Hybrid Multilogistic Regression by Means of Evolutionary Radial Basis Functions: Application to Precision Agriculture*

P.A. Gutiérrez[1,**], C. Hervás-Martínez[1], J.C. Fernández[1], and F. López-Granados[2]

[1] Department of Computer Science and Numerical Analysis, University of Cordoba, Spain
Tel.: +34 957 218 349; Fax: +34 957 218 630
pagutierrez@uco.es
[2] Institute of Sustainable Agriculture, CSIC, Córdoba, Spain

Abstract. In this paper, a previously defined hybrid multilogistic regression model is extended and applied to a precision agriculture problem. This model is based on a prediction function which is a combination of the initial covariates of the problem and the hidden neurons of an Artificial Neural Network (ANN). Several statistical and soft computing techniques have been applied for determining these models such as logistic regression, ANNs and Evolutionary Algorithms (EAs). This paper proposes the use of Radial Basis Functions (RBFs) transformations for this model. The estimation of the coefficients of the model is basically carried out in two phases. First, the number of RBFs and the radii and centers' vector are determined by means of an EA. Afterwards, the new RBF nonlinear transformations obtained for the best individual in the last generation are added to the covariate space. Finally, a maximum likelihood optimization method determines the rest of the coefficients of the multilogistic regression model. In order to determine the performance of this approach, it has been applied to a problem of discriminating cover crops in olive orchards affected by its phenological stage using their spectral signatures obtained with a high-resolution field spectroradiometer. The empirical results for this complex real agronomical problem and the corresponding Dunnet statistical test carried out show that the proposed model is very promising in terms of classification accuracy and number of wavelengths used by the classifier.

Keywords: evolutionary neural networks, radial basis functions, multilogistic regression, multiclassification, spectral signatures, cover crops.

1 Introduction

The traditional statistical approach to pattern recognition is a natural application of the Bayesian decision theory. In order to achieve a correct classification, either the

* This work has been partially subsidized by the TIN2008-06681-C06-03 project of the Spanish Inter-Ministerial Commission of Science and Technology (MICYT), FEDER funds and the P08-TIC-3745 project of the "Junta de Andalucía" (Spain).
** Corresponding author.

E. Corchado et al. (Eds.): HAIS 2009, LNAI 5572, pp. 244–251, 2009.

class densities or the posterior class probability distributions are required. This distinction results in two types of techniques: those that model the class densities, such as kernel methods (Gaussian and Gaussian mixture classifiers), and those that model the posterior probabilities directly, such as classification trees, linear and nonparametric Logistic Regression (LR), multilayer neural networks with sigmoidal or product units (global functions), or Radial Basis Functions (RBFs, local functions)… [1] These last classifiers simultaneously estimate the minimum-error Bayesian *a posteriori* probabilities for all classes. However, in spite of the great number of these techniques developed to solve classification problems, there is no optimum methodology or technique to solve specific problems. This point has encouraged the combination or hybridization of different types of classifiers [2]. The LR model has been widely used in statistics for many years and has recently been the object of extensive study in the machine learning community. In general the prediction function of this model is linear. Hastie, Tibshirani and Friedman [3] proposed the use of non-linear functions in the prediction function for solving complex real problems of classification. Based on that proposal, a recently multilogistic methodology [4] hybridizes a linear model and nonlinear Product-Unit Neural Network (PUNN) models for binary and multiclassification, LIPUNN. Following these ideas, we present in this paper an extension and a different aspect of the methodology by designing a multilogistic learning algorithm which combines different elements such as LR, RBF Neural Networks (RBFNNs), an Evolutionary Programming (EP) algorithm and a specific multilogistic algorithm for simplifying the models obtained. The methodology involves augmenting/replacing the input covariate vector with RBFs, which are transformations of the input variables, and then using linear models in this new space of derived covariates. The approach is named Logistic Initial and Radial Basis Function regression (LIRBF). RBFNNs, as an alternative to multilayer perceptrons, have been found to be very helpful to many engineering problems because: (1) they are universal approximators [5]; (2) they have more compact topology than other neural networks; and (3) their learning speed is fast because of their locally tuned neurons.

The basic differences of LIPUNN methodology and the LIRBF presented here are: (1) in this paper we are using RBFs hidden units (which are *local* approximators), while in [4] we used product unit basis functions (which are *global* approximators); ii) the initialization of the population procedure is different as it incorporates a standard *k*-means process for the initialization of the centers of the Gaussian RBFs; iii) the LR algorithms used in this paper are more advanced and perform an automatic structural simplification of the model (MLogistic [6] and SLogistic [7]); iv) a more exhaustive 10-fold cross-validation with 10 repetitions per fold has been performed.

This methodology is applied to a real problem of discriminating cover crops in olive orchards as affected by its phenological stage using a high-resolution field spectroradiometer. Olive (Olea europaea L.) is the main perennial Spanish crop where soil management in olive orchards is mainly based on intensive and tillage operations. These operations have a great relevancy in terms of increase of atmospheric CO_2, desertification, erosion and land degradation [8]. Due to these negative environmental impacts, the European Union only subsidizes those cropping systems which require the implementation of conservation techniques, which mainly consist of altering as little as possible the natural soil and protecting it with cover crops. Remotely sensed

data may offer the ability to efficiently identify and map crops and cropping methods over large areas [9]. These techniques can signify lower costs, faster work and better reliability than ground visits. The accuracy of the thematic map is extremely important since this map could be used to help the administrative follow-up for making the decision on conceding or not the subsidy. To map olive trees and cover crops, it is necessary that suitable differences exist in spectral reflectance among them and bare soil and this is the main objective of this study. The paper is aimed to establish the misclassification percentage of the spectral signature discrimination problem and to validate the classification accuracy of the methodologies presented by using a 10-fold cross-validation procedure. Three models were tested: a) Evolutionary RBFNNs, ERBFNN, b) MultiLogistic using Initial and RBF covariates (without structural simplification, MLIRBF), and c) SimpleLogistic using Initial and RBF covariates (with structural simplification, SLIRBF).

The rest of the paper is structured as follows: The LIRBF model is defined in Section 2. Section 3 presents the different steps of the learning algorithm. Section 4 explains the experimental design applied for the real agronomical problem. Finally, the work is summarized and the conclusions are drawn in Section 5.

2 LIRBF Model

In classification problems, measurements $\mathbf{x}_i$, $i = 1, 2, ..., k$, are taken on a single individual (or object), and the individuals are to be classified into one of J classes on the basis of these measurements. It is assumed that J is finite, and the measurements $\mathbf{x}_i$ are random observations from these classes. Based on the training sample $D = \{(\mathbf{x}_n, \mathbf{y}_n);\ n = 1, ..., N\}$, where $\mathbf{x}_n = (x_{1n}, ..., x_{kn})$ is the vector of measurements taking values in $\Omega \subset R^k$, and $\mathbf{y_n}$ is the class level of the n-th individual, we wish to find a decision function $C : \Omega \to \{1, 2, ..., J\}$ for classifying the individuals. A misclassification occurs when the decision rule C assigns an individual of the training sample to a class j when it is actually coming from a class $l \neq j$. To evaluate the performance of the classifiers we define the Correct Classification Rate (CCR) by $CCR = \frac{1}{N}\sum_{n=1}^{N} I(C(\mathbf{x}_n) = \mathbf{y}_n)$, where $I(.)$ is the zero-one loss function. We adopt the common technique of representing the class levels using a "1-of-J" encoding vector $\mathbf{y} = \left(y^{(1)}, y^{(2)}, ..., y^{(J)}\right)$, such as $y^{(l)} = 1$ if $\mathbf{x}$ corresponds to an example belonging to class l and $y^{(l)} = 0$ otherwise. Usually it is assumed that the training data are independent and identically distributed samples from an unknown probability distribution.

Suppose that the conditional probability that $\mathbf{x}$ belongs to class l verifies: $p_l(\mathbf{x}) = p\left(y^{(l)} = 1 \middle| \mathbf{x}\right) > 0,\ l = 1, 2, ..., J,\ \ \mathbf{x} \in \Omega$. Under a multinomial logistic regression, the probability that $\mathbf{x}$ belongs to class l is based in the equation:

$$p_l(\mathbf{x},\boldsymbol{\beta}_l)=\frac{\exp(f_l(\mathbf{x},\boldsymbol{\beta}_l))}{\sum_{l=1}^{J}\exp(f_l(\mathbf{x},\boldsymbol{\beta}_l))},\quad l=1,\ldots,J, \tag{1}$$

where $\boldsymbol{\beta}_l=(\beta_{l,0},\beta_{l,1},\ldots,\beta_{l,m})$, the $x_0=1$ value has been added to the input covariates vector $\mathbf{x}$ and the prediction function $f_l(\mathbf{x},\boldsymbol{\beta}_l)=\boldsymbol{\beta}_l^{\mathrm{T}}\mathbf{x}$ is linear in the covariates. The vector components $\boldsymbol{\beta}_l$ are estimated from the training data set D. If we use the normalization probability axiom, we have that $f_J(\mathbf{x},\boldsymbol{\theta}_J)=0$ and, in this way, it is not necessary to estimate the parameters $\boldsymbol{\beta}_J$.

Our Logistic Regression model proposal is based on the combination of the standard Initial covariates and nonlinear RBF transformed covariates (LIRBF). The general expression of the model is given by the following equation:

$$f_l(\mathbf{x},\boldsymbol{\theta}_l)=\boldsymbol{\alpha}_l^{\mathrm{T}}\mathbf{x}+\boldsymbol{\beta}_l^{\mathrm{T}}\mathbf{B}(\mathbf{x},\mathbf{W}),\quad l=1,2,\ldots,J\text{-}1, \tag{2}$$

where $\mathbf{x}=(1,x_1,\ldots,x_k)$ and $\mathbf{B}(\mathbf{x},\mathbf{W})=(B_1(\mathbf{x},\mathbf{w_1}),\ldots,B_m(\mathbf{x},\mathbf{w_m}))$, with $B_j(\mathbf{x},\mathbf{w}_j)$ being a Gaussian RBF:

$$B_j(\mathbf{x},\mathbf{w}_j)=\exp\left(-\frac{\|\mathbf{x}-\mathbf{c}_j\|^2}{r_j^2}\right). \tag{3}$$

Then, the prediction function of the LIRBF model is:

$$f_l(\mathbf{x},\boldsymbol{\theta}_l)=\alpha_0^l+\sum_{i=1}^{k}\alpha_i^l x_i+\sum_{j=1}^{m}\beta_j^l\exp\left(-\frac{\|\mathbf{x}-\mathbf{c_j}\|^2}{r_j^2}\right),\quad l=1,2,\ldots,J-1, \tag{4}$$

Let $\boldsymbol{\theta}_l=(\boldsymbol{\alpha}^l,\boldsymbol{\beta}^l,\mathbf{W})$ be, where $\boldsymbol{\alpha}^l=(\alpha_0^l,\alpha_1^l,\ldots,\alpha_k^l)$, $\boldsymbol{\beta}^l=\left(\beta_1^l,\ldots,\beta_m^l\right)$ are the coefficients of the LR model and $\mathbf{W}=(\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_m)$ are the parameters of the RBFs, where $\mathbf{w}_j=(w_{j0},w_{j1},\ldots,w_{jk})$, $\mathbf{c}_j=(w_{j1},\ldots,w_{jk})$ is the centre of the j-th Gaussian RBF and $r_j=w_{j0}$ is the corresponding radius. The probability that $\mathbf{x}$ belongs to class l is given by substituting $f_l(\mathbf{x},\boldsymbol{\theta}_l)$ of Eq. 1 by the value of Eq. 4.

To perform the Maximum Likelihood (ML) estimation of $\boldsymbol{\theta}=(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2,\ldots,\boldsymbol{\theta}_{J\text{-}1})$, one can minimize the negative log-likelihood function:

$$L(\boldsymbol{\theta})=\frac{1}{N}\sum_{n=1}^{N}\left[-\sum_{l=1}^{J-1}y_n^{(l)}f_l(\mathbf{x}_n,\boldsymbol{\theta}_l)+\log\sum_{l=1}^{J-1}\exp f_l(\mathbf{x}_n,\boldsymbol{\theta}_l)\right]. \tag{5}$$

The classification rule coincides with the optimal Bayes' rule. In other words, an individual should be assigned to the class which has the maximum probability:

$$C(\mathbf{x})=\hat{l},\ \text{where}\ \hat{l}=\arg\max_l\, p_l(\mathbf{x},\boldsymbol{\theta}_l^*),\ \text{for}\ l=1,2,\ldots,J$$

The non-linearity of the model with respect to the $\boldsymbol{\theta}_l$ parameters and the indefinite character of the associated Hessian matrix do not recommend the use of gradient-based methods to maximize the log-likelihood function. Moreover, the optimal number of basis functions of the model (i.e. the RBFs in the RBFNN) is unknown. Thus, the estimation of the vector $\hat{\boldsymbol{\theta}}$ is carried out by means of a combination of an Evolutionary Programming (EP) algorithm and a standard ML optimization method.

