PERGAMON

Expert Systems with Applications 20 (2001) 201–206

Expert Systems with Applications

www.elsevier.com/locate/eswa

Improving the performance of neural networks in classification using fuzzy linear regression

Chun-Hung Cheng[a,*], Boon-Toh Low[a], Pak-Kei Chan[a], J. Motwani[b]

[a]*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong*
[b]*Department of Management, Grand Valley State University, Grand Rapid, MI 49504, USA*

Abstract

In this paper, we apply the fuzzy linear regression (FLR) with fuzzy intervals analysis into a neural network classification model. The FLR works as a data handler and separates the data sample into two groups. By training two independent neural works with these two groups, we can better describe the distribution space of the corresponding data sample with two different functions, rather than using only one function. The experimental result shows that our approach improves the accuracy of classification. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Neural network; Fuzzy linear regression; Classification

1. Introduction

The classification problem is concerned with categorizing observations into different groups. 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 prediction. In this research, we consider a binary classification problem. In a binary problem, observations are classified into two groups. Although our result is applied to a binary pattern, they can be easily generalized to deal with the general case.

Many methods have been developed for classification problem. They include neural network (NN), multivariate discriminant analysis (MDA), decision tree, logistic regression, k nearest neighbor, among others (Bhattacharyya & Pendharkar, 1998; Han, Chandler & Liang, 1996; Tam & Kiang, 1992). Applications to different domains, such as managerial decisions, financial forecasting, bankruptcy prediction, image recognition, text-to-speech matching, medical diagnosis have been applied and tested (Han et al., 1996; Markham & Ragsdale, 1995; Salchenberger, Cinar & Lash, 1992; Tam & Kiang, 1992; Wilson & Sharda, 1994). Although several research studies suggest that the neural network approach have higher classification ability (Archer & Wang, 1993; Bhattacharyya & Pendharkar, 1998; Tam & Kiang, 1992; Wilson & Sharda, 1994) than many other methods, the predictive capability of the neural network approach still has potential for further improvement. Han et al. (1996) indicate that the relative performance of different classification techniques may depend on the data conditions of the training data set. Specifically, in a training data set, some observations may have similar input vector values but different output vector values. These observations are referred to as "bad" observations. If "bad" observations are used in the training process, they may adversely affect the performance of the resulting neural network.

The objective of this study is to propose a way to improve the accuracy of the neural network by separating "good" data from "bad" data for training. Our model comprises of two phases. In Phase I, fuzzy regression method with fuzzy interval analysis is applied. In Phase II, two simple back-propagation neural network constructions as the final classification engine are provided.

By using the fuzzy linear regression with the fuzzy interval analysis, we separate the training data into two groups based on the fuzzy interval. The separated training data sets are used to generalize two neural networks accordingly. With two neural networks, we formulate two different functions to describe the distribution space of the data. In our experiment, our model is compared with the conventional backpropagation neural network. This result shows that using two different functions to describe the distribution space of the observations promises a more accurate classification result.

* Corresponding author. Tel.: +852-2609-8322; fax: +852-2603-5505.
E-mail address: chcheng@se.cuhk.edu.hk (C.-H. Cheng).

0957-4174/01/$ - see front matter © 2001 Elsevier Science Ltd. All rights reserved.
PII: S0957-4174(00)00059-2

The paper is organized as follows. In Section 2, Tanaka's model (Tanaka, Uejima & Asai, 1982), the modified Tanaka's model (Tanaka & Ishibuchi, 1992) with fuzzy interval analysis, and the multilayer feedforward backpropagation neural network are introduced. In Section 3, the sample data and the methodology used are described. In Section 4, the details of our model are explained. In Section 5, the results and the comparison between our model and the conventional back-propagation neural network are reported. In Section 5, we use synthetic data to show the advantages of the new model. Finally, a conclusion is given in Section 6.

2. Literature review

The FLR approach was introduced by Tanaka et al. (1982; Tanaka & Ishibuchi, 1992) in order to deal with a vague phenomenon. The assumptions of Tanaka's model include (i) the input and output data of fuzzy linear model are fuzzy, (ii) the relationship between the input and output data is given by a fuzzy function, and (iii) the distribution of the data is possibilistic (Peters, 1994). The fuzzy linear regression has been applied to forecasting in an uncertain environment for finding an interrelationship between the linear interval model and the output intervals of the given data. Tanaka's model assumes a linear function as follows:

$$Y = A_0\chi_0 + A_1\chi_1 + \cdots + A_N\chi_N = \mathbf{Ax}\ (\chi_0 \leqslant 1) \tag{1}$$

where Y is the dependent variable, $\mathbf{x}$ is the vector of the independent variables, and $\mathbf{A}$ is the vector of a fuzzy set on the product space of parameters.

The fuzzy parameters A_j are represented in the form of triangular fuzzy numbers:

$$A_j(a_h) = \begin{cases} 1 - \dfrac{|\alpha_j - a_j|}{c_j} & \text{if } \alpha_j - c_j \le a_j \le \alpha_j + c_j \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $A_j(a_j)$ is the membership function of the fuzzy set of a_j, α_j is the center, and c_j is the spread of the fuzzy number.

According to Peters (1994), the membership degree of Y can be obtained:

$$Y(y) = \begin{cases} 1 - \dfrac{|y - x^{t}\alpha}{c^{t}|x|} & \text{for} \quad x \neq 0 \\ 1 & \text{for} \quad x = 0, y \neq 0 \\ 0 & \text{for} \quad x = 0, y = 0 \end{cases} \tag{3}$$

.

We minimize the total vagueness using

$$\text{MIN}\sum_{j=0}^{N}\left(c_j\sum_{i=0}^{M}|\chi_{ij}|\right) \tag{4}$$

where M is the number of training samples.

As the membership value of each observation y_j is greater than an imposed threshold, we have:

$$Y(y_i) \ge h \quad \text{for} \ \ i = 1,\ 2,\ \ldots,\ M. \tag{5}$$

A linear programming problem is constructed:

$$\text{MIN}\sum_{j=0}^{N}\left(c_j\sum_{i=0}^{M}|\chi_{ij}|\right)$$

Subject to

$$\begin{aligned} &\sum_{j=0}^{N}\alpha_j\chi_{ij} + |L^{-1}(h)|\sum_{j=0}^{N}c_j\chi_{ij}| \ge y_i \\ &\sum_{j=0}^{N}\alpha_j\chi_{ij} + |L^{-1}(h)|\sum_{j=0}^{N}c_j\chi_{ij}| \le y_i \end{aligned} \tag{6}$$

$$c \ge 0, \alpha \in \Re, \chi_{i0} \leqslant 1$$

$$0 \le h \le 1$$

$$i = 1,\ 2,\ \ldots,\ M$$

where $|L^{-1}(h)| = 1 - h$, $h \in [0,1]$, and the choice of the h value influences the widths c_j of the fuzzy parameters.

Tanaka's model can be used to analyze the interval of the dependent variable y, however, the drawback is that a few values may dominate the estimation of the bounds of the crisp interval. Therefore, the model is very sensitive to outliers (Peters, 1994). Peters (1994) provided a modification to Tanaka's model. In this model, the bounds of the interval are assumed fuzzy rather than crisp, so that each of the dependent data y has a membership degree of belonging to the interval. Peters' fuzzy linear programming model is formulated as follows:

$$\text{MAX}\bar{\lambda} = \frac{1}{M}\sum_{i=1}^{M}\lambda_i$$

Subject to

$$\begin{aligned} &(1 - \bar{\lambda})p_0 - \sum_{i=0}^{M}\sum_{j=0}^{N}c_j|\chi_{ij}| \ge -d_0 \\ &(1 - \lambda_i)p_i + \sum_{j=0}^{N}\alpha_j\chi_{ij} + \sum_{j=0}^{N}c_j|\chi_{ij}| \ge y_i \\ &(1 - \lambda_i)p_i + \sum_{j=0}^{N}\alpha_j\chi_{ij} + \sum_{j=0}^{N}c_j|\chi_{ij}| \ge -y_i \end{aligned} \tag{7}$$

$$-\lambda_i \ge -1$$

$$\lambda_i, c \ge 0, \alpha \in \Re, \chi_{i0} \leqslant 1, |L^{-1}(h)| \leqslant 1$$

$$i = 1,\ 2,\ \ldots,\ M$$

where λ represents the membership degree to which the solution belongs to the set "good solution". The value of λ can be determined by a trade-off between the objective

دریافت فوری ← متن کامل مقاله

ISIArticles
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