



ANÁLISE DO EFEITO *LEAD-LAG* ENTRE BOLSAS DE VALORES DO JAPÃO E DO BRASIL UTILIZANDO REDES NEURAIAS ARTIFICIAIS E INDICADORES TÉCNICOS

ANALYSIS OF LEAD-LAG EFFECT BETWEEN JAPANESE AND BRAZILIAN STOCK MARKET USING ARTIFICIAL NEURAL NETWORK AND TECHNICAL INDICATORS

6 Finanças

Douglas Rorie Tanno, UEM, Brasil, ra104156@uem.br

Wagner Igarashi, UEM, Brasil, wigarashi@uem.br

Deisy Cristina Corrêa Igarashi, UEM, Brasil, dccigarashi@uem.br

Resumo

O mercado de ações tem atraído cada vez mais os investidores brasileiros, necessitando de estratégias de investimento cada vez mais sofisticadas, buscando maximizar o lucro obtido nos investimentos. Com essa ideia, este trabalho propõe avaliar um possível efeito *lead-lag* entre bolsas do Japão e do Brasil, utilizando simuladores com base em redes neurais artificiais e análise técnica, buscando prever as ações brasileiras para a obtenção de lucro. Para isso, foram escolhidas a SoftBank Corp. e a Telefônica Brasil, principais empresas do ramo de Telecomunicações das principais bolsas de ambos os países. Os experimentos demonstraram que modelos de investimento utilizando os indicadores técnicos de ambas as empresas, em conjunto com as redes neurais artificiais, podem indicar um efeito *lead-lag* entre as bolsas.

Palavras-chave: Efeito *lead-lag*; Redes neurais artificiais; Análise técnica; Mercado de ações.

Abstract

The stock market has increasingly attracted Brazilian investors, requiring increasingly sophisticated investment strategies, seeking to maximize the profit obtained from investments. With this idea in mind, this work aims to evaluate a possible lead-lag effect between stock exchanges in Japan and Brazil, using simulators based on artificial neural networks and technical analysis, trying to predict Brazilian stocks to obtain profit. For this, SoftBank Corp. and Telefônica Brasil, the main companies in the Telecommunications sector on the main exchanges in both countries, were chosen. The experiments showed that the investment models using the technical indicators of both companies, with the use of artificial neural networks, can indicate a lead-lag effect between the exchanges.

Keywords: Lead-lag effect; Artificial neural networks; Technical analysis; Stock market.

1. INTRODUÇÃO

O mercado financeiro continua crescendo e se popularizando, tornando cada vez mais comum que investidores busquem o suporte de sistemas especializados que auxiliem nas tomadas de decisões sobre investimentos, identificando as empresas e os períodos mais favoráveis para que haja maior retorno e segurança do investimento.

Para a análise de comportamento do mercado e a realização de previsões, podem ser utilizadas duas classes de análise: a fundamentalista e a técnica. Na análise fundamentalista, são estudadas informações diversas, podendo ser subjetivas e dificilmente tratadas computacionalmente, como o contexto das empresas e do mercado, relatórios, balanços, noticiários, entre outros, sendo mais apropriada para uma análise de médio a longo prazo. Na análise técnica, são utilizados os dados de preço e volume de transações, sendo mais relacionada à oferta e demanda de curto prazo. Contudo, é mais fácil automatizar a análise técnica.

Outro aspecto relevante a ser discutido, é que o período de atividade diário das bolsas de valores no mundo pode variar de acordo com sua localização geográfica. De acordo com o WorldTimeZone (2021), as primeiras bolsas a abrir e fechar diariamente são as dos países mais próximos à Linha Internacional de Data, iniciando com a bolsa da Nova Zelândia, seguido de Sidney, Tóquio, Hong Kong, Cingapura, Mumbai e Moscou. Devido a esta defasagem espacial e temporal, acredita-se que uma pequena parte das ações seja um indicador (*leading*) antecedente do mercado, ao se considerar que a flutuação dessas ações poderia causar um efeito cascata e, conseqüentemente, influenciar as demais ações. As ações seguidoras (*lagging*) podem chegar a duplicar os movimentos de preço das ações em momentos futuros. Este fenômeno é conhecido como efeito *lead-lag* (YUAN; JIN; LIAN, 2021).

Nos últimos anos, foram publicados diversos estudos dedicados à análise das correlações entre os mercados de ações de diferentes países, buscando identificar um efeito *lead-lag*. De acordo com Brandao (2020) parte das pesquisas foram motivadas pelo colapso mundial quase simultâneo dos mercados de ações em outubro de 1987, que forneceu evidências de fortes ligações entre movimentos de preços dos principais mercados de ações mundiais. Muitas vezes, o impacto do desempenho de outras bolsas de valores é maior em países em desenvolvimento, como é o caso do Brasil.

A maioria dos estudos realizados sobre as correlações internacionais do mercado de ações e possíveis efeitos *lead-lag* envolveu os principais mercados de ações mundiais, como EUA e China. Por exemplo, Brooks (2008) examinou a relação *lead-lag* entre o índice Financial Times Stock Exchange 100 (FTSE 100), da bolsa de valores de Londres, e o preço do índice futuro com base em dados de alta frequência. Kavussanos, Visvikis e Alexakis (2008) investigaram a relação *lead-lag* entre os futuros do índice de ações FTSE/Athex-20 e FTSE/Athex Mid-40 e os índices em dinheiro subjacentes na Grécia. Gong, Ji, Su, Li e Ren (2016) estudaram a relação *lead-lag* entre os índices China Securities Index 300 (CSI 300), Hang Seng Index, Standard and Poor's 500 (SP 500) e seus futuros associados para revelar a variação de sua relação ao longo do tempo. Entre os trabalhos associados à bolsa brasileira, pode-se destacar o de Oliveira e Medeiros (2009), que investigaram um possível efeito *lead-lag* entre o índice Dow Jones, da bolsa de Nova Iorque, com o índice Ibovespa, da bolsa de São Paulo, utilizando diversos modelos econométricos. Foi observado que o índice Ibovespa é, em grande parte, influenciado pelo movimento do Dow Jones, mostrando indícios de um efeito *lead-lag*.

Estudos recentes demonstraram resultados favoráveis na previsão, como os de Yuan, Jin e Lian (2021), que, por meio do método Caminho Térmico Ótimo, conhecido como Thermal Optimal Path (TOP), utilizaram dados da bolsa de Hong Kong para prever as ações da bolsa da China Continental, onde, mesmo a primeira liderando as ações em apenas 1 minuto, ainda é possível

observar um efeito *lead-lag* sobre a segunda. Ren, Ji, Cai e Jiang (2019) utilizaram testes de causalidade Granger e o método TOP entre três índices da Shanghai Stock Exchange, bolsa de Shanghai, e identificou maior efeito *lead-lag* analisando um intervalo de 1 hora entre as ações.

Com base nos estudos apresentados, foi levantada a hipótese de que seria possível utilizar os valores das ações desses mercados para a predição das ações no ocidente. Para isso, seria necessário verificar o efeito *lead-lag* entre essas bolsas, utilizando uma bolsa do oriente como o papel de *lead*, que lidera uma tendência, e uma do ocidente com o papel de *lag*, que reagem ao *lead* com defasagem temporal. Portanto, este estudo estabelece a seguinte pergunta de pesquisa: o efeito *lead-lag* em diferentes mercados, aliado a indicadores técnicos, podem ser utilizados na modelagem de redes neurais artificiais para a predição de movimento de preços de ações na B3? Para responder esta questão de pesquisa, foram utilizados dados da SoftBank Corp. (no Japão) para verificar se há um efeito *lead-lag* na empresa Telefônica Brasil.

Este estudo se justifica no âmbito de se estudar e encontrar formas mais eficientes para realizar predições na bolsa de valores, com o auxílio de informações provenientes de outras bolsas. Algumas bolsas de valores podem permitir identificar possíveis padrões de movimento de preço, antes mesmo que possam ocorrer em outras bolsas, visto que, as diferentes bolsas de valores possuem diferentes horários de funcionamento, por estarem localizadas em fusos horários diferentes.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Mercado de Ações

Os mercados financeiros causaram impacto significativo em muitas áreas como negócios, educação, empregos, tecnologia e, portanto, na economia (Hiransha, Gopalakrishnan, Vijay & Soman, 2018). Ao longo dos anos, investidores e pesquisadores têm se interessado em desenvolver e testar modelos de comportamento do preço das ações (Fama, 1995). No entanto, analisar os movimentos do mercado de ações e o comportamento dos preços é desafiador devido à natureza dinâmica, não linear, não estacionária, não paramétrica, ruidosa e caótica dos mercados (ABU-MOSTAFA; ATIYA, 1996).

De acordo com Zhong e Enke (2017) os mercados de ações são afetados por fatores inter-relacionados que incluem variáveis econômicas, políticas, psicológicas e específicas da empresa. A análise técnica e fundamentalista são as duas abordagens principais para analisar os mercados financeiros (PARK; IRWIN, 2007; NGUYEN; SHIRAI; VELCIN, 2015). Para investir em ações e obter lucros com baixo risco, os investidores têm usado essas duas abordagens principais para tomar decisões nos mercados financeiros (Arevalo, García, Guijarro & Peris, 2017).

2.2 Análise Técnica

Três premissas para a abordagem da análise técnica foram apresentadas por Murphy (1999). A primeira - o mercado assimila tudo - todos os elementos e fatos que podem afetar o preço refletem no preço. Portanto, entender e analisar os preços é o suficiente para compreender o comportamento do mercado. A segunda premissa - os preços seguem tendências - uma vez em

movimento, tendem a seguir aquele movimento, devendo a análise buscar identificar as tendências em seus primeiros sinais. A terceira premissa - a história se repete - para entender e prever o futuro é necessário entender o passado. Com base nesses conceitos, pode-se considerar que a análise técnica busca extrair indicadores a partir de resultados passados para estimar o comportamento e movimentos futuros no preço.

De acordo com Oriani e Coelho (2016) os indicadores podem ser categorizados em *leading* e *lagging*, o primeiro indica mudança, e o segundo confirma uma tendência existente. Dentre os indicadores do tipo *lagging*, destacam-se a Média Móvel Simples e a Média Móvel Exponencial, e entre os do tipo *leading*, o Indicador de Força Relativa e a Média Móvel Convergente/Divergente. A Taxa de Variação, também, pode ser utilizada como indicador, mensurando o deslocamento dos preços com base no preço de fechamento de datas anteriores (SACHETIM, 2006). A seguir são detalhados os cálculos dos indicadores técnicos utilizados neste estudo: média móvel simples, média móvel exponencial, indicador de força relativa, média móvel convergente/divergente e taxa de variação.

A Média Móvel Simples (MMS ou *Simple Moving Average* - SMA) é um indicador do tipo *lagging* que agrega um intervalo de resultados em um valor suavizado, indicando a direção da tendência e gerando potenciais sinais de venda e compra. Esse indicador se tem por base a tendência atual, não antecipando tendências futuras (MURPHY, 2021). Pode apresentar problemas quando o mercado não possui uma tendência definida. O indicador SMA pode ser utilizado de diversas formas, mas em geral sua utilização é realizada por meio do cálculo do mesmo indicador várias vezes com base em períodos diferentes, como visto em Baba, Wang, Kawachi, Xu e Deng (2003) e Lauren e Harlili (2014). O cálculo é dado pela seguinte expressão matemática:

$$SMA(n) = \frac{\sum_{i=1}^n P_i}{n} \quad (1)$$

onde P_i é o preço do fechamento do dia i e n o período em dias.

A Média Móvel Exponencial (MME ou *Exponential Moving Average* - EMA), como o indicador anterior, também, é um indicador *lagging*. O EMA é utilizado, interpretado e tem os mesmos objetivos que o indicador SMA, ou seja, identificar a direção da tendência e gerar potenciais sinais de vendas e compras. Porém, o EMA cobre uma fraqueza da SMA, atribuindo um peso maior aos resultados mais recentes, se adaptando melhor às mudanças em mercados voláteis. Entretanto, seu uso possui maior probabilidade da geração de *whipsaw* (sinais incorretos) (MURPHY, 2021). O EMA utiliza o valor das médias móveis exponenciais dos dias anteriores para o seu cálculo. Entretanto, devido à definição recursiva, é necessária a definição do caso base, para evitar a realização de recursão indefinidamente. Sendo assim, para o primeiro dia deve ser utilizado o SMA (MURPHY, 2021). Para o cálculo do indicador, primeiramente é calculado o valor do fator de ponderação k com a fórmula matemática a seguir:

$$k = \frac{2}{n + 1} \quad (2)$$

Baseando-se no valor de k e levando em consideração que na fórmula de EMA é considerado o primeiro intervalo como um SMA na posição de EMA($n-1$), representando o EMA do dia anterior, tem-se a seguinte expressão matemática:

$$EMA(n) = (P_i - EMA(n - 1)) * k + EMA(n - 1) \quad (3)$$

O Indicador de Força Relativa (*Relative Strength Index* - RSI) é um indicador do tipo *leading*. Ele mostra a velocidade da mudança nos preços e o status de sobrecomprado e sobrevendido, ou seja, quando estão sendo negociadas ações que se encontram fora da situação entendida como real, acima ou abaixo do valor entendido, respectivamente. Segundo Drakopoulou (2016), este indicador é utilizado com a faixa de valores entre 0 e 100, indicando que uma ação está sobrecomprada quando o valor do índice se encontra acima de 70 ou 80, ou sobrevendida, com o valor do índice abaixo de 30 ou 20. O indicador serve ao mesmo tempo para gerar sinais de compra e de venda, e para indicar o início de uma nova tendência, algo que ocorre quando o indicador cruza o valor de 50, caso este cruzamento seja ascendente a tendência é de alta nas ações, e caso seja descendente, a tendência é de queda nas mesmas. Para o cálculo do RSI, é utilizada a seguinte expressão matemática:

$$RSI = 100 - \frac{100}{1 - RS} \quad (4)$$

onde

$$RS = \frac{\text{GanhoMédioNoPeríodo}}{\text{PerdaMédiaNoPeríodo}} \quad (5)$$

A Média Móvel Convergente/Divergente (*Moving Average Convergence/Divergence* - MACD) é um indicador do tipo *leading*. É calculado utilizando dois indicadores EMA que, ao se cruzarem, podem representar um sinal de compra ou venda. É comum a utilização de EMA com valores de 12 e 26 dias no cálculo, sintetizando a movimentação de 2 semanas e 1 mês. Outros valores podem ser utilizados de acordo com a necessidade (MURPHY, 2021; KANG, 2021). O MACD pode ser calculado a partir da expressão matemática a seguir:

$$MACD = EMA(12) - EMA(26) \quad (6)$$

A Taxa de Variação (*Rate of Change* - ROC) mensura o deslocamento percentual dos preços com base na análise do preço de fechamento de um período de n dias passados (SACHETIM, 2006). Representa um possível otimismo e pessimismo do mercado. A ROC pode ser calculada a partir da expressão matemática a seguir:

$$ROC = \frac{\text{PreçoDeFechamento} - \text{PreçoDeFechamento}(n)}{\text{PreçoDeFechamento}(n)} \quad (7)$$

2.3 Redes Neurais Artificiais

Redes neurais artificiais (RNAs) são sistemas que simulam o comportamento e a estrutura do sistema nervoso humano e de seus componentes mínimos, os neurônios. De acordo com Hebb (1949), os neurônios são células especializadas na transmissão de informação e são compostos

simplificadamente por dendritos, núcleo e axônio. Os dendritos são como braços dos neurônios que recebem os impulsos de outros neurônios, e por sinapses químicas ou elétricas processam a informação, e a retransmitem por seu axônio a outros neurônios. As sinapses, quando frequentes, se fortalecem e podem enfraquecer quando não são frequentemente ativadas. Este comportamento sugerido por Hebb (1949) é uma das bases para a aprendizagem em RNAs.

Como no cérebro humano, apenas um processamento limitado é realizado por pequenos componentes, mas graças ao grande número de componentes e a associação entre eles ser de forma paralelamente distribuída em larga escala, o todo é capaz de realizar tarefas complexas. Além disso, o sistema é capaz de se adaptar e responder a entradas diversas. Essa qualidade torna as RNAs versáteis para serem utilizados em diversas áreas do conhecimento.

O segredo da aprendizagem (treinamento) estaria na calibração da ponderação entre as entradas de acordo com a checagem entre o resultado produzido e o resultado esperado por uma unidade de processamento que é um neurônio artificial, a menor unidade de uma RNA. De acordo com Haykin (2008), um neurônio artificial é constituído de: Um número fixo de entradas n simula os dendritos; Uma única saída y simula o axônio; Pesos w modulam o comportamento das sinapses, podendo ser um estímulo inibitório ou excitatório; Um bias, ou tendência, b representa um erro sistemático, agindo como um modulador da saída; A função de ativação realiza o comportamento do neurônio artificial.

Ao associar vários destes neurônios artificiais em camadas de nós, a capacidade do aprendizado aumentaria, podendo representar diferentes modelos para diferentes tipos de tarefas: regressão, classificação, predição, etc. Uma RNA com várias camadas compreende: uma camada de entrada, que recebe os sinais das variáveis de um problema a ser modelado, distribuindo para camadas intermediárias, chamadas de camadas ocultas, e uma camada de saída, onde a solução do problema é obtida. O número de camadas intermediárias pode variar, também aumentando o tempo necessário para o treinamento da rede neural (HAYKIN, 2008).

Especificamente, neste estudo, a RNA foi treinada como um previsor para transações de: compra (+1), venda (-1) e manutenção do ativo (0).

3 MATERIAIS E MÉTODOS

Para o desenvolvimento deste estudo, foram utilizados recursos computacionais, os quais são descritos na sequência. Como plataforma para o desenvolvimento, testes e execução do algoritmo foi utilizado um *desktop* com: Processador Intel(R) Core(TM) i5 CPU 760 2.80 GHz; Memória RAM 4GB; Disco rígido de 1000GB; Sistema operacional Windows 10 Pro 64 bits.

As bases de dados referentes às ações das bolsas de valores da Tokyo Stock Exchange e da B3 foram obtidas a partir da plataforma do Yahoo Finance. Para o pré-processamento dos dados da Tokyo Stock Exchange e da B3, e subsequente extração de métricas e análise estatística, foi utilizada a linguagem Python 3.9.5 e bibliotecas específicas: Pandas, Numpy e o Scikit-learn.

A biblioteca Pandas foi utilizada para facilitar a manipulação e a análise dos dados, pois oferece meios para trabalhar com tabelas numéricas e séries temporais. Para a geração dos indicadores técnicos propostos, foi utilizada a biblioteca Numpy, que auxilia a manipular os dados brutos,

possuindo suporte à *arrays* e matrizes de forma eficiente, aproveitando a estrutura de dados que a biblioteca Pandas disponibiliza.

O modelo de aprendizagem de máquina proposto, incluindo o treinamento, execução e avaliação, foi executado utilizando a biblioteca Scikit-learn, que é popularmente utilizada em pesquisas na área de inteligência artificial, a qual dispõe de diversos algoritmos de classificação, regressão, agrupamento e extração.

Em relação ao método, o desenvolvimento do trabalho foi dividido nas seguintes etapas: 1 – Extração e análise dos dados; 2 – Modelo geral de simulação; 3 – Simulador de pré-avaliação; 4 – Simulador com aplicação de análise técnica; 5 – Simulador com redes neurais artificiais.

3.1 Extração e Análise de Dados

Para o desenvolvimento do trabalho, foi proposto analisar o efeito *lead-lag* entre duas bolsas, sendo que seria utilizada uma bolsa presente em um país do oriente, possuindo horário de abertura e fechamento primeiro, e uma do ocidente, com horário de abertura e fechamento defasado temporalmente em comparação à primeira. Para isso seriam utilizados índices gerais e dados de uma empresa que se encontraria presente em ambas bolsas.

Para analisar o efeito de *lead*, foi escolhida a bolsa Tokyo Stock Exchange (TSE), localizada em Tóquio, Japão, pelo fato de ser considerada a maior bolsa do oriente. Já para o efeito *lag*, foi escolhida a bolsa B3, antiga Bovespa, localizada em São Paulo, Brasil, também pelo fato de ser a principal bolsa de valores do país.

Após a escolha das bolsas, foram escolhidos os maiores indicadores de desempenho das ações negociadas na TSE e na B3, os índices Nikkei 225 e Ibovespa, respectivamente. As informações dos índices, contendo o nome das empresas participantes, código, categoria de atuação e o peso das ações das mesmas, foram extraídos a partir da plataforma Yahoo Finance.

Com base nas informações dos índices, não foi identificada nenhuma empresa que possua ações negociadas em ambas as bolsas de valores. Portanto, foram escolhidas as principais empresas que compõem o mesmo setor de atuação nas duas bolsas. A escolha foi realizada através do peso de relevância das ações das empresas do setor de Tecnologia e/ou de Telecomunicações. O peso indica a representatividade de uma ação dentro de um índice.

A partir da Nikkei 225, foi selecionada a empresa SoftBank Group Corp. (9984), que compõe o setor de Tecnologia e ramo de Comunicações, com peso de 6.33% no momento da cotação em 2021. A partir da Ibovespa, foi selecionada a empresa Telefônica Brasil (VIVT3), que compõe o setor de Telecomunicações, com peso de 1.19% no momento da cotação. Após a seleção dos índices e das empresas, foram extraídas bases de dados históricos das ações de cada uma, sendo equivalente a um intervalo de 5 anos (01 de janeiro de 2017 a 31 de dezembro de 2021). As bases possuem dados referentes à data, preço de abertura, preço de fechamento, preço de fechamento ajustado, máximo, mínimo e volume.

3.2 Modelo Geral de Simulação

Foram desenvolvidos simuladores para compra e venda de ações, que, com base na análise de indicadores técnicos, observam as variações nos preços das ações no decorrer de um período, e emitem sinais para tomada de decisões de compra e venda.

A princípio, os simuladores iniciam com um montante de US\$100.000,00 (cem mil dólares), onde, ao primeiro sinal de compra, todo o valor do montante seria utilizado para comprar uma quantidade de ações que depende do preço do fechamento do dia. Havendo ações em posse, ao primeiro sinal de venda, todas as ações seriam vendidas, havendo o retorno de dinheiro para o montante, valor que também depende do preço de fechamento da ação na data da venda. Ambos os processos de compra e venda são repetidos até o final de cada experimento. Caso ao fim da execução hajam ações em posse, as mesmas são vendidas com base no preço de fechamento do último dia. Os ganhos de cada indicador foram comparados entre si, assim como a porcentagem de acertos na predição das variações das ações para cada indicador, onde espera-se que o preço das ações caia após uma venda e que suba após uma compra. Para todos os experimentos, foram atribuídos os valores +1 como sinal de compra, -1 como sinal de venda e 0 como sinal para não haver nenhuma movimentação.

Antes do desenvolvimento dos simuladores, a partir dos dados históricos extraídos, foram calculados os seguintes indicadores técnicos: Média móvel simples rápida, de 5 dias (SMA5); Média móvel simples média, de 10 dias (SMA10); Média móvel simples lenta, de 20 dias (SMA20); Média móvel exponencial rápida, de 5 dias (EMA5); Média móvel exponencial média, de 10 dias (EMA10); Média móvel exponencial lenta, de 20 dias (EMA20); Média móvel convergente/divergente com EMA de 12 e 26 dias (MACD); Indicador de força relativa de 14 dias (RSI); Taxa de variação de 15 dias (ROC).

3.3 Simulador de Pré-avaliação

Foram realizados experimentos utilizando apenas a aplicação de análise técnica, servindo como uma pré-análise para os experimentos com redes neurais artificiais. Os experimentos foram executados para a VIVT3 e SFTY separadamente. As estratégias de investimento foram simuladas com base em cada indicador técnico e o preço de fechamento. Como resultado verificou-se que entre o SMA, EMA e o MACD, o único que permitiu predições de movimento com lucro, nas 2 empresas foi o MACD, sendo que ele funcionou melhor na SFTY.

Para o RSI, os melhores resultados foram obtidos com o limite máximo de 70 e mínimo de 30, tanto para a SFTY como para a VIVT3, e para o ROC, de 10 e -5 para a SFTY e para a VIVT3. Verificou-se também que o RSI e o ROC funcionaram melhor na VIVT3 do que na SFTY.

3.4 Simulador com Aplicações de Análise Técnica

Como pré-avaliação do efeito *Lead-lag* entre as bolsas da SFTY e VIVT3, foram realizados experimentos utilizando indicadores técnicos de ambas as bolsas como estratégia de investimento na VIVT3. Os simuladores foram denominados AT-LL1 e AT-LL2, onde AT vem da abreviação de "análise técnica", e LL de "*lead-lag*". Os seguintes dados serviram de características de entrada para os experimentos com seus respectivos simuladores:

- AT-LL1: Preço de fechamento da VIVT3 e indicadores técnicos da SFTY;
- AT-LL2: Preço de fechamento da VIVT3, indicadores técnicos da VIVT3 e indicadores técnicos da SFTY.

Primeiramente o período de tempo utilizado para os experimentos foi ajustado, equiparando as datas de operações das duas bolsas, pois uma bolsa pode não operar em determinada data devido a algum feriado nacional, acontecimento ou determinação da bolsa do país.

Os sinais de compra, venda e de não movimentação, emitidos por cada indicador, foram combinados através da técnica de votação (*voting*). A votação baseia-se em diversos resultados para a tomada de decisões. Nos experimentos, foram utilizados diferentes limiares de decisão para a votação, havendo resultados superiores com a utilização da decisão com base na votação da maioria dos indicadores. Portanto, foram eleitos os experimentos que obtiveram melhores resultados com base no lucro do investimento.

3.4.1 AT-LL1

Na primeira etapa dos experimentos, foram utilizados os indicadores técnicos da SFTY como fatores de decisão de compra e venda na VIVT3. Foram realizados experimentos utilizando o RSI, MACD, ROC, SMA e EMA da SFTY e o preço de fechamento da VIVT3, com limiar de decisão 3, denominados E1, e utilizando apenas do RSI, MACD e ROC da SFTY, e o preço de fechamento da VIVT3, com limiares de decisão 3 e 4, denominados E2 e E3, respectivamente.

Os experimentos foram executados fazendo uso de diferentes variações de limites superiores e inferiores para o RSI e ROC, com limites superiores e inferiores de 70 e 30, 75 e 30, e 80 e 20 para o RSI, e 10 e -5, 10 e -10, e 15 e -10 para o ROC, respectivamente. Os resultados dos experimentos estão presentes nas Tabelas 1, 2 e 3.

	Retorno (U\$)		
	RSI 70/30 e ROC 10/-5	RSI 75/30 e ROC 10/-5	RSI 80/20 e ROC 10/-5
E1	144892.70	135317.01	142674.35
E2	157258.13	154942.23	174793.43
E3	133136.19	127511.25	121719.63

Tabela 1 – Resultados dos experimentos E1, E2 e E3, com variados limites superiores e inferiores para o RSI e de 10 e -5 para o ROC.

	Retorno (U\$)		
	RSI 70/30 e ROC 10/-10	RSI 75/30 e ROC 10/-10	RSI 80/20 e ROC 10/-10
E1	180510.30	168580.71	144500.98
E2	134923.30	134916.47	109909.22
E3	134593.60	135334.05	184153.07

Tabela 2 – Resultados dos experimentos E1, E2 e E3, com variados limites superiores e inferiores para o RSI e de 10 e -10 para o ROC.

	Retorno (U\$)

	RSI 70/30 e ROC 15/-10	RSI 75/30 e ROC 15/-10	RSI 80/20 e ROC 15/-10
E1	170579.83	166800.16	149616.03
E2	132250.94	137886.38	170195.45
E3	137827.79	142148.91	109405.07

Tabela 3 – Resultados dos experimentos E1, E2 e E3, com variados limites superiores e inferiores para o RSI e de 15 e -10 para o ROC.

Pode-se verificar que não houveram prejuízos em relação ao montante inicial de investimento, tendo sido geradas diversas variações de lucro. Os melhores resultados ocorreram para E3, com RSI de 80 e 20, e ROC de 10 e -10, havendo 84.15% de lucro, seguidos de E1, com RSI de 70 e 30, e ROC de 10 e -10, havendo 80.51% de lucro.

3.4.2 AT-LL2

Na segunda etapa dos experimentos, foram utilizados os indicadores técnicos da VIVT3 e da SFTY, combinando os seus resultados para a tomada de decisões através da votação. Foram realizados experimentos utilizando o RSI, MACD, ROC, SMA e EMA da SFTY e da VIVT3 e preço de fechamento da VIVT3, com limiares de decisão 4 e 5 para a votação, denominados E4 e E5, respectivamente, e utilizando apenas o RSI, MACD e o ROC da SFTY e VIVT3, e o preço de fechamento da VIVT3, com limiares de decisão 3 e 4, denominados E6 e E7, respectivamente.

Os resultados dos experimentos estão descritos nas Tabelas 4, 5 e 6. É possível verificar que os experimentos que obtiveram maior lucro foram os de E6, com RSI de 80 e 20, gerando lucro de 115.77% para ROC de 10 e -10, de 96.75% para ROC de 10 e -5, e de 86.98% para ROC de 15 e -10.

	Retorno (U\$)		
	RSI 70/30 e ROC 10/-5	RSI 75/30 e ROC 10/-5	RSI 80/20 e ROC 10/-5
E4	169752.82	139914.53	117767.17
E5	153263.91	151881.64	151267.44
E6	150629.72	159408.07	196753.10
E7	124707.31	130820.80	140269.70

Tabela 4 – Resultados dos experimentos E4, E5, E6 e E7, com variados limites superiores e inferiores para o RSI e de 10 e -5 para o ROC.

	Retorno (U\$)		
	RSI 70/30 e ROC 10/-10	RSI 75/30 e ROC 10/-10	RSI 80/20 e ROC 10/-10
E4	169502.01	138300.63	131398.97
E5	162943.69	160546.50	159159.07
E6	156898.01	169540.67	215770.22
E7	143567.31	143567.31	145180.06

Tabela 5 – Resultados dos experimentos E4, E5, E6 e E7, com variados limites superiores e inferiores para o RSI e de 10 e -10 para o ROC.

	Retorno (U\$)		
	RSI 70/30 e ROC 15/-10	RSI 75/30 e ROC 15/-10	RSI 80/20 e ROC 15/-10
E4	156318.11	122180.71	131673.69
E5	163008.81	158280.53	145854.35
E6	145099.09	155002.33	186976.47
E7	159857.42	164115.51	147887.31

Tabela 6 – Resultados dos experimentos E4, E5, E6 e E7, com variados limites superiores e inferiores para o RSI e de 15 e -10 para o ROC.

3.5 Simulador com Redes Neurais Artificiais

Para o desenvolvimento das redes neurais artificiais, foi utilizado um perceptron multicamadas (MLP), que, de acordo com Haykin (2008), se trata de um modelo de rede neural artificial com uma ou mais camadas ocultas entre a camada de entrada e de saída, sendo amplamente utilizado em previsões e classificações. Como características de entrada, foram utilizados os dados do período do ano de 2017 a 2020 como conjunto para o treinamento da RNA, e os do ano de 2021 como conjunto para a realização dos testes com o simulador de investimento com o modelo de RNA já treinado.

A Figura 1 apresenta a representação geral da fase de treinamento da rede neural artificial utilizada. O detalhamento deste processo será detalhado a seguir.

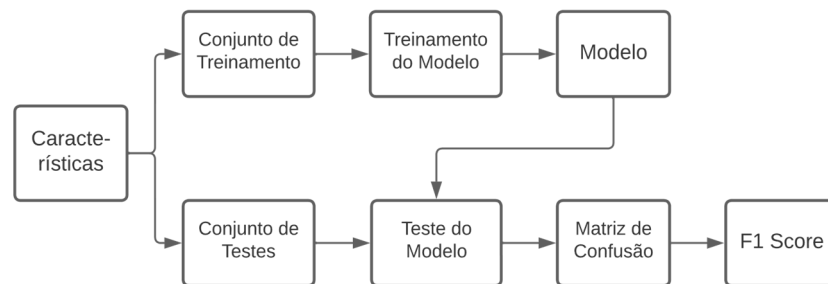


Figura 1 – Esquema da fase de treinamento do modelo de rede neural atificial.

Para a fase de treinamento do modelo de RNA, os dados de entrada foram separados os conjuntos de treinamento e testes, com tamanho de 10% para o conjunto de teste. Foi realizada a separação e a normalização das características de entrada, e então foram treinadas e testadas diversas configurações de rede (número e tamanho de camadas) e de entradas. Durante os testes do modelo, foi possível observar que muitas vezes não foi gerada convergência, isto é, o modelo não aprendeu os comportamentos que lhe eram esperados, onde em alguns casos o modelo só aprendia a vender sem nunca comprar, ou vice-versa. Muitas vezes a causa do problema poderia ser a *seed* escolhida aleatoriamente, que acabava não gerando um modelo útil. Com objetivo de mitigar este problema e haver reprodutibilidade dos experimentos, foi utilizado um intervalo fixo de *seeds* de 1 a 20 para a separação do conjunto de treinamento.

Cada configuração testada resultou em uma matriz de confusão, sendo calculado o F1 Score a partir da mesma, para avaliar o desempenho do modelo. Baseando-se nas estratégias de investimento que obtiveram os melhores resultados para os simuladores AT-LL1 e AT-LL2,

foram desenvolvidos dois simuladores de investimento utilizando redes neurais artificiais, denominados RNA-AT-LL1 e RNA-AT-LL2, onde RNA equivale a redes neurais artificiais.

3.5.1 RNA-AT-LL1

O RNA-AT-LL1 utilizou como características de entrada os preços de fechamento da VIVT3 e os indicadores técnicos da SFTY. Os indicadores técnicos utilizados foram o RSI de limites 80 e 20, ROC de limites 10 e -10, e MACD. Para a decisão de compra e venda, foi utilizada a votação com limiar de decisão 2.

Foram realizados 20 experimentos utilizando RNA-AT-LL1 devido a natureza aleatória do treinamento da RNA, e foram selecionados apenas os modelos com *F1 Score* maior do que 0,6 para os movimentos de compra (1), venda (-1) e manutenção (0) do ativo. O saldo final médio dos modelos selecionados foi de US\$ 109.659,46, com taxa de acerto médio nas transações de 91.74% e retorno médio de 9,66% em relação ao saldo inicial.

3.5.2 RNA-AT-LL2

O RNA-AT-LL2 utilizou como características os preços de fechamento da VIVT3 e os indicadores técnicos da VIVT3 e da SFTY. Os indicadores técnicos utilizados, de ambas empresas, foram o RSI de limites 80 e 20, ROC de limites 10 e -10, e MACD. Para a decisão de compra e venda, foi utilizada a votação com limiar de decisão 3.

Foram realizados 20 experimentos utilizando RNA-AT-LL2 devido a natureza aleatória do treinamento da RNA, e foram selecionados apenas os modelos com *F1 Score* maior do que 0,25 para os movimentos de compra (1), venda (-1) e manutenção (0) do ativo. O saldo final médio dos modelos selecionados foi de US\$ 119.427,85, com taxa de acerto médio nas transações de 83,61% e retorno médio de 19,43% em relação ao saldo inicial.

4 ANÁLISE COMPARATIVA

Com objetivo de verificar se ocorreu o efeito *lead-lag* entre a bolsa do Japão e do Brasil, utilizando as bolsas VIVT3 e SFTY como representantes, esta seção analisa e compara o desempenho dos experimentos utilizando redes neurais artificiais (RNA-AT-LL1 e RNA-AT-LL2), análise técnica com dados da VIVT3 e SFTY (AT-LL1 e AT-LL2), e análise técnica apenas com dados da VIVT3 e da SFTY separadamente (denominados AT-VIVT3 e AT-SFTY, respectivamente).

Foram realizados experimentos com os simuladores AT-LL1, AT-LL2, AT-VIVT3 e AT-SFTY utilizando apenas o período de 2021 para o investimento, visto que é necessário que os experimentos equipararem-se aos resultados de RNA-AT-LL1 e RNA-AT-LL2, onde foram realizados investimentos no ano de 2021 com base no treinamento da rede neural artificial com os dados de 2017 a 2020.

Buscando ter mais critérios de comparação, foram implementados dois outros simuladores, um oráculo e um *hold*. O oráculo conta com uma estratégia de investimento fictícia, sempre acertando em suas decisões, realizando movimentos de acordo com o preço de fechamento do dia seguinte. Por exemplo, são compradas as ações no dia D apenas se as mesmas valorizarem

em D+1, ou as ações são vendidas no dia D apenas se houver a desvalorização em D+1. O *hold* por sua vez simula um investimento de longo prazo, onde é esperado que a ação comprada se valorize com o tempo. Por exemplo, o *hold* compra as ações no primeiro dia do período utilizado e as vende apenas no último dia do período, não efetuando nenhuma movimentação de compra e venda durante o intervalo.

Para o período de 2021, o oráculo obteve 240.6% de lucro, representando um resultado máximo possível a ser obtido por meio do modelo de simulação utilizado. O *hold* obteve lucro de 4.9%, representando um investimento sem as estratégias utilizadas pelos simuladores, apenas comprando as ações no primeiro dia e vendendo-as no último dia.

Durante as análises, os simuladores AT-VIVT3, AT-SFTY, oráculo e *hold* não foram modificados, sendo alterado apenas o limiar de decisão da votação para AT-VIVT3 e AT-SFTY, alinhando-os aos experimentos com RNA.

4.1 Análise 1

Os resultados dos experimentos com os simuladores RNA-AT-LL1, AT-LL1, AT-VIVT3, AT-SFTY, oráculo e *hold*, estão apresentados na Figura 2. Foram apresentados dois valores para RNA-AT-LL1, onde RNA-AT-LL1(15) se refere ao melhor resultado com as redes neurais artificiais, que foi o de *seed* 15, e RNA-AT-LL1(média) se refere à média dos resultados com as 20 *seeds*. É preciso relembrar que foi considerado o limiar de decisão 2 para a votação, RSI de 80 e 20, e ROC de 10 e -10.

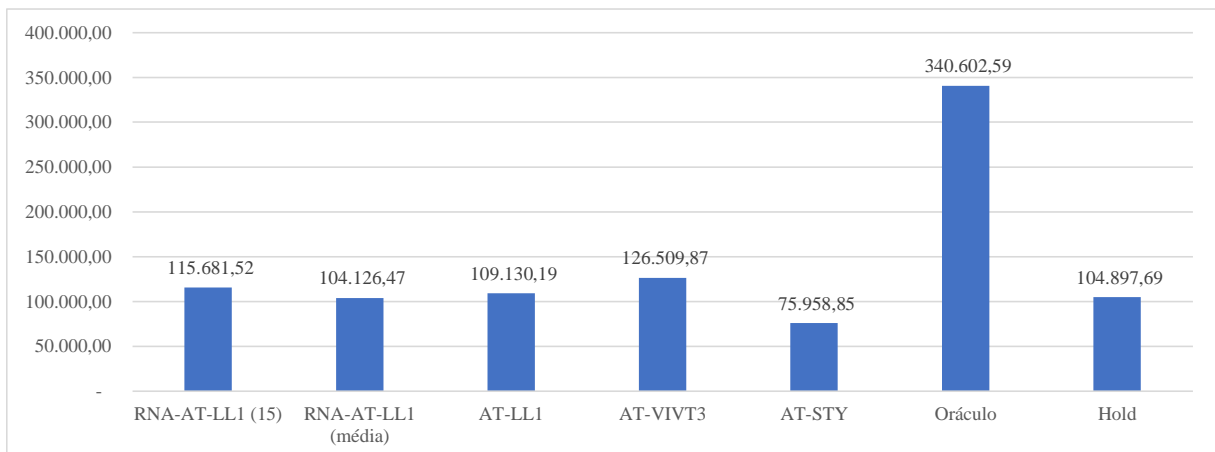


Figura 2 – Resultados dos investimentos, em US\$, dos experimentos utilizando os simuladores RNA-AT-LL1, AT-LL1, AT-VIVT3, AT-SFTY, oráculo e *hold*.

Pode-se verificar que o RNA-AT-LL1(15), que representa a melhor simulação de RNAAT-LL1, obteve melhor retorno financeiro do que RNA-AT-LL1(média) e AT-LL1, onde houve 15.98% de lucro. Também é possível observar que o RNA-AT-LL1(média), com lucro de 4.13%, não obteve retorno superior ao de AT-LL1, de 9.13%, sendo possível que, em média, a utilização de redes neurais artificiais em estratégias de investimento pode não gerar maior lucro do que a utilização da análise técnica sem aprendizagem de máquina. O RNA-AT-LL1(média) também não obteve mais lucro do que o *hold*, de 4.9%.

Além disso, nenhuma estratégia de investimento de *lead-lag*, com ou sem RNA, superou o lucro resultante de AT-VIVT3, de 26.51%, podendo ser justificado pelos resultados negativos de AT-SFTY, que obteve prejuízo de 24.04%. O baixo desempenho de AT-SFTY pode representar que a análise técnica não se deu bem para investimentos na empresa SFTY. Portanto, como os indicadores técnicos da SFTY foram utilizados como base para a estratégia de investimento na VIVT3, o baixo desempenho da análise técnica da própria SFTY pode ter afetado o efeito *lead-lag*. Entretanto, o lucro obtido por AT-VIVT3 ainda se encontra bem inferior ao do obtido pelo oráculo, de 240.6%.

4.2 Análise 2

A Figura 3 apresenta os resultados dos experimentos com os simuladores RNA-ATLL2, AT-LL2, AT-VIVT3, AT-SFTY, oráculo e *hold*. Também apresentados dois valores para RNA-AT-LL2, onde RNA-AT-LL2(3) se refere ao melhor resultado com as redes neurais artificiais, que foi o de *seed* 3. Foi considerado o limiar de decisão 3 para a votação, RSI de 80 e 20, e ROC de 10 e -10.

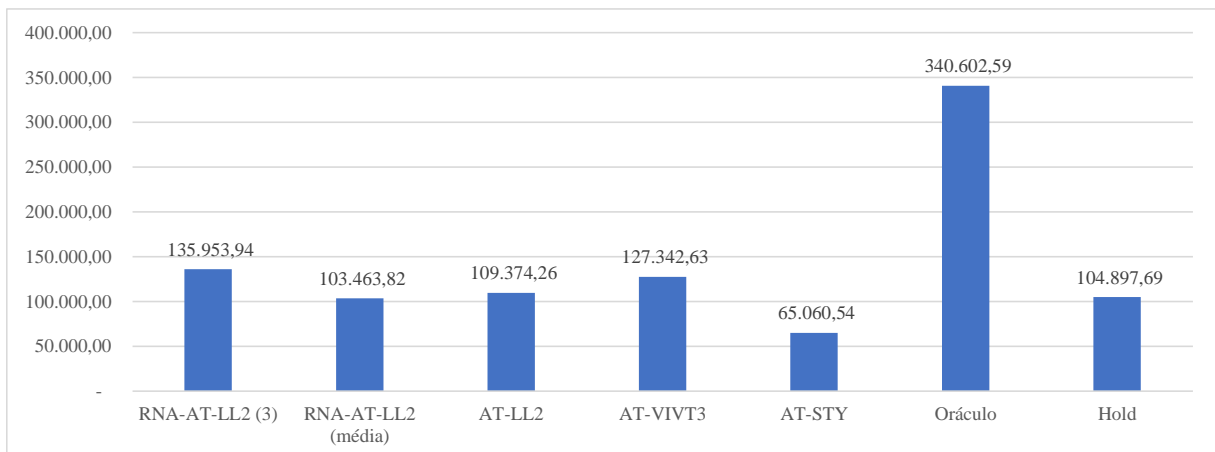


Figura 3 – Resultados dos investimentos, em US\$, dos experimentos utilizando os simuladores RNA-AT-LL2, AT-LL2, AT-VIVT3, AT-SFTY, oráculo e *hold*.

Pode-se verificar que o RNA-AT-LL2(3), melhor experimento utilizando redes neurais artificiais, obteve retorno financeiro superior ao RNA-AT-LL2, AT-LL2 e AT-VIVT3, havendo lucro de 35.95%, indicando que a utilização das redes neurais artificiais, em conjunto com a análise técnica da VIVT3 e SFTY, pode resultar em lucros nos investimentos. Entretanto, a média dos resultados de investimentos de RNA-AT-LL2, RNA-AT-LL2(média), com lucro de 3.46%, ainda não obteve retorno financeiro superior ao de AT-LL2, AT-

VIVT3, oráculo, e do *hold*, que obtiveram lucro equivalente a 9.37%, 27.34%, 240.6% e 4.9%, respectivamente.

Os resultados demonstraram que a utilização dos indicadores técnicos de ambas empresas pode gerar um efeito *lead-lag*, considerando o melhor caso, sendo uma estratégia mais viável do que a exposta na Análise 1.