3 Learning the LIRBF Model Coefficients

The process is structured in different steps. The first step obtains an RBFNN using and EP algorithm (designing its structure and obtaining the weights), and the second obtains the LIRBF model. The EP algorithm determines the number m of RBFs in the model, and the corresponding vector $\mathbf{W}$ of centres and radii. Once the basis functions have been determined by the EP algorithm, we consider a transformation of the input space by adding the nonlinear transformations given by these RBFs. The model is now linear in these new variables and the initial covariates. The remaining coefficient vector $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are calculated by the ML optimization method.

To apply evolutionary neural network techniques, we consider a population of RBFNN models. There are no connections between the nodes of a layer and none between the input and output layers either. The activation function of the j-th node in the hidden layer is given by Eq. 3. The activation function of the output node l is given by $g_l(\mathbf{x}, \boldsymbol{\beta}^l, \mathbf{W}) = \beta_0^l + \sum_{j=1}^{m} \beta_j^l B_j(\mathbf{x}, \mathbf{w}_j)$, where β_j^l is the weight of the connection between the hidden node j and the output node l. The transfer function of all output nodes is the identity function. The main objective of the algorithm is to design a RBFNN with the better structure and weights for the classification problem tackled. The search begins with an initial population of RBFNNs, to which a population-update algorithm is applied in each iteration. The algorithm shares many characteristics and properties with other previous algorithms [10, 11]. Individuals are subject to the operations of replication and mutation, but crossover is not used due to its potential disadvantages in evolving ANNs [10]. Although in this step a concrete value for the $\boldsymbol{\beta}$ vector is obtained, we only consider the estimated weight vector $\hat{\mathbf{W}}$, which builds the RBFs. The parameters of the EP algorithm are the following: the centres w_{ji} and the coefficients β_j are initialized in the $[-2, 2]$ interval and the radii in the $[0,1]$ interval. The number of RBF nodes is $m = 3$. The population size is $1{,}000$ and the maximum number of generations is 50.

In a second step, we consider a transformation of the input space, by including the nonlinear RBFs obtained for the best model in the last generation by the EP algorithm, that is, $z_1 = B_1(\mathbf{x}, \hat{\mathbf{w}}_1)$,..., $z_m = B_m(\mathbf{x}, \hat{\mathbf{w}}_m)$. Then, we minimize the negative log-likelihood function for N observations (Eq. 5). Now, the Hessian matrix associated to this expression using the new variables $x_1, x_2, ..., x_k, z_1, z_2, ..., z_m$ is semi-definite positive and exist a global optimum. In this final step, the MLogistic and

SLogistic algorithms have been used for obtaining the parameters matrix $\boldsymbol{\theta}_l = (\boldsymbol{\alpha}^l, \boldsymbol{\beta}^l, \mathbf{W})$, where the coefficients $\mathbf{W}$ are given by the EP algorithm. Both algorithms are available in the WEKA machine learning workbench [12].

4 Experiments

The study was conducted in Andalusia, southern Spain, in a location named "Cortijo del Rey", in early spring and early summer. Forty spectral signatures of live cover crop, twenty of dead cover crops, ten of olive trees, and ten of bare soil were taken on spring and summer. Measurements were collected using an ASD Handheld FieldSpec Spectroradiometer. The hyperspectral range was between 400 and 900 nm. Then, these collected hyperspectral measurements were reduced and averaged to represent 25 nm-wide measurements between 575 and 725 nm, as previous experiments determined that these seven measurements are sufficient for the determination of the original spectral curves. The experimental design for the two phenological stages was conducted using a 10-fold cross-validation procedure with 10 repetitions for fold. The experiments were carried out using a software package developed in JAVA by the authors, as an extension of the JCLEC framework (http://jclec.sourceforge.net/) [13].

