Data mining for the diagnosis of type 2 diabetes

Alexis Marcano-Cedeño
Group for Automation in Signals and Communications.
Technical University of Madrid (UPM).
Madrid, Spain.
a.marcano@gc.ssr.upm.es

Diego Andina
Group for Automation in Signals and Communications.
Technical University of Madrid (UPM).
Madrid, Spain.
d.andina@gc.ssr.upm.es

***Abstract*—Diabetes is the most common disease nowadays in all populations and in all age groups. diabetes contributing to heart disease, increases the risks of developing kidney disease, blindness, nerve damage, and blood vessel damage. Diabetes disease diagnosis via proper interpretation of the diabetes data is an important classification problem. Different techniques of artificial intelligence has been applied to diabetes problem. The purpose of this study is apply the artificial metaplasticity on multilayer perceptron (AMMLP) as a data mining (DM) technique for the diabetes disease diagnosis. The Pima Indians diabetes was used to test the proposed model AMMLP. The results obtained by AMMLP were compared with decision tree (DT), Bayesian classifier (BC) and other algorithms, recently proposed by other researchers, that were applied to the same database. The robustness of the algorithms are examined using classification accuracy, analysis of sensitivity and specificity, confusion matrix. The results obtained by AMMLP are superior to obtained by DT and BC.**

***Keywords*-ANNs, Artificial Metaplasticity, Data Mining, Decision trees, Diabetes**

I. Introduction

Diabetes is a major health problem in both industrial and developing countries, and its incidence is rising. It is a disease in which the body does not produce or properly use insulin, the hormone that "unlocks" the cells of the body, allowing glucose to enter and fuel them [1]. Diabetes increases the risks of developing kidney disease, blindness, nerve damage, blood vessel damage and it contributes to heart disease. The World Health Organization in 2000 indicated there were approximately 170 million people with diabetes, and estimated that the number of cases of the disease worldwide will be more than doubled to 366 million by 2030 [2]. Diabetes occurs in two major forms: type 1, or insulindependent diabetes, and type 2, or non-insulin-dependent diabetes. The type 1 diabetes, is characterized by an absolute deficiency of insulin secretion. Individuals at increased risk of developing this type of diabetes can often be identified by serological evidence of an autoimmune pathologic process occurring in the pancreatic islets and by genetic markers [3][4]. The most common form of diabetes is Type 2 [3][4][5]. It is believed that some of the causes of type 2 diabetes are associated with: diet changes, aging, urbanization, and increasing prevalence of obesity and physical inactivity [3][4]. This type diabetes results from insulin resistance (a condition in which the body fails to properly use insulin), combined with relative insulin deficiency. In Type 2 diabetes, either the body does not produce enough insulin or the cells ignore the insulin [3][4]. Although detection of diabetes is improving, about half of the patients with Type 2 diabetes are undiagnosed and the delay from disease onset to diagnosis may exceed 10 years. Thus, earlier detection of Type 2 diabetes and treatment of hyperglycaemia and related to metabolic abnormalities is of vital importance [3][4][6].

There are many factors to analyze to diagnose the diabetes of a patient, and this makes the physician's job difficult. There is no doubt that evaluation of data taken from patient and decisions of experts are the most important factors in diagnosis. But, this is not easy considering the number of factors that has to evaluate [7]. To help the experts and helping possible errors that can be done because of fatigued or inexperienced expert to be minimized, classification systems provide medical data to be examined in shorter time and more detailed. Expert systems and different artificial intelligence techniques for classification systems in medical diagnosis is increasing gradually. As for other clinical diagnosis problems, classification systems have been used for diabetes diagnosis problem [8], [9], [10].

The purpose of this study is apply the artificial metaplasticity on multilayer perceptron (AMMLP) as a data mining (DM) technique for the diabetes disease diagnosis. The Pima Indians diabetes was used to test the proposed model AMMLP. The results obtained by AMMLP were compared with decision tree (DT), Bayesian classifier (BC) and other algorithms, recently proposed by other researchers, that were applied to the same database. The DT and BC algorithms adopted in this research are based on weka (waikato environment for knowledge analysis) package [11]. The robustness of the algorithms are examined using classification accuracy, analysis of sensitivity and specificity and confusion matrix.

