

Área Temática:
FINANÇAS

Título:
RETORNO DE AÇÕES UTILIZANDO *DEEP LEARNING*: UM ESTUDO SOBRE A CAPACIDADE PREDITIVA DA REDE LSTM E O AJUSTE DO NÚMERO DE ÉPOCAS

RESUMO

Este estudo teve por objetivo identificar um intervalo do número de épocas no treinamento de uma rede neural recorrente LSTM que minimiza o erro de predição sobre séries de retornos de ações brasileiras. A perspectiva teórica que embasa o artigo é a Hipótese dos Mercados Adaptáveis, em que mercados dinâmicos e estocásticos, oportunidades de lucro e a competição entre investidores enviesados e eficazes sujeitariam as estratégias de investimentos a ciclos de ganhos e perdas, favorecendo o desenvolvimento de modelos de gestão eficiente de ativos. A hipótese de pesquisa é existência de um intervalo de épocas que minimizaria o erro de predição de uma rede LSTM sobre retornos de ações brasileiras. O constructo dependente é a capacidade preditiva mensurada pelo erro de predição (acurácia). A regularização por *early stopping*, como constructo independente, é expressa pelo número de épocas de treinamento. A abordagem do trabalho é quantitativa, baseada em dados secundários em recorte longitudinal dos preços de fechamento de ações coletados na B3, no período de 2008 a 2018, em amostra com 2718 pontos de dados para cada ativo. Como economia emergente, as diferenças do mercado brasileiro em relação aos desenvolvidos implicam em maior nível de risco e portanto em retornos esperados mais elevados. As ações foram escolhidas pela liquidez, capitalização de mercado e continuidade do *ticker* ao longo do recorte. Os preços foram transformados em retornos e uma regressão quadrática foi utilizada para identificar a existência de ponto de mínimo. O modelo gerou predições de um dia e o erro de predição foi medido em função do número de épocas. Os resultados não indicaram a existência de pontos de inflexão mínimos do erro de predição, para ambas as ações, refutando a hipótese deste estudo. Embora o comportamento preditivo possa variar em função do modelo utilizado, os resultados não confirmaram o padrão comumente descrito na literatura. O uso de retornos, em vez de preços, pode ter atenuado os componentes não estacionários, limitando o desempenho preditivo. A contribuição deste trabalho foi evidenciar um comportamento não esperado do erro de predição dos retornos de ações de uma rede LSTM, submetida à regularização por *early stopping*, no mercado brasileiro. O ajuste de épocas deve ser criterioso na ausência de regiões de mínimos na curva do erro de predição para evitar o risco de *overfitting*.

Palavras-chave

Retorno de ações, predição, *Machine Learning*, LSTM, regularização, *early-stopping*.

Abstract

This study aimed to identify a range in the number of epochs in the training of an LSTM recurrent neural network that minimizes the prediction error on series of Brazilian stock returns. The theoretical perspective that underpins the article is the Adaptive Markets Hypothesis, in which dynamic and stochastic markets, profit opportunities and competition between biased and effective investors would subject investment strategies to cycles of gains and losses, favoring the development of efficient asset models. The research hypothesis is the existence of an interval of epochs that would minimize the prediction error of an LSTM network on Brazilian stock returns. The dependent construct is the predictive capacity measured by the prediction error (accuracy). Regularization by early stopping, as an independent construct, is expressed by the number of training epochs. The work approach is quantitative, based on secondary data in a longitudinal section of closing share prices collected at B3, from 2008 to 2018, in a sample with 2718 data points for each asset. As an emerging

economy, the differences between the Brazilian market and developed markets imply a higher level of risk and therefore higher expected returns. The shares were chosen based on their liquidity, market capitalization and ticker continuity throughout the period. Prices were transformed into returns and a quadratic regression was used to identify the existence of a minimum point. The model generated one-day predictions and the prediction error was measured as a function of the number of epochs. The results did not indicate the existence of minimum inflection points in the prediction error for both actions, refuting the hypothesis of this study. Although the predictive behavior may vary depending on the model used, the results did not confirm the pattern commonly described in the literature. Using returns rather than prices may have attenuated nonstationary components, limiting the predictive performance. The contribution of this work was to highlight an unexpected behavior of the prediction error of stock returns in an LSTM network, subject to regularization by early stopping, in the Brazilian market. The epoch adjustment must be careful in the absence of minimum regions in the curve of prediction error to avoid the risk of overfitting.

Keywords

Stock return, forecast, Machine Learning, LSTM, regularization, early-stopping.

1. INTRODUÇÃO

O desenvolvimento de algoritmos de *Machine Learning* (ML) para a predição de retornos de ativos tem sido objeto de interesse tanto da pesquisa em finanças como da indústria financeira. A Hipótese dos Mercados Eficientes (EMH) considera que a evolução das preferências racionais dos investidores resulta em equilíbrio nas distribuições dos retornos de ações ao longo do tempo (*weak form*), assim não seria possível uma predição baseada em preços históricos, pois o preço atual já incorporaria o passado (FAMA, 1965; FAMA, 1970). Lo e MacKinlay (1988), entretanto, demonstraram que os preços de carteiras de ações e índices não seguem necessariamente um Passeio Aleatório e exibem correlação serial. Encontraram explicações na teoria comportamental sobre vieses humanos, como o apetite ao risco e a aversão à perda, contrariando o comportamento racional (KAHNEMAN; TVERSKY, 2013). Assim, os preços passados poderiam prever retornos por meio de algoritmos que detectassem padrões nas séries temporais financeiras (LO; MAMAYSKY; WANG, 2000). Inicialmente, a negociação de ações e futuros baseada em algoritmos, *Algorithmic Trading* (AT), resultou em maior escalabilidade, velocidade e menor propensão a erros e vieses humanos (HENDERSHOTT; JONES; MENKVELD, 2011; HASBROUCK; SAAR, 2013; MENKVELD, 2013; BROGAARD; HENDERSHOTT; RIORDAN, 2014). Buscando identificar anomalias e padrões mais complexos, as estratégias quantitativas passaram a utilizar ML (PATEL; SHAH; THAKKAR; KOTECHA, 2015; ARNOTT; HARVEY; MARKOWITZ, 2019; HARVEY, 2021). Na última década, os avanços do *Deep Learning* (DL) (LECUN; BENGIO; HINTON, 2015) estimularam a sua exploração em finanças, especialmente no uso da *Long Short-Term Memory* (LSTM), em crescente uso na predição de séries financeiras (HOCHREITER; SCHMIDHUBER, 1997; KRAUS; FEUERRIEGEL, 2017; KIM; WON, 2018; FISCHER; KRAUSS, 2018).

Todavia, a configuração de algoritmos que reduzam significativamente o erro de predição sem aumentar demasiadamente o viés de aprendizado vem concentrando boa parte dos esforços de pesquisa nessa área. Um dos maiores desafios é evitar o *overfitting*, que ocorre quando o modelo se ajusta excessivamente aos dados de treinamento às custas de uma degradação no desempenho preditivo, perdendo a capacidade de generalização. Para reduzir esse risco, utilizam-se estratégias de regularização de parâmetros, como o *early stopping*, que consiste em interromper o treinamento a partir do momento (época) em que uma medida de desempenho, como o erro de predição, atinge um mínimo. O número de épocas é o número de vezes que o processo de aprendizado é realizado sobre o conjunto de dados de treinamento. Entretanto, não existe um valor pré-determinado para o número de épocas, restando ao pesquisador seguir uma estratégia e definir os valores iniciais baseado na experiência (PRECHELT, 1998a; PRECHELT, 1998b; GOODFELLOW et al., 2016; CHONG; HAN; PARK, 2017; AGGARWAL, 2018). Até o momento, não foram identificados trabalhos na literatura sobre o comportamento preditivo de uma rede LSTM em função do número de épocas, em especial para o retorno de ações no mercado brasileiro de ações. Nesse contexto, haveria um intervalo do número de épocas que minimizaria o erro de predição de uma rede LSTM sobre retornos de ações brasileiras?

O propósito deste estudo é identificar um intervalo do número de épocas no treinamento de uma rede neural recorrente LSTM, sobre séries de retornos de ações brasileiras, em que o erro de predição possa ser minimizado. Para tanto, utiliza-se a estratégia de regularização *early stopping*. O experimento analisa o comportamento

do erro de predição em função da variação do número de épocas. O erro de predição foi obtido pela Raiz quadrada do Erro Quadrático Médio (RMSE). Foi utilizada uma regressão quadrática para identificar um ponto de mínimo, porém os resultados não foram conclusivos quanto à existência de um intervalo de número de épocas que minimizasse o erro de predição.

Este trabalho contribui com a literatura sobre regularização de redes neurais recorrentes LSTM em finanças (PRECHELT, 1998a; GOODFELLOW et al., 2016; DENG et al., 2016; CHONG; HAN; PARK, 2017; BAEK; KIM, 2018; FISCHER; KRAUSS, 2018; GU, KELLY; XIU, 2021). Oferece também aos profissionais da gestão quantitativa de investimentos insights e recomendações para a configuração de parâmetros de modelos de DL direcionados ao mercado local, em especial para o tratamento do retorno de ações, ainda com poucos estudos publicados (AQR, 2015; KRAUSS, DO; HUCK, 2017; ISRAEL, KELLY; MOSKOWITZ, 2020; GU, KELLY; XIU, 2020; TOP1000FUNDS, 2020). Além disso, fornece aos investidores individuais material informativo sobre estratégias baseadas em ML (HARVARD, 2017; IEEE SPECTRUM, 2017; J.P. MORGAN, 2017; BLACKROCK, 2019; BARRON'S, 2020; BLOOMBERG, 2022; ITAÚ 2023).

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Mercados Eficientes e Mercados Adaptáveis

A Hipótese dos Mercados Eficientes (EMH) assume que não seria possível prever o preço de uma ação com base nos preços passados, pois estes já estariam refletidos no seu preço atual (*weak form*). A EMH utiliza o conceito de passeio aleatório (*random walk*) para argumentar que a evolução das preferências do investidor e o surgimento de novas informações se combinariam para construir situações de equilíbrio. Assim, os preços de mercado sinalizariam realocações de recursos que, por sua vez, se refletiriam nas distribuições dos retornos ao longo do tempo. Por esse raciocínio, os padrões de preços passados não poderiam ser utilizados simplesmente porque eles não seguiriam uma tendência (FAMA, 1965; FAMA, 1970).

Não obstante, Lo e MacKinlay (1988) apresentaram evidências empíricas de que o mercado de ações não obedeceria ao passeio aleatório, encontrando correlações seriais positivas significativas para retornos mensais e semanais. Por esta razão, os preços passados poderiam, até certo ponto, prever retornos futuros (JEGADEESH, 1990). Em linha com esse entendimento, estudos sobre finanças comportamentais identificaram vieses psicológicos que afetam as preferências e o comportamento racional esperado dos investidores (TVERSKY; KAHNEMAN, 1974; SHILLER, 1995; THALER, 1999; KAHNEMAN, 2011; KAHNEMAN; TVERSKY, 2013). Tais vieses, combinados às condições do ambiente econômico se refletiriam na formação de preços. Assim, sob uma perspectiva evolucionária, a Hipótese dos Mercados Adaptáveis (AMH) propõe que as estratégias de investimentos enfrentariam ciclos de ganhos e perdas em resposta às mudanças nas condições do mercado, à magnitude das oportunidades de lucro e à competição entre investidores sujeitos a vieses e aqueles com heurísticas eficazes (LO, 2004). Embora a EMH represente uma boa aproximação em mercados estacionários e economicamente previsíveis, a AMH oferece melhores explicações em ambientes mais dinâmicos e estocásticos, favorecendo o desenvolvimento de modelos de gestão e alocação eficiente de investimentos (LO, 2012).

Desde sua proposição, estudos empíricos sobre previsibilidade de retornos obtiveram resultados compatíveis com a AMH para várias classes de ativos e em diversos mercados (NEELY; WELLER; ULRICH, 2009; KIM, SHAMSUDDIN; LIM, 2011; CHARLES, DARNÉ; KIM, 2012; URQUHART; HUDSON, 2013; URQUHART; MCGROARTY, 2014; URQUHART; MCGROARTY, 2016). Estudos recentes têm confirmado a AMH em bolsas de países desenvolvidos (NODA, 2016; BOYA, 2019; TRIPATHI; DIXIT, 2020), em economias emergentes (XIONG et al., 2019; SOUZA, SILVA; LIMA, 2022), como também em mercados de criptoativos (KHUNTIA; PATTANAYAK, 2018; CHU; ZHANG; CHAN, 2019). Embora mais complexa, a implementação da AMH foi viabilizada não somente pela evolução da tecnologia de *trading* e de ferramentas estatísticas, mas também pelo desenvolvimento de algoritmos de execução. Tais algoritmos podem detectar padrões, regularidades e movimentos específicos em séries temporais financeiras, além de atenuar o componente de ruído pela filtragem de flutuações aleatórias (LO; MAMAYSKY; WANG, 2000; LO, 2012).

2.2. Capacidade preditiva da rede LSTM

No mercado de ações, o desenvolvimento de estratégias sistemáticas de negociação de ativos contribuiu para a eficiência dos preços e a melhoria da capacidade preditiva. Como consequência, prevaleceu sobre as estratégias puramente discricionárias, resultando em diminuição dos *spreads*, maior liquidez e redução da seleção adversa (HENDERSHOTT; JONES; MENKVELD, 2011; HASBROUCK; SAAR, 2013; BROGAARD; HENDERSHOTT; RIORDAN, 2014).

A predição de séries temporais financeiras, porém, exige especial atenção pela forte presença de ruído no sinal de dados (ZHANG; PATUWO; HU, 1998; DUNIS; WILLIAMS, 2003). O ruído resulta da atuação de inúmeros participantes do mercado, comprando e vendendo em diferentes momentos e para variados propósitos. Apresenta características das sequências randômicas mas pode resultar também de eventos inesperados, como notícias de impacto econômico, com alguma persistência no tempo (KAUFMAN, 2013).

Para lidar com as especificidades das séries financeiras, propõe-se a *Long Short-Term Memory* (LSTM) (HOCHREITER; SCHMIDHUBER, 1997). Por sua natureza recorrente, sobrepõe camadas neurais na dimensão temporal conseguindo extrair características das relações não lineares dos dados, memorizar séries temporais longas, reter seletivamente informações passadas relevantes, e descartar dados não relevantes, como ruídos. Os circuitos de realimentação favorecem a memorização extensa, adequada ao aprendizado por experiência, mesmo para intervalos de tempo arbitrários. Além disso, a capacidade de aprendizado se sustenta durante o treinamento, resultante de gradientes da função de perda estáveis pelo uso de funções sigmóides e tangentes hiperbólicas, portanto com menor risco de *vanishing* e *exploding gradients* (HOCHREITER; SCHMIDHUBER, 1997; XIONG, NICHOLS; SHEN, 2015; BAO; YUE; RAO, 2017; KRAUS; FEUERRIEGEL, 2017; FISCHER; KRAUSS, 2018; KIM; WON, 2018).

A LSTM se estrutura em unidades de chaveamento e armazenamento como: comporta de entrada (*input gate*), comporta de esquecimento (*forget gate*), comporta de saída (*output gate*) e célula de estado (*state cell*). A comporta de entrada i_t especifica qual informação será adicionada à célula de estado. A comporta de esquecimento f_t define qual parte da informação será removida. A de saída o_t escolhe qual informação da célula de estado será utilizada como saída. Estas três comportas utilizam a função sigmoide σ para emitir valores entre zero e um, quantificando o que

será passado adiante. A célula de estado C_t atua adicionando ou removendo informações com a ajuda das comportas, operada por uma função tangente hiperbólica \tanh (BAO; YUE; RAO, 2017; KRAUS; FEUERRIEGEL, 2017; SELVIN et al., 2017).

A representação matemática dos vetores das comportas e saídas é mostrada a seguir:

$$\text{Comporta de entrada:} \quad i_t = \sigma (W_i \cdot [h_{t-1}, x_t] + b_i) \quad (1)$$

$$\text{Comporta de esquecimento:} \quad f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

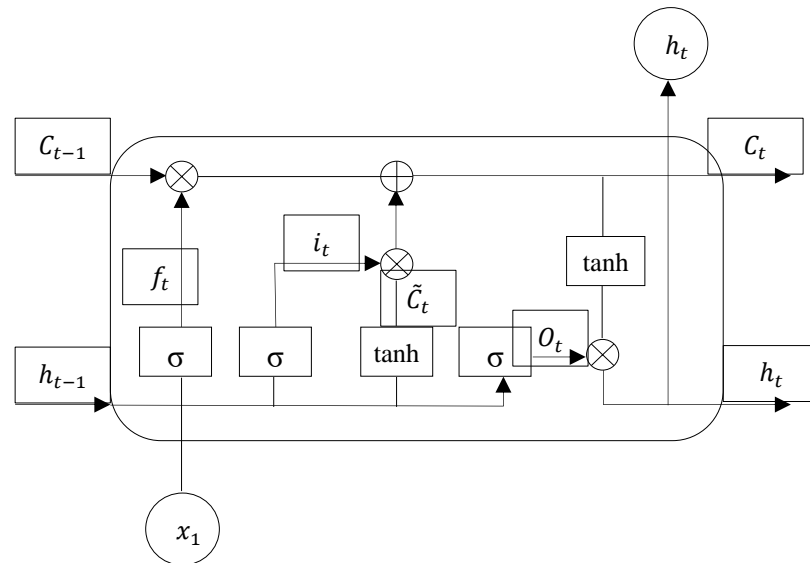
$$\text{Comporta de saída:} \quad o_t = \sigma (W_o \cdot [h_{t-1}, x_t] + b_o) \quad (3)$$

$$\text{Célula de estado:} \quad C_t = \tanh (W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

$$\text{Vetor de saída:} \quad h_t = o_t * \tanh (C_t) \quad (5)$$

onde i , f e o são os vetores das comportas de entrada, esquecimento e saída, C é o vetor da célula de estado, x e h os vetores de entrada e saída, W representa as matrizes de pesos, e b os vetores de erros (*bias*) para as respectivas comportas f , i e o . Os subscritos t e $t-1$ se referem aos instantes atual e anterior, σ é a função sigmoide, e \tanh a tangente hiperbólica, conforme a Figura 1 (BAO; YUE; RAO, 2017; KRAUS; FEUERRIEGEL, 2017; SELVIN et al., 2017).

Figura 1: Estrutura de uma unidade da LSTM



Fonte: Baseado em Bao, Yue e Rao (2017)

A partir dos anos 2000, a pesquisa em finanças quantitativas vem se desenvolvendo, particularmente em estratégias preditivas baseadas em *Machine Learning* (ML), como resultado dos avanços em capacidade computacional, redução nos custos de armazenamento de dados e oferta de ferramentas de software especializadas (ARNOTT; HARVEY; MARKOWITZ, 2019; HARVEY, 2021). Na década seguinte, o *Deep Learning* (DL), uma abordagem de ML baseada em redes neurais constituídas de um grande número de camadas internas, obteve avanços significativos em domínios como o reconhecimento de imagens e a interpretação da fala, ampliando as perspectivas de aplicação em outras áreas (LECUN; BENGIO; HINTON, 2015). Como métodos adaptativos orientados a dados, as redes de DL são

adequadas a situações em que há poucos pressupostos sobre o fenômeno em estudo (GOODFELLOW et al., 2016; XIONG; NICHOLS; SHEN, 2015). Ao longo da década de 2010, a LSTM apresentou resultados significativos em pesquisas sobre séries temporais e reconhecimento da fala (SUTSKEVER; VINYALS; LE, 2014; SCHMIDHUBER, 2015). Nos últimos cinco anos, tem sido frequentemente utilizada na previsão de séries financeiras (DENG et al., 2016; KRAUS; FEUERRIEGEL, 2017; KIM; WON, 2018; FISCHER; KRAUSS, 2018; SEZER; OZBAYOGLU, 2018; HOSEINZADE; HARATIZADEH, 2019; LONG; LU; CUI, 2019; CAO et al., 2020).

2.3. Parametrização, overfitting, regularização e *early stopping*

Apesar da adequação das redes LSTM a estudos dessa natureza, a concepção de modelos de redes neurais não é trivial. Como modelos de reconhecimento de padrões, dependem de uma bem elaborada representação de dados e envolvem preocupações com parametrização, generalização, desempenho e *overfitting* (ZHANG; PATUWO; HU, 1998; DUNIS; WILLIAMS, 2003; AMINIMEHR et al., 2022).

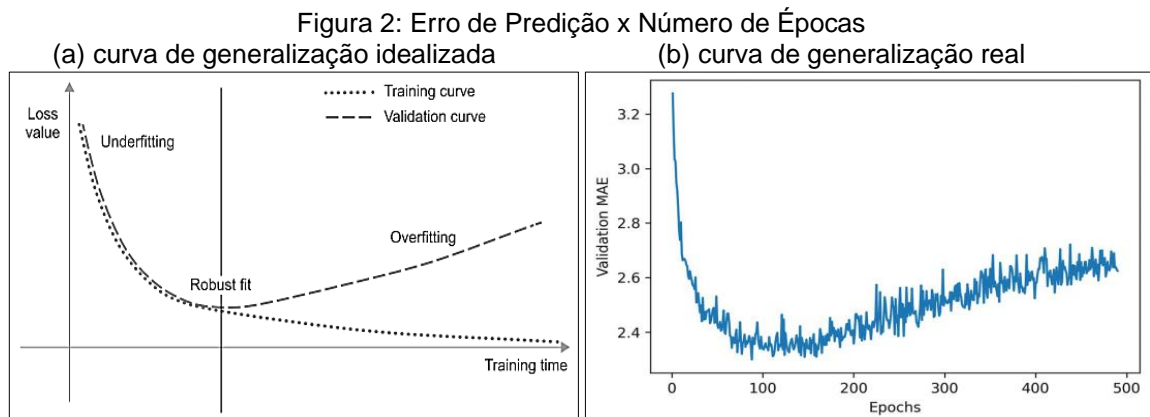
A configuração de parâmetros em abordagens de DL representa um grande desafio para os pesquisadores. Embora haja um razoável consenso nos estudos sobre abordagens de ML, não se pode dizer o mesmo quanto às redes neurais profundas. Modelos de DL podem ter de dez a cinquenta hiperparâmetros, tornando complexa a minimização da função de perda no espaço de configurações. Diante disso, estudos propondo métodos de otimização de parâmetros exerceram um papel relevante no desenvolvimento de algoritmos e novas modelagens (BERGSTRA et al., 2011; GOODFELLOW et al., 2016).

O *overfitting* ocorre quando um modelo se ajusta excessivamente aos dados de treinamento, mas perde capacidade preditiva diante de dados novos. Esse fenômeno permanece como um problema a ser resolvido, pois grandes redes neurais apresentam processamento lento e de difícil tratamento. Isso se deve à combinação ou co-adaptação das previsões no ajuste de pesos das unidades neurais, tendo como consequência a perda de generalização (SRIVASTAVA et al., 2014; LÓPEZ DE PRADO, 2018).

Os métodos de redução do *overfitting* também constituem formas de configuração e otimização de parâmetros. As estratégias de regularização, por exemplo, buscam reduzir a complexidade de uma rede neural estabelecendo limites no espaço de variação de parâmetros, ainda que às custas de um possível aumento do erro de previsão. Por isso, significativos esforços de pesquisa têm se concentrado no desenvolvimento de estratégias mais eficazes (GOODFELLOW et al., 2016; AGGARWAL, 2018; SIRIGNANO, 2019). O *dropout*, uma técnica que previne o excesso de co-adaptação pela desconexão aleatória de unidades neurais da rede durante o treinamento, tem sido bem-sucedido na prevenção do *overfitting* em redes neurais *feed-forward* e *Boltzman machines* (SRIVASTAVA et al., 2014).

O *early stopping*, uma técnica alternativa, consiste na interrupção do processo de treinamento quando a variância do erro de previsão atinge um nível mínimo ao se variar o número de épocas. É frequentemente utilizado por sua facilidade de implementação e resultados superiores em diversas abordagens. Ao se interromper o ajuste a partir do momento (época) em que a capacidade de generalização se deteriora ou se estabiliza, evita-se o ajuste excessivo dos pesos da rede neural, diminuindo a chance de ocorrência de *overfitting*. Embora os primeiros estudos mostrem curvas de generalização idealizadas, como na Figura 2(a), tais curvas podem assumir configurações mais complexas, podendo apresentar mais de um mínimo local e comportamento errático ou intermitente, Figura 2(b) (PRECHELT, 1998a;

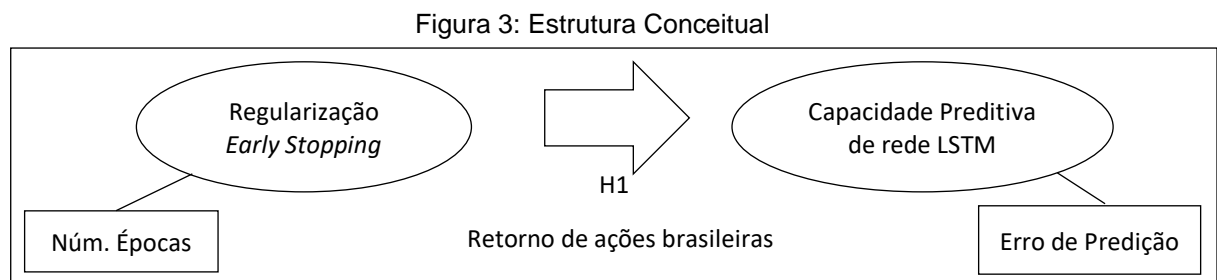
GOODFELLOW et al., 2016; AGGARWAL, 2018; FISCHER; KRAUSS, 2018, CHOLLET, 2021).



Fonte: Chollet (2021)

Neste estudo, é testada a hipótese da existência de um intervalo de épocas que minimiza o erro de predição de uma rede LSTM sobre retornos de ações brasileiras. Na relação entre constructos, a regularização por *early stopping* é representada pelo número de épocas de treinamento, e a capacidade preditiva é mensurada pelo erro de predição (acurácia) RMSE (MAKRIDAKIS, 1993; AMINIMEHR et al., 2022) (Figura 3).

H1. *Existe um intervalo de épocas que minimiza o erro de predição de uma rede LSTM sobre retornos de ações brasileiras.*



Fonte: Elaborada pelos autores

3. METODOLOGIA

Este trabalho utilizou uma abordagem quantitativa para verificar a existência de um intervalo de épocas que minimize o erro de predição de uma rede neural recorrente LSTM.

Os dados secundários são os preços de fechamento de ações negociadas na B3 (B3 PREÇOS HISTÓRICOS, 2023). O recorte foi longitudinal, compreendendo o período de 2008 a 2018, constituindo uma série temporal diária (dias em que ocorreram negociações) de 10 anos, resultando em amostra com 2718 pontos de dados para cada ativo. O tamanho da amostra foi maior do que alguns estudos de referência (HEATON; POLSON; WITTE, 2017; BAO; YUE; RAO, 2017; KIM; WON, 2018), porém menor do que o de Fischer e Krauss (2018). Os preços foram transformados em retornos aritméticos simples e agrupados em sequências de 10

pontos de dados e a base foi segmentada em: 80% para o treinamento, 10% para a validação, e 10% para os testes (FISCHER; KRAUSS, 2018; XIONG; NICHOLS; SHEN, 2015; KRAUS; FEUERRIEGEL, 2017).

O mercado de ações brasileiro foi escolhido por sua diferenciação em relação aos mercados desenvolvidos. Mercados emergentes requerem uma análise específica, seja pela menor integração aos mercados globais, como pelas diferenças entre a participação do produto interno e a capitalização de mercado, como também pelo maior nível de risco comparado ao dos desenvolvidos, com retornos esperados mais altos (BEKAERT; HARVEY, 2017). A escolha das ações foi não probabilística, seguindo critérios de presença no índice de referência Ibovespa (B3 IBOVESPA, 2023), portanto dando preferência para a liquidez e capitalização de mercado, bem como pela continuidade do *ticker* (código da ação) ao longo do período de recorte, evitando ativos com descontinuidades decorrentes de privatizações (PINHEIRO, 1999) e fusões e aquisições (NAKANE; WEINTRAUB, 2005; BARROS; WANKE, 2014).

A capacidade preditiva da rede LSTM, o constructo dependente, foi mensurada pelo erro de predição RMSE (acurácia) (MAKRIDAKIS, 1993; HEATON; POLSON; WITTE, 2017; KRAUS; FEUERRIEGEL, 2017; KIM; WON, 2018; AMINIMEHR et al., 2022). A regularização por *early stopping*, o constructo independente, foi expressa pelo número de épocas de treinamento (PRECHELT, 1998a; PRECHELT, 1998b; GOODFELLOW et al., 2016; AGGARWAL, 2018).

Os valores da acurácia em função do número de épocas foram plotados em gráfico para observação e análise de seu comportamento. A forma funcional quadrática foi utilizada para mensurar as variações marginais e identificar a existência de ponto de inflexão (WOOLDRIDGE, 2019), conforme a equação:

$$A_i = \alpha + \beta_1 E_i + \beta_2 E_i^2 + \varepsilon_i \quad (6)$$

Onde:

A_i – acurácia da i -ésima ação

α – intercepto da função quadrática

β_1 – coeficiente do termo linear

β_2 – coeficiente do termo quadrático

E_i – número de épocas da i -ésima ação

ε_i – termo de erro que concentra a parcela de acurácia não explicada pela função

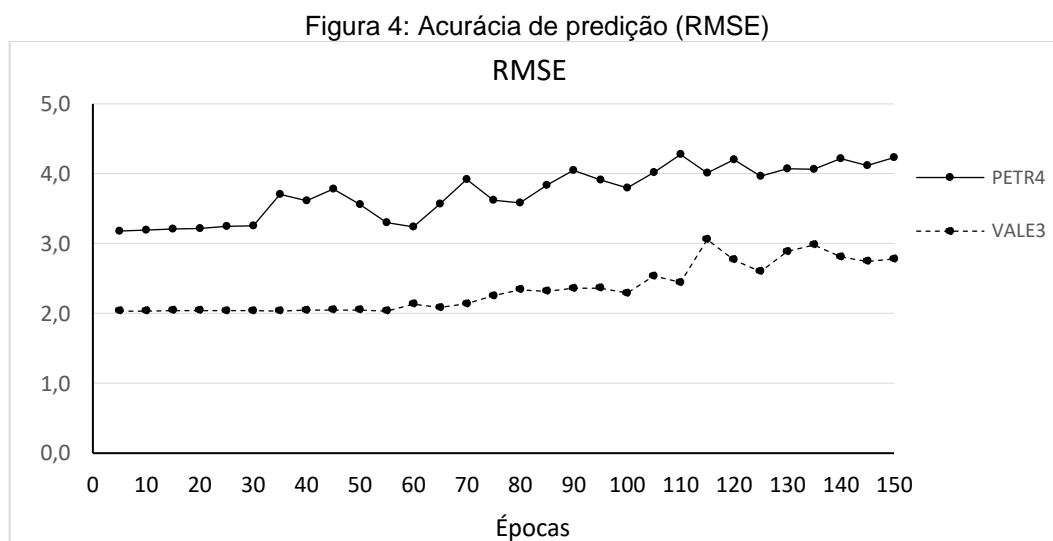
Para que o número de épocas apresente um valor mínimo, o coeficiente β_1 deve ser negativo, o coeficiente β_2 positivo, e ambos devem ser estatisticamente significativos (WOOLDRIDGE, 2019; SILVEIRA; BARROS; FAMÁ, 2003). A validade dos resultados é mensurada pela própria acurácia, pois os retornos obtidos na predição são comparados com os retornos que ocorreram de fato para cada ativo.

Para a análise do modelo preditivo, foi elaborado um artefato de software em linguagem Python, utilizando as bibliotecas computacionais: Numpy, Pandas, Scikit-Learn e TensorFlow (GÉRON, 2019; CHOLLET, 2021; PEDREGOSA et al., 2011). A rede é composta por duas camadas, cada uma com 50 unidades neuronais. O otimizador de gradiente ADAM (*Adaptative Moment Estimation*) foi escolhido pela facilidade de implementação, eficiência computacional e pela adequação a dados não estacionários, gradientes esparsos e baixa relação sinal/ruído. A taxa de aprendizado foi fixada em 0,001 e tamanho do lote de 50 (KINGMA; BA, 2014; RUDER, 2016). Estes parâmetros foram mantidos fixos para a observação do comportamento do erro

de predição em função do número de épocas. O código-fonte está publicamente disponível na plataforma GitHub no endereço: <https://github.com/Camilollzo/Stock-returns>, no arquivo “modLSTM.py”.

4. RESULTADOS

Foram utilizadas as duas ações de maior peso no índice Ibovespa a VALE3 e a PETR4. O modelo LSTM gerou predições de um dia e o comportamento da acurácia de predição, medida pelo RMSE, foi analisado em função do número de épocas. As épocas foram ajustadas para variarem de 5 até 150, com incrementos de 5, resultando em 30 pontos de dados para cada ação (Figura 4).



Fonte: Elaborada pelos autores

Para identificar um ponto de mínimo, foi realizada a regressão por função quadrática conforme equação (6). Para PETR4, o termo quadrático β_2 não resultou significativo e o termo linear β_1 foi significativamente positivo, indicando não existir um intervalo de épocas com valor mínimo do erro de predição, restando porém uma relação positiva com o número de épocas. Para VALE3, o termo quadrático resultou positivo e significativo, e o termo linear também positivo, não caracterizou uma inflexão do erro de predição. Assim, não foi possível identificar um intervalo de número de épocas que minimizaria o erro de predição tanto para PETR4 como para VALE3 (Tabela 1).

