

# Universal Approximation Property of Weighted Approximate Identity Neural Networks

---

**Saeed PanahianFard<sup>1</sup>**

*School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Pulau Pinang, Malaysia*

*E-mail: saeedpanahian@usm.my*

**Zarita Zainuddin**

*School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Pulau Pinang, Malaysia*

*E-mail: zarita@usm.my*

Universal approximation property of feedforward artificial neural networks is one of the important research topics in the approximation theory of artificial neural networks. In this paper, we examine the universal approximation property of one-hidden-layer feedforward weighted approximate identity neural networks in the weighed spaces of continuous functions. To do this end, we firstly introduce the notion of weighted approximate identity. Secondly, we present two theorems by using this notion in order to show universal approximation property of weighted approximate identity neural networks in the weighed spaces of continuous functions. The obtained results develop the existence theoretical results.

*CENet2015*

*12-13 September 2015*

*Shanghai, China*

---

<sup>1</sup>Speaker

Corresponding Author

### 1. Introduction

In recent decades, the investigation focusing on the problem of the universal approximation property of one-hidden-layer feedforward artificial neural networks has been escalated [1]. It has been verified that feedforward artificial neural networks are fit for approximating continuous functions [2]. In other words, feedforward artificial neural networks has universal approximation property. The over view of the universal approximation property of this general feedforward neural networks can be found in some publications such as [3,4]. In the same context, approximation property of one-hidden-layer feedforward approximate identity neural networks have been introduced by Turchetti et al. in 1998. See [1-4] for reading about the universal approximation property of different subclasses of feed forward approximate identity neural networks.

In the exhibit study, we aim at considering the universal approximation property of feedforward weighted approximate identity artificial neural networks as a subclass of approximate identity neural networks. On this stage, we ought to clarify what we mean by weighted approximate identity neural networks. These networks can be regarded as extension of feedforward approximate identity neural networks by utilizing weighted approximate identity as activation functions. To our knowledge, such study has not been explored.

The objective of this study is to investigate the universal approximation property of one hidden layer feedforward weighted approximate identity artificial neural networks in the weighted spaces of continuous functions.

The methods of this study can be described as follows: firstly, we present the basic notion of weighted approximate identity in order to construct one hidden layer feedforward weighted approximate identity neural networks. Second, we prove two theorems that show that one hidden layer feedforward weighted approximate identity neural networks have universal approximation property in the weighted spaces of continuous functions.

The framework of the present paper is formed in the following. In Section 2, some basic definitions are introduced; in Section 3, two theorems are presented. Finally, our conclusion is given in Section 4.

### 2. Basic Definitions

Some basic definitions are presented in this section which will be used in the next section.

**Definition1:** let  $w(x):R \rightarrow R^+$  be a weight function.  $\{\varphi_n(x)\}_{n \in \mathbb{N}}, \varphi_n(x):R \rightarrow R$   $\mathbb{R}$  is a weighted approximate identity with the following properties:

$$1) \int_{\mathbb{R}} \varphi_n(x)w(x)dx = 1;$$

2) for a positive number  $\varepsilon$  and a positive number  $\delta > 0$ , there is  $N$  so that if  $n \geq N$  then

$$\int_{|x|>\delta} |\varphi_n(x)|w(x)dx \leq \varepsilon.$$

We now recall the notion of  $\varepsilon$ -net and finite epsilon-net in the weighted spaces of continuous functions  $C(R, w)$

**Definition 2:** Let  $\varepsilon > 0$  A set  $V_\varepsilon \subset C(R, w)$  is an  $\varepsilon$ -net of a set  $V$ , if  $f \in V_\varepsilon$  can be found for  $\forall f \in V$  such that  $\|f - f\|_{C(\mathbb{R}, w)} < \varepsilon$ . [5]

**Definition:4** The  $\varepsilon$ -net is called finite if the  $\varepsilon$ -net be a finite set of elements.[5]

POS (CENet 2015) 007

### 3. Theoretical Results for Weighted Spaces of Continuous Functions

In this section, we present a theoretical result which states that the convolution of weighted approximate identity with a function  $f$  in the weighted spaces of continuous converges uniformly to on  $C(\mathbb{R}, w)$ .

**Theorem 1.** Let  $C(R, w)$  be the linear weighted spaces of all continuous functions on  $\mathbb{R}$  with a compact support. Let  $\{\varphi_n(x)\}_{n \in \mathbb{N}}$ ,  $\varphi_n(x): R \rightarrow R$  be a weighted approximate identity. Let  $f \in C(R, w)$ , then  $\varphi_n * f$  converges uniformly to  $f$  on  $C(R, w)$ .

**Proof.** Let  $x \in \mathbb{R}$  and  $\varepsilon > 0$ . For a positive number  $\delta > 0$ , we have  $|f(x) - f(y)| < \frac{\varepsilon}{2\|\varphi\|_1}$ ,  $\forall y, |x - y| < \delta$ .  $\{\varphi_n * f\}_{n \in \mathbb{N}}$  can be defined by  $\varphi_n(x) = n\varphi(nx)$ . After that,

$$\begin{aligned} \varphi_n * f(x) - f(x) &= \int_{\mathbb{R}} n\varphi(ny)\{f(x - y) - f(y)w(x)dy \\ &= (\int_{|y| < \delta} + \int_{|y| \geq \delta}) n\varphi(ny)\{f(x - y) - f(y)\}w(x)dy \\ &= \int_{|y| < \delta} n\varphi(ny)\{f(x - y) - f(y)\}w(x)dy + \\ &\quad \int_{|y| \geq \delta} n\varphi(ny)\{f(x - y) - f(y)\}w(x)dy \end{aligned} \quad (3.1)$$

we calculate

$$I_1 = \int_{|y| < \delta} n\varphi(ny)\{f(x - y) - f(y)\}w(x)dy, \quad (3.2)$$

$$I_2 = \int_{|y| \geq \delta} n\varphi(ny)\{f(x - y) - f(y)\}w(x)dy \quad (3.3)$$

as follows:

$$\begin{aligned} |I_1| &\leq \int_{|y| < \delta} n|\varphi(ny)|\{f(x - y) - f(y)\}w(x)dy \\ &< \frac{\varepsilon}{2\|\varphi\|_1} \int_{|y| < \delta} n|\varphi(ny)|w(x)dy \\ &= \frac{\varepsilon}{2\|\varphi\|_1} \int_{t < n\delta} |\varphi(t)|w(x)dt \\ &\leq \frac{\varepsilon}{2\|\varphi\|_1} \int_{\mathbb{R}} |\varphi(t)|w(x)dt = \frac{\varepsilon}{2}. \end{aligned} \quad (3.4)$$

For  $I_2$ , we have

$$\begin{aligned} |I_2| &\leq 2\|f\|_{C(\mathbb{R}, w)} \int_{|y| \geq \delta} n|\varphi(ny)|w(x)dy \\ &= 2\|f\|_{C(\mathbb{R}, w)} \int_{t \geq n\delta} |\varphi(t)|w(x)dt. \end{aligned} \quad (3.5)$$

Since

$$\lim_{n \rightarrow \infty} \int_{t \geq n\delta} |\varphi(t)|w(x)dt = 0, \quad (3.6)$$

There is  $n_0 \in \mathbb{N}$ , so that for all  $n \geq n_0$ ,

$$\int_{t \geq n\delta} |\varphi(t)|w(x)dt < \frac{\varepsilon}{4\|f\|_{C(\mathbb{R}, w)}}. \quad (3.7)$$

Combining  $I_1$  and  $I_2$  for  $n \geq n_0$ , we have

$$\|\varphi_n * f(x) - f(x)\|_{C(\mathbb{R}, w)} < \varepsilon. \quad (3.8)$$

Using Theorem 1, we now give Theorem 2 as our main result of this section. Theorem 2 states that weighted approximate identity neural networks has universal approximation property. The proof is essentially the same as the proof of Theorem 1 [6].

