

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE COMPUTAÇÃO
CURSO DE CIÊNCIA DA COMPUTAÇÃO

TALYSON RODRIGUES DE MELLO

**COMPARATIVO ENTRE REDES NEURAIS RECORRENTES GRU E
LSTM PARA A PREDIÇÃO DE INSTRUMENTOS FINANCEIROS**

TRABALHO DE CONCLUSÃO DE CURSO

MEDIANEIRA

2021

TALYSON RODRIGUES DE MELLO

**COMPARATIVO ENTRE REDES NEURAS RECORRENTES GRU E
LSTM PARA A PREDIÇÃO DE INSTRUMENTOS FINANCEIROS**

Trabalho de Conclusão de Curso apresentado ao Departamento Acadêmico de Computação da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Bacharel em Ciência da Computação”.

Orientador: Prof. Dr. Paulo Lopes Menezes

MEDIANEIRA

2021



TERMO DE APROVAÇÃO

COMPARATIVO ENTRE REDES NEURAIS RECORRENTES GRU E LSTM PARA A PREDIÇÃO DE INSTRUMENTOS FINANCEIROS

Por

TALYSON RODRIGUES DE MELLO

Este Trabalho de Conclusão de Curso foi apresentado às 10:00h do dia 4 de maio de 2021 como requisito parcial para a obtenção do título de Bacharel no Curso de Ciência da Computação, da Universidade Tecnológica Federal do Paraná, Câmpus Medianeira. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Prof. Paulo Lopes de Menezes
UTFPR - Câmpus Medianeira

Prof. Cesar Angonese
UTFPR - Câmpus Medianeira

Prof. Nelson Miguek Betzek
UTFPR - Câmpus Medianeira

A folha de aprovação assinada encontra-se na Coordenação do Curso.

RESUMO

MELLO, Talyson Rorigues de. COMPARATIVO ENTRE REDES NEURAIAS RECORRENTES GRU E LSTM PARA A PREDIÇÃO DE INSTRUMENTOS FINANCEIROS. 65 f. Trabalho de Conclusão de Curso – Curso de Ciência da Computação, Universidade Tecnológica Federal do Paraná. Medianeira, 2021.

Este trabalho visa comparar dois modelos de redes neurais recorrentes, para a predição de cotação de instrumentos financeiros, considerando critérios como coeficiente de determinação e coeficiente de correlação dos dados previstos versus dados reais. Se realizaram a análise comparativa do modelo LSTM (Long short-term memory) e GRU (Gated Recurrent Unit) baseados em uma rede neural recorrente. O trabalho fez a comparação dos dois modelos em uma arquitetura idêntica de rede neural recorrente para ambos os modelos. Foram utilizados algoritmos baseados em Python que usam as bibliotecas Keras, Numpy Pandas e Scikit-learn. A cotação do Índice Ibovespa para o período de 28 de abril de 1993 a 01 de abril de 2021, discriminadas pela cotação diária, serviu como dado de entrada. Os resultados da pesquisa apontaram um desempenho melhor do modelo GRU, o qual utilizando os 200 fechamentos anteriores do Índice Ibovespa como previsores, 300 épocas para treinamento e utilizando a função de ativação *Relu* conseguiu explicar cerca de 98% das variações dos dados reais.

Palavras-chave: mercado financeiro, previsão de séries temporais, redes neurais recorrentes

ABSTRACT

MELLO, Talyson Rorigues de. COMPARISON BETWEEN RECURRENT NEURAL NETWORKS GRU AND LSTM FOR THE PREDICTION OF FINANCIAL INSTRUMENTS. 65 f. Trabalho de Conclusão de Curso – Curso de Ciência da Computação, Universidade Tecnológica Federal do Paraná. Medianeira, 2021.

This work aims to compare two models of recurrent neural networks, for the prediction of quotation of financial instruments, considering criteria such as coefficient of determination and correlation coefficient of predicted versus actual data. Comparative analysis of the LSTM (Long short-term memory) and GRU (Gated Recurrent Unit) model was performed based on a recurrent neural network. The work compared the two models in an identical recurrent neural network architecture for both models. Python-based algorithms using the Keras, Numpy Pandas and Scikit-learn libraries were used. The quotation of the Ibovespa Index for the period from April 28, 1993 to April 1, 2021, broken down by the daily quotation, served as input. The research results showed a better performance of the GRU model, which using the 200 previous closings of the Ibovespa Index as predictors, 300 times for training and using the Relu activation function was able to explain about 98% of the variations of the real data.

Keywords: financial market, recurrent neural networks, time series prediction

LISTA DE FIGURAS

FIGURA 1	– Representação em diagrama de blocos do sistema neural	13
FIGURA 2	– Representação do neurônio biológico	14
FIGURA 3	– Função limiar	15
FIGURA 4	– Ilustração do modelo limiar	16
FIGURA 5	– Exemplo de rede <i>feed-forward</i> de camada única	18
FIGURA 6	– Exemplo de rede <i>feed-forward</i> de camadas múltiplas	19
FIGURA 7	– Rede neural recorrente sem auto realimentação e sem neurônios ocultos ..	20
FIGURA 8	– Rede neural recorrente com auto realimentação e com neurônios ocultos .	21
FIGURA 9	– Célula de memória Long short-term memory	22
FIGURA 10	– Portão de esquecimento de uma célula de memória Long short-term memory	22
FIGURA 11	– Portão de entrada de uma célula de memória Long short-term memory ...	23
FIGURA 12	– Atualização da célula de memória Long short-term memory	23
FIGURA 13	– Portão de saída de uma célula de memória Long short-term memory	24
FIGURA 14	– Célula de memória Gated Recurrent Unit	24
FIGURA 15	– Exemplo das três fases primárias da tendência de alta	36
FIGURA 16	– Exemplo das três fases primárias da tendência de baixa	37
FIGURA 17	– Exemplo movimento lateral como uma linha.	37
FIGURA 18	– Exemplo de tendência de alta, baixa e lateral.	39
FIGURA 19	– Suportes e resistências para tendência de alta, baixa e lateral.	39
FIGURA 20	– Página Yahoo Finanças de onde foram obtido os dados.	42
FIGURA 21	– Arquitetura do modelo para comparação entre Long short-term memory e Gated Recurrent Unit.	47
FIGURA 22	– Gráfico do teste de normalidade Kolmogorov-Smirnov para o conjunto de dados utilizado.	50
FIGURA 23	– Histograma atributo Close.	52
FIGURA 24	– Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 200 dados anteriores, 300 épocas de treinamento e com a função de ativação <i>Relu</i>	57
FIGURA 25	– Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 200 dados anteriores, 600 épocas de treinamento e com a função de ativação <i>Relu</i>	57
FIGURA 26	– Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 200 dados anteriores, 900 épocas de treinamento e com a função de ativação Linear.	58
FIGURA 27	– Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 100 dados anteriores, 300 épocas de treinamento e com a função de ativação <i>Selu</i>	58
FIGURA 28	– Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 100 dados anteriores, 900 épocas de treinamento e com a função de ativação <i>Relu</i>	59

LISTA DE TABELAS

TABELA 1	– Formato padrão dos dados do Yahoo Finanças.	43
TABELA 2	– Formato adequado para a análise dos dados.	43
TABELA 3	– Quantidade de dados e épocas utilizados nos testes.	48
TABELA 4	– Parâmetros utilizado para os testes com as funções de ativação Linear, Relu, Selu e Sigmóide	48
TABELA 5	– Estatística descritiva de cada atributo do conjunto de dados.	51
TABELA 6	– Testes que passaram por underfit.	53
TABELA 7	– Resultados Gated Recurrent Unit e Long short-term memory para função de ativação Linear ordenados de forma decrescente pelo R^2	53
TABELA 8	– Resultados Gated Recurrent Unit e Long short-term memory para função de ativação Relu ordenados de forma decrescente pelo R^2	54
TABELA 9	– Resultados Gated Recurrent Unit e Long short-term memory para função de ativação Selu ordenados de forma decrescente pelo R^2	55
TABELA 10	– Resultados Gated Recurrent Unit e Long short-term memory para função de ativação Sigmóide ordenados de forma decrescente pelo R^2	56
TABELA 11	– Cinco melhores resultados obtidos nos testes ordenados de forma decrescente pelo R^2	56
TABELA 12	– Testes de Kruskal-Wallis para homogeneidade e semelhança de entre os dados reais de fechamento do Ibovespa e os previstos pelos cinco melhores modelos testados.	59

LISTA DE SIGLAS

B3	Brasil, Bolsa, Balcão
BOT	Diminutivo de robot
CETIP	Central de Custódia e de Liquidação Financeira de Títulos Privados
CSV	Comma Separated Values
CVM	Comissão de Valores Mobiliários
ETF	Exchange traded fund
GRU	Gated Recurrent Unit
Ibovespa	Índice Bovespa
LSTM	Long short-term memory
MSE	Mean squared error
ON	Ordinária
PN	Preferencial
RNAR	Rede neural artificial recorrente
RNA	Rede neural artificial
RNR	Rede neural recorrente
RPM	Rede Perceptron multicamadas

SUMÁRIO

1 INTRODUÇÃO	9
1.1 OBJETIVO GERAL	10
1.2 OBJETIVOS ESPECÍFICOS	10
1.3 JUSTIFICATIVA	10
2 REFERENCIAL TEÓRICO	12
2.1 REDES NEURAIS ARTIFICIAS	12
2.1.1 Neurônio biológico	13
2.1.2 Neurônio artificial	15
2.1.3 Funções de ativação	16
2.1.4 Arquiteturas de redes neurais artificiais	17
2.1.5 Rede neural feed-forward	18
2.1.6 Rede neural feed-forward multicamadas	18
2.1.7 Rede neural recorrente	19
2.1.8 Long short-term memory	20
2.1.9 Gated Recurrent Unit	23
2.1.10 Processos de aprendizagem	25
2.1.11 Rede perceptron	25
2.1.12 Rede perceptron multicamadas	26
2.2 MERCADO DE CAPITAIS	26
2.2.1 Ações	27
2.2.2 Ações ordinárias	28
2.2.3 Ações preferenciais	29
2.2.4 A bolsa do Brasil	30
2.2.5 Índices do mercado de ações	30
2.2.6 Índice Bovespa	31
2.3 ANÁLISE TÉCNICA	32
2.3.1 Análise Técnica versus Análise Fundamentalista	33
2.3.2 Princípios da teoria de Dow	34
2.3.3 Conceitos básicos de tendência	38
3 MATERIAIS E MÉTODOS	41
3.1 MATERIAIS	41
3.1.1 Conjunto de dados	41
3.1.2 Formato dos dados	42
3.1.3 Linguagem de programação Python	43
3.1.4 Biblioteca Keras	43
3.1.5 Biblioteca Numpy	44
3.1.6 Biblioteca Pandas	44
3.1.7 Biblioteca Scikit-learn	45
3.1.8 Ambiente de desenvolvimento e hardware	45
3.2 MÉTODOS	45
3.2.1 Pré-processamento de dados	45

3.2.2	Conjunto de dados de treinamento e validação	46
3.2.3	Análise e comparação dos modelos utilizados	46
3.2.4	Arquitetura do modelo	46
3.2.5	Testes realizados	47
3.2.6	Parâmetros utilizados nos modelos	49
3.2.7	Análise dos resultados	49
4	RESULTADOS E DISCUSSÃO	51
4.1	ANÁLISE EXPLANATÓRIA DO CONJUNTO DE DADOS	51
4.2	DESEMPENHO DOS MODELOS	52
5	CONSIDERAÇÕES FINAIS	60
5.1	CONCLUSÕES	60
5.2	TRABALHOS FUTUROS	61
	REFERÊNCIAS	62

1 INTRODUÇÃO

A tecnologia teve e ainda tem impacto profundo em quase todos os aspectos da sociedade, e investir no mercado financeiro certamente não é exceção. De fato, a tecnologia democratizou o investimento nas últimas décadas e também exerceu pressão significativa sobre as taxas cobradas pelas corretoras.

Anos atrás, se uma pessoa quisesse negociar algum instrumento financeiro, precisaria ligar para sua corretora para realizar uma ordem de venda ou compra. As taxas de corretagem para compra ou venda de ações eram praticamente fixas entre as corretoras e altas devido à falta de informações, falta de alternativas e concorrência entre as corretoras. Os investidores não sabiam como seus investimentos estavam se saindo quando quisessem, precisavam comprar um jornal ou ligar para sua corretora, para saber a cotação atual de seus instrumentos financeiros.

Uma das maiores inovações dos últimos anos foi o advento dos robôs de investimentos (EDWARDS, 2018). Empresas como Betterment, Wealthfront e outras usaram a tecnologia para permitir que estes robôs construam e gerenciem portfólios de clientes usando algoritmos (JI, 2017). Tirar o elemento humano da equação de investimento pode reduzir drasticamente o custo do investimento (PACKIN, 2018).

Atualmente em 2021 os investidores podem pesquisar na Internet para ver qual corretora tem as melhores taxas de corretagem, e comprar ou vender seus instrumentos financeiros a um clique do mouse. Mesmo com todo esse avanço da tecnologia a bolsa de valores no Brasil ainda é pouco conhecida pela população em geral e segundo histórico de perfil dos investidores pessoas físicas de Abril de 2021 (B3, 2021b), em 09/04/2021 tinha 3.561.296 cadastros de pessoas físicas ativos da bolsa de valores, isso representa menos de 2% da população estimada do Brasil, de 212,7 milhões em 2021 (IBGE, 2021).

Com a entrada de novos investidores e especuladores, essas pessoas podem estar se beneficiando com ferramentas que podem auxiliá-las a tomar a melhor decisão na hora de investir ou especular. A utilização de redes neurais artificiais recorrentes (RNAR) na previsão de valores de instrumentos financeiros na bolsa vem como uma ferramenta para isso.

Uma RNAR tem a capacidade de “simular” o funcionamento do cérebro do humano por meio de neurônios artificiais utilizando algoritmos de aprendizado. Aplicando-se esses

algoritmos sobre um conjunto de dados de entrada, após um processamento desses dados e por meio de uma função de ativação, a rede neural retorna os valores calculados.

1.1 OBJETIVO GERAL

Comparar dois tipos de redes neurais recorrentes, na previsão do valor de fechamento diário de algum instrumento financeiro negociado na B3 ¹ (Brasil, Bolsa, Balcão).

1.2 OBJETIVOS ESPECÍFICOS

1. Obter dados históricos do Índice Bovespa (Ibovespa) para utilizar nas RNAR, para fazer as previsões de valores;
2. Utilizar RNAR do tipo Long Short-term Memory e Gated Recurrent Unit para comparação;
3. Fazer previsão do valor de fechamento do Ibovespa utilizando os modelos escolhidos;
4. Analisar e comparar os resultados obtidos pelos modelos.

1.3 JUSTIFICATIVA

A utilização de RNAR pode auxiliar pessoas, na tomada de decisão, no momento de criar uma ordem de venda ou compra de um instrumento financeiro, ou até mesmo ajudar BOTs (Diminutivo de *robot*) investidores a terem resultados otimizados (GURESEN et al., 2011). Se as pessoas entenderem os benefícios e as limitações das RNAR, elas poderão ter um novo conjunto de ferramentas que poderão permitir que tenham investimentos melhores no longo

¹<http://www.b3.com.br/>

prazo ou curto prazo.

2 REFERENCIAL TEÓRICO

2.1 REDES NEURAS ARTIFICIAS

Redes neurais artificiais (RNA) são sistemas computacionais baseado no funcionamento do sistema nervoso central de seres vivos, que têm o processamento de informações de uma forma diferente de sistemas computacionais convencionais, são compostos por neurônios artificiais e interconexões (sinapses artificiais), que processam valores de entradas simulando o funcionamento de redes neurais biológicas (JAIN et al., 1996).

Ao contrário dos sistemas de computação convencionais, que têm instruções fixas para realizar cálculos específicos, a rede neural artificial precisa ser ensinada e treinada para funcionar corretamente. A vantagem de uma RNA é que podem aprender novos padrões de entrada e saída e ajustar os parâmetros do sistema. Em vez disso, os usuários simplesmente fornecem padrões de entrada / saída de amostra apropriados para a rede (HAYKIN, 2007).

As RNA têm aplicabilidade em diversas áreas como: carros autônomos (TIAN et al., 2018), detecção de câncer de pele (ESTEVA et al., 2017), reconhecimento de padrões (LI et al., 2015), previsão de valores (DOUCOURE et al., 2016), entre outras diversas aplicações. Algumas áreas de aplicação citadas por NUNES e SILVA (2018) são:

Aproximador universal de funções que o objetivo é mapear o relacionamento funcional entre variáveis de um sistema a partir de um conjunto de valores conhecido;

Controle de processos que consiste em reconhecer ações de controle que possibilitem o alcance de requisitos de qualidade, eficiência e segurança do processo;

Reconhecimento / classificação de padrões que consiste de associar um padrão de entrada (amostra), com classes previamente definidas, como acontece por exemplo com o reconhecimento de faces;

Agrupamento de dados o objetivo é a identificação de similaridades entre diversos padrões de entrada a fim de conseguir o seu agrupamento ou identificação de sua classe;

Sistemas de previsão o objetivo é estimar valores futuros a partir de series temporais ou a partir de outras medidas observadas.

2.1.1 Neurônio biológico

O sistema neural do corpo humano pode ser visto como um sistema de três estágios: receptores, rede neural e atuadores (HAYKIN, 2007). Os receptores recebem os estímulos internamente ou do mundo externo, depois passam a informação para os neurônios em uma forma de impulsos elétricos. A rede neural então processa as entradas e toma a decisão apropriada das saídas. Finalmente, os atuadores traduzem os impulsos elétricos da rede neural em respostas ao ambiente externo (ARBIB, 2012). A Figura 1 representa a comunicação bidirecional entre os estágios.

