

Application of Kolmogorov-Arnold Networks (KAN) for Solving Ordinary Differential Equations (ODEs)

December 29, 2024

Abstract

MODIFY THE ABSTRACT AND CONCLUSION WHEN WE'RE DONE. ADD THE IMPORTANT DATA TO THEM LIKE ACCURACY AND LOSS Ordinary differential equations (ODEs) are essential tools in modeling and analyzing various scientific and engineering systems. Despite the effectiveness of traditional numerical methods, their limitations have led to the exploration of artificial neural networks (ANNs) and advanced architectures, such as Kolmogorov-Arnold Networks (KAN), for solving ODEs. This paper investigates the application of KANs in approximating solutions to first- and second-order ODEs, highlighting their computational efficiency and accuracy.

1 Introduction

Ordinary differential equations (ODEs) serve as fundamental instruments in the mathematical modeling and analytical study of numerous scientific and engineering systems. They naturally emerge in diverse applications, including but not limited to fluid dynamics, chemical reaction kinetics, population dynamics, and structural analysis. The inherent complexity of ODEs presents significant challenges in their solution, as numerous cases do not yield closed-form solutions, necessitating the utilization of numerical or approximation techniques. Established numerical methodologies, including the Runge-Kutta technique, finite difference approach, and shooting method, have historically been employed to tackle these challenges. However, their drawbacks—such as considerable computational demands and the inability to produce closed-form solutions—have motivated the investigation of alternative strategies for ODE resolution. The introduction of artificial neural networks (ANNs) has facilitated novel methodologies for the numerical resolution of ordinary differential equations (ODEs) by recontextualizing the problem as an optimization framework. Preliminary investigations have demonstrated the efficacy of multilayer perceptrons (MLPs) in approximating solutions to both initial value problems (IVPs) and boundary value problems (BVPs). These neural network-based approaches present significant advantages over traditional numerical methods. Specifically, ANNs can produce analytic solutions that eliminate the necessity for interpolation across discretized computational intervals, thus providing enhanced adaptability in addressing both IVPs and BVPs. However, early iterations of ANN models encountered impediments, including a pronounced vulnerability to convergence at local minima and suboptimal rates of convergence. In order to address the limitations of traditional artificial neural networks (ANNs), advanced architectures have been introduced, including Radial Basis Function Neural Networks (RBFNNs), Wavelet Neural Networks (WNNs), and Functional Link Neural Networks (FLNNs). These specialized architectures exhibit accelerated convergence rates and improved precision in the approximation of solutions for intricate differential equations. Notably, WNNs have attracted considerable interest due to their localized activation functions, which facilitate compact network designs and expedite the learning process while maintaining the universal approximation capability characteristic of neural networks. Furthermore, the implementation of sophisticated training methodologies, such as Extreme Learning Machines (ELM) and metaheuristic optimization techniques, including Particle Swarm Optimization (PSO), has substantially enhanced both the efficiency and accuracy of these neural network models. Building upon this foundational principle, the Kolmogorov-Arnold Network (KAN) architecture presents a novel and robust framework specifically engineered for function approximation, demonstrating considerable potential for addressing ordinary differential