The remainder of this paper is organized as follows. Section 2 presents a brief introduction to data mining. Section 3 the algorithms and the data base used in this research are presented. Section 4 show the experimental results obtained. Finally, Section 5 presents the summarized conclusions.

II. Literature review

A. knowledge discovery in databases and Data mining

Fayyad et al., 1996 defined knowledge discovery in databases (KDD) as the nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data. It consists of the following main steps: (i) data preparation, (ii) data preprocessing, (iii) data mining (DM), (iv) evaluation and interpretation, (v) implementation. DM is only a step in the KDD process that consists of applying data analysis and discovery algorithms. DM tasks can be classified into two groups [13]: descriptive and predictive. These tasks can be accomplished by using various methods based on DM functions, used to specify the types of patterns to be mined. These functions include summarization (characterization), clustering, association, classification [14],[15].

Data mining classification technology contains two parts: construction of classification model, and evaluation of model classification efficiency. In the first part, the adopted classification algorithm is trained by a classified training data set in order to build classification predictive model. In the second part, testing data set is used to test classification efficiency of this model [15]. This study averaged 50 training for construction and efficiency evaluation. We use the following algorithms: artificial metaplasticity on multilayer perceptron (AMMLP), decision tree (DT) and bayesian classifier.

III. Materials and methods

A. Algorithms used

1) Artificial metaplasticity in MLP training: AMMLP: In neuroscience and other fields "metaplasticity" indicates a higher level of plasticity, expressed as a change or transformation in the way synaptic efficacy is modified. Metaplasticity is defined as *the induction of synaptic changes, that depends on prior synaptic activity* [16].

The proposed model defines artificial metaplasticity as a probabilistic learning procedure [17],[18] that produces greater modifications in the synaptic weights with less frequent patterns than frequent patterns, as a way of extracting more information from the former than from the latter. As Biological metaplasticity, AMP then favors synaptic strengthening for low-level synaptic activity, while the opposite occurs for high level activity[30]. The model is applicable to general ANNs. Andina *et al.* proposed general AMP concepts for ANNs, and demonstrate them over Radar detection data[18]. In this paper it has been implemented and tested for a MLP over the Australian and German credit approval data sets. The AMP implementation applied tries to improve results in learning convergence and performance by capturing information associated with significant rare events. *It is based on the idea of modifying the ANN learning procedure such that infrequent patterns which can contribute heavily to the performance, are considered with greater relevance during learning without changing the convergence of the error minimization algorithm.* It is has been proposed on the hypothesis that biological metaplasticity property maybe significantly due to an adaptation of nature to extract more information from infrequent patterns (low synaptic activity) that, according to Shannon's Theorem, implicitly carry more information[18].

AMP is analytically introduced in an arbitrary MLP training, by the application of a weighting function $w_X^*(x)$[16]:

$$f_X^*(x) = \frac{A}{\sqrt{(2\pi)^N}.e^{B\sum_{i=1}^{N} x_i^2}} = \frac{1}{w_X^*(x)} \quad (1)$$

where $w_X^*(x)$ is defined as $1/f_X^*(x)$, being $f_X^*(x)$ an approximation of the input patterns probability density function (*pdf*), N is the number of neurons in the MLP input layer, and parameters A and $B \in R^+$ are algorithm optimization values which depend on the specific application of the AMLP algorithm. Values for A and B have been empirically determined. Eq. (1) is a gaussian distribution. Then, $w_X^*(x)$ has high values for infrequent x values and close to 1 for the frequent ones and can therefore be straightforwardly applied in weights updating procedure for the AMP model of biological metaplasticity during learning.

