Medical Disease Prediction Using Artificial Neural Networks

Dimitrios H. Mantzaris, George C. Anastassopoulos and Dimitrios K. Lymberopoulos, *Member IEEE*

***Abstract*–This study examines a variety of Artificial Neural Network (ANN) models in terms of their classification efficiency in an orthopedic disease, namely osteoporosis. Osteoporosis risk prediction may be viewed as a pattern classification problem, based on a set of clinical parameters. Multi-Layer Perceptrons (MLPs) and Probabilistic Neural Networks (PNNs) were used in order to face the osteoporosis risk factor prediction. This approach is the first computational intelligence technique based on ANNs for osteoporosis risk study on Greek population.**

MLPs and PNNs are both feed-forward networks; however, their modus operandi is different. Various MPL architectures were examined after modifying the number of nodes in the hidden layer, the transfer functions and the learning algorithms. Moreover, PNNs were implemented with spread values ranging from 0.1 to 50, and 4 or 2 neurons in output layer, according to coding of osteoporosis desired outcome.

The obtained results lead to the conclusion that the PNNs outperform to MLPs, thus they are proved as appropriate computation intelligence technique for osteoporosis risk factor prediction. Furthermore, the overfitting problem was more frequent to MLPs, contrary to PNNs as their spread value increased.

The aim of proposed PNN is to assist specialists in osteoporosis prediction, avoiding unnecessary further testing with bone densitometry.

I. Introduction

Artificial Neural Networks (ANNs) are subfield of Artificial Intelligence (AI) systems. Their ability to correlate input and corresponding output data [1] – [4], based on vector mapping, has established themselves as a powerful tool in various applications. ANNs have been applied in various medical fields, constituting themselves as a useful technique in clinical practice [5] – [7], such as cardiology [8], oncology [9], pathology, endocrinology [10], radiology [11], urology [12] – [15], pneumonology [16], pediatrics [12], [17] and pediatric surgery [17], [18]. Medicine is a field that ANNs can be proven as a powerful tool to enhance current medical techniques [17] – [21].

This study focus on the development and assessment of ANN pattern recognition models based on both Multi-Layer Perceptrons (MLPs), as well as Probabilistic Neural Networks (PNNs) and the application of these models to the problem of osteoporosis risk factor prediction. More specifically a decision support tool has been developed to help clinicians identify which people are at increased risk for osteoporosis and should therefore undergo further testing with bone densitometry. This application area is considered as extremely important since early detection of osteoporosis is vital for the prevention of osteoporotic fractures, which are associated with increased morbidity and mortality and high socio-economic costs. The proposed ANN architectures and their performance in clinical data are presented in this paper.

Manuscript received July 5, 2008.

D. H. Mantzaris is with the Medical Informatics Laboratory, Democritus University of Thrace, GR-68100, Alexandroupolis, Greece (e-mail: dmantzar@med.duth.gr)

G. C. Anastassopoulos is with the Medical Informatics Laboratory, Democritus University of Thrace, GR-68100, Alexandroupolis, Greece, and Hellenic Open University, GR-26222, Patras, Greece (corresponding author to provide phone: +30-2551030503; fax: +30-2551030503; e-mail: anasta@med.duth.gr).

D. K. Lymperopoulos is with the Electrical and Computer Engineering Department, University of Patras, GR-26504, Patras, Greece (e-mail: dlympero@wcl.ee.upatras.gr).

II. Materials and Methods

A. Osteoporosis

Osteoporosis is the prevailing bone's disease, and its features are low bone density mass and the modification of their micro-architecture structure, so that bones' tolerance is reduced and the risk of fracture is increased. Apart from the direct physical implications of a fracture, such as pain and inconvenience, osteoporotic fractures involving the hip or the spine are a major cause of morbidity and mortality. Studies show that two out of five people over seventy-five who fracture a hip will die within a year as a direct result. There is an enormous public health problem with huge recourses required to deal with the immediate and long-term effects of fractures like hospitalization, loss of independence, support at home or in institutions etc. In the European Union one person breaks a bone because of osteoporosis every fifteen seconds.

Often the first apparent symptom of osteoporosis is a broken bone, which is why the condition is also known as "the silent crippler", as people do not realize they have osteoporosis until it's too late. However early detection and treatment of osteoporosis can decrease the fracture risk of a person to a minimum. For these reasons, there are studies [6], [7] where NNs were used for predicting whether a person has osteoporosis or not.

