Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 37(1), 2010, Pages 43–54
ISSN: 1223-6934

Evolutionary Conditional Rules versus Support Vector Machines Weighted Formulas for Liver Fibrosis Degree Prediction

Ruxandra Stoean, Catalin Stoean, Monica Lupsor, Horia Stefanescu, and Radu Badea

Abstract. Present paper brings together two novel evolutionary techniques designed for classification and applied for the differentiation among five possible degrees of liver fibrosis within chronic hepatitis C. A purely evolutionary method - the cooperative coevolutionary classifier - endowed with a hill climbing algorithm for the selection of influential attributes is put in opposition to a hybridized approach for the task - the evolutionary support vector machine. Each of the two exhibits interesting resulting features as regards additional information on the importance of each indicator and the interaction among these for the final predicted outcome. The medical experts can eventually benefit from both methodologies as a support for their decision making and decide what further knowledge they need to extract from them, i.e., either in the form of conditional rules, weighted formulas or both.

2010 Mathematics Subject Classification. Primary 68T05; Secondary 68T20.
Key words and phrases. classification, evolutionary algorithms, cooperative coevolution, support vector machines, liver fibrosis staging.

1. Introduction

The correct prediction of the degree of liver fibrosis provides crucial information about the appropriate diagnosis and treatment in patients with chronic hepatitis C. New non-invasive ways of assessment in combination with intelligent computational support are the primary current concern of this area in medicine. The current paper responds to this demand by presenting two possible evolutionary learning techniques that can achieve both a competitive accuracy and provide valuable extra information on the influence and interaction of predictive attributes.

The cooperative coevolutionary algorithm (CCEA) has been previously successfully embedded into a classification scheme [21, 22] and is now used for distinguishing between five possible stages of fibrosis. Conditional rules for each degree are evolved by parallel subpopulations and interact at the evaluation phase in order to form a complete set to classify the data. Moreover, a hill climbing algorithm is further inserted to the preamble of the CCEA and conducts a dynamic feature selection to aid the decision process.

Complementarily, a support vector learning (SVM) scheme endowed with an evolutionary algorithm (EA) engine for solving the inherent optimization problem is put to the same task. The evolutionary support vector machine approach (ESVM) [23, 24] outputs weighted formulas of the relations between the medical indicators towards a certain fibrosis outcome.

Received December 16, 2009. Revision received February 07, 2010.

The paper is structured in the following manner. Section 2 describes the fibrosis staging problem in detail and outlines the available data for experimentation. Section 3 defines the chosen standpoint on the classification task, while sections 4 and 5 present the two approaches in turn. Section 6 gives the experimental results while a comparison of the advantages of each considered technique is discussed in the enclosing section.

2. Liver Fibrosis Staging

The prognosis and management of chronic liver diseases largely depend on the amount and progression of liver fibrosis. In patients with chronic hepatitis C, the precise stage of liver fibrosis is the most important predictor of disease progression and determines the need for antiviral therapy [17]. Until recently, liver biopsy has been the only way to evaluate fibrosis and it has traditionally been considered as the gold standard [3]. Liver fibrosis is evaluated semi-quantitatively according to the METAVIR 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.

However, liver biopsy is an invasive and painful procedure, often with poor patient compliance, also carrying a significant, although small risk of life-threatening complications [5]. An obvious trend in the clinical practice observed in the latest years consists of finding a correct method for liver fibrosis evaluation in a non-invasive way, both by biochemical tests as well as imaging methods, as an alternative to liver biopsy. The latest technological discovery in the evaluation of liver fibrosis is the Fibroscan, a specially adapted ultrasound device using the principle of the one-dimension transient elastography for the assessment of liver stiffness. This gained popularity as a user-friendly, non-invasive technique for measuring liver stiffness, which has been shown to be significantly correlated with liver fibrosis stages in a variety of clinical conditions including chronic hepatitis C [6, 19]. Although a good binary threshold differentiation among degrees of fibrosis has been obtained from investigating solely stiffness [8, 29] (even on the particular patients involved in this study [14]), the interaction between itself and the other complementary medical examinations could further improve the performance while increasing the complexity of the problem through a simultaneous distinction between more degrees of fibrosis.

