Input Space Partitioning for Neural Network Learning

Shujuan Guo, School of Electronic & Information Engineering, Xi’an Jiaotong University, Xi’an, China

Sheng-Uei Guan, School of Electronic & Information Engineering, Xi’an Jiaotong University, Xi’an, China

Weifan Li, Dept. of Computer Science and Software Engineering, Xi’an Jiaotong-Liverpool University, Suzhou, China

Ka Lok Man, Dept. of Computer Science and Software Engineering, Xi’an Jiaotong-Liverpool University, Suzhou, China

Fei Liu, Department of Computer Science & Computer Engineering, La Trobe University, Melbourne, VIC, Australia

A. K. Qin, School of Computer Science and Information Technology, RMIT University, Melbourne, VIC, Australia

ABSTRACT

To improve the learning performance of neural network (NN), this paper introduces an input attribute grouping based NN ensemble method. All of the input attributes are partitioned into exclusive groups according to the degree of inter-attribute promotion or correlation that quantifies the supportive interactions between attributes. After partitioning, multiple NNs are trained by taking each group of attributes as their respective inputs. The final classification result is obtained by integrating the results from each NN. Experimental results on several UCI datasets demonstrate the effectiveness of the proposed method.

Keywords: Correlation, Input Attributes, Neural Network, Partitioning, Promotion

INTRODUCTION

Neural Network (NN) is a supervised machine learning technique, which is typically employed to solve classification problems. When solving a classification problem with the conventional NN, the input data fed into the NN often consists of multiple attributes of various properties. However, training the NN with all of the available input attributes may not lead to the desirable performance considering the curse of dimensionality.

DOI: 10.4018/jaec.2013040105
Dimension reduction was proposed for solving higher dimensional problems but it has its limitations. The distribution of input attributes has also been studied in various research such as the neuro-fuzzy approach (Konstantaras, Varley, Vallianatos, Collins, & Holifield, 2006) and the random subspace method (Ho, 1998). Unfortunately, these research works did not investigate the relationship between input attributes.

Ang et al. (Ang, Guan, Tan, & Mamun, 2008) developed an algorithm for computing the relationship among input attributes. Although Ang et al. partitioned the input space, all input attributes were trained together in the final step. That is, interference was avoided by partitioning in the beginning yet reintroduced in the final stage of training.

In fact, there might exist supportive or interfering interactions amongst input attributes with supportive interactions likely to improve the performance and interfering interactions likely to degrade the performance. Intuitively, if input attributes with positive interactions are trained together, good performance could be achieved. To improve the performance of neural network learning, the input attributes are grouped according to the degree of promotion or correlation. A set of constructive back-propagation (CBP) neural networks are trained by taking each group as their respective inputs with training results being integrated using some ensemble learning technique. By this way, attributes with interference will never be trained together in the whole procedure, and the training accuracy can be greatly improved.

This paper is structured as follows. Section 2 defines some terms related to the measurement of promotion interaction between input attributes and related concepts. Sections 3 presents our approach to attribute partitioning and grouping. Constructive neural network is used in our training of neural sub-networks for each group while some ensemble approach is adopted for results integration from sub-networks. Section 4 provides our experimental results, analysis and comparison to related work. Section 5 concludes the paper with some suggestions for future research.

TERMINOLOGY AND CONCEPTS

Let \( \mu(i, j) \) \((i \neq j) \) denote the promotion rate of two attributes \( i \) and \( j \), which is defined by:

\[
\mu(i, j) = \begin{cases} 
1, & \text{where } E_{ij} \leq \min\{E_i, E_j\} \\
0, & \text{where } E_{ij} > \min\{E_i, E_j\}
\end{cases}
\]  

where \( E_i \) represents the classification error obtained by training with single attribute \( i \), and \( E_{ij} \) represents the classification error obtained by training with two attributes \( i \) and \( j \). When the promotion rate of two attributes is 1, these two attributes are considered to be mutually supportive for classification. Otherwise, they are considered to be mutually interfered for classification.

To take full advantage of inter-attribute promotions, we compute the average value of all \( E_{ij} \) with \( \mu(i, j) = 1 \). Any pair of attributes whose classification error is less than this average value is considered to have significant promotion to each other. The smaller the corresponding classification error is, the more significant the promotion is.

In statistics, correlation measures the strength and direction of the linear relationship between two random variables. There exist many ways of calculating correlation. This paper employs the Pearson’s correlation coefficients (Sedgwick, 2012).

We use \( P \) value and significance level (\( \alpha \)) (Sedgwick, 2012) to identify whether two attributes are correlated to each other. If \( P < \alpha \) (usually taking the value of 0.1 or 0.05), it is considered that two attributes are correlated to each other. To distinguish the degree of correlation, the mean value of correlation coefficients of all the correlated attributes is calculated and employed as the threshold. Any pair of correlated attributes whose correlation coefficient is greater than this threshold is considered to have strong correlation to each other. The larger the corresponding correlation coefficient is, the stronger the correlation is.