

UNIVERSIDADE PRESBITERIANA MACKENZIE

SÉRGIO TORRES

**PREVISÃO DO PREÇO DE AÇÕES BRASILEIRAS UTILIZANDO REDES
NEURAS ARTIFICIAIS**

São Paulo
2021

SÉRGIO TORRES

**PREVISÃO DO PREÇO DE AÇÕES BRASILEIRAS UTILIZANDO REDES
NEURAS ARTIFICIAIS**

Dissertação apresentada ao Programa de Pós Graduação em Administração de Empresas da Universidade Presbiteriana Mackenzie como requisito parcial para a obtenção do título de Mestre em Administração de Empresas.

Orientador: Prof. Dr. Eli Hadad Júnior

São Paulo
2021

T693p Torres, Sergio.

Previsão do preço de ações brasileiras utilizando redes neurais artificiais. / Sergio Torres.

51 f.; il. ; 30 cm

Dissertação (Mestrado em Administração de Empresas) –
Universidade Presbiteriana Mackenzie, São Paulo, 2021.

Orientador: Eli Hadad Junior.

Referências bibliográficas: f. 49-51.

1. GRU, 2. LSTM, 3. Machine learning, 4. MDA,
5. Redes Neurais Artificiais. I. Hadad Junior, Eli, *orientador*.
II. Título.

CDD 332.6322

Bibliotecária Responsável: Jaqueline Bay Inacio Duarte - CRB 8/9509

SÉRGIO TORRES

**PREVISÃO DO PREÇO DE AÇÕES BRASILEIRAS UTILIZANDO REDES
NEURAS ARTIFICIAIS**

Dissertação apresentada ao Programa de Pós Graduação em
Administração de Empresas da Universidades Presbiteriana
Mackenzie como requisito parcial para à obtenção do título
de Mestre em Administração de Empresas.

Aprovada em 06 de Agosto de 2021.

BANCA EXAMINADORA



Prof. Dr. Eli Hadad Junior

Universidade Presbiteriana Mackenzie



Prof. Dr. Leonardo Fernando da Crnz Basso

Universidade Presbiteriana Mackenzie



Prof. Dr. Diogo de Prince Mendonça

Universidade Federal de São Paulo

AGRADECIMENTOS

A Deus, por ter permitido que eu tivesse saúde e determinação para não desanimar durante a realização deste trabalho, visto que passamos por um período extremamente delicado devido a pandemia.

Agradeço ao Prof. Eli Hadad Junior pela paciência e pela contribuição no desenvolvimento desta pesquisa.

Aos professores do programa de Administração de Empresas da Universidade Presbiteriana Mackenzie pelas contribuições e ensinamentos que possibilitaram novas visões para as realizações de trabalhos científicos.

Aos meus colegas de curso por compartilharem comigo tantos momentos de aprendizado, conhecimento e descobertas que solidificaram ainda mais nossos conceitos em finanças.

A empresa Capricórnio Têxtil S.A, por ter aberto as portas para essa nova fase na minha carreira profissional e acadêmica, patrocinando grande parte do curso. Empresa com visão estratégica e com grande crescimento no cenário nacional na qual tive orgulho de trabalhar e contribuir com o meu conhecimento. Gostaria de fazer um agradecimento especial ao CEO da empresa Gustavo Manfredini.

A minha família por toda paciência, ajuda e apoio nestes momentos que fiquei ausente mas que foram importantes para a finalização deste trabalho.

“Todos podem ver as táticas de minhas conquistas, mas ninguém consegue discernir a estratégia que gerou as vitórias”

(Sun Tzu)

RESUMO

A análise do mercado de ações é influenciada pela tendência de alta e de queda do preço dos ativos em relação a curtos períodos de tempo, muitas vezes em virtude de sentimentos e expectativas de investidores, que podem afetar a movimentação das ações. Uma das possibilidades de maximizar os retornos na previsão de ações é prever o movimento direcional dos ativos em vez de prever os preços. Este artigo analisa o uso de redes neurais artificiais para prever o movimento direcional dos ativos e compara o desempenho de duas arquiteturas de redes neurais: LSTM e GRU. O índice Ibovespa e algumas ações são analisados em diversos horizontes de tempo visando a descoberta de padrões específicos. O desempenho das referidas arquiteturas é muito semelhante entre si: para horizontes de tempo curtos a arquitetura LSTM apresentou melhor precisão, e para horizontes de tempo longos a GRU é mais precisa que a LSTM.

Palavras-chave: GRU, LSTM, *machine learning*, MDA, Redes Neurais Artificiais

ABSTRACT

The stock market analysis is influenced by the upward and downward trend related to the asset prices in relation to short periods. Due to the feelings and the expectations of investors, which can affect the movement of shares. One of the possibilities to maximize the returns in stock forecast is to predict the directional asset movement instead of forecasting the prices. This paper analyses the use of artificial neural networks to preview the directional asset movement and compares the performances of two architectures: LSTM and GRU. Shares and Ibovespa index are analyzed in several time horizons aiming at the observations of specific patterns. The performance of the said architectures is very similar to each other: for short time horizons, the LSTM architecture presented a better accuracy, and for long time horizons, the GRU is more accurate than LSTM.

Keywords: Artificial Neural Network, GRU, LSTM, MDA, machine learning

LISTA DE ILUSTRAÇÕES

Figura 1: Diagrama de uma rede perceptron	19
Figura 2: Arquitetura LSTM	20
Figura 3: Arquitetura GRU.....	21
Figura 4: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 1 ano	25
Figura 5: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 2 anos.....	25
Figura 6: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 5 anos.....	25
Figura 7: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 10 anos.....	26
Figura 8: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 20 anos.....	26
Figura 9: Previsão BBAS3 para 60 dias - horizonte de tempo de 1 ano.....	27
Figura 10: Previsão BBAS3 para 60 dias - horizonte de tempo de 2 anos	28
Figura 11: Previsão BBAS3 para 60 dias - horizonte de tempo de 5 anos	28
Figura 12: Previsão BBAS3 para 60 dias - horizonte de tempo 10 anos.....	28
Figura 13: Previsão ITUB4 para 60 dias - horizonte de tempo de 1 ano.....	29
Figura 14: Previsão ITUB4 para 60 dias - horizonte de tempo de 2 anos	29
Figura 15: Previsão ITUB4 para 60 dias - horizonte de tempo de 5 anos	30
Figura 16: Previsão ITUB4 para 60 dias - horizonte de tempo de 10 anos	30
Figura 17: Previsão ITUB4 para 60 dias - horizonte de tempo de 20 anos	30
Figura 18: Previsão BBDC4 para 60 dias - horizonte de tempo de 1 ano	31
Figura 19: Previsão BBDC4 para 60 dias - horizonte de tempo de 2 anos.....	31
Figura 20: Previsão BBDC4 para 60 dias - horizonte de tempo de 5 anos.....	31
Figura 21: Previsão BBDC4 para 60 dias - horizonte de tempo de 10 anos.....	32
Figura 22: Previsão BBDC4 para 60 dias - horizonte de tempo de 20 anos.....	32
Figura 23: Previsão PETR4 para 60 dias - horizonte de tempo de 1 ano	33
Figura 24: Previsão PETR4 para 60 dias - horizonte de tempo de 2 anos.....	33
Figura 25: Previsão PETR4 para 60 dias - horizonte de tempo de 5 anos.....	34
Figura 26: Previsão PETR4 para 60 dias - horizonte de tempo de 10 anos.....	34
Figura 27: Previsão PETR4 para 60 dias - horizonte de tempo de 20 anos.....	34
Figura 28: Previsão BRKM5 para 60 dias - horizonte de tempo de 1 ano	35
Figura 29: Previsão BRKM5 para 60 dias - horizonte de tempo de 2 anos.....	35
Figura 30: Previsão BRKM5 para 60 dias - horizonte de tempo de 5 anos.....	35
Figura 31: Previsão BRKM5 para 60 dias - horizonte de tempo de 10 anos.....	36
Figura 32: Previsão BRKM5 para 60 dias - horizonte de tempo de 20 anos.....	36
Figura 33: Previsão VALE3 para 60 dias - horizonte de tempo de 1 ano.....	37
Figura 34: Previsão VALE3 para 60 dias - horizonte de tempo de 2 anos	37
Figura 35: Previsão VALE3 para 60 dias - horizonte de tempo de 5 anos	37
Figura 36: Previsão VALE3 para 60 dias - horizonte de tempo de 10 anos	37
Figura 37: Previsão GGBR4 para 60 dias - horizonte de tempo de 1 ano	38
Figura 38: Previsão GGBR4 para 60 dias - horizonte de tempo de 2 anos.....	39
Figura 39: Previsão GGBR4 para 60 dias - horizonte de tempo de 5 anos.....	39
Figura 40: Previsão GGBR4 para 60 dias - horizonte de tempo de 10 anos.....	39
Figura 41: Previsão GGBR4 para 60 dias - horizonte de tempo de 20 anos.....	39
Figura 42: Previsão EGIE3 para 60 dias - horizonte de tempo de 1 ano	40
Figura 43: Previsão EGIE3 para 60 dias - horizonte de tempo de 2 anos.....	41
Figura 44: Previsão EGIE3 para 60 dias - horizonte de tempo de 5 anos.....	41
Figura 45: Previsão EGIE3 para 60 dias - horizonte de tempo de 10 anos.....	41
Figura 46: Previsão EGIE3 para 60 dias - horizonte de tempo de 20 anos.....	41
Figura 47: Previsão CPLE6 para 60 dias - horizonte de tempo de 1 ano	42
Figura 48: Previsão CPLE6 para 60 dias - horizonte de tempo de 2 anos.....	43
Figura 49: Previsão CPLE6 para 60 dias - horizonte de tempo de 5 anos.....	43
Figura 50: Previsão CPLE6 para 60 dias - horizonte de tempo de 10 anos.....	43
Figura 51: Previsão CPLE6 para 60 dias - horizonte de tempo de 20 anos.....	43

Figura 52: Previsão VIVT4 para 60 dias - horizonte de tempo de 1 ano.....	44
Figura 53: Previsão VIVT4 para 60 dias - horizonte de tempo de 2 anos	45
Figura 54: Previsão VIVT4 para 60 dias - horizonte de tempo de 5 anos	45
Figura 55: Previsão VIVT4 para 60 dias - horizonte de tempo de 10 anos	45
Figura 56: Previsão VIVT4 para 60 dias - horizonte de tempo de 20 anos	45
Figura 57: Previsão TIMP4 para 60 dias - horizonte de tempo de 1 ano.....	46
Figura 58: Previsão TIMP4 para 60 dias - horizonte de tempo de 2 anos	47
Figura 59: Previsão TIMP4 para 60 dias - horizonte de tempo de 5 anos	47
Figura 60: Previsão TIMP4 para 60 dias - horizonte de tempo de 10 anos	47
Figura 61: Previsão TIMP4 para 60 dias - horizonte de tempo de 20 anos	47

LISTA DE TABELAS

Tabela 1: Horizonte de tempo	23
Tabela 2: Comparativo das RNAs por horizonte de tempo	24
Tabela 3: Ações por segmento	26
Tabela 4: Comparativo das ações BBAS3 por horizonte de tempo	27
Tabela 5: Comparativo das ações ITUB4 por horizonte de tempo	28
Tabela 6: Comparativo das ações BBDC4 por horizonte de tempo	31
Tabela 7: Comparativo das ações PETR4 por horizonte de tempo	33
Tabela 8: Comparativo das ações BRKM5 por horizonte de tempo	35
Tabela 9: Comparativo das ações VALE3 por horizonte de tempo	36
Tabela 10: Comparativo das ações GGBR4 por horizonte de tempo	38
Tabela 11: Comparativo das ações EGIE3 por horizonte de tempo	40
Tabela 12: Comparativo das ações CPLE6 por horizonte de tempo	42
Tabela 13: Comparativo das ações VIVT4 por horizonte de tempo	44
Tabela 14: Comparativo das ações TIMP4 por horizonte de tempo	46

LISTA DE SIGLAS

IA	Inteligência Artificial
IBOV	Índice Ibovespa
GRU	Gated Recorrent Unit
LSTM	Long Short Term Memory
ML	Machine Learning
MDA	Mean Directional Accuracy
MLP	Multi Layer Perceptron
MSE	Mean Square Error
RNA	Redes Neurais Artificiais
SVM	Suport Vector Machine

SUMÁRIO

1 INTRODUÇÃO	12
2 REFERENCIAL TEÓRICO	14
3 METODOLOGIA.....	17
3.1 REDES NEURAS ARTIFICIAIS	17
3.2 REDES NEURAS RECORRENTES.....	19
3.3 REDES NEURAS LONG SHORT MEMORY TERM (LSTM).....	19
3.4 ARQUITETURA DE REDES NEURAS GATED RECURRENT UNIT (GRU)	20
3.5 ANÁLISE DE EFICIÊNCIA	21
3.6 MÉTODOS DE OTIMIZAÇÃO	22
3.7 SÉRIES FINANCEIRAS	23
4 RESULTADOS	24
5 CONCLUSÕES.....	48
REFERÊNCIAS	49

1 INTRODUÇÃO

O mercado de ações é um ambiente dinâmico, público, organizado e regulamentado em que potenciais investidores negociam ações e outros valores mobiliários. As ações são títulos negociáveis que representam a parcela mínima de uma empresa, que busca financiar seus investimentos no mercado secundário, diminuindo o custo do capital e alavancando as atividades que desenvolvem.

No curto prazo a tendência de alta e baixa do preço dos ativos é influenciada por sentimentos e expectativas dos investidores que acabam impactando no movimento direcional das ações. As séries financeiras apresentam períodos nos quais os preços sofrem grandes oscilações por um período de tempo prolongado seguidos por períodos nos quais há relativa calma (GUJARATI e PORTER, 2006). Ainda segundo o autor o desafio em modelar uma série temporal dos preços das ações é que elas não são estacionárias, ou seja, as propriedades estatísticas da série como média, desvio padrão e correlação mudam no transcorrer do tempo.

Uma das possibilidades para se maximizar o retorno na previsão de séries é prever o movimento direcional acentuado ao invés do preço das ações. Para Markowitz (2002) um investidor racional busca minimizar o risco para um dado nível de retorno. Uma previsão correta, do movimento direcional, pode auxiliar o investidor a maximizar lucros e minimizar possíveis riscos de perdas.

O que podemos observar é que as variações futuras dos preços das ações são independentes em relação as variações passadas. Em um mercado eficiente todos os investidores são racionais e avaliam os investimentos de forma racional. Se existirem investidores não racionais as negociações devem ser aleatórias e não devem alterar os preços dos títulos. No caso dos investidores não racionais percebe-se a existência dos arbitradores que acabam por eliminar a influência de investidores não racionais nos preços dos títulos de mercado (SILVA, 2001).

Ao longo dos anos modelos, novas ferramentas, e tecnologias estão sendo propostos para prever tendências no mercado financeiro. A análise técnica e fundamentalista são algumas das estratégias de negociação utilizadas no mercado financeiro para realizar previsões de séries. Segundo Medeiros (2018), o avanço da inteligência artificial (IA) possibilitou o desenvolvimento de estruturas cognitivas de aprendizado. Essas estruturas consideram o ambiente, a capacidade de resolução de problemas complexos, e a orientação de metas. No

nosso dia a dia notamos, esse avanço, em muitas atividades como por exemplo, o reconhecimento de voz e de escrita, jogos, robôs, etc.

O sistema lógico dessas máquinas são cadeias e conjuntos de regras desenvolvidos para resolver tarefas complexas. Os modelos de aprendizado de máquina analisam grandes quantidades de informações de maneira rápida e automática e gera resultados em grande escala. As redes neurais artificiais (RNA) são algoritmos computacionais inspirados na estrutura do cérebro humano e são capazes de realizar o aprendizado bem como o reconhecimento de padrões.

As RNAs podem ser consideradas um dos adventos mais significativos da área da inteligência artificial e o seu funcionamento possibilita o aprendizado de padrões com base na complexidade de elementos que simulam o comportamento dos neurônios (MEDEIROS, 2018). Uma de suas principais características é a capacidade de aprender por meio de aprendizado e de generalizar a informação aprendida, gerando um modelo não-linear, tornando sua aplicação na análise bastante eficiente. Para um bom desempenho das redes neurais o treinamento é fundamental.

Segundo Tkác e Verner (2016) ao contrário das técnicas convencionais de previsão, as RNAs têm a capacidade de resolver problemas não lineares, com várias soluções. Os modelos são reconhecidos pela eficiência, robustez e adaptabilidade na classificação e suporte nas análises financeiras.

Diante do grande número de algoritmos e diferentes tipos de RNAs disponíveis, este trabalho contribui com os estudos de *machine learning* no sentido de encontrar tipologias que melhoram a performance dos modelos. As séries financeiras são voláteis e estudos relacionados com o movimento direcional podem ser úteis para os investidores quando comparado com a previsão dos preços. Os modelos, resultantes dessas séries, tendem a ser no futuro rápidos, otimizados, precisos e com baixo uso de memória computacional. Esse trabalho, investiga essas tendências, com a possibilidade de novos estudos para a melhoria contínua das previsões.

