



# Multinomial logistic regression and product unit neural network models: Application of a new hybrid methodology for solving a classification problem in the livestock sector

Mercedes Torres<sup>a,\*</sup>, Cesar Hervás<sup>b</sup>, Carlos García<sup>a</sup>

<sup>a</sup> Department of Management and Quantitative Methods, ETEA, Spain

<sup>b</sup> Department of Computing and Numerical Analysis, University of Córdoba, Spain

## ARTICLE INFO

### Keywords:

Evolutionary computation  
Neural networks  
Product units  
Classification  
Livestock application

## ABSTRACT

This work presents a new approach for multi-class pattern recognition based on the hybridization of a linear and nonlinear model. We propose multinomial logistic regression where some new covariates are defined by a product unit neural network, where in turn, the nonlinear basis functions are constructed with the product of the inputs raised to arbitrary powers. The application of this methodology involves, first of all, training the coefficients and the basis structure of product unit models using techniques based on artificial neural networks and evolutionary algorithms, followed by the application of multinomial logistic regression to both the new derived features and the original ones. To evaluate the efficacy of our technique we pose a difficult problem, the classification of sheep with respect to their milk production in different lactations, using covariates that only involve the first weeks of lactation. This enables the productive capacity of the animal to be identified more rapidly and leads to a faster selection process in determining the best producers. The results obtained with our approach are compared to other classification methodologies. Although several of these methodologies offer good results, the percentage of cases correctly classified was higher with our approach, which shows how instrumental the potential use of this methodology is for decision making in livestock enterprises, a sector relatively untouched by the technological innovations in business management that have been appearing in the last few years.

© 2009 Elsevier Ltd. All rights reserved.

## 1. Introduction

Classification problems are often encountered in many different fields, such as biology (Hajmeer & Basheer, 2003), medicine (Schwarzer, Vach, & Schumacher, 2000; Youngdai & Sunghoon, 2006), computer vision (Subasi, Alkan, & Koklukaya, 2005), artificial intelligence and remote sensing (Yuan-chin & Sung-Chiang, 2004). In the business world, applications of this type are becoming more and more frequent: in finance (Parag & Pendharkar, 2005; Rada, 2008; Tian-Shyug et al., 2006), marketing (Kaefer, Heilman, & Ramenofsky, 2005), and human resource management (Sexton & McMurtrey, 2005). There has been a renewed interest in this type of technique in the last few years due to the difficulties inherent in such new problems as dealing with data mining, document classification, financial forecasts, web-mining, etc. This great practical interest in classification problems has motivated researchers to develop a huge number of methods as quantitative models for classification purposes (i.e. see Bernadó & Garrell, 2003; Duda & Hart,

2001). Linear Discriminant Analysis (LDA) (Johnson & Wichern, 2002) was the first method developed to address the classification problem from a multidimensional perspective. LDA has been used for decades as the main classification technique and it is still being used, at least as a reference point, to compare the performance of new techniques. Another widely used parametric classification technique, developed to overcome LDA's restrictive assumptions (multivariate normality, equality of dispersion matrices between groups), is Quadratic Discriminant Analysis (QDA). The Logistic Regression model (LR) has also been widely used in statistics for many years and has recently been the object of extensive study in the machine learning community (Dreiseitl & Ohno-Machado, 2002; Duda & Hart, 2001; Hosmer & Lemeshow, 2000; Yuan-chin & Sung-Chiang, 2004). During the last two decades several alternative non-parametric classification techniques have also been developed, including, among others, mathematical programming techniques (Fredd & Glover, 1981), multicriteria decision aid methods (Dourmos, Zopounidis, & Pardalos, 2000), neural networks (Patuwo, Hu, & Hung, 1993; Widrow, 1962) and machine learning approaches (Kordatoff & Michlaski, 1990).

However, in spite of the great number of techniques developed to solve classification problems, there is no optimum methodology

\* Corresponding author. Tel.: +34 957 222100; fax: +34 957 222101.

E-mail addresses: [mtorres@etea.com](mailto:mtorres@etea.com) (M. Torres), [chervas@uco.es](mailto:chervas@uco.es) (C. Hervás), [garcia@etea.com](mailto:garcia@etea.com) (C. García).

or technique to resolve a specific problem and this is why the comparison and combination of different types of classification is a common practice today (Lin, Lee, & Lee, 2008; Major & Ragsdale, 2001; Martínez, Hervás, et al., 2006).

As a matter of fact, the present work deals with the application of a new hybrid methodology that combines multinomial logistic regression and unit-product network models as an alternative to other well known techniques (some relatively recent and others more traditional) for solving a real classification problem in the livestock sector. It is an extension to more than two classes of a recent work which propose this method for two classes (Hervás & Martínez, 2007).

The combination of the two techniques is justified by the fact that, although LR is a simple and useful procedure, it poses problems when applied to a real problem of classification, where we frequently cannot make the stringent assumption that there are additive and purely linear effects of the covariates. These difficulties are usually overcome by augmenting or replacing the input vector with new variables, basis functions, which are transformations of the input variables, and then by using linear models in this new space of derived input features. Methods like sigmoidal feed-forward neural networks (Bishop, 1995), projection pursuit learning (Friedman & Stuetzle, 1981), generalized additive models (Hastie & Tibshirani, 1990) and PolyMARS (Kooperberg, Bose, & Stone, 1997), and a hybrid of multivariate adaptive splines (Friedman, 1991; Tian-Shyug et al., 2006), specifically designed to solve classification problems, can be seen as different non-linear basis function models.

