REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

ANÁLISE TOPOLÓGICA DE SÉRIES TEMPORAIS PARA PREVISÃO EM PEQUENA ESCALA

TOPOLOGICAL ANALYSIS OF TIME SERIES FOR SMALL-SCALE FORECASTING

ANÁLISIS TOPOLÓGICO DE SERIES TEMPORALES PARA PRONÓSTICOS A PEQUEÑA ESCALA

Jerson Leite Alves
Doutor, Instituto Federal do Piauí, Brasil
E-mail: jerson@ifpi.edu.br

Paulo Victor Silva Sousa
Tecnólogo, Instituto Federal do Piauí, Brasil
E-mail: paulovictormsn@gmail.com

Rui Marques Carvalho
Doutor, Instituto Federal do Piauí, Brasil
E-mail: rui.marques@ifpi.edu.br

Reneé Rodrigues Lima
Doutor, Instituto Federal do Piauí, Brasil
E-mail: renee@ifpi.edu.br

Atecio Alves
Doutor, Instituto Federal do Piauí, Brasil
E-mail: atecio.alves@ifpi.edu.br

Lucas Pereira Viana
Mestre, Instituto Federal do Maranhão, Brasil
E-mail: lucas.viana@ifma.edu.br

Antônio Aguiar Freitas
Doutor, Universidade Estadual Vale do Acaraú, Ceará
E-mail: aguiar_antonio@uvanet.br

Ana Gabriela Lima Pacífico
Mestra, Instituto Federal do Piauí, Brasil
E-mail: anagabrielalimap@gmail.com

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

Resumo

A previsão de séries temporais constitui uma área de grande relevância em diferentes setores, porém métodos estatísticos tradicionais e modelos de aprendizado profundo podem apresentar limitações diante de séries complexas, não lineares e sujeitas a ruídos. Nesse contexto, este trabalho investiga a aplicação da Análise Topológica de Dados (TDA) como recurso complementar à previsão de séries temporais em pequena escala. Foram analisadas três séries temporais sintéticas que simulam situações reais, referentes ao período de 2021 a 2025: índice de mercado financeiro, temperatura máxima diária de uma cidade e consumo diário de energia de uma cidade. A metodologia envolveu o pré-processamento dos dados, a reconstrução das séries por meio da incorporação de Takens, a filtração por complexos de Vietoris–Rips e a extração de descritores a partir de diagramas de persistência. Esses atributos foram incorporados aos modelos ARIMA e LSTM, permitindo comparar seu desempenho com configurações baseadas apenas nos valores originais das séries. Os resultados, avaliados por meio das métricas MAE e RMSE, indicaram melhora no desempenho preditivo com a inclusão dos atributos topológicos em diferentes séries e modelos. Os resultados evidenciam o potencial da TDA para capturar estruturas e padrões da dinâmica temporal que podem não ser facilmente identificados por abordagens tradicionais, contribuindo para modelos de previsão mais robustos em cenários de pequena escala.
Palavras-chave: Análise Topológica de Dados; Séries Temporais; Previsão de Dados; ARIMA; LSTM.

Abstract

Time series forecasting is a highly relevant field across various sectors; however, traditional statistical methods and deep learning models can face limitations when dealing with complex, non-linear, and noisy time series. In this context, this study investigates the application of Topological Data Analysis (TDA) as a complementary tool for small-scale time series forecasting. Three synthetic time series simulating real-world scenarios for the 2021–2025 period were analyzed: a financial market index, a city's daily maximum temperature, and a city's daily energy consumption. The methodology involved data preprocessing, series reconstruction via Takens' embedding, filtration using Vietoris–Rips complexes, and the extraction of descriptors from persistence diagrams. These attributes were incorporated into ARIMA and LSTM models, allowing for a performance comparison against configurations based solely on the original series values. The results, evaluated using MAE and RMSE metrics, indicated improved predictive performance when topological attributes were included across the different series and models. These findings highlight TDA's potential to capture temporal dynamic structures and patterns that may not be easily identified by traditional approaches, thereby contributing to more robust forecasting models in small-scale scenarios.
Keywords: Topological Data Analysis; Time Series; Data Forecasting; ARIMA; LSTM.

Resumen

La predicción de series temporales es un área de gran relevancia en diversos sectores; sin embargo, los métodos estadísticos tradicionales y los modelos de aprendizaje profundo pueden presentar limitaciones al tratar con series complejas, no lineales y propensas al ruido. En este contexto, este trabajo investiga la aplicación del Análisis Topológico de Datos (ATD) como recurso complementario para la predicción de series temporales a pequeña escala. Se analizaron tres series temporales sintéticas que simulan situaciones reales, abarcando el período de 2021 a 2025: un índice del mercado financiero, la temperatura máxima diaria de una ciudad y el consumo energético diario de una ciudad. La metodología incluyó el preprocesamiento de datos, la reconstrucción de series mediante la incorporación de Takens, el filtrado utilizando complejos de Vietoris-Rips y la extracción de descriptores de diagramas de persistencia. Estos atributos se incorporaron a los modelos ARIMA y LSTM, lo que permitió comparar su rendimiento con configuraciones basadas únicamente en los valores originales de las series. Los resultados, evaluados mediante las métricas MAE y RMSE, indicaron una mejora en el rendimiento predictivo con la inclusión de atributos topológicos en diferentes series y modelos. Los resultados ponen de relieve el potencial del análisis de dinámica temporal (TDA) para capturar estructuras y patrones en la

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

dinámica temporal que no se identifican fácilmente con los enfoques tradicionales, lo que contribuye a modelos de pronóstico más robustos en escenarios a pequeña escala.
Palabras clave: Análisis topológico de datos; Series temporales; Pronóstico de datos; ARIMA; LSTM.

1. Introdução

A previsão de séries temporais é uma área de grande relevância científica e prática, com aplicações em setores como economia, meteorologia, saúde e engenharia. Tradicionalmente, modelos estatísticos, como ARIMA e GARCH, têm sido amplamente utilizados devido à sua formulação matemática sólida. Entretanto, tais métodos apresentam limitações quando aplicados a séries complexas, não lineares ou sujeitas a ruídos significativos.

Nos últimos anos, o avanço do aprendizado de máquina e, em especial, das redes neurais profundas, ampliou as possibilidades de modelagem de séries temporais. Arquiteturas como LSTM e Transformers demonstraram elevada capacidade de capturar dependências temporais de curto e longo prazo. Contudo, essas abordagens priorizam aspectos estatísticos e sequenciais, deixando de explorar a estrutura geométrica e topológica intrínseca aos dados.

Nesse contexto, a Análise Topológica de Dados (TDA) surge como uma ferramenta inovadora ao fornecer descritores robustos, invariantes a ruído e capazes de revelar padrões ocultos em séries temporais. A integração entre TDA e aprendizado profundo representa, portanto, uma abordagem promissora para melhorar a precisão de modelos preditivos, especialmente em cenários de pequena escala.

2. Revisão da Literatura

A previsão de séries temporais constitui um dos problemas centrais da ciência de dados, com aplicações que abrangem desde economia e meteorologia até biomedicina e engenharia (SINGH, 2025). Métodos estatísticos clássicos, como modelos autorregressivos (AR), médias móveis (MA) e suas extensões mais sofisticadas, como ARIMA e GARCH, consolidaram-se ao longo das últimas décadas como ferramentas de referência para modelagem

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

temporal (RUBIO et al., 2023). Apesar de sua relevância, tais abordagens assumem premissas de linearidade e estacionariedade, o que limita sua eficácia em sistemas dinâmicos complexos.

Com os avanços da inteligência artificial, modelos de aprendizado profundo passaram a ocupar papel de destaque na modelagem temporal. Redes neurais recorrentes (RNN), suas variantes LSTM e GRU (NGUYEN, 2025) e arquiteturas baseadas em atenção, como os Transformers (VASWANI et al., 2017), demonstraram elevado poder de generalização e capacidade de capturar dependências de longo prazo (ORESHKIN et al., 2019). Estudos recentes têm mostrado que tais modelos frequentemente superam os métodos estatísticos tradicionais em termos de precisão preditiva, especialmente em séries não lineares e de alta variabilidade (LIU, 2025).

Entretanto, essas abordagens permanecem fortemente ancoradas na estrutura estatística dos dados e exploram de forma limitada a geometria e topologia subjacentes às séries temporais. Nesse contexto, a Análise Topológica de Dados (TDA) surge como alternativa promissora (GOEL et al., 2020). Fundamentada em conceitos de topologia algébrica, a TDA permite extrair atributos robustos, invariantes a ruídos e capazes de capturar propriedades estruturais globais dos dados. Técnicas como a incorporação de Takens (TAKENS, 1981), a construção de complexos de Vietoris–Rips e o cálculo da homologia persistente produzem descritores que refletem a conectividade, ciclos e cavidades presentes na evolução temporal dos sistemas (BUBENIK, 2020).

Trabalhos recentes confirmam a aplicabilidade da TDA em diferentes domínios. (ALVES, 2025) mostram que atributos derivados de diagramas de persistência, quando integrados ao modelo N-BEATS, elevam significativamente a precisão preditiva em séries financeiras. RAI et al. (2024) utilizam TDA para identificar eventos extremos em mercados de ações, enquanto HAN et al. (2024) destacam sua eficácia na previsão de carga elétrica. Em outra direção, SHEKHAR et al. (2024) aplicam TDA à ciência dos materiais, demonstrando sua versatilidade para caracterização de estruturas complexas.

A combinação entre TDA e aprendizado profundo, como discutido em ALVES (2025), representa um caminho inovador e de grande potencial para a modelagem

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

temporal. Ao capturar atributos topológicos e integrá-los a arquiteturas neurais, é possível construir modelos mais robustos, menos sensíveis a ruídos e capazes de generalizar para diferentes domínios. Essa literatura fundamenta a presente proposta, que busca explorar, em pequena escala, como descritores topológicos podem melhorar previsões quando adicionados a modelos estatísticos como ARIMA e de deep learning como LSTM.

3. Metodologia

O desenvolvimento do trabalho está estruturado em quatro etapas principais, que se articulam de forma progressiva. Inicialmente, foi realizada a geração de séries temporais sintéticas a partir de funções que simulam situações reais, em seguida a coleta de dados, garantindo a natureza prática e contextualizada. Esses dados passaram por um processo de pré-processamento, que inclui a normalização, o tratamento de valores ausentes e a segmentação em janelas temporais adequadas para análise.

As séries foram construídas para reproduzir características típicas observadas em dados reais, preservando sazonalidade, tendência e ruído aleatório. Para facilitar o entendimento, denominou-se TS a série temporal que descreva a temperatura, ES a que descreve o consumo de energia e FS a do índice de mercado financeiro, e ainda, será denotado com “+TDA" os modelos que tiverem a implementação das características topológicas no vetor concatenado.

A temperatura (1) apresenta comportamento sazonal anual, pequena tendência de aquecimento e variabilidade diária, com média = 29 °C, amplitude = 3 °C, tendência = 0,0008 °C/dia e ε_T~N(0,0,8).

$$\mathrm{T}(t) = 29 + 3sen\left(\frac{2\pi t}{365}\right) + 0{,}0008t + \varepsilon_T \quad (1)$$

O consumo de energia (2) apresenta tendência crescente, sazonalidade anual e oscilações semanais, com média = 850 MWh, crescimento gradual, sazonalidade anual, sazonalidade semanal e ruído normal.

$$\mathrm{E(t)} = 850 + 0{,}12t + 45sen\left(\frac{2\pi t}{365}\right) + 18sen(72\pi t) + \varepsilon_E \quad (2)$$

A série financeira (3) foi simulada com crescimento exponencial moderado e oscilações típicas do mercado, com valor inicial ≈ 100.000 pontos,