It is in this respect that the association between the elastographical value, together with the other clinical and biochemical medical attributes, and the corresponding different stages of liver fibrosis must be artificially discovered and intelligently learnt in order to offer a reliable automated assistance within the modern medical decision-making in the field. Therefore, we study a chronic hepatitis C data set coming from the 3rd Medical Clinic, University of Medicine and Pharmacy, Cluj-Napoca, Romania, that contains 722 samples of 24 indicators. The first one is the stiffness indicator from the Fibroscan, while the others represent the usual hematological and biochemical (illustrative of the hepatic function) exams that are required in a patient with chronic hepatitis C. The fibrosis stage is confirmed by the result of the liver biopsy procedure.

Previous works on a primary division of the current data set included 125 patients and 26 attributes and pursued the same goal of differentiating among the 5 distinct liver fibrosis stages. It involved the application of neural networks, naïve Bayesian classification and the k-nearest neighbor algorithm, while investigating the importance of the stiffness indicator within the discrimination [9]. On the same initial collection,

[1] employed a naïve Bayes classifier and a probabilistic neural network model; these were also used in conjunction with a feature selection algorithm, but there was gain only in runtime, without improving the accuracy. Overall, the best reported accuracy of prediction on this smaller data set was of 70%.

In comparison to the previous research inquiries, the two proposed evolutionary-powered learning engines aim to demarcate among the 5 different classes on the significantly larger data set and additionally provide some automatically detected relations (either in the form of rules or weighted formulas) between the indicators that were considered for classification.

3. A Viewpoint on the Formulation of the Classification Task

A classification problem may be defined by a training set of m pairs of the form $(\mathbf{x}_i, y_i)$, where each couple holds the information related to a data sample and its confirmed outcome, and a test set of p pairs of the type $(\mathbf{x}_i^{'}, y_i^{'})$, where the target is hidden to the learning machine and must be predicted. Every example is described by n attributes, $\mathbf{x}_i, \mathbf{x}_i^{'} \in [a_1, b_1] \times [a_2, b_2] \times ... \times [a_n, b_n]$, where a_i, b_i denote the bounds of definition for attribute i of a sample, $i = 1, 2, ..., n$, and each corroborated outcome $y_i, y_i^{'} \in \{0, 1, ..., k-1\}$, where there are k possible classes.

Learning then pursues two steps: training and testing. A chosen classifier learns the associations between each training sample and the acknowledged output. Either in a black-box manner or explicitly, the obtained classifier then takes each test sample and makes a forecast on its probable class, according to what has been learnt. For generalization ability testing and objectivity, it is customary that the training/test stages are repeated for a (sufficient) number of times over different arrangements of the training/test sets and this is called cross-validation. One common type of cross-validation (and the one experimentally chosen for this work) is the repeated random sub-sampling validation, which presumes the successive random splitting of the given data into training and test sets. The prediction accuracy of the technique is eventually assessed as the average result concerning the number of correctly labeled cases over the total number of test samples. The information regarding the disposal of the misclassified examples – the confusion matrix – brings additional insight into the problem and possible limitations of the considered method.

4. Cooperating Evolutionary Rules for Degree Differentiation

Within a possible evolutionary treatment of a classification problem [21, 22, 26], the aim is to perform a generation of rules for each class. Rules are first randomly created, subsequently tested against the training set and continually adjusted in order to increase the training accuracy they provide. A means through which rules with different labels communicate must be consequently further created. Recent research showed that cooperative coevolution can be successfully applied for classification purposes [21, 22].

According to the Darwinian principles, an individual evolves through the interaction with the environment. An important part of its environment is, however, represented by other individuals so, as a consequence, evolution can be viewed as coevolution. On the one hand, the individuals could collaborate for the same purpose and thus construct the solution together within a cooperative coevolution or, on the contrary, they could compete against each other for the same resources as part of a

competitive framework. The CCEA [18] requires that any candidate solution of the problem at hand is decomposed into several subsolutions and each of these separate components is evolved by a distinct EA. The only interaction between the different populations takes place when an individual is evaluated: its quality (or adaptation to the environment) cannot be measured separately because it represents only a part of a potential solution, but individuals from all the other populations have to be selected and brought together in order to construct a complete candidate solution that can be assigned a fitness score. The numerical value that stands for its evaluation is assigned as the fitness of the initial individual; naturally, its fitness value directly depends on the collaborators that are chosen, as an individual might form a good solution with some individuals from the other populations, while it might yield very poor results with others.

