Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 32 (2007) 1–11
Expert Systems with Applications
www.elsevier.com/locate/eswa

Artificial neural networks classification and clustering of methodologies and applications – literature analysis from 1995 to 2005

Shu-Hsien Liao [a], Chih-Hao Wen [b]

[a] Department of Management Sciences and Decision Making, Tamkang University, No. 151, Yingjuan Road, Danshuei Jen, Taipei 251, Taiwan, ROC
[b] Graduate School of Resource Management, National Defence University, Management College, P.O. Box 90046-17 Jon-Ho, Taipei County, Taiwan, ROC

Abstract

Based on a scope of 10,120 articles on ANNs, this paper uses data mining including association rules and cluster analysis, to survey these ANNs papers through keyword classification and clustering of articles from 1995 to 2005, exploring the ANNs methodologies and application developments during that period. The four decision variables of keywords, author's nationality, research category, and year of publication, are implemented for data mining with total of 110,080 data items. The research findings show that some specific ANNs methodologies and applications pattern can be extracted from the mining results, and these describe the ANNs development over this period. In addition, using more data mining approaches for analysis could provide different explanations for ANNs development. Finally, discussion and brief conclusion are presented.

Keywords: Artificial neural networks; Artificial neural networks methodologies and applications; Data mining; Association rule; Cluster analysis; Literature analysis

1. Introduction

Artificial neural networks (ANNs) are computational modeling tools that have been extensively used in many disciplines to model complex real-world problems. ANNs may be defined as structures comprised of densely interconnected adaptive simple processing elements (called artificial neurons or nodes) that are capable of performing massive parallel computations for data processing and knowledge representation (Schalkoff, 1997). Although ANNs are drastic abstractions of their biological counterparts, the idea of ANNs is not to replicate the operation of the biological systems but to use what is known about the functionality of the biological networks in order to solve complex problems. Neural networks discover patterns and relationships in massive amounts of data by using hardware and software that simulates the processing patterns of the human brain. Contrary to conventional computing methods, ANNs are 'trained' to produce the desired input–output relationships. During the training (learning) phase, examples of data are presented to the network and, using a learning algorithm, the parameters are tuned to adjust the network behavior. According to available knowledge of the problem, and the objective of the operator, the learning procedure employed can be either 'supervised', 'unsupervised', or both. The supervised learning procedure is performed with pairs of known input–output patterns, whereas unsupervised learning consists of presenting training examples to the network input while the network organizes itself progressively to reach maximal separation between the naturally occurring classes of examples. The principal applications of ANNs have been in the area of pattern recognition. Here, a pattern is turned into a

E-mail addresses: michael@mail.tku.edu.tw (S.-H. Liao), 6060101zq-az@yahoo.com.tw (C.-H. Wen).

0957-4174/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2005.11.014

feature vector, which is then used as ANN input, and the output is interpreted as identifying the input to be a member of one of a number of classes of possible inputs. An important quality of neural networks (referred to as generalization) is that when they are correctly trained, neural networks can appropriately process data that have not been used for training. The most popular neural networks are multilayer perceptions, which are generally supervised–trained with the error back-propagation algorithm (Rumelhart, Hinton, & Williams, 1986). One major property of these networks is their ability to find non-linear surfaces separating the underlying patterns, which is generally considered to be an improvement on conventional methods.

On the other hand, radial basis function (RBF) network is a particular class of multilayer networks (Poggio & Girosi, 1989), where learning occurs usually in two stages: learning in the hidden layer (usually by an unsupervised bottom–up self-organizing method such as *K*-means clustering) followed by the output layer (a top–down supervised method such as least squares estimation). RBF networks have two important advantages: finding the input to output map using local approximators, and rapid learning requiring fewer examples. Another popular class of networks is the self-organizing map, or Kohonen network (Kohonen, 1988). A Kohonen network consists of two fully connected-unit layers. The output layer is generally ordered in a low-dimensional framework (a one-dimensional array or a two-dimensional matrix) of units. The objective of this network is to build a map where units of an area are activated when inputs with similar characteristics are presented. Among the other popular networks are adaptive resonance theory (ART) networks and their derivates (ART1, ART2, fuzzy ART, etc.) (Carpenter & Grossberg, 1992) and Hopfield models (Hopfield, 1982). One quality of neural networks is that they can be considered as non-linear statistical methods. Nevertheless, a large amount of data is required to overcome the existing non-linearities in the data structure. Therefore, the attractiveness of ANNs comes from the remarkable information processing characteristics of a biological system such as non-linearity, high parallelism, robustness, fault and failure tolerance, learning, ability to handle imprecise and fuzzy information, and their capability to generalize.

This research began on June 2005, and it was first based on a search in the keyword index of article for 'neural networks' on the Elsevier SDOS, IEEE Xplore, Blackwell Synergy, Wiley InterScience, and Taylor & Francis online database, for the period from 1995 to 2005, in which 10,120 articles were found on 30 June 2005. Based on a scope of 10,120 articles on ANNs, this paper uses a data mining approach including association rules and clustering analysis to survey the ANNs through keyword classification and clustering of articles from 1995 to 2005. Based on this, the ANNs methodologies and applications development during that period are explored. The rest of the paper is organized as follows. Section 2 introduces research design. Sections 3 and 4 describe the data mining process and its results. Section 5 discusses our findings, and Section 6 contains a brief conclusion.

2. Research design

2.1. Why keywords are used as data source for literature analysis?

Basically, there are three aspects for research, including research problem, theory, and methods/tools. Sometimes, the research objective and motivation is described as the research problem, which contains the nature of a specific problem domain. In addition, a problem domain might be a person, a group of people, an organization, an object, or an event, which can be estimated and evaluated. This means that a problem domain has specific domain knowledge with certain complexity and that this knowledge may exist either tacitly or explicitly under different situations and conditions. This form of domain knowledge could be a threshold to researchers who have limited understanding of this problem domain. This means that researchers should understand or learn something on research problem domain knowledge before starting their research. On the other hand, a research problem may describe a process or procedure of research experiment, case study, system development, or modeling on a specific research field. These processes or procedures focus on a limited scope that social science or natural science can manipulate and observe the research problem by collecting and analyzing data from the problem domain. In this regard, a study must define the research process or procedure in order to illustrate the research scope from real-world situations or conditions. A research problem is a logical description to show why a study is critical or what is its contribution and originality in a specific problem domain. Without good enough domain knowledge, the research problem can be the most difficulty part of work at the initial stage of research. Because a researcher must spend great amount of time to understand the domain knowledge deeply and then can provide a basis to understand and define the research scope.

On the other hand, once the scope of a research problem domain has been determined, researcher can describe research problem by citing specific theory. A theory is generalized phenomena, which presents results or findings from specific research experiment or long term observation with systematic approach. However, a phenomenon can be extended to different theories on different research problem domain. For example, in physics, the principle of inertia states that objects continue in a state of rest or of uniform motion unless acted upon by forces. In knowledge management, the knowledge inertia describes that people may stem from the use of routine problem solving procedures, stagnant knowledge sources, and following past experience or knowledge (Liao, 2002). Due to diversity of research problem domains, researchers must to cite theory or read literature in relation to the research problem in order to fit how people refer to a phenomenon in a specific problem domain.