Journal of Engineering and Applied Sciences
MEDWELL PUBLICATIONS SINCE 2000

Constructive Approach of Neural Network Approximation of Trigonometric Activation Function

Eman Samir Bhaya and Zaineb Hussain Abd Al-Sadaa
Department of Mathematics, College of Education for Pure Sciences, University of Babylon, Hillah, Iraq

Key words: Image, processing, approximation, networks

Corresponding Author:
Eman Samir Bhaya
Department of Mathematics, College of Education for Pure Sciences, University of Babylon, Hillah, Iraq

Page No.: 2343-2347
Volume: 15, Issue 10, 2020
ISSN: 1816-949x
Journal of Engineering and Applied Sciences

Abstract: There are many uses for approximation using neural networks including astronomy, image processing and robots, this is due to its ease of use in approximation.

INTRODUCTION

In recent years, many researchers have studied the issue of approximation using neural networks. There have been many research papers on the possibility of approximation using neural networks which we call density problem. For more you can read[1-7]. All researchers in this research focused on estimating the degree of approximation of neural networks. The more complex issue is complexity problem how to determine the number of neurons necessary for the appropriate approximation. Research has been conducted to study the relationship between the degree of approximation and the counting of neurons in the hidden layer of the neural network. From this study, article[8] which developed a work of[2] and gave a way to find a neural network with a single hidden layer using the step function in the neurons and presented a direct theorem about the error of the best approximation. By[9] studied approximation using neural networks in sigmoidal activation function where he presented a direct theorem for approximation using neural networks whose inputs were real numbers using step functions. There have recently presented important facts about the L_p, p<1 approximation for more read[10-15], they studied the approximation using neural networks of functions in smooth classes and of error rate c/n where the number of neurons in the hidden layer.

In this study, we presented direct estimates of the upper bound for the degree of approximation using feed forward neural networks with one hidden layer and linear output. That is we have studied the complex problem of neural networks approximation which made us present a upper bound approximation method and identified the number of neurons in the hidden layer and that in terms of the first order modulus of smoothness for functions in the L_p spaces, it mean, we present a kind of Jackson's approximation-theorem. The rth symmetric difference of f along direction h is given by Johnen and Scherer[16].

$$\Delta_h^{(r)} f(x) = \sum_{i=0}^{r} \binom{r}{i} (-1)^i f\left(x+\left(\frac{r}{2}-i\right)h\right)$$

In terms of $\Delta_h^r f(.)$ the r-th modulus of smoothness of f is defined by Johnen and Scherer[16]:

$$\omega_r(f, t)_p = \sup_{0<\|h\|\leq t} \|\Delta_h^{(r)} f(.)\|_p$$

The main results: In this section, we introduce our main results begin with:

Theorem 2.1: Let $\varnothing$ is bounded, monotone and odd trigonometric real function. If $f \in L_p[a, b]$ then for any natural number $n \in \mathbb{N}$ there exists one hidden layer neural network satisfies:

$$\|N_n(x) - f(x)\|_p \leq c(p)\omega(f, \delta)_p$$

Proof: Define N_n: $[a, b] \to \mathbb{R}$, as:

$$N_n(x) = c_o + \sum_{i=1}^{n} c_i \varnothing(w_i x + \theta_i)$$

where parameters c_i's and w_i's are define as following:

$$c_o = f(a) - \sum_{i=1}^{n} c_i \varnothing(w_i a + \theta_i). \text{ For } 1 \leq i \leq n, \text{ we get}$$

$$c_i = \frac{1}{2m}(f(x_i) - f(x_{i-1})),\ w_i = \frac{2nd_n}{b-a},\ \theta_i = -\frac{nd_n}{b-a}(x_i + x_{i-1})$$

Define a partion for [a,b] consists of notes of length b-a/n as:

$$a = x_o < x_1 < x_2 <, ..., < x_n = b, \text{ let } m = \sup_{x \in R} \varnothing(x) \text{ and}$$
$$d_n = \varnothing^{-1}\left(m - \frac{m}{2n}\right)$$

According to the choice of c_o we have $N_n(a) = f(a)$. Also, we have $-m \leq \varnothing(x) \leq m$, for any real number x, we fix m. For any $x \in [a, b]$ there exists $j \in \mathbb{N}$ and $0 < j \leq n$, such that $x \in [x_{j-1}, x_j]$ that note that:

$$N_n(x) = c_o + \sum_{i=1}^{n} c_i \varnothing(w_i x + \theta_i)$$

$$N_n(x) = f(a) - \sum_{i=1}^{n} c_i \varnothing(w_i a + \theta_i) + \sum_{i=1}^{n} c_i \varnothing(w_i x + \theta_i)$$

$$= f(a) + \sum_{i=1}^{n} c_i (\varnothing(w_i x + \theta_i) - \varnothing(w_i a + \theta_i))$$

$$= f(a) + \sum_{i=1}^{n} \frac{1}{2m}(f(x_i) - f(x_{i-1})(\varnothing(w_i x + \theta_i) - \varnothing(w_i a + \theta_i))$$

Suppose $E_i(x) = \varnothing(w_i x + \theta_i)$ then:

$$N_n(x) = f(a) + \sum_{i=1}^{n} \frac{1}{2m}(f(x_i) - f(x_{i-1}))(E_i(x) - E_i(a))$$

$$= f(a) + \sum_{i=1}^{j-1} \frac{1}{2m}(f(x_i) - f(x_{i-1}))(E_i(x) - E_i(a)) + \frac{1}{2m}$$
$$(f(x_j) - f(x_{j-1}))(E_j(x) - E_j(a)) + \sum_{i=j+1}^{n} \frac{1}{2m}$$
$$(f(x_i) - f(x_{i-1})(E_i(x) - E_i(a))$$

For $i > j$, we have $x \leq x_j \leq x_{i-1}$. So, the properties of $\varnothing$ give $0 < E_i(x) - E_i(a) \leq E_i(x_j) - E_i(a) \leq E_i(x_{i-1}) - E_i(a) \leq E_i(x_{i-1}) - m = \varnothing(w_i x_{i-1} + \theta_i) + m = \varnothing(-d_n) + m = -\varnothing(\varnothing^{-1}(m - m/2n)) = m/2n$. So,

$$\sum_{i=j+1}^{n} \left\| \frac{1}{2m}(f(x_i) - f(x_{i-1})(E_i(x) - E_i(a)) \right\|_p \leq$$
$$\sum_{i=j+1}^{n} \left\| \frac{1}{2m}(f(x_i) - f(x_{i-1})) \right\|_p (E_i(x) - E_i(a)) \leq$$
$$\sum_{i=j+1}^{n} \left\| \frac{1}{2m}(f(x_i) - f(x_{i-1})) \right\|_p \left(\frac{m}{2n}\right)$$

$$\leq \frac{1}{2m}\left(\frac{m}{2n}\right) \sum_{i=j+1}^{n} \|(f(x_i) - f(x_{i-1}))\|_p$$

$$\leq \frac{1}{4n} \sum_{i=j+1}^{n} \|(f(x_i) - f(x_{i-1}))\|_p$$

$$\leq c(p)\omega\left(f, \frac{b-a}{n}\right)_p$$

$$\left\| \frac{1}{2m}(f(x_j) - f(x_{j-1}))(E_j(x) - E_j(a)) \right\|_p \leq \frac{1}{2m} \|f(x_j) - f$$
$$(x_{j-1})\|_p \left(\frac{m}{2n}\right) \leq c(p)\omega\left(f, \frac{b-a}{n}\right)_p$$

And:

$$\sum_{i=1}^{j-1} \frac{1}{2m}(f(x_i) - f(x_{i-1})(E_i(x) - E_i(a)) =$$
$$\frac{1}{2m} \sum_{i=1}^{j-1} (f(x_i) -$$
$$f(x_{i-1}) E_i(x) - E_i(a) - 2m) + (f(x_{j-1}) - f(a))$$

