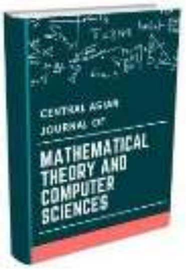

CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES

https://cajmtcs.casjournal.org/index.php/CAJMTCS

Volume: 07 Issue: 02 | April 2026 ISSN: 2660-5309

.

Article

Generalized Stechkin-Marchaud Inequalities via Deep Neural Network Approximation: Theorems on Convergence, Stability, and Error Bounds

Nagham Ali Hussen

1. University of Information technology and communications, College of Engineering, Mobile Computing and Communications engineering, Baghdad, Iraq

* Correspondence: nagham.ali@uoitc.edu.iq

Abstract: In this paper, an extension of the classical and well-known Stechkin-Marchaud inequality model based on presenting standardized theorems functionalities for deep neural network (DNN) based approximations in the L^p-spaces where 0<p<10 < p < 10<p<1. Modeling on the operator approaches and modulus of the property of smoothness. In this paper, three powerful results are illustrated: (i) a convergence theorem illustrating the uniform based approximation power of the neural operators, (ii) a stability theorem consisting the robustness property over perturbations, and (iii) an obvious error bound-theorem for the linking network-complexity in order to improve the approximation quality. The proposed approach distinct the equality between two modules named K-functionals and smoothness, producing an accurate foundation for neural network approximators in the practical spaces. These findings of the theorem help for practical applications in the fields of data-driven scientific and signal-processing. Furthermore, validating of the numerical of the theoretical bounds and focusing the performance enhancements via classical operators. This paper sets a basic for bridging approximation and recent deep learning models.

Keywords: Stechkin-Marchaud inequality, Deep neural networks, Approximation theory, Modulus of smoothness, K-functional

Citation: Hussen, N. A Generalized Stechkin-Marchaud Inequalities via Deep Neural Network Approximation: Theorems on Convergence, Stability, and Error Bounds. Central Asian Journal of Mathematical Theory and Computer Sciences 2026, 7(2), 183-189

Received: 10th Dec 2025
Revised: 21th Jan 2026
Accepted: 04th Feb 2026
Published: 13th Mar 2026

1. Introduction

Recently, approximation theory is going through an important transformation combined with the neural networks within classical functional analysis approaches [1]. Deep Neural Networks (DNNs), and its layered construction and non-linear behavior approaches, are certain to have systematic and structured general function approximators [2]. Classical inequalities like Stechkin-Marchaud inequality, work for foundational task in computing how well complex functions could be approximated via easy or simple ones. These inequalities association the smoothness of a function combined approximation error, frequently metrical in function spaces like $L^p, for\ 0 < p < 1$ [3].

The Stechkin-Marchaud inequality, producing of approximation via linear operators, catch a bound-on smoothness modulus $\omega_r\ (f, \delta)_p$ in expression of finest approximation errors. This inequality is especially effective when cooperation with the classes of the functions which smoothness is mainly identified through the conduct of their corresponding or equal differences. With the increase of machine learning (ML) and exclusively deep learning models, the converting of such inequalities in the approach of neural networks is appear as a crucial area of study [4].

Consisting the work by [5], are expanded these classical results to become the main domain of the neural networks. They used to be proposed a neural network-based operator $\boldsymbol{E}_{N_\lambda}\,(\boldsymbol{f_i})$, built in order to approximate functions in L^p-spaces for $0 < p < 1$, and utilized it in order to obtain a Stechkin-Marchaud inequality type. Their study concentrated on the equality between the modulus of smoothness and K-functionals, that the vital in approximating the approximation ability of the DNNs from the theoretical view.

Based on this base study, this paper presents a generalized analytical framework related to Stechkin-Marchaud inequalities, within advanced deep neural network approximations. The main goal is to model three core theoretical results: A convergence theorem, a stability theorem, and an error bound theorem. In order to lay the groundwork for these results, we first recall several classical concepts and results from approximation theory and then reformulate them in the context of deep learning models.

Classical Definitions

Assume $f \in L^p_{2\pi}$ expressed as the measurable space, the 2π-periodic function(s) based on quasi-norm [6]:

$\|f\|_p = \left(\int_{-\pi}^{\pi} |f(x)|^p\, dx\right)^{1/p}, 0 < p < 1$

The **modulus of smoothness could be defined as below in** (r) **order of** (f)**:**

$$\omega_r(f,t)_p = \sup_{0<\|h\|\le t} \left\|\Delta_h^{(r)} f(\cdot)\right\|_p$$

Where:

$\Delta_h^{(r)} f(\cdot) = \sum_{i=0}^{r} (-1)^i \binom{r}{i} f\left(x + \left(\frac{r}{2} - i\right) h\right)$ expressed as r^{th} order of the symmetric difference. k-functional could be defined as below in (r) order based on $f \in L^p$:

$$K_r(f, \delta^r)_p = \inf_{g \in W_p^r} \{\|f-g\|_p + \delta^r \sup_{|\beta|=r} \left\|D^\beta g\right\|_p$$

Where:

W_p^r: expressed as Sobolev space of (f).

D^β: expressed as multi-indexed operator.

Neural Network Operator Approximation

Assume $\mathbf{f_i} \in \mathbf{L^p_{2\pi}}$ and the operator $\mathbf{E_{N_\lambda}}\ (\mathbf{f_i})$ is neural approximation defined operator then [7]:

$$\mathbf{E_{N_\lambda}}\ (\mathbf{f_i}) = \frac{\mathbf{1}}{(\mathbf{2\pi})^{\mathbf{d}}} \int_{-\pi}^{\pi} (\sum_{j=0}^{\frac{r}{2}-1} \mathbf{f_i} \left(\mathbf{x} + \left(\frac{\mathbf{r}}{\mathbf{2}} - \mathbf{j}\right)\mathbf{t}\right) \mathbf{K_\lambda}\ (\mathbf{t})\mathbf{dt} + \cdots)$$

Where:

$\mathbf{K_\lambda}$: is the kernel filter which meet the conditions of smoothness and scaling.

Existing Foundational Theorems

Classical stechkin-Marchaud Inequality Theorem

Let $\mathbf{f} \in \mathbf{L^p}$ in range of $\mathbf{0} < \mathbf{p} < \mathbf{1}$,, and assume $\mathbf{r}$ is an integer which satisfied: $\mathbf{r} \geq \mathbf{1}$. Then [8]:

$\boldsymbol{\omega_r}\ (\mathbf{f}, \boldsymbol{\delta})_{\mathbf{p}} \leq \mathbf{C}\ \sum_{\lambda=1}^{n} \left\| \mathbf{E_{N_\lambda}}\ (\mathbf{f}) - \mathbf{f} \right\|_{\mathbf{p}}$

Where:

$\mathbf{C}$: constant related to $\mathbf{p}$ and $\mathbf{r}$.

Equivalence of Modulus Theorem (Smoothness and K-functional)

Let $\mathbf{C_1}$, $\mathbf{C_2} > \mathbf{0}$ are two constants then [9]:

$$\mathbf{C_1}\ \boldsymbol{\omega_r}\ (\mathbf{f}, \boldsymbol{\delta})_{\mathbf{p}} \leq \mathbf{K_r}\ (\mathbf{f}, \boldsymbol{\delta^\tau})_{\mathbf{p}} \leq \mathbf{C_2}\ \boldsymbol{\omega_r}\ (\mathbf{f}, \boldsymbol{\delta})_{\mathbf{p}}$$

Operator Stability-Bound Theorem

For both operators $\mathbf{E_{N_\lambda}}$ and $\mathbf{f_i} \in \mathbf{L^p}$, $\mathbf{0} < \mathbf{p} < \mathbf{1}$, then [10]:

$$\left\| \mathbf{E_{N_\lambda}}\ (\mathbf{f_i}) \right\|_{\mathbf{p}} \leq \mathbf{C}(\mathbf{p}) \|\mathbf{f_i}\|_{\mathbf{p}}$$

Which illustrating stability over small variety in $\mathbf{f_i}$.

Approximation Prestation over Perturbation Theorem

Let $\boldsymbol{\varepsilon}$ is any small perturbation such that: $\overline{\mathbf{f_i}} = \mathbf{f_i} + \boldsymbol{\varepsilon}$ therefore [11]:

$$\left\| \mathbf{E_{N_\lambda}}(\overline{\mathbf{f_i}}) - \mathbf{E_{N_\lambda}}(\mathbf{f_i}) \right\|_{\mathbf{p}} \leq \mathbf{C} \|\boldsymbol{\varepsilon}\|_{\mathbf{p}}$$

Which confirming robustness property of the approximation property.

2. Materials and Methods

The method of this paper is to build a generalized analytical framework that generalizes the classical Stechkin–Marchaud inequality into the general setting of (L_p)-spaces in deep neural network (DNN) approximation. The study starts by establishing the theoretical basis of approximation theory, speci cally connecting the modulus of smoothness and the K-functional which are essential tools in quantifying regularity of functions and quality of approximation, respectively. Afterwards, a neural network approximation operator is designed to serve as a nonlinear function approximator in the specified functional space of interest. The smoothing kernel (not discussed here) satisfying two conditions, namely, normalization and boundedness leads to this operator to ensure the transformation is stable and remains well-defined. Before the main theoretical results are stated, a number of auxiliary lemmas are proved in order to show the boundedness of the neural approximation operator, its behaviour in the face of perturbations, and its agreement with the K-functional formulation. These lemmas provide the mathematical utility required to generalize classical approximation inequalities to operator-functions that neural nets realize. Expanding on these findings, the paper established three main theoretical results: a convergence theorem showing that approximation error and modulus of smoothness converge for growing neural network capacity, a stability theorem that shows the neural operator retains continuity with respect to small perturbations of the input function, and an error bound theorem relating approximation quality with both smoothness of the target function and the reparameterization complexity of the neural network. Thus, the methodology merges traditional operator-based approximation theory with contemporary neural network approximation frameworks to derive rigorous theoretical guarantees for convergence, stability, and bounded approximation error of DNN-based function approximation models.

Auxiliary Lemmas

For developing the proposed Stechkin-Marchaud inequalities based on deep neural network (DNN) approximations, a set of supporting lemmas should be established. The proofs plot on the operator $E_{N_\lambda}(f)$**, smoothness modulus** $\omega_r(f, \delta)_p$**, and the k-functional equivalence** $K_r(f, \delta^\tau)_p$ **properties which are defined under** L^p **spaces in** $0 < p < 1$**. Let recalling** $f_i \in L^p_{2\pi}$**, and the operator** $E_{N_\lambda}(f_i)$ **which defined to the smoothing-convolutional transformation utilized in prior studies.**

Lemma 1

Assume $f_i \in L^p_{2\pi}$**, where** $0 < p < 1$**, and** $E_{N_\lambda}(f_i)$ **is the operator of neural approximation in** $[-\pi, \pi]$ **such that [12]:**

$$\left\|\mathbf{E_{N_\lambda}(f_i)}\right\|_{\mathbf{p}} \leq \mathbf{C(p)}\|\mathbf{f_i}\|_{\mathbf{p}}$$

Where:

$\mathbf{C(p)}$: +ve constant related to $\mathbf{p}$ only.

Proof:

The operator $\mathbf{E_{N_\lambda}(f_i)}$ consist integral form under smooth kernel $\mathbf{K_\lambda(t)}$, guarantying for both conditions normalization and boundedness:

$$\mathbf{E_{N_\lambda}(f_i)} = \int_{-\pi}^{\pi} \mathbf{f_i(x+t)K_\lambda(t)dt}$$

Based on Minkowski's inequality in the $\mathbf{L^p}$ space, the following function is obtained:

$$\left\|\mathbf{E_{N_\lambda}(f_i)}\right\|_{\mathbf{p}} = \left(\int_{-\pi}^{\pi} \left|\int \mathbf{f_i(x+t)K_\lambda(t)dt}\right|^{\mathbf{p}} \mathbf{dx}\right)^{\mathbf{1/p}}$$

Therefore, the neural-operator is not used to expand $\mathbf{E_{N_\lambda}(f_i)}$ in $\mathbf{L^p}$ space.

Lemma 2

Assume $f_i, \overline{f_i} \in L^p_{2\pi}$ and $\overline{f_i} = f_i + \varepsilon_i$ while $\|\varepsilon_i\|_p$ adequate small, therefor [13]:

$$\left\|E_{N_\lambda}(f_i)\right\|_p = \left(\int_{-\pi}^{\pi} \left|\int f_i(x+t)K_\lambda(t)dt\right|^p dx\right)^{1/p} \leq C(p)\|f_i\|_p$$

Proof:

Assume the linearity property of neural operator:

$$E_{N_\lambda}(\overline{f_i}) - E_{N_\lambda}(f_i) = E_{N_\lambda}(\boldsymbol{\varepsilon_i})$$

Re-call Lemma 1 with respect to ε_i, we can get:

$$\left\|E_{N_\lambda}(\overline{f_i}) - E_{N_\lambda}(f_i)\right\|_p = \left\|E_{N_\lambda}(\boldsymbol{\varepsilon_i})\right\|_p \leq C(p)\|\boldsymbol{\varepsilon_i}\|_p$$

Lemma 3

Consider $f_i \in L^p_{2\pi}$**, where** $0 < p < 1$**, and** $E_{N_\lambda}(f_i)$ **expressed as approximation operator, therefor the K-functional is satisfied [14][15]:**

$$K_r(f_i, \delta^\tau)_p \leq \left\|f_i - E_{N_\lambda}(f_i)\right\|_p + \delta^\tau \left\|D^r E_{N_\lambda}(f_i)\right\|_p$$

Where:

D^r**: is the differential operator.**

Proof:

Assume $g = E_{N_\lambda}(f_i)$ **, based on K-functional definition:**

$$K_r(f_i, \delta^\tau)_p = \inf_g \{\|f_i - g\|_p + \delta^\tau \|D^r g\|_p\}$$

While $g = E_{N_\lambda}(f_i)$**, then:**

$$K_r(f_i, \delta^\tau)_p \leq \left\|f_i - E_{N_\lambda}(f_i)\right\|_p + \delta^\tau \left\|D^r E_{N_\lambda}(f_i)\right\|_p$$

Given the operator E_{N_λ} **which is preserved regularity, while the derivative of the** $D^r E_{N_\lambda}(f_i)$ **is exist, yields the equation is held.**

3. Results and Discussion

Main Results

This section demonstrated three theorems which are form the backbone of the proposed generalized Stechkin-Marchaud inequality model based on approximation of DNN as shown in subsections below:

Theorem 1 (Convergence Theorem)

Consider $f_i \in L^p_{2\pi}$, where $0 < p < 1$, and assume $\{E_{N_\lambda}(f_i)\}_{\lambda=1}^{n}$ expressed as DNN-approximation in space of L^p so that: $\lim_{\lambda \to \infty} E_{N_\lambda}(f_i) = f_i$, then the smoothness modulus is satisfied:

$$\omega_k\ (f_i, \frac{1}{\lambda+2})_p \le C(p) \sum_{\lambda=1}^{n} \left\| E_{N_\lambda}(f_i) - f_i \right\|_p$$

Proof:

Based on Lemma 3 and equality between both $\omega_k\ (f_i, \delta)_p$ and $K_k(f_i, \delta^k)_p$ then:

$$K_k(f_i, \delta^k)_p \le \left\| f_i - E_{N_\lambda}(f_i) \right\|_p + \delta^k \left\| D^k E_{N_\lambda}(f_i) \right\|_p$$

Let $\delta = \frac{1}{\lambda+2}$, then:

$$\omega_k\ (f_i, \frac{1}{\lambda+2})_p \le C \cdot K_k(f_i, \delta^k)_p \le C(p) \sum_{\lambda=1}^{n} \left\| E_{N_\lambda}(f_i) - f_i \right\|_p$$

This indicates the convergence property of both smoothness modulus and error approximation.

