KANs: Kolmogorov-Arnold Networks

Atou Koffi KOUGBANHOUN

AIMS Senegal-AMMI

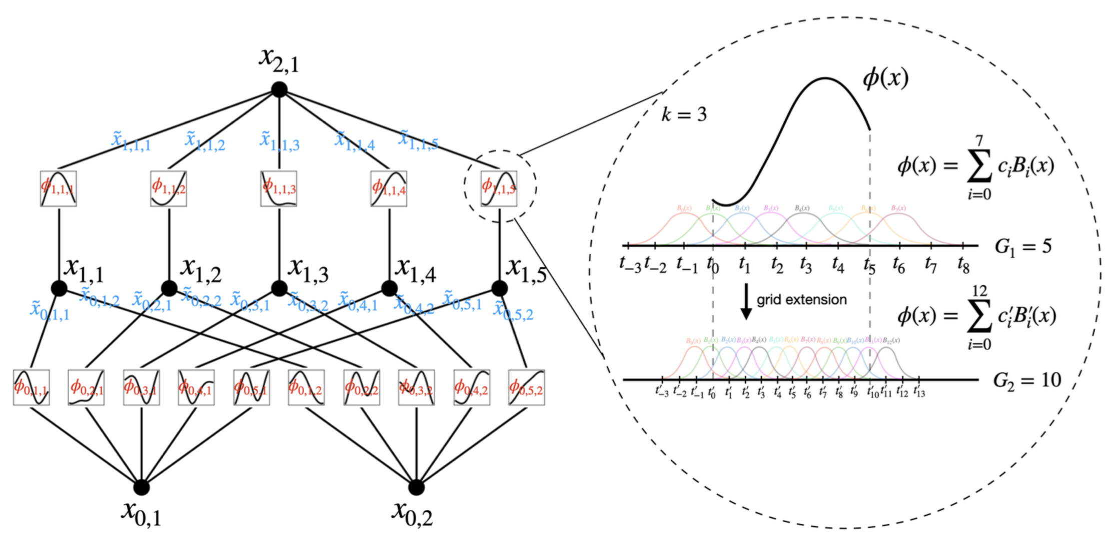
akouganhoun@aimsammi.org

1. Introduction

Kolmogorov-Arnold Networks (KANs) represent a novel neural network architecture inspired by the Kolmogorov-Arnold representation theorem. This theorem posits that any multivariable continuous function can be decomposed into a superposition of continuous univariate functions and addition. By leveraging this theoretical foundation, KANs can potentially achieve more efficient function approximation with fewer parameters compared to traditional neural networks. This poster delves into the architecture of KANs, highlighting their use of univariate continuous functions in hidden layers and the structured combination of these functions to produce accurate outputs. The advantages of KANs, such as improved efficiency, better generalization, and flexibility across various applications, are explored. Applications in scientific computing, finance, and engineering demonstrate the practical impact of KANs in modeling complex systems. By presenting both the theoretical and practical aspects, this poster aims to showcase the potential of KANs in advancing neural network design and function approximation.

2. Theoretical Background

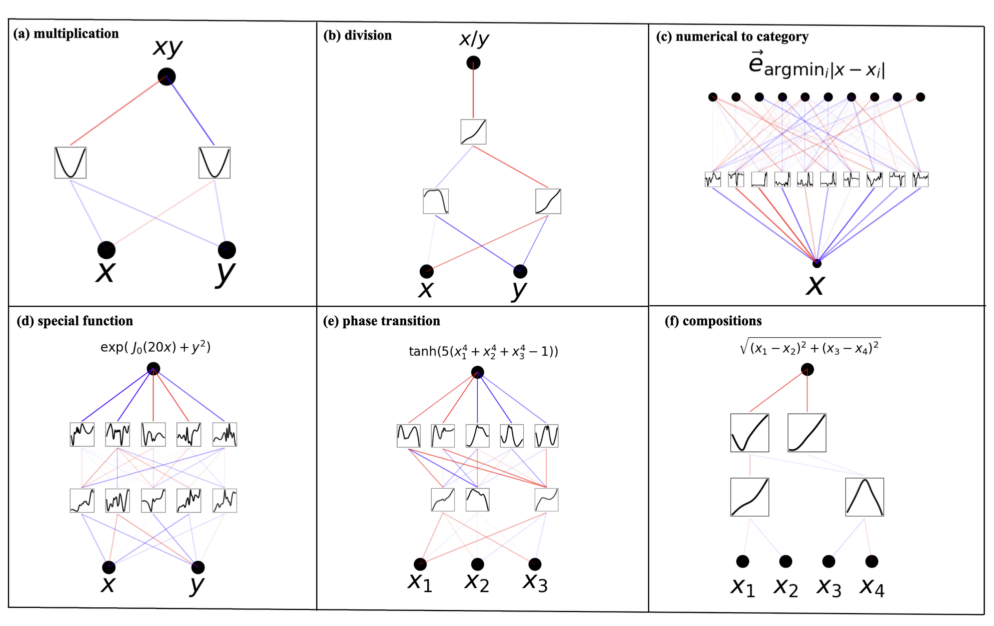
The Kolmogorov-Arnold Representation Theorem asserts that any continuous function of multiple variables can be represented using a finite number of continuous functions of a single variable. This implies that KANs can decompose complex multivariable functions into simpler, univariate functions, potentially leading to more efficient learning and generalization.