O objetivo é usar as RNAs para prever o movimento direcional dos ativos financeiros, utilizando um modelo que será treinado com séries de tempo de ações listadas na B3 S.A. Brasil, Bolsa, Balcão (B3) com o intuito de reconhecer padrões de direção dos ativos. A otimização no processo de predição, usando Python e as bibliotecas Keras e Tensor Flow são diferenciais que permitem agilidade nas estimativas e nas análises do movimento direcional do preço das ações.

2 REFERENCIAL TEÓRICO

Moghaddam, Moghaddam e Esfandyaric (2006) investigaram a capacidade da rede neural artificial (RNA) em prever a taxa de câmbio diária da NASDAQ. As redes neurais *feed forward* foram treinadas com o algoritmo *backpropagation*, considerando um histórico de curto prazo para o treinamento das RNAs. Guresen, Kayakutlu e Daimb (2011) realizaram um comparativo de diferentes estruturas de redes neurais artificiais para a previsão do fechamento diário do índice NASDAQ. Os autores analisaram os modelos *perceptron* multicamadas (MLP), rede neural artificial dinâmica (DAN2) e redes neurais híbridas que usam heteroscedasticidade condicional autoregressiva generalizada (GARCH). O modelo MLP apresentou os melhores resultados.

Choudhury e Sen (2017) comparam diferentes técnicas de aprendizado de máquina para a previsão do preço das ações e estabeleceram um comparativo entre os pros e contras da utilização dessas técnicas. As redes neurais artificiais utilizadas para desenvolver os modelos foram treinadas usando o algoritmo *backpropagation*. Persio e Honchar (2017) compararam diferentes arquiteturas de redes neurais recorrentes: redes neurais multicamadas, LSTM e *gated recurrent unit* (GRU) na previsão dos movimentos dos preços das ações do Google, em diferentes horizontes temporais. Os resultados apontaram que a arquitetura LSTM superou as outras técnicas em um horizonte de análise de 5 dias.

Gu, Kelly, e Xiu (2018), realizaram uma pesquisa comparativa de métodos de aprendizado de máquina para resolver o problema canônico da precificação de ativos de acordo com o prêmio de risco. Os métodos de aprendizado de máquina podem ajudar a melhorar a compreensão empírica dos preços dos ativos. As redes neurais, em menor grau, e *random forest* foram os métodos que apresentaram o melhor desempenho. Dametto (2018) comparou e combinou três tipos de algoritmos de aprendizagem de máquina: método *ensemble* de RNA com redes *Multilayer Perceptrons* (MLP), auto regressiva com entradas exógenas (NARX) e *Long Short-Term Memory* (LSTM) para predição do Índice Bovespa. Os resultados das métricas indicam bom desempenho de predição pelo modelo *ensemble* obtendo 70% de acerto no movimento do índice, porém, não conseguiu atingir melhores resultados que as redes MLP e NARX, ambas com 80% de acerto.

Atkins, Niranjana e Gerding (2018) utilizaram informações extraídas de fontes de notícias para prever a direção do movimento dos ativos subjacentes e suas estatísticas de segunda ordem nos índices de ações norte americano. O estudo compara modelos de

aprendizado de máquina *Latent Dirichlet Allocation* para representar informações de *feeds* de notícias e classificadores de Bayes para prever a direção dos movimentos. Os autores concluíram que os movimentos de volatilidade são mais previsíveis do que os movimentos dos preços dos ativos ao usar notícias financeiras como insumo de aprendizado de máquina.

Rasekhschaffe e Jones (2019), utilizaram técnicas de aprendizado de máquina para prever o retorno das ações limitando o risco de *overfitting*. Segundo os autores os algoritmos de aprendizado de máquina podem descobrir relacionamentos contextuais e não lineares. O *overfitting* representa um desafio quando se tenta extrair informações de dados históricos com ruído. Selvamuthu, Kumar e Mishra (2019) utilizaram três arquiteturas de redes neurais: Levenberg-Marquardt, gradiente conjugado escalonado e regularização Bayesiana para previsão do mercado de ações com base em dados observados em 15 minutos de uma empresa indiana. As três técnicas apresentaram a mesma performance.

Chin-Sheng Huang (2019) comparou quatro modelos de previsão: redes neurais artificiais, *support vector machine* (SVM), *random forest* e naive-Bayes para prever a direção do movimento do índice de preços das ações de Taiwan, com duas abordagens de entrada. As redes neurais apresentaram performance superior aos demais modelos. Os resultados mostram que o desempenho dos modelos melhora quando os parâmetros são representados como dados determinísticos de tendência binária. Chen, Liu e Hsu (2019) analisaram a previsão do movimento do preço das ações de Taiwan no desempenho de investimentos. Com base no índice da bolsa de valores de Taiwan, o estudo comparou quatro modelos de aprendizado de máquina: RNA, SVM, *Random Forest* e Naive-Bayes. Os modelos de *machine learning* superaram o índice de referência do mercado.

Kong, Zhu, e Azencoff (2020) utilizaram métodos de aprendizagem de máquina para prever saltos intra diários das ações. Para realizar o estudo os autores utilizaram 1271 ações listadas na bolsa de valores de Shenzhen na China e incorporaram indicadores técnicos e medidas de liquidez. O comparativo das metodologias: *random forest*, redes neurais artificiais, *support vector machine* e KNN apontaram eficácia, na previsão, utilizando uma combinação de medidas de liquidez e indicadores financeiros. O *random forest* apresentou a melhor performance quando comparado com os demais métodos de aprendizagem de máquina. Ndikum (2020) compara o desempenho de algoritmos e ferramentas de *machine learning* com o tradicional modelo *capital asset pricing model* (CAPM), em ações listadas na bolsa dos Estados Unidos. Os modelos ML superaram o CAPM nos testes realizados.

Han e Kim (2021) utilizaram técnicas de aprendizado de máquina para prever pagamentos futuros nas ofertas pública inicial de ações (IPO). O modelo utiliza informações contábeis e preço das ações. Os resultados indicam uma melhora no retorno sobre o investimento. Zhang, Chu, e Shen (2021) usaram as redes LSTM para prever o movimento dos preços das ações e compararam com outras arquiteturas de redes neurais artificiais. Os resultados apontaram que as redes LSTM são mais adequadas para processar as séries temporais financeiras não lineares.

3 METODOLOGIA

3.1 REDES NEURAIIS ARTIFICIAIS

De acordo com Haykin (2001) o trabalho com RNAs é motivado por reconhecer que suas tipologias são semelhantes ao do cérebro humano que apresenta uma engrenagem altamente complexa e não linear para processar as informações de uma forma inteiramente diferente de um computador. Ele tem a capacidade de organizar os seus constituintes estruturais, os neurônios, de forma a realizar certos processamentos tais como reconhecimento de padrões, percepção e controle motor, de forma mais rápida que o mais rápido computador existente. Assim é possível conceituar uma rede neural artificial como um processador maciçamente e paralelamente distribuído, constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ainda, segundo o autor existem seis propriedades que caracterizam uma rede neural artificial: não linearidade; mapeamento de entrada e saída; adaptabilidade; resposta e evidências; informação contextual; tolerância e falhas.

De acordo com Medeiros (2018) a aprendizagem pode ser caracterizada como supervisionada (em que há *feedback* que retorna a rede para orientar o treinamento) e não supervisionada (a rede aprende de forma auto organizada). Entre as tarefas básicas que podem ser executadas por uma RNA estão: associação e reconhecimento de padrões; aproximação de funções; controle e filtragem.

As RNAs tem topologias diferentes e a implementação depende de variáveis como o número de nós da camada de entrada, o número de camadas escondidas e o número de neurônios da camada de saída. Recebendo entradas de n neurônios (y_i), o neurônio k calcula a sua saída através da equação:

$$y_k = g \left(\sum_{i=0}^n (y_i w_{ki}) + b_k \right) \quad (1)$$

Em que:

g função de ativação

y_i é a saída calculada pelo neurônio i ,

w_{ki} representa o peso sináptico entre o neurônio i e o neurônio k e

b_k é o peso entre um valor constante e diferente de zero ao neurônio k , conhecido como bias.

De acordo com Kovács (2002), a função $g(v)$ referida como função de ativação do neurônio, possui comportamento monotônico sobre o argumento v , conhecida como faixa dinâmica, e saturação fora dessa faixa. A função de ativação $g(v)$ pode ser aproximada por uma variedade de funções que procuram preservar as características com maior ou menor fidelidade.

As funções de ativações mais utilizadas são as funções sigmoide, também conhecida como função logística e a tangente hiperbólica devido a assimetria e continuidade das derivadas. As equações dessas duas funções são as seguintes:

- Função sigmoide

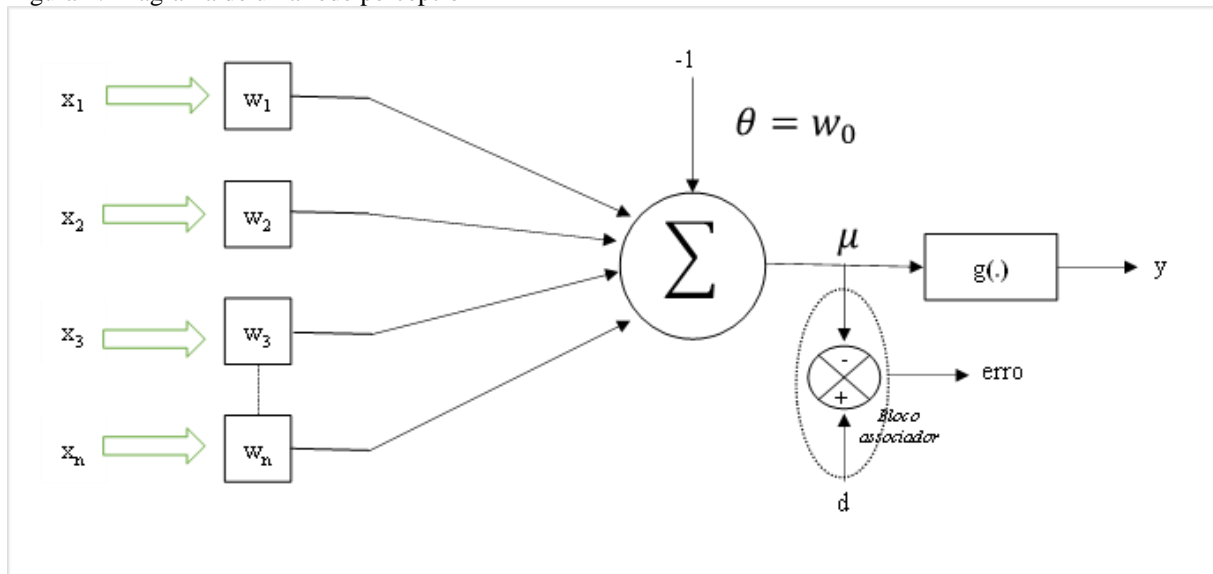
$$g(v) = (1 + \exp(-v))^{-1} \quad (2)$$

- Função tangente hiperbólica

$$g(v) = \tanh(v) = \frac{1 - \exp(-v)}{1 + \exp(-v)} \quad (3)$$

Na década de 50, Rosenblat criou uma rede de múltiplos neurônios do tipo discriminadores lineares chamado de rede *perceptron*. Um *perceptron* é uma rede com os neurônios dispostos em várias camadas. Os neurônios que recebem diretamente as entradas da rede são chamados de camada de entrada. Os neurônios que recebem como entrada as saídas daquela camada de entrada constituem a segunda camada e assim sucessivamente até a camada final, que é a camada de saída. As camadas internas que não são nem a entrada e nem a saída são chamadas de camadas ocultas. Um rede neural *perceptron* pode ser representada pela figura 1.

Figura 1: Diagrama de uma rede perceptron



Fonte: adaptado pelo autor (2021)

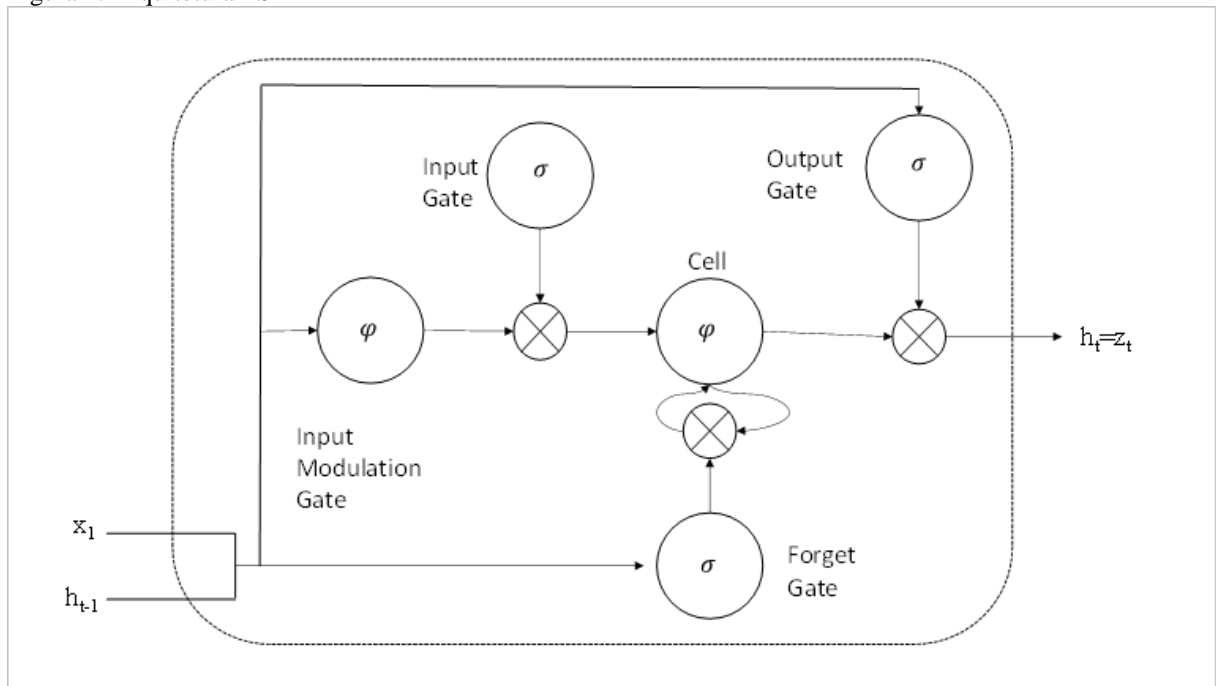
3.2 REDES NEURAIAS RECORRENTES

Uma rede neural recorrente (RNN) é uma classe de redes neurais que inclui conexões ponderadas. Como as RNNs incluem loops, elas podem armazenar informações ao processar novas entradas. Essa memória as torna ideais para tarefas de processamento onde as entradas anteriores devem ser consideradas. As atuais redes *deep learning* são baseadas em RNNs (HAYKEN, 2001).

3.3 REDES NEURAIAS LONG SHORT MEMORY TERM (LSTM)

A LSTM é uma arquitetura de rede neural recorrente (RNN) com valores em intervalos arbitrários. “A LSTM é bem adequada para classificar, processar e prever séries temporais com intervalos de tempo de duração desconhecida” (Data Science Academy, 2021, online). Essa rede possui uma estrutura em cadeia que contém quatro redes neurais e diferentes blocos de memória chamados células, conforme a figura 2.

Figura 2: Arquitetura LSTM



Fonte: Adaptado pelo autor

A informação é retida pelas células e as manipulações de memória são feitas pelos portões (*gates*): as informações que não são úteis são removidas das células (*forget gate*); novas informações são adicionadas (*input gate*), e essas informações são extraídas e apresentadas nas células de saída (*output gate*).

3.4 ARQUITETURA DE REDES NEURAIAS GATED RECURRENT UNIT (GRU)

A GRU é uma nova geração de RNN e é bastante semelhante a uma LSTM. O estado oculto das GRUs transferem informações. Essa arquitetura possui apenas dois portões, um portão de redefinição (*reset gate*) e um portão de atualização (*update date*). Esses portões podem ser treinados para manter informações antigas, sem dissipá-las com o tempo ou remover informações irrelevantes para a previsão.

