

Impacto do horizonte histórico de dados no desempenho preditivo e financeiro de redes neurais do tipo LSTM: um estudo de caso com ativos bancários brasileiros

 <https://doi.org/10.47236/2594-7036.2026.v10.2228>

Caio Santos Silva¹
Rafael Lima de Carvalho²

Data de submissão concluída: 19/5/2026. Data de aprovação: 6/7/2026. Data de publicação: 24/8/2026.

Resumo – Este trabalho investigou o impacto do horizonte histórico de dados no desempenho de modelos preditivos baseados em redes neurais do tipo Long Short-Term Memory (LSTM) aplicados ao mercado financeiro. Foram analisados cinco ativos do setor bancário brasileiro, utilizando dados diários obtidos por meio da biblioteca Yahoo Finance. O estudo comparou duas configurações experimentais, com cinco e dez anos de dados históricos, avaliadas por divisão cronológica entre conjuntos de treinamento e teste. O desempenho preditivo foi analisado por métricas de regressão, incluindo RMSE, MAE, MAPE e R^2 , enquanto o desempenho financeiro foi avaliado por meio de backtesting com diferentes estratégias de negociação e comparação com a estratégia Buy & Hold. Os resultados indicaram que, para os ativos com histórico completo, o horizonte de dez anos esteve associado a melhor desempenho preditivo na maior parte das métricas avaliadas. No entanto, o impacto financeiro variou conforme o ativo e a estratégia adotada, indicando que maior aderência preditiva não garante, isoladamente, maior retorno operacional. Conclui-se que a ampliação do horizonte histórico pode favorecer a capacidade de generalização dos modelos LSTM no cenário analisado, embora não seja possível isolar, com o desenho experimental adotado, se esse efeito decorre apenas do maior volume de observações ou também da maior diversidade de regimes de mercado presente no período mais longo.

Palavras-chave: Backtesting. LSTM. Negociação algorítmica. Previsão financeira. Séries temporais.

Impact of historical data horizon on the predictive and financial performance of LSTM neural networks: a case study with brazilian banking stocks

Abstract – This study investigated the impact of the historical data horizon on the performance of predictive models based on Long Short-Term Memory (LSTM) neural networks applied to the financial market. Five assets from the Brazilian banking sector were analyzed using daily data obtained through the Yahoo Finance library. The study compared two experimental configurations, using five and ten years of historical data, evaluated through chronological train-test splits. Predictive performance was assessed using regression metrics, including RMSE, MAE, MAPE, and R^2 , while financial performance was evaluated through backtesting with different trading strategies and comparison against the Buy & Hold benchmark strategy. The results indicated that, for

¹ Bacharel em Ciência da Computação pela Universidade Federal do Tocantins. Palmas, Tocantins, Brasil.  caio.santos@uft.edu.br  <https://orcid.org/0009-0003-7000-6213>  <http://lattes.cnpq.br/6894608338552623>.

² Doutor em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro. Professor do Programa de Pós-Graduação em Governança e Transformação Digital do *Campus Palmas*, da Universidade Federal do Tocantins. Palmas, Tocantins, Brasil.  rafael.lima@uft.edu.br  <https://orcid.org/0000-0002-5296-8641>  <http://lattes.cnpq.br/0175648235036864>.

assets with complete historical records, the ten-year horizon was associated with better predictive performance in most of the evaluated metrics. However, the financial impact varied according to the asset and the trading strategy, indicating that better predictive adherence does not, by itself, guarantee higher operational returns. The study concludes that extending the historical data horizon may improve the generalization capability of LSTM models in the analyzed scenario, although the adopted experimental design does not allow isolating whether this effect is due only to the larger number of observations or also to the greater diversity of market regimes contained in the longer period.

Keywords: Algorithmic trading. Backtesting. Financial forecasting. LSTM. Time series.

Impacto del horizonte histórico de datos en el desempeño predictivo y financiero de redes neuronales del tipo LSTM: un estudio de caso con activos bancarios brasileños

Resumen – Este trabajo investigó el impacto del horizonte histórico de datos en el desempeño de modelos predictivos basados en redes neuronales Long Short-Term Memory (LSTM) aplicadas al mercado financiero. Se analizaron cinco activos del sector bancario brasileño utilizando datos diarios obtenidos mediante la biblioteca Yahoo Finance. El estudio comparó dos configuraciones experimentales, con cinco y diez años de datos históricos, evaluadas mediante división cronológica entre conjuntos de entrenamiento y prueba. El desempeño predictivo fue analizado mediante métricas de regresión, incluyendo RMSE, MAE, MAPE y R^2 , mientras que el desempeño financiero fue evaluado mediante backtesting con diferentes estrategias de negociación y comparación con la estrategia Buy & Hold. Los resultados indicaron que, para los activos con historial completo, el horizonte de diez años estuvo asociado con mejor desempeño predictivo en la mayoría de las métricas evaluadas. Sin embargo, el impacto financiero varió según el activo y la estrategia adoptada, lo que indica que una mayor adherencia predictiva no garantiza, por sí sola, mayor retorno operacional. Se concluye que la ampliación del horizonte histórico puede favorecer la capacidad de generalización de los modelos LSTM en el escenario analizado, aunque el diseño experimental adoptado no permite aislar si este efecto se debe únicamente al mayor volumen de observaciones o también a la mayor diversidad de regímenes de mercado presente en el período más largo.

Palabras clave: Backtesting. LSTM. Negociación algorítmica. Predicción financiera. Series temporales.

Introdução

A previsão de séries temporais financeiras é um problema amplamente estudado, caracterizado pela presença de alta volatilidade, não linearidade e dependências temporais complexas (Tsay, 2005). Métodos tradicionais, como ARIMA e GARCH, apresentam limitações na captura de padrões não lineares e dependências de longo prazo (Zhang; Patuwo; Hu, 1998), o que tem motivado o uso de técnicas de aprendizado profundo, especialmente redes neurais recorrentes do tipo *Long Short-Term Memory* - LSTM (Hochreiter; Schmidhuber, 1997), capazes de modelar dependências temporais complexas em dados sequenciais. Estudos empíricos também reforçam a aplicação de LSTM em previsões financeiras.

Fischer e Krauss (2018), por exemplo, utilizaram redes LSTM para previsão no mercado financeiro e contribuíram para consolidar essa arquitetura como alternativa relevante aos modelos tradicionais em problemas de séries temporais financeiras.

Em revisão recente sobre modelos de deep learning aplicados à previsão de preços de séries financeiras, Zhang, Sjarif e Ibrahim (2024) destacam que arquiteturas profundas vêm sendo amplamente utilizadas nesse domínio, em razão da capacidade de aprender padrões complexos, lidar com relações não lineares e explorar diferentes configurações de entrada e arquitetura. Os autores também apontam que aspectos como estrutura do modelo, disponibilidade de dados e escolha das variáveis de entrada permanecem relevantes para a avaliação do desempenho preditivo.