The unit-product networks are similar to standard sigmoidal neural networks but are based on multiplicative nodes instead of additive ones. These functions correspond to a special type of neural networks called product-unit neural networks (PUNN) introduced by Durbin and Rumelhart (1989), and developed by Ismail and Engelbrecht (1999) and Schmitt (2002). The nonlinear basis functions of the model are constituted by the product of the variables initially included in the problem formulation raised to arbitrary powers. We estimate the variables' exponents and determine the optimum number of product units in several steps (solving one of the main problems in the use of this type of models). In a first step an evolutionary algorithm (EA) is applied that optimizes a loss function. However, these algorithms are relatively poor at finding the precise optimum solution in the region that the algorithm converges to. So a local optimization algorithm was used, in a second step, to improve the EA's lack of precision. Once the basis functions have been determined by the EA, the model is linear in these new variables together with the initial covariates, and the fitting proceeds with the standard maximum likelihood optimization method for multinomial logistic regression. Finally, a backward-step procedure is applied, pruning variables sequentially to the model obtained previously until further pruning does not improve the fit. In this way the models obtained can be simpler and more comprehensible for researchers in the livestock sector.

The performance of the proposed methodology was evaluated in a real problem which consists of classifying a sheep flock into three classes, according to its milk production capacity, by using solely the first milk controls, and thus shortening the current evaluation process that uses the selection schema of the Manchegan breed (Montoro & Pérez-Guzmán, 1996). Three classes are established: the best productive ones (called "good"), the worst (called "bad") and the intermediate (called "normal"). With the results of the classification, the stock farmer would be able to identify the most productive animals in the flock with a minimum of necessary information and could then contribute to the genetic progress of the breed. Moreover, these models could lead to a decrease in the great differences in production that have been found in the last few years between different Spanish sheep breeds,

like the Manchegan, with respect to other breeds (the French Lacune, for example) (Buxadé, 1998; Gallego & Bernabeu, 1994; Serano & Montoso, 1996).

So we have here an application of new computational methodologies for the management of a dairy, a sector relatively untouched by the technological innovations in business management that have been appearing in the last few years (Torres, Hervás, & Amador, 2005). In general, the greater part of operational researchers' and agrarian economists' attention has been concentrated on the area of animal feed, due more to their connection with the animal food industry than to any connection with the dairy establishments themselves.

Simultaneously we compare our model results with those obtained by a standard multinomial logistic regression that uses only the original input variables, to verify the advantages of our approach. Furthermore, other classification algorithms based on artificial neural networks were applied (Dreiseitl & Ohno-Machado, 2002). Specifically we use a standard multilayer perceptron model (MLP) that uses a back-propagation learning algorithm (Hayken, 1994; Williams & Minai, 1990), and another MLP model, where an evolutionary algorithm is coupled with a pruning one to eliminate non-significant model coefficients (Bebis & Georgopoulos, 1997; Honaver & Balakrishnan, 1998) (from now on we will call this model MLPEA). Thus we attempt to achieve the neuronal network architecture that will allow us to predict what sheep productive capacity will be relying on the least possible amount of information. We have also applied the second most popular choice of network in classification problems; the radial basis function network (RBF). This type of network has a very strong mathematical foundation and uses normalized Gaussian radial basis functions (Orr et al., 1996; Oyang, Hwang, Ou, Chen, & Chen, 2005).

Finally we apply other well known classification methods to our problem (some of them of statistical origin and others from the computational field) to compare their classification capacity with ours. We have used: the classical decision tree C4.5 (Quinlan, 1993) with pruning (<http://www.cse.unsw.edu.au/quinlan/>); three statistical algorithms: a Linear Discriminant Analysis, LDA, where hypothetically the instances within each class are normally distributed with a common covariance matrix; a quadratic discriminant analysis, QDA, where each covariance matrix is different and estimated by the corresponding sample covariance matrix; and, finally, the K-Nearest Neighbour algorithm (KNN) (Dasarathy, 1991; Hervás & Martínez, 2007; Kaefer et al., 2005).

The rest of the paper is organized as follows. Section 2 describes the proposed logistic regression model and the other methodologies applied, Section 3 explains the process to obtain the data set as well as the procedure to select the variables to include in the models. The results of the experiment are tabulated and discussed in Section 4. Finally, conclusions are presented in Section 5.

## 2. Methodology

### 2.1. Multinomial logistic regression with product unit covariates

Logistic regression methods are common statistical tools for modelling discrete response variables such as binary, categorical and ordinal responses. If the values of the response of these variables refer to different categories where a specific group of elements (called a sample), can be grouped into a class according to a series of characteristics which have previously been measured for each element, we are confronted with what is called a classification problem. So, in a classification problem we find the following elements: a number of features  $x_i$ ,  $i = 1, \dots, p$  which are measured from each element in the sample; a finite number of classes  $K$  where the elements have to be classified as well as the

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات