

ANÁLISE DE ATIVOS FINANCEIROS POR MEIO DE REDE NEURAL RECORRENTE DO TIPO LSTM

VINÍCIOS ANTÔNIO PASSOS BALDUINO

UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO (UFTM)

DANIEL VITOR TARTARI GARRUTI

UNIVERSIDADE FEDERAL DE UBERLÂNDIA (UFU)

LEANDRO CRUVINEL LEMES

UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO (UFTM)

OSMAR ALÉSSIO

UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO (UFTM)

ANÁLISE DE ATIVOS FINANCEIROS POR MEIO DE REDE NEURAL RECORRENTE DO TIPO LSTM

1 Introdução

O mercado de ações é parte importante da economia de um país e quanto mais desenvolvido, maior é seu impacto econômico (LEVINE, 1991). Não é de se surpreender a quantidade de estudos acerca deste objeto, inclusive de áreas indiretamente relacionadas.

Teorias recentes abordam o comportamento dos investidores com certo viés, abrindo margem para que o mercado seja potencialmente previsível (LO, 2004). Mais especificamente, seria possível prever a movimentação dos preços em algum determinado momento. Caso o mercado seja de fato previsível significa que o investidor pode obter lucros anormais (acima do mercado) e geralmente em um menor nível de risco.

Em uma sociedade altamente tecnológica como a atual, a possibilidade do desenvolvimento de ferramentas antes impensáveis ou consideradas impossíveis de serem construídas, tornam-se realidade. São diversos os métodos estatísticos pelos quais é possível estudar séries temporais (como é o caso do histórico de preços de ações). No entanto, uma abordagem que vem se destacando ao longo dos anos e justamente substituindo parte destes métodos são as Redes Neurais Artificiais (RNAs) (ANGELO *et al.*, 2011; BULÍN *et al.*, 2019).

Uma vez que o mercado de ações pode ser previsível, explorar esta ferramenta pode ser útil para praticantes do mercado financeiro. Para explorar esta possível previsibilidade, faz sentido analisar a evolução do preço da ação sob um olhar das RNAs, permitindo aliar assim a complexidade advinda de inúmeros dados e informações disponíveis e padrões dos agentes de mercado, na tentativa de captar anomalias nas flutuações dos preços provenientes da atividade dos agentes (ZHANG *et al.*, 2021).

Garruti *et al.* (2022) destacam a importância de utilizar técnicas de *Machine Learning*, como é o caso das RNAs, para previsão de preços de ativos financeiros, ressaltando a superioridade dessas abordagens em relação ao modelo teórico, o qual se baseia na eficiência de mercado. Além disso, Gomes *et al.* (2019) afirmam que há evidências no mercado brasileiro de certa mudança no comportamento dos investidores associadas a questões políticas, favorecendo a previsibilidade em alguns casos.

Diante disso, o objetivo deste trabalho é examinar a capacidade preditiva de uma rede neural recorrente do tipo *Long Short Term Memory* (LSTM) – que tem demonstrado resultados interessantes, como relatado por Zhang *et al.* (2021) – aplicada a séries temporais financeiras no Brasil, comparando tais resultados com o modelo teórico e trabalhos anteriores.

Este trabalho justifica-se em vários aspectos, dentre eles: (i) a previsão de preços de ações no contexto brasileiro é essencial para ajudar os investidores, pois auxilia a tomada decisões de investimentos e no gerenciamento de riscos; (ii) melhorar a precisão dessas previsões pode fornecer um recurso lucrativo para os investidores; e, (iii) o uso de ferramentas sofisticadas como a LSTM pode ajudar a identificar padrões e tendências ocultos nos dados (ou melhor, invisível a percepção humana), fornecendo uma visão mais clara e informada para os investidores.

Os resultados mostram que LSTM foi capaz de superar o modelo teórico na maioria dos casos investigados, o que sinaliza a abrangência da HMA no mercado brasileiro. Foi preciso que a configuração da rede fosse ajustada em cada caso. Uma técnica de otimização de parâmetros foi aplicada, mas seus resultados foram inferiores aos demais, o que mostra a necessidade de maiores avanços nessa direção, especialmente ao manipular dados do mercado financeiro e contrapõe os resultados de Bulín, Šmídl e Švec (2019) e Rogachev e Melikhova (2020).

Assim, o estudo pode fornecer noções úteis sobre as características e dinâmicas específicas desse mercado, bem como sobre a eficácia das técnicas de LSTM nesse contexto, contribuindo para o avanço científico de pesquisas na área de previsões de preços por meio de RNAs.

Como trabalhos futuros, recomenda-se a implementação de ferramentas adicionais para a otimização dos hiperparâmetros, baseando-se por exemplo em Liu, Suzuki e Du (2023), bem como a análise de outros ativos.

Este trabalho está dividido em Introdução, que foi o tópico acompanhado agora, em seguida, a Fundamentação teórica, que visa explicar a base para a execução do trabalho. Metodologia, auxiliando a entender como o trabalho foi desenvolvido. Na sequência, os resultados acompanhados da discussão; e, por fim, as conclusões.

2 Literatura Relacionada

2.1 Hipóteses de Mercado Eficiente e evoluções

Se o mercado for previsível e uma vez devidamente previsto, o investidor poderá obter sempre o maior retorno possível, isso seria a utopia do mercado financeiro. Abordando esta possibilidade, se faz necessário discutir duas teorias principais, a Hipótese de Mercado Eficiente (HME) e a Hipótese de Mercado Adaptativo (HMA).

Segundo Mussa *et al.* (2008) a HME surge em meados de 1970. Tal teoria é significativa para o mercado até os dias atuais. A sua proposta é descrever o mercado como eficiente, sendo assim, tendo sempre refletido em seu preço todas as informações referentes a ele. Ou seja, não é possível obter retornos acima da média do mercado (FAMA, 1970).

De acordo com Gomes *et al.* (2019) a teoria da HME trata o ser humano como “*Homo Economics*” e com isso, este tomaria sempre as melhores decisões frente ao mercado financeiro. O que na prática pode nem sempre ser verdade, pois essa forma de pensar ignora no ser humano suas decisões com base na emoção. Em contrapartida (LO, 2004), intersecta a psicologia humana à economia, e propõe a HMA, que tem por objetivo verificar a previsibilidade do mercado por ciclos, ou seja, entre eventos semelhantes, poderia haver a possibilidade do mercado repetir a iniciativa tomada anteriormente, o que garantiria uma certa previsibilidade.

De acordo com Dourado e Tabak (2014) países com a economia ainda em desenvolvimento tendem a possuir um mercado menos eficiente, além disso afirma que apesar de não concluir que o mercado brasileiro segue a tendência de mercado adaptativo, não é possível descartar essa possibilidade.

Até o momento foi possível compreender que o mercado, pode apresentar comportamento previsível e que os dados obtidos pela movimentação deste, podem ser interpretados como séries temporais. Dentre os vários estudos que envolvem séries temporais, a análise por meio de redes neurais tem tomado espaço e se destacando frente aos temas de pesquisas atuais.

2.2 Redes neurais artificiais e seu emprego em Finanças

Segundo Rauber (2005) uma Rede Neural Artificial (RNA) tem seu funcionamento baseado nos neurônios do cérebro, e de acordo com Angelo *et al.* (2011), a rede pode ser observada em pequenas partes que são responsáveis por tarefas de atribuições de pesos. O conceito matemático de atribuir pesos a uma informação que é feito pela RNA é o que aproxima seu funcionamento ao do cérebro humano.

Existem diferentes modelos de redes neurais, essa variedade se justifica tanto pelos vários problemas computacionais quanto pela evolução destas. Não cabe ao escopo deste trabalho um detalhamento de diferentes modelos de redes neurais ou da evolução histórica destes estudos, o que se segue é uma apresentação didática do funcionamento básico de uma rede neural com os

modelos primitivos e a discriminação do modelo de rede neural memória de curto e longo prazo (*Long Short Term Memory* - LSTM).

O modelo simplificado de uma RNA é o *perceptron* (RAUBER, 2005), representado na Figura 1, sua arquitetura consiste em uma camada de neurônios de entrada, responsável pelo recebimento de cada informação. A quantidade de neurônios da camada é equivalente com a quantidade de informações que serão passadas à rede por vez. Além desta, uma camada de neurônios de saída, responsável pela resposta da rede neural, esta camada pode conter um ou mais neurônios, dependendo do problema computacional que a rede foi desenvolvida para resolver. Todos os neurônios da camada de entrada são ligados a cada neurônio da camada de saída. Estas ligações recebem o nome de sinapses. O funcionamento desta rede acontece justamente pela capacidade das sinapses atribuírem diferentes pesos para cada ligação, e através do resultado destas atribuições com a soma de um determinado número b , chamado bias, geram resultados acerca de cada informação.

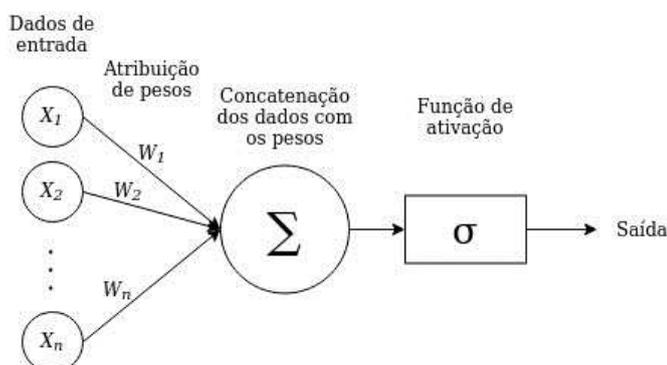


FIGURA 1 – Rede neural artificial do tipo *perceptron*. Fonte: Adaptado de Rauber (2005).

Um exemplo da aplicação deste modelo seria na previsão de preço de um determinado ativo. Existem fatores que influenciam o preço, como o nível de risco (GOMES *et al.*, 2019), o histórico de preços (GARRUTI *et al.*, 2021; ZHANG *et al.*, 2021), dentre vários outros. Desta forma a rede poderia encontrar um valor justo para o ativo, baseado nos padrões encontrados na sua elaboração.

Caso seja necessário, há a possibilidade de construir a rede com camadas intermediárias, as quais recebem o nome de camadas ocultas. As camadas ocultas não possuem um número pré-definido de neurônios. Sendo assim, pode ser ajustado para melhor desempenho da RNA. A função destas camadas é a de auxiliar a distribuição de pesos para cada dado/informação da camada de entrada. A Figura 2 apresenta a forma primitiva desta estrutura da rede conhecida como *Multilayer Perceptron*, que tem sido amplamente aplicada a problemas de previsão de preços (LABOISSIERE *et al.*, 2015).

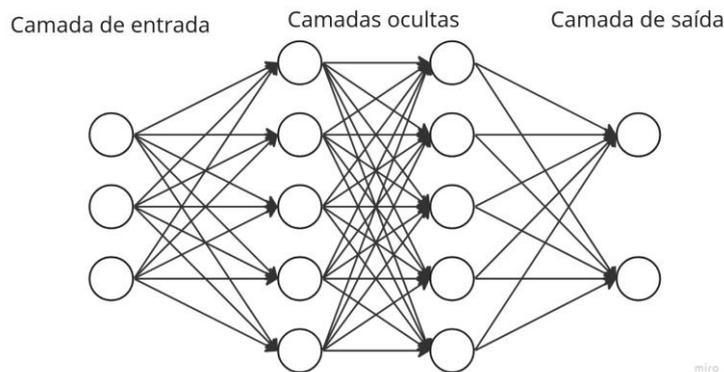


FIGURA 2 – Rede neural artificial do tipo *Multilayer Perceptron*. Fonte: Adaptado de Angelo *et al.* (2011).

Com a evolução dos estudos de redes neurais outros modelos puderam ser desenvolvidos, o que permitiu a criação de redes neurais mais eficazes para a resolução de determinados problemas.

Com o desenvolvimento de novas arquiteturas de redes um novo conceito também surgiu, as redes neurais recorrentes (RNR). Enquanto no exemplo de rede anterior a informação era tratada linearmente, entrando pela camada de entrada e seguindo as várias camadas ocultas que a rede pode apresentar até a camada de saída, na RNR a informação não segue necessariamente esse caminho. Os neurônios da camada de entrada podem receber informação da camada de saída, bem como os neurônios das camadas ocultas podem trocar informação entre si. Arquiteturas assim expõe a rede a mais dados, o que pode implicar na melhor distribuição de pesos para as informações (ZHANG *et al.*, 2021; GÜLMEZ, 2023).

Uma RNR que tem sido objeto de estudos atualmente é a rede do tipo memória de curto e longo prazo, popularmente conhecidas como *Long Short-Term Memory* (LSTM) (BULÍN *et al.*, 2019; GARRUTI *et al.*, 2021; ZHANG *et al.*, 2021; GÜLMEZ, 2023). A rede do tipo LSTM se diferencia das outras RNAs, por não possuir camadas de neurônios e sim, células de memória (Figura 3). Estas células possuem sinapses assim como os neurônios das redes anteriores, porém tais sinapses não possuem a tarefa de fazer atribuição de pesos e sim transmitir a informação.

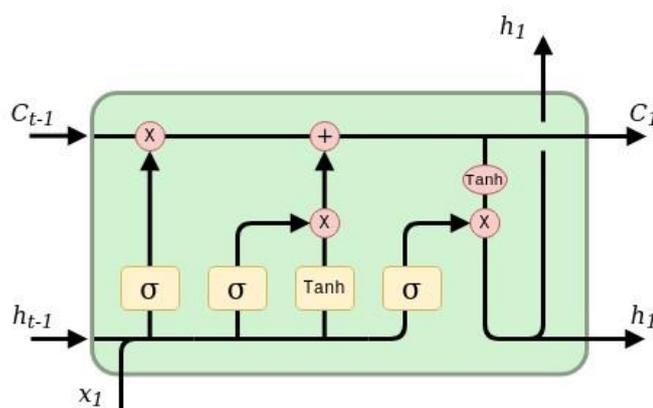


FIGURA 3 – Arquitetura de célula do tipo LSTM. Fonte: adaptado de Zhang, Chu e Shen (2021).

Cada célula de memória desta está conectada com a sua anterior e posterior, isso faz com que a informação seja transmitida entre elas. O caminho superior na célula, destacado na figura 4 é o estado da memória, ele é responsável por armazenar todas as conclusões que a rede tira a partir dos dados. O estado da memória de uma célula pode ser passado para a seguinte, se passado isso faz com que a célula atual utilize das conclusões da célula anterior para chegar nas próprias. Podemos observar que o estado da memória está conectado com todos os demais

processos da célula, logo cada informação que a célula recebe e devolve influência e é influenciada por ele.

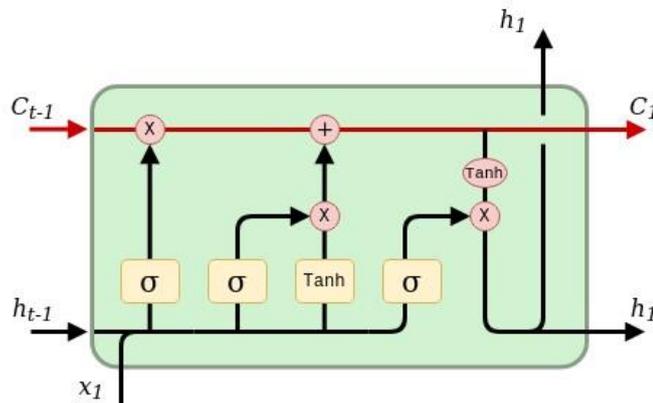


FIGURA 4 - Estado da memória da célula LSTM. Fonte: adaptado de Zhang, Chu e Shen (2021).

Na parte inferior esquerda da arquitetura existe entrada de dados, tanto dos novos, quanto os dados de saída da célula anterior e são concatenados. A primeira interação dos dados com a célula acontece no portão de esquecimento, ilustrado na figura 5, para este portão é atribuído a letra f de *forget*. O portão de esquecimento determina o quanto de informação será perdida do estado da memória anterior, para que assim possa receber novas informações.

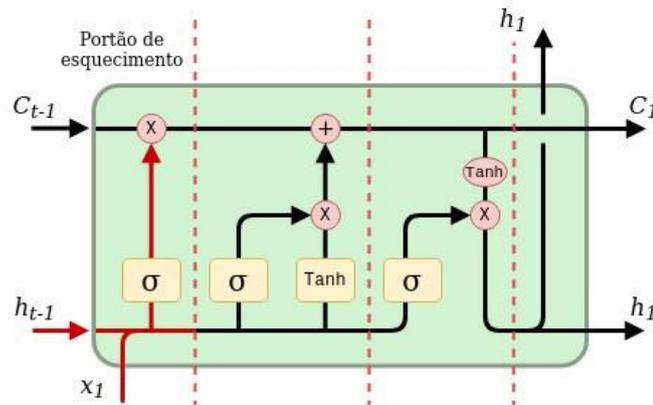


FIGURA 5 – Entrada de dados e portão de esquecimento. Fonte: adaptado de Zhang, Chu e Shen (2021).

Após o portão de esquecimento tem o de entrada apresentado na figura 6. Este portão possui duas partes, para a primeira é aplicada uma função *sigmoid* nos dados, esta parte é representada pela letra i de input. Na segunda parte é aplicada uma função tangente hiperbólica ("Tanh") nos mesmos dados esta parte é representada pela letra C por se tratar de um estado de memória parcial. Neste portão são selecionadas as informações que serão armazenadas no estado da memória e o quanto elas influenciarão o estado.

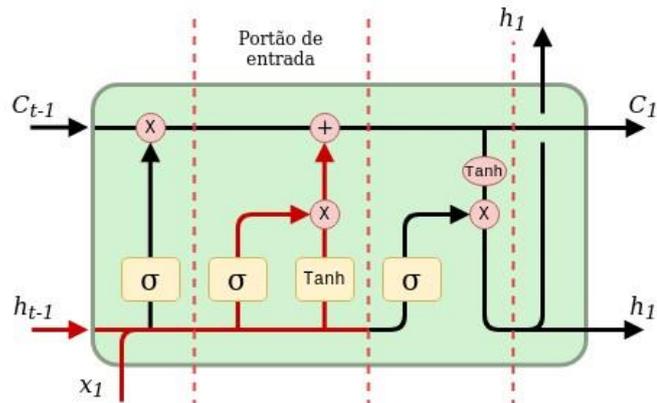


FIGURA 6 – Portão de entrada. Fonte: adaptado de Zhang, Chu e Shen (2021).

Para concluir a interação com a célula, os dados precisam passar pelo portão de saída, apresentado na figura 7. Neste momento a informação que entrou no começo da célula sofre todas as distribuições de peso conforme o estado da célula atual, gerando assim novos dados a serem passados para a próxima célula ou a conclusão do processo, caso seja a última célula.

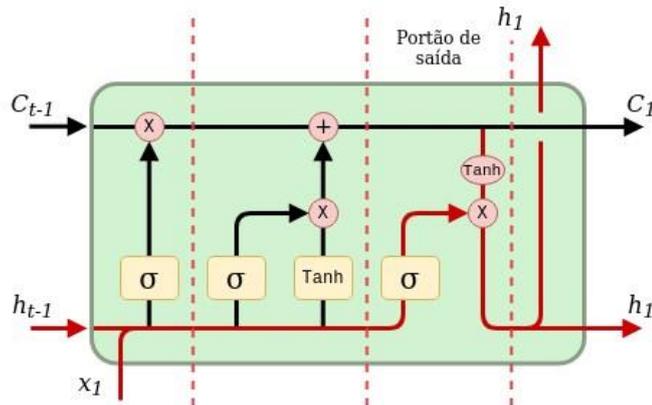


FIGURA 7 – Portão de saída. Fonte: adaptado de Zhang, Chu e Shen (2021).

Quando a rede neural é iniciada ela precisa passar por um processo de treinamento, visto que os pesos ideais para cada informação são desconhecidos. A princípio é necessário que os dados sejam divididos entre o conjunto de treino e o conjunto de testes. Geralmente entre 70% e 90% dos dados são separados para o conjunto de treino, ou seja, são expostos à rede para que seja feito o processo de balanceamento dos pesos. Os demais dados são separados para o conjunto de testes, úteis para avaliação do seu desempenho.

A atribuição de pesos da rede neural por meio das suas interações com o conjunto de dados de treino. A parte responsável dentro da RNA por controlar os pesos da rede, é a chamada função de perda, ou *loss*. A partir dela, a rede altera alguns valores de pesos (viés) e repete o processo. O ideal é que o valor se aproxime de 0, indicando assim que a rede está conseguindo fazer uma melhor distribuição de valores a cada interação.

Após estabilização de pesos é preciso avaliar o desempenho da rede, fazendo as previsões e comparando com os dados de teste. Ao final do processo existem três possíveis diagnósticos para o aprendizado da rede. Ao observar o comportamento da função de perda e este seja um decaimento progressivo, significa que a rede aprendeu. Com isso, suas previsões serão baseadas nos padrões observados no treino. Caso contrário, a rede pode estar sub ou sobre ajustada (conhecidos como *underfitting* e *overfitting*, respectivamente). Numa rede com *underfitting*, há uma taxa de erro sem decaimento tanto no conjunto de treino como no conjunto de teste. Já

numa rede com *overfitting*, as previsões possuem baixo nível de erro na amostra de treino, porém isso não acontece com a amostra de teste.

3 Metodologia

Para o desenvolvimento do trabalho utilizou-se das metodologias de pesquisa quantitativa, fazendo uso de coleta de dados, análise e tratamento, bem como a análise dos resultados. Toda a parte de código deste trabalho foi desenvolvida na linguagem de programação Python por meio da ferramenta Colab do Google. A utilização desta ferramenta permite que o código seja armazenado e execute em nuvem, trazendo maior praticidade.

3.1 Séries analisadas e descrição dos dados

Os conjuntos de dados analisados neste trabalho consistem em dados reais da Brasil, Bolsa, Balcão (B3). Foram extraídos do site YahooFinance as séries temporais referentes aos preços de abertura, máximo, mínimo, fechamento e volume de negociações das ações de 5 empresas brasileiras por meio da biblioteca *yfinance*, quais sejam: Ambev (ABEV3), do setor alimentício; Banco do Brasil (BBAS3) e Banco Bradesco (BBDC4) do setor bancário; e, duas do setor elétrico, Taesa (TAEE3) e Companhia Elétrica de Minas Gerais (CMIG4). O período analisado consiste na janela entre os dias 01 de janeiro de 2008 a 31 de dezembro de 2022, totalizando assim 15 anos. A escolha destes ativos se faz pela possibilidade de comparar os resultados com o trabalho de Garruti *et al.* (2022) que se dedicou ao setor elétrico, além de poder testar setores diferentes e, portanto, avaliar o potencial de generalização dos modelos de previsão.

3.2 Pré processamento dos dados

O objetivo é a previsão dos preços máximos e mínimos diários da ação no dia seguinte, para isso foram utilizadas as variáveis exógenas de abertura, baixa e fechamento além da própria série de alta.

Foi estipulada uma janela temporal de três dias para a previsão de preço do quarto dia. No caso, foram desenvolvidos dois modelos para cada ativo, sendo um predictor de preços de máximo (*High*) e outro para preços de mínimo (*Low*). Alinhada com uma estratégia de investimento com operações *intraday*, pode ser mais lucrativo do que prever somente o preço de fechamento. Por exemplo, Garruti *et al.* (2022) mostra que usar tais previsões podem gerar lucros anormais expressivos, chegando a um ganho 10 vezes maior que o índice de mercado.

Foram separados de maneira linear 80% dos dados para treinamento dos modelos, representando o intervalo entre os dias 01 de janeiro e 2008 e 02 de janeiro de 2020. O restante ficou reservado para teste do desempenho dos mesmos, sendo assim, o conjunto de treino contou com um total de 2.978 amostras, cada amostra com três passos no tempo e quatro variáveis.

Devido ao modelo ser desenvolvido seguindo uma arquitetura LSTM, é necessário que os dados estejam escalonados. O método utilizado foi o *MinMaxScaler* e as métricas de avaliação também são da biblioteca *Scikit Learn*, o qual retorna sempre um valor entre 0 e 1.

Para a arquitetura dos modelos foi utilizada a biblioteca *TensorFlow* por meio da interface da biblioteca *Keras*, ambas as arquiteturas possuem a mesma estrutura (vide Tabela 1), o que as diferencia são alterações em parâmetros das camadas. Ambos os modelos foram construídos utilizando o método *Sequential*, método que desenvolve uma arquitetura de empilhamento de camadas, o que permite trabalhar com o conceito de aprendizado profundo. Tal proposta segue o modelo empregado por Gülmez (2023).

TABELA 1 – Arquitetura Base do Modelo de Previsão utilizado neste estudo. Os dados são dimensionados considerando x observações, o tamanho da janela de dados j e a quantidade n_i de neurônios. *Dense* é um tipo de RNA nos quais os neurônios estão todos conectados.

Camada	Tipo	Estrutura de Dados	Parâmetros
Entrada	Dados em Lista (Sinais)	(x, j, v)	Não há
1a. Camada Oculta	LSTM	$(n.d., j, n1)$	Ativação
	<i>Dropout</i>	$x \in [0,1)$	
2a. Camada Oculta	LSTM	$(n.d., j, n2)$	Ativação
	<i>Dropout</i>	$x \in [0,1)$	
3a. Camada Oculta	<i>Dense</i>	$(n.d., n3)$	Ativação
Camada de Saída	<i>Dense</i>	$(n.d., 1)$	Ativação = Linear

Fonte: Elaborada pelos autores.

O primeiro modelo proposto, denominado "LSTM 1", foi construído com $n1 = n2 = 100$, sem *dropout*, $n3 = 25$. Após as duas camadas LSTM, foi construída uma camada densa com 25 neurônios todos conectados entre si e por fim uma camada densa de saída. As funções de ativação das camadas LSTM para todos os modelos foram a tangente hiperbólica, enquanto na camada densa foi linear.

Como o problema computacional trabalhado aqui é de regressão [do preço], a saída deve conter um único neurônio. Ainda, a evolução do treino baseou-se em 300 iterações (*epochs* = 300) e amostragem para ajuste de pesos (*batch size* = 100) e taxa de aprendizagem fixada em 0,01. O segundo modelo, o "LSTM 2" utilizou a mesma arquitetura, alterando-se apenas a taxa de aprendizagem para 0,001. Para a otimização do modelo foi utilizado o algoritmo de estimativa de momento adaptativo ADAM (KINGMA; BA, 2014) e a função de perda MAPE, definida a seguir.

Um terceiro modelo foi construído na tentativa de identificar os hiperparâmetros ótimos para toda a arquitetura, denominando-o por "LSTM *tuned*". Para tanto, aplicou-se a técnica de *Hyperband*, proposta por Li *et al.* (2017). Assim, buscou-se identificar a melhor combinação de $n1$, $n2$, $n3$, funções de ativação, taxa de *dropout* e taxa de aprendizagem que fornecesse a menor taxa de erro no treino da rede. Tal processo foi aplicado separadamente para cada ativo. Isso se justifica pela possível singularidade de cada ativo que, ao testar uma rede com parâmetros fixos, simplifica o entendimento de que o padrão de oscilação dos preços seria o mesmo para todos eles.

De acordo com Rogachev e Melikhova (2020), as tecnologias de redes neurais têm sido bem-sucedidas na área da economia. No entanto, a seleção de hiperparâmetros pode aprimorar a qualidade do treinamento. Neste estudo referenciado, foi observada uma melhoria de 9% na análise de informações textuais após a utilização de seleção automatizada.

Ao todo para cada série (uma de preços máximos e outra para preços mínimos, seguindo Garruti *et al.*, 2021 e Laboissiere *et al.*, 2015) foi possível recolher três resultados de modelos de redes neurais e dois de modelos teóricos. Foram realizadas as predições e comparadas com o tradicional modelo teórico, o de passeio aleatório (*random walk*), que é considerado o melhor resultado num mercado eficiente, segundo Fama (1995); e a média móvel de 7 dias, considerando que o preço da ação no momento t é a média do preço dos 7 dias anteriores a ele.

3.3 Métricas de desempenho

As métricas para avaliação das predições da rede foram a erro médio absoluto *Mean Absolute Error* (MAE); o erro percentual médio *Mean Absolute Percentage Error* (MAPE); a média do

erro quadrado *Mean Squared Error* (MSE), é o mesmo erro avaliado pela função de perda, no entanto aqui ele não possui a função de ajudar na configuração da rede, somente auxiliar na avaliação do desempenho desta; e por fim a raiz quadrada do quadrado da média de erro *Root Mean Squared Error* (RMSE).

O MAE é dado pela seguinte fórmula:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

É utilizado o módulo dos erros, pois há casos em que o erro negativo pode zerar o positivo fazendo com que o valor deste erro seja menor que o real, passando uma ideia de que o modelo é mais preciso do que de fato é. Esta métrica aponta o quão distante está a previsão do valor real, logo, quanto mais próximo de 0 o resultado for, melhor é o modelo, porém, não se pode afirmar a sua eficácia com base em somente uma métrica.

O cálculo do MAPE é feito pela seguinte fórmula:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (2)$$

Assim como a maioria das métricas quanto mais próximo de 0 significa que melhor foi o aprendizado do modelo. Enquanto a métrica anterior trazia a média da distância de previsão em valores absolutos, aqui essa distância é representada em valores percentuais.

O MSE segue a fórmula:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

O uso desta métrica deixa o modelo sensível a valores discrepantes, permitindo uma análise mais aprofundada do erro absoluto. O RMSE (raiz quadrada de MSE) é mais apropriado, já que é comparável a MAE.

4 Resultados e discussões

Ao analisar a previsão dos preços máximos de cinco ativos da bolsa brasileira, observa-se um desempenho superior dos modelos de redes neurais em quatro deles em relação aos modelos teóricos. A exceção ocorre no caso do ativo TAEE3, que apresentou melhores resultados com o modelo de *random walk*. Nenhum dos ativos apresentou os melhores resultados ao utilizar a média móvel de 7 dias. Esses resultados estão alinhados com a Hipótese do Mercado Adaptativo proposta por Lo (2004), o qual considera o mercado previsível, e divergem da Hipótese do Mercado Eficiente proposta por Fama (1970), com exceção de um único ativo.

Em relação aos modelos de redes neurais, observa-se um equilíbrio entre os dois primeiros em termos dos melhores resultados, pois cada um deles apresentou melhor desempenho em dois ativos. No entanto, o modelo LSTM *tuned* não obteve o melhor resultado em nenhum dos ativos e, pelo contrário, teve o pior desempenho em três dos cinco ativos, enquanto o modelo MA7 teve os piores resultados nos outros dois ativos. Esses resultados contradizem as conclusões de Bulín, Šmídl e Švec (2019) e Rogachev e Melikhova (2020), pois o modelo com parâmetros otimizados não demonstrou um desempenho superior em relação aos demais modelos.

TABELA 2 - Métricas de desempenho na previsão de preços máximos.

Ativo	Modelo	HIGH			
		MAE	MAPE	MSE	RMSE
ABEV3	LSTM 1	0,20	1,31	0,07	0,26
	LSTM 2	0,18	1,18	0,07	0,25
	LSTM <i>tuned</i>	0,20	1,31	0,08	0,28
	RW	0,20	1,35	0,08	0,29
	MA7	0,31	2,05	0,18	0,42
BBAS3	LSTM 1	0,53	1,53	0,61	0,78
	LSTM 2	0,51	1,46	0,58	0,76
	LSTM <i>tuned</i>	1,09	3,14	1,74	1,32
	RW	0,55	1,60	0,73	0,86
	MA7	0,82	2,38	1,41	1,19
BBDC4	LSTM 1	0,26	1,33	0,14	0,38
	LSTM 2	0,38	1,94	0,22	0,47
	LSTM <i>tuned</i>	1,00	5,28	1,24	1,11
	RW	0,29	1,53	0,18	0,43
	MA7	0,45	2,34	0,37	0,60
CMIG4	LSTM 1	0,12	1,30	0,03	0,18
	LSTM 2	0,16	1,67	0,05	0,22
	LSTM <i>tuned</i>	0,13	1,39	0,03	0,18
	RW	0,14	1,55	0,04	0,21
	MA7	0,23	2,44	0,09	0,31
TAEE3	LSTM 1	0,45	3,76	0,55	0,74
	LSTM 2	1,19	9,31	2,26	1,50
	LSTM <i>tuned</i>	1,97	15,51	6,23	2,50
	RW	0,11	0,96	0,03	0,17
	MA7	0,17	1,38	0,05	0,23

Fonte: Dados dos autores.

Para a previsão de preços mínimos, os resultados são semelhantes aos preços máximos. É possível observar um desempenho superior dos modelos de redes neurais em todos os ativos, com exceção novamente da TAEE3. Isso confirma mais uma vez a concordância deste estudo com a HMA e contrapõe a HME.

Observa-se ainda a baixa performance do LSTM *tuned*, que não apresentou o melhor desempenho para nenhum dos ativos, contrariando novamente os resultados de Bulín, Šmídl e Švec (2019) e Rogachev e Melikhova (2020). No entanto, diferentemente da previsão de preços máximos, no caso dos preços mínimos, foi o modelo teórico da média móvel de 7 dias que obteve os piores resultados para a maioria dos ativos.

TABELA 3 - Métricas de desempenho na previsão de preços mínimos.

Ativo	Modelo	LOW			
		MAE	MAPE	MSE	RMSE
ABEV3	LSTM 1	0,22	1,55	0,10	0,31
	LSTM 2	0,18	1,25	0,07	0,26
	LSTM <i>tuned</i>	0,18	1,29	0,08	0,27
	RW	0,20	1,38	0,08	0,27
	MA7	0,30	2,08	0,17	0,42
BBAS3	LSTM 1	0,48	1,48	0,67	0,82
	LSTM 2	0,53	1,60	0,80	0,90
	LSTM <i>tuned</i>	0,60	1,81	0,74	0,86
	RW	0,55	1,67	0,72	0,85
	MA7	0,83	2,51	1,48	1,22
BBDC4	LSTM 1	0,25	1,38	0,14	0,38
	LSTM 2	0,32	1,71	0,20	0,45
	LSTM <i>tuned</i>	0,26	1,43	0,15	0,39
	RW	0,28	1,53	0,16	0,41
	MA7	0,45	2,43	0,38	0,61
CMIG4	LSTM 1	0,12	1,36	0,03	0,18
	LSTM 2	0,14	1,56	0,04	0,21
	LSTM <i>tuned</i>	0,14	1,60	0,04	0,19
	RW	0,14	1,63	0,04	0,21
	MA7	0,23	2,64	0,10	0,32
TAEE3	LSTM 1	0,41	3,61	0,42	0,65
	LSTM 2	1,11	8,64	2,06	1,44
	LSTM <i>tuned</i>	2,42	19,84	11,85	3,44
	RW	0,13	1,10	0,04	0,21
	MA7	0,16	1,41	0,06	0,24

Fonte: Dados dos autores.

De maneira geral, infere-se a ineficiência do mercado de ações brasileiras, ou ao menos para quatro dos cinco ativos analisados, indo de encontro a outros estudos brasileiros que utilizaram técnicas de inteligência artificial para a previsão de preços, como Garruti *et al.* (2021), Garruti *et al.* (2022) e Laboissiere (2015). No entanto, ambos analisaram apenas companhias do setor de energia, enquanto este estudo abrangeu demais setores, apresentando assim uma evolução nos estudos.

5 Conclusões

O objetivo deste estudo foi realizar a previsão diária de preços máximos e mínimos de 5 ativos da bolsa brasileira. Para isso foi utilizado 3 modelos de redes neurais LSTM e dois baseados na teoria de investimentos. Por fim, utilizou-se métricas de erro para verificação dos resultados obtidos.

Verificou-se uma vantagem dos modelos de redes neurais em geral, não sendo verdade para apenas um dos ativos que obteve melhor desempenho quando utilizado o modelo de *random walk*. Desta maneira, nota-se uma aproximação do estudo com a Hipótese de Mercados Adaptativos e contraposição em relação a Hipótese de Mercados Eficientes, exceto no ativo TAEE3.

Nota-se também uma contraposição aos estudos de Bulín, Šmídl e Švec (2019) e Rogachev e Melikhova (2020), os quais obtiveram melhores resultados ao utilizar os parâmetros automatizados. Pois o modelo LSTM *tuned*, implementado neste artigo, não obteve os melhores resultados para nenhuma das previsões, e pelo contrário, foi o pior em algumas delas.

Apesar de obter resultados interessantes na previsão dos preços diários, nota-se possíveis fragilidades no estudo, como a utilização dos hiperparâmetros. Diante disso, sugere-se em estudos futuros a utilização de outras ferramentas para a otimização, como as utilizadas por Liu, Suzuki, e Du (2023). Além disso, é possível estender a análise para outros ativos da bolsa brasileira, ou de outros países.

Referências

ANGELO, Claudio Felisoni de et al. Séries temporais e redes neurais: uma análise comparativa de técnicas na previsão de vendas do varejo brasileiro. **BBR-Brazilian Business Review**, v. 8, n. 2, p. 1–21, 2011.

BULÍN, Martin; ŠMÍDL, Luboš; ŠVEC, Jan. On using stateful LSTM networks for key-phrase detection. In: **Text, Speech, and Dialogue: 22nd International Conference, TSD 2019, Ljubljana, Slovenia, September 11–13, 2019, Proceedings 22**. Springer International Publishing, 2019. p. 287-298.

DOURADO, Glener DE Almeida; TABAK, Benjamin Miranda. Teste da Hipótese de Mercados Adaptativos para o Brasil. **Revista Brasileira de Finanças**, v. 12, n. 4, p. 517, 2014.

FAMA, Eugene F. Efficient capital markets: A review of theory and empirical work. **The Journal of Finance**, v. 25, n. 2, p. 383-417, 1970.

FAMA, Eugene F. Random walks in stock market prices. **Financial analysts journal**, v. 51, n. 1, p. 75-80, 1995.

GARRUTI, Daniel Vitor Tartari et al. Aplicação de LSTM e MLP para prever preço de ações no Brasil. In: **EGEN - Encontro de Gestão e Negócios - 2021**. [S.l.: s.n.], 2021.

GARRUTI, Daniel Vitor Tartari et al. Qualidade de previsões dos preços máximos e mínimos: uma análise aplicada a ações do setor elétrico. In: **XXIX SIMPEP**. [S.l.: s.n.], 2022.

GOMES, Lara Ferreira de Almeida; SOUZA, Paulo Vitor Souza de; SILVA, César Augusto Tibúrcio. Hipótese dos mercados adaptativos: uma análise da eficiência de títulos do mercado acionário brasileiro. **Revista de Contabilidade do Mestrado em Ciências Contábeis da UERJ**, v. 24, n. 3, p. 25, 2019.

GÜLMEZ, Burak. Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. **Expert Systems with Applications**, v. 227, p. 120346, 2023.

KINGMA, Diederik P.; BA, Jimmy Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

- LABOISSIERE, Leonel A.; FERNANDES, Ricardo A.S.; LAGE, Guilherme G. Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks. **Applied Soft Computing**, v. 35, p. 66-74, 2015.
- LEVINE, Ross. Stock markets, growth, and tax policy. **The Journal of Finance**, v. 46, n. 4, p. 1445-1465, 1991.
- LI, Lisha et al. Hyperband: A novel bandit-based approach to hyperparameter optimization. **The journal of machine learning research**, v. 18, n. 1, p. 6765-6816, 2017.
- LIU, Wei; SUZUKI, Yoshihisa; DU, Shuyi. Forecasting the Stock Price of Listed Innovative SMEs Using Machine Learning Methods Based on Bayesian optimization: Evidence from China. **Computational Economics**, p. 1-34, 2023.
- LO, A. W. The adaptive markets hypothesis. **The Journal of Portfolio Management, Institutional Investor Journals Umbrella**, v. 30, n. 5, p. 15–29, 2004.
- MUSSA, Adriano et al. Hipótese de mercados eficientes e finanças comportamentais: as discussões persistem. **FACEF pesquisa**, v. 11, n. 1, p. 5-17, 2008.
- RAUBER, Thomas Walter. Redes neurais artificiais. **Universidade Federal do Espírito Santo**, v. 29, 2005.
- ROGACHEV, A. F.; MELIKHOVA, E. V. Automation of the process of selecting hyperparameters for artificial neural networks for processing retrospective text information. In: **IOP Conference Series: Earth and Environmental Science**. IOP Publishing, 2020. p. 012012.
- YAHOO Finance. Acessado em: 28/02/2023. Disponível em: <https://finance.yahoo.com/>.
- 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**, v. 38, p. 101484, 2021.