Analysis in Theory and Applications
Anal. Theory Appl., Vol. **39**, No. 1 (2023), pp. 93-104

DOI: 10.4208/ata.OA-2021-0006
March 2023

On Approximation by Neural Networks with Optimized Activation Functions and Fixed Weights

Dansheng Yu[1,*], Yunyou Qian[1] and Fengjun Li[2]

[1] *Department of Mathematics, Hangzhou Normal University, Hangzhou, Zhejiang 310036, China*
[2] *School of Mathematics and Statistics, Ningxia University, Yinchuan, Ningxia 750021, China*

Received 25 March 2021; Accepted (in revised version) 23 February 2022

Abstract. Recently, Li [16] introduced three kinds of single-hidden layer feed-forward neural networks with optimized piecewise linear activation functions and fixed weights, and obtained the upper and lower bound estimations on the approximation accuracy of the FNNs, for continuous function defined on bounded intervals. In the present paper, we point out that there are some errors both in the definitions of the FNNs and in the proof of the upper estimations in [16]. By using new methods, we also give right approximation rate estimations of the approximation by Li's neural networks.

Key Words: Approximation rate, modulus of continuity, modulus of smoothness, neural network operators.

AMS Subject Classifications: 41A35, 41A25, 41A20

1 Introduction

Feed-forward neural networks (FNNS) have been investigated entensively and deeply because of their universal approximation capbilities on compact input sets and approximation in a finite set. In the present paper, we deal with the FNNS with one hidden layer, which can be mathematically expressed as

$$N_n(x) = \sum_{j=0}^{n} c_j \sigma\left(\langle a_j \cdot x\rangle + b_j\right), \quad x \in \mathbb{R}^s, \quad s \in \mathbb{N},$$

where for $0 \le j \le n, b_j \in \mathbb{R}$ are the thresholds, $a_j \in \mathbb{R}^s$ are the connection weights, $c_j \in \mathbb{R}$ are the coefficients, $\langle a_j \cdot x\rangle$ is the inner product of a_j and x, and σ is the activation

*Corresponding author. *Email addresses:* `dsyu@hznu.edu.cn` (D. Yu), `yunyouqian@yeah.net` (Y. Qian), `fjli@nxu.edu.cn` (F. Li)

function. In many fundamental network models, the activation function σ is usually taken to be a sigmoidal function.

As we know, FNNS are universal approximators. Theoretically, any continuous function defined on a compact set can be approximated to any desired degree of accuracy by increasing the number of hidden neurons. A lot of results concerning the existence of an approximation and determining the number of neurons required to guarantee that all functions (belong to a certain class) can be approximated to the prescribed degree of accuracy, have been achieved by many authors (see [1-21] and [23-27]).

Let $\sigma : \mathbf{R} \to [0,c]$ be the ramp function defined by

$$\sigma(x) := \begin{cases} 0, & x \leq -\mu_0, \\ c, & x \geq \mu_0, \\ \dfrac{x+\mu_0}{2\mu_0}c, & -\mu_0 < x < \mu_0, \end{cases} \qquad c \in \mathbf{R}^+, \quad 0 < \mu_0 \leq \frac{1}{2}.$$

Define

$$\begin{aligned} \varphi_1(x) &:= \sigma(x+\mu_0) - \sigma(x-\mu_0) \\ &= \begin{cases} 0, & |x| \geq 2\mu_0, \\ \left(1 - \dfrac{1}{2\mu_0}|x|\right) c, & |x| < 2\mu_0, \end{cases} \\ \varphi_2(x) &:= \sigma(x+2\mu_0) - \sigma(x-2\mu_0) \\ &= \begin{cases} 0, & |x| \geq 3\mu_0, \\ \left(\dfrac{3}{2} - \dfrac{1}{2\mu_0}|x|\right) c, & \mu_0 < |x| < 3\mu_0, \\ c, & |x| \leq \mu_0. \end{cases} \end{aligned}$$

Obviously, $\varphi_1(x)$ and $\varphi_2(x)$ are triangle function and trapezoidal function, respectively. Furthermore, $\varphi_1(x)$ and $\varphi_2(x)$ are nonnegative even functions, and are nonincreasing for $x > 0$. By using $\varphi_j(x)$ as the activation functions, Li [16] introduced the following single-hidden layer feed-forward neural network operators:

$$N_{n,j}(f,x) := \frac{\sum\limits_{k=0}^{n} f(x_k)\varphi_j\left(\frac{1}{h}(x-x_k)\right)}{\sum\limits_{k=0}^{n} \varphi_j\left(\frac{1}{h}(x-x_k)\right)}, \tag{1.1}$$

where $x_k = a + kh, k = 0,1,\cdots,n$, are the uniform space nodes on the interval $[a,d]$, with $h = \frac{d-a}{n}$.

In [16], Li obtained the following approximation rate of $N_{n,j}(f,x)$ for functions $f(x) \in C[a,d]$:

$$\|N_{n,j}(f) - f\| \leq 4\omega_2(f,h), \quad n \in \mathbf{Z}^+, \tag{1.2}$$