Table 1. Statistical results (Mean±Standard Deviation) in CCR_G for the different classifiers

	(1)	(2)	(3)	(4)	(5)
Cortijo spring	91.63±9.74	95.25±7.70	95.25±7.70	87.50±11.51	84.50±11.53
Cortijo summer	89.60±14.06	91.00±12.19	91.40±11.81	77.40±14.95	72.00±12.39
	(6)	(7)	(8)	(9)	(10)
Cortijo spring	91.00±8.90	91.25±8.79	86.00±5.11	92.13±7.88	**96.13±7.04**
Cortijo summer	80.00±14.49	79.60±14.77	86.60±15.06	91.40±10.73	**92.80±10.45**

(1) MLogistic; (2) SLogistic; (3) LMT; (4) C4.5; (5) NBTree; (6) ABoost10; (7) ABoost100; (8) ERBFNN; (9) MLIRBF; (10) SLIRBF.

We compare our LIRBF approaches (without structural simplification, MLIRBF, and with structural simplification, SLIRBF) to recent algorithms [14]: LR without characteristic selection (MLogistic); LR with characteristics selection (SLogistic); Logistic Model Trees, LMT [14]; C4.5 trees; Naïve Bayesian Trees (NBTree); boosted C4.5 trees using AdaBoost.M1 with 10 and 100 boosting iterations (ABoost(10) and ABoost(100)) and our Evolutionary RBFNN algorithm (ERBFNN). Table 1 shows the average and standard deviation of the Correct Classification Rates for the generalization set (CCR_G) of the learners generated by the previously mentioned algorithms. A descriptive analysis of the results leads to the following remark: the SLIRBF method obtains the best mean CCR_G ($CCR_G = 96.13\%$ in spring and 92.80% in summer). To ascertain if there are significant differences in mean CCR_G, and, under the hypothesis of normality of the results, we statistically compare all the classifiers. Our decisions on the comparisons will be determined by using confidence intervals for mean differences. We choose the Dunnett [15] procedure to build these intervals; it is a multiple comparison test with a control method,

the SLIRBF methodology, to identify which of these algorithms are worse (Win, W or 1), similar (Draw, D or 0) or better (Lose, L or -1) than SLIRBF for each phonological stage. The results can be seen in table 2, the Dunnett test applied for a 95% confidence level. We conclude that, SLIRBF obtains significant better results than MLogistic, C4.5, NBTree, ABoost10, ABoost100, ERBFNN and MLIRBF for $\alpha =$ 0.05 in spring. The differences are significant with respect to C4.5, NBTree, ABoost10, ABoost100 and ERBFNN in summer. Computational requirements for training MLIRBF models were nearly insignificant once the ERBFNN models are built. If we need to produce a very accurate thematic map ready to be used for decision-making procedures by administrations, the criteria for selecting the models should be based on the accuracy of the classifications and these more sophisticated and accurate models would be highly recommended.

Table 2. Results of the Dunnet test comparing SLIRBF(10) to MLogistic(1), SLogistic(2), LMT(3), C4.5(4), NBTree(5), ABoost10(6), ABoost100(7), ERBFNN(8) and MLIRBF(9)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	W	D	L
C. spring	1	0	0	1	1	1	1	1	1	-	7	2	0
C. summer	0	0	0	1	1	1	1	1	0	-	5	4	0

Finally, Table 3 presents the best models obtained in each phenological stage. In spring, the best model uses three RBFs centred in the λ_{625}*, λ_{650}* and λ_{725}* wavelengths, where the form of the spectral signatures is non linear. In summer, the discriminant functions are very simple with only two RBFs located around the λ_{575}* and λ_{6oo}* wavelengths. Therefore, we have very simple classifiers with high CCR_G values, which allow reducing the cost of future remote sensed images, because a low number of wavelengths will be necessary.

Table 3. Discrimination equations provided by the best SLIRBF models in spring and summer for the classification of the spectral signatures