For the classification problem, each population may evolve rules for a certain class and thus the number of species equals the number of outcomes of the classification problem. A complete candidate solution may therefore represent an entire set of rules that optimally associates the indicators with the fibrosis levels.

Each rule has the same representation as the samples in the data set to be classified, i.e., it has the same number of features and one outcome. The value for each attribute is initially randomly generated following a uniform distribution between the definition bounds for that specific feature, that is between the minimum value that exists for that attribute in the data set and the maximum one. Individuals can be interpreted as simple IF-THEN rules having the condition part in the attributes space and the conclusion in the classes space (1). This representation is consistent with the considerations within the area of learning classifier systems [4, 11, 28].

However, here an individual does not specifically encode the decision class, as this is implicit from the population to which the rule belongs.

$$\text{IF } attr_1 = val_1 \text{ AND } attr_2 = val_2 ... \text{ AND } attr_n = val_n \text{ THEN } y_i \tag{1}$$

In order to evaluate an individual (rule) from a certain population, a complete set of rules has to be formed, in the sense that one rule from each of the other classes has to be selected. The entire rule set is then applied to the training data: for every sample, similarities between each rule in the set and the current object are computed and the found class is concluded to be the one of the rule that is closest. A prediction accuracy over all training samples is obtained and assigned as the fitness of the rule to be evaluated.

In order to calculate how close the current rule is to a sample from the training/test set, a distance measure has to be employed. In the experiments conducted within the current paper, it is the Manhattan distance (2) that is considered in this respect; c and s represent an individual and a sample, respectively, and c_i is the i-th component of the potential solution.

$$d(c, s) = \sum_{i=1}^{n} \mid c_i - s_i \mid \tag{2}$$

However, there is obviously no obstacle in using any other desired distance measure.

At the end of the run, CCEA provides several populations of rules, each one with the prototypes that define a certain class. In order to apply these rules for samples in the test set, individuals are selected once more from each population, objects are labeled accordingly and the prediction accuracy is achieved.

It is customary that techniques for automated diagnosis in biology and medicine [15, 20] make an a priori use of some mechanism of selecting the most relevant indicators in the data set. The reason lies in the assumption that some attributes might only hinder the search for the accurate solutions or even block the entire method under the curse of dimensionality. This proves especially true for the CCEA, as it otherwise takes all indicators in the conditional part of a rule.

In this respect, the classification method can be aided by an incorporated dynamic feature selection mechanism. In order to make the additional procedure efficient, a limited number of applications of the CCEA technique is desired. For that reason, a hill climbing algorithm is used with the purpose of choosing the attributes that make the CCEA perform more beneficially.

An individual is represented as binary, has a number of genes equal to the number of features that exist in the data set and a value of 1 means that the corresponding attribute is taken into consideration, while 0 that the attribute is skipped. An individual is randomly constituted and the selected attributes are considered for the CCEA. The algorithm generates rules based on the newly defined training set and then applies them to the test set; an accuracy is obtained and that value represents the fitness evaluation of the hill climber. Perturbation is then applied for the individual, a new climber is obtained, it is evaluated and, if fitter, it replaces the previous one.

Once evolution is finished, the best individual from each population is collected into the rule set, which is applied to the test samples. It is therefore one rule that "votes" for each class.

5. Evolutionary Support Vector Machines for Stage Classification

A modern and powerful way to tackle classification is provided by the original, though internally complex technique of SVMs. The methodology considers the distinct classes of examples being divided by geometrical surfaces – separating hyperplanes – whose optimal behavior is determined by an extension of the method of Lagrange multipliers.

