Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI-25)
Doctoral Consortium Track

Toward Interpretable Time Series Modeling: A Kernel Representation Perspective

Kunpeng Xu
University of Sherbrooke, Sherbrooke, Quebec, Canada
kunpeng.xu@usherbrooke.ca

Abstract

Time series modeling is essential in finance, healthcare, and environmental science, yet nonlinear patterns, noise, and concept drift pose challenges. Although deep learning models, such as Transformer-based and recent pre-trained models, have achieved good performance across various time series tasks, they often lack interpretability, especially in co-evolving time series. This work introduces a kernel representation learning (KRL) perspective, rethinking time series modeling through kernel-induced self-representation to effectively capture temporal structures and dynamic transitions. Additionally, we establish theoretical connections between KRL and advanced deep-network models, demonstrating how kernel methods provide a principled approach to capturing complex time series behaviors.

1 Introduction

Time series analysis is crucial in fields like finance, healthcare, and meteorology. The complexities of time series data – such as nonlinearity, noise, and concept drift – make accurate modeling and prediction challenging [Chen *et al.*, 2022a; Chen *et al.*, 2022b; **?**]. Recent deep learning advances, particularly Transformer-based architectures, have improved forecasting by capturing long-range dependencies. However, most models focus on prediction accuracy at the expense of interpretability. This lack of interpretability is a concern in many fields. In finance, for instance, shifts in market regimes and nonlinear relationships are as important as prediction accuracy. Detecting and understanding these shifts is vital for adapting to changing market conditions. Similarly, in healthcare, understanding a model's reasoning is crucial, as clinical decisions can significantly impact patient outcomes.

An additional challenge arises with co-evolving time series [Xu *et al.*, 2024c], where multiple time series exhibit interdependent behavior over time. In such cases, understanding the interactions and dynamics among these series becomes crucial for both prediction and interpretation. A more holistic approach is required – one that not only forecasts future values but also explains the underlying behaviors and relationships that govern the evolution of these time series [Xu, 2025]. Accordingly, we explore the following research question:

> **RQ** – *Can we model co-evolving or large-scale time series as an* ***ecosystem*** 🌍 *and track its evolution to balance predictive power with interpretability?*

2 Our Contributions

Rethinking time series modeling from a kernel perspective, we introduce an innovative approach to capture complex, nonlinear relationships within data. Our goal is to construct a robust representation that not only enhances predictive accuracy but also offers deeper interpretability. In the subsections that follow, we present our key contributions:

2.1 KRL for Time Series Modeling

Kernel-induced Representation. We introduce CORAL [Xu *et al.*, 2025a], a method designed to address the dynamic complexities of co-evolving time series. By leveraging kernel-induced self-representation, CORAL captures intricate interdependencies and adapts to concept drift through a representation matrix transformation of the data. This paradigm offers several advantages: (1) a novel approach for modeling complex interdependencies that integrates easily with deep learning frameworks, (2) adaptive detection and response to concept/distribution drift without prior assumptions, and (3) a transformation of time series concept/distribution drift into a matrix optimization problem that enhances interpretability.

Domain-agnostic Twin Learning Framework. We propose a unified twin learning framework [Xu *et al.*, 2024a] for domain-agnostic time series analysis that simultaneously identifies behavioral patterns and optimal segmentation boundaries. Our KRL-based approach determines pattern durations and captures dynamic switches in non-linear time series data. Leveraging the learned kernel representation, we introduce a heuristic segmentation method based on a segment-score function that quantifies the *behavior-consistency* of the time series data. Optimizing this score enables our method to adaptively identify segmentation boundaries that preserve and elucidate the core behaviors in the time series.

Adaptive Kernel Learning. We propose a data-driven kernel subspace learning method [Xu *et al.*, 2022a; Xu *et al.*, 2025b] that adaptively learns kernel mappings from self-representation, eliminating the need for predefined kernel functions. The approach preserves the intrinsic manifold structure of the time series while promoting a block-diagonal structure in the representation matrix. Under certain conditions, the representation behaves as adaptive-weighted kernel learning and satisfies the triangular inequality, providing a theoretical lower bound and enhancing learning efficiency.

2.2 Theoretical Connection

We explore how kernel representation learning (KRL) offers insights into modern deep learning models, improving both interpretability and theoretical understanding.

Transformer-based models. The self-attention mechanism in Transformers can be viewed as a kernel operation, where the Query-Key dot product acts as a kernel mapping inputs into a high-dimensional space. This perspective clarifies attention dynamics and links Transformers to kernel methods.

Kolmogorov-Arnold Networks. We further show that KAN's learnable activations and nonlinear mappings align with kernel-induced self-representation under certain settings [Xu *et al.*, 2024e]. Building on this, our WormKAN model [Xu *et al.*, 2024d] leverages KAN to detect concept drift in time series, drawing inspiration from physics (e.g., wormholes) to model dynamic transitions. This bridges kernel theory and interpretable deep learning.

2.3 Real-world Applications

Financial Market. We tackle the task of discovering and forecasting regime switches in financial markets, which reflect changing market behaviors over time. To this end, we propose RHINE—a kernel-based model that captures nonlinear interactions and dynamic shifts [Xu *et al.*, 2024b]. It also employs eigengap thresholding to automatically infer the number of regimes, enhancing adaptability.

DNA/RNA Sequences. We propose a scalable multi-view kernel learning framework for categorical sequences (e.g., DNA/RNA) [Xu *et al.*, 2022b]. Each view is encoded as a kernel matrix, and instance weights are jointly learned with data partitioning. The model adaptively constructs kernels without relying on predefined functions and employs a kernel approximation to reduce the $O(N^2)$ space complexity.

3 Ongoing and Future Works

We will continue refining the theoretical connections between KRL and deep models. We are also developing a kernel-based diffusion model to generate a public synthetic time series dataset—preserving privacy while adapting to distribution shifts. To support the community, we will launch an open-access website hosting this dataset to facilitate research and collaboration in time series representation learning.

References

[Chen *et al.*, 2022a] Rongbo Chen, Mingxuan Sun, Kunpeng Xu, Jean-Marc Patenaude, and Shengrui Wang. Clustering-based cross-sectional regime identification for financial market forecasting. In *International Conference on Database and Expert Systems Applications*. Springer, 2022.

[Chen *et al.*, 2022b] Rongbo Chen, Kunpeng Xun, Jean-Marc Patenaude, and Shengrui Wang. Dynamic cross-sectional regime identification for financial market prediction. In *2022 IEEE 46th Annual Computers, Software, and Applications Conference*. IEEE, 2022.

[Xu *et al.*, 2022a] Kunpeng Xu, Lifei Chen, and Shengrui Wang. Data-driven kernel subspace clustering with local manifold preservation. In *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 876–884. IEEE, 2022.

[Xu *et al.*, 2022b] Kunpeng Xu, Lifei Chen, and Shengrui Wang. A multi-view kernel clustering framework for categorical sequences. *Expert Systems with Applications*, 197:116637, 2022.

[Xu *et al.*, 2024a] Kunpeng Xu, Lifei Chen, Jean-Marc Patenaude, and Shengrui Wang. Kernel representation learning with dynamic regime discovery for time series forecasting. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2024.

[Xu *et al.*, 2024b] Kunpeng Xu, Lifei Chen, Jean-Marc Patenaude, and Shengrui Wang. Rhine: A regime-switching model with nonlinear representation for discovering and forecasting regimes in financial markets. In *Proceedings of the 2024 SIAM International Conference on Data Mining (SDM)*, pages 526–534. SIAM, 2024.

[Xu *et al.*, 2024c] Kunpeng Xu, Lifei Chen, and Shengrui Wang. Are kan effective for identifying and tracking concept drift in time series? *arXiv preprint arXiv:2410.10041*, 2024.

[Xu *et al.*, 2024d] Kunpeng Xu, Lifei Chen, and Shengrui Wang. Kan4drift: Are kan effective for identifying and tracking concept drift in time series? In *NeurIPS Workshop on Time Series in the Age of Large Models*, 2024.

[Xu *et al.*, 2024e] Kunpeng Xu, Lifei Chen, and Shengrui Wang. Kolmogorov-arnold networks for time series: Bridging predictive power and interpretability. *arXiv preprint arXiv:2406.02496*, 2024.

[Xu *et al.*, 2025a] Kunpeng Xu, Lifei Chen, and Shengrui Wang. Drift2matrix: Kernel-induced self representation for concept drift adaptation in co-evolving time series. *arXiv preprint arXiv:2501.01480*, 2025.

[Xu *et al.*, 2025b] Kunpeng Xu, Lifei Chen, and Shengrui Wang. Towards robust nonlinear subspace clustering: A kernel learning approach. *arXiv preprint arXiv:2501.06368*, 2025.

[Xu, 2025] Kunpeng Xu. Kernel representation learning for time sequence: Algorithm, theory, and applications. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 29309–29310, 2025.