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

crescimento médio ao longo do período, ciclos de mercado e elevada volatilidade.

$$\mathrm{F}(t) = 100000e^{0,00018t} + 3500sen\left(\frac{2\pi t}{180}\right) + \varepsilon_F \quad (3)$$

Na sequência, foi implementada a extração de atributos topológicos por meio de técnicas da Análise Topológica de Dados (TDA) em toda a série. Para tanto, as séries temporais foram reconstruídas no espaço de fase utilizando a incorporação de Takens, que transforma a série em uma nuvem de pontos multidimensional. Sobre essa representação, foram construídos os complexos de Vietoris–Rips e calculados os diagramas de persistência, dos quais foram obtidas medidas topológicas como ciclos, conectividade e persistência.

Esses descritores foram posteriormente vetorizados em formatos compatíveis com algoritmos de aprendizado de máquina, como paisagens e imagens de persistência.

Concluída a etapa de extração de atributos, o tanto a série temporal quanto os dados topológicos extraídos foram concatenados em um único vetor e aplicado nos dois modelos preditivos distintos.

O primeiro foi o ARIMA, e a configuração adotada foi ARIMA(2,1,2), que serve como referência estatística clássica e pode ser visto em Figura 1 (a). O segundo consiste em uma rede neural recorrente do tipo LSTM, escolhida por sua capacidade de lidar com dependências temporais de curto e longo prazo, com hiperparâmetros: 1 camada LSTM, 60 neurônios, janela temporal de 30 dias, função de ativação *tanh*, otimizador Adam e 100 épocas, e pode ser visto em Figura 1 (b).

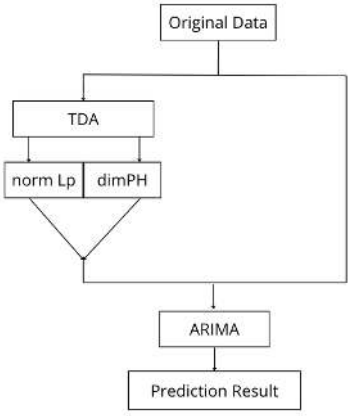
Em ambos os casos, os modelos são treinados em duas configurações: uma considerando apenas os valores originais das séries e outra incorporando os descritores topológicos extraídos pela TDA.

CC BY

REMUNOM
ISSN 2178-6925

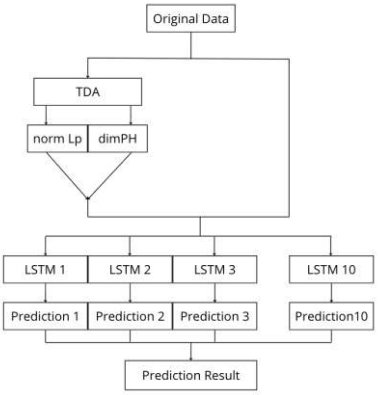
Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

Figura 1 (a).



Fonte: autor

Figura 1 (b)



Fonte: autor

Por característica o modelo que utiliza LSTM é repetido 10 vezes, e coletado a mediana da previsão para diminuir ruídos.

Por fim, os resultados obtidos são avaliados por meio de métricas quantitativas amplamente utilizadas na literatura, como o Erro Médio Absoluto (MAE) e a Raiz do Erro Quadrático Médio (RMSE), de modo a comparar o impacto da inclusão dos atributos topológicos no desempenho dos modelos. A análise será complementada com representações gráficas das previsões em relação aos dados observados, permitindo uma interpretação visual do efeito da TDA no processo de previsão.

4. Resultados e Discussão

Esta seção apresenta os resultados obtidos a partir da aplicação dos modelos ARIMA e LSTM desenvolvidos para a tarefa de previsão de séries temporais. As análises foram conduzidas com conjuntos de dados sintéticos, de 2021 a 2025, que representam dados reais de: Série financeira pública (IBOVESPA – pontos), Temperatura máxima diária (°C) e Consumo diário de energia elétrica (MWh) permitindo uma avaliação comparativa da performance dos modelos.

Os experimentos buscaram evidenciar a eficácia de arquiteturas amplamente utilizadas e o impacto da incorporação de medidas topológicas extraídas por meio da TDA no desempenho preditivo. Os resultados estão organizados de forma a destacar os aspectos quantitativos das previsões, além de

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

possibilitar uma interpretação crítica sobre as capacidades e limitações dos modelos considerados.

Foram analisados 1.826 dias, e os períodos de treino e teste dos modelos foi dividido na proporção 2021 a 2024 e 2025, respectivamente, ou seja, 1.460 dias para treinamento, 366 dias para teste. Os modelos baseados em LSTM foi aplicado por 10 vezes e tomado a mediana dos resultados como medida de centralidade, já os modelos baseados em ARIMA uma única vez e os resultados de previsão nos termos das métricas MAE e RMSE estão na Tabela 1.

Tabela 1 – Avaliação do Resultado dos modelos ARIMA e LSTM.

Série / Modelo	ARIMA		LSTM		Melhor desempenho
	MAE	RMSE	MAE	RMSE	
TS	0,79	1,32	**0,77**	**1,14**	LSTM
TS+TDA	0,65	1,09	**0,65**	**0,99**	LSTM
ES	0,78	1,29	**0,63**	**1,21**	LSTM
ES+TDA	0,66	1,09	**0,55**	**1,05**	LSTM
FS	**0,91**	**1,29**	0,99	1,31	ARIMA
FS+TDA	**0,73**	**1,06**	0,91	1,12	ARIMA

Fonte: Autor

Em seguida, pode-se ver nas figuras 1 a 3, onde é mostrado o comportamento das séries temporais em estudo e suas respectivas estruturas geométricas e topológicas associadas. Na parte de cima em cada figura está ilustrado o Complexo de Vietoris-Rips com filtração por raio ε, e na parte de baixo o gráfico da série temporal descrevendo o comportamento temporal da série ao longo do período estudado.

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

Figura 1: Complexo de Vietoris-Rips e Gráfico do IBOVESPA



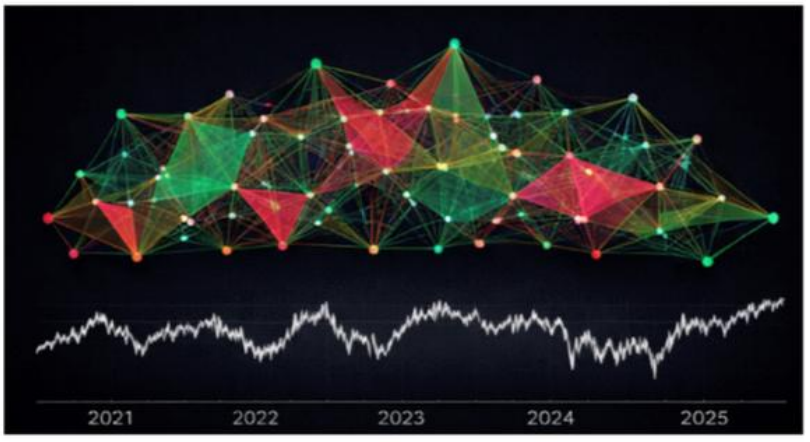
Fonte: Autor

Figura 2: Complexo de Vietoris-Rips e Gráfico da temperatura máxima diária



Fonte: Autor

Figura 3: Complexo de Vietoris-Rips e Gráfico do consumo diário de energia



Fonte: Autor

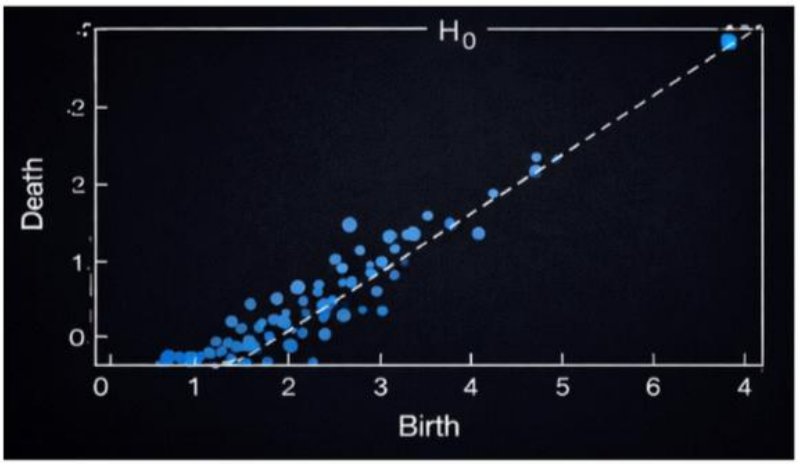
Estão ilustrados nas figuras 4 a 6 temos os Diagramas de Persistência das séries temporais geradas a partir dos momentos de nascimento e morte de cada estrutura topológica dentro dos Complexos de Vietoris-Rips.

CC BY

REMUNOM
ISSN 2178-6925

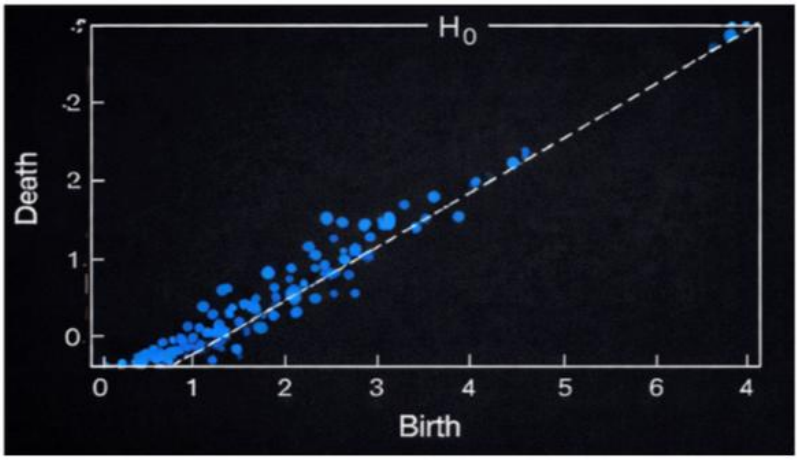
Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

Figura 4: Diagramas de Persistência do IBOVESPA



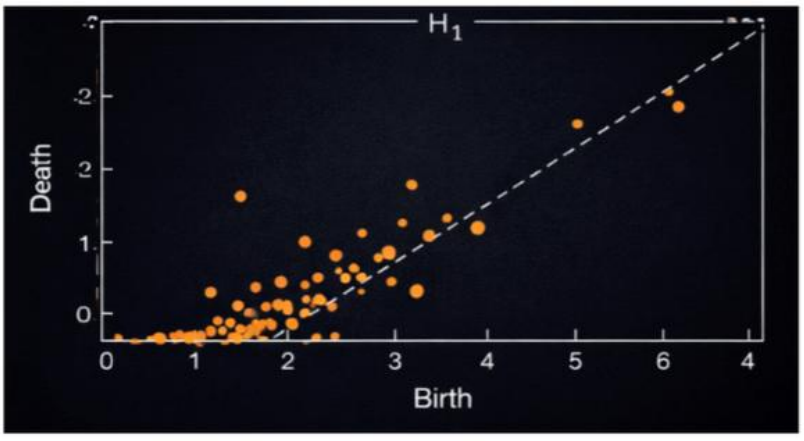
Fonte: Autor

Figura 5: Diagramas de Persistência da temperatura máxima diária



Fonte: Autor

Figura 6: Diagramas de Persistência do consumo diário de energia



Fonte: Autor

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

5. Conclusão

Os experimentos indicam que ambos os modelos foram capazes de reproduzir a dinâmica geral da série temporal sintética. O ARIMA apresentou desempenho satisfatório na modelagem da tendência e da sazonalidade, porém teve limitações para representar mudanças abruptas induzidas pela volatilidade da série financeira. Por outro lado, o modelo LSTM apresentou menor erro de previsão, evidenciando maior capacidade de aprender relações não lineares entre as três séries consideradas.

O comportamento observado é consistente com a literatura recente e demonstra que a infraestrutura computacional necessária para os experimentos do estudo foi implementada com sucesso.

Os resultados evidenciam o potencial de evolução da abordagem combinada de janelas temporais, incorporação por atraso e análise topológica. A principal contribuição reside na capacidade de capturar propriedades qualitativas da dinâmica que não são facilmente detectáveis por métodos tradicionais. A presença de estruturas topológicas persistentes sugere que as séries possuem organização interna significativa, mesmo em contextos aparentemente ruidosos, como o mercado financeiro.

Referências

ALVES, J. L. Modelos de aprendizado profundo guiados por Análise Topológica de Dados para previsão de séries temporais / Jerson Leite Alves. – Campinas, SP : [s.n.], 2025.

ANSARI, M.; ALAM, M. Advances in time series forecasting: a survey of statistical and deep learning methods. Applied Intelligence, v. 54, p. 2231-2250, 2024.

BUBENIK, P. Persistent homology and topological data analysis. Acta Applicandae

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

Mathematicae, v.170, p. 1-42, 2020.

GOEL, A. et al. Topological data analysis: concepts, algorithms, and applications. ACM Computing Surveys, v. 53, n. 4, p. 1-36, 2020.

HAN, Yang et al. Research on fusing topological data analysis with convolutional neural network. arXiv preprint arXiv:2407.09518, 2024.

LIU, Chenxi et al. Towards cross-modality modeling for time series analytics: A survey in the llm era. arXiv preprint arXiv:2505.02583, 2025.

NGUYEN, Huu Duy. Salinity Intrusion Prediction in the Estuary Using Machine Learning: Vietnam's Mekong Delta Tested for Global Study. Iranian Journal of Science and Technology, Transactions of Civil Engineering, p. 1-19, 2025.

ORESHKIN, Boris N.; CARPOV, Dmitri; CHERNOGLAZOV, Alexey; AHMAD, Syed. N-BEATS: neural basis expansion analysis for interpretable time series forecasting. arXiv preprint, arXiv:1905.10437, 2019. DOI: 10.48550/arXiv.1905.10437.

RAI, Anish et al. Identifying extreme events in the stock market: A topological data analysis. Chaos: An Interdisciplinary Journal of Nonlinear Science, v. 34, n. 10, 2024.

RUBIO, Lihki et al. Forecasting volatility by using wavelet transform, ARIMA and GARCH models. Eurasian Economic Review, v. 13, n. 3, p. 803-830, 2023.

SHEKHAR, Shivanshu; CHOWDHURY, Chandra. Topological data analysis enhanced prediction of hydrogen storage in metal–organic frameworks (MOFs). Materials Advances, v. 5, n. 2, p. 820-830, 2024.

SINGH, Uphar et al. GA-LSTM: Performance Optimization of LSTM driven Time

CC BY

REMUNOM
ISSN 2178-6925

Received: 13/08/2026 - **Accepted:** 11/09/2026
Vol: 13.19
DOI: 10.66104/nvr97c32
Pages: 1-13

Series Forecasting: U. Singh et al. Computational Economics, v. 66, n. 4, p. 2873-2908, 2025.

TAKENS, F. Detecting strange attractors in turbulence. In: RAND, D.; YOUNG, L.-S. (Ed.). Dynamical Systems and Turbulence. Berlin: Springer, 1981. p. 366-381.

VASWANI, A. et al. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS). Long Beach: Curran Associates, 2017. p. 5998-6008.

WANG, Y. et al. A survey of deep learning for time series forecasting. IEEE Transactions on Neural Networks and Learning Systems, v. 32, n. 4, p. 1297-1318, 2021.

CC BY