Application of Differential Equations in Neural Networks

Majid Bagheri [1*]

[1] Applied Mathematics Department, Ahar Branch, Islamic Azad University, Ahar, Iran.

*Corresponding Author: majid.bagheri5391@gmail.com

Receive Date: 07 January 2025 Revise Date: 15 January 2025 Accept Date: 19 April 2025

Abstract

Generative artificial intelligence (AI) has evolved dramatically, encompassing many techniques to create novel and diverse data. While models like Generative Adversarial Networks (GANs) and Variational Auto Encoders (VAEs)have taken center stage, a lesser-explored but incredibly intriguing avenue is the realm of Neural Differential Equations (NDEs) which tries to consider to the unknown territory of NDEs in the generation of artificial intelligence. In this research, the relationship and interaction of mathematical branches, especially integral differential calculus and fractional calculus, with the learning of electric machines is stated. The main aim of the research is that, firstly, the modeling of neural network phenomena in the form of differential equations of fractional order is superior to its classical form in terms of time memory mechanism and system dynamics. Secondly, the method of converting the above-mentioned equation into an ordinary differential equation in terms of finding its solutions is provided.

Keyword: artificial intelligence; Neural Differential Equations; differential equation;

1. Introduction

NDEs[1] fuse principles of differential equations and NNs[2], resulting in a dynamic framework that can generate continuous and smooth data. Traditional generative models often generate discrete samples, limiting their expressive power and making them unsuitable for applications that require continuous data, such as time-series predictions, fluid dynamics, and realistic motion synthesis. NDEs bridge this gap by introducing a continuous generative process, enabling data creation that seamlessly evolves over time.

The research at intersection area between two scientific disciplines, promotes a cross fertilization and eventually produces remarkable results for both. Physics and Artificial Intelligence are an example of this. The use of AI techniques in the characterization or forecast of systems in physics has been a constant in recent years. In return, physics has collaborated in this symbiosis explaining, from the perspective of statistical mechanics, the training mechanism of neural networks, among other contributions.

In the specific case of Dynamic Systems and Machine Learning, although there are many strategies that can be inscribed in its intersection, the mixed strategies between the theory of differential equations and neural networks stand out [1-2]. The research in this area, suggest that there exists a duality between differential equations and many deep neural networks that is worth trying to understand.

[1] .Neural Differential Equations

[2] . Neural Networks

2. Mathematical relational with artificial intelligence

What is artificial intelligence mathematics? Artificial intelligence mathematics is an important field in computer science that tries to simulate and imitate the intelligent power of humans in machines and computer systems by using mathematical concepts and techniques. The mathematics of artificial intelligence is used in some subfields of this science, such as machine learning, neural networks, and natural language processing.

For example, you may attend a training course and learn topics related to machine learning and artificial intelligence, but when you enter the real world, you will see that most of the problems are solved based on mathematical and statistical principles, and the algorithms follow the principles. are math. In this connection, there are important theories that are used directly and indirectly in the world of artificial intelligence.

2.1. Differential calculus

The most important goal of integral in artificial intelligence is to optimize and adjust machine learning parameters. If you are in the field of data science, learning calculus is essential. Among the most important tasks that are used with the help of differential and integral calculus in artificial intelligence are:

1) Create charts or visuals
2) Simulation
3) Creating programs to solve mathematical problems
4) Coding in programs
5) Algorithm design and analysis

2.2. Fractional calculus

Searching in the articles published in the last few decades in the fields of science and engineering, we more or less come across the topics of fractional calculus, differential equations with fractional derivatives and concepts of this type. These articles are both theoretical and practical, and they have made a significant contribution to research. So far, many books have been written in this field with theoretical and practical aspects [3-5]. Some writers introduced fractional derivative and integral as a generalization of classical concepts and tried to prove their claim.

2.3. Methodology

This research presents the latest and most up-to-date developments in fractional differential integral and fractional differential equations applied to neural networks, including many different potentially useful operators of fractional calculus. The subject of fractional calculus and its applications has gained considerable popularity and importance over the past three decades, therefore, largely because of its demonstrated applications in many seemingly diverse and vast fields of science and engineering. Some areas of current application of fractional models are: fluid flow, solute transport or dynamic processes in self-similar and porous structures, diffusion. Diffusion-like transfer, viscoelastic theory of materials, electromagnetic theory, earthquake dynamics, control theory of dynamic systems, optics and signal processing, biological sciences, economics, geology, astrophysics, probability and statistics, chemical physics and especially neural networks. Their complex microscopic behavior and macroscopic dynamics cannot be described classically. Fractional models express some solutions that such features are impossible for conventional models. What are the useful properties of these fractional operators that help in modeling many unusual processes? From the authors' point of view and from the known experimental results, most of the processes related to complex systems have non-local dynamics that include long memory in time, and fractional derivative and integral operators have some of these characteristics.

2.4. Definition of derivative of fractional order

Abdon Atangana suggested a beta derivative [6]. The proposed version has several properties that are a limitation for fractional derivatives and have been used to model some physical problems. These derivatives may not be considered as fractional derivatives, but they can be considered as a natural extension of classical derivatives [7]. The derivative is defined as follows:

$$D_x^{\alpha} f(x) = \lim_{\theta \to 0} \frac{f\left(t + \theta\left(t + \frac{1}{\mathcal{T}(\alpha)}\right)^{1-\alpha}\right) - f(x)}{\theta} \quad (1)$$

$$t > 0 \;, \alpha \in (0,1], D = \frac{d}{dx}$$

where T(.) is the gamma function.

Its inverse (integral) operator is in the following form:

$$I_x^{\alpha} f(x) = \int_a^x \left(t + \frac{1}{\mathcal{T}(\alpha)}\right)^{\alpha-1} f(t)dt \quad (2)$$

In the above definition, by placing $h = \theta\left(t + \frac{1}{\mathcal{T}(\alpha)}\right)^{1-\alpha}$ and if $\theta \to 0$ then $h \to 0$, so $D_t^{\alpha} f(x) = \left(t + \frac{1}{\mathcal{T}(\alpha)}\right)^{1-\alpha} \frac{df(f(t))}{dt}$

Considering:

$$\eta = \frac{l}{\alpha}\left(t + \frac{1}{\mathcal{T}(\alpha)}\right)^{\alpha} \quad (3)$$

where l is an arbitrary constant, according to the chain rule in the derivative we have:

$$D_t^{\alpha} f(\eta) = l\frac{df(\eta)}{d\eta} \quad (4)$$

2.5. Application of differential equations in artificial intelligence

Differential equations are one of the basic tools in mathematics that are used in many different scientific fields, including computer science and artificial intelligence. These equations are in the form of functions that govern the behavior and changes of the systems and can be used to accurately predict the state and temporal changes of the systems. In the field of artificial intelligence, the use of differential equations is very useful for modeling and recognizing complex patterns and behaviors in systems. These equations can be used as a tool to optimize algorithms, predict and manage the behavior of systems, and even design and simulate neural networks.

Also, differential equations are used as an important tool in creating complex and automatic artificial intelligence systems. These equations can help to detect complex patterns, control dynamic systems and facilitate the understanding of systems behavior. In general, the use of differential equations in artificial intelligence helps machines to act intelligently and make automatic decisions.

2.6. Neural Differential Equations (NDE)

Eq. NDE: $D(Z(t)) = f_{\omega}(Z(t), t)$, $D = \frac{d}{dt}$ (5)

In this equation:

•z(t), represents the state of the system (or the "hidden state") at any time t. In the context of neural networks, this can be thought of as the activations at any "depth" or point in the processing of the input data.
•f, is a function parameterized by ω (the parameters or weights of the neural network), which specifies the rate of change of the hidden state. This function is typically modeled using neural network layers themselves, which are designed to be differentiable with respect to both the hidden state and the parameters.
•t, explicitly denotes time or depth within the model, introducing a notion of continuous transformation depth, unlike traditional models where depth is implicitly defined by discrete layers [9-12].

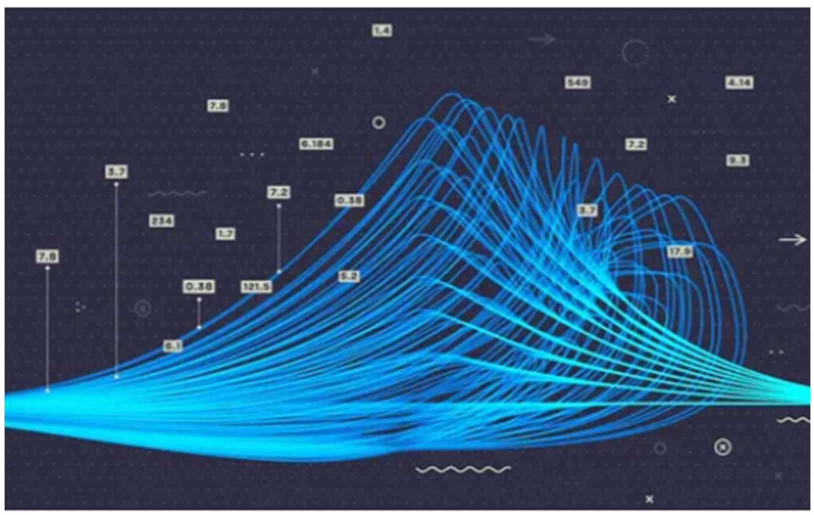

Fig.1. Eq. NDE in NNs

2.7. Fractional Neural Differential Equation (FNDE)

$$D^{\alpha}(Z(t)) = f_{\omega}(Z(t), t) \quad , \quad D^{\alpha} = \frac{d^{\alpha}}{dt} , \quad 0 < \alpha \leq 1 \tag{6}$$

The reasons for using fractional derivatives in computer artificial intelligence algorithms and neural differential equations are that they have non-local dynamics and include long memory in time, so it is more consistent with reality.

2.8. The method of converting FDE[3] to ODE[4]

Using a transformation:

$$\eta = \frac{l}{\alpha}\left(t + \frac{1}{\Upsilon(\alpha)}\right)^{\alpha}, Z(t) = U(\eta),$$

And considering relationship (4), equation (6) becomes the following ODE equation:

$$lD(U(\eta)) = g_{w}(U(\eta), \eta) \tag{7}$$

[3] . Ordinary Differential Equations
[4] . Fractional Differential Equations

which is the same equation (5) in terms of time η in the role of t variable and it is solved in the form of an integral differential equation through analytical methods and numerical methods, which is beyond the scope of this research.

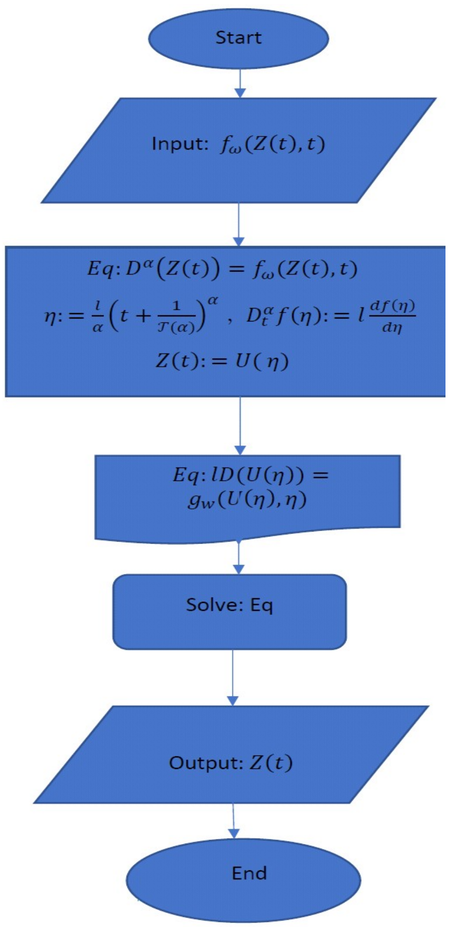

Fig.2. Algorithm flowchart

3. Algorithm analysis

The neural network differential equation formation algorithm has been modeled based on the researches that have been presented in the form of classical differential equation, which have created problems such as lack of information and data storage and dynamic locality in machine learning [13]. In this article, the differential equation governing neural networks is discussed and investigated, and through the derivative of the beta fraction and its properties, it has been converted into a classical differential equation, which has solved the above problems by obtaining the solution of this equation [8].

3.1. Case Study: Forecasting with AI

Consider the dividends prediction problem. It is a complex system influenced by countless factors, which can be modeled using differential equations. By analyzing historical data and past information with a parameterized function $f_{\omega}(Z(t),t)$, a differential equation can becreated that describes how stock prices change over time(t). The solution to this equation provides a function Z(t) that can predict future prices under certain conditions.

For example, suppose $f_{\omega}(Z(t),t) = -2tZ$ and $\alpha = \frac{1}{2}$, $l = 1$,$\mathcal{T}\left(\frac{1}{2}\right) = \sqrt{\pi}$, that the solution equation (7)($D^{1/2}(Z(t)) = -2tZ$)can be obtained graphically as follows:

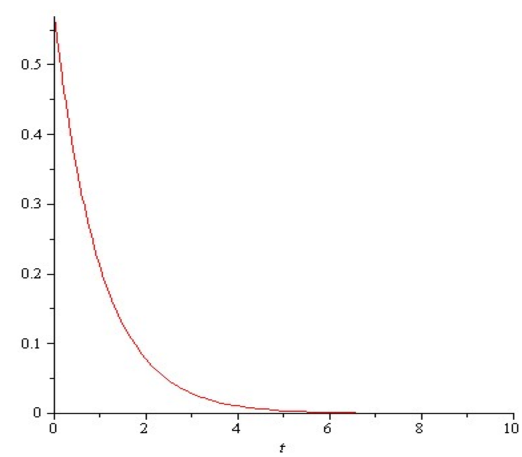

Fig.3. $Z(t) = e^{-\left(t+\frac{1}{\sqrt{\pi}}\right)}$

4. Conclusion

The interplay between mathematics, particularly differential equations, and AI is a testament to the multidisciplinary nature of solving complex problems. As we continue to push the boundaries of what's possible with AI, the foundational role of differential equations in understanding and predicting the behavior of complex systems remains undiminished. For enthusiasts and professionals alike, the journey through this mathematical landscape is both challenging and rewarding, offering insights that are crucial for advancements in AI and beyond.

References

[1] Shin.Y.H. et al. Hopfield-type neural ordinary differential equation for robust machine learning ,Pattern Recognit Lett, 2021.

[2] Lei. D. et al.Neural ordinary differential grey model and its applications, Expert Syst Appl, 2021.

[3] Diethelm. K., The Analysis of Fractional Differential Equations, Springer-Verlag, Berlin,2010.

[4] Li .C. andZeng. F., Numerical Methods for Fractional Calculus, CRC press, New York, 2015.

[5] Podlubny. I., Fractional Differential Equations. Academic Press, San Diego and London,1999.

[6] Atangana, A.,Baleanu, D., and Alsaed . A.(2016),Analysis of time-fractional Hunter–Saxton equation: a model of neumatic liquid crystal, inOpen Phys.,14:145-149,2016.

[7] Khalil, R.,AlHorani, M.,Yousef, A., and Sababheh, M.(2014),A new definition of fractional derivative, 264(2014) 65–70. InJ. Comput. Appl. Math.,264:65-70, 2014.

[8] Yusuf, A.,Inc,M., Aliyu, A.I., and Baleanu, D.Optical Solitons Possessing Beta Derivative of the ChenLee-Liu Equation in Optical Fibers,in Front. Phys., 7:1-7, 2019.

[9] Chen, Ricky T. Q.; Rubanova, Yulia; Bettencourt, Jesse; Duvenaud, David K. neural Ordinary differential equations In Bengio, 2018.

[10] Wallach .S., Larochelle .H., Grauman, K.,Cesa-Bianchi, N., Garnett, R, Advances in Neural Information Processing Systems". Vol. 31. Curran Associates, Inc,arXiv:1806.07366.

[11] Ruiz-Balet, D; Zuazua, E.Neural ODE Control for Classification, Approximation, and Transport. SIAM Review. 65 (3): 735–773.2023.

[12] Patrick. K., On Neural Differential Equations (PhD). Oxford, United Kingdom: University of Oxford, Mathematical Institute,2021.

[13] Raissi.M. and Karniadakis. G. E.. Hidden physics models: Machine learning of nonlinear partial differential equations. Journal of Computational Physics, pages 125–141, 2018.