Multiobjective Pareto Ordinal Classification for Predictive Microbiology

M. Cruz-Ramírez[1], J.C. Fernández[1], A. Valero[2],
P.A. Gutiérrez[1], and C. Hervás-Martínez[1]

[1] Dept. of Computer Science and Numerical Analysis, U. of Córdoba, Rabanales Campus,
Albert Einstein Building, 14071, Córdoba, Spain
{mcruz,jcfernandez,pagutierrez,chervas}@uco.es
[2] Dept. of Food Science and Technology, U. of Córdoba, Rabanales Campus,
Darwin Building, 14071, Córdoba, Spain
bt2vadia@uco.es

Abstract. This paper proposes the use of a Memetic Multiobjective Evolutionary Algorithm (MOEA) based on Pareto dominance to solve two ordinal classification problems in predictive microbiology. Ordinal classification problems are those ones where there is order between the classes because of the nature of the problem. Ordinal classification algorithms may take advantage of this situation to improve its classification. To guide the MOEA, two non-cooperative metrics have been used for ordinal classification: the Average of the Mean Absolute Error, and the Maximum Mean Absolute Error of all the classes. The MOEA uses an ordinal regression model with Artificial Neural Networks to classify the growth classes of microorganisms such as *Listeria monocytogenes* and *Staphylococcus aureus*.

1 Introduction

Knowing how different food product properties, environments and their history can influence the micro-flora developed when food is stored is an important first step towards forecasting its commercial shelf-life, alterations and safety. In order to be able to predict microbial behavior in each new situation and estimate its consequences with respect to the safety and quality of food, there has to be an exact definition of the food environment and how it will influence microbial growth and survival. The need to learn more about microbial behavior in limiting conditions that prevent growth could be met by using mathematical models. Binary logistic regression has frequently been applied to determine the probability of growth under a given set of conditions. However, microbial responses in limiting environmental conditions (i.e. low pH, temperature, water activity, etc.) are subject to several variable sources, often not experimentally controlled. This can result in biased growth/no growth estimations when the probability of growth approaches 0.5 [9].

One of the first published works related to the development of multiclassification models showed a high degree of accuracy when estimating growth/no growth boundaries of *S. aureus* [11]. However, although this paper provided a categorical classification into three classes, by adding new information about the probability of growth

V. Snasel et al. (Eds.): SOCO Models in Industrial & Environmental Appl., AISC 188, pp. 153–162.
springerlink.com

associated to stringent conditions, growth/no growth models can provide alternative and more accurate estimations.

For this reason, this paper addresses growth/no growth models from the perspective of ordinal classification into four classes. A classification problem occurs when an object needs to be assigned to a predefined group or class based on a number of observed attributes related to that object. Ordinal classification is the problem where the variable to be predicted is not of a numeric or nominal type but is instead ordinal, so that the categories have a logical order. In our case, the predictive microbiology problem to be solved has clearly ordinal behavior (see Section 3). Artificial Neural Networks (ANNs) [1] have been an important tool for classification since recent research activities identified them as a promising alternative to traditional classification methods such as logistic regression. On the other hand, Evolutionary Algorithms (EAs) [16] are global search heuristics and one of the main alternatives to local search algorithms for training ANNs. Obtaining ANN models using EAs is known as Evolutionary Artificial Neural Networks [16]. These methodologies maintain a population of ANNs that are subject to a series of transformations in the evolutionary process so as to obtain acceptable solutions to the problem.

Often a great number of objectives must be processed to obtain a viable solution to a problem, usually without any a priori knowledge of how the objectives interact with each other. This is known as a Multiobjective Optimization Problem, and the most popular methods are based on Pareto dominance. The training of ANNs by evolutionary Pareto-based algorithms is known as Multiobjective Evolutionary Artificial Neural Networks, and has been in use for the last few years in the resolving classification tasks [13]. Hybridization of intelligent techniques, coming from different computational intelligence areas, is a common solution, because of the growing awareness that such combinations frequently perform better than the individual techniques coming from computational intelligence [6].

This study deals with learning and the improvement in the generalization of classifiers designed using a MOEA with ANNs to determine growth limits in two important microorganisms in predictive microbiology, *Listeria monocytogenes* and *Staphylococcus aureus*.

The rest of the paper is organized as follows. Section 2 covers background materials, explaining the classification model used, the MOEA and the ordinal metrics used for guiding the MOEA. Section 3 presents the *L. monocytogenes* and *S. aureus* microorganisms. Section 4 shows the experimental design and results and finally Section 5 shows the conclusions obtained.

2 Background

2.1 Ordinal Model

The big problem in ordinal classification is that there is no notion of the precise distance between classes. The samples are labeled by a set of ranks with different categories given an order. Nominal classification algorithms can also be applied to prediction problems involving ordinal information but obviating the order of the classes. However, this process loses information that could improve the predictive ability of the classifier.

Although there are some other approaches for ordinal regression [14] (mainly based on reducing the problem to binary classification, or simplifying it to regression or cost-sensitive classification), the majority of proposals can be grouped under the term *threshold methods*. These methods are based on the idea that, in order to model ordinal ranking problems from a regression perspective, one can assume that some underlying real-valued outcomes exist, although they are unobservable. Consequently, two different things are estimated:

- A function $f(\mathbf{x})$ that predicts the real-valued outcomes and tries to discover the nature of the assumed underlying outcome.
- A threshold vector $\mathbf{b} \in \mathbb{R}^{J-1}$ to represent the intervals in the range of $f(\mathbf{x})$, where $b_1 \leq b_2 \leq \ldots \leq b_{Q-1}$ (possible different scales around different ranks).

We propose an adaption of the classical Proportional Odd Model (POM) model [18] for ANNs. Since we are using the POM model and ANNs, our proposal does not assure monotonicity. The POM model works based on two elements: the first one is a linear layer with only one node (see Fig. 1) whose inputs are the non-linear transformations of a first hidden layer. The task of this node is to stamp the values into a line, to give them an order, which facilitates ordinal classification. After this one node linear layer, an output layer is included with one bias for each class whose objective is to classify the patterns into their corresponding class. This classification structure corresponds to the POM model, which, like the majority of existing ordinal regression models, can be represented in the following general form:

$$C(\mathbf{x}) = \begin{cases} c_1, & \text{if } f(\mathbf{x},\theta) \leq \beta_0^1 \\ c_2, & \text{if } \beta_0^1 < f(\mathbf{x},\theta) \leq \beta_0^2 \\ \cdots & \\ c_J, & \text{if } f(\mathbf{x},\theta) > \beta_0^{J-1} \end{cases}, \tag{1}$$

where $\beta_0^1 < \beta_0^2 < \cdots < \beta_0^{J-1}$ (this will be the most important constraint in order to adapt the nominal classification model to ordinal classification), J is the number of classes, $\mathbf{x}$ is the input pattern to be classified, $f(\mathbf{x},\theta)$ is a ranking function and θ is the vector of parameters of the model. Indeed, the analysis of (1) reveals the general idea previously presented: patterns, $\mathbf{x}$, are projected to a real line by using the ranking function, $f(\mathbf{x},\theta)$, and the biases or thresholds, β_0^i, separating the ordered classes, where $\beta_0^0 = -\infty$ and $\beta_0^J = \infty$.

The standard POM model approximates $f(\mathbf{x},\theta)$ by a simple linear combination of the input variables, while our model considers a non-linear basis transformation of the inputs. Let us formally define the model for each class as $f_l(\mathbf{x},\theta,\beta_0^l) = f(\mathbf{x},\theta) - \beta_0^l$; $1 \leq l \leq J$, where the projection function $f(\mathbf{x},\theta)$ is estimated with the following S sigmoidal basis functions $f(\mathbf{x},\theta) = \beta_0 + \sum_{j=1}^{S} \beta_j B_j(\mathbf{x},\mathbf{w}_j)$, replacing $B_j(\mathbf{x},\mathbf{w}_j)$ by sigmoidal basis functions:

$$B_j(\mathbf{x},\mathbf{w}_j) = \frac{1}{1 + exp\left(-\left(\sum_{i=1}^{k} w_{ji} x_i\right)\right)}.$$

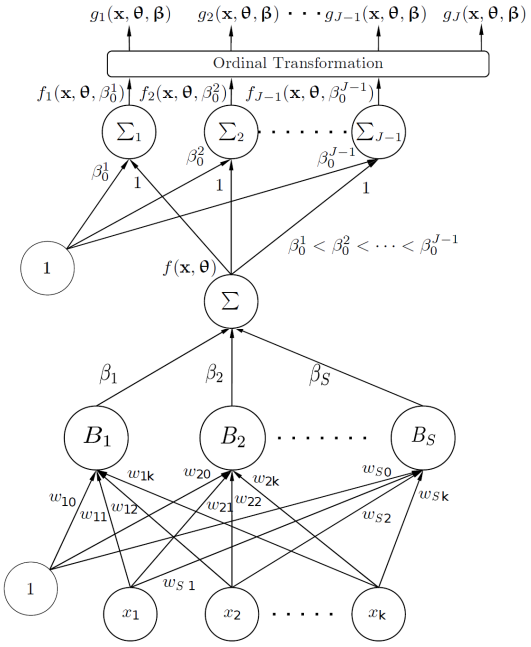

Fig. 1. Proposed sigmoidal model for ordinal regression

By using the POM model, this projection can be used to obtain cumulative probabilities, cumulative odds and cumulative logits of ordinal regression in the following way [18]:

$$P(Y \leq l) = P(Y = 1) + \cdots + P(Y = l),$$

$$odds(Y \leq l) = \frac{P(Y \leq l)}{1 - P(Y \leq l)},$$

$$logit(Y \leq l) = \ln\left(\frac{P(Y \leq l)}{1 - P(Y \leq l)}\right) = f(\mathbf{x},\theta) - \beta_0^l \equiv f_1(\mathbf{x},\theta,\beta_0^1),$$

$$P(Y \leq l) = \frac{1}{1 + \exp(f(\mathbf{x},\theta) - \beta_0^l)} \equiv \frac{1}{1 + \exp(f_l(\mathbf{x},\theta,\beta_0^l))};\ 1 \leq l \leq J,$$

where $P(Y = j)$ is the probability a pattern $\mathbf{x}$ has of belonging to class j, $P(Y \leq l)$ is the probability a pattern $\mathbf{x}$ has of belonging to class 1 to l and the logit is modeled by using the ranking function, $f(\mathbf{x},\theta)$, and the corresponding bias, β_0^l. We can come back to $P(Y = l)$ from $P(Y \leq l)$:

$$P(Y = l) = g_l(\mathbf{x},\theta,\beta) = P(Y \leq l) - P(Y \leq l-1),\ l = 1,\ldots,J,$$

and the final model can be expressed as:

$$g_l(\mathbf{x},\theta,\beta) = \frac{1}{1+\exp(f_l(\mathbf{x},\theta,\beta_0^l))} - \frac{1}{1+\exp(f_{l-1}(\mathbf{x},\theta,\beta_0^{l-1}))},\ l = 1,\ldots,J.$$

2.2 Metrics and Methodology

There are many ordinal measures to determine the efficiency of a g classifier, but not all pairs formed by these metrics are valid to guide a MOEA. To guide the learning of our MOEA for designing ANN models to determine the growth limits of microorganisms in predictive microbiology, the two metrics used are the Average Mean Absolute Error (*AMAE*) and the Maximum Mean Absolute Error (*MMAE*). These metrics are explained in detail in [7]. In general, these two ordinal measures are non-cooperative [7]: when the value of one of them increases the value of the other decreases. Thus the use of a MOEA based on Pareto dominance is justified.