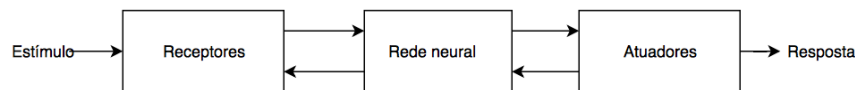


Figura 1 – Representação em diagrama de blocos do sistema neural

Fonte: Adaptado de Haykin (2007)

O elemento fundamental da rede neural é chamado de neurônio. Como ilustrado na Figura 2, um neurônio consiste principalmente de três partes: dendritos, soma e axônio (JAIN et al., 1996). Os dendritos são as estruturas que formam a árvore dendrital que recebe o sinal dos neurônios ao redor, onde cada linha está conectada a um neurônio. O axônio é um cilindro fino que transmite o sinal de um neurônio para outro. No final do axônio, o contato com os dendritos é feito através de uma sinapse. O sinal inter-neuronal na sinapse é impulsos elétricos advindos de reações físico-químicas. Um neurônio dispara um impulso elétrico somente se determinada condição for satisfeita (GURNEY, 2014).

O sinal de impulso de entrada de cada sinapse para o neurônio é excitatório ou inibitório, o que significa que ajuda ou dificulta o disparo do impulso elétrico (KOVÁCS, 2002). A condição de causar o disparo do impulso elétrico é que o sinal excitatório deve exceder o sinal inibitório em certa quantidade em um curto período de tempo, chamado período de soma latente. Quando atribuímos um peso a cada sinal de impulso de entrada, o sinal excitatório tem um peso positivo e o sinal inibitório tem um peso negativo. Dessa forma, podemos dizer que:

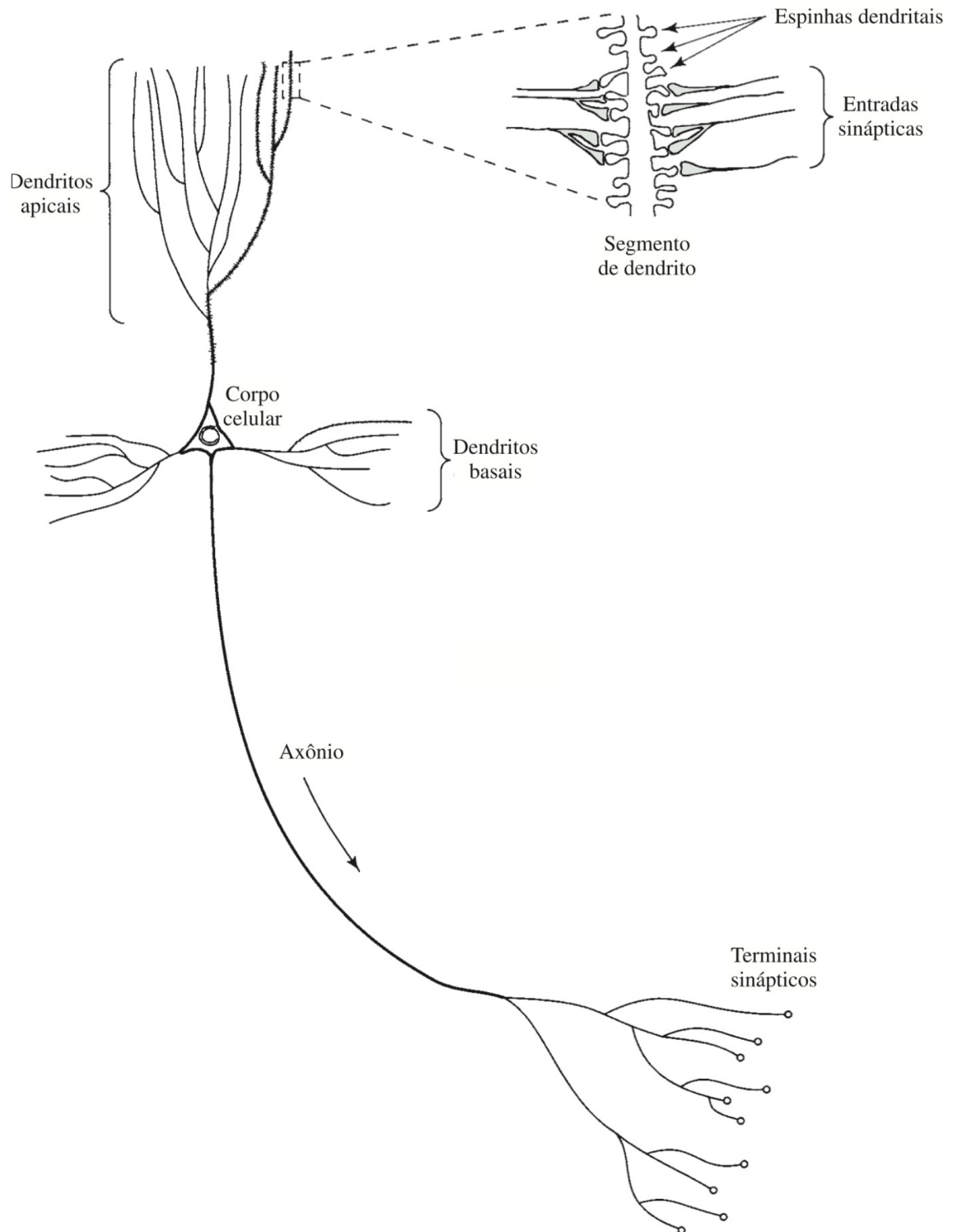


Figura 2 – Representação do neurônio biológico

Fonte: Fonte: (HAYKIN, 2007)

“Um neurônio dispara um impulso elétrico apenas se o peso total das sinapses que recebem impulsos no período de soma latente excede o limite” (ARBIB, 2012).

2.1.2 Neurônio artificial

O modelo inicial de um neurônio artificial foi introduzido por Warren McCulloch e Walter Pitts em 1943 (MCCULLOCH; PITTS, 1943). O modelo neural de McCulloch-Pitts é referido como uma unidade limiar (*threshold*) (KROGH, 2008). É um neurônio de um conjunto de entradas $x_1, x_2, x_3, \dots, x_m$ e uma saída y . O limiar simplesmente classifica o conjunto de entradas em duas classes diferentes. Assim a saída y é binária. Tal função pode ser descrita matematicamente usando estas equações 1 e 2.

$$u = \sum_{i=1}^m x_i w_i \quad (1)$$

$$y = f(u) \quad (2)$$

No qual $w_1, w_2, w_3, \dots, w_m$ são valores de peso normalizados na faixa de $(0,1)$ ou $(-1,1)$ e associado a cada linha de entrada, u é a soma ponderada e T é uma constante de limiar (MCCULLOCH; PITTS, 1943). A função f é uma função de passo linear no limite T como mostrado na Figura 3. A representação simbólica da porta de limiar é mostrada na figura 4.

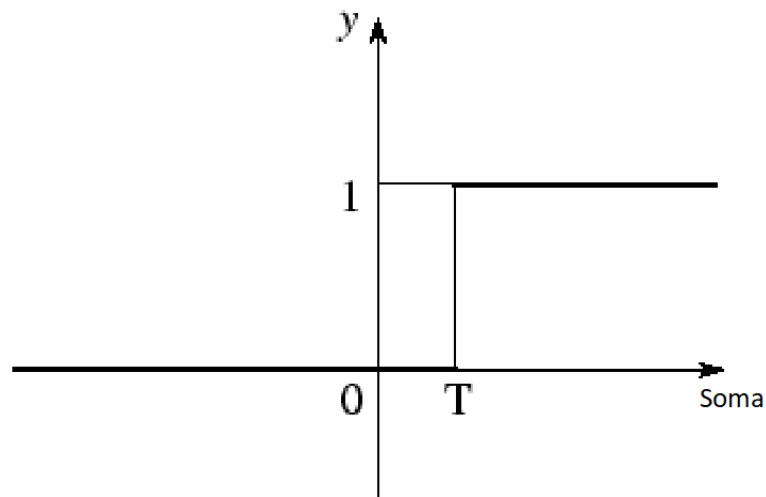


Figura 3 – Função limiar

Fonte: Adaptado de Chakraborty (2016)

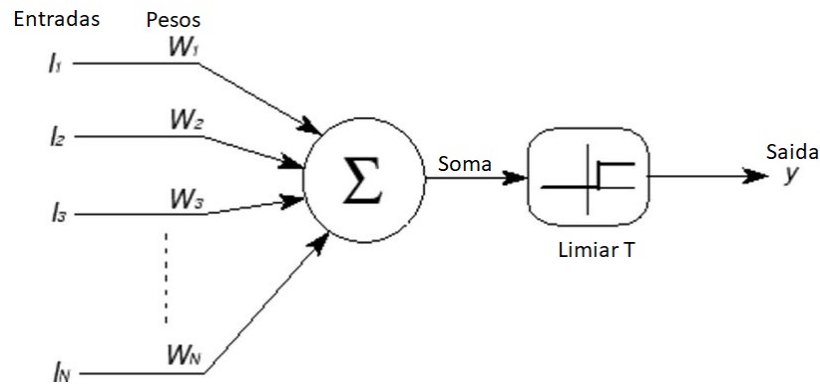


Figura 4 – Ilustração do modelo limiar

Fonte: Adaptado de Chakraborty (2016)

2.1.3 Funções de ativação

Em RNA, a função de ativação de um neurônio define a saída do mesmo dada uma entrada ou conjunto de entradas. Existem funções lineares e não lineares, as funções de ativação não linear permitem que RNA computem problemas não triviais. Comparado com redes neurais biológicas, a função de ativação é geralmente uma abstração que representa a taxa de ativação do potencial de ação na célula. Em sua forma mais simples, essa função é binária, ou seja, o neurônio dispara o impulso elétrico ou não (HAYKIN, 2007; NUNES; SILVA, 2018).

Alguns tipos de funções de ativação que são comumente usadas:

Função limiar A saída é binária, dependendo se a entrada é maior ou menor que algum valor limite.

$$f(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

Função sigmóide A função sigmoide consiste em 2 funções, logísticas e tangenciais . Os valores da função logística variam de 0 e 1 e -1 a +1 para função tangencial.

$$f(x) = \frac{1}{1 + \exp(-x)}$$

Função linear por partes A saída é proporcional à saída total ponderada.

$$f(x) = \begin{cases} 1, & x \geq \frac{1}{2} \\ x, & -\frac{1}{2} < x < \frac{1}{2} \\ 0, & x \leq -\frac{1}{2} \end{cases}$$

Função Relu consiste entre o maior valor entre x e 0

$$f(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } x \geq 0. \end{cases}$$

Função Selu é definida como:

$$f(x) = \begin{cases} \text{se } x > 0 \implies \text{scale} \cdot x \\ \text{se } x < 0 \implies \text{scale} \cdot \text{alpha} \cdot (e^x - 1) \end{cases}$$

Função linear consiste em retornar o próprio x .

$$f(x) = x$$

2.1.4 Arquiteturas de redes neurais artificiais

A arquitetura de uma RNA define a forma como os seus neurônios são arranjados em relação aos outros dessa mesma rede. Os arranjos da RNA são organizados por meio do sentido das conexões sinápticas dos neurônios. A estrutura mais comum de conectar neurônios a uma RNA é por camadas. Segundo NUNES e SILVA (2018) uma RNA pode ser dividida em três partes ou camadas, as quais são nomeadas desta forma:

1. Camada de entrada fornecem informações (dados) que geralmente são normalizadas do mundo externo para a rede. Nenhum processamento é executado em nenhum dos neurônios desta camada, eles apenas transmitem as informações para a próxima camada;
2. Camada oculta não têm conexão direta com o mundo exterior (daí o nome “oculto”). Eles executam cálculos e transferem informações da camada de entrada para a camada de saída, quase todo o processamento interno da rede é realizado nesta camada. Uma coleção de neurônios ocultos forma uma “camada oculta”;
3. Camada de saída é o conjunto de neurônios que são coletivamente chamados de “camada de saída” e são responsáveis por cálculos e transferência de informações da rede para o mundo externo.

As principais arquiteturas de RNAs podem ser divididas em: redes *feed-forward* (alimentação à frente) de camada simples, redes *feed-forward* de camadas múltiplas, redes recorrentes e rede reticuladas (NUNES; SILVA, 2018).

2.1.5 Rede neural feed-forward

A rede neural *feed-forward* foi o primeiro e mais simples tipo de rede neural artificial criado. Ela contém vários neurônios dispostos em uma única camada de neurônios, que é a própria camada de saída (NUNES; SILVA, 2018). Esse tipo de rede é amplamente utilizado para problemas separáveis linearmente, a rede de camada simples não é capaz de classificar conjuntos de dados que não são separáveis linearmente (BOURQUIN et al., 1997). Uma maneira de resolver esse problema é usar a arquitetura de camadas múltiplas. A Figura 5 representa uma rede *feed-forward* de camada simples.

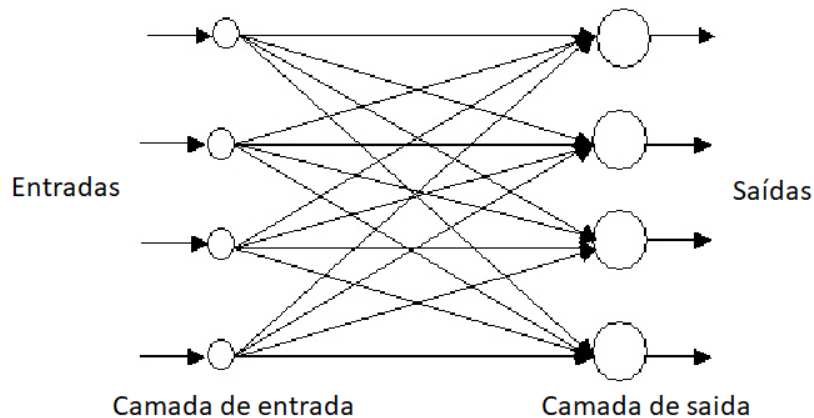


Figura 5 – Exemplo de rede *feed-forward* de camada única

Fonte: Adaptado de Sazli (2006)

2.1.6 Rede neural feed-forward multicamadas

As redes multicamadas resolvem o problema de classificação para conjuntos não lineares, empregando camadas ocultas, cujos neurônios não estão diretamente conectados à saída. As camadas ocultas adicionais podem ser interpretadas geometricamente como hiperplanos adicionais, que aumentam a capacidade de separação da rede. A RNA *feed-forward* multicamadas se distingue da rede de camada única por ter uma ou mais camadas ocultas. Nesta estrutura multicamadas, os neurônios de entrada passam as informações para os neurônios

na primeira camada oculta, depois as saídas da primeira camada oculta são passadas para a próxima camada e assim por diante (NUNES; SILVA, 2018; HAYKIN, 2007). A Figura 6 é uma representação.

A RNA *feed-forward* multicamadas também pode ser vista como um conjunto de redes de camada única. O nível de complexidade de processamento neste tipo de RNA é maior e isso pode ser visto pelo fato de que muitas RNAs de camada única são combinadas nessa RNA multicamada. O projetista de uma RNA deve considerar quantas camadas ocultas são necessárias, dependendo da complexidade do problema a ser resolvido (KARSOLIYA, 2012).

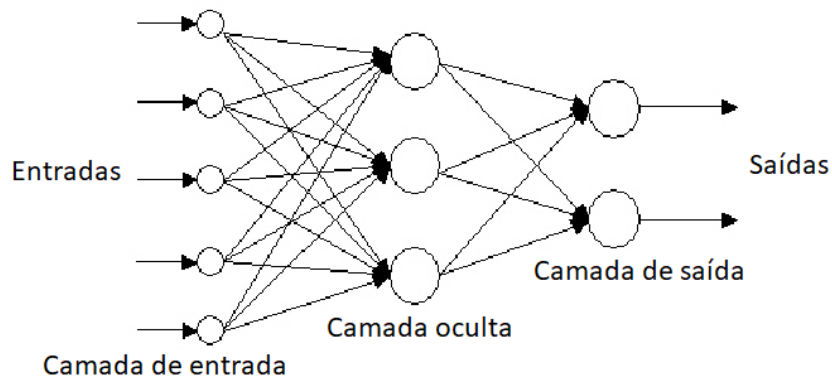


Figura 6 – Exemplo de rede *feed-forward* de camadas múltiplas

Fonte: Adaptado de Sazli (2006)

2.1.7 Rede neural recorrente

Uma rede neural recorrente (RNR) é uma rede que tem a capacidade de armazenar as saídas de um neurônio ou mais em sua memória para uso futuro (NUNES; SILVA, 2018). Na Figura 7 é apresentado um exemplo de arquitetura de uma RNR sem auto realimentação e sem neurônios ocultos, a auto-realimentação acontece quando dados de saída de um neurônios volta para a sua própria entrada e neste exemplo os dados de saída estão indo para entradas de neurônios diferentes da origem da saída. Já na Figura 8 se tem o exemplo de uma RNR com auto realimentação e que possui neurônios ocultos (NUNES; SILVA, 2018; HAYKIN, 2007).

A capacidade de armazenar as saídas de um neurônio de uma RNR traz a capacidade dela de aprender conhecimento sequencial. Por exemplo, para uma certa sequência de entrada,

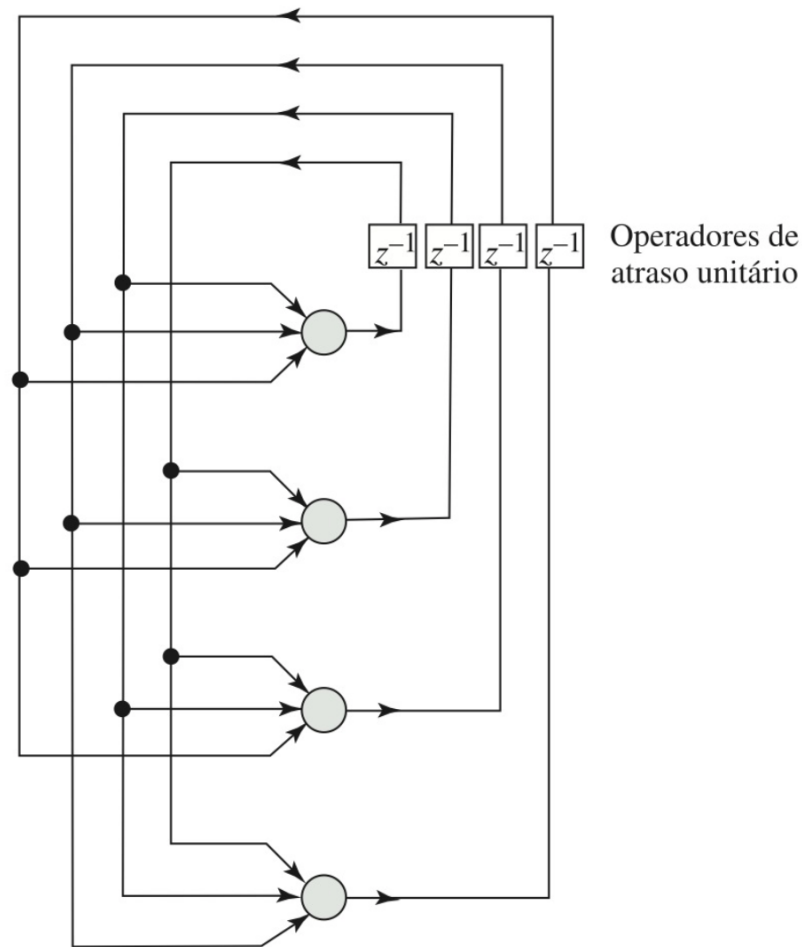


Figura 7 – Rede neural recorrente sem auto realimentação e sem neurônios ocultos

Fonte: Haykin (2007)

podemos usar o RNR para prever a próxima saída (NUNES; SILVA, 2018).

2.1.8 Long short-term memory

Em 1997 Hochreiter e Schmidhuber introduziu o *Long short-term memory* (LSTM) (HOCHREITER; SCHMIDHUBER, 1997), que é um tipo de RNR capaz de aprender dependências de longo prazo nos dados e lidar com o problema do desaparecimento gradiente. (GERS et al., 1999).

A ideia central é uma célula de memória que pode manter seu estado ao longo do tempo, consistindo em uma memória explícita e unidades de passagem que regulam o fluxo de

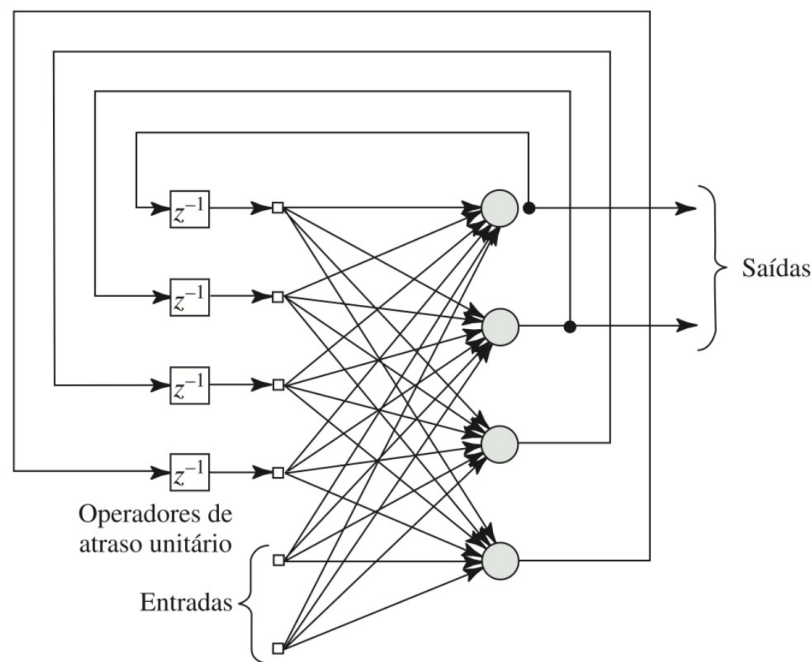


Figura 8 – Rede neural recorrente com auto realimentação e com neurônios ocultos

Fonte: Haykin (2007)

informações para dentro e para fora da memória (COLAH, 2015).

Uma célula de memória representa a memória do LSTM (Figura 9) que sofre mudanças por meio do esquecimento da memória antiga (portão de esquecimento) e adição de novas memória (portão de entrada) (COLAH, 2015).

Portão de esquecimento controla quais informações a célula deve jogar fora da memória, ou seja esquecer (COLAH, 2015) (Figura 10).

Portão de entrada controla quais novas informações são adicionadas ao estado da célula de memória a partir da entrada atual (COLAH, 2015) (Figura 11).

Atualização da célula de memória O vetor de estado da célula agrega os dois componentes (memória antiga por meio do portão de esquecimento e nova memória por meio da portão de entrada) (COLAH, 2015) (Figura 12).

Portão de saída decide condicionalmente o que sai da célula de memória (COLAH, 2015) (Figura 13).

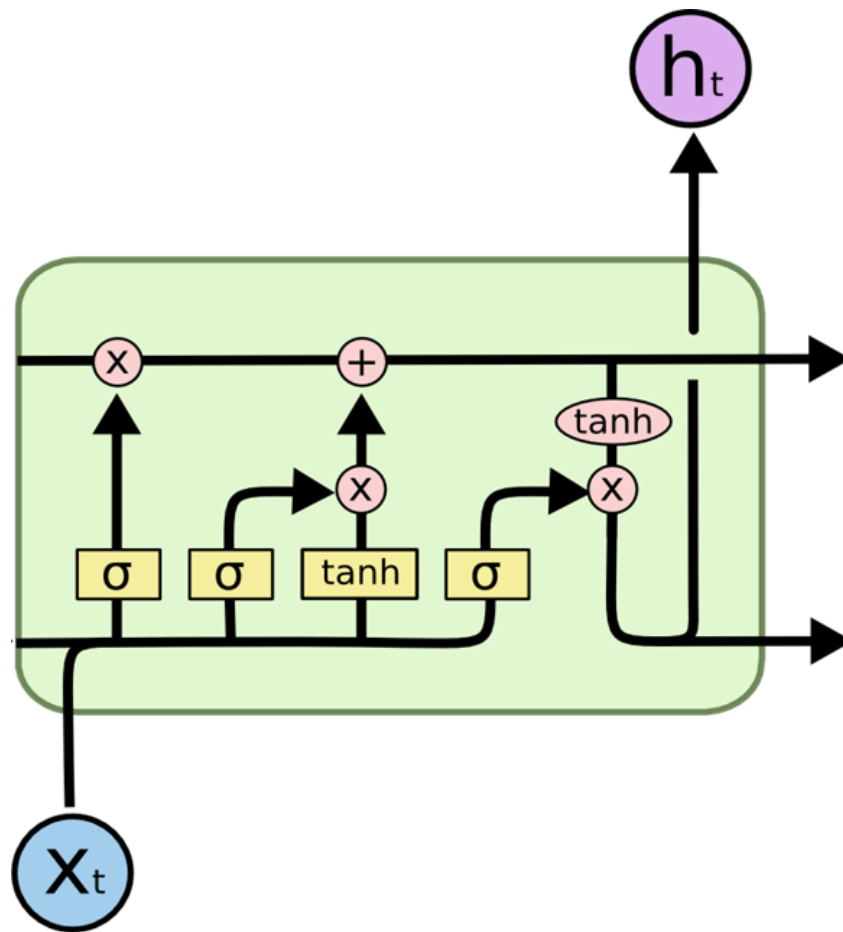


Figura 9 – Célula de memória Long short-term memory

Fonte: Colah (2015)

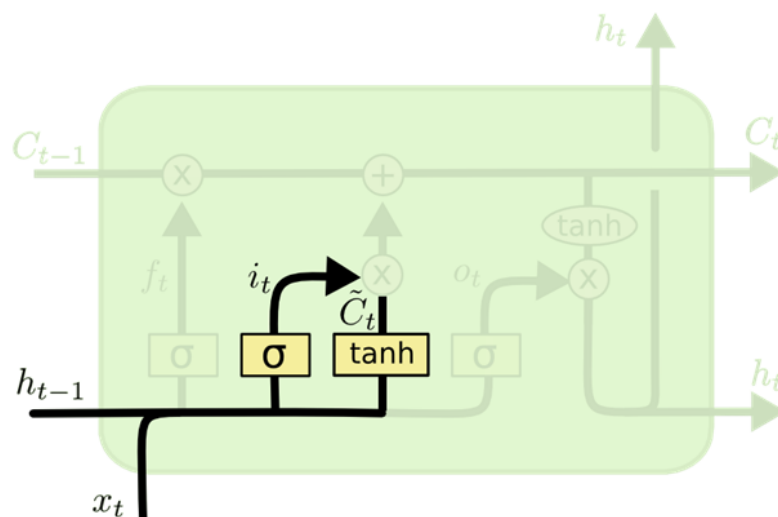


Figura 10 – Portão de esquecimento de uma célula de memória Long short-term memory

Fonte: Colah (2015)

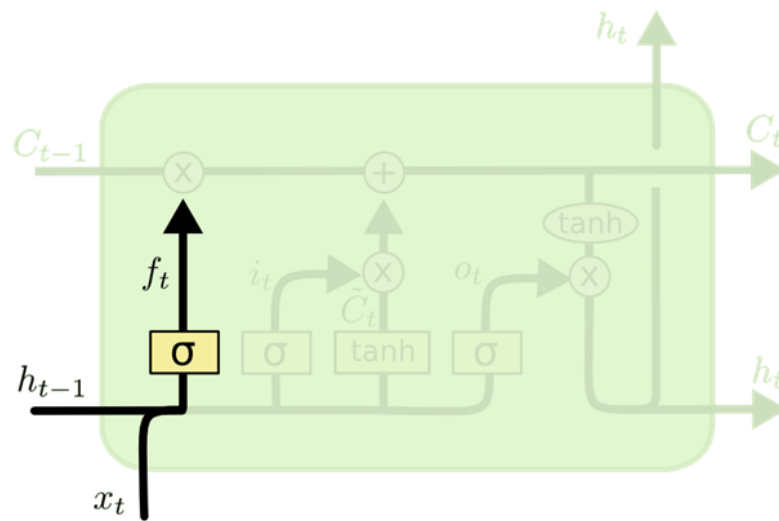


Figura 11 – Portão de entrada de uma célula de memória Long short-term memory

Fonte: Colah (2015)

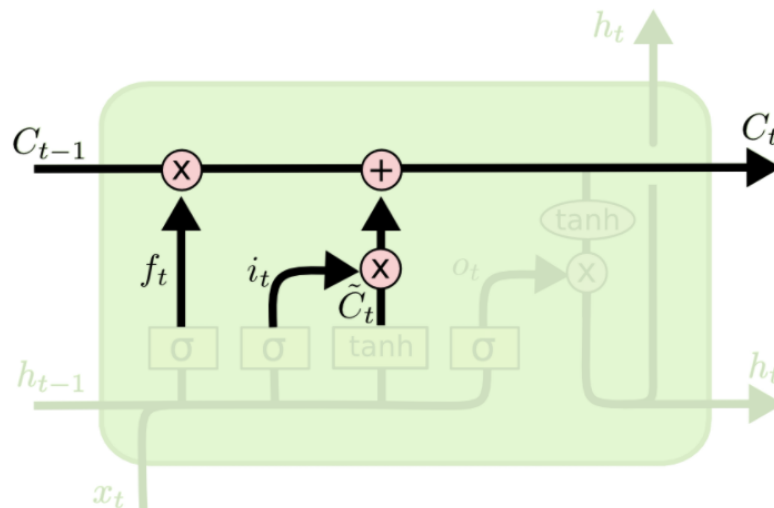


Figura 12 – Atualização da célula de memória Long short-term memory

Fonte: Colah (2015)

2.1.9 Gated Recurrent Unit

O *Gated Recurrent Unit* (GRU) é uma variação do modelo LSTM que foi introduzida por Cho et al. (2014). Essa variação combina os portões de esquecimento e de entrada em um único "portão de atualização". Ele também mescla o estado da célula de memória e faz algumas outras alterações. O modelo resultante é mais simples do que os modelos LSTM padrão e tem

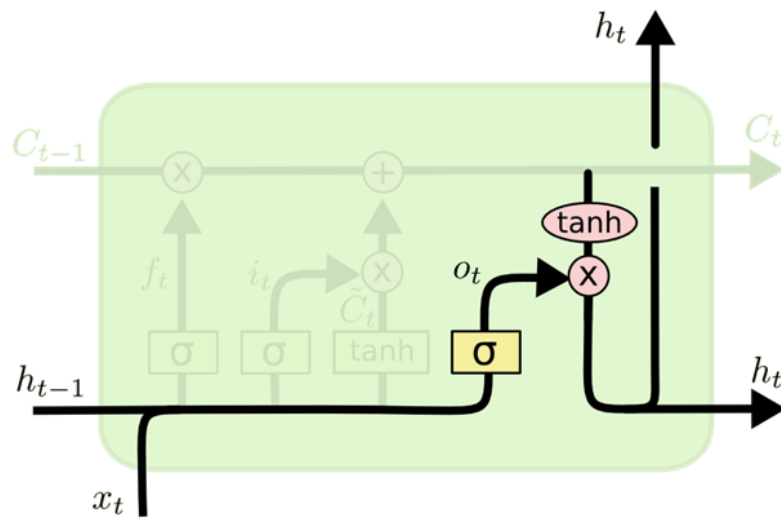


Figura 13 – Portão de saída de uma célula de memória Long short-term memory

Fonte: Colah (2015)

se tornado cada vez mais popular (COLAH, 2015).

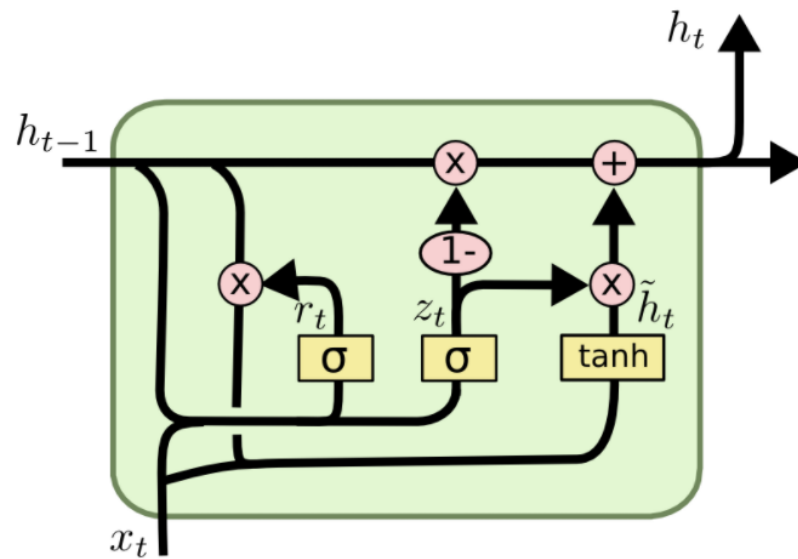


Figura 14 – Célula de memória Gated Recurrent Unit

Fonte: Colah (2015)

2.1.10 Processos de aprendizagem

As RNA se utilizam de algoritmos de aprendizagem, esses algoritmos são formados por um conjunto de regras que permitem que a rede neural aprenda com a entrada de dados que são fornecidos a RNA, assim alterando os pesos sinápticos das conexões, entre os neurônios. Os dados usados para treinar a rede geralmente são fornecidos de duas maneiras, aleatoriamente e sequencialmente. Dependendo de como o aprendizado é realizado pode ser dividido em três tipos, que são: supervisionados, não supervisionados e reforço (NUNES; SILVA, 2018; HAYKIN, 2007; NORVIG; RUSSELL, 2013).

Aprendizagem supervisionada consiste em ter um conjunto de dados de entrada da rede e um conjunto de dados de saída. As saídas calculadas pela rede são então comparadas com um conjunto de dados de saída verdadeiro. Se a qualquer momento houver diferenças entre os dados calculados pela a rede e os dados de saída, a rede será ajustada (NORVIG; RUSSELL, 2013);

Aprendizagem não supervisionada consiste em dar o conjunto de dados de entrada para a rede e a única coisa que é indicada para a rede é se as saídas geradas estão corretas ou incorretas (NORVIG; RUSSELL, 2013);

Aprendizagem por reforço não há nenhum tipo de guia para a rede, dessa forma, a única coisa que a rede pode fazer é reconhecer padrões nos dados de entrada e criar categorias a partir desses padrões. Então após o treinamento, você insere um conjunto de dados na rede e a rede poderá classificá-lo e indicar em qual categoria ele foi classificado (NORVIG; RUSSELL, 2013).

2.1.11 Rede perceptron

No final da década de 1950, Frank Rosenblatt introduziu uma rede composta pelas unidades que foram aprimoradas na versão do modelo de limiar de McCulloch-Pitts (ROSENBLATT, 1958). O modelo de neurônio de Rosenblatt, o perceptron, foi o resultado da fusão entre dois conceitos da década de 1940, o modelo de McCulloch-Pitts de um neurônio artificial é a regra de aprendizado de Hebbian de ajustar pesos (BOSE; LIANG, 1996; HEBB, 2005). Além dos valores de pesos variáveis, o modelo perceptron adicionou uma entrada extra

que representa o Bias. Assim, a equação modificada de $u = \sum_{i=1}^m x_i w_i$ mudou agora para a seguinte:

$$u = \sum_{i=1}^n x_i w_i + b,$$

onde b representa o valor de Bias. (HAYKIN, 2007).

No perceptron o disparo do impulso elétrico são representados como valores numéricos. Nas sinapses entre o dendrito e os axônios, os disparo do impulso elétrico são modulados em várias quantidades. Isso também é modelado no perceptron multiplicando cada valor de entrada por um peso. Um neurônio real dispara um impulso elétrico de saída somente quando a intensidade total dos sinais de entrada excede um certo limite (NUNES; SILVA, 2018).

