

# An Evolutionary Technique for Medical Diagnostic Risk Factors Selection

Dimitrios Mantzaris<sup>1</sup>, George Anastassopoulos<sup>1,2</sup>, Lazaros Iliadis<sup>2,3</sup>, Adam Adamopoulos<sup>2,4</sup>

<sup>1</sup> Medical Informatics Laboratory, Democritus University of Thrace, GR-68100, Alexandroupolis, Hellas

dmantzar@med.duth.gr    anasta@med.duth.gr

<sup>2</sup> Hellenic Open University, GR-26222, Patras, Greece

<sup>3</sup> Department of Forestry & Management of the Environment and Natural Resources, Democritus University of Thrace, GR-68200, Orestiada, Hellas

liliadis@fmenr.duth.gr

<sup>4</sup> Medical Physics Laboratory, Democritus University of Thrace, GR-68100, Alexandroupolis, Hellas

adam@med.duth.gr

**Abstract** This study proposes an Artificial Neural Network (ANN) and Genetic Algorithm model for diagnostic risk factors selection in medicine. A medical disease prediction may be viewed as a pattern classification problem based on a set of clinical and laboratory parameters. Probabilistic Neural Networks (PNNs) were used to face a medical disease prediction. Genetic Algorithm (GA) was used for pruning the PNN. The implemented GA searched for optimal subset of factors that fed the PNN to minimize the number of neurons in the ANN input layer and the Mean Square Error (MSE) of the trained ANN at the testing phase. Moreover, the available data was processed with Receiver Operating Characteristic (ROC) analysis to assess the contribution of each factor to medical diagnosis prediction. The obtained results of the proposed model are in accordance with the ROC analysis, so a number of diagnostic factors in patient's record can be omitted, without any loss in clinical assessment validity.

## 1 Introduction

Artificial Intelligence (AI) engineering is a relatively modern scientific field and has been reinforced by computer technology advancement. Artificial Neural Net-

works (ANNs), Genetic Algorithms (GAs) and Fuzzy Logic are non symbolic approaches of AI.

ANNs have been proved as a powerful tool for solving a variety of problems [1, 2] The problems' categories, where ANNs have been applied, are bioengineering [3], signal processing [4], environmental subjects [5, 6] and other fields.

While the results of medical statistic models are satisfactory, there are non-linear models that may contribute to the enhancement of medical decision support. In particular, ANNs have the ability to correlate input data to corresponding output data. Especially, ANNs have effectively contributed to disease diagnosis [7-14], like oncology [9], pediatrics [10], urology [10, 11], pediatric surgery [12], orthopedics [14], etc.

In an ANN design stage, the implementation of the best ANN architecture to solve a real-world problem is relatively complex. A neural network with few neurons implies inadequate lore, while a big one leads to poor generalization ability, presenting overfitting [15]. Early works for the investigation of appropriate ANN structure is achieved by trial and error method. However, in the last few years, more efficient methods for designing ANN architectures, automatically, have been developed.

This study presents a GA for the ANN pruning and the detection of the essential smallest input data of diagnostic factors for ANN training. The obtained ANN architecture uses diminished number of diagnostic factors and has an evolved structure, without any loss in terms of its performance and functionality ability.

The abdominal pain diagnosis except of the traditional methods (clinical, laboratory, imaging) could also be supported by numerically scoring systems, fuzzy logic techniques, etc [16]. The aim of implemented method was the detection of essential diagnostic factors for construction of PNNs to estimate the abdominal pain in childhood.

The obtained results were compared with the Receiver Operating Characteristic (ROC) analysis outcome. There was a high level of convergence between the two methods and some of the diagnostic factors have proven to be essential for clinical evaluation and prognosis, whereas, some other factors could be excluded during clinical estimation, without any loss in ANNs' prognosis accuracy. From a technical point of view, the detection of essential diagnostic factors gives the ability for the design of an ANN with simpler structure and improved performance, because ANN training is based on smaller data sets.

## **2 Artificial Neural Network's Pruning With Genetic Algorithm**

Both ANNs and GAs are inspired by biological processes. However, ANNs' learning is based on individuals (phenotypic learning), while GAs adapt a population to changing environment (genotypic learning). In recent years, there is a large

body of literature in combination of GAs and ANNs to produce evolutionary ANNs, with improved performance and simplified architectures [17].

Trial and error method, which is used for the implementation of ANNs, is computationally complex and does not ensure that the proposed architecture is the best one. These restrictions conducted to the development of more efficient methods which are divided in constructive and pruning (destructive) algorithms [18, 19]. Constructive algorithms start with a minimum number of neurons for ANNs and dynamically add neurons, generating more complex ANNs to achieve a satisfactory ANNs solution. On the other hand, a pruning algorithm starts from maximal ANNs and cuts nodes, layers and synaptic connections during the training based on already collected data.

The constructive and destructive algorithms might investigate restricted topological subsets instead of complete space of ANN's architectures [20]. GAs are another approach to solve the problem of ANN's implementation. GAs are an optimization tool when the search space is of great complexity and large size. The determination of optimal ANN structure that solves a specific problem can be achieved by a GA search [15, 21, 22].

### 3 Receiver Operating Characteristic

A Receiver Operating Characteristic (ROC) is a graphical plot of the True Positive Rate (or sensitivity) versus False Positive Rate ( $1 - \text{specificity}$ ) for a classifier system as its discrimination threshold is varied.

Sensitivity and specificity are statistical measures of the performance of a binary classification test. The sensitivity measures the proportion of actual positives which are correctly identified as such. The percentage of children with abdominal pain, who need to be operated, is the sensitivity. The mathematical equation of sensitivity or True Positive Rate (TPR) is

$$\text{Sensitivity (TPR)} = \frac{TP}{TP + FN} \quad (1)$$

where TP is the number of true positive patterns and FN is the number of false negative patterns.

The specificity measures the proportion of negatives which are correctly identified. In particular, the percentage of children having abdominal pain and have not to be operated is the specificity. The mathematical equation of specificity is

$$\text{Specificity (SPC)} = \frac{TN}{TN + FP} \quad (2)$$

where TN is the number of true negative patterns and the FP is the number of false positive patterns.

The correlation between specificity and False Positive Rate (FPR) is

$$\begin{aligned} SPC &= 1 - FPR \Rightarrow \\ FPR &= 1 - SPC \end{aligned} \quad (3)$$

ROC analysis provides tools to select possibly optimal models and to discard suboptimal ones independently from the cost context or the class distribution. ROC analysis is used for diagnostic decision making in medicine, radiology, psychology and other areas for many decades. Nowadays, it has been introduced in other areas like machine learning and data mining [23].

The essential statistic from ROC is the Area Under Curve (AUC), which is a measure of overall accuracy that is not dependent upon a particular threshold [24]. The mathematical equation of AUC is

$$AUC = \frac{1}{n_+ n_-} \sum_{x \in \{\text{set of all test results}\}} \left( n_{- = j} \times n_{+ > j} + \frac{n_{- = j} \times n_{+ = j}}{2} \right) \quad (4)$$

where  $x_+$  is each value of a diagnostic factor for cases with positive actual states,  $x_-$  is each value of the same diagnostic factor for cases with negative actual states,  $n_+$  is the sample size of data set ( $D_+$ ) contained cases with positive actual states, and  $n_-$  the sample size of data set ( $D_-$ ) contained cases with negative actual states.

The AUC can be utilized as an estimator of the discriminatory performance of the diagnostic factors of a system. The AUC for a system without resolving power equals to 0.5, while the AUC for a system with perfect discrimination equals to 1. It is clear that the AUC for a system with satisfactory resolving power is between 0.5 and 1. The greater of AUC, the best discrimination ability for the system [24].

## 4 Data Collection

The abdominal pain data was obtained from the Pediatric Surgery Clinical Information System of Alexandroupolis' University Hospital, Greece. The appendicitis diagnosis is based on 15 clinical and biochemical factors which are sex, age, religion, demographic data, duration of pain, vomitus, diarrhea, anorexia, tenderness, rebound, leucocytosis, neutrophilia, urinalysis, temperature and constipation. The possible diagnosis stages are discharge, observation, no findings, focal appendicitis, phlegmonous or suppurative appendicitis, gangrenous appendicitis and peritonitis. These factors and the diagnosis stages are well described in Table 2 and Table 1, correspondingly, in [12]. As presented in [12], the possible stages of abdominal pain examination are seven, whereof four stages demand operative treatment and three are referred for conservative treatment.

The present study is based on a data set consisted of 516 cases, whereof 422 (81.78%) normal and 94 (18.22%) underwent operative treatment. The pruned data set, used in the proposed model, was divided into a set of 400 (77.52%) records for construction of PNNs and another set of 116 (22.48%) records for assessment of PNNs' performance.

## 5 Evolutionary PNN Architecture

The aim of the present study is the elimination of abdominal pain in childhood diagnostic factors and the determination of essential diagnostic factors for evolved ANN implementation.

A Probabilistic Neural Network (PNN) is the selected ANN architecture among a great variety of ANN topologies that was used in this study. A PNN which is based on Parzen's Probabilistic Density Function (PDF) estimator, is a three-layer feed-forward network, consisting of an input layer, a radial basis and a competitive layer [25].

As it was mentioned in section 4, the number of abdominal pain diagnostic factors equals to 15. The possible combinations of input data subsets are given by following mathematical notation,

$$C = \frac{N!}{k!(N-k)!} \quad (4)$$

where N is the maximum number of diagnostic factors (in this problem N = 15) and k is an integer for the number of diagnostic factors of each input data subset. The k variable's values are range from 1 to 15.

The chromosome's length of each individual was equal to the total number of diagnostic factors, so that the population of the used GA consisted of binary strings of 15 bits. The GA used is two-objective, thus the GA has to search for diagnostic data sets that at the same time: (a) minimize the number of diagnostic factors used during the training phase and therefore minimize the number of nodes in ANN input and hidden layers, and (b) minimize the Mean Square Error of the testing phase. For this purpose, the following fitness function was used:

$$f = MSE + \frac{I}{N} \quad (5)$$

where I is the number of ANN input nodes and N is the maximum number of diagnostic factors in the original, full-sized training data set.

To find out the essential diagnostic factors that can be used for evolutionary PNN, different experiments were performed using the scattered crossover, single-point crossover, two-point crossover and uniform crossover, as well as gaussian mutation and uniform mutation binary GA operators.

## 6 Experimental Results

The aim of the present study is the elimination of abdominal pain diagnostic factors based on GA search for the essential and optimal combination of necessary factors for PNN construction.

Whereas, the PNNs architecture is constrained by the available features of specific problem, 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 100.

An extensive investigation was performed, to assess the PNNs' performance for training and testing full-sized data set consisted of all the 15 diagnostic factors. The obtained results for best-implemented PNNs are presented in Table 1. The radbas and compet were the transfer functions for hidden and output layers, correspondingly.

The number of neurons for input and hidden layers of PNNs was specified by the number of diagnostic parameters and the cases of training set, correspondingly. The number of neurons for PNNs' output layer is seven and is based on coding of possible diagnosis according to [12]. The values of spread and the MSE for the PNNs with the best performance are recorded, correspondingly, in the 1<sup>st</sup> and 2<sup>nd</sup> column of the Table 1.

Table 1. Spread and MSE for the best implemented PNNs with full-sized data set.

Spread	MSE
0.1	0.0025
1.0	0.0025
10.0	0.14
100	0.5375

The obtained results by executing the GA are presented in Table 2. The 2<sup>nd</sup> column of this table depicts the fitness value of the optimal individual, the 3<sup>rd</sup> column, the MSE of the optimal individual and the 4<sup>th</sup> column the independent diagnostic factors that were used as inputs for PNNs construction. It is mentioned that the forenamed values were recorded for the same values of spread as in Table 1.

As it is presented Table 2, the GA managed to converge to PNNs that used 7 up to 8 of the 15 diagnostic factors. At the same time, the MSE of PNNs that were trained with the pruned input data sets is significantly decreased in accordance of MSE of trained PNNs with 15 diagnostic factors. Consequently, the diagnostic ability of evolved PNNs is improved in compare with full-sized trained PNNs.

The decrease in genetically trained PNNs is of the order of 3.15% to 18.9%, depending on the value of spread. The diagnostic factors which are more effective on PNNs training and testing are Demographic Data, Duration of Pain, Leucocytosis, Neutrophilia and Temperature.

## 7 ROC Curves

The available data set of appendicitis' records was processed by ROC analysis. The aim of this data processing was evaluation of importance role of each diagnostic factor for appendicitis estimation. The obtained results are summarized in Table 3. The diagnostic factors are presented in 2<sup>nd</sup> column of the Table 3, while the Area Under Curve of ROC for each factor is recorded in the 3<sup>rd</sup> column of this Table. Each PNN network was evolved genetically using different crossover and mutation operators so the Table 2 records two different results for each spread value.

Table 2. MSE of PNNs trained with pruned sets of diagnostic input factors

Spread	Fitness Value	MSE	Diagnostic Factors
0.1	0.03196126	0.00242131	Demographic data, Vomitus, Rebound, Leucocytosis, Neutrophilia, Temperature, Constipation
	0.03753027	0.00000000	Demographic data, Duration of pain, Anorexia, Tenderness, Leucocytosis, Temperature, Constipation
1.0	0.00823245	0.00242131	Age, Duration of pain, Diarrhea, Anorexia, Tenderness, Leucocytosis, Neutrophilia, Temperature
	0.00435835	0.00000000	Age, Religion, Duration of pain, Tenderness, Lecocytosis, Neutrophilia, Temperature, Constipation
10.0	0.16319613	0.11380145	Age, Demographic data, Duration of pain, Diarrhea, Leucocytosis, Neutrophilia, Temperature
100	1.75484262	0.43583535	Sex, Age, Demographic data, Duration of pain, Tenderness, Leucocytosis, Urinalysis
	1.58002421	0.47457627	Sex, Demographic data, Vomitus, Anorexia, Tenderness, Leucocytosis, Urinalysis

Table 3. Area of Diagnostic factors' curves

Nr	Diagnostic Factors	AUC
1	Sex	0.511
2	Age	0.357
3	Religion	0.560
4	Demographic data	0.572
5	Duration of Pain	0.508
6	Vomitus	0.787
7	Diarrhea	0.535
8	Anorexia	0.690
9	Tenderness	0.947
10	Rebound	0.894
11	Leucocytosis	0.949
12	Neutrophilia	0.977
13	Urinalysis	0.073
14	Temperature	0.929
15	Constipation	0.518

The AUC is an important statistic of ROC analysis. A value of area larger than 0.5, proves the importance role of a diagnostic factor for appendicitis estimation. As it is shown in Table 3, the most important diagnostic factors are Religion, Demographic Data, Vomitus, Anorexia, Tenderness, Rebound, Leucocytosis,

Neutrophilia and Temperature.

The Sex, Duration of Pain, and Constipation are parameters without significant contribution to appendicitis prediction as their values of AUC are equal to 0.5. The Age and Urinalysis are diagnostic factors that have not the ability to discriminate true positive patients as their value are 0.357 and 0.073, correspondingly, which are smaller than 0.5.

The results as obtained by genetically evolved PNNs and ROC analysis were further processed. The aim of this processing is the investigation of the convergence in terms of proposed diagnostic factors for appendicitis estimation. It is concluded that Tenderness, Leucocytosis, Neutrophilia and Temperature are essential factors for appendicitis prediction, so these parameters are strongly recommended have to be recorded for each patient.

## 8 Discussion

This study presents a specific GA to evolve the subsets of patients' data that are used as inputs for PNN construction and testing. After adequate steps of genetic evolution, the GA converged to diagnostic factors subsets that were consisted from 7 or 8 over a total of 15 diagnostic factors. The evolved PNNs overperformed the full-trained PNNs in terms of the MSE. Consequently, the implementation of PNNs based on specifically selected diagnostic factors, instead of all of them resulted to increase of PNNs' performance and prognostic ability while at the same time the training procedure was speed up. The comparison of genetically evolved PNNs' outcomes with those of ROC analysis concludes that the medical diagnostic factors present a high level of redundancy and overlapping. Therefore, a number of the diagnostic factors for appendicitis estimation may be omitted with no compromise to the fidelity of clinical evaluation.

**Acknowledgment** We are grateful to the Pediatric Surgeon of the Pediatric Surgery Clinic of the University Hospital of Alexandroupolis, Greece for their valuable contribution.

## References

1. Dayhoff J., and DeLeo J., "Artificial Neural Networks Opening the Black Box", *CANCER Supplement*, 2001, Vol. 91, No. 8, pp. 1615-1635.
2. Huang D., "A Constructive Approach for Finding Arbitrary Roots of Polynomials by Neural Networks", *IEEE Transactions on Neural Networks*, 2004, Vol. 15, No. 2, pp. 477 - 491.
3. Levano M., and Nowak H., "Application of New Algorithm, Iterative SOM, to the Detection of Gene Expressions", *Proceedings 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 141-147.
4. Zaknick A., "Introduction to the Modified Probabilistic Neural Network for General Signal Processing", *IEEE Transactions on Signal Processing*, 1998, Vol. 46, No.7, pp. 1980-1990.
5. Iliadis L., "An Intelligent Artificial Neural Network Evaluation System Using Fuzzy Set Hedges: Application in Wood Industry" *Proceedings 19th IEEE Annual International Conference on Tools with Artificial Intelligence (ICTA)*, pp. 366-370.