The MOEA used in this work is called MPENSGA2 (Memetic Pareto Evolutionary NSGA2). This algorithm is based on the original algorithm NSGA2, and is described in detail in [10]. The framework of the MPENSGA2 is shown in Fig. 2, and the main differences with respect to [10] are:

- The metrics for guiding the algorithm are *AMAE* and *MMAE*.
- The delete links mutator and the parametric mutator have been modified to take into account the constraints of the ordinal model POM, β_0^i bias.
- The local search algorithm, *iRprop*$^+$ is used in all generations, not only in three generations of the evolutionary process. This is because the datasets used in this work have a relatively small size, so that the computational cost does not increase dramatically. The function with respect to which the local search is made is the cross-entropy function, as well as *MMAE* and *AMAE*, which are not derivable. Since patterns, according to the ordinal model, are projected onto a straight line, an improvement in entropy is an improvement in *AMAE* and the *MMAE* metrics.

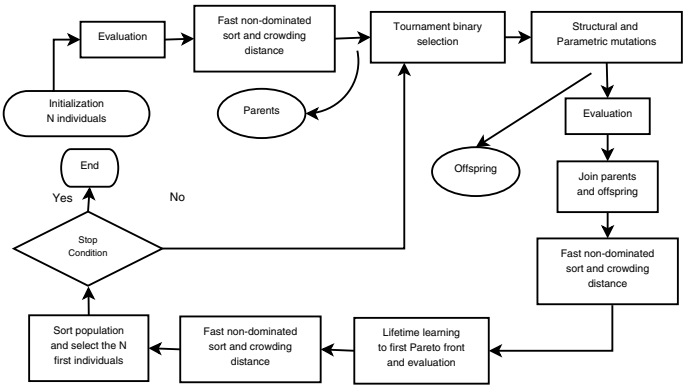

Fig. 2. Framework of the MPENSGA2 Algorithm

3 Description Problem

In this paper, four observed microbial responses are obtained based on the growth probability of a microorganism: $p = 1$ (growth); $0.5 \leq p < 1$ (high growth probability); $0 < p < 0.5$ (low growth probability); and $p = 0$ (no-growth). These four observed microbial responses were used in [8]. There is an intrinsic ranking of the different classes, so this paper tackles the problem of microbial growth from an ordinal point of view. This paper addresses the following two microbiological problems:

Listeria monocytogenes have been a serious problem for food industries due to their ubiquity in the natural environment and the specific growth conditions of the pathogen that lead to its high prevalence in different kinds of food products. One impetus for this research has been the problem of listeriosis, and different strategies have been proposed to limit the levels of contamination at the time of consumption to less than 100 CFU/g (European Commission [5]). *L. monocytogenes* data were gathered at 7°C in Nutrient Broth with different combinations of environmental factors pH (5.0–6.0, six levels at regular intervals), water activity (0.960–0.990, six levels) and acetic acid concentration (0–0.4% (w/w), three levels), as can be seen in [20]. Thus, 108 different conditions were tested with 20 replicates per condition.

Staphylococcus aureus has been recognized as an indicator of deficient food and processing hygiene and is a major cause of food gastroenteritis worldwide [19]. A fractional factorial design was followed in order to ascertain the growth limits of *S. aureus* [19] by carefully choosing a subset (fraction) of the experimental runs of a full factorial design in order to reduce experimental time and resources. The selection was based on restricting the levels of the environmental factors studied for the growth/no-growth domain of *S. aureus*. Since no growth was detected at 7.5°C or below, data were collected at 8°, 10°, 13°, 16°and 19°C, at pH levels from 4.5 to 7.5 (0.5 intervals) and at 19 a_w levels (from 0.856 to 0.999 at regular intervals). In this study, there are 30 replicates per condition, more than other studies obtaining the growth/no-growth transition.

Table 1. Features of the datasets

Dataset	#Training patterns	#Generalisation patterns	#Patterns per class in training	#Patterns per class in generalisation
L. monocytogenes	72	36	(29-10-10-23)	(15-3-6-12)
L. monocytogenes (SMOTE)	92	36	(29-20-20-23)	(15-3-6-12)
S. aureus	146	141	(60-22-7-57)	(57-23-5-56)
S. aureus (SMOTE)	167	141	(60-22-28-57)	(57-23-5-56)

4 Experiments

4.1 Experimental Design

A fractional factorial design matrix form was used in this study (in [20] the fractional factorial design for *L. monocytogenes* is presented and in [19], for *S. aureus*). This type of experimental design is common in predictive microbiology. These designs consider values close to the frontiers to train, and central values to generalize.

Table 1 shows the characteristics of the original datasets, where their unbalanced nature can be appreciated. In order to deal with this unbalanced nature, a pre-processing method has been applied to each dataset. The method applied is the well-known Synthetic Minority Over-sampling Technique algorithm (SMOTE) [2]. This method is applied twice to the minority class in such a way that each application of SMOTE duplicates its number of patterns. The synthetic generated patterns are only used to train the model, not to test it, as they cannot be considered real data. These synthetic patterns were generated using information from the five nearest neighbors. This method has been configured and run using WEKA software [15]. The final distribution of the datasets after applying this pre-processing can be seen in Table 1. This study only presents the results obtained after applying the SMOTE procedure.

Once the Pareto front is built into the last generation of training, two selection strategies are used to choose the individuals. The first strategy selects the best model in *AMAE*, which is the upper individual from the Pareto front. This selection method is called MPENSGA2-A. The second strategy selects the best model in *MMAE*, which is the bottom individual from the Pareto front. This selection method is called MPENSGA2-M. Because the MPENSGA2 algorithm is stochastic, the algorithm was run 30 times and the mean and standard deviation was obtained from the 30 individuals for the upper and lowest extremes.

In all experiments, the population size for the MPENSGA2 algorithm is established at 100. The probability of choosing a type of mutator and applying it to an individual is equal to 1/5. For the *iRprop*$^+$ algorithm, the number of epochs is established at 20. The other configured parameters are $\eta^+ = 1.2$, $\eta^- = 0.5$, $\Delta_0 = 0.0125$, $\Delta_{\min} = 0$ and $\Delta_{\max} = 50$, based on previous works.

4.2 Comparison Methods

For comparison purposes, different nominal and ordinal classification methods from the literature have been included in the experimentation.

The nominal classification methods are: MLogistic, SLogistic, MLP, C4.5 and LibSVM. These methods have been configured and run using WEKA software [15].

The ordinal classification methods are: SVMRank [17], ASAOR(C4.5) [12], GPOR [3], SVOR-EX and SVOR-IM [4]. These methods have been configured and run using the code provided by the authors. The corresponding hyper-parameters for these methods were adjusted using a grid search with a 10-fold cross-validation.

4.3 Results

Table 2 shows the results obtained in generalization for the two pre-processed datasets. For the MPENSGA2 method, the results presented correspond to mean values and standard deviation (mean$_{\text{SD}}$) for the 30 extreme models of the Pareto fronts generated in 30 runs (one Pareto front for each run). In addition, the best models obtained in each of the Pareto front extremes are presented in order to make comparisons with other methods. The measures used for this comparison are: the Correctly Classified Rate, *CCR*; the Mean Absolute Error, *MAE*; the Average *MAE*, *AMAE*; the Maximum *MAE*, *MMAE*; and the Kendall's τ_b. The *CCR* is the classic nominal metric to evaluate a classifier,