For $1 \le i < j$, we have $x_i \le x_{j-1} \le x$ and so, $2m \ge E_i(x)-E_i(a) > E_i(x_i)-E_i(x_{i-1}) = \varnothing(d_n)-\varnothing(-d_n) = 2m-m/n$ which implies $\|E_i(x)-E_i(a)-2m\|_p \le m/n$. So:

$$N_n(x)-f(x_{j-1}) = \frac{1}{2m}\sum_{i=1}^{j-1}\big(f(x_i)-f(x_{i-1})\big)$$
$$\big(E_i(x)-E_i(a)-2m\big)+\frac{1}{2m}\Big(f(x_j)-f(x_{j-1})\Big)$$
$$\Big(E_j(x)-E_j(a)+\sum_{i=j+1}^{n}\frac{1}{2m}f(x_i)-\big(f(x_{i-1})\big(E_i(x)-E_i(a)\big)\Big)$$

Consequently,

$$\|N_n(x)-f(x_{j-1})\|_p \le c(p)\sum_{i=1}^{j-1}\frac{1}{2m}\|(f(x_i)-f(x_{i-1})(E_i(x)-$$
$$E_i(a)-2m)+\frac{1}{2m}\big(f(x_j)-(x_{j-1})\big)\big(E_j(x)-E_j(a)\big)+$$
$$\sum_{i=j+1}^{n}\frac{1}{2m}\big(f(x_i)-f(x_{i-1})\big(E_i(x)-E_i(a)\big)\|_p \le$$
$$\frac{c(p)}{2m}\omega\left(f,\frac{b-a}{n}\right)\sum_{i=1}^{j-1}\|\big(E_i(x)-E_i(a)-2m\big)\|_p+\frac{1}{2m}$$
$$\|f(x_j)-f(x_{j-1}))\big(E_j(x)-E_j(a)\|_p+\left\|\sum_{i=j+1}^{n}\frac{1}{2m}\big(f(x_i)-\right.$$
$$f(x_{i-1})\big(E_i(x)-E_i(a)\big)\|_p \le c(p)\,\omega\left(f,\frac{b-a}{n}\right)_p$$

Using the direct theorem:

$$\|f(x)-f(x_{j-1})\|_p \le \omega\left(f,\frac{b-a}{n}\right)_p$$

We have:

$$\|N_n(x)-f(x)\|_p \le \|N_n(x)-f(x_{j-1})\|_p+\|f(x)-f(x_{j-1})\|_p$$
$$\le c(p)\,\omega\left(f,\frac{b-a}{n}\right)_p$$

Theorem 2.2: Let $\varnothing$ is bounded, monotone and odd trigonometric real function. $f \in Lip(\alpha)$, $\alpha \in (0, 1)$ if and only if there is one hidden layer neural network N_n, satisfying:

$$\|N_n(x)-f(x)\|_p \le c(p)(\delta)^{\alpha}$$

where, $\delta = \frac{b-a}{n}$.

Proof: Define $N_n:[a, b] \to \mathbb{R}$, as:

$$N_n(x) = c_0+\sum_{i=1}^{n}c_i\varnothing(w_i x+\theta_i)$$

where parameters c_i's and w_i's are define as following:

$$c_o = f(a)-\sum_{i=1}^{n}c_i\varnothing(w_i x+\theta_i)$$

For $1 \le i \le n$, we get:

$$c_i = \frac{1}{2m}\big(f(x_i)-f(x_{i-1})\big),\; w_i = \frac{2nd_n}{b-a},\; \theta_i = -\frac{nd_n}{b-a}(x_i+x_{i-1})$$

Let $m = \sup_{x\in\mathbb{R}}\varnothing(x)$ and $d_n = \varnothing^{-1}(m-m/2n)$.

Since, $f \in Lip(\alpha)_k$ then $\omega(f, \delta)_p = O(\delta)^{\alpha}$ $\|N_n(x)-f(x)\|_p = o(\delta^{\alpha})$

Let $f \in L^p_{2\pi}$, $0<p<1$, then $\|N_n(x)-f(x)\|_p \le c(p)(\delta)^{\alpha}$.

We must prove that $f_i \in Lip(\alpha)_k$. Now, $\|EN_\lambda[f_i]-f_i\|_p \le c(p)(\delta^{\alpha})$ and by using Theorem 2.1:

$$\|N_n(x)-f(x)\|_p \le c(p)\,\omega(f, \delta)$$

Then:

$$c(p)\,\omega(f, \delta)_p \le c(p)(\delta^{\alpha})$$

$$\omega(f, \delta)_p \le 0(\delta^{\alpha})$$

Therefore, the definition of Lipschitian function is conclude we get.

Examples 3: In this section let us demonstrate our theorems.

Example 3.1; Cao *et al*.[15]: Let $f(x) = \sin x$, $x \in [0, \pi]$. Choose $\varnothing(x) = 2/\pi \tan^{-1}x$, $x \in \mathbb{R}$. It is clear $f \in Lip_1(1)$ and also we have $f(a) = f(0) = f(b) = 0$. Using the properties of the $\tan^{-1}x$, if $m = 1$, $\varnothing^{-1}(x) = \tan(\pi/2)\, x$ and $d_n = \tan(\pi/2(1-1/2n))$. So,

$$c_i = \frac{1}{2m}\big(f(x_i)-f(x_{i-1})\big)$$

$$c_i = \frac{1}{2}\left(\sin\frac{i\pi}{n}-\sin\frac{(i-1)\pi}{n}\right)$$

$$\theta_i = -\frac{nd_n}{b-a}\left(2a+(2i-1)\frac{b-a}{n}\right)$$

$$= \frac{-n\tan\left(\frac{\pi}{2}\left(1-\frac{1}{2n}\right)\right)}{\pi}(0+2i-1)\frac{\pi}{n}$$
$$= -\tan\left(\frac{\pi}{2}\left(1-\frac{1}{2n}\right)\right)(2i-1)$$

$$w_i = \frac{2nd_n}{b\text{-}a} = \frac{2n}{\pi}\tan\left(\frac{\pi}{2}\left(1\text{-}\frac{2}{2n}\right)\right)$$

$$c_o = f(a)\text{-}\sum_{i=1}^{n} c_i \varnothing(w_i a + \theta_i)$$

$$= 0\text{-}\sum_{i=1}^{n}\frac{1}{2}\left(\sin\frac{i\pi}{n}\text{-}\sin\frac{(i\text{-}1)\pi}{n}\right)\frac{2}{\pi}\arctan\left((1\text{-}2i)\tan\left(\frac{\pi}{2}\left(1\text{-}\frac{1}{2n}\right)\right)\right).$$

So, we can define the following neural network having one hidden layer and n neural:

$$N_n(x) = c_o + \sum_{i=1}^{n} c_i \varnothing(w_i x + \theta_i),\ x \in [0, \pi]$$

From Theorem 2.2, we get:

$$\|N_n(x)\text{-}\sin x\|_p \le c(p)\left(\frac{\pi}{n}\right)^{\alpha}$$

Example 3.2: If our target function is f(x) = cos x, x∈[0, π] we choose the sigmoidal activation function $\varnothing(x) = 2/\pi \tan^{-1}x$, $x \in \mathbb{R}$. And f(a) = cos (0)= f(b) = cos(π) = 1, m = $\sup \varnothing^{-1}(x) = \tan(\pi/2)x$:

$$d_n = \tan\left(\frac{\pi}{2}\right)x\left(1\text{-}\frac{1}{2n}\right)$$

$$c_i = \frac{1}{2m}\left(f(x_i)\text{-}f(x_{i-1})\right)$$

$$= \frac{1}{2m}\left(\cos\frac{i\pi}{n}\text{-}\cos\frac{(i\text{-}1)\pi}{n}\right)$$

$$\theta_i = \text{-}\frac{nd_n}{b\text{-}a}\left(2a + (2i\text{-}1)\left(\frac{b\text{-}a}{n}\right)\right)$$

