

Aplicação de Redes Neurais Recorrentes do tipo GRU na Previsão de Preços de Ações

Application of GRU-Type Recurrent Neural Networks in Stock Price Forecasting

Amanda de Sena Souza^{*}

Alberto Abrantes Gomes de Amorim^{**}

Gabriel Oliveira Salvador de Souza^{***}

Rita de Cassia Souza Paz^{****}

Resumo: A previsão de preços de ações é um desafio relevante devido à volatilidade, à não linearidade e à influência de diversos fatores econômicos sobre o mercado financeiro. Nesse contexto, as redes neurais artificiais, especialmente os modelos recorrentes, destacam-se pela capacidade de aprender padrões temporais complexos a partir de dados históricos. Este estudo teve como objetivo avaliar a aplicação de uma rede neural recorrente do tipo GRU (*Gated Recurrent Unit*) na previsão do preço de fechamento de ações. A hipótese central é que uma arquitetura GRU, associada a técnicas de controle de *overfitting*, é capaz de capturar padrões relevantes em séries temporais financeiras e produzir previsões de curto prazo com desempenho satisfatório. Foram utilizados dados históricos de aproximadamente 3 anos e 11 meses, de janeiro de 2021 a novembro de 2024, considerando como variáveis de entrada os preços de abertura, fechamento, máximo, mínimo e volume negociado. Os preços foram corrigidos pelo IPCA e organizados em janelas temporais de cinco dias para prever o preço de fechamento do dia seguinte. A base foi dividida em 63% para treinamento, 7% para validação e 30% para teste, respeitando a ordem temporal dos dados. O modelo apresentou MAE de 1,75, MSE de 7,18, RMSE de 2,68, MAPE de 4% e R^2 de 92%, indicando boa aderência entre valores reais e previstos. Os resultados sugerem que a GRU possui capacidade satisfatória de generalização e pode auxiliar na análise de tendências do mercado financeiro.

Palavras-chave: redes neurais artificiais; GRU; previsão de ações; séries temporais; mercado financeiro.

Abstract: Stock price forecasting is a relevant challenge due to volatility, nonlinearity, and the influence of several economic factors on financial markets. In this context, artificial neural networks, especially recurrent models, stand out for their ability to learn complex temporal patterns from historical data. This study aimed to evaluate the application of a recurrent neural network model based on GRU (*Gated Recurrent Unit*) to forecast stock closing prices. The central hypothesis is that a GRU architecture, combined with overfitting control techniques, is capable of capturing relevant patterns in financial time series and producing satisfactory short-term forecasts. Historical data covering approximately 3 years and 11 months, from January 2021 to

^{*}Graduanda em Ciências Econômicas, Universidade Federal Fluminense. ORCID: <https://orcid.org/0009-0006-2132-3812>. E-mail: amandasena@id.uff.br.

^{**}Graduando em Ciências Econômicas, Universidade Federal Fluminense. ORCID: <https://orcid.org/0009-0000-4584-0298>. E-mail: albertoaga@id.uff.br.

^{***}Graduando em Ciências Econômicas, Universidade Federal Fluminense. ORCID: <https://orcid.org/0009-0000-9053-3676>. E-mail: gsalvador@id.uff.br.

^{****}Doutora em Engenharia e Ciências dos Materiais, Departamento de Ciências Econômicas de Campos. Universidade Federal Fluminense. ORCID: <https://orcid.org/0000-0002-9564-1008>. E-mail: ritapaz@id.uff.br.

November 2024, were used, considering opening price, closing price, highest price, lowest price, and trading volume as input variables. Prices were adjusted by the Brazilian IPCA inflation index and organized into five-day time windows to predict the following day's closing price. The dataset was divided into 63% for training, 7% for validation, and 30% for testing, preserving the temporal order of the data. The model achieved an MAE of 1.75, MSE of 7.18, RMSE of 2.68, MAPE of 4%, and R^2 of 92%, indicating good agreement between actual and predicted values. The results suggest that the GRU model has satisfactory generalization capacity and may support trend analysis in financial markets.

Keywords: artificial neural networks; GRU; stock forecasting; time series; financial market.

1 Introdução

A previsão de preços de ações é uma das tarefas mais complexas no campo da análise financeira, pois envolve séries temporais caracterizadas por ruído, volatilidade, sazonalidades, choques externos e relações não lineares entre variáveis. A dinâmica dos mercados financeiros é influenciada por fatores microeconômicos, macroeconômicos, políticos, institucionais e comportamentais, o que torna difícil a construção de modelos preditivos baseados apenas em relações lineares ou em pressupostos tradicionais de estacionariedade (Hyndman; Athanasopoulos, 2021).

Nos últimos anos, o avanço das técnicas de inteligência artificial e aprendizado profundo ampliou as possibilidades de modelagem de séries temporais financeiras. Entre essas técnicas, destacam-se as redes neurais artificiais, que são modelos computacionais capazes de aprender padrões a partir dos dados. Em vez de serem programadas com regras fixas, essas redes ajustam seus parâmetros internos durante o processo de treinamento, buscando reduzir o erro entre os valores previstos e os valores observados (Goodfellow; Bengio; Courville, 2016).

As redes neurais recorrentes (RNN), em particular, são adequadas para problemas que envolvem dados sequenciais, pois permitem considerar informações passadas no processo de previsão. No caso de séries financeiras, essa característica é importante porque o preço de uma ação em determinado momento pode estar relacionado ao seu comportamento em períodos anteriores. Contudo, as redes recorrentes tradicionais apresentam limitações associadas ao

desaparecimento do gradiente, especialmente quando precisam lidar com dependências temporais mais longas (Hochreiter; Schmidhuber, 1997).

Para enfrentar esse problema, foram desenvolvidas arquiteturas mais avançadas, como LSTM (*Long Short-Term Memory*) e GRU (*Gated Recurrent Unit*). A rede LSTM, proposta por Hochreiter e Schmidhuber (1997), introduziu mecanismos de memória capazes de preservar informações relevantes por períodos mais longos. Posteriormente, Cho et al. (2014) propuseram a arquitetura GRU, que é uma rede recorrente que utiliza mecanismos de controle chamados portas, responsáveis por regular quais informações devem ser mantidas, atualizadas ou descartadas ao longo do tempo. Em comparação com a LSTM, a GRU possui estrutura mais simples, menor número de parâmetros e maior eficiência computacional, mantendo, em muitos casos, desempenho semelhante em tarefas de previsão temporal (Cho et al., 2014).

O presente estudo tem como propósito avaliar a aplicação de uma rede neural recorrente do tipo GRU na previsão de preços de ações, utilizando dados históricos de mercado. O problema investigado pode ser expresso pela seguinte questão: uma rede neural GRU, treinada com dados históricos de preços e volume negociado, é capaz de prever o preço de fechamento de ações no curto prazo com desempenho estatisticamente satisfatório?

A hipótese que orienta esta pesquisa é que uma arquitetura GRU, alimentada por variáveis históricas de mercado e treinada com técnicas de validação e controle de *overfitting*, apresenta capacidade de capturar padrões temporais relevantes e produzir previsões próximas aos valores reais observados.

O objetivo geral deste artigo é analisar o desempenho de uma rede neural GRU aplicada à previsão do preço de fechamento de ações. Como objetivos específicos, busca-se: apresentar os fundamentos teóricos das redes neurais artificiais e das redes recorrentes; descrever a arquitetura GRU empregada no estudo; detalhar os procedimentos metodológicos de coleta, preparação, divisão e treinamento dos dados; avaliar o desempenho do modelo por meio das métricas MAE (*Mean Absolute Error*), MSE (*Mean Squared Error*), RMSE (*Root*

Mean Square Error), MAPE (*Mean Absolute Percentage Error*) e R^2 (coeficiente de determinação); e discutir a aderência entre os valores reais e previstos.

Ressalta-se que o estudo não pretende afirmar a superioridade definitiva da GRU em relação aos modelos estatísticos tradicionais, mas avaliar sua viabilidade como alternativa de modelagem para séries financeiras. A comparação sistemática com modelos como ARIMA, suavização exponencial, regressões e outras técnicas de aprendizado de máquina é indicada como etapa de aprofundamento em pesquisas futuras.

2 Revisão de Literatura

As redes neurais artificiais são modelos matemáticos inspirados, de forma simplificada, no funcionamento do cérebro humano. Sua unidade básica é o neurônio artificial, que recebe valores de entrada, multiplica cada entrada por um peso, soma os resultados, adiciona um termo de viés e aplica uma função de ativação para produzir uma saída. Quando vários neurônios são organizados em camadas e conectados entre si, forma-se uma rede neural artificial capaz de modelar relações complexas entre variáveis (Goodfellow; Bengio; Courville, 2016).

A estrutura básica de uma rede neural é composta por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. A camada de entrada recebe os dados, as camadas ocultas realizam transformações matemáticas sucessivas e a camada de saída apresenta o resultado final. No processo de treinamento, a rede realiza previsões, compara essas previsões com os valores reais, calcula uma função de perda e ajusta seus pesos por meio de algoritmos de otimização (Goodfellow; Bengio; Courville, 2016).

No caso de séries temporais, as RNNs apresentam uma vantagem importante: elas incorporam a noção de memória. Isso significa que a previsão em determinado instante pode considerar informações de instantes anteriores. Essa característica torna as RNNs adequadas para problemas envolvendo preços de ações, séries econômicas, temperatura, vendas e demais dados organizados sequencialmente (Hyndman; Athanasopoulos, 2021).

Entretanto, as RNNs tradicionais podem apresentar dificuldade para preservar informações por muitos períodos, devido ao problema do desaparecimento do gradiente (*vanishing gradiente*). Quando o gradiente se torna muito pequeno durante o processo de retropropagação, a rede passa a ter dificuldade em aprender relações temporais mais longas. Para contornar essa limitação, foram desenvolvidas arquiteturas como a LSTM e a GRU (Hochreiter; Schmidhuber, 1997; Cho et al., 2014).

A LSTM utiliza três portas principais: porta de entrada, porta de esquecimento e porta de saída. Esses mecanismos regulam o fluxo de informação e permitem que a rede armazene ou descarte informações ao longo do tempo. A GRU, por sua vez, simplifica essa estrutura ao utilizar duas portas principais: a porta de atualização e a porta de redefinição. A porta de atualização controla quanto da informação passada será mantida, enquanto a porta de redefinição regula quanto da informação anterior será esquecida no cálculo do novo estado (Cho et al., 2014).

Essa simplificação faz com que a GRU possua menor número de parâmetros em comparação com a LSTM, o que tende a reduzir o custo computacional e acelerar o treinamento. Em muitos casos práticos, esse tipo de arquitetura apresenta bom equilíbrio entre desempenho preditivo e eficiência computacional, sobretudo em problemas envolvendo séries temporais (Cho et al., 2014; Goodfellow; Bengio; Courville, 2016).

Na literatura sobre previsão financeira, modelos baseados em aprendizado profundo têm sido utilizados para identificar padrões não lineares em séries temporais. Como os preços de ações não seguem necessariamente relações lineares estáveis, técnicas como LSTM e GRU podem oferecer ganhos em relação a métodos tradicionais, sobretudo quando o objetivo é capturar tendências, oscilações e dependências temporais de curto prazo (Hyndman; Athanasopoulos, 2021).

No entanto, o uso dessas técnicas exige cautela, pois redes neurais podem sofrer *overfitting*, especialmente quando aprendem excessivamente os dados de treinamento e perdem capacidade de generalização para dados novos.

Por isso, estratégias como validação, *dropout*, regularização e *early stopping* são fundamentais. O *dropout*, por exemplo, desativa aleatoriamente parte dos neurônios durante o treinamento, reduzindo a dependência excessiva de determinadas conexões e contribuindo para melhorar a generalização do modelo (Srivastava et al., 2014). Já o *early stopping* monitora o desempenho do modelo no conjunto de validação e interrompe o treinamento quando não há melhora após determinado número de épocas, evitando que a rede continue ajustando ruídos específicos da amostra de treinamento.

Além disso, algoritmos de otimização desempenham papel central no treinamento das redes neurais. O otimizador *Adam*, proposto por Kingma e Ba (2015), combina propriedades dos métodos *Momentum* e *RMSProp*, ajustando adaptativamente a taxa de aprendizado para cada parâmetro. Essa característica torna o *Adam* amplamente utilizado em aplicações de aprendizado profundo, pois favorece estabilidade e eficiência no processo de treinamento.

3 Metodologia

Esta pesquisa caracteriza-se como um estudo aplicado, de abordagem quantitativa e natureza experimental, voltado à avaliação de um modelo de rede neural recorrente do tipo GRU para previsão de preços de ações. O estudo utilizou dados históricos de empresas como Tesla, Apple, Vale e Petrobrás, considerando como variáveis de entrada os preços de abertura, fechamento, máximo, mínimo e o volume negociado. O período analisado compreendeu aproximadamente 3 anos e 11 meses, de janeiro de 2021 a novembro de 2024. Os preços das ações foram corrigidos pelo IPCA com o objetivo de transformar a série nominal em uma série real, descontando o efeito da inflação acumulada no período analisado. Esse procedimento permite maior comparabilidade temporal dos dados e evita que o modelo de rede neural aprenda padrões associados apenas à variação nominal dos preços. Como a previsão realizada é de curtíssimo prazo, utilizando cinco dias para prever o sexto, o impacto diário da correção inflacionária tende a ser limitado; entretanto, considerando uma série histórica de quase quatro anos, a correção contribui para uma interpretação econômica mais consistente dos resultados. Nesta etapa, não foi realizado teste

comparativo entre preços corrigidos e preços nominais, o que constitui uma possibilidade de aprimoramento metodológico.

Após a coleta e preparação dos dados, foram desenvolvidos *scripts* em *Python* para estruturação da base, treinamento do modelo e avaliação dos resultados. A variável-alvo do modelo foi o preço de fechamento do dia seguinte. Para isso, foram utilizadas janelas temporais de cinco dias, isto é, a rede recebia os dados dos cinco dias anteriores para prever o preço de fechamento do sexto dia. A escolha dessa janela se justifica por representar aproximadamente uma semana útil de negociação, permitindo que o modelo utilize informações recentes do mercado para uma previsão de curto prazo. Contudo, não foram testadas sistematicamente janelas maiores, como 10, 20 ou 30 dias, o que poderia indicar se horizontes temporais mais longos melhorariam ou prejudicariam o desempenho do modelo.

A divisão dos dados respeitou a natureza temporal da série. Inicialmente, a base foi separada em 70% para treinamento e 30% para teste. Em seguida, 10% do conjunto de treinamento foi reservado para validação, resultando em uma divisão efetiva de 63% dos dados para treinamento, 7% para validação e 30% para teste. Essa separação foi realizada sem embaralhamento dos dados, uma vez que, em séries temporais, a mistura aleatória entre passado e futuro pode gerar vazamento de informação e comprometer a validade da avaliação do modelo (Hyndman; Athanasopoulos, 2021).