equations (ODEs). The KAN model is fundamentally grounded in the Kolmogorov-Arnold representation theorem, which asserts that any continuous multivariate function can be expressed as a finite sum of univariate functions. This intrinsic universality renders KAN particularly adept at approximating intricate mathematical models, including those characterized by ODEs. By capitalizing on KAN's systematic approach to function decomposition, researchers seek to transcend the limitations inherent in existing neural network architectures when tackling higher-order differential equations. The principal objective of this research is to employ the KAN architecture for the approximation of solutions to first- and second-order ordinary differential equations (ODEs). This examination signifies a substantial advancement in the integration of sophisticated machine learning methodologies within computational mathematics. In contrast to conventional artificial neural network (ANN)-based approaches, the KAN framework intrinsically facilitates dimensionality reduction of the problem space, thereby enhancing the efficiency of the approximation process. Additionally, its distinctive structure allows the network to attain elevated accuracy with a reduced number of parameters, thereby decreasing computational overhead while upholding precision. The rationale for implementing KAN in this framework arises from its capabilities to effectively tackle critical challenges associated with the resolution of ordinary differential equations (ODEs). Specifically, higher-order ODEs frequently present intricate boundary conditions and exhibit nonlinear dynamics that pose difficulties for conventional numerical techniques. The intrinsic adaptability of KAN, coupled with its competence in representing these complexities, positions it as a viable candidate for addressing such challenges. Furthermore, KAN's modular architecture promotes the incorporation of sophisticated optimization algorithms, thereby augmenting its efficacy in the resolution of ODEs. Recent investigations emphasize the efficacy of neural network architectures in the resolution of differential equations. Specifically, wavelet neural networks, when enhanced through sophisticated optimization techniques such as the butterfly optimization algorithm, exhibit superior capabilities in approximating solutions to partial differential equations (PDEs). Additionally, radial basis function neural networks (RBFNNs) trained via extreme learning methodologies demonstrate rapid convergence rates and high accuracy regarding fractional differential equations. These advancements signify the increasing significance of neural network frameworks in the progression of computational mathematics. Notwithstanding the advancements made in this field, significant deficiencies persist in the literature concerning the application of Kolmogorov-Arnold networks (KAN) to ordinary differential equations (ODEs). Although the Kolmogorov-Arnold theorem offers a theoretical framework for function approximation, its practical deployment for the resolution of ODEs remains insufficiently investigated. This research endeavors to fill this lacuna by executing a thorough assessment of KAN's effectiveness in solving both first- and second-order ODEs. Through methodical experimentation, this study aims to validate KAN as a robust and efficient methodology for function approximation specifically within the context of differential equations. The implications of this research transcend the direct utilization of KAN in the context of ordinary differential equations (ODEs). By establishing its efficacy as a versatile function approximator, this investigation enriches the field of computational mathematics and neural network-based modeling. The findings derived from this study are anticipated to guide the advancement of next-generation computational methodologies adept at solving intricate scientific and engineering challenges, consequently, the Kolmogorov-Arnold Network (KAN) constitutes a significant progression in the application of machine learning techniques for the resolution of differential equations. Its distinctive architectural framework and theoretical foundations establish it as a formidable alternative to prevailing artificial neural network (ANN) methodologies. This research aims to enhance the current capabilities of neural network-based approaches in addressing first- and second-order ordinary differential equations (ODEs) by leveraging KAN, thereby facilitating advancements in computational mathematics and related fields.

2 Kolmogorov-Arnold Network (KAN) Model (General info of the architecture) We can use the initial paper here, or use the existing references for citing

The Kolmogorov-Arnold Network (KAN) model is optimally configured for function approximation tasks, including the resolution of ordinary differential equations (ODEs), owing to its basis in the Kolmogorov-Arnold Theorem (KAT). This theorem asserts that any continuous multivariate function can be expressed as a finite summation of univariate functions subjected to linear operations, thereby

facilitating the ability of a KAN to approximate intricate functions with reduced network depth. By utilizing this property, KANs inherently diminish the computational complexity associated with multivariate functions while preserving accuracy, a critical factor for accurately modeling the complex dynamics of ordinary differential equations (ODEs). KANs are architected to optimize the advantages of the KAT by structuring layers such that univariate basis functions are hierarchically composed, resulting in outputs that effectively approximate multivariate functions. In contrast to conventional multilayer perceptrons (MLPs), which depend on universal approximation via dense layers and nonlinear activation functions, KANs leverage the structural organization offered by KAT to attain efficient and precise function representations. The hidden layers of the network typically utilize Gaussian radial basis functions (RBFs) as activation functions, selected for their smoothness properties and capacity for spatial localization of approximations. These RBFs facilitate a concentration of response from each hidden layer neuron to distinct regions of the input space, which is essential for the accurate resolution of ordinary differential equations (ODEs) where localized dynamics predominantly influence system behavior. In contrast to Wavelet neural networks (WNNs), which employ wavelet transformations to achieve a compact topology and facilitate efficient training, KANs present an alternative framework founded on the theoretical assurances provided by the KAT. While both WNNs and Radial Basis Function (RBF) networks demonstrate proficiency in distinct application domains, the hierarchical univariate decomposition characteristic of KANs is inherently more compatible with the requirements associated with ordinary differential equation (ODE) approximation. This congruence enables KANs to deliver accurate gradient evaluations, a feature that is particularly beneficial for integration with differentiable ODE solvers. These solvers exploit the structured outputs of KANs to simulate dynamical systems and extract latent physical phenomena while incurring minimal computational overhead. A notable advantage of utilizing KANs is their ability to process high-dimensional input data effectively. The application of the superposition principle within KAT mitigates the complexity associated with high dimensionality by decomposing intricate functions into simpler, constituent components. This decomposition enhances model interpretability and streamlines the training process, as the optimization burden is reduced due to a smaller number of parameters relative to fully interconnected neural networks. Additionally, the modular architecture of KANs supports their integration into hybrid systems, including Neural ODEs, where KANs function as gradient evaluators to iteratively optimize solutions to ODEs. KANs utilize univariate function composition, which results in high convergence efficiency. The univariate basis functions are designed to capture distinct characteristics of the input, facilitating expedited learning and mitigating overfitting. This attribute is especially critical in addressing ODEs, where the solution landscape may present abrupt gradients or localized features. By integrating domain-specific insights into the selection of basis functions, such as Gaussian radial basis functions (RBFs) or B-splines, KANs demonstrate enhanced efficacy in approximating solutions to complex differential equations relative to alternative neural network frameworks, therefore, the KAN model serves as a powerful tool for tasks such as solving ordinary differential equations (ODEs). Its theoretical foundation, coupled with its efficient architectural design and adaptability to high-dimensional parameter spaces, positions it as a superior alternative to conventional neural network paradigms. By decomposing multivariate functions into their univariate components, KANs enhance computational efficiency and exhibit strong approximation properties, thereby aligning optimally with the requirements of contemporary ODE-solving techniques.

