Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 40(2), 2013, Pages 237–248
ISSN: 1223-6934

Evolutionary-based intelligent decision model to optimize the liver fibrosis stadialization

Smaranda Belciug, Mircea-Sebastian Serbanescu, Florin Gorunescu, and Radu Badea

Abstract. This paper presents a novel approach to build an intelligent decision system (IDS) inspired by the evolutionary paradigm in order to solve the automatic liver fibrosis stadialization by optimizing the decision-making process. The evolutionary paradigm was used to answer the basic question: how to distinguish between machine learning algorithms facing a medical decision issue, and how to integrate the most effective of them into IDS, able to provide an optimum decision? In the proposed IDS, a set of well-known neural networks are regarded as the initial population of solutions, and an appropriate hierarchy of algorithms is established fitness-proportionally based on a statistically built fitness measure. Then, the IDS framework is built using the best algorithms and the paradigm of a weighted voting system. In a concrete application, the degrees of liver fibrosis, ranging from F0 (no fibrosis) to F4 (cirrhosis), have been automatically identified in 722 patients with chronic hepatitis C infection using 25 main medical attributes. The decision performance proved significantly superior to the classical approach using standalone algorithms. This approach showed a way to directly and easily optimize the medical decision-making by using the evolutionary paradigm.

2010 Mathematics Subject Classification. Primary 68T20; Secondary 68U35.
Key words and phrases. intelligent decision system, evolutionary computing, neural networks, heuristic methods, liver fibrosis.

1. Introduction

From a computational point of view, by a decision support system (DSS) we understand a computer-based information system assisting the decision-making process, and used to solve a large variety of real-life problems. Basically, DSSs are developed to support the solution of unstructured management issues in order to improve the decision-making process [15]. Recent advances in artificial intelligence (AI) and statistical learning (SL) enhanced these systems, giving rise to intelligent decision systems (IDS) [10]. Among the most popular approaches, one can mention the expert systems and well-known AI models, such as neural networks (NNs), genetic algorithms (GAs), support vector machines (SVMs), cluster analysis, intelligent agents, swarm intelligence, random forests, etc. [25], [11], [7]. The IDSs development has been encouraged by their effectiveness when applied in a large variety of real-world decision issues, such as: medical decision-making, business intelligence, customers' relationship management, etc.

The majority of IDSs based on machine learning (ML) techniques is built around a single algorithm and solves a specific problem only. There are various such approaches based on natural computing algorithms applied to different real-life problems. Recent studies propose the use of structured frameworks, usually known as committees of

Received September 20, 2013.

machines, involving more than one algorithm (e.g., NNs, SVMs) working together to solve a given problem [4], [12], [23].

The medical decision-making is nowadays one of the most promising fields to use IDSs. Thus, NNs, SVMs and classification trees have been proposed as standalone algorithms to solve medical decision problems, such as: prediction of severe acute pancreatitis at admission to hospital [1], seizure prediction with spectral power of EEG in epilepsy [14], improving the accuracy of early diagnosis of Alzheimer-type dementia [18]. On the other hand, competitive/collaborative neural computing systems and hybrid neural network-genetic algorithms have been used as structured computational frameworks involving more than one algorithm to solve decision issues, such as: early detection of pancreatic cancer, or breast cancer detection and recurrence [9], [3].

Inspired by the evolutionary paradigm, the current work proposes a flexible strategy to design and implement IDS in order to optimize the liver fibrosis stadialization. Thus, based on the evolutionary metaphor, this strategy envisages the integration in an evolutionary manner of different well-performing NNs, competing and collaborating with each other in straight relation to the environment, hence making an "intelligent" global decision. The main contributions of the paper are twofold: first, the novel evolutionary-inspired strategy was used to develop IDS to support medical decision, and, secondly, the validation of IDS in a real-life application regarding the liver fibrosis stadialization.

2. Methods and Material

2.1. IDS design. The evolutionary-based concept underlying the design and functionality of IDS is inspired by the ideas behind both the committees of machines and the competitive/collaborative computation paradigm. In this respect, firstly a committee of ML techniques is formed by combining several algorithms into a single system, and, secondly, the overall decision of this committee is achieved in a synergic evolutionary way by using a fitness-based weighted voting system (WVS).

The basic idea is to think about efficient types of natural computing algorithms to form the initial population of solutions of a decision problem. Afterward, they will "evolve" this way:

- Obtaining new variants by a 'mutation' operator, used to maintain the genetic diversity from one generation to the next one;
- Obtaining hybrids by a crossover operator, using pairs of selected algorithms to breed offspring.

The original population will be enriched, and, after applying the selection mechanism, the next population will be obtained. Basically, after the algorithms have been tested, the best of them will be retained, based on their fitness given by the individual decision performance, to give birth to the next generation. After a certain number of generations, the most performing algorithms will be kept to form IDS. The population of potential solutions (algorithms) is subject to a problem-dependent selection process, followed by the establishment of a fitness proportional hierarchy.

Secondly, based on WVS and inspired by the evolutionary metaphor, the synergism of this decision structure is involved by both the competition and the collaboration between the component algorithms in making the overall decision. In this context, the 'crossover' operator refers to the fitness-based 'recombination' of the individual decisions, and the 'offspring' represents the overall IDS decision.

Because the algorithms forming IDS are of stochastic nature, they have to be independently run a certain number of times to obtain a reliable result regarding their robustness and effectiveness. In this respect, a sample size estimation procedure (two-tailed type of null hypothesis with default statistical power goal $P \geq 95\%$, and type I error $\alpha = 0.05$) has been proposed. The average accuracy computed as the percentage of correctly classified cases represented the decision performance of each competitor. From a statistical point of view, the classification accuracy obtained during the multiple independent computer runs of each algorithm constitutes a statistical sample of decision performance. The model validation has been achieved by using the standard 10-fold cross-validation.

3. Fitness measure

For the sake of simplicity, the corresponding fitness measure was defined by four standard performance metrics, integrated in the following structured functional framework:

(1) *Decision accuracy* (training/testing) computed as the percentage of correctly classified cases in both training and testing phases.
(2) *Accuracy standard deviation* (SD) (training/testing).
(3) *Accuracy 95% confidence interval* (CI) (training/testing).
(4) *CPU time (running time)* (training/testing) over a specific number of independent computer runs, measuring the computation speed of each algorithm.

3.1. Benchmarking methodology. The fitness-based comparison of the decision components is solved by means of a thorough statistical analysis, using the independent samples of computer runs, and consisting of:

(1) The parametric t-test for independent samples.
(2) The nonparametric alternatives given by Mann-Whitney U test (M-W U) and Kolmogorov-Smirnov two-sample test (K-S).
(3) Two-sided z-test for comparing proportions.

In addition, in conjunction with the above tests, and in order to analyze the long-run behavior of each algorithm, the following statistical tests have been also used:

(1) Normality test: Kolmogorov-Smirnov & Lilliefors (K-S & L).
(2) Homogeneity of variances: Levene's test.

3.2. Variation operators and selection. Based on the fitness measure used in conjunction with the benchmark methodology presented above, the initial and the "offspring" algorithms are statistically compared, and the fittest are chosen to seed the next generations by applying the 'variation' approach. Inspired by the idea underlying the classical variation operators from the evolutionary computing field, we have considered the hybridization of two algorithms as 'recombination', and the development of new variants as 'mutation'.

The survivor selection mechanism used in this approach is based on a steady-state model, in which the entire population is not replaced at once, but just a number of old individuals are replaced by offspring. Technically, once each algorithm has been individually evaluated and the comparison process ended, two scenarios might be possible regarding either the selection of individuals to form the new population, or the IDS final decision-making:

- There is no statistically significant difference between two or more algorithms, measured using the corresponding p-level value (default threshold $p = 0.05$), regardless of the comparison tests;
- Some tests reveal statistically significant difference while others do not identify such a difference.

In the former case, the algorithms in question will be kept or removed together from the population, or will be *ex aequo* ranked, receiving the same decision weight when used to make the overall decision. In the latter case, depending on the ranking order and taking into account the concrete problem to be solved, some of them will be kept while others will be removed.

3.3. Decision-making synergic mechanism. In the decision-making process, the IDS components are involved in a weighted collaborative operating mode in the following way. The best algorithms selected to form IDS are applied to new data and an overall decision is being made based on a standard WVS. From mathematical point of view, WVS is represented as a system $\{q\colon w_1, w_2, ..., w_k\}$, where q is the quota, and algorithm P_k has W_k votes, associated with the corresponding weight w_k. It is worth mentioning that a thorough approach involves, in addition, the use of the power index [24].