Osteoporosis is presented after the age of 50 years and its frequency increases in proportion with the age. It is most common for women than men. The ordinary type of osteoporosis arises after menopausal. A percentage of 75% of women with osteoporosis, doesn't known their trouble.

The main factor for osteoporosis growth is the high osteal density mass loss between 45 to 50 years of a person. The osteal absorption is greater than osteal production, specific for women elder than 50 years. Thereby, the osteal density mass loss is prospective.

The osteoporosis diagnosis, both a priori without symptomatic findings or in case of a bone fracture, is based on laboratory and osteal bone densitometry examination. This examination is applied to specific bones, using Dual Energy X-ray absorptiometry. This technique is based on radiation absorption from the patient, so it is not recommended for entire orthopedic cases.

The osteoporosis data, which were used at the design of ANN models, were obtained from the Orthopedic Clinical Information System of Alexandroupolis' University Hospital, Greece. For each case, there were 4 clinical parameters that have been considered. These parameters were: sex, age, height and weight. The estimation of osteoporosis risk factor was based on T-score value, which is the patient's bone density compared with the normally expected in a healthy young adult of the same sex.

The present study is based on data set consisted of 3426 cases. This data set was divided into a set of 2426 records and another set of 1000 records. The former was used for training of MLPs and the construction of PNNs, whereas the latter for performance testing of neural networks.

B. Neural Network Models for Osteoporosis Risk Prediction

The proposed pattern recognition models for osteoporosis risk factor classification are based on a non-symbolic computational intelligence method implemented by ANN [12]. The development of such an ANN demands the determination of a number of parameters, such as the type of ANN, the number of neurons in each layer and the applied learning algorithm.

MLPs are feed-forward networks with back-propagation learning rule and are used in majority of ANN models [22], [23]. The correlation of dependent and independent variables constitutes an important feature for MLPs, so they can be used in medical data processing.

A MLP consists of an input layer, where the number of input nodes equals to the number of variables of the problem, and an output layer with a number of nodes defined by the problem's requirements [2]. The hidden layers and a large number of characteristics of MPL are varied.

Although MLPs have been used successfully in a wide range of medical applications, they are faced with suspicious by many researcher [24], [25]. The reason of this confrontation proceeds by the heuristic ("black-box") feature of MLPs, as they can detect hidden correlations into data. In contrast to the "black-box" feature of MLPs, PNNs which approximate Bayesian statistical technique, combine new input vectors with existed storage data in order to classify correctly the input data; a process familiar to human behavior [26].

PNNs are based on Parzen's Probabilistic Density Function (PDF) estimator [27] and their aim is the correct classification of input vectors to one of the available target classes of the problem. A PNN is a three-layer feed-forward network, consisting of an input layer, a radial basis and a competitive layer. The radial basis layer computes distances from the input vector to the training input vectors and produces a vector whose elements indicate how close the input is to a training input. The third layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a competitive transfer function on the output of the third layer classifies the input vector into a specific class because that class has the maximum probability of being correct.

The number of input nodes of PNN equals to the number of variables of the problem and the number of nodes for output layer equals to the number of classes, as they are defined by the problem. The number of nodes for hidden layer is the number of patterns during the PNN's construction.

The PNNs do not require iterative learning process, so that may managed magnitude of data faster that MLP neural networks. This PNNs' feature results by the Bayesian technique's behavior.

The osteoporosis risk factor prediction is based on four variables; consequently, in this study, the input layer of implemented ANN models consists of 4 neurons. The neurons' number of output layer is defined by both the desired number of problem's variables and the type of ANN. Specifically, the MLP demands a neuron in output layer for estimation of osteoporosis risk factor's stages, as the MLP's result is an integer value of 1 to 4, whereas the PNN architecture uses so many neurons as the number of osteoporosis' stages, that constitute the number of classes of input data.

The development of ANNs demands the data preprocessing. For this study, the sex variable was coded as 1 for female and 2 for male persons, whereas age, height and weight (measured in years, cms and kgs correspondingly) were obtained as recorded in the database. The input parameters of ANNs, as well as their coding are presented in Table I. The 1st column of the Table I corresponds the ANNs' inputs with osteoporosis variables that are presented in the 2nd column. The 3rd column depicts the coding of each variable. The parentheses report the physical correspondence for each variable coding. The values of variables age, height and weight are numerical that correspond to age, height and weight of each patient.