6. Paschalidou A., Iliadis L., Kassomenos P., and Bezirtzoglou C., "Neural Modelling of the Tropospheric Ozone Concentrations in an Urban Site", Proc. 10th International Conference on Engineering Applications of Neural Networks, 2007 Thessaloniki, Greece, pp. 436-445.
7. Keogan M., Lo J., Freed K., Raptopoulos V., Blake S., Kamel I., Weisinger K., Rosen M., and Nelson R., "Outcome Analysis of Patients with Acute Pancreatitis by Using an Artificial Neural Network", Academic Radiology, 2002, Vol. 9, No. 4, pp. 410-419.
8. Brause R., Hanisch E., Paetz J., and Arlt B., "Neural Networks for Sepsis Prediction - the MEDAN-Project1", Journal für Anästhesie und Intensivbehandlung, 2004, Vol. 11, No. 1, pp. 40-43.
9. Gómez-Ruiz J., Jerez-Aragónés J., Muñoz-Pérez J., and Alba-Conejo E., "A Neural Network Based Model for Prognosis of Early Breast Cancer", Applied Intelligence, 2004, Vol. 20, No. 3, pp. 231-238.
10. Mantzaris D., Anastassopoulos G., Tsalkidis A., and Adamopoulos A., "Intelligent Prediction of Vesicoureteral Reflux Disease", WSEAS Transactions on Systems, 2005, Vol. 4, Issue 9, pp. 1440-1449.
11. Anagnostou T., Remzi M., and Djavan B., "Artificial Neural Networks for Decision-Making in Urologic Oncology", Reviews In Urology, 2003, Vol. 5, No. 1, pp.15-21.
12. Mantzaris D., Anastassopoulos G., Adamopoulos A., and Gardikis S., "A Non-Symbolic Implementation of Abdominal Pain Estimation in Childhood", Information Science, 2008, Vol. 178, pp. 3860-3866.
13. Economou G., Lymperopoulos D., Karavatselou E., and Chassomeris C., "A New Concept Toward Computer-Aided Medical Diagnosis - A Prototype Implementation Addressing Pulmonary Diseases" IEEE Transactions in Information Technology in Biomedicine, 2001 Vol. 5, Issue 1, pp. 55-66.
14. Mantzaris D., Anastassopoulos C., and Lymperopoulos K., "Medical Disease Prediction Using Artificial Neural Networks", Proceedings IEEE International Conference on Bioinformatics and BioEngineering, 2008, Athens, Greece
15. Georgopoulos E., Likothanassis S., and Adamopoulos A., "Evolving Artificial Neural Networks Using Genetic Algorithms", Neural Network World, 2000, Vol. 4, pp.565-574
16. Blazadonakis M., Moustakis V., and Charissis G., "Deep Assessment of Machine Learning Techniques Using Patient Treatment in Acute Abdominal Pain in Children", Artificial Intelligence in Medicine, 1996, Vol. 8, pp. 527-542
17. Branke J., "Evolutionary Algorithms for Neural Network Design and Training", Proceedings 1st Nordic Workshop on Genetic Algorithms and its Applications, 1995, Vaasa, Finland.
18. Yao X., "Evolving Artificial Neural Networks", Proceedings of the IEEE, 1999, Vol. 87, No. 9, pp. 1423-1447.
19. Burgess N., "A Constructive Algorithm that Converges for Real-Valued Input Patterns", International Journal on Neural Systems, 1994 Vol. 5, No. 1, pp. 59-66.
20. Angeline P., Sauders G., and Pollack J., "An Evolutionary Algorithm that Constructs Recurrent Neural Networks", IEEE Transaction on Neural Networks, 1994 Vol. 5, pp. 54-65.
21. Adamopoulos A., Georgopoulos E., Manioudakis G., and Likothanassis S., "An Evolutionary Method for System Structure Identification Using Neural Networks", Neural Computation '98, 1998
22. Billings S., and Zheng G., "Radial Basis Function Network Configuration Using Genetic Algorithms", Neural Networks, 1995, Vol. 8, pp. 877-890.
23. Swets J., "Signal Detection Theory and Roc Analysis in Psychology and Diagnostics: Collected Papers", Lawrence Erlbaum Associates, 1996, Mahwah NJ.
24. Streiner D., and Cairney J., "What's Under the ROC? An Introduction to Receiver Operating Characteristics Curves", The Canadian Journal of Psychiatry, 2007, Vol. 52, No. 2, pp. 121-128.
25. Parzen E., "On Estimation of a Probability Density Function and Mode", Annals of Mathematical Statistics, 1962, Vol. 33, No.3, pp. 1065-1076.