As the *pdf* weighting function proposed is the distribution of the input patterns that does not depend on the network parameters, the AMMLP algorithm can then be summarized as a weighting operation for updating each weight in each MLP learning iteration as:

$$\Delta^* w = w^*(x)\,\Delta w \quad (2)$$

being $\Delta w = w(t+l) - w(t)$ the weight updating value obtained by usual BPA and $w^*(x)$ the realization of the described weighting function $w^*(x)$ for each input training pattern x. Note that for eqs. (8-10) the proposed implementation of metaplasticity in MLPs is analytically an implementation of importance sampling [20].

2) Decision tree algorithm: Decision trees (DT) are powerful classification algorithms and prediction, it that are widely used in data mining. The most commonly used DT algorithms include Quinlans ID3, C4.5, C5 [21] and Breimans classification and regression tree [22]. As the name implies, this technique recursively separates observations in branches to construct a tree for the purpose of improving the prediction accuracy. In doing so, they use mathematical algorithms (e.g., information gain, Gini index, and chi-squared test) to identify a variable and corresponding threshold for the variable that splits the input observation into two or more subgroups. This step is repeated at each leaf node until the complete tree is constructed. The objective of the splitting algorithm is to find a variable-threshold pair that maximizes the homogeneity (order) of the resulting two or more subgroups of samples [21].

The most commonly used mathematical algorithm for splitting includes Entropy based information gain (used in ID3, C4.5, C5), Gini index (used in CART), and chi-squared test (used in CHAID). Based on the favorable prediction results we have obtained from the preliminary runs, in this study we chose to use J48 algorithm as our decision tree method. The J48 algorithm is an implementation of the C4.5 algorithm [11] in the program WEKA [23].

3) Bayesian classifier: Theory of Bayesian classifier (BC) stems from Bayesian theorem in statistics, while presetting a hypothesis, i.e., every attribute is independent, so that the classifier can be simple and fast. According to Bayesian theorem, the probability of a set of data x_t belonging to c is:

$$P\left(C/X_t\right) = \frac{p\left(C\right) p\left(X_t/C\right)}{p\left(X_t\right)} \tag{3}$$

where

- $p\left(C\right)$ is the prior probability of C: the probability that C is correct before the data X_t are seen.
- $p\left(X_t/C\right)$ is the conditional probability of seeing the data X_t given that the hypothesis C is true. This conditional probability is called the likelihood.
- $p\left(X_t\right)$ is the marginal probability of X_t.
- $p\left(C/X_t\right)$ is the posterior probability: the probability that the hypothesis is true, given the data and the previous state of belief about the hypothesis

Based on above formula, BC calculates conditional probability of an instance belonging to each class, and based on such conditional probability data, the instance is classified as the class with the highest conditional probability. In knowledge expression, it has the excellent interpretability same as decision tree, and is able to use previous data to build analysis model for future prediction or classification [24] [25]. If the eigenvalues of data are continuous, there are two ways to process [24][25]:

- Suppose it be normal distribution and find (means, variances) of eigenvalues as likelihood.
- Use splitting method to transfer continuous data into discrete data.

B. Data description

The Pima Indians Diabetes data set studies the influence of diabetes on the American population of Pima Indians. A population of women of Pima Indians was tested for diabetes in accordance withWorld Health Organization criteria. These data belongs to the National Institute of Diabetes and Digestive and Kidney Diseases and is part of the UCI database [12]. The original data set is composed of 768 instances, with 8 numeric attributes and a class variable labeled 1 or 0 showing whether diabetes was present. There are 268 examples belonging to class 1 and 500 belonging to class 0.

All samples have eight features. These features are:

- Feature 1: Number of times pregnant (Pregn).
- Feature 2: Plasma glucose concentration a 2 h in an oral glucose tolerance test (Gluco).
- Feature 3: Diastolic blood pressure (mm Hg) (Diasp).
- Feature 4: Triceps skin fold thickness (mm) (Thick).
- Feature 5: 2-h serum insulin (lU/ml)(Insul).
- Feature 6: Body mass index (weight in $kg/(heightinm)^2$ (Massi).
- Feature 7: Diabetes pedigree function (Predf).
- Feature 8: 2-h serum insulin (lU/ml) (Predf).
- Feature 8: Age (years) (Age).
- Class: Diabetes onset within five years (0 or 1).

