

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA - FEELT
GRADUAÇÃO EM ENGENHARIA BIOMÉDICA

JOÃO PAULO AMADO CARRIJO BORGES

Análise de preços de ações bancárias no Brasil: uma abordagem com redes neurais LSTM

Uberlândia

2023

JOÃO PAULO AMADO CARRIJO BORGES

Análise de preços de ações bancárias no Brasil: uma abordagem com redes neurais LSTM

Trabalho apresentado como requisito parcial de avaliação na disciplina Trabalho de Conclusão de Curso de Engenharia Biomédica da Universidade Federal de Uberlândia.

Orientador: Flávio Luiz de Moraes Barboza

Uberlândia

2023

AGRADECIMENTOS

Primeiramente gostaria de expressar minha gratidão à minha família, cujo suporte, amor e motivação foram peças fundamentais para que eu perseguisse meus objetivos acadêmicos. Agradeço também aos meus amigos pelo companheirismo e troca de experiências durante a faculdade, vocês foram essenciais para me manter motivado e focado durante toda a jornada.

Agradeço a Universidade Federal de Uberlândia que proporcionou ambientes e recursos propícios para o crescimento intelectual, foi um privilégio estudar nessa instituição.

Aos professores do curso, agradeço pela dedicação, sabedoria e conhecimento compartilhados, peça essencial para meu crescimento profissional. Um agradecimento especial ao meu orientador Flávio Luiz de Moraes Barboza que desde o início prestou toda a assistência e atenção necessárias para que eu pudesse realizar o trabalho da melhor forma possível, usando de sua competência para fornecer sugestões e informações importantes para o desenvolvimento do mesmo.

RESUMO

A imprevisibilidade das flutuações nos preços das ações decorre de uma intrincada interação de diversos fatores nos bastidores. O desafio reside na compilação de informações multifacetadas, consolidando-as em um conjunto único e na construção de um modelo confiável para previsões precisas. Um modelo matemático bem desenvolvido, acompanhado por um conjunto ideal de atributos, pode eficientemente prever os preços das ações e proporcionar uma visão mais clara da situação de mercado. A abordagem proposta utiliza dados históricos disponíveis de uma ação e gera previsões sobre um ativo específico, considerando características como preço de abertura, máxima do dia, mínima do dia anterior, preço de fechamento, data de negociação, quantidade total de comércio e volume de negócios. O modelo sugerido emprega a análise de séries temporais para prever o preço das ações ao longo de um período, utilizando uma abordagem mista, que combina aspectos qualitativos e quantitativos, baseada na estratégia de estudo de caso. Optou-se por analisar os bancos Itaú, Banco do Brasil, BTG, Bradesco e Santander, devido à relevância e representatividade dessas instituições no cenário financeiro. Sendo assim, analisando os cenários de ações os resultados indicam que a inclusão dos valores dos índices BVSP, IFNC e ISE como variáveis de entrada na rede neural LSTM construída não resultou em uma melhoria significativa de desempenho, exceto para o caso do Banco do Brasil. Este resultado evidencia a complexidade na modelagem das relações entre os índices e as flutuações nas ações dos bancos brasileiros.

Palavras-chave: Mercado de ações; redes neurais; LSTM.

ABSTRACT

The unpredictability of stock price fluctuations arises from an intricate interplay of several behind-the-scenes factors. The challenge lies in compiling multifaceted information, consolidating it into a single set, and building a reliable model for accurate predictions. A well-developed mathematical model, accompanied by an ideal set of attributes, can efficiently predict stock prices and provide a clearer view of the market situation. The proposed approach uses available historical data for a stock and generates predictions about a specific asset, considering characteristics such as opening price, high of the day, low of the previous day, closing price, trading date, total trade quantity and trading volume. The suggested model employs time series analysis to predict stock prices over a period, using a mixed approach that combines qualitative and quantitative aspects, based on the case study strategy. We chose to analyze the banks Itaú, Banco do Brasil, BTG, Bradesco and Santander, due to the relevance and representativeness of these institutions in the financial scenario. Therefore, analyzing the stock scenarios, the results indicate that the inclusion of the values of the BVSP, IFNC and ISE indexes as inputs variables in the constructed LSTM neural network did not result in a significant improvement in performance, except for the case of Banco do Brasil. This result highlights the complexity in modeling the relationships between indices and fluctuations in Brazilian bank shares.

Keywords: Stock market; neural networks; LSTM.

LISTA DE ILUSTRAÇÕES

Figura 1 - LSTM aplicada as ações do Banco do Brasil.	24
Figura 2 - Gráfico de perda por época para a ação do Banco do Brasil.	25
Figura 3 - Dados reais, Random Walk e valores previstos pela rede LSTM.	26
Figura 4 - Rede neural LSTM Teste sem ISE.	27
Figura 5 - Dados reais, Random Walk e valores previstos pela rede LSTM sem ISE.	27
Figura 6 - Rede neural LSTM Teste sem IFNC.	28
Figura 7 - Dados reais, Random Walk e valores previstos pela rede LSTM sem INFC.	28
Figura 8 - Rede neural LSTM Teste sem BVSP.	29
Figura 9 - Dados reais, Random Walk e valores previstos pela rede LSTM sem BVSP.	29
Figura 10 - Rede neural LSTM Teste sem Índices.	30
Figura 11 - Dados reais, Random Walk e valores previstos pela rede LSTM sem índices. ...	30
Figura 12 - LSTM aplicada as ações do Banco Bradesco.	32
Figura 13 - Gráfico de perda por época para ações do Banco Bradesco.	33
Figura 14 - Dados reais, Random Walk e valores previstos pela rede LSTM.	33
Figura 15 - Rede neural LSTM Teste sem ISE.	34
Figura 16 - Dados reais, Random Walk e valores previstos pela rede LSTM sem ISE.	34
Figura 17 - Rede neural LSTM Teste sem IFNC.	35
Figura 18 - Dados reais, Random Walk e valores previstos pela rede LSTM sem INFC.	35
Figura 19 - Rede neural LSTM Teste sem BVSP.	36
Figura 20 - Dados reais, Random Walk e valores previstos pela rede LSTM sem BVSP.	36
Figura 21 - Rede neural LSTM Teste sem Índices.	37
Figura 22 - Dados reais, Random Walk e valores previstos pela rede LSTM sem índices. ...	37
Figura 23 - LSTM aplicada as ações do Banco BTG Pactual.	38
Figura 24 - Gráfico de perda por época para ações do Banco BTG Pactual.	39
Figura 25 - Dados reais, Random Walk e valores previstos pela rede LSTM.	40
Figura 26 - Rede neural LSTM Teste sem ISE.	40
Figura 27 - Dados reais, Random Walk e valores previstos pela rede LSTM sem ISE.	41
Figura 28 - Rede neural LSTM Teste sem IFNC.	41
Figura 29 - Dados reais, Random Walk e valores previstos pela rede LSTM sem INFC.	42
Figura 30 - Rede neural LSTM Teste sem BVSP.	42
Figura 31 - Dados reais, Random Walk e valores previstos pela rede LSTM sem BVSP.	43

Figura 32 - Rede neural LSTM Teste sem Índices.....	43
Figura 33 - Dados reais, Random Walk e valores previstos pela rede LTSM sem índices. ...	44
Figura 34 - LTSM aplicada as ações do Banco Itaú.....	45
Figura 35 - Gráfico de perda por época para ações do Banco Itaú.....	46
Figura 36 - Dados reais, Random Walk e valores previstos pela rede LTSM.	46
Figura 37 - Rede neural LSTM Teste sem ISE.	47
Figura 38 - Dados reais, Random Walk e valores previstos pela rede LTSM sem ISE.....	47
Figura 39 - Rede neural LSTM Teste sem IFNC.	48
Figura 40 - Dados reais, Random Walk e valores previstos pela rede LTSM sem INFC.....	48
Figura 41 - Rede neural LSTM Teste sem BVSP.	49
Figura 42 - Dados reais, Random Walk e valores previstos pela rede LTSM sem BVSP.....	49
Figura 43 - Rede neural LSTM Teste sem Índices.....	50
Figura 44 - Dados reais, Random Walk e valores previstos pela rede LTSM sem índices. ...	50
Figura 45 - LTSM aplicada as ações do Banco Santander.....	51
Figura 46 - Gráfico de perda por época para ações do Banco Santander.....	52
Figura 47 - Dados reais, Random Walk e valores previstos pela rede LSTM.	52
Figura 48 - Rede neural LSTM Teste sem ISE.	53
Figura 49 - Dados reais, Random Walk e valores previstos pela rede LSTM sem ISE.....	53
Figura 50 - Rede neural LSTM Teste sem IFNC.	54
Figura 51 - Dados reais, Random Walk e valores previstos pela rede LSTM sem IFNC.....	54
Figura 52 - Rede neural LSTM Teste sem BVSP.	55
Figura 53 - Dados reais, Random Walk e valores previstos pela rede sem BVSP.	56
Figura 54 - Rede neural LSTM Teste sem Índices.....	57
Figura 55 - Dados reais, Random Walk e valores previstos pela rede LTSM sem índices. ...	57

LISTA DE ABREVIATURAS E SIGLAS

ANBIMA	Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais
ARIMA	Média Móvel Integrada Auto-Regressiva
BVSP	Ibovespa
IFNC	Índice Financeiro
ISE	Índice de Sustentabilidade Empresarial
LSTM	Long Short Term Memory
MAE	<i>Mean Absolute Error</i>
MAPE	Erro Percentual Médio Absoluto
RMSE	<i>Mean Squared Error</i>
RNA	Rede Neural Artificial
RNN	Recurrent Neural Network
SVM	Support Vector Machine

SUMÁRIO

1	INTRODUÇÃO.....	12
2	FUNDAMENTAÇÃO TEÓRICA	14
3	METODOLOGIA.....	22
4	RESULTADOS E DISCUSSÃO	23
5	CONCLUSÃO.....	60

1 INTRODUÇÃO

De acordo com a Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais (ANBIMA), no ano de 2022, o percentual de investidores no Brasil teve um aumento de 5% em relação ao ano anterior, representando um acréscimo de 8 milhões de brasileiros. No âmbito do mercado de ações, companhias de capital aberto transacionam partes de seus ativos, experimentando flutuações diárias em seus valores, além de alterações na liquidez ao longo de um determinado período. Nesse cenário, investidores buscam ativamente por títulos que demonstrem desempenho superior à média do mercado a longo prazo, especialmente durante fases sensíveis da economia. Isso evidencia a importância de compreender os critérios de análise e as métricas empregadas por analistas na identificação das ações com maior potencial de retorno no mercado (IZIDORO,2015).

Prever os preços das ações é uma tarefa complexa e desafiadora para empresas, investidores e *traders* que buscam antecipar retornos futuros. Os mercados de ações são inerentemente ruidosos, não lineares e caoticamente determinísticos o que representa um desafio para otimizar a previsão dos preços futuros. A seleção de características a partir dos dados financeiros também se revela uma tarefa árdua na previsão de ações. Ainda, pode-se mencionar que o desempenho do modelo preditivo pode ser afetado pelo uso de recursos limitados ou pela inclusão excessiva de características, levando a modelos complexos e de difícil interpretação, com possível piora devido à colinearidade entre variáveis.

Um modelo bem desenvolvido, com um conjunto ideal de atributos, pode prever de forma eficaz os preços das ações e oferecer uma visão mais clara da situação do mercado. Diversas pesquisas foram realizadas para estudar a correlação entre variáveis específicas e o comportamento dos preços das ações, embora o sucesso desses modelos varie consideravelmente em termos de precisão e robustez. A seleção criteriosa de variáveis é crucial para o desempenho do modelo, sendo uma contribuição significativa deste estudo.

Devido à natureza volátil do mercado de ações, faz-se necessário uma análise extensiva, fundamentada em dados históricos. Algoritmos anteriores para previsão de tendências nas ações utilizam dados históricos de séries temporais. Os métodos convencionais de previsão de preços de ações concentram-se na análise estatística desses dados.

1.1 Objetivo geral

Desenvolver um programa capaz de prever o preço de ações de bancos brasileiros considerando dados históricos qualitativos e quantitativos. A abordagem proposta utiliza dados históricos disponíveis da ação e fornece previsões sobre o valor da ação, considerando características como preço de abertura, alta do dia, baixa do dia anterior, preço de fechamento, data de negociação, quantidade total de comércio e volume de negócios. O modelo proposto emprega a análise de séries temporais para prever o preço das ações ao longo de um período.

1.2 Objetivos específicos

- Testar a influência do preço dos índices ao longo do tempo como variáveis de entrada na rede neural;
- Verificar a performance da rede neural modelada com o objetivo de prever o preço das ações de cinco grandes bancos brasileiros.

2 FUNDAMENTAÇÃO TEÓRICA

Os seres humanos começam um pensamento do zero a cada segundo, as redes neurais tradicionais não conseguem fazer isso, e parece ser uma grande limitação. Por exemplo, imagine que você queira classificar que tipo de evento está ocorrendo a cada ponto em um filme. Não está claro como uma rede neural tradicional poderia usar seu raciocínio sobre eventos anteriores no filme para informar os posteriores. As redes neurais recorrentes abordam essa questão. São redes com laços, permitindo que a informação persista (PATTERSON, 2017).

O setor bancário desempenha um papel crucial no panorama econômico brasileiro, exercendo uma influência substancial no mercado financeiro. Ao longo das últimas décadas, os bancos brasileiros têm enfrentado desafios e crises, mas, notavelmente, demonstraram resiliência e capacidade de adaptação. A estabilidade do setor bancário nacional, mesmo em face de turbulências econômicas globais, tem suscitado um interesse crescente entre pesquisadores e profissionais do mercado financeiro. A complexidade das interações entre os bancos, o ambiente regulatório e as condições macroeconômicas no Brasil oferecem uma área fértil para investigações mais aprofundadas (PEREIRA, 2013).

Com a crescente globalização e digitalização dos serviços financeiros, a compreensão dos fatores que influenciam a dinâmica dos bancos no contexto brasileiro torna-se fundamental para orientar estratégias eficazes, tanto para investidores quanto para formuladores de políticas econômicas. Nesse cenário, a análise científica do papel dos bancos no mercado financeiro brasileiro contribui significativamente para o desenvolvimento de abordagens mais informadas e sustentáveis para o setor (PEREIRA, 2013).

Os guias no mercado de ações desempenham um papel vital ao oferecer orientação aos novos investidores, mas mesmo esses especialistas estão sujeitos a equívocos, muitas vezes decorrentes de crises imprevistas. Por essa razão, os investidores devem conduzir uma análise criteriosa ao decidir em quais ações investir. As ações no setor bancário são reconhecidas como uma das classes mais lucrativas na bolsa de valores, apesar das adversidades econômicas globais que levaram à falência de vários bancos internacionais. O setor bancário brasileiro demonstrou resiliência, apresentando resultados financeiros expressivos e consolidando-se como uma opção sólida. A análise fundamentalista, baseada no pressuposto de um atraso que afeta os preços das ações, acrescenta uma perspectiva importante para os investidores ao avaliar oportunidades no mercado. Em última análise, a relevância desse tema transcende os âmbitos acadêmico, científico e profissional, justificando plenamente a investigação contínua nesse campo (MODRO; SANTOS, 2015).

A teoria do passeio aleatório desafia a hipótese de mercado eficiente, a qual sugere que os preços dos ativos financeiros refletem todas as informações disponíveis. A origem dessa ideia remonta a trabalhos anteriores, e sua aplicação em modelos teóricos e estudos empíricos na área de finanças tem sido amplamente discutida. A hipótese de mercado eficiente, em particular na forma fraca, implica que os preços seguem um passeio aleatório, onde o conhecimento dos preços e retornos passados não oferece vantagem na previsão do preço futuro. Esta visão levou ao teste do modelo de passeio aleatório, uma análise crítica necessário para avaliar a eficiência do mercado (KIM; SHAMSUDDING, 2008).

A literatura em finanças empíricas tem explorado intensivamente o teste do modelo de passeio aleatório, crucial para verificar a hipótese de mercado eficiente na forma fraca. A análise se concentra na imprevisibilidade dos retornos das ações com base apenas em informações de preços passados, bem como na associação linear entre a variância do retorno e o período de análise. A evolução desses estudos revela uma mudança de perspectiva ao longo do tempo, desde a primeira geração, que apontou para a correlação serial nula na mudança de preço, até a segunda geração, que, utilizando metodologias mais avançadas, rejeitou a hipótese de passeio aleatório em alguns casos (BRUNI, 2004).

A natureza volátil do mercado de ações o torna uma área que requer uma abundância de análise com base em dados antigos. Os algoritmos de previsão de tendências de ações anteriores utilizam dados históricos em séries temporais. Os procedimentos típicos de previsão de preços de ações cientificamente orientados concentram-se na análise estatística dos dados de ações (LOKE, 2017). A previsão da bolsa de valores tornou-se uma área de interesse aguçado. A avaliação específica é uma delas, mas não fornece resultados específicos de forma confiável, sendo essencial desenvolver estratégias para previsões mais precisas de maneira progressiva (PRAMOD, 2020).

No passado, vários pesquisadores propuseram abordagens inovadoras para prever os preços das ações. Pesquisadores têm utilizado regressão linear, Máquinas de Vetores de Suporte (SVM), ARIMA (Média Móvel Integrada Auto-Regressiva) e técnicas tradicionais de aprendizado de máquina para a previsão de preços de ações. As limitações dessas abordagens são a falta de precisão. Há uma necessidade de uma abordagem melhor que possa prever variações significativas nos preços com precisão (WANH et al., 2005).

Métodos convencionais para lidar com a análise do mercado financeiro e a previsão de preços das ações incluem uma revisão significativa do desempenho passado das ações e da credibilidade geral da própria empresa, bem como uma análise estatística que se concentra exclusivamente no cálculo e na identificação de padrões nos preços das ações. Além disso, uma

análise estatística é realizada com foco exclusivo no cálculo e na identificação de padrões nos valores das ações (KIM; HA, 2019).

Os preços das ações estão correlacionados pela natureza do mercado; portanto, será difícil prevê-los. O algoritmo de previsão utiliza dados de mercado para prever o preço das ações usando técnicas de aprendizado de máquina, como a rede neural recorrente conhecida como Long Short Term Memory (LSTM). Nesse processo, os pesos são ajustados para cada ponto de dados usando o método do gradiente estocástico. Este sistema fornecerá resultados precisos em comparação com os algoritmos de previsão de preços de ações atualmente disponíveis. A rede é treinada e avaliada com vários tamanhos de dados de entrada para obter resultados gráficos (PRAMOD, 2020).

Um classificador de aprendizado supervisionado pode ser utilizado para prever o movimento dos preços das ações com base em dados de índices financeiros e determinar sua capacidade. No mercado financeiro, abordagens computacionais analíticas têm sido utilizadas para modelagem de portfólio. Uma discussão sobre a metodologia estatística de IA aborda que o uso da metodologia SVM (Support Vector Machine) demonstra que metodologias táticas podem ser aplicadas para prever os preços das ações (DEEPAK; MALATHI, 2017).

Outra questão significativa com as Redes Neurais Artificiais básicas para previsão de ações é o fenômeno de explosão de gradientes, onde os pesos de uma rede muito grande se tornam excessivamente grandes ou muito pequenos (respectivamente), retardando drasticamente sua convergência para o valor ideal. Isso é frequentemente causado por dois fatores: os pesos são inicializados aleatoriamente e os pesos mais próximos ao final da rede tendem a mudar significativamente mais do que aqueles no início (KIM; HA, 2019).

A natureza volátil do mercado de ações é uma área que demanda extensa análise com base em dados históricos. Algoritmos tradicionais de previsão de tendências de ações utilizam dados históricos em séries temporais, e os procedimentos tradicionais de previsão técnica para os preços das ações são fundamentados na análise estatística de dados. Muitas pesquisas estão sendo conduzidas para encontrar um modelo preciso para prever os preços das ações, e ainda não há uma solução universal disponível para aplicação. Portanto, os dados históricos de uma ação serão considerados para a previsão do preço das ações (LOKE, 2017).

A aprendizagem profunda é uma técnica muito popular para a previsão do mercado de ações. Vários índices do mercado de ações, como S&P 500, Dow Jones Industrial Average, NASDAQ, NYSE, BSE, são mencionados como os mais comumente usados pelos pesquisadores além de que diversas métricas de avaliação, MAE, RMSE, Precision, Recall, F1

Score, são discutidas. As soluções existentes de aprendizado profundo estão focadas em classificação ou regressão para a previsão de preços de ações (FENG et al., 2019).

Abordagens tradicionais, como Regressão Linear e Regressão de Vetores de Suporte, foram utilizadas, mas a precisão não foi adequada, pois essas soluções não são otimizadas para o objetivo de investimento (BATHLA, 2020).

Modelos estatísticos tais como ARIMA são utilizados para prever ações futuras em que erro percentual médio é aprimorado, enquanto o SVR é usado para prever o movimento dos preços das ações, diferentes técnicas são aplicadas para pré-processar os dados. É mencionado no estudo de BATHLA (2020) que se os parâmetros forem determinados eficientemente, os preços podem ser previstos com precisão adequada. O Erro Percentual Médio Absoluto (MAPE) é utilizado como métrica de avaliação, já o SVR, SVR híbrido-ANN e SVR-RF são aplicados para a previsão do mercado de ações. Observou-se que o SVR possui propriedades semelhantes às do SVM, pois o objetivo é reduzir o erro entre os valores reais e previstos. Os valores de MAPE, MAE e RMSE foram utilizados como métricas de avaliação (BATHLA, 2020).

Uma rede neural do tipo Long Short-Term Memory (LSTM) é uma arquitetura de rede recorrente, empregando um algoritmo de aprendizado baseado em gradiente apropriado. Sua concepção visa a superação de desafios relacionados à retropropagação de erros. De acordo com Hoochereither e Schmidhuber (1997), a capacidade da LSTM em abranger intervalos temporais superiores a 1000 passos, mesmo diante de sequências de entrada ruidosas e incompressíveis, preservando suas habilidades de detecção de defasagens a curto prazo, é alcançada por meio de um algoritmo eficiente que opera com base em gradiente. Esse algoritmo é aplicado a uma arquitetura que assegura um fluxo de erro constante, evitando problemas de explosão ou desaparecimento, ao direcionar o erro de maneira constante por meio dos estados internos de unidades especializadas. Essa constância no fluxo de erro é mantida desde que o cálculo do gradiente seja truncado em pontos específicos da arquitetura, sem impacto no fluxo de erro a longo prazo (HOCHEREITER; SCHMIDHUBER, 1997).

O modelo de *Recurrent Neural Network* (RNN) e sua arquitetura LSTM são comumente usadas para a previsão do preço de fechamento. A RNN é um tipo de arquitetura de rede neural profunda que possui uma estrutura profunda na dimensão temporal e tem sido amplamente empregada na modelagem de séries temporais. Ao contrário de uma rede neural tradicional, que pressupõe a independência entre as unidades dos vetores de entrada, o modelo RNN adiciona um estado oculto gerado pelas informações sequenciais de uma série temporal, sendo a saída dependente desse estado oculto (BAO et al., 2017).

Na arquitetura LSTM, ocorre uma substituição das camadas ocultas usuais por células LSTM. Essas células, por sua vez, são constituídas por diversos portões que exercem controle sobre o fluxo de entrada. Cada célula LSTM é composta por um portão de entrada, estado da célula, portão de esquecimento e portão de saída. Adicionalmente, inclui uma camada sigmoidal, uma camada tangente hiperbólica (tanh) e uma operação de multiplicação ponto a ponto. O portão de entrada tem a responsabilidade de gerenciar a entrada, enquanto o estado da célula percorre toda a rede, podendo adicionar ou remover informações com a assistência dos portões. O portão de esquecimento desempenha a função de decidir a fração da informação a ser permitida, e o portão de saída consiste na saída gerada pelo LSTM. As camadas sigmoidal e tanh geram números entre zero e um, descrevendo a quantidade de cada componente que deve ser permitida a passar. A camada tanh, adicionalmente, gera um novo vetor que será adicionado ao estado. O estado da célula é então atualizado com base nas saídas dos portões, uma representação matemática do qual pode ser formulada por meio de equações (SELVIN et al., 2017).

Uma LSTM representa uma variação das Redes Neurais Recorrentes (RNN) e se destaca por sua capacidade de extrair informações de estágios anteriores para orientar previsões futuras. De maneira geral, uma Rede Neural Artificial (RNA) compreende três camadas principais, sendo a camada de entrada, as camadas ocultas e a camada de saída (PATTERSON, 2017).

Em uma Rede Neural que contém apenas uma camada oculta, o número de nós na camada de entrada sempre depende da dimensão dos dados. Os nós da camada de entrada se conectam à camada oculta por meio de links chamados 'sinapses'. A relação entre cada par de nós (da entrada para a camada oculta) possui um coeficiente chamado peso, que é o tomador de decisões para os sinais. O processo de aprendizado é naturalmente um ajuste contínuo dos pesos, e após concluir o processo de aprendizado, a Rede Neural Artificial terá pesos ótimos para cada sinapse. Os nós da camada oculta aplicam uma função de ativação, em que se pode destacar, a função sigmoidal ou tangente hiperbólica (tanh) na soma dos pesos provenientes da camada de entrada, chamada função de ativação. Essa transformação gera valores, com uma taxa de erro minimizada entre os dados de treino e teste, utilizando a função *SoftMax* que transforma as saídas de cada classe para 0 e 1 e divide pela soma das saídas. Isso essencialmente proporciona a probabilidade de a entrada estar em uma determinada classe (PATTERSON, 2017).

As fases anteriores dos dados devem ser lembradas para prever e estimar valores futuros, nesse caso, a camada oculta atua como um estoque para as informações passadas da sequência de dados. O termo "recorrente" é usado para descrever o processo de usar elementos de

sequências anteriores para prever dados futuros. As Redes Neurais Recorrentes (RNN) não conseguem armazenar memória por longos períodos, então o uso da Memória de Longo Prazo (LSTM) baseada em uma "linha de memória" mostrou-se muito útil em casos de previsão com dados de longo prazo. Em uma LSTM, a memorização de estágios anteriores pode ser realizada por meio de portões com uma linha de memória alongada incorporada (PATTERSON., 2017).

A previsão de dados de séries temporais é um tema importante nas áreas de economia, negócios e finanças. Tradicionalmente, existem várias técnicas para prever efetivamente o próximo intervalo de dados de séries temporais, como Autoregressivo Univariado (AR), Média Móvel Univariada (MA), Suavização Exponencial Simples (SES) e, mais notavelmente, Autoregressivo Integrado de Médias Móveis (ARIMA) com suas muitas variações. Em particular, o modelo ARIMA demonstrou sua superioridade em precisão e acurácia na previsão dos próximos intervalos de séries temporais. Com o avanço recente na capacidade computacional dos computadores e, mais importante ainda, o desenvolvimento de algoritmos e abordagens de aprendizado de máquina mais avançados, como o aprendizado profundo, novos algoritmos foram desenvolvidos para prever dados de séries temporais (SIAMÍ-NAMINI; NAMIM, 2018).

Redes de Memória de Longo Prazo (LSTM) são redes neurais recorrentes equipadas com um mecanismo especial de controle de acesso às células de memória. Os portões nessas redes podem evitar que o restante da rede modifique o conteúdo das células de memória por vários passos de tempo, permitindo que as redes LSTM preservem sinais e propaguem erros por um período muito mais longo do que as redes neurais recorrentes convencionais. Ao ler, gravar e apagar conteúdo de forma independente das células de memória, os portões também podem aprender a focar em partes específicas dos sinais de entrada e ignorar outras partes. Essas propriedades capacitam as redes LSTM a processar dados com interdependências complexas e separadas, destacando-se em uma variedade de domínios de aprendizado de sequências, como reconhecimento de fala, reconhecimento de escrita à mão offline, tradução automática e geração de legenda para imagens (HOCHREITER; SCHMIDHUBER, 1997).

Mesmo para dados não sequenciais, o sucesso recente de redes profundas mostrou que longas cadeias de computação sequencial são fundamentais para encontrar e explorar padrões complexos. Redes profundas enfrentam exatamente os mesmos problemas que as redes recorrentes aplicadas a sequências longas, ou seja, a informação das computações passadas rapidamente se atenua à medida que avança pela cadeia, portanto, parece atrativo generalizar as vantagens das LSTM para a computação profunda (HOCHREITER; SCHMIDHUBER, 1997).

Após análise de 5 artigos de revisão sistêmica sobre o Aprendizado de Máquina os algoritmos relacionados ao tema conseguiram resultados satisfatórios ao realizar esse tipo de previsão. O uso de redes LSTM nesse cenário, visando prever tendências futuras nos preços das ações com base no histórico de preços, juntamente com indicadores de análise técnica. (NELSON et al., 2017).

De acordo com Nelson et al. (2017) quando se trata de mercados de ações, além de sua complexidade e dinamismo inerentes, há um debate constante sobre a previsibilidade dos retornos das ações. A hipótese de Mercado Eficiente, que define que o preço atual de um ativo sempre reflete instantaneamente todas as informações anteriores disponíveis para ele. Há também a hipótese denominada *Random walk* que afirma que o preço de uma ação muda independentemente de sua história, em outras palavras, o preço de amanhã dependerá apenas das informações de amanhã, independentemente do preço de hoje. Essas duas hipóteses determinam que não há meios de prever com precisão o preço de uma ação. Além disso, uma estratégia aleatória pode superar alguns dos métodos mais clássicos de negociação técnica, como Moving Average Convergence Divergence (MACD) e Relative Strength Index (RSI).

Por outro lado, há outros autores que afirmam que, de fato, os preços das ações podem ser previstos, pelo menos em algum grau. Uma variedade de métodos para prever e modelar o comportamento das ações tem sido objeto de estudo em diversas disciplinas, como economia, estatística, física e ciência da computação. Vale ressaltar que, em 2012, estimava-se que aproximadamente 85% das negociações nos mercados de ações dos Estados Unidos eram realizadas por algoritmos (GLANTZ; KISSEL, 2013).

Um método popular de modelagem e previsão do mercado de ações é a análise técnica, que é um método baseado em dados históricos do mercado, principalmente preço e volume. Ela segue algumas suposições: (1) os preços são definidos exclusivamente pela relação oferta-demanda; (2) os preços mudam seguindo tendências; (3) mudanças na oferta e demanda causam reversões de tendências; (4) mudanças na oferta e demanda podem ser identificadas em gráficos; e (5) padrões em gráficos tendem a se repetir (6). Em outras palavras, a análise técnica não leva em consideração quaisquer fatores externos, como políticos, sociais ou macroeconômico (NELSON et al., 2017).

LSTM pode ser um modelo profundo e recorrente de redes neurais. Redes recorrentes diferem das tradicionais redes feed-forward no sentido de que não possuem apenas conexões neurais em uma única direção, em outras palavras, os neurônios podem enviar dados para uma camada anterior ou para a mesma camada. Nesse caso, os dados não fluem em uma única direção, e o efeito prático disso é a existência de uma memória de curto prazo, além da memória

de longo prazo que as redes neurais já possuem em decorrência do treinamento. As LSTM visavam melhor desempenho ao abordar o problema do gradiente desaparecer do que as redes recorrentes enfrentariam ao lidar com sequências de dados longas. Isso é feito mantendo o fluxo de erro constante por meio de unidades especiais chamadas "portões", que permitem ajustes nos pesos, assim como a truncagem do gradiente quando suas informações não são necessárias. (HOCHREITER; SCHMIDHUBER, 1997).

Para a previsão de preços de ações, dada a notória performance que as redes LSTM têm demonstrado em Processamento de Linguagem Natural (NLP), elas têm sido amplamente utilizadas principalmente ao usar dados de texto de notícias como entrada para prever tendências de preços. No entanto, também há trabalhos que utilizam dados de preços para antecipar movimentos de preços, empregando dados históricos de preços, juntamente com índices de ações, para prever se o preço de uma ação subirá, cairá ou permanecerá o mesmo no dia (NELSON; PEREIRA, DE OLIVEIRA, 2017).

Ainda, pode-se mencionar a classificação CAMELS que se mostrou uma ferramenta eficaz de supervisão interna para avaliar a robustez de uma instituição financeira, identificando aquelas que demandam atenção ou cuidado especial. Essa classificação garante as condições saudáveis de um banco, analisando diferentes aspectos de um banco com base em demonstrações financeiras, fontes de financiamento, dados macroeconômicos, orçamento e fluxo de caixa. CAMELS é um acrônimo para os componentes de segurança e solidez bancária: adequação de capital, qualidade dos ativos, qualidade de gestão, resultado, liquidez e sensibilidade (ROZZANI, 2013).

3 METODOLOGIA

A metodologia de pesquisa, deste trabalho, teve abordagem mista, qualitativa e quantitativa. Para o desenvolvimento do projeto foram escolhidos os bancos Itaú, Banco do Brasil, BTG, Bradesco e Santander como objetos de estudo, pela relevância e representatividade dessas instituições no cenário financeiro. Essa seleção busca proporcionar uma análise abrangente e representativa do setor bancário, considerando diferentes perfis e estruturas organizacionais. Para construção do conjunto de dados leva-se em consideração os preços das ações no período compreendido entre 01 de janeiro de 2020 e 31 de julho de 2023.

Cabe ressaltar, também que esses bancos também são incluídos no ISE (Índice de Sustentabilidade Empresarial), BVSP (Ibovespa) e IFNC (Índice Financeiro). O ISE é um índice composto por empresas que atendem a critérios de sustentabilidade e responsabilidade social. A inclusão no ISE destaca o compromisso das empresas com práticas ambientais, sociais e de governança. O Ibovespa é o principal índice de referência da Bolsa de Valores de São Paulo (B3) e representa uma média ponderada das ações das empresas mais negociadas no mercado acionário brasileiro. Já o IFNC é um índice setorial da B3 que engloba ações de empresas do setor financeiro.

Em relação aos dados macroeconômicos, o projeto inclui dados relacionados às taxas de juros, representadas pela taxa Selic, e ao valor do câmbio, expresso na relação entre o Real brasileiro e o Dólar americano. Além disso são incluídos os dados das variáveis CAMELS de cada banco estudado nesse projeto.

A abordagem da proposta do trabalho é pautada em desenvolver um programa preditor de preços ao incorporar ou não o preço dos índices como uma variável de entrada na arquitetura da rede neural. A inclusão dos índices na modelagem busca capturar tendências macroeconômicas e padrões do mercado que podem influenciar diretamente o comportamento das ações bancárias. Ao comparar os resultados obtidos com e sem a inclusão dos índices, o estudo pretende fornecer *insights* sobre a relevância dessa informação adicional na melhoria da precisão das previsões.

Nesse trabalho, os resultados gerados foram realizados por meio de um código Python que utiliza a biblioteca Keras para construir uma rede neural recorrente do tipo LSTM (Long Short-Term Memory). Primeiramente, é definida uma semente para garantir a reprodutibilidade dos resultados e o conjunto de dados é dividido em treinamento (80%) e teste (20%), sendo esses dados os preços diários das ações dos bancos Itaú, Banco do Brasil, BTG, Bradesco e Santander. Em seguida, um modelo sequencial é criado, e uma camada de entrada é adicionada para

processar séries temporais com janelas de tamanho 5 e 37 dimensões por ponto, dimensões essas que são as variáveis. A arquitetura inclui uma camada LSTM com 128 unidades e ativação ReLU para aprender dependências temporais de longo prazo. Uma camada de dropout com taxa de 20% é inserida para evitar o overfitting ao desativar aleatoriamente neurônios durante o treinamento. Finalmente, uma camada densa de saída com ativação linear é adicionada para gerar previsões contínuas. Depois disso a rede é treinada ao longo de 100 épocas com objetivo de minimizar a função de perda. Em resumo, o modelo é projetado para tarefas de regressão, onde a entrada é uma série temporal e a saída é um valor contínuo.

4 RESULTADOS E DISCUSSÃO

Para desenvolvimento do projeto o código foi implementado em Python utilizando bibliotecas como TensorFlow para a construção do modelo, biblioteca scikit-learn para avaliação de métricas de desempenho e a biblioteca yfinance para baixar dados financeiros dos ativos. Foram analisados o desempenho de cinco redes neurais de acordo com as métricas MAE (*Mean Absolute Error*), MAPE (*Mean Absolute Percentage Error*) e RMSE (*Mean Squared Error*).

A primeira rede é uma rede neural com a inclusão de todos os índices financeiros considerados nesse estudo (ISE, IFNC e BVSP) como variáveis. A segunda é uma rede neural sem a inclusão do índice ISE, ou seja, considera-se somente os índices IFNC e BVSP. A terceira é uma rede neural sem a inclusão do índice IFNC. A quarta é uma rede neural sem a inclusão do índice BVSP. A quinta é uma rede neural sem a inclusão dos três índices mencionados nesse estudo (ISE, IFNC, e BVSP). E por último temos o desempenho do *Random Walk* usado como base de comparação. Importante ressaltar que à medida que o valor das métricas se aproximam de 0, melhor é o desempenho da rede.

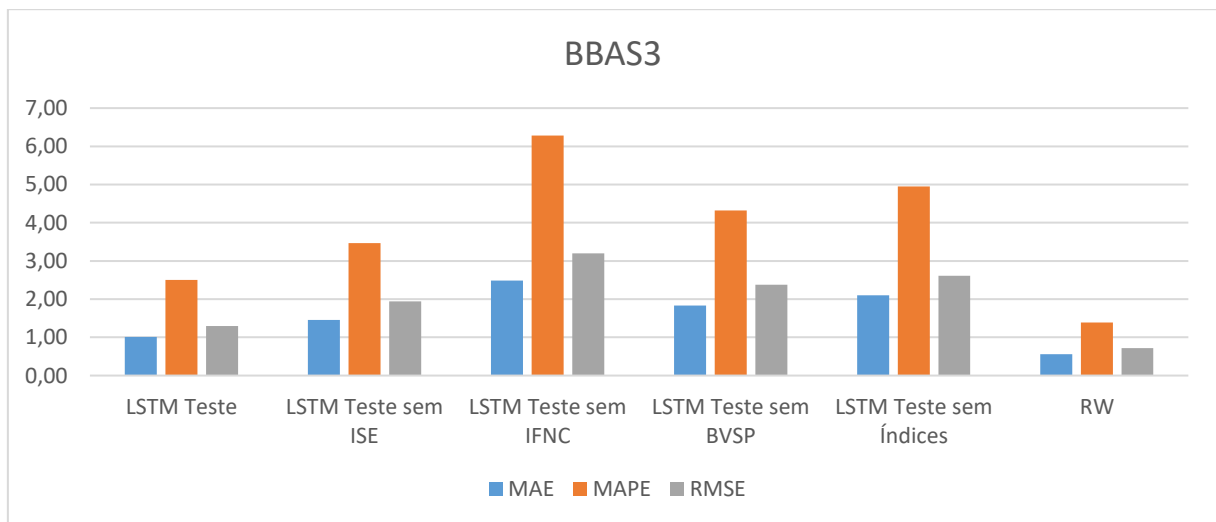
Importante salientar que o modelo da rede é modificado de 37 dimensões por ponto para 36 nas redes em que não se é considerado algum índice, e é alterado para 34 na rede em que não se considera nenhum dado dos índices ISE, IFNC e BVSP como variável de entrada na rede.

Para a avaliação do desempenho da rede neural LSTM construída foi utilizada a métrica *Mean Squared Error* (MSE) como função de perda. Dessa forma, o objetivo do treinamento da rede é minimizar a função de perda, ajustando-a para reduzir a diferença entre as previsões da rede e os valores verdadeiros do conjunto de treinamento. Portanto, foi analisado os valores de perda das redes neurais dos bancos mencionados nesse estudo: Banco do Brasil, Banco

Bradesco, Banco BTG Pactual, Banco Itaú e Banco Santander. Sendo assim, foram reunidos em um gráfico os valores previstos pela rede neural modelada, intitulada como “LSTM”, os valores do Random Walk usados como referência para avaliar o desempenho da redes e os valores dos dados reais.

4.1 – LTSM aplicada as ações do Banco do Brasil

Figura 1 - LTSM aplicada as ações do Banco do Brasil.



Fonte: O Autor.

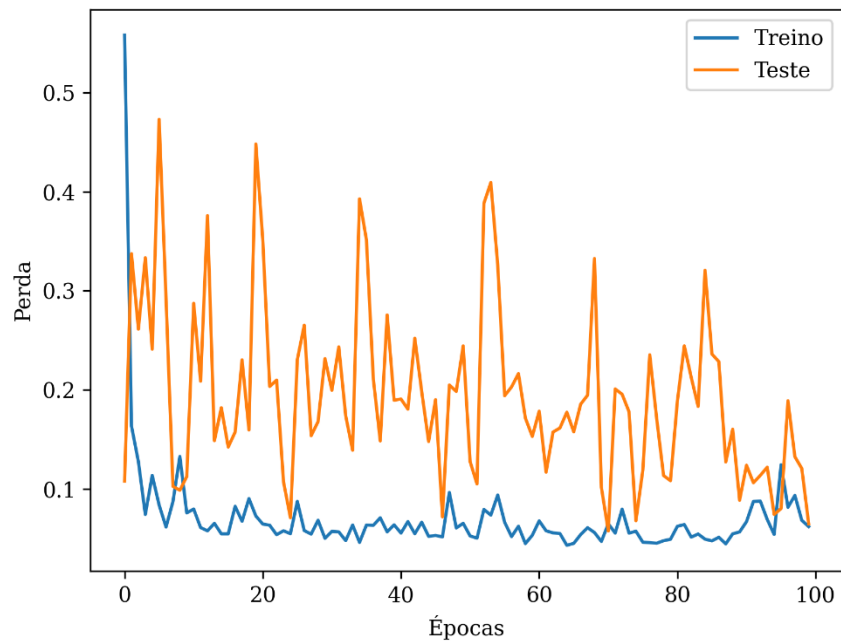
No contexto específico do Banco do Brasil, observou-se que a rede neural alcançou seu desempenho mais destacado quando consideramos todas as métricas avaliadas (MAE, MAPE e RMSE), sendo essa avaliação realizada com a inclusão de todos os índices pertinentes (ISE, IFNC e BVSP). A segunda melhor performance foi registrada pela rede neural que excluiu apenas o índice ISE, seguida pela rede que omitiu o índice BVSP, sem a inclusão de quaisquer índices, e, por último, a rede que desconsiderou o índice IFNC.

4.1.1 – Resultados individuais das redes para ações do Banco do Brasil

4.1.2.1 – Rede neural LSTM Teste

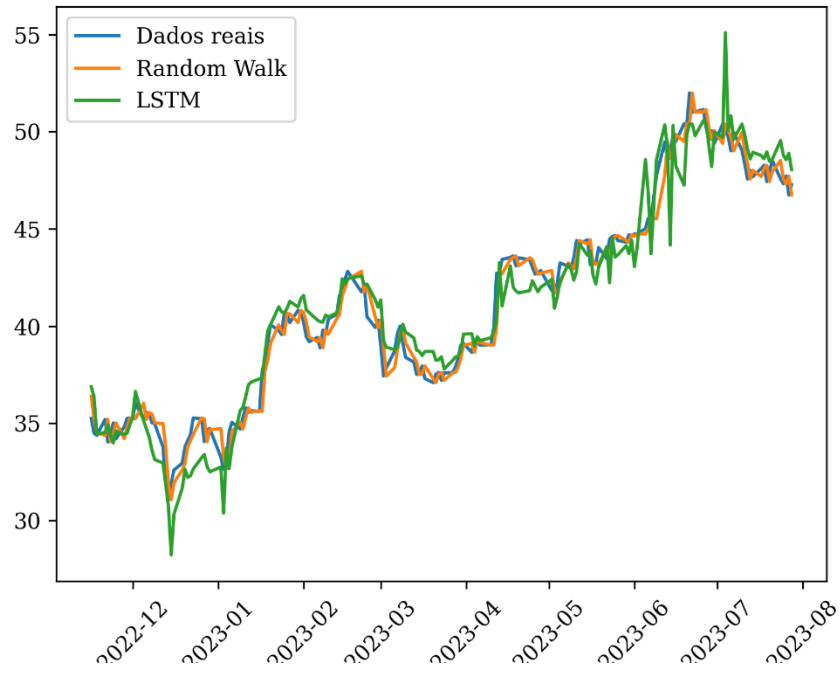
Em redes neurais, a análise do gráfico de perda por época revela informações cruciais sobre o processo de treinamento. Ao observar a evolução da perda ao longo das épocas, é possível identificar padrões, ajustar hiperparâmetros e avaliar a convergência do modelo. Em particular, picos de perda podem indicar *overfitting*, enquanto uma redução consistente sugere que o modelo está aprendendo de maneira eficaz.

Figura 2 - Gráfico de perda por época para a ação do Banco do Brasil.



Fonte: O Autor

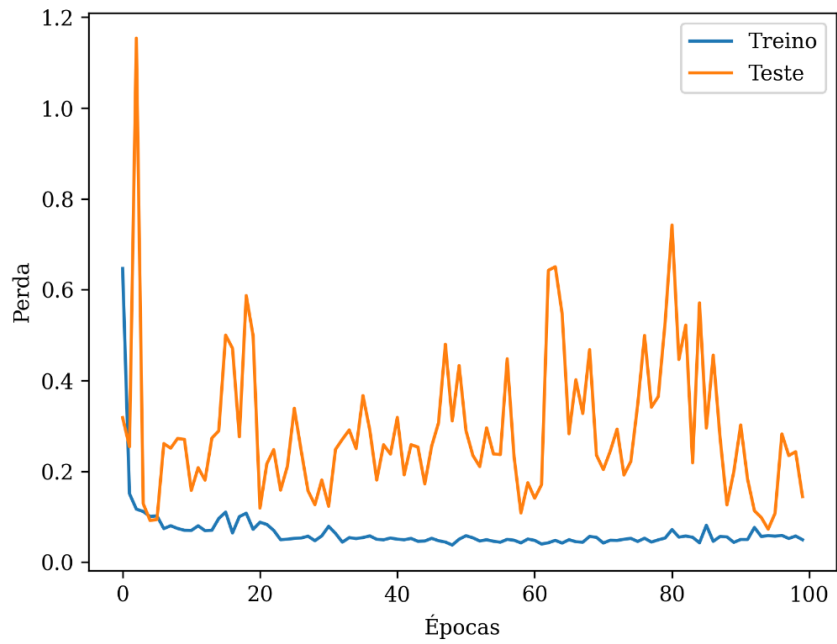
Figura 3 - Dados reais, Random Walk e valores previstos pela rede LSTM.



Fonte: Autor.

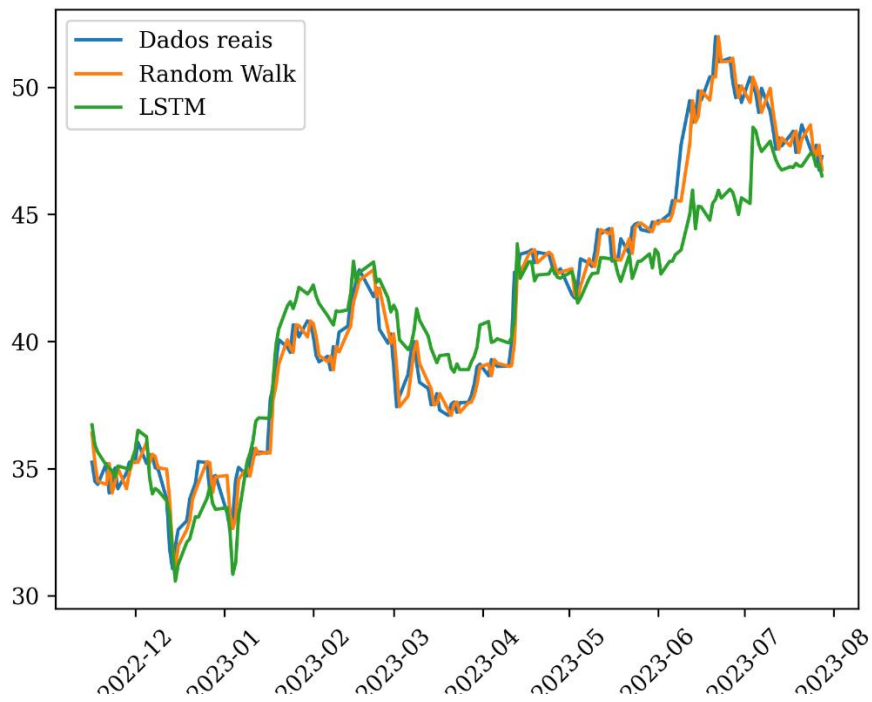
4.1.2.2 – Rede neural LSTM Teste sem ISE

Figura 4 - Rede neural LSTM Teste sem ISE.



Fonte: O Autor.

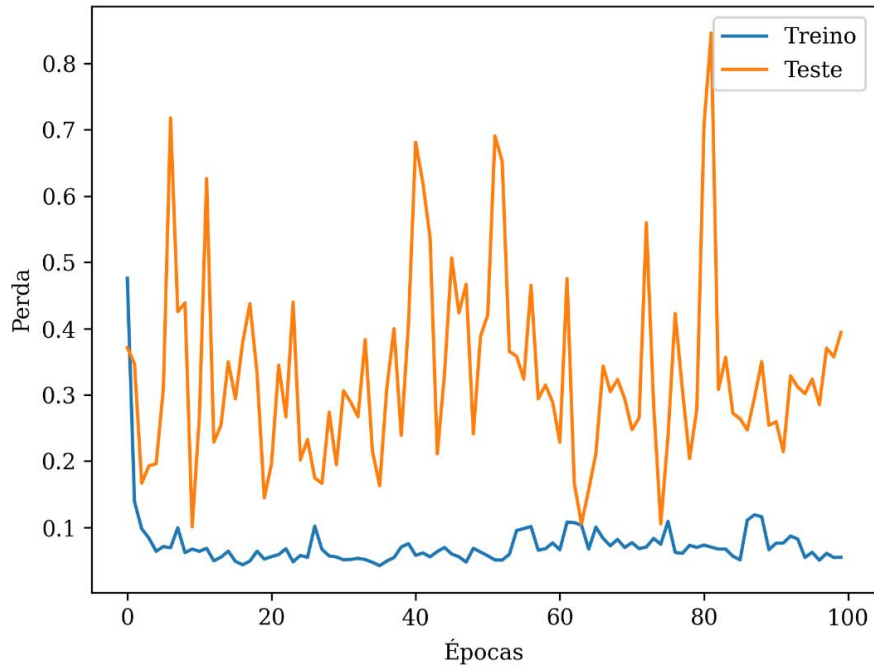
Figura 5 - Dados reais, Random Walk e valores previstos pela rede LSTM sem ISE.



Fonte: Autor

4.1.2.3 – Rede neural LSTM Teste sem IFNC

Figura 6 - Rede neural LSTM Teste sem IFNC.



Fonte: O Autor.

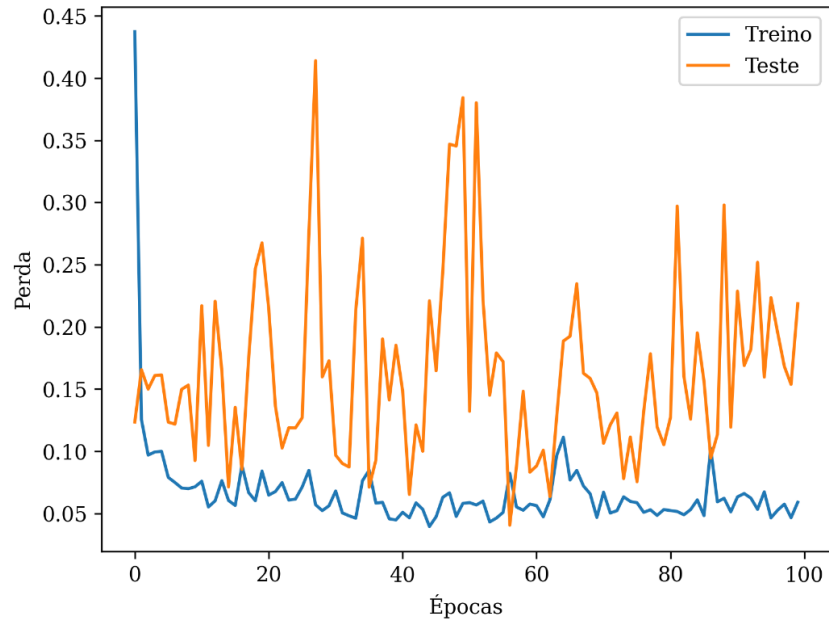
Figura 7 - Dados reais, Random Walk e valores previstos pela rede LSTM sem INFC.



Fonte: Autor.

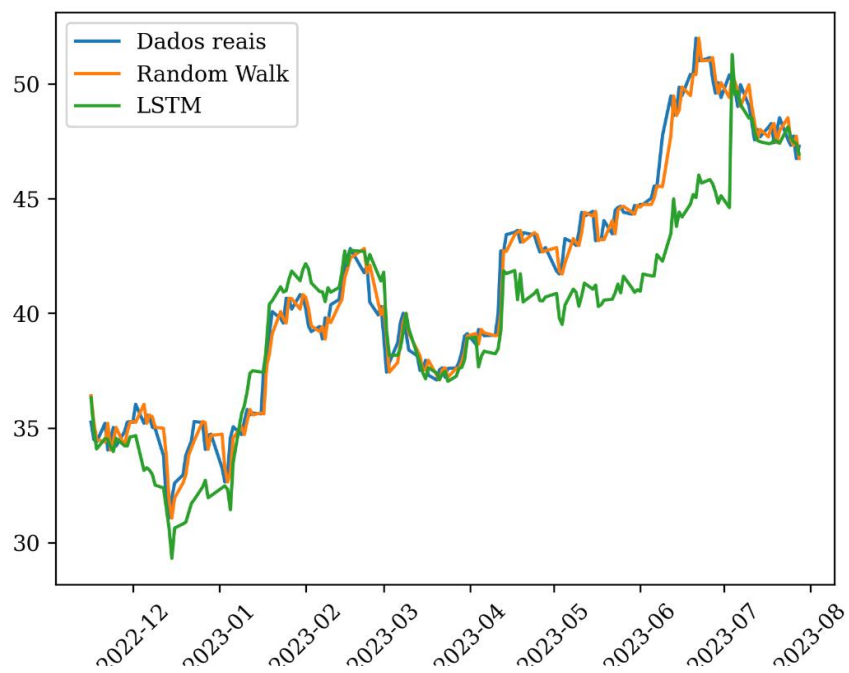
4.1.2.4 – Rede neural LSTM Teste sem BVSP

Figura 8 - Rede neural LSTM Teste sem BVSP.



Fonte: O Autor.

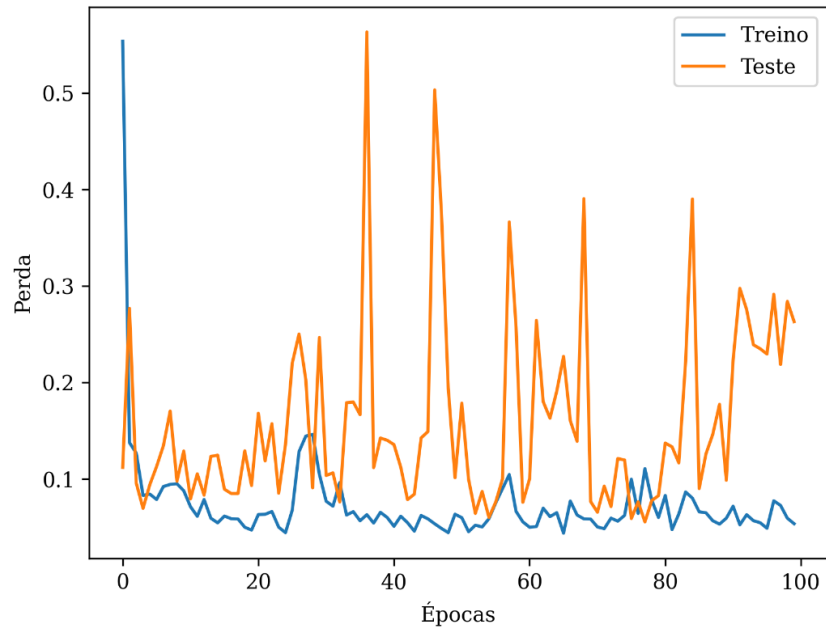
Figura 9 - Dados reais, Random Walk e valores previstos pela rede LSTM sem BVSP.



Fonte: Autor.

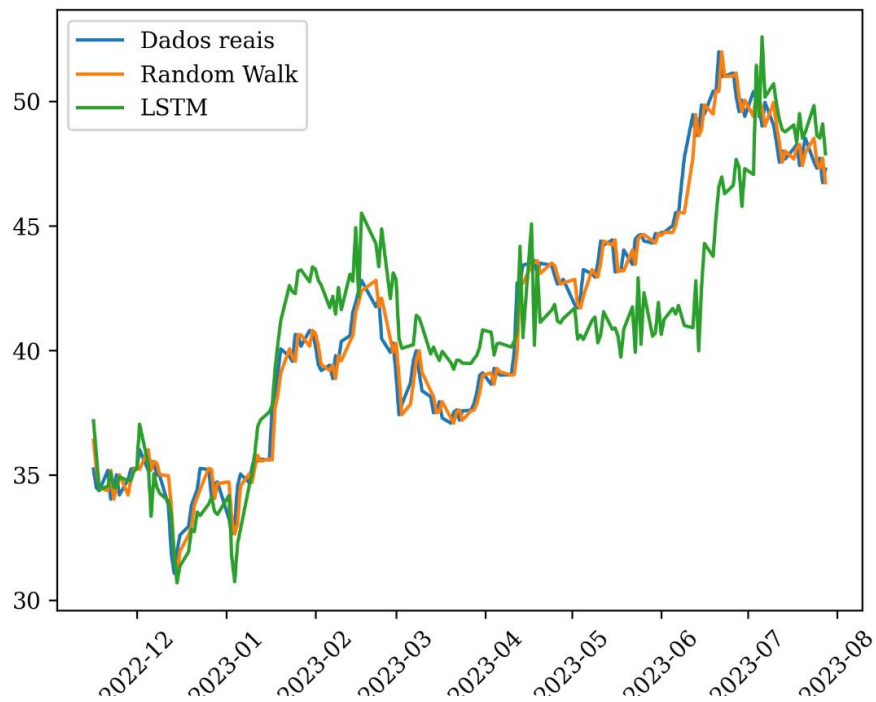
4.1.2.5 – Rede neural LSTM Teste sem Índices

Figura 10 - Rede neural LSTM Teste sem Índices.



Fonte: O Autor.

Figura 11 - Dados reais, Random Walk e valores previstos pela rede LSTM sem índices.



Fonte: Autor

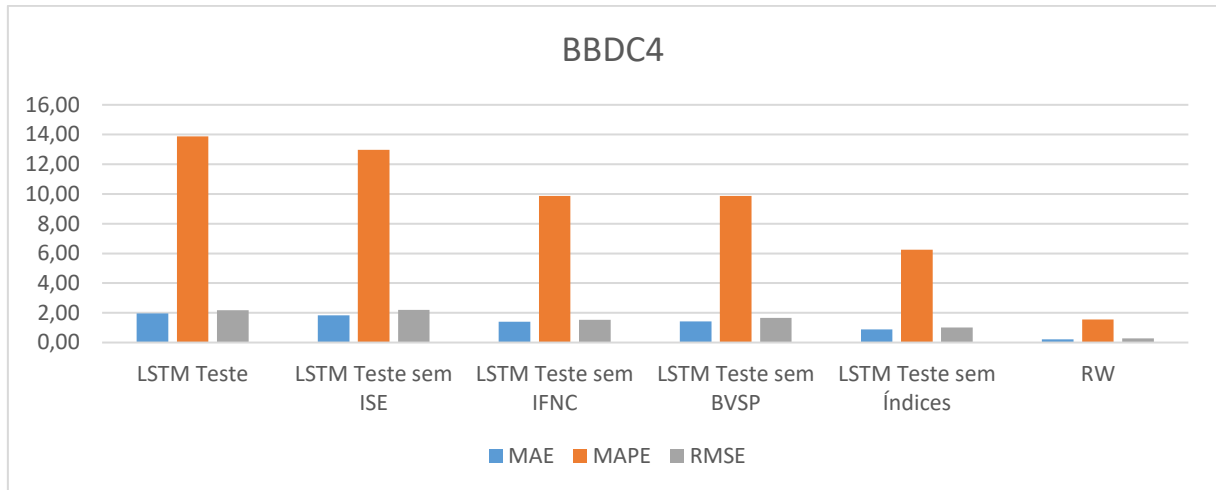
A avaliação da Rede Neural revelou um padrão significativo de elevado índice de perda em função da época, ao compararmos os conjuntos de treino (em azul) e teste (em laranja). Essa disparidade entre as curvas indica que o modelo está enfrentando desafios substanciais ao generalizar para dados não utilizados no treinamento, evidenciando uma possível dificuldade na capacidade de aprendizado do modelo. O aumento persistente do erro no conjunto de teste ao longo das épocas sugere a presença de *overfitting*, no qual a rede neural se adapta excessivamente aos dados de treino, resultando em uma performance deficiente em novos dados. Essa observação sublinha a importância de estratégias de regularização, ajustes na complexidade do modelo ou, possivelmente, a inclusão de outras variáveis relevantes para aprimorar a generalização da Rede Neural LSTM no contexto.

Na rede incluindo todos os índices nota-se que a perda da rede segue uma tendência de queda ao longo das épocas. Isso indica que a rede está aprendendo de forma eficaz a representação dos dados e com isso fazendo previsões mais precisas. Além disso a linha do valor de preço da ação seguiu próxima ao valor real do ativo na maior parte do tempo. A análise comparativa entre os dados reais (representados em azul), o Random Walk (em laranja) e as previsões geradas pela rede neural LSTM (em verde) revela uma notável semelhança entre os conjuntos. A sobreposição visual dos dados reais e as previsões da LSTM indica a habilidade do modelo em capturar padrões e tendências presentes nos dados observados. Surpreendentemente, a rede LSTM demonstra uma capacidade significativa de superar a aleatoriedade do Random Walk em alguns momentos, evidenciando sua habilidade em discernir padrões mais complexos e representar de maneira mais precisa a dinâmica subjacente dos dados reais. Essa consistência entre os dados reais e as previsões da LSTM valida a robustez do modelo, sugerindo que a rede neural é capaz de realizar projeções coerentes em relação aos comportamentos passados e atuais dos dados.

Já nas redes com a exclusão dos índices ISE, IFNC e BVSP e a rede sem a inclusão de qualquer um desses índices, observa-se que não há uma tendência clara de baixa na perda da rede. Pode-se dizer então que a rede não está aprendendo de forma eficaz o que prejudica o objetivo da rede prever o valor da ação. Isso pode ser notado nos gráficos de comparação em que o valor de previsão de preço da rede LSTM está mais distante do valor real da ação em comparação a rede LSTM com a inclusão dos índices.

4.2 – LTSM aplicada as ações do Banco Bradesco

Figura 12 - LTSM aplicada as ações do Banco Bradesco.



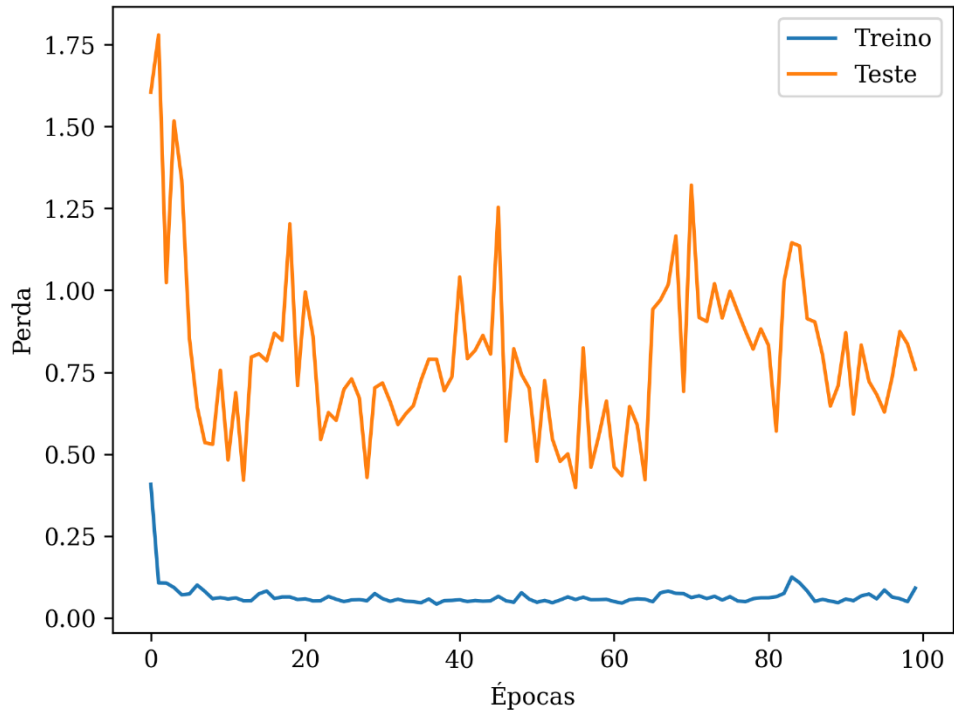
Fonte: O Autor.

Na análise do Banco Bradesco o melhor desempenho entre as métricas foi a rede neural sem a inclusão dos índices. Observa-se também que a exclusão do índice IFNC permite um maior desempenho, seguido por BVSP e ISE. Por último a rede com a inclusão de todos os índices teve o pior desempenho.

4.2.2 – Resultados individuais das redes para ações do Banco Bradesco

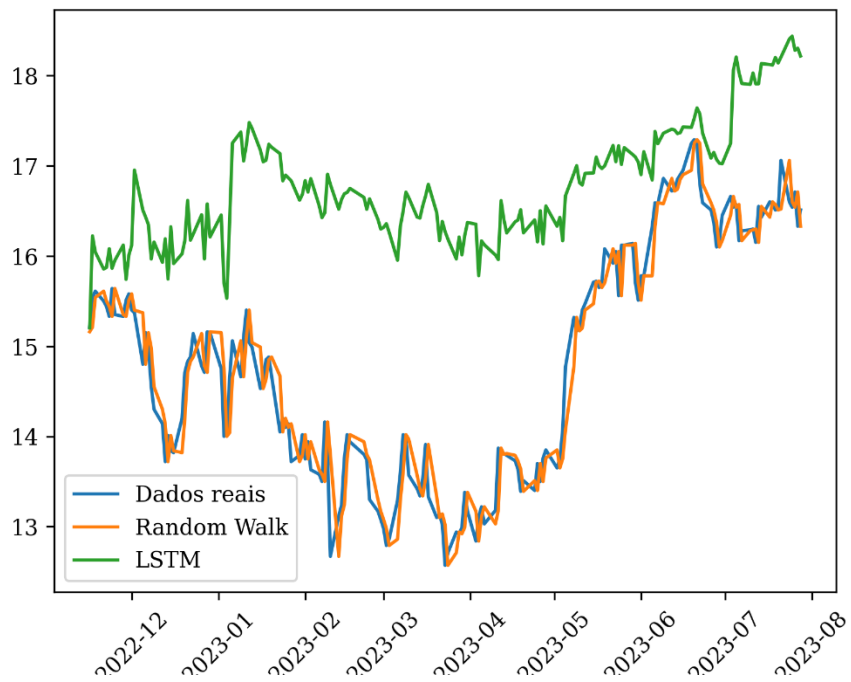
4.2.3 – Rede neural LSTM Teste

Figura 13 - Gráfico de perda por época para a ação do Banco Bradesco.



Fonte: O Autor.

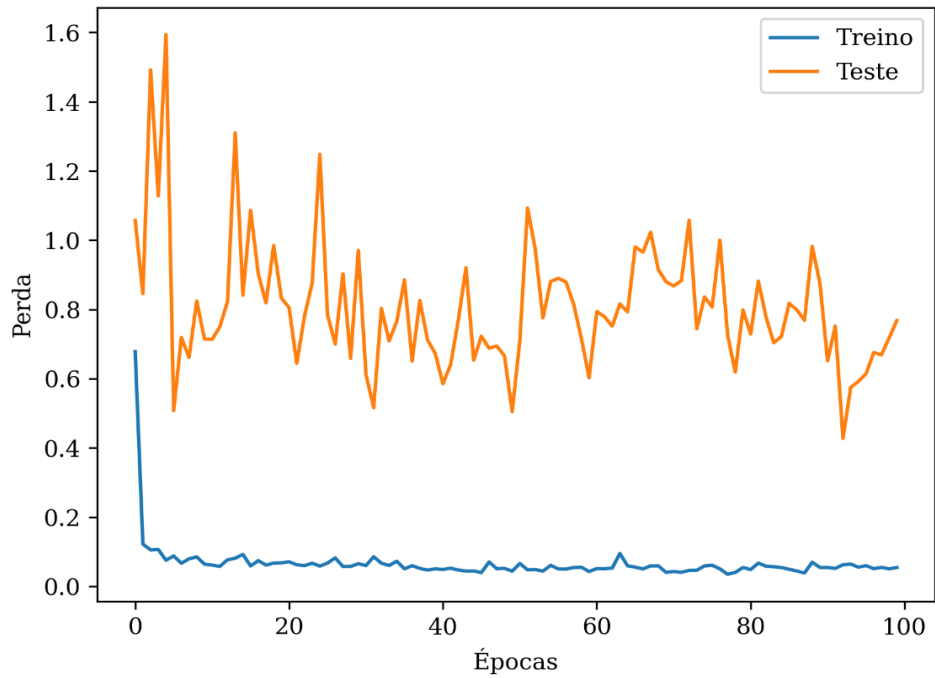
Figura 14 - Dados reais, Random Walk e valores previstos pela rede LSTM.



Fonte: Autor.

4.2.4 – Rede neural LSTM Teste sem ISE

Figura 15 - Rede neural LSTM Teste sem ISE.



Fonte: O Autor.

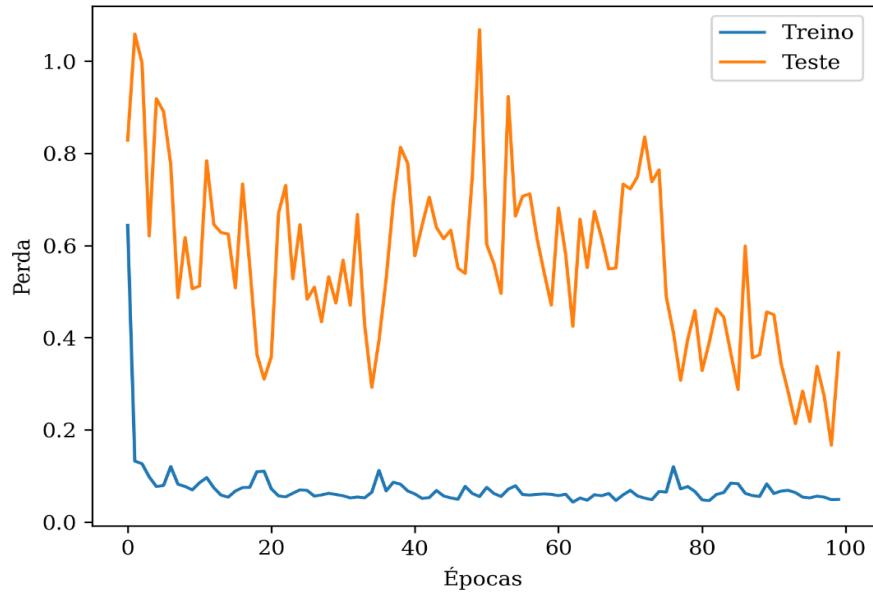
Figura 16 - Dados reais, Random Walk e valores previstos pela rede LSTM sem ISE.



Fonte: O Autor.

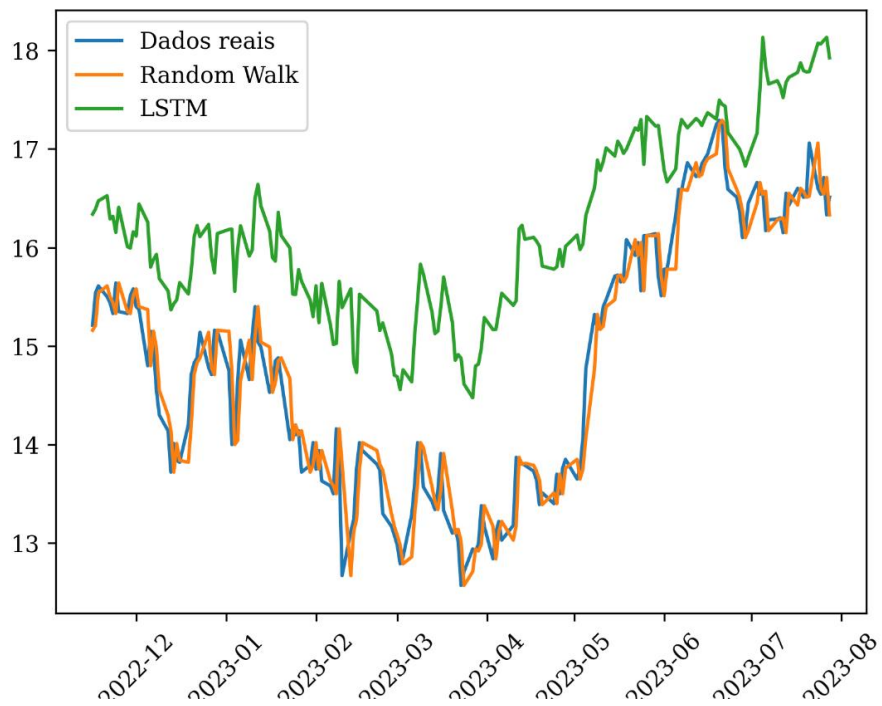
4.2.5 – Rede neural LSTM Teste sem IFNC

Figura 17 - Rede neural LSTM Teste sem IFNC.



Fonte: O Autor.

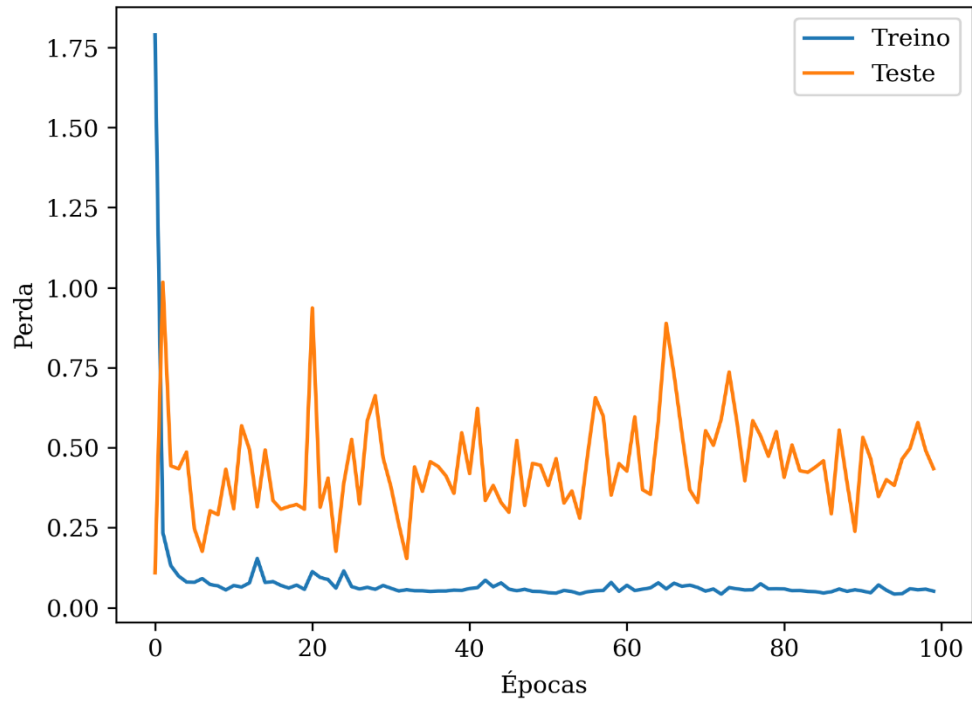
Figura 18 - Dados reais, Random Walk e valores previstos pela rede LSTM sem INFC.



Fonte: O Autor.

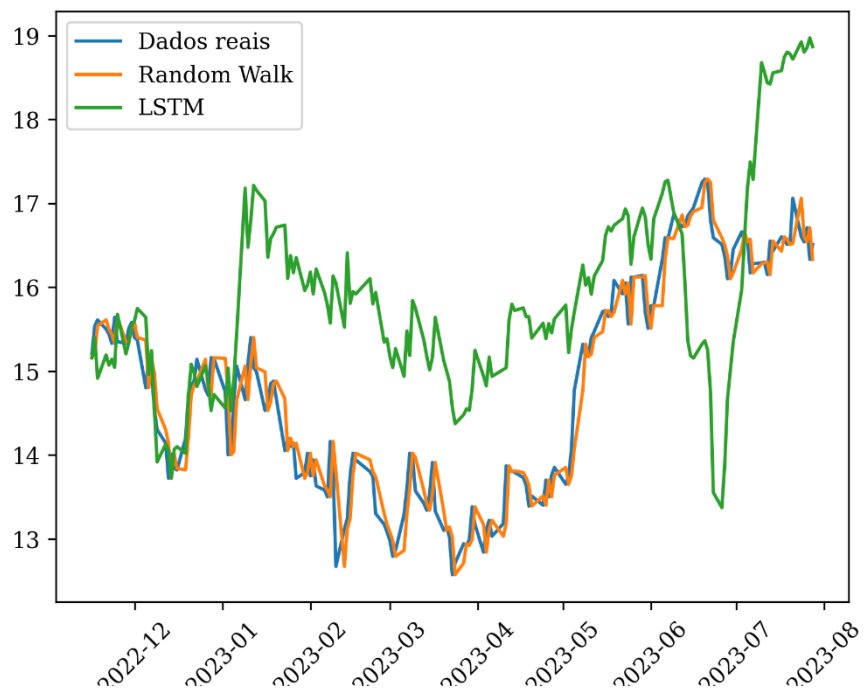
4.2.6 – Rede neural LSTM Teste sem BVSP

Figura 19 - Rede neural LSTM Teste sem BVSP.



Fonte: O Autor.

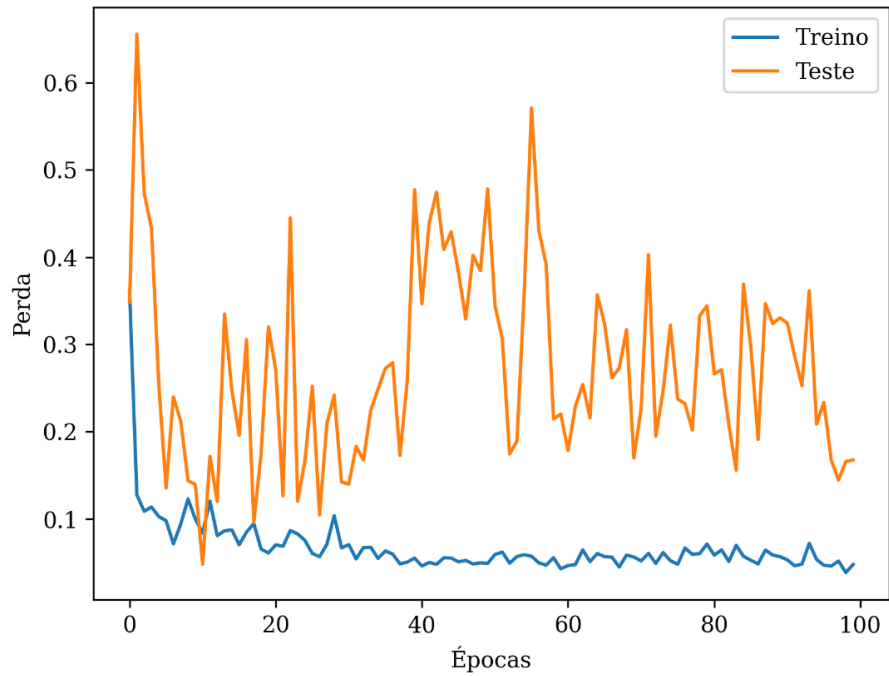
Figura 20 - Dados reais, Random Walk e valores previstos pela rede LSTM sem BVSP.



Fonte: o Autor.

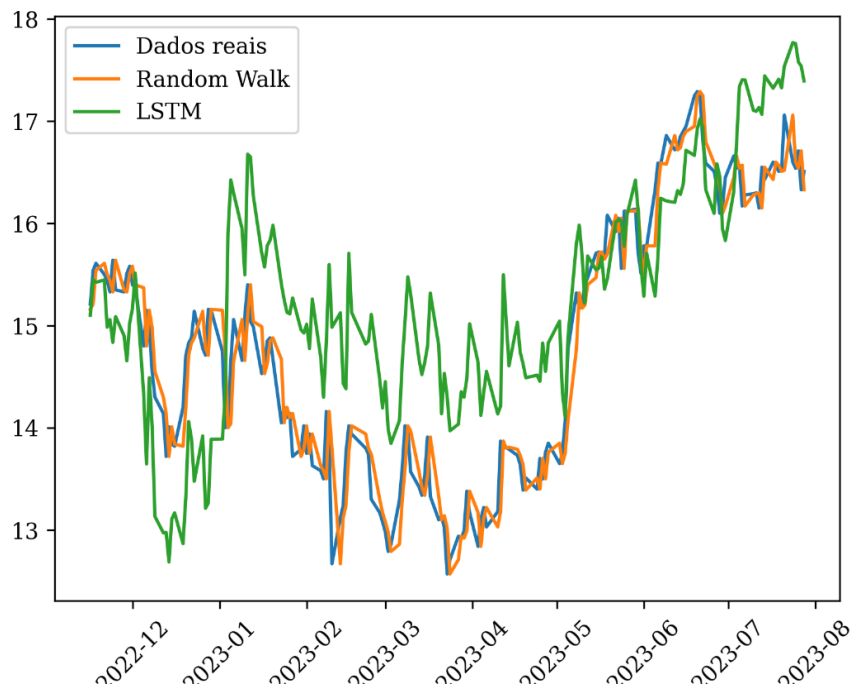
4.2.7 – Rede neural LSTM Teste sem Índices

Figura 21 - Rede neural LSTM Teste sem Índices.



Fonte: O Autor.

Figura 22 - Dados reais, Random Walk e valores previstos pela rede LSTM sem índices.



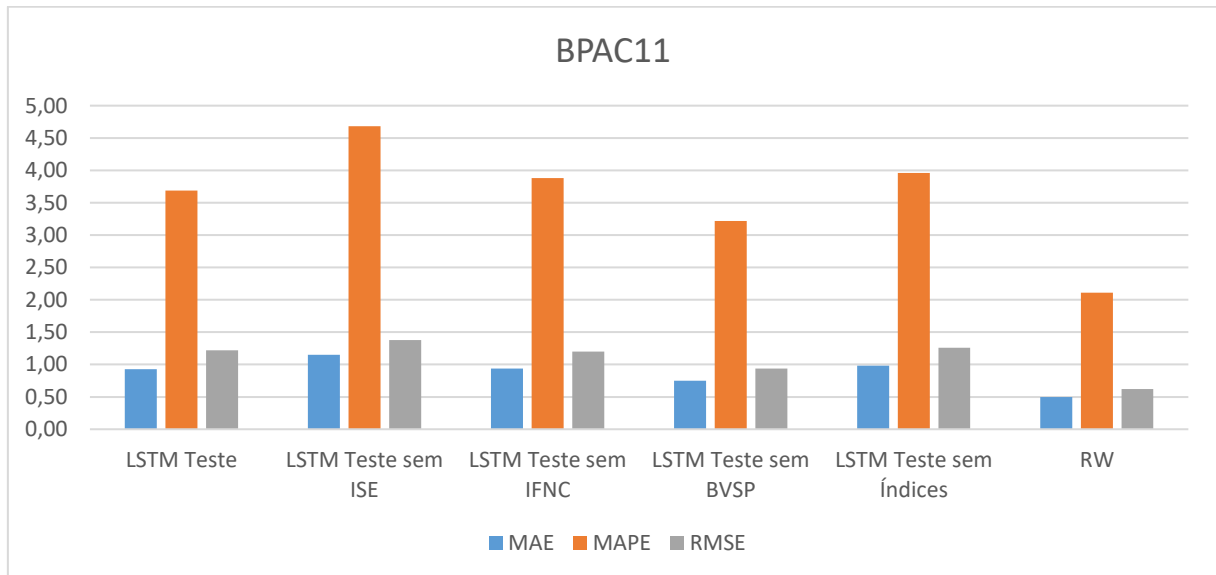
Fonte: Autor.

Em todos os casos observou-se que depois de uma certa quantidade de épocas passadas o valor de perda da rede tendeu-se a permanecer lateralizado, o que foi refletido em valores significativamente distantes entre o valor que a rede previu e o valor real da ação. Porém nota-se que a rede sem a inclusão dos índices BVSP, IFNC e ISE obteve os menores valores de perda ao longo das épocas chegando ao um valor abaixo de 0.2 após 100 épocas, resultado esse refletido em valores previsto mais próximos dos valores reais da ação.

O elevado índice de perda reflete dificuldades na capacidade dessa rede neural em aprender eficientemente a relação entre os dados de entrada e as previsões desejadas. Essa situação pode ser causada por diversos fatores, como uma arquitetura inadequada da rede, hiperparâmetros mal ajustados, overfitting, underfitting ou até mesmo a falta de representatividade dos dados de treino em relação aos dados de teste.

4.3 – LSTM aplicada as ações do Banco BTG Pactual

Figura 23 - LSTM aplicada as ações do Banco BTG Pactual.



Fonte: Autor.

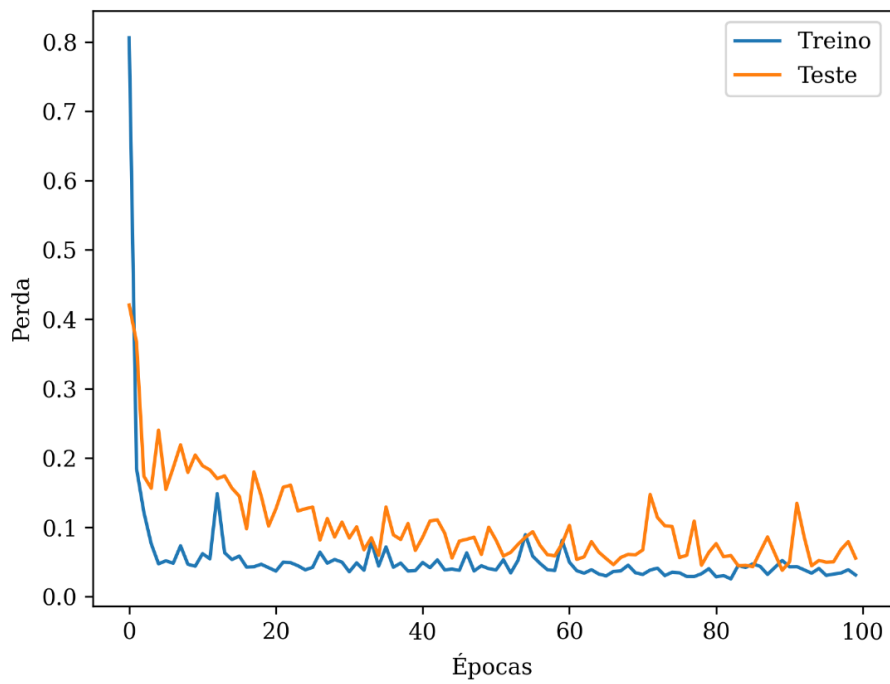
Nesse caso a exclusão do índice BVSP permitiu um maior desempenho da rede, seguido pela rede com a inclusão de todos os índices. Logo após a rede com a exclusão do IFNC teve um desempenho melhor que a rede sem a inclusão de nenhum índice, e por último a rede com exclusão do índice ISE teve o pior desempenho. Observa-se também que a rede com todos os

índices obteve um resultado mediano, obtendo o terceiro melhor resultado atrás da rede neural sem o índice BVSP e o Random Walk.

4.3.1 – Resultados individuais das redes para ações do Banco BTG Pactual

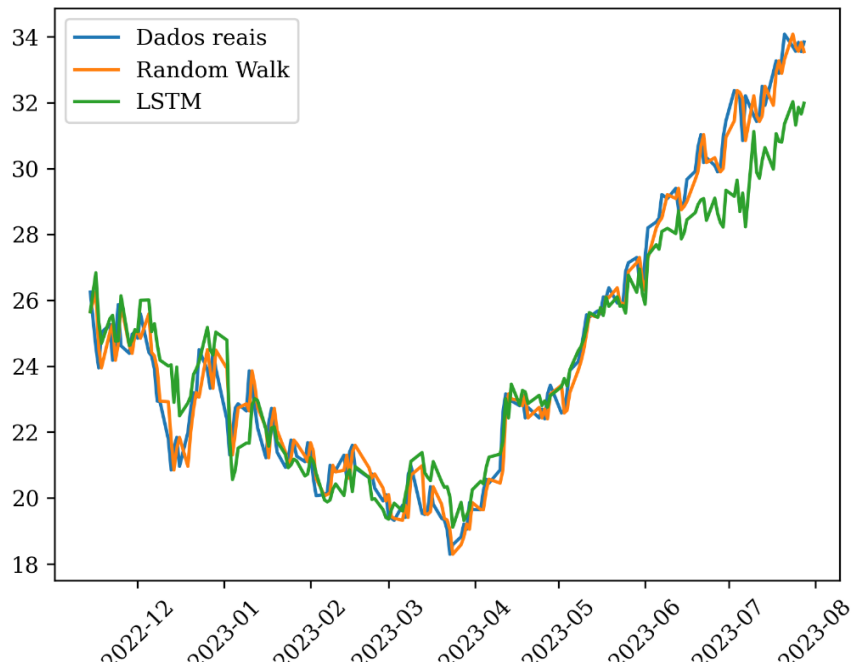
4.3.2 – Rede neural LSTM Teste

Figura 24 - Gráfico de perda por época para a ação do Banco BTG Pactual.



Fonte: O Autor.

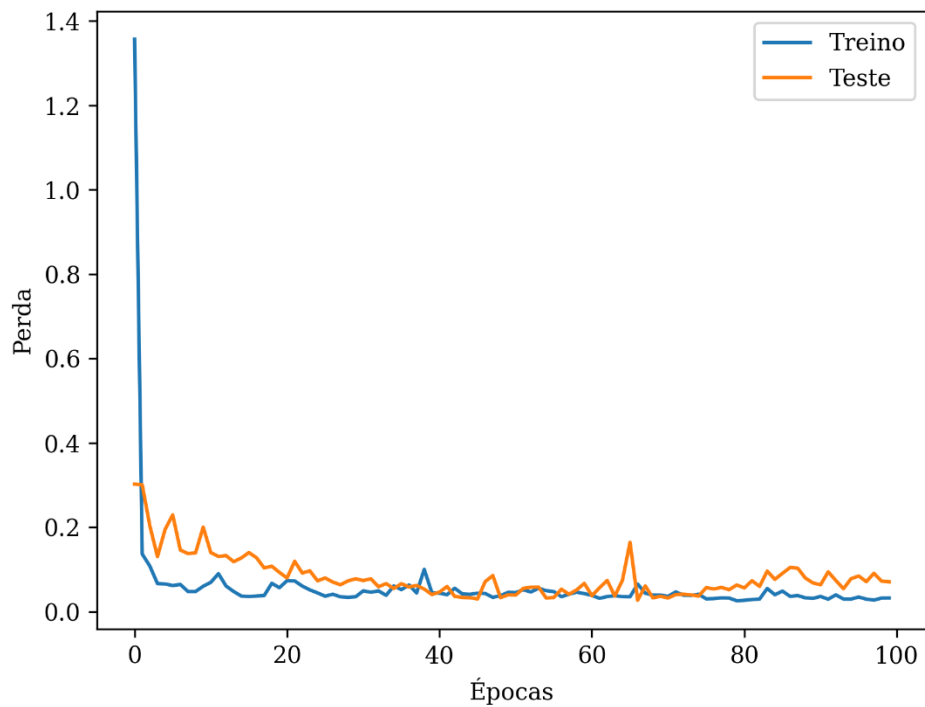
Figura 25 - Dados reais, Random Walk e valores previstos pela rede LSTM.



Fonte: Autor.

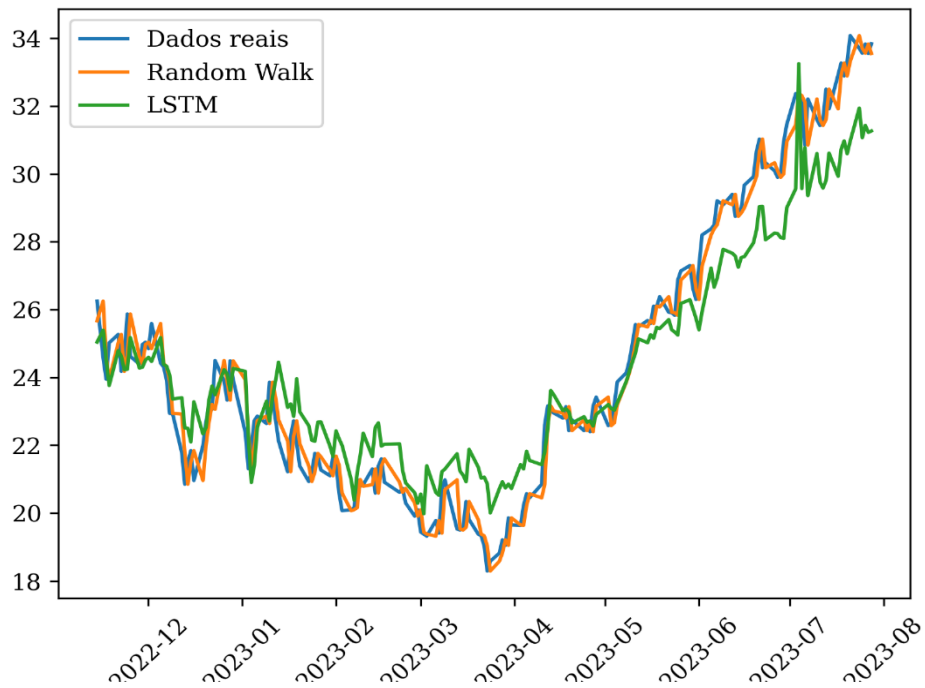
4.3.3 – Rede neural LSTM Teste sem ISE

Figura 26 - Rede neural LSTM Teste sem ISE.



Fonte: o Autor.

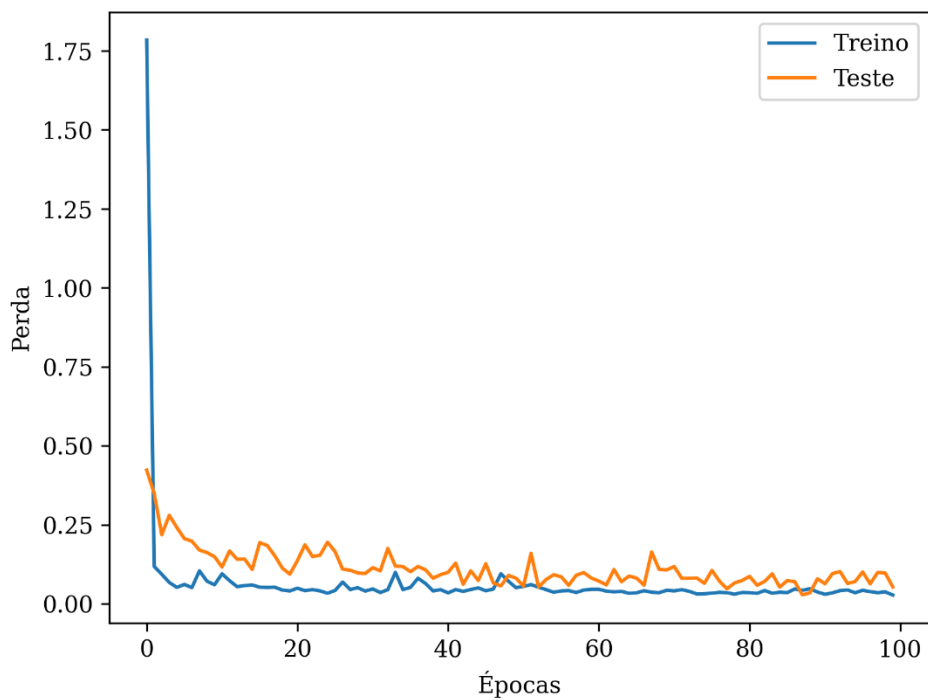
Figura 27 - Dados reais, Random Walk e valores previstos pela rede LSTM sem ISE.



Fonte: Autor.

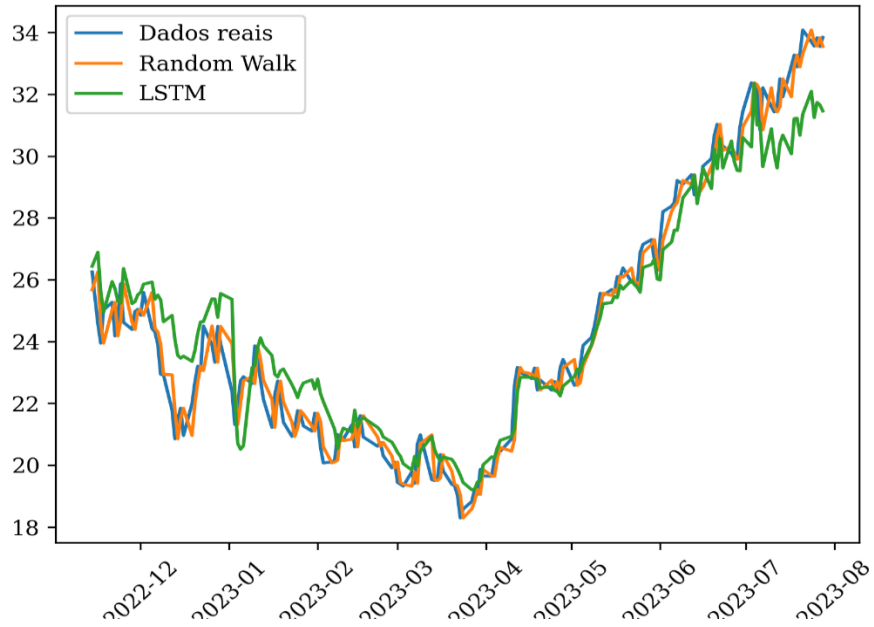
4.3.4 – Rede neural LSTM Teste sem IFNC

Figura 28 - Rede neural LSTM Teste sem IFNC.



Fonte: O Autor.

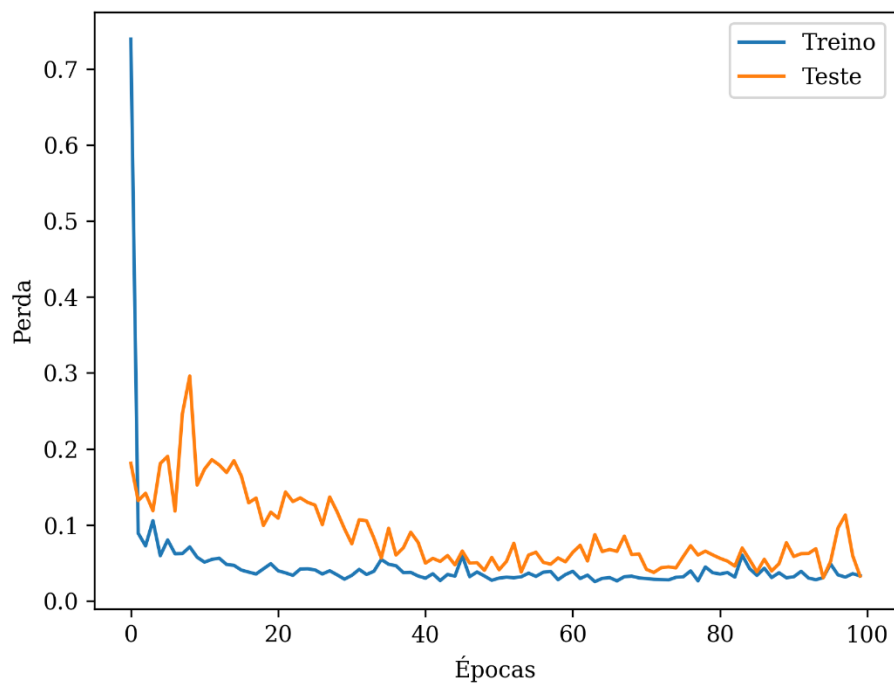
Figura 29 - Dados reais, Random Walk e valores previstos pela rede LSTM sem INFC.



Fonte: o Autor.

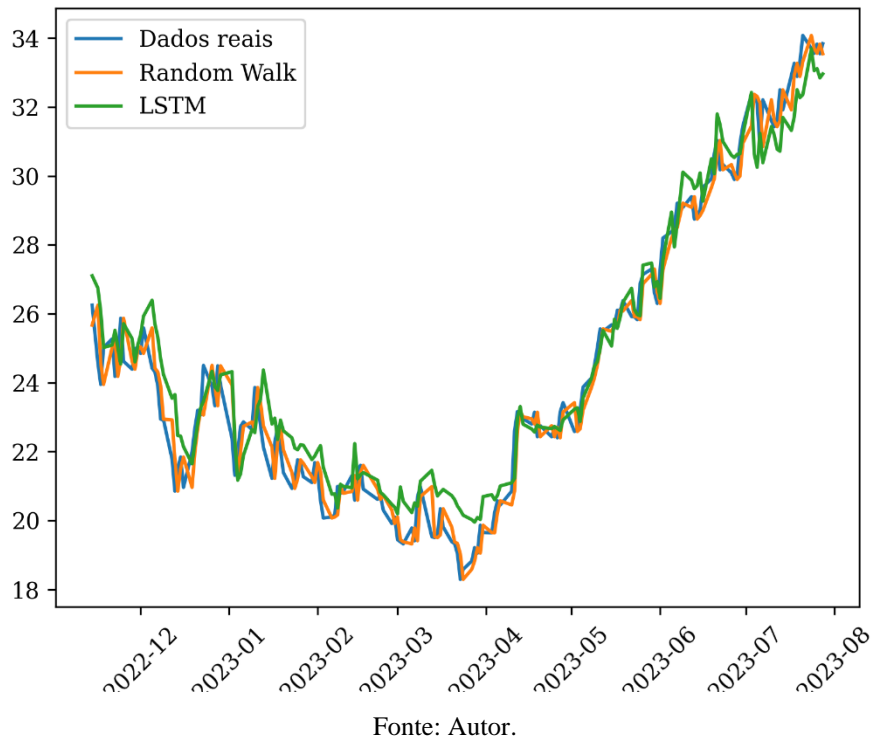
4.3.5 – Rede neural LSTM Teste sem BVSP

Figura 30 - Rede neural LSTM Teste sem BVSP.



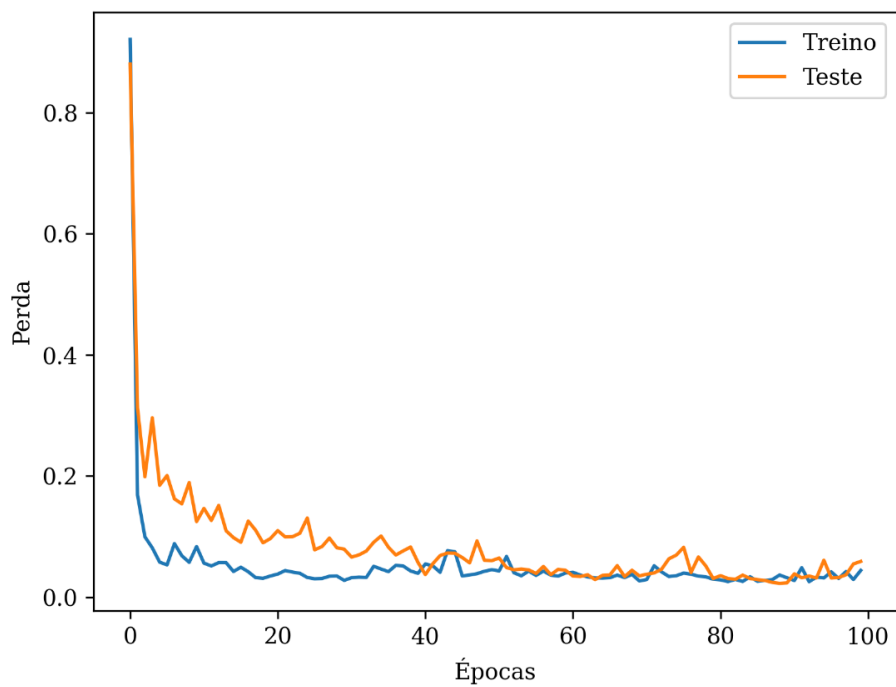
Fonte: O Autor.

Figura 31 - Dados reais, Random Walk e valores previstos pela rede LSTM sem BVSP.



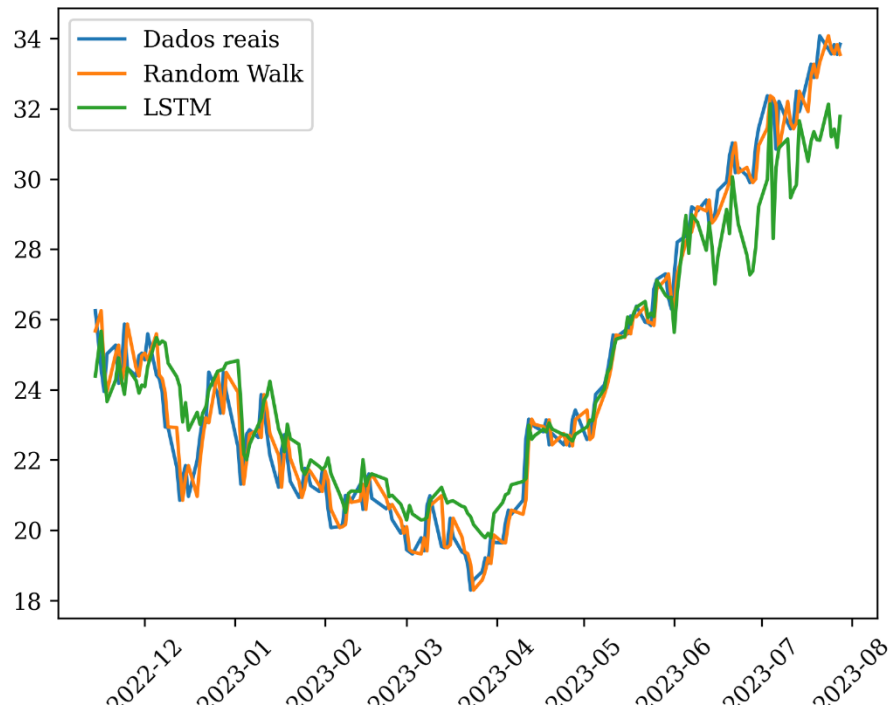
4.3.6 – Rede neural LSTM Teste sem Índices

Figura 32 - Rede neural LSTM Teste sem Índices.



Fonte: O Autor.

Figura 33 - Dados reais, Random Walk e valores previstos pela rede LSTM sem índices.

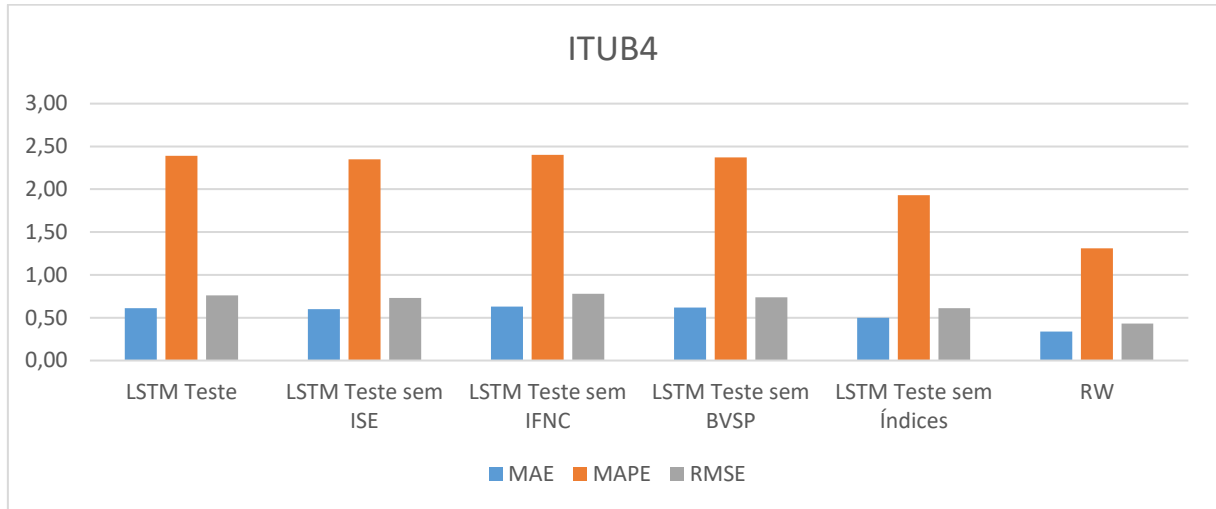


Fonte: Autor.

Verificou-se que em todos os casos a perda da rede tende-se a diminuir ao longo das épocas, o que foi refletido em valores previstos mais próximos dos valores reais da ação. Isso indica que o modelo está aprendendo de forma eficaz com os dados fornecidos e portanto se tornando mais capaz de realizar previsões mais precisas.

4.4 – LTSM aplicada as ações do Banco Itaú

Figura 34 - LTSM aplicada as ações do Banco Itaú.

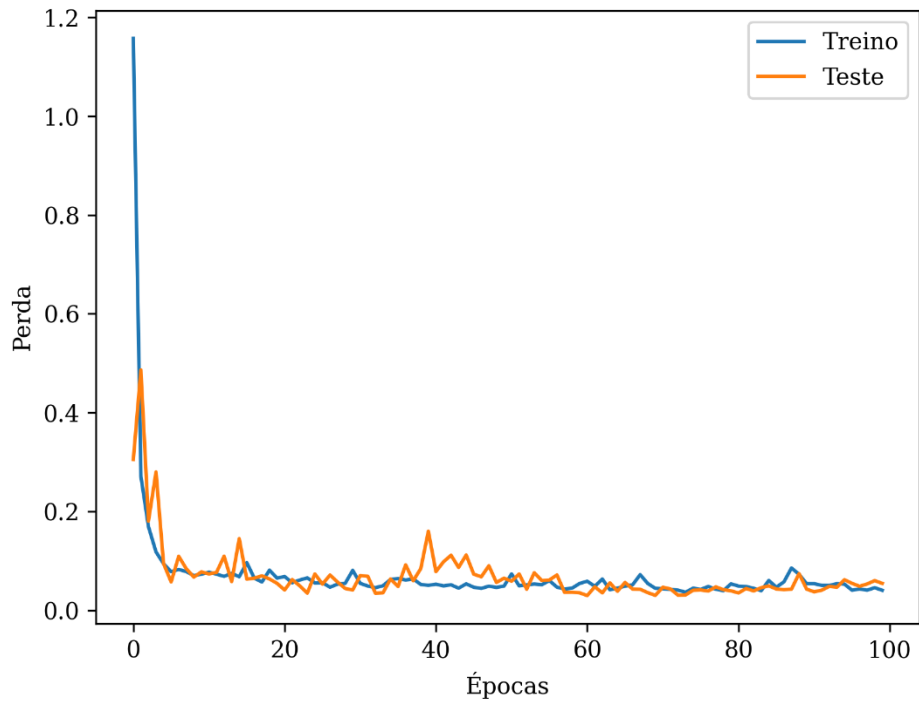


Fonte: o Autor.

Nos testes referentes ao banco Itaú, observou-se que o desempenho das redes com e sem os índices separadamente obtiveram resultados bem semelhantes, porém a rede com a exclusão dos três índices obteve o melhor resultado, perdendo apenas para o Random Walk.

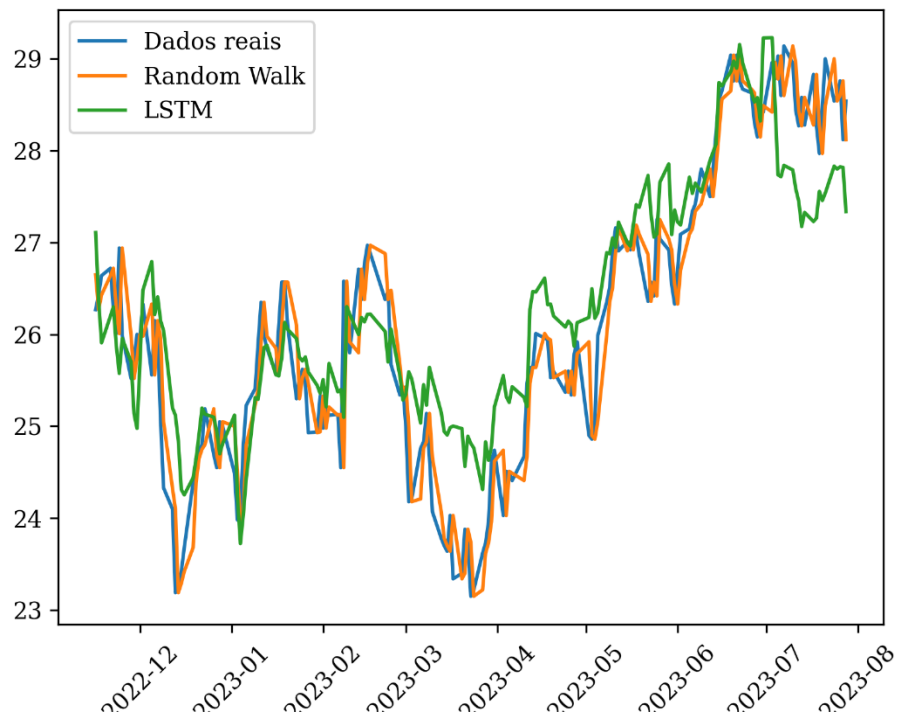
4.4.1– Rede neural LSTM Teste

Figura 35 - Gráfico de perda por época para a ação do Banco Itaú.



Fonte: O Autor.

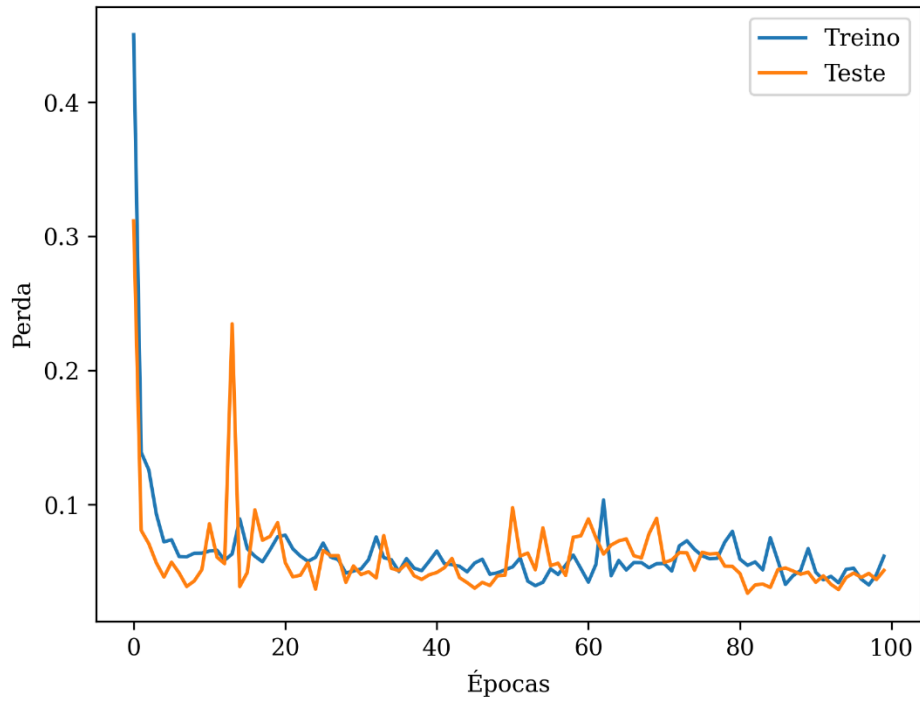
Figura 36 - Dados reais, Random Walk e valores previstos pela rede LSTM.



Fonte: Autor

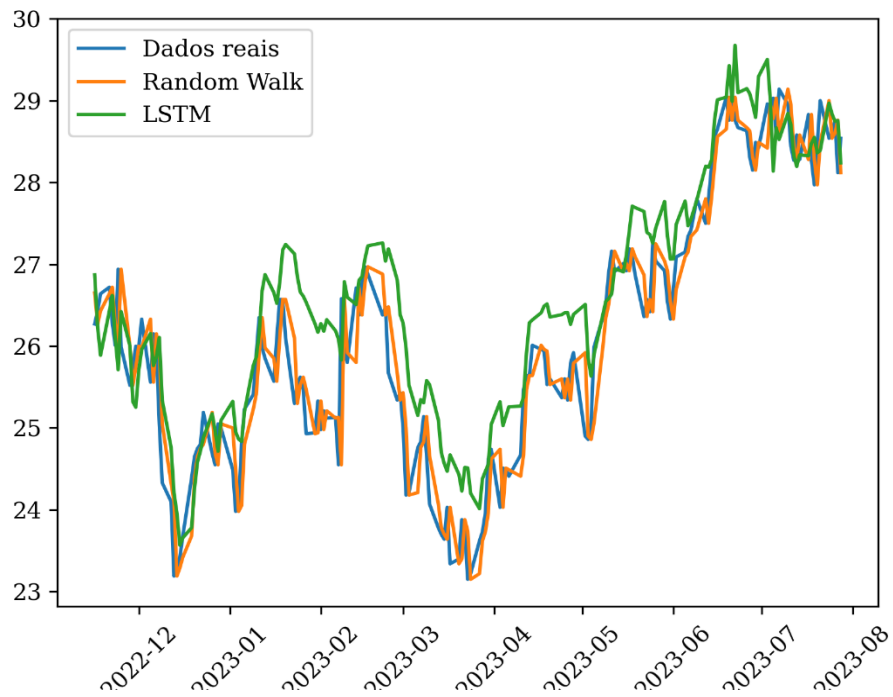
4.4.2– Rede neural LSTM Teste sem ISE

Figura 37 - Rede neural LSTM Teste sem ISE.



Fonte: O Autor.

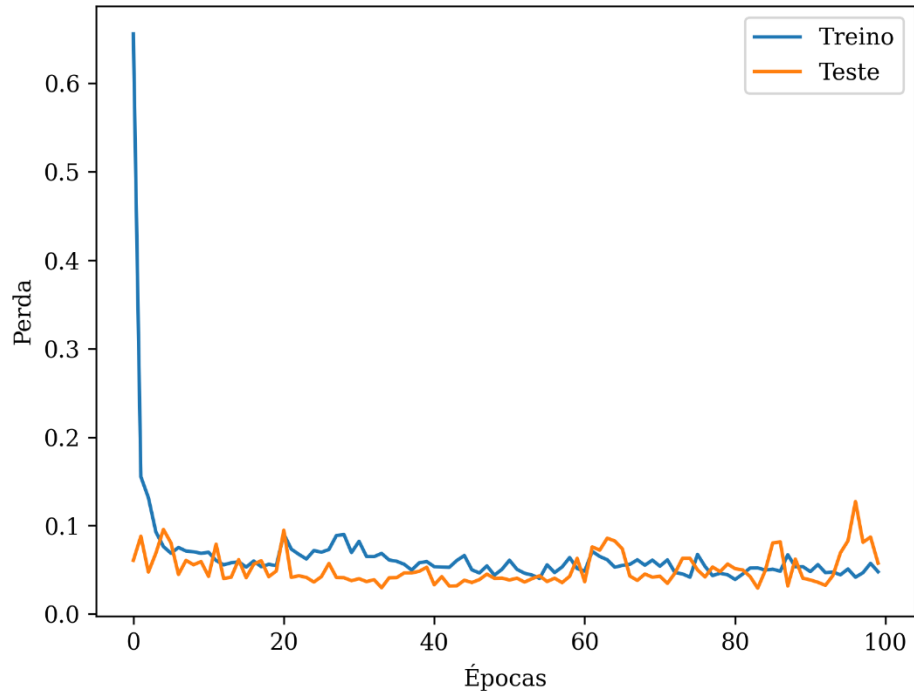
Figura 38 - Dados reais, Random Walk e valores previstos pela rede LSTM sem ISE.



Fonte: o Autor.

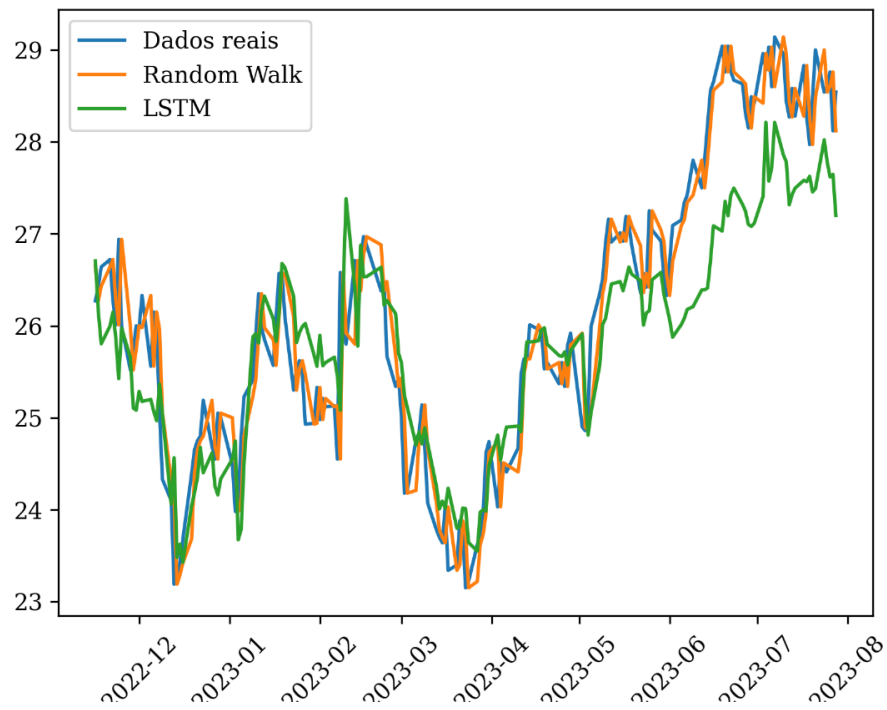
4.4.3 – Rede neural LSTM Teste sem IFNC

Figura 39 - Rede neural LSTM Teste sem IFNC.



Fonte: O Autor.

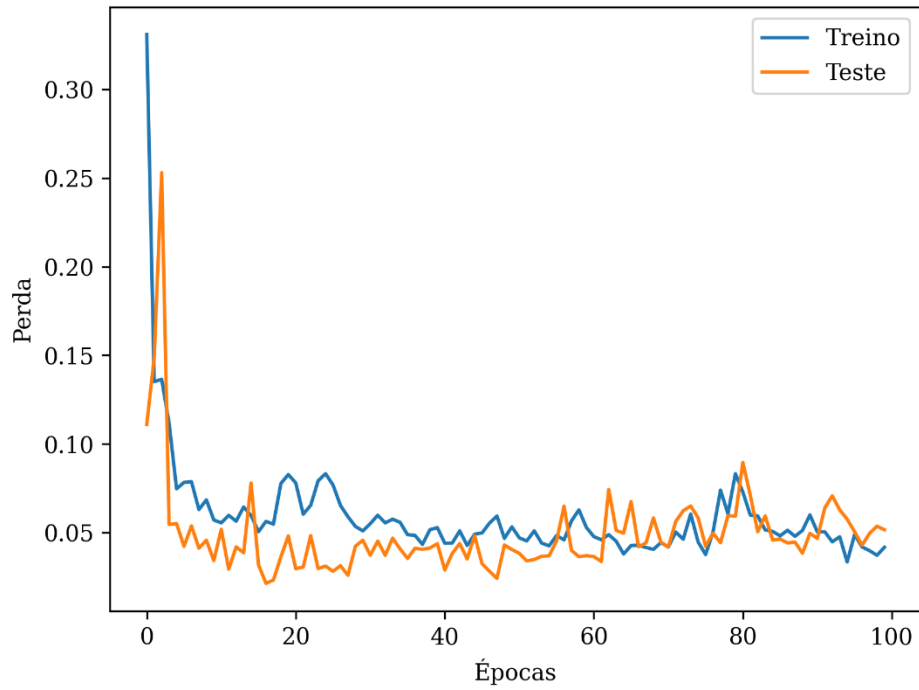
Figura 40 - Dados reais, Random Walk e valores previstos pela rede LSTM sem INFC.



Fonte: o Autor.

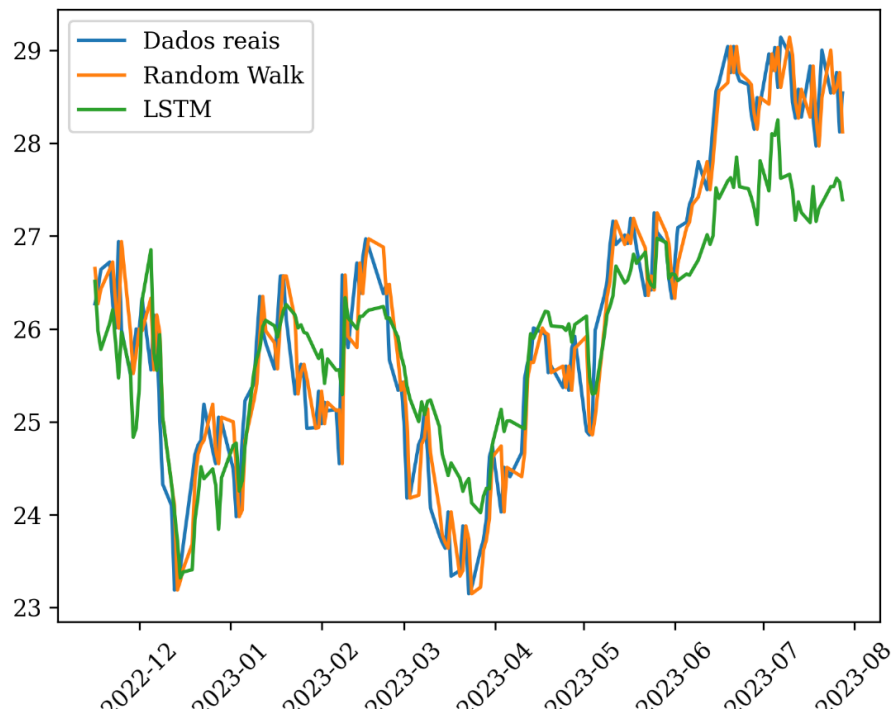
4.4.4 – Rede neural LSTM Teste sem BVSP

Figura 41 - Rede neural LSTM sem BVSP.



Fonte: O Autor.

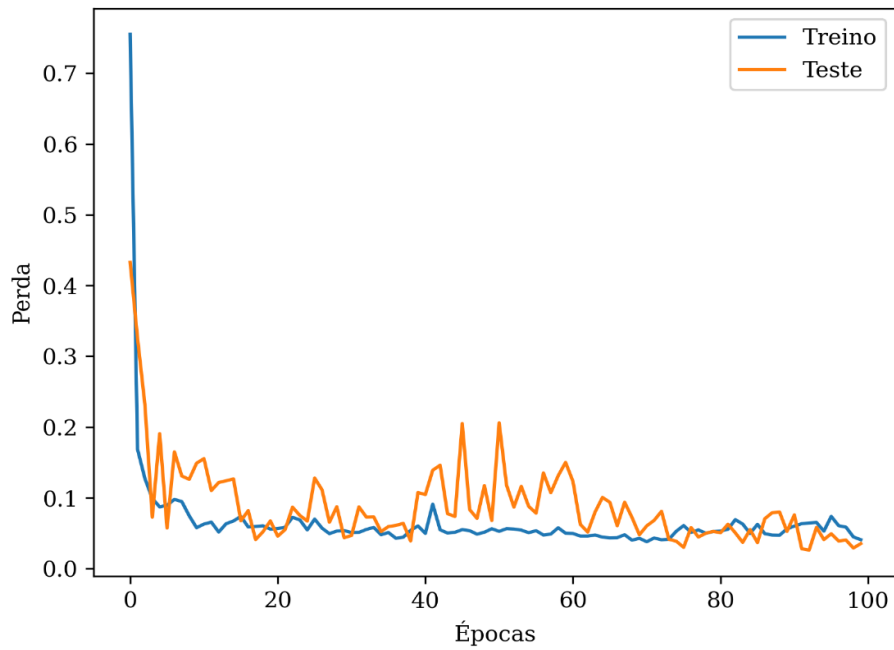
Figura 42 - Dados reais, Random Walk e valores previstos pela rede LSTM sem BVSP.



Fonte: o Autor.

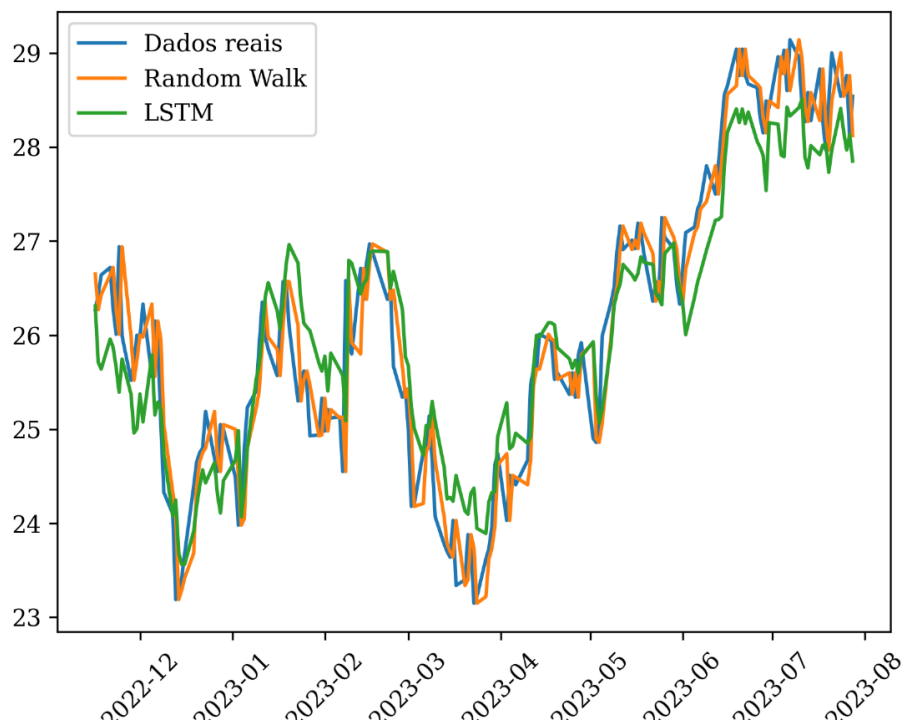
4.4.5– Rede neural LSTM Teste sem Índices

Figura 43 - Rede neural LSTM Teste sem Índices.



Fonte: O Autor.

Figura 44 - Dados reais, Random Walk e valores previstos pela rede LSTM sem índices.

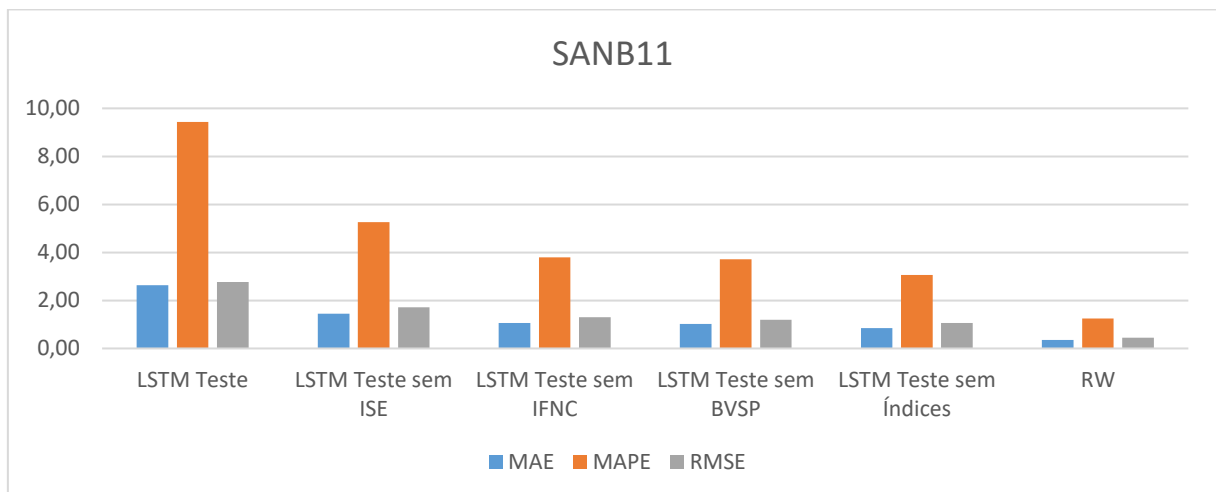


Fonte: o Autor.

Nesse caso constata-se que em todos os casos a rede apresentou uma tendência de baixa de perda durante as épocas iniciais, porém depois tendeu-se a estabilizar lateralmente. Isso foi refletido em valores relativamente próximos entre o valor de previsão e valor real da ação. No entanto nota-se que em alguns momentos o valor de previsão se distancia do valor real da ação, isso pode ser verificado com alguns pontos de elevação na linha de perda da rede de teste.

4.5 – LTSM aplicada as ações do Banco Santander

Figura 45 - LTSM aplicada as ações do Banco Santander.



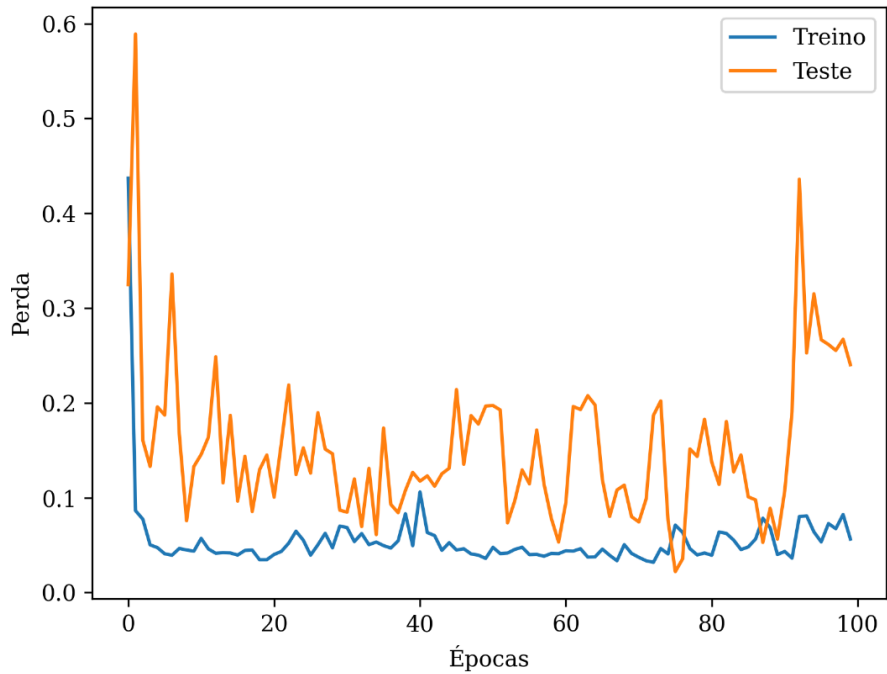
Fonte: Autor.

Em relação ao Banco Santander, a rede que obteve melhor resultado foi a rede neural sem a inclusão dos índices, comportamento esse bem semelhante ao que ocorreu com o Banco Bradesco. A exclusão do índice BVSP da rede obteve melhor desempenho em relação a exclusão do IFNC e ISE, respectivamente, e por último a rede neural com a inclusão de todos os índices obteve o pior desempenho.

4.5.1 – Resultados individuais das redes

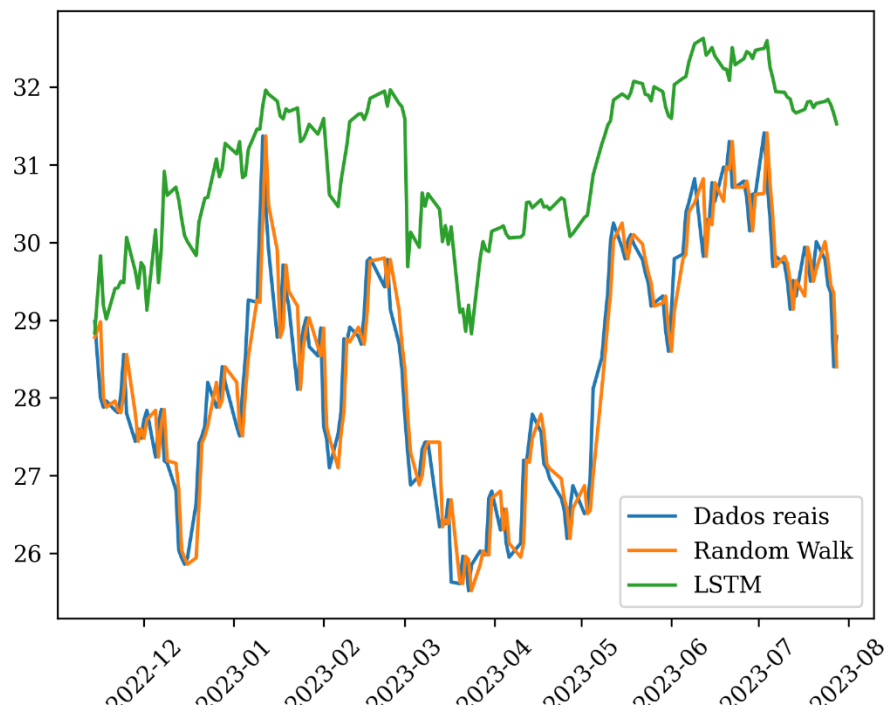
4.5.2 – Rede neural LSTM Teste

Figura 46 - Gráfico de perda por época para ações do Banco Santander



Fonte: O Autor.

