International Journal of Mathematics and
Computer Science, **20**(2025), no. 1, 179–182
DOI: https://doi.org/10.69793/ijmcs/01.2025/sharba

Bernstein Neural Network Approximation

Zaineb Hussain Abd Al-sadaa[1], Zainab Abdulmunim Sharba[2], Hawraa Abbas Almurieb[1]

[1]Department of Mathematics
College of Education for Pure Sciences
University of Babylon
Hillah, Iraq

[2]Department of Computer Science
College of Science for Women
University of Babylon
Babil, Iraq

email: zainab.hussain@uobabylon.edu.iq, zainab.abd@uobabylon.edu.iq, pure.hawraa.abbas@uobabylon.edu.iq

(Received July 20, 2024, Accepted August 21,2024, Published August 23, 2024)

Abstract

Bernstein Polynomials were one of the first mathematical polynomials that were used for approximation. Moreover, Neural networks were used for almost applicable targets. In this paper, we introduce the Bernstein neural network as a mathematical operator that combines the concepts of neural networks and Bernstein polynomials. It is used to approximate measurable functions by representing them as a weighted sum of Bernstein polynomials. Moreover, a special activation function that is derived from Bernstein Polynomials is defined and used to approximate functions. Furthermore, we estimate the degree of approximation with equivalent bounds to modulus of smoothness. Finally, we estimate the degree of approximation using comparable limitations to the modulus of smoothness.

Key words and phrases: Approximation, Modulus of Smoothness, Bernstein.
AMS (MOS) Subject Classifications: 03F10, 03B45.
ISSN 1814-0432, 2025, https://future-in-tech.net

1 Introduction

Neural networks have an important benefit in approximation theory, as functions can be approximated using neural networks [1, 2]. Many authors have presented different models of neural networks with different activation functions. Coscarelli and Spigler [3] have also constructed artificial (NNs) and examined the behavior of his neural network in terms of Bernstein polynomials. New neural networks using multivariate square Bernstein polynomials with a positive integer coefficient were introduced by the authors in [4]. In [5], Bernstein neural networks (BNNs) were defined as a type of neural network architecture that utilizes Bernstein polynomials. The parameters of NNs are activation functions. These networks are specifically designed for function approximation tasks. Continuous functions were approximated by those BNNs based on modulus of continuity. For every given weight $\mathbf{w}$, and input $\mathbf{x}$ such that $\mathbf{x}, \mathbf{w} \in [-1,1]^d$, the neural network is defined as

$$BN_n\left(\mathbf{x}\right) = \sum_{j=0}^{d} \sum_{k=0}^{n} c_j \sigma_j \left(x_j w_j + b_j\right), \tag{1.1}$$

where $c_{j,i} = \sum_{i=1}^{k} (-1)^{k-i} \binom{k}{i} f_j\left(\frac{k}{i}\right)$ and $\sigma_j(x) = x_j^k \left(1_j - x_j^k\right)^{n-k}$. Later, in a similar way to $N_n\left(\mathbf{x}\right)$,the authors in [6] applied Bernstein polynomials to train NNs and get better accuracy compared to ReLU networks. Let $L_p^d(I), I = [-1,1], 0 < p < 1$, be the space of measurable functions equipped with the norm $\|f\|_p = (\int_{-1}^{1} |f(x)|^p dx)^{\frac{1}{p}}$. The $k-$th symmetric difference of f along direction h is given by

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

In terms of $\Delta_h^k(f(x))$, define the $k-$th modulus of smoothness of f by

$$\Lambda_k(f,t)_p = \sup_{0<\|h\|\leq t} \|\Delta_h^k(f(x))\|_p. \tag{1.2}$$

2 Auxiliary Results

In this section, we present some useful lemmas for our work. From [6], we obtain the similar lemma below:

Lemma 2.1. *Let $f \in L_p^d(I)$, and $n \in \mathbb{N}$. Then*

$$\|BN_n(f)\|_p \leq \|BN_n(f) - f\|_p. \tag{2.3}$$

The following lemma gives the property of boundness for the BNN.