A arquitetura da rede foi composta por uma camada de entrada com cinco variáveis, correspondentes aos dados de preço e volume; uma primeira camada GRU com 64 neurônios e *dropout* de 20%; uma segunda camada GRU com 32 neurônios e *dropout* de 20%; uma camada densa com 16 neurônios; e uma camada final com um neurônio responsável pela previsão do preço de fechamento do dia seguinte. As camadas GRU utilizaram a função de ativação tangente hiperbólica, enquanto a camada densa utilizou a função ReLU. A configuração com duas camadas GRU foi definida empiricamente, a partir de uma arquitetura intermediária, suficientemente complexa para capturar padrões temporais, mas sem excesso de parâmetros que pudesse ampliar o risco de

overfitting. A escolha da GRU, em vez da LSTM, foi motivada pelo horizonte de previsão de curtíssimo prazo e pela janela temporal reduzida, situação em que uma arquitetura com menor número de parâmetros tende a oferecer bom equilíbrio entre desempenho e custo computacional. Nesta etapa, não foi aplicada uma técnica automática de otimização de hiperparâmetros, como *Grid Search*, *Random Search* ou otimização bayesiana.

O treinamento foi configurado com taxa de aprendizado de 0,1%, *batch size* igual a 16 e máximo de 100 épocas. A função de perda adotada foi o erro quadrático médio, MSE. O otimizador utilizado foi o *Adam*, escolhido por sua capacidade de adaptar automaticamente o tamanho dos ajustes nos pesos da rede, combinando características dos métodos *Momentum* e *RMSProp* (Kingma; Ba, 2015). A taxa de aprendizado de 0,001 foi adotada por ser amplamente utilizada com o *Adam* e por ter apresentado convergência estável da função de perda. Entretanto, não foi realizada uma busca sistemática com valores alternativos, como 0,01, 0,0005 ou 0,0001. O *dropout* de 20% foi empregado como valor moderado de regularização; percentuais mais altos, como 30% ou 40%, poderiam reduzir ainda mais o *overfitting*, mas também poderiam limitar a capacidade do modelo de capturar oscilações relevantes nas séries financeiras.

Também foi utilizado *EarlyStopping* com paciência de 20 épocas, visando interromper o treinamento quando não houvesse melhora relevante no desempenho de validação. Nesse procedimento, as 100 épocas funcionam como limite máximo de treinamento, mas o algoritmo pode encerrar o processo antes caso a perda de validação deixe de melhorar por 20 épocas consecutivas, restaurando os melhores pesos obtidos durante o treinamento. Essa estratégia é coerente com a necessidade de reduzir o risco de *overfitting* em modelos de aprendizado profundo, especialmente em séries temporais financeiras (Goodfellow; Bengio; Courville, 2016).

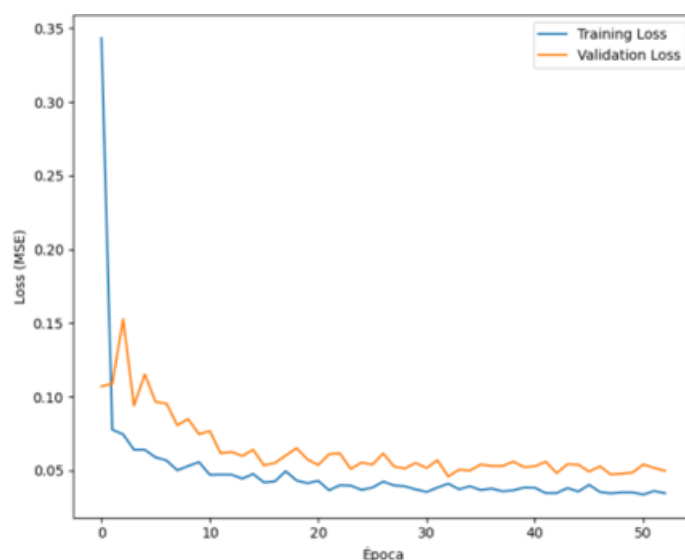
A avaliação do modelo foi realizada por meio das métricas MAE, MSE, RMSE, MAPE e R^2 . O MAE mede o erro absoluto médio entre valores previstos e reais. O MSE calcula a média dos erros elevados ao quadrado, penalizando erros maiores. O RMSE corresponde à raiz quadrada do MSE, expressando o

erro na mesma unidade da variável analisada. O MAPE mede o erro absoluto médio em termos percentuais, permitindo avaliar a magnitude relativa do erro em relação ao valor real da ação, o que é especialmente importante quando se comparam ativos com faixas de preço distintas. O R^2 indica a proporção da variação dos dados reais explicada pelo modelo, sendo que valores mais próximos de 1 indicam melhor ajuste (Hyndman; Athanasopoulos, 2021).

4 Resultados e Discussões

O treinamento da rede GRU apresentou comportamento considerado estável. De acordo com o gráfico de evolução da função de perda (*loss*), apresentado na Figura 1, a *loss* de treinamento diminuiu ao longo das épocas, enquanto a *loss* de validação permaneceu relativamente estável. Esse comportamento sugere que o modelo não apresentou sinais fortes de *overfitting*, pois não houve divergência acentuada entre o erro de treinamento e o erro de validação.

Figura 1 – Evolução da *loss* durante o treinamento e validação.



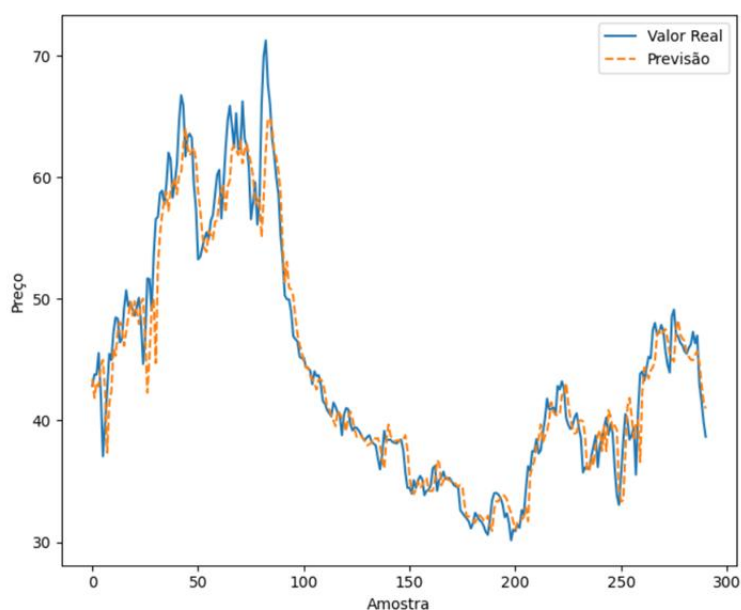
Fonte: Elaborado pelos autores.

As métricas obtidas foram: MAE de 1,75, MSE de 7,18, RMSE de 2,68, MAPE de 4% e R^2 de 92%. Esses resultados indicam que o modelo apresentou bom desempenho preditivo no conjunto avaliado. O MAE de 1,75 significa que, em média, as previsões diferiram dos valores reais em aproximadamente 1,75 unidade monetária. O RMSE de 2,68, por ser mais sensível a erros elevados,

indica a existência de alguns desvios maiores, possivelmente associados a períodos de maior volatilidade da série. O MAPE de 4% indica que o erro percentual médio absoluto foi reduzido, o que sugere bom desempenho preditivo para uma série financeira caracterizada por volatilidade.

A diferença entre MAE e RMSE é relevante para a interpretação dos resultados. Quando o RMSE é maior que o MAE, há indício de que alguns erros pontuais foram mais expressivos, uma vez que o RMSE penaliza de forma mais intensa os desvios maiores. Em séries de ações, esse comportamento é esperado, pois movimentos bruscos de mercado, notícias corporativas, alterações macroeconômicas e choques externos podem provocar oscilações abruptas nos preços (Hyndman; Athanasopoulos, 2021). Observando a comparação entre previsão e valor real (gráfico da Figura 2), os maiores desvios tendem a ocorrer justamente em momentos de virada repentina de tendência ou de maior volatilidade, quando o comportamento futuro deixa de seguir o padrão imediatamente anterior.

Figura 2 – Comparação valores previstos vs valores reais



Fonte: Elaborado pelos autores.

O coeficiente de determinação R^2 de 92% indica que o modelo conseguiu explicar parcela expressiva da variação observada nos dados. Embora o R^2 não deva ser interpretado isoladamente como garantia de capacidade preditiva perfeita, seu valor elevado, associado aos baixos erros médios, sugere que a rede GRU capturou padrões importantes da série analisada.

A análise visual entre os valores reais e previstos reforça essa interpretação. O gráfico da Figura 2 mostra que a linha prevista (em laranja) acompanha de forma consistente a linha dos valores reais (em azul), especialmente em relação à tendência geral da série. As maiores diferenças ocorrem em momentos de maior volatilidade, o que é coerente com a natureza instável do mercado financeiro.

Esses resultados sugerem que a rede GRU foi capaz de aprender padrões temporais relevantes a partir dos dados históricos. A utilização de janelas de cinco dias permitiu ao modelo considerar informações recentes do comportamento da ação para prever o preço de fechamento do dia seguinte. Além disso, o uso de *dropout*, validação e *early stopping* contribuiu para reduzir o risco de *overfitting* e melhorar a capacidade de generalização, conforme recomendado na literatura sobre aprendizado profundo (Srivastava et al., 2014; Goodfellow; Bengio; Courville, 2016). Ainda assim, a estabilização da curva de validação e o fato de o gráfico de treinamento não se estender até as 100 épocas planejadas sugerem que o *EarlyStopping* pode ter interrompido o treinamento antes do limite máximo, preservando a configuração de pesos associada ao menor erro de validação.

Apesar dos resultados positivos, é importante destacar algumas limitações. Em primeiro lugar, séries financeiras são altamente sensíveis a eventos externos que podem não estar representados apenas nos dados históricos de preço e volume. Notícias, decisões de política monetária, crises geopolíticas, resultados corporativos e mudanças regulatórias podem alterar rapidamente o comportamento dos preços. Em segundo lugar, o desempenho obtido em determinado período não garante que o modelo manterá a mesma qualidade preditiva em outros contextos de mercado, especialmente se o período

de teste conter eventos atípicos. Além disso, a arquitetura foi definida empiricamente, sem comparação sistemática com diferentes taxas de aprendizado, percentuais de *dropout*, tamanhos de janela, preços nominais e corrigidos, modelos estatísticos tradicionais ou outras arquiteturas de aprendizado profundo. Por essa razão, modelos de previsão financeira devem ser constantemente reavaliados e recalibrados (Hyndman; Athanasopoulos, 2021).

Outro ponto relevante é que o modelo não deve ser interpretado como instrumento isolado de decisão de investimento. Embora a GRU tenha apresentado boa aderência aos dados e capacidade de acompanhar tendências, previsões de preços de ações envolvem incerteza. Assim, seus resultados devem ser utilizados como apoio à análise, combinados com indicadores fundamentalistas, técnicos, macroeconômicos e critérios de gestão de risco. Do ponto de vista computacional, uma rede GRU com duas camadas e janelas curtas é relativamente leve; portanto, a atualização diária com novos dados pode ser viável em computadores comuns ou em ambientes como *Google Colab*. A maior limitação para uso prático não está apenas no custo computacional, mas na qualidade dos dados, na atualização contínua, no monitoramento do desempenho e na interpretação adequada das previsões.

