FAN: Fourier Analysis Networks

Yihong Dong[1*], **Ge Li**[1*], **Yongding Tao**[1], **Xue Jiang**[1], **Kechi Zhang**[1], **Jia Li ♂**[1], **Jing Su**[2], **Jun Zhang**[2], **Jingjing Xu**[2]
[1]School of Computer Science, Peking University [2]ByteDance
dongyh@stu.pku.edu.cn, lige@pku.edu.cn

Abstract

Despite the remarkable success achieved by neural networks, particularly those represented by MLP and Transformer, we reveal that they exhibit potential flaws in the modeling and reasoning of periodicity, i.e., they tend to memorize the periodic data rather than genuinely understanding the underlying principles of periodicity. However, periodicity is a crucial trait in various forms of reasoning and generalization, underpinning predictability across natural and engineered systems through recurring patterns in observations. In this paper, we propose FAN, a novel network architecture based on Fourier Analysis, which empowers the ability to efficiently model and reason about periodic phenomena. By introducing Fourier Series, the periodicity is naturally integrated into the structure and computational processes of the neural network, thus achieving a more accurate expression and prediction of periodic patterns. As a promising substitute to multi-layer perceptron (MLP), FAN can seamlessly replace MLP in various models with fewer parameters and FLOPs. Through extensive experiments, we demonstrate the effectiveness of FAN in modeling and reasoning about periodic functions, and the superiority and generalizability of FAN across a range of real-world tasks, including symbolic formula representation, time series forecasting, and language modeling.

1 Introduction

The flourishing of modern machine learning and artificial intelligence is inextricably linked to the revolutionary advancements in the foundational architecture of neural networks. For instance, multi-layer perceptron (MLP) (Rosenblatt, 1958; Haykin, 1998) plays a pivotal role in laying the groundwork for current deep learning models, with its expressive power guaranteed by the universal approximation theorem (Hornik et al., 1989). Recent claims about the impressive performance of large models on various tasks are typically supported by Transformer architecture (Vaswani et al., 2017; Touvron et al., 2023; OpenAI, 2023). In this context, the community's enthusiasm for research on neural networks has never diminished. Some emerged neural networks demonstrate notable capabilities in specific fields (Gu & Dao, 2023; Liu et al., 2024), sparking widespread discussion within the community.

Beneath the surface of apparent prosperity, we uncover a critical issue that remains in existing neural networks: *they struggle to model the periodicity from data*. We showcase this issue through an empirical study as illustrated in Figure 1. The results indicate that existing neural networks, including MLP (Rosenblatt, 1958), KAN (Liu et al., 2024), and Transformer (Vaswani et al., 2017), face difficulties in fitting periodic functions, even on a simple sine function. Although they demonstrate

[*]Equal Contribution

[†]This work was supported by a cooperation project between Peking University and ByteDance Company. During this time, Yihong was also an intern at ByteDance.

[‡]The code is available at `https://github.com/YihongDong/FAN`

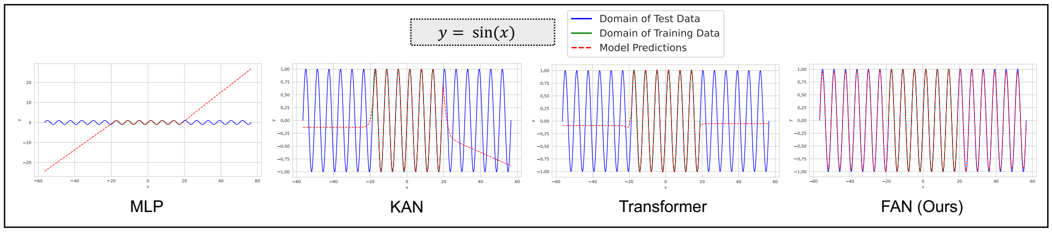

Figure 1: The performance of different neural networks within and outside the domain of their training data for the sine function, where x is a scalar variable.

proficiency in interpolation within the domain of training data, they tend to falter when faced with extrapolation challenges of test data, especially in the out-of-domain (OOD) scenarios. Therefore, their generalization capacity is primarily dictated by the scale and diversity of the training data, rather than by the learned principles of periodicity to perform reasoning. We argue that periodicity is an essential characteristic in various forms of reasoning and generalization, as it provides a basis for predictability in many natural and engineered systems by leveraging recurring patterns in observations.

In this paper, we investigate a key research problem: *How to enable neural networks to model periodicity?* One core reason existing neural networks fail to model periodicity is that they heavily rely on data-driven optimization without explicit mechanisms to understand the underlying principles in the data. To this end, we propose a Fourier Analysis Network (FAN), a novel neural network framework based on Fourier Analysis. By leveraging the power of Fourier Series, we explicitly encode periodic patterns within the neural network, offering a way to model the general principles from the data. FAN holds great potential as a substitute to traditional MLP, which not only exhibits exceptional capabilities in periodicity modeling but also demonstrates competitive or superior effects on general tasks.

To verify the effectiveness of FAN, we conduct extensive experiments from two main aspects: periodicity modeling and application of real-world tasks. 1) For periodicity modeling, FAN achieves significant improvements in fitting both basic and complex periodic functions, compared to existing neural networks (including MLP, KAN, and Transformer), particularly in OOD scenarios. 2) FAN demonstrates superior performance in real-world tasks, including symbolic formula representation, time series forecasting, and language modeling. The experimental results indicate that FAN outperform baselines (including MLP, KAN, and Transformer) for symbolic formula representation task, and Transformer with FAN surpasses the competing models (including Transformer, LSTM (Hochreiter & Schmidhuber, 1997), and Mamba (Gu & Dao, 2023)), for time series forecasting and language modeling tasks. As a promising substitute to MLP, FAN improves the model's generalization performance meanwhile reducing the number of parameters and floating point of operations (FLOPs) employed. We believe FAN is promising to be an important part of the fundamental model backbone.

2 Preliminary Knowledge

Fourier Analysis (Stein & Weiss, 1971; Duoandikoetxea, 2024) is a mathematical framework that decomposes functions into their constituent frequencies, revealing the underlying periodic structures within complex functions. At the heart of this analysis lies Fourier Series (Tolstov, 2012), which expresses a periodic function as an infinite sum of sine and cosine terms. Mathematically, for a function $f(x)$, its Fourier Series expansion can be represented as:

$$f(x) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2\pi n x}{T}\right) + b_n \sin\left(\frac{2\pi n x}{T}\right) \right), \tag{1}$$

where T is the period of the function, and the coefficients a_n and b_n are determined by integrating the function over one period: