International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 8 (2016) pp. 247-256

Performance evaluation of classification algorithms by excluding the most relevant attributes for dipper/non-dipper pattern estimation in Type-2 DM patients

Zehra Aysun Altikardes[1], Hasan Erdal[2], Ahmet Fevzi Baba[3], Ali Serdar Fak[4] and Hayriye Korkmaz[5]

[1] Marmara University, Vocational School of Technical Sciences,
Dept. of Computer Technologies, Istanbul, Turkey
aaltikardes@marmara.edu.tr

[2, 3, 5] Marmara University, Faculty of Technology,
Department of Electrical & Electronics Engineering, Istanbul, Turkey
herdal@marmara.edu.tr , fbaba@marmara.edu.tr , hkorkmaz@marmara.edu.tr

[4] Marmara University,
Hypertension and Atherosclerosis Center, Istanbul, Turkey
serdarfak@marmara.edu.tr

Abstract: **Diabetes Mellitus (DM) is a high prevalence disease that causes cardiovascular morbidity and mortality. On the other hand, the absence of physiologic night-time blood pressure decrease can further lead to morbidity problems such as target organ damage both in diabetics and non-diabetics patients. However, the Non-dipping pattern can only be measured by the 24-hour ambulatory blood pressure monitoring (ABPM) device. ABPM has certain challenges such as insufficient devices to distribute to patients, lack of trained staff or high costs. Therefore, in this study, it is aimed to develop a classifier model that can achieve a sufficiently high accuracy percentage for Dipper/non-Dipper blood pressure pattern in patients by excluding ABPM data.**

The study was conducted with 56 Turkish patients in Marmara University Hypertension and Atherosclerosis Center and School of Medicine Department of Internal Medicine, Division of Endocrinology between the years 2010 and 2012. Our purpose was to find out if the proposed method would be able to detect non-dipping/dipping pattern through various data mining algorithms in WEKA platform such as J48, NaïveBayes, MLP, RBF. All algorithms were run to get accurate Dipper/non-Dipper pattern estimation excluding the attributes of ABPM data.

The results show that Neural Network (MLP and RBF) algorithms mostly produced reasonably high classification accuracy, sensitivity and specificity percentages reaching up to 90.63% when the attributes were reduced. However in medical sciences, sensitivity is taken as a valid and reliable indication for diagnosis. Therefore, MLP had a higher sensitivity percentage (83.3%) than others. Also, ROC values, which had the closest values to 1, were achieved by RBF for each selection mode. ROC was 0.872 for 10 fold CV mode and 0.856 for percentage split mode.

Finally, ANN MLP and RBF algorithms were used, and it was observed that RBF algorithm had the highest success rate regarding sensitivity that was 83.3%. In medical diagnosis, a higher sensitivity performance is regarded as a more valid indication of metric than a higher specificity.

The proposed model could represent an innovative approach that might simplify and fasten the diagnosis process by skipping some steps in Dipper/non-Dipper diagnosis/prognosis.

Keywords: Diabetes, Blood pressure, Ambulatory monitoring, Classification, Attribute reduction.

I. Introduction

Patients with diabetes or hypertension are at high risk of developing cardiovascular diseases; however co-existence of these disorders increases the risk enormously. As an insidious prognosis, Diabetes Mellitus (DM) is a serious metabolic disease that creates difficulties in determining its prevalence causes for cardiovascular morbidity and mortality [1, 2]. According to the calculations made by WHO, while the prevalence of DM in 2000 was 171 million in the world, this number is expected to reach 366 million by 2030 [3-8]. Approximately 97% of these patients fall into the classification of Type-2 DM [9, 10]. 30% of Type-2 DM patients are also diagnosed with hypertension [11, 12]. It makes DM and hypertension interrelated diseases. Particularly in Turkey, 63% of the population is pre-hypertensive, and 25% of deaths are caused by hypertension [13]. Both diseases are a common and significant reason for both morbidity and mortality, and require high costs both in diagnosis and treatment procedures; therefore, it is vital to address these two diseases specifically.

The absence of physiologic night-time blood pressure decrease is called a non-dipping pattern, which is associated with a poorer cardiovascular prognosis. A non-dipping blood pressure pattern is accepted as hypertensive target organ damage and is proven to increase cardiovascular morbidity

Table 1. Comparison of related studies

References	Methodology	Disease of interested	Instances	Attributes	Reduction
[14]	ANN, Logistic Regression	Left Ventricular Hypertrophy	101 cases 21<age< 85	19	yes
[15]	Genetic Algorithms (GA), Weighted k-Nearest Neighbors (WkNN)	Type-2 DM (T2DM), coronary heart disease (CHD)	352 cases	18	no
[16]	Multilayer Perceptron (MLP)	Hypertension, Coronary artery disease, Rheumatic valvular heart disease, Chronic cor pulmonale, Congenital heart disease	352 cases Ages: 86, 82, 71, 60, 53	40	no
[17]	Fuzzy Weighted Preprocessing, and Artificial Immune Recognition System (AIRS)	The heart disease, hepatitis disease	270 cases	19	yes
[18]	Fuzzy Weighted Preprocessing, and Artificial Immune Recognition System (AIRS)	The heart disease, hepatitis disease	270 cases	19	yes
[19]	Principal Component Analysis (PCA), k-NN Based Weighting Pre-processing and Artificial Immune Recognition System (AIRS)	Atherosclerosis	114 cases:68 male, 46 female 20<age<69	61	yes
[20]	PCA-ANFIS Hybrid Learning Algorithms	Hearth valve disease	215 cases:132 male, 83 female 15 <age< 80	12	yes
[21]	ANFIS Hybrid Learning Algorithms	Heart valve disease	215 cases;132 male, 83 female 15 <age< 80	91	yes
[22]	Fuzzy k-NN Algorithms	Valvular heart diseases	215 cases:132 male, 83 female 15 <age< 80	251	yes
[23]	Fuzzy k-NN Algorithms	Heart valve disease	215 cases: 132 male, 83 female 15 <age< 80	91	no
[24]	PCA-ANFIS	Diabetes	215 cases 21<age	8	yes
[25]	PCA LS-SVM	ECG Arrhythmia	215 cases	279	yes
[26]	Neuro Fuzzy Method	Thyroid	215 cases	5	no
[27]	Fuzzy Expert System	Prostate Cancer	200 cases 43<age<76	1PSA	no
[28]	Complementary Learning Fuzzy Neural Network	Breast Cancer	78cases 27<age<90	more than 5	no
[29]	Fuzzy	Coronary Artery Disease	199 cases	19	no
[30]	Proposed System in the initial phase	Dipper/non-Dipper	65 cases 44<age<58	47	yes

both in the diabetic and non-diabetic patients [12]. That means that patients with a non- dipping pattern have a higher cardiovascular risk. Identifying non-Dipper pattern requires the measurement and recording of blood pressure, a diagnostic procedure called Ambulatory Blood Pressure Monitoring (ABPM) for 24 hours. Although ABPM is a non-invasive and simple procedure, handling a device for 24 hours, including bedtime may pose several difficulties such as at least two visits to the outpatient clinic, physical discomfort, psychological anxiety and distress of carrying device, high cost for medical insurance companies and also the state, insufficient number of devices and trained staff.

Studies conducted using artificial intelligence (AI) techniques in the medical field in diabetes, cancer, and cardiovascular diseases are widespread as listed in Table-1. These studies are compared by the following parameters: Methodology, a disease of interest (DoI), properties (case number, gender, and age) of the instances, the number of attributes, data reduction [14-30]. An important point is whether the number of attributes is reduced or not. This is presented in the last two columns of Table 1. It was seen that there is no study to estimate Dipper/non-Dipper pattern classification by using data mining algorithms.

Additionally, many studies [14, 17-22, 24, 25, 31-34] show that the use of most relevant attributes gives a higher classification accuracy percentage than the using of all attributes. For this reason, similar studies aimed to increase the accuracy percentage by using the most relevant attributes obtained through attribute selection algorithm outcomes and by reducing the least relevant ones. Unlike these studies, this research aimed to achieve reasonably high accuracy percentage close to the previous percentages rates by reducing the most relevant attributes under the guidance and knowledge of the physician. Therefore, the proposed study started with 94 patients as a pilot project. The initial results of the study were published in [30].

II. Materials and Methods

Classification is the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known [35]. Examples of some classification algorithms used in similar works are linear classifiers (Fisher's linear discriminant, Logistic regression, NaiveBayes classifier, Perceptron), Decision Trees (J48), Artificial Neural Networks (MLP, RBF) [14-16, 31-34], so on.

This study has two parts: Attribute selection and classification. First of all, the importance of attribute rankings were calculated for 6 different cases through running various attribute selection algorithms in WEKA platform. Then, four different classification algorithms were run by excluding ABPM data. For each case, classification accuracy, specificity, and sensitivity were calculated and compared.

A. Groups

The study population initially consisted of 94 Turkish patients in Marmara University Hypertension and Atherosclerosis Center and School of Medicine Department of Internal Medicine, Division of Endocrinology. Patients who were over 18 years old and who did not use antihypertensive medicine and also patients with Type-2 DM and normotensive were chosen from the initial group. Thus, the number of patients who participated in this study decreased from 94 to 65. Classification of blood pressure and diagnosis of Dipper/non-Dipper pattern was based on USA Joint National Committee VII guideline. Patients with acute coroner syndrome and cardiac insufficiency diagnosis or history were excluded from the research. Lastly, patients with missing data were also not included in the study group. Hence, this research was conducted with 56 patients in total - 22 males and 34 females- with a mean age of 51±7.5. All the necessary clinical assessment and laboratory data of each patient were completed within four weeks. Thus, a forming database for all participants lasted 2 years, from 2010 to 2012.

B. Attributes and Preprocess

The patients' data were prepared to be processed by WEKA Explorer [36]. Estimation models can be created for both nominal and numeric attributes. The total number of attributes including ABPM data was originally 47 as shown in Table 2.

All attributes types were numeric data. Only four (Gender, OrH, EF, mask HT) were nominal. Their types were converted into binary ones by using supervised filter NominaltoBinary.

C. Separation Methods of Test and Train set

The focused groups must be separated into two segments as train and test sets. One of the separation methods is to divide by using default percentage values (66 % train set, 34 % test set for WEKA) [37]. Another method is to use Cross-Validation (CV). CV is a statistical method for evaluating and comparing learning algorithms by dividing data into two segments: the first one is used to learn or train a model and the second one is used to validate the model. In typical cross-validation, the training and validation sets must cross-over in successive rounds so that each data point can have a chance of being cross-validated. The basic form of cross-validation is k-fold cross-validation [38-40]. CV is widely accepted in data mining and machine learning community and serves as a standard procedure for performance estimation and model selection [37]. The consensus in the data mining community seems to be that k = 10 is a good compromise. This value of k is particularly attractive because it helps estimations using 90% of the data, thus making it more likely to generalize the full data [38-41].

CV method was used to eliminate two problems arising from unbalanced data set and memorization of the model due to a limited number of instances (in data set) [41]. Because, as explained in section Experiment, the data set used in this study was already balanced. Therefore, CV method was used to address the aforementioned second problem.

In the process of determining the appropriate algorithms, two different evaluations have been conducted by frequently used learning algorithms in medical informatics. The first evaluation aimed to emphasize the importance of ABPM data among all attributes whereas the second evaluation was performed to compare the algorithms' success in the classification of Dipper/non-Dipper patterns. These evaluations are presented respectively in the next sections.

Table 2. Original attributes

Demographic data	Laboratory data-1	Autonomic Tests	Laboratory data-2			Total number of Attributes
			ECG	***ECHO***	***ABPM***	
Gender	HbA1C	V1	S V1	SWth	dt sis	
Age	Fasting glucose	V2	R V5	PWth	dt dias	
Height	Creatinin	Vm		LVED D	nt sis	
Weight	Microalbumi nuria	DSD		LVES D	nt dias	
Waist circumference	LDL cholesterol	AKH		ME	24h sis	
Body Surface Area	Ankle-brachi al index	OrH		MA	24h dias	
Body Mass Index		HDKBF		MDT	mask HT	
		PH		AoIVR T	%sisD	
		SH		LVM	%diasD	
		Ewing		LVMI		
				EF		
				LA		
				Ao		
7	6	10	2	13	9	**47**

D. Significance of ABPM data

The novelty of the study is to design a decision support system (DSS) in order to estimate Dipper/non-Dipper pattern by excluding ABPM data. It is proven that the clinicians have to use ABPM during the Dipper/non-Dipper diagnosis [12]. For this reason, listed classification algorithms shown in Table 3 were run one by one, and the results supported the idea that ABPM plays a vital role in medical prognosis [14, 42].

The decrease in the accuracy percentages for each algorithm as listed in Table 3 shows the significance of ABPM data in the classification of Dipper/non-Dipper pattern. This was supported by the results in Table 3. For instance, in the experiment conducted without ABPM data, J48 algorithm calculated the highest accuracy percentage (Table 3: 67.86%).

Now, the next section explains the various attribute selection algorithms.

E. Attribute Selection and Reduction

Although there may be a difference in the field of application, the increase in the number of attributes in a classification problem leads a decrease both in real time execution and classification performance [33, 43]. Various attribute selection algorithms were run on WEKA Explorer Platform with the aim of emphasizing the importance of ABPM data in classification and reducing the data that are not included in ABPM. The attribute ranking evaluation for six different cases is presented in Fig. 1. Two different selection modes were used for each algorithm: 'a' stands for the full training set, and 'b' is for CV mode. The methods are explained as follows:

m1: Ranker search method has been used from the Classifier type.

m2, m3, m4: Respectively Bestfirst, Linear Forward selection, greedy stepwise search method has been used from CfsSubsetEval type.

m5: Ranker Search Method has been used from Gain Attribute Eval type.

m6: Ranker Search Method has been used from Correlation Attribute Eval type.

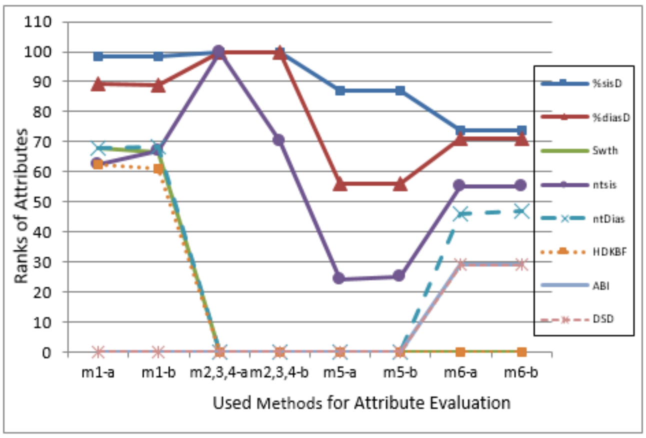

Figure 1. Attributes Ranking for Various Evaluation Algorithms

Ranker was identified as a common method for m1, m5, and m6 cases. However, in attribute evaluation algorithm, respectively Classifier, CfsSubsetEval and Correlation Attribute Eval were selected.

The Fig. 1 shows that the attributes with highest ranks are common for all methods. The first three data (%sisD, %diasD, and ntsis) are also ABPM data. It is observed that for all attributes, except for ntsis, selection of "full training set" or "cross validation" has no influence on attribute ranking. Since the other attributes' ranks are 0 or close to 0, they are not included in Fig. 1 [33].

Table 3. Accuracy percentages of different algorithms

Algorithms	***Accuracy (%)***			
	including ABPM data (number of attributes: 47)		***excluding ABPM data (number of attributes: 38)***	
	10-fold CV	***%66 Train Set Percentage Split***	***10-fold CV***	***%66 Train Set Percentage Split***
J48	95.31	100.00	67.86	36.84
NaiveBayes	81.25	72.73	57.14	47.37
MLP	76.56	72.73	51.79	36.84
RBF	71.88	77.27	44.64	57.89

The attributes interrelated with each other were reduced to only one attribute in order to find the relation between the most distanced attributes. These results were also filtered by a physician. Thus, the number of attributes was reduced from 47 to 6 [44].

III. Experiments

J48, NaiveBayes, MLP, and RBF algorithms were run one by one in Classifier Interface. RBF, which has the highest accuracy for the prediction of Dipper/non-Dipper pattern in k=10 fold CV mode. The evaluation also includes the actual and predicted classification results of instances and output entropy evaluation measures.

