Diagnostic Feature Extraction on Osteoporosis Clinical Data Using Genetic Algorithms

George C. Anastassopoulos[1], Adam Adamopoulos[2], Georgios Drosos[3], Konstantinos Kazakos[3], and Harris Papadopoulos[4,5]

[1] Medical Informatics Laboratory, Medical School, Democritus University of Thrace, Greece
anasta@med.duth.gr
[2] Medical Physics Laboratory, Medical School, Democritus University of Thrace, Greece
adam@med.duth.gr
[3] Department of Orthopedics, University Hospital of Alexandroupolis, Medical School, Democritus University of Thrace, GR-68100,
drosos@otenet.gr, kazakosk@yahoo.gr
[4] Frederick Research Center, Filokyprou 7-9, Palouriotisa, Nicosia 1036, Cyprus
[5] Computer Science and Engineering Department, Frederick University, 7 Y. Frederickou St., Palouriotisa, Nicosia 1036, Cyprus
h.papadopoulos@frederick.ac.cy

Abstract. A medical database of *589* women thought to have osteoporosis has been analyzed. A hybrid algorithm consisting of Artificial Neural Networks and Genetic Algorithms was used for the assessment of osteoporosis. Osteoporosis is a common disease, especially in women, and a timely and accurate diagnosis is important for avoiding fractures. In this paper, the 33 initial osteoporosis risk factors are reduced to only 2 risk factors by the proposed hybrid algorithm. That leads to faster data analysis procedures and more accurate diagnostic results. The proposed method may be used as a screening tool that assists surgeons in making an osteoporosis diagnosis.

Keywords: osteoporosis, risk factor, artificial neural networks, genetic algorithm.

1 Introduction

Physicians use medical protocols in order to diagnose diseases. Medical diagnosis is the accurate decision of upon the nature of a patient's disease, the prediction of its likely evolution and the chances of recovery, based on a set of clinical and laboratorial data. When dealing with clinical data provided by a medical protocol that takes account of a large number of clinical features, one has to consider methods for feature selection. These methods could tentatively decrease the number of clinical features that are used for clinical assessment of the patients [1, 2]. Four are the main benefits of a successful feature selection: detection of clinical data redundancy, smaller clinical data sets, faster data analysis procedures, more accurate diagnostic

H. Papadopoulos et al. (Eds.): AIAI 2013, IFIP AICT 412, pp. 302–310, 2013.

results. As far as the first benefit, clinical features that do not play any significant role in patients' assessment could be detected as redundant and therefore could be ignored, or omitted thereafter. As a result, we step to the second benefit: the elimination of redundant clinical features that are not considered for the clinical assessment of the patients could lead to smaller clinical protocols that could incorporate fewer clinical parameters. Consequently, since the number of the clinical parameters that are considered is decreased, the time for data analysis and assessment is decreased in an analogous fashion. That is the third benefit we obtain. Finally, a robust and efficient feature extraction procedure, not only minimizes the size of the data set that is considered for analysis, but at the same time leads to more accurate diagnostic results, since it helps practitioners to take account of the diagnostic parameters that are essential for patients' assessment and prevent any confusion that could be caused from redundant data. At its best, a feature extraction method could lead to an optimum, by providing an optimal combination of the smallest set of clinical data that, when analyzed, could give the highest diagnostic results.

In the present paper, the method that was developed and applied for the detection of the optimal set of clinical osteoporosis data is a hybrid algorithm that incorporated two main algorithmic tools provided by the field of Computational Intelligence: Artificial Neural Networks (ANN) and Genetic Algorithms (GA) [3]. The main idea behind the proposed hybrid algorithm is based on two steps: First, to apply ANN for the osteoporosis data classification. This could provide an estimation of the classification error of the ANN when the whole clinical data set, with all of the clinical parameters that it incorporates, is considered as input for the ANN. Second, to apply a GA that could perform a heuristic search, in order to investigate for the optimally minimized clinical data set, that could be used as an input to the ANN, leading to even higher performance and even smaller classification error. Specifically, the GA is designed to accomplish a three-fold task: to find the optimally minimum set of clinical parameters that could be used as input to the ANN, to find the optimal ANN architecture by investigation for the minimum number of neurons in the input and the hidden layers of the ANN, and finally, to investigate for the combination of the input data set and ANN architecture that provides the optimal classification error, that is the minimum classification error possible.

When the GA investigation procedure comes to a successful end, the input parameters that are not included in the optimal (smallest) subset of the essential input parameters of the ANN, are considered as redundant and therefore can be omitted and eliminated during ANN training and testing.

The paper is structured as follows: In the next section the osteoporosis disease, as well as the risk factors that affect osteoporotic fractures are presented. In section 3 the used data and method are described. Section 4 details the results, while in section 5 a comparison with the other screening tools is made. Finally, section 6 gives our conclusions.

2 Osteoporosis

Osteoporosis is defined as a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue, with a consequent increase in

bone fragility and susceptibility to fracture [4]. Fractures related to osteoporosis usually affect the hip, spine, distal forearm and proximal humerus and account for a relatively high percentage of the total number of fractures [5, 6]. These fractures are associated with a high societal and personal cost and some of these fractures –hip and vertebra fractures- are also associated with increased mortality, morbidity or permanent disability [5 – 7].

2.1 Diagnosis of Osteoporosis

Osteoporosis is characterized by low Bone Mineral Density (BMD); this is the amount of bone mass per unit volume (volumetric density), or per unit area (areal density). BMD is measured by different techniques like Dual-Energy X-ray Absorptiometry (DEXA), Quantitative UltraSound (QUS), Quantitative Computed Tomography (QCT) [8 – 10]. DEXA of the proximal femur and lumbar spine (central DEXA) is the most widely used bone densitometric technique [8, 11] since there are many prospective studies that have documented a strong gradient of risk for fracture prediction [7].

According to the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis, the objectives of bone mineral measurements are to provide diagnostic criteria, prognostic information on the probability of future fractures, and a baseline on which to monitor the natural history of the treated or untreated patient [7].

2.2 Risk Factors for Low BMD

The identification of the high-risk subset of patients (women) is of great importance for the prevention of osteoporotic fractures, but routine BMD measurement of all women is not feasible for most populations [7].

Several osteoporosis risk assessment instruments have been proposed to identify women who are likely to have osteoporosis and should therefore undergo bone densitometry [12 – 20]. Nevertheless at present there is no universally accepted policy for population screening in Europe to identify patients with osteoporosis or those at high risk of fracture.

3 Patients and Methods

This is a prospective study including *589* women that underwent a measurement of bone mineral density (BMD) with Dual-Energy X-ray Absorptiometry (DEXA) of the lumber spine. Patients receiving treatment for osteoporosis were excluded. Also, race is not included as a parameter, since all the women were Caucasian. The data set consisted of *33* parameters, which are presented in Table 1.

One more parameter, the estimated T-score was used for the classification of each subject. The T-score data were preprocessed and expressed to integer values ranging from *0* (absence of osteoporosis) to *3* (severe osteoporosis).

For the specific problem described above, two steps were considered: At the first step, the data set was divided in two parts: 80% of the data were used for ANN training and the rest 20% was used for ANN testing. A large number of computer experiments were performed in this step in order to investigate the effectiveness of some ANN parameters on the obtained performance in terms of the classification error. The main aim of this step was to conclude the internal architecture of the ANN in terms of the number of hidden layers, the number of neural nodes per hidden layer and the type of the transfer function of neurons. The obtained classification error that was provided by this procedure was kept for the purpose of comparison with the corresponding results that were obtained by the GA pruned ANN in the second step of the method.

Table 1. Parameters of the data set

A/A	Osteoporosis risk factor
1	Occupation
2	Allergies
3	Age
4	Body weight
5	Height
6	Menarche
7	Menopause
8	Number of pregnancies
9	Smoking
10	Alcohol consumption
11	Coffee intake
12	History of fracture
13	Spinal fracture
14	Carpal fracture
15	Sports
16	Parents with osteoporosis
17	Loss of height more than 3 cm
18	Kyphosis
19	Amenorrhea (pause of menstruation) for more than 12 months
20	Rheumatoid arthritis
21	Dairy consumption
22	History of diarrhea
23	Cortisone intake
24	Thyroxin intake
25	Estrogen therapy
26	Anorexia nervosa
27	Hyper parathyroidism
28	Insulin depended diabetes
29	Ovariectomy
30	Paget disease
31	Steroids intake
32	Diuretics intake
33	Chemotherapy

In the second step a GA was invoked for the selection of input parameters and investigation of the internal architecture of the ANN. The GA was utilized to search for the optimal subset of the input parameters that should be used for training and testing the pruned ANN. The individuals of the GA population consisted of binary strings of length equal to the number of parameters of the original data set plus six more bits that represented the binary expression of the number of neurons in the hidden layer. Since we considered a total of $N = 33$ clinical parameters for the evaluation of the patients, the GA individuals' chromosome consisted of $33 + 6 = 39$ genes, with each gene to be represented by a binary digit (bit). For the first *33* genes of the chromosome, the allele *0* denotes that the corresponding clinical parameter is not included in the subset of the parameters that will be used for ANN training and testing and therefore is omitted. On the opposite, the allele *1* for the first *33* genes of the chromosome denotes that the corresponding parameter will be considered for ANN training and testing. The total number of the considered parameters was denoted by *I*. The last *6* genes of the chromosome were considered as the binary representation of the number *H* of neurons in the hidden layer. Starting counting from *1*, the *6*-bit binary strings could represent up to a number of *64* neurons in the hidden layer. Since binary representation is adopted for the chromosome of the individuals of the GA, all the well known genetic operators for selection, crossover and binary mutation could be applied on the GA population.