C. Data preparation

The quality of the data is the most important aspect as it influences the quality of the results from the analysis. The data should be carefully collected, integrated, characterized, and prepared for analysis. In this study, we applied the techniques of data preprocessing in order to improve the quality of the mining result and the efficiency of the mining process. In this study we analyze a data set composed of 768 data instances. A preliminary analysis of the data indicates the usage of zero for missing data. Since, it does not make sense to have the value of a variable such as plasma-glucose concentration 0 in living people; all the observations with zero entries are removed. After removing all the above said values and variables, only 763 instances remain from the data in our study.

To comparatively evaluate classifiers performance, all the classifiers presented in this particular case were trained with the same training data set and tested with the same evaluation data set. Each data set was divided into training and testing data randomly, in which there are $60 - 40\%$ training and testing sets per data set (The data split was made according to recent researches applying metaplasticity algorithms)[27],[28]

IV. Results

For validation methods the algorithm AMMLP, DT and BC were trained 50 times, respectively, to construct classification model and evaluate classification efficiency. In each attribute input mode, the same training and testing processes were repeated 10 times. Finally, the average classification efficiency of each classification model was calculated.

1) Network selection: To select the optimal configuration for each classifiers used in this study was tested different network structures and parameters, the results are presented in this session. We present only the top three results of each classifier. Table 1, Table 2 and Table 3 show different network structures and parameters for AMMLP, DT and BC respectively.

For this research, 50 AMMLPs were generated, with different weights and random values following a normal distribution (mean 0 and variance 1). In each experiment 50 networks were trained in order to achieve an average result that did not depend on the initial random value of the weights of the ANN. Two different criteria were applied to stop the training: in one case it was stopped when the error reached 0.01 (the error diminishes but cannot converge to 0), and in the other the training was conducted with a fixed number of 2.000 epochs. Different network structures and parameters A and B of metaplasticity were tested empirically, in order to obtain the parameters and network structures appropriate for this case. Structure with 8 neurons in the hidden layer and parameter values A=8 and B=0.2, mean squared error = 0.01 and epoch = 2000 where found valid for this case.

Table 1 shows the network structure, metaplasticity parameters, epochs and MSE used in the training and testing phases of the classifiers AMMLP.

Table I
AMMLP RESULTS OBTAINED FOR DIFFERENT NETWORK STRUCTURES AND PARAMETERS OF METAPLASTICITY ALGORITHM

Classifier	Network Structure			Metaplasticity Parameters		Mean Squared Error	Epoch	Accuracy
	I	HL	O	A	B			
AMMLP1	**8**	**4**	**1**	**37**	**0.2**	**0.01**	**2000**	**89.93%**
AMMLP2	8	7	1	38	0.3	0.01	2000	88.50%
AMMLP2	8	8	1	39	0.5	0.01	2000	87.85%

For the DT was used J48 algorithm with different confidence factor: 0.25, 0.4 and 0.6. Table 2 show the tree best results obtained by DT.

Table II
THE THREE BEST RESULTS OBTAINED BY DECISION TREE WITH DIFFERENT CONFIDENCE FACTOR

Classifier	Confidence factor	Minimum number objects	Accuracy (%)
Decision Tree, DT1	**0.25**	**2**	**77.60**
Decision Tree, DT2	0.1	2	72.08
Decision Tree, DT3	0.6	2	74.68

The BC was applied with and without kernel estimation (the attributes are modeled using a normal distribution). Table 3 show the 2 best results obtained by BC.

Table III
RESULTS OBTAINED BY BAYESIAN CLASSIFIER WITH AND WITHOUT USING KERNEL ESTIMATION

Classifier	Using kernel estimation	Accuracy (%)
Bayesian classifier, BC1	False	**78.57**
Bayesian classifier, BC2	**True**	78.51

A. Performance measures

The classification efficiency averages of three classification models were analyzed and compared to obtain the optimal model for predict, the improvement of patients with TBI and also know the features o profile of the patients that improved. In general, measures of quality of classification are built from a confusion matrix which records correctly and incorrectly recognition such as the true positive (TP), false positive (FP), false negative (FN) and the true negative (TN) in binary classification. In order to extend the usage of confusion matrix, we define the TP, FP, FN and TN in this paper as follows:

- True positive (TP): the TP denotes the number of patients having a diseases specify and are diagnosed as sick individuals
- False positive (FP): the FP denotes the number of patients having a diseases specify and are diagnosed as healthy individuals.
- False negative (FN): the FN denotes the number of patients no having a diseases specify and are diagnosed as sick individuals.
- True negative (TN): the TN denotes the number of patients no having a diseases specify and are diagnosed as healthy individuals.

1) Sensitivity, specificity and accuracy: The definitions of the sensitivity, specificity and accuracy prediction of classification are are defined by:

$$Sensitivity = \frac{TP}{TP + FN}(\%) \tag{4}$$

$$Specificity = \frac{TN}{FP + TN}(\%) \tag{5}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}(\%) \tag{6}$$

2) Performance evaluation: The accuracy measure assesses the overall effectiveness of the algorithm, while, the other two indices, specificity and sensitivity estimate the classifier performance on different classes. The two indices are often employed in medical applications. Table 6 shows the classification results obtained in the best simulation for each classifiers used in this study in a confusion matrix.

Table IV
CONFUSION MATRIX GENERATED WITH THE CLASSIFICATION RESULTS OBTAINED FOR EACH CLASSIFIER IN THE BEST SIMULATION.

Classifiers	Desired Result	Prediction	
		Diabetes	No Diabetes
AMMLP	Diabetes	82	26
	No Diabetes	5	195
DT	Diabetes	85	23
	No Diabetes	46	154
BC	Diabetes	70	38
	No Diabetes	28	172

The performance measures in terms of specificity, sensitivity and accuracy perdition of the classifiers applied in this research were validated used 10-fold cross-validation. These results are shows in table 4.

Table V
CLASSIFICATION ACCURACIES OF CLASSIFIERS APPLIED IN THIS RESEARCH VALIDATED USED 10-FOLD CROSS-VALIDATION.

Classifiers	Performance measures		
	Specificity (SD)	Sensitivity (SD)	Accuracy (SD)
AMMLP	**75.62% ± (0.3)**	**97.10% ± (0.5)**	**89.53% ± (0.4)**
DT	75.8% ± (2.5)	76.72% ± (2.3)	75.3% ± (2.3)
BC	64.01% ± (0.5)	86.6% ± (0.6)	77.77% ± (0.5)

SD: standard deviation.

As seen in the table 4 and table 5, the results obtained by AMMLP are superior to obtained by DT and BC in all cases. Also, comparing standard deviation, with the more stable classification efficiency, decision tree is the best classification algorithm.

3) Comparation: The prediction accuracy obtained by AMMLP was compared with the best results obtained by other researchers using the same database, are summarized in Table 6.

Table VI
CLASSIFICATION ACCURACIES OF PROPOSED MODEL AMMLP AND OTHER CLASSIFIERS FOR THE PIMA INDIANS DIABETES

Author (year)	Method	Accuracy (%)
Carpenter and Markuzon, (1998)	ARTMAP-IC	81.0
Deng and Kasabov, 2001	ESOM	78.4 ±1.6
Kayaer and Yildirim, 2003	GRNN	80.21
Abdel-Aal, (2005)	T-MC	77.6
Luukka and Lepplampi, (2006)	PCA-Entropy	80.47
Polat and Gunes, (2007)	PCA-ANFIS	89.47
Srinivasa,Venugopal and Patnaik, (2007)	SAMGA	74.6
Ji and Carin, (2007)	POMDP	71.43
Polat, Gunes, and Aslan, (2008)	LS-SVM	82.05
Kahramanli and Allahverdi, (2008)	FNN	84.2
Ghazavi and Liao, (2008)	FUZZY MODELS	77.65
Termutas *et al.* (2009)	MLNN-LM	82.37
Lekkas and Mikhailov (2010)	eClass	79.37
Aibinu *et al.,* (2010)	CVNN-GDA	81.00
Dogantekin *et al.,* (2010)	LDA-ANFIS	84.61
Aibinu, Salami and Shafie, (2011)	CAR	81.28
Castro, Nebot and Mugica ,(2011)	LR-FIR	75,39
Marcano-Cedeño *et al.,* (2011)	AMMLP	89.93
In this study, 2011	**AMMLP**	**89.53 ± 0.4**