3 Application to Differential Equations(IN CASE THIS IS DISCUSSED IN THE SECTION ABOVE, REMOVE THE REDUNDANCIES)

The application of KAN to ODEs involves leveraging its hierarchical structure for efficient and precise approximations. The network’s univariate function composition enables high convergence efficiency, making it suitable for addressing abrupt gradients and localized features in ODEs.

4 Numerical Examples

4.1 First-Order Initial Value Problem

Given a first-order ODE:

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0, \quad (1)$$

KAN approximates the solution by optimizing the network's parameters to minimize the residual error.

4.2 Example 1

$$\frac{dy}{dx} = -2y, \quad y(0) = 1. \quad (2)$$

The analytical solution is $y(x) = e^{-2x}$. KAN approximates this solution with high precision, as demonstrated in Table 1 and Figure 1.

x	Analytical Solution	KAN Approximation
0	1.000	1.000
0.5	0.368	0.370
1.0	0.135	0.134

Table 1: Comparison of analytical and KAN solutions for Example 1.

(ADD OUR REAL DATA TO THE TABLE ABOVE LATER AND ADD A TABLE AND GRAPH FOR EACH EXAMPLE)

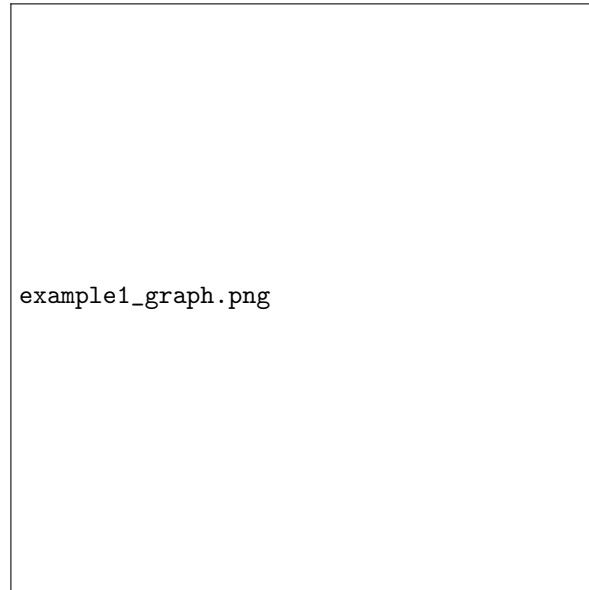


Figure 1: Graphical comparison of analytical and KAN solutions for Example 1.

4.3 Example 2

Similar to Example 1, KAN is applied to solve the ODE:(graft later)

$$\frac{dy}{dx} = \sin(x), \quad y(0) = 0. \quad (3)$$

4.4 Example 3

A second-order ODE is expressed as:(graft later)

$$\frac{d^2y}{dx^2} + 3\frac{dy}{dx} + 2y = 0, \quad y(0) = 1, \quad y'(0) = -1. \quad (4)$$

4.5 THIS IS NOT A SUBSECTION, REMOVE THIS LINE LATER. Examples above are all from papers 1, 2, 3, and any new reference we may add After each examples we must also place the related tables, then the graphs

5 Conclusion

This study demonstrates the efficacy of KAN in solving ordinary differential equations, emphasizing its computational efficiency and accuracy. The findings establish KAN as a robust alternative to traditional neural network architectures for function approximation.(graft later)

Acknowledgements(if any)

6 REFERENCES(ADD AFTER MAKING THE DECISION FOR THE FUTURE WORKS AND WHICH WORKS AND SUBTYPES TO INCLUDE)