The SVM learning scheme specifically leads to the optimization problem in (3). A hyperplane of coefficients w and b is required to simultaneously achieve the separation with a minimal training error expressed by the constraints and by preserving the ability to generalize on new instances, specified by the objective function. The constraints may be relaxed by introducing allowed deviations ξ_i, $i = 1, 2, ..., m$, to the formulation, with the additional requirement that they are kept to a minimum (penalized by C). If the samples cannot be delimited in a linear fashion, then a nonlinear surface is obtained by the kernel trick [27]: data are mapped into a higher dimensional space where a linear surface is able to do the separation. Kernels are commonly taken in either a polynomial expression (with a certain degree) or Gaussian (radial) formulation (of parameter γ).

$$\min_{\mathbf{w},b} \text{such that the objective } \|\mathbf{w}\|^2 + C\sum_{i=1}^{m}\xi_i, C > 0, \text{ holds}$$
$$\text{subject to constraints } y_i(\langle \mathbf{w}, \mathbf{x}_i\rangle - b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, ..., m\ . \tag{3}$$

The resulting unimodal optimization statement in (3) may be standardly resolved by relying on a mathematically complex extension of the Lagrange multipliers technique [27]. A dual formulation is derived and the optimal Lagrange multipliers are considered as the solutions of the system resulting by setting the gradient of the new objective function to zero. Once the Lagrange multipliers are found, several underlying conditions may be used to further compute the coefficients of the hyperplane. Existing software programs, however, only output the class of a test case directly after the black-box training; no formula describing the relations among indicators can be visualized, at least not straightforwardly. However, (3) may be alternatively plainly addressed by the adaptable and general algorithms of evolutionary computation. Based on principles of evolution and hereditary [7], EAs are able to inherently and unrestrainedly determine the coefficients that lead to an optimal separation into classes. The novel evolutionary technique constructed as an alternative of the SVM architecture thus adopts the learning strategy of the latter but aims to simplify and generalize its training, by offering a transparent substitute optimization. Contrary to the canonical technique, the evolutionary approach can at all times explicitly acquire the coefficients of the decision function, without any further constraints. Moreover, in order to converge, the evolutionary method does not require the positive (semi-)definition properties for kernels within nonlinear learning.

The present EA methodology addresses the primal optimization problem, however there are other evolutionary computation approaches that either solve the dual problem [16] or target further SVM-related issues like kernel evolution [12], parameter approximation [10] or selection of best features [13].

ESVMs [23, 24, 25] maintain a population of potential solutions. Each such individual encodes a candidate array of $\mathbf{w}$ and b and the whole set interacts towards the creation of enhanced solutions and the survival of the fittest against the training examples. After a number of generations, the EA converges to an optimal solution, which represents the best decision hyperplane that separates samples from divergent classes and is both accurate and general enough. Since the coefficients $\mathbf{w}$ and b are encoded in the structure of individuals, the equation of the evolving separating hyperplane is available at all times and especially useful at the end of the process. The evolved mathematical combination of chronic hepatitis C indicators may prove to be helpful to understand the weight and interaction of each medical attribute on the liver fibrosis grade. The accuracy of prediction of the model is computed using the available test samples, while the formula is stored for future reference and employment when a new case appears.

The considered EA features the following specific elements and behavior. An individual c is represented as an array of the coefficients of the hyperplane, $\mathbf{w}$ and b (4). Individuals are randomly generated, such that $w_i \in [-1, 1], i = 1, 2, ..., n$ (recall that n is the number of features of a sample), and $b \in [-1, 1]$.

$$c = (w_1, ..., w_n, b). \tag{4}$$

The fitness expression derives from the objective function of the optimization problem (3) and is subject to minimization. Constraints are handled by penalizing the infeasible individuals through appointing a function $t : R \to R$ which returns the value of the argument, if negative, and zero otherwise. Consequently, the expression of the fitness function for determining the optimal coefficients of the decision hyperplane is defined as in (5).

$$f(\mathbf{w}, b) = \|\mathbf{w}\|^2 + C \sum_{i=1}^{m} \xi_i + \sum_{i=1}^{m} [t(y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle - b) - 1 + \xi_i)]^2. \quad (5)$$

Additionally, all deviations ξ_i, $i = 1, 2, ..., m$ (m is the number of samples), are computed [2] in order to be referred in the fitness function. If the sign of the deviation equals that of the class, the corresponding $\xi_i = 0$; else, the (normalized) absolute deviation is returned as the indicator for error. Normalization may be necessary as the sum of deviations is added to the expression of the fitness function. As a consequence, in the early generations, when the generated coefficients lead to high deviations, their sum, considered from 1 to the number of training samples, takes over the whole fitness value and the evolutionary process is driven off the course to the optimum.