Lemma 2.2. *For any $f \in L_p^d(I)$ and $n \in \mathbb{N}$, there exists a constant M that satisfies*

$$\|BN_n(f)\|_p \leq M\|f\|_p. \tag{2.4}$$

3 Main Results

In this section, we prove the existence of best approximation, and then study the bounds of the degree of approximation that is equivalent to modulus of smoothness

Theorem 3.1. *Let $f \in L_p^d(I)$, then, for $n \in \mathbb{N}$ and $d \leq n$, there exist a BNN of the form (1.1) that satisfies $E_n(f)_p \leq c(p)\Lambda_k(f,\delta)_p$.*

Proof. By (1.1), (1.2) and the quasi triangle inequality

$$\begin{aligned}
\|BN_n f - f\|_p &= \left\| \sum_{j=0}^{d} \sum_{i=1}^{k} \sum_{k=0}^{n} c_{j,i}\sigma_j\left(w_i x_i + b_j\right) - f \right\|_p \\
&\leq \left\| \sum_{j=0}^{d} \sum_{i=0}^{k} \sum_{k=0}^{n} (-1)^{k-i} \binom{k}{i} f_j\left(\frac{k}{i}\right) x_j^k \left(1_j \ - \ x_j^k\right)^{n-k} - f \right\|_p \\
&\leq c(p) \sum_{j=0}^{d} \sum_{k=0}^{n} \left\| \sum_{i=1}^{k} (-1)^{k-i} \binom{k}{i} f_j\left(\frac{k}{i}\right) x_j^k \left(1_j \ - \ x_j^k\right)^{n-k} - f \right\|_p \\
&\leq c(p) \sum_{j=0}^{d} \sum_{k=0}^{n} \left\| \sum_{i=1}^{k} (-1)^{k-i} \binom{k}{i} x_j^k \left(1_j \ - \ x_j^k\right)^{n-k} \left(f_j\left(\frac{k}{i}\right) - f\left(y_i\right)\right) \right\|_p \\
&\leq c\Lambda_k(f,\delta)_p. \quad \square
\end{aligned}$$

Now, we study the lower bound of the degree of best approximation.

Theorem 3.2. *Let $f \in L_p^d(I)$ and $n \in \mathbb{N}$. Then $\Lambda_k(f,\delta)_p \leq c(p)\|N_n(f) - f\|_p$.*

Proof. By using (2.3), (2.4) and Theorem 3.1, we have

$$\begin{aligned}
\Lambda_k(f,\delta)_p &\leq \Lambda_k(f - BN_n(f),\delta)_p + \Lambda_k(BN_n(f),\delta)_p \\
&\leq c\,(p)\left(\|f - BN_n(f)\|_p + \|BN_n(f)\|_p\right), \\
&\leq c\,(p)\,\|BN_n(f) - f\|_p. \quad \square
\end{aligned}$$

4 Conclusions

Bernstein polynomials still play an active role in both theoretical and practical approaches such as neural networks. In this paper, we dealt with an operator depending on neural networks and Bernstein polynomials. We defined a mathematical formula of BNNs. We studied some mathematical properties of our operator, especially boundness. Moreover, we used the modulus of smoothness to estimate the degree of approximation with the following bounds: Direct theorem (Theorem 3.1) was proved to get the upper bound of the degree of approximation with modulus of smoothness. It approaches zero as fast as the smoothness of the function itself. Inverse theorem (Theorem 3.2) was proved to get the lower bound of the degree of approximation with the same modulus above.

References

[1] E.S. Bhaya, Z.A. Sharba, L_p approximation by ReLU neural networks, Karbala International Journal of Modern Science, **6**, no. 4, (2020), 414–419.

[2] H.A. Almurieb, E.S. Bhaya, Monotone approximation by quadratic neural network of functions in L_p spaces for $p < 1$, Iraqi Journal of science, **61**, no.4, (2020), 870–874.

[3] S.R. Costarelli, Approximation results for neural network operators activated by sigmoidal functions, Neural Networks, **44**, (2013), 101–106.

[4] J.M. Ibtihal, A.J. Mohammad, Neural Network of Multivariate Square Rational Bernstein Operators with Positive Integer Parameter, European Journal of Pure and Applied Mathematics, **15**, no. 3, (2022), 1189–1200.

[5] H.A. Almurieb, A.A. Hamody, Best neural network approximation by using Bernstein polynomials with GRNN learning application, Al-Bahir Journal for Engineering and Pure Sciences, **1**, no. 1, (2022), 5–9.

[6] H.A.S.Y. Khedr, Deepbern-nets: Taming the complexity of certifying neural networks using Bernstein polynomial activations and precise bound propagation, Proceedings of the AAAI Conference on Artificial Intelligence, **38**, no. 19, (2024), 21232–21240.