Cortijo spring. Discriminant equations
F_1= -18.83-17.64(λ_{575}*)-1.80(λ_{675}*) + 9.96 (λ_{700}*)+19.43 (RBF_1)
F_2= -16.53-10.58 (λ_{575}*)+4.10((λ_{675}*) + 9.02((λ_{700}*)-4.77(λ_{725}*)+4.71(RBF_1) +21.68(RBF_3)
F_3= 7.44 -4.83 ((λ_{575}*)+9.02 ((λ_{700}*)+4.71 (RBF_1)-30.17 (RBF_2)
<u>RBFs</u>
RBF_1= (exp(-0.5((((λ_{725}* -1.18)^2)$^{0.5}$/ (1.28)^2))
RBF_2= (exp(-0.5*(((λ_{625}* +0.67)^2)$^{0.5}$/ (0.93)^2))
RBF_3= (exp(-0.5*(((λ_{625}* +0.39) ^{2} + (λ_{650}* +0.21) ^{2}) $^{0.5}$/ (1.13) ^{2}))
Cortijo summer. Discriminant equations
F_1= 0.68 +1.53 ((λ_{675}*)+ 1.28 (λ_{725}*)
F_2= -1.19+1.53 * ((λ_{675}*)+ 5.57 (RBF_1)
F_3= -1.44+1.53 * (λ_{675}*)+4.30 (RBF_2)
<u>RBFs</u>
RBF_1= (exp(-0.5 (((λ_{600}* +0.72)^2)$^{0.5}$ / (0.66)^2))
RBF_2= (exp(-0.5(((((λ_{575}*-1.23)^2)$^{0.5}$/ (0.87) ^{2}))
F_i= log odd p_i, λ_i*$\in$(-2,2)

5 Conclusions

The end goal of this work was to assess the potential of the methodologies presented for discriminating cover crops in two different phenological stages (spring and summer). From the analysis of the results obtained, several conclusions can be drawn. The covariate selection process incorporated in the SLIRBF methodology is necessary. In both phonological states, our methodology obtains mean CCR_G values significantly better or similar than the rest of methodologies. From the observation of the best models, it can be concluded that the models presented are able to reduce substantially the covariates but with a high classification accuracy. Thus, the interpretability of the models is enhanced and the cost of future remote sensed images is reduced. The obtained results would provide information to program the suitable wavelengths of airborne hyperspectral sensors such as Compact Airborne Spectrographic Imager (CASI) for administrative follow-up of agro-environmental measures in olive orchards under conservation agriculture.

References

1. Hastie, T.J., Tibshirani, R.J.: Nonparametric regression and classification. From statistics to neural networks. Nato ASI Series 136, 70–82
2. Major, R.L., Ragsdale, C.T.: Aggregating expert predictions in a networked environment. Computers & Operations Research 28(12), 1231–1244 (2001)
3. Hastie, T.J., Tibshirani, R.J., Friedman, J.: The elements of statistical learning. In: Data mining, inference and prediction. Springer, Heidelberg (2001)
4. Hervás-Martínez, C., Martínez-Estudillo, F., Carbonero, M.: Multilogistic Regression by means of Evolutionary Product-Unit Neural Networks. Neural Networks 21, 951–961 (2008)
5. Park, J., Sandberg, I.W.: Universal approximation using radial basis function networks. Neural Computation 3(2), 246–257 (1991)
6. le Cessie, S., van Houwelingen, J.: Ridge estimators in logistic regression. Applied Statistics 41(1), 191–201 (1992)
7. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression; a statistical view of boosting. The Annals of Statistics 38(2), 337–374 (2000)
8. Schlesinger, W.H.: Carbon sequestration in soils: some cautions amidst optimism. Agric. Ecosyst. Environ. 82, 121–127 (2000)
9. South, S., Qi, J., Lusch, D.P.: Optimal classification methods for mapping agricultural tillage practices. Remote Sensing of Environment 91, 90–97 (2004)
10. Angeline, P.J., Sauders, G.M., Pollack, J.B.: An evolutionary algorithm that constructs recurren neural networks. IEEE Transactions on Neural Nets 5, 54–65 (1994)
11. Martínez, F.J., Hervás-Martínez, C., Gutiérrez, P.A., Martínez-Estudillo, A.C.: Evolutionary product-unit neural networks classifiers. Neurocomputing 72(1-2), 548–561 (2008)
12. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. In: Data Management Systems, 2nd edn. Morgan Kaufmann, Elsevier (2005)
13. Ventura, S., Romero, C., Zafra, A., Delgado, J.A., Hervas, C.: JCLEC: a Java framework for evolutionary computation. Soft Computing 12(4), 381–392 (2008)
14. Landwehr, N., Hall, M., Frank, E.: Logistic model trees. Machine Learning 59(1-2), 161–205 (2005)
15. Hochberg, Y., Tamhane, A.: Multiple Comparison Procedure. J. Wiley and Sons, Chichester (1987)