] Lin, C.C., Ou, Y.K., Chen, S.H., et al. (2010) Comparison of artificial neural network and logistic regression models for predicting mortality in elderly patients with hip fracture. *International Journal of the Care of the Injured*, **41**, 869-873.
- [14] Frank, A. and Asuncion, A. (2010) UCI Machine Learning Repository. University of California, School of Information and Computer Science. Irvine. <http://archive.ics.uci.edu/ml>
- [15] Kleinbaum, D.G., Kupper, L.L. and Muller, K.E. (1998) Applied regression analysis and other multivariate methods. 2nd Edition, PSW-Kent, Boston.
- [16] Yeh, Y.C. (2003) Application and practice of artificial neural network. Scholar Books Co., Ltd., Taipei.
- [17] Lu, T., Chen, X. and Zhou, S. (2010) Optimization for impact factors of dam deformation based on BP neural network model. *International Conference on Intelligent Computation Technology and Automation*, Changsha, 11-12 May 2010, 854-857.
- [18] Paola, J.D. (1994) Neural network classification of multispectral imagery. Master's Thesis, University of Arizona, Tucson.
- [19] Wang, C. (1994) A theory of generalization in learning machine with neural application. Ph.D. Thesis, University of Pennsylvania, Philadelphia.
- [20] Yanguy, T. and Kanarkard, W. (2011) A radial basis function committee machine for cardiotocography classification. *The 12th Graduate Research Conference*, Khon Kaen University, Khon Kaen, 262-267.
- [21] Huang, M.L., Hung, Y.H. and Chen, W.Y. (2010) Neural network classifier with entropy based feature selection on breast cancer diagnosis. *Journal of Medical Systems*, **34**, 865-873. [doi:10.1007/s10916-009-9301-x](https://doi.org/10.1007/s10916-009-9301-x)
- [22] Huang, M.L. and Chen, H.Y. (2011) Glaucoma classification model based on GDx VCC measured parameter by decision tree. *Journal of Medical Systems*, **34**, 1141-1147. [doi:10.1007/s10916-009-9333-2](https://doi.org/10.1007/s10916-009-9333-2)