

Chapter 55

Theoretical Analyses of the Universal Approximation Capability of a Class of Higher Order Neural Networks Based on Approximate Identity

Saeed Panahian Fard
Universiti Sains, Malaysia

Zarita Zainuddin
Universiti Sains, Malaysia

ABSTRACT

One of the most important problems in the theory of approximation functions by means of neural networks is universal approximation capability of neural networks. In this study, we investigate the theoretical analyses of the universal approximation capability of a special class of three layer feedforward higher order neural networks based on the concept of approximate identity in the space of continuous multivariate functions. Moreover, we present theoretical analyses of the universal approximation capability of the networks in the spaces of Lebesgue integrable multivariate functions. The methods used in proving our results are based on the concepts of convolution and epsilon-net. The obtained results can be seen as an attempt towards the development of approximation theory by means of neural networks.

1. INTRODUCTION

The function approximation can be explained as follows: Let f be a continuous function on a compact set. We intend to find a simple function g such that $\|f - g\| < \epsilon$. This problem has attracted many researchers' attentions in the last century. According to Tikk et al. (2003), in 1900, Hilbert presented his 23 conjectures at the second international congress of mathematicians in Paris. Based on the 13th conjecture, there exist continuous functions with multi variables which cannot be represented as the finite superposition

DOI: 10.4018/978-1-5225-0788-8.ch055

of continuous functions with fewer variables. In 1957, Arnold rejected this conjecture. In the same year, Kolmogorov proved his representation theorem with a constructive proof. This theorem shows that a continuous function with multi variables can be decomposed as the finite superposition of continuous functions with one variable. In 1965, Sprecher improved the Kolmogorov's representation theorem. In 1966, Lorentz further improved this theorem.

In the approximation theory of artificial neural networks (ANNs) that problem reduces to find an artificial neural network such that approximate f , i.e. $\|f - \text{ANN}\| < \epsilon$. In 1980, De Figueriedo generalized this theorem for multilayer feedforward artificial neural networks. In 1989, Poggio and Girosi showed that this theorem is irrelevant for artificial neural networks because in a Kolmogorov networks, nodes have wild and complex functions. Then, many researchers have been tried to solve the problem of approximation function by artificial neural networks such as Cybenko (1989), Funuhashi (1989), Park and Sandberg, (1991, 1993), Mhaskar, (1993), Leshno (1993), Suzuki (1998), Hahm and Hong, (2004), Li (2008), Ismailov (2012), Wang et al. (2012), Lin et al. (2013), and Arteaga and Marrero (2013).

The standard form of universal approximation capability of feedforward neural networks states that under what kind of conditions an arbitrary continuous function can be approximated by a single-hidden-layer feedforward neural networks to any degree of accuracy. Comprehensive surveys of the universal approximation capability of feedforward neural networks can be found in Nong (2013); Bouaziz et al. (2014); Wang (2010); Ismailov (2014); Arteaga and Marrero (2014), Costarelli (2014). Recently, the history of the development of universal approximation by artificial neural networks has been presented by Pricipe and Chen (2015).

In the present chapter, we are motivated to extend the scheme of the univariate universal approximation capability to the scheme of the multivariate universal approximation capability. In other words, the motivation of this chapter is to develop the theory of the universal approximation capability of a class of feedforward higher order neural networks based on approximate identity in multivariate functions spaces. We should address what we can expect from higher order neural networks based on approximate identity. In fact, higher order neural networks based on approximate identity are merging higher order neural networks and approximate identity neural networks.

Higher order neural networks extend standard feedforward neural networks by using a higher correlation of input neurons for better fitting properties. Higher order neural networks was first studied by Giles and Maxwell (1987) and further analysed by Pao (1989). The above networks have also been used in pattern recognition (Schmit & Davis, 1993), financial data simulation (Zhang et al., 2002). In recent years, universal approximation capability of different models of higher order neural networks have been investigated (Long et al., 2007; Long et al., 2007a).

Approximate identity neural networks are neural networks models which employ approximate identity as neuron activation functions. We briefly review the literature of this field of research. Turchetti et al. published the first paper on approximate identity neural networks (Turchetti, 1998). In recent years, universal approximation capability of several different types of approximate identity neural networks models have been investigated (Panahian & Zainuddin, 2013, 2014, 2012, 2014a, 2013a, 2014b, 2015, 2014c, 2014d).

The objective of this study is to describe an approximate representation of multivariate functions by a class of three layer feedforward higher order neural networks based on approximate identity in the space of continuous multivariate functions and in the spaces of Lebesgue integrable multivariate functions. In other words, we aim to provide an answer to the question that what kind of three layer feedforward

13 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/theoretical-analyses-of-the-universal-approximation-capability-of-a-class-of-higher-order-neural-networks-based-on-approximate-identity/161078

Related Content

A Mathematical Model for a Vibrating Human Head

J. C. Misra, S. Dandapat and S. Adhikary (2010). *International Journal of Artificial Life Research* (pp. 29-42).
www.irma-international.org/article/mathematical-model-vibrating-human-head/44669

Role of Artificial Intelligence and Machine Learning in Drug Discovery and Drug Repurposing

Sameer Quazi and Zarish Fatima (2023). *Structural and Functional Aspects of Biocomputing Systems for Data Processing* (pp. 182-197).
www.irma-international.org/chapter/role-of-artificial-intelligence-and-machine-learning-in-drug-discovery-and-drug-repurposing/318557

Comparison of Zipper and Non-Zipper Merging Patterns Near Merging Point of Roads

Y. Wakita, T. Iguchi, H. Shimizu, T. Tamaki and E. Kita (2010). *International Journal of Natural Computing Research* (pp. 19-39).
www.irma-international.org/article/comparison-zipper-non-zipper-merging/49123

Gait Analysis Using Principal Component Analysis and Long Short Term Memory Models

Maheswari R., Pattabiraman Venkatasubbu and A. Saleem Raja (2023). *Structural and Functional Aspects of Biocomputing Systems for Data Processing* (pp. 79-97).
www.irma-international.org/chapter/gait-analysis-using-principal-component-analysis-and-long-short-term-memory-models/318552

Computational Analysis of Symmetric and Cambered Blade Darrieus Vertical Axis Wind Turbine With Bio-Mimicked Blade Design

Punit Prakash, Praveen Laws, Nishant Mishra and Santanu Mitra (2022). *Applications of Nature-Inspired Computing in Renewable Energy Systems* (pp. 225-240).
www.irma-international.org/chapter/computational-analysis-of-symmetric-and-cambered-blade-darrieus-vertical-axis-wind-turbine-with-bio-mimicked-blade-design/294394