2.1.12 Rede perceptron multicamadas

Em uma rede perceptron multicamadas (RPM), há uma camada de neurônios de entrada, uma camada de neurônios de saída e uma ou mais camadas intermediárias de neurônios que são conhecidas como camadas oculta ou escondida.(NUNES; SILVA, 2018) Questões importantes no projeto de uma RPM incluem a especificação do número de camadas ocultas e o número de neurônios nessas camadas.(HAYKIN, 2007)

2.2 MERCADO DE CAPITAIS

O mercado de capitais é o local onde títulos e fundos são centralizadas entre entidades que possuem capital e aqueles que precisam de capital (PINHEIRO, 2018). As entidades que possuem capital incluem investidores institucionais e de varejo, enquanto aqueles que buscam capital são empresas e governos. Os mercados de capitais buscam melhorar as eficiências transacionais. Esses mercados trazem aqueles que detêm capital e aqueles que buscam capital juntos e fornecem um lugar onde as entidades podem trocar valores mobiliários (PINHEIRO, 2018).

Os mercados de capitais são compostos de mercados primários e secundários. No mercado primário é onde novos títulos são emitidos e vendidos, e no mercado secundário, onde

os títulos já emitidos são negociados entre os investidores. A maioria dos mercados primários e secundários modernos são plataformas eletrônicas e são operados utilizando computadores (PORTAL DO INVESTIDOR, 2019; PINHEIRO, 2018).

Mercado primário está aberto a investidores que compram títulos diretamente da empresa emissora. Esses títulos são considerados ofertas primárias ou ofertas públicas iniciais (IPOs). Quando uma empresa abre seu capital, ela vende seus títulos e ações para investidores institucionais (BOEHME; ÇOLAK, 2012);

Mercado secundário é um local supervisionado por um órgão regulamentador, como a Comissão de Valores Mobiliários (CVM), onde os títulos existentes ou já emitidos são negociados entre os investidores (SHI et al., 2013). As empresas emissoras não participam do mercado secundário. A B3 e a bolsa de Nova York são exemplos do mercado secundário. Além disso, o mercado secundário serve a uma finalidade importante no mercado de capitais porque cria liquidez, dando aos investidores a confiança para compra de títulos (PINHEIRO, 2018; INFOMONEY, 2006; LEE; WHANG, 2002).

O termo mercado de capitais define, de forma ampla, o local onde várias entidades negociam diferentes ativos financeiros. Esses locais podem incluir o mercado de ações, o mercado de renda fixa e os mercados de moeda e de câmbio. A maioria dos mercados está concentrada nos principais centros financeiros, como a B3, bolsa de valores de Nova York, NASDAQ (Associação Nacional de Corretores de Títulos de Cotações Automáticas), bolsa de valores de Tóquio (PINHEIRO, 2018).

Os mercados primários estão abertos a investidores específicos que compram títulos diretamente da companhia emissora. Esses títulos são considerados ofertas primárias ou ofertas públicas iniciais (IPOs) (SHI et al., 2013). Quando uma empresa abre seu capital, ela vende seus títulos e ações para investidores institucionais (PINHEIRO, 2018; INFOMONEY, 2006).

2.2.1 Ações

A primeira ação foi criada em 1602 pela Companhia Holandesa das Índias Orientais (*Verenigde Oostindische Compagnie*) assim acabou dando origem a Bolsa de Valores de Amsterdã, que por ser a primeira bolsa de valores conhecida é considerada a bolsa de valores mais antiga da mundo (GELDERBLOM; JONKER, 2004; CHAUDHURI, 1999; LEMOS,

2018). Uma ação é um título de propriedade de uma parte do capital social da empresa que a emitiu, representa a propriedade de uma empresa e representa uma reivindicação proporcional sobre seus ativos (o que a empresa possui) e lucros (o que a empresa gera de lucro) (NETO, 2018).

A compra de ações outorga ao comprador a oportunidade de participar do sucesso da empresa por meio de aumentos no preço da ação e dividendos ¹ que a empresa possa vir a pagar (PINHEIRO, 2018). Os acionistas têm uma reivindicação sobre os ativos da empresa em caso de liquidação, mas não possuem esses ativos (PINHEIRO, 2018).

Os detentores de ações ordinárias (ON) têm direitos de voto nas assembleias de acionistas e o direito de receber dividendos se forem declarados (BODIE et al., 2014). Os detentores de ações preferenciais (PN) não têm direito a voto, mas recebem preferência em termos do pagamento de quaisquer dividendos sobre os detentores de ações ordinárias (BODIE et al., 2014). Eles também têm uma reivindicação maior sobre os ativos da empresa do que os detentores de ações ordinárias (PINHEIRO, 2018).

2.2.2 Ações ordinárias

Os detentores de ações ordinárias exercem controle de uma empresa elegendo um conselho de administração e votando em políticas corporativas (PINHEIRO, 2018; NETO, 2018). Os investidores geralmente obtêm um voto por ação ordinária para eleger membros do conselho que supervisionam as principais decisões tomadas pela administração da empresa. Os acionistas ordinários estão na parte inferior da escada prioritária em termos de estrutura de propriedade. Em um caso de liquidação dos ativos da empresa, os acionistas ordinários têm direitos sobre os ativos da empresa somente depois que os credores, acionistas preferenciais e outros detentores de dívidas sejam pagos integralmente (COMISSÃO, 2014).

Com as ações ordinárias, se uma empresa for à falência, os acionistas ordinários não recebem seu dinheiro até que os credores e os acionistas preferenciais da empresa tenham recebido sua respectiva parte dos ativos remanescentes (PINHEIRO, 2018). Isso torna as ações ordinárias mais arriscadas do que títulos de dívidas ou ações preferenciais (PINHEIRO, 2018; BODIE et al., 2014). A vantagem das ações ordinárias é que elas geralmente superam os títulos

¹Um dividendo é a distribuição de uma parte dos lucros da empresa e é paga a uma classe de seus acionistas. Os dividendos são decididos e administrados pelo conselho de administração da empresa, embora devam ser aprovados pelos acionistas através de seus direitos de voto. Dividendos podem ser emitidos como pagamentos em dinheiro, novas ações ou outras propriedades, embora os dividendos em dinheiro sejam os mais populares (PINHEIRO, 2018; PENMAN; SOUGIANNIS, 1998).

de dívidas e ações preferenciais no longo prazo. Muitas empresas emitem todos os três tipos de títulos.

2.2.3 Ações preferenciais

Os detentores de ações preferenciais têm prioridade sobre os detentores de ações ordinárias quando se trata de recebimento de dividendos, que geralmente geram mais do que ações ordinárias e podem ser pagos mensalmente ou trimestralmente. A decisão de pagar o dividendo é a critério do conselho de administração da empresa (COMISSÃO, 2014).

Ao contrário dos acionistas ordinários, os acionistas preferenciais têm direitos limitados que geralmente não incluem o voto (PINHEIRO, 2018). As ações preferenciais combinam características de dívida, na medida em que pagam dividendos fixos e capital próprio, na medida em que têm o potencial de se valorizarem. Isso atrai investidores que buscam estabilidade em fluxos de caixa futuros potenciais (BODIE et al., 2014).

Os acionistas preferenciais têm uma reivindicação prévia sobre os ativos de uma empresa se ela for liquidada, embora permaneçam subordinados aos detentores de títulos. As ações preferenciais são patrimônio, mas, em muitos aspectos, elas são ativos híbridos que se encontram entre ações e títulos (COMISSÃO, 2014). Eles oferecem renda mais previsível do que ações ordinárias e são classificados pelas principais agências de classificação de crédito. Ao contrário dos detentores de bônus, deixar de pagar um dividendo aos acionistas preferenciais não significa que a empresa esteja inadimplente. Como os detentores de ações preferenciais não desfrutam das mesmas garantias que os credores, os *ratings*² das ações preferenciais são geralmente inferiores aos dos mesmos bônus do emissor, com os rendimentos sendo, portanto, mais altos (PINHEIRO, 2018).

As ações preferenciais geralmente não possuem direitos de voto, embora sob alguns acordos esses direitos possam reverter para os acionistas que não tenham recebido seus dividendos (BODIE et al., 2014). As ações preferenciais têm menos potencial para se valorizar do que as ações ordinárias (PINHEIRO, 2018).

²*Ratings* são notas de crédito emitidas por agências de classificação de risco sobre a qualidade de crédito, ou seja, as agências avaliam a capacidade de um emissor (bancos, financeiras, empresas) de honrar suas obrigações financeiras integralmente e no prazo determinado (INFOMONEY, 2005).

2.2.4 A bolsa do Brasil

A história do mercado acionário brasileiro data do início do século 19, quando a primeira bolsa de valores brasileira foi inaugurada em 1817 (BARCELLOS, 2018). A B3 foi criada em março de 2017 pela incorporação da BMFBOVESPA e da Central de Custódia e de Liquidação Financeira de Títulos Privados (CETIP), o que deu origem a uma empresa de infraestrutura de mercado financeiro de classe mundial (BMFBOVESPA, 2018). Consolidou as atividades da BMFBOVESPA na negociação e pós-negociação de produtos listados e as atividades da CETIP em serviços de registro e depósito para títulos e valores mobiliários. Em maio de 2019, a B3 tinha mais de 400 empresas listadas. A bolsa de valores no Brasil é monitorada e regulamentada pela Comissão de Valores Mobiliários³ (CVM), que é uma agência federal que faz parte do Ministério da Economia (CVM, 2019).

A B3 é uma das maiores empresas, em valor de mercado, detendo uma posição global proeminente na indústria do mercado de ações. A B3 é uma empresa de capital aberto, negociada sob o código B3SA3. Os índices Ibovespa, IBrX-50 e IBrX-100 são os principais índices que acompanham as ações. A B3 tem uma sessão de pré-abertura entre 9h45 e 10h, seguida de um pregão padrão entre 10h e 17h00. Há também uma sessão de pós-venda entre as 17:30 e 18:00 de segunda-feira a sexta-feira, exceto em feriados (B3, 2021a). O seguinte formato indica um código de uma ação: XXXXY

- XXXX = quatro letras maiúsculas que representam o nome do emissor da ação;
- Y = um número que representa o tipo de ação (ON ou PN).

O número três no final do código da ação representa que esta ação é ordinária, o quatro é uma ação preferencial, e cinco, seis, sete e oito indicam a classe de uma ação preferencial “A”, “B”, “C”, “D” etc. Os direitos de cada classe constam do estatuto social da empresa emissora das ações. Os preços das ações são cotados em reais, com duas casas decimais e liquidados em D + 2 dias (B3, 2019).

2.2.5 Índices do mercado de ações

Um índice do mercado de ações é uma carteira de ações hipotética que representa

³<http://www.cvm.gov.br/>

um segmento do mercado financeiro, mostrando seu desempenho atual e histórico. O cálculo do valor do índice é calculado com todas as ações desta carteira de ações, ou seja é uma média das ações que a compõem, ponderadas conforme os fatores definidos no seu regulamento (COMISSÃO, 2014). Segundo COMISSÃO (2014) os índices cumprem três funções principais:

- São indicadores de variação de preços do mercado;
- Servem de parâmetros para avaliação de performance de portfólios;
- São instrumentos de negociação no mercado futuro.

Os índices de mercado medem o valor de uma carteira de participações com características específicas do mercado. Cada índice tem sua própria metodologia, que é calculada conforme seu regulamento (COMISSÃO, 2014). Uma ampla variedade de investidores usa índices de mercado para acompanhar os mercados financeiros e administrar suas carteiras de investimento. Os índices estão profundamente arraigados no negócio de administração de investimentos, com os fundos os usam como referência para comparações de desempenho e os gerentes que os utilizam como base para a criação de fundos de índices de investimento.

Os investidores não podem investir diretamente em um índice, portanto, esses portfólios são usados amplamente como média de desempenho ou para desenvolver fundos de índices. Os investidores seguem diferentes índices de mercado para avaliar os movimentos do mercado (COMISSÃO, 2014). Os principais índices de ações para acompanhar o desempenho do mercado na B3 são:

- Índice Bovespa - Ibovespa;
- Índice Brasil 50 - IBrX 50;
- Índice Brasil 100 - IBrX 100.

2.2.6 Índice Bovespa

O Índice Bovespa (Ibovespa) é o principal indicador de desempenho das ações negociadas na B3. O índice lista empresas importantes no mercado de capitais brasileiro, tem como objetivo ser um indicador do desempenho da média das cotações dos ativos de maior volume de negociação e representatividade do mercado de ações brasileiro (COMISSÃO, 2014). Desde sua criação em 1968, o Ibovespa atua como parâmetro para investidores de todo o mundo interessados em ações do mercado brasileiro (BMFBovespa, 2019).

Segundo COMISSÃO (2014) a seleção das ações para a carteira teórica do Ibovespa inclui as ações de empresas listadas na B3 que atendam, a todos os seguintes critérios:

- Estar entre os ativos elegíveis que, no período de vigência das três carteiras anteriores, em ordem decrescente de Índice de Negociabilidade ⁴, representem em conjunto 85% do somatório total desses indicadores;
- Ter presença em pregão de 95% no período de vigência das três carteiras anteriores;
- Ter participação em termos de volume financeiro maior ou igual a 0,1% no período de vigência das três carteiras anteriores;
- Não ser classificada como *Penny Stocks* ⁵.

A ponderação de cada ação é feita pelo valor de mercado de seus ativos em circulação, na espécie pertencente à carteira, com limite de participação baseado na liquidez. Se um ativo deixar de atender a dois dos critérios de inclusão citados, forem classificados como *Penny Stocks*, entrarem em situação especial ou ainda se enquadrarem em outros critérios estabelecidos pela B3, será excluído do Ibovespa. A cada quatro meses é feita uma reavaliação do Ibovespa, alterando-se a composição e o peso da carteira, para que a sua representatividade se mantenha ao longo do tempo.

2.3 ANÁLISE TÉCNICA

A análise técnica é um estudo, que visa avaliar investimentos e identificar oportunidades de negociação, analisando as tendências estatísticas coletadas da atividade de negociação, como movimento e volume de preços (LEMOS, 2018). Ao contrário dos analistas fundamentalistas, que tentam avaliar o valor intrínseco de um ativo, os analistas técnicos se concentram em padrões de movimentos de preços, sinais de negociação e várias outras ferramentas de gráficos analíticos para avaliar ativo (LEMOS, 2018).

Ela se resume a uma análise da oferta e da demanda no mercado para determinar para onde a tendência de preço está caminhando. Em outras palavras, a análise técnica tenta entender o sentimento (momento de comprar ou momento de vender) do mercado por trás das tendências de preço, em vez de analisar os indicadores fundamentalistas de um ativo. A análise técnica

⁴“Índice de Negociabilidade é uma média geométrica entre a participação de uma ação no número de negócios (peso de 1/3) e em volume financeiro negociado (peso de 2/3), sempre em relação ao total do mercado” (COMISSÃO, 2014).

⁵“*Penny Stocks* é um ativo cuja cotação seja inferior a R\$1,00” (BMFBOVESPA, 2014).

pode ser usada em qualquer instrumento financeiro com dados históricos de negociação. Isso inclui ações, contratos futuros, comódites, títulos de renda fixa, troca de moedas e outros títulos (LEMOS, 2018).

A análise técnica como se conhece hoje foi introduzida por Charles Dow e pela Teoria de Dow no final do século XIX (EDWARDS et al., 2018). Vários pesquisadores notáveis, incluindo William Peter Hamilton, Robert D. Edwards e John Magee contribuíram ainda mais para os conceitos da Teorias de Dow, ajudando a formar sua base (EDWARDS et al., 2018). Nos dias atuais, a análise técnica evoluiu para incluir centenas de padrões e indicadores desenvolvidos através de anos de pesquisa (EDWARDS et al., 2018).

Analistas técnicos partem da premissa que negociações passadas e as mudanças de preço de um título podem ser indicadores valiosos dos movimentos futuros de preço do título (EDWARDS et al., 2018). Eles podem usar análise técnica independente de outros esforços de pesquisa ou em combinação com alguns conceitos de considerações de valor intrínseco, mas na maioria das vezes suas convicções são baseadas somente nos gráficos estatísticos de um título (EDWARDS et al., 2018).

Existem muitas formas diferentes de análise técnica: algumas dependem de padrões gráficos, outras usam indicadores técnicos e osciladores, e a maioria usa uma combinação de técnicas (EDWARDS et al., 2018). De qualquer forma, o uso exclusivo de dados históricos e de volume de dados por parte dos analistas técnicos é o que os separa de suas contrapartes fundamentais. Ao contrário dos analistas fundamentalista, os analistas técnicos não se preocupam com a avaliação de uma empresa, a única coisa que importa são os dados históricos de negociação passadas e quais informações estes dados podem fornecer sobre movimentos futuros de preços.

2.3.1 Análise Técnica versus Análise Fundamentalista

Análise técnica e análise fundamentalista são as duas principais escolas de pensamento quando se trata de analisar os mercados financeiros (LEMOS, 2018). A análise técnica examina o movimento de preço de um título e usa esses dados para prever movimentos futuros de preços (LEMOS, 2018). A análise fundamentalista, ao contrário da análise técnica, analisa fatores econômicos e financeiros que influenciam um ativo (LEMOS, 2018).

Os analistas técnicos geralmente começam sua análise com gráficos, enquanto os

analistas fundamentais começam com os relatórios financeiros de uma empresa. A análise técnica e a análise fundamentalista têm objetivos diferentes. Analistas técnicos tentam identificar muitas negociações de curto a médio prazo, onde podem virar uma ação, enquanto analistas fundamentais tentam fazer investimentos de longo prazo no negócio subjacente de uma ação (LEMOS, 2018).

Analistas fundamentalistas tentam determinar o valor de uma empresa observando seu relatório de resultados, balanço e demonstração de fluxo de caixa. Em termos financeiros, o analista tenta medir o valor intrínseco de uma empresa, descontando o valor dos fluxos de caixa futuros projetados para um valor presente líquido (LEMOS, 2018). Um preço de um ativo que seja negociado abaixo do valor intrínseco dele é considerado uma boa oportunidade de investimento e vice-versa.

Já os analistas técnicos acreditam que não há razão para analisar os relatórios financeiros de um ativo, uma vez que o preço das ações já inclui todas as informações relevantes (EDWARDS et al., 2018). Em vez disso, o analista se concentra em analisar o próprio gráfico de ações para sugestões sobre a tendência do preço.

2.3.2 Princípios da teoria de Dow

Charles Henry Dow delineou os princípios básicos da teoria, que posteriormente recebeu o seu nome, nos editoriais que escrevia para o Wall Street Journal (LEMOS, 2018). Após sua morte em 1902, seu sucessor, William Peter Hamilton, como editor do Wall Street Journal, adotou os princípios de Dow e, ao longo de 27 anos continuou escrevendo sobre o mercado de ações, organizou e formulou a Teoria de Dow como se conhece hoje (LEMOS, 2018). Nesta seção estão os princípios básico que regem a teoria de Dow:

O mercado tem três tendências:

- Tendência primária, são os movimentos gerais, para cima e para baixo que geralmente duram mais de um ano e podem durar até vários anos e resultam em valorização geral ou depreciação de mais de 20% (LEMOS, 2018);
- Tendência secundária, são reações importantes que interrompem o progresso dos preços da direção tendência primária e duram de três semanas a três meses. Normalmente, eles retrocedem de um terço a dois terços do ganho (ou perda, conforme o caso) nos preços registrados no balanço anterior na direção da tendência primária (LEMOS, 2018);
- Tendência terciária, são pequenas flutuações breves de até três semanas, que reforçam

ou contrariam a tendência principal. É a única das três tendências que pode ser “manipuladas” (LEMOS, 2018).

O volume deve acompanhar a tendência, as negociações tendem a se expandir à medida que os preços se movem na direção da tendência primária predominante (LEMOS, 2018). Desta forma o volume aumenta quando os preços sobem e diminuem à medida que os preços caem em uma tendência primária de alta, já em uma tendência primária de baixa, o volume de negociações aumenta quando os preços caem e diminuem quando o preço sobe (EDWARDS et al., 2018).

Tendências primárias de alta têm três fases:

1. Acumulação, fase em que o grupo do *insiders* realizam suas ordens de compra. Este grupo geralmente possui informações fundamentais que possam afetar o preço de uma ação, mas que não estão disponíveis para o público (EDWARDS et al., 2018). Um exemplo seria um indivíduo que sabe sobre uma fusão de duas empresas antes de ser anunciado ao público;
2. Subida sensível, se tem um avanço estável no preço das ações e melhora nos resultados das empresas. No entanto o público, ainda não está totalmente convencido de que a melhora dos negócios é para valer (LEMOS, 2018). Apenas os investidores mais experientes e atentos começam a emitir ordens de compra, fazendo o volume de negociações começar a aumentar gradativamente nas subidas de preço e a diminuir nas quedas (LEMOS, 2018);
3. Estouro ou excesso, a maioria dos investidores já se convenceu de que existe uma alta nos preços (LEMOS, 2018). Os preços das ações sobem muito, junto com o volume de negociações, acompanhando o ritmo das altas nos preços. Nessa fase, geralmente, saem boatos sobre takeovers, fusões entre empresas e acordos entre empresas que prometem gerar receitas e lucros crescentes (LEMOS, 2018). Na mídia em geral o tom é amplamente favorável a um aumento no ritmo de crescimento econômico do país (LEMOS, 2018).

Tendências primárias de baixa têm três fases:

1. Distribuição, durante esta fase, os investidores mais atentos sentem o fato de que os lucros dos negócios atingiram uma altura anormal e começar emitir ordens de vendas de suas participações em um ritmo crescente Edwards et al. (2018). O volume de negociações ainda é alto, embora tenda a diminuir e o público ainda está ativo, mas começando a mostrar sinais de frustração, à medida que os lucros esperados desaparecem Edwards et al. (2018);
2. Pânico, durante esta fase, o número de compradores começa a diminuir e o número de vendedores aumenta. A tendência de queda dos preços de repente acelera em uma queda quase vertical, enquanto o volume atinge grandes proporções Edwards et al. (2018);

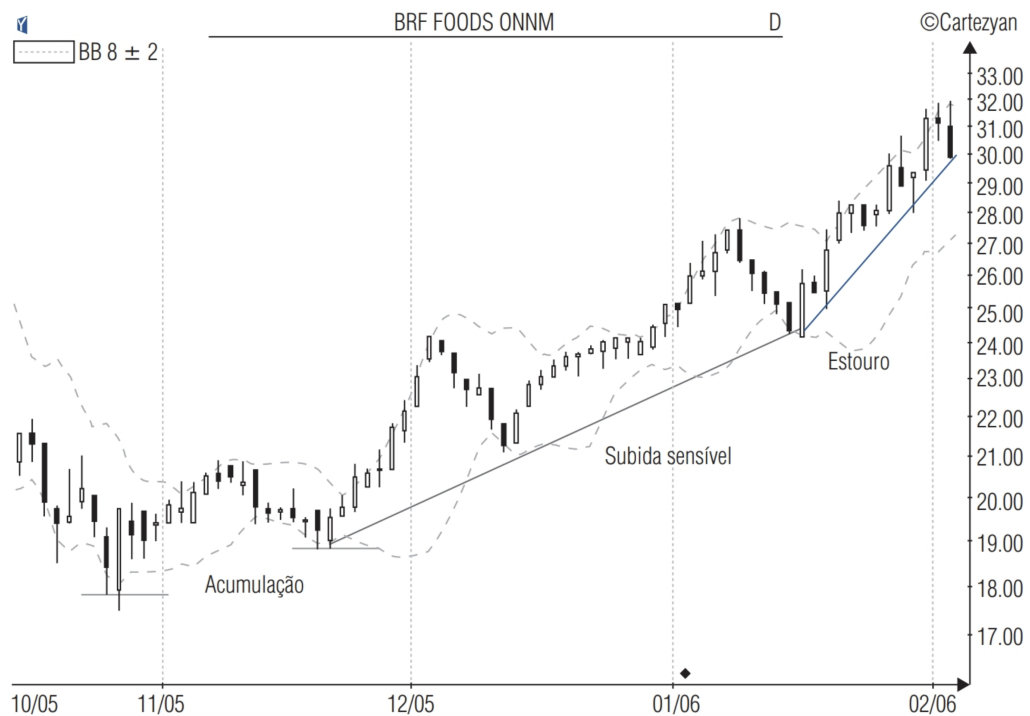


Figura 15 – Exemplo das três fases primárias da tendência de alta

Fonte: Lemos (2018)

3. Baixa lenta, durante esta fase os preços já atingem um patamar muito baixo, desencorajando os investidores que não venderam durante a fase de pânico a vender neste momento (LEMOS, 2018). O volume de negociações diminuiu, atingindo nível muito baixo.

As médias descontam tudo, os preços refletem as atividades de mercado combinadas de milhares de investidores, incluindo aqueles que possuem mais experiência e as melhores informações (*insiders*) sobre tendências e eventos, as médias em suas flutuações cotidianas descontam tudo o que é conhecido, tudo que é previsível e todas as condições que podem afetar a oferta ou a demanda do mercado Edwards et al. (2018). Mesmo calamidades naturais imprevisíveis, quando acontecem, são rapidamente avaliadas e seus possíveis efeitos são descontados (LEMOS, 2018).

As duas médias devem-se confirmar, este é o princípio mais frequentemente questionado e o mais difícil de racionalizar de todos os princípios da Teoria de Dow Edwards et al. (2018). No entanto, resistiu ao teste do tempo. Define que as duas médias (industrial e de transportes) devem andar na mesma direção. A falha de uma média em confirmar a direção da outra deve ser encarada como um alerta ao investidor de que pode existir uma possível mudança na tendência principal (LEMOS, 2018).

O mercado pode-se desenvolver em linha, uma linha na Teoria de Dow é um

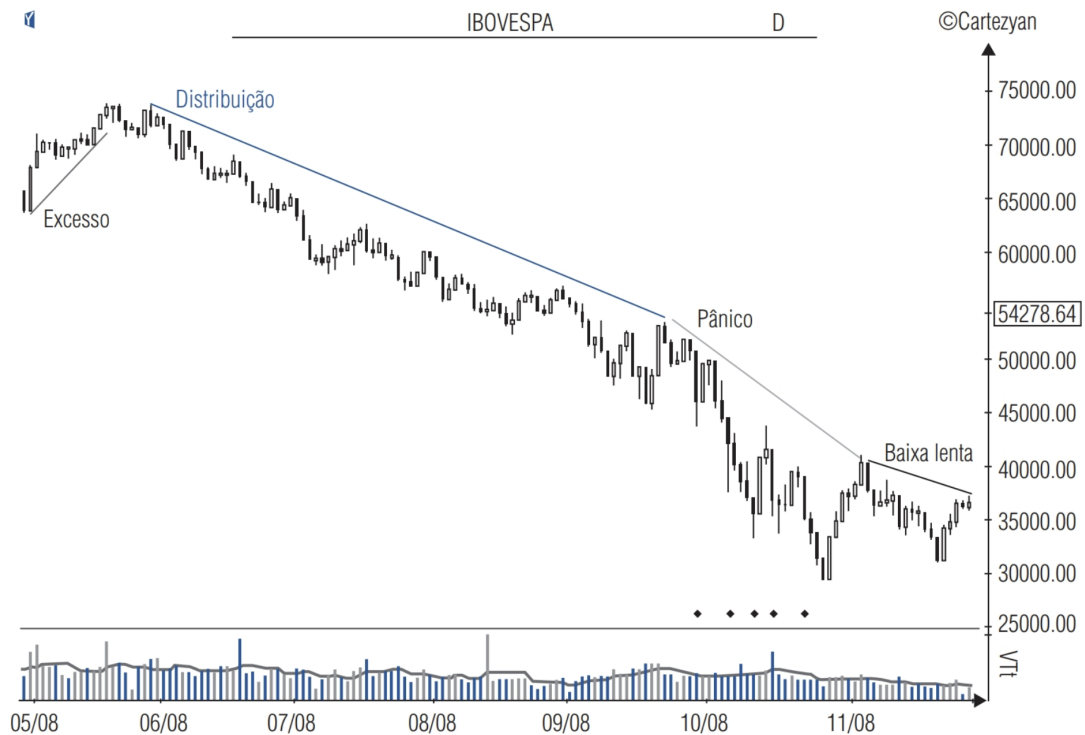


Figura 16 – Exemplo das três fases primárias da tendência de baixa

Fonte: Lemos (2018)

movimento lateral (como aparece na Figura 17) que normalmente a direção do movimento seguinte à linha se dá na direção da tendência primária. Pode durar de poucas semanas até por vários meses, durante os quais os preços flutuam dentro um intervalo de aproximadamente 5% ou menos (do seu valor médio) (LE MOS, 2018).



Figura 17 – Exemplo movimento lateral como uma linha.

Fonte: Lemos (2018)

As médias devem ser calculadas com preços de fechamento, a Teoria de Dow não leva em consideração quaisquer preços de altas ou baixas e abertura, que possam ser registradas

durante um dia de negociação, mas leva em consideração apenas os valores de fechamento (LEMOS, 2018).

A tendência está valendo até que haja sinais de reversão, este princípio da Teoria de Dow é aquele que, talvez mais do que qualquer outro, suscita críticas, por causa do atraso com que esse sinal é percebido pelos investidores (LEMOS, 2018). O que afirma é uma probabilidade, um aviso contra a mudança da posição de mercado muito cedo de uma tendência antes que ela tenha sido confirmada pela ação do mercado (LEMOS, 2018).

2.3.3 Conceitos básicos de tendência

Uma tendência é na verdade nada mais do que a direção geral para a qual um ativo ou mercado irá (LEMOS, 2018). O objetivo principal de quem utiliza a análise é técnica é o de descobrir a tendência e assim pode tirar proveito dela e obter lucro (LEMOS, 2018).

As tendências nem sempre são fáceis de detectar porque os preços quase nunca se movem em linhas retas. Em vez disso, os preços tendem a se mover em uma série de altos e baixos ao longo do tempo. Na análise técnica, busca a direção geral desses altos e baixos que constituem uma tendência. Uma tendência de alta é classificada como uma série de máximos e mínimos mais altos, enquanto uma tendência de baixa consiste em mínimos mais baixos e máximos mais baixos (EDWARDS et al., 2018).

Na Figura 18 se tem um exemplo de uma tendência de alta. Cada um dos pontos altos da tendência é maior do que o anterior, enquanto cada um dos pontos baixos da tendência é maior do que o anterior. Ainda na Figura 18 se tem mais dois exemplos tipos de tendência, de baixa e horizontal. Uma tendência de baixa é o inverso de uma tendência de alta, já uma tendência lateral é o movimento horizontal de preços que ocorre quando as forças de oferta e demanda são quase iguais.

Suporte e resistência são conceitos importantes de se entender sobre uma tendência. Os níveis de suporte são onde a demanda é percebida como forte o suficiente para evitar que o preço caia ainda mais, enquanto os níveis de resistência são preços nos quais a venda é considerada forte o suficiente para evitar que os preços subam ainda mais (EDWARDS et al., 2018).

Os níveis de suporte e resistência são níveis psicologicamente importantes em que muitos compradores e / ou vendedores estão dispostos a negociar uma ação (LEMOS, 2018). Quando as linhas de tendência são quebradas, a psicologia do mercado muda e novos níveis de

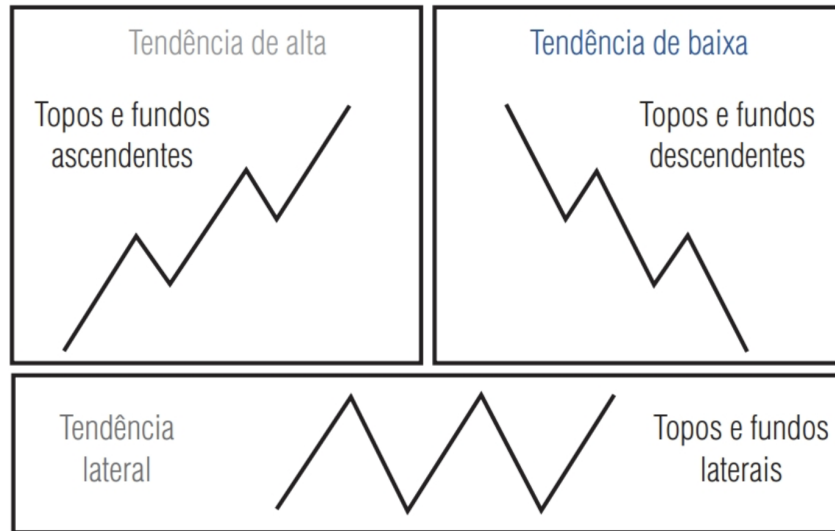


Figura 18 – Exemplo de tendência de alta, baixa e lateral.

Fonte: Lemos (2018)

apoio e resistência são estabelecidos (LEMOS, 2018). Na Figura 19, se tem um exemplo de suporte e resistência nos três tipo de tendência.

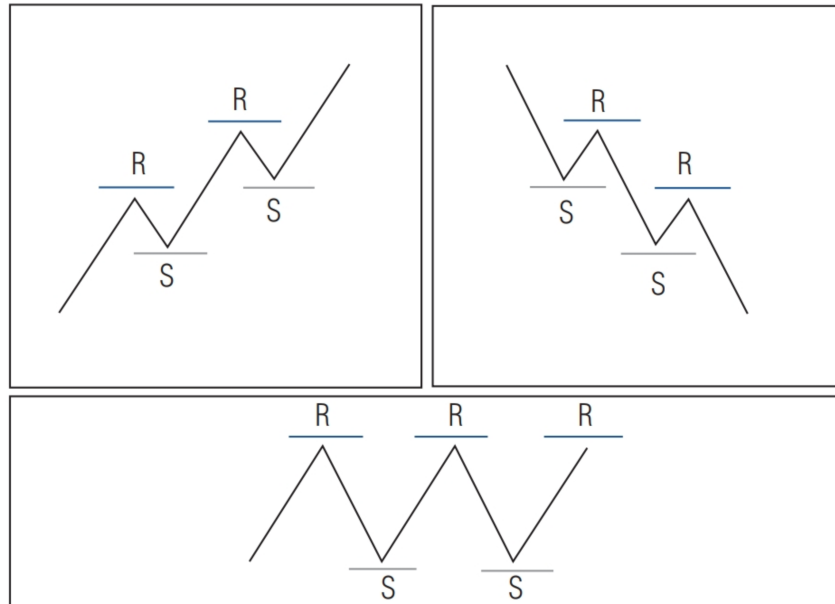


Figura 19 – Suportes e resistências para tendência de alta, baixa e lateral.

Fonte: Lemos (2018)

Os investidores devem estar cientes dos níveis de suporte e resistência e evitar colocar ordens de compra ou venda nesses pontos importantes, já que eles geralmente são caracterizados por muita volatilidade.

A análise técnica se baseia no pressuposto de que todas as informações já estão refletidas no preço de um título, o que significa que a análise desse preço é tudo o que importa. Ao analisar os padrões de preços e as estatísticas, os analistas técnicos tentam avaliar o sentimento geral do mercado e determinar para onde os preços podem ser direcionados.

3 MATERIAIS E MÉTODOS

Neste capítulo é apresentado ferramentas e métodos que ajudaram a chegar no objetivo proposto neste trabalho. Serão apresentadas as ferramentas que auxiliaram na coleta e análise de dados históricos do Índice Bovespa e na criação de RNR do tipo GRU e LSTM.

3.1 MATERIAIS

3.1.1 Conjunto de dados

O conjunto de dados utilizado foram os dados históricos do Índice Bovespa, que estão presentes na B3. Os valores históricos coletados foram do valor diário de cada dia de pregão, coletando apenas alta diária, baixa diária, valor de fechamento e volume de negociação.

A ferramenta utilizada para a coleta dos dados históricos foi o site do Yahoo Finanças¹, que permite a coleta de dados históricos de diversos instrumentos financeiros do mundo inteiro. Esta ferramenta permite selecionar um período e frequência dos dados que se deseja obter, a frequência pode ser definida como diariamente, semanalmente ou mensalmente. Quando definido os parâmetros dos dados que se precisa, o Yahoo Finanças permite o download destes dados em formato comma separated values (CSV), como pode ser visto na Figura 20. No caso do Ibovespa, esta ferramenta tem dados apenas posterior à data de 28/04/1993.

¹<https://br.financas.yahoo.com/>

IBOVESPA (^BVSP)

Sao Paulo - Sao Paulo Preço Adiado. Moeda em BRL.

☆ Adicionar à lista

96.976,35 +409,80 (+0,42%)

A partir de 10:31AM BRT. Mercado aberto.

Resumo Gráfico Conversas **Dados Históricos** Opções Componentes

Período: 25 de mai de 2019 - 30 de ma		Mostrar: Preços históricos		Frequência: Diariamente		Aplicar
Câmbio em BRL						Fazer download dos dados
Data	Abrir	Alto	Baixo	Fechamento*	Fechamento ajustado**	Volume
29 de mai de 2019	96.389,00	96.986,00	95.690,00	96.567,00	96.567,00	4.559.100
28 de mai de 2019	94.863,00	96.552,00	94.707,00	96.393,00	96.393,00	6.383.100
27 de mai de 2019	93.626,00	95.444,00	93.626,00	94.864,00	94.864,00	2.544.800

Figura 20 – Página Yahoo Finanças de onde foram obtido os dados.**Fonte: Autoria própria.**

3.1.2 Formato dos dados

O arquivo CSV retornado pelo o Yahoo Finanças possui sete atributos separados por vírgula, que são:

- Date, se refere a data dos valores coletados e estão do seguinte formato AAAA-MM-DD;
- Open, se refere ao valor de abertura do ativo para as negociações;
- High, se refere ao valor mais alto que foi negociado o ativo no dia;
- Low, se refere ao valor mais baixo que foi negociado o ativo no dia;
- Close, se refere ao valor de fechamento do ativo no dia;
- Adj Close, se refere ao valor de fechamento ajustado do ativo no dia;
- Volume, se refere a quantidade de negociações que o ativo teve no dia.

Na maior parte das vezes não se tem os dados no formato esperado, neste caso não foi diferente, o Yahoo Finanças retorna os dados igual como está na Tabela 1, com isso foi necessário a conversão dos dados para o formato adequado para a análise. Os atributos *Date*, e *Adj Close* são desnecessário para a análise, por isso será necessário removê-los e deixar os dados igual ao formato da Tabela 2.

Tabela 1 – Formato padrão dos dados do Yahoo Finanças.

Date	Open	High	Low	Close	Adj Close	Volume
2021-03-09	110611	112525	109343	111331	111331	12135300
2021-03-10	111331	112928	109999	112776	112776	12511800
2021-03-11	112782	115127	112776	114984	114984	12092100
2021-03-12	114984	114984	113253	114160	114160	8110700

Fonte: Autoria própria.

Tabela 2 – Formato adequado para a análise dos dados.

Open	High	Low	Close	Volume
110611	112525	109343	111331	111331
111331	112928	109999	112776	112776
112782	115127	112776	114984	114984
114984	114984	113253	114160	114160

Fonte: Autoria própria.

3.1.3 Linguagem de programação Python

Python² é uma linguagem de programação multi paradigma de alto nível, e que possui grande documentação disponível e comunidade de desenvolvedores ativa. A versão 3.7.9 foi utilizada neste trabalho.

3.1.4 Biblioteca Keras

Keras³ é uma das principais bibliotecas de RNA de alto nível atualmente. Ela é escrita na linguagem de programação Python e tem suporte para redes neurais convolucionais e recorrentes. A versão 2.4.3 do Keras foi utilizada neste trabalho, para a implementação das camadas GRU e LSTM.

²<https://www.python.org/>

³<https://keras.io/>

3.1.5 Biblioteca Numpy

Como os dados utilizados para a previsão estão em um formato que pode ser representado em forma de matriz, será necessário a utilização de alguma ferramenta que possa auxiliar em todas as operações que serão necessárias a serem realizadas com o conjunto de dados. Para tal tarefa foi selecionado o NumPy⁴ que é uma biblioteca que foi criada para permitir cálculos numéricos eficientes em matrizes multidimensionais para a linguagem de programação Python. É derivado da fusão de duas bibliotecas anteriores chamadas Numarray e Numeric. As operações com as matrizes são feitas por chamadas para funções escritas nas linguagens de programação Fortran e C, para assim ter um melhor desempenho em suas operações.

O NumPy define um novo tipo de dados chamado *ndarray* ou *n-dimensional*. Também define funções que trabalham em cada elemento da matriz do seu tipo de dado criado. Os elementos do tipo *ndarrays* podem conter dados do tipo inteiro, decimal e números complexos de vários tamanhos. Quando a biblioteca NumPy é carregada, o tipo de dados *ndarrays* se torna parte da linguagem Python tanto quanto os tipos de dados padrão do Python, como inteiros, strings, listas, dicionários, etc. A versão 1.19.5 do Numpy foi utilizada neste trabalho.

3.1.6 Biblioteca Pandas

Semelhante ao NumPy, o Pandas⁵ é uma das bibliotecas Python mais usadas para trabalhos com matrizes. Ela fornece estruturas e ferramentas de análise de dados de alto desempenho e fáceis de usar. Ao contrário da biblioteca NumPy, que fornece objetos para matrizes multidimensionais, o Pandas fornece um objeto de tabela 2D na memória chamado *Dataframe*. É como uma planilha com nomes de colunas e rótulos de linhas. Portanto, com tabelas 2d, o Pandas é capaz de fornecer muitas funcionalidades adicionais, como a criação de tabelas dinâmicas, colunas de computação baseadas em outras colunas e gráficos de plotagem. A versão 1.2.0 do Pandas foi utilizada neste trabalho.

⁴<https://www.numpy.org/>

⁵<https://pandas.pydata.org/>

3.1.7 Biblioteca Scikit-learn

Scikit-learn⁶ é uma biblioteca em Python que fornece muitos algoritmos de aprendizagem não supervisionados, aprendizagem supervisionados, regressão, classificação e pré-processamento. A versão 0.24.1 do Scikit-learn foi utilizada neste trabalho para a realização da normalização dos dados.

3.1.8 Ambiente de desenvolvimento e hardware

Tanto o ambiente de desenvolvimento utilizado para a criação e execução dos testes realizadas neste trabalho quanto o hardware utilizado para o treinamento dos modelos é o disponibilizado pela plataforma Kaggle⁷ que possui as seguintes configurações: sistema operacional Ubuntu 18.04.5 LTS; Tesla P100-PCIE com 16GB de memória; Intel® Xeon® E5-2699 v3 @ 2.30GHz; memória ram 16GB; 20GB de capacidade de armazenamento.

3.2 MÉTODOS

3.2.1 Pré-processamento de dados

Muitos algoritmos de inteligência artificial têm melhor desempenho quando as variáveis de entrada numéricas são dimensionadas para um intervalo padrão. Para este trabalho foi adotado a normalização dos dados, que dimensiona cada variável de entrada separadamente para o intervalo 0 entre 1, para valores de ponto flutuante onde pode se obter melhor desempenho dos algoritmos.

⁶<https://scikit-learn.org/>

⁷<https://www.kaggle.com/>

Foi utilizado o método `MinMaxScaler` da biblioteca `Scikit-learn 3.1` com a configuração padrão e os valores de escala de intervalo entre 0 e 1. Uma instância do `MinMaxScaler` foi definida com parâmetro `feature_range` com os valores 0 e 1.

3.2.2 Conjunto de dados de treinamento e validação

Os dados utilizados para treinamento são de 28/04/1993 até 17/08/2015, os dados utilizados para a validação dos modelos são de 18/08/2015 até 01/04/2021. Os dados utilizados no treinamento representam cerca de 80% dos dados disponíveis e os dados para os testes representam cerca de 20%. Sendo esses dados divididos em dois arquivos, um para treinamento e outro para os testes. O atributos destes dados utilizado como entrada para o treinamento dos modelos GRU e LSTM foi o *Close*, mostrado na Tabela 2.

3.2.3 Análise e comparação dos modelos utilizados

Para definir qual dos modelos (GRU ou LSTM) se saiu melhor para o objetivo deste trabalho, foi utilizado análise estatística, utilizando-se de critérios como coeficiente de determinação quanto coeficiente de correlação dos dados previsto versus dados reais.

3.2.4 Arquitetura do modelo

A arquitetura do modelo definida para a comparação é a mostrada na Figura 21. A camada de entrada consiste em n unidades do modelo testado (GRU ou LSTM), nas camadas ocultas se tem $\frac{n}{2}$ unidades do modelo testado e na camada de saída se tem uma camada densa, com uma unidade de saída. Onde n é quantidade de fechamentos diários anteriores que vai ser

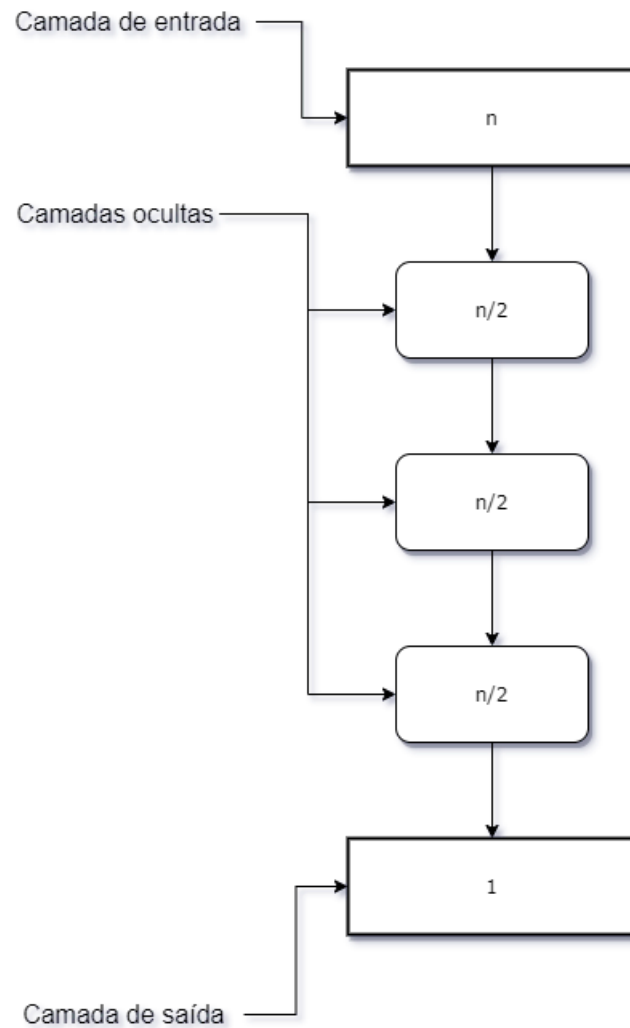


Figura 21 – Arquitetura do modelo para comparação entre Long short-term memory e Gated Recurrent Unit.

Fonte: Autoria própria.

utilizado para treinar o modelo.

3.2.5 Testes realizados

Foram definidos nove testes para cada modelo utilizado, consistindo em mudar a quantidade de dados (fechamentos anteriores) utilizados para previsão e a quantidade de épocas de treinamento para cada teste, como mostrado na Tabela 3. Além da variação da quantidade de dados e quantidade de épocas, também foi mudado a função de ativação na camada de saída de cada modelo, sendo utilizando as funções de ativação: *Linear*, *Relu*, *Selu* e *Sigmóide* (o

funcionamento de cada função pode ser visto em 2.1). Assim totalizando 36 ($9 \cdot 4$) testes para cada modelo 72 ($36 \cdot 2$) testes no total.

Os tipos de testes realizados com a função de ativação Linear, Relu, Selu e Sigmóide podem ser vistos na Tabela 4.

Tabela 3 – Quantidade de dados e épocas utilizados nos testes.

Dados	Épocas
100	300
100	600
100	900
200	300
200	600
200	900
300	300
300	600
300	900

Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

Tabela 4 – Parâmetros utilizado para os testes com as funções de ativação Linear, Relu, Selu e Sigmóide

GRU		LSTM	
Dados	Épocas	Dados	Épocas
100	300	100	300
100	600	100	600
100	900	100	900
200	300	200	300
200	600	200	600
200	900	200	900
300	300	300	300
300	600	300	600
300	900	300	900

Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

3.2.6 Parâmetros utilizados nos modelos

O parâmetro de regularização de dados *dropout* e o parâmetro para definição do tamanho do lote para treinamento *batch_size* foram fixos em todos os testes, tendo valores de 0,2 para o *dropout* e 100 *batch_size*. A função de perda utilizada foi a *Mean squared error* (MSE), também fixa para todos os testes.

Além dos parâmetros citados anteriormente, também foi necessário definir um otimizador para o modelo, o otimizador escolhido foi o *rmsprop* por ser recomendado pela documentação da biblioteca Keras, para aplicações em RNR.

3.2.7 Análise dos resultados

Para avaliar os resultados dos modelos comparados para previsão, foi utilizado testes estatísticos não paramétricos, pois os dados não seguiam uma distribuição normal de probabilidade (PINHEIRO et al., 2013), como pode ser visto na Figura 22. O coeficiente de determinação R^2 será usado para testar o quanto os valores previstos pelo modelo se ajustaram aos valores reais e o coeficiente de correlação de Spearman para quantificar a correlação dos dados reais e os obtidos pelos modelos.

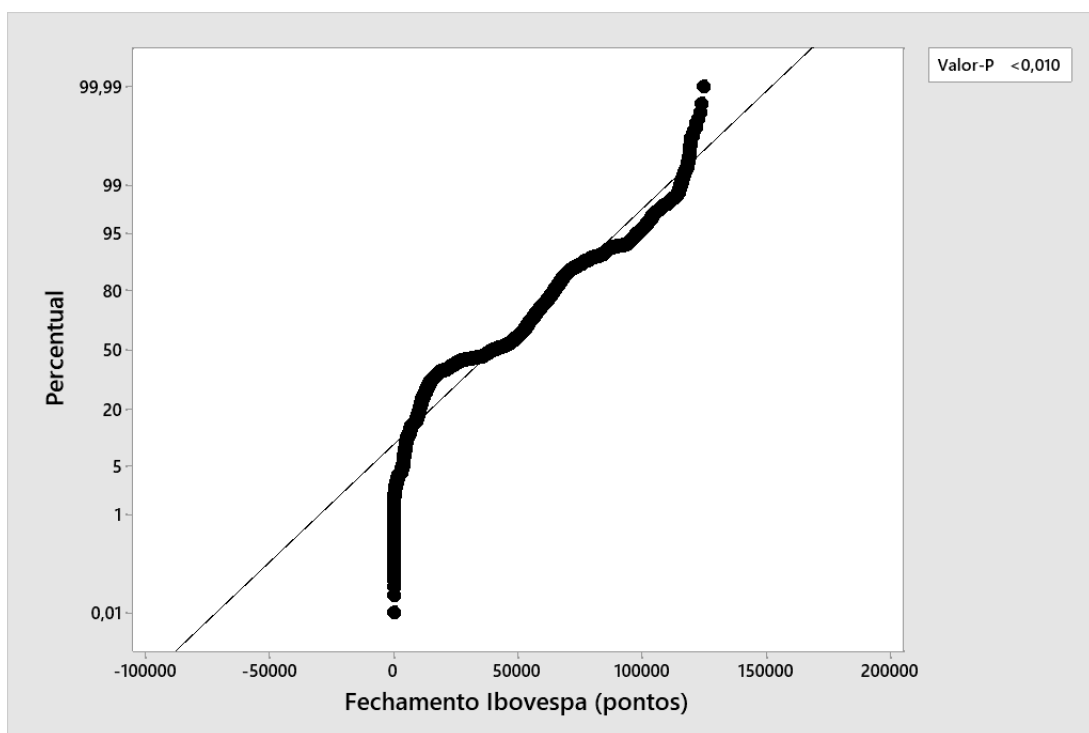


Figura 22 – Gráfico do teste de normalidade Kolmogorov-Smirnov para o conjunto de dados utilizado.

Fonte: Autoria própria.

4 RESULTADOS E DISCUSSÃO

Neste capítulo serão apresentados os resultados obtidos pelos modelo GRU e LSTM nos testes realizados, bem como a discussão sobre qual modelo teve o melhor desempenho para prever o valor de fechamento do Índice Ibovespa.

4.1 ANÁLISE EXPLANATÓRIA DO CONJUNTO DE DADOS

O conjunto de dados utilizado para treino e teste dos modelos, foram analisados previamente, antes de se realizar testes estatísticos para entendimento básico dos dados. A Tabela 5 mostra a estatística descritiva da variável *Open*, *High*, *Low* e *Close*. A variável *Volume* que também pertence ao conjunto de dados não foi utilizada nesta análise, pelo fato do Ibovespa ser um índice e não ser negociado diretamente¹.

Tabela 5 – Estatística descritiva de cada atributo do conjunto de dados.

Variável	N	Q1	Me	Q3	Min	Max	Assimetria
Open	6906	11706	40188	61293	24	125075	0,51
High	6906	11838	40777	61974	24	125324	0,50
Low	6906	11554	39631	60760	24	123227	0,51
Close	6906	11700	40236	61289	24	125077	0,51

N - Quantidade de dados; Q1 - Primeiro quartil; Me - Mediana; Q3 - Terceiro quartil; Min - Menor valor observado; Max- Maior valor observado.

Fonte: Autoria própria.

Analisando os valores do coeficiente assimetria dos atributos da Tabela 5 se pode afirmar que todos os atributos possuem assimetria positiva, o que indica que os dados são assimétricos para a direita segundo (PINHEIRO et al., 2013), o que pode ser observado pelo

¹É possível negociá-lo por meio de contratos futuros ou *exchange traded fund* (ETF) e assim ter um volume de negociação.

histograma do atributo *Close* na Figura 23.

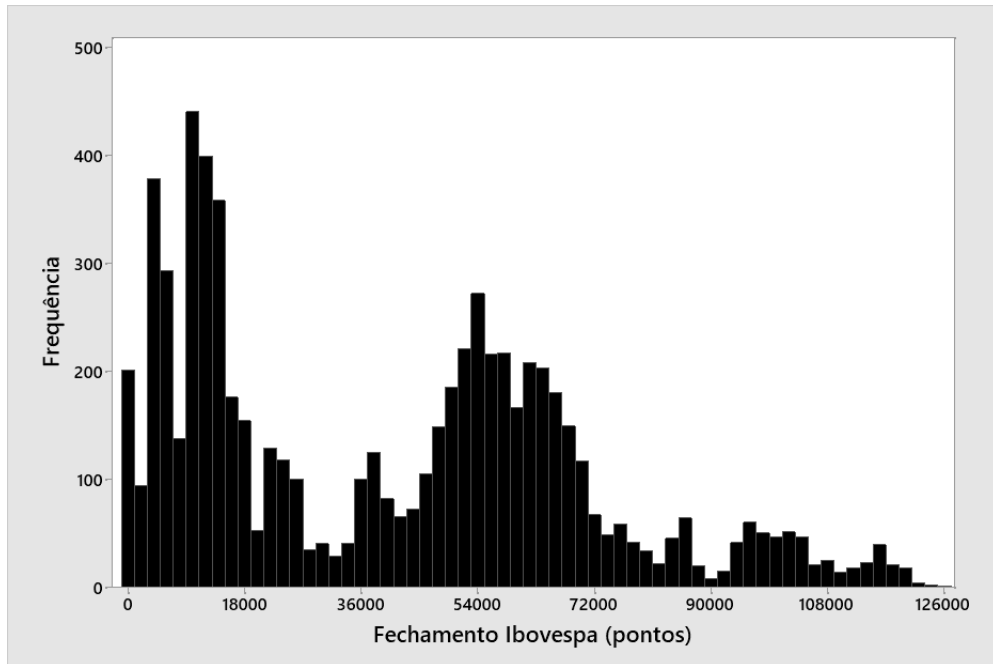


Figura 23 – Histograma atributo Close.

Fonte: Autoria própria.

4.2 DESEMPENHO DOS MODELOS

Com o auxílio do biblioteca Keras 3.1 criou-se uma RNR do tipo GRU e LSTM, com topologia baseada na arquitetura explicada em 3.2. A camada de entrada dos modelos recebeu como entrada o atributo *Close*, normalizado entre 0 e 1.

Alguns testes passaram por *underfit*², o que acarretou na impossibilidade de serem comparados, eles podem ser vistos na Tabela 6. As Tabelas 7, 8, 9 e 10 apresentam os resultados obtidos por cada modelo com funções de ativações distintas. Nos 4 resultados apresentados nas tabelas, o modelo GRU foi o que obteve o melhor desempenho para prever o valor de fechamento do Índice Ibovespa.

O coeficiente de determinação R^2 foi utilizado para avaliar a qualidade de ajuste dos modelos, pois indica quanto um determinado modelo foi capaz de explicar o valor de

²Underfit ocorre quando o modelo é incapaz de capturar a relação entre os exemplos de entrada e os valores de destino.

fechamento do Ibovespa, quanto mais próximo esse valor é de 100% melhor o modelo se saiu nos testes.

Tabela 6 – Testes que passaram por underfit.

Modelo	F	Dados	Épocas
GRU	Relu	300	300
GRU	Relu	300	600
GRU	Relu	300	900
LSTM	Relu	300	300
LSTM	Relu	300	600
LSTM	Relu	300	900

F - Função de ativação; Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

Tabela 7 – Resultados Gated Recurrent Unit e Long short-term memory para função de ativação Linear ordenados de forma decrescente pelo R^2 .

Modelo	Dados	Épocas	R^2	S
GRU	200	900	95,10%	97,96%
GRU	300	900	94,52%	97,29%
GRU	100	600	94,42%	98,52%
GRU	300	600	94,17%	97,22%
GRU	200	600	94,13%	98,03%
GRU	300	300	93,36%	96,53%
GRU	100	900	92,99%	97,76%
GRU	100	300	91,59%	97,32%
GRU	200	300	91,19%	95,48%
LSTM	200	900	85,06%	98,88%
LSTM	100	900	83,73%	96,96%
LSTM	200	300	79,74%	92,18%
LSTM	300	900	78,37%	89,22%
LSTM	100	600	74,11%	93,31%
LSTM	300	300	72,69%	87,17%
LSTM	200	600	63,42%	86,55%
LSTM	300	600	61,88%	63,21%
LSTM	100	300	61,38%	66,95%

R^2 - Coeficiente de determinação; S - Coeficiente de correlação de Spearman; Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

Apesar de terem sido utilizados 100, 200 e 300 fechamentos anteriores como preditores, na maior parte dos casos os melhores resultados foram obtidos com 200 fechamentos anteriores. Tal fato só foi diferente para os testes realizado com a função de ativação *Selu*,

Tabela 8 – Resultados Gated Recurrent Unit e Long short-term memory para função de ativação Relu ordenados de forma decrescente pelo R^2 .

Modelo	Dados	Épocas	R^2	S
GRU	200	300	98,00%	99,16%
GRU	200	600	97,19%	99,02%
GRU	100	900	94,65%	98,44%
GRU	100	300	93,14%	98,22%
GRU	100	600	90,02%	98,13%
GRU	200	900	86,68%	95,80%
LSTM	100	900	83,28%	95,57%
LSTM	100	600	80,13%	98,22%
LSTM	200	300	74,31%	96,77%
LSTM	100	300	71,48%	94,84%
LSTM	200	600	70,83%	88,16%
LSTM	200	900	58,99%	67,42%
GRU	300	300	*	*
GRU	300	600	*	*
GRU	300	900	*	*
LSTM	300	300	*	*
LSTM	300	600	*	*
LSTM	300	900	*	*

R^2 - Coeficiente de determinação; S - Coeficiente de correlação de Spearman; Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

onde o melhor caso aconteceu com apenas 100 fechamentos anteriores, como pode ser visto na Tabela 9. No caso da quantidade de épocas de treinamento, também se tem uma predominância principal entre os melhores resultados obtidos, sendo 300 épocas a predominância. Apenas os testes realizados com a função de ativação Sigmóide diferem, se tendo o melhor resultado com 900 épocas de treinamento.

A Tabela 11 mostra os cinco melhores resultado entre todos os testes realizados, sendo todos esses resultados obtidos pelo modelo GRU, conseguindo chegar a um coeficiente de determinação de 98,00%, utilizando de 200 fechamentos para predição, 300 épocas para treinamento e com a função de ativação *Relu*.

As Figuras 24, 25, 26, 27 e 28 apresentam o gráfico de dispersão para os valores previsto (eixo das abcissas) e os dados reais de fechamento do Ibovespa (eixo das ordenadas), pelos cinco modelos com o melhor coeficiente de determinação.

O coeficiente de correlação linear de Spearman foi utilizado para medir a relação linear entre os dados reais e os previstos pelos modelos. Onde um valor mais próximo de 1 significa uma associação de classificação perfeita positiva, um valor mais próximo de 0 significa que

Tabela 9 – Resultados Gated Recurrent Unit e Long short-term memory para função de ativação Selu ordenados de forma decrescente pelo R^2 .

Modelo	Dados	Épocas	R^2	S
GRU	100	300	94,73%	98,36%
GRU	300	300	93,06%	96,60%
GRU	300	600	92,83%	97,01%
GRU	200	600	92,61%	97,33%
LSTM	300	900	92,24%	99,01%
GRU	100	600	90,82%	98,41%
GRU	200	900	90,22%	95,46%
GRU	300	900	89,95%	95,42%
GRU	200	300	89,95%	94,53%
LSTM	100	600	88,58%	99,00%
LSTM	200	600	84,59%	96,81%
LSTM	100	900	83,04%	91,20%
GRU	100	900	81,35%	96,65%
LSTM	200	900	77,20%	87,79%
LSTM	300	300	75,96%	97,98%
LSTM	100	300	75,95%	95,87%
LSTM	300	600	70,98%	87,04%
LSTM	200	300	58,58%	61,81%

R^2 - Coeficiente de determinação; S - Coeficiente de correlação de Spearman; Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

não há uma associação de classificação e um valor mais próximo de -1 significa uma associação negativa perfeita entre os intervalos, segundo (VIEIRA, 2003). De acordo com a classificação de (VIEIRA, 2015) os melhores resultados obtidos podem ser classificados como uma correlação forte positiva, por terem um coeficiente Spearman entre 75% e 100%, como apresentado na Tabela 11.

Para fazer um comparativo dos dados reais de fechamento do Ibovespa e os previstos pelos melhores modelos e verificar o quanto os dados são homogêneos entre os dados reais e os previstos, realizou-se o teste não paramétrico de Kruskal-Wallis (PONTES; CORRENTE, 2001), devido aos dados não seguirem uma distribuição normal de probabilidade (Tabela 12).

Pode-se observar na Tabela 12, que todos os cinco melhores modelos testados obtiveram Valor-p $> 0,05$, o que significa que os testes de Kruskal-Wallis aceitam a hipótese de homogeneidade entre os dados reais e os previstos com um nível de 5% de significância.

Portanto, o GRU teve melhor desempenho geral, tendo obtido os cinco melhores resultados (Tabela 11), com cada função de ativação testada e 29 resultados com coeficiente de correlação acima de 75%. O modelo LSTM chegou a ter 92,24% de coeficiente de determinação

Tabela 10 – Resultados Gated Recurrent Unit e Long short-term memory para função de ativação Sigmóide ordenados de forma decrescente pelo R^2 .

Modelo	Dados	Épocas	R^2	S
GRU	200	300	74,19%	91,46%
LSTM	200	300	73,00%	99,54%
LSTM	300	300	72,49%	99,32%
LSTM	300	600	71,39%	99,60%
LSTM	300	900	70,78%	97,15%
LSTM	100	900	70,54%	90,97%
GRU	300	900	70,20%	87,90%
GRU	300	300	69,76%	90,82%
LSTM	100	300	69,61%	87,79%
LSTM	200	900	69,13%	96,95%
LSTM	200	600	68,28%	97,29%
GRU	300	600	68,24%	80,99%
GRU	100	300	68,07%	79,61%
GRU	100	900	64,84%	62,64%
GRU	200	600	64,25%	71,57%
GRU	100	600	63,55%	68,13%
LSTM	100	600	58,74%	67,53%
GRU	200	900	57,95%	69,38%

R^2 - Coeficiente de determinação; S - Coeficiente de correlação de Spearman; Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

Tabela 11 – Cinco melhores resultados obtidos nos testes ordenados de forma decrescente pelo R^2 .

Modelo	Dados	Épocas	F	R^2	S
GRU	200	300	Relu	98,00%	99,16%
GRU	200	600	Relu	97,19%	99,02%
GRU	200	900	Linear	95,10%	97,96%
GRU	100	300	Selu	94,73%	98,36%
GRU	100	900	Relu	94,65%	98,44%

F - Função de ativação; R^2 - Coeficiente de determinação; S - Coeficiente de correlação de Spearman; Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

em um dos resultados e 28 resultados com coeficiente de correlação acima de 75%.

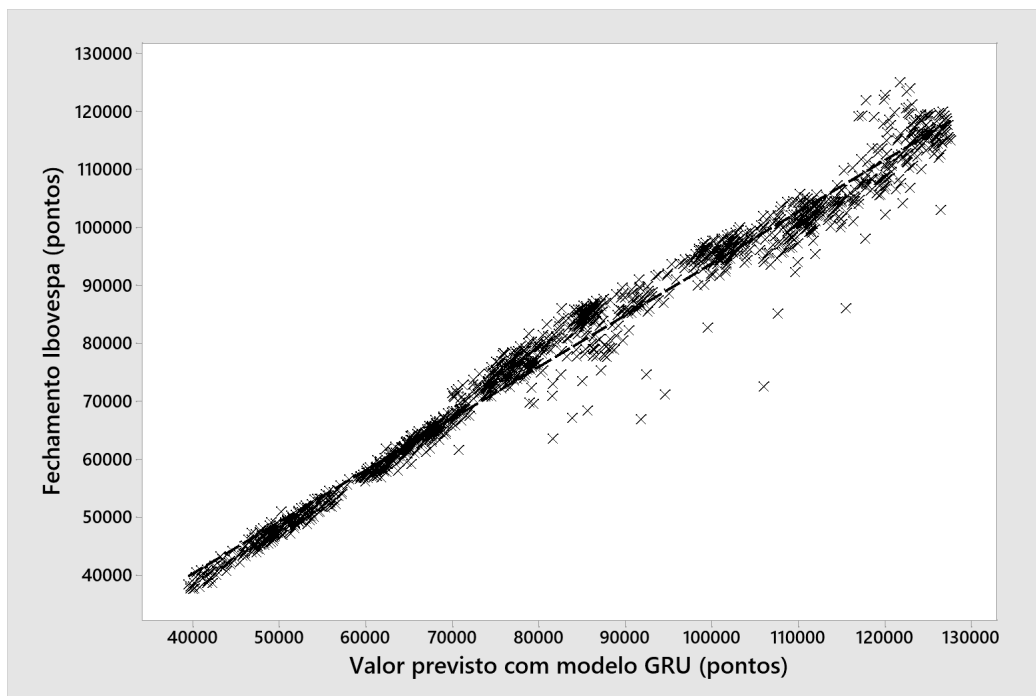


Figura 24 – Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 200 dados anteriores, 300 épocas de treinamento e com a função de ativação *Relu*.

Fonte: Autoria própria.

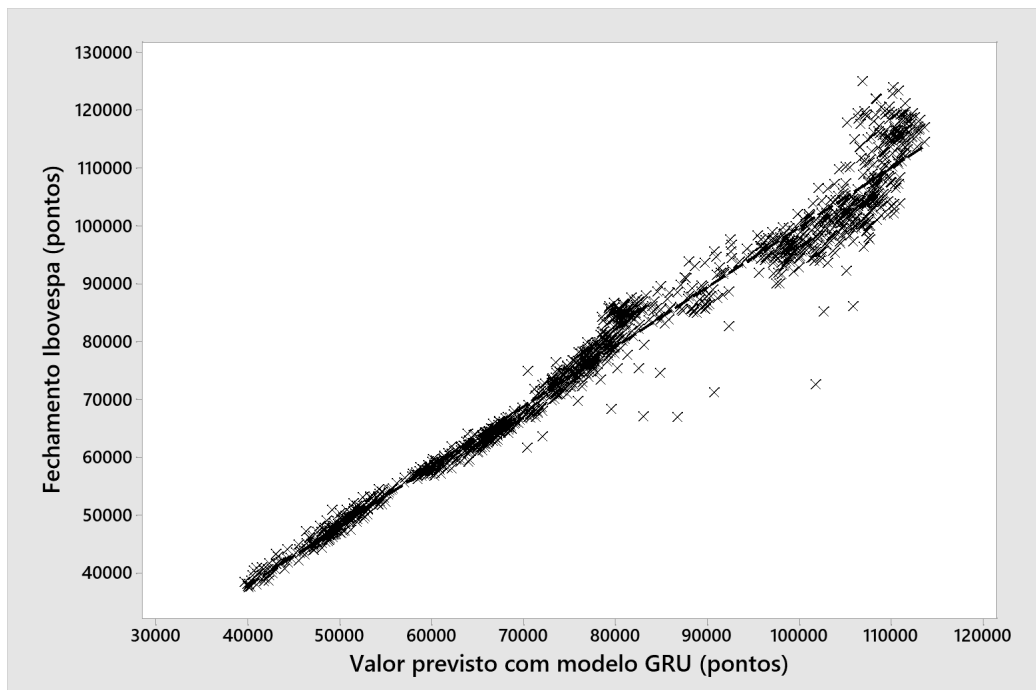


Figura 25 – Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 200 dados anteriores, 600 épocas de treinamento e com a função de ativação *Relu*.

Fonte: Autoria própria.

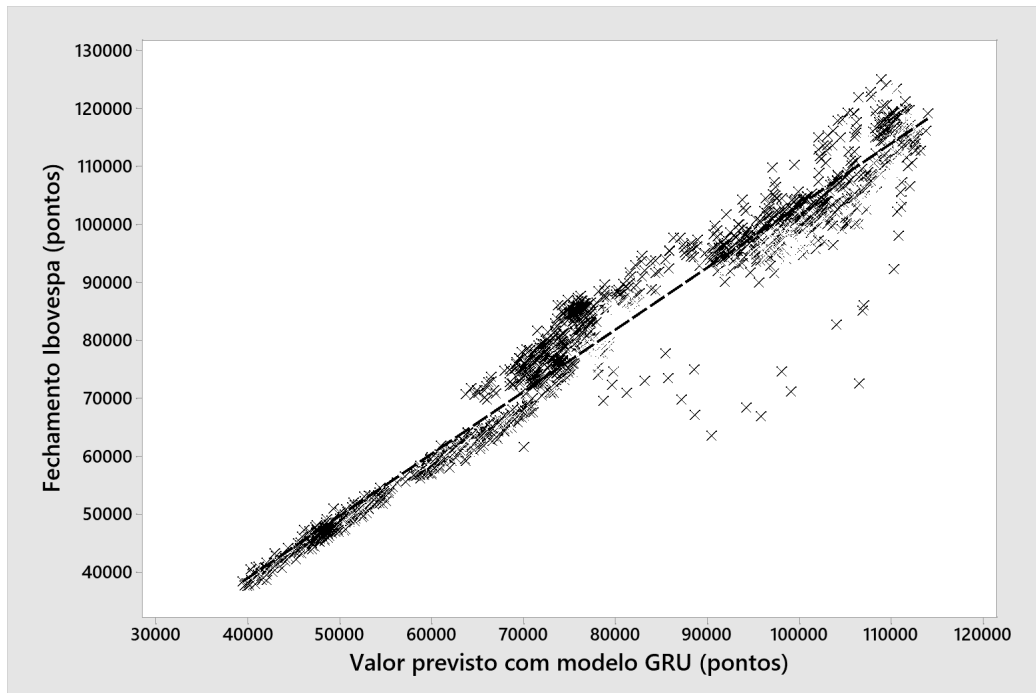


Figura 26 – Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 200 dados anteriores, 900 épocas de treinamento e com a função de ativação Linear.

Fonte: Autoria própria.

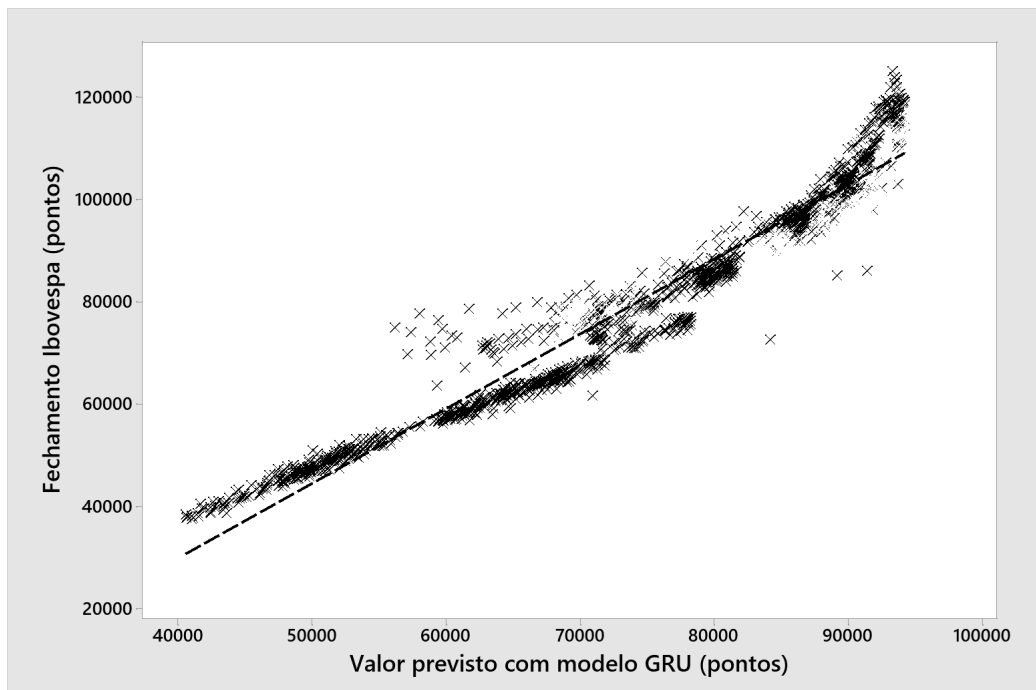


Figura 27 – Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 100 dados anteriores, 300 épocas de treinamento e com a função de ativação *Selu*

Fonte: Autoria própria.

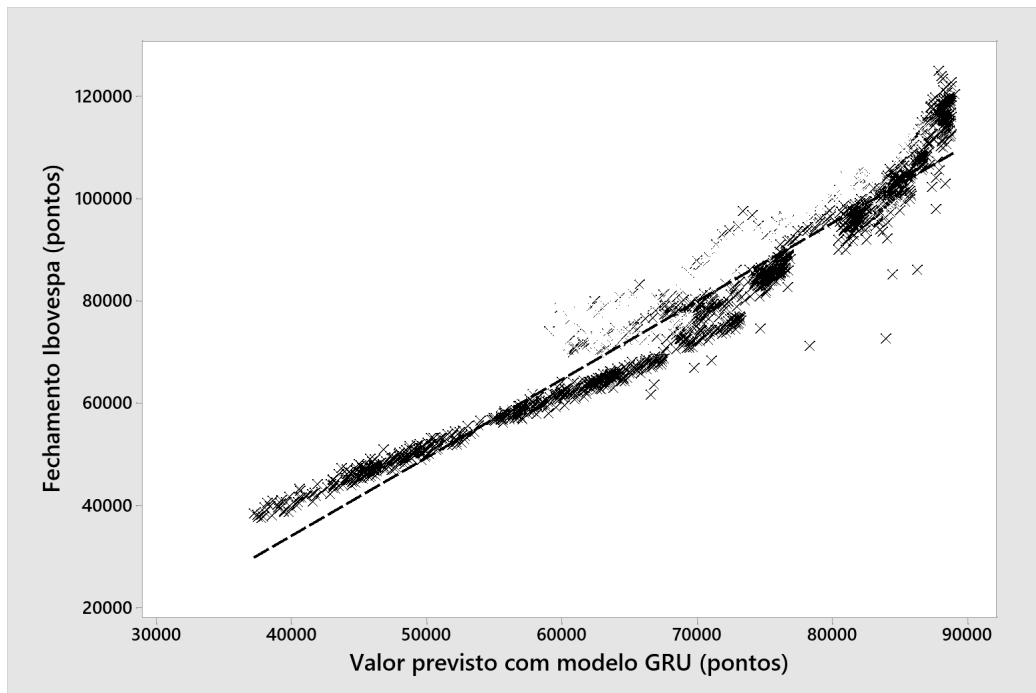


Figura 28 – Dados estimados do fechamento do Ibovespa, com o modelo Gated Recurrent Unit com 100 dados anteriores, 900 épocas de treinamento e com a função de ativação *Relu*.

Fonte: Autoria própria.

Tabela 12 – Testes de Kruskal-Wallis para homogeneidade e semelhança de entre os dados reais de fechamento do Ibovespa e os previstos pelos cinco melhores modelos testados.

Modelo	Dados	Épocas	F	Valor-H	Valor-p
GRU	200	300	Relu	1385,91	0,348
GRU	200	600	Relu	1385,83	0,348
GRU	200	900	Linear	1385,67	0,349
GRU	100	300	Selu	1385,8	0,348
GRU	100	900	Relu	1385,68	0,349

F - Função de ativação; Dados - Quantidade de fechamentos anteriores utilizados para predição.

Fonte: Autoria própria.

5 CONSIDERAÇÕES FINAIS

Neste capítulo estão as considerações finais do trabalho resumidas, da mesma forma que suas possíveis contribuições a investidores. Também é abordado possíveis trabalhos futuros para melhorar e dar continuidade ao objetivo de utilizar técnicas de inteligência artificial para prever o valor de fechamento de instrumentos financeiros.

5.1 CONCLUSÕES

Neste trabalho se mostrou promissora a hipótese de obtenção de um modelo para a predição do valor de fechamento do Índice Ibovespa, uma vez que uma rede neural recorrente do tipo *Gated Recurrent Unit* (GRU) utilizando os 200 fechamentos anteriores como previsores, 300 épocas para treinamento e utilizando a função de ativação *Relu* conseguiu explicar cerca 98% das variações dos dados reais.

Além da rede neural recorrente GRU, também se conseguiu com uma rede neural recorrente do tipo *Long short-term memory* (LSTM) utilizando os 300 fechamentos anteriores como previsores, 900 épocas para treinamento e utilizando a função de ativação *Selu* conseguisse explicar cerca 92% das variações dos dados reais.

A correlação do coeficiente de Spearman dos dados reais e os previstos pelo melhor modelo obtido GRU ficou classificada como forte com um coeficiente de 99,16%. Para o melhor modelo LSTM se teve um coeficiente de Spearman de 99,01%.

Assim sendo, conclui-se que com estes modelos é possível realizar uma estimativa de possíveis valores de fechamento para um determinado instrumento financeiro negociado em bolsa, já que ambos tiveram fortes coeficientes de determinação e coeficientes de correlação.

5.2 TRABALHOS FUTUROS

Ao longo desta pesquisa notou-se a possibilidade de se realizar testes com mais atributos preditores, assim como com instrumentos financeiros que seja possível se obter o volume de negociações. Assim a seleção de novos atributos pode auxiliar no desempenho dos modelos GRU e LSTM. Existem muitas técnicas para a seleção de atributos que podem ser aplicadas em trabalhos futuros para a previsão de fechamento de instrumentos financeiros.

Os dados utilizados neste trabalho foram obtidos do Yahoo Finanças, que tem a limitação de menor frequência dos dados obtidos ser a diária, porém existem outras fontes que podem fornecer dados de frequências menores de tempo como 1h e 15min por exemplo. Então uma proposta para trabalho futuro é a utilização de dados com uma frequência menor, para treinamento e testes dos modelos.

Outro sugestão de trabalho futuro é a utilização de um dos modelos em conjunto com operações automatizadas com o auxílio de robôs de investimento. Um robô de investimento é uma ferramenta que realiza a negociação de instrumentos financeiros de forma automatizada, sendo assim, um modelo GRU ou LSTM poderia indicar momentos de compra ou venda para um robô de investimentos.

REFERÊNCIAS

- ARBIB, M. A. **Brains, machines, and mathematics**. [S.l.]: Springer Science & Business Media, 2012.
- B3. **Projeto D+2**. 2019. Disponível em: <http://www.b3.com.br/pt_br/projeto-d-2/sobre-o-projeto/>. Acesso em: 21 mai. 2019.
- B3. **Horário de negociação**. 2021. Disponível em: <http://www.b3.com.br/pt_br/solucoes/plataformas/puma-trading-system/para-participantes-e-traders/horario-de-negociacao/acoes/>. Acesso em: 07 mai. 2021.
- B3. **Perfil pessoas físicas**. 2021. Disponível em: <http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/consultas/mercado-a-vista/perfil-pessoas-fisicas/faixa-etaria/>. Acesso em: 4 abr. 2021.
- BARCELLOS, M. **Histórias do Mercado de Capitais no Brasil: Depoimentos inéditos de personalidades que marcaram a trajetória das bolsas de valores no país**. [S.l.]: Alta Books Editora, 2018.
- BMFBOVESP. **MANUAL DE DEFINIÇÕES E PROCEDIMENTOS DOS ÍNDICES DA BMFBOVESPA**. [S.l.], 2014.
- BMFBOVESP. **Perfil e Histórico**. 2018. Disponível em: <<https://web.archive.org/web/20190909171156/http://ri.bmfbovespa.com.br/static/ptb/perfil-historico.asp?idioma=ptb>>. Acesso em: 22 mai. 2021.
- BMFBOVESP. **Índice Bovespa (Ibovespa)**. 2019. Disponível em: <https://web.archive.org/web/20200505062128/http://www.bmfbovespa.com.br/pt_br/produtos/indices/indices-amplos/indice-bovespa-ibovespa.htm>. Acesso em: 22 mai. 2021.
- BODIE, Z.; KANE, A.; MARCUS, A. **Fundamentos de investimentos**. [S.l.]: AMGH Editora, 2014.
- BOEHME, R.; ÇOLAK, G. Primary market characteristics and secondary market frictions of stocks. **Journal of Financial Markets**, Elsevier, v. 15, n. 2, p. 286–327, 2012.
- BOSE, N. K.; LIANG, P. **Neural network fundamentals with graphs, algorithms, and applications (mcgraw-hill series in electrical computer engineering)**. McGraw-Hill USA, 1996.
- BOURQUIN, J.; SCHMIDLI, H.; HOOGEVEST, P. van; LEUENBERGER, H. Basic concepts of artificial neural networks (ann) modeling in the application to pharmaceutical development. **Pharmaceutical development and technology**, Taylor & Francis, v. 2, n. 2, p. 95–109, 1997.
- CHAKRABORTY, S. Optimized routing approach using mcculloch-pitts neuron model. 04 2016.
- CHAUDHURI, K. N. **The English East India Company: The study of an early joint-stock company 1600-1640**. [S.l.]: Taylor & Francis, 1999.

CHO, K.; MERRIËNBOER, B. V.; GULCEHRE, C.; BAHDANAU, D.; BOUGARES, F.; SCHWENK, H.; BENGIO, Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. **arXiv preprint arXiv:1406.1078**, 2014.

COLAH. **Understanding LSTM Networks**. 2015. Disponível em: <<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>>. Acesso em: 02 mai. 2021.

COMISSÃO, D. V. M. O mercado de valores mobiliários brasileiro. **Rio de Janeiro: Comissão de Valores Mobiliários**, 2014.

CVM. **Sobre a CVM**. 2019. Disponível em: <http://www.cvm.gov.br/menu/aceso_informacao_institucional/sobre/cvm.html>. Acesso em: 22 mai. 2019.

DOUCOURE, B.; AGBOSSOU, K.; CARDENAS, A. Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data. **Renewable Energy**, Elsevier, v. 92, p. 202–211, 2016.

EDWARDS, B. P. The rise of automated investment advice: Can robo-advisers rescue the retail market. **Chi.-Kent L. Rev.**, HeinOnline, v. 93, p. 97, 2018.

EDWARDS, R. D.; MAGEE, J.; BASSETTI, W. C. **Technical analysis of stock trends**. [S.l.]: CRC press, 2018.

ESTEVA, A.; KUPREL, B.; NOVOA, R. A.; KO, J.; SWETTER, S. M.; BLAU, H. M.; THRUN, S. Dermatologist-level classification of skin cancer with deep neural networks. **Nature**, Nature Publishing Group, v. 542, n. 7639, p. 115, 2017.

GELDERBLOM, O.; JONKER, J. Completing a financial revolution: the finance of the dutch east india trade and the rise of the amsterdam capital market, 1595–1612. **The journal of economic history**, Cambridge University Press, v. 64, n. 3, p. 641–672, 2004.

GERS, F. A.; SCHMIDHUBER, J.; CUMMINS, F. Learning to forget: Continual prediction with lstm. IET, 1999.

GURESEN, E.; KAYAKUTLU, G.; DAIM, T. U. Using artificial neural network models in stock market index prediction. **Expert Systems with Applications**, Elsevier, v. 38, n. 8, p. 10389–10397, 2011.

GURNEY, K. **An introduction to neural networks**. [S.l.]: CRC press, 2014.

HAYKIN, S. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2007.

HEBB, D. O. **The organization of behavior: A neuropsychological theory**. [S.l.]: Psychology Press, 2005.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural computation**, MIT Press, v. 9, n. 8, p. 1735–1780, 1997.

IBGE. **Projeção população do Brasil**. 2021. Disponível em: <<https://www.ibge.gov.br/apps/populacao/projecao/>>. Acesso em: 4 abr. 2021.

- INFOMONEY. **Conheça as diferenças entre mercado primário e secundário.** 2006. Disponível em: <<https://www.infomoney.com.br/educacao/guias/noticia/568515/conheca-diferencas-entre-mercado-primario-secundario>>. Acesso em: 1 mai. 2019.
- INFOMONEY, E. **Entenda a diferença entre o rating e a recomendação dada por analistas.** 2005. Disponível em: <<https://www.infomoney.com.br/educacao/guias/noticia/363532/entenda-diferenca-entre-rating-recomendacao-dada-por-analistas>>. Acesso em: 7 mai. 2019.
- JAIN, A. K.; MAO, J.; MOHIUDDIN, K. Artificial neural networks: A tutorial. **Computer, IEEE**, n. 3, p. 31–44, 1996.
- Jl, M. Are robots good fiduciaries: Regulating robo-advisors under the investment advisers act of 1940. **Colum. L. Rev.**, HeinOnline, v. 117, p. 1543, 2017.
- KARSOLIYA, S. Approximating number of hidden layer neurons in multiple hidden layer bpnn architecture. **International Journal of Engineering Trends and Technology**, v. 3, n. 6, p. 714–717, 2012.
- KOVÁCS, Z. L. **Redes neurais artificiais.** [S.l.]: Editora Livraria da Física, 2002.
- KROGH, A. What are artificial neural networks? **Nature biotechnology**, Nature Publishing Group, v. 26, n. 2, p. 195, 2008.
- LEE, H.; WHANG, S. The impact of the secondary market on the supply chain. **Management science, INFORMS**, v. 48, n. 6, p. 719–731, 2002.
- LEMO, F. **Análise técnica dos mercados financeiros : um guia com-pleto e definitivo dos métodos de negociação de ativos.** [S.l.]: Saraiva, 2018. ISSN 9788553131082.
- LI, H.; LIN, Z.; SHEN, X.; BRANDT, J.; HUA, G. A convolutional neural network cascade for face detection. In: **Proceedings of the IEEE conference on computer vision and pattern recognition.** [S.l.: s.n.], 2015. p. 5325–5334.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.
- NETO, A. A. Mercado financeiro. 2018.
- NORVIG, P.; RUSSELL, S. **Inteligência Artificial.** 3th. ed. Brasil: Elsevier, 2013.
- NUNES, I.; SILVA, H. S. D. **Artificial neural networks: a practical course.** [S.l.]: Springer, 2018.
- PACKIN, N. G. **Why We Still Want Humans, And Not (Just) Robots, To Invest Our Money.** 2018. Disponível em: <<https://www.forbes.com/sites/nizangpackin/2018/12/18/why-we-still-want-humans-and-not-just-robots-to-invest-our-money/>>. Acesso em: 7 mai. 2019.
- PENMAN, S. H.; SOUGIANNIS, T. A comparison of dividend, cash flow, and earnings approaches to equity valuation. **Contemporary accounting research**, Wiley Online Library, v. 15, n. 3, p. 343–383, 1998.
- PINHEIRO, J.; CUNHA, S.; GOMES, G.; CARVAJAL, S. **Probabilidade e estatística: quantificando a incerteza.** [S.l.]: Elsevier Brasil, 2013.

PINHEIRO, J. L. **Mercado de Capitais**. [S.l.]: Atlas, 2018.

PONTES, A. C. F.; CORRENTE, J. E. Comparações múltiplas não-paramétricas para o delineamento com um fator de classificação simples. **Rev. Mat. Estat., São Paulo**, v. 19, p. 179–197, 2001.

PORTAL DO INVESTIDOR. **Mercado Primário X Mercado Secundário**. 2019. Disponível em: http://www.investidor.gov.br/menu/Menu_Investidor/funcionamento_mercado/mercado_primario.html>. Acesso em: 1 mai. 2019.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958.

SAZLI, M. A brief review of feed-forward neural networks. **Communications, Faculty Of Science, University of Ankara**, v. 50, p. 11–17, 01 2006.

SHI, C.; PUKTHUANHONG, K.; WALKER, T. Does disclosure regulation work? evidence from international ipo markets. **Contemporary Accounting Research**, Wiley Online Library, v. 30, n. 1, p. 356–387, 2013.

TIAN, Y.; PEI, K.; JANA, S.; RAY, B. Deeptest: Automated testing of deep-neural-network-driven autonomous cars. In: ACM. **Proceedings of the 40th international conference on software engineering**. [S.l.], 2018. p. 303–314.

VIEIRA, S. **Bioestatística: tópicos avançados. Rev. E atual**. [S.l.]: Rio de Janeiro: Elsevier: Ed. Campus, 2003.

VIEIRA, S. **Introdução à bioestatística**. [S.l.]: Elsevier, 2015.