Table 2. Statistical results in generalization for the pre-processing datasets

Dataset	Method	*CCR*	*MAE*	*AMAE*	*MMAE*	τ_b
L. monocytogenes	Mlogistic	0.750	0.250	0.279	0.333	0.844
	Slogistic	0.750	0.305	0.454	0.833	0.817
	MLP	0.722	0.277	0.233	0.333	0.845
	C4.5	0.805	0.222	0.195	0.333	0.862
	LibSVM	0.833	*0.166*	0.287	0.666	0.892
	SVMRank	0.805	0.194	0.175	0.333	0.884
	ASAOR(C4.5)	0.750	0.277	0.300	0.333	0.823
	GPOR	0.750	0.250	0.279	0.333	0.874
	SVOR-EX	0.750	0.250	0.191	0.333	0.871
	SVOR-IM	0.861	*0.166*	0.141	0.333	0.883
	Best MPENSGA2-A	**0.916**	**0.111**	**0.079**	*0.250*	*0.920*
	Best MPENSGA2-M	*0.888*	**0.111**	*0.100*	**0.166**	**0.922**
	MPENSGA2-A	$0.837_{0.037}$	$0.168_{0.037}$	$0.161_{0.048}$	$0.302_{0.070}$	$0.895_{0.021}$
	MPENSGA2-M	$0.839_{0.038}$	$0.165_{0.038}$	$0.159_{0.049}$	$0.297_{0.074}$	$0.896_{0.022}$
S. aureus	Mlogistic	0.687	0.475	0.750	1.217	0.718
	Slogistic	0.673	0.553	0.766	1.304	0.656
	MLP	**0.773**	*0.340*	0.646	1.200	0.794
	C4.5	0.716	0.496	0.732	1.400	0.649
	LibSVM	0.758	0.347	0.476	0.800	0.781
	SVMRank	0.695	0.347	0.450	0.652	0.794
	ASAOR(C4.5)	0.702	0.390	0.497	0.800	0.760
	GPOR	0.581	0.929	1.137	1.600	0.382
	SVOR-EX	0.680	0.361	0.459	0.652	0.791
	SVOR-IM	0.723	**0.319**	**0.419**	0.600	**0.806**
	Best MPENSGA2-A	*0.764*	**0.319**	*0.426*	*0.571*	*0.802*
	Best MPENSGA2-M	0.652	0.425	0.500	**0.414**	0.745
	MPENSGA2-A	$0.695_{0.053}$	$0.366_{0.054}$	$0.486_{0.048}$	$0.849_{0.138}$	$0.783_{0.029}$
	MPENSGA2-M	$0.554_{0.073}$	$0.518_{0.083}$	$0.519_{0.092}$	$0.641_{0.117}$	$0.722_{0.043}$

The best result is in **bold** face and the second best result in *italics*.

the *MAE*, *AMAE* and *MMAE* are ordinal metrics that depend on the distance between the ranking of two consecutive classes, and the τ_b measure is another ordinal metric independent on the values chosen for the ranks representing the classes.

The best models are generated by the MPENSGA2-A method, this method obtaining the best results in all metrics. Thus, we propose the MPENSGA2 algorithm to solve the two problems of predictive microbiology, specifically the upper best model of the Pareto front, which maximizes the value of *AMAE*.

EAs, and more specifically the MOEA, are computationally expensive especially when compared to local search algorithms, but the evolution of architectures enables ANNs to adapt their topologies to the different datasets without human intervention. This thus provides an approach to automatic ANN design as both ANN connection weights and structures can be evolved [21]. It is clear that the different non-evolutionary methods considered in this study demand a lower computational cost than the MPENSGA2 algorithm. However, the obtained models benefit clearly from the optimized structure learned by the EA, which allows achieve better results in generalization.

5 Conclusions

The proposed models for obtaining ordinal classification by using a generalized POM model with sigmoidal basis functions present competitive results with respect to other nominal and ordinal classifiers considering the five metrics used.

This modeling approach can provide a new insight for the predictive microbiology field, since it could directly determine, and with a high confidence level, if a pathogenic or spoilage microorganism could flourish. For stakeholders, the application of this kind of tools could be very useful in order to set microbiological criteria or to determine microbial shelf life under a given set of conditions. Implementation of risk management measures in food industries based on qualitative approaches (i.e. a combination of factors which limit microbial growth to below 0.01 or which fall into the "low growth probability" class) will suppose a breakthrough in guaranteeing microbial food safety.

Acknowledgement. This work was subsidized in part by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the "Junta de Andalucía" (Spain), under Project P08-TIC-3745. M. Cruz-Ramírez's research has been subsidized by the FPU Predoctoral Program (Spanish Ministry of Education and Science), grant reference AP2009-0487.

References

1. Bishop, C.: Neural Networks for Pattern Recognition. Oxford University Press (1995)
2. Chawla, N., Bowyer, K., Hall, L., Kegelmeyer, W.: SMOTE: Synthetic minority oversampling technique. Journal of Artificial Intelligence Research 16, 321–357 (2002)
3. Chu, W., Ghahramani, Z.: Gaussian processes for ordinal regression. Journal of Machine Learning Research 6, 1019–1041 (2005)
4. Chu, W., Keerthi, S.S.: Support Vector Ordinal Regression. Neural Computation 19(3), 792–815 (2007)
5. Commission, E.: Opinion of the scientific committee on veterinary measures relating to public health on listeria monocytogenes (1999), `http://www.europa.eu.int/comm/food/fs/sc/scv/out25`
6. Corchado, E., Abraham, A., Ponce, A.C., Ferreira de Carvalho, L.: Hybrid intelligent algorithms and applications. Information Sciences 180(14), 2633–2634 (2010)
7. Cruz-Ramírez, M., Hervás-Martínez, C., Sánchez-Monedero, J., Gutierrez, P.A.: A preliminary study of ordinal metrics to guide a multi-objective evolutionary algorithm. In: 11th International Conference on Intelligent Systems Design and Applications, ISDA 2011, Córdoba, Spain, pp. 1176–1181 (2011)
8. Cruz-Ramírez, M., Sánchez-Monedero, J., Fernández-Navarro, F., Fernández, J., Hervás-Martínez, C.: Memetic Pareto differential evolutionary artificial neural networks to determine growth multi-classes in predictive microbiology. Evol. Intelligence 3(3-4), 187–199 (2010)
9. Fernández, J., Hervás, C., Martínez-Estudillo, F., Gutierrez, P.: Memetic pareto evolutionary artificial neural networks to determine growth/no-growth in predictive microbiology. Applied Soft Computing 11, 534–550 (2011)
10. Fernández, J., Martínez, F., Hervás, C., Gutiérrez, P.: Sensitivity versus accuracy in multiclass problems using memetic pareto evolutionary neural networks. IEEE Transactions on Neural Networks 21(5), 750–770 (2010)

11. Fernandez-Navarro, F., Valero, A., Hervás-Martínez, C., Gutiérrez, P.A., Gimeno, R.G., Cosano, G.Z.: Development of a multi-classification neural network model to determine the microbial growth/no growth interface. I. J. of Food Microbiology 141(3), 203–212 (2010)
12. Frank, E., Hall, M.: A simple approach to ordinal classification. In: Proceedings of the 12th European Conference on Machine Learning, EMCL 2001, pp. 145–156 (2001)
13. Goh, C., Teoh, E., Tan, K.: Hybrid multiobjective evolutionary design for artificial neural networks. IEEE Transactions on Neural Networks 19(9), 1531–1548 (2008)
14. Gutiérrez, P.A., Pérez-Ortíz, M., Fernández-Navarro, F., Sánchez-Monedero, J., Hervás-Martínez, C.: An experimental study of different ordinal regression methods and measures, Salamanca, Spain, pp. 296–307 (2012)
15. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. Special Interest Group on Knowledge Discovery and Data Mining Explorer Newsletter 11(1), 10–18 (2009)
16. Kordík, P., Koutník, J., Drchal, J., Kovárik, O., Cepek, M., Snorek, M.: Meta-learning approach to neural network optimization. Neural Networks 23, 568–582 (2010)
17. Li, L., Lin, H.T.: Ordinal regression by extended binary classification. Advances in Neural Information Processing Systems 19, 865–872 (2007)
18. McCullagh, P.: Regression models for ordinal data. Journal of the Royal Statistical Society, Series B (Methodological) 42(2), 109–142 (1980)
19. Valero, A., Pérez-Rodríguez, F., Carrasco, E., Fuentes-Alventosa, J., García-Gimeno, R., Zurera, G.: Modelling the growth boundaries of staphylococcus aureus: Effect of temperature, ph and water activity. International Journal Food Microbiology 133, 186–194 (2009)
20. Vermeulen, A., Gysemans, K.P.M., Bernaerts, K., Geeraerd, A.H., Impe, J.F.V., Debevere, J., Devlieghere, F.: Influence of ph, water activity and acetic acid concentration on *Listeria monocytogenes* at 7^{o}c: Data collection for the development of a growth/no growth model. International Journal of Food Microbiology 114(3), 332–341 (2007)
21. Yao, X.: Evolving artificial neural networks. Proceedings of the IEEE 87, 1423–1447 (1999)