In essence, the steps of the synergic decision-making process concern:

(1) The weights estimation for the selected algorithms. By default, a weight is directly proportional to the individual decision performance;
(2) The choice of a quota between 51% and 100% of the total number of votes ('legitimate' WVS);
(3) Computation of the IDSS decision and estimation of the corresponding confidence level.

The estimation of a confidence level of the IDS decision is based on a probabilistic approach, described in [9]. Moreover, if a consensus is obtained, an upper bound can also be estimated.

3.4. Case study: liver fibrosis stadialization. Hepatic fibrosis is the major indicator of progressive liver disease and, in case of patients with chronic hepatitis C, its precise stage is the most important predictor of disease progression and determines the need for antiviral therapy. Technically, liver fibrosis is evaluated semi-quantitatively according to the METAVIR F scoring system as follows: F0 - no fibrosis, F1 - portal fibrosis without septa, F2 - portal fibrosis and few septa, F3 - numerous septa without cirrhosis, and F4 - cirrhosis.

For 60 years liver biopsy was thought of as the "gold standard" diagnosis for assessing the progression of fibrosis in chronic hepatitis C patients [5]. An obvious trend in clinical practice observed in the latter years consists in finding non-invasive ways to obtain reliable methods for liver fibrosis evaluation by using both biochemical tests as well as medical imaging methods, seen as a viable alternative. In this regard, one of the last technological approaches in the evaluation of liver fibrosis is the Fibroscan®(Echosens, Paris, France - http://www. echosens.com), a specially adapted ultrasound device using the principle of the one-dimension transient elastography for the assessment of liver stiffness [19]. On the other hand, ML techniques have been often used as decision-making tools in hepatic fibrosis stadialization. In this respect, one can mention the use of NNs [6], [8], [16], [17], SVMs, evolutionary SVMs, and cooperative coevolutionary classifier [20], [21], [22].

In this study, the evolutionary-inspired IDS has been applied for liver fibrosis stadialization to assess its practical effectiveness. Using a concrete medical database regarding the chronic HCV infection, IDS has proven its efficacy against the standalone utilization of its 'intelligent' components, outlining thus the benefits of this novel approach.

3.5. Dataset. The proposed IDS has been applied on a medical database consisting of 722 patients with chronic HCV infection, examined at the 3^{rd} Medical Clinic within the University of Medicine and Pharmacy "Iuliu Haţieganu" Cluj-Napoca, Romania, between May 2007 and August 2008. The patients were referred to liver stiffness measurement (LSM) one day prior to percutaneous liver biopsy, and besides the epidemiological, anthropometric and clinical parameters, the biological parameters were determined for all patients on the same day as LSM. The data refer to: healthy volunteers, chronic viral C hepatitis, liver cirrhosis, and fatty alcoholic liver disease. The main medical attributes that are chosen by doctors to trigger a certain degree of liver fibrosis are outlined in Table 1, and consist of 25 noninvasive attributes related to the fibrosis stages F0-F4. The fibrosis stage was confirmed by the result of the liver biopsy procedure.

The study was approved by the local Ethical Committee of the University of Medicine and Pharmacy "Iuliu Hatieganu" Cluj-Napoca, and the nature of the study was explained to the patients, each of whom provided written informed consent before the beginning of the study, in accordance with the principles of the *Declaration of Helsinki* (revision of Edinburgh, 2000).

3.6. Evolutionary-based design of the model. The main components of the evolutionary-based design of IDS are the following ones.

A. Initial population of algorithms

For the sake of simplicity, the initial population of algorithms taking into account consisted of neural networks (NNs) only, known as well-performing classifiers, and a genetic algorithm (GA), seen as efficient optimization algorithm. Concretely, we considered three well-known NNs in use today: three-layer perceptron (3-MLP), radial basis function neural network (RBF), and probabilistic neural network (PNN) [4], [12].

Note.