**Theorem 2.** Let  $C(\mathbb{R}, w)$  be the linear weighted spaces of all continuous functions on  $\mathbb{R}$  with a compact support and  $V \subset C(\mathbb{R}, w)$ . Let  $\{\varphi_n(x)\}_{n \in \mathbb{N}}, \varphi_n(x): \mathbb{R} \rightarrow \mathbb{R}$  be a weighted approximate identity. Let  $\{\sum_{j=1}^M \lambda_j \varphi_j(x) w(x) | \lambda_j \in \mathbb{R}, x \in \mathbb{R}, M \in \mathbb{N}\}$  be dense in  $C(\mathbb{R}, w)$ ; given  $\varepsilon > 0$ . There is  $N \in \mathbb{N}$  which depends on  $V$  and but  $\varepsilon$  not on  $f$ , so that to any  $f \in V$  there will be weights  $C_k = C_k(f, V, \varepsilon)$  satisfying

$$\left( f(x) - \sum_{i=1}^N c_k \varphi_i(x) w(x) \right)_{C(\mathbb{R}, w)} < \varepsilon \quad (3.9)$$

moreover, every  $c_k$  is a continuous function of  $f \in V$ .

**Proof.** Because  $V$  is a compact set, for any positive number  $\varepsilon > 0$ , there exists a finite  $\frac{\varepsilon}{2}$ -net  $(f^1, \dots, f^M)$  for  $V$ . This indicates that for any  $f \in V$ , there is an  $f^j$  so that  $(f - f^j)_{C(\mathbb{R}, w)} < \frac{\varepsilon}{2}$ . For any  $f^j$ , by assumption of the theorem, there are  $\lambda_i^j \in \mathbb{R}, N_j \in \mathbb{N}$  and  $\varphi_i^j(x)$  such that

$$\left\| f^j(x) - \sum_{i=1}^{N_j} \lambda_i^j \varphi_i^j(x) w(x) \right\|_{C(\mathbb{R}, w)} < \frac{\varepsilon}{2}. \quad (3.10)$$

For any  $f \in V$ , we can define

$$F_-(f) = \{j \mid \|f - f^j\|_{C(\mathbb{R}, w)} < \frac{\varepsilon}{2}\}, \quad (3.11)$$

$$F_0(f) = \{j \mid \|f - f^j\|_{C(\mathbb{R}, w)} = \frac{\varepsilon}{2}\}, \quad (3.12)$$

$$F_+(f) = \{j \mid \|f - f^j\|_{C(\mathbb{R}, w)} > \frac{\varepsilon}{2}\}. \quad (3.13)$$

According to the definition of  $\frac{\varepsilon}{2}$ -net,  $F_-(f)$  is not a null set. If  $f^j \in V$  comes near to  $f$  such that  $\|f^j - f\|_{C(\mathbb{R}, w)}$  is small enough, then  $F_-(f) \subset F_-(f^j)$  and  $F_+(f) \subset F_+(f^j)$ . Thus,  $F_-(f) \cap F_+(f) \subset F_-(f^j) \cap F_+(f^j) = \emptyset$ , which indicates that  $F_-(f) \subset F_-(f) \cup F_0(f)$ . Now, we can obtain the following.

$$F_-(f) \subset F_-(f) \cup F_0(f). \quad (3.14)$$

Define

$$d(f) = \left[ \sum_{j \in F_-(f)} \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R}, w)} \right) \right]^{-1} \quad (3.15)$$

and

$$f_h = \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f) \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R}, w)} \right) \lambda_i^j \varphi_i^j(x) w(x) \quad (3.16)$$

Then  $f_h \in \{\sum_{i=1}^M \lambda_i \varphi_i(\mathbf{x})\}$  approximates  $f$  with accuracy  $\varepsilon$ :

$$\begin{aligned}
 & \|f - f_h\|_{C(\mathbb{R},w)} \\
 &= \left\| \sum_{j \in F_-(f)} d(f) \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R},w)} \right) \left( f - \sum_{i=1}^{N_j} \lambda_i^j \varphi_i^j(x) w(x) \right) \right\|_{C(\mathbb{R},w)} \\
 &= \left\| \sum_{j \in F_-(f)} d(f) \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R},w)} \right) \left( f - f^j + f^j - \sum_{i=1}^{N_j} \lambda_i^j \varphi_i^j(x) w(x) \right) \right\|_{C(\mathbb{R},w)} \\
 &\leq \sum_{j \in F_-(f)} d(f) \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R},w)} \right) \left( \|f - f^j\|_{C(\mathbb{R},w)} + \left\| f^j - \sum_{i=1}^{N_j} \lambda_i^j \varphi_i^j(x) w(x) \right\|_{C(\mathbb{R},w)} \right) \\
 &< \sum_{j \in F_-(f)} d(f) \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R},w)} \right) \left( \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \right) = \varepsilon.
 \end{aligned} \tag{3.17}$$

For the following step, the continuity of  $c_k$  is proved. For the proof, (14) can be used to obtain

$$\begin{aligned}
 & \sum_{j \in F_-(f)} \left( \frac{\varepsilon}{2} - \|f^j - f^j\|_{C(\mathbb{R},w)} \right) \\
 &\leq \sum_{j \in F_-(f)} \left( \frac{\varepsilon}{2} - \|f^j - f^j\|_{C(\mathbb{R},w)} \right) \\
 &\leq \sum_{j \in F_-(f)} \left( \frac{\varepsilon}{2} - \|f^j - f^j\|_{C(\mathbb{R},w)} \right) + \\
 & \sum_{j \in F_0(f)} \left( \frac{\varepsilon}{2} - \|f^j - f^j\|_{C(\mathbb{R},w)} \right).
 \end{aligned} \tag{3.18}$$

Let  $f^j \rightarrow f$  in (18), we have the following

$$\sum_{j \in F_-(f)} \left( \frac{\varepsilon}{2} - \|f^j - f^j\|_{C(\mathbb{R},w)} \right) \rightarrow \sum_{j \in F_-(f)} \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R},w)} \right) \tag{3.19}$$

This shows that  $d(f^j) \rightarrow d(f)$ . Thus,  $f^j \rightarrow f$  results

$$d(f^j) \left( \frac{\varepsilon}{2} - \|f^j - f^j\|_{C(\mathbb{R},w)} \right) \lambda_i^j \rightarrow d(f) \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R},w)} \right) \lambda_i^j. \tag{3.20}$$

Let  $N = \sum_{j \in F_-(f)} N_j$  and then we define  $c_k$  in terms of

$$\begin{aligned}
 f_h &= \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f) \left( \frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R},w)} \right) \lambda_i^j \varphi_i^j(x) w(x) \\
 &\equiv \sum_{k=1}^N c_k \varphi_k(x) w(x)
 \end{aligned} \tag{3.21}$$

From (3.20), it follows that  $c_k$  is a continuous functional of  $f$ .

#### 4. Conclusion

We have investigated universal approximation property of one-hidden-layer feedforward weighted approximate identity neural networks in the weighted spaces of continuous functions. We have introduced the basic notion of weighted approximate identity. Then, we have proved Theorem 1 shows that the convolution of weighted approximate identity with a continuous function  $f$  in the weighted spaces of continuous functions converges uniformly to  $f$  on  $C(R, w)$ . We have used Theorem 1 to present Theorem 2. Theorem 2 shows that one-hidden-layer feedforward weighted approximate identity neural networks has universal approximation property in the weighted spaces of continuous functions. An interesting idea as pursued is to investigate the universal approximation property of those networks in the weighted Lebesgue spaces.

#### References

- [1] Zainuddin Z, Panahian Fard S. (2014) *Spherical approximate identity neural networks are universal approximators*. The 10th International Conference on Natural Computation (ICNC), IEEE. Xiamen, pp. 72-76
- [2] Panahian Fard S, Zainuddin Z (2014) *Toroidal approximate identity neural networks are universal approximators*. Lecture Notes in Computer Science, Springer International Publishing, 8834:135-142
- [3] Zainuddin Z, Panahian Fard, S (2012) *Double approximate identity neural networks universal approximation in real Lebesgue space*. Lecture Notes in Computer Science, Springer Berlin Heidelberg, 7663:409-415
- [4] Panahian Fard S, Zainuddin, Z (2015) *Almost everywhere approximation capabilities of double Mellin approximate identity neural networks* *Soft Computing*, 1-9. DOI: 10.1007/s00500-015-1753-y
- [5] Lebedev V (1997) *An introduction to functional analysis and computational mathematics*. Birkhauser, Boston
- [6] Wu W, Nan, D, Li Z, Long J (2007) *Approximation to compact set of functions by feedforward neural networks*. International Joint Conference on Neural Networks. IEEE. Orlando, FL, 20:1222-1225