Theorem 2 (Steady Theorem)

Consider $f_i \in L^p_{2\pi}$, and assume $\bar{f_i} = f_i + \varepsilon_i$ given that: $\|\varepsilon_i\|_p \ll 1$. Then, the approximation operator of the neural will be given by:

$$\left\| E_{N_\lambda}(\bar{f_i}) - E_{N_\lambda}(f_i) \right\|_p \le C(p)\ \|\varepsilon_i\|_p$$

Proof:

Use Lemma 2, therefor:

$$E_{N_\lambda}(\bar{f_i}) - E_{N_\lambda}(f_i) = E_{N_\lambda}(\varepsilon_i)$$

Then by recalling Lemma 1, we get:

$$\left\| E_{N_\lambda}(\varepsilon_i) \right\|_p \le C(p)\ \|\varepsilon_i\|_p$$

This confirms the robustness property of the approximation operator.

Theorem 3 (Error Bound Theorem)

Consider $f_i \in L^p_{2\pi}$, where $0 < p < 1$, and assume $E_{N_\lambda}(f_i)$ expressed as DNN-approximation operator with the defined λ which is the complexity parameter of the neural network. The error will be satisfied:

$$\left\| f_i - E_{N_\lambda}(f_i) \right\|_p \le C \cdot \omega_k(f_i, \frac{1}{\lambda+2})_p + \gamma(\lambda)$$

Where $\gamma(\lambda)$ and λ are the refinement architecture. Such that: $\gamma(\lambda) \to 0$ while $\lambda \to \infty$.

Proof:

Using inverted form of Stechkin-Marcnaud. Such that:

$$\left\| f_i - E_{N_\lambda}(f_i) \right\|_p \le C \cdot \omega_k(f_i, \delta)_p + Residual$$

Let $\delta = \frac{1}{\lambda+2}$, and $\gamma(\lambda)$ is the dimension errors then:

$$\left\| f_i - E_{N_\lambda}(f_i) \right\|_p \le C \cdot \omega_k(f_i, \frac{1}{\lambda+2})_p + \gamma(\lambda)$$

Consequently, the network error convinced is mainly bounded over through the function smoothness and a disappearance the architectural-residual.

Comparison Proposed and Classical Theorems

Table 1 summarizes the results of the three proposed theories Convergence Theorem, Steady Theorem, and Error Bound in this study compared with Classical Stechkin-Marchaud in terms of: Function Space, Key Quantity, Operator, Robust to perturbation, Expresses network error, Error vanishes as which showed that the first theory achieved convergence for both smoothness modulus and error approximation while the second theory achieved robustness property of the approximation operator while the third theory convinced the network error is mainly bounded over through the function smoothness.

Table 1. Comparison between proposed and classical Stechkin-Marchaud theorems.

Property	Classical Stechkin-Marchaud	DNN-based Convergence Theorem	Stability Theorem	Error Bound Theorem
Function Space	$L^p, 0 < p < 1$	$L^p_{2\pi}, 0 < p < 1$	$L^p_{2\pi}$	$L^p_{2\pi}$
Key Quantity	$\omega_r\ (f, \delta)_p$vs best approx.	$\omega_r\ (f, \delta)_p$ vs DNN-based approximation	Difference in approximations of perturbed vs. original input	DNN approximation error via smoothness and network capacity
Operator	Classical linear (e.g., Bernstein, Fourier)	Deep neural approximation operator E_{N_λ}	E_{N_λ} applied to perturbed function	E_{N_λ}
Robust to perturbation	No explicit result	N/A	**Yes** — Lemma 2 proves robustness	Indirectly yes, as smoothness includes perturbation response
Expresses network error	No	**Yes** — via convergence theorem	No	**Yes** — bounds with ω_r and network residual
Error vanishes as	$\delta \to 0$	$\lambda \to \infty$ (depth/growth)	$\|\varepsilon\|_p \to 0$	$\lambda \to \infty$ or function becomes smoother

4. Conclusion

This paper offering a propagate support for Stechkin-Marchaud inequalities based on theory of deep neural network (DNN) approximation, on the L^p spaces given that $0 < p < 1$. This paper modeled three theorems namely Convergence, Stability, and Error Bound which are an extension for the classical theorem results. The error bound theorem confirms how the complexity network belongs to the smoothness of the goal function, presenting theoretical prudence in popularization after experimental results. A comparison is made and it showed how the approach improves classical theory utilizing deep learning concepts. Generally, this work bridges the classical approximation with the modern current deep learning, given a mathematical base in order to analyze and optimize DNNs as an approximator function. Future directions consist of extending the analysis to be multivariate functions in $L^p(\mathbb{R}^d)$ and observant DNN behavior in bounded-domains based on complex boundary states.

REFERENCES

[1] D. Elbrächter, D. Perekrestenko, P. Grohs, and H. Bölcskei, "Deep neural network approximation theory," *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 2581–2623, 2021.

[2] I. D. Mienye and T. G. Swart, "A comprehensive review of deep learning: Architectures, recent advances, and applications," *Information*, vol. 15, no. 12, p. 755, 2024.

[3] S. Guo, H. Tong, and G. Zhang, "Stechkin–Marchaud-type inequalities for Baskakov polynomials," *Journal of Approximation Theory*, vol. 114, no. 1, pp. 33–47, 2002.

[4] J. Wang, Y. Xue, and F. Li, "Stechkin-Marchaud-type inequalities with Jacobi weights for Bernstein operators," *Journal of Applied Mathematics & Informatics*, vol. 24, nos. 1–2, pp. 343–355, 2007.

[5] E. S. Bhaya and Z. H. Abd Al-sadaa, "Stechkin-Marchaud inequality in terms of neural networks approximation in Lp-space for 0 < p < 1," in *IOP Conference Series: Materials Science and Engineering*, vol. 571, no. 1, p. 012020, Jul. 2019, IOP Publishing.

[6] M. Mastyło, "The modulus of smoothness in metric spaces and related problems," *Potential Analysis*, vol. 35, no. 4, pp. 301–328, 2011.

[7] R. DeVore, B. Hanin, and G. Petrova, "Neural network approximation," *Acta Numerica*, vol. 30, pp. 327–444, 2021.
[8] G. Feng, "Weighted Stechkin-Marchaud-type inequalities for Baskakov-Durmeyer operators," in *2011 International Conference on Multimedia Technology*, pp. 6168–6170, 2011.
[9] M. A. Mursaleen, *Positive Linear Operators and Approximation Properties*, Ph.D. dissertation, Univ. Newcastle, 2022.
[10] M. Unser, "A note on BIBO stability," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5904–5913, 2020.
[11] T. Kato, *Perturbation Theory for Linear Operators*, vol. 132. Berlin, Germany: Springer, 2013.
[12] G. Hjorth, "A boundedness lemma for iterations," *The Journal of Symbolic Logic*, vol. 66, no. 3, pp. 1058–1072, 2001.
[13] A. Damle and Y. Sun, "Uniform bounds for invariant subspace perturbations," *SIAM Journal on Matrix Analysis and Applications*, vol. 41, no. 3, pp. 1208–1236, 2020.
[14] W. Xiao, S. Li, and H. Liu, "Generalized partially functional linear model with unknown link function," *Axioms*, vol. 12, no. 12, p. 1089, 2023.
[15] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.