(1) For a standardization of the decision performance, we used the implementation within the Statistica 7 software package (Statistica Neural Networks (SNN) package -StatSoft. Inc., Tulsa, OK 74104, USA).
(2) Although the evolution process can be theoretically iterated in many steps (generations), we stopped it after the second generation, and, therefore, IDS consisted of one parent (3-MLP), one offspring (hybrid -MLP/GA) and one mutant (partially connected neural network -PCNN).
(3) According to the universal approximation theorem, a single hidden layer (i.e., 3-layer MLP) is sufficient to uniformly approximate any given continuous function [12]

B. Crossover: hybrid MLP/GA algorithm

We considered the hybridization of MLP with GA to enhance the classification performance of the former, since MLP proved to provide the best decision performance during the training/testing process. Starting from the fact that training MLP means

TABLE 1. Description of the fibrosis data average values and standard deviation

Feature	F0	F1	F2	F3	F4
Stiffness	4.28± 0.88	6.06± 2.02	7.8± 3.6	11.92± 6.93	32.28± 18
Sex	Male: 8	Male: 78	Male: 55	Male: 32	Male: 118
Body Mass Index	25.23±5	23.76±8.05	24.94±6.27	24.31±8.2	24±9.6
Glycemia	110±79.83	83.4±41.39	80.45±47.42	99±44.62	101.93±49.3
Triglycerides	125.66±90	92.62±79.96	87.48±64.66	98.85±64	106.6±74.92
Cholesterol	207.48±73.92	154.44±94.38	150±86.23	164±73.8	147.27±74.5
HLD Cholesterol	39.21±28.49	33.42±32.45	37.4±32.93	33.97±32.6	27.9±29.46
Aspartate aminotransferase	32.41±20.42	39±29	47.47±36.5	68.56±54.37	77.15±54.9
Alanin aminotransferase	53.72±37.85	63.77±51.24	75.62±60.63	102.75±94.93	77.2±59.82
Gama glutamyl transpeptidase	63.59±63.52	49.19±46.74	62.45±71	78.34±84.3	137.4±208.44
Total bilirubin	0.57±0.31	0.56±0.4	0.58±0.46	0.7±0.4	2.15±7.68
Alkaline phosphatase	188.83±146.31	148.17±97.01	150.33±110.45	178.97±106.77	248.95±163.3
Prothrombin index	87±42.82	76.64±47.5	63.84±50.45	70.15±44.28	69.4±36.8
TQS (quiq time)	12.36±5.93	10.78±10.37	9.06±7.76	10.6±7.9	16±7
Prothrombin time ratio	0.84±0.36	0.75±0.43	0.68±0.48	0.84±0.46	1.09±0.48
Prolonged activ. partial thromboplastin time	24.8±12.26	20.43±14.23	17.85±15.08	21±14.42	28.26±11.66
Haematids	4.05±1.72	3.87±2.08	29.97±336.66	4.16±1.85	6.26±29.13
Hemoglobin	11.89±5.07	11.5±5.95	11.05±6.36	12.62±5.37	12.85±4.32
Hematocrit	35.18±14.86	31±19.68	25.65±21.44	30.9±19.87	36.06±14.38
Medium erytrocity volume	69±36.23	57.7±41.63	48.92±44.11	58.8±42	72.82±35.73
Avg. erytrocitary hemoglobin	22.16±12.83	18.7±14.38	16.43±15.04	19.2±14.62	26.3±15.8
Avg. concetration of hemoglobin in a red blood cell	16.31±17.2	17.26±17	15.17±16.8	18.6±17.05	21.68±17.22
Thrombocytes	209.69±73.35	184.6±105.54	170.85±105.3	175.84±85.19	116.33±76.3
Sideraemia	56.59±56	69.38±58.45	73.31±62.69	86±74.2	85.25±77.72
Interquartile range	0.55±0.2	0.89±0.6	1.16±1.07	1.75±1.68	4.91±4.76

optimizing its synaptic weights, on the one hand, and GAs represent efficient optimization techniques, we replaced the classical back-propagation (BP) learning algorithm with the GA approach. The recently developed method of evolutionary-driven MLP [3], which has been thoroughly investigated and validated on several benchmark real-world medical datasets regarding the breast cancer detection and recurrence, has provided the motivation to consider employing the new learning strategy towards the automatic identification of the liver fibrosis degrees.

The proposed methodology unfolds in the following manner. The hybrid algorithm consists of:

- The NN component -the classifier, designed as MLP;
- The GA component -the MLP's weights optimizer, designed as GA.

