

Chapter XIX

A New Topology for Artificial Higher Order Neural Networks: Polynomial Kernel Networks

Zhao Lu

Tuskegee University, USA

Leang-san Shieh

University of Houston, USA

Guanrong Chen

City University of Hong Kong, China

ABSTRACT

Aiming to develop a systematic approach for optimizing the structure of artificial higher order neural networks (HONN) for system modeling and function approximation, a new HONN topology, namely polynomial kernel networks, is proposed in this chapter. Structurally, the polynomial kernel network can be viewed as a three-layer feedforward neural network with a special polynomial activation function for the nodes in the hidden layer. The new network is equivalent to a HONN; however, due to the underlying connections with polynomial kernel support vector machines, the weights and the structure of the network can be determined simultaneously using structural risk minimization. The advantage of the topology of the polynomial kernel network and the use of a support vector kernel expansion paves the way to represent nonlinear functions or systems, and underpins some advanced analysis of the network performance. In this chapter, from the perspective of network complexity, both quadratic programming and linear programming based training of the polynomial kernel network are investigated.

INTRODUCTION

As an important neural processing topology, artificial higher order neural networks (HONNs) have demonstrated great potential for approximating unknown functions and modeling unknown systems (Kosmatopoulos et al., 1995; Kosmatopoulos and Christodoulou, 1997). In particular, HONNs have been adopted as basic modules in the construction of dynamic system identifiers and also controllers for highly uncertain systems (Rovithakis, 1999; Lu et al., 2006). Nevertheless, as an important factor that affects the performance of neural networks, the structure of a network is usually hard to determine appropriately in any specific application. It is possible to reduce modeling errors by increasing the complexity of the network; however, increasing the complexity may overfit the data leading to a degradation of its generalization ability. As a consequence, in practice the choice of network structure is often a compromise between modeling errors and the network complexity. Some efforts have been made in the attempt to determine the optimal topological structure of HONN by using for example genetic algorithms (Rovithakis et al., 2004).

Recently, there has been a trend in the machine learning community to construct a nonlinear version of a linear algorithm using the so-called ‘kernel method’ (Schölkopf and Smola, 2002; Vert et al., 2004). As a new generation of learning algorithms, the kernel method utilizes techniques from optimization, statistics, and functional analysis to achieve maximal generality, flexibility, and performance. The kernel machine allows high-dimensional inner-product computations to be performed with very little overhead and brings all the benefits of the mature linear estimation theory. Of particular significance is the Support Vector Machine (SVM) that forms an important subject in the learning theory. SVM is derived from statistical learning theory (Evgeniou et al., 2000; Cristianini and Shawe-Taylor, 2000), which is a two-layer network with inputs transformed

by the kernels corresponding to a subset of the input data, while its output is a linear function of the weights and kernels. The weights and the structure of the SVM are obtained simultaneously by a constrained minimization at a given precision level of the modeling errors. For all these reasons, the kernel methods have become more and more popular as an alternative to neural-network approaches. However, due to the fact that SVM is basically a non-parametric technique, its effective use in dynamical systems and control theory remains to be seen.

Actually, SVM includes a number of heuristic algorithms as special cases. The relationships between SVM and radial basis function (RBF) networks, neuro-fuzzy networks and multilayer perceptron have been accentuated and utilized for developing new learning algorithms (Chan et al., 2001; Chan et al., 2002; Suykens and Vandewalle, 1999). Of particular interest is a recent observation that Wiener and Volterra theories, which extend the standard convolution description of linear systems by a series of polynomial integral operators with increasing degrees of nonlinearity, can be put into a kernel regression framework (Franz and Schölkopf, 2006).

Inspired by the unifying view of Wiener and Volterra theories and polynomial kernel regression, provided by (Franz and Schölkopf, 2006), and by the fact that the Wiener expansion decomposes a signal according to the order of interaction of its input elements, in this chapter a new topology for HONN, called the polynomial kernel network, is proposed and investigated, which bridges the gap between the parametric HONN model and the non-parametric support vector regression model.

Structurally, the polynomial kernel network can be viewed as a three-layer feedforward neural network with a special polynomial activation function for the nodes in the hidden layer. Due to the equivalence between the proposed polynomial kernel network and the support vector machine with an inhomogeneous polynomial

10 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/new-topology-artificial-higher-order/5293

Related Content

Deep Clustering

M. Parimala Boobalan (2019). *Handbook of Research on Deep Learning Innovations and Trends* (pp. 164-179).

www.irma-international.org/chapter/deep-clustering/227851

Bipolar Neutrosophic Cubic Graphs and Its Applications

C. Antony Crispin Sweetey, K. Vaiyomathiand F. Nirmala Irudayam (2020). *Handbook of Research on Advanced Applications of Graph Theory in Modern Society* (pp. 492-536).

www.irma-international.org/chapter/bipolar-neutrosophic-cubic-graphs-and-its-applications/235550

Hybrid-Learning Methods for Stock Index Modeling

Yuehui Chenand Ajith Abraham (2006). *Artificial Neural Networks in Finance and Manufacturing* (pp. 64-79).

www.irma-international.org/chapter/hybrid-learning-methods-stock-index/5349

Optimum Design of Carbon Fiber-Reinforced Polymer (CFRP) Beams for Shear Capacity via Machine Learning Methods: Optimum Prediction Methods on Advance Ensemble Algorithms – Bagging Combinations

Melda Yucel, Aylin Ece Kayabekir, Sinan Melih Nigdeliand Gebrail Bekda (2020). *Artificial Intelligence and Machine Learning Applications in Civil, Mechanical, and Industrial Engineering* (pp. 85-103).

www.irma-international.org/chapter/optimum-design-of-carbon-fiber-reinforced-polymer-cfrp-beams-for-shear-capacity-via-machine-learning-methods/238140

Deep Reinforcement Learning for Optimization

Md Mahmudul Hasan, Md Shahinur Rahmanand Adrian Bell (2019). *Handbook of Research on Deep Learning Innovations and Trends* (pp. 180-196).

www.irma-international.org/chapter/deep-reinforcement-learning-for-optimization/227852