TABLE I.
CODING OF OSTEOPOROSIS CLINICAL PARAMETERS.

NN inputs	Variables	Coding
1	Age	*Numeric value (years)*
2	Sex	*1 (Woman) 2 (Man)*
3	Height	*Numeric value (cm)*
4	Weight	*Numeric value (kg)*

The output variable of ANNs was the T-score parameter. These values of bone densitometry were classified into 4 stages, as presented in Table II, so their coding is a number according to T-score result.

The used transfer functions of the MLP structure, were two, nominally hyperbolic tangent sigmoid for hidden layer and linear for output layer [28], [29]. The radial basis and competitive transfer functions were applied for hidden and output layer of PNN, correspondingly [28], [29]. Mathematical equations of these transfer functions are depicted in Table III.

TABLE II.
CODING OF T-SCORE VALUES

T-score Value	Coding
≤ - 2.5	1
-2.5 – -1.5	2
-1.5 – 0	3
≥0	4

The determination of numbers of neurons for MLP's hidden layer was achieved by trial and error. The Levenberg-Marquardt back-propagation learning algorithm was selected for MLPs' training, as it is a robust algorithm, appropriate for non-linear least-squared problems [28].

TABLE III.
TRANSFER FUNCTIONS

Transfer Function	Mathematical Equation
Competitive (compet)	Calculates a layer's maximum output from its net input
Hyperbolic tangent sigmoid (tansig)	$f(x) = \frac{2}{1+e^{-2x}} - 1$
Linear (purelin)	$f(x) = x$
Radial basis (radbas)	$f(x) = e^{-x^2}$

The structure of PNNs has only one hidden layer, contrary to MLPs, where the number of hidden layers is not completely defined. Moreover, the number of neurons for PNN's hidden layer depends by the number of patterns during PNN's construction. Consequently, the proposed PNN had 2426 neurons for hidden layer, as the available data set for PNN implementation, consisted of 2426 cases. PNNs' design is straightforward and does not depend on training, thus no learning algorithm was selected during PNN's implementation [28].

The mean squared error (MSE) [28] was used as evaluation criterion of performance of MLPs which mathematical notation is

$$MSE = \frac{1}{N}\sum_{k=1}^{N} e(k)^2 = \frac{1}{N}\sum_{k=1}^{N} \left[t(k) - a(k)\right]^2 \qquad (1)$$

where N is the number of patterns, t(k), a(k) and e(k) are the desired, the MLP's calculated and the error value for pattern k, respectively.

As mentioned above, the neurons' number of input and output layers is defined by the problem. It was clarified in section II that input parameters are 4 and output parameter is one; consequently, in this study, the input layer consists of 4 neurons and the output layer has one neuron that determines the patients' osteoporosis risk factor. The definition of neurons in hidden layer was achieved by a computational process that modifies the number of these neurons and calculates the performance for each of ANN topologies.

The PNNs architecture is constrained by the available features of specific problem, however, the width of the calculated Gaussian curve for each probability density function have to be defined. In the present study, this spread factor varied from 0.1 to 50.

III. RESULTS

The development and performance assessment of ANN models were based on MATLAB Neural Network Toolbox, due to its effectiveness and user-friendly interface [28].

The results of implemented MLPs are summarized in Table IV. The 2nd and 3rd columns describe the MLP architectures and the transfer functions of each ANN's model, correspondingly. The desired MSE for each MLP structure is represented in the 4th column. The percentages of successful prognosis over testing, training and overall data set are presented in 5th, 6thand 7th columns, correspondingly. In other words, the 5th, 6th and 7th columns represent the correct classified cases over 1000, 2426 and 3425 cases correspondingly. The 8th, 9th and 10th columns depict the percentages of successful prognosis over pathological cases for testing, training and overall data set, respectively. The values in parenthesis exhibit the real number of cases that were detected correctly by implemented MLPs.