De forma geral, os resultados indicam que redes neurais recorrentes do tipo GRU constituem uma alternativa promissora para previsão de curto prazo em séries financeiras, especialmente quando o objetivo é capturar tendências e padrões não lineares. A boa performance observada, expressa pelo R^2 superior a 90% e pelos baixos valores de erro médio, reforça a pertinência do uso dessa arquitetura em aplicações de análise financeira automatizada (Cho et al., 2014).

5 Considerações Finais

Este artigo analisou a aplicação de uma rede neural recorrente do tipo GRU na previsão de preços de ações, utilizando dados históricos de mercado como preços de abertura, fechamento, máximo, mínimo e volume negociado. O estudo partiu da hipótese de que a GRU seria capaz de capturar padrões

temporais relevantes e produzir previsões de curto prazo com desempenho satisfatório.

Os resultados obtidos confirmam a hipótese proposta. O modelo apresentou MAE de 1,75, MSE de 7,18, RMSE de 2,68, MAPE de 4% e R^2 de 92%, indicando boa capacidade de ajuste e aderência entre valores reais e previstos. A análise da curva de perda durante o treinamento e validação também indicou estabilidade, sem evidências expressivas de *overfitting*. Além disso, a comparação gráfica entre os valores reais e previstos mostrou que a rede acompanhou adequadamente a tendência da série, embora tenha apresentado diferenças em momentos de maior volatilidade.

A arquitetura GRU mostrou-se eficiente para lidar com dados sequenciais e adequada à previsão de curto prazo em séries financeiras. Sua estrutura mais simples em relação à LSTM, associada ao menor custo computacional, representa uma vantagem relevante em aplicações práticas que demandam treinamento mais rápido e boa capacidade de generalização, corroborando os resultados de Cho et al. (2014).

Entretanto, os resultados devem ser interpretados com cautela. A previsão de preços de ações envolve incertezas significativas e não depende apenas de padrões históricos. Assim, recomenda-se que trabalhos futuros ampliem o conjunto de variáveis explicativas, incorporando indicadores macroeconômicos, indicadores técnicos, dados fundamentalistas e variáveis de sentimento de mercado. Também seria relevante comparar o desempenho da GRU com outros modelos, como LSTM, ARIMA/SARIMA, *Random Forest*, *XGBoost* e *Transformers*, de modo a verificar se a arquitetura recorrente apresenta superioridade estatística em diferentes cenários. Além disso, recomenda-se testar preços nominais e corrigidos pelo IPCA, diferentes janelas temporais, taxas de aprendizado e percentuais de *dropout*.

Conclui-se que a rede neural GRU apresentou desempenho satisfatório e pode ser utilizada como ferramenta de apoio à análise de séries temporais financeiras, especialmente para identificação de tendências de curto prazo. Seu uso, contudo, deve estar associado a uma estratégia analítica mais ampla, que

considere o risco, a volatilidade e as múltiplas dimensões que influenciam o comportamento do mercado de ações.

6 Agradecimentos

Agradecemos à Universidade Federal Fluminense – UFF pelo apoio financeiro concedido por meio da bolsa PIBINOVA/UFF para o desenvolvimento desta pesquisa.

7 Referências

CHO, Kyunghyun et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: **Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing**. Doha: Association for Computational Linguistics, 2014.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. Cambridge: MIT Press, 2016.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. **Neural Computation**, v. 9, n. 8, p. 1735-1780, 1997.

HYNDMAN, Rob J.; ATHANASOPOULOS, George. **Forecasting: Principles and Practice**. 3. ed. Melbourne: OTexts, 2021.

KINGMA, Diederik P.; BA, Jimmy. Adam: a method for stochastic optimization. In: **International Conference on Learning Representations — ICLR**, 2015.

SRIVASTAVA, Nitish et al. Dropout: a simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, p. 1929-1958, 2014.

Recebido em: 31/mai/2026.

Aprovado em: 10/jun/2026.