



## INFLUÊNCIA DE MERCADO DA CHINA NA PREDIÇÃO DO MOVIMENTO DE PREÇO COM BASE EM REDES NEURAIIS ARTIFICIAIS

### INFLUENCE OF CHINA MARKET ON PREDICTION OF PRICE MOVEMENT BASED ON ARTIFICIAL NEURAL NETWORKS

#### 6. Finanças

João Vitor Malvestio da Silva, UEM, Brasil, [ra93089@uem.br](mailto:ra93089@uem.br)

Wagner Igarashi, UEM, Brasil, [wigarashi@uem.br](mailto:wigarashi@uem.br)

Deisy Cristina Corrêa Igarashi, UEM, Brasil, [dccigarashi@uem.br](mailto:dccigarashi@uem.br)

#### Resumo

Por conta da queda da taxa Selic e consequentemente o baixo retorno nos investimentos de renda fixa, o mercado de ações vem se aquecendo e se tornando cada vez mais popular na vida dos brasileiros. Esse tipo de investimento pode ser lucrativo, mas possui riscos elevados, muitos fatores podem levar a aumentar ou diminuir os valores das ações durante um mesmo dia. Diante disso, esse trabalho se propõe a utilizar indicadores técnicos de mercados defasados, no caso o Chinês, para diminuir os riscos dos investidores. Foram utilizados os preços e indicadores técnicos da empresa China Telecom Corp Ltd em uma rede neural artificial de modo a simular investimentos na empresa Telefônica Brasil. Como resultado verificou-se uma diferença de ganho positiva ao se utilizar dados de uma empresa em um mercado defasado no tempo.

**Palavras-chave:** Indicadores técnicos; Mercado defasados; Redes neurais artificiais; Predição de preço.

#### Abstract

Due to the drop in the Selic rate and consequently the low return on fixed income investments, the stock market has been heating up and becoming increasingly popular in the lives of Brazilians. This type of investment can be profitable, but it has high risks, many factors can lead to an increase or decrease in stock values during the same day. Therefore, this work proposes to use technical indicators of lagged markets, in this case the Chinese, in an artificial neural network model to reduce the risks of investors. The prices and technical indicators of the company China Telecom Corp Ltd were used in order to simulate investments in the company Telefônica Brasil. As a result, there was a positive gain difference when using data from a company in a market lagged in time.

**Keywords:** Technical indicators; Lagged markets; Artificial neural networks; Price prediction.

#### 1. INTRODUÇÃO

Em 2019, o índice Ibovespa pela primeira vez atingiu 100 mil pontos (“Bovespa alcança”, 2019), devido ao alto número de pessoas que estão experimentando esse tipo de mercado. Muitos são atraídos pelas possibilidades de lucro, porém, não se atentam a volatilidade desse tipo de mercado.

A inteligência artificial está crescendo dentro e fora da área de tecnologia e vem se mostrando útil em outras áreas de atuação completamente diferentes da tecnologia, como por exemplo a área financeira. Esse tipo de aplicação vem ajudando pessoas ao redor do mundo que buscam um lucro razoável e mais segurança em seus investimentos, principalmente quando se trata de bolsa de valores.

A justificativa do desenvolvimento deste estudo, se deve ao fato de investir na bolsa de valores estar vinculado a riscos, de cada vez mais investidores serem atraídos para este tipo de investimento, desde os mais agressivos, isto é, investidores que entendem perdas a curto prazo visando aproveitar lucros mais altos a longo prazo, até investidores mais conservadores, que prezam por conservar seu patrimônio e se sentem desconfortáveis com qualquer tipo de perda. Então, este trabalho tem como objetivo analisar dados de bolsas defasadas no tempo para a predição de preço na bolsa brasileira utilizando técnicas de inteligência artificial.

A motivação desse estudo se dá por analisar a bolsa da China, por se tratar de uma bolsa defasada no tempo, e que temporalmente abre mais cedo em relação à bolsa brasileira. Deste modo, esta informação pode ser utilizada para melhorar os modelos de predição na bolsa do Brasil, visando otimizar o lucro e minimizar os riscos.

Contudo, este estudo apresenta limitações. Segundo Gong e Cortese (2017) a China é um país socialista com características chinesas e por conta disso não possui muitas empresas globais com ações em suas bolsas, isso impede comparações diretas de ações da mesma empresa na bolsa de valores brasileira. Mesmo assim, a China ainda é referência mundial em alguns ramos e para contornar isso, serão realizadas comparações de empresas do mesmo ramo para identificar pontos de queda e ascensão de ações.

O presente trabalho apresenta, além desta seção de introdução: a seção 2, o qual versa sobre a revisão bibliográfica relativa aos conceitos teóricos dos indicadores técnicos, redes neurais que auxiliam ou mesmo delinham o trabalho e também os trabalhos correlatos; a seção 3, no qual são descritos os métodos e ferramentas utilizadas durante a pesquisa; a seção 4; no qual é delineado o desenvolvimento do trabalho e apresentada a estrutura e funcionamento dos dois modelos de predição criados; a seção 5; no qual apresenta e discute os resultados obtidos a partir de cada um dos modelos criados; a seção 6, no qual são delineadas as conclusões e as possibilidades de trabalhos futuros; por fim, são descritas as referências bibliográficas.

## **2. REVISÃO TEÓRICA**

Para discorrer sobre investimentos na bolsa de valores, é necessário entender as motivações desse tipo de investimento, do que se trata e também compreender seus indicadores.

### **2.1 Análise do mercado de investimentos e mercados defasados**

Antes de tudo é fundamental falar sobre a taxa básica de juros da economia brasileira, conhecida como taxa SELIC, esta é utilizada pelo Banco Central do Brasil (n.d.) para controlar a inflação. Em 2020 houve uma queda brusca da taxa para 2%, onde atingiu seu menor patamar na história, porém, em 2021 é notável um crescente aumento desse valor. Como dito anteriormente, essa taxa tem um papel importante no controle da inflação, quanto menor a taxa, menor é a rentabilidade das aplicações financeiras, conseqüentemente, incentiva mais a

população a investir em rendas variáveis e colocar mais dinheiro em circulação. Assim, quando a taxa está alta, aplicações fixas trazem mais segurança e um maior retorno, evitando que tenha muito dinheiro em circulação e também tendendo a diminuir a inflação.