V. CONCLUSION

In this research the artificial metaplasticity on multilayer perceptron (AMMLP) was applied as a data mining (DM) technique for the diabetes disease diagnosis. AMMLP is an improvement over the classical MLP trained with the BP algorithm. The AMMLP is based on the biological metaplasticity property of neurons. The Pima Indians diabetes was used to test the proposed model AMMLP. The performance measures obtained by AMMLP classifier are significantly superior to the obtained by decision tree and Bayesian classifier. When we compared the results obtained by AMMLP with other algorithm, recently proposed by other researchers, the AMMLP was only exceeded by one algorithm, but in this case the author presented the best result obtained in one simulation. The AMMLP obtained an excellent accuracy of 89.53%. From the results obtained in this study, we can conclude that the AMMLP can serve as a second opinion for physicians when making their final diagnostic decisions.

REFERENCES

[1] Mohamed E.I., Linderm R., Perriello G., Di Daniele N., Poppl S.J. and De Lorenzo A., Predicting type 2 diabetes using an electronic nose-base artificial neural network analysis. *Diabetes Nutrition & Metabolism*, Vol. 15(4), pp.215-221, 2002.

[2] Shaw J.E.,Sicree R.A.and Zimmet P.Z.. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Research and Clinical Practice*. Vol. 87, pp. 4-14, 2010. doi:10.1016/j.diabres.2009.10.007.

[3] American diabetes asociation, http://www.diabetes.org/diabetes-basics/

[4] International Diabetes Federation, http://www.idf.org

[5] Stavros L. and Ludmil M., Evolving fuzzy medical diagnosis of Pima Indians diabetesnext term and of dermatological diseases. *Artif. Intell. Med.* Vol. 50(2), pp. 117-126, 2010. doi:10.1016/j.artmed.2010.05.007

[6] Temurtas H., Yumusak N. and Temurtas F., A comparative study on diabetes disease diagnosis using neural networks. *Expert Syst. Appl.* Vol. 36(4), pp. 8610-8615, 2009. doi:10.1016/j.eswa.2008.10.032.

[7] Polat K., Gunes S. and Aslan A., A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine. *Expert Syst. Appl.*, Vol.34(1), pp. 214-221, 2008. doi:10.1016/j.eswa.2006.09.012.

[8] Jevtic A., Quintanilla-Dominguez J., Cortina-Januchs M.G. and Andina A., Edge detection using ant colony search algorithm and multiscale contrast enhancement.in *Proc. Systems, Man and Cybernetics*. pp. 2193 -2198 (2009). doi: 10.1109/ICSMC.2009.5345922

[9] Torres J., Vega A., Torres S. and Andina D., Chirp Detection Through Discrete Wavelet Transform. in *Proc. Engineering Academy and Society Conference on Signal Processing, Robotics And Automation*. pp. 1971-1975 (2002)

[10] Polat K. and Gunes S., An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. *Digital Signal Processing*. Vol. 17(4), pp. 702-710, 2007.

[11] Witten I.H. and Frank E., Data Mining Practical Machine Learning Tools and Techniques, pp. 525. Elsevier, San Francisco, CA. USA. 2nd. ed. (2005).

[12] UCI machine learning respiratory. *http://archive.ics.uci.edu/ml/datasets.html*.