As the fittest coefficients of the separating hyperplane, $\mathbf{w}^{opt}$ and b^{opt}, are found, the target for a new, unseen test data instance $\mathbf{x}_i^{'}$ can be determined directly following the function in (6).

$$f(\mathbf{x}) = \langle \mathbf{w}^{opt}, \mathbf{x}_i^{'} \rangle - b^{opt}, i = 1, 2, ..., p. \quad (6)$$

If the classification task is intrinsically binary, then the class is determined from the sign of the function as positive or negative. If the problem has several ($k > 2$) classes, then a voting mechanism, based on the values of the current sample as a parameter of the different resulting decision functions, is applied. In the *one-against-all* mechanism that is chosen to act on the present prediction challenge, a distinct ESVM is created to separate one fibrosis grade from the united others. Each ESVM algorithm determines the optimal expression of a corresponding decision function (6). In order to establish the label for a test sample, the values of the k functions for that example are compared and the highest estimate is taken to point to the class that triggered that ESVM. The accuracy of the ESVM classifier is calculated by counting the percent of correctly appointed test cases.

6. Rules versus Formulas for Fibrosis Staging

On the one hand, experimentation targets to obtain the ESVM model for the separation of the chronic hepatitis C patients into the five distinct classes of fibrosis in the additional attempt to provide the formula of indicators interaction for each outcome as opposed to the others. On the other hand, a hill climbing algorithm is initially used to dynamically pick the proper attributes from the data set and then the CCEA is employed to classify the selected data into the same five different stages. Our task is to put the two approaches against each other with the aim of an objective comparison in terms of both accuracy and further support for the medical decision making.

6.1. Experimental setup. Both methodologies obey the following conditions:

- The 722 patient examples available for fibrosis degree determination are split into 542 training and 180 test samples, in the proportion 75-25% that is customary in machine learning experimentation. Samples are normalized.
- Both the ESVM and the CCEA are run 30 times on the randomly generated training and test sets of the established percentage, for reasons of both statistical validity and mandatory analysis within EA trials. The reported performance is computed as the average of the accuracy results in 30 applications of the

algorithm on the test samples. Therefore, we employ repeated random subsampling cross-validation.
- The values for the variables to be tuned are manually generated, therefore showing the ease of parameterization of both algorithms.
- We employ common operators for real encoding, i.e., intermediate recombination and mutation with normal perturbation [7]; selection is chosen as the well performing binary tournament type.

The hill climber starts from a randomly generated binary configuration of 24 genes (the number of attributes from the data set), the indicators that have a corresponding value of 1 are considered as selected and the CCEA is applied 30 times to the data set referring only the chosen features. The average accuracy obtained over the 30 repeats represents the fitness evaluation of the hill climber. In each of the 30 runs, the training and test sets are randomly chosen in order to have a more objective evaluation.

Mutation is then applied and the genes values can be flipped depending on the probability set for this purpose. The generated offspring is evaluated and, if fitter, it replaces the parent hill climber. The process continues and, if there is no improvement in fitness for a number of iterations (20), a new individual is generated and the process restarts. A fixed budget of fitness evaluations is set for the hill climber as a stop condition for the algorithm (1000). The mutation probability is decided to be small (0.1) in order to have only slight steps from one configuration to another (about 2-3 bits are changed when mutation occurs).

Each involved CCEA population is set to a size of 50 and, since there exist five classes, there are 250 individuals evolved overall. A high probability is set for recombination (0.9) with the purpose of bringing homogeneity within each species, while for the mutation probability a small value is chosen (0.2) in order to gradually explore the search space and avoid the rapid change of the entire genotype. As the values of the genes are set between 0 and 1, the mutation strength is set to a small number that allows the search to perform fine tuning (0.1). A number of 80 generations is set as a stop condition, as it has been observed during pre-experimentation that it provides sufficient time to reach an optimum.

As for the parameters of the ESVM, the population of potential hyperplanes has a size of 150. Probabilities of recombination (0.2) and mutation (0.2) have small values for a slow change in individual structure, while mutation strength (0.9) makes large jumps for the selected genes. The stop condition for the algorithm is set to 300 generations. Following pre-experimental conclusions, the SVM specific variables consist of a polynomial kernel of degree 1 and a penalty for errors $C = 1$.

The obtained results and specific outcome behavior of the two approaches for fibrosis staging are shown separately in the next subsections.