Historicamente, houve uma queda acentuada na taxa SELIC no ano de 2020, o que possivelmente impulsionou parte da população brasileira a se interessar em aplicações na bolsa de valores. No Brasil existe a Brasil, Bolsa, Balcão (B3, n.d) que é o resultado da combinação entre a BM&FBovespa e a Cetip, se tornando uma companhia de infraestrutura de mercado financeiro de classe mundial. A bolsa de valores se resume a um ambiente de negociações de valores mobiliários como ações, títulos públicos e *commodities*. Ela vem para ser um ponto de encontro entre os investidores, trazendo regras, segurança e transparência entre as partes envolvidas. Essa bolsa agrega vários tipos de ativos, entre os quais o mais conhecido são as ações que, por sua vez, são uma pequena parte de uma empresa, precificada pelas expectativas dos investidores. Então, ao investir dinheiro nestes papéis, se compra uma parte de uma companhia, ou seja, se torna um sócio e acionista.

Além do Brasil, existem vários países com suas próprias bolsas de valores. A China possui três das maiores bolsas do mundo, a *Shanghai Stock Exchange* (SSE) (Overview<sub>1</sub>, n. d.), que segundo seu site oficial contém em torno de 1,7 mil companhias listadas e valor de mercado de aproximadamente US\$ 6 trilhões, e a *Shenzhen Stock Exchange* (SZSE) (Overview<sub>2</sub>, n. d.), que segundo seu site oficial, possui o valor de mercado de cerca de US\$ 4,5 trilhões e quase 2,3 mil empresas listadas. Essas duas bolsas se localizam na parte continental da China. Também existe a *Hong Kong Stock Exchange* (HKEX) (About HKEX, n.d), acompanhada pelo índice Hang Seng, que se localiza fora do continente e segundo seu site oficial, é responsável por monitorar as 50 maiores e mais líquidas empresas negociadas na ilha. Essa é a bolsa mais antiga das três.

A bolsa de HKEX possui o horário de funcionamento, com base no horário de Brasília, iniciando com a pré-abertura às 22h30m, funcionamento das 23h até 05h e o pós-mercado das 05h25m até 06h. A bolsa de Xangai faz sua pré-abertura às 22h, o mercado vai das 22h30m às 04h e o pós-mercado das 04h25m às 05h. Já a bolsa de Shenzhen *Stock Exchange* possui os horários mais diferentes, com a pré-abertura às 22h15m, o funcionamento é das 22h30m até às 4h. Porém possui uma pausa para almoço das 00h30m até as 2h. Todas essas informações foram colhidas dos próprios sites oficiais de cada bolsa. Dado que o horário de funcionamento da B3 é das 10h até às 17h segundo seu site oficial (B3, n. d.). Essas bolsas funcionam de segunda a sexta, e não abrem em feriados. É notável também que as bolsas da China se encerram na madrugada antes que a bolsa B3 do Brasil inicie suas operações.

Hoje existem duas operações principais que os *traders*, assim são chamados os especuladores da bolsa, fazem, o *day trade*, que se resume a operações de compra e venda no mesmo dia, e o *swing trade*, são operações que duram normalmente dois dias, mas podem chegar a mais de duas semanas.

O principal objetivo de quem investe na bolsa de valores é o lucro a curto prazo, porém, por se tratar de um mercado muito volátil, os riscos de perda são altos. Para chegar nesse lucro, existem várias formas de prever valores das ações, como análise técnica e inteligência artificial. Dentro da inteligência artificial, temos a aprendizagem de máquina ou *machine*

*learning* que é um sistema que pode modificar seu comportamento autonomamente tendo como base a sua própria experiência com a mínima interferência humana.

Para a criação de modelos de predição que possam simular como obter esse lucro e diminuir os riscos, a estratégia que será tomada é utilizar dados históricos de uma empresa na bolsa de valores da China, que é um mercado defasado no tempo, pois por conta do fuso horário, a bolsa da China encerra as suas operações antes que a bolsa do Brasil inicie suas operações, e com isso é possível utilizar análise técnica e técnicas de aprendizagem de máquina, como a rede neural, para realizar a predição do movimento dos preços de ações de uma empresa na bolsa brasileira.

## 2.2 Análise de indicadores técnicos

Para auxiliar as análises dentro da bolsa de valores, existe um grande aliado que são os indicadores técnicos. Esses indicadores são fórmulas matemáticas que são aplicadas a dados históricos de preços de fechamento de alguma ação e o resultado evidencia regiões ou tendências de movimentos que são críticas para o processo decisório de compra e venda. Existem vários indicadores técnicos no mercado, esses são divididos em três principais classes que são conhecidos como indicadores de tendência, osciladores e indicadores de volume. É válido ressaltar que todos esses indicadores retornam valores para serem plotados em gráficos para ajudar a visualizar tendências, mas no caso, essas informações serão utilizadas pelo modelo de rede neural implementado.

## 2.3 Indicador de volume

Os indicadores de volume são aqueles que representam o número de transações que aparecem em um intervalo de tempo, que também é conhecido como volume. Nas ações, o volume significa o volume de negócios realizados. Dentro desse indicador temos alguns mais utilizados ou conhecidos como o indicador de acumulação/distribuição, índice de fluxo de dinheiro, volume e o mais famoso, o *On Balance Volume*, também chamado de OBV. Segundo Lemos (2015) este, que foi criado por Joe Granville, faz uma relação de preço e volume. Para o cálculo do OBV, é utilizada a seguinte fórmula (Lemos, 2015):

$$OBV_{atual} = OBV_{anterior} + \begin{cases} \text{volume,} & \text{se fechamento} > \text{Fechamento}_{anterior} \\ 0, & \text{se fechamento} = \text{Fechamento}_{anterior} \\ -\text{volume,} & \text{se fechamento} < \text{Fechamento}_{anterior} \end{cases}$$

Onde o valor de OBV é o valor anterior dele mesmo somado a variação de volume, lembrando que esse valor deve ser inicializado em 0, e possui três situações possíveis: caso o valor de fechamento seja maior que o valor de fechamento anterior, é adicionado o volume, caso o valor de fechamento for o mesmo que o fechamento anterior, é somado a 0 e caso o valor de fechamento for menor que o fechamento passado, é subtraído o volume.