The implemented MLPs underwent to further processing in terms of their predicting abilities. As mentioned in section II, the possible stages of osteoporosis examination are four, whereof three are referred in pathological situations and one is referred for patients without osteoporosis. Consequently, the osteoporosis risk factor's stages classified into two groups, one for pathological cases and another group for persons without osteoporosis. These results were encoded to 0 for normal cases and 1 for pathological cases, thus it was attempted a binary coding of the desired results, in order to be used by artificial neural models.

The results of MLPs' simulating phase for two stages coding are summarized in Table IV. In particular, the columns 11th to 13th depicts the percentage of cases that were classified correctly, whereas 14th to 16th columns presents the pathological cases that were detected true positive cases.

During the implemented phase, the initial weights and biases of MLP neural networks were varied, keeping the other parameters unchangeable. In particular, Nr. 1 and Nr. 2 neural network models have the same architecture, but their initial conditions were different. Similar configurations were applied to Nr. 4 – 7 and Nr. 8 – 10 in order to train and construct ANNs for osteoporosis risk factor prediction. It is obvious that different initial conditions for MLPs training imply variation of neural networks' performance.

The obtained results on Table IV present that as the

number of neurons in hidden layer is increased, the improvement of MLPs' performance is achieved. This behavior is prospective, as the more neurons in hidden layer, the more weights and bias, so the MLP's ability is improved to store acquired knowledge. The performance's comparison of MLPs' 4-3-1 with 4-5-1 topology proves this statement. However, there is a maximum number of neurons in hidden layer, which its overstepping implies the decrement of MLP's performance. The limitation of number of hidden neurons results from overfitting problem, as the neural network has memorized the training patterns, but it has not learnt to generalize to new data [12]. The neural networks from Nr. 2 and Nr. 11 have difficulty in recognizing normal patterns and sorting them in appropriate classification area, as they have not the ability to learn from normal input data.

It is obvious that MLPs' performance with 2-stages osteoporosis coding has been improved, contrary to MLPs that had to classify cases into four stages. For example, Nr. 8 MLP with 2-stages osteoporosis coding outperforms to this with 4-stages. The first MLP had not the ability to discriminate patients without osteoporosis, classifying all records as pathological cases; however, the later MLP classified correctly more cases.

ANNs present insufficient ability to distinguish the different desired outcomes because of inadequate training process. In this study, an important feature for implemented MLPs is the number of normal cases that were sorted correctly. The performance of Nr. 6 MLP with 4-stages osteoporosis coding is satisfactory, however, only a small number of the normal cases distinguished rightly. Specifically, the correct classified normal cases for testing set is calculated by subtraction of real numbers of the 8th by the 5th column. The obtained result is 1, so that only one normal case was distinguished correctly. The rest normal cases were misclassified as pathological (false positive). Similarly, 4 and 5 normal cases classified correctly for training and overall data set, correspondingly.

This study investigates the implementation of neural network models in order to incorporate in medical decision support systems. The aim of a MLP with 4-5-1 topology, and the utilization of competitive transfer function in hidden layer, is the combination of types, topologies and transfer functions of neural networks in order to construct more efficient neural networks. Nevertheless, the aforementioned MLP presents diminished performance contrary to other MLP's structures. Moreover, it is clear that this MLP have overfitting problem as its ability to distinguish new normal and pathological patterns is inexistent.

TABLE IV
EXPERIMENTAL RESULTS USING MLP

Nr	Architecture of Artificial Neural Network	Transfer Functions	Goal MSE	4-Stages Osteoporosis Coding						2-Stages Osteoporosis Coding					
				Percentage of Successful Prognosis			Percentage of Successful Prognosis Over Pathological Situations			Percentage of Successful Prognosis			Percentage of Successful Prognosis Over Pathological Situations		
				Testing Set	Training Set	Overall Set	Testing Set	Training Set	Overall Set	Testing Set	Training Set	Overall Set	Testing Set	Training Set	Overall Set
1	4–3–1	tansig purelin	0.5	29.30 (293)	28.98 (703)	29.07 (996)	34.40 (291)	34.31 (701)	34.33 (992)	83.70 (837)	83.80 (2033)	83.77 (2870)	97.75 (827)	98.43 (2011)	98.23 (2838)
2	4–3–1	tansig purelin	0.5	29.40 (294)	31.12 (755)	30.62 (1049)	34.75 (294)	36.96 (755)	36.31 (1049)	84.90 (849)	83.88 (2035)	84.18 (2884)	99.76 (844)	99.31 (2029)	99.45 (2873)
3	4–3–1	tansig purelin	0.9	31.00 (310)	31.20 (781)	31.84 (1091)	36.64 (310)	38.23 (781)	37.77 (1091)	19.30 (193)	19.21 (466)	19.24 (659)	4.61 (39)	4.16 (85)	4.29 (124)
4	4–5–1	tansig purelin	0.5	31.10 (311)	31.74 (770)	31.55 (1081)	36.76 (311)	37.69 (770)	37.42 (1081)	67.60 (676)	70.20 (1703)	69.44 (2379)	77.66 (657)	81.60 (1667)	80.44 (2324)
5	4–5–1	tansig purelin	0.5	30.70 (307)	30.87 (749)	30.82 (1056)	36.29 (307)	36.56 (747)	36.48 (1054)	81.60 (816)	81.41 (1975)	81.47 (2791)	92.91 (786)	91.48 (1869)	91.90 (2655)
6	4–5–1	tansig purelin	0.5	31.90 (319)	32.40 (786)	32.25 (1105)	37.59 (318)	38.28 (782)	38.08 (1100)	74.20 (742)	73.45 (1782)	73.67 (2524)	84.63 (716)	84.39 (1724)	84.46 (2440)
7	4–5–1	tansig purelin	0.5	32.30 (323)	33.35 (800)	33.04 (1132)	37.59 (318)	38.96 (796)	38.56 (1114)	84.90 (849)	84.30 (2045)	84.47 (2894)	100.00 (846)	99.95 (2042)	99.97 (2888)
8	4–5–1	tansig purelin	0.9	30.70 (307)	30.87 (749)	30.82 (1056)	36.29 (307)	36.56 (747)	36.48 (1054)	81.60 (816)	81.41 (1975)	81.47 (2791)	92.91 (786)	91.48 (1869)	91.90 (2655)
9	4–5–1	tansig purelin	0.9	32.30 (323)	33.35 (809)	33.04 (1132)	37.59 (318)	38.96 (796)	38.56 (1114)	74.20 (742)	73.45 (1782)	73.67 (2524)	84.63 (716)	84.39 (1724)	84.46 (2440)
10	4–5–1	tansig purelin	0.9	31.60 (316)	32.03 (777)	31.90 (1093)	37.00 (313)	37.59 (768)	37.42 (1081)	35.80 (358)	34.00 (825)	34.53 (1183)	31.44 (266)	29.56 (604)	30.11 (870)
11	4–5–1	compet purelin	0.5	29.30 (293)	28.24 (685)	28.55 (978)	34.63 (293)	33.53 (685)	33.85 (978)	68.10 (681)	65.66 (1593)	66.37 (2274)	70.80 (599)	67.99 (1389)	68.81 (1988)
12	4–7–1	tansig purelin	0.5	31.60 (316)	32.03 (777)	31.90 (1093)	37.00 (313)	37.59 (768)	37.42 (1081)	41.60 (416)	45.18 (1096)	44.13 (1512)	44.92 (380)	48.56 (992)	47.49 (1372)

The Table V presents the obtained results of PNNs. The radbas and compet were the transfer functions for hidden and output layers, correspondingly. Whereas the number of neurons for input and hidden layers of PNNs was constant, the number of output neurons was variable according to the coding of desired values. The spread of radial basis function, which is used in second layer, is the only parameter that can be modified. The values of spread for PNNs with the best performance are presented in the 2nd column of the Table V. Initially, the implementation of PNNs based on the 4-stages of osteoporosis risk factor prediction, so the output layer consisted of 4 neurons. The obtained results of 4-2426-4 results. The increase of spread's value implies the decrease of PNN's performance for estimation of normal and pathological cases for overall and training data set. Moreover, it is mentioned that the percentage of successful prognosis for overall and pathological cases of testing set increases as the spread's value increases. However, there is a value of spread that constitutes limit for the performance improvement for testing set. The limit for spread parameter of the implemented PNNs equals to 7.3, whereas values greater than aforementioned number involves the decrement of PNNs' predicting ability for testing set. The Nr. 1 and Nr. 2 PNNs have not sufficient generalization ability, contrary to

TABLE V
EXPERIMENTAL RESULTS USING PNN ARCHITECTURES

Nr	Spread of radbas	4 neurons for output layer						2 neurons for output layer					
		Percentage of Successful Prognosis			Percentage of Successful Prognosis Over Pathological Situations			Percentage of Successful Prognosis			Percentage of Successful Prognosis Over Pathological Situations		
		Testing Set	Training Set	Overall Set	Testing Set	Training Set	Overall Set	Testing Set	Training Set	Overall Set	Testing Set	Training Set	Overall Set
1	0.5	30.40 (304)	99.04 (2233)	74.05 (2537)	32.74 (277)	93.10 (1902)	75.42 (2179)	78.80 (788)	97.28 (2360)	91.89 (3148)	89.95 (761)	99.51 (2033)	96.71 (2794)
2	0.6	30.40 (304)	99.04 (2233)	74.05 (2537)	32.74 (277)	93.10 (1902)	75.42 (2179)	78.80 (788)	97.28 (2360)	91.89 (3148)	89.95 (761)	99.51 (2033)	96.71 (2794)
3	2.7	34.80 (348)	58.16 (1411)	51.34 (1759)	38.42 (325)	61.58 (1258)	54.79 (1583)	84.70 (847)	86.73 (2104)	86.14 (2951)	98.46 (833)	99.70 (2037)	99.34 (2870)
4	3.6	36.20 (362)	50.62 (1228)	46.41 (1590)	40.67 (344)	54.19 (1107)	50.22 (1451)	85.00 (850)	85.20 (2067)	85.14 (2917)	99.29 (840)	99.90 (2041)	99.72 (2881)
5	7.3	40.30 (403)	41.63 (1010)	41.24 (1413)	46.10 (390)	48.21 (985)	47.59 (1375)	84.80 (848)	84.46 (2049)	84.56 (2897)	99.88 (845)	100.00 (2043)	99.97 (2888)
6	13.1	38.20 (382)	39.74 (964)	39.29 (1346)	44.92 (380)	46.84 (957)	46.28 (1337)	84.70 (847)	84.21 (2043)	84.35 (2890)	100.00 (846)	100.00 (2043)	100.00 (2889)
7	19.2	37.50 (375)	38.13 (925)	37.95 (1300)	44.21 (374)	45.28 (925)	44.96 (1299)	84.60 (846)	84.21 (2043)	84.33 (2889)	100.00 (846)	100.00 (2043)	100.00 (2889)
8	40.0	32.10 (321)	32.28 (783)	32.22 (1104)	37.94 (321)	38.26 (783)	38.21 (1104)	84.60 (846)	84.21 (2043)	84.33 (2889)	100.00 (846)	100.00 (2043)	100.00 (2889)

PNN topology, after the simulating phase, underwent to similar processing with those of MLPs. Consequently, the 3rd, 4th and 5th columns exhibit the percentage of successful classification of patterns, whereas the 6th, 7th and 8th columns represent the performance of PNNs over pathological patterns for testing, training and overall data set, correspondingly.

Moreover, the desired stages of osteoporosis risk factor were classified into two groups, one for pathological cases and another group for persons without osteoporosis. In this case, the PNN topology was modified, leading the number of neurons for output layer to two. The corresponding results that were obtained by execution of 4-2426-2 PNNs topology are presented in 9th to 14th columns of Table V.

A small variation of radial basis spread does not affect the PNNs' performance, as it is shown by Nr. 1 and Nr. 2 PNNs in Table V. The difference between 0.5 and 0.6 is impalpable, so it does not occurred alteration of obtained Nr. 3 PNN, which classifies correct new patterns that present in the input layer of the neural network.

It is pointed out that Nr. 4 PNN outperforms Nr. 3 PNN, as concluded by obtained results. Nevertheless, it is important to compare the normal cases that were classified correctly, as calculated by the subtraction of pathological by entire cases for testing, training and overall data set. Thus, the successful categorized normal cases are 23, 153 and 176 for testing, training and overall data of Nr. 3 PNN with 4 neurons for output layer, whereas for Nr. 4 PNN corresponding values are 18, 121, and 139. Moreover, the Nr.3 PNN with 2 neurons for output layer classified 14, 67 and 81 normal cases contrary to 10, 26 and 36 normal cases of Nr. 4 PNN. It is obvious that Nr. 3 PNN has the ability to classify more normal cases than Nr. 4 PNN. Consequently, Nr. 3 PNN outperforms Nr. 4 PNN, in terms of generalization ability.

IV. CONCLUSION

ANNs, as a subfield of computational intelligence, are used widely in industrial and medical applications. Despite of the ANN's architectures, learning algorithms and transfer functions variety, the basic function of ANNs is the presence of an input data set, and the generation of corresponding outputs based on vector mapping.

In this paper, the possibility of applying artificial neural models in medical making decision, and in particular, the osteoporosis risk factor estimation has been examined, because it is an important medical problem for public health. Its frequency and the serious consequences for patients are the reasons for the vivid interesting for development of computational and accurate techniques which do not expose the patients in radiation.

The development of artificial neural techniques was based on MLPs with back-propagation algorithm, as well as PNNs, which are both feed-forward neural networks. The MLPs has been characterized as black box, because the internal connections are highly non-linear and not subject to the usual statistics. On the other hand, PNNs approximate Bayesian function; however, their output is clearly not a probability, as several steps are required to osteoporosis risk factor prediction.

As it was found, the PNNs outperformed the MLPs, in terms of the successful prognosis of cases. Therefore the proposed methodology unveiled the PNN artificial models' behavior contrary to MLPs artificial networks' behavior is much better, or in other words, the prognostic ability of PNNs is enhanced compared to MLPs categorization performance.

ACKNOWLEDGMENT

We are grateful to the orthopedics of the Orthopedic Clinic of the University Hospital of Alexandroupolis, Greece for their valuable contribution.

REFERENCES

[1] J. Dayhoff and J. DeLeo, "Artificial neural networks opening the black box," *Cancer Supplement*, vol. 91, No. 8, pp. 1615-1635, 2001.

[2] V. Piuri and F. Scotti, "Morphological classification of blood leucocytes by microscope images," in *2004 IEEE Int. Conf. Computational Intelligence for Measurement Systems and Applications*, pp. 103-108.

[3] S. J. Perantonis, N. Ampazis, S. Varoufakis, and G. Antoniou, "Constrained learning in neural networks: Application to stable factorization of 2-D polynomials," *Neural Processing Lett.*, vol. 7, pp. 5–14, 1998.

[4] D. S. Huang and S. D. Ma, "Linear and nonlinear feedforward neural network classifiers: A comprehensive understanding," *J. Intelligent Syst.*, vol. 9, pp. 1–38, 1999.

[5] W. G. Baxt, "Application of artificial neural networks to clinical medicine," *Lancet,* vol. 346, no. 113, pp. 5-8, 1995.

[6] J. S. Chiu, Y. C. Li, F. C. Yu, and Y. F. Wang, "Applying an artificial neural network to predict osteoporosis in the elderly," *Studies in Health Technology and Informatics,* vol. 124, pp. 609-614, 2006.

[7] G. Lemineur, R. Harba, N. Kilic, O. N. Ucan, O. Osman, and L. Benhamou, "Efficient estimation of osteoporosis using artificial neural networks," *In 33rd Annual Conf. IEEE Industrial Electronics Society (IECON)*, Taipei, Taiwan 5-8, 2007, pp. 3039-3044.

[8] R. Silipo and C. Marchesi, "Artificial neural networks for automatic ECG analysis", *IEEE Trans. Signal Processing*, vol. 46, pp. 1417-1425, 1998.

[9] A. Taktak, A. Fisher, and B. Damato, "Modelling survival after treatment of intraocular melanoma using artificial neural networks and Bayes theorem," *Phys. Med. Biol.*, vol. 49, pp. 87-98, 2004.

[10] G. Zhang and V. Berardi, "An investigation of neural networks in thyroid function diagnosis," *Health Care Management Science*, vol. 1, pp. 29–37, 1998.

[11] L. Lanzarini, M. V. Camacho, A. Badran, and I. D. G. Armando, "Images compression for medical diagnosis using neural networks," *J. Computer Science and Technology*, Vol.2, No. 1, 1999, pp. 78-80.

