



A novel hybrid classification model of artificial neural networks and multiple linear regression models

Mehdi Khashei*, Ali Zeinal Hamadani, Mehdi Bijari

Department of Industrial Engineering, Isfahan University of Technology, Isfahan, Iran

ARTICLE INFO

Keywords:

Classification
Pattern recognition
Artificial neural networks (ANNs)
Multiple linear regression models (MLR)
Discriminant analysis (DA)

ABSTRACT

The classification problem of assigning several observations into different disjoint groups plays an important role in business decision making and many other areas. Developing more accurate and widely applicable classification models has significant implications in these areas. It is the reason that despite of the numerous classification models available, the research for improving the effectiveness of these models has never stopped. Combining several models or using hybrid models has become a common practice in order to overcome the deficiencies of single models and can be an effective way of improving upon their predictive performance, especially when the models in combination are quite different. In this paper, a novel hybridization of artificial neural networks (ANNs) is proposed using multiple linear regression models in order to yield more general and more accurate model than traditional artificial neural networks for solving classification problems. Empirical results indicate that the proposed hybrid model exhibits effectively improved classification accuracy in comparison with traditional artificial neural networks and also some other classification models such as linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), *K*-nearest neighbor (KNN), and support vector machines (SVMs) using benchmark and real-world application data sets. These data sets vary in the number of classes (two versus multiple) and the source of the data (synthetic versus real-world). Therefore, it can be applied as an appropriate alternate approach for solving classification problems, specifically when higher forecasting accuracy is needed.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Classification is an important area of research that concerned with assigning an object to one of a set of classes, based upon attributes of that object. The performance of the classification process is dependent on how well the discriminant function for the specific problem performs. A discriminant function is developed to minimize the misclassification rate, according to some given samples of input and output vector couples, which are referred to as training data set. This discriminant function is then used for classifying new observations into previously defined groups and for testing the accuracy of the classification. Classification problems have been examined in fields as diverse as business, medicine, biology, image recognition, etc. and using of these models has become more indispensable in aforementioned areas, especially in business and finance. Several different classification approaches have been proposed in the literature since the earliest work of Fisher (1936). The classification approaches are generally categorized in two main categories, linear and nonlinear approaches.

Linear classification approaches partition the input space into a collection of disjoint regions, separated by linear decision boundaries. Notable examples of linear classification techniques that have been widely used in classification include those by multiple linear regression (MLR), linear discriminant analysis (LDA), logistic regression, separating hyper planes, etc. These classification techniques work well when the classes are linearly separable. However, in many real world problems the data may not be linearly separable and also data are very closely spaced and therefore a highly nonlinear decision boundary is required in order to separate the data (Satapathy et al., 2009). Several classes of nonlinear classification techniques have been proposed in the literature in order to overcome the linear limitation of the linear classification techniques. These techniques include those by the classical techniques such as quadratic discriminant analysis (QDA), *K*-nearest neighbor (KNN), etc. and artificial neural networks approaches such as neural trees, multilayer perceptrons (MLPs), probabilistic neural networks (PNNs), support vector machines (SVMs), etc.

Artificial neural networks are one of the most accurate and widely used classification techniques that have enjoyed fruitful applications in many areas. Several distinguishing features of artificial neural networks make them valuable and attractive for

* Corresponding author. Tel.: +98 311 3912550 1; fax: +98 311 3915526.
E-mail address: khashei@in.iut.ac.ir (M. Khashei).

classification tasks. First, as opposed to the traditional model-based techniques, artificial neural networks are data-driven self-adaptive methods in that there are few a priori assumptions about the models for problems under study. Second, artificial neural networks can generalize. After learning the data presented to them (a sample), artificial neural networks can often correctly infer the unseen part of a population even if the sample data contain noisy information. Third, artificial neural networks are universal functional approximators. It has been shown that a network can approximate any continuous function to any desired accuracy. Finally, artificial neural networks are nonlinear (Khashei, Hejazi, & Bijari, 2008).

Given the advantages of artificial neural networks, it is not surprising that this methodology has attracted overwhelming attention in classification (Maulik & Mukhopadhyay, 2010). Artificial neural networks have been found to be a viable contender to various traditional classification models in many different areas (Dubois, Bohling, & Chakrabarti, 2007; Kara & Okandan, 2007). Castellani and Rowlands (2009) address the design and the training of a multilayer perceptron classifier for identification of wood veneer defects from statistical features of wood sub-images. Kruzlicova et al. (2009) demonstrate the possibility of using artificial neural networks for the Slovak white wines classification. Olmez and Dokur (2003) propose using the artificial neural networks in order to handle the heart sounds classification problems and to increase the classification performance.

Banerjee, Kiran, Murty, and Venkateswarlu (2008) present an artificial neural system for classification and identification of Anopheles mosquito species based on the information content of ribosomal DNA sequences. Acharya, Bhat, Iyengar, Rao, and Dua (2003) deal with the classification of certain diseases using artificial neural network and fuzzy equivalence relations. The heart rate variability is used as the base signal from which certain parameters are extracted and presented to the artificial neural network for classification. Guven and Kara (2006) concentrate on the diagnosis of subnormal eye through the analysis of electro-oculography (EOG) signals with using of the artificial neural network. Fingerprints classification (Nagaty, 2001), cervical cancer classification (Qiu, Tao, Tan, & Wu, 2007), protein structure classification (Karci & Demir, 2009), gait classification in post-stroke patients (Kaczmarczyk, Wit, Krawczyk, & Zaborski, 2009), bankruptcy prediction problem (Pendharkar, 2005), Tamil documents classification (Rajan, Ramalingam, Ganesan, Palanivel, & Palaniappan, 2009) are some other successful applications of the artificial neural networks in comparison with other those of the traditional classification models.

Although artificial neural networks are flexible computing frameworks and universal approximators that can be applied to a wide range of forecasting problems with a high degree of accuracy, their performance in some specific situations such as linear problems is inconsistent (Khashei & Bijari, 2010). In the literature, several papers are devoted to comparing artificial neural networks with linear models. Despite of the several studies, which have shown artificial neural networks are significantly better than the conventional linear models and their results considerably and consistently more accurately, some other studies have reported inconsistent results (Zhang, Patuwo, & Hu, 1998). Some researchers believe that in some specific situations where artificial neural networks perform worse than linear statistical models, the reason may simply be that the data is linear without much disturbance, therefore; cannot be expected that artificial neural networks to do better than linear models for linear relationships (Khashei & Bijari, 2010). However, for whatever reason, using artificial neural networks to model linear problems have yielded mixed results; and hence, it is not wise to apply neural networks blindly to any type of data.

Both multiple linear regression and artificial neural networks models have achieved successes in their own linear or nonlinear domains. However, none of them is a universal model that is suitable for all circumstances. The approximation of the multiple linear regression models to complex nonlinear problems as well as artificial neural networks to model linear problems may be totally inappropriate, and also, in problems that consist both linear and nonlinear correlation structures. Using hybrid models or combining several models has become a common practice in order to overcome the limitations of each component model (Khashei, Bijari, & Raissi, 2009). The basic idea of these multi-model approaches is the use of each component model's unique capability to better capture different patterns in the data. In addition, since it is difficult to completely know the characteristics of the data in a real problem, hybrid methodology that has both linear and nonlinear modeling capabilities can be a good strategy for practical use.

In the literature, different combination techniques have been proposed in order to overcome the deficiencies of single classification models and yield more accurate results (Hur & Kim, 2008). The combination techniques can be generally categorized in two main categories, competitive and cooperative architectures. In a competitive architecture, the aim is to build appropriate modules to represent different parts, and to be able to switch control to the most appropriate model. In a cooperative modular, the aim is to combine models to build a complete picture from a number of partial solutions. The assumption is that a model may not be sufficient to represent the complete behavior of a under study system.

In recent years, several hybrid classification models have been proposed, using artificial neural networks and applied to the classification problems with good performance. Chakraborty (2009) proposes an integrated approach for classification and variable selection using the Bayesian K -nearest neighbor and stochastic search variable selection technique for simultaneous cancer classification and gene selection. This model provides a full probabilistic treatment for K -nearest neighbor along with adaptive variable selection. Connolly, Granger, and Sabourin (2010) propose an adaptive classification system (ACS) that combines a fuzzy ART-MAP neural network classifier suitable for incremental learning, and a dynamic particle swarm optimization (DPSO) algorithm capable of finding and tracking several local optima in the optimization space, for video-based face recognition. Ostermark, 2000 proposes a flexible hybrid genetic fuzzy neural network (GFNN) algorithm for multigroup classification problems that combines genetic computation with those on fuzzy neural networks. Aci, Inan, Avci, and neighbor (2010) propose a hybrid method by using K -nearest neighbor, Bayesian models and genetic algorithms in order to achieve successful results on classifying by eliminating data that make difficult to learn.

Polat and Gunes (2009) propose a novel hybrid classification system based on C4.5 decision tree classifier and one-against-all approach to classify the multi-class problems in order to improve the classification accuracy in the case of multi-class classification problems. Tagluk, Akin, and Sezgin (2010) describe a new hybrid model by combining wavelet transforms and artificial neural networks to classify sleep apnea syndrome (SAS). Wang, Li, Zhang, Gui, and Huang (2010) propose a novel ensemble method which combines base probabilistic neural networks (PNNs) classifiers with neighborhood rough set model based gene reduction. Pendharkar (2001) propose a hybrid evolutionary-neural approach for binary classification that incorporates a special training data over-fitting minimizing selection procedure for improving the prediction accuracy on holdout sample. This approach integrates parallel global search capability of genetic algorithms (GAs) and local gradient-descent search of the back-propagation algorithm. Sinha and Fieguth (2006) propose a new neuro-fuzzy classifier that

Download English Version:

<https://daneshyari.com/en/article/10322518>

Download Persian Version:

<https://daneshyari.com/article/10322518>

[Daneshyari.com](https://daneshyari.com)