Dioubi *et al.* (2024) argumentam que a natureza não linear e não estacionária dos mercados financeiros dificulta a capacidade preditiva de modelos tradicionais, razão pela qual arquiteturas híbridas baseadas em decomposição de séries temporais e LSTM tendem a apresentar melhor desempenho ao lidar com diferentes padrões históricos de mercado. Na mesma direção, Bousbaa, Sanchez-Medina e Bencharef (2023) sustentam que modelos estáticos treinados sobre distribuições históricas fixas sofrem deterioração progressiva de desempenho à medida que o comportamento do mercado muda, especialmente em cenários de *concept drift* (quando o comportamento do mercado muda drasticamente) e alterações de regime financeiro. Os mesmos autores demonstram que mecanismos adaptativos, como o mecanismo de janelas deslizantes (*sliding windows*) dinâmicas e aprendizado incremental, tornam-se necessários para preservar a capacidade de generalização em séries financeiras altamente voláteis. Neste trabalho, a generalização é a capacidade de o modelo manter desempenho adequado em dados não utilizados durante o treinamento, ou seja, conseguir produzir previsões consistentes mesmo diante de cenários diferentes daqueles observados na etapa de aprendizagem, como períodos de alta, queda abrupta, crise e recuperação do mercado.

De forma semelhante, Shabe, Engelbrecht e Anderson (2025) observaram que modelos recorrentes baseados em LSTM apresentaram queda significativa de desempenho durante a pandemia da COVID-19, evidenciando que mudanças abruptas de regime comprometem a estabilidade preditiva quando os modelos não conseguem adaptar-se continuamente aos novos padrões históricos. O trabalho de Balamohan *et al.* (2025) destaca que grandes volumes de dados financeiros podem introduzir ruídos e instabilidades capazes de prejudicar a convergência e favorecer o sobreajuste, motivo pelo qual os autores propõem a utilização do otimizador *AdamW* para melhorar a capacidade de generalização das redes LSTM em diferentes condições de mercado.

De modo semelhante, Lahbabi *et al.* (2025) afirmam que ambientes financeiros não estacionários exigem mecanismos estocásticos adicionais para evitar convergência prematura e baixa robustez dos modelos, demonstrando que a introdução controlada de ruído no treinamento melhora a generalização em cenários de elevada volatilidade. Kahra e Kahra (2025), por sua vez, observam que modelos profundos aplicados a séries financeiras enfrentam problemas recorrentes de *overfitting* (quando o modelo se adapta demais aos dados de treinamento e falha na generalização) e baixo desempenho diante de mudanças rápidas de mercado, reforçando que a diversidade temporal dos dados e a capacidade de adaptação do modelo são fatores centrais para previsões mais estáveis.

Em suma, essas evidências reforçam a existência de uma lacuna relacionada ao impacto do horizonte histórico utilizado no treinamento de modelos LSTM aplicados ao mercado financeiro. Em séries financeiras, a ampliação do período histórico pode

aumentar o número de observações disponíveis e, simultaneamente, expor o modelo a diferentes condições de mercado, como períodos de crescimento, crise, queda abrupta e recuperação. Entretanto, quando esses fatores ocorrem de forma conjunta, torna-se metodologicamente difícil separar o efeito do maior volume de dados do efeito da diversidade de regimes presentes no período analisado.

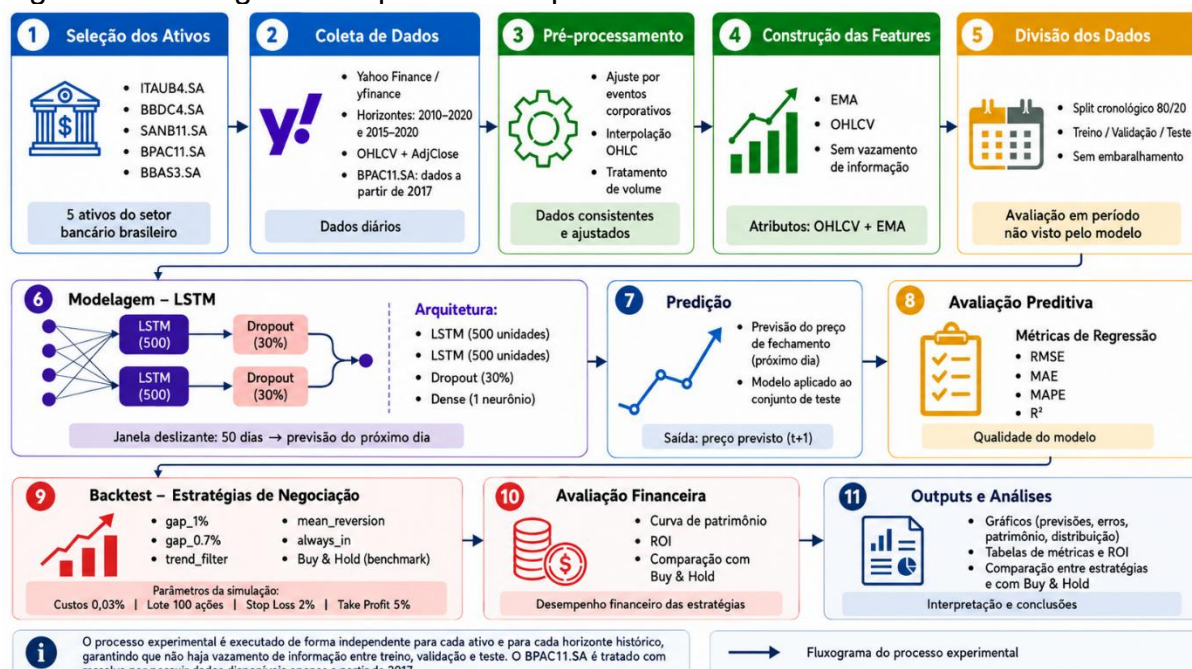
Para contribuir com essa discussão, este trabalho analisa o impacto de diferentes horizontes históricos de treinamento no desempenho de modelos LSTM aplicados ao mercado financeiro, com foco em ativos do setor bancário brasileiro. Para isso, foram realizados experimentos computacionais comparando configurações com cinco e dez anos de dados históricos, avaliadas por métricas de regressão e por simulações de estratégias de negociação baseadas nas previsões do modelo. A análise busca verificar se a ampliação do horizonte histórico está associada a ganhos de desempenho preditivo e financeiro, sem assumir, entretanto, que o desenho experimental adotado permite isolar completamente o efeito do volume de observações em relação à diversidade de regimes de mercado presente no período mais longo.

Materiais e métodos

Os dados históricos de preços de ações foram coletados por meio da biblioteca *yfinance* (Aroussi, 2019), que realiza consultas automatizadas à *API* do *Yahoo Finance*. O estudo contempla cinco ativos do setor bancário brasileiro, todos negociados na bolsa brasileira B3 (Brasil, Bolsa, Balcão): ITUB4.SA (Itaú Unibanco), BBDC4.SA (Bradesco), BBAS3.SA (Banco do Brasil), SANB11.SA (Santander Brasil) e BPAC11.SA (BTG Pactual). A escolha por bancos de grande capitalização justifica-se pela liquidez elevada dos papéis e pela representatividade do setor no índice Ibovespa.

Para cada ativo, foram obtidas as seguintes variáveis diárias: preço de abertura (Open), máxima (High), mínima (Low), fechamento (Close), fechamento ajustado (Adj Close) e volume negociado (Volume). Neste trabalho, o termo **ativo** se refere a cada ação ou unidade negociada na B3 analisada no estudo. Já o termo **pregão** corresponde a um dia útil de negociação no mercado financeiro; assim cada observação diária da base representa um pregão para determinado ativo. O estudo foi estruturado a partir da comparação entre dois horizontes históricos de treinamento: uma configuração com aproximadamente dez anos de dados, abrangendo o período de janeiro de 2010 a dezembro de 2020, e uma configuração com aproximadamente cinco anos de dados, abrangendo o período de janeiro de 2015 a dezembro de 2020. Essa comparação foi aplicada aos cinco ativos selecionados: ITUB4.SA (Itaú Unibanco Holding S.A.), BBDC4.SA (Banco Bradesco S.A.), BBAS3.SA (Banco do Brasil S.A.), SANB11.SA (Banco Santander Brasil S.A.) e BPAC11.SA (BTG Pactual S.A.). Em todos os casos, os dados foram divididos de forma cronológica, utilizando 80% das observações para treinamento e 20% para teste, sem embaralhamento das séries temporais. Os períodos analisados incluem eventos econômicos e políticos relevantes, como crise hídrica e energética, impeachment, recessão econômica e o crash provocado pela pandemia de COVID-19 em março de 2020. Assim, os conjuntos de teste representam cenários desafiadores para avaliar tanto a qualidade preditiva dos modelos quanto o desempenho das estratégias de negociação. A Figura 1 apresenta o processo experimental adotado no estudo, organizado em etapas sequenciais de coleta, pré-processamento, modelagem LSTM, backtesting e avaliação.

Figura 1 – Fluxograma do processo experimental



Fonte: elaborado pelos autores (2026).

Para cada ativo, foram treinados modelos independentes nas configurações de cinco e dez anos. Em ambas as configurações, a divisão entre treino e teste foi realizada de forma estritamente cronológica, aplicando-se a proporção fixa de 80% para treinamento e 20% para teste sobre o total de observações disponíveis, sem ajuste manual das datas de corte. Dessa forma, os períodos de teste diferem entre as configurações, pois decorrem diretamente do tamanho do horizonte histórico utilizado. Essa escolha preserva a ordem temporal dos dados e evita vazamento de informação futura para o treinamento. No caso específico do BPAC11.SA, a comparação deve ser interpretada com cautela, pois o ativo possui dados disponíveis apenas a partir de 2017, não compondo uma série completa de dez anos.

Pré-processamento dos dados

O pré-processamento é baseado na metodologia proposta por Zanotto e Hölbíg (2026), com adaptações voltadas à consistência das séries temporais. Inicialmente, os preços de abertura, máxima, mínima e fechamento foram ajustados por um fator de correção calculado a partir da razão entre o preço de fechamento ajustado e o preço de fechamento original. Esse fator representa a diferença entre o preço bruto negociado no pregão e o preço ajustado após eventos corporativos, como desdobramentos de ações, agrupamentos e distribuição de dividendos. Ao aplicar essa razão às variáveis de preço, busca-se colocar a série histórica em uma base comparável ao longo do tempo, reduzindo distorções artificiais que poderiam ser interpretadas pelo modelo como variações reais de mercado. Em seguida, valores ausentes nas variáveis de preço foram tratados por interpolação linear bidirecional, enquanto registros de volume nulo foram substituídos pelo último valor válido disponível, com preenchimento retroativo aplicado quando necessário.

Como variável adicional, foi calculada a Média Móvel Exponencial (EMA) de cada série de fechamento. A EMA é um indicador técnico que suaviza série de preços, atribuindo maior peso aos valores mais recentes; por isso, é utilizada para representar tendências de curto ou médio prazo. O período da EMA foi definido dinamicamente

em função do comprimento da série, respeitando o intervalo [5, 60] dias, pela expressão $span = \max(5, \min(\text{int}(n \times 0,2), 60))$, em que n representa o número total de observações. Para preservar a integridade temporal dos dados, a EMA foi calculada separadamente para os conjuntos de treino e de teste: no treino, aplicou-se a função de suavização exponencial diretamente sobre a série histórica; no teste, o cálculo partiu do último valor da EMA do treino, propagando a atualização ponto a ponto sem acesso a observações futuras.

A normalização das variáveis foi conduzida por meio do *StandardScaler* da biblioteca *scikit-learn*³, método que transforma cada *feature*, isto é, cada variável de entrada do modelo, para média zero e desvio padrão unitário. Essa etapa evita que variáveis com escalas maiores, como volume negociado, tenham influência desproporcional durante o treinamento. O ajuste do normalizador foi realizado exclusivamente sobre o conjunto de treino, e a transformação inversa aplicada às previsões para recuperar os valores na escala original de preços.

Arquitetura da rede neural LSTM

A arquitetura da rede neural baseia-se no trabalho de Zanotto e Hölbig (2026), e é composta por duas camadas recorrentes do tipo Long Short-Term Memory (LSTM), cada uma com 500 unidades ocultas, seguidas de camadas de regularização Dropout com taxa de 30%, e uma camada densa de saída com um único neurônio responsável pela predição do preço de fechamento do pregão seguinte. A escolha pela arquitetura LSTM deve-se à sua capacidade de capturar dependências temporais em sequências de dados, característica relevante para séries financeiras, nas quais o comportamento atual dos preços pode estar associado a padrões observados em períodos anteriores.

As variáveis de entrada do modelo foram formadas pelas características diárias de mercado após o pré-processamento, incluindo preços de abertura, máxima, mínima, fechamento, volume negociado e a Média Móvel Exponencial (EMA). Para cada previsão, utilizou-se uma janela temporal de 50 observações consecutivas, de modo que o modelo recebesse como entrada os dados dos 50 pregões anteriores e produzisse como saída um único valor escalar, correspondente à previsão do preço de fechamento do pregão seguinte. Assim, a tarefa preditiva foi formulada como um problema supervisionado de regressão em séries temporais.

Os mesmos hiperparâmetros foram mantidos em todos os ativos e nas duas configurações de horizonte histórico, cinco e dez anos. Nesse contexto, hiperparâmetros são configurações definidas antes do treinamento do modelo, como número de camadas, quantidade de unidades ocultas, tamanho da janela temporal, taxa de Dropout, número máximo de épocas e tamanho do lote. Essa decisão foi adotada para preservar a comparabilidade entre os experimentos, evitando que diferenças de desempenho fossem atribuídas a ajustes específicos da arquitetura para determinado ativo ou horizonte temporal. Dessa forma, o objetivo do estudo não foi otimizar a arquitetura da rede neural, mas avaliar se a alteração do horizonte histórico de treinamento estaria associada a mudanças no desempenho preditivo e financeiro.

O modelo foi compilado com o otimizador Adam e função de perda Mean Squared Error (MSE). O MSE mede o erro quadrático médio entre os valores reais e os previstos pelo modelo, e é ajustado ao problema de regressão por penalizar desvios maiores entre as previsões e os preços observados. O treinamento foi

³ https://scikit-learn.org/stable/supervised_learning.html

executado por até 50 épocas, com tamanho de lote igual a 64, validação interna de 10% dos dados de treinamento e interrupção antecipada ou Early Stopping com paciência de 10 épocas. O Early Stopping interrompe o treinamento quando a perda de validação deixa de melhorar, restaurando os pesos da época com melhor desempenho. Esse procedimento foi utilizado para reduzir o risco de sobreajuste e melhorar a capacidade de generalização do modelo.

Tabela 1 - Hiperparâmetros utilizados na arquitetura LSTM

Componente	Configuração adotada
<i>Tipo de Modelo</i>	Rede neural LSTM
<i>Camadas recorrentes</i>	2 camadas LSTM
<i>Unidades ocultas</i>	500 unidades por camada
<i>Regularização</i>	Dropout de 30%
<i>Janela temporal</i>	50 pregões
<i>Variáveis de entrada</i>	Open, High, Low, Close, Volume e EMA
<i>Variável de saída</i>	Preço de fechamento do pregão seguinte
<i>Função de perda</i>	Mean Squared Error (MSE)
<i>Otimizador</i>	Adam
<i>Épocas máximas</i>	50
<i>Tamanho do lote</i>	64
<i>Validação interna</i>	10% dos dados de treino
<i>Critério de parada</i>	Early Stopping com paciência de 10 épocas

Fonte: elaborado pelos autores (2026).

Estratégias de geração de sinais

A partir das predições do modelo, cinco estratégias ativas de negociação foram implementadas para gerar sinais operacionais. Diferentemente de uma estratégia passiva, que apenas compra e mantém o ativo, uma estratégia ativa altera sua posição ao longo do tempo com base em regras de decisão. Os sinais gerados foram: compra (BUY), venda a descoberto (SELL_SHORT) ou neutralidade (HOLD). O sinal de compra indica a abertura de posição comprada, buscando ganho com a valorização do ativo. A venda a descoberto representa uma posição vendida, usada para simular ganho em cenários de queda do preço. O sinal HOLD indica ausência de nova operação naquele pregão. Em todas as estratégias, o sinal para o dia i é calculado com base no preço real do dia anterior ($i-1$) e na predição para o dia i , garantindo que nenhuma informação do pregão corrente seja utilizada na tomada de decisão.

Gap percentual (gap_1% e gap_0,7%): emite sinal de compra quando a predição supera o preço anterior em pelo menos 1% (ou 0,7%), e sinal de venda a descoberto quando a predição é inferior a esse limiar. Casos intermediários recebem o sinal *HOLD*.

Filtro de tendência (trend_filter): combina a predição do modelo com a EMA do período. O sinal de compra é emitido quando tanto a predição quanto o preço real estão acima da EMA; o sinal de venda a descoberto, quando ambos estão abaixo da EMA.

Reversão à média (mean_reversion): baseia-se no desvio do preço em relação à EMA. Quando o preço está 2% abaixo da EMA, emite sinal de compra; quando 2% acima, emite sinal de venda a descoberto.

Always-in (always_in): compra sempre que a predição supera o preço anterior e vende a descoberto em caso contrário, permanecendo posicionada em todos os pregões.

Backtest financeiro

O desempenho financeiro de cada estratégia foi avaliado por meio de *backtesting*, procedimento que simula como uma estratégia teria se comportado em dados históricos, sem envolver operações reais no mercado, utilizado em sistemas de negociação para validação de estratégias quantitativas (López de Prado, 2018). A simulação usou um capital inicial de R\$ 5.000,00, taxa de corretagem de 0,03% por operação, lote mínimo de 100 ações e regras de controle de risco. O lote mínimo representa a menor quantidade de ações considerada para abertura de uma operação. O *stop loss* foi definido em 2% e encerra automaticamente a posição quando a perda atinge esse limite, enquanto o *take profit* foi definido em 5% e encerra a posição quando o ganho atinge esse percentual. Ao término do período de teste, as posições abertas são liquidadas pelo preço do último pregão.

Como referência de comparação, calculou-se o retorno da estratégia *Buy & Hold*, que consiste na compra do maior lote possível no primeiro pregão do período de teste e manutenção da posição até o último pregão, sem operações intermediárias. A métrica principal de comparação entre estratégias foi o Retorno sobre o Investimento (ROI), dado por $ROI = (\text{patrimônio final} / \text{capital inicial} - 1) \times 100$; ele mostra a variação percentual do patrimônio final em relação ao capital inicial.

Avaliação do modelo preditivo

A calibração do modelo consistiu no treinamento supervisionado da rede LSTM a partir das janelas temporais construídas no conjunto de treino. Em cada amostra, os valores normalizados dos 50 pregões anteriores foram utilizados como entrada, e o preço de fechamento do pregão seguinte foi utilizado como valor-alvo. Após o treinamento, as previsões foram transformadas de volta para a escala original dos preços, permitindo a comparação direta com os valores reais e o cálculo das métricas de regressão. A separação cronológica entre treino, validação e teste foi mantida para evitar vazamento de informação futura.

A qualidade das previsões foi mensurada por quatro métricas de regressão calculadas sobre os preços na escala original: Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Percentual Absoluto Médio (MAPE) e Coeficiente de Determinação (R^2).

O RMSE penaliza erros maiores de forma quadrática, sendo sensível a outliers e expresso na mesma unidade de preço (reais). O MAE representa o erro absoluto médio sem penalização diferenciada, oferecendo uma leitura direta da magnitude do desvio típico das previsões. O MAPE mostra o erro em termos percentuais relativos ao valor real, facilitando a comparação entre ativos com diferentes faixas de preço. O coeficiente de determinação (R^2) mede a proporção da variância dos preços reais explicada pelo modelo. Valores próximos de 1,0 indicam alto poder preditivo, enquanto valores próximos de zero sugerem desempenho semelhante à previsão baseada na média da série. Valores negativos podem ocorrer quando o modelo apresenta erro superior ao de um preditor baseado apenas na média dos dados.

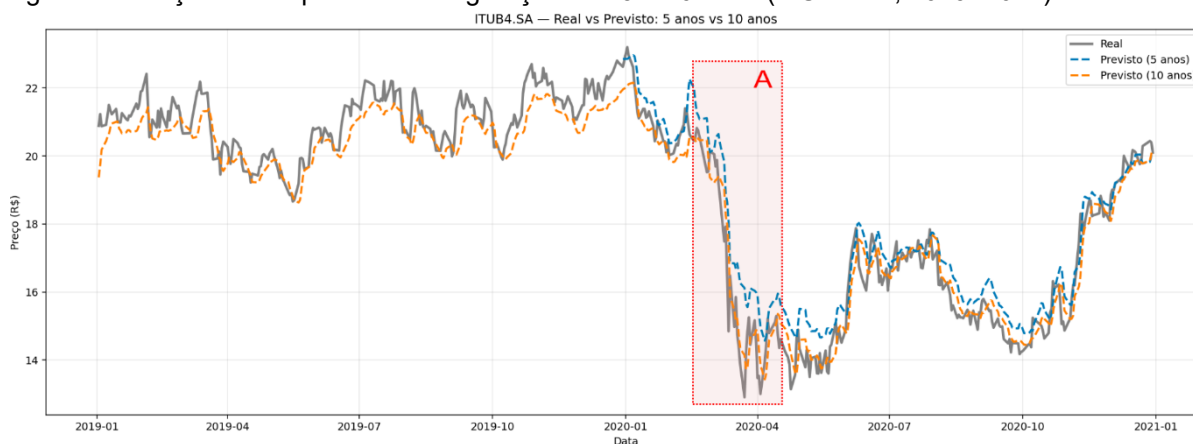
Essas métricas foram comparadas entre as configurações de cinco e dez anos para todos os ativos analisados, permitindo avaliar se a ampliação do horizonte histórico esteve associada a diferenças na acurácia preditiva e no desempenho financeiro das estratégias. No caso do BPAC11.SA, os resultados foram interpretados separadamente, devido à limitação de disponibilidade histórica do ativo, com dados disponíveis apenas a partir de 2017.

Resultados e discussões

Qualidade preditiva do modelo

A Figura 2 apresenta, de forma ilustrativa, as curvas de preço real e previsto para o ITUB4.SA nas configurações de cinco e dez anos, com as datas reais do período de teste no eixo horizontal. O ITUB4.SA foi utilizado como exemplo visual por apresentar comportamento representativo da comparação entre horizontes históricos nos ativos com série completa. Ambos os modelos acompanham a dinâmica geral dos preços, incluindo a queda provocada pelo crash da COVID-19 e a posterior recuperação. No entanto, o modelo treinado com dez anos apresentou maior proximidade em relação à série real nesse ativo, enquanto o modelo de cinco anos mostrou maior defasagem em parte do período de queda. Essa diferença sugere que o maior horizonte histórico pode ter favorecido a generalização do modelo nesse caso específico.

Figura 2 – Preço real vs. previsto: configurações de 5 e 10 anos (ITUB4.SA, 2019–2021)



Fonte: elaborado pelos autores (2026).

A Tabela 2 apresenta as métricas de regressão calculadas para os cinco ativos nas configurações de cinco e dez anos. Considerando os ativos com histórico completo, o horizonte de dez anos apresentou menor MAPE e maior R^2 em ITUB4.SA, BBAS3.SA, BBDC4.SA e SANB11.SA. Em ITUB4.SA, por exemplo, o MAPE foi reduzido de 4,24% para 2,48%, enquanto o R^2 aumentou de 0,8770 para 0,9550. Em BBAS3.SA, o MAPE passou de 6,02% para 3,63%, e o R^2 aumentou de 0,8508 para 0,9410. Esses resultados indicam que, nos ativos com série histórica completa, o maior horizonte temporal esteve associado a melhor aderência preditiva. O BPAC11.SA apresentou comportamento distinto, com pior desempenho na configuração denominada de dez anos; esse resultado, contudo, deve ser interpretado com cautela em razão da disponibilidade histórica limitada do ativo, iniciada apenas em 2017.

Tabela 2 – Métricas de regressão por ativo e horizonte histórico

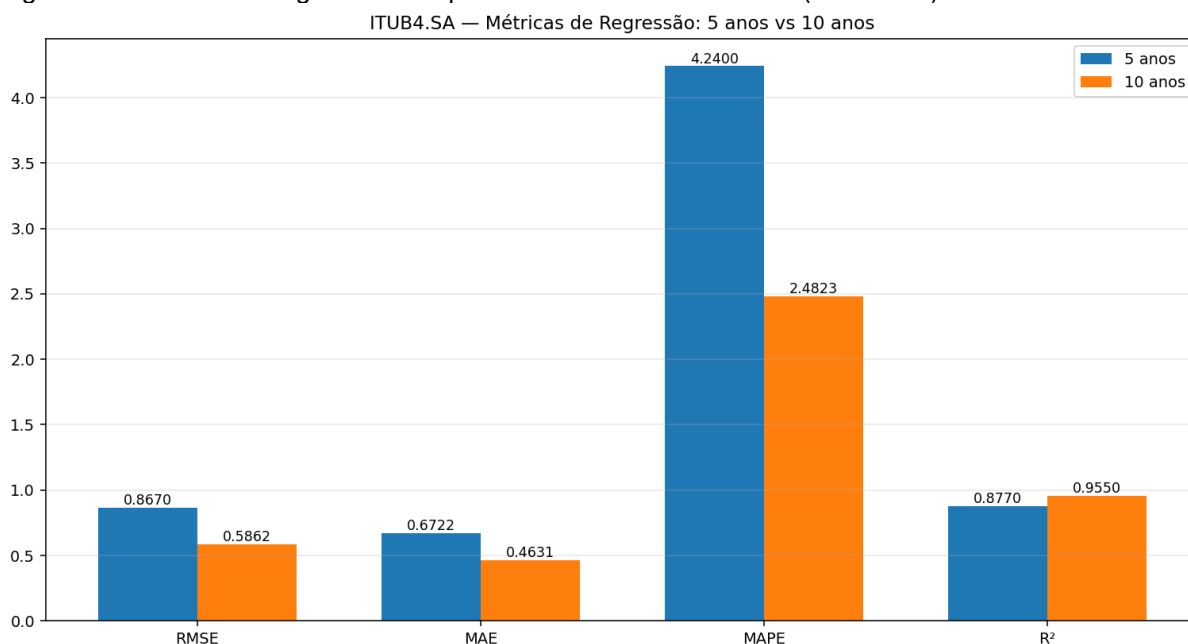
Ativo	Horizonte	RMSE	MAE	MAPE (%)	R^2
ITUB4.SA	5 anos	0,8670	0,6722	4,2400	0,8770
ITUB4.SA	10 anos	0,5862	0,4631	2,4823	0,9550
BBAS3.SA	5 anos	0,8904	0,6197	6,0229	0,8508
BBAS3.SA	10 anos	0,6212	0,5003	3,6272	0,9410
BBDC4.SA	5 anos	0,8432	0,6642	5,3528	0,8673
BBDC4.SA	10 anos	0,7537	0,6029	4,0737	0,9213
SANB11.SA	5 anos	1,2569	0,9050	4,5186	0,9255
SANB11.SA	10 anos	1,2362	0,9619	3,9302	0,9277
BPAC11.SA*	5 anos	1,4864	1,2728	7,1084	0,0872
BPAC11.SA*	10 anos	2,4186	2,1963	12,4497	-1,4167

Fonte: elaborado pelos autores (2026)

Nota: BPAC11.SA possui limitação de disponibilidade histórica, com dados disponíveis apenas a partir de 2017; por isso, sua configuração de dez anos não representa uma série completa de dez anos.

A Figura 3 ilustra, especificamente para o ITUB4.SA, a comparação entre as métricas de regressão nas configurações de cinco e dez anos. Nesse ativo, o MAPE foi reduzido de 4,2400% para 2,4823%, enquanto o RMSE diminuiu de 0,8670 para 0,5862. Como a Tabela 2 sintetiza os resultados para todos os ativos avaliados, a Figura 3 deve ser interpretada como exemplo visual da comparação, e não como evidência única do estudo.

Figura 3 – Métricas de regressão comparativas: 5 anos vs. 10 anos (ITUB4.SA)

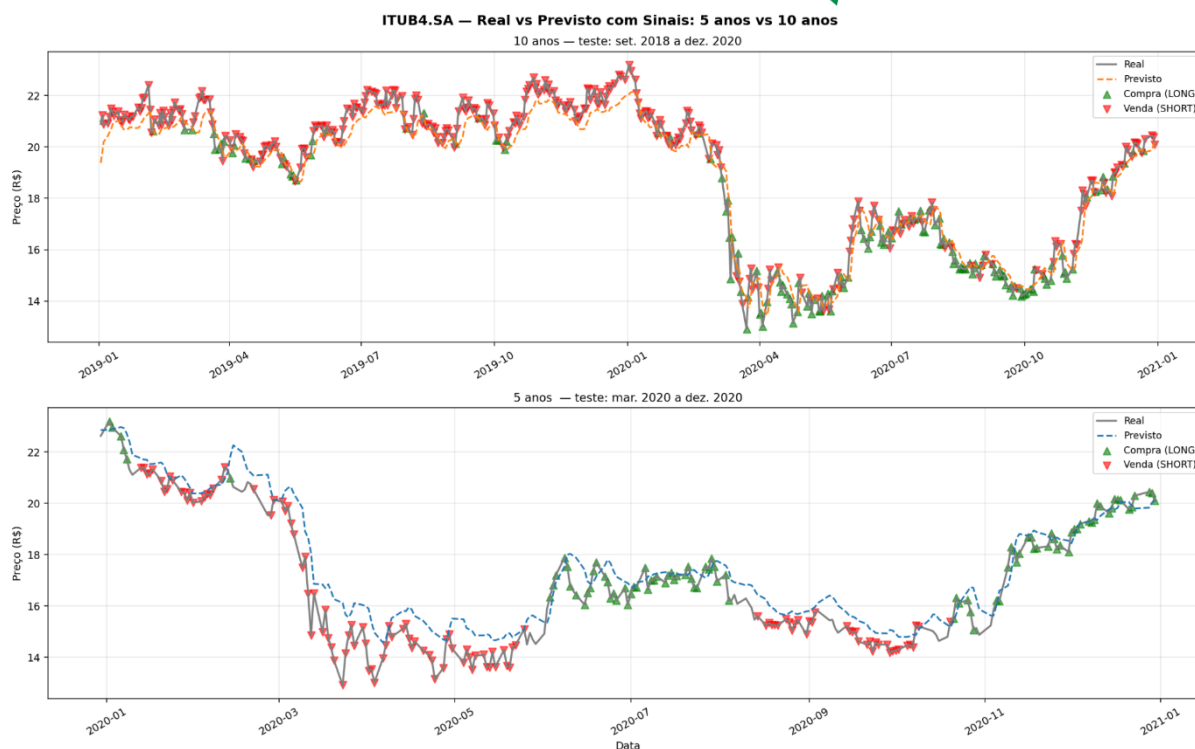


Fonte: elaborado pelos autores (2026).

Qualidade dos sinais gerados para negociação

A Figura 4 apresenta os sinais de compra (triângulos verdes) e venda a descoberto (triângulos vermelhos) gerados pela melhor estratégia de cada configuração, sobrepostos às curvas de preço real e previsto. A análise visual dos sinais permite compreender a lógica operacional do modelo e identificar as diferenças de comportamento entre as duas configurações.

Figura 4 – Preço real vs. Previsto com sinais de negociação: 5 e 10 anos (ITUB4.SA)



Fonte: elaborado pelos autores (2026).

Na parte superior da Figura 4, referente ao experimento de dez anos, os sinais estão bem distribuídos ao longo de todo o período e demonstram coerência com a dinâmica do mercado. Durante a fase de alta entre 2019 e início de 2020, observa-se alternância entre sinais de compra e de venda a descoberto, refletindo oscilações de curto prazo em torno da tendência principal. Na queda do *crash*, em março de 2020, há concentração de sinais de venda a descoberto em parte do movimento descendente, o que sugere que o modelo captou momentos relevantes da reversão de tendência. Na fase de recuperação, os sinais de compra voltam a aparecer com maior frequência. Esse comportamento ajuda a explicar o desempenho financeiro superior das estratégias no caso do ITUB4.SA com dez anos de dados.

No painel inferior, referente ao experimento de cinco anos, o comportamento dos sinais apresenta maior instabilidade durante o *crash* da COVID-19. Entre fevereiro e abril de 2020, observa-se alternância frequente entre sinais de compra e venda a descoberto, sem a mesma consistência direcional observada no painel superior. Esse padrão é coerente com a menor precisão do modelo de cinco anos, com base no MAPE de 4,2400% frente a 2,4823% no experimento de dez. Assim, a maior oscilação dos sinais em períodos de alta volatilidade pode ter contribuído para o desempenho inferior da maioria das estratégias no experimento de cinco anos.

Desempenho financeiro das estratégias

A Tabela 3 apresenta o ROI da melhor estratégia ativa obtida para cada ativo nas configurações de cinco e dez anos, juntamente com o retorno da estratégia Buy & Hold. Essa comparação permite verificar se a melhora preditiva observada nas métricas de regressão se traduz em ganhos financeiros nas estratégias simuladas. Ambos os experimentos cobrem o *crash* da COVID-19, porém com extensões distintas: o experimento de dez anos compreende janeiro de 2019 a dezembro de 2020, cobrindo um período de alta sustentada anterior ao *crash* seguido da queda e da recuperação; o experimento de cinco anos compreende janeiro de 2020 a

dezembro de 2020, iniciando poucos dias antes do *crash* e capturando a queda completa e a recuperação. Esses períodos resultam diretamente da aplicação da divisão cronológica 80/20 sobre cada horizonte histórico, sem ajuste manual das datas de corte.

Tabela 3 - ROI da melhor estratégia ativa e do Buy & Hold por ativo e horizonte histórico

Ativo	Horizonte	Melhor estratégia ativa	ROI melhor estratégia (%)	ROI Buy & Hold (%)
ITUB4.SA	5 anos	trend_filter	+4,60	-10,15
ITUB4.SA	10 anos	always_in	+87,77	-3,21
BBAS3.SA	5 anos	mean_reversion	+6,42	-15,91
BBAS3.SA	10 anos	mean_reversion	+26,39	-11,11
BBDC4.SA	5 anos	mean_reversion	-21,81	-11,16
BBDC4.SA	10 anos	always_in	+0,41	-3,49
SANB11.SA	5 anos	trend_filter	-5,08	-1,07
SANB11.SA	10 anos	gap_1%	-1,13	+5,92
BPAC11.SA*	5 anos	trend_filter	+15,93	+44,73
BPAC11.SA*	10 anos	trend_filter	+7,65	+44,73

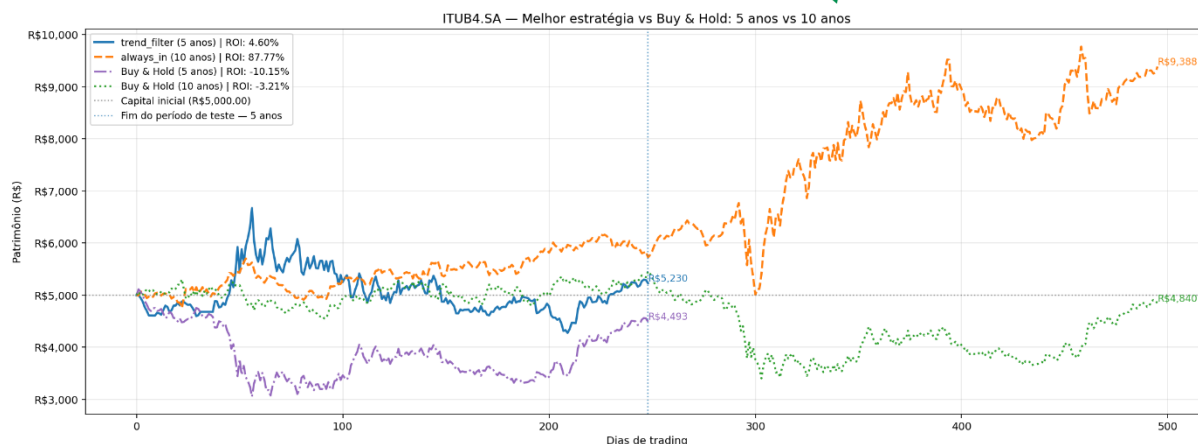
Fonte: elaborado pelos autores (2026)

Nota: BPAC11.SA possui limitação de disponibilidade histórica, com dados disponíveis apenas a partir de 2017.

A análise financeira mostra que a melhora nas métricas preditivas não se converteu de forma uniforme em maior retorno operacional. Em ITUB4.SA e BBAS3.SA, o horizonte de dez anos apresentou os melhores resultados financeiros entre as estratégias ativas, com ROI de +87,77% e +26,39%, respectivamente, superando em ambos os casos a estratégia Buy & Hold. Em BBDC4.SA, a configuração de dez anos também melhorou o desempenho em relação à configuração de cinco anos, mas o ganho foi modesto, com ROI de +0,41%. Em SANB11.SA, a melhor estratégia ativa permaneceu negativa nas duas configurações, embora o resultado de dez anos tenha sido menos desfavorável que o de cinco anos. Nesse ativo, o Buy & Hold de dez anos superou as estratégias ativas simuladas. O BPAC11.SA apresentou comportamento distinto: a configuração de cinco anos superou a de dez anos nas métricas preditivas e no ROI da melhor estratégia ativa, mas ambas ficaram abaixo do Buy & Hold. Esse resultado deve ser analisado como caso particular, pois o ativo possui histórico limitado e não apresenta série completa desde 2010.

A Figura 5 apresenta a evolução patrimonial da melhor estratégia de cada configuração para o ITUB4.SA, comparada às estratégias Buy & Hold dos horizontes de cinco e dez anos. A linha vertical pontilhada indica o encerramento do período de teste de cinco anos no dia 250. A estratégia *always_in* de dez anos encerrou com R\$ 9.388, enquanto a *trend_filter* de cinco anos encerrou com R\$ 5.230. As estratégias *Buy & Hold* apresentaram desempenho inferior em ambos os cenários, encerrando com R\$ 4.840 no experimento de dez anos e R\$ 4.493 no experimento de cinco anos, ambos abaixo do capital inicial de R\$ 5.000.

Figura 5 – Melhor estratégia de cada configuração vs. Buy & Hold (ITUB4.SA)

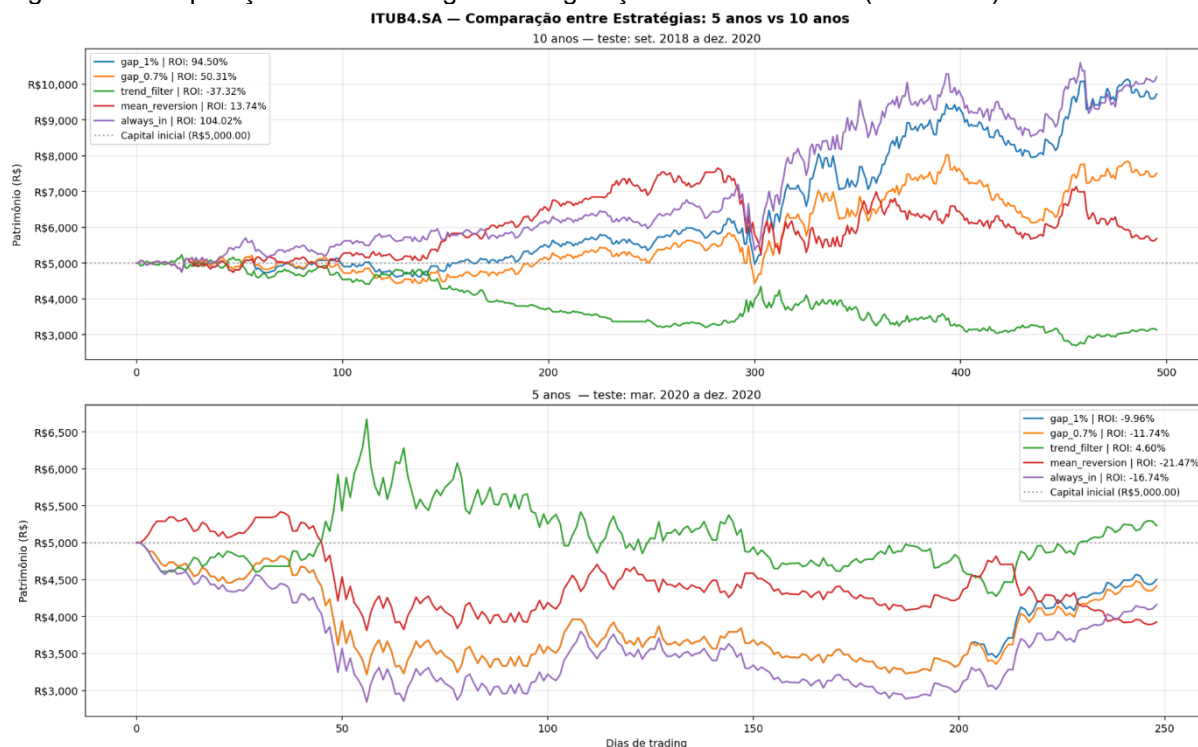


Fonte: elaborado pelos autores (2026).

Análise comparativa entre horizontes históricos

A Figura 6 apresenta a comparação entre todas as estratégias em cada configuração em painéis separados. No painel superior, referente ao experimento de dez anos, a maioria das curvas apresenta trajetória crescente na primeira metade do período, correspondente à fase de alta do mercado entre 2019 e início de 2020, seguida de queda abrupta ao redor do dia 280, quando ocorre o *crash*, e recuperação diferenciada por estratégia. O *always_in* e o *gap_1%* destacam-se por manter patrimônio acima do capital inicial mesmo após o *crash*, graças à geração de sinais vendidos durante a queda. No painel inferior, referente ao experimento de cinco anos, as curvas apresentam trajetória predominantemente decrescente na fase inicial, correspondente ao *crash*, com apenas o *trend_filter* conseguindo encerrar o período em território positivo.

Figura 6 – Comparação entre estratégias: configurações de 10 e 5 anos (ITUB4.SA)



Fonte: elaborado pelos autores (2026).

A análise conjunta das métricas de regressão e do desempenho financeiro indica que o horizonte histórico utilizado no treinamento influencia a qualidade preditiva dos modelos e pode afetar o desempenho das estratégias de negociação, embora esse efeito não seja uniforme entre todos os ativos. Nos ativos com histórico completo, o modelo treinado com dez anos apresentou, em geral, melhor desempenho preditivo do que o modelo treinado com cinco anos. Uma possível explicação é que o horizonte mais longo expõe o modelo a maior quantidade de observações e a diferentes condições de mercado, incluindo períodos de crescimento, crise e recuperação. Entretanto, como esses dois fatores ocorrem simultaneamente no desenho experimental adotado, não é possível afirmar de forma causal qual deles foi predominante. Na configuração de cinco anos, os modelos apresentaram maior dificuldade em alguns períodos de alta volatilidade, especialmente no crash da COVID-19. Esse comportamento pode estar relacionado ao menor horizonte histórico disponível para treinamento, mas também pode refletir as características específicas do período de teste de cada ativo.