CONCLUSÃO

Os investidores compram ações visando retorno financeiro no futuro, seja pelo pagamento de dividendos, pelos lucros alcançados ou vendendo as ações após a valorização da empresa em que investiu. Para alcançar o melhor resultado possível, o investidor busca definir uma melhor estratégia de acordo com o seu perfil de investimento, a partir do que conhece e espera do mercado, tentando prever qual empresa irá valorizar mais.

Dentro desse contexto, o presente trabalho teve como objetivo analisar o efeito *lead-lag* entre bolsas do Japão e do Brasil. Para isso, foram eleitas as empresas de maior peso de relevância, dentro da mesma área de atuação, das principais bolsas desses países, e foram realizados experimentos em cima de seus históricos de ações. As empresas escolhidas foram a SoftBank Group Corp. (SFTY), da Tokyo Stock Exchange, e a Telefônica Brasil (VIVT3), da B3, ambas da área de Telecomunicações.

Para alcançar o objetivo proposto, foram utilizados simuladores de análise técnica propriamente dita (AT), e de análise técnica utilizando redes neurais artificiais (RNA). Os experimentos consistiram em simular os investimentos na bolsa VIVT3 utilizando apenas os indicadores técnicos da SFTY, e simular os investimentos na bolsa VIVT3 utilizando os indicadores técnicos da VIVT3 e da SFTY em conjunto.

Os resultados do primeiro experimento, tanto com RNA quanto com AT, comparados investimento com apenas análise técnica da VIVT3, obtiveram baixo lucro financeiro. Entretanto, para o segundo experimento, o melhor resultado dentre os investimentos utilizando o simulador com RNA gerou lucro financeiro superior ao do investimento com análise técnica para VIVT3.

A verificação do efeito *lead-lag* entre as bolsas do Japão e do Brasil está longe de ser simples, porém, os resultados obtidos no segundo experimento podem indicar a existência desse efeito. Contudo, como a utilização de modelos de simulação com base em redes neurais artificiais depende de fatores aleatórios, dificulta demonstrar com garantia a obtenção de lucro financeiro ao se utilizar estratégias de investimento se apoiando em um possível efeito *lead-lag*.

Como possibilidades de trabalhos futuros, podem ser consideradas as seguintes opções: ampliar o estudo em relação ao efeito *lead-lag* considerando outras bolsas e diferentes setores de atuação das empresas a analisar; analisar outros indicadores técnicos e outros tipos de redes neurais para a predição do movimento de preços de ações.

REFERÊNCIAS

- Abu-Mostafa, Y. S., Atiya, A. F. (1996). Introduction to financial forecasting. *Applied Intelligence*, v. 6, p. 205–213.
- Arevalo, R., García, J., Guijarro, F., Peris, A. (2017). A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting. *Expert Systems with Applications*, v. 81, p. 177–192.
- Baba, N., Wang, Y., Kawachi, T., Xu, L. e Deng, Z. (2003). Utilization of ai and gas to improve the traditional technical analysis in the financial markets. *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, Springer, Berlin, Heidelberg, p. 1095–1099.

- Brandao, V. (2020). Correlação entre as bolsas de valores mundiais e a bovespa. caminho para riqueza. Disponível em: <<https://www.caminhoparariqueza.com.br/correlacao-bolsas-de-valores>>. Acesso em: 20 de set. de 2021. 22
- Brooks, C. *Introductory Econometrics for Finance*. (2008) Cambridge: Cambridge university press, 2nd ed.
- Drakopoulou, V. (2016). A review of fundamental and technical stock analysis techniques. *Journal of Stock Forex Trading*, v. 5, n. 1, p.16.
- Gong, C., Ji, S. Su, L., Li, S., & Ren, F. (2016). The lead-lag relationship between stock index and stock index futures: A thermal optimal path method. *Physica A: Statistical Mechanics and Its Applications*, v. 444, p. 63–72.
- Haykin, S. (2008). *Neural networks and learning machines*. [S.l.]: Prentice Hall.
- Hebb, D. (1949). *The organization of behavior*. [S.l.]: Wiley, New York.
- Hiransha, M., Gopalakrishnan, E.A., Vijay, K. M., & Soman, K.P. (2018). NSE Stock Market Prediction Using Deep-Learning Models. *Procedia Computer Science*, v. 132, p. 1351–1362.
- Kang, B. K. (2021). Improving macd technical analysis by optimizing parameters and modifying trading rules: Evidence from the japanese nikkei 225 futures market. *Journal of Risk and Financial Management*, v. 14, n. 1, p. 1–21.
- Kavussanos, M. G., Visvikis, I. D., Alexakis, P. D. (2008). The lead-lag relationship between cash and stock index futures in a new market. *European Financial Management*, v. 14, n. 5, p. 1007–1025.
- Lauren, S., Harlili, S. D. (2014). Evaluating the impact of technical indicators on stock forecasting stock trend prediction using simple moving average supported by news classification. *2014 International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA)*, p. 135–139.
- Murphy, J. J. (2021). *Análise técnica do mercado financeiro: Um guia abrangente de aplicações e métodos de negociação*. [S.l.]: Alta Books.
- Murphy, J. J. (1999). *Study Guide to Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. [S.l.]: Prentice Hall Press.
- Nguyen, T. H., Shirai, K., Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, v. 42.
- Oliveira, G. R., Medeiros, O. R. (2009). Testando a existência de efeitos lead-lag entre os mercados acionários norte-americano e brasileiro. *Brazilian Business Review*, v. 6, n. 1, p. 1–21.
- Oriani, F. B., Coelho, G. P. (2016). Evaluating the impact of technical indicators on stock forecasting. *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, p. 1–8.
- Park, C., Irwin, S. H. (2007). What do we know about the profitability of technical analysis? *Journal of Economic Surveys*, v. 21, p. 786–826.
- Ren, F., Ji, S., Cai, M., Li, S., & Jiang, X. (2019). Dynamic lead–lag relationship between stock indices and their derivatives: A comparative study between chinese mainland, hong kong and us stock markets. *Physica A: Statistical Mechanics and its Applications*, v. 513, n. 1, p. 709–723, 2019.
- Sachetm, H. M. (2006). *Análise técnica: estudo da confiabilidade dos principais indicadores de análise técnica, aplicados as ações mais negociadas na bovespa no período de 1995 a 2005*.
- Worldtimezone. (2021) *World Time Markets Maps*. Disponível em: <<https://www.worldtimezone.com/>>. Acesso em: 01 set. 2021. 11
- Yuan, X., Jin, L., Lian, F. (2021). The lead–lag relationship between chinese mainland and hong kong stock markets. *Physica A: Statistical Mechanics and its Applications*, v. 574, p. 125999.