6.2. Results of cooperative coevolution for classification. Depending on the selected attributes, the best accuracy result obtained as the average over 30 repeated runs of random cross-validation of the CCEA is of 62.11% correctly classified patients. Over the 1000 fitness evaluations of the hill climber, the average accuracy is of 55.93%, while the worst test accuracy is of 47.92%. The individual that yields the best obtained accuracy only selected 9 attributes out of the 24 available and these are the following: stiffness, triglycerides, HDL cholesterol, aspartate aminotransferase, gama glutamyl transpeptidase, alkaline phosphatase, prothrombin index, prolonged activated partial thromboplastin time and hematocrit.

		Actual				
		F0	F1	F2	F3	F4
Predicted	F0	**0**	0	2	0	0
	F1	7	**59**	27	7	1
	F2	0	0	**4**	0	8
	F3	0	0	0	**8**	8
	F4	0	0	3	2	**46**

TABLE 1. Example of a confusion matrix of CCEA for differentiating the correct labeling from the misclassifications: predicted outcomes on the rows, actual classes on the columns, number of correct predictions on the diagonal.

As concerns the quality of the results, a confusion matrix is outlined in Table 1, thus illustrating how far the misclassified samples are from the actual classes and which degrees are better interpreted. It is taken from a run that achieved an accuracy of 65% on the test set. The correctly classified samples are the ones on the principal diagonal. The test set comprises of 180 samples, all from the F1 class are correctly classified, but, unfortunately many from F2 are also labeled as F1. Most of the data are split between levels F1 (59) and F4 (62) and the CCEA basically concentrates on recognizing these two classes as good as possible. However, it has to be underlined that, in each run, the training and test data are different and the confusion matrices may change considerably.

Naturally, besides the discussed configuration, there are several others found by the hill climbing algorithm which also produce good results. It is very interesting to observe that there are not always the same attributes that are selected; the hill climbing algorithm rather discovers sets of features that in connection perform better. However, there are some attributes that are included more often into many successful configurations. Overall, among the attributes that seem more decisive are the following: stiffness, sex, cholesterol, glycemia, prothrombin index and alkaline phosphatase. However, the attributes were not necessarily considered in this combination and there are others that immediately follow in importance, like prolonged activated partial thromboplastin time or haematids; the enumeration could continue in the order of weight, according to the automated artificial intelligence composed methodology used in the current paper. It has to be underlined however that the most significant feature, the one that has been chosen in most of the successful combinations, was the liver stiffness indicator, fact that is also acknowledged by the medical experts. What is finally important about the approach is that, beside establishing the most important attributes, the CCEA also provides thresholds for the values of these remaining indicators.

The computational effort has also been investigated, in order to fully reach a comparison between the two classifiers. The two approaches are mechanically very different, still in order to bring experimentation at this level somewhat closer, we have run the CCEA only once when evaluating the hill-climber. The necessary number of fitness evaluations was found to be 1880000, while the computational time reached 1288 seconds. However, it must be remembered that the coevolutionary approach additionally embeds a preprocessing hill-climber which, naturally, besides its feature selection advantage also leads to increased computations.

		Actual				
		F0	F1	F2	F3	F4
Predicted	$F0$	**0**	0	0	0	0
	$F1$	7	**63**	32	9	1
	$F2$	1	1	**2**	1	0
	$F3$	0	0	4	**4**	3
	$F4$	0	0	2	2	**48**

TABLE 2. Example of a resulting confusion matrix of the ESVM for the 5-degree fibrosis differentiation: Rows exhibit predicted outcomes, columns show actual classes. The diagonal outlines the amount of correctly classified samples, while each intersection of coordinates signals the number of misclassified cases.

6.3. Results of evolutionary support vector machines. The average over the 30 performed trials of random cross-validation resulted in a test accuracy of 62.03%. The distribution of correctly predicted versus misclassified cases per class can be visualized in the confusion matrix of Table 2, in an example run that obtained 65% accuracy.

The solution provided by the ESVM is also important for the capability to output the formula for the differentiation of each class from all the others. The formulas bring evidence of the high significance of the Fibroscan-derived stiffness for classes $F0$, $F1$ and $F4$, while the usual medical exams exhibit different smaller influences as each class is concerned. Once the expressions of the relationships between the medical indicators and each fibrosis degree are found as in this example, the labeling of any new case can be instantly conducted: Its values for every attribute ($A1$-$A24$) are introduced in each stored formula and the maximum obtained number indicates the predicted grade from whose expression it was computed.

As regards the computational effort of the ESVM, one run is terminated in 42.3 seconds and needs 248179 fitness evaluations.

7. Conclusions

It can be observed that both approaches that are undertaken for offering support for liver fibrosis degree differentiation behave similarly as concerns the accuracy of prediction. Each of them may be therefore chosen to be employed for the computationally aided staging, depending on the preference for the accompanying helpful information. The ESVM and CCEA with hill-climbing each possess unique features in this respect that are summarized below:

- The hill-climber inside the CCEA is additionally able to select the most important indicators of fibrosis. This proves to be important for the given task, as learning can concentrate only on the most influential features from all given attributes.
- The ESVM inherits the learning roots from SVMs and can thus circumvent the curse of dimensionality for highly multidimensional data sets.
- The ESVM is able to provide a formula of the relationship between indicators and fibrosis degree. This is highly significant for an immediate testing of new cases and for understanding relationships between medical variables that may be

difficult to grasp even for experienced physicians. The weights also offer some insight on the importance of every indicator for the staging process.
- The CCEA can provide the thresholds under which the previously considered essential attributes make the distinction between one fibrosis level and the others.
- Although the CCEA needs more computational effort, this is explainable for its preprocessing abilities.

Acknowledgements

Present work was supported by the research grant No. 41071/2007, Diagnostic algorithm for the stage and prediction of evolution of hepatic fibrosis using non-invasive ultrasound techniques, optimized by stochastic analysis and image analysis (SONOFIBROCAST), financed by the Romanian Ministry of Education and Research and the National Authority for Scientific Research (MEdC & ANCS).

The authors would like to thank the editor and the reviewers for their helpful comments and suggestions that brought insight and improvement to the paper during the revision.

References

[1] S. Belciug, M. Lupsor and R. Badea, Features selection approach for non-invasive evaluation of liver fibrosis, *Annals of the University of Craiova, Mathematics and Computer Science Series* **35** (2008), 15–20.

[2] R.A. Bosch and J.A. Smith, Separating hyperplanes and the authorship of the disputed federalist papers, *Amer Math Mon* **105** (1998), no. 7, 601-608.

[3] A.A. Bravo, S.G. Sheth and S. Chopra, Liver biopsy, *N Engl J Med* **344** (2001), 495–500.

[4] M.V. Butz and S.W. Wilson, An Algorithmic Description of XCS. *Soft Comput* **6** (2002), 144–153.

[5] J.F. Cadranel, P. Rufat and F. Degos, Practices of liver biopsy in France: results of a prospective nationwide survey. For the Group of Epidemiology of the French Association for the Study of the Liver (AFEF), *Hepatology* **32** (2000), 477–481.

[6] L. Castera, J. Vergniol and J. Foucher, Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C, *Gastroenterology* **128** (2005), 343–350.

[7] A.E. Eiben and J.E. Smith, *Introduction to Evolutionary Computing*, Springer-Verlag, Berlin, Heildelberg, New York, 2003.

[8] J. Foucher, E. Chanteloup, J. Vergniol, L. Castera, B. Le Bail, X. Adhoute et al, Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study, *Gut* **55** (2006), 403–408.

[9] M. Gorunescu, S. Belciug, A.-B. Salem, M. Lupsor, R. Badea and H. Stefanescu, A machine learning-based diagnosis for liver diseases using the Fibroscan medical ultrasound technique, *4rd ACM International Conference on Intelligent Computing and Information Systems ICICIS09* (2009), 243–248.

[10] F. Friedrichs and C. Igel, Evolutionary tuning of multiple svm parameters. In: Verleysen M (ed), *12th European Symposium on Artificial Neural Networks* (2004), 519-524.

[11] J.H. Holland, Escaping brittleness: the possibilities of general-purpose learning algorithms applied to parallel rule-based systems, R.S. Michalski, J.G. Carbonell, T.M. Mitchell (Eds.), *Machine learning, an Artificial Intelligence Approach* (1986).

[12] T. Howley and M.G. Madden, The genetic kernel support vector machine: Description and evaluation, *Artif Intell Rev* **24** (2005), no. 3-4, 379-395 .

[13] O. Kramer and T. Hein, Stochastic Feature Selection in Support Vector Machine Based Instrument Recognition, *32nd Annual Conference on Artificial Intelligence KI 2009: Advances in Artificial Intelligence* **5803** (2009), 727–734.

[14] M. Lupsor, R. Badea, H. Stefanescu, M. Grigorescu, Z. Sparchez, A. Serban et al, Analysis of histopathological changes that influence liver stiffness in chronic hepatitis C. Results from a cohort of 324 patients, *J Gastrointestin Liver Dis* (2009) **17**, no. 2, 155–163.

[15] Y. Marinakis, G. Dounias and J. Jantzen, Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification, *Comput. Biol. Med.* **39** (2009), no. 1, 69–78.
[16] I. Mierswa, Making indefinite kernel learning practical, Technical Report TR 41-06, University of Dortmund (2006).
[17] NIH Consensus Statement on Management of Hepatitis C, *NIH Consens State Sci Statements 2002* **19** (2002), 1–46.
[18] M.A. Potter and K. de Jong, Cooperative coevolution: An architecture for evolving coadapted subcomponents, *Evol. Comput.* **8** (2000), no. 1, 1–29.
[19] L. Sandrin, B. Fourquet, J.M. Hasquenoph, S. Yon, C. Fournier, F. Mal et al, Transient elastography: a new noninvasive method for assessment of hepatic fibrosis, *Ultrasound Med Biol* **29** (2003), 1705–1713.
[20] Q. Shen, Z. Meia and B.-X. Ye, Simultaneous genes and training samples selection by modified particle swarm optimization for gene expression data classification, *Comput. Biol. Med.* **39** (2009), no. 7, 646–649.
[21] C. Stoean, Mike Preuss, D. Dumitrescu and R. Stoean, Cooperative Evolution of Rules for Classification, *IEEE Intl. Conf. on Symbolic and Numeric Algorithms for Scientific Computing* (2006), 317–322.
[22] C. Stoean and R. Stoean, Evolution of Cooperating Classification Rules with an Archiving Strategy to Underpin Collaboration, H.-N. Teodorescu, J. Watada, L.C. Jain (Eds.), *Intelligent Systems and Technologies, Studies in Computational Intelligence* **217** (2009), 47–65.
[23] R. Stoean, M. Preuss, C. Stoean, E. El-Darzi and D. Dumitrescu, Support Vector Machine Learning with an Evolutionary Engine, *J Oper Res Soc* **60** (2009), no. 8, 1116–1122.
[24] R. Stoean, M. Preuss, C. Stoean, E. El-Darzi and D. Dumitrescu, An Evolutionary Approximation for the Coefficients of Decision Functions within a Support Vector Machine Learning Strategy, A. E. Hassanien, A. Abraham (Eds.), *Foundations on Computational Intelligence, Studies in Computational Intelligence* (2009), 315–347.
[25] R. Stoean and D. Dumitrescu, Linear Evolutionary Support Vector Machines for Separable Training Data, *Annals of the University of Craiova, Mathematics and Computer Science Series* **33** (2006), 141–146.
[26] C. Stoean and D. Dumitrescu, Elitist Generational Genetic Chromodynamics as a Learning Classifier System, *Annals of University of Craiova, Mathematics and Computer Science Series* **33** (2006), 132–140.
[27] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, Berlin, Heildelberg, New York, 1995.
[28] S.W. Wilson, Classifier Fitness Based on Accuracy. *Evol Comput* **3** (1995), no. 2, 149–176.
[29] M. Ziol, A. Handra-Luca and A. Kettaneh, Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C, *Hepatology* **41** (2005), 48–54.

(Ruxandra Stoean, Catalin Stoean) Department of Computer Science, University of Craiova,
A. I. Cuza Str., No. 13, 200585, Craiova, Romania
E-mail address: {ruxandra.stoean, catalin.stoean}@inf.ucv.ro

(Monica Lupsor, Horia Stefanescu, Radu Badea) Department of Ultrasonography, 3rd Medical Clinic,
University of Medicine and Pharmacy Cluj - Napoca,
Croitorilor Str., No. 19-21, 400162, Cluj-Napoca, Romania
E-mail address: mmlupsor@yahoo.com, {hstefanescu, rbadea}@umfcluj.ro