MLP consists of a number of inputs equaling the number of predictive attributes (25), one hidden layer with a number of processing units equaling the number of liver fibrosis stages (5), and one output unit representing the class attribute (F0-F4). The classical logistic sigmoid:

$$f(x) = 1/(1 + e^{-x}) \tag{1}$$

has been considered as activation function since it satisfies the condition imposed by the universal approximation theorem on the approximation function. It is worth noting that the choice of the number of hidden processing units has been heuristically chosen in order to produce competitive performance with low computational costs.

Next, a weight vector has been encoded as a vector of real numbers (MLP weights), and thus, in evolutionary computing terms, the weight vector is represented by a chromosome which contains a number of genes equaling the number of neurons from the input layer multiplied by the number of neurons from the hidden layer. The binary tournament selection [7] has been used, while the blend crossover (BLX-α) [13] has been considered for recombination. Technically, two offspring were (uniformly) randomly generated from the interval [chromosomemin - Iα, chromosomemax + Iα], $I = chromosome_{max} - chromosome_{min}$, using the formulas:

$$chromosome_{min} = min\{chromosome_i^1, chromosome_i^2\} \tag{2}$$

$$chromosome_{max} = max\{chromosome_i^1, chromosome_i^2\} \tag{3}$$

The mutation process consisted of two steps:

(a) One establishes whether an addition or a subtraction is being made by randomly generating a number between 0 and 1 per gene for each chromosome chosen for mutation. If the number is smaller than the mutation probability p_m, then that gene is being mutated;

(b) Using the chromosome's error, given by the formula:

$$chromosome_{error} = (100 - chromosome_{accuracy})/100, \tag{4}$$

each gene is being mutated according to step (a), subtracting or adding to the initial value the chromosome's error.

Regarding the appropriate choice of the population size and number of generations, we have heuristically considered a population of 100 individuals and a number of generations equaling 100, given that the best performance has been experimentally obtained in this case. At each iteration, all the 100 existing chromosomes have been evaluated (the fitness of a given chromosome was computed by running the network on the training set and considering the corresponding accuracy), and the best 40 of them were kept for reproduction and mutation. Then, they have been mated with each other, replenishing the population with 40 new offspring. The corresponding parameters were heuristically chosen as $\alpha = 0.3$ and $p_m = 0.35$ to provide the best accuracy possible in this case.

C. Mutation: partially connected MLP

The 'mutation' proposed an artificial replica of the way the human brain works, commonly known as partially connected neural network (PCNN). Originally, when a signal is being processed by the brain, only certain neurons participate to that course of action, those who have been excited, the other synapses being inhibited. Thus, inspired by the methodology presented in [2], we developed a PCNN especially designed to handle this type of data. Technically, after a certain number of training samples have been presented to the network, the weights that did not suffer major modifications (i.e., did not surpass a certain threshold throughout the BP learning algorithm) are erased from the network's architecture, being inhibited (i.e., set to 0). In this study, we have heuristically used a MLP with two hidden layers trained by the BP algorithm. The first hidden layer contains 7 neurons, heuristically chosen for optimal performance, the second hidden layer contains a number of processing units equaling the number of liver fibrosis stages (5), and there is one output unit

representing the class attribute (F0-F4). In addition, the models parameters have also been heuristically chosen as N = 50 training samples, and $\tau = 0.005$, in order to produce higher performance with less computational costs.

Note. Both the hybrid MLP/GA and PCNN algorithms have been implemented in Java by the authors.

4. Results

Apart from the theoretical argument based on the philosophy behind the evolutionary strategy that this novel technique would surpass its component algorithms, the assumption that it also outperforms standard ML classifiers as regards the decision accuracy on the liver fibrosis prediction task is further on investigated.

Technically, to assess the performance of the decision models envisaged in this study, each algorithm has been executed in 100 independent computer runs (i.e., each model has been run 100 times in a complete 10-fold cross-validation cycle), thus obtaining a statistical power greater than 95%, with type I error $\alpha = 0.05$, for the statistical comparison tests subsequently used. The average accuracy obtained over the 100 complete cross-validation cycles represented the decision performance of each competitor. The CPU time was measured for each model averaged over 100 independent runs using a relatively high-level system with Intel(R) Core(TM)2 Extreme CPU X9000, 2.80 GHz and 4GB (RAM).

4.1. IDS construction. As it was stated above, the fitness measure considered in this case consisted of the training/testing accuracy, standard deviation (SD), 95% confidence interval (CI), and CPU time, averaged over 100 independent complete cross-validation cycles. The experimental results are displayed in Table 2.