By the *33* input parameters, there were performed *10* individual computer experiments. For each experiment, the training and the testing set were constructed randomly (80% and 20% of the cases respectively). The mean value of these 10 experiments denoted by $<MSE_t>$ was kept for comparison with the obtained results of the GA pruned ANN.. This mean value is denoted as $<MSE_t>$.On the other hand, the application of the GA concluded to pruned ANN with performance that was related to the MSE that in this case is denoted as MSEp. For the evaluation of the individuals of the GA the fitness function that was used consisted of three terms, referring to the performance, the size of the input subset and the size of the hidden layer, that is, MSE_p over $<MSE_t>$, *I* over *N* and *H* over *64*, respectively. Thus, the fitness function can be written as:

$$f = \frac{MSE_p}{< MSE_t >} + \frac{I}{N} + \frac{H}{64} \tag{1}$$

As it is obvious in Eq. (1), the GA searches for the optimal combination of performance, input parameter data set size and number of hidden neurons. This is done by trying to minimise the fitness function of Eq. (1) by using the Matlab GA tool For each individual of the GA, an ANN is constructed, with the number *I* of input nodes that is indicated by the individual's chromosome first *33* genes and the number of neurons *H* that is denoted by the decimal representation of the last *6* genes of the chromosome. The input data subset is also constructed by considering the alleles *1* of the first *33* genes of the chromosome. Subsequently, the constructed ANNs are trained using the 80% of the cases and tested using the rest 20% of the cases, and the generated MSE is recorded.

4 Results

The proposed methodology was applied on the osteoporosis clinical data of *589* patients that were described in Section 3. In the first step of our methodology, ten individual computer experiments were performed with ANN that were trained and tested by the full clinical data set, that is, ANN with *33* input nodes in the input layer that correspond to the 33 clinical parameters of the medical protocol used to collect the data and *12* nodes in the hidden layer. This step resulted to a mean value of the MSE of these ten individual experiments which is $<MSE_t> = 3.75 \cdot 10^{-4}$. The results that were obtained by the computer experiments of the application of the GA are presented in Table 2, whereas a typical evolution of the fitness function in those experiments is shown in Fig. 1. Each row on Table 2 presents the obtained results of a specific computer experiment. The first column of Table 2 refers to the experiment number, the second column refers to the generation number, the third column refers to the MSE_p of the fittest individual (pruned ANN) of that generation, the fourth column refers to the number of input nodes *I* of the ANN of the fittest individual, the fifth column refers to the features used by that ANN (as numbered in Table 1), and finally, the sixth column refers to the number *H* of neurons in the hidden layer of the ANN of the fittest individual.

Table 2. Results of GA search and ANN pruning

#Exp.	#Gen.	MSEp	I	Features	H
1	1	$1.63 \cdot 10^{-4}$	15	1 3 7 8 10 11 13 15 17 20 23 24 28 30 32	2
1	17	$1.43 \cdot 10^{-4}$	5	2 6 7 10 25	2
1	29	$1.37 \cdot 10^{-4}$	3	3 7 10	1
1	49	$1.37 \cdot 10^{-6}$	2	7 10	1
2	1	$1.85 \cdot 10^{-6}$	12	3 7 8 10 14 15 16 18 23 25 28 32	6
2	19	$1.65 \cdot 10^{-6}$	2	7 10	1

The first four rows in Table 2 refer to experiment Nr. 1, presenting four instances of the GA evolution at four different generations. For each generation it is shown the performance of the best individual. Thus, the GA starts with *15* inputs in generation Nr. 1, to settle to just 2 inputs in generation Nr. 49. It is noteworthy that the number of hidden nodes (*H*) is relatively small, ranging from *2* for the 1st generation, to *1* for the 29th and 49th generation. The last two rows of Table 2 refer to the results of a second computer experiment. In that experiment the GA starts with *12* inputs at generation Nr. 1, to settle down to *2* inputs at generation Nr. 19. The results presented in the third column of Table 2 indicate that assuming the performance of the pruned ANN is also improved, in terms of the MSE, since MSE_p is decreased down to the *36.5%* of the original $<MSE_t>$ of the non-pruned ANN. According to the results presented in fourth column, the smallest number of inputs is just *2*, which is only the *6%* of the total number of diagnostic features, (which is $N = 33$). Even by using these small data subsets as input, and only one neuron in the hidden layer (as it is shown in the last column of Table 2), the pruned ANN maintain improved performance in terms of the MSE that they achieve, which is roughly the *1/3* of the $<MSE_t>$.