[13] Data mining: Concepts and techniques 2nd ed. San Francisco: Morgan Kaufmann Publishers (2006). ISBN 1-55860-901-6

[14] Köksa G.,Batmaz I. and Caner M., A review of data mining applications for quality improvement in manufacturing industry.*Expert Syst. Appl.*, 2011. doi:10.1016/j.eswa.2011.04.063

[15] Duen-Yian Y., Ching-Hsue C. and Yen-Wen C., A predictive model for cerebrovascular disease using data mining. *Expert Syst. Appl.*, Vol 38(7), pp. 8970-8977, 2011. doi:10.1016/j.eswa.2011.01.114

[16] Marcano-Cedeño A., Quintanilla-Dominguez J., Andina D., Breast cancer classification applying artificial metaplasticity algorithm. *Neurocomputing*. Vol 74(8), pp. 1243-1250. 2011. doi: 10.1016/j.neucom.2010.07.019.

[17] Ropero-Pelaez F.J. and Andina D., Probabilistic versus Incremental Presynaptic Learning in Biologically Plausible Synapses. *LNCS*. Vol 6686, pp. 80-89. 2011. doi:10.1007/978-3-642-21344-1_9

[18] Andina D., Álvarez-Vellisco A., Jevtić A. and Fombellida J., Artificial metaplasticity can improve artificial neural networks learning. *Intell. Autom. Soft Co*. Vol 15(4), pp 683-696. (2009).

[19] D. Andina and Ropero-Pelaez F.R., On the Biological Plausibility of Artificial Metaplasticity. *LNCS*. Vol 6686, pp. 119-128. 2011. doi:10.1007/978-3-642-21344-1_13

[20] Andina D., Martinez-Antorrena J. and Melgar I., Importance sampling in neural detector training phase Importance sampling in neural detector training phase. in *Proc. of the World Automation Congress*. pp. 43-48 (2004). doi: 10.1109/WAC.2004.185387

[21] Quinlan J.R., C4.5 programs for machine learning. In Morgan Kaufmann Publishers. San Francisco, CA, USA. 1993, pp. 302. ISBN:1-55860-238-0.

[22] Breiman L., Friedman J.H., Olshen R.A. and Stone C.J., Classification and regression trees. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software. (1984).

[23] Weka. http://www.cs.waikato.ac.nz/ ml/index.html 23th January 2011

[24] Loether H.J. and McTavish D.G., Descriptive and inferential statistics: An introduction (4th ed.). Needham Heights, MA: Allyn and Bacon. (1993)

[25] Yan S., Yiyuan T., Shuxue D., Shipin L., Yifen C., Diagnose the mild cognitive impairment by constructing Bayesian network with missing data. *Expert Syst. Appl.*, Vol 38(1), pp. 442-449, 2011. doi:10.1016/j.eswa.2010.06.084

[26] Humar K. and Novruz A., Design of a hybrid system for the diabetes and heart diseases. *Expert Syst. Appl.*, Vol 35(1), pp. 82-89, 2008. doi:10.1016/j.eswa.2007.06.004

[27] Marcano-Cedeño A., Álvarez-Vellisco A. and Andina D., Artificial metaplasticity MLP applied to image classification. in *Proc. 7th Int. Conf. on Industrial Informatics*. 650-653 (2009). doi:10.1109/INDIN.2009.5195879.

[28] Marcano-Cedeño A., Quintanilla-Dominguez J., Andina D., Wood Defects Classification Using Artificial Metaplasticity Neural Network. in *Proc. 35th Annua Conf. on of the IEEE Industrial Electronics Society*, 3422-3427 (2009) doi:10.1109/IECON.2009.5415189.

[29] D. Andina and Ropero-Pelaez F.R., On the Biological Plausibility of Artificial Metaplasticity. *LNCS*. Vol 6686, pp. 119-128. 2011. doi:10.1007/978-3-642-21344-1_13

[30] Marcano-Cedeño A., Joaquín Torres and Andina D., A Prediction Model to Diabetes Using Artificial Metaplasticity . *LNCS*. Vol 6687, pp. 418-425. 2011. doi:10.1007/978-3-642-21326-7_45