O valor do indicador sozinho não é muito útil, como visto na fórmula, deve ser utilizado seu histórico para utilizar todo o seu potencial. Para que este indicador tenha utilidade, é possível interpretar das seguintes formas:

- Indicador OBV crescente: demanda sobre o ativo está aumentando;

- Indicador OBV decrescente: demanda sobre o ativo está caindo;
- Se o volume aumenta com o preço da ação, é sinal de valorização do ativo;
- Se o volume começa a cair enquanto a ação sobe, é sinal de que a alta está perto do fim;
- Se o volume cai com o preço da ação, é sinal de tendência de queda;
- Se o volume cresce, mas o preço continua caindo, é sinal de que pode haver reversão para um movimento de alta.

Todas essas situações são de grande importância para as tomadas de decisões para quem analisa os resultados, pois mostra uma possível tendência para a ação analisada.

## 2.4 Indicador de tendência

Já os indicadores de tendência, tem como objetivo apresentar tendências como o próprio nome indica. Eles indicam pontos de tendência de alta ou pontos de tendência de baixa. Dentro desta classe existe um indicador muito conhecido que é chamado de *moving average*, ou médias móveis, que é consultado antes de ser realizada ordens de compra ou de venda por ter uma assertividade alta. O indicador utilizado neste trabalho é a média móvel simples (*simple moving average* – SMA).

Segundo Murphy (1999) SMA é uma das técnicas de análise mais utilizadas, ele tem como função exibir o preço médio de uma ação em determinado período, além de sua mudança com o decorrer do tempo. Normalmente são utilizados períodos de 9, 10, 21, 34 dias. Nesse estudo foi utilizado o período de 14 dias. A SMA possibilita a eliminação de ruídos de flutuação permitindo focar na tendência da ação. Na sequência pode-se observar a fórmula da SMA a partir de um período  $n$  do preço de fechamento ( $C - close$ ) de uma ação (Lemos, 2015):

$$SMA = \frac{C_1 + C_2 + \dots + C_n}{n}$$

### 2.4.1. Indicador De Oscilação

Por fim, os osciladores são utilizados para avaliar a variação dos preços dos ativos, utilizados em mercados lateralizados, ou seja, mercados que ficam estabilizados dentro de um intervalo de preço. Utilizando este tipo de indicador é possível identificar pontos de compras e vendas. Neste contexto, o mais conhecido e utilizado é o índice de força relativa (*relative strength index* - *RSI*), que segundo Murphy (1999), é um bom indicador de sobrecompra e sobrevenda. Com essas informações é possível determinar que é um bom momento de vender as ações se há sobrecompra, pois elas estarão com preço elevado, e que é um bom momento de comprar ações se há sobrevenda, pois elas estarão com preços baixos. De modo geral, este indicador varia de 0 a 100, e em geral é utilizado como padrão o valor de 70 para sobrecompra e o valor de 30 para sobrevenda. Para o cálculo do RSI é utilizada a seguinte fórmula (Lemos, 2015):

$$RSI = 100 - \frac{100}{1 + RS}$$

Em que:  $RS = \text{Média}(n) \text{ fechamentos de alta (A)} / \text{Média}(n) \text{ fechamentos de baixa}$ .

Para  $n$ , frequentemente é usado o período de 9 a 15 dias.

## 2.5 Análise de redes neurais

As redes neurais artificiais são modelos matemáticos inspirados em um neurônio biológico. Por causa disso, é possível fazer um paralelo entre eles para facilitar o entendimento de ambos.

### 2.5.1. Neurônio biológico

Segundo Santos (2022), o neurônio biológico é dividido em três partes: os dendritos que são terminais pelos quais os neurônios recebem informações das sinapses, essas informações vêm através de substâncias químicas; o corpo celular que é onde os neurônios processam todas as suas informações que chegam nele; e, por fim, o axônio que é a parte mais extensa do neurônio que tem a função de repassar as informações para outros neurônios através de suas terminações. Na Figura 1 temos um exemplo visual de como é a anatomia do neurônio.

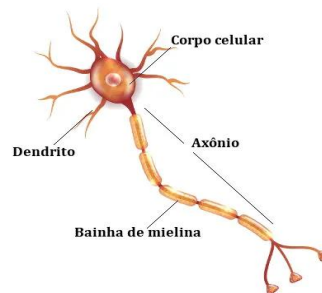


Figura 1 - Anatomia de um neurônio

Fonte: Santos (2022)

### 2.5.2. Rede neural artificial

É possível fazer um paralelo com a estrutura biológica, pois segundo Haykin (2008), uma rede neural artificial possui estruturas parecidas com a citada anteriormente, possuindo entradas de dados, onde cada uma dessas entradas são ponderadas por pesos sinápticos, esses pesos são relacionados ao problema a ser resolvido, assim esses dados seguem para o neurônio de fato, onde chegam na função agregadora (Figura 2).