Tabela 1: Acurácia de Predição como função quadrática do Número de Épocas

	PETR4			VALE3		
	Acurácia	t	P-value	Acurácia	t	P-value
Termo linear (β_1)	0.00962 (0.00289)	3.33	0.003	0.00024 (0.00227)	0.11	0.916
Termo quadrático (β_2)	-0.000014 (0.00002)	-0.76	0.453	0.00004 (0.00001)	3.00	0.006

Fonte: Elaborada pelos autores

A regressão linear confirmou a relação positiva e significativa, tanto estatística como econômica, entre o número de épocas e o erro de predição, para ambas as ações (Tabela 2).

Tabela 2: Acurácia de Predição como função linear do Número de Épocas

	PETR4			VALE3		
	Acurácia	<i>t</i>	<i>P-value</i>	Acurácia	<i>t</i>	<i>P-value</i>
Termo linear (β_1)	0.00748 (0.00069)	10.77	0.000	0.00685 (0.00062)	11.01	0.000

Fonte: Elaborada pelos autores

5. DISCUSSÃO

A ausência um ponto inflexão do erro de predição para as ações analisadas indica que não haveria um intervalo do número de épocas que minimizaria o erro de predição, refutando a hipótese deste estudo. Além disso, a relação positiva significativa da regressão linear sugere que o aumento do número de épocas implica em perda de capacidade preditiva para o intervalo estudado. Portanto, a estratégia de *early stopping* não se aplicaria para a regularização da rede LSTM sobre retornos de ações, ao menos para os ativos deste estudo.

Os resultados não confirmam o comportamento comumente descrito na literatura (GOODFELLOW et al., 2016; AGGARWAL, 2018), embora muitos estudos admitam que tal padrão possa variar em função do modelo, apresentando até mesmo vários mínimos locais. Parte destes estudos se concentraram em redes *feedforward*, e não em redes recorrentes, como a LSTM. Ainda assim, reconhecem que os pontos de mínimo, se existirem, poderiam não apresentar resultados significativos (BISHOP, 1995; SJÖBERG; LJUNG, 1995; PRECHELT, 1998a; PRECHELT, 1998b).

Vale destacar que estudos em finanças que utilizaram preços, em vez de retornos, obtiveram resultados preditivos mais expressivos, portanto mais propensos a resultados que confirmam a literatura (YOSHIHARA; FUJIKAWA, SEKI; UEHARA, 2014; DING; ZHANG, LIU; DUAN, 2015; PATEL; SHAH, THAKKAR; KOTECHA, 2015). Por outro lado, a transformação das séries de preços em retornos, ao atenuar os componentes não estacionários, reduziria o desempenho de preditores baseados em tendências (CHEN; PELGER; ZHU, 2023), o que por sua vez explicaria os resultados obtidos.

Até meados da década passada, não havia estudos aplicando DL a retornos de ações (CHONG; HAN; PARK, 2017), e desde então poucos trabalhos adotaram esse critério (FISCHER; KRAUSS, 2018; AMINIMEHR et al., 2022; CHEN; PELGER; ZHU, 2023). Assim, este estudo contribui para a literatura ao analisar o comportamento preditivo de um modelo de DL, a rede neural recorrente LSTM, no contexto de uma economia emergente, como a brasileira, sob um processo de regularização por *early stopping*, evidenciando um comportamento não esperado do erro de predição do retorno de ações em função do número de épocas (PRECHELT, 1998a; GOODFELLOW et al., 2016; DENG et al., 2016; CHONG; HAN; PARK, 2017; BAEK; KIM, 2018; FISCHER; KRAUSS, 2018; GU, KELLY; XIU, 2021). Sob tais condições, recomenda-se um ajuste criterioso no número de épocas para evitar o *overfitting*, particularmente quando não for possível identificar regiões de mínimos na curva do erro de predição.

Para os gestores quantitativos, o estudo contribui com a modelagem de estratégias de negociação que utilizem o modelo pesquisado, auxiliando a escolha de métodos de regularização que melhorem o desempenho preditivo e reduzam o risco de *overfitting*, tendo em vista a escassez de trabalhos sobre o comportamento do retorno de ações. Os gestores podem também aproveitar as curvas de generalização deste trabalho para comparação com outros modelos sobre capacidade preditiva (AQR, 2015; KRAUSS; DO; HUCK, 2017; ISRAEL; KELLY; MOSKOWITZ, 2020; GU; KELLY; XIU, 2020; TOP1000FUNDS, 2020). Os métodos utilizados podem ser

também aplicados à predição da volatilidade para a precificação de derivativos de ações e outros ativos financeiros (XIONG; NICHOLS; SHEN, 2015; KIM; WON, 2018).

Aos investidores individuais, oferece conteúdo informativo sobre estratégias de investimentos baseadas em ML, apoiando a tomada de decisão na escolha de fundos quantitativos (HARVARD, 2017; IEEE SPECTRUM, 2017; J.P. MORGAN, 2017; BLACKROCK, 2019; BARRON'S, 2020; BLOOMBERG, 2022; ITAÚ 2023).

6. CONCLUSÃO

Este estudo questionou a existência um intervalo do número de épocas que minimizaria o erro de predição de uma rede LSTM sobre retornos de ações brasileiras. Os resultados não confirmaram a existência desse intervalo, mas identificaram uma relação positiva significativa entre o aumento do número de épocas e a perda de capacidade preditiva.

As séries temporais se limitaram a dois ativos, não obstante serem os mais líquidos e dentre os de maior capitalização de mercado. Ainda assim, tal amostra tem sua representatividade limitada perante a referência de mercado (B3 IBOVSPA, 2023). O recorte longitudinal, embora compatível com outros trabalhos na literatura, poderia ser ampliado para um período de quinze ou vinte anos visando predições mais robustas, mantidos os cuidados com ativos que atendam aos requisitos de continuidade. Poderia ainda se buscar uma maior granularidade dos dados, com a totalidade das cotações de preços ao longo dos pregões (*tick-by-tick*), com cotações de compra e venda (*bid-ask*) e as filas de cotações dos livros de ofertas limitadas (*limited order book*), entretanto, com o devido ajuste da capacidade de processamento em razão do volume adicional de dados.

Futuras pesquisas podem se beneficiar de uma escolha mais ampla de ações, de diversos setores e com diferentes níveis de capitalização de mercado, no intuito de melhor representar o mercado local e generalização dos resultados. Uma maior granularidade dos dados pode oferecer novos *insights* sobre a capacidade preditiva. A regularização de parâmetros como taxa de aprendizado, tamanho do lote, tamanho da sequência, número de camadas e de unidades neurais, entre outros, são também condicionantes a serem exploradas. Em se tratando de área do conhecimento com amplas perspectivas de evolução, sugerem-se estudos preditivos combinando diferentes abordagens de ML e comparações com abordagens estatísticas de séries temporais e com mercados desenvolvidos.

REFERÊNCIAS

- AGGARWAL, C. C. Neural networks and deep learning. Springer, 10, 978-3. 2018.
- AMINIMEHR, A.; RAOOFI, A.; AMINIMEHR, A.; AMINIMEHR, A. A Comprehensive Study of Market Prediction from Efficient Market Hypothesis up to Late Intelligent Market Prediction Approaches. Computational Economics, 60(2), 781-815. 2022.
- AQR. How Can a Strategy Everyone Knows About Still Work? 2015. Disponível em <www.aqr.com/Insights/Perspectives/How-Can-a-Strategy-Still-Work-If-Everyone-Knows-About-It>. Acesso em 08 jun 2023.
- ARNOTT, R.; HARVEY, C. R.; MARKOWITZ, H. A backtesting protocol in the era of machine learning. The Journal of Financial Data Science, 1(1), 64-74. 2019.
- B3 IBOVESPA. 2023. Disponível em <www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm> Acesso em 06 mai 2023.
- B3 PREÇOS HISTÓRICOS. 2023. Disponível em <http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/historico/mercado-a-vista/series-historicas/> Acesso em 06 mai. 2023.
- BAEK, Y.; KIM, H. Y. ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. Expert Systems with Applications, 113, 457-480. 2018.
- BAO, W.; YUE, J.; RAO, Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PloS one, 12(7). 2017.
- BARRON'S. 2020. How Artificial Intelligence Could Lead to Better Investment Decisions. Disponível em <<https://www.barrons.com/articles/how-artificial-intelligence-could-lead-to-better-investment-decisions-51594048894>>. Acesso em 10 jun 2023.
- BARROS, C. P.; WANKE, P. Banking efficiency in Brazil. Journal of International Financial Markets, Institutions and Money, 28, 54-65. 2014.
- BEKAERT, G.; HARVEY, C. R. Emerging equity markets in a globalizing world. Available at SSRN 2344817. 2017.
- BERGSTRA, J.; BARDENET, R.; BENGIO, Y.; KÉGL, B. Algorithms for hyper-parameter optimization. Advances in Neural Information Processing Systems, 24. 2011.
- BISHOP, C. M. Regularization and complexity control in feed-forward networks. 1995.
- BLACKROCK. Artificial intelligence and machine learning in asset management. 2019. Disponível em <www.blackrock.com/corporate/literature/whitepaper/viewpoint-artificial-intelligence-machine-learning-asset-management-october-2019.pdf>
- BLOOMBERG. Quant Hedge Funds Post Historic Returns in Ugly Year for Wall Street. 2022. Disponível em <<https://www.bloomberg.com/news/articles/2022-12-16/quant-traders-are-big-winners-in-this-year-s-market-turmoil#xj4y7vzkg>> Acesso em 10 jun 2023.
- NODA, C. M. From efficient markets to adaptive markets: Evidence from the French stock exchange. Research in International Business and Finance, 49, 156-165. 2019.
- BROGAARD, J.; HENDERSHOTT, T.; RIORDAN, R. High-frequency trading and price discovery. The Review of Financial Studies, 27(8), 2267-2306. 2014.
- CAO, B.; ZHAO, J.; LV, Z.; GU, Y.; YANG, P.; HALGAMUGE, S. K. Multiobjective evolution of fuzzy rough neural network via distributed parallelism for stock prediction. IEEE Transactions on Fuzzy Systems, 28(5), 939-952. 2020.

CHARLES, A.; DARNÉ, O.; KIM, J. H. Exchange-rate return predictability and the adaptive markets hypothesis: Evidence from major foreign exchange rates. *Journal of International Money and Finance*, 31(6), 1607-1626. 2012.

CHEN, L.; PELGER, M.; ZHU, J. Deep learning in asset pricing. *Management Science*. 2023.

CHOLLET, F. *Deep learning with Python*. Simon and Schuster. 2021.

CHONG, E.; HAN, C.; PARK, F. C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 83, 187-205. 2017.

CHU, J.; ZHANG, Y.; CHAN, S. The adaptive market hypothesis in the high frequency cryptocurrency market. *International Review of Financial Analysis*, 64, 221-231. 2019.

DENG, Y.; BAO, F.; KONG, Y.; REN, Z.; DAI, Q. Deep direct reinforcement learning for financial signal representation and trading. *IEEE transactions on neural networks and learning systems*, 28(3), 653-664. 2016.

DING, X.; ZHANG, Y.; LIU, T.; DUAN, J. Deep learning for event-driven stock prediction. In *Twenty-fourth international joint conference on artificial intelligence*. 2015.

DUNIS, C. L.; WILLIAMS, M. Applications of advanced regression analysis for trading and investment. *Applied quantitative methods for trading and investment*, 1-40. 2003.

FAMA, E. F. Random Walks in Stock Market Prices. *Financial Analysts Journal*, 21(5), 55-59. 1965.

FAMA, E. F. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417. 1970.

FISCHER, T.; KRAUSS, C. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654-669. 2018.

GÉRON, A. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, Inc. 2019.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. MIT press. 2016

GU, S.; KELLY, B.; XIU, D. Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223-2273. 2020

GU, S.; KELLY, B.; XIU, D. Autoencoder asset pricing models. *Journal of Econometrics*, 222(1), 429-450. 2021

HARVARD. An economy of algorithms. 2017. Disponível em <https://seas.harvard.edu/news/2017/01/economy-algorithms> Acesso em 10 jun 2023.

HARVEY, C. R. Why Is Systematic Investing Important? *Journal of Systematic Investing*. EQDerivatives, Inc. v I, issue 1. 2021.

HASBROUCK, J.; SAAR, G. Low-latency trading. *Journal of Financial Markets*, 16(4), 646-679. 2013.

HEATON, J. B.; POLSON, N. G.; WITTE, J. H. Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry*, 33(1), 3-12. 2017.

HENDERSHOTT, T.; JONES, C. M.; MENKVELD, A. J. Does algorithmic trading improve liquidity? *The Journal of finance*, 66(1), 1-33. 2011.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural Computation*, 9(8), 1735-1780. 1997.

HOSEINZADE, E.; HARATIZADEH, S. CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, 129, 273-285. 2019.

IEEE SPECTRUM. Hedge Funds Look to Machine Learning, Crowdsourcing for Competitive Advantage. 2017. Disponível em <spectrum.ieee.org/fintech-trends-hedge-funds-look-to-machine-learning-crowdsourcing-for-competitive-advantage#toggle-gdpr>. Acesso em 10 jun 2023.

ISRAEL, R.; KELLY, B. T.; MOSKOWITZ, T. J. Can Machines' Learn' Finance? *Journal of Investment Management*. 2020.

ITAU. The Weekly Globe. Gestão de portfólio e Inteligência Artificial. 2023. Disponível em <<https://blog.itau.com.br/privateinsights/gestao-de-portfolio-e-inteligencia-artificial/blt889dbaad1b2e40dc>>. Acesso em 08 jun 2023.

JEGADEESH, N. Evidence of predictable behavior of security returns. *The Journal of finance*, 45(3), 881-898. 1990.

J.P. MORGAN. Innovations in Finance with Machine Learning, Big Data and Artificial Intelligence. 2017. Disponível em <www.jpmorgan.com/insights/research/machine-learning>. Acesso em 10 jun 2023.

KAHNEMAN, D. Thinking, fast and slow. Macmillan. 2011.

KAHNEMAN, D.; TVERSKY, A. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I* (pp. 99-127). 2013.

KAUFMAN, P. J. Trading Systems and Methods,+ Website (Vol. 591). Wiley Sons. 2013.

KHUNTIA, S.; PATTANAYAK, J. K. Adaptive market hypothesis and evolving predictability of bitcoin. *Economics Letters*, 167, 26-28. 2018.

KIM, J. H.; SHAMSUDDIN, A.; LIM, K. P. Stock return predictability and the adaptive markets hypothesis: Evidence from century-long US data. *Journal of Empirical Finance*, 18(5), 868-879. 2011.

KIM, H. Y.; WON, C. H. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, 103, 25-37. 2018.

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

KRAUS, M.; FEUERRIEGEL, S. Decision support from financial disclosures with deep neural networks and transfer learning. *Decision Support Systems*, 104, 38-48. 2017.

KRAUSS, C.; DO, X. A.; HUCK, N. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, 259(2), 689-702. 2017.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, 521(7553), 436-444. 2015.

LO, A. W.; MACKINLAY, A. C. Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies*, 1(1), 41-66. 1988.

LO, A. W.; MAMAYSKY, H.; WANG, J. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *The journal of finance*, 55(4), 1705-1765. 2000.

LO, A. W. The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*, Forthcoming. 2004.

LO, A. W. Adaptive markets and the new world order (corrected May 2012). *Financial Analysts Journal*, 68(2), 18-29. 2012.

LONG, W.; LU, Z.; CUI, L. Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, 164, 163-173. 2019.

LÓPEZ DE PRADO, M. L. *Advances in financial machine learning*. John Wiley Sons. 2018.

MAKRIDAKIS, S. Accuracy measures: theoretical and practical concerns. *International journal of forecasting*, 9(4), 527-529. 1993.

MENKVELD, A. J. High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4), 712-740. 2013.

NAKANE, M. I.; WEINTRAUB, D. B. Bank privatization and productivity: Evidence for Brazil. *Journal of Banking Finance*, 29(8-9), 2259-2289. 2005.

NEELY, C. J.; WELLER, P. A.; ULRICH, J. M. The adaptive markets hypothesis: evidence from the foreign exchange market. *Jrnl of Financial Quantitative Analysis*, 44(2), 467-488. 2009.

NODA, A. A test of the adaptive market hypothesis using a time-varying AR model in Japan. *Finance Research Letters*, 17, 66-71. 2016.

PATEL, J.; SHAH, S.; THAKKAR, P.; KOTECHEA, K. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert systems with applications*, 42(1), 259-268. 2015.

PEDREGOSA, F. et al. *Scikit-learn: Machine learning in Python*. *The Journal of Machine Learning Research*, 12, 2825-2830. 2011.

PINHEIRO, A. C. Privatização no Brasil: por quê? Até onde? Até quando? A economia brasileira nos anos 90. Capítulo 3, p. 180-182. Rio de Janeiro. Banco Nacional de Desenvolvimento Econômico e Social. 1999.

PRECHELT, L. Automatic early stopping using cross validation: quantifying the criteria. *Neural Networks*, 11(4), 761-767. 1998a.

PRECHELT, L. Early stopping – but when? In *Neural Networks: Tricks of the trade* (pp. 55-69). Springer, Berlin, Heidelberg. 1998b.

RUDER, S. An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747. 2016.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117. 2015.

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

SEZER, O. B.; OZBAYOGLU, A. M. Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. *Applied Soft Computing*, 70, 525-538. 2018.

SILVEIRA, A. D. M. D.; BARROS, L. A. B. D. C.; FAMÁ, R. Estrutura de governança e valor das companhias abertas brasileiras. *Revista de Administração de Empresas*, 43, 50-64. 2003.

SIRIGNANO, J. Deep learning for limit order books. *Quantitative Finance*, 19, 549-570. 2019.

SJÖBERG, J.; LJUNG, L. Overtraining, regularization and searching for a minimum, with application to neural networks. *International Journal of Control*, 62(6), 1391-1407. 1995.

SHILLER, R. J. Conversation, information, and herd behavior. *The American Economic Review*, 85(2), 181-185. 1995.

SOUZA, P. V. S.; SILVA, C. A. T.; LIMA, F. G. Evidence of the adaptive market hypothesis in shares traded by B3 listed banking companies. *Managerial Finance*, 48(1), 113-125. 2022.

SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958. 2014.

SUTSKEVER, I.; VINYALS, O.; LE, Q. V. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27. 2014.

THALER, R. H. Mental accounting matters. *Journal of Behavioral Decision Making*, 12(3), 183-206. 1999.

TOP1000FUNDS The future is quant. 2020. Disponível em <www.top1000funds.com/2020/02/the-future-is-quant/> Acesso em 10 jun 2023.

TRIPATHI, A.; DIXIT, A. Adaptive market hypothesis and investor sentiments: global evidence. *Managerial Finance*, 46(11), 1407-1436. 2020.

TVERSKY, A.; KAHNEMAN, D. Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124-1131. 1974.

URQUHART, A.; HUDSON, R. Efficient or adaptive markets? Evidence from major stock markets using very long run historic data. *International Review of Financial Analysis*, 28, 130-142. 2013.

URQUHART, A.; MCGROARTY, F. Calendar effects, market conditions and the Adaptive Market Hypothesis: Evidence from long-run US data. *International Review of Financial Analysis*, 35, 154-166. 2014.

URQUHART, A.; MCGROARTY, F. Are stock markets really efficient? Evidence of the adaptive market hypothesis. *International Review of Financial Analysis*, 47, 39-49. 2016.

WOOLDRIDGE, J. M. *Introductory Econometrics: A Modern Approach*. 7th ed. Boston, MA. Cengage Learning. 2019.

XIONG, X.; MENG, Y.; LI, X.; SHEN, D. An empirical analysis of the Adaptive Market Hypothesis with calendar effects: Evidence from China. *Finance Research Letters*, 31. 2019.

XIONG, R.; NICHOLS, E. P.; SHEN, Y. Deep Learning Stock Volatility with Google domestic trends. arXiv preprint arXiv:1512.04916. 2015.

YOSHIHARA, A.; FUJIKAWA, K.; SEKI, K.; UEHARA, K. Predicting stock market trends by recurrent deep neural networks. In *PRICAI 2014: Trends in Artificial Intelligence: 13th Pacific Rim International Conference on Artificial Intelligence*, Gold Coast, QLD, Australia, December 1-5, 2014. Proceedings 13 (pp. 759-769). Springer International Publishing. 2014.

ZHANG, G.; PATUWO, B. E.; HU, M. Y. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14(1), 35-62. 1998.