A. J48 algorithm

J48 Algorithm is a C4.5 implementation in Decision Tree. It is an iterative algorithm that splits the subjects where the information gain is the greatest [30, 46-48]. The output is based on IF-THEN rule and membership function sets [48, 49]. The tree is built in a top-down approach by splitting the subjects. It starts with the selection process of the best variable in the root of the tree [48]. J48 is able to cut the poor or non-meaningful branches into an efficient pruning process [50]. An entropy calculation for the best root of DSD is explained according to the values in Table 4:

Step1: 37 patients (66 %) were selected as a train set out of 56 patients. 18 of them had Dipper pattern while 19 had a non-Dipper pattern. The data set to be modelled is balanced.

Entropy of train set was calculated in order to select the attribute for classification. The system's entropy was calculated by using Eq. 1.

$$\text{Entropy\_before} = -p(D)*\log2(p(D)) - p(N)*\log2(p(N)) \quad (1)$$
$$= -0{,}4865*(-1{,}0395) - (0{,}5135)*(-0{,}9616)$$
$$= 0{,}9995 \text{ bits}$$

Here;

p(D) = 18/37 = 0.4865
p(N) = 19/37 = 0.5135
Total number of Instance (class object) =D+N=18+19 = 37

Step 2: Class Object (N: non-Dipper, D: Dipper as seen in Table 5) was written under each corresponding DSD column.

Step 3: No change was observed in class object type after DSD took the value of 18. If DSD >18 then the class object is Dipper; If DSD <=18 then a further attribute selection is employed to determine whether it is D or N.

Step 4: A single attribute value was calculated for best classifier by using gain information as shown in Eq. 2.

$$\text{Entropy\_left}=-(11/30)*\log2(11/30)-(19/30)*\log2(19/30)$$
$$= 0.9480$$
$$\text{Entropy\_right} = -(7/7)*\log2(7/7) - (0/7)*\log2(0/7)$$
$$= 0$$
$$\text{Entropy\_after}=(30/37)*\text{Entropy\_left}+(7/37)*\text{Entropy\_right}$$
$$\text{Entropy\_after} = 0.8108 * 0.9480+0.1892 * 0 = 0.7680 \text{ bits}$$
$$\text{Information Gain (DSD)} = \text{Entropy\_before} - \text{Entropy\_after} \quad (2)$$
$$= 0.9995 - 0.7680 = 0.2315 \text{ bits}$$

B. NaiveBayes Algorithms

NaïveBayes Algorithms are based on Bayes probability theorem in Statistical methods. It is a both predictive and descriptive classifier method and analyzes the relationship between the target and dependent/ independent variables to derive a conditional probability on each relation. In this regard, the probability with naïve approach can be calculated by using Eq. 3 [51].

$$P(V|C_i) = \prod_{j=1}^{n} P(V_j|C_i) \quad (3)$$

NaiveBayes uses a different methodology than J48 and NN algorithms. It was useful in terms of providing variety in the performance comparison. However, it did not show higher performance metrics in Dipper/non-Dipper pattern classification.

C. NN (MLP, RBF) Learning algorithms

MLP and RBF are the most popular algorithms in Neural networks (NN) [52]. In NN, a number of neurons in the input layer is defined by the numbers of attributes. Each class corresponds a neuron on the output layer. The most important thing in NN is to define the number of hidden layers and the number of neurons in that layer. That is a complex question to answer. In the literature, the number of hidden layers is usually chosen somewhere between the number of input and output layers. In hidden layer configuration, just two following rules must be considered: (i) the number of hidden layers is equal to one, and it is sufficient for the majority of problems and (ii) the number of neurons in that layer is the mean of the neurons in the input and output layers. According to the WEKA, the default number of neurons in the hidden layer is also the mean value of input and output neurons as a common heuristic approach [47, 53]. In the experiments, the number of the neurons in the hidden layer is set to 4 as seen in Fig.2.

Table 4. An example of an entropy calculation for the best root of DSD.

	Values of DSD for different patients																					
	3	***4***	***5***	***6***	***7***	***8***	***9***	***9.5***	***11***	***12***	***13***	***14***	***15***	***17***	***18***	***19***	***19.6***	***20***	***22***	***26***	***31***	***41***
Class Object	D	D	N	D	N	N	D	N	N	N	D	D	D	D	N	D	D	D	D	D	D	D
		N	N	N	N	N	N		N	N	N	N										
					N	D			D	D												
										N												

The model has three layers as seen in Fig.2. Each node in a previous layer is fully connected to a node in the following layer. In this network, the flow of information moves forward from input to output thus it is called “feed-forward network”. MLP uses a variety of learning techniques. The most popular one is back-propagation. As training and learning algorithms, a back propagation algorithm is preferred [54]. Here, the output values are compared with the actual outcomes and the error between them is reflected to the weights. In back propagation algorithms, training starts with a random set of weights. The success of network in many applications depends on the appropriate selection of initial weights. Gradient descent, which is a standard mathematical optimization algorithm to find MLP weights, reaches the best results. Each weight in the network is found by calculating the derivatives of squared error related to each parameter.

RBF NN has the advantages of adaptive and self-learning ability [47]. It is superior to MLP as it calculates the first cluster of parameters independently from the second cluster of parameters and produces high-performance classifiers, as seen in this study [52].

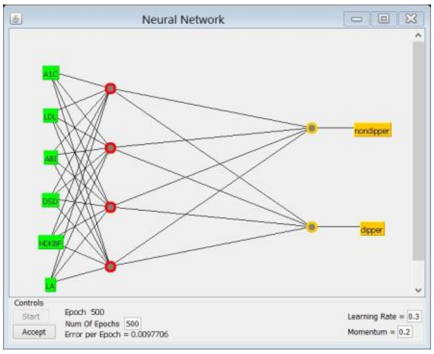

Figure 2. NN Models

IV. Results

In this study, a decision support system (DSS) which can estimate dipping/non-dipping pattern without using ABPM data was developed. By the opinion of the medical expert, the number of attributes was reduced from 47 to 6.

When foreseeing dipping/non-dipping pattern through 24-hour blood pressure data, the algorithms in Table 3 provided accuracy reaching to 100%. This ratio indicates the importance of ABPM device. If the comfort of the patient and the total costs are concerned, it becomes clear that this study is important in estimating non-dipping pattern without using this device.

Performance evaluation parameters were calculated by using Eq. 4, 5, 6.

$$\text{Accuracy} = ((TP + TN) / (P+N)) \quad (4)$$

$$\text{Sensitivity} = ((TP) / (TP+FN)) \quad (5)$$

$$\text{Specificity} = ((TN) / (TN+FP)) \quad (6)$$

Here;

True Positive (TP): A case is detected with non-Dipping pattern both by expert clinicians and DSS.

True Negative (TN): A case is detected with Dipper pattern both by expert clinicians and DSS.

False Positive (FP): A case is detected with a non-Dipping pattern by DSS whereas it was previously labeled as a Dipper pattern by expert clinicians.

False Negative (FN): A case is detected with Dipping pattern by DSS whereas it was previously labeled as a non-Dipper pattern by expert clinicians.

In Table 5, an example is given for calculation of RBF classifier’s performance while selection mode is set to 10-fold CV.

Table 5. Confusion matrix and calculation example for RBF in table 7 (k=10 fold cv)

a	**b**	**<-- classified as**
17	7	a = non-Dipper
3	29	b = Dipper
% Accuracy = ((17+29)/(17+7+3+29))x100 = 82,142		
% Sensitivity = (17/(17+7))x100 =70,833		
% Specificity = (29/(29+3))x100 =90,625		

As it can be seen in Table 6, when the selection mode was set to k=10 fold CV, the highest accuracy, and specificity percentage was achieved by RBF while the highest sensitivity percentage was achieved by using MLP. When the other selection mode, Percentage split, was set; all percentages were the highest and equal for both MLP and RBF. For both selection modes, J48 gave the sensitivity percentages below the acceptable limit. Also, ROC values, which had the closest values to 1, were achieved by RBF for each selection mode. ROC was 0.872 for 10 fold CV mode and 0.856 for percentage split mode.

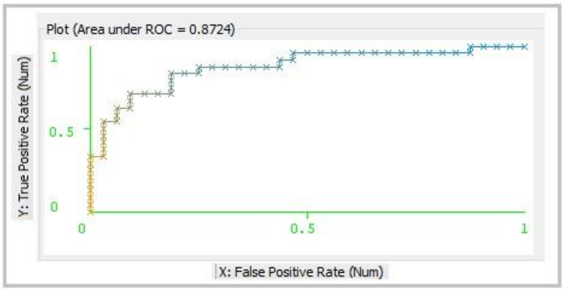

Figure 3. ROC of RBF Model (Selection mode is set to k=10-fold CV)

ROC curve is a graphical plot that illustrates the performance of the classifier. The curve is plotted with the true positive rate against the false positive rate. ROC analysis provides an optimal model selection while discarding suboptimal ones independently from the class distribution for diagnostic decision making. An example graph showing ROC for Table 6 is given in Fig. 3.

V. Discussion

This study has validated the hypothesis that a learning model with over 80% accuracy in Dipper/non-Dipper diagnosis without ABPM data can be developed. Many studies [12, 14, 40] emphasize the importance of ABPM in the diagnosis and mention the challenges involved in this process. The use of proposed model removes these challenges referred to in the introduction.

This model skips some of the steps in the process of diagnosis, thus, it simplifies and accelerates the process and reduces the costs. Moreover, the life quality of the patient is maintained. This pilot project suggests that the model can be applied to a larger group of people.

During the model development, performances of J48, NaiveBayes, and RBF and MLP methods are compared. It could be seen that NN algorithms display a higher classification metrics. On the other hand, in data mining, there is no single best method. Since it is an experimental science, the best method is identified by the best scores on the focused problem.

It was observed that in the rankings of selection algorithms, ABPM data showed a high-rank value. It was also identified that DSD attribute takes place in every model trial. This study also supported the finding that the data obtained through autonomic tests was significant in the classification of Dipper/non-Dipper pattern.

In medical diagnosis/prognosis, the damages of false classification of FN were greater than the false classification of FP. Therefore, it is better to have a higher sensitivity performance rather than a higher specificity [56]. Achieving a sensitivity value over 70% indicates that the model is successful since 70% is defined as an acceptable success limit by medical expert [44].

Further research can focus on developing a model that can be used in the classification of Dipper/non-Dipper without the data of autonomic dysfunction and orthostatic hypotension which requires physical activity. Another model can also be developed to predict Dipper / non-Dipper pattern for 29 patients with non-diabetic and hypertensive features (who are excluded in this study) by following the same steps explained in this study. The models developed for each two groups with opposite characteristics will, then, be compared. Finally, the number of patients can be increased in future research to emphasize the robustness of the model.

VI. CONCLUSION

In this study, decision trees (DT) were preferred due to its exhibition of intuitive approach and closeness to human decision-making mechanism. It emerged that DTs gave the highest percentage of success until the number of attributes were reduced to 13 and the accuracy of classification is 78%. When the class pattern of a new patient is not known as dipper or non-dipper, the found model of J48 can be applied on the patient. They were also able to predict the right pattern with an accuracy of 78 % for a new patient [45]. The DTs algorithms is more successful than the ANN algorithms to predict the class of new patients. However, this method ignored the branches with low gains; it may block the new model formation that could produce higher gains in later steps. Also, Random Forest (RF) that is a kind of algorithm for Decision Trees suggested by Breiman. The algorithm purpose that is during the classification process it is to raise the value of using multiple classification decision trees. However, the classification results of the RF is not enough high value as an expected in Table 6.

Table 6. Compared to NN algorithms, J48 and performance evaluation of classification algorithms

Algorithms		**6 Attributes *, K=10 Fold CV**				**6 Attributes *, Percentage Split is 66%**			
		Accuracy (%)	***Sensitivity (%)***	***Specificity (%)***	***ROC (AUC)***	***Accuracy (%)***	***Sensitivity (%)***	***Specificity (%)***	***ROC (AUC)***
DTs	J48	75,00	66,67	81,25	0,743	57,89	30,00	88,89	0,733
	Random Forest	69,64	66,67	71,88	0,781	68,42	70	66,67	0,844
Bayes	NaiveBayes	69,64	70,83	68,75	0,757	63,16	80,00	44,44	0,756
ANN	MLP	76,79	83,33	71,88	0,771	78,95	80,00	77,78	0,789
	RBF	82,14	70,83	90,63	0,872	78,95	80,00	77,78	0,856

*** HbA1C, LDL, ABI, DSD, HDKBF, LA**

NaïveBayes, a statistical analysis method, could create a model giving approximately 78% accuracy by using a relatively few number of patients and reduced attributes. If the number of patients is (input vector number) increased, the size of the matrix will also increase. So, the algorithm is expected to produce better results [54, 55].

Finally, ANN MLP and RBF algorithms were used for classification, and it was observed that RBF algorithm had the highest success rate in terms of sensitivity that was 83.3%. In medical diagnosis, a higher sensitivity performance is regarded as a more valid indication of metric than a higher specificity [56].

Moreover, compared to NN, DTs and NaiveBayes algorithms showed higher accuracy, sensitivity and specificity when more attributes were used for classification. However, as the number of attributes was reduced, the success of these algorithms also decreased. For instance, J48 gave the sensitivity percentages below the acceptable limit of 70% [57].

To minimize the memorization ability of the model as much as possible due to the usage of the small data set, Cross Validation was employed for all above-mentioned algorithms.

Acknowledgment

This study has been supported by Scientific Research Project of Marmara University (Project No. FEN-C-DRP-110908-0225).

The local ethics committee approved the study (approval no: MAR-YÇ-2009-0166, date 08.05.2009 and 90.09.2011.0108).

References

[1] Türker H. (MD.) Comparison of Insulin Resistance Hypertensive and Normotensive Patients with Type-2 Diabetes. PhD. Dissertation, Fatih Sultan Mehmet Training and Research Hospital, Department of Internal Medicine, Istanbul. 2006

[2] Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projection for 2030. Diabetes Care. 2004; 27: 1047-1053.

[3] Satman İ. Epidemiology Study. Population-based study of diabetes and related risk characteristics of Turkey. Results of the Turkish Diabetes Care. 2002; 25: 1551–1556.

[4] King H, Aubert RE, Herman WH. Global burden of diabetes. 1995–2025: prevalence, numerical estimates, and projections. Diabetes Care. 1998; 21: 1414–1431.

[5] King H, Rewers M. Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults: the WHO AdHoc Diabetes Reporting Group. Diabetes Care. 1993; 16: 157–177.

[6] Satman İ. Epidemiology of Diabetes Mellitus. All Aspects of Diabetes Mellitus (2nd ed.), Nobel Tıp Press, Istanbul, 2001; p.69-84. ISBN: 9754200521

[7] The Expert Committee: Report on the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2003; 26(1):5-19.

[8] Ertürk T. (MD.) The effect on diabetic foot scar of the learning levels of diabetic patient and nurse. Dissertation, Istanbul Training and Research Hospital, General Surgery Clinic, Istanbul. 2005.

[9] Bozkurt BO. (MD.) Tracing of overt renal nephropathy indicators of Type-2 diabetic patients of prognosis at the Haseki training and research hospital internal diseases nephrology polyclinic. Dissertation, Haseki Training and Research Hospital, Department of Family Medicine, Istanbul. 2007

[10] Amos AF, McVarty DJ, Zimmot P. The rising global burden of diabetes and its complications: estimates and projections to the year 2010. Diabetic Medicine. 1997; 14: 7–55.

[11] Leitão CB, Canani LH, Kramer CK, Moehlecke M, Pinto LC, Ricardo ED, Gross JL. Blood pressure means rather than nocturnal dipping pattern are related to complications in Type 2 diabetic patients. Diabetic Medicine. 2008; 25(3): 308-313.

[12] Bilgin D. (MD.) Masked with and without hypertension, hypertensive organ damage in patients with Type-2 Diabetes Mellitus. Dissertation, Marmara University School of Medicine, Istanbul. 2010

[13] Turgan C. The study of hypertension prevalence in Turkey.

http://www.turkhipertansiyon.org/kongre2009/pdf/14_mayis/Cetin_Turgan.pdf Accessed 29.4.2011

[14] Filho VD, Chaves Hilton de C Jr, Valença MJS, Campello de Souza FMA. Constructive Neural Network for Detecting Left Venticular Hypertrophy. Proceedings of the Fifth International Conference on Hybrid Intelligent Systems, IEEE. 2005.

[15] Giardiana M, Azuaje F, McCullagh P, HarperR. A Supervised Learning Approach to Predicting Coronary Heart Disease Complications in Type-2 Diabetes Mellitus Patients. Sixth Symposium on Bioinformatics and Bioengineering, IEEE.2006.

[16] Zheng J, Jiang Y, Yan H. Committee Machines with Ensembles of Multilayer Perceptron for the Support of Diagnosis of Heart Diseases. International Conference on Communications, Circuits and Systems Proceedings, Guilin. 2006; 2046 – 2050.

[17] Polat K, Güneş S. A hybrid approach to medical decision support systems: Combining feature selection, fuzzy weighted pre-processing and AIRS. Computer Methods and Programs in Biomedicine. 2007; 88(2):164-174.

[18] Polat K, Şahan S, Güneş S. Automatic detection of heart disease using an artificial immune recognition system with fuzzy resource allocation mechanism and k-nn (nearest neighbour) based weighting preprocessing. Expert Systems with Applications. 2007; 32(2):625-631.

[19] Latifoğlu F, Polat K, Kara S, Güneş S. Medical diagnosis of atherosclerosis from Carotid Artery Doppler Signals using principal component analysis, k-NN based weighting pre-processing and Artificial Immune Recognition System. Journal of Biomedical Informatics. 2008; 41(1):15-23.

[20] Avcı E, Turkoglu İ. An Intelligent Diagnosis System Based On Principle Component Analysis and ANFIS for the Heart Valve Diseases. Expert Systems with Applications. 2009; 36:2873–2878.

[21] Şengür A. An Expert System Based on Linear Discriminant Analysis and Adaptive Neuro-Fuzzy Inference System to Diagnosis Heart Valve Diseases. Expert Systems with Applications. 2008; 35(1-2): 214-222.

[22] Şengür A. An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnosis of valvular heart diseases. Computers in Biology and Medicine. 2008; 38 (3):329-338. Top 25 hottest article.

[23] Şengür A, Türkoğlu İ. A Hybrid Method Based on Artificial Immune System and Fuzzy k-NN Algorithm for Diagnosis of Heart Valve Diseases. Expert systems with applications. 2008; 35(3): 1011-1020.

[24] Polat K, Güneş S. An Expert System Approach Based on Principal Component Analysis and Adaptive Neuro-Fuzzy Inference System to Diagnosis of Diabetes Disease. Digital Signal Processing. 2007; 17(4): 702-710.

[25] Polat, K, Güneş S. Detection Of ECG Arrhythmia Using A Differential Expert System Approach Based On Principal Component Analysis And Least Square Support Vector Machine. Applied Mathematics and Computation. 2007; 186: 898-906.

[26] Keleş A, Keleş A. ESTDD: Expert system for thyroid diseases diagnosis. Expert Systems with Applications: An International Journal. 2008; 34: 1.

[27] Castanho MP, Barros LC, Yamakami A, Vendite LL. Fuzzy expert system: An example in prostate cancer. Applied Mathematics and Computation, 2008; 202(1): 78-85.

[28] Tan TZ, Quek C, Ng GS, Ng EYK. A novel cognitive interpretation of breast cancer thermography with complementary learning fuzzy neural memory structure. Expert Systems with Applications: An International Journal. 2007; 33: 3.

[29] Tsipouras MG, Exarchos TP, Fotiadis DI, Kotsia AP, Vakalis KV, Naka KK, Michalis LK. Automated Diagnosis of Coronary Artery Disease Based on Data Mining and Fuzzy Modeling. IEEE Transactions on Information Technology in Biomedicine. 2008; 12(4): 447-458.

[30] Altikardes ZA, Erdal H, Baba A, Tezcan H, Fak AS, Korkmaz H. A Study to Classify Dipper/non-Dipper Blood Pressure Pattern of Type-2 Diabetes Mellitus Patients without Holter Device. World Congress on Computer and Information Systems International Conference on Hammamet, Tunisia. IEEE. 2014. doi: 10.1109/WCCAIS.2014.6916555

[31] Salama MA, El-Bendary N, Revett K, Fahmy AA. Interval-based attribute evaluation algorithm. In Computer Science and Information Systems, Federated Conference on IEEE. 2011; 153-156

[32] Kannan SS, John SM, Sundaravadivel P, Ilangovan S, Kumar AVA. Performance Evaluation of Classification Algorithms C4. 5 and SMO on Microarray Gene-set Data. International Journal of Research and Technology. 2013; 2:5

[33] Ren W, Hu L, Zhao K, Chu J, Jia B. Intrusion Classifier based on Multiple Attribute Selection Algorithms. Journal of Computers. 2013; 8(10): 2536-2543

[34] Parveen AN, Inbarani HH, Kumar ES Performance analysis of unsupervised feature selection methods. In Computing, Communication and Applications, 2012 International Conference on IEEE. 2012; 1-7. doi:10.1109/ICCCA.2012.6179181

[35] Alpaydin E Introduction to Machine Learning. MIT Press. 2010; 9. ISBN 978-0-262-01243-0

[36] David SK, Saeb AT, Al Rubeaan K. Comparative Analysis of Data Mining Tools and Classification Techniques using WEKA in Medical Bioinformatics. Computer Engineering and Intelligent Systems. 2013; 4(13), 28-38.

[37] Gupta R. Lecture No: 3. AU-KBC Research Centre, MIT Campus, Anna University. 2008

[38] Refaeilzadeh P, Tang L, Liu H. Cross-validation. In Encyclopedia of database systems, Springer, US. 2009; 532-538

[39] Bramer M. Principles of data mining. Springer, US. 2013

[40] Gupta R, Mittal A, Singh K. A time-series-based feature extraction approach for prediction of protein structural class. EURASIP Journal on Bioinformatics and Systems Biology. 2008; 2.