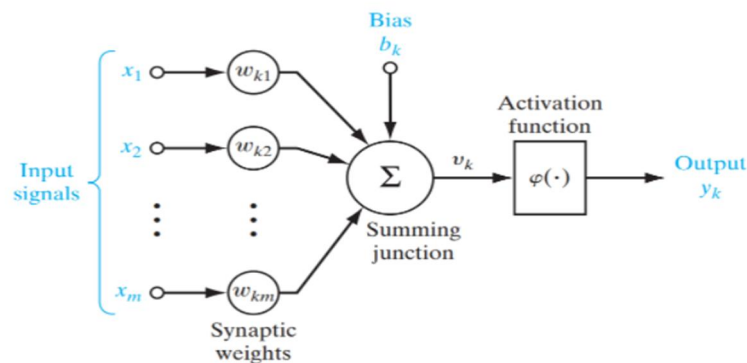


Figura 2 - Esquema de um neurônio artificial

Fonte: Haykin (2008, p.11)

De acordo com a Figura 2, essa função recebe todas as informações que chegam no neurônio, de modo a calcular o produto e a soma dessas informações e o resultado disso é ponderado pelo limiar de ativação, que é uma constante junto a função agregadora. O valor de saída é chamado de potencial de ativação e é dado pelo somatório do produto de cada sinal de entrada e o seu peso de ponderação subtraído pelo limiar de ativação. Esse resultado é enviado para a função de ativação e o seu resultado é uma condição se o sinal deve ser transmitido ou não. (Haykin, 2008).

### 2.5.3. Aplicabilidade de uma rede neural artificial

Para demonstrar o funcionamento de uma rede neural artificial (RNA) é possível usar como exemplo o trabalho de Nielsen (2015), que criou uma rede neural que recebe como entrada uma imagem de um dígito manuscrito e tem como objetivo dizer qual dígito representa aquela imagem. Uma imagem é composta por vários pixels, cada um desses pixels carrega uma informação, quanto mais preto é esse pixel, mais próximo de 0 é seu valor, e o mesmo vale para a cor branca, quanto mais branco for o pixel mais seu valor se aproxima de 1. Cada um desses “ $n \times n$ ” pixels devem ser passados como parâmetro para “ $n \times n$ ” neurônios de entrada da RNA, os quais serão ligados à 10 neurônios de saída correspondentes aos dígitos de 0 a 9 que irão atribuir pesos a cada pixel de entrada de acordo com o dígito (Figura 3). Esses pesos podem ser positivos ou negativos, o que é importante para dígitos que possuem uma semelhança como o 4 e o 9.

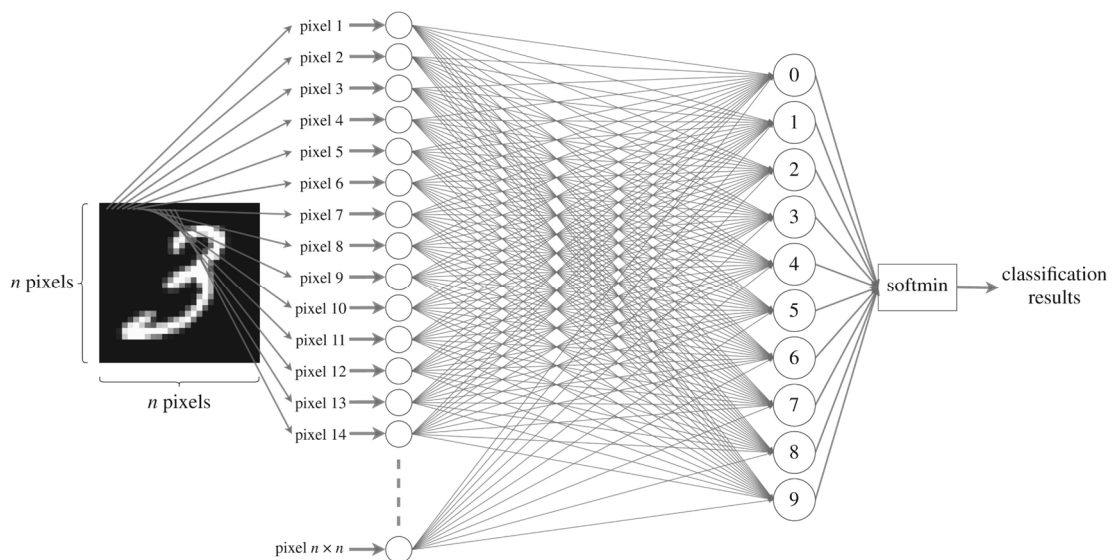


Figura 3 - Diagrama de uma rede neural para reconhecimento de dígitos

Fonte: Nielsen (2015)

Na Figura 3 é possível visualizar como ficam as entradas e também os neurônios. Após cada neurônio receber o valor do pixel e seu peso, ele irá então fazer a multiplicação das entradas com os pesos e depois somar todos os valores entre si e com uma constante. Essa calibração com os pesos é de grande importância para a rede neural para que não aconteça o caso de

número parecidos causem classificação errada. E o resultado irá passar por uma função de normalização que vai transformar o valor em um número entre 0 e 1.

A calibração, também conhecida como treinamento, é outro ponto importante para uma rede neural, pois na inicialização da RNA esses pesos são aleatórios. Para fazer esses ajustes nos pesos, criamos uma função de custo que vai somar todos os erros da rede neural. Em uma primeira tentativa de processar o dígito 3, essa rede neural gera um falso positivo para outro número pois seus pesos vão estar descalibrados, então devemos somar todos os valores dos neurônios que deveriam se aproximar de 0 e ter então o resultado dessa função de custo.

A partir desse resultado a rede neural vai fazer pequenas modificações nos pesos para que possa novamente processar o dígito 3 e verificar se houve um aumento ou diminuição no valor dessa função. Nessa fase de treinamento é importante colocar que a rede neural possui um gabarito, para que possa fazer as comparações e aprimorar seus pesos de acordo com os resultados obtidos da função de custo.

Para não viciar a em apenas um dígito, é importante apresentar várias vezes as imagens de dígitos à entrada da RNA. As imagens devem ser de vários dígitos diferentes para que a RNA consiga fazer um treinamento eficaz. Com uma amostra de 10.000 dígitos, pode-se fazer uma somatória da função custo de todas essas entradas e assim ela possa fazer a calibração dos pesos e verificar se houve uma mudança positiva ou negativa nessa soma.

Depois de repetir este processo várias vezes, deve-se fazer passar para a etapa de teste e avaliar a eficiência dessa rede, de modo que ela acerte os dígitos.

## **2.6 Trabalhos correlatos**

O trabalho “Análise de técnicas de predição financeira utilizando Aprendizagem de Máquina e uma técnica ad-hoc” conforme Menegazzo (2021) serviu como base para esse trabalho. Enquanto o trabalho dele é focado nos indicadores técnicos e na técnica ad-hoc, comparando um modelo que utiliza apenas indicadores, outro que utiliza a técnica ad-hoc e um terceiro modelo oráculo que possui o gabarito das ações para controle. Na análise de Menegazzo (2021) o modelo ad-hoc se mostrou mais eficaz. Deste modo, a diferenciação do presente trabalho em relação ao de Menegazzo (2021) se dá na utilização ou não de indicadores de uma empresa do mercado chinês. Contudo, o experimento de Menegazzo (2021) contribui como subsídio para o presente estudo.

Seguindo na linha de análise do mercado financeiro utilizando inteligência artificial temos o trabalho de “Análise comparativa de técnicas de predição de preços de ações” do autor Pereira (2020), que aborda tanto a parte de indicadores técnicos e também as redes neurais, mas o foco do estudo é em predições de ações em geral da bolsa de valores.

Outro trabalho, com enfoque mais financeiro é o de Mazzochi (2013) que se assemelha na parte de indicadores técnicos para a análise de investimentos na bolsa de valores. Neste trabalho, o autor cria um modelo de predição para ajudar investidores iniciantes ou não, a investir na bolsa de valores, aplicando o modelo a várias ações verificando os ganhos e perdas. Ele utiliza indicadores diferentes e também um pacote de software diferente para construir a rede neural.



### 3. MATERIAIS E MÉTODOS

Visando adquirir conhecimentos e entender como criar os modelos de predição e como identificar momentos de compra e venda de ações, foi realizada uma revisão bibliográfica sobre o tema de indicadores técnicos e redes neurais, descritas na seção 2. Nessa seção são abordados as linguagens e plataformas utilizadas nos modelos, como foi realizada a extração, coleta e manipulação dos dados coletados, a configuração da rede neural e o funcionamento do modelo.

#### 3.1 Linguagem e plataforma utilizada

Os modelos foram desenvolvidos na linguagem *Python*, uma linguagem interpretada, relativamente simples e de fácil compreensão, que vem se popularizando a cada dia, sendo utilizado por engenheiros, matemáticos, cientistas de dados, pesquisadores, não se limitando apenas a programadores.

Junto com o *Python* foi utilizado o *Jupyter Notebook* (NumFocus Foundation<sup>2</sup>, n. d.), uma interface de desenvolvimento de aplicativos em código aberto que permite que você crie e compartilhe documentos com códigos, equações, visualizações e textos explicativos. Além de ser rápido e prático para a criação de protótipos, o *Jupyter* possui uma interface intuitiva.

#### 3.2 Extração, coleta e manipulação de dados

Para desenvolvimento do modelo, foi construído uma estrutura para coleta e extração dos dados das ações, sendo as principais informações o preço atual, histórico de preços de abertura e fechamento, volume de ativos negociados e indicadores técnicos.

Foi utilizada a biblioteca *pandas DataReader*, uma sub biblioteca do *Pandas*, que permite que sejam buscados dados de várias fontes da internet e armazenados localmente em uma estrutura de dados chamada *DataFrame*. De todas as fontes disponíveis, foi utilizada as informações financeiras da fonte *Yahoo! Finances*. Para cada uma das ações, foi realizada a coleta em dois períodos, de 01 de janeiro de 2016 até 31 de dezembro de 2020, e também de 01 de janeiro de 2021 até 31 de dezembro de 2021. Os primeiros quatro anos foram utilizados para treinar a inteligência artificial e o segundo período de um ano foi utilizado para simulações de compra e venda de ações.

Para o armazenamento e manipulação de todas essas informações, foi utilizado o *Pandas*, uma biblioteca de código livre do *Python* para a análise de dados e como ferramenta de manipulação (NumFocus Foundation<sup>1</sup>, n. d.). As informações históricas das ações são coletadas via *pandas DataReader*, e armazenadas em *DataFrames*.

Foi utilizada também a biblioteca *NumPy* (NumFocus Foundation<sup>3</sup>, n. d.) para manipulações de vetores e matrizes, usados no treinamento da rede neural artificial.

#### 3.3 Rede neural artificial

Para a rede neural artificial (RNA) foi utilizado o pacote *Scikit Learn*, que segundo Pedregosa et al. (2011) é um pacote de aprendizado de máquina em *python*, simples e eficiente para análise de dados e predições. Esse pacote é utilizado para resolver diversos problemas como por exemplo classificação, regressão, clusterização, seleção de modelo, e pré-processamento.

### 3.4 Funcionamento dos Modelos

Foram criados 2 modelos de RNA: um com indicadores técnicos da empresa Telefônica Brasil apenas e outro com indicadores técnicos da empresa Telefônica Brasil e indicadores técnicos da China Telecom.

Os indicadores técnicos da China Telecom foram utilizados como dados adicionais para realizar as tomadas de decisões nas simulações de investimento. Foram realizadas simulações de compra e venda, onde sempre foi realizada a compra e venda total do valor disponível.

O valor de caixa inicial estipulado para o início das simulações foi de 100.000 dólares. As comparações são realizadas com o dia anterior, sempre levando em consideração a atenção com os horários de funcionamento de ambas as bolsas, principalmente feriados.

Os dois modelos de RNA foram criados a fim de avaliar se a consideração de um mercado, defasado em relação ao tempo, pode auxiliar na predição de preços de ações de outro mercado. Nessa avaliação, foram verificados o caixa final de cada modelo, ou seja, verificado se o balanço foi positivo, negativo ou neutro, e analisado se realmente existe uma relação do mercado chinês com o mercado brasileiro.

## 4. DESENVOLVIMENTO

Na sequência são delineados maiores detalhes sobre os dois modelos de RNA desenvolvidos no estudo, de modo a tentar demonstrar se as informações de uma China Telecom Corp podem auxiliar a melhorar o modelo de predição de preço da Telefônica Brasil, representado pelo código VIVT3 na B3.

### 4.1 Funcionamento RNA utilizando apenas indicadores da Telefônica Brasil

Para ter uma base de comparação, foi desenvolvido um modelo usando a IDE Jupyter Notebook, que utilizou o valor de fechamento, volume negociado e a variação, todos esses do dia anterior, os indicadores técnicos OBV, RSI e médias móveis para comparação com o modelo que utiliza as informações do mercado chinês. É importante colocar que foram realizadas várias simulações, por questões da inicialização da RNA.

Utilizando o *pandas DataFrame*, foi realizada a extração dos dados da ação VIVT3 para dois *DataFrames* distintos, o primeiro contemplava os valores da bolsa no período de janeiro de 2016 até dezembro de 2020 e o segundo os valores entre janeiro de 2021 até dezembro de 2021.

Após a extração dos dados, foi criada uma função para calcular o indicador técnico RSI, pois dos três utilizados, é o único que possuía uma complexidade onde apenas com as funções disponibilizadas pelo *pandas* não era suficiente.

Em ambos os *DataFrames*, foram criados alguns campos novos para cada linha de modo a auxiliar nas tomadas de decisões da rede neural. São esses o valor de fechamento anterior, a variação que é a diferença entre o preço de fechamento e o de abertura para um dia, a variação do dia anterior, o volume do dia anterior e também a decisão que é se a variação for positiva é retornado o valor 1 e se a variação for negativa é retornado -1, se não houver variação é retornado o valor 0.

Além desses campos novos, foi criado um novo campo para cada um dos três indicadores, dado as duas devidas entradas. Para o RSI, foi utilizado o valor de abertura do dia e também o valor 14, que é o período que o indicador vai utilizar. As médias móveis também utilizaram o valor da abertura, e os valores de abertura de 14 dias anteriores para chegar no resultado. E o OBV utilizou a data de fechamento anterior, o OBV anterior e também o volume negociado para os cálculos. Tanto o RSI como as médias móveis só fornecem resultados após 14 dias de investimento, pois necessitam dos valores anteriores a essa data para funcionar. O OBV precisa apenas de dois dias para começar a retornar resultados.

Por conta dessas primeiras linhas que não continham algumas informações dos indicadores no *DataFrame*, foi utilizado sempre intervalos de tempo maiores do que o que se propunha, mas que fossem utilizados nos cálculos desses valores iniciais e então esta ‘sobra’ inicial foi descartada.

Na etapa de treinamento, foram repassados os dados de treinamento para a rede neural. Na Figura 4, temos um exemplo das cinco primeiras linhas que foram informadas como parâmetro, temos os campos selecionados, a data, o valor de abertura da ação, e os valores do dia anterior, valor de fechamento, volume e decisão. Também temos os indicadores OBV, SMA e RSI.

	Open	CloseAnt	VolAnt	DecAnt	OBV	Sma14	Rsi14
<b>Date</b>							
<b>2016-01-22</b>	29.950001	29.500000	35900.0	1.0	39300.0	29.450714	44.240711
<b>2016-01-26</b>	30.100000	30.400000	12600.0	1.0	28300.0	29.425000	45.999736
<b>2016-01-27</b>	29.809999	29.629999	11000.0	-1.0	82800.0	29.410000	43.164584
<b>2016-01-28</b>	30.120001	30.500000	54500.0	1.0	67900.0	29.425714	46.930049
<b>2016-01-29</b>	29.830000	29.650000	14900.0	-1.0	78000.0	29.397857	43.993682

Figura 4 - Exemplo das linhas iniciais repassadas a rede neural

Na etapa de simulações, foi criado um algoritmo para realizar as compras e vendas, este era bem simples, o caixa inicial era de 100.000 dólares, quando o valor da ação predito pela rede neural era maior que da abertura e tinha valor em caixa, era feita a compra, a ação de compra era realizada fazendo o valor de caixa dividido pelo valor da ação. Caso o valor predito fosse menor que o valor de abertura e tivesse algum valor investido, ele fazia a venda. Na operação de venda era feito o valor de ações compradas vezes o valor de abertura, e o resultado era atribuído no caixa. Caso não fosse nenhum desses dois casos, ele mantinha o dinheiro em caixa ou o valor investido.

## 4.2 Funcionamento do modelo de RNA com base na China Telecom

Foi criado também o modelo que utiliza dados do mercado chinês, seguindo um padrão bem parecido com o anterior. Foi feita a extração de dados da ação VIVT3 nos dois períodos de 2016 até 2020 e também de 2021 para dois *dataframes* distintos, mas também foi realizada a extração dos dados da ação da China Telecom Corp Ltd, representado pelo código 0728.HK para dois *DataFrames* diferentes.

Após a extração dos dados, partimos para a fase de tratamentos de dados. Assim como no outro modelo, foram criados campos novos para todos os *dataframes* que são o valor de fechamento do dia anterior, a variação, a variação do dia anterior, o volume do dia anterior e também a decisão. Também foi feito o cálculo dos três indicadores técnicos, OBV, RSI e SMA para todos os *DataFrames*.

Ainda no tratamento de dados, nesse modelo foi necessário incluir os dados extraídos das ações da bolsa 0728.HK para o da bolsa VIVT3. Como as bolsas estão situadas em países diferentes, cada país possui seus feriados e dias que as bolsas não abrem, o que necessitou ser feito um mapeamento entre essas datas.

Para contornar o problema, foi criado um *DataFrame* apenas com as datas que houve funcionamento da bolsa B3 e após isso foi realizada uma consulta de conjunto de dados *outer join left*, que na biblioteca *pandas* do *python* é feito utilizando a função *merge*, passando como primeiro parâmetro o *DataFrame* que será a tabela a esquerda, em seguida o *DataFrame* direito, a operação *left*, para sinalizar qual lado será o mandatório e, por fim, a chave dos *DataFrames*, onde foi utilizada a data. Essa operação foi aplicada em ambos os *DataFrames* que continham os valores da bolsa 0728.HK, assim removendo os dias que houve operação na bolsa da China, mas não houve na B3.

Algumas das linhas desses *DataFrames* ficaram sem informações pois algumas datas ocorreram na bolsa B3 mas não houveram na bolsa da HKEX. A solução desse problema foi bem simples, a biblioteca *pandas* possui uma função chamada *ffill* que identifica as linhas que possuem informações vazias e atribui o valor da linha anterior para preencher essas lacunas. Foi aplicado essa função em ambas os *DataFrames* para corrigir essas linhas vazias.

Dado os *DataFrames* com informações da bolsa chinesa preparados, foram criados os campos novos nos *DataFrames* com os dados da bolsa da B3 com os dados da bolsa da China, sendo esses a data de abertura e fechamento, o volume e decisão e também os três indicadores OBV, RSI e SMA.

Passando para a etapa de treinamento, foram repassados os dados de treinamento com as informações adicionais do mercado chinês para a rede neural. Diferente do modelo anterior, esse então contempla campos adicionais de valor de abertura e fechamento, decisão e volume da ação chinesa e também os indicadores os indicadores OBV, SMA e RSI (Figura 6).

Date	Open	CloseAnt	VolAnt	DecAnt	OBV	Sma14	Rsi14	ChineseOpen	ChineseClose	ChineseVolume	ChineseDecision	ChineseOBV	ChineseSm
2016-01-04	32.520000	NaN	NaN	NaN	0.0	32.520	NaN	3.60	3.51	32810565.0	-1.0	0.0	3.600
2016-01-05	30.459999	30.350000	37400.0	-1.0	36100.0	31.490	NaN	3.50	3.51	47503523.0	0.0	0.0	3.550
2016-01-06	30.020000	30.400000	36100.0	0.0	43800.0	31.000	NaN	3.50	3.49	37401888.0	0.0	-37401888.0	3.533
2016-01-07	29.900000	30.450001	7700.0	1.0	-1500.0	30.725	NaN	3.49	3.49	0.0	0.0	-37401888.0	3.522
2016-01-08	30.219999	30.440001	45300.0	1.0	-26100.0	30.624	NaN	3.40	3.46	58165495.0	1.0	-95567383.0	3.498

Figura 6 - Exemplo das linhas iniciais com dados da ação 0728.HK repassadas à rede neural

A etapa de simulações segue o mesmo algoritmo criado para o modelo anterior para a compra e vendas de ação utilizando o mesmo caixa inicial de 100.000 dólares. Como a rede neural que cria suas próprias regras para realizar as predições, os valores da bolsa chinesa não são utilizados no algoritmo.

## 5. RESULTADOS E DISCUSSÃO

Foram realizadas nove simulações de ambos os modelos, e foi verificado nos dois modelos qual teve o maior ganho. Para facilitar a discussão, vamos chamar o modelo que utiliza apenas indicadores de modelo 1 e o modelo que possui os dados da ação chinesa será chamado de modelo 2. Cada uma das execuções fez a predição de 233 dias.

Simulações	Modelo 1	Modelo 2
1	0,0%	0,0%
2	0,0%	0,0%
3	15,63%	4,65%
4	0,0%	12,48%
5	0,0%	-0,15%
6	0,0%	8,73%
7	0,0%	0,0%
8	2,97%	17,18%
9	0,0%	-3,17%

Tabela 1 - Ganhos por simulação de ambos os modelos avaliados

Na tabela 1 é possível visualizar os ganhos de cada um dos modelos em cada simulação, o ganho é a quantidade de dinheiro ganho menos o caixa inicial, dividido pela caixa inicial vezes cem. O modelo 1 conseguiu apenas dois modelos com lucro e o restante não obteve lucro nem perda, enquanto que o modelo 2 teve quatro modelos que conseguiram resultados positivos, dois com resultados negativos e o restante não obteve lucro nem perda.

De todas as simulações realizadas, a simulação 3 obteve o melhor resultado com 15,63% de ganho no modelo 1, enquanto que a simulação 8 obteve o melhor resultado no modelo 2 com 17,18% de ganho, sendo melhor que o modelo 1.

Simulação	Modelo 1		Modelo 2	
	Acertos	Erro	Acertos	Erro
1	152	81	152	81
2	79	154	87	146
3	149	84	145	88
4	152	81	77	156
5	152	81	75	158
6	152	81	146	87
7	152	81	152	81
8	151	82	68	165
9	48	185	153	80

Tabela 2 - Total de acertos e erros por simulação de ambos os modelos avaliados

Na tabela 2 conseguimos ver a quantidade de acertos e erros de cada dia de investimento em cada modelo dado a simulação utilizada. Para que seja contabilizado um acerto no dia, é verificado se a decisão de compra ou venda predita pela rede neural é a mesma decisão que foi definida no *DataFrame* que possui as informações do ano de 2021. Caso contrário, é contabilizado um erro para aquele dia.

No modelo 1, a simulação 3 foi a que conseguiu o melhor resultado, mas não foi a que conseguiu a maior taxa de acertos. O mesmo ocorre no modelo 2, onde a simulação 8 foi a que conseguiu o melhor resultado, mas foi a que teve maior quantidade de erros dentre as simulações. Ainda no modelo 2, conseguimos ver que a simulação 9 foi a que teve mais acerto, porém, foi a que teve o pior resultado com -3,17% de ganho.

## 6. CONSIDERAÇÕES FINAIS

Os investidores buscam o lucro com o menor risco de perda possível nos investimentos da bolsa de valores. Neste cenário, o presente trabalho permitiu demonstrar que redes neurais artificiais (RNA) e os indicadores técnicos, aliado às informações de bolsas de outros países, podem ser uma ferramenta para investidores que buscam uma direção ou uma estratégia de investimento, conforme os resultados obtidos por ambos os modelos.

Tal resultado foi possível por meio do levantamento bibliográfico para identificar e escolher os indicadores técnicos para aplicar nos modelos. Dentre os indicadores analisados foram escolhidos o *relative strength index* (RSI), *on-balance volume* (OBV) e o *simples moving average* (SMA).

