Available online at www.sciencedirect.com
ScienceDirect
Neurocomputing 70 (2007) 2959–2968
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Functional-link net with fuzzy integral for bankruptcy prediction

Yi-Chung Hu[a,*], Fang-Mei Tseng[b]

[a]Department of Business Administration, Chung Yuan Christian University, Chung-Li, Taiwan, ROC
[b]Department of International Business, Yuan Ze University, Chung-Li, Taiwan, ROC

Received 1 December 2005; received in revised form 4 October 2006; accepted 21 October 2006
Communicated by A. Abraham
Available online 11 November 2006

Abstract

The classification ability of a single-layer perceptron could be improved by considering some enhanced features. In particular, this form of neural networks is called a functional-link net. In the output neuron's activation function, such as the sigmoid function, an inner product of a connection weight vector with an input vector is computed. However, since the input features are not independent of each other for the enhanced pattern, an assumption of the additivity is not reasonable. This paper employs a non-additive technique, namely the fuzzy integral, to aggregate performance values for an input pattern by interpreting each of the connection weights as a fuzzy measure of the corresponding feature. A learning algorithm with the genetic algorithm is then designed to automatically find connection weights. The sample data for bankruptcy analysis obtained from Moody's Industrial Manuals is considered to examine the classification ability of the proposed method. The results demonstrate that the proposed method performs well in comparison with traditional functional-link net and multivariate techniques.

Keywords: Functional-link net; Fuzzy integral; Bankruptcy prediction; Genetic algorithm; Neural networks

1. Introduction

Functional-link net methodology was proposed by Pao [27,28]. It has been applied to many real problems, such as recognition of English language script [30], MPEG-4 video compression [29], the determination of transport modes [13], and the design of robot controller [12]. A functional-link net is a single-layer perceptron (SLP); however, an enhanced pattern with the tensor representation can be presented to a functional-link net rather than the original input pattern. For instance, for an input vector $\mathbf{x} = (x_1, x_2)$ with two features, its tensor is (x_1, x_2, x_1x_2) with an enhanced feature (i.e., x_1x_2), where values of x_1 and x_2 are treated as performances on the first and second features for $\mathbf{x}$, respectively. Thus, in an exclusive-OR problem, which is a non-linearly separable problem, consisting of four patterns, $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$, the corresponding enhanced patterns with the tensor representation of these patterns are $(0, 0, 0)$, $(0, 1, 0)$, $(1, 0, 0)$, and $(1, 1, 1)$, respectively. It is seen that an exclusive-OR problem is thus transformed as a linearly separable problem. In other words, it is possible that the classification performance of a SLP is improved by considering some enhanced features. Although original inputs could be enhanced by other complex methods such as functional expansion, they are not the focus of this paper.

For two-class classification problems, such as the bankruptcy analysis considered in this paper, since a desired output of an input pattern is either zero or one, a functional-link net with single output node is concerned. Furthermore, the sigmoid function whose output ranged between 0 and 1 is usually used as the output node's activation function. The advantage of using the sigmoid function is that errors between actual and desired outputs of individual training patterns can be easily measured [16]. Undoubtedly, the output of the sigmoid function is obtained by computing a weighted sum or an inner product of the connection weights with an enhanced pattern. That is, an aggregated value for an enhanced

*Corresponding author. Tel.: +886 3 2655130; fax: +886 3 2655199.
E-mail address: ychu@cycu.edu.tw (Y.-C. Hu).

0925-2312/$ - see front matter
doi:10.1016/j.neucom.2006.10.111

pattern is obtained by computing the sum of product terms in the sigmoid function. This indicates that the additivity property [25] of the interaction among individual features in an enhanced pattern is assumed. However, it is clear that an assumption of additivity for the tensor representation is not reasonable.

It is considered that the fuzzy integral proposed by Sugeno [35,36] is an appropriate technique for aggregating the performance values of individual features for an enhanced pattern, since it does not assume the independence of one feature from another, and considers the implicit interrelation between features. The main contribution of this paper is to propose a novel fuzzy integral-based functional-link net (FIFLN) by replacing the weighted sum with the fuzzy integral in the sigmoid function. In FIFLN, since a fuzzy measure representing the grade of importance of the corresponding feature is used with fuzzy integral to aggregate performance values [12,13,16,25,35], each connection weight is considered to denote a fuzzy measure. In order to determine the appropriate connection weights in a functional-link net with high classification performance and low training error, we design a genetic algorithm (GA) [10,22,32] to automatically find these parameters. Since bankruptcy or financial distress analysis has long been an important classification problem for a business, a focus of this paper is to compare the classification performance of the FIFLN with that of the traditional functional-link net and SLP for bankruptcy prediction.

The rest of this paper is organized as follows. The traditional functional-link net, the FIFLN and the FIFLN learning algorithm are introduced in Sections 2, 3, and 4, respectively. In Section 5, the proposed FIFLN is evaluated by computer simulations on the sample data for bankruptcy analysis obtained from Moody's Industrial Manuals during the period 1975–1982. It can be seen that the FIFLN outperforms the traditional functional-link net and SLP. Also, it performs well in comparison with some widely used multivariate techniques. The future research and conclusions are presented in Sections 6 and 7, respectively.

2. Functional-link net with weighted sum

The input pattern of a functional-link net is either a functional expansion representation (e.g., sinusoidal function), tensor representation or a combination of these two representations. Since Pao [27] demonstrated the effectiveness of the tensor representation in classification, the tensor model is considered in this paper. In the tensor model, a component of $\mathbf{x}$ (i.e., $(x_1, x_2, \ldots, x_n)$), say x_i, can be enhanced as x_ix_j, $x_ix_jx_k$, and $x_ix_jx_kx_l$ and so on, where $i \leqslant j \leqslant k \leqslant l$. It is seen that an original tensor representation contains many higher-order terms. For instance, an enhanced pattern of $\mathbf{x} = (x_1, x_2, x_3)$ with the tensor representation can be generated as $(x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3, x_2x_3, x_1^2x_2, x_1^2x_3, x_1x_2^2, x_2^2x_3, x_1x_3^2, x_2x_3^2, x_1^3, x_2^3, x_3^3, x_1x_2x_3)$. It can be seen that a large number of terms in the tensor representation will be generated as the dimensions of x increase. Hence, Pao [27] suggested that higher-order terms beyond the second order such as $x_2x_3^2$ and $x_1x_2x_3$, are not required. In addition, two or more equal indices in the enhanced pattern should be omitted. In other words, $i<j$ is considered for x_ix_j. For instance, an acceptable tensor representation of (x_1, x_2, x_3) is $(x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3)$.

As depicted in Fig. 1, a functional-link net with single output node and $n+n(n-1)/2$ input nodes is a one-layer feed-forward network, where y is the actual output corresponding to the input pattern $\mathbf{x}$, and n is the number of input features. It is seen that a tensor of $\mathbf{x}$, say $\mathbf{x}_t$ (i.e., $(x_1, x_2, \ldots, x_n, x_1x_2, x_1x_3, \ldots, x_{n-1}x_n)$), is generated from $\mathbf{x}$ through the functional link. Subsequently, this enhanced pattern (i.e., $\mathbf{x}_t$) with $n+n(n-1)/2$ terms is presented to a functional-link net.

Let f_h denote the output node's activation function or sigmoid function, and θ be a bias in f_h. Without losing generality, f_h is defined as follows:

$$f_h(u) = \frac{1}{1+e^{-u}}, \tag{1}$$

where u is equal to $\mathbf{wx}_t - \theta$ such that $y = f_h(u)$. For $\mathbf{x}_t$, $\mathbf{wx}_t = w_1x_1 + \cdots + w_nx_n + w_{n+1}x_1x_2 + \cdots + w_{n+n(n-1)/2}x_{n-1}x_n$ is an aggregated value which is the inner product of $\mathbf{w}$ with $\mathbf{x}_t$. It is seen that the additivity assumption of the interaction among $x_1, x_2, \ldots, x_n, x_1x_2, x_1x_3, \ldots,$ and $x_{n-1}x_n$ holds. In other words, these terms are assumed to be independent of each other. Whether an input belongs to one of two classes can be decided by computing $f_h(u)$. In principle, $\mathbf{x}$ can be categorized as one class or the other class if the sigmoid function's output value is not below or below 0.5 [13]. Thus, square errors denoted by E between the actual and desired outputs of individual training patterns can be measured as

$$E = \frac{1}{2}\sum_{i=1}^{m}(d_i - y_i)^2, \tag{2}$$

where d_i and y_i are the desired output and the actual output of the ith input training pattern, say $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in})$,

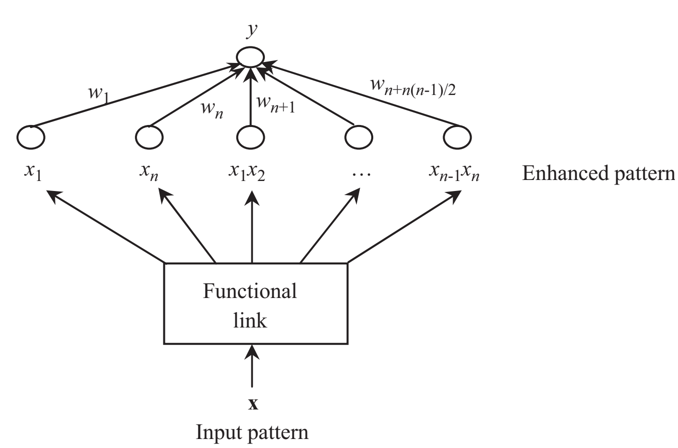

Fig. 1. An architecture of a functional-link net.