Table 2. Fitness measure

Model	Training accuracy/SD (%)	Testing accuracy/SD (%)	95% CI-training/testing	CPU time
3-MLP	63.31/3.67	59.19/2.12	(62.58, 64.04)/(58.77, 59.61)	3'58"
RBF	60.78/2.84	55.35/3.45	(60.22, 61.34)/(54.66, 56.03)	0'6"
PNN	59.93/7.66	53.64/3.11	(58.42, 61.46)/(53.03, 54.26)	0'6"
Hybrid	63.50/3.14	61.16/4.54	(62.87, 64.13)/(60.26, 62.06)	3'15"
PCNN	62.02/3.28	57.92/3.85	(61.38, 62.66)/(57.17, 58.67)	3'21"

At the first glance, the competitive phase generated a fitness-based hierarchy, in terms of classification performance, given as follows: (1) hybrid MLP/GA, (2) 3-MLP, (3) PCNN, (4) RBF, (5) PNN. Taking into account the difference in performance between the first three algorithms above and the last two, despite their higher computation speed, we kept for the subsequent comparative statistical analysis and IDS construction MLP/GA, 3-MLP and PCNN only. In this context, it is worth comparing the performance of the selected algorithms with the performance, ranging between 51.8% and 62.03%, obtained by the use of other ML methods applied on the same dataset and reported in the literature [8], [21], [22], showing the effectiveness of the IDS components.

To compare the performance, we considered the following four statistical tests: (a) t-test for independent samples, (b) Mann-Whitney U test (c) Kolmogorov-Smirnov two-sample test, (d) two-sided z-test, while to analyze the long-run behavior of each

algorithm we used: (a) Kolmogorov-Smirnov & Lilliefors, and (b) Levene tests. The reason to use both parametric and nonparametric comparison tests resides in the fact that the data are not entirely normally distributed, and the non-parametric alternatives could provide valuable information, in addition. The corresponding results are displayed in Table 3.

Table 3. Benchmark results

	Comparison tests			
Competitors	t-test p-level	*Mann-Whitney* U/ p-level	*Kolomogorov* -Smirnov max_min/max_pos (p-level)	z -test (p - level)
3-MLP vs. hybrid	3.947/0.00	3000/0.00	-0.43/0.09 (< 0.001)	0.78
3-MLP vs. PCNN	2.86/0.00	4017/0.2	-0.1/0.26 (< 0.005)	0.88
Hybrid vs. PCNN	5.43/0.00	2771/0.00	-0.01/0.42/(< 0.001)	0.65

Thus, while there are highly significant differences in testing accuracy between the three competitors when using both the parametric t-test for independent samples, the nonparametric alternative given by the Kolmogorov-Smirnov two-sample test and the Mann-Whitney U test, excepting the case 3-MLP vs. PCNN (M-W U), the two-sided z-test did not reveal any statistically significant difference, which is not surprising taking into account the nature of this test comparing the proportions only.

The normality test applied to data revealed different results. Thus, 3-MLP and PCNN behave "normally", while the data corresponding to the hybrid MLP/GA were not completely normally distributed. In addition, the variances are slightly different (Levene's test). This is not problematic for the t-test since the sample size is large enough and the samples have the same size.

4.2. Synergic decision process. Taking into account the above selection process, the WVS considered in this study was given by {51: 34.3, 33.2, 32.5}.

In order to evaluate the confidence level of the overall diagnosis, the corresponding decision partition B_i, i =1, 2, 3, with $P(B_1) = 34.3\%$, $P(B_2) = 33.2\%$ and $P(B_3) = 32.5\%$, respectively, has been considered. Since the percentages of successful diagnosis each algorithms produces is 61.16%, 59.19% and 57.92%, respectively, then the confidence level of the IDS decision equals 59.45%. For a consensus between the three "voters" ("virtual doctors"), obtained in certain cases, the upper bound confidence level can reach 93.33%.