Para a etapa de criação dos modelos de RNA, foi considerada a implementação base do trabalho de Menegazzo (2021)”, modificando-o para buscar os dados das ações desejadas, aplicando os indicadores selecionados no estudo e também fazendo as devidas alterações para incluir os dados das ações do mercado chinês.

Por fim, pode-se inferir que o objetivo principal deste trabalho foi alcançado, ao ser identificado que existe uma relação entre o mercado defasado no tempo da China com a bolsa de valores da B3. Enquanto o modelo que utiliza apenas indicadores técnicos conseguiu seu melhor resultado sendo 15,63% de ganho, o modelo que possui as informações das ações da China, conseguiu um resultado de 17,18%. Comparando também com as outras simulações executadas, apesar de ter dois resultados negativos, o modelo conseguiu melhores resultados gerais se comparado com o primeiro.

Colocando ao lado de outros investimentos mais tradicionais, como a poupança por exemplo, quando esse estudo foi iniciado, em 2021, a poupança estava com um rendimento negativo de 6,37% ao ano segundo Castro (2022), e atualmente esse número está em 6,17% ao ano segundo Alvarenga (2022), é notável para um investidor mais agressivo ou até mesmo um moderado que esse resultado de 17,18% ao ano é um valor interessante.

Contudo, é importante ressaltar que o mercado financeiro é relativamente imprevisível, mesmo em um ambiente controlado foi difícil de simular um investidor se aventurando nesse universo do mercado financeiro. Os anos analisados também foram bem anormais para os padrões do mercado, em 2020 tivemos o início da pandemia da Covid19, o que tornou as ações ainda mais

voláteis, com diversas quedas durante o ano. E essas anomalias se repetiram durante 2021 com as novas ondas dessa doença.

Dados os obstáculos e indagações geradas durante o desenvolvimento do trabalho tem-se como possibilidade de trabalhos futuros: realização de experimentos em outras ações da China, ou mais ações simultâneas com períodos históricos maiores; realização de experimentos com outras configurações de redes neurais; implementação do modelo aplicado a outros mercados defasados no tempo como o Japão; utilizar algoritmo genético para buscar melhorias nos parâmetros da rede neural.

## REFERÊNCIAS

About HKEX. (n. d.). HKEX. Disponível em: [https://www.hkexgroup.com/about-hkex/about-hkex?sc\\_lang=en](https://www.hkexgroup.com/about-hkex/about-hkex?sc_lang=en).

Alvarenga, D. (2021). *Com inflação de 10,06% em 2021, poupança tem pior rentabilidade desde 1990*. G1, Grupo Globo. Disponível em: <https://g1.globo.com/economia>.

Alvarenga, D. (2022). *Selic a 13,25%: quanto rendem os investimentos e onde colocar o dinheiro*. G1, Grupo Globo. Disponível em: <https://g1.globo.com/economia>.

B3. (n.d.). *Ações*. Disponível em: <https://www.b3.com.br>

Banco Central do Brasil. (n. d.). *Taxa Selic*. (Disponível em: <https://www.bcb.gov.br>.)

*Bovespa alcança 100 mil pontos pela 1ª vez, perde patamar, mas bate novo recorde*. (2019).G1, Grupo Globo. Disponível em: <https://g1.globo.com/economia/noticia/2019/03/18/bovespa-cotacao-18-03-2019.ghtml>.

Castro, F. de. (2020). *Poupança: Quem deixou R\$ 1.000 na caderneta perdeu R\$ 63,70 em 2021*, UOL *Economia*. Disponível em: <https://economia.uol.com.br>.

Gong, X., Cortese, C. (2017). A socialist market economy with Chinese characteristics: The accounting annual report of China Mobile. *Accounting Forum*, Volume 41, Issue 3, September, Pages 206-220.

Haykin, S. S. *Redes Neurais: Princípios e Prática*. 2. ed. Porto Alegre: Bookman, 2008.

Lemos, F. (2015). *Análise Técnica dos Mercados Financeiros*. São Paulo: Saraiva Educação.

Mazzochi, R. do C. (2020). *Análise de investimentos na bolsa de valores*. Trabalho de conclusão de curso, graduação em Ciência da computação, Centro Universitário Unifacvest, Lages.

Menegazzo, A. (2020). *Análise de técnicas de previsão financeira utilizando Aprendizagem de Máquina e uma técnica ad-hoc*.51 f. Trabalho de conclusão de curso, graduação em Informática, Universidade Estadual de Maringá, Maringá.

Murphy, J. J. (2021). *Análise técnica do mercado financeiro: Um guia abrangente de aplicações e métodos de negociação*. Rio de Janeiro: Alta Books.

Nielsen, M. A. (2015). Using neural nets to recognize handwritten digits. Em: *Neural Networks and Deep Learning*, Determination Press.

NumFocus Foundation<sub>1</sub>. (n. d.). About pandas. *Pandas*. Disponível em: <https://pandas.pydata.org/about/>.

NumFocus Foundation<sub>2</sub>. (n. d.). About Us: Project Jupyter's origins and governance. *Project Jupyter*. Disponível em: <https://jupyter.org/about>.

NumFocus Foundation<sub>3</sub>. About us. *NumPy project and Community*. Disponível em: <https://numpy.org/about/>.

Overview<sub>1</sub>. (n. d.). SSE. Disponível em: <http://english.sse.com.cn/aboutsse/overview/>.

Overview<sub>2</sub>. (n. d.). Disponível em: <https://www.szse.cn/English/about/overview/index.html>.

Pedregosa et al. (2011). *Scikit-learn: Machine Learning in Python*, JMLR 12, pp. 2825-2830, 2011.

Python Software Foundation. (n. d.). About Python. *Python*. Disponível em: <https://www.python.org/about/>.

Santos, V. S. dos. (n.d.). O que é neurônio? *Brasil Escola*. Disponível em: <https://brasilecola.uol.com.br/o-que-e/biologia/o-que-e-neuronio.htm>.