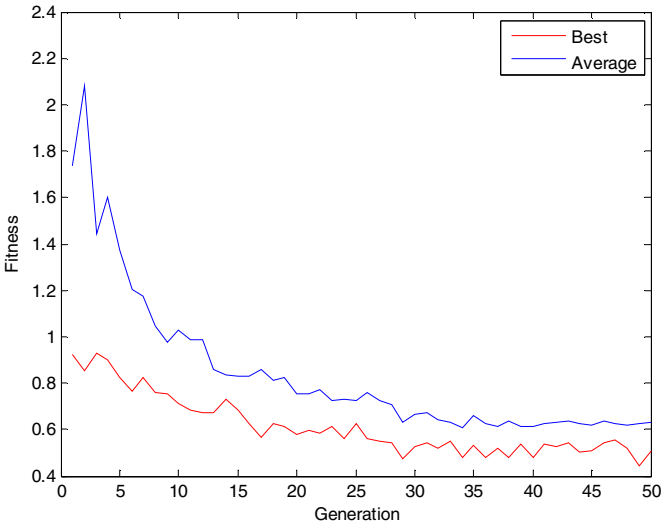

Fig. 1. GA evolution of mean and best individual fitness for 50 generations

5 Discussion

A common feature that the vast majority of the obtained results share is that they reveal a high order of redundancy in the original osteoporosis data set. The results presented in Table 2 indicated that just *2* to *3* input parameters, out of a total of *33*, that is a portion that varies from *6%* down to *9%* of the original data set, are essential in order of high performance to be achieved. In other words, the ANNs that were used for the classification of the osteoporosis data were pruned to less than *90%*. At the same time, the performance of these ANNs was highly improved, in terms of the MSE that is achieved, which for the pruned ANN is decreased almost to the *1/3* of the corresponding one that was achieved by the non-pruned ANN.

The two parameters that were common in all experiments is menopause (*7*) and alcohol intake (*10*), followed by a third one, the age (*3*). The osteoporosis risk assessment instruments mentioned earlier have been used to identify women that should undergo bone densitometry [12 - 17]. Apart from one instrument that uses only one parameter (body weight) [12] all other instruments include age and body weight, and some also include hormone replacement therapy [13 – 15]. History of fracture is included in two [13, 15] while one of them includes also history of rheumatoid arthritis and race [15].

In all but one osteoporosis risk assessment instruments, age is one of the most significant parameter. In our study, menopause, alcohol intake and age are the most common parameters. Menopause is related to age; actually it is related to a certain age group. Alcohol intake is not an important parameter in any of these instruments. It is surprising that alcohol intake is an important factor in our study since in Mediterranean countries alcohol intake is not common in women; unless alcohol intake is related to other dietary habits. One could expect that other parameters (like weight, previous fracture) to be more significant.

6 Conclusion

Simpler ANN architecture with decreased number of neural nodes and synaptic connections may result in less complicated and less time-consuming training and testing procedures and at the same time to performance improvement. Nowadays increased computer power and the contemporary development of ANN training algorithms, provide the essential means for fast and accurate implementation of ANN techniques to solve problems of various types, in different scientific fields, (classification, prediction, system identification, to name a few). Despite the fact that ANN training and testing even for complicated problems and large data sets is accomplished with low computational cost, it is legitimate, if not desirable, to investigate for even faster, more reliable and more accurate ANN training and testing methods. The present work was focused on the detection of any kind of redundancy in the data sets that are used as inputs during ANN training and testing. Redundancy and overlapping in data sets that are used for ANN training, when exists, definitely increase computational cost for ANN training, while at the same time, may mislead, or suppress the training procedure. This could generate problems in terms of ANN ability to accomplish successfully and with the desired accuracy the task that it was designed for.

From the clinical point of view, the results of our study are novel. Apart from menopause, age (well known parameters that are related to bone mineral density) and alcohol intake; we would expect some other parameters to be more common in our experiments. We will continue the study with more cases, in an effort to extract the most significant risk factors that affect the osteoporosis.