4.3. Case study. A concrete example consisting in three different testing cases is displayed in Table 4.

The model decision was given by the weighted sum of the three classifiers' diagnosis, rounded to the nearest integer, determining the medical IDS diagnosis (fibrosis stage). Thus, while in the first case, the first two standalone algorithms provided the same diagnosis (F1), which is the real one, contrary to the third one (F2), the WVS collaborative mechanism provided the right diagnosis. In the second case, the second best algorithm provided the correct diagnosis (F3), while the first and the third disagreed. Still the WVS system successfully solved the situation, providing the right diagnosis. The third case, the most interesting one, showed that all the three

TABLE 4. Concrete application (Decisions for three testing samples)

True diagnosis	Hybrid MLP/GA	3-MLP	PCNN	WVS diagnosis	IDSS
F1	F1	F1	F2	$0.343 \cdot 1 + 0.332 \cdot 1 + 0.325 \cdot 2 = \mathbf{1.32}$	**F1**
F3	F4	F3	F2	$0.343 \cdot 4 + 0.332 \cdot 3 + 0.325 \cdot 2 = \mathbf{3.01}$	**F3**
F2	F3	F3	F0	$0.343 \cdot 3 + 0.332 \cdot 3 + 0.325 \cdot 0 = \mathbf{1.99}$	**F2**

decision components were wrong. Still, the WVS mechanism, by balancing the standalone decisions, provided the correct one. To conclude, this example illustrates once again the effectiveness of the synergic competitive/collaborative diagnosis system in comparison to separate standalone competitors.

5. Conclusions and Future Work

As opposed to the traditional paradigm, an evolutionary metaphor-based strategy is proposed to build an intelligent decision model for medical purpose. Applied to a real-world medical database regarding the liver fibrosis stadialization, its design and functionality proved to be straightforward and efficient. Due to the well-known flexibility and adaptability of its "intelligent" components, it is easily adaptable to a wide variety of decision problems.

Future work has to propose:

(1) The design and deploy of a management module, seen as the "engine control unit (ECU)", automatically running the decision system;
(2) The enrichment of the initial population of candidate algorithms by considering other powerful bio-inspired algorithms instead of NNs only.
(3) The application of IDS to other medical or non-medical decision issues.

Finally, let us mention the only major limitation in using this decision model, given by this particular database size used in the learning phase. Basically, for large databases, one can abandon the 10-fold cross-validation, using instead different subsets for training and testing (e.g., 75% vs. 25%), and increasing thus the reliability of the results.

The present work was supported by the research grant no.41071/2007-2010, entitled "SONOFIBROCAST" (Diagnostic algorithm for the stage and prediction of evolution of hepatic fibrosis using noninvasive ultrasound techniques, optimized by stochastic analysis and image analysis), financed by the Romanian Ministry of Education and Research and the National Authority for Scientific Research (MEdC & ANCS), and involving University of Craiova, University of Medicine and Pharmacy of Craiova, and University of Medicine and Pharmacy "Iuliu Hatieganu" of Cluj-Napoca.

References

[1] B. Andersson, R. Andersson, M. Ohlsson and J.J. Nilsson, Prediction of Severe Acute Pancreatitis at Admission to Hospital Using Artificial Neural Networks, *Pancreatology* **11**, (2011), 328–335.

[2] S. Belciug and El-Darzi E., A partially connected neural network-based approach with application to breast cancer detection and recurrence, *Proceedings of the 5th IEEE Conference on Intelligent Systems-IS2010*, London, UK, (2010), 191–196.
[3] S. Belciug and F. Gorunescu, A hybrid neural network/genetic algorithm applied to breast cancer detection and recurrence. *Expert Systems* **30**, (2013), no. 3, 243–254.
[4] C. Bishop, *Neural networks for pattern recognition*, Oxford University Press, 1995.
[5] A.A. Bravo, S.G. Sheth and S. Chopra, Liver biopsy. *N. Engl. J. Med.* **344** (2001), 495–500.
[6] M. Cazzaniga, F. Salerno, G. Borroni et al., Prediction of asymptomatic cirrhosis in chronic hepatitis C patients: accuracy of artificial neural networks compared with logistic regression models. *Eur. J. Gastroenterol. Hepatol.* **21**, (2009), no. 6, 681–687.
[7] F. Gorunescu, *Data Mining: Concepts, Models and Techniques*, Springer-Verlag Berlin Heidelberg, 2011.
[8] F. Gorunescu, S. Belciug, M. Gorunescu, M. Lupsor, R. Badea and H. Stefanescu, Radial basis function network-based diagnosis for liver fibrosis estimation, *Proc. 2nd Intl. Conf. on e-Health and Bioengineering*, 17-18 September 2009, Iasi, Romania, University Press, (2009), 209–212.
[9] F. Gorunescu, M. Gorunescu, A. Saftoiu, P. Vilmann and S. Belciug, Competitive/collaborative neural computing system for medical diagnosis in pancreatic cancer detection. *Expert Systems* **28** (2011), no. 1, 33–44.
[10] S. Guerlain, D.E. Brown and C. Mastrangelo, Intelligent Decision Support Systems, *Proc. IEEE International Conference on Systems, Man, and Cybernetics* **3** (2000), 1934–38.
[11] T. Hastie, R. Tibshirani and R. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.)*, Springer, 2009.
[12] S. Haykin, *Neural networks. A comprehensive foundation* (2nd ed.), Prentice-Hall. Inc., 1999.
[13] F. Herrera, M. Lozano and A.M. Sanchez, A taxonomy for the Crossover Operator for Real-Coded Genetic Algorithms: An Experimental Study, *International Journal of Intelligent Systems* **18** (2003), 309–338.
[14] Y. Park, L. Luo, K.K. Parhi and T. Netoff, Seizure prediction with spectral power of EEG using cost- sensitive support vector machines, *Epilepsia* **52** (2011), 1761–1770.
[15] D.J. Power, *A Brief History of Decision Support Systems*, http://dssresources.com/history/dsshistory.html, November 2013.
[16] M.R. Raoufy, P. Vahdani, S.M. Alavian, S. Fekri, P. Eftekhari and S. Gharibzadeh, A novel method for diagnosing cirrhosis in patients with chronic hepatitis B: artificial neural network approach, *J. Med. Syst.* **35** (2011), no. 1, 121–126.
[17] S. Resino, J.A. Seoane, J.M. Bellon et al., An artificial neural network improves the non-invasive diagnosis of significant fibrosis in HIV/HCV coinfected patients, *J. Infect.* **62** (2011), no. 1, 77–86.
[18] D. Salas-Gonzalez, J.M. Gorriz, J. Ramirez et al., Computer-aided diagnosis of Alzheimer's disease using support vector machines and classification trees, *Phys. Med. Biol.* **55**(10) (2010), no. 10, 2807–2817.
[19] L. Sandrin, B. Fourquet, J.M. Hasquenoph et al., Transient elastography: a new noninvasive method for assessment of hepatic fibrosis, *Ultrasound Med. Biol.* **29**, (2003), 1705-1713.
[20] Y. Sela, M. Freiman, E. Dery et al., fMRI-based hierarchical SVM model for the classification and grading of liver fibrosis, *IEEE Trans. Biomed. Eng.* **58** (2011), no. 9, 2574–2581.
[21] C. Stoean, R. Stoean, M. Lupsor, H. Stefanescu and R. Badea, Feature selection for a cooperative coevolutionary classifier in liver fibrosis diagnosis, *Computers in Biology and Medicine* **41** (2011), 238–246.
[22] R. Stoean, C. Stoean, M. Lupsor, H. Stefanescu and R. Badea, Evolutionary-driven support vector machines for determining the degree of liver fibrosis in chronic hepatitis C, *Artificial Intelligence in Medicine* **51** (2011), 53–65.
[23] D. Valinicius, A. Verikas, M. Bacauskiene and A. Gelzinis, Evolving committees of support vector machines, *LNCS Machine Learning and Data Mining in Pattern Recognition*, vol. **4571** (2007), 263–275.
[24] P. Tannenbaum, *Excursions in Modern Mathematics (6th ed)*, Prentice Hall, 2006.
[25] E. Turban, J.E. Aronson and T-P. Liang, *Decision Support Systems and Intelligent Systems* (7th ed.), Prentice-Hall of India, 2006.

(Smaranda Belciug) Department of Computer Science, University of Craiova, Romania
E-mail address: smaranda.belciug@inf.ucv.ro

(Mircea-Sebastian Serbanescu) Department of Computer Science, University of Pitesti, Romania
E-mail address: mircea_serbanescu@yahoo.com

(Florin Gorunescu) Department of Biostatistics and Computer Science, University of Medicine and Pharmacy of Craiova, Romania
E-mail address: florin.gorunescu@webmail.umfcv.ro

(Radu Badea) Department of Ultrasonography, University of Medicine and Pharmacy "Iuliu Hatieganu", Cluj-Napoca, Romania
E-mail address: rbadea@umfcluj.ro