

# Chapter 41

## Utilizing Feature Selection on Higher Order Neural Networks

**Zongyuan Zhao**

*University of Tasmania, Australia*

**Shuxiang Xu**

*University of Tasmania, Australia*

**Byeong Ho Kang**

*University of Tasmania, Australia*

**Mir Md Jahangir Kabir**

*University of Tasmania, Australia*

**Yunling Liu**

*China Agricultural University, China*

**Rainer Wasinger**

*University of Tasmania, Australia*

### ABSTRACT

*Artificial Neural Network has shown its impressive ability on many real world problems such as pattern recognition, classification and function approximation. An extension of ANN, higher order neural network (HONN), improves ANN's computational and learning capabilities. However, the large number of higher order attributes leads to long learning time and complex network structure. Some irrelevant higher order attributes can also hinder the performance of HONN. In this chapter, feature selection algorithms will be used to simplify HONN architecture. Comparisons of fully connected HONN with feature selected HONN demonstrate that proper feature selection can be effective on decreasing number of inputs, reducing computational time, and improving prediction accuracy of HONN.*

### INTRODUCTION

Artificial Neural Network (ANN) is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use (Haykin, 1999). It was motivated by inspecting the human brain, which has high efficiency in computing and recognizing (West, 2000). ANN has been successfully applied to applications involving pattern classification and function approximation (Shin & Ghosh, 1991).

Although some training algorithms of ANN, such as back propagation (BP) have shown great performance, ANNs often take long time to converge and may stuck in local minima (Fulcher, Zhang, & Xu, 2006). Also, ANNs are not suitable for discontinuous data (Zhang, 2008). Moreover, the explanations

DOI: 10.4018/978-1-5225-0788-8.ch041

for ANNs' output is not obvious (Spirkovska & Reid, 1990). These shortages of ANNs are the motivations of the development of Higher Order Neural Network (HONN).

At first, HONN was designed to promote ANN's computational, storage and learning capabilities, as the order or structure of HONN can be tailored to the order or structure of a problem (Giles & Maxwell, 1987). They also prove that when a priori knowledge, such as geometric invariances, is encoded in HONN, the network becomes more efficient in solving problems that utilize this knowledge.

This character of HONN provides a solution for invariant pattern recognition problems (Redding, Kowalczyk, & Downs, 1993). As a preprocessing for BP ANN, HONN can be designed to be invariant to changes in scale, translation, and in-plane rotation (Schmidt & Davis, 1993). As invariances are built directly into the architecture of HONN and do not need to be learned, the training time can be shortened and a smaller training set is required (Spirkovska & Reid, 1993). Some researchers also provide pruning algorithms to decrease the complexity of HONN by reducing the number of network weights (Kosmatopoulos, Polycarpou, Christodoulou, & Ioannou, 1995; Li, Wang, Li, Zhang, & Jinyan, 1998).

Unfortunately, proper higher order attributes can only be chosen manually by expert knowledge. This can be achieved in visual object recognition, but not fit for other tasks such as function approximation and stock prediction, as the higher order features in these dataset are meaningless. Feature Selection methods for ANN provides a solution of this problem by reducing the number of attributes of HONN.

Feature Selection (FS), also known as attribute selection, or variable selection, is the process of selecting a subset of relevant features for use in computational model construction (Chakraborty & Pal, 2015). Actually, it can be treated as a searching procedure: search for an acceptable feature subset evaluated by certain criterion from the original dataset (Yukyee & Yeungsam, 2010). Benefits provided by FS include improved model interpretability, shorter training times, and enhanced generalization by reducing over-fitting (Luping, Lei, & Chunhua, 2010).

Feature Selection models can be divided into filters and wrappers according to whether it depends on data mining structure or purely by datasets. Filters focus on general characters of instances and take no consider about data mining algorithms (Huan & Lei, 2005). Filters choose only "good" features which are more representative or contain more information. It is commonly used as a pre-process of large scale data. Wrappers use predetermined data mining algorithm to evaluate the performance of each feature subsets. They are more time consuming than filters but also has greater influence on the data mining performance (Kohavi & John, 1997).

In this paper, we will present the utilization of FS filters on HONN by comparing the networks' performance. The second and third sections describe the background research of HONN and FS respectively. After that we will present our experiment and the end the conclusion and some future research will be discussed.

## **BACKGROUND OF HIGHER ORDER NEURAL NETWORK**

Among ANNs, Multilayer Perceptron (MLP) has been shown to be capable of approximating generic classes of functions. The research in paper (Leshno, Lin, Pinkus, & Schocken, 1993) has proved that a three layer MLP with one hidden layer can perform well enough as a universal approximation. A typical three layer MLP can be described in Figure 1.

We use the following notations

14 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/utilizing-feature-selection-on-higher-order-neural-networks/161063](http://www.igi-global.com/chapter/utilizing-feature-selection-on-higher-order-neural-networks/161063)

## Related Content

---

### Real-Time Anomaly Detection Using Facebook Prophet

Nithish T., Geeta R. Bharamagoudar, Karibasappa K. G. and Shashikumar G. Totad (2021). *International Journal of Natural Computing Research* (pp. 29-40).

[www.irma-international.org/article/real-time-anomaly-detection-using-facebook-prophet/298998](http://www.irma-international.org/article/real-time-anomaly-detection-using-facebook-prophet/298998)

### Application of the Cloud Computing for the Effective Implementation of the Medical Information System

Ekaterina Kldiashvili (2014). *International Journal of Natural Computing Research* (pp. 52-68).

[www.irma-international.org/article/application-of-the-cloud-computing-for-the-effective-implementation-of-the-medical-information-system/118157](http://www.irma-international.org/article/application-of-the-cloud-computing-for-the-effective-implementation-of-the-medical-information-system/118157)

### Learning: A Psychological Perspective

Frank van der Velde (2011). *System and Circuit Design for Biologically-Inspired Intelligent Learning* (pp. 41-58).

[www.irma-international.org/chapter/learning-psychological-perspective/48890](http://www.irma-international.org/chapter/learning-psychological-perspective/48890)

### Superior Koch Curve

Sanjeev Kumar Prasad (2011). *International Journal of Artificial Life Research* (pp. 24-31).

[www.irma-international.org/article/superior-koch-curve/62070](http://www.irma-international.org/article/superior-koch-curve/62070)

### Swarm Intelligence for Biometric Feature Optimization

Santosh Kumar, Deepanwita Datta and Sanjay Kumar Singh (2017). *Nature-Inspired Computing: Concepts, Methodologies, Tools, and Applications* (pp. 830-863).

[www.irma-international.org/chapter/swarm-intelligence-for-biometric-feature-optimization/161053](http://www.irma-international.org/chapter/swarm-intelligence-for-biometric-feature-optimization/161053)