[1]

3. Architecture of KANs

- **Input Layer:** Receives the input variables.
- **Hidden Layers:** Utilize univariate continuous functions to transform the input.
- **Output Layer:** Combines the outputs of the hidden layers to produce the final result.
- **Activation Functions:** KANs typically employ activation functions that are continuous and can handle transformations required by the theorem.



[1]

4. Advantages of KANs

- **Efficiency:** Reduced complexity in representing multivariable functions can lead to fewer parameters and faster training times.
- **Generalization:** The theoretical foundation of KANs suggests better generalization to unseen data due to the structured decomposition of functions.
- **Flexibility:** KANs can be applied to various domains where continuous functions need to be approximated, including regression tasks and complex system modeling.

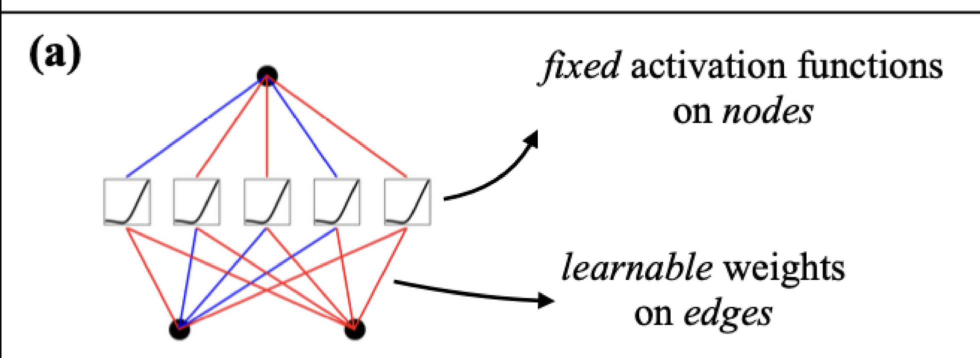
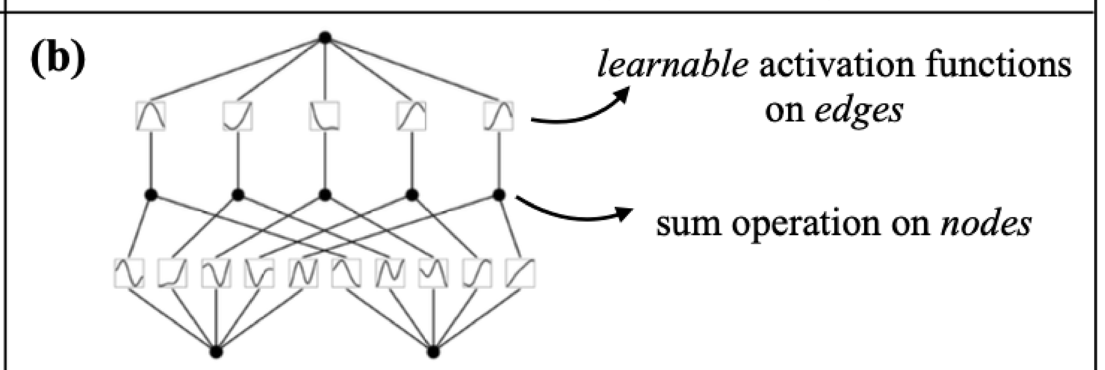
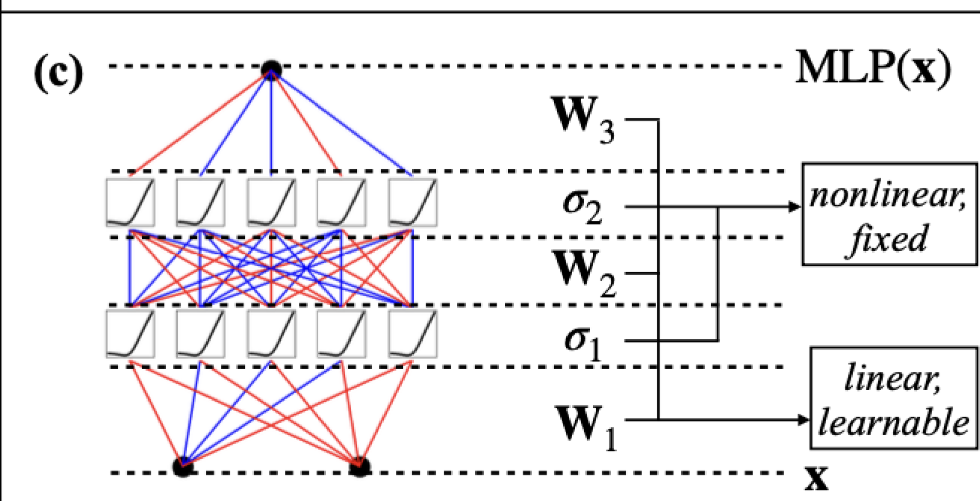
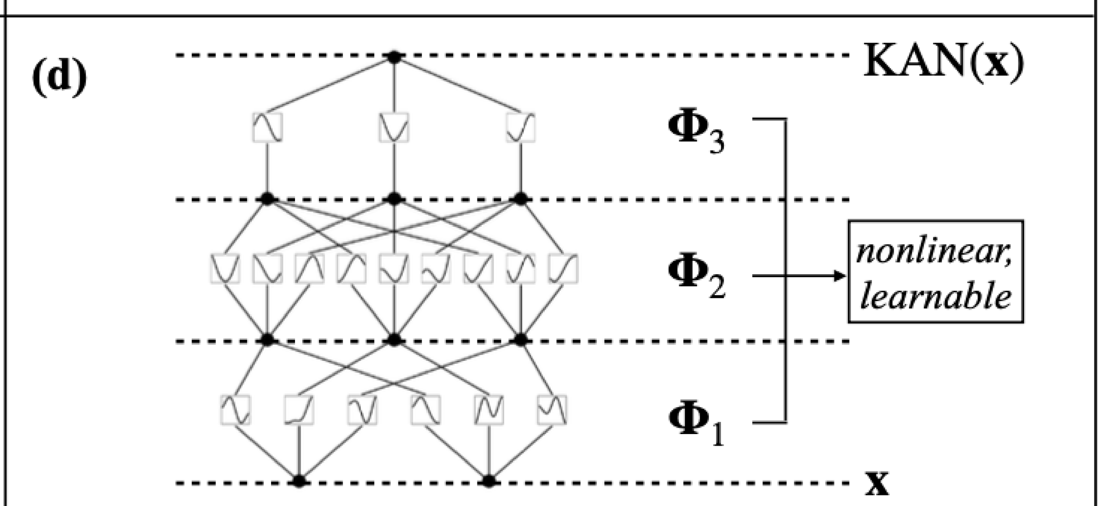
5. Applications of KANs

- **Scientific Computing:** Modeling complex physical systems where underlying functions are continuous.
- **Finance:** Predicting trends and modeling market behaviors based on continuous functions of multiple variables.
- **Engineering:** Control systems and signal processing where accurate function approximation is crucial.

6. Conclusion

Kolmogorov-Arnold Networks offer a promising approach to neural network architecture by leveraging a powerful mathematical theorem. Their efficiency and potential for better generalization make them a valuable tool in various fields requiring complex function approximation.

MLP vs KANs

Model	Multi-Layer Perceptron (MLP)	Kolmogorov-Arnold Network (KAN)
Theorem	Universal Approximation Theorem	Kolmogorov-Arnold Representation Theorem
Formula (Shallow)	$f(\mathbf{x}) \approx \sum_{i=1}^{N(\epsilon)} a_i \sigma(\mathbf{w}_i \cdot \mathbf{x} + b_i)$	$f(\mathbf{x}) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right)$
Model (Shallow)		
Formula (Deep)	$\mathrm{MLP}(\mathbf{x}) = (\mathbf{W}_3 \circ \sigma_2 \circ \mathbf{W}_2 \circ \sigma_1 \circ \mathbf{W}_1)(\mathbf{x})$	$\mathrm{KAN}(\mathbf{x}) = (\mathbf{\Phi}_3 \circ \mathbf{\Phi}_2 \circ \mathbf{\Phi}_1)(\mathbf{x})$
Model (Deep)		

7. References

[1] Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y. Hou, and Max Tegmark. Kan: Kolmogorov-arnold networks, 2024.