$$= \frac{-n\tan\left(\frac{\pi}{2}\right)\left(1\text{-}\frac{1}{2n}\right)}{\pi}\left(1 + (2i\text{-}1)\frac{\pi}{n}\right.$$

$$= \text{-}\tan\left(\frac{\pi}{2}\left(1\text{-}\frac{1}{2n}\right)\right)(2i)$$

$$w_i = \frac{2nd_n}{b\text{-}a} = \frac{2n}{\pi}\tan\left(\frac{\pi}{2}\left(1\text{-}\frac{2}{2n}\right)\right)$$

$$c_0 = f(a)\text{-}\sum_{i=1}^{n} c_i \varnothing(w_i a + \theta_i)$$

$$= 1\text{-}\sum_{i=1}^{n}\frac{1}{2}\left(\cos\frac{i\pi}{n}\text{-}\cos\frac{(i\text{-}1)\pi}{n}\right) * \frac{2}{\pi}\arctan\left((1\text{-}2i)\tan\left(\frac{\pi}{2}\left(1\text{-}\frac{1}{2n}\right)\right)\right)$$

So, we can define the neural network approximate as:

$$N_n(x) = c_o + \sum_{i=1}^{n} c_i \varnothing(w_i x + \theta_i),\ x \in [0, \pi].$$

From Theorem 2.2. we get:

$$\|N_n(x) - \cos x\|_p \le c(p)(\frac{\pi}{n})a$$

CONCLUSION

The main aim of this study is to introduce a saturation problem for the approximation of function in $L_p, p<1$ quasi normed spaces using neural network with trigonometric activation function, in a constructive approach.

REFERENCES

01. Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Mathe Control Signals Syst., 2: 303-314.
02. Funahashi, K.I., 1989. On the approximate realization of continuous mappings by neural networks. Neural Networks, 2: 183-192.
03. Hornik, K., M. Stinchcombe and H. White, 1990. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Network, 3: 551-560.
04. Li, X., 1998. On simultaneous approximations by radial basis function neural networks. Appl. Math. Comput., 95: 75-89.
05. Leshno, M., V. Lin, A. Pinks and S. Schochen, 1993. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. Neural Netw., 6: 861-867.
06. Chen, T. and H. Chen, 1995. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its
07. Chui, C.K. and X. Li, 1992. Approximation by ridge functions and neural networks with one hidden layer. J. Approximation Theor., 70: 131-141. application to dynamical systems. IEEE. Trans. Neural Networks, 6: 911-917.
08. Korain, P., 1993. On the complexity of approximating mappings using feedforward networks. Neural Networks, 6: 449-653.
09. Bulsari, A., 1993. Some analytical solutions to the general approximation problem for feedforward neural networks. Neural Network, 6: 991-996.

10. Jones, L.K., 1992. A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and neural network training. Ann. Stat., 20: 608-613.
11. Barron, A.R., 1993. Universal approximation bounds for super positions of a sigmoidal function. IEEE Trans. Inform. Theory, 39: 930-945.
12. Murata, N., 1996. An integral representation of functions using three-layered networks and their approximation bounds. Neural Networks, 9: 947-956.
13. Xu, Z. and F. Cao, 2004. The essential order of approximation for neural networks. Sci. China Ser. F. Inf. Sci., 47: 97-112.
14. Xu, Z.B. and F.L. Cao, 2005. Simultaneous Lp-approximation order for neural networks. Neural Networks, 18: 914-923.
15. Cao, F., T. Xie and Z. Xu, 2008. The estimate for approximation error of neural networks: A constructive approach. Neurocomputing, 71: 626-630.
16. Johnen, H. and K. Scherer, 1977. On the Equivalence of the K-Functional and Moduli of Continuity and Some Applications. In: Constructive Theory of Functions of Several Variables, Schempp, W. and K. Zeller (Eds.). Springer, Berlin, Germany, ISBN: 978-3-540-08069-5, pp: 119-140.