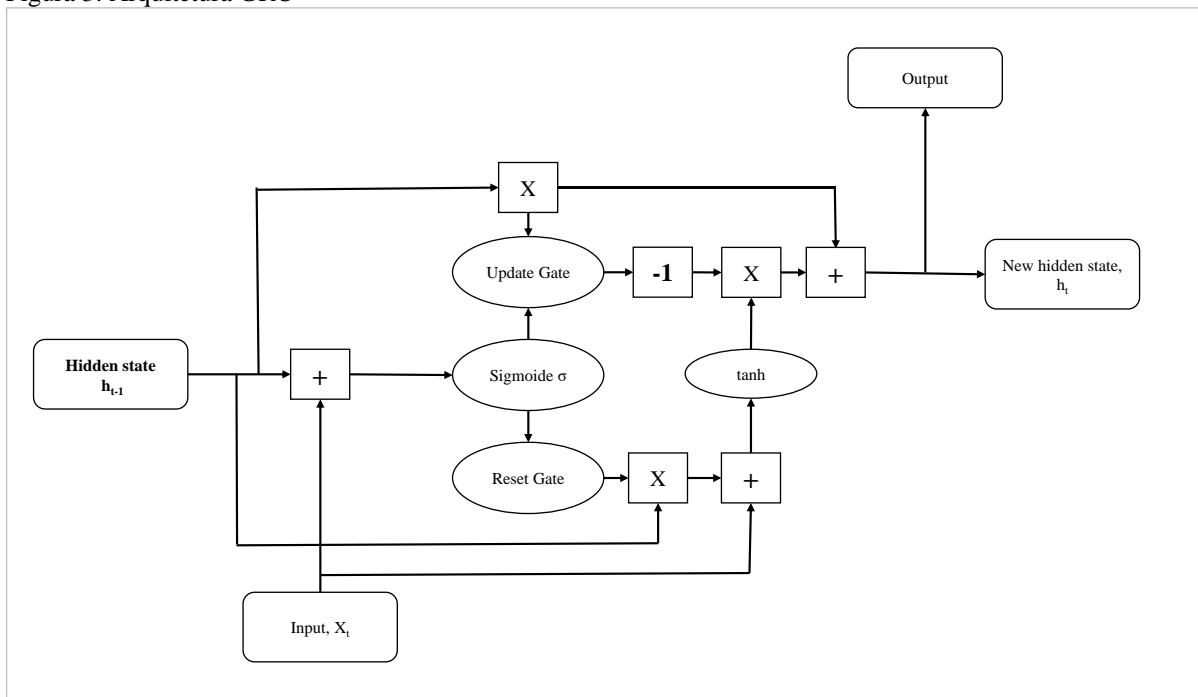
A estrutura da GRU permite capturar adaptativamente dependências de grandes sequências de dados sem descartar informações de partes anteriores da sequência. “Isso é alcançado através de suas unidades de portões. Esses portões são responsáveis por regular as informações a serem mantidas ou descartadas a cada etapa do tempo” (Data Science Academy, 2021, online).

A GRU tem a capacidade de manter dependências ou memória de longo prazo, os cálculos ocorre nas células da GRU com o objetivo de gerar o estado oculto. As GRUs têm

apenas um estado oculto transferido entre as etapas do tempo. “Esse estado oculto é capaz de manter as dependências de longo e curto prazo ao mesmo tempo, devido aos mecanismos de restrição e cálculos pelos quais o estado oculto e os dados de entrada passam” (Data Science Academy, 2021, online).

As GRUs são treinadas para filtrar qualquer informação irrelevante, mantendo o que é útil. Esses filtros ocorrem em vetores contendo valores entre 0 e 1 que serão multiplicados com os dados de entrada e de estado oculto. Um valor 0 indica que os dados de entrada ou oculto não são importantes e retornarão como zero. O valor 1 significa que os dados são importantes e serão utilizados. A representação gráfica dessa estrutura é representada pela figura 3.

Figura 3: Arquitetura GRU



Fonte: adaptado pelo autor

3.5 ANÁLISE DE EFICIÊNCIA

A performance de um RNA é avaliada comparando o desempenho da predição em relação aos valores observados. Neste trabalho, a medida utilizada para medir o desempenho dos modelos de redes neurais é a *mean directional accuracy* (MDA). Para o cálculo da MDA utiliza-se o conceito da primeira derivada que representa a inclinação da curva em qualquer ponto e a segunda derivada que indica a concavidade. Logo, a primeira derivada representa a mudança na inclinação ou a diferença no preço do ativo com relação ao dia anterior. A segunda

derivada é a diferença nessa mudança de preço. A seguinte fórmula é utilizada para calcular a MDA:

$$MDA = \frac{1}{N} \sum_T \mathbf{1}_{sign}(X_t - X_{t-1}) = sign((F_t - X_{t-1})) \quad (6)$$

Em que:

X_t : valor observado da série,

F_t : valor previsto

N: número total de observações

$sign(.)$ é o sinal da função

$\mathbf{1}$ é uma função indicadora

3.6 MÉTODOS DE OTIMIZAÇÃO

Em RNAs os algoritmos de otimização tem por objetivo minimizar a função de perda, com o intuito de reduzir o erro entre os dados previstos e os valores esperados. A minimização do erro torna as previsões das RNAs mais precisas, porém o custo computacional para otimizar esses parâmetros, é muito alto devido a atualização do “*bias*”. O gradiente estocástico descendente é um dos processos utilizados para contornar esse gasto computacional, pois utiliza o gradiente da função objeto para encontrar a melhor solução com a menor variação possível.

O tamanho dos passos, para atingir o mínimo local, é denominado taxa de aprendizagem. Para encontrar esse mínimo o otimizador utiliza um conjunto de lotes. Quanto maior o número de lotes mais precisa a taxa de aprendizagem e mais demorado o processo para encontrar a melhor solução. Para otimizar esse processo alguns algoritmos podem ser utilizados, entre eles:

- Adam: o algoritmo é baseado no gradiente de primeira ordem. Em que o primeiro momento da função é normalizado pelo segundo momento gerando a direção da atualização.
- RMShop: o gradiente de funções complexas tendem a desaparecer ou explodir de acordo com a propagação dos dados. Para contornar essa tendência, o RMShop utiliza mini lotes, como técnica, de aprendizagem dos modelos. O gradiente é normalizado por uma média móvel diminuindo o risco de explosão ou desaparecimento dos gradientes
- Adagrad: é uma técnica de otimização que procura adaptar a taxa de aprendizado de acordo com frequência dos parâmetros. As atualizações são maiores para os parâmetros menos frequentes e menores para os mais frequentes. A vantagem é um aumento na velocidade, escalabilidade, robustez, e pode ser usado para treinar redes neurais em grande escala. A

principal desvantagens é que o denominador é elevado ao quadrado tornando esse valor muito pequeno comprometendo o aprendizado da RNA.

- *Stochastic gradient descent* (SGD): o algoritmo procura encontrar o erro mínimo ou máximo por interações dos parâmetros que ocorre uma de cada vez, tornando o algoritmo muito rápido. A desvantagem é que as atualizações são realizadas com alta variância o que leva a função objeto a grande flutuações.

3.7 SÉRIES FINANCEIRAS

O conjunto de dados reais utilizado neste trabalho, são séries históricas das ações listadas na B3. O preço de fechamento do ativo foi extraído da Bloomberg e da Economática, com periodicidade diária. O horizonte de tempo corresponde ao período de 01 janeiro de 2000 a 30 de dezembro de 2019, perfazendo um histórico de 20 anos com 4949 registros para realização dos estudos.

O objetivo é trabalhar com recortes de tempo e comparar a previsão das RNAs em diferentes períodos de tempo. A tabela 1 apresenta os recortes analisados neste trabalho:

Tabela 1: Horizonte de tempo

Amostra	Horizonte de tempo		Dias	Mês	Ano
	Mínimo	Máximo			
1	01/01/2019	30/12/2019	363	11	1
2	01/01/2018	30/12/2019	728	23	2
3	01/01/2015	30/12/2019	1824	59	5
4	01/01/2010	30/12/2019	3650	119	10
5	01/01/2000	30/12/2019	7303	239	20

Fonte: Autor

Para o desenvolvimento dos modelos foi utilizado o software Python e a bibliotecas: Keras, Tensor Flow e Scikit Learn.

Todos os conjuntos de dados foram divididos em amostras: de treino, teste e validação. Para o desenvolvimento dos modelos utilizou-se 80% dos dados disponíveis, e 20% para teste. Os últimos 60 períodos de cada série foram utilizados para análise *out-of-sample*.

4 RESULTADOS

O índice Ibovespa (IBOV), que é medido em pontos, representa uma carteira teórica com as ações mais representativas negociadas na B3 e é um termômetro do mercado acionário brasileiro. Este índice foi escolhido para analisar o desempenho das redes neurais em diferentes períodos.

As redes neurais foram desenvolvidas com 30 camadas de 50 neurônios, gerando várias combinações possíveis de sinais. Cada combinação, é estimada 100 vezes até atingir a convergência ideal (*epochs*). A tabela 2 apresenta um comparativo, entre as duas arquiteturas analisadas, de acordo com o horizonte de tempo:

Tabela 2: Comparativo das RNAs por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	47,9%	45,8%
01/01/2018	30/12/2019	493	335	98	60	50,5%	49,5%
01/01/2015	30/12/2019	1234	928	246	60	46,1%	47,8%
01/01/2010	30/12/2019	2472	1918	494	60	50,5%	51,5%
01/01/2000	30/12/2019	4948	3899	989	60	48,7%	49,5%

Fonte: Autor

A melhor performance em relação ao movimento direcional foi obtido com a arquitetura GRU, considerando o horizonte de tempo de 10 anos. O MDA aponta que em 52% dos casos a rede neural acerta o movimento direcional.

Outro ponto importante que podemos notar, na tabela 2, é que para períodos de até 2 anos a arquitetura LSTM apresenta melhores resultados quando comparados com a GRU. Quando analisamos períodos de tempo mais longos (acima de 5 anos) a arquitetura GRU é mais precisa que a LSTM.

O número de épocas (*epochs*) e *batch size* representam, respectivamente, os resultados esperados e a taxa de aprendizagem dos modelos. A utilização desses parâmetros permite rastrear o processo de aprendizagem. Para cada passo é aplicado uma nova descida do gradiente e os pesos e bias são cálculos novamente.

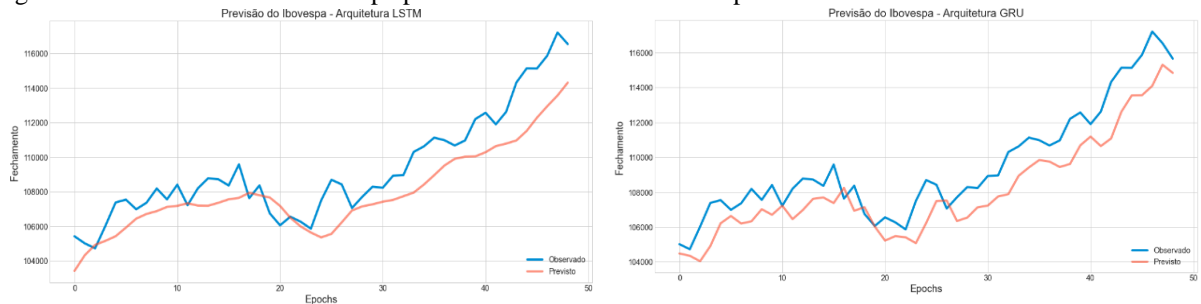
Diferentes *epochs* e *batch size* foram utilizados para o treinamento das duas arquiteturas e constatamos que o número de épocas (*epochs*) não interfere nos resultados. Quanto maior o número de *epochs* menor a precisão do movimento direcional dos modelos.

Os otimizadores desempenham papel importante na precisão dos modelos. Eles são responsáveis pela atualização dos pesos que minimizam a função erro. Os otimizadores

calculam os gradientes da função de erro em relação aos demais parâmetros da rede atualizando-os. Neste trabalho foram utilizados os otimizadores: adam, adagrad, SGD, RMSprop. O otimizador adam apresentou os melhores resultados e foi escolhido para o cálculo do gradiente.

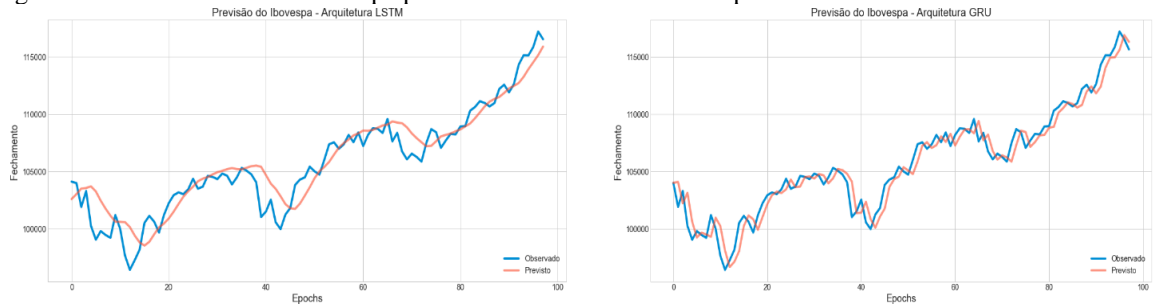
Após todos os testes os modelos foram treinados utilizando-se o otimizador adam e a função de perda *mean square error* (MSE). As figuras de 4 a 11 exibem os resultados gráficos dos modelos, considerando as duas arquiteturas de redes neurais estudadas.

Figura 4: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 1 ano



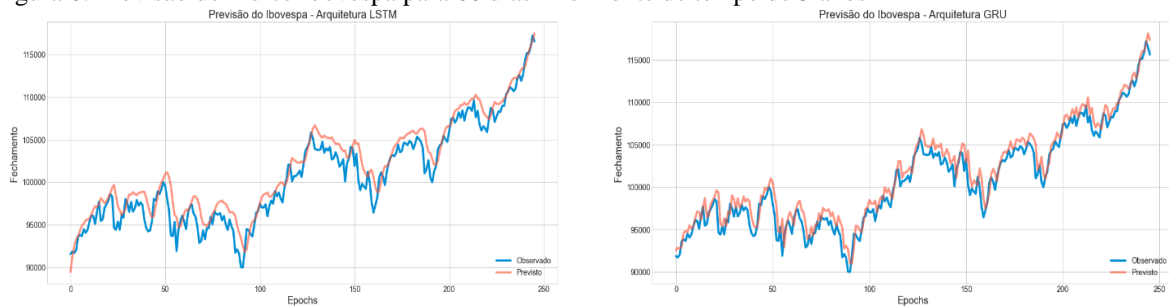
Fonte: Autor (2021)

Figura 5: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 2 anos



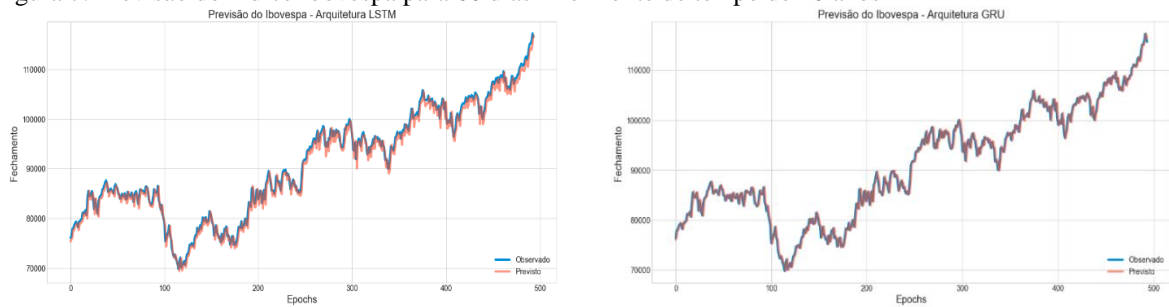
Fonte: Autor (2021)

Figura 6: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 5 anos



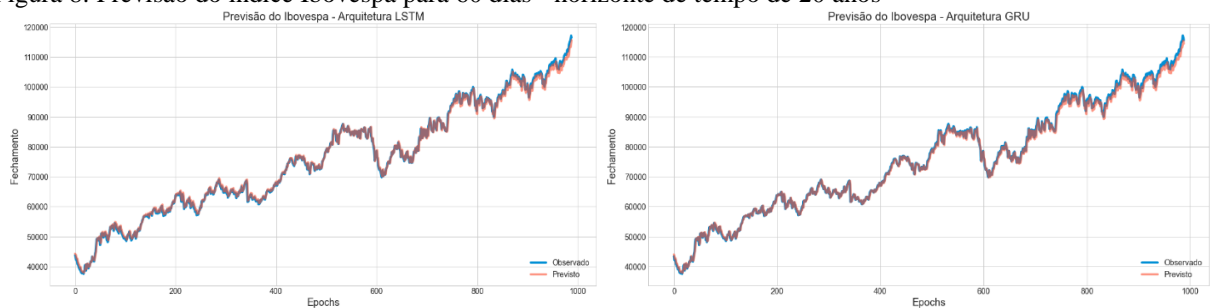
Fonte: Autor (2021)

Figura 7: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 8: Previsão do índice Ibovespa para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

Além do IBOV, também analisamos o MDA por ação e segmento. O critério de seleção dos ativos (tabela 3) leva em consideração algumas ações que apresentaram negociações em praticamente todos os pregões no histórico de 20 anos.

Tabela 3: Ações por segmento

id.	Empresa	Ação	Segmento
1	Banco Bradesco S.A.	ITUB4	Financeiro
2	Banco do Brasil S.A.	BBAS3	Financeiro
3	Banco Itaú S.A.	BBDC4	Financeiro
4	Braskem S.A.	BRKM5	Petroquímico
5	Petróleo Brasileiro S.A. - Petrobras	PETR4	Petróleo e Gás
6	Gerdau S.A.	GGBR4	Siderúrgico
7	Vale S.A.	VALE3	Mineração
8	Companhia Paranaense de Energia	CPLE6	Energia e Saneamento
9	Engie Brasil Energia S.A	EGIE3	Energia e Saneamento
10	Telefonica S.A.	VIVT4	Telecomunicações
11	TIM Participações S.A.	TIMP4	Telecomunicações

Fonte: Autor (2021)

- Ações do Banco do Brasil S.A. (BBAS3)

A tabela 3 apresenta um comparativo, entre as duas arquiteturas de RNA, de acordo com o horizonte de tempo para as ações BBAS3:

- A arquitetura LSTM, apresenta resultados superiores a GRU na análise do movimento direcional.
- O horizonte de tempo de até dois anos apresenta maior precisão quando comparado com horizontes de tempo mais longos.
- O desempenho da GRU não conseguiu superar o desempenho da LSTM na análise do BBAS3.
- Tanto a LSTM quanto a GRU não conseguiram treinar o modelo no período de 20 anos.

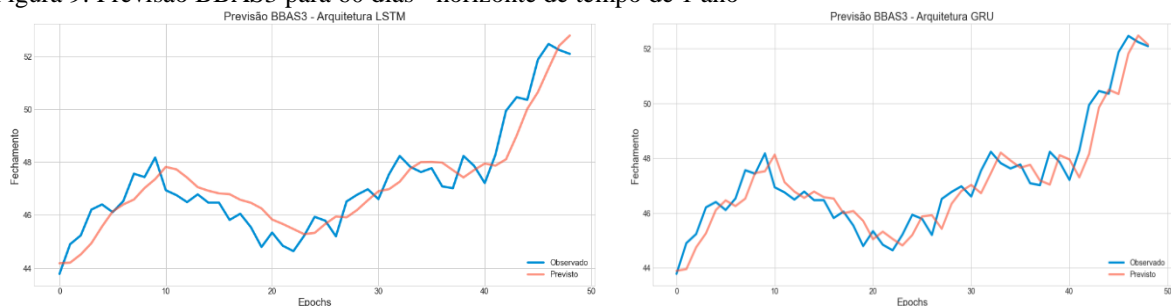
Tabela 4: Comparativo das ações BBAS3 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	56,25%	45,83%
01/01/2018	30/12/2019	493	335	98	60	50,52%	50,52%
01/01/2015	30/12/2019	1234	928	246	60	48,98%	46,53%
01/01/2010	30/12/2019	2472	1918	494	60	51,52%	50,10%
01/01/2000	30/12/2019	4948	3899	989	60		

Fonte: Autor (2021)

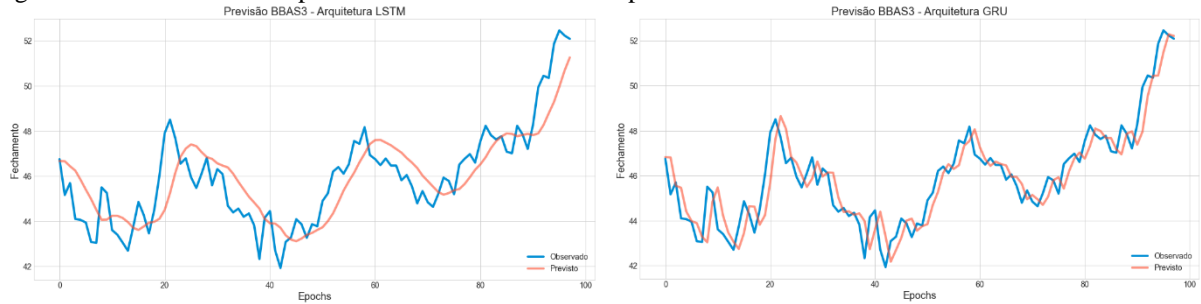
As figuras de 9 a 12 exibem os gráficos referentes as previsões do BBAS3 considerando o período de 60 dias em diferentes horizontes de tempo.

Figura 9: Previsão BBAS3 para 60 dias - horizonte de tempo de 1 ano



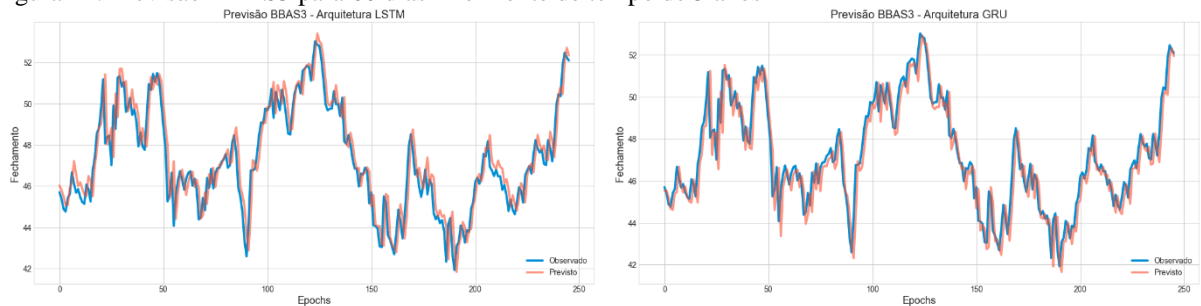
Fonte: Autor (2021)

Figura 10: Previsão BBAS3 para 60 dias - horizonte de tempo de 2 anos



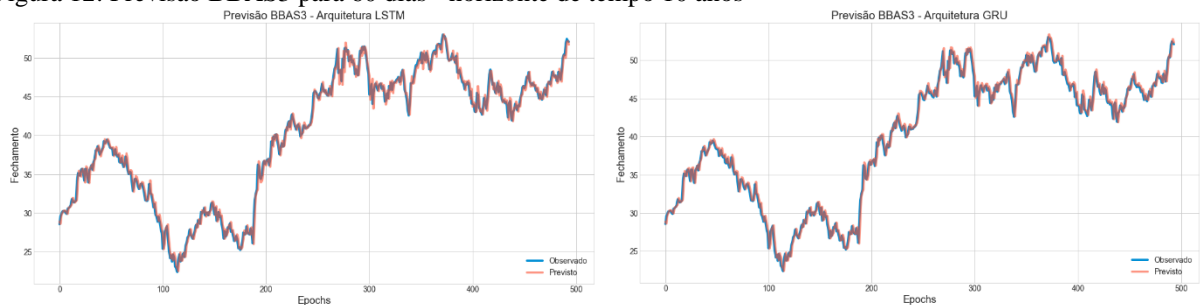
Fonte: Autor (2021)

Figura 11: Previsão BBAS3 para 60 dias - horizonte de tempo de 5 anos



Fonte: Autor (2021)

Figura 12: Previsão BBAS3 para 60 dias - horizonte de tempo 10 anos



Fonte: Autor (2021)

- Ações do Banco Itaú S.A. (ITUB4)

As ações do ITUB4 apresentam comportamento similar ao do BBAS3. Na tabela 5 podemos observar:

- A arquitetura LSTM teve um desempenho superior quando comparada com a arquitetura GRU.
- Nos períodos curtos (até 2 anos) a performance da LSTM é superior a GRU.
- Em período longos as duas arquiteturas tem performance similar.

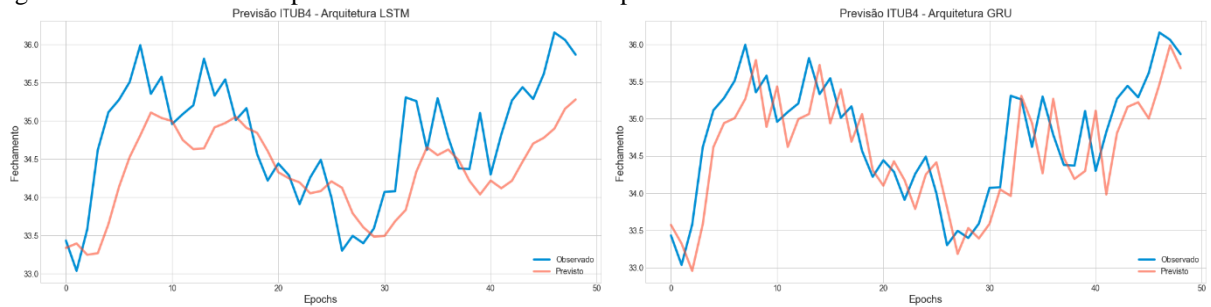
Tabela 5: Comparativo das ações ITUB4 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	50,00%	45,83%
01/01/2018	30/12/2019	493	335	98	60	49,48%	47,42%
01/01/2015	30/12/2019	1234	928	246	60	47,35%	48,57%
01/01/2010	30/12/2019	2472	1918	494	60	51,52%	49,70%
01/01/2000	30/12/2019	4948	3899	989	60	48,79%	48,28%

Fonte: Autor (2021)

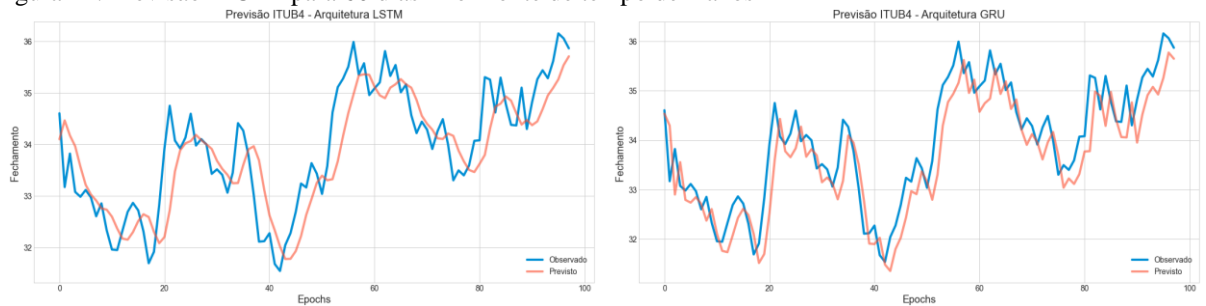
As figuras de 13 a 17 apresentam os resultados gráficos referentes as previsões do ITUB4 considerando o período de 60 dias em diferentes horizontes de tempo,

Figura 13: Previsão ITUB4 para 60 dias - horizonte de tempo de 1 ano



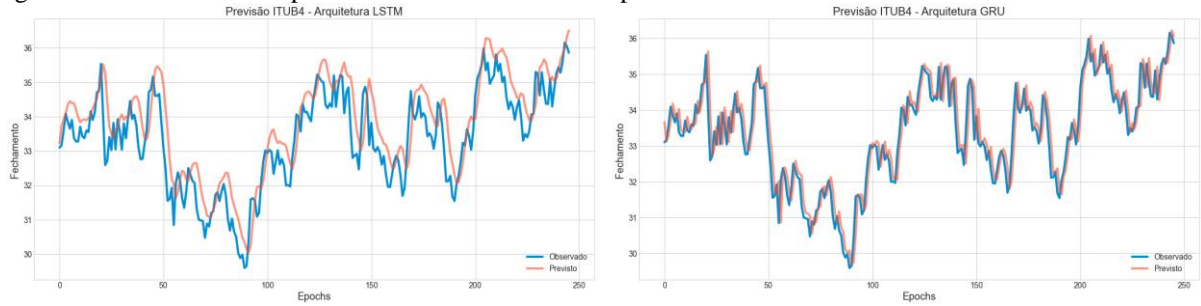
Fonte: Autor (2021)

Figura 14: Previsão ITUB4 para 60 dias - horizonte de tempo de 2 anos



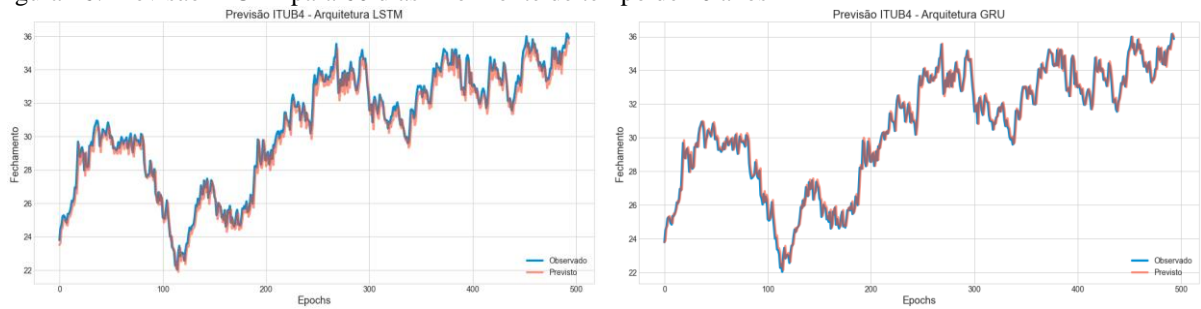
Fonte: Autor (2021)

Figura 15: Previsão ITUB4 para 60 dias - horizonte de tempo de 5 anos



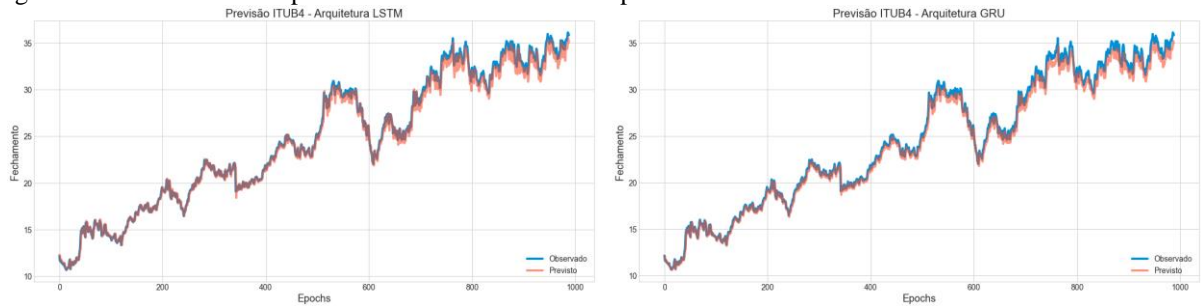
Fonte: Autor (2021)

Figura 16: Previsão ITUB4 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 17: Previsão ITUB4 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

- Ações do Banco Bradesco S.A. (BBDC4)

As ações do BBDC4 também apresentaram comportamento similar a BBAS3 e ITUB4. A tabela 6 apresenta os resultados e notamos que apenas para o período de 5 anos a eficácia da GRU foi superior a LSTM na análise do MDA.

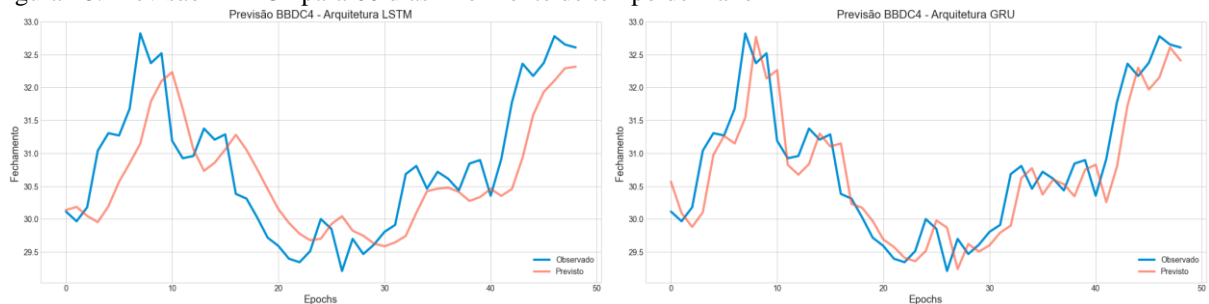
Tabela 6: Comparativo das ações BBDC4 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	50,00%	50,00%
01/01/2018	30/12/2019	493	335	98	60	54,64%	52,58%
01/01/2015	30/12/2019	1234	928	246	60	50,20%	54,29%
01/01/2010	30/12/2019	2472	1918	494	60	52,74%	51,72%
01/01/2000	30/12/2019	4948	3899	989	60	51,11%	50,40%

Fonte: Autor (2021)

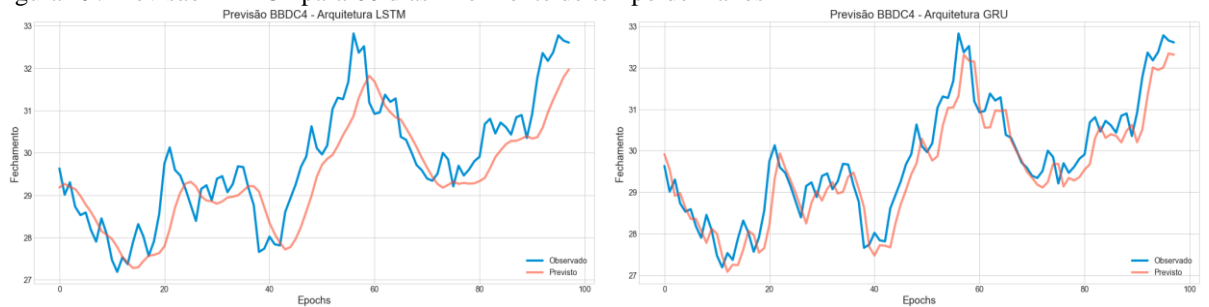
As figuras 18 a 22 exibem os resultados gráficos.

Figura 18: Previsão BBDC4 para 60 dias - horizonte de tempo de 1 ano



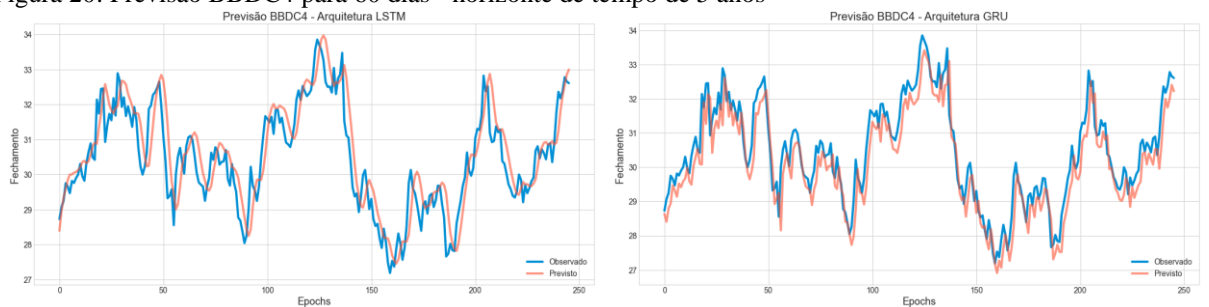
Fonte: Autor (2021)

Figura 19: Previsão BBDC4 para 60 dias - horizonte de tempo de 2 anos



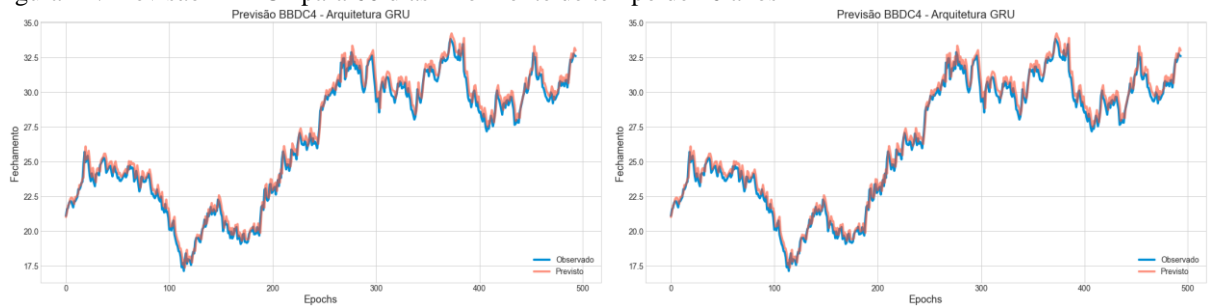
Fonte: Autor (2021)

Figura 20: Previsão BBDC4 para 60 dias - horizonte de tempo de 5 anos



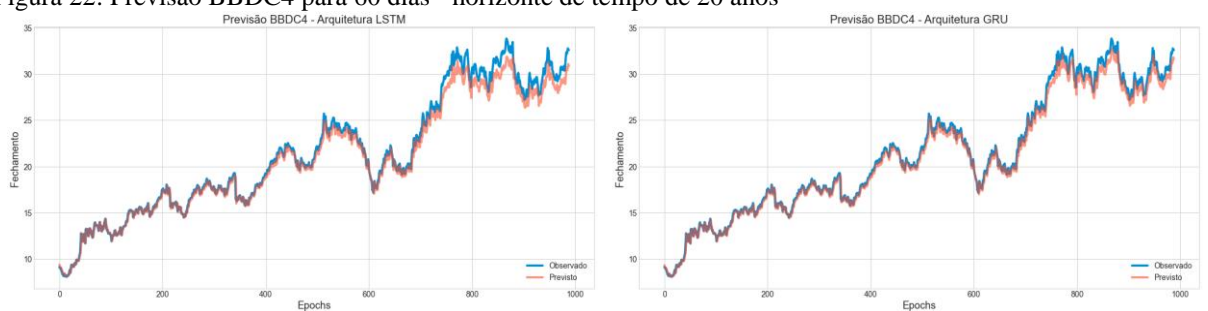
Fonte: Autor (2021)

Figura 21: Previsão BBDC4 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 22: Previsão BBDC4 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

- Ações da Petróleo Brasileiro S.A. - Petrobras (PETR4)

A tabela 7 apresenta o comparativo das análises do MDA para as ações PETR4. Podemos observar o seguinte:

- No horizonte de um ano o MDA apresenta performance superior aos demais períodos e a arquitetura GRU é bem mais precisa que a arquitetura LSTM.
- No horizonte de dois anos a eficácia da arquitetura LSTM é superior a GRU.
- Quanto maior o horizonte de análise menor a precisão do MDA.
- No longo prazo, o movimento direcional, tende a ficar bem próximo no comparativo entre as duas arquiteturas.

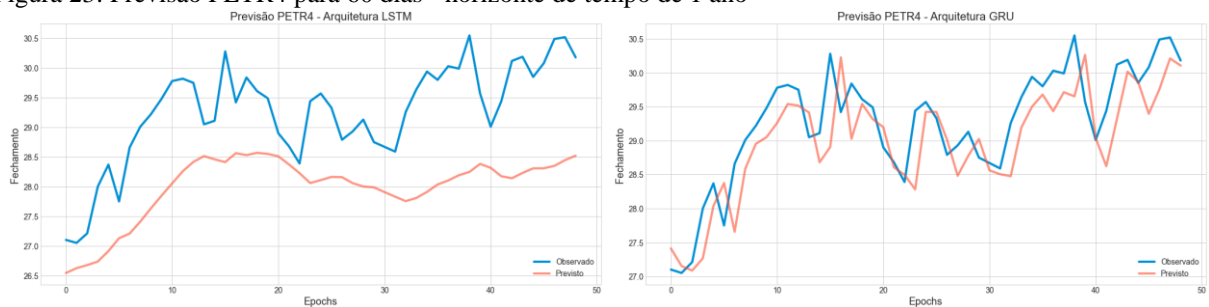
Tabela 7: Comparativo das ações PETR4 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	56,25%	62,50%
01/01/2018	30/12/2019	493	335	98	60	53,61%	49,48%
01/01/2015	30/12/2019	1234	928	246	60	46,12%	47,76%
01/01/2010	30/12/2019	2472	1918	494	60	49,09%	48,48%
01/01/2000	30/12/2019	4948	3899	989	60	47,87%	47,06%

Fonte: Autor (2021)

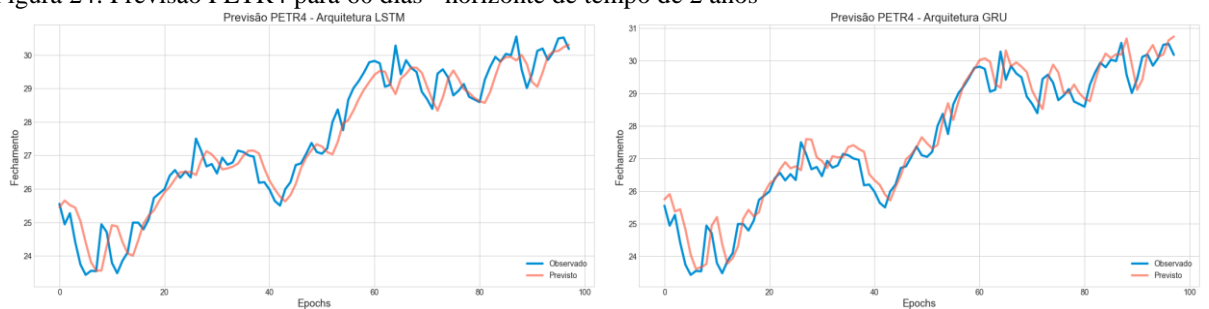
Graficamente (figuras 23 a 27) também observamos uma diferença entre as curvas dos valores reais e previstos pelo modelo, Uma das possíveis explicações é que devido as oscilações internas e externas do mercado de petróleo as previsões da PETR4 tendem a ser afetadas.

Figura 23: Previsão PETR4 para 60 dias - horizonte de tempo de 1 ano



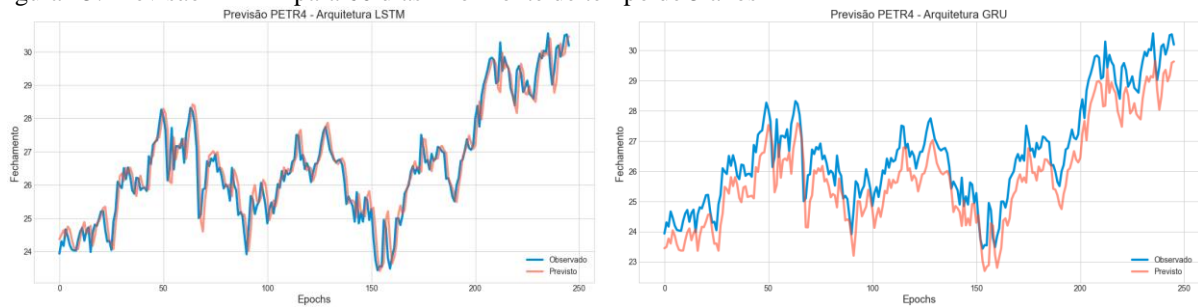
Fonte: Autor (2021)

Figura 24: Previsão PETR4 para 60 dias - horizonte de tempo de 2 anos



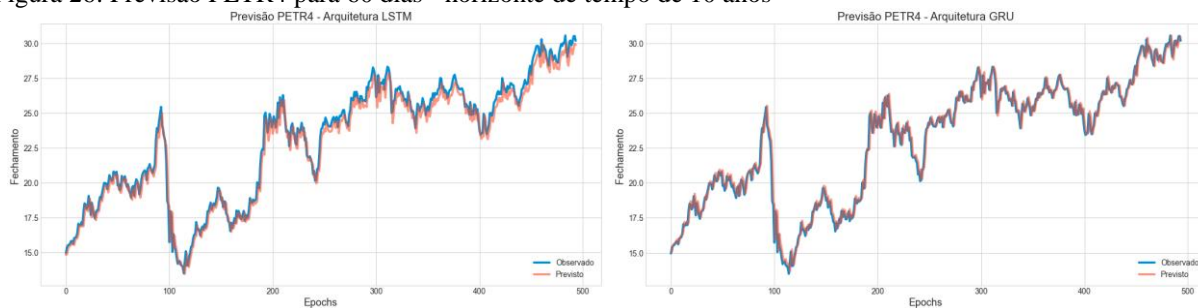
Fonte: Autor (2021)

Figura 25: Previsão PETR4 para 60 dias - horizonte de tempo de 5 anos



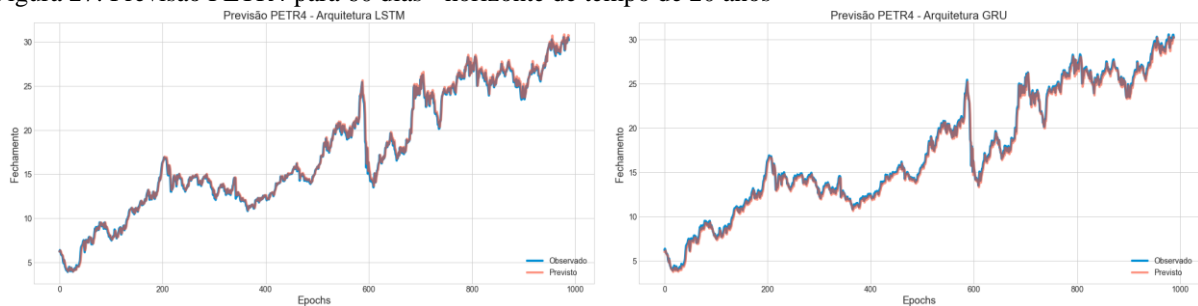
Fonte: Autor (2021)

Figura 26: Previsão PETR4 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 27: Previsão PETR4 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

- Ações da Braskem S.A. (BRKM5)

As ações das BRKM5 apresentam resultados mais precisos quando comparados com a PETR4. Na tabela 8 notamos que:

- No horizonte de tempo de um ano a eficácia da arquitetura LSTM (64,5%) é superior a GRU (52,1%).
- Em todos os horizontes de tempo a LSTM apresenta resultados superiores a GRU.
- Quanto maior o período de análise menor a precisão nas duas arquiteturas.

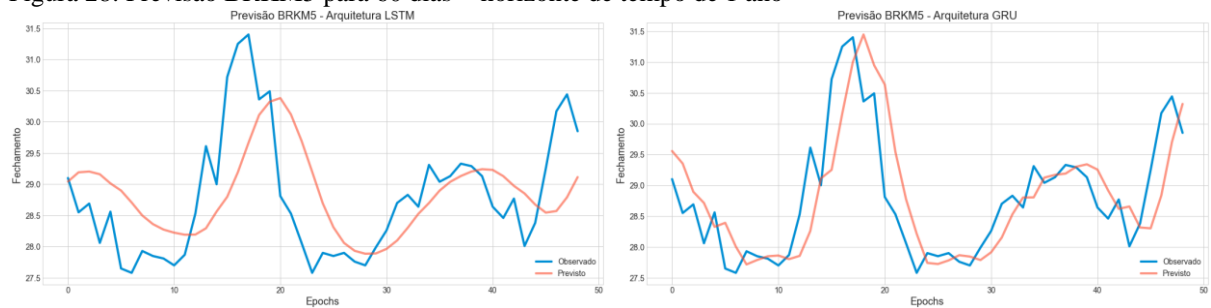
Tabela 8: Comparativo das ações BRKM5 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	64,58%	52,08%
01/01/2018	30/12/2019	493	335	98	60	59,79%	57,73%
01/01/2015	30/12/2019	1234	928	246	60	60,00%	55,51%
01/01/2010	30/12/2019	2472	1918	494	60	52,94%	52,33%
01/01/2000	30/12/2019	4948	3899	989	60	51,32%	50,71%

Fonte: Autor (2021)

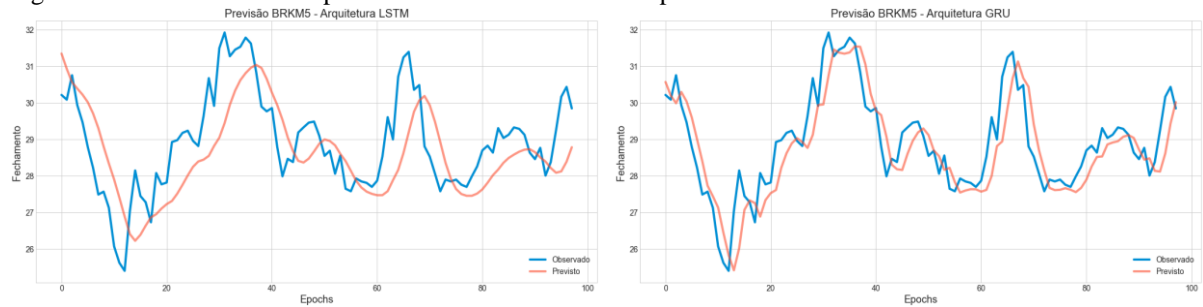
As próximas figuras (28 a 32) exibem os resultados gráficos das ações da BRKM5,

Figura 28: Previsão BRKM5 para 60 dias - horizonte de tempo de 1 ano



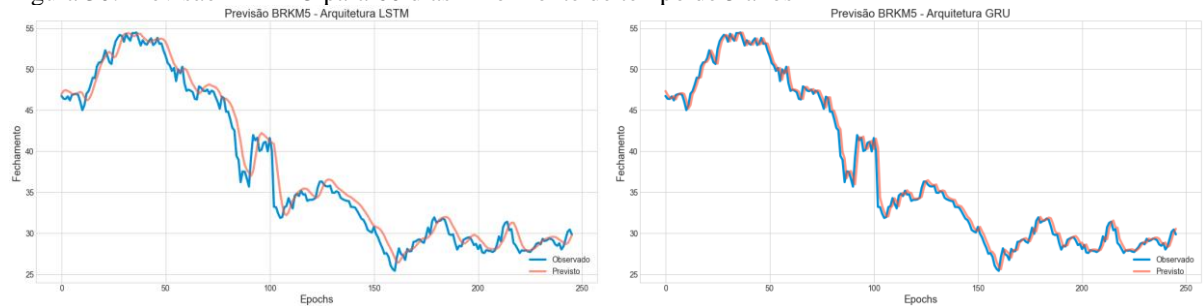
Fonte: Autor (2021)

Figura 29: Previsão BRKM5 para 60 dias - horizonte de tempo de 2 anos



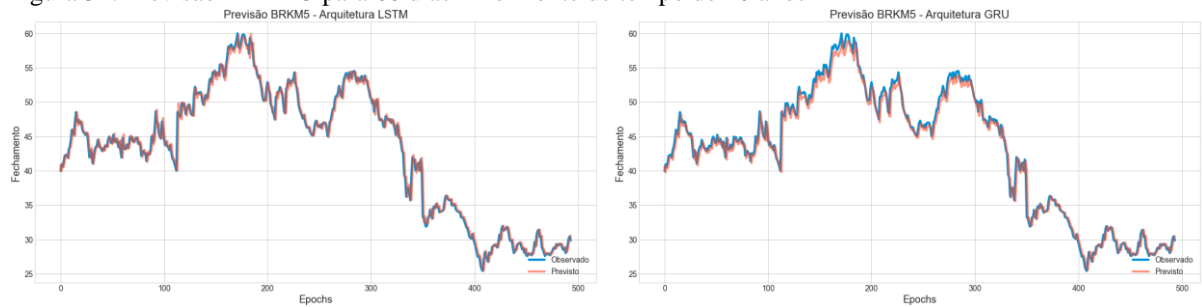
Fonte: Autor (2021)

Figura 30: Previsão BRKM5 para 60 dias - horizonte de tempo de 5 anos



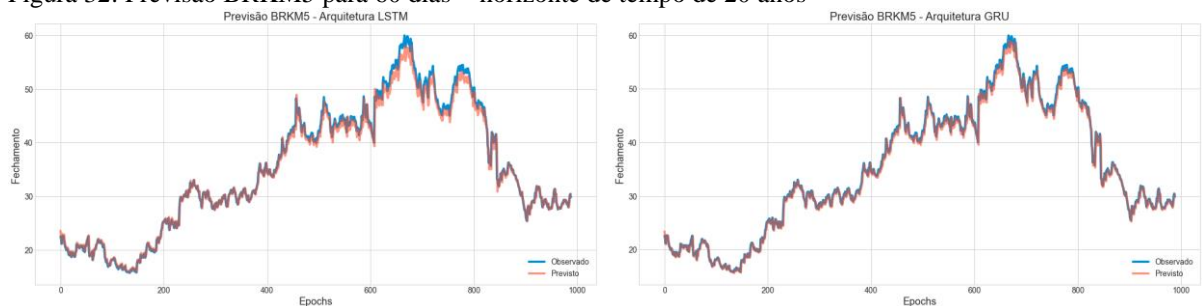
Fonte: Autor (2021)

Figura 31: Previsão BRKM5 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 32: Previsão BRKM5 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

- Ações da Vale S.A. (VALE3)

As RNAs não convergiram para a VALE3 no horizonte de 20 anos e nos demais horizontes os resultados foram distintos:

- No período de um ano a performance da GRU (56,2%) foi superior a LSTM (50%).
- No período de dois anos a GRU (50,5%) também foi superior a LSTM (48,4%), mas com uma diferença menor.
- No horizonte de três anos a eficácia da LSTM é maior a GRU, 49,4% contra 46,1% respectivamente.
- No horizonte de 10 anos as duas arquiteturas apresentaram o mesmo resultado.