Considerações finais

Os resultados obtidos indicam que o horizonte histórico de dados influencia o desempenho preditivo de modelos LSTM aplicados à previsão de ativos bancários brasileiros. Nos ativos com série histórica completa, ITUB4.SA, BBAS3.SA, BBDC4.SA e SANB11.SA, a configuração com dez anos apresentou menor MAPE e maior R^2 em comparação à configuração com cinco anos. Esse resultado sugere que a ampliação do horizonte temporal de treinamento pode favorecer a aderência das previsões em séries financeiras, especialmente em períodos que envolvem mudanças abruptas de mercado. Entretanto, o estudo não permite afirmar que esse efeito decorre exclusivamente da diversidade de regimes de mercado, pois o horizonte mais longo também implica maior número de observações disponíveis para treinamento.

A avaliação financeira mostrou resultados mais heterogêneos. Em ITUB4.SA e BBAS3.SA, o horizonte de dez anos esteve associado a melhor desempenho financeiro das estratégias ativas. Em BBDC4.SA, houve melhora em relação à configuração de cinco anos, mas com retorno próximo ao capital inicial. Em SANB11.SA, as estratégias ativas não superaram o Buy & Hold na configuração de dez anos. Já o BPAC11.SA apresentou comportamento distinto, com melhor desempenho preditivo na configuração de cinco anos e forte superioridade do Buy & Hold, resultado que deve ser interpretado com cautela em razão da limitação de dados históricos disponíveis apenas a partir de 2017. Assim, os achados indicam que melhor desempenho preditivo pode contribuir para melhores estratégias, mas não garante, isoladamente, maior retorno financeiro.

Como limitação principal, destaca-se que o desenho experimental compara horizontes históricos distintos, mas não isola estatisticamente o efeito do volume de observações em relação ao efeito da diversidade de regimes de mercado. Para que essa distinção seja avaliada com maior precisão, estudos futuros podem construir experimentos controlados nos quais o número de observações seja reduzido, mantendo-se a diversidade de regimes, ou, inversamente, em que se comparem bases com volumes semelhantes e diferentes composições históricas. Recomenda-se também investigar mecanismos adaptativos de treinamento, como janelas deslizantes dinâmicas e aprendizado incremental, além da incorporação de dados alternativos, como indicadores macroeconômicos e métricas de sentimento, e da otimização automática de hiperparâmetros.

Referências

AROUISSI, Ran. yfinance: **Yahoo! finance market data downloader**. Version 0.1.54, 2019. Disponível em: <https://pypi.org/project/yfinance/>. Acesso em: 10 abr. 2026.

BALAMOCHAN, S. *et al.* Advancement of stock market forecasting with adamW and LSTM: a scalable deep learning system leveraging technical indicators. In: **International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering**. IEEE, 2025. DOI: 10.1109/ICECONF65644.2025.11379417.

BOUSBAA, Z.; SANCHEZ-MEDINA, J.; BENCHAREF, O. Financial time series forecasting: a data stream mining-based system. **Electronics**, v. 12, n. 9, p. 2039, 2023. DOI: 10.3390/electronics12092039.

DIOUBI, F. *et al.* Enhancing stock market predictions via hybrid external trend and internal components analysis and long short term memory model. **Journal of King Saud University - Computer and Information Sciences**, v. 36, 2024. DOI: 10.1016/j.jksuci.2024.102252.

FISCHER, Thomas; KRAUSS, Christopher. Deep learning with long short-term memory networks for financial market predictions. **European Journal of Operational Research**, Amsterdam, v. 270, n. 2, p. 654-669, 2018. DOI: 10.1016/j.ejor.2017.11.054.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. **Neural Computation**, Cambridge, v. 9, n. 8, p. 1735-1780, 1997. DOI: 10.1162/neco.1997.9.8.1735.

KAHRA, D.; KAHRA, V. VSF-Net: A variational encoder and Sharpness-Aware attention framework for robust forecasting of non-stationary financial time series. **International Journal of Information Technology**, 2025. DOI: 10.1007/s41870-025-02933-y.

LAHBABI, H. *et al.* Stock price forecasting using ito-ADAM optimized long Short-Term memory neural networks. **Statistics, Optimization and Information Computing**, v. 14, p. 3282–3298, 2025. DOI: 10.19139/soic-2310-5070-2876.

LÓPEZ DE PRADO, Marcos. **Advances in financial machine learning**. Hoboken: Wiley, 2018.

SHABE, R.; ENGELBRECHT, A.; ANDERSON, K. Incremental reinforcement learning for portfolio optimisation. **Computers**, v. 14, n. 242, 2025. DOI: 10.3390/computers14070242.

TSAY, Ruey S. **Analysis of financial time series**. 2. ed. Hoboken: Wiley-Interscience, 2005.

ZANOTTO, Eduardo Luiz; HÖLBIG, Carlos Amaral. Previsão de preços de ações e ETF na bolsa de valores B3 aplicando técnicas de machine learning. **Revista Sítio Novo**, Palmas, v. 10, p. e1879, 2026. DOI: 10.47236/2594-7036.2026.v10.1879.

ZHANG, Cheng; SJARIF, Nilam Nur Amir; IBRAHIM, Roslina. Deep learning models for price forecasting of financial time series: a review of recent advancements: 2020–2022. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, v. 14, n. 1, e1519, 2024. DOI: 10.1002/widm.1519.

ZHANG, Guoqiang; PATUWO, B. Eddy; HU, Michael Y. Forecasting with artificial neural networks: the state of the art. **International Journal of Forecasting**, Amsterdam, v. 14, n. 1, p. 35-62, 1998. DOI: 10.1016/S0169-2070(97)00044-7.

Informações complementares

Descrição		Declaração
Financiamento		Não se aplica.
Aprovação ética		Não se aplica.
Conflito de interesses		Não há.
Disponibilidade dos dados de pesquisa subjacentes		O trabalho não é um <i>preprint</i> e o código dos experimentos computacionais usados na pesquisa está disponibilizado em: https://github.com/CaioSantdev/FinanceNeuralNetwork.git
Uso de Inteligência Artificial		Sim. Uso para checagem das referências bibliográficas, correção de texto e auxílio na tradução do resumo para inglês e espanhol.
CrediT	Caio Santos Silva	Funções: investigação, escrita – rascunho original.
	Rafael Lima Carvalho	Funções: supervisão, escrita – revisão e edição.

Avaliadores: Dr. Edson Leite Araújo* (Universidade Federal do Vale do São Francisco, Pernambuco, Brasil). O avaliador “A” optou pela avaliação fechada e pelo anonimato.

Revisor do texto em português: Marco Aurélio Mello.

Revisora do texto em inglês: Patrícia Luciano de Farias Teixeira Vidal.

Revisora do texto em espanhol: Graziani França Claudino de Anicézio.

Como citar (ABNT):

SILVA, Caio Santos; CARVALHO, Rafael Lima de. Impacto do horizonte histórico de dados no desempenho preditivo e financeiro de redes neurais do tipo LSTM: um estudo de caso com ativos bancários brasileiros. **Revista Sítio Novo**, Palmas, v. 10, p. e2228, 2026. DOI: 10.47236/2594-7036.2026.v10.2228. Disponível em:
<https://sitionovo.iftto.edu.br/index.php/sitionovo/article/view/2228>.

* Optou pela avaliação fechada e autorizou somente a divulgação da identidade como avaliador no trabalho publicado.