[41] Wan, X., Liu, J., Cheung, W. K., & Tong, T. (2014). Learning to improve medical decision making from imbalanced data without a priori cost. BMC medical informatics and decision making, 14(1), 111.

[42] Magdás A, Szilágyi L, Belényi B, Incze A. Ambulatory monitoring derived blood pressure variability and cardiovascular risk factors in elderly hypertensive patients. Bio-medical materials and engineering. 2014; 24(6), 2563-2569.

[43] Nan, Liu, Koh Zhi Xiong, Goh Junyang, Lin Zhiping, Haaland Benjamin, Ting Boon Ping, et al. 2014. "Prediction of adverse cardiac events in emergency department patients with chest pain using machine learning for variable selection." BMC medical informatics and decision making 14 (1): 75. doi: 10.1186/1472-6947-14-75.

[44] Fak AS (MD.). 2009-2012 Interview. Marmara University, Istanbul, Turkey.

[45] Altikardes ZA. Expert system design to predict 24-hour ambulatory blood pressure features in Type-2 Diabetes Mellitus patients. Dissertation, Marmara University. 2012

[46] Doğan YA. Data Mining Based Target Classification for Tactical Underwater Sensor Networks. Dissertation, Turkish Naval Academy, Naval Sciences and Engineering Institute, Department of Computer Engineering, Istanbul. 2004.

[47] Witten IH, Frank E. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, New Zealand. 2005.

[48] Mollazade K, Ahmadi H, Omid M, Alimardani R. An intelligent combined method based on power spectral density, decision trees and fuzzy logic for hydraulic

pumps fault diagnosis. International Journal of Intelligent Systems and Technologies. 2008; 3(4):251-263

[49] Zhou Z, Jiang Y. Medical diagnosis with C4.5 rule preceded by artificial neural network ensemble. IEEE Transactions on Information Technology in Biomedicine. 2003; 1: 37-42

[50] Sitar-Taut DA, Sitar-Taut AV Overview on how data mining tools may support cardiovascular prediction. Journal of Applied Computer Science & Mathematics. 2010; 8 (4).

[51] Hudairy H. Data mining and decision making support in the governmental sector. Dissertation, Louisville University; 2004.

[52] Jia W, Zhao D, Shen T, Su C, Hu C, Zhao Y. A new optimized GA-RBF neural network algorithm. Computational intelligence and neuroscience. 2014.

[53] Heaton J. Introduction to Neural Networks for Java. 2nd edn, Heaton Research, Inc. 2005; ISBN: 1604390085

[54] Türkoğlu F. Author Recognition in Turkish Documents with Hybrid Approach. Dissertation, Yıldız Technical University. 2006

[55] Gaga L, Moustakis V, Vlachakis Y, Charissis G. ID+: Enhancing Medical Knowledge Acquisition with Machine Learning. Applied Artificial Intelligence: An International Journal. 1996; 10:79- 94.

[56] Mena L, Gonzalez JA, Maestre G. Extracting new patterns for cardiovascular disease prognosis. Expert Systems. 2009; 26(5):364-377.

[57] Altikardes ZA, Erdal H, Baba A, Fak AS, Korkmaz H. Performance evaluation of classification algorithms by excluding the most relevant attributes for dipper/non-dipper pattern estimation in Type-2 DM patients. 15th The International Conference on Intelligent Systems Design and Applications, ISDA 2015, Marrakesh, Morocco. IEEE Catalog Number: CFP15394-ART, ISBN: 978-1-4673-8712-5

Author Biographies

Zehra Aysun Altikardes received the B.E. degree from Marmara University Technical Education Faculty, Department of Electronics Education in 1993. She received MSc. and PhD. degrees in Electronics & Computer Education in Marmara University, Institute for Graduate Studies in Pure and Applied Sciences, Istanbul, Turkey in 2001, and 2012, respectively. She is currently the Head of Computer Technologies Department in Vocational School of Technical Sciences in Marmara University. Her research fields include medical informatics, artificial intelligence, distance education, and web programming.

Hasan Erdal received B.Sc. degrees from Marmara University Technical Education Faculty, Department of Electrical Education, Istanbul, and the M.Sc. and the Ph.D. degrees in Electrical Education from Marmara University, Institute for Graduate Studies in Pure and Applied Sciences, in 1990 and 1998 respectively. He has been Associate Professor in the Electrical and Electronics Engineering Department, since July 2010. His research interests focus on automatic control applications, electrical machines, control education, control algorithms and intelligent control.

Ahmet Fevzi Baba received B.Sc. degree from Marmara University Technical Education Faculty, Department of Electrical & Electronics Education in 1986. He received M.Sc. and the Ph.D. degrees in Electrical & Electronics Education from Marmara University, Institute for Graduate Studies in Pure and Applied Sciences, in 1989 and 1995 respectively. He has been the Head of Electrical and Electronics Engineering Department, Faculty of Technology in Marmara University from 2011. His research interests focus on automatic control applications, control algorithms and intelligent control.

Ali Serdar Fak received the MD degree from Istanbul Faculty of Medicine, Istanbul University, in 1987. He is currently a professor of the Department of Internal Medicine, and the Director of Hypertension and Atherosclerosis Center in Marmara University.

Hayriye Korkmaz received the B.E. degree in Electronics Education, the M.S. and the Ph.D. degree in Electronics and Communication Education from Marmara University, Istanbul, Turkey in 1993, 1995, and 2002, respectively. She is currently an Associate Professor with the Faculty of Technology of Electrical and Electronics Engineering, Marmara University. Her research interests include PC based data acquisition for industrial and biomedical applications, sensors and transducers, graphical development platforms such as LabVIEW, and embedded programming.