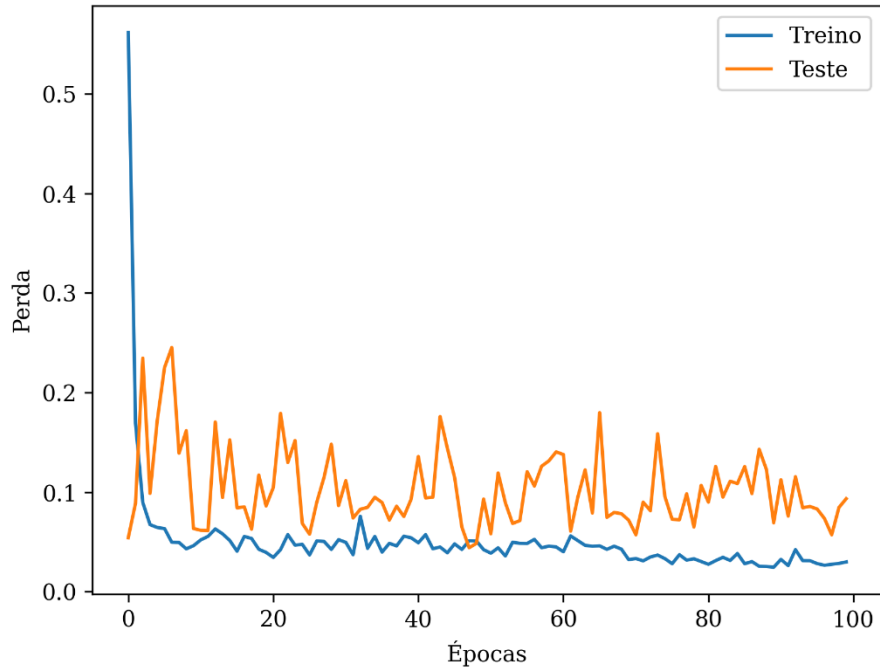
Figura 47 - Dados reais, Random Walk e valores previstos pela rede LSTM.



Fonte: Autor.

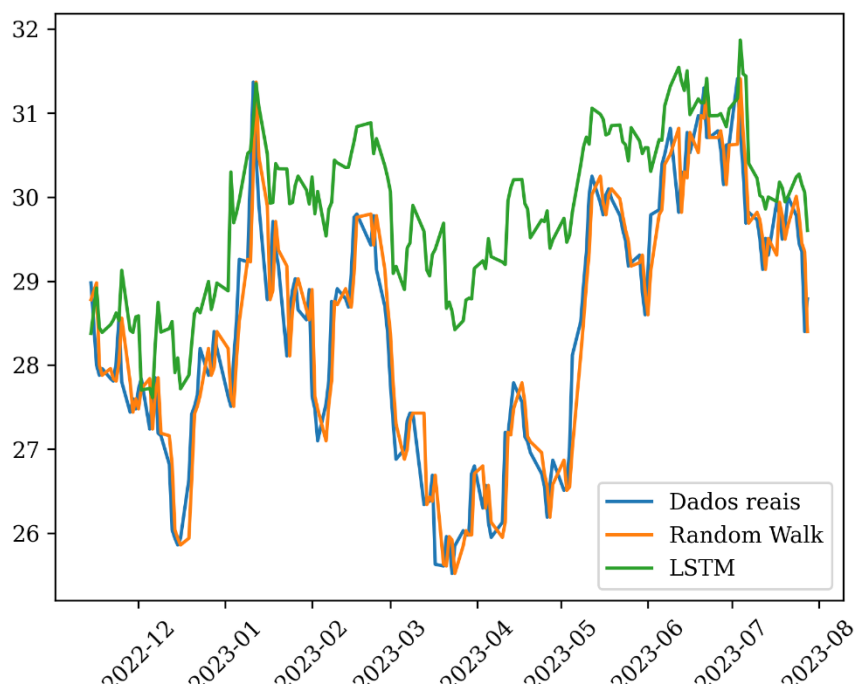
4.5.3 – Rede neural LSTM Teste sem ISE

Figura 48 - Rede neural LSTM Teste sem ISE.



Fonte: O autor.

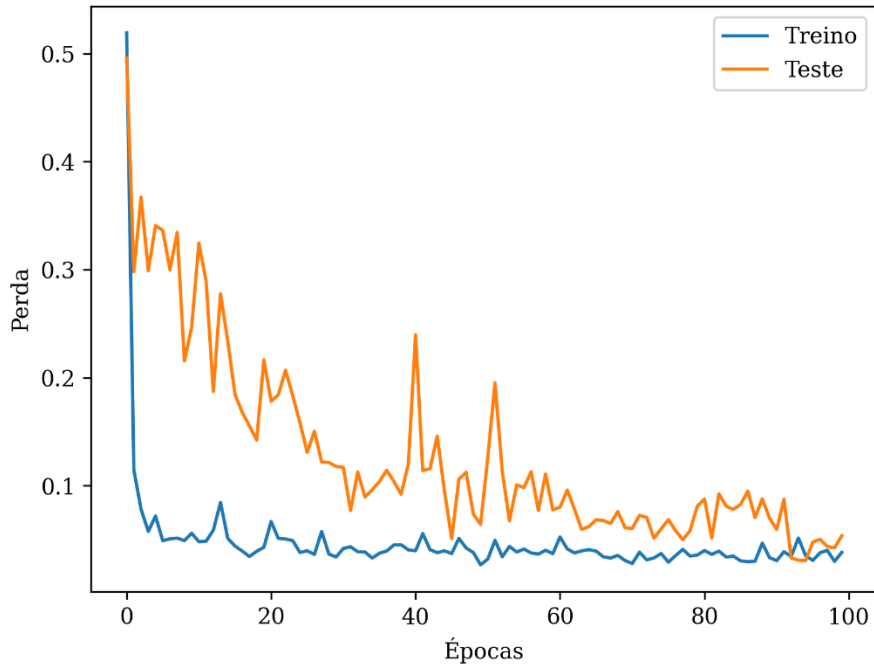
Figura 49 - Dados reais, Random Walk e valores previstos pela rede LSTM Teste sem ISE.



Fonte: Autor.

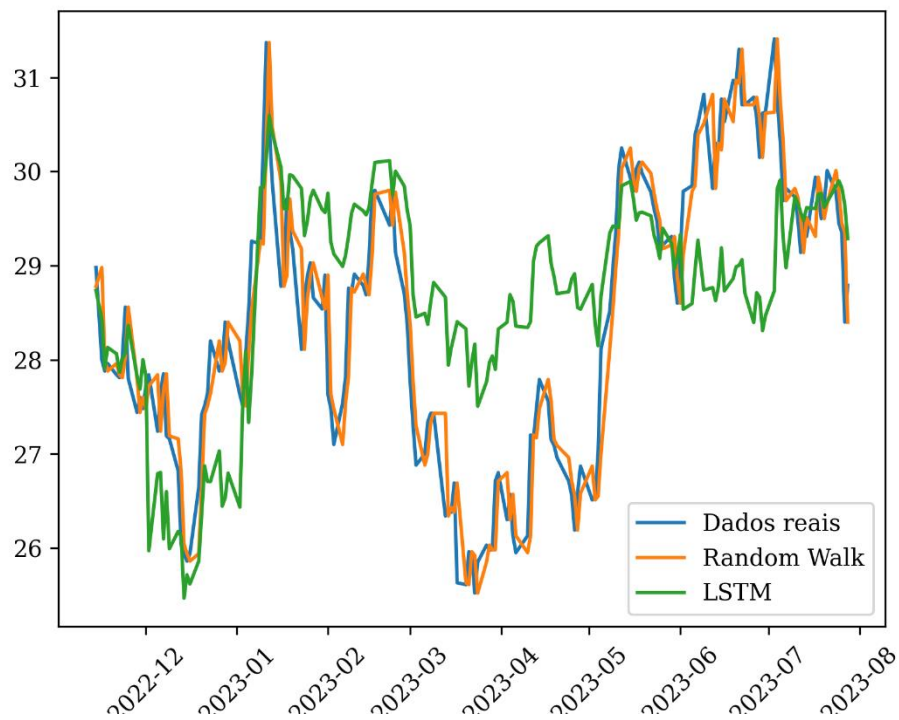
4.5.4 – Rede neural LSTM Teste sem IFNC

Figura 50 - Rede neural LSTM Teste sem IFNC.



Fonte: O Autor.

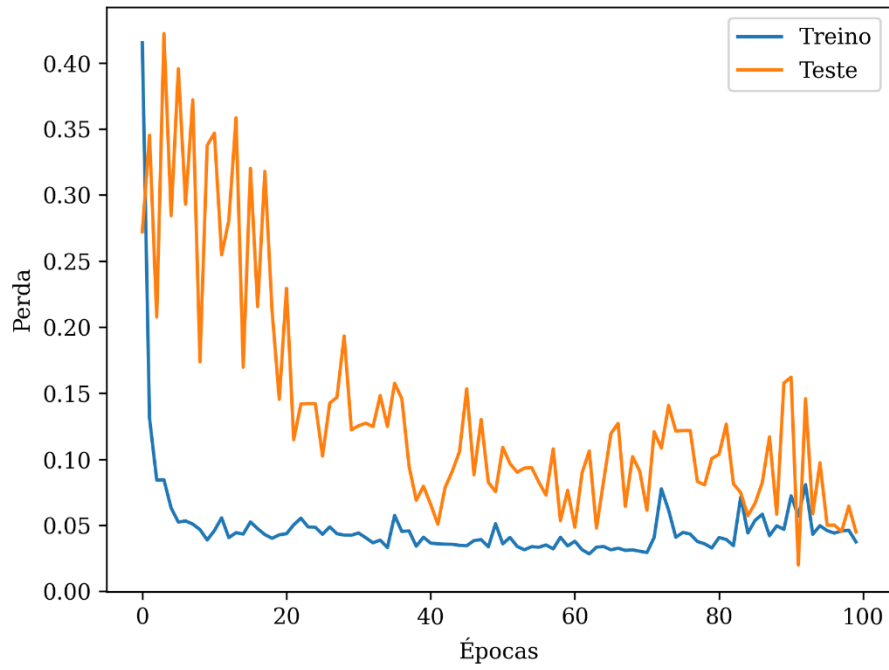
Figura 51 - Dados reais, Random Walk e valores previstos pela rede LSTM Teste sem IFNC.



Fonte: Autor.

4.5.4 – Rede neural LSTM Teste sem BVSP

Figura 52 - Rede neural LSTM Teste sem BVSP.



Fonte: O Autor.

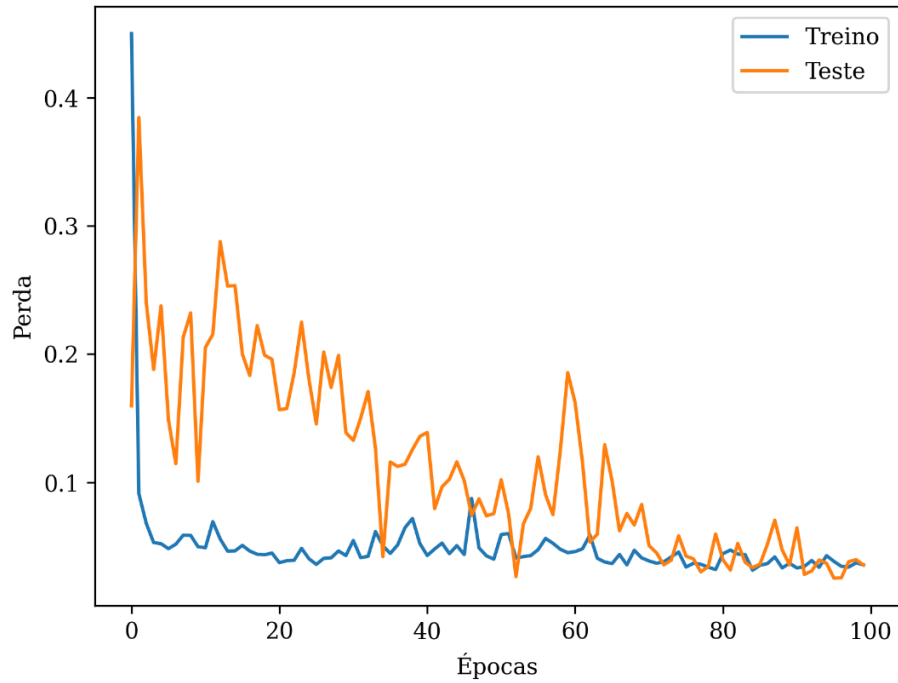
Figura 53 - Dados reais, Random Walk e valores previstos pela rede sem BVSP.



Fonte: Autor

4.5.5 – Rede neural LSTM Teste sem Índices

Figura 54 - Rede neural LSTM Teste sem Índices.



Fonte: O Autor.

Figura 55 - Dados reais, Random Walk e valores previstos pela rede LSTM sem índices.



Fonte: Autor

No caso do Banco Santander é possível notar que os valores de perda da rede não apresentam uma tendência clara de baixa, apresentando até mesmo momentos de alta relevante ao longo das épocas. Além disso o valor de perda se manteve alto durante boa parte do período analisado. Isso pode ser notado na discrepância relevante entre o valor previsto e o valor real da ação, visto que em somente um momento os dois valores se aproximam significativamente.

Já na rede que exclui somente o índice ISE, o valor de perda da rede tendeu-se a seguir lateralmente, sem melhoras significativas, isso é refletido em valores notavelmente distantes entre preço previsto e preço real da ação.

Por fim nas redes que excluem o índice IFNC, BVSP e todos os índices, verificou-se que o índice de perda apresentou uma tendência de queda ao longo das épocas, atingindo valores consideravelmente mais baixos em comparação as duas redes citadas anteriormente. Isso foi refletido em uma maior proximidade entre valor previsto e valor real da ação.

4.6 Resultados

Os resultados obtidos por meio da implementação de uma rede neural recorrente do tipo LSTM utilizando a biblioteca Keras em Python são notáveis e indicam a eficácia desse modelo em lidar com sequências temporais complexas. A flexibilidade e facilidade de uso oferecidas pelo Keras permitem uma construção intuitiva e rápida da arquitetura LSTM, proporcionando uma análise aprofundada de dados sequenciais. Ao treinar o modelo com conjuntos de dados apropriados, observou-se uma capacidade notável de capturar padrões temporais de longo prazo, demonstrando a adequação das LSTMs para tarefas que envolvem dependências temporais significativas.

Além disso, a natureza modular do Keras facilita a experimentação com diferentes arquiteturas e hiperparâmetros, otimizando assim o desempenho da rede. Esses resultados destacam a efetividade da combinação entre a simplicidade do Keras e a potência das LSTMs na modelagem de dados sequenciais, oferecendo uma abordagem robusta para problemas complexos de séries temporais.

Porém, constatou-se que nenhuma rede neural obteve valores de métricas (MAE, MAPE e RMSE) melhores que as do *benchmark* Random Walk. Essa constatação sugere que o modelo não está adequadamente ajustado para atingir o objetivo final de prever o preço da ação, considerando os cinco bancos estudados nesse trabalho. Essa inadequação é evidenciada nos valores de perda das redes, que não alcançaram valores consideravelmente baixos e constantes.

Verificou-se também que apenas no modelo considerando os dados do Banco do Brasil a inclusão dos índices ISE, IFNC e BVSP como variável de entrada melhoraram de forma relevante o desempenho da rede neural, atingindo valores menores das métricas MAPE, MPAE, RMSE e valores relativamente próximos entre valor previsto e valor real da ação.

5 CONCLUSÃO

Em conclusão, os resultados obtidos sugerem que a inclusão dos valores dos índices BVSP, IFNC e ISE como variáveis de entrada na rede neural LSTM construída não proporcionou uma melhoria de desempenho significativa, exceto para o caso do Banco do Brasil. Este resultado indica a complexidade na modelagem das relações entre os índices e as flutuações nas ações dos bancos brasileiros. Contudo, é importante ressaltar que o desempenho superior do Random Walk em relação à rede neural em diversos casos destaca a necessidade de uma avaliação mais cuidadosa da abordagem adotada.

Para trabalhos futuros, recomenda-se explorar diferentes arquiteturas e configurações para a rede neural LSTM, a fim de identificar a modelagem mais adequada para o contexto específico das ações dos bancos brasileiros. Além disso, considerar a inclusão de outras variáveis relevantes pode contribuir para melhorar a capacidade preditiva do modelo. Estudos comparativos entre diferentes algoritmos de previsão também podem oferecer insights valiosos sobre a escolha do método mais eficaz para esse tipo de previsão. Em suma, a continuidade da pesquisa exploratória é essencial para aprimorar a precisão e a confiabilidade das previsões em cenários dinâmicos como o mercado de ações.

REFERÊNCIAS

- BAO, W.; YUE, J.; RAO, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLOS ONE*, 12(7), e0180944.
- BATHA, Gourav. "Stock Price prediction using LSTM and SVR." **In 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)**, pp. 211-214. IEEE, 2020.
- BRUNI, A. L. "A eficiência informacional do mercado ADRs Brasileiros: Uma análise com teste de autocorrelação, raiz unitária e cointegração." In Revista de Gestão e Planejamento. Salvador, BA, 2004. Jan/Jun. n. 9, p. 53-65.
- DEEPAK, R. S. e MALATHI, D.. "Machine learning approach in stock market prediction." *International Journal of Pure and Applied Mathematics*. 115, no. 8 (2017): 71-77.
- FENG, F.; WANG, X. He, X.; LUO, C.; LIU, Y.; CHUA, T.S. "Temporal relational ranking for stock prediction", *ACM Transactions on Information Systems (TOIS)*, 37(2), pp.1-30,2019.
- GLANTZ, M.; KISSELL, R. **Multi-asset Risk Modeling**, 1st ed. Academic Press, 2013.
- HOCHREITER, S., & SCHMIDHUBER, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- HUANG, Y. Nakamori; WANG, S.Y. "Forecasting stock market movement direction with support vector machine", *Computers & operations research*, 32(10), pp.2513-2522, 2005.
- Kim, T. e Ha Y. K. "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data." *PloS one* 14, no. 2 (2019): e0212320.
- KIM, J. H.; SHAMSUDDIN, A. "Are Asian Stock Markets efficient ? Evidence from new multiple variance ratio tests." In Journal of Empirical Finance. Elsevier Inc. 2008. vol. 15, p. 518-532
- LOKE, K. S. "Impact of financial ratios and technical analysis on stock price prediction using random forests." **In 2017 International Conference on Computer and Drone Applications (IConDA)**, pp. 38-42. IEEE, 2017.
- MODRO, Wilton Moisés; SANTOS, José Odálio dos. A Relação Entre o Retorno das Ações Ordinárias, Métricas de Desempenho e Fatores Econômicos: Um Estudo dos três Principais Bancos Brasileiros entre 2001 e 2010. 2015. 26 f. Monografia (Especialização) - Curso de Administração, Pontifícia Universidade Católica de São Paulo, São Paulo, 2013.
- NASCIMENTO, Luiz Paulo do. Elaboração de Projeto de Pesquisa. São Paulo: Cengage Learning Ltda, 2012.
- NELSON, D. M. Q.; PEREIRA, A. C. M.; DE OLIVEIRA, R. A. (2017). "Stock market's price movement prediction with LSTM neural networks." **In: International Joint Conference on Neural Networks (IJCNN)**, 2017. IEEE.

PATTERSON, J. 2017. Deep Learning: A Practitioner's Approach, **O'Reilly Media**.

PEREIRA, Cleverson Luiz. Mercado de Capitais. Curitiba: **Editora Intersaberes**, 2013

PRAMOD, B. S.; MALLIKARJUNA, Shastry Pm. "Stock price prediction using LSTM." *Test Engineering and Management*.83 (2020): 5246-5251.

ROZZANI, Nabilah; RAHMAN, Rashidah Abdul. **Camels and performance evaluation of banks in Malaysia: conventional versus Islamic**. Journal of Islamic Finance and Business Research, v. 2, n. 1, p. 36-45, 2013.

SELVIN, S.; VINAYAKUMAR, R.; GOPALAKRISHNAN, E. A.; MENON, V. K.; SOMAN, K. P. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. 2017. **International Conference on Advances in Computing, Communications and Informatics (ICACCI)**. pp. 1643-1647. IEEE, 2017.

SIAMÍ-NAMINI, S.; NAMIN, A. S. (2018). "Forecasting economics and financial time series: Arima vs. lstm." arXiv preprint arXiv:1803.06386.

IZIDORO, Cleyton. Mercado de Capitais. São Paulo: **Pearson Education do Brasil**, 2015.