

## Topological Data Analysis (TDA)-based Feature Engineering for Reinforcement Learning-based Trading Strategies in Financial Markets

Luiz de Jesus<sup>1</sup>, Francisco Fernandez-Navarro<sup>2</sup>, Mariano Carbonero-Ruz<sup>3</sup>

<sup>1,3</sup> *Department of Quantitative Methods, Universidad Loyola, Córdoba, Spain*

<sup>2</sup> *Department of Computer Languages and Computer Science, Universidad de Málaga, Málaga, Spain*

### Abstract

This research investigates a novel approach to feature engineering for reinforcement learning (RL)-based trading strategies by leveraging Topological Data Analysis (TDA). As machine learning techniques increasingly transform financial trading strategies, RL has emerged as a robust methodology for developing automated trading systems. However, the effectiveness of these systems heavily depends on the quality of features extracted from financial market data.

TDA, situated at the intersection of mathematics and data science, applies algebraic topology to analyse complex datasets and has demonstrated significant potential in uncovering hidden structures within high-dimensional data. Financial markets, characterised by complex interactions, non-linear relationships, and high dimensionality, present an ideal domain for TDA application. The topological features extracted through TDA—such as persistence diagrams, bar codes, and mapper graphs—can reveal market regimes, identify stability or turbulence periods, and capture intricate relationships between financial instruments.

Our research addresses three primary questions:

1. What are the most relevant topological features of financial market data for TDA-based feature engineering?
2. Can TDA-based feature engineering improve the performance of RL-based trading strategies?
3. How can TDA-based feature engineering be optimally combined with other feature engineering techniques?

The methodology follows five systematic steps:

1. **Data Collection:** Gathering historical financial market data from public sources, including price and volume information and economic indicators spanning multiple years to capture various market regimes.
2. **Topological Data Analysis:** Applying TDA techniques to analyse topological features, with particular attention to persistence (indicating market regime stability), non-linearity (capturing complex market dynamics), and connectivity (revealing relationships between financial instruments).
3. **Feature Engineering:** Extracting TDA-based features and combining them with traditional features such as technical indicators, fundamental factors, and market sentiment measures.
4. **Reinforcement Learning:** Employing various RL algorithms (DQN, Policy Gradient methods, PPO, SAC) to learn trading strategies based on the engineered feature set.

5. Performance Evaluation: Comparing TDA-enhanced strategies against conventional approaches using metrics such as returns, Sharpe ratio, maximum drawdown, and performance consistency across different market conditions.

The expected outcomes include identifying relevant topological features for trading strategies, demonstrating the benefits of incorporating these features, and gaining insights into synergistic combinations of TDA-based and traditional feature engineering techniques.

This research intersects three cutting-edge fields—topological data analysis, reinforcement learning, and quantitative finance—potentially reshaping financial data analysis and algorithmic trading approaches. By capturing the underlying topological structure of financial markets, we aim to provide RL algorithms with more informative representations of market states, leading to more robust and effective trading decisions while contributing to a broader understanding of financial market dynamics through the lens of topology.

**Keywords:** Topological Data Analysis, Reinforcement Learning, Trading Strategies

## References

- [1] L. A. Leaverton, 'Analysis of Financial Time Series using TDA: Theoretical and Empirical Results'.
- [2] M. Gidea and Y. Katz, 'Topological Data Analysis of Financial Time Series: Landscapes of Crashes', *Physica A: Statistical Mechanics and its Applications*, vol. 491, pp. 820–834, Feb. 2018, doi: 10.1016/j.physa.2017.09.028.
- [3] N. Ravishanker and R. Chen, 'Topological Data Analysis (TDA) for Time Series'. arXiv, Sep. 23, 2019. Accessed: May 26, 2023. [Online]. Available: <http://arxiv.org/abs/1909.10604>
- [4] P. T.-W. Yen and S. A. Cheong, 'Using Topological Data Analysis (TDA) and Persistent Homology to Analyze the Stock Markets in Singapore and Taiwan', *Front. Phys.*, vol. 9, p. 572216, Mar. 2021, doi: 10.3389/fphy.2021.572216.
- [5] S. W. Akingbade, M. Gidea, M. Manzi, and V. Nateghi, 'Why Topological Data Analysis Detects Financial Bubbles?' arXiv, Apr. 13, 2023. Accessed: May 26, 2023. [Online]. Available: <http://arxiv.org/abs/2304.06877>

Corresponding Author Email: [luiscarlos@al.u Loyola.es](mailto:luiscarlos@al.u Loyola.es)