[12] D. H. Mantzaris, G. C. Anastassopoulos, A. D. Tsalkidis, and A. V. Adamopoulos, "Intelligent prediction of vesicoureteral reflux disease," *WSEAS Trans. Systems*, Issue 9, vol. 4, pp. 1440-1449, 2005

[13] D. Tasoulis, P. Spyridonos, N. Pavlidis, D. Cavouras, P. Ravazoula, G. Nikiforidis, et all, "Urinary bladder tumor grade diagnosis using on-line trained neural networks," in *Proc. Knowledge Based Intelligent Information Eng. Systems Conf.*, Heidelberg, 2003, pp. 199-206.

[14] M. Tanthanuch and S. Tanthanuch, "Prediction of upper urinary tract calculi using an artificial neural network," *J. Medical Association of Thailand*, vol. 87, No. 5, pp. 515-518, 2004.

[15] F. Dieterlea, S. Müller-Hagedorn, H. Liebich, and G. Gauglitza, "Urinary nucleosides as potential tumor markers evaluated by learning vector quantization," *Artificial Intelligence in Medicine*, vol. 28, issue 3, pp. 265–279, 2003.

[16] M. Munley, J. Lo, G. Sibley, G. Bentel, M. Anscher, and L. Marks, "A neural network to predict symptomatic lung injury," *J. Physics in Medicine and Biology*, vol. 44, pp. 2241–2249, 1999.

[17] D. Mantzaris, G. Anastassopoulos, A. Adamopouos, I. Stephanakis, K. Kambouri, and S. Gardikis, "Selective clinical estimation of childhood abdominal pain based on pruned artificial neural networks," in *Proc. 3rd WSEAS Int. Conf. on Cellular and Molecular Biology*, *Biophysics and Bioengineering*, Athens, 2007, pp. 50-55.

[18] D. Mantzaris, G. Anastassopoulos, A. Adamopouos, I. Stephanakis, K. Kambouri, and S. Gardikis, "Abdominal pain estimation in childhood based on artificial neural network classification," in *Proc. the 10th Int. Conf. on Engineering Applications of Neural Networks*, Thessaloniki, 2007, pp. 129-134.

[19] K. Papik, B. Molnar, R. Schaefer, Z. Dombovari, Z. Tulassay, and J. Feher, "Application of neural networks in medicine - a review," *Medical Scince Monitor*, vol. 4, no. 3, pp.538-546, 1998.

[20] M. Lundina, J. Lundina, H. Burked, S. Toikkanenb, L. Pylkkänenc, and H. Joensuua, "Artificial neural networks applied to survival prediction in breast cancer," *Oncology*, vol.57 no.4 pp.281-286, 1999.

[21] E. I. Mohamed, C. Maiolo, R. Linder, S. J. Pöppl, and A. De Lorenzo, "Artificial neural network analysis: a novel application for predicting site-specific bone mineral density," *Acta Diabetologica*, vol. 40, pp. 19-22, 2003.

[22] E. Tafeit and G. Reibnegger, "Review artificial neural networks in laboratory medicine and medical outcome prediction," *Clin. Chem. Lab. Med.*, vol. 37, no. 9, pp. 845–853, 1999.

[23] J. Shieh, S. Fan, and W. Shi, "The intelligent model of a patient using artificial neural networks for inhalational anaesthesia," *J. Chinese Institute of Chemical Engineers*, vol. 33, no. 6, pp. 609-620,2002.

[24] J. W. Gurney, "Neural networks at the crossroads: caution ahead," *J. Radiology*, vol. 193, no. 1, pp. 27-28, 1994.

[25] J. Wyatt, "Nervous about neural networks," *Lancet*, vol. 346, pp. 1175-1176, 1995.

[26] R. K. Orr, "Use of a probabilistic neural network to estimate the risk of mortality after cardiac surgery", *J. Medical Decision Making*, vol. 17, no. 2, pp. 178-185, 1997.

[27] E. Parzen, "On estimation of a probability density function and mode," *Annals of Mathematical Statistics*, vol. 33, no.3, 1962, pp. 1065–1076.

[28] D. Howard and B. Mark (2008), "Neural network toolbox User's Guide", *The Math Works, Inc.* [Online]. Availiable: http://www.mathworks.com/access/helpdesk/help/pdf_doc/nnet/nnet.pdf.

[29] D. Anderson and G. McNeil, "Artificial neural networks technology," New York: Kaman Science Corporation, 1992, ch. 5.