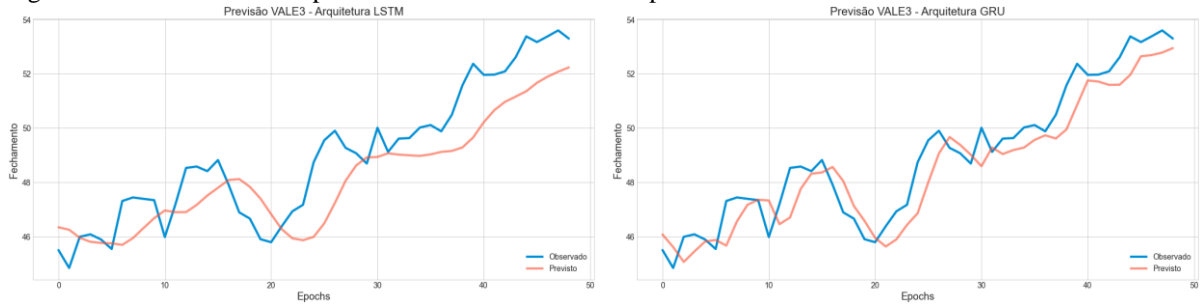
Tabela 9: Comparativo das ações VALE3 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	50,00%	56,25%
01/01/2018	30/12/2019	493	335	98	60	48,45%	50,52%
01/01/2015	30/12/2019	1234	928	246	60	49,39%	46,12%
01/01/2010	30/12/2019	2472	1918	494	60	48,68%	48,68%
01/01/2000	30/12/2019	4948	3899	989	60	0,00%	0,00%

Fonte: Autor (2021)

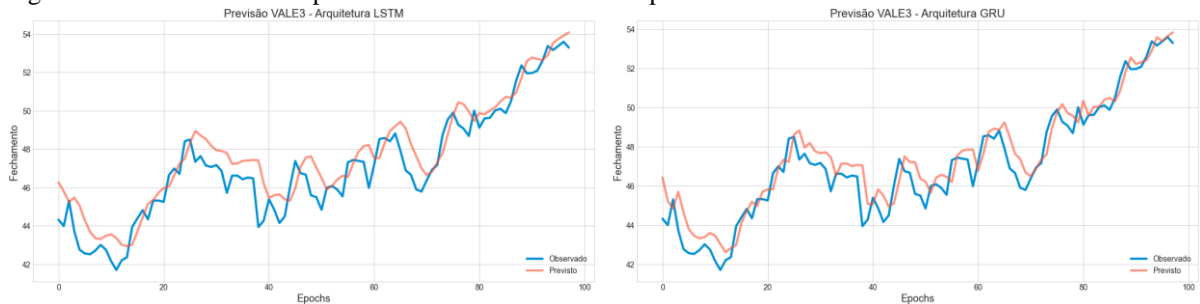
As figuras de 33 a 36 apresentam os resultados gráficos observados para a VALE3.

Figura 33: Previsão VALE3 para 60 dias - horizonte de tempo de 1 ano



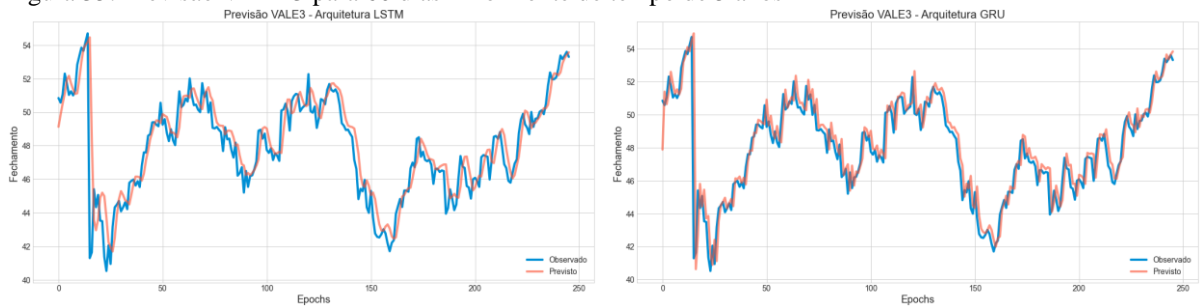
Fonte: Autor (2021)

Figura 34: Previsão VALE3 para 60 dias - horizonte de tempo de 2 anos



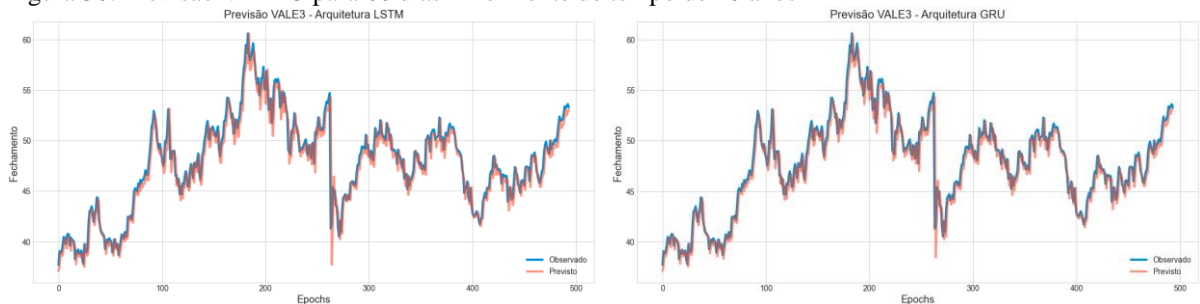
Fonte: Autor (2021)

Figura 35: Previsão VALE3 para 60 dias - horizonte de tempo de 5 anos



Fonte: Autor (2021)

Figura 36: Previsão VALE3 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

- Ações Metalúrgica Gerdau S.A. (GGBR4)

A tabela 10 exibe os resultados das previsões para as ações da GGBR4.

- Para o horizontes de tempo de até 5 anos a performance da LSTM é superior a GRU.
- Para períodos de tempo mais longos a performance da GRU supera a LSTM.
- A arquitetura GRU mantém, praticamente, o mesmo percentual de eficácia (46%) para períodos acima de 5 anos.
- O padrão de quanto maior o horizonte de tempo menor a precisão não se configura para as ações da GGBR4.

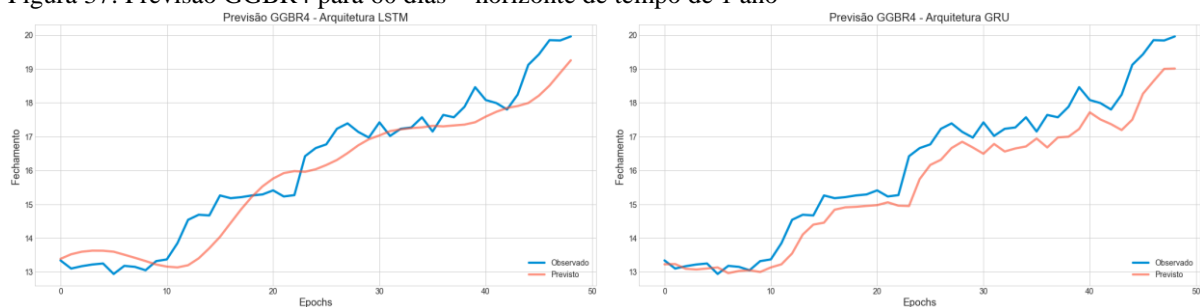
Tabela 10: Comparativo das ações GGBR4 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	58,33%	52,08%
01/01/2018	30/12/2019	493	335	98	60	45,36%	44,33%
01/01/2015	30/12/2019	1234	928	246	60	46,94%	46,53%
01/01/2010	30/12/2019	2472	1918	494	60	44,62%	46,45%
01/01/2000	30/12/2019	4948	3899	989	60	45,95%	46,05%

Fonte: Autor (2021)

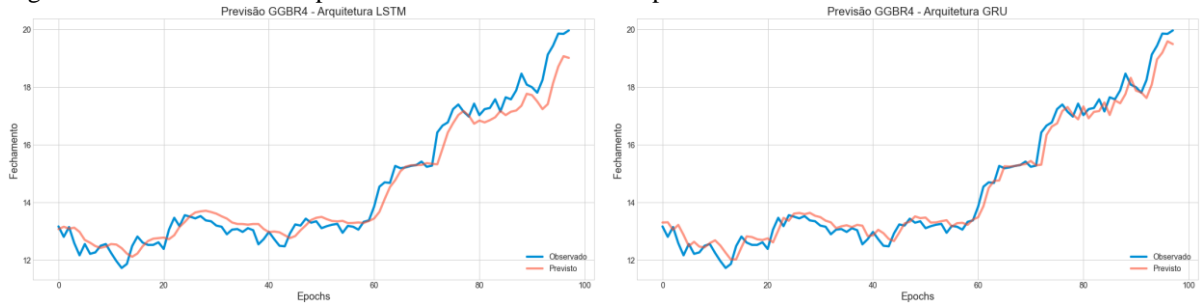
Graficamente (figuras 37 a 41) as curvas dos valores reais e previstos tendem a ficar mais próximas. Para períodos superiores a 2 anos o MDA fica em torno de 45%.

Figura 37: Previsão GGBR4 para 60 dias - horizonte de tempo de 1 ano



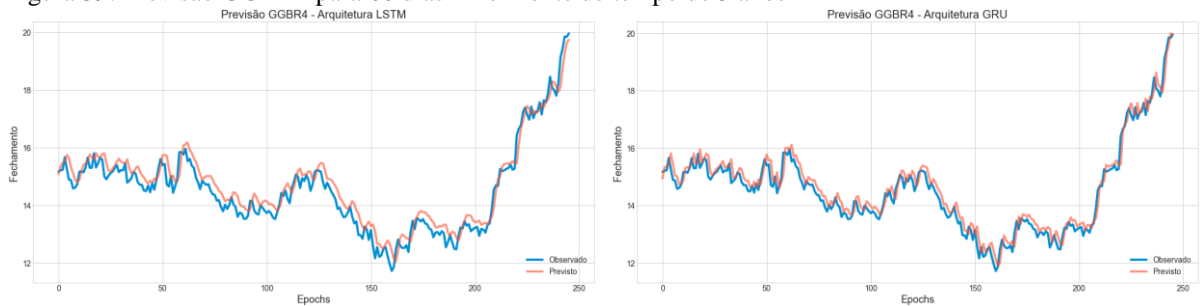
Fonte: Autor (2021)

Figura 38: Previsão GGBR4 para 60 dias - horizonte de tempo de 2 anos



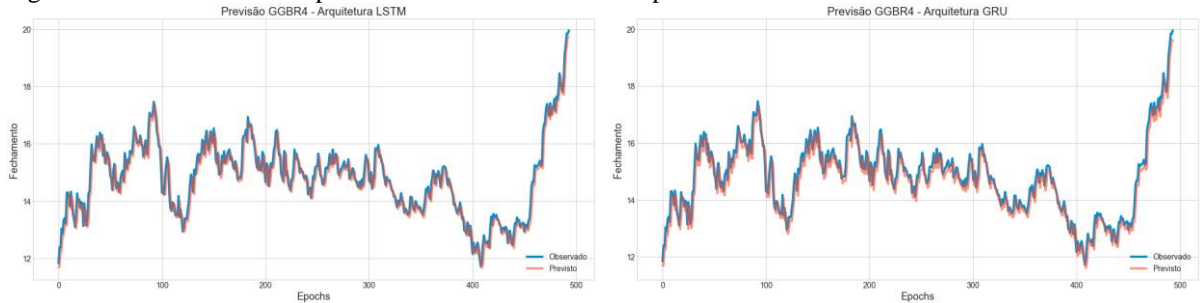
Fonte: Autor (2021)

Figura 39: Previsão GGBR4 para 60 dias - horizonte de tempo de 5 anos



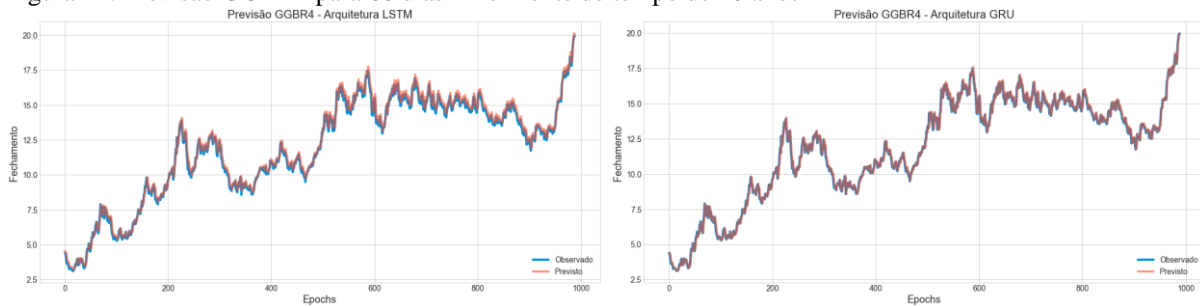
Fonte: Autor (2021)

Figura 40: Previsão GGBR4 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 41: Previsão GGBR4 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

- Ações da Engie Brasil Energia S.A. (EGIE3)

A Engie Brasil é a maior geradora privada de energia no Brasil e considerada uma ação defensiva. Na análise do MDA para a EGIE3 (tabela 3) podemos identificar:

- Para horizontes de até 2 anos a arquitetura LSTM apresenta maior precisão em relação a GRU.
- Para horizontes de tempo acima de 5 anos a GRU supera a LSTM.
- No período de 1 ano a GRU teve um desempenho abaixo da média e foi superada pela LSTM em 5 pontos percentuais.
- O padrão de quanto maior o horizonte de tempo menor a precisão é observado somente para a arquitetura LSTM.

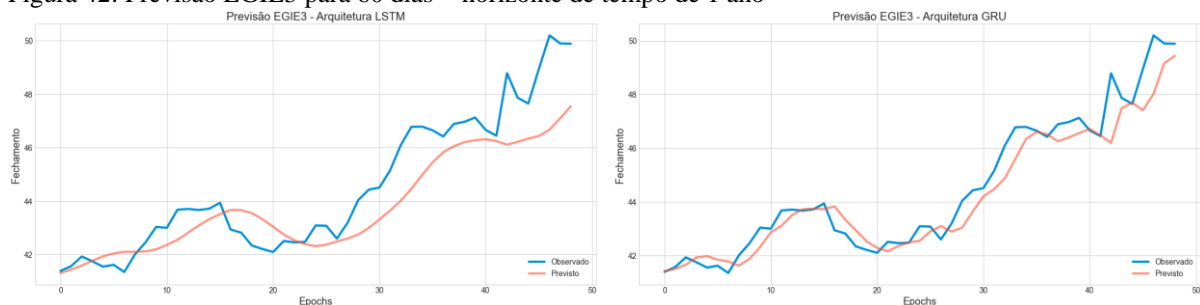
Tabela 11: Comparativo das ações EGIE3 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	58,33%	54,17%
01/01/2018	30/12/2019	493	335	98	60	51,55%	46,39%
01/01/2015	30/12/2019	1234	928	246	60	50,61%	52,65%
01/01/2010	30/12/2019	2472	1918	494	60	49,09%	52,94%
01/01/2000	30/12/2019	4948	3899	989	60	47,57%	48,38%

Fonte: Autor (2021)

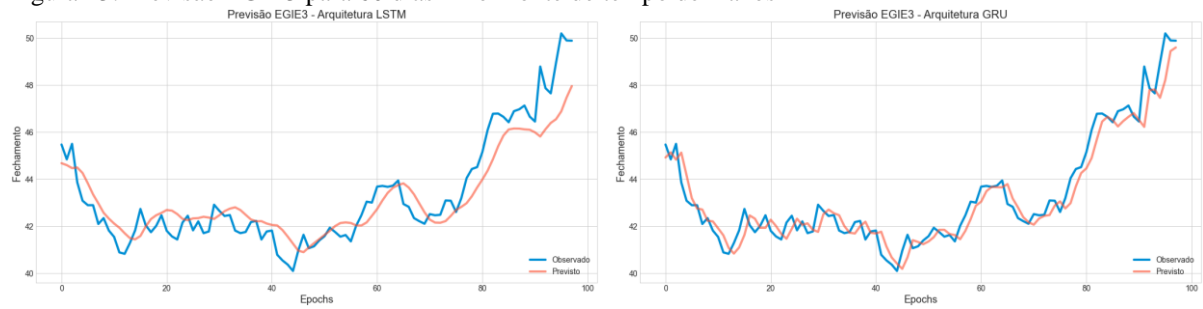
As figuras 42 a 46 apresentam os resultados das previsões para diferentes horizontes de tempo.

Figura 42: Previsão EGIE3 para 60 dias - horizonte de tempo de 1 ano



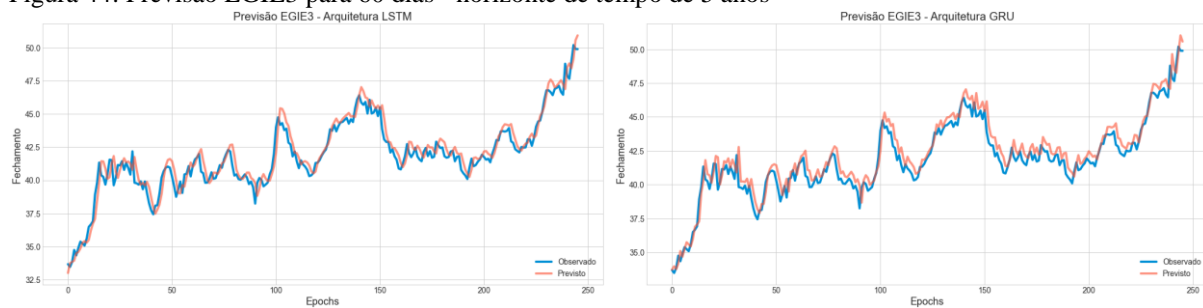
Fonte: Autor (2021)

Figura 43: Previsão EGIE3 para 60 dias - horizonte de tempo de 2 anos



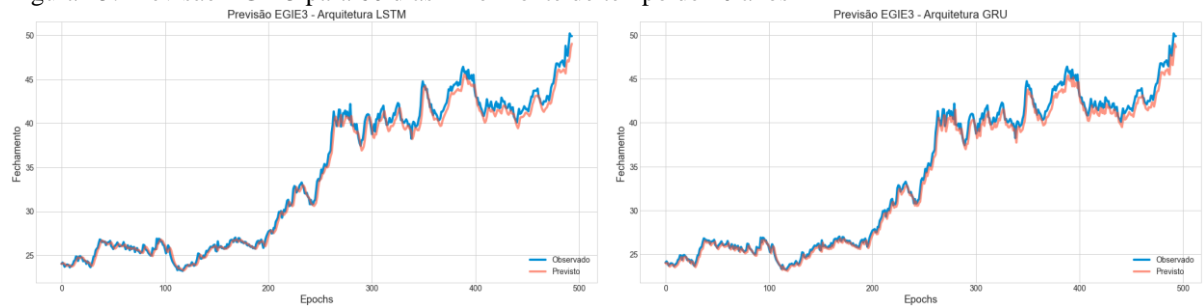
Fonte: Autor (2021)

Figura 44: Previsão EGIE3 para 60 dias - horizonte de tempo de 5 anos



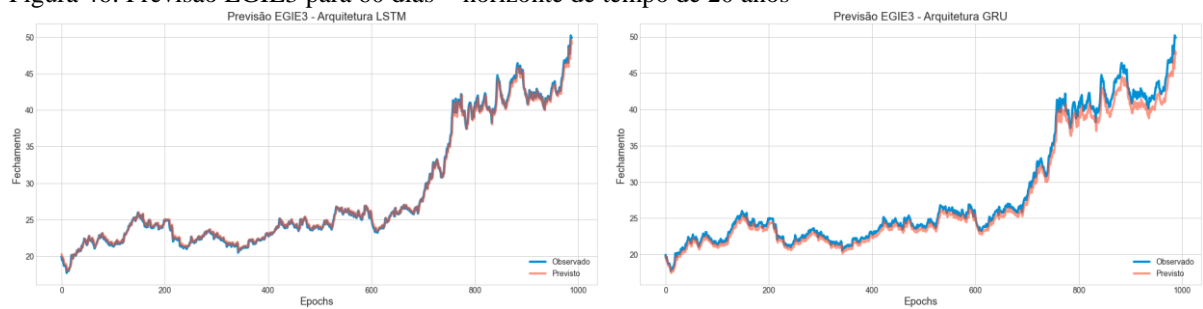
Fonte: Autor (2021)

Figura 45: Previsão EGIE3 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 46: Previsão EGIE3 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

- Companhia Paranaense de Energia (CPLE6)

A Copel é uma companhia de distribuição, geração, transmissão e distribuição de energia. Também é considerada uma empresa defensiva e os resultados da previsão da CPLE3 são exibidos na tabela 3. Podemos notar o seguinte:

- A arquitetura GRU supera a performance da LSTM em todos os horizontes de tempo.
- O horizonte de 5 anos apresenta a melhor performance para o ativo (56,7%) na arquitetura GRU.
- As RNAs não apresentam o padrão de quanto maior o horizonte de tempo menor a precisão. Porém, o período de 20 anos apresenta resultados inferiores aos demais períodos.

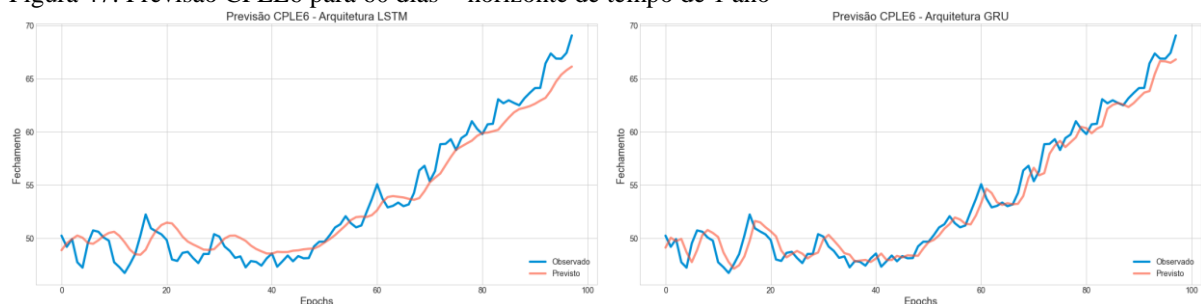
Tabela 12: Comparativo das ações CPLE6 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	49,48%	51,55%
01/01/2018	30/12/2019	493	335	98	60	50,52%	52,58%
01/01/2015	30/12/2019	1234	928	246	60	53,47%	56,73%
01/01/2010	30/12/2019	2472	1918	494	60	52,94%	53,14%
01/01/2000	30/12/2019	4948	3899	989	60	48,58%	48,89%

Fonte: Autor (2021)

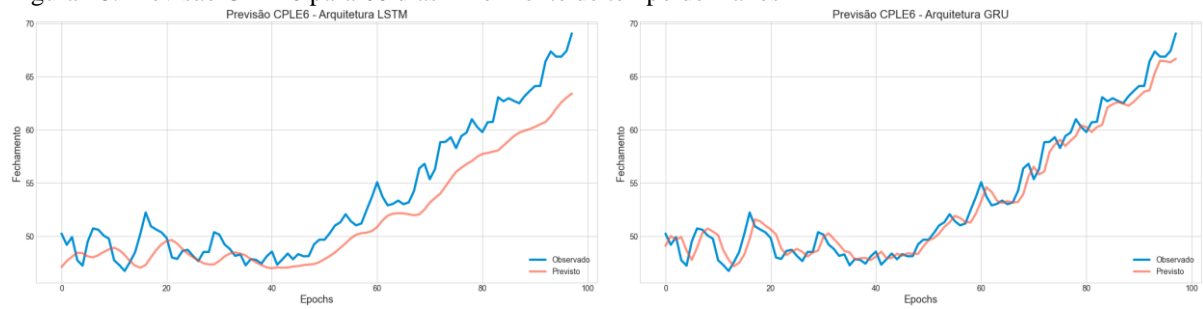
Graficamente (figuras 47 a 51) as curvas referentes aos valores reais e previsto pelo modelo ficam mais próximas, porém o movimento direcional da análise sofre variações,

Figura 47: Previsão CPLE6 para 60 dias - horizonte de tempo de 1 ano



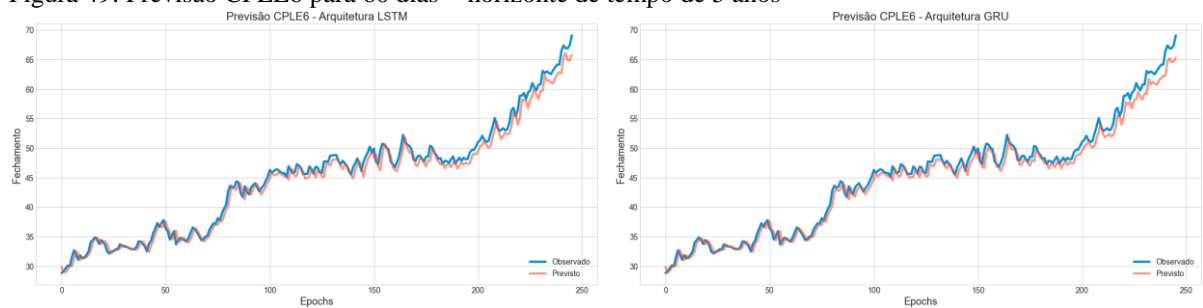
Fonte: Autor (2021)

Figura 48: Previsão CPLE6 para 60 dias - horizonte de tempo de 2 anos



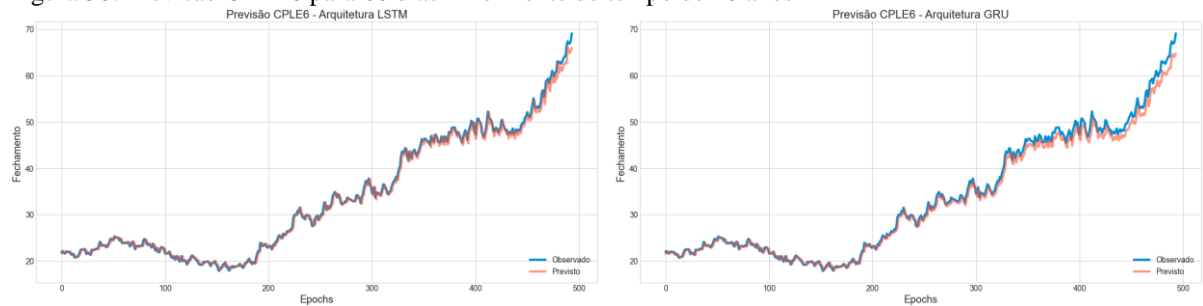
Fonte: Autor (2021)

Figura 49: Previsão CPLE6 para 60 dias - horizonte de tempo de 5 anos



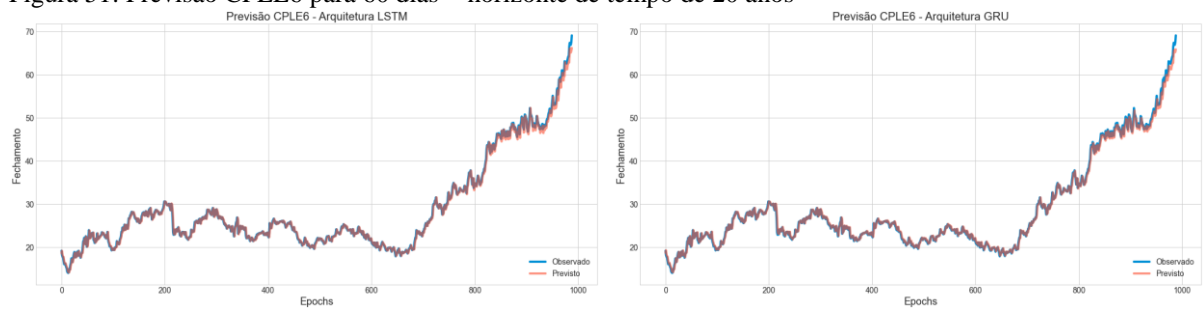
Fonte: Autor (2021)

Figura 50: Previsão CPLE6 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 51: Previsão CPLE6 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

- Ações Telefonica S.A. (VIVT4)

A Telefonica S.A. trabalha com vários produtos de telecomunicações e tem a operadora de celulares Vivo como uma de suas subsidiárias. Na tabela 13 observamos:

- Desempenho abaixo do esperado para a arquitetura LSTM no período de um ano (39,6%).
- A arquitetura GRU apresentou resultados superiores a LSTM para períodos de até dois anos e com boa diferença percentual.
- Na GRU quanto maior o período de observação menor a precisão no movimento direcional.
- Não observamos uma tendência na arquitetura LSTM.

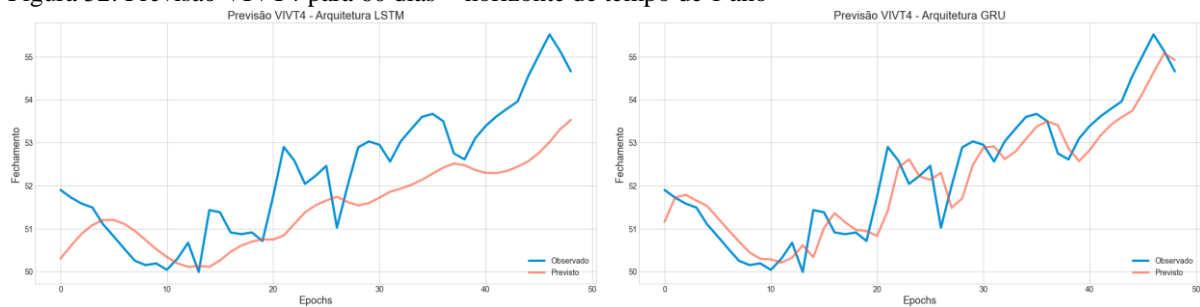
Tabela 13: Comparativo das ações VIVT4 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	39,58%	54,17%
01/01/2018	30/12/2019	493	335	98	60	45,36%	53,61%
01/01/2015	30/12/2019	1234	928	246	60	49,39%	49,39%
01/01/2010	30/12/2019	2472	1918	494	60	49,49%	48,07%
01/01/2000	30/12/2019	4948	3899	989	60	48,48%	47,77%

Fonte: Autor (2021)

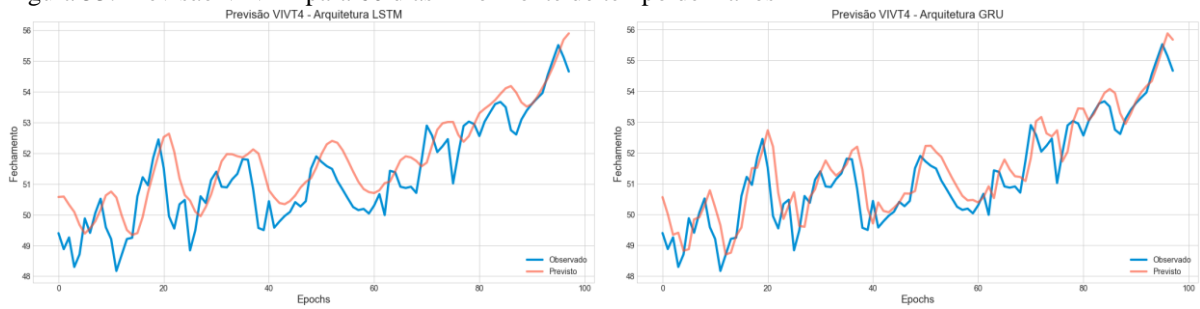
Nas figuras 52 a 56 observamos que os valores reais e os previstos, pelo modelo, apresentam diferenças quando comparamos as duas arquiteturas.

Figura 52: Previsão VIVT4 para 60 dias - horizonte de tempo de 1 ano



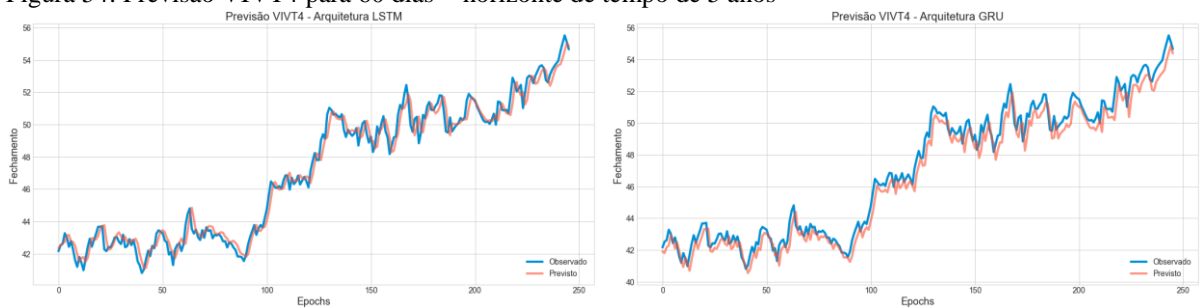
Fonte: Autor (2021)

Figura 53: Previsão VIVT4 para 60 dias - horizonte de tempo de 2 anos



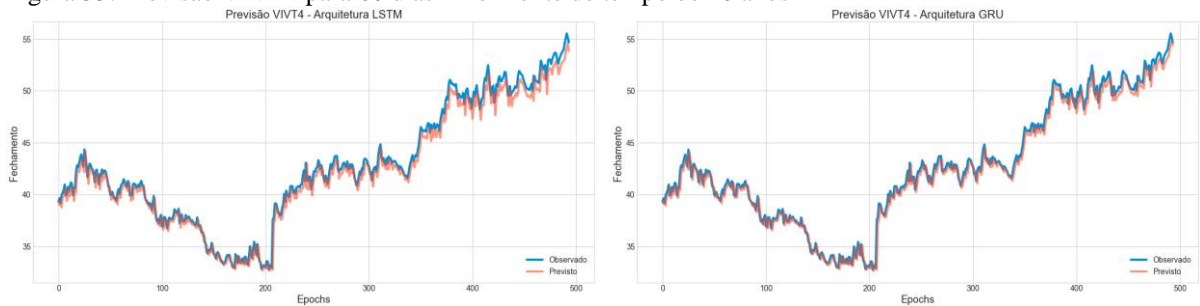
Fonte: Autor (2021)

Figura 54: Previsão VIVT4 para 60 dias - horizonte de tempo de 5 anos



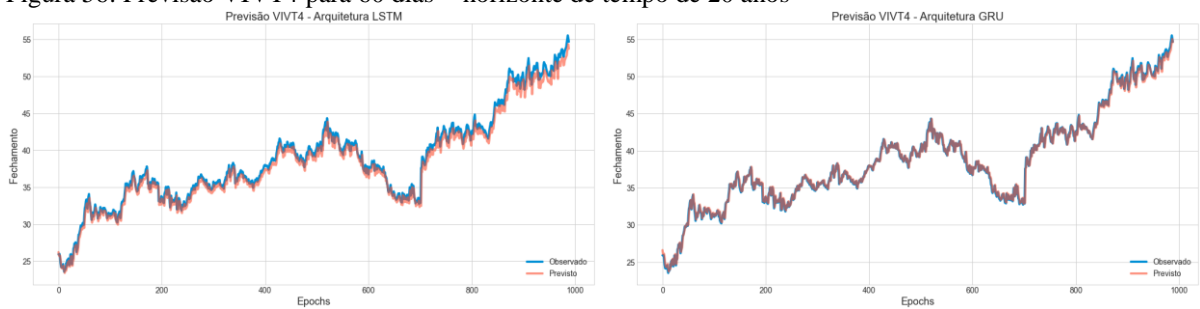
Fonte: Autor (2021)

Figura 55: Previsão VIVT4 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 56: Previsão VIVT4 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

- TIM Participações S.A. (TIMP4)

A TIM é uma empresa de telecomunicações que atua no segmento de telefonia móvel e apresenta os seguintes resultados (tabela 14):

- No período de 1 ano a precisão da LSTM (66,7%) é bem superior a GRU (52%).
- Para períodos de tempo superiores a 5 anos a precisão do indicador é bem menor quando comparado com períodos menores.
- Na arquitetura GRU observamos que quanto maior o período de análise menor a eficácia. Na LSTM, também observamos esse padrão exceto para o período de 5 anos,

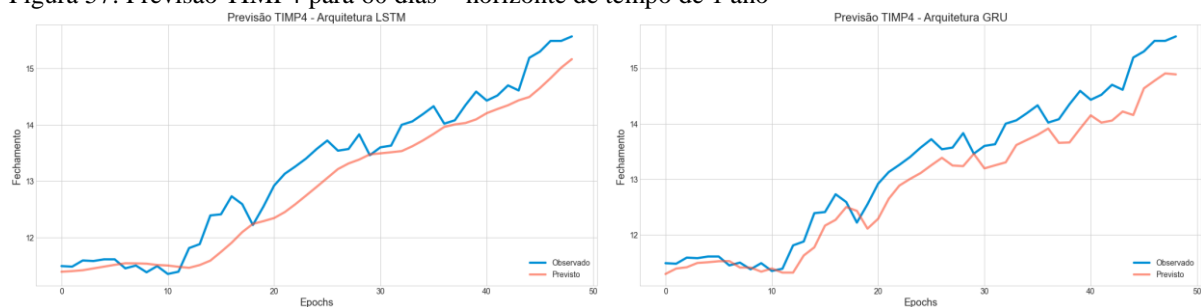
Tabela 14: Comparativo das ações TIMP4 por horizonte de tempo

Horizonte de tempo		# Amostra	Treino	Teste	Validação	MDA	
Inicial	Final					LSTM	GRU
01/01/2019	30/12/2019	247	138	49	60	66,67%	52,08%
01/01/2018	30/12/2019	493	335	98	60	47,42%	50,52%
01/01/2015	30/12/2019	1234	928	246	60	51,02%	49,80%
01/01/2010	30/12/2019	2472	1918	494	60	46,65%	46,45%
01/01/2000	30/12/2019	4948	3899	989	60	44,74%	45,34%

Fonte: Autor (2021)

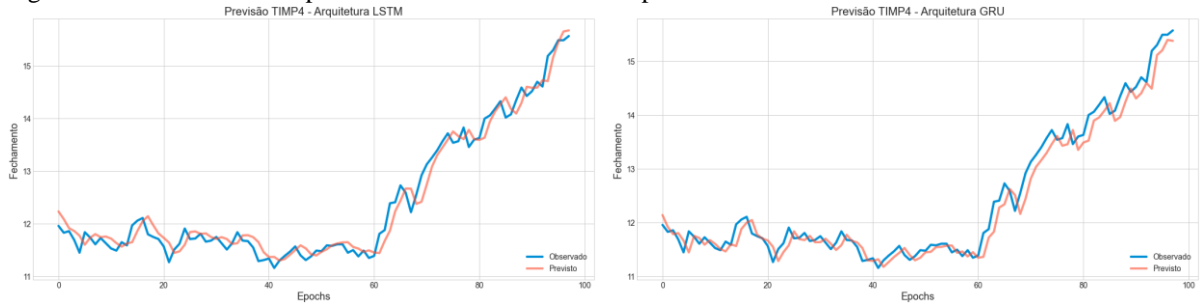
As figuras de 57 a 61 apresentam os comparativos entre as previsões da TIMP para o período de 60 dias em diferentes horizontes de tempo.

Figura 57: Previsão TIMP4 para 60 dias - horizonte de tempo de 1 ano



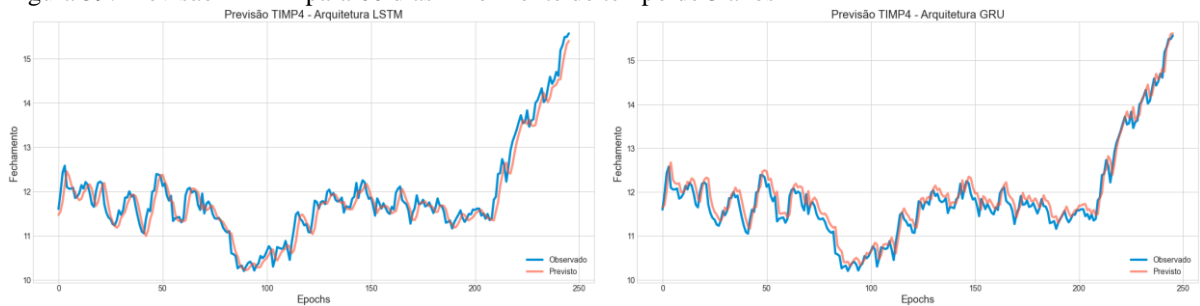
Fonte: Autor (2021)

Figura 58: Previsão TIMP4 para 60 dias - horizonte de tempo de 2 anos



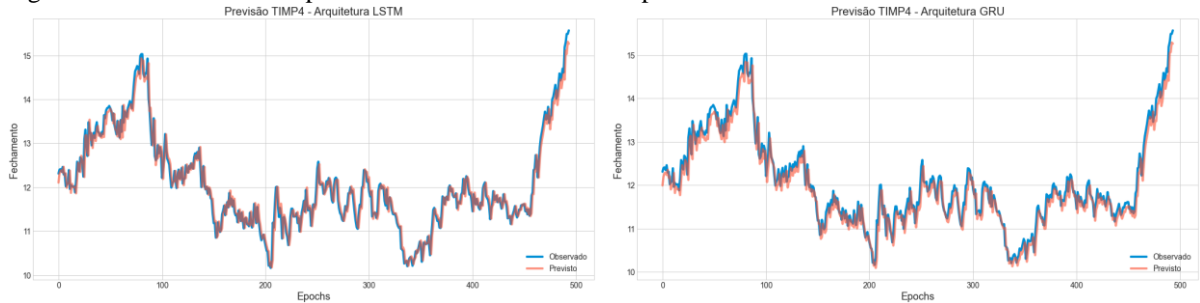
Fonte: Autor (2021)

Figura 59: Previsão TIMP4 para 60 dias - horizonte de tempo de 5 anos



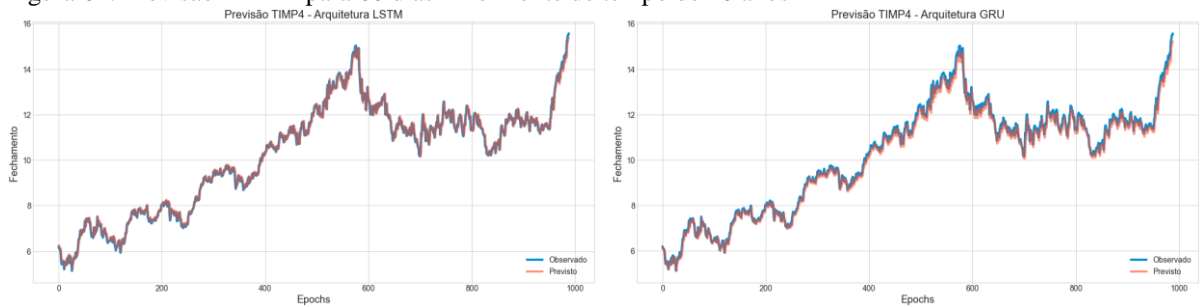
Fonte: Autor (2021)

Figura 60: Previsão TIMP4 para 60 dias - horizonte de tempo de 10 anos



Fonte: Autor (2021)

Figura 61: Previsão TIMP4 para 60 dias - horizonte de tempo de 20 anos



Fonte: Autor (2021)

5 CONCLUSÕES

Quando analisamos o movimento direcional de um ativo financeiro, esperamos como resposta, a flutuação do ativo, ou seja, o movimento de alta ou de baixa. Prever esse movimento, não é fácil e os resultados obtidos, neste trabalho, demonstram essa dificuldade.

As duas arquiteturas de redes neurais analisadas apresentaram resultados bem similares, na análise do índice Ibovespa, mas existe evidências de algumas variações que são significativas. O movimento direcional apresentou performance em torno de 50% demonstrando que o índice sofre muitas oscilações decorrentes de contextos políticos e econômicos. Com relação aos horizontes de tempo analisados, os resultados também são similares indicando que o tamanho da amostra não interfere na performance do movimento direcional.

Quando analisamos horizontes de tempo menores, os resultados evidenciam que as arquiteturas de redes neurais LSTM apresentam melhores resultados. Para séries mais longas a performance da arquitetura GRU é superior quando comparada com a LSTM. Alguns parâmetros utilizados nas redes neurais, também, não influenciaram na análise do movimento direcional.

A análise por ações evidencia que para horizontes de tempo mais curtos os resultados do movimento direcional são mais precisos quando comparados com horizontes de tempo mais longos. O horizonte de tempo correspondente a 20 anos apresentou menor eficácia quando comparado com os demais períodos de tempo, ou seja, quando analisamos as ações individualmente, o tamanho da amostra influencia nos resultados. Por segmento também observamos padrões similares. As ações do segmento financeiro sofreram maiores oscilações em relação aos demais segmentos.

Na teoria, outras variáveis podem ser adicionadas ao conjunto de dados para aumentar a sua precisão. A inflação pode causar variações nas análises, bem como, análise de ações individuais, segmento, dados de inflação e câmbio. O período de pandemia, que não foi contemplado neste trabalho, também pode afetar o resultado.

Para avaliar a precisão do MDA também é possível comparar o índice brasileiro com outros índices, mais estáveis, no mesmo horizonte de tempo analisado neste trabalho.

REFERÊNCIAS

ATKINS, Adam; NIRANJAN, Mahesan; GERDING, Enrico. Financial news predicts stock market volatility better than close price. *Journal of Finance and Data Science*, Amsterdam, v.4, n.2, 120–137, 2018.

HUANG, Sheng; LIU, Sheng. Machine learning on stock price movement forecast: the sample of the Taiwan stock exchange. *International Journal of Economics and Financial Issues*, Stockton, v. 9, n. 2, 189–201. 2019.

CHIA-CHENG, Chen; LIU, Yisheng; HSU, Ting-Hsin. An analysis on investment performance of machine learning: an empirical examination on Taiwan stock market, *International Journal of Economics and Financial Issues*, Taiwan, v. 9, n. 4, 1–10. 2019.

CHOUDHURY, Sasmita S.; SEN, Moumita, Trading in Indian stock market using ANN: a decision review. *Advances in Modelling and Analysis A*, Stockton, n. 54, v. 2, 252–262, 2017.

GU, Shihao; KELLY, Bryan; XIU, Dacheng. Empirical asset pricing via machine learning. *The Review of Financial Studies*, Chicago, v. 33, n. 5, 2223-2273, 2018.

DAMETTO, Ronaldo César. *Estudo da aplicação de redes neurais artificiais para a predição de séries temporais financeiras*. 2018. Dissertação (Mestrado em Engenharia de Produção) - Universidade Estadual Paulista Júlio de Mesquita Filho, Bauru, 2018. Disponível em <http://hdl.handle.net/11449/157058>

Data Science Academy. Deep Learning Book. [S.I.] 2021. Disponível em: <<https://www.deeplearningbook.com.br/>>. Acesso em: 10 Abril. 2021.

GUJARATI, Damodar N.; PORTER, Dawn C. *Econometria básica*. 5. ed. Rio de Janeiro: Mac Grow Hill. 2006.

GURESEN, Erkam; KAYAKUTLU, Gulgun; DAIMB, Tugrul U. Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, Amsterdam, v. 38, n. 8, 2011.

HAN, Jae Jung; KIM, Hyun-jung, Stock price prediction using multiple valuation methods based on artificial neural networks for KOSDAQ IPO companies, *Investment Analysts Journal*, United Kingdom, v. 50, n.1, 17–31, 2021.

HAYKIN, Simon S. *Redes neurais*. 2. ed. Porto Alegre: Bookman. 2001.

KONG, Ao; ZHU, Hongliang; AZENCOTT, Robert. Predicting intraday jumps in stock prices using liquidity measures and technical indicators. *Journal of Forecasting*, New Jersey, v. 40, n. 3, 416-438, 2020.

KOVÁCS, Zsolt László. *Redes neurais artificiais*. 1. ed. São Paulo: Editora Livraria da Física. 2002.

RUBINSTEIN, Mark. Markowitz's "Portfolio Selection": a fifty-year retrospective The *Journal of Finance*, New York, v. 57, n. 3, 1041-1045, 2002.

MEDEIROS, Luciano Frontino de. *Inteligência artificial aplicada: uma abordagem introdutória*, 1. ed., Curitiba: Editora Intersaberes, 2018.

MOGHADDAM, Amin Hedayati; MOGHADDAM, Moein Hedayati; ESFANDYARI, Morteza.. Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science*, Amsterdam, v. 21, n. 41, p. 89-93, 2016.

NDIKUM, Philip. Machine learning algorithms for financial asset price forecasting. *arXiv preprint arXiv:2004.01504*, Nova Iorque, 2020.

DI PERSIO, Luca; HONCHAR, Oleksandr. Recurrent neural networks approach to the financial forecast of Google assets. *International Journal of Mathematics and Computers in Simulation*, United States, v. 11, p. 7-13, 2017.

RASEKHSCHAFFE, Keywan Christian; JONES, Robert C. Machine Learning for Stock Selection. *Financial Analysts Journal*, Oxfordshire, v. 75, n. 3, p. 70-88, 2019.

SELVAMUTHU, Dharmaraja; KUMAR, Vineet; MISHRA, Abhishek. Indian stock market prediction using artificial neural networks on tick data. *Financial Innovation*, New York, v. 5, n. 1, p. 1-12, 2019.

SILVA, Reinaldo Oliveira da. *Teorias da administração*. 1. ed. São Paulo: Prentice Hall. 2001.

TKÁČ, Michal; VERNER, Robert; Artificial neural networks in business: Two decades of research. *Applied Soft Computing*, Amsterdam, v. 38, p. 788-804, 2016.

ZHANG, Yongjie; CHU, Gang; SHEN, Dehua. The role of investor attention in predicting stock prices: The long short-term memory networks perspective. *Finance Research Letters*, Amsterdam, v. 38, p. 101484, 2021.