Acknowledgements. This work was supported by the European Regional Development Fund and the Cyprus Government through the Cyprus Research Promotion Foundation "DESMI 2009-2010", research contract TPE/ORIZO/0609(BIE)/24 ("Development of New Venn Prediction Methods for Osteoporosis Risk Assessment").

References

1. Papatheocharous, E., Papadopoulos, H., Andreou, A.S.: Feature Selection Techniques for Software Cost Modelling and Estimation: A Comparative Approach. Engineering Intelligent Systems 18(3-4), 233–246 (2010)
2. Papatheocharous, E., Papadopoulos, H., Andreou, A.S.: Software Effort Estimation with Ridge Regression and Evolutionary Attribute Selection. In: Proceedings of the 3rd Workshop on Artificial Intelligence Techniques in Software Engineering, AISEW 2010, CoRR abs/1012.5754 (2010)
3. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. Springer (1996)
4. Consensus Development Conference: Diagnosis, prophylaxis and treatment of osteoporosis. Am. J. Med. 94, 646–650 (1993)

5. Cooper, C., Atkinson, E.J., Jacobsen, S.J., O'Fallon, W.M., Melton, L.J.: A population based study of survival after osteoporotic fractures. Am. J. Epidemiol. 137, 1001–1005 (1993)
6. Johnell, O., Kanis, J.A.: An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos Int. 17, 1726–1733 (2006)
7. Kanis, J.A., Burlet, N., Cooper, C., Delmas, P.D., Reginster, J.Y., Borgstrom, F., Rizzoli, R.: European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). European Guidance for the Diagnosis and Management of Osteoporosis in Postmenopausal Women. Osteoporos Int. 19(4), 399–428 (2008)
8. World Health Organization: Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Technical Report Series 843. WHO, Geneva (1994)
9. Blake, G.M., Fogelman, I.: Role of dual-energy X-ray absorptiometry in the diagnosis and treatment of osteoporosis. J. Clin. Densitom. 10, 102–110 (2007)
10. Engelke, K., Gluer, C.C.: Quality and performance measures in bone densitometry. I. Errors and diagnosis. Osteoporos Int. 17, 1283–1292 (2006)
11. Marshall, D., Johnell, O., Wedel, H.: Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. Br. Med. J. 312, 1254–1259 (1996)
12. Michaëlsson, K., Bergström, R., Mallmin, H., Holmberg, L., Wolk, A., Ljunghall, S.: Screening for osteopenia and osteoporosis: selection by body composition. Osteoporos Int. 6(2), 120–126 (1996)
13. Lydick, E., Cook, K., Turpin, J., Melton, M., Stine, R., Byrnes, C.: Development and validation of a simple questionnaire to facilitate identification of women likely to have low bone density. Am. J. Manag. Care. 4(1), 37–48 (1998)
14. Cadarette, S.M., Jaglal, S.B., Kreiger, N., McIsaac, W.J., Darlington, G.A., Tu, J.V.: Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. CMAJ 162(9), 1289–1294 (2000)
15. Sedrine, W.B., Chevallier, T., Zegels, B., Kvasz, A., Micheletti, M.C., Gelas, B., Reginster, J.Y.: Development and assessment of the Osteoporosis Index of Risk (OSIRIS) to facilitate selection of women for bone densitometry. Gynecol. Endocrinol. 16(3), 245–250 (2002)
16. Koh, L.K., Sedrine, W.B., Torralba, T.P., Kung, A., Fujiwara, S., Chan, S.P., Huang, Q.R., Rajatanavin, R., Tsai, K.S., Park, H.M., Reginster, J.Y.: Osteoporosis Self-Assessment Tool for Asians (OSTA) Research Group. A Simple Tool to Identify Asian Women at Increased Risk of Osteoporosis. Osteoporos Int. 12(8), 699–705 (2001)
17. Geusens, P., Hochberg, M.C., van der Voort, D.J., Pols, H., van der Klift, M., Siris, E., Melton, M.E., Turpin, J., Byrnes, C., Ross, P.: Performance of risk indices for identifying low bone density in postmenopausal women. Mayo. Clin. Proc. 77(7), 629–637 (2002)
18. Weinstein, L., Ullery, B.: Identification of at-risk women for osteoporosis screening. Am. J. Obstet. Gynecol. 183(3), 547–549 (2000)
19. McLeod, K.M., Johnson, C.S.: Identifying Women with Low Bone Mass: A Systematic Review of Screening Tools. Geriatric Nursing 30(3), 164–173 (2009)
20. Anastassopoulos, G., Mantzaris, D., Iliadis, L., Kazakos, K., Papadopoulos, H.: Osteoporosis Risk Factor Estimation Using Artificial Neural Networks. Engineering Intelligent Systems 18(3/4), 205–211 (2010)