

Improved Wavelet Neural Networks and Its Applications in Function Approximation

N

Zarita Zainuddin

School of Mathematical Sciences, Universiti Sains Malaysia, Malaysia

Ong Pauline

Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Malaysia

INTRODUCTION

The advent of the computer is a blessing to the emergence of artificial neural networks (ANNs) – a research area which has significant links to the effort of scientists to emulate the complex functions of the human brain. The ANNs, which conceptually model the human brain metaphor with vast number of inter-connecting artificial neurons, experience an upsurge in popularity due to their effectiveness and suitability for solving large-scale problems where the physical processes are highly complex and ill-defined. Furthermore, in diverse domains of science and engineering applications, the problems at hand usually do not have an algorithmic solution, or the solution is too complex to be obtained by using the conventional methods. The ANNs with the properties of adaptive learning ability and nonparametric characteristic are thus extremely vital in this undertaking.

Extensive and intensive research in ANNs have spawned a variety of different network paradigms, from which the integration of wavelets and feedforward neural networks has fruits in a special variant of ANNs, namely, the wavelet neural networks (WNNs) (Zhang & Benveniste, 1992). Constituting of wavelet activation functions as the nonlinear transfer function in its modeling framework, the WNNs feature the eye-catching characteristics of: (1) preserving the universal approximation property; (2) establishing the explicit link between the networks coefficients and wavelet transform; and (3) achieving same quality of approximation with a network of reduced size. Moreover, its superiority in alleviating the deficiencies of the popular multilayer perceptrons, which are subject to slow learning and local minima problems, has been asserted. Given this capabilities, the WNNs are unceas-

ingly gaining their prevalence and popularity in a large range of practical situations, for instance, forecasting the draught conditions (Belayneh, Adamowski, Khalil, & Ozga-Zielinski, 2014), monitoring the harmonic pollution in power-electronics-based devices (Jain & Singh, 2014), predicting the pulp and paper properties (Zainuddin, Daud, Ong, & Shafie, 2011), detecting microbial contaminants in clinical samples (Kodogiannis, 2013), approximating the wind speed (Yao & Yu, 2013), classifying the heterogeneous cancer profiles (Zainuddin & Ong, 2011b), and diagnosing the early stage of diseases (Dheeba, Singh, & Singh, 2014; San, Ling, & Nguyen, 2013; Zainuddin & Ong, 2012), are some of the real world uses for which the WNNs have demonstrated their excellent viability.

Since numerous application problems are highly dependent on the accuracy of WNNs, in view of the fact that the accuracy is always a key factor in decision making, improving the predictive competence of WNNs therein forms the basis of this study. A host of approaches have been put forward in this regard; including the utilization of different learning algorithms in the search for perfection in optimizing the weight vectors of WNNs (Chen, 2011; Tzeng, 2010), the alteration of network's architecture (Wan, Li, & Cheng, 2004), the modification of the types of activation functions used in the hidden layer (Zainuddin & Ong, 2011a), the exploitation of fuzzy rules in knowledge representation during the learning process of WNNs (Abiyev, 2011; Liu, Wu, Zhang, Wang, & Chen, 2011), and the initialization of wavelet translation vectors (Zainuddin & Ong, 2011b, 2012), which is the core interest of this work. For a plausible convergence, as well as for superb generalization performance; a good initialization of the appropriate number of wavelet functions and their locations (translation vectors) are

DOI: 10.4018/978-1-4666-5888-2.ch627

extremely crucial. Determining the initial locations of the translation vectors judiciously will effectively reflect upon the essential attributes residing within the input space, and thus, the WNNs can model the underlying mapping between the input-output spaces progressively from a set of good starting points, which potentially improve the validity of modeling. Moreover, with proper initialization of the network's architecture, a better configuration of WNNs with less computational complexity can be achieved.

To serve for this intention, a novel clustering algorithm – the symmetry similarity fuzzy C-means (SS-FCM) was proposed in this study, as an alternative to the existing WNNs translation vectors initialization approaches. The main thrust of this study is thus geared towards an improved WNNs model, incorporating the SS-FCM algorithm in the network preprocessing framework, for the purpose of function approximation. The objectives of this study can be summarized as follows:

- To introduce the WNNs as an alternative tool to function approximation problems
- To develop and apply a novel technique, i.e., the SS-FCM algorithm in the translation vectors initialization of WNNs for the purpose of optimizing its generalization ability

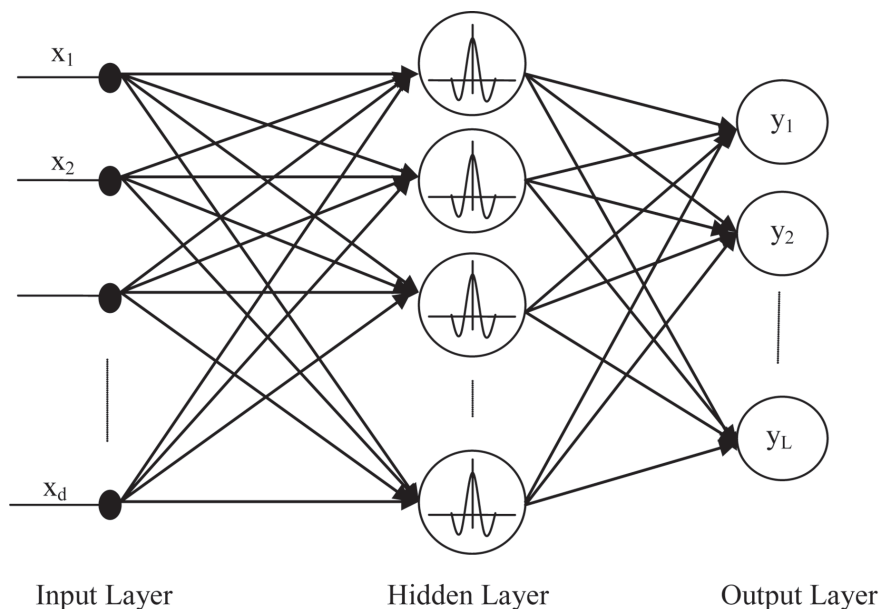
- To assess the potential benefits of the proposed WNNs in the context of function approximation, through empirical approaches with simulated as well as real world application problem, i.e., prediction of the blood glucose level for diabetics.

BACKGROUND

Introduction to Wavelet Neural Networks

Inspired by the similarities between the single hidden layer feedforward neural networks with the wavelet decomposition, Zhang and Benveniste (1992) proposed an alternative neural network model, namely, the WNNs for approximating functions. In fact, WNNs are merely different from the traditional ANNs in its adoption of wavelet functions as the activation functions in the hidden layer. The commonly used wavelet activation functions are Gaussian wavelet, Mexican Hat, and Morlet. In the following sections, a brief introduction of the WNNs structure, learning algorithm, and parameter initialization is given.

Figure 1. The schematic architecture of wavelet neural networks



16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/improved-wavelet-neural-networks-and-its-applications-in-function-approximation/113094

Related Content

Use of Grey Numbers for Evaluating a System's Performance Under Fuzzy Conditions

Michael Voskoglou (2021). *Encyclopedia of Information Science and Technology, Fifth Edition* (pp. 315-331).

www.irma-international.org/chapter/use-of-grey-numbers-for-evaluating-a-systems-performance-under-fuzzy-conditions/260194

Early Warning of Companies' Credit Risk Based on Machine Learning

Benyan Tanand Yujie Lin (2023). *International Journal of Information Technologies and Systems Approach* (pp. 1-21).

www.irma-international.org/article/early-warning-of-companies-credit-risk-based-on-machine-learning/324067

Understanding and Assessing Quality of Models and Modeling Languages

John Krogstie (2018). *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 4810-4821).

www.irma-international.org/chapter/understanding-and-assessing-quality-of-models-and-modeling-languages/184185

An Optimal Routing Algorithm for Internet of Things Enabling Technologies

Amol V. Dhumane, Rajesh S. Prasad and Jayashree R. Prasad (2017). *International Journal of Rough Sets and Data Analysis* (pp. 1-16).

www.irma-international.org/article/an-optimal-routing-algorithm-for-internet-of-things-enabling-technologies/182288

Clique Size and Centrality Metrics for Analysis of Real-World Network Graphs

Natarajan Meghanathan (2018). *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 6507-6521).

www.irma-international.org/chapter/clique-size-and-centrality-metrics-for-analysis-of-real-world-network-graphs/184347