

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”**  
**FACULDADE DE ENGENHARIA DE BAURU**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**RONALDO CÉSAR DAMETTO**

**ESTUDO DA APLICAÇÃO DE REDES NEURAS ARTIFICIAIS PARA PREDIÇÃO DE  
SÉRIES TEMPORAIS FINANCEIRAS**

**BAURU - SP**

**2018**

RONALDO CÉSAR DAMETTO

**ESTUDO DA APLICAÇÃO DE REDES NEURAS ARTIFICIAIS PARA PREDIÇÃO DE  
SÉRIES TEMPORAIS FINANCEIRAS**

Dissertação de mestrado apresentada como exigência para obtenção do título de Mestre em Engenharia de Produção pelo Programa de Pós-Graduação em Engenharia de Produção da Universidade Estadual Paulista “Júlio de Mesquita Filho”, UNESP, Campus Bauru, sob orientação do Prof. Dr. Antonio Fernando Crepaldi.

**BAURU - SP**

**2018**

Dametto, Ronaldo César.

Estudo da aplicação de redes neurais artificiais  
para predição de séries temporais financeiras /  
Ronaldo César Dametto, 2018

64 f. : il.

Orientadora: Antonio Fernando Crepaldi

Dissertação (Mestrado)-Universidade Estadual  
Paulista. Faculdade de Engenharia, Bauru, 2018

1. Redes neurais artificiais. 2. Previsão séries  
temporais financeiras. 3. Ensemble de redes neurais.  
4. Long short-term memory (LSTM). I. Universidade  
Estadual Paulista. Faculdade de Engenharia. II.  
Título.

**ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE RONALDO CESAR DAMETTO, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO, DA FACULDADE DE ENGENHARIA - CÂMPUS DE BAURU.**

Aos 06 dias do mês de agosto do ano de 2018, às 14:30 horas, no(a) Anfiteatro da Seção Técnica de Pós-graduação da FEB, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Prof. Dr. ANTONIO FERNANDO CREPALDI - Orientador(a) do(a) Departamento de Engenharia de Produção / Faculdade de Engenharia de Bauru - UNESP, Prof. Dr. ROGERIO ANDRADE FLAUZINO do(a) Departamento de Engenharia Elétrica e de Computação / Universidade de São Paulo/São Carlos, Prof. Dr. KELTON AUGUSTO PONTARA DA COSTA do(a) Departamento de Computação / Faculdade de Tecnologia de Bauru, sob a presidência do primeiro, a fim de proceder a arguição pública da DISSERTAÇÃO DE MESTRADO de RONALDO CESAR DAMETTO, intitulada **ESTUDO DA APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS PARA PREDIÇÃO DE SÉRIES TEMPORAIS FINANCEIRAS**. Após a exposição, o discente foi arguido oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: aprovado. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.

Prof. Dr. ANTONIO FERNANDO CREPALDI

Prof. Dr. ROGERIO ANDRADE FLAUZINO

Prof. Dr. KELTON AUGUSTO PONTARA DA COSTA

## RESUMO

O aprendizado de máquina vem sendo utilizado em diferentes segmentos da área financeira, como na previsão de preços de ações, mercado de câmbio, índices de mercado e composição de carteira de investimento. Este trabalho busca comparar e combinar três tipos de algoritmos de aprendizagem de máquina, mais especificamente, o método *Ensemble* de Redes Neurais Artificiais com as redes Multilayer Perceptrons (MLP), auto-regressiva com entradas exógenas (NARX) e *Long Short-Term Memory* (LSTM) para previsão do Índice Bovespa. A amostra da série do Ibovespa foi obtida pelo Yahoo!Finance no período de 04 de janeiro de 2010 a 28 de dezembro de 2017, de periodicidade diária. Foram utilizadas as séries temporais referentes a cotação do Dólar, além de indicadores numéricos da Análise Técnica como variáveis independentes para compor a previsão. Os algoritmos foram desenvolvidos através da linguagem Python usando *framework* Keras. Para avaliação dos algoritmos foram utilizadas as métricas de desempenho MSE, RMSE e MAPE, além da comparação entre as previsões obtidas e os valores reais. Os resultados das métricas indicam bom desempenho de previsão pelo modelo *Ensemble* proposto, obtendo 70% de acerto no movimento do índice, porém, não conseguiu atingir melhores resultados que as redes MLP e NARX, ambas com 80% de acerto.

**Palavras-chave:** Redes neurais artificiais. Previsão séries temporais financeiras. Ensemble de redes neurais. Long short-term memory (LSTM).

## ABSTRACT

Different segments of the financial area, such as the forecast of stock prices, the foreign exchange market, the market indices and the composition of investment portfolio, use machine learning. This work aims to compare and combine two types of machine learning algorithms, the Artificial Neural Network Ensemble method with Multilayer Perceptrons (MLP), auto-regressive with exogenous inputs (NARX) and Long Short-Term Memory (LSTM) for prediction of the Bovespa Index. The Bovespa time series samples were obtained daily, using Yahoo! Finance, from January 4th, 2010 to December 28th, 2017. Dollar quotation, Google trends and numerical indicators of the Technical Analysis were used as independent variables to compose the prediction. The algorithms were developed using Python and Keras framework. Finally, in order to evaluate the algorithms, the MSE, RMSE and MAPE performance metrics, as well as the comparison between the obtained predictions and the actual values, were used. The results of the metrics indicate good prediction performance by the proposed Ensemble model, obtaining a 70% accuracy in the index movement, but failed to achieve better results than the MLP and NARX networks, both with 80% accuracy.

**Keywords:** Artificial neural networks. Financial time series forecast. Ensemble of neural networks. Long short-term memory (LSTM).

## LISTA DE ABREVIATURAS E SIGLAS

IA	Inteligência Artificial
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Scaled Error
MLP	Multilayer Perceptron
MSE	Mean Squared Error
RI	Retorno de Investimento
RN	Rede Neural
RNA	Rede Neural Artificial
SVM	Support Vector Machine

## LISTA DE FIGURAS

Figura 01 – Representação de um candlestick.....	26
Figura 02 – Índice Bovespa: Fechamento e Retorno Logarítmico.....	28
Figura 03 – Neurônio Artificial.....	31
Figura 04 – Arquitetura de uma rede neural multilayer perceptron (MLP) com duas camadas escondidas.....	35
Figura 05 – Arquitetura de uma rede neural NARX.....	36
Figura 06 – Representação de uma célula LSTM.....	37
Figura 07 – Representação de uma célula LSTM sendo replicado no tempo.....	38
Figura 08 – Exemplo de combinação de três RNAs.....	40
Figura 09 – Metodologia.....	43
Figura 10 – Série temporal financeira representando os valores de fechamento do índice Bovespa no período entre 04/01/2010 até 28/12/2017.....	46
Figura 11 – Gráfico de correlação Pearson e Spearman das variáveis.....	47
Figura 12 – Exemplo do subconjunto de Teste.....	49
Figura 13 – <i>Ensemble</i> das três RNAs .....	53
Figura 14 – Comparativo Real x Previsto pela RNA MLP.....	54
Figura 15 – Comparativo Real x Previsto pela RNA NARX.....	55
Figura 16 – Comparativo Real x Previsto pela RNA LSTM.....	56
Figura 17 – Comparativo Real x Previsto pelo método Ensemble.....	57

## LISTA DE TABELAS

Tabela 01 – Subconjunto da amostra com os 5 primeiros registros do arquivo Ibovespa.....	44
Tabela 02 – Subconjunto da amostra com os 5 primeiros registros do arquivo cotação dólar.....	44
Tabela 03 – Subconjunto da amostra com os exemplos das variáveis independentes e dependente.....	45
Tabela 04 – Análise estatística da série temporal com o valor do fechamento do Índice Bovespa.....	46
Tabela 05 – Coeficiente de correlação (Pearson, Spearman) e coeficiente de determinação das variáveis independentes.....	47
Tabela 06 – Exemplo dos valores normalizados usando Equação 14.....	48
Tabela 07 – Comparativo Real x Previsto pela RNA MLP.....	54
Tabela 08 – Comparativo Real x Previsto pela RNA NARX.....	55
Tabela 09 – Comparativo Real x Previsto pela RNA LSTM.....	56
Tabela 10 – Comparativo Real x Previsto pelo método Ensemble.....	57
Tabela 11 – Comparativo das RNAs e Ensemble através das métricas.....	58

## LISTA DE QUADROS

Quadro 01 – Descrição das variáveis independentes da abertura do pregão.....	45
Quadro 02 – Avaliação das funções de ativação sigmoid, tanh e relu.....	51

## LISTA DE EQUAÇÕES

Equação 01 – Média Móvel Simples (MMS).....	24
Equação 02 – Índice de Força Relativa.....	25
Equação 03 – Força Relativa.....	25
Equação 04 – Retorno Logarítmico.....	28
Equação 05 – Função Logística.....	31
Equação 06 – Função Tangente Hiperbólica.....	32
Equação 07 – Função Gaussiana.....	32
Equação 08 – Função Linear.....	32
Equação 09 – Função predição rede neural NARX.....	36
Equação 10 – Erro Médio Percentual Absoluto .....	41
Equação 11 – Erro Quadrático Médio.....	41
Equação 12 – Raiz do Erro Quadrático Médio.....	42
Equação 13 – Transformação Linear MinMax.....	48

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>12</b>
1.1 Justificativa .....	14
1.2 Objetivos .....	15
1.2.1 Objetivo Geral .....	15
1.2.2 Objetivos Específicos .....	15
1.3 Organização do Trabalho .....	16
<b>2. FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>17</b>
2.1 O Mercado Financeiro .....	17
2.1.1 Mercado de Ações .....	19
2.1.2 Índice Bovespa .....	21
2.1.3 Preços .....	21
2.1.4 Análise Fundamentalista .....	23
2.1.5 Análise Técnica .....	24
2.1.6 Representação dos preços em <i>Candlesticks</i> .....	25
2.2 Séries Temporais .....	26
2.2.1 Séries Temporais Financeiras - Retorno .....	28
2.3 Aprendizagem de Máquina .....	29
2.3.1 Redes Neurais Artificiais .....	30
2.3.1.1 Rede Neural <i>MultiLayer Perceptron</i> .....	34
2.3.1.2 Rede Neural NARX .....	35
2.3.1.3 Rede Neural com Long-Short Term Memory .....	36
2.3.2 <i>Ensembles</i> de Rede Neurais .....	38
2.4 Métricas de Desempenho .....	40
2.4.1 Erro Médio Percentual Absoluto .....	41
2.4.2 Erro Quadrático Médio .....	41
2.4.3 Raiz do Erro Quadrático Médio .....	42
<b>3. MÉTODO DE PESQUISA .....</b>	<b>43</b>
3.1 Séries financeiras analisadas .....	44
3.2 Análise descritivas dos dados .....	45
3.3 Normalização e Tratamento de Dados .....	48
3.4 Python .....	48
3.5 Técnicas de Aprendizado de Máquina e Configurações Experimentais .....	49
3.6 Método de previsão .....	50

3.6.1	Definições dos Modelos de Previsão.....	53
<b>4.</b>	<b>ANÁLISE DE RESULTADOS .....</b>	<b>54</b>
<b>5.</b>	<b>CONCLUSÃO.....</b>	<b>59</b>
<b>6.</b>	<b>REFERÊNCIAS .....</b>	<b>60</b>

## 1. INTRODUÇÃO

O mercado de ações é considerado como parte vital de qualquer economia, pois quanto mais ativo, mais se traduz em oportunidades para que pessoas e empresas apliquem seus recursos e contribuam para o aquecimento da economia. A compra e venda de ações é um tipo de investimento de curto e longo prazo que podem ser aplicados por qualquer pessoa. Porém este tipo de investimento envolve riscos e por isso o investidor deve ter conhecimento básico sobre o comportamento do mercado de ações (BM&FBOVESPA, 2017).

Dada a sua complexidade e dinamismo inatos, às previsões no mercado de ações, apresenta constante debate quanto à possibilidade de prever mudanças de preços em ações. Fama & Malkiel (1970) introduziu a hipótese do Mercado Eficiente que diz que o preço atual de um ativo sempre reflete toda a informação previamente existente, instantaneamente, ou seja, não é possível utilizar nenhuma informação conhecida para prever os preços futuro. Em conjunto a essa, existe também a hipótese do passeio aleatório (MALKIEL, 1973) que alega que o preço de uma ação varia independentemente do seu histórico, em outras palavras, o preço de amanhã irá ser influenciado apenas pela informação de amanhã, independente do preço de hoje.

Essas duas hipóteses juntas determinam que não existem meios de prever precisamente o preço de uma ação, todavia, existem diversos outros autores que alegam por meio de experimentos, que de fato, variações de preços podem sim ser previstas até certo nível (LO; MACKINLAY, 1999). São objeto de estudo de muitas disciplinas uma variedade de métodos para previsão e modelagem do comportamento de ações. Segundo Oliveira (2007), em 2002 aproximadamente 250 bilhões de dólares eram gerenciados em fundos de investimentos por modelos quantitativos, tais como lógica *fuzzy*, redes neurais artificiais, algoritmos genéticos, fractais e modelos de Markov. Importante notar que em 2012 era estimado que aproximadamente 85% das negociações de ações feitas nas bolsas de valores dos Estados Unidos eram feitas por meio de algoritmos (GLANTZ; KISSEL, 2013).

A previsão do comportamento das ações no mercado financeiro é uma análise de difícil tratamento devido ao comportamento não linear de sistemas como estes, gerando a “[...] necessidade de utilização de uma ferramenta capaz de tratar esse tipo de problema de forma automática e eficiente” (SILVA; SAPATTI; FLAUZINO, 2010, p.

302). Dessa forma, o avanço tecnológico elevou o grau de sofisticação das operações oferecidas pelo mercado financeiro aos investidores, o que eleva as possibilidades de ganhos e reduz riscos nas atividades de investimento em ações (MARTINS; METTE; MACEDO, 2008). Existem diversos métodos para auxiliar na tarefa de modelar e estimar a volatilidade e previsão de preços de ativos, um deles é a utilização de Redes Neurais Artificiais (RNAs) (MARTINS; METTE; MACEDO, 2008; ROCA; MÓL, 2015). Entre as aplicações possíveis de redes neurais, tais como reconhecimento de caracteres, diagnóstico médico, entre outras, provavelmente a previsão de séries temporais é uma das mais estudadas. As Redes Neurais Artificiais têm um bom comportamento para problemas não-lineares que possuem alto número de variáveis.

Vários trabalhos da literatura mostram que as Redes Neurais Artificiais trazem resultados positivos e mais aderentes que outros métodos de previsão de séries temporais, como ações e índices financeiros negociados na BM&FBovespa (MARTINS; METTE; MACEDO, 2008; OLIVEIRA, 2007; ROCA; MÓL, 2015). Redes Neurais Artificiais do tipo *perceptron* de multicamadas (ROSENBLATT, 1958) entregam melhores resultados do que outras técnicas de inteligência artificial na previsão de séries temporais (CAO; LEGGIO; SCHNIEDERJANS, 2005; SAAD; PROKHOROV; WUNSCH, 1998). As redes neurais estão entre as técnicas mais utilizadas para prever valores futuros ou tendências dos mercados de ações ao redor do mundo (ATSALAKIS; VALAVANIS, 2009). Importante ressaltar que a utilização de RNA serve tanto na abordagem linear como não linear. Hsu e Chen (2003), Kurt e Oktay (2010), Li *et al.* (2014), Chang *et al.* (2014), Mattar *et al.* (2015), entre outros, utilizaram Redes Neurais Artificiais na previsão de séries temporais. Entretanto, métodos baseados em redes neurais sozinhas têm sido largamente usados na previsão de séries temporais, com tudo, *ensembles* de redes neurais ainda não possuem a mesma popularidade nessa área, embora apresentem melhores resultados do que redes neurais sozinhas, dependendo da aplicação. Já foi demonstrado que *ensembles* simples podem realizar melhores generalizações do que uma rede neural sozinha (DIETTERICH; 2000; HANSEN; LIISBERG; SALAMON, 1992), e diversos estudos também mostram que um subconjunto de todas as redes neurais possíveis podem obter melhores resultados que uma rede neural sozinha (WEST; DELLANA; QIAN, 2005; WU; CHEN, 2001; ZHOU; WU; TANG, 2002).

Por se tratar de um problema caracterizado por uma série temporal, as redes neurais recorrentes e profundas podem ser mais apropriadas para este tipo de problema do que outros algoritmos em que não haja memória de curto prazo (GRAVES, 2012; NELSON, 2017).

Esse trabalho elege a combinação da análise técnica e a cotação do Dólar juntamente com a avaliação dos algoritmos de Aprendizagem de Máquina, mais especificamente, comparando e combinando os resultados de 3 (três) modelos de redes neurais, sendo: *perceptrons* de múltiplas camadas (*Multilayer Perceptrons - MLP*), rede neural auto-regressiva com entradas exógenas (*nonlinear autoregressive network with exogenous inputs - NARX*) e a rede neural com *Long Short-Term Memory* (LSTM) para a tarefa de prever o comportamento da série de índices diários da bolsa de valores de São Paulo (Ibovespa).

## 1.1 Justificativa

O mercado de capitais possui como uma das principais funções, através do mercado primário, possibilitar que as companhias realizem a emissão de valores mobiliários com o intuito de viabilizar projetos de investimento em condições mais vantajosas do que as oferecidas pelos empréstimos e financiamentos bancários.

Entretanto, alguns desses valores mobiliários, como as ações, representam frações patrimoniais da companhia e, dessa forma, não são resgatáveis em data pré-definida. Essa característica, entre outras, poderia afastar investidores do mercado de capitais caso eles não tivessem como negociar com terceiros os valores subscritos. Todavia, o mercado secundário cumpre essa função, sendo o local onde os investidores negociam valores mobiliários emitidos entre si. Portanto, o mercado secundário oferece liquidez aos títulos emitidos no mercado primário.

Além disso, os fundos de pensão, com sua importância no sistema previdenciário brasileiro e cumprindo um relevante papel no cenário econômico e social, podem por lei, possuir 50% dos investimentos feitos em renda variável.

Porém, apesar da importância do mercado de ações, ele apresenta uma complexidade inerente ao seu dinamismo, e por isso, as tarefas de prevê-lo ou modelá-lo, torna-se uma difícil, mas extremamente importante ferramenta para

tomada de decisões diante as previsões de tendências ou preços das ações no mercado.

Diante deste cenário, este trabalho tem como justificativa principal contribuir com as pesquisas relacionadas ao tema que se utiliza de algoritmos como os relacionados à Inteligência Computacional, mais especificamente, Aprendizagem de Máquina e sua subárea denominada Aprendizagem Profunda. Contribuindo assim com a avaliação de modelos que apresentem melhores resultados, tendo em vista que uma das grandes dificuldades nas pesquisas é avaliar qual o modelo e suas respectivas configurações melhor atende ao problema, tendo em vista a grande variedade de algoritmos disponíveis e em constante evolução.

## **1.2 Objetivos**

### **1.2.1 Objetivo Geral**

Este trabalho tem como objetivo geral o desenvolvimento e avaliação de 3 (três) diferentes tipos de redes neurais, bem como, um processo de *ensemble* dessas redes para predição e comportamento do índice Bovespa para “um-passo-a-frente”, durante o período de 10 (dez) dias, e utilizando-se das métricas de distância (previsto/real) para avaliá-los. Será avaliada a variável exógena de cotação do Dólar em seu valor de compra para compor os dados de entradas das redes, assim como, alguns indicadores técnicos.

### **1.2.2 Objetivos Específicos**

Os objetivos específicos deste trabalho são os seguintes:

- I. Estudar a teoria na qual são baseadas as redes neurais perceptrons de múltiplas camadas (MLP), rede neural recorrente auto-regressiva com entradas exógenas (NARX – *nonlinear autoregressive network with exogenous inputs*), redes neurais recorrentes LSTM e *ensemble* de redes neurais;
- II. Combinar e avaliar as séries temporais financeiras do Ibovespa com os indicadores da análise técnica, assim como, a cotação do dólar;

- III. Propor um modelo *ensemble* de Redes Neurais Artificiais e compará-lo aos demais modelos estudados neste trabalho, através de métricas de desempenho, identificando o melhor desempenho;
- IV. Analisar o desempenho do modelo com relação aos demais modelos em sua capacidade de acompanhar o movimento do índice Bovespa.

### **1.3 Organização do Trabalho**

No capítulo 2, introduz conceitos do mercado financeiro necessários para a compreensão do trabalho; mostra o que são as séries temporais, dando ênfase às séries temporais financeiras. Descreve as redes neurais artificiais, bem como, o algoritmo de aprendizagem de máquina e os modelos de redes neurais: MLP, NARX e LSTM. Apresenta, também, as principais métricas usadas na literatura para medir o desempenho das redes neurais.

No capítulo 3 apresenta a metodologia utilizada neste trabalho para avaliação dos modelos apresentados.

No capítulo 4 demonstra os resultados obtidos com o processamento dos modelos identificados no trabalho.

Por fim, no capítulo 5 as conclusões, assim como, sugestões de trabalhos futuros.

## 2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os conceitos utilizados ao longo dessa dissertação. Muitos dos conceitos utilizados nos trabalhos aqui pesquisados são inerentes ao mercado financeiro, e nem sempre são conhecidos pelo público em geral.

O objetivo desta seção é esclarecer os termos financeiros comumente utilizados em tais trabalhos, uma vez que o seu entendimento é essencial para a melhor compreensão do trabalho apresentado. Os conceitos relacionados as técnicas e/ou modelos de Inteligência Artificial, bem como, Aprendizagem de Máquina e Aprendizagem Profunda com suas aplicações também fazem parte desta seção.

A descrição desses conceitos em profundidade não faz parte do escopo deste trabalho, uma vez que já estão consolidados na literatura.

### 2.1 O Mercado Financeiro

Podemos dizer que o mercado financeiro é o meio onde se realizam as transferências de recursos, ou fluxo dos recursos monetários, entre os agentes financeiros/econômicos de forma direta ou indireta (MELLAGI; ISHIKAWA, 2000). Ele é a forma como captações viram empréstimos ou investimentos (BODIE *et al.* 2009).

O mercado de investimentos é parte do sistema financeiro. Nele um agente superavitário empresta seu capital a um agente deficitário com o objetivo de compensação financeira. Este empréstimo ocorre por meio de uma instituição financeira, devidamente regularizada no sistema financeiro. Dessa forma, analogamente para o mercado de investimentos, um investidor empresta seu capital a outro agente com o objetivo de obter lucro. O investimento é visto como um produto, ofertado pela instituição financeira que intermediará a operação.

Os investimentos sempre foram um desafio para o setor financeiro no que tange à oferta de produtos. Com uma grande variedade, cada um com vantagens e desvantagens, comparar todos os produtos e sugerir uma combinação que seja a melhor para o investidor é uma tarefa complexa. A complexidade envolvida se deve a fatores como a grande quantidade de produtos, a análise das condições econômicas, a análise das preferências de cada investidor, além das questões legais envolvidas.

Dentre as instituições que regulam nacionalmente os investimentos, temos a Comissão de Valores Mobiliários (CVM). Conforme Mellagi e Ishikawa (2000) a CVM é uma autarquia federal vinculada ao Ministério da Fazenda, instituída pela Lei nº 6.385, de 7-12-1976, cujas atribuições são a normatização, regulamentação, desenvolvimento, controle e fiscalização do mercado de valores mobiliários do país. Conforme regulamentação da CVM de 2013, as instituições financeiras do Brasil tiveram de implementar mecanismos que garantissem a conformidade entre os interesses do cliente investidor, seus perfis de investimento e as aplicações que efetuassem. Estes mecanismos aumentaram a complexidade na sugestão de produtos de investimentos. Atualmente, uma sugestão de aplicação deve melhorar a rentabilidade dos investimentos do cliente, estar em conformidade com os interesses do mesmo e com os dispositivos dessa regulamentação.

A principal característica da regulamentação é normatizar o que muitas instituições já faziam de um modo menos formal. Tais instituições procuravam: alinhar o comportamento do cliente no que tange a sua capacidade de assumir riscos, seu conhecimento do cenário econômico e de investimentos, sua situação financeira e seus interesses em curto, médio e longo prazo. O resultado da análise dessas características é conhecido como perfil de investidor. Com o resultado da identificação do perfil do investidor é possível saber se os investimentos estão enquadrados com esse perfil.

Os investimentos ou ativos financeiros, possuem um risco, uma rentabilidade e uma liquidez associados. Existe uma relação entre risco e rentabilidade para os investimentos (BODIE *et al.* 2009). Ela normalmente é proporcional, embora alguns fatores externos como a política, podem distorcer tal relação.

Os conceitos de risco, rentabilidade e liquidez podem ser definidos como:

**Risco:** é a variação da rentabilidade de um investimento. Esta variação influencia na probabilidade de se obter um rendimento esperado. Quando há muita variação em sua rentabilidade, um investimento pode não trazer a rentabilidade esperada ou até trazer prejuízo. Quanto mais previsível o valor de resgate/venda, maior a segurança e menor o risco do investimento (CAVALCANTI; MISUMI, 2001). Ainda segundo Cavalcanti e Misumi (2001), em termos estatísticos, risco é a dispersão de valores esperados em torno de uma média (desvio padrão). Quanto mais dispersos

os valores esperados da aplicação em relação a uma média, mais arriscado é o investimento.

**Rentabilidade:** é o retorno do capital investido, ou seja, se o investimento fosse retirado, subtraindo-se o capital após a retirada do capital inicialmente investido, teríamos a rentabilidade. Representa o resultado financeiro da operação (CAVALCANTI; MISUMI, 2001).

**Liquidez:** é a facilidade com que um investidor pode converter seu investimento novamente em capital, isso é, elevado volume de negociação (HOJI, 2007). Pode ter ligação, tanto com a natureza do investimento quanto com o seu valor para o mercado.

Uma forma para se minimizar o risco é diversificar os investimentos. A diversificação não pode ser feita de qualquer forma, deve-se analisar a natureza de cada um e sua relação com os outros. Normalmente avalia-se a variação de um ativo em relação ao outro a fim de detectar o grau de proximidade entre eles (MARKOWITZ, 1952).

Os ativos financeiros (comercializados no Mercado Financeiro) estão divididos em 3 grupos:

**Renda Fixa:** títulos de baixo risco cuja rentabilidade pode ser conhecida na sua contratação, tipicamente composta por produtos como títulos do governo, poupança, certificado de depósito bancário, debêntures, entre outros.

**Ações:** títulos de empresas de capital aberto comercializada em Bolsa de Valores. Por ser um mercado de alta volatilidade é considerado de elevado risco;

**Derivativos:** cujo preço deriva de outro ativo, comumente atrelado a uma operação real. Por exemplo: um agricultor que queira assegurar que quando sua plantação estiver prestes a ser exportada o preço dela se mantenha nos patamares atuais de câmbio, dessa forma cria-se um derivativo para segurar o preço de câmbio.

### 2.1.1 Mercado de Ações

O Mercado de Ações está diretamente relacionado com a economia de um país. Quanto mais desenvolvida é uma economia, mais ativo é seu mercado de capitais, o que se traduz em mais oportunidades para as pessoas, empresas e instituições aplicarem suas poupanças (SILVA, 2015). Quando um país possui várias empresas de capital aberto, estas empresas precisam de um lugar em comum para negociar

suas ações (ou ativos): a bolsa de valores. Neste local, entre as outras atividades, são feitas as compras e vendas das ações das empresas lá cadastradas. O Brasil possui atualmente uma única bolsa de valores em operação, a B3 S.A., resultado da fusão da BM&FBOVESPA S.A. (Bolsa de Valores, Mercadorias e Futuros) e Cetip S.A. (Mercados Organizados) em março de 2017. A B3 está entre as maiores bolsas do mundo em valor de mercado e é líder na América Latina. Atualmente negocia os ativos de aproximadamente 450 empresas as quais integram os índices Ibovespa, IBrX-50, IBrX e Itag, entre outros (BM&FBOVESPA, 2017).

A BM&FBovespa (2017) define o conceito de ações como: “Ações são pequenas partes de uma empresa. Quando uma companhia se torna aberta, o patrimônio dela é dividido em várias cotas, que são distribuídas para os investidores.”. Dentro da realidade brasileira, as ações são classificadas basicamente entre dois tipos: ordinárias (ON) e preferencias (PN). Os tipos PN, têm prioridade no recebimento de dividendos (antes dos acionistas ordinários) e os tipos ON, possuem como principal característica o direito de voto, ou seja, a capacidade de influenciar nas decisões da empresa (ASSAF NETO, 2001). O possuidor de ações pode comprá-las ou vendê-las a qualquer momento, porém, para lucrar neste tipo de mercado recomenda-se seguir algumas precauções, como conhecer a empresa que receberá o investimento, conhecer o mercado onde ela atua, certificar-se da credibilidade desta e avaliar a possibilidade de lucratividade no futuro. Uma das formas de obter lucro no mercado de ações é com a alta das ações adquiridas. O preço da ação é formado pelo fluxo de oferta e procura de cada papel. (BM&FBOVESPA, 2017).

Existem, basicamente, dois modelos diferentes que são utilizados para apoiar a estratégia de compra e venda de ações no mercado financeiro: modelo fundamentalista e modelo técnico. O fundamentalista analisa declarações fiscais e relatórios da companhia (NISON, 2003). Já o modelo técnico tem seu alicerce nos históricos de dados das ações e depois deduzindo desse histórico, provável tendência futura (EDWARDS *et al* 2006). O histórico de preços de uma determinada ação, portanto, produz uma série temporal não-linear, “[...] cuja previsão de comportamento é de fundamental importância para negociações futuras.” (SILVA; SAPATTI; FLAUZINO, 2010, p. 302). Nas seções 2.1.4 e 2.1.5, serão explanados um pouco melhor as análises mencionadas.

### 2.1.2 Índice Bovespa

O Índice Bovespa (Ibovespa), criado em 1968, é o indicador do desempenho médio das cotações das ações de maior negociabilidade e representatividade do mercado brasileiro negociadas na B3, resultado de uma carteira teórica de ativos, composto atualmente por 59 ações de empresas (BM&FBOVESPA, 2017). Leite e Sanvicente (1994) afirmam que a negociabilidade é o principal fator para a utilização do Ibovespa como índice de mercado, devendo ser considerado para a análise da rentabilidade a variação percentual entre dois pontos históricos quaisquer.

Em janeiro de 2018 as três ações com maiores participações na composição do Ibovespa foram ITUB4 do tipo PN, com mais de 10% de participação, a VALE3 do tipo ON com mais de 10% e a BBDC4 do tipo PN, com mais de 8% de participação (BM&FBOVESPA, 2018). Tais papéis são disponibilizados pelas empresas Itaú Unibanco S/A, VALE S/A e Bradesco S/A, respectivamente.

### 2.1.3 Preços

O preço pelo qual uma ação é negociada em bolsa, é fruto das condições de mercado (oferta e demanda) que refletem as condições estruturais e comportamentais da economia do País e específicas da empresa e de seu setor econômico (FORTUNA, 2011).

Ao final do dia, quando todas as negociações já foram concluídas, a forma mais comum de se analisar o desempenho de uma ação é observando o seu preço final. Essa métrica é usada pelos meios de comunicação e pelo público em geral, porém muitas vezes o preço final por si só não mostra como a ação se comportou ao longo do dia, já que desde a abertura até o fechamento do pregão, inúmeras operações de compra e venda desta ação podem acontecer, jogando seu preço para cima e para baixo durante o dia inteiro até esse preço chegar no seu valor final. As métricas utilizadas por investidores mais especializados para realizar uma avaliação mais profunda do preço de uma ação são as seguintes:

- **Fechamento** – mostra o valor que a ação atingiu ao final do período. É a métrica mais conhecida e utilizada na análise do mercado;

- **Abertura** – é o preço da ação no momento da abertura do pregão. Não é necessariamente igual ao valor de fechamento do dia anterior, visto que geralmente uma ação começa o dia já em alta ou em queda, devido a operações de *pre-market*, feitas antes do horário de abertura do pregão, e de *after-market*, feitas após o final do pregão do dia;
- **Máximo** – geralmente, ao longo do dia, o preço de uma ação sofre variações até chegar no valor de fechamento. Esta métrica mostra o valor máximo atingido pela ação ao longo do período. Este valor ajuda a interpretação de um preço porque quando o valor máximo de uma ação é muito maior do que o valor de fechamento, significa que o preço da ação caiu ao longo do período e a tendência é de que continue caindo no período seguinte. O inverso pode ser interpretado para um valor máximo muito perto do valor de fechamento;
- **Mínima** – mostra o menor valor atingido por uma ação durante o dia (sua interpretação é inversa à do valor máximo);
- **Volume** – é a soma do valor envolvido em todas as operações de compra e venda de determinada ação ao longo do dia. Pode variar entre alguns milhares de reais para empresas menores até centenas de milhões de reais para as maiores empresas do Brasil. O volume por si só não permite deduções sobre o comportamento do mercado, mas a análise de seu histórico permite identificar o início ou o final de tendências. Por exemplo, quando o volume é alto por um certo período e depois começa a cair, significa que o interesse do mercado pela tendência adquirida pela ação está acabando.

Todas as métricas de preços podem ser utilizadas isoladamente, em conjunto ou para criar indicadores de ações. Logo, é interessante que uma rede neural, cujo objetivo seja prever séries temporais financeiras, receba de parâmetros de entrada todas ou uma parte dessas métricas, lhe fornecendo dessa forma mais ferramentas para atingir seu objetivo de previsão.

#### 2.1.4 Análise Fundamentalista

A análise fundamentalista consiste em estudar a situação financeira da empresa, considerando os elementos internos, como balanço patrimonial, demonstração do resultado de exercício, fluxo de caixa, entre outros; e externos à empresa, como dados sobre o setor econômico de atuação e sobre o mercado acionário (KERR, 2011).

Analistas fundamentalistas utilizam as demonstrações financeiras para estimar um valor de mercado para as empresas. Sabendo-se esse valor estimado, é possível estimar um valor para a ação da empresa sem especulação. Caso esse valor esteja mais baixo que o valor de mercado das ações, é um indicativo de que está num bom momento de vender, caso esteja acima do valor de mercado, significa que está em um bom momento de comprar. Entretanto esse não é o único critério para a análise fundamentalista. Também se leva em conta a perspectiva de crescimento do mercado na área do ativo que está sendo comprado, no histórico da empresa, na sua credibilidade. A análise é feita em cima da empresa em si, normalmente para um investimento de médio a longo prazo.

Segundo Hoji (2007) alguns indicadores podem ser utilizados para apoiar as decisões na análise fundamentalista, a seguir alguns desses indicadores:

- **P/L** – É o índice Preço/Lucro por ação, que estima o prazo de retorno do capital investido e indica em quanto tempo haverá o retorno do investimento. O termo P é a cotação atual da ação e o termo L é o lucro futuro esperado, sendo o índice calculado pela fórmula de perpetuidade. Exemplo: se o preço atual da ação é R\$ 100,00 e o lucro anual por ação esperado é de R\$ 3,57, o índice P/L será de 28,0. Esse índice significa que o retorno do investimento ocorrerá num prazo de 28 anos.
- **P/VPA** – É o preço (P) da ação dividido pelo valor patrimonial da ação (VPA). Suponha que o patrimônio líquido da companhia seja de R\$ 150 milhões e o capital social da empresa esteja composto por 2 milhões ações, o que resulta em VPA de R\$ 75,00. Se o preço da ação é R\$ 100,00, o índice resultante é de 1,33, o que significa que a ação está sendo negociada por um preço 33% superior ao valor patrimonial.
- **Dividend Yield** - O índice é obtido com a divisão dos dividendos pagos pelo preço da ação e indica se a companhia é uma boa pagadora de dividendos. Suponha

que a “Companhia A” distribui R\$ 5,00 por ano e sua ação está cotada a R\$ 100,00 e a “Companhia B” distribui R\$ 6,00 por ano e sua ação está cotada também a R\$ 100,00. O *dividend yield* da A é 0,05 e o da B é 0,06, que é melhor, pois, quanto maior, melhor.

Outros indicadores podem ser utilizados na análise fundamentalista, entre eles, podemos citar: volume médio de negócios, volatilidade, crescimento de LPA (lucro por ação), margem operacional, nível de endividamento, etc.

### 2.1.5 Análise Técnica

Segundo Cavalcanti (2001), para a escola técnica, os gráficos traduzem o comportamento do mercado, e avaliam a participação de massas de investidores que influenciam a formação dos preços. Portanto, uma característica básica da análise técnica é a defesa da análise comportamental. O gráfico de barras é o instrumento mais utilizado pelos grafistas, permitindo visualizar os preços máximo, mínimo, médio (em alguns casos) e de fechamento, o volume negociado de determinada ação, em determinado dia, numa determinada Bolsa.

O comportamento de determinada ação e a tendência que essa ação está seguindo, apenas observando o histórico de seus preços não é suficiente, pois esse histórico não mostra a análise psicológica do que está acontecendo. Para melhorar o estudo dos preços de determinada ação, foram desenvolvidos indicadores técnicos que capturam comportamentos que não são visíveis quando apenas os preços são observados. Estes indicadores são geralmente modelos matemáticos que analisam valores passados de uma ação e tentam mostrar de alguma forma a tendência que os preços estão seguindo. O objetivo dos indicadores técnicos não é prever o que vai acontecer no futuro, como uma rede neural, mas fornecer ao investidor uma visão diferente da tendência do mercado, que não seria possível de se obter apenas com a observação dos seus preços.

Embora existam centenas de indicadores disponíveis na literatura, são apresentados os indicadores utilizados nesse trabalho, sendo eles: Médias Móveis, conforme Equação 01 e o Índice de Força Relativa (IFR), conforme Equação 02.

$$MMS = \frac{1}{N} (x_i + x_{i-1} + x_{i-2} + \dots + x_{i-N+1}) = \frac{1}{N} \sum_{k=0}^N x_{i-k} \quad (01)$$

Onde  $x_i$  são os termos da série temporal original,  $N$  é a ordem da média móvel.

$$IFR = 100 - \frac{100}{(1+RM)} \quad (02)$$

$$RM = \frac{MA}{MB} \quad (03)$$

Onde:

IFR = Índice de força relativa de  $n$  dias

RM = Relação entre as médias

MA = Média de alta

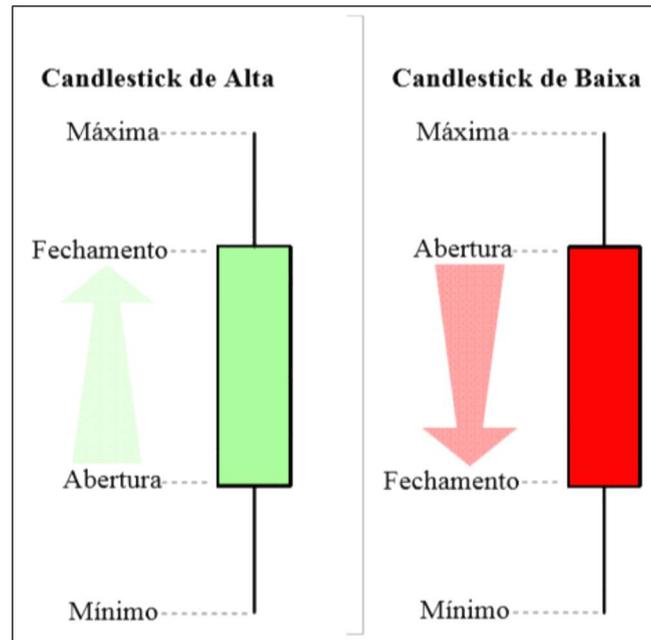
MB = Média de baixa

O valor RM é composto pela média dos ( $n$  pregões) fechados em alta sobre a média dos ( $n$  pregões) fechados em baixa. O IFR varia numa escala entre “0 à 100”, quando o valor for zero significa que a ação está completamente sem força de alta, e quando o valor for cem por cento significa que a ação está dominada pelas altas, sinalizando um venda.

### 2.1.6 Representação dos preços em *Candlesticks*

Um *candlestick* representa a variação dos preços de um determinado ativo em uma unidade de tempo (por exemplo, 15 minutos). Os preços de abertura, fechamento, máximo e mínimo de uma ação podem ser representados graficamente em um *candlestick*, conforme Figura 01, que tem esse nome porque seu formato lembra o de uma vela. Os valores de abertura e fechamento do período de tempo escolhido são representados por linhas horizontais, e o retângulo delimitado por estas linhas é pintado de verde se o preço do fechamento é maior que o de abertura (indicando que houve um aumento do preço), e pintado de vermelho se o preço de fechamento é menor do que o de abertura (indicando que houve uma queda no preço). Os limites superior e inferior mostram os valores máximo e mínimo, respectivamente, atingidos durante o período escolhido. Todos os valores atingidos pela ação ao longo deste período estão contidos na linha vertical delimitada pelos valores de máxima e mínima.

**Figura 01** – Representação de um *candlestick*.



Fonte: autor.

Os valores representados em um *candlestick* permitem algumas interpretações sobre o comportamento de uma ação. Por exemplo, se o tamanho da linha vertical inferior for muito grande em relação ao resto do *candlestick*, significa que o valor da ação caiu muito durante o período, mas que esta queda se reverteu, o que pode significar o início de uma tendência de alta. Por esta razão, os valores representados em *candlestick* estão geralmente entre os mais utilizados como entrada para redes neurais de previsão ou para indicadores de ações.

## 2.2 Séries Temporais

Uma série temporal é qualquer conjunto de observações ordenadas no tempo (MORETTIN; TOLOI, 2006). Para Souza (1989) uma série temporal é uma classe de fenômenos cujo processo observacional e consequente quantificação numérica geram uma sequência de observações distribuídas ao longo do tempo. Podemos citar alguns exemplos, tais como: valores diários de poluição em uma cidade, índices diários da Bolsa de Valores de São Paulo, acidentes ocorridos nas rodovias da cidade de São Paulo durante um mês, etc. Ao se estudar séries temporais univariadas, pode-se ter vários objetivos, conforme Morettin e Tolo (2006), como: i) avaliar a natureza e a estrutura de seu mecanismo gerador, os quais estão relacionados com o intervalo de

ocorrência das observações no tempo, podendo ser contínua ou discreta; ii) realizar previsões com um ou vários passos à frente; iii) simplesmente descrever o comportamento da série, utilizando gráficos, verificando a existência de tendências, ciclos, variações sazonais, pontos influentes entre outros; iv) procurar a periodicidade relevante nos dados.

Os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas. Em estatística, um *Processo* é um evento qualquer que evolui ao longo do tempo. Uma variável que tem seu valor alterado ao longo do tempo de forma aleatória é dita seguir um *Processo Estocástico*, ou seja, um conjunto de observações indexadas pelo tempo. As alterações sofridas em seus valores podem ter uma natureza contínua ou discreta ao longo do tempo. Além disso, os valores em si possíveis de serem assumidos por ela também podem ter uma natureza contínua ou discreta (RESENDE, 2016). Ao se analisar uma ou mais séries temporais a representação gráfica dos dados sequencialmente ao longo do tempo pode revelar padrões de comportamento importantes, sendo eles: tendência de crescimento (ou decrescimento), que diz respeito ao sentido do deslocamento da série durante um longo espaço de tempo; padrões cíclicos, que são os movimentos ondulatórios que a série temporal apresenta a cada intervalo de tempo; sazonalidade, que são os efeitos que a série temporal apresenta em decorrência das variações climáticas e o ruído aleatório, que são os dados que não podem ser modelados, pois são gerados por influências externas ao processo, é também chamado de erro. (MORETTIN; TOLOI, 2006).

Além dos elementos apresentados no parágrafo anterior, as seguintes características também podem ser identificadas nas séries temporais financeiras, sendo elas:

**Pontos influentes** – são valores incomuns, que fogem do padrão do restante da série temporal. Em séries financeiras, podem ser considerados pontos influentes momentos de alta volatilidade no mercado, gerando grandes altas ou quedas nos preços, que logo depois voltam ao seu patamar normal;

**Heteroscedasticidade condicional** – a variância para os valores de entrada e saída da série temporal não é constante com o passar do tempo, tornando o comportamento da série temporal mais aleatório;

**Não-linearidade** – devido à sua complexidade e ao seu comportamento estocástico, não é possível modelar este tipo de série temporal com uma função linear.

### 2.2.1 Séries Temporais Financeiras - Retorno

A maior parte dos estudos financeiros concentra-se na análise da série de retornos ao invés do uso da série de preços. A razão desta preferência, conforme Campbell *et al* (1997), está relacionada a dois fatos. Em primeiro lugar o retorno de um ativo financeiro contém as informações que atendem aos interesses dos investidores. Em segundo lugar a série de retornos possui propriedades estatisticamente mais atrativas que a série de preços.

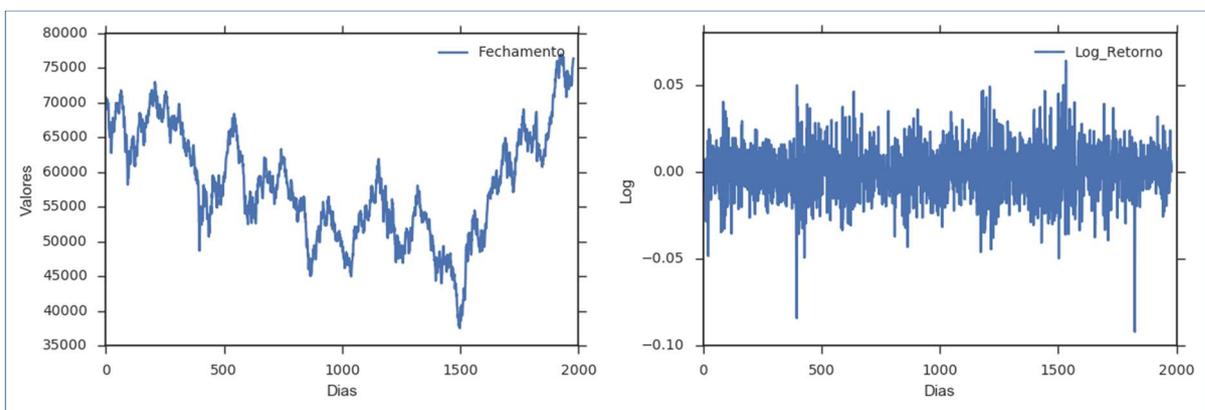
O retorno de um ativo entre os instantes de tempo entre  $t-1$  e  $t$  é dado por:

$$r_{i,t} = \frac{\ln(P_{i,t})}{\ln(P_{i,t-1})} = p_{i,t} - p_{i,t-1} \quad (04)$$

Onde  $r_{i,t}$  é o retorno logarítmico do ativo  $i$ , no período  $t$ ;  $P_{i,t}$  é o preço do ativo  $i$ , no período  $t$ ;  $P_{i,t-1}$  é o preço do ativo  $i$ , no período  $t-1$ ;  $p_{i,t}$  é o logaritmo do preço do ativo  $i$ , no período  $t$ ; e  $p_{i,t-1}$  é o logaritmo do preço do ativo  $i$ , no período  $t-1$ .

Para exemplificar, a Figura 02 mostra a série temporal do índice Bovespa apresentando o gráfico com o valor de fechamento do dia e o retorno logarítmico para o mesmo período.

**Figura 02** – Índice Bovespa: Fechamento e Retorno Logarítmico.



Fonte: Autor

## 2.3 Aprendizagem de Máquina

Após o final da Segunda Guerra Mundial, o conceito de Inteligência Artificial (IA) conforme Norvig e Russel (2004) abrange uma série de subcampos, sistematizando e automatizando tarefas intelectuais, da criação de sistemas que pensam como seres humanos.

Geralmente, para implementar a Inteligência Artificial (IA), usamos o Aprendizado de Máquina, sendo representado por diversos algoritmos, baseados na teoria de álgebra, probabilidades, estatística e otimização, capazes de extrair conhecimento dos dados, por exemplo: *decision tress*, *random forests*, *artificial neural networks*, entre outros (MOHRI *et al.*, 2012). Aprendizado de máquina é uma área que trabalha com otimizações matemáticas com o objetivo de reconhecimento e detecção de padrões. Em termos gerais, técnicas de aprendizado de máquina tentam modelar um sistema, ou prever resultados futuros, com informações de saídas ou estados passados.

Existem diversos algoritmos de aprendizado de máquina desenvolvidos para diferentes objetivos. É importante ressaltar que Aprendizagem de Máquina segue um raciocínio de metodologias orientadas aos dados. Podemos pensar na Aprendizagem de Máquina como um ramo da estatística e da computação que reúne uma série de métodos que possuem dois objetivos principais. O primeiro é o desempenho preditivo de modelos e o segundo é automatizar o processo de modelagem das bases de dados observadas ou aprendizado com os dados observados. Além disso, boa parte envolve métodos de calcular o erro de validação, e selecionar ou ponderar modelos com base nesses erros de previsão (VASCONCELOS, 2017).

Os algoritmos de Aprendizagem de Máquina, ou métodos estatísticos, podem ser divididos em aprendizado supervisionado, aprendizado não supervisionado e aprendizagem por reforço. Na aprendizagem supervisionada reúne métodos de estimação onde se tem pelo menos uma variável dependente, ou variável de resposta, que é explicada. Isso quer dizer que o que se está querendo modelar é objetivo e mensurável e pode ser previsto ou projetável. O aprendizado supervisionado envolve problemas de classificação, estimação da distribuição a posteriori interpolação (métodos não lineares) e regressão linear. Por outro lado, o aprendizado não supervisionado estuda métodos onde não se tem uma variável objetivo ou variável de

resposta que é usada para supervisionar o modelo, como análise de *cluster* por exemplo (BISHOP, 2006; SUGIYAMA, 2016).

O aprendizado por reforço se baseia no aprendizado animal que ocorre por meio de tentativas e erros, isto é, o aprendizado é realizado através da interação com o ambiente. Enquanto o aprendizado supervisionado depende do especialista, através do conjunto de exemplos de entradas em saídas, o aprendizado por reforço não precisa de um “supervisor” e aprende através de suas próprias experiências (KAELBLING; LITTMAN; MOORE, 1996).

### **2.3.1 Redes Neurais Artificiais**

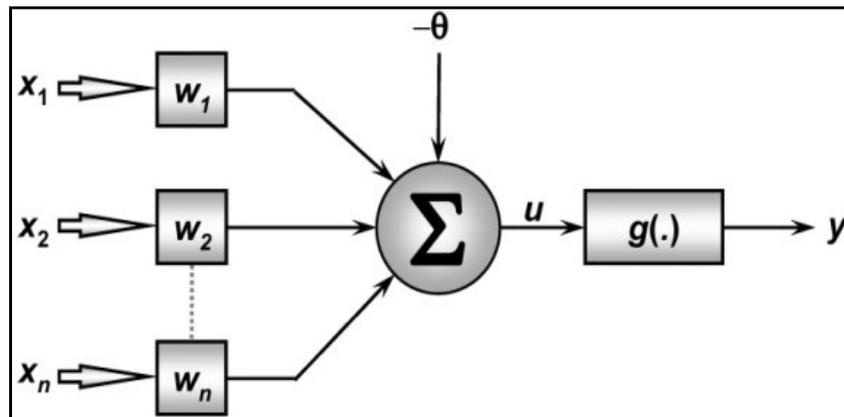
As Redes Neurais Artificiais (RNAs) foram desenvolvidas na década de 1940, por Walter Pitts e McCulloch, o primeiro matemático e o segundo neurofisiologista, com o objetivo de desenvolver um modelo computacional análogo a neurônios biológicos e circuitos eletrônicos, capazes de simular conexões sinápticas pelo uso de resistores variáveis e amplificadores (LUDWIG; COSTA, 2007).

Os primeiros trabalhos acadêmicos sobre RNAs foram publicados na década de 1960, porém foi a partir da década de 1990 que as RNAs começaram a ser largamente estudadas e aplicadas em diversas áreas (MARTINS; METTE; MACEDO, 2008). Em sua grande maioria, as RNAs são aplicadas a problemas relacionados à clusterização, ao reconhecimento e à previsão de séries temporais. Uma das aplicações nesta área é a previsão de ações no mercado financeiro. A sua maior vantagem é em virtude de sua capacidade de capturar os padrões não lineares normalmente existentes nas séries temporais (KHASHERI; BIJARI, 2010; ZHANG, 2003). Desta forma as RNAs vêm auxiliar em problemas que envolvem sistemas complexos como este, pois são capazes de extrair características relevantes de grandes conjuntos de dados, baseando-se em conhecimentos prévios (OLIVEIRA, 2007).

As RNAs são formadas por neurônios interligados entre si simulando sinapses artificiais estruturadas, baseado em modelos biológicos. Desta forma, as RNAs possuem características como capacidade de aprendizagem através de treinamento, habilidade de generalização e tolerância a falhas (MARTINS; METTE; MACEDO, 2008; HAYKIN, 2001). Um neurônio pode ser visualizado conforme a Figura 03, com

os  $x_i$  sinais de entrada externos e os  $w_i$  pesos sinápticos, que ponderam e variam de acordo com a relevância de cada  $x_i$ . Em seguida é realizada a soma ponderada dos sinais de entrada, realizada pelo combinador linear  $\Sigma$ , e subtrai-se a variável  $\theta$  chamada de limiar de ativação. Desta forma, a saída de cada neurônio é denominada de potencial de ativação  $u$ . As saídas dos neurônios artificiais são enviadas para uma função de ativação, denominada por  $g()$ , a qual tem a função de limitar a saída do neurônio dentro de um intervalo determinado que possa ser assumido pela rede e gera o sinal de saída  $y$  (SILVA; SAPATTI; FLAUZINO, 2010).

**Figura 03** – Neurônio Artificial.



Fonte: Silva, Sapatti e Flauzino (2010).

As principais funções de ativação  $g()$ , são as funções logística, Equação 05, sendo que o resultado de saída produzido assumirá sempre valores reais entre zero e um; a função tangente hiperbólica, Equação 06, apresentará o resultado de saída sempre valores reais entre -1 e 1; a função gaussiana, Equação 07, que a saída do neurônio produzirá resultados iguais para aqueles valores de potencial de ativação  $\{\mu\}$  que estejam posicionados a uma mesma distância de seu centro (média) e a função linear, Equação 08, produz resultados de saída idênticos aos valores do potencial de ativação  $\{\mu\}$  (SILVA; SAPATTI; FLAUZINO, 2010).

$$g(u) = \frac{1}{1+e^{-\beta \cdot u}} \quad (05)$$

Onde  $\beta$  é uma constante real associada ao nível de inclinação da função logística frente ao seu ponto de inflexão.

$$g(u) = \frac{1-e^{-\beta \cdot u}}{1+e^{-\beta \cdot u}} \quad (06)$$

Onde  $\beta$  também está associado ao nível de inclinação da função tangente hiperbólica em relação ao seu ponto de inflexão.

$$g(u) = e^{-\frac{(u-c)^2}{2\sigma^2}} \quad (07)$$

Onde  $c$  é um parâmetro que define o centro da função gaussiana e  $\sigma$  denota o desvio padrão associado à mesma, isto é, a quão espalhada (dispersada) está a curva em relação ao seu centro.

$$g(u) = u \quad (08)$$

Uma RNA pode ser dividida em três camadas básicas: entrada, oculta e saída. A camada de entrada é responsável por receber os sinais  $x_i$ . As camadas ocultas ou intermediárias são compostas pelos neurônios artificiais responsáveis pelo processamento interno da rede. A camada de saída é composta pelos neurônios artificiais responsáveis por gerar o sinal  $y$  (HAYKIN, 2001).

Os possíveis arranjos de neurônios são chamados de arquitetura da RNA. As diferentes formas de composição estrutural que uma arquitetura pode assumir é a topologia da rede (HAYKIN, 2001). A forma como as interconexões de neurônios serão estruturadas pode variar de acordo com o número de camadas, o número de neurônios em cada camada, a função de ativação dos neurônios em uma camada e a forma como as camadas são conectadas, podendo ser de forma total ou parcial (ROCA; MÓL, 2015). O mecanismo de funcionamento das RNAs consiste em aprender a relação intrínseca existente nos dados através de um número de neurônios que se comunicam entre si, distribuídos em diferentes camadas (ADHIKARI; AGRAWAL, 2014).

Muitas redes, com diferentes topologias e formas de treinamento, foram desenvolvidas, ao longo do tempo, para se adaptar ao problema a ser resolvido nas

mais diversas áreas das ciências. As aplicações potenciais das RNA podem ser classificadas da seguinte forma, segundo Silva *et al.* (2010): Aproximador universal de funções; Controle de processos; Reconhecimento/classificação de padrões; Agrupamento de dados (clusterização); Sistemas de previsão; Otimização de sistemas e Memórias associativas.

As principais arquiteturas de redes neurais artificiais, considerando a disposição de seus neurônios, assim como suas formas de interligação entre eles e a constituição de suas camadas, podem ser divididas em: redes *feedforward* (alimentação à frente) de camada simples, redes *feedforward* de camadas múltiplas, redes recorrentes e redes reticuladas.

Uma rede *feedforward* (FF) de camada simples não possui camadas intermediárias, tendo apenas uma camada de entrada e uma de saída. Já a arquitetura *feedforward* de camadas múltiplas possui uma ou mais camadas escondidas. As arquiteturas *feedforward* possuem um fluxo de dados em um único sentido, da camada de entrada para a camada de saída (ROCA; MÓL, 2015; SILVA; SAPATTI; FLAUZINO, 2010). Para uma rede FF resolver um problema, deve-se fornecer a essa rede os dados de entrada para os neurônios da primeira camada. Estes dados vão sendo propagados ao longo da rede até chegarem à camada de saída, que vai informar o resultado obtido.

O grande diferencial de uma rede neural em relação a outros modelos matemáticos é a sua capacidade de aprender com o passado. Uma rede neural é capaz de, ao longo de um processo de aprendizado, realizar tal feito por meio do ajuste dos seus pesos sinápticos até obter uma resposta satisfatória.

Normalmente, na construção da RNA divide-se aleatoriamente a série temporal disponível em duas partes: uma parte maior (entre 60% e 90%) que será utilizada para treinar a rede e uma parte menor (entre 10% e 40%) que será utilizada para avaliações e testes. Na etapa de treinamento, a rede recebe uma amostra de valores de entradas e os respectivos resultados esperados, no caso do treinamento supervisionado. Os pesos das correlações e os limiares são estabelecidos nesta etapa. Tal processo visa balancear a RNA de forma que suas respostas estejam próximas aos valores desejados. Na etapa de teste e validação consiste em prever os próximos valores da série temporal. Dessa forma os valores históricos são comparados com os valores

previstos pela RNA e pôde-se calcular o erro obtido pelas previsões da rede. (ROCA; MÓL, 2015; HAYKIN, 2001).

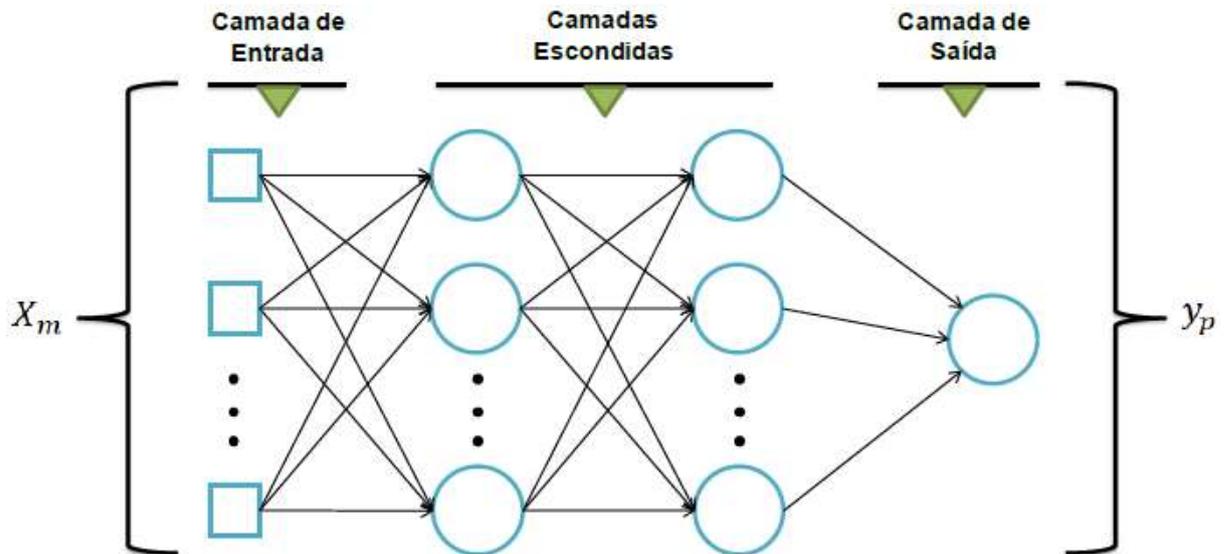
Existem várias técnicas matemáticas voltadas para a previsão (ou generalização) de séries temporais e valores, como por exemplo: regressões lineares, regressões logísticas e o método dos mínimos quadrados. Estes métodos, contudo, não se comportam bem quando a série apresenta algum ruído, ou não alcançam desempenho satisfatório quando o número de variáveis de entrada é muito grande. As redes neurais artificiais surgem como uma alternativa para este tipo de problema, pois além de contornarem as limitações apresentadas, conseguem modelar problemas não lineares e identificar características de sazonalidade, tendência e periodicidade, comumente presentes em séries temporais financeiras.

Em trabalhos voltados à previsão de ações por meio do uso de redes neurais, os tipos de entrada mais fornecidos a essas redes são entradas relacionadas às próprias séries temporais financeiras, séries temporais externas ao mercado sendo analisado (OLIVEIRA *et al.*, 2011) e as mais variadas combinações de indicadores técnicos (KARA; BOYACIOGLU; BAYKAN, 2011). Cada uma destas entradas é fornecida para um nodo diferente na camada de entrada de uma rede neural, que é treinada para retornar uma previsão de valor ou tendência futura.

### **2.3.1.1 Rede Neural *MultiLayer Perceptron***

A rede neural *MultiLayer Perceptron* (MLP) são caracterizadas pela presença de pelo menos uma camada intermediária (escondida) de neurônios, situada entre a camada de entrada e a respectiva camada neural de saída. Tem a vantagem de ser um método simples, em termos de complexidade computacional. A Figura 04 mostra a arquitetura de uma rede neural MLP com uma camada de entrada, duas camadas escondidas, e uma camada de saída.

**Figura 04** – Arquitetura de uma rede neural *multilayer perceptron* (MLP) com duas camadas escondidas.



Fonte: adaptado de Silva *et al.* (2010)

Na rede neural MLP, no que diz respeito a sua topologia, as definições do número de camadas escondidas, bem como, ao número de neurônios nessas camadas são definidos de forma empírica.

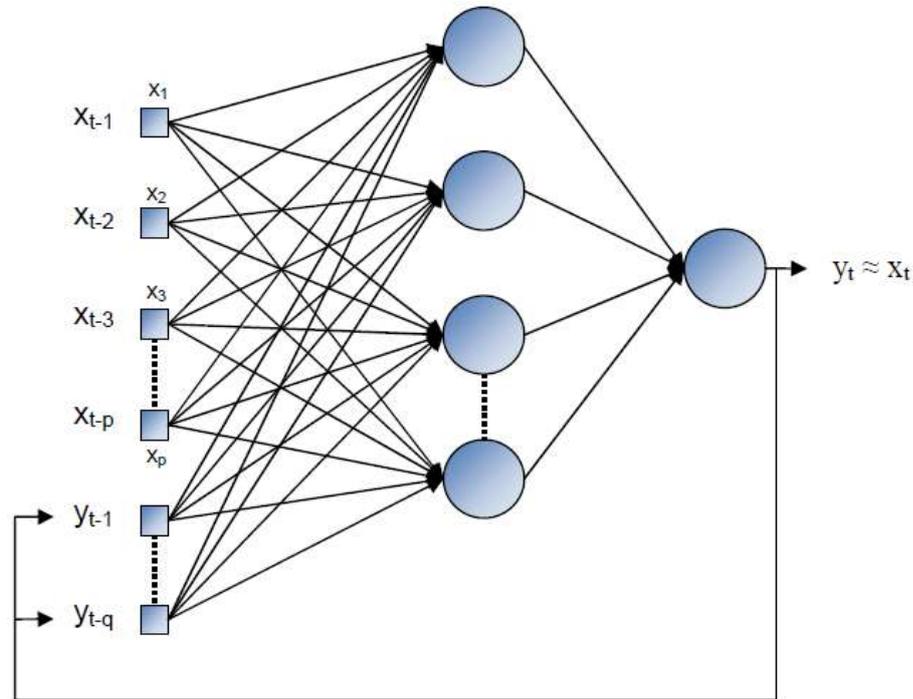
Um dos algoritmos mais utilizados no treinamento desse tipo de rede é o *backpropagation* (retropropagação). Esse algoritmo é baseado na regra de aprendizagem por correção de erro e consiste na aplicação de dois passos através das diferentes camadas: um passo à frente, a propagação, e um passo atrás, a retropropagação.

### 2.3.1.2 Rede Neural NARX

A rede neural com o modelo auto-regressivo não-linear com entradas exógenas (*Nonlinear Auto-Regressive with exogenous inputs* – NARX), é identificada como uma rede recorrente cuja aplicabilidade direciona-se ao mapeamento de sistemas dinâmicos tipicamente não-lineares (SILVA *et al.*, 2010). Assim como a MLP com entradas atrasadas no tempo, o treinamento da NARX é também efetuado de maneira similar ao MLP convencional. Considerando a topologia apresentada na Figura 05,

pode-se observar que o modelo recebe reentradas do próprio sistema, aumentando a capacidade de mapeamento não-linear da rede (HAYKIN, 2001).

**Figura 05** – Arquitetura de uma rede neural NARX.



Fonte: adaptado de Haykin, 2001 e Silva et al., 2010.

Os valores previstos pela rede NARX, portanto, não apenas se baseiam em  $p$  valores passados, mas também em  $q$  valores previstos anteriormente, ou seja:

$$x(t) = f(x(t-1), x(t-2), \dots, x(t-n_p), y(t-1), y(t-2), \dots, y(t-n_q)) \quad (09)$$

Onde  $n_p$  é a ordem do predição e indica o número de medidas (amostras) anteriores que serão necessárias para a estimação de  $x(t)$ . O valor  $n_q$  expressa a ordem de contexto, ou seja, a quantidade de saídas passadas que serão também utilizadas na estimação de  $x(t)$ .

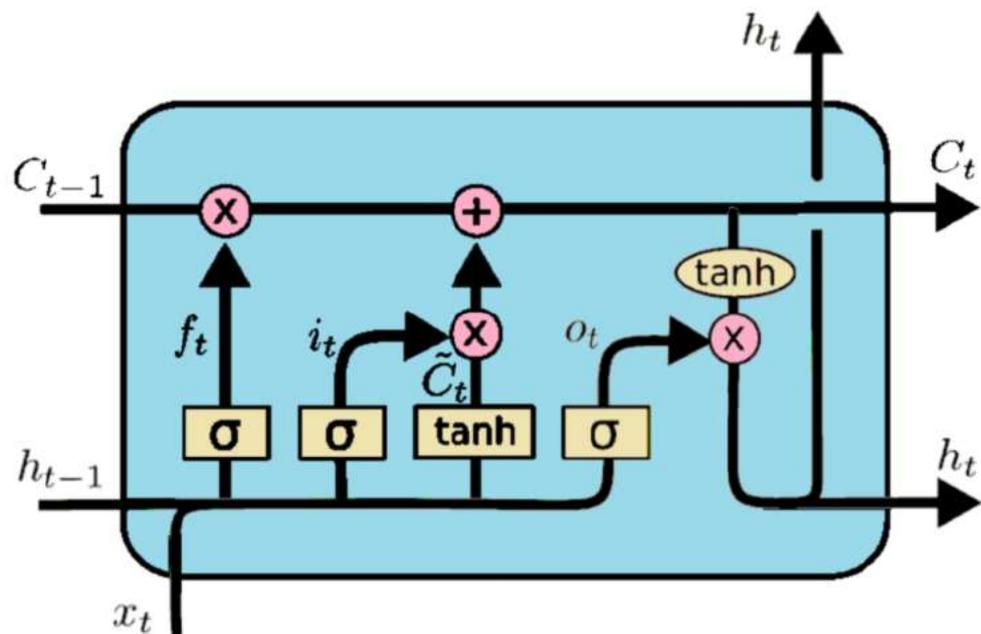
### 2.3.1.3 Rede Neural com Long-Short Term Memory

As redes neurais com células *Long-Short Term Memory* (LSTM) é uma variação das redes recorrentes que busca dirimir a dificuldade quando se trata de aprender dependências após vários estágios de processamento. Este problema está relacionado ao fato de que após diversas interações provocando atualizações no

estado oculto da rede neural recorrente, muita informação importante do começo de uma sequência pode se perder no final ou em algum momento em que ela será necessária (GOODFELLOW, *et al.* 2016; GREFF, *et al.* 2016; OLAH, C. 2015).

Uma rede neural LSTM possui as mesmas propriedades que uma rede neural recorrente padrão, contudo é capaz de armazenar informações por longos períodos de tempo ao processar uma sequência. Os nós de memória de uma LSTM são chamados de “células”. Células possuem propriedades mais complexas que suas equivalentes nas redes neurais recorrentes tradicionais. Estas são capazes de carregar informações até o final do processamento de uma sequência bem como selecionar informações que devem ser “esquecidas” a partir de um certo ponto. As Figuras 06 e 07 representam o funcionamento de uma célula LSTM.

**Figura 06** – Representação de uma célula LSTM.



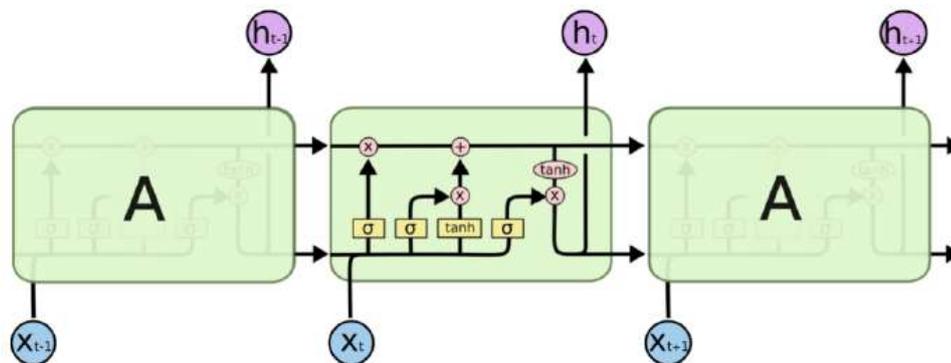
Fonte: adaptado de OLAH, C. 2015

No instante  $t$ ,  $C_t$  representa o estado de célula,  $h_t$  a saída da célula,  $x_t$  a amostra da sequência,  $f_t$  o *gate* de esquecimento,  $i_t$  o *gate* de entrada e  $o_t$  o *gate* de saída. Todos esses valores acabam sendo concatenados, multiplicados ou somados, conforme mostra o circuito desta ilustração.

Ainda analisando a Figura 06, percebe-se que existem cinco pontos de conexão com o exterior da célula, representados por:  $C_{t-1}$ ,  $h_{t-1}$ ,  $x_t$ ,  $C_t$  e  $h_t$ . Considerando que a célula em questão representa um momento  $t$  na sequência a ser processada pela

rede.  $C_t$  representa o estado da célula no instante  $t$  e este estado representa as informações que chegaram até este passo em instante de tempo passados, ao contrário de apenas a última atualização do estado. Conseguir guardar essa informação é o grande diferencial das redes LSTMs, pois ela consegue guardar informações relevantes do passado de uma sequência que podem só fazer sentido no futuro da mesma. Essa informação pode ser alterada através das Portas (ou *Gates*) de esquecimento e de aprendizado.

**Figura 07** – Representação de uma célula LSTM sendo replicada no tempo.



Fonte: adaptado de OLAH, C. 2015

### 2.3.2 Ensembles de Rede Neurais

*Ensemble* é uma palavra francesa que significa junto, comitê. A ideia geral é representar um conjunto de soluções combinadas para um único problema. Uma forma de melhorar os resultados fornecidos por uma rede neural é colocá-la em um *ensemble* de redes neurais. *Ensemble* é um paradigma de aprendizado em que propostas alternativas de solução para um problema, denominadas componentes, têm suas saídas individuais combinadas na obtenção de uma solução final.

Podemos dizer que um classificador *ensemble* é um conjunto de classificadores cujas decisões individuais são combinadas de alguma forma (normalmente por votação ponderada ou não ponderada) para classificar novos exemplos. Métodos *ensemble* usam essa combinação para melhorar o desempenho preditivo que poderia ser obtido de qualquer um dos classificadores individuais (DIETTERICH, 2000).

A ideia é que redes neurais diferentes entre si (tanto na organização de sua arquitetura quanto na inicialização dos pesos iniciais de seus neurônios) retornem resultados diferentes para o mesmo conjunto de entradas. O uso de várias redes neurais para calcular o mesmo problema adiciona uma redundância ao sistema, tornando possível que redes neurais que façam previsões erradas tenham seus resultados ignorados, pois o conjunto geral obtém um resultado diferente. Muitos métodos para construir *ensembles* têm sido desenvolvidos, os quais podem ser aplicados a diferentes algoritmos de aprendizagem.

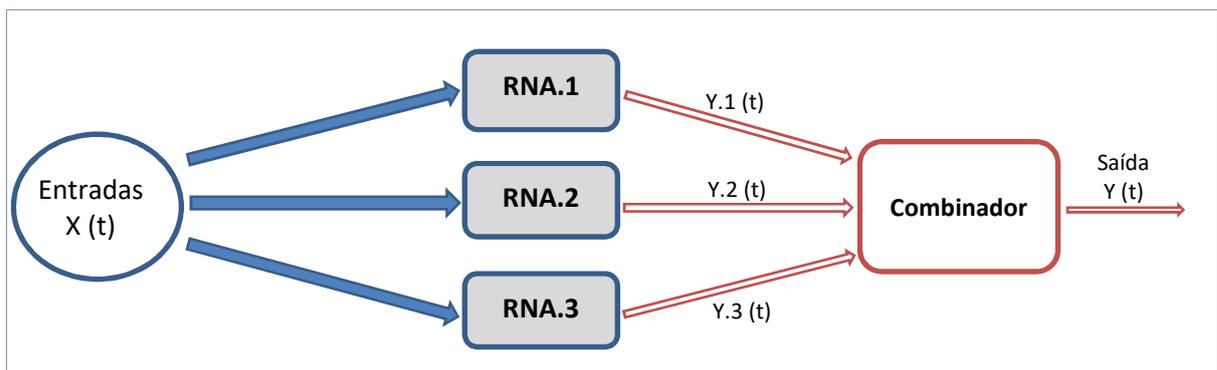
A abordagem de *ensembles* tem sido amplamente utilizada nas últimas duas décadas, tanto para problemas de regressão quanto para problemas de classificação de padrões, já que os *ensembles* são comprovadamente capazes de aumentar a capacidade de generalização e, conseqüentemente, o desempenho geral do sistema (HANSEN; SALAMON, 1990; HASHEN et al. 1994; SHARKEY, 1999). Atualmente, abordagens que utilizam fusão de informação vêm sendo utilizadas amplamente em problemas como: classificação (STEFKA; HOLENA, 2015; MOUSAVI; EFTEKHARI, 2015; SEOKHO; CHO; KANG, 2015), tomada de decisão (POLIKAR, 2006; TAPIA-ROSETO *et al.*, 2016), predição de séries temporais (MERIGÓ; PALACIOS-MARQUES; RIBEIRO-NAVARRETE, 2015; KOURENTEZES; PETROPOULOS, 2015), dentre outros. No entanto, tal melhora na capacidade de generalização se apoia na qualidade de seus componentes e na diversidade do erro apresentada por eles (PERRONE; COOPER, 1993).

Cada um dos componentes em um *ensemble* deve apresentar um bom desempenho quando aplicado isoladamente ao problema e, ao mesmo tempo, deve “cometer erros” distintos quando comparados aos demais componentes. Essa necessidade de diversidade do erro dos componentes é, de certa forma, intuitiva. Se forem combinados componentes que apresentam um mesmo padrão de erro, claramente não haverá nenhum incremento de desempenho. Esta combinação trará apenas um aumento no custo computacional, sem resultados práticos de desempenho.

Os estudos iniciais sobre a combinação de componentes para problemas de regressão foram feitos paralelamente por Perrone (1993) e Hashem (1993; 1997), e tornaram-se um tópico intensamente investigado nos anos subsequentes. Esse interesse acabou por contribuir muito para o amadurecimento do conceito de

diversidade de erros em regressores. A Figura 08 ilustra a aplicação do método *Ensemble* para a combinação de três redes neurais artificiais produzindo resultados diferentes para uma mesma solução. Os resultados são combinados para produzir uma única solução com alta capacidade de generalização (HANSEM; SALAMON, 1990).

**Figura 08** - Exemplo de combinação de três RNAs.



Fonte: autor.

A ideia é combinar redes que constituem soluções diferentes obtendo diferentes padrões de erro quando apresentadas a um mesmo conjunto de teste (SHARKEY, 1999). Existe um número de parâmetros de treinamento que pode ser manipulado para a obtenção de tais redes, como exemplo: condições iniciais, dados de treinamento, topologia das redes e o algoritmo de treinamento.

Os classificadores *ensemble* predizem as classes por meio de uma combinação dos votos dos classificadores: média, média ponderada, votação simples e ponderação das saídas (ROKACH, 2005).

## 2.4 Métricas de Desempenho

Nesta seção estão as funções mais comuns utilizadas para avaliar o desempenho de modelos de regressão. Na regressão buscamos prever um valor numérico, como exemplo, o valor de fechamento uma ação para o dia seguinte. Para quantificar o desempenho de uma rede neural na previsão de séries temporais costuma-se utilizar métricas que analisam as diferenças entre a série de valores previstos e a série de valores reais. A seguir, são mostradas algumas das métricas mais utilizadas para

estes cálculos nos estudos pesquisados. Existem métricas que são utilizadas com maior frequência, porém não há uma que seja considerada melhor que as outras. A maioria dessas métricas é amplamente usada no campo da estatística e possuem autores desconhecidos. O objetivo dessa seção é listar algumas dessas métricas pesquisadas.

#### 2.4.1 Erro Médio Percentual Absoluto

O Erro Médio Percentual Absoluto (ou MAPE – mean absolute percentage error) obtém as diferenças percentuais entre todos os valores reais e previstos obtidos e faz uma média simples destes valores. Como todos os elementos da série temporal têm igual peso no resultado final, resultados isolados muito diferentes dos demais não fazem tanta diferença. Essa métrica é útil para se ter uma visão geral do erro médio gerado pelo algoritmo de previsão escolhido. O MAPE é calculado conforme Equação 10, na qual  $P_{real,t}$  é o valor real da série no dia  $t$  e  $P_{previsto,t}$  é o valor previsto para série, também no dia  $i$ .

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left( \frac{|P_{real,t} - P_{previsto,t}|}{P_{real,t}} \right) * 100 \quad (10)$$

#### 2.4.2 Erro Quadrático Médio

O erro quadrático médio, ou MSE (*Mean Squared Error*) é talvez a métrica mais utilizada para calcular o desempenho de modelos de previsão. É calculado a partir da soma da variância e dos quadrados das diferenças obtidas entre os valores reais e previstos, e é dado pela Equação 11 onde  $N$  é o número de dias da série temporal sendo analisada,  $real_t$  é o valor da série no dia  $t$  e  $previsto_t$  é o valor previsto para a série também no dia  $t$ .

$$MSE = \frac{1}{N} \sum_{t=1}^N (real_t - previsto_t)^2 \quad (11)$$

### 2.4.3 Raiz do Erro Quadrático Médio

Essa métrica, também chamada de RMS ou RMSE (Root Mean Squared Error), é semelhante ao MAPE no sentido de que obtêm o erro de todos os valores calculados. A diferença está no fato do erro calculado não ser um percentual, mas sim um valor numérico que indica o tamanho do erro médio obtido. O quanto mais próximo de zero for o erro, melhor são os resultados obtidos. O RMS é calculado a partir da Equação 12, onde  $N$  é o número de dias da série temporal sendo analisada,  $real_t$  é o valor real da série no dia  $t$  e  $previsto_t$  é o valor previsto para a série também no dia  $t$ .

$$RMSE = \sqrt{\frac{1}{N} \sum (real_i - previsto_i)^2} \quad (12)$$

Neste capítulo foram apresentados conceitos de fundamental importância para o decorrer deste trabalho. Destacam-se as formas de análise de ações e as diferenças entre as duas escolas: fundamentalista e técnica. Com relação aos modelos de aprendizagem de máquina, mas especificamente, redes neurais artificiais, apresentamos as arquiteturas que foram utilizadas neste trabalho e suas características essenciais. E, por fim, as métricas a serem adotadas para medição do desempenho dos modelos avaliados. Nos próximos capítulos apresentamos o método de pesquisa adotado, e, em seguida, os resultados e a conclusão do trabalho.

### 3. MÉTODO DE PESQUISA

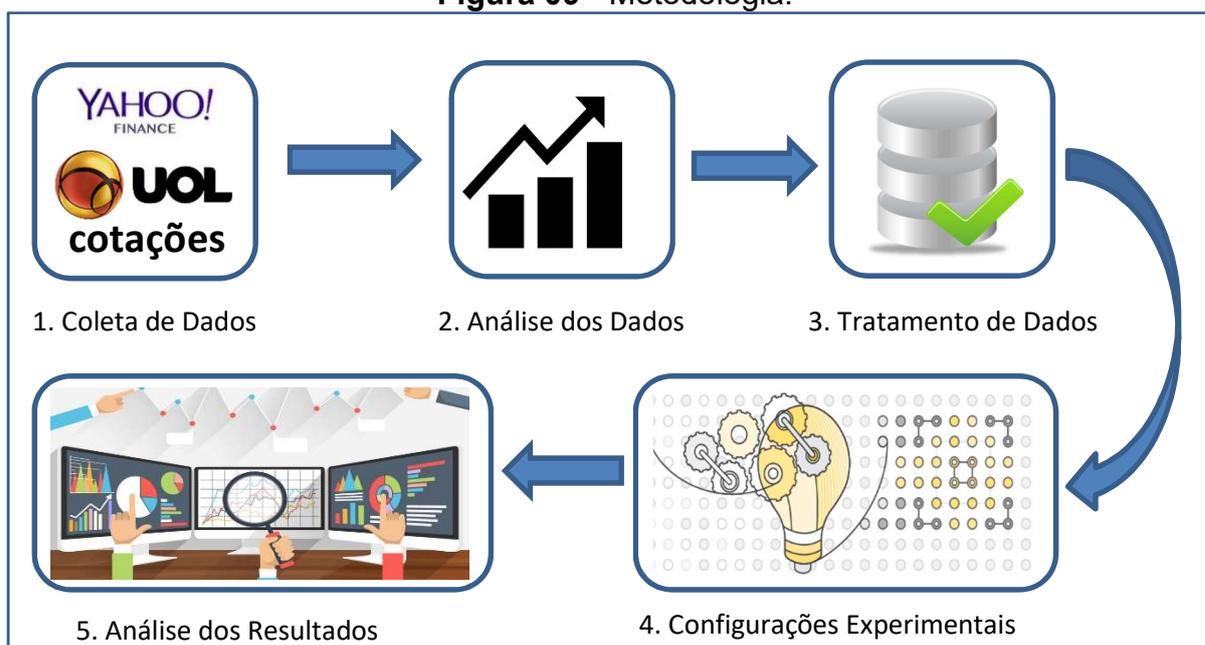
Quanto à abordagem do problema, utilizou-se a pesquisa quantitativa por fazer uso das modalidades de coleta de informações e também de seus tratamentos. Richardson (1999, p. 70 citado por Beuren, 2006, p.92) enfatiza que a abordagem quantitativa caracteriza-se pelo emprego de quantificação tanto nas modalidades de coleta de informações, quanto no tratamento delas por meio de técnicas estatísticas, desde as mais simples, como percentual, média, desvio padrão, às mais complexas, como coeficiente de correlação, análise de regressão, etc.

A pesquisa quantitativa normalmente, mostra-se apropriada quando existe a possibilidade de medidas quantificáveis de variáveis e inferências a partir de amostras de uma população, usando instrumentos específicos. Esse tipo de pesquisa utiliza medidas numéricas para testar dados científicos e hipóteses, ou buscar padrões numéricos relacionados a conceitos cotidianos (GIL, 1999).

Para entendimento da operação dos algoritmos e dos resultados gerados são comentadas a seguir as séries financeiras analisadas e os métodos para previsão e negociação.

Este capítulo descreve a metodologia aplicada neste trabalho, a qual é ilustrada pela Figura 09. Cada etapa será explicada nas próximas seções.

**Figura 09 - Metodologia.**



Fonte: autor.

### 3.1 Séries financeiras analisadas

O conjunto de dados utilizado neste trabalho consiste de dados reais da Bolsa de Valores de São Paulo (B3). Foi extraído do Yahoo! Finance<sup>1</sup> (YAHOO, 2017) as séries temporais com periodicidade diária do Ibovespa referentes ao período de 04 de janeiro de 2010 a 27 de dezembro de 2017, perfazendo uma janela de estudos possível para até 8 anos.

**Tabela 01** - Subconjunto da amostra com os 5 primeiros registros do arquivo – Ibovespa.

Data	Abertura	Mínima	Máxima	Fechamento	Fechamento Ajustado	Volume
04-01-2010	68587	70081	68587	70045	70045	1655400
05-01-2010	70046	70595	69928	70240	70240	1984200
06-01-2010	70237	70937	70016	70729	70729	2243600
07-01-2010	70723	70723	70045	70451	70451	1555000
08-01-2010	70455	70766	70158	70263	70263	1634400

Fonte: autor.

Segundo Keiel e Bender (2017) a utilização de variáveis exógenas na predição apresenta maior eficácia do que modelos sem variáveis externas. Nesse sentido, que a escolha do Dólar foi definida como variável externa para incluir nos modelos a serem avaliadas e confirmada a sua correlação negativa conforme Tabela 05. As cotações do Dólar foram obtidas através do site UOL Economia Cotações (UOL, 2018), referentes a 04 de janeiro de 2010 a 27 de dezembro de 2017, perfazendo a mesma janela de estudo do Índice Bovespa.

**Tabela 02** – Subconjunto da amostra com os 5 primeiros registros do arquivo cotação dólar.

Data	Compra	Venda	Variação %	Variação Valor	Status
04-01-2010	1,719	1,721	- 1,260	- 0,022	Baixa
05-01-2010	1,730	1,732	0,640	0,011	Alta
06-01-2010	1,736	1,738	0,350	0,006	Alta
07-01-2010	1,743	1,745	0,400	0,007	Alta
08-01-2010	1,728	1,730	- 0,860	- 0,015	Baixa

Fonte: autor.

<sup>1</sup> <https://finance.yahoo.com>

A próxima etapa foi unir as séries temporais em uma única, nessa etapa foi necessário trabalhar somente com os dados que possuem relação através do índice data contido na tabela Ibovespa.

**Tabela 03** - Subconjunto da amostra com os exemplos das variáveis

Abertura	Mínima	Máxima	Log_Ret	MA5	IFR	Dolar	Fechamento
70267	71068	70158	0,0024	70423,2	64,6969	1,735	70433
70429	70429	69284	-0,0051	70390,4	48,3510	1,746	70076
70081	70626	69535	0,0044	70321,6	59,4386	1,759	70385
70377	70508	69661	-0,0083	70191,6	39,4376	1,764	69801
69780	69788	68695	-0,0119	69934,6	25,7605	1,771	68978

Fonte: autor.

### 3.2 Análise descritivas dos dados

A seleção e testes das variáveis de entrada para as redes destinadas à previsão da série Ibovespa faz parte dos objetivos específicos deste trabalho. Conforme apresentado no Quadro 01 as seguintes variáveis foram selecionadas:

**Quadro 01** - Descrição das variáveis independentes e dependente (fechamento).

Variável	Descrição
Abertura	É o valor do índice no momento da abertura do pregão.
Mínima	Mostra o menor valor atingido do índice durante o dia.
Máxima	Mostra o valor máximo atingido do índice ao longo do período.
Log_Ret	Essa variável é apresentada como o Retorno Logarítmico, que consiste no logaritmo natural do retorno aritmético, conforme Equação 04.
MA5	Média Móvel para 5 dias, conforme Equação 01.
IFR	O IFR (Índice de Força Relativa), conforme Equação 02, está entre os indicadores mais utilizados na análise técnica.
Dolar	Cotação do Dólar para compra.
Fechamento	Mostra o valor que o índice atingiu ao final do dia do pregão.

Fonte: autor.

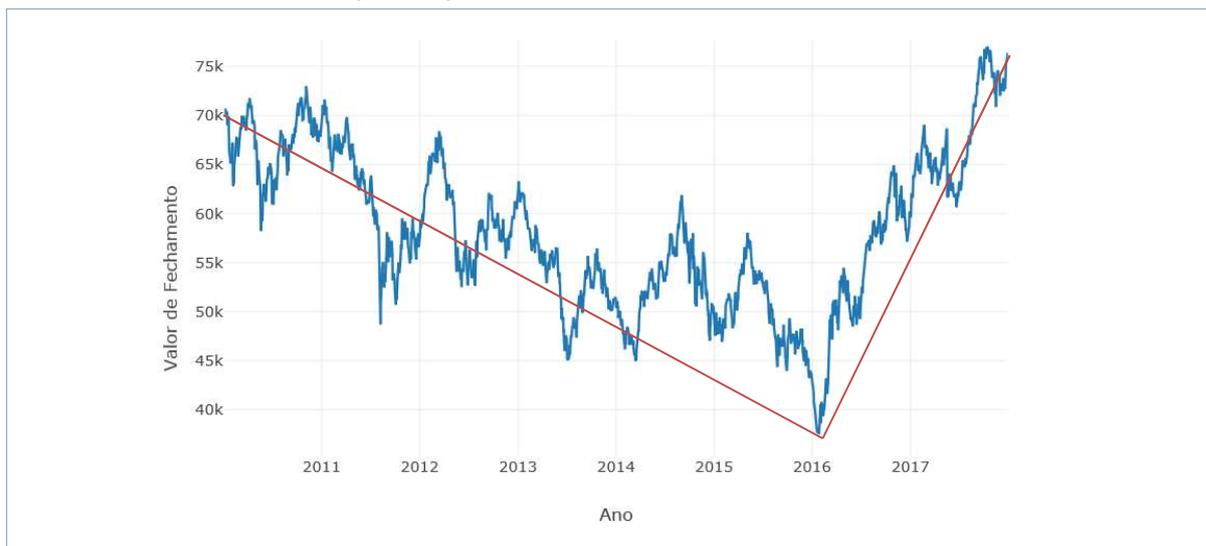
**Tabela 04** – Análise estatística da série temporal com o valor do fechamento do Índice Bovespa.

<b>Fechamento</b>	
Contagem	1969
Média	58124,672
Desvio Padrão	7956,766
Mínimo	37497
Máximo	76990

Fonte: autor.

A seguir serão apresentados gráficos para avaliação do comportamento do índice financeiro Ibovespa.

**Figura 10** - Série temporal financeira representando os valores de fechamento do índice Bovespa no período entre 04/01/2010 até 28/12/2017.



Fonte: autor.

É possível observar na Figura 10 que o índice Bovespa apresenta a partir de 2010 até 2016 uma tendência de baixa de longo prazo (superior a um ano), com picos de alta de curto (menos de um mês) e médio prazo (entre um a três meses); depois iniciou um período de altas. Vale ressaltar que em 31/08/2016 culminou no impeachment da Presidência da República do Brasil.

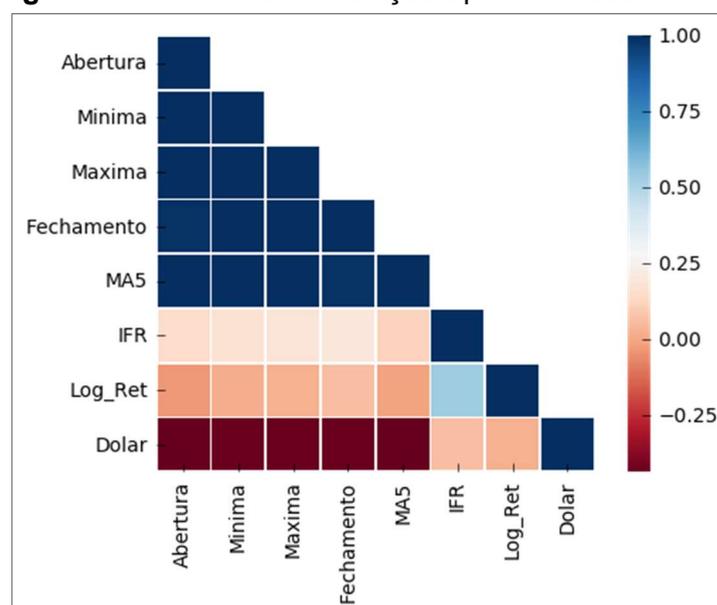
**Tabela 05** – Coeficiente de correlação (Pearson, Spearman) e coeficiente de determinação das variáveis independentes.

Variáveis	Variável Fechamento		
	Pearson	Spearman	Kendal
Abertura	0,994808	0,994322	0,936288
Mínima	0,998025	0,997820	0,963172
Máxima	0,998245	0,998080	0,964403
Log_Ret	0,056984	0,068992	0,045460
MA5	0,993826	0,993202	0,929872
IFR	0,276923	0,255475	0,174684
Dolar	-0,350720	-0,422588	-0,311758

Fonte: autor.

Conforme pode ser observado na Tabela 05 (demonstração da relação linear de Pearson e não linear de Spearman e Kendal), que representa o coeficiente de correlação, sendo 1 totalmente correlacionado e zero sem correlação, e na Figura 11, que representa em cores uma matriz de correlação, sendo que a diagonal principal representa a autocorrelação sendo igual a 1 e a variável mais representativa é aquela com o maior número de pixels escuros. Podemos notar as variáveis independentes Abertura, Máxima, Mínima, MA5 (Média Móvel para 5 dias) apresentam alto coeficiente de correlação positivo com a variável Fechamento, assim como, a variável exógena Dolar (representa o valor da cotação da Compra do Dólar) apresenta coeficiente negativo, para ambos os testes.

**Figura 11** - Gráfico de correlação Spearman das variáveis.



Fonte: autor.

### 3.3 Normalização e Tratamento de Dados

Para realizar as previsões, um conjunto de valores são usados como entradas para os modelos de previsão. Os valores estão disponíveis somente nos dias úteis. Os dados foram, então, normalizados (transformação linear) usando a Equação 13, onde  $N$  é o valor final normalizado,  $x_i$  é o valor original a ser normalizado, e  $Min$  e  $Max$  são, respectivamente, os valores máximo e mínimo do conjunto de dados a ser normalizado. Todos os valores das séries temporais são então transformados para a escala entre 0 e 1, conforme Tabela 06.

$$N = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (13)$$

**Tabela 06** – Exemplo dos valores normalizados usando Equação 14.

	Abertura	Mínima	Máxima	Log_Ret	MA5	IFR	Dolar	Fechamento
$\min(x)$	37501	38031	37046	-0,0921	37789,4	3,6903	1,550	37497
$\max(x)$	76985	78024	76659	0,0638	76728,0	96,4035	41,644	76990
$x_i$	70267	71068	70158	0,0024	70423,2	64,6969	1,735	70433
$N$	<b>0,8299</b>	<b>0,8260</b>	<b>0,8358</b>	<b>0,6061</b>	<b>0,8380</b>	<b>0,6580</b>	<b>0,0046</b>	<b>0,8339</b>

As series temporais passaram por alguns tratamentos de dados, sendo:

- I. Conversão das informações relacionadas a variável data que estavam no tipo *string* para o tipo *datetime* utilizando a função do framework pandas denominada *to\_datetime()*;
- II. Exclusão de todas as linhas que estavam com as variável Abertura com o valor *Null* (sem valor).
- III. Exclusão das linhas iniciais com valor *Null* nas variáveis IFR e MA5.

### 3.4 Python

A linguagem Python<sup>2</sup> foi escolhida para desenvolvimento dos scripts dos algoritmos com os modelos de Redes Neurais Artificiais e *Ensembles*. Python foi criada por Guido van Rossum, no Instituto Nacional de Pesquisa para Matemática e

<sup>2</sup> <https://www.python.org>

Ciência da Computação da Holanda (CWI), em 1991. É uma linguagem de *script* orientada a objetos, interpretada, *open-source* (desenvolvida sob licença de código aberto, de livre utilização e distribuição, sem nenhum custo), suportada por várias plataformas e escrita para ser de fácil aprendizado (PYTHON, 2018), além disso, dispões de vários módulos bastante convenientes o que faz dela uma opção que provê alta produtividade no desenvolvimento de aplicações.

### 3.5 Técnicas de Aprendizado de Máquina e Configurações Experimentais

Os dados gerados na etapa anterior conforme a Tabela 03, foram separados em subconjuntos de treinamento (2/3) e validação (1/3), sendo que desta amostra já foram retirados os dias a serem utilizados para teste (10 dias). Essa técnica tem como propósito avaliar a aptidão de generalização de cada modelo, apresentando dados para treinamento isolando os dados de teste.

A Figura 12 mostra como o conjunto de entrada é criado, utilizando uma janela deslizante de dias. Seguindo este exemplo, a primeira configuração do experimento usaria os 10 dias  $D_0, D_1, D_2, D_3 \dots D_9$  para compor o conjunto de teste.

**Figura 12** - Exemplo do subconjunto de Teste

Valores de Entrada				Valor Previsto
	Abertura	Alta	Baixa	Fechamento
$D_0$	$AB_0$	$AL_0$	$BA_0$	$FE_0$
$D_1$	$AB_1$	$AL_1$	$BA_1$	$FE_1$
$D_2$	$AB_2$	$AL_2$	$BA_2$	$FE_2$
$D_3$	$AB_3$	$AL_3$	$BA_3$	$FE_3$
$D_4$	$AB_4$	$AL_4$	$BA_4$	$FE_4$

Fonte: autor.

Foram realizados diversos testes com a amostra sem e com as variáveis acrescentadas a série original do Ibovespa. Com os indicadores técnicos, foram no total 13 indicadores avaliados, no entanto, somente o indicador IFR e MMS

apresentaram resultados satisfatórios. A séries temporais com a cotação do dólar apresentam os valores de compra e venda, mas somente uma das variáveis foi suficiente e em destaque o valor de compra.

Durante os testes foi avaliado também a utilização da série temporal com a quantidade de buscas realizadas no *Google Trends*<sup>3</sup> pelo filtro de pesquisa “Bovespa”, porém, nos testes não apresentaram resultados significativos ou até mesmo aumentaram os indicadores das métricas de avaliação dependendo do modelo de rede neural em estudo.

### 3.6 Método de previsão

O método de previsão foi dividido em duas fases. Na primeira fase foram realizados os tratamentos dos dados das séries, o treinamento das 3 (três) redes neurais identificadas por MLP, NARX e LSTM. Na segunda fase ocorre a combinação dos valores, como valor previsor. As previsões são, por fim, avaliadas seguindo as mesmas métricas de desempenho.

A determinação do modelo neural é uma tarefa fundamental em qualquer aplicação. Várias métricas são descritas na literatura que oferecem uma indicação das melhores estimativas dos parâmetros da rede, tais como os encontrados em Charytoniuk e Chen (2000), Hippert, Pedreira e Souza (2001) e Crone e Dhawan (2007) para aplicações com predição de séries temporais e Zang (2007) para aplicações diversas. No entanto, uma metodologia sistemática para obtenção dos valores ótimos não é disponível. Weigend e Gershefeld (1994) afirmam que não é possível construir um algoritmo de predição que seja universal, isto é, que seja capaz de prever qualquer tipo de série temporal. Essa afirmação contribui para que as aspirações de construir modelos de predição devam ser mais modestas. Em vez de buscar o conhecimento completo do futuro, o que se pode ter como objetivo é encontrar o modelo mais adequado para certos tipos de dados, e a definição de “mais adequado” pode ser explicado como o modelo que exige o mínimo de informação para descrever os dados (MENEZES JÚNIOR, 2012).

Tendo em vista o exposto, as configurações referentes ao número de neurônios, quantidade de dados para treinamento e validação, número de épocas,

---

<sup>3</sup> <https://trends.google.com.br/trends/>

quantidade de camadas escondidas e funções de ativação em cada camada, foram selecionados através de testes, através das quais, permitiu eliminar configurações que apresentassem baixos resultados avaliados nas métricas propostas (MSE, RMSE e MAPE).

A simulação foi realizada utilizando-se o ano de 2010 a 2017 (com exceção de 10 dias finais da amostra) como janela de calibração dos preditores. Para configuração das funções de ativação nas 3 camadas da rede MLP, foi submetido a testes a combinação de 3 funções de ativação que apresentam performance distintas (FACURE, 2018), sendo elas: Sigmoide (sigmoid), Tangente Hiperbólica (tanh) e Unidade Linear Retificada (relu). Conforme Quadro 02, que representa a combinação das 3 camadas pelas 3 funções de ativação resultando em 27 linhas ( $3^3$ ), segue os resultados dessa sequência de avaliações das funções de ativação.

A rede foi treinada utilizando o método de otimização *Adam*, que é um algoritmo para otimizar funções estocásticas, baseado em estimativas de baixa ordem. Também é apropriado para objetivos não estacionários e problemas com muito ruído e/ou gradientes esparsos. Os resultados empíricos demonstram que o *Adam* funciona bem na prática e o compara favoravelmente com outros métodos de otimização estocástica (KINGMA, D; BA, J, 2015). O modelo foi projetado utilizando a implementação de redes presente na biblioteca *TensorFlow*<sup>4</sup> sendo utilizada sob a interface da biblioteca *Keras*<sup>5</sup>.

**Quadro 02** – Avaliação das funções de ativação sigmoid, tanh e relu.

Ordenado pelo resultado da métrica MAPE				
	Função ativação Entrada	Função ativação Camada Escondida	Função ativação Saída	MAPE
1	relu	tanh	relu	6,66902
2	tanh	relu	relu	18,12816
3	relu	relu	relu	35,80039
4	tanh	relu	tanh	42,88810
5	sigmoid	relu	tanh	46,53802
6	tanh	sigmoid	relu	54,10297
7	relu	relu	tanh	63,29215
8	sigmoid	tanh	sigmoid	67,72382

<sup>4</sup> <https://www.tensorflow.org>

<sup>5</sup> <https://keras.io>

9	sigmoid	relu	relu	71,07043
10	sigmoid	tang	relu	74,64350
11	tanh	tanh	sigmoid	75,44560
12	tanh	sigmoid	tanh	77,71258
13	relu	sigmoid	relu	93,87918
14	relu	sigmoid	tanh	121,95566
15	relu	tanh	sigmoid	129,29390
16	relu	relu	sigmoid	146,05280
17	tanh	relu	sigmoid	152,17850
18	relu	sigmoid	sigmoid	154,27665
19	tanh	sigmoid	sigmoid	155,29464
20	tanh	tanh	tanh	161,05143
21	sigmoid	tanh	tanh	163,80048
22	tanh	tanh	tanh	182,97971
23	relu	tanh	relu	189,83298
24	sigmoid	relu	sigmoid	198,91475
25	sigmoid	sigmoid	relu	334,80545
26	sigmoid	sigmoid	tanh	353,39476
27	sigmoid	sigmoid	sigmoid	435,26919

Fonte: autor.

Esses testes foram submetidos a mesma base de dados com o índice Bovespa e utilizou-se 80% da base para treinamento e 20% para validação com 5 épocas de treino.

Para avaliação da quantidade de épocas a serem submetidas para treinamento das redes foram feitos testes e adotou 100 épocas como padrão.

Importante destacar que os 3 (três) modelos de redes neurais avaliados, sendo o MLP, NARX e LSTM foram configurados em sua topologia de maneira que todos ficassem o mais semelhante possível e assim fossem avaliados, sendo que o intuito não foi buscar a configuração individual de cada modelo que apresentasse melhor performance.

Também é importante ressaltar que a previsão é feita para  $(d+1)$ . E que isso foi executado recorrentemente por 10 dias, com os dados reais sendo rerepresentados, tornando-se parte das entradas.

### 3.6.1 Definições dos Modelos de Previsão

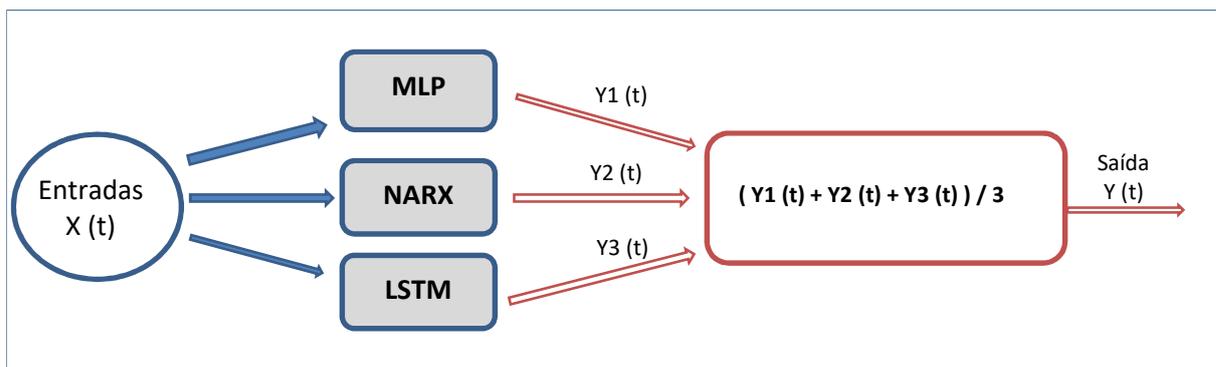
O primeiro modelo desenvolvido é baseado em uma rede neural MLP com topologia 7:14:1, onde 7 são os números de neurônios na entrada, 14 são os neurônios na camada intermediária e 1 é o número de neurônio na saída. As funções de ativação definidas respectivamente entre as camadas foram Unidade Linear Retificada (relu), Tangente Hiperbólica (tanh) e Unidade Linear Retificada (relu).

O segundo modelo desenvolvido é baseado em uma rede neural NARX com a mesma topologia do modelo MLP. Neste modelo foi acrescentada uma reentrada com dois períodos de atraso.

O terceiro modelo desenvolvido é baseado em uma rede neural LSTM com o mesmo número de neurônios de entrada que os modelos MLP e NARX.

O método *ensemble* adotado neste trabalho consiste em realizar o cálculo da média simples entre os resultados das redes MLP, NARX e LSTM, conforme Figura 13 e Equação 15.

**Figura 13** – *Ensemble* das três RNAs.



Fonte: autor.

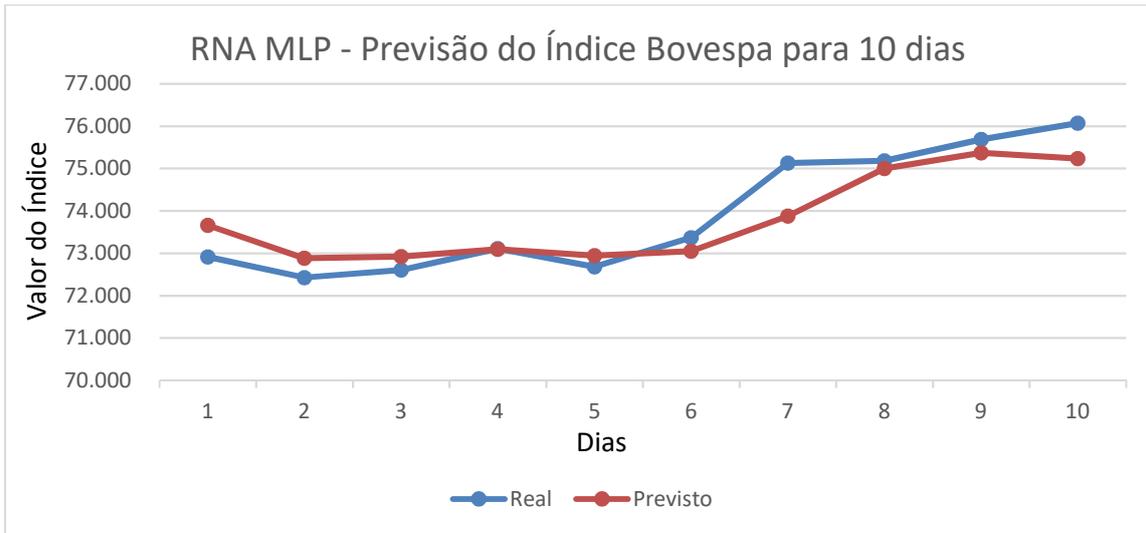
$$y(t) = \frac{y1(t) + y2(t) + y3(t)}{3} \quad (15)$$

Onde  $y1(t)$ ,  $y2(t)$  e  $y3(t)$  são resultados de cada modelo de rede neural, 3 é o total de modelos propostos neste trabalho e  $y(t)$  é o resultado apresentado de previsão para o modelo proposto.

#### 4. ANÁLISE DE RESULTADOS

Conforme as figuras e quadros apresentados na sequência, podemos observar a performance de cada modelo de rede neural, bem como, o modelo Ensemble proposto.

**Figura 14** – Comparativo Real x Previsto pela RNA MLP.



Fonte: autor.

A Figura 14 apresenta os resultados estimados de  $y(t)$ , valor previsto, comparado com os valores reais do índice Bovespa com a utilização da rede neural MLP durante 10 dias.

**Tabela 07** - Comparativo Real x Previsto pela RNA MLP.

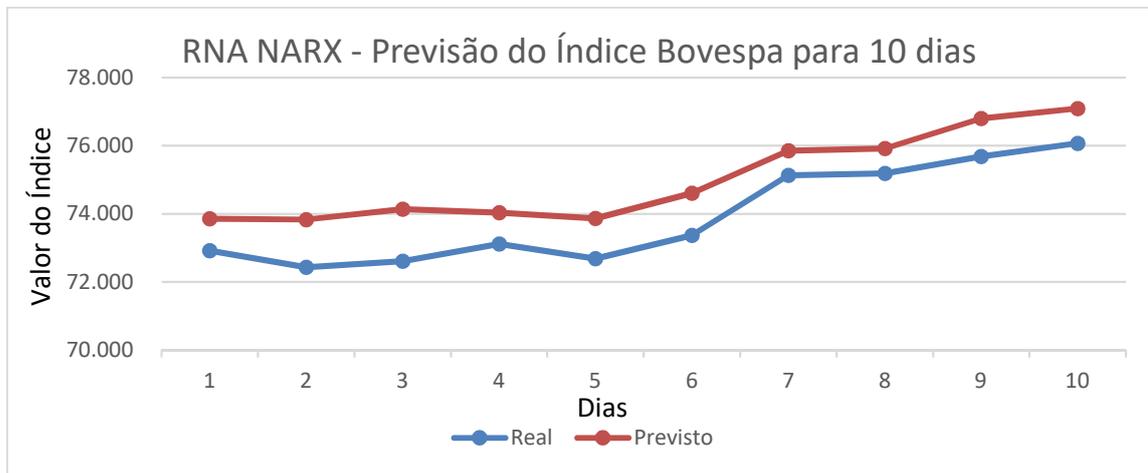
MLP						
Nº Dias	Data	Real	Movimento	Previsto	Movimento	Precisão Movimento
1	13/12/2018	72.914	Baixa	73.661	Alta	Errado
2	14/12/2018	72.429	Baixa	72.884	Baixa	Certo
3	15/12/2018	72.608	Alta	72.921	Alta	Certo
4	18/12/2018	73.115	Alta	73.096	Alta	Certo
5	19/12/2018	72.680	Baixa	72.949	Baixa	Certo
6	20/12/2018	73.367	Alta	73.054	Alta	Certo
7	21/12/2018	75.133	Alta	73.884	Alta	Certo
8	22/12/2018	75.187	Alta	75.005	Alta	Certo
9	26/12/2018	75.688	Alta	75.371	Alta	Certo
10	27/12/2018	76.073	Alta	75.237	Baixa	Errado

Fonte: autor.

Analisando a Tabela 07, é possível observar o alto índice de acerto dos movimentos do índice Bovespa acompanhando em 80%. Utilizando-se os valores Previsto e o Real, chega-se aos seguintes resultados nas métricas propostas, sendo MSE igual a 342.620,40; RMSE igual a 585,34 e MAPE igual a 0,63.

Na Figura 15, podemos analisar que a rede NARX apresentou indícios de não ter aprendido corretamente com os dados de treinamento e que apenas “decorou” os valores, tendo em vista que a linha da previsão está acompanhando os valores reais, porém em nenhum momento ocorreu de cruzá-la, o que pode dar indícios da não generalização do modelo, porém não é conclusivo tendo em vista que a janela de predição foi de apenas 10 dias.

**Figura 15** – Comparativo Real x Previsto pela RNA NARX.



Fonte: autor.

**Tabela 08** - Comparativo Real x Previsto pela RNA NARX.

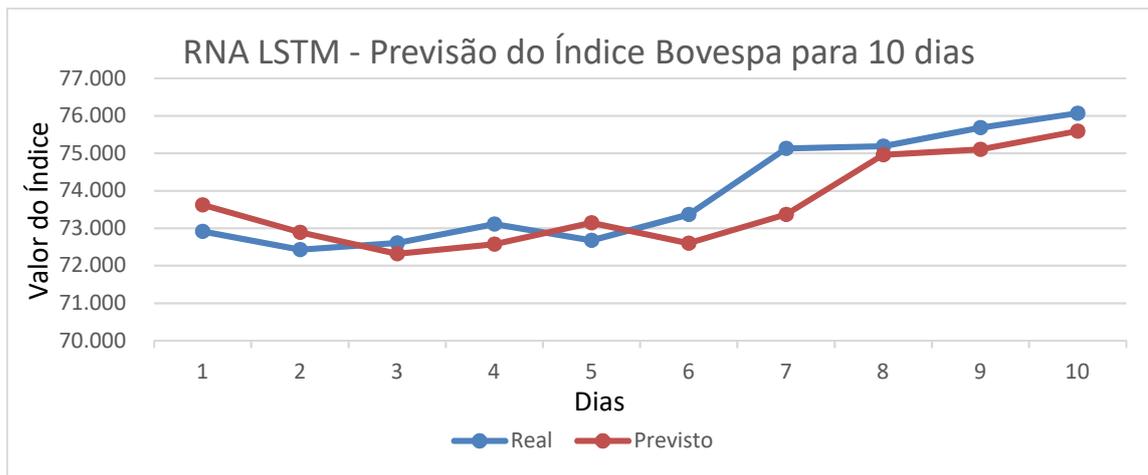
NARX						
Nº Dias	Data	Real	Movimento	Previsto	Movimento	Precisão Movimento
1	13/12/2018	72.914	Baixa	73.853	Alta	Errado
2	14/12/2018	72.429	Baixa	73.831	Baixa	Certo
3	15/12/2018	72.608	Alta	74.132	Alta	Certo
4	18/12/2018	73.115	Alta	74.036	Baixa	Errado
5	19/12/2018	72.680	Baixa	73.861	Baixa	Certo
6	20/12/2018	73.367	Alta	74.607	Alta	Certo
7	21/12/2018	75.133	Alta	75.854	Alta	Certo
8	22/12/2018	75.187	Alta	75.919	Alta	Certo
9	26/12/2018	75.688	Alta	76.796	Alta	Certo
10	27/12/2018	76.073	Alta	77.094	Alta	Certo

Fonte: autor.

Analisando a Tabela 08, referente as previsões realizadas pela rede neural NARX, também é possível observar o alto índice de acerto dos movimentos do índice Bovespa acompanhando em 80%. Utilizando-se os valores Previsto e o Real, chega-se aos seguintes resultados nas métricas propostas, sendo MSE igual a 1.227.627,30; RMSE igual a 1.107,98 e MAPE igual a 1,46.

É possível observar na Figura 16 que a rede neural LSTM, que está enquadrada na subárea de aprendizagem profunda, apresentou bons resultados para previsão de series temporais. Vale ressaltar que o modelo de arquitetura dessa rede não foi alvo de avaliação em todos os seus aspectos, e que o objetivo nesse trabalho foi avaliá-las no aspecto de estrutura de maneira que ficasse o mais próximo possível das demais em termos de estrutura.

**Figura 16** – Comparativo Real x Previsto pela RNA LSTM.



Fonte: autor.

**Tabela 09** - Comparativo Real x Previsto pela RNA LSTM.

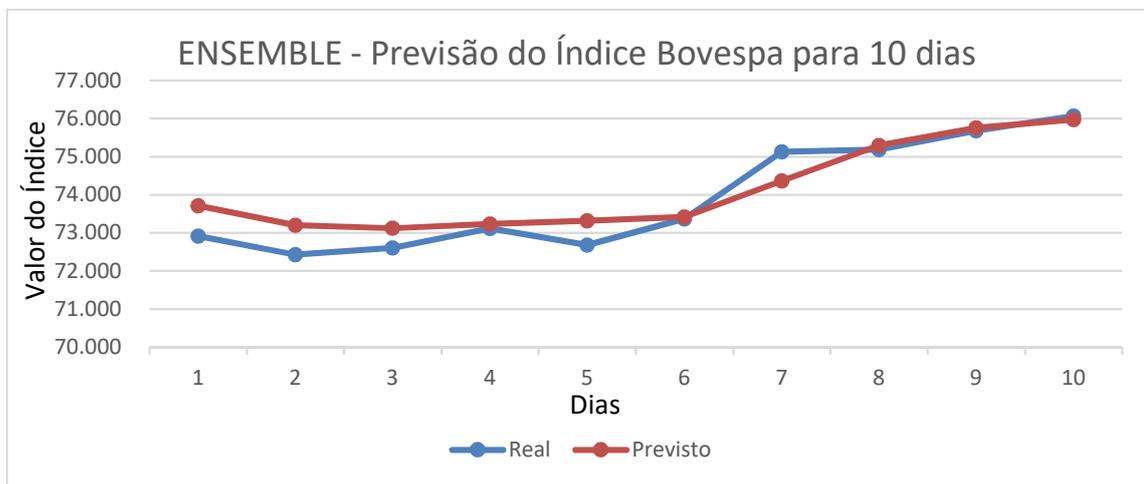
LSTM						
Nº Dias	Data	Real	Movimento	Previsto	Movimento	Precisão Movimento
1	13/12/2018	72.914	Baixa	73.625	Alta	Errado
2	14/12/2018	72.429	Baixa	72.890	Baixa	Certo
3	15/12/2018	72.608	Alta	72.317	Baixa	Errado
4	18/12/2018	73.115	Alta	72.572	Alta	Certo
5	19/12/2018	72.680	Baixa	73.143	Alta	Errado
6	20/12/2018	73.367	Alta	72.598	Baixa	Errado
7	21/12/2018	75.133	Alta	73.368	Alta	Certo
8	22/12/2018	75.187	Alta	74.961	Alta	Certo
9	26/12/2018	75.688	Alta	75.109	Alta	Certo
10	27/12/2018	76.073	Alta	75.596	Alta	Certo

Fonte: autor.

Analisando a Tabela 09, referente as previsões realizadas pela rede neural LSTM, é possível observar que apesar do bom resultado, acertando 60% do movimento, e que as métricas de precisão também foram boas, podemos observar que a mesma apresentou valores próximos dos valores reais, mas em direções opostas, conforme é possível constatar no gráfico. Utilizando-se os valores Previsto e o Real, chega-se aos seguintes resultados nas métricas propostas, sendo MSE igual a 563.237,30; RMSE igual a 750,49 e MAPE igual a 0,85.

Conforme valores apresentados na Tabela 10, o modelo *Ensemble* apresentou bons resultados, o que pode ser observado também na Figura 17, que compara as previsões com os valores reais do índice Bovespa.

**Figura 17** – Comparativo Real x Previsto pelo método *Ensemble*.



Fonte: autor.

**Tabela 10** - Comparativo Real x Previsto pelo método *Ensemble*.

ENSEMBLE						
Nº Dias	Data	Real	Movimento	Previsto	Movimento	Precisão Movimento
1	13/12/2018	72.914	Baixa	73.713	Alta	Errado
2	14/12/2018	72.429	Baixa	73.202	Baixa	Certo
3	15/12/2018	72.608	Alta	73.123	Baixa	Errado
4	18/12/2018	73.115	Alta	73.235	Alta	Certo
5	19/12/2018	72.680	Baixa	73.318	Alta	Errado
6	20/12/2018	73.367	Alta	73.420	Alta	Certo
7	21/12/2018	75.133	Alta	74.369	Alta	Certo
8	22/12/2018	75.187	Alta	75.295	Alta	Certo
9	26/12/2018	75.688	Alta	75.759	Alta	Certo
10	27/12/2018	76.073	Alta	75.976	Alta	Certo

Fonte: autor.

Analisando a Tabela 10, referente as previsões realizadas pelo método Ensemble, é possível observar o índice de 70% de acerto dos movimentos do índice Bovespa. Utilizando-se os valores Previsto e o Real, chega-se aos seguintes resultados nas métricas propostas, sendo MSE igual a 253.503,29; RMSE igual a 503,49 e MAPE igual a 0,54.

Como análise dos resultados, podemos observar que o método *Ensemble* apresentou bons resultados comparados com as RNAs individualmente, conforme pode ser observado nas métricas MSE, RMSE e MAPE, na Tabela 11, também podemos verificar que o  $R^2$  (coeficiente de correlação) apresentou um alto valor, corroborando com a conclusão da eficiência do método proposto. Todavia, ao analisar o movimento da previsão para Alta ou Baixa do índice, podemos constatar que as redes MLP e NARX, apresentaram 80% de acertos, enquanto a LSTM ficou com 60% e o modelo *Ensemble* com 70%.

**Tabela 11** – Comparativos das RNAs e *Ensemble* através das métricas.

<b>RNA</b>	<b>MSE</b>	<b>RMSE</b>	<b>MAPE</b>	<b>R<sup>2</sup></b>
MLP	342.620,40	585,34	0,63	0,87
NARX	1.227.627,30	1.107,98	1,46	0,97
LSTM	563.237,30	750,49	0,85	0,74
<b>ENSEMBLE</b>	<b>253.503,29</b>	<b>503,49</b>	<b>0,54</b>	0,92

Fonte: autor.

Com isso, avaliamos que apesar dos bons resultados apresentados nas métricas, o método proposto não sobressaiu com relação a identificar a movimentação do Ibovespa.

## 5. CONCLUSÃO

Conforme o objetivo principal desse trabalho de desenvolver e comparar 3 modelos de redes neurais, sendo os modelos MLP, NARX e LSTM. E, após obtidos os resultados de estimação e predição dos modelos, realizar a combinação dos resultados no método *Ensemble* das redes neurais propostas e avaliar os resultados. Podemos concluir que o método *Ensemble*, através da média aritmética das previsões das redes neurais MLP, NARX e LSTM, apresentou os melhores resultados nas métricas adotadas para avaliação de precisão.

Porém, não provou ser o mais eficaz comparando com os modelos de rede neural MLP e NARX, mas ficou à frente da rede LSTM nesse experimento para acompanhamento da movimentação do índice Bovespa para uma janela de previsão de 10 dias, mesmo apresentando os melhores resultados nas métricas de precisão, o que demonstra que apesar das métricas de avaliação dos erros serem bons, as métricas adotadas não apresentam o direcionamento das previsões.

Com isso, sugere-se para próximos trabalhos, avaliar outras combinações de *Ensemble*, como por exemplo, sugere-se combinar os resultados de entradas para o *Ensemble* com outros métodos de previsão, como GARCH e/ou suas variações, assim como, utilizar outras técnicas de combinação linear mais avançadas para apoiar em seu resultado.

Outros métodos para gerar os resultados do *Ensemble* também podem ser avaliados, por exemplo, a escolha do resultado por votação. Sugere-se também pesquisar outras métricas que permitam avaliar a direção das previsões e erros.

## 6. REFERÊNCIAS

- ADHIKARI, R.; AGRAWAL, R. K. A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications*, v. 24, n. 6, p. 1441–1449, 2014.
- AMBEV. Ambev. Disponível em: <<http://www.ambev.com.br/>>. Acesso em: 3 jul. 2016.
- ASSAF NETO, A. Mercado financeiro. 4 ed. São Paulo: Atlas, 2001.
- ATSALAKIS, G. S.; VALAVANIS, K. P. Surveying stock market forecasting techniques—part ii: Soft computing methods. *Expert Systems with Applications*, Elsevier, v. 36, n. 3, p. 5932–5941, 2009.
- BEUREN, I. M. COLAUTO, R.D. Coleta, análise e interpretação dos dados. In: Beuren, Ilse Maria. (Org). Como elaborar trabalhos monográficos em contabilidade. 3 ed. São Paulo: Atlas, 2006.
- BM&FBOVESPA. BM&F Bovespa. Disponível em: <<http://ri.bmfbovespa.com.br/>>. Acesso em: 1 jul. 2017.
- BISHOP, C. A. Pattern Recognition and Machine Learning. Information Science and Statistics. Springer, New York, 2006.
- BM&FBOVESPA. BM&F Bovespa. 2018. Disponível em: <<http://www.bmfbovespa.com.br>>. Acesso em: 20 jan. 2018.
- BODIE, Z.; KANE, A.; MARCUS, A. J. Investments, volume 8. The MacGraw-Hill Companies, New York, EUA, 2009.
- CAO,Q.;LEGGIO,K.B.;SCHNIEDERJANS,M.J.A comparison between fama and french's model and artificial neural networks in predicting the chinese stock market. *Computers & Operations Research*, Elsevier, v. 32, n. 10, p. 2499–2512, 2005.
- CAVALCANTI, F; MISUMI, J.Y. Mercado de capitais. Rio de Janeiro: Campus, 2001.
- CHANG, F.-J.; CHEN, P.-A.; LU, Y.-R.; HUANG, E.; CHANG, K.-Y. Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. *Journal of Hydrology*, Elsevier, v. 517, p. 836–846, 2014.
- CHARYTONIUK, W.; CHEN, M. Neural network design for short-term load forecasting. In: [S.l.], p. 554-561, 2000.
- CRONE, S.; DHAWAN, R. Forecasting seasonal time series with neural networks: a sensitivity analysis of architecture parameters. In: International Joint Conference on Neural Networks. [S.l.], p. 2009-2014, 2007.
- DIETTERICH, T. G. Ensemble methods in machine learning. In: INTERNATIONAL WORKSHOP ON MULTIPLE CLASSIFIER SYSTEMS, 1., 2000. Anais... New York: Springer Verlag, 2000. p. 1-15. (Lecture notes in computer science). Disponível em: <<http://www.cs.orst.edu/~tgd/>>.
- EDWARDS, R.D.; MAGEE, J.; BASSETTI, W.H.C. Technical analysis of stock trends. Flórida: CRC Press, 2006.
- FACURE, Matheus. Disponível em: < <https://matheusfacure.github.io//>>. Acesso em: 13 mar. 2018.
- FAMA, E. F.; MALKIEL, B. G. Efficient capital markets: a review of theory and empirical work. *The Journal of Finance*, vol 25, n. 2, p.383-417, 1970.
- FORTUNA, E. Mercado Financeiro: Produtos e serviços. 18. ed. Rio de Janeiro: Qualitymark, 2011.

- GIL, A. C. Como elaborar projetos de pesquisa. 4. ed. São Paulo: Atlas, 2007.
- GLANTZ, M.; KISSELL, R. Multi-asset Risk Modeling. Academic Press, 2013. ISBN 978-0-12-401690-3.
- GRAVES, A. Supervised Sequence Labelling with Recurrent Neural Networks. New York: Springer, 2012.
- GREFF, K.; SRIVASTAVA, R.K.; KOUTNÍK, J.; STEUNEBRINK, B.R.; SCHMIDHUBER, J. Lstm: a search space odyssey. IEEE transactions on neural networks and learning systems, IEEE, 2016.
- GOODFELLOW, I., BENGIO, Y., COURVILLE, A. Deep Learning. MIT Press, 2016.  
<http://www.deeplearningbook.org>.
- GOOGLE TRENDS. Google Trends. Disponível em: <<https://trends.google.com.br/trends/>>. Acesso em: 4 jan. 2018.
- HANSEN, L.K.; SALAMON, P. Neural networks ensembles. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990. p.993–1001
- HANSEN, L. K.; LIISBERG, C.; SALAMON, P. Ensemble methods for handwritten digit recognition. In: IEEE. Neural Networks for Signal Processing, 1992. p. 333–342.
- HASHEM, S.; SCHMEISER, B.; TIH, Y. Optimal linear combinations of neural networks: An overview. In Proceedings of the IEEE International Conference on Neural Networks, 1994. p. 1507–1512. Orlando, USA.
- HAYKIN, Simon. Redes Neurais: princípios e práticas. 2 ed. Porto Alegre: Bookman, 2001.
- HIPPERT, H.S.; PEDREIRA, C.E.; SOUZA, R.C. Neural networks for short-term load forecasting: a review and evaluation. IEEE Transactions on Power Systems. v.16, n. 1, p.44-55, 2001.
- HOJI, M. Os riscos e os prazeres do mercado de ações: maximize os ganhos em bolsa de valores com o home broker. São Paulo: ProfitBooks, 2007.
- HSU, C.-C.; CHEN, C.-Y. Regional load forecasting in Taiwan – applications of artificial neural networks. Energy Conversion and Management, v. 44, p. 1941–1949, 2003.
- HSU, F-H. Behind Deep Blue: Building the Computer That Defeated the World Chess Champion. Princeton, NJ, USA, Princeton University Press, 2002. ISBN: 0691090653.
- KAEHLING, L.P.; LITTMAN, M.L.; MOORE, A.W. Reinforcement learning: a survey. Journal of Artificial Intelligence Research 4, 1996.
- KARA, Y.; BOYACIOGLU, M. A.; BAYKAN, Ö. K. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. Expert systems with Applications, Elsevier, v. 38, n. 5, p. 5311–5319, 2011.
- KERR, R. B. Mercado Financeiro e de Capitais. São Paulo: Pearson Prentice Hall, 2011.
- KEIEL, G.; BENDER, F.A. Identificação de processos geradores de séries temporais financeiras: uso de entradas exógenas na modelagem do IBovespa. V Congresso de Pesquisa e Extensão da FSG. ISSN 2318-8014. 2017.
- KINGMA, D.; BA, J. Adam: a method for stochastic optimization. Published as a Conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015.
- KOURENTZES, N.; PETROPOULOS, F. Forecasting with multivariate temporal aggregation: The case of promotional modelling. International Journal of Production Economics, in press, 2015.

- KURT, A.; OKTAY, A. B. Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. *Expert Systems with Applications*, Elsevier, v. 37, p. 7986–7992, 2010.
- LEITE, H.P.; SANVICENTE, A.Z. Índice Ibovespa: um padrão para os investimentos brasileiros. São Paulo: Atlas, 1994.
- LUDWIG JR, O.; COSTA, E.M.M. *Redes neurais: fundamentos e aplicações com programas em c*. Rio de Janeiro: Ciência Moderna, 2007.
- LO, A.; MACKINLAY, A. *A non-random walk down Wall Street*. Princeton Univ. Press, Princeton, 1999. ISBN 0691057745. <http://www.jstor.org/stable/j.ctt7tccx>
- MALKIEL, B. G. *A random walk down wall street*. WW Norton & Company, New York, 1973.
- MARTINS, M. A. DOS S.; METTE, F.; MACEDO, G. R. DE. A Utilização de Redes Neurais Artificiais para a Estimacão dos Preços da Petrobrás PN na Bovespa. *ConTexto*, v. 8, n. 14, p. 1–16, 2008.
- MATTAR, M. A.; ALAZBA, A. A.; ZIN EL-ABEDIN, T. K. Forecasting furrow irrigation infiltration using artificial neural networks. *Agricultural Water Management*, Elsevier, v. 148, p. 63–71, 2015.
- MELLAGI, F. A.; ISHIKAWA, S. *Mercado financeiro e de capitais*. São Paulo: Atlas, 2000.
- MENEZES JUNIOR, J. M. P. Contribuições ao problema de predição recursiva de séries temporais univariadas usando redes neurais recorrentes. Tese (doutorado) - Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-graduação em Engenharia de Teleinformática, Fortaleza, 2012.
- MERIGÓ, J. M.; PALACIOS-MARQUÉS, D.; RIBEIRO-NAVARRETE, B. Aggregation systems for sales forecasting. *Journal of Business Research*, v. 68, n. 11, p. 2299–2304, 2015.
- MO, H.; WANG, J.; NIU, H. Exponent back propagation neural network forecasting for financial cross-correlation relationship. *Expert Systems with Applications*, v. 53, n. 2016, p. 106–116, 2016.
- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. *Foundations of machine learning*. The MIT Press, 2012.
- MORETTIN, P.A. e TOLOI, C. M. C. *Análise de séries temporais*. 2. ed. São Paulo: Egard Blucher, 2006.
- MOUSAVI, R.; EFTEKHARI, M. A new ensemble learning methodology based on hybridization of classifier ensemble selection approaches. *Applied Soft Computing*, v. 37, p. 652–666, 2015.
- NISON, Steve. *The candlestick course*. New York: John Wiley & Sons, 2003.
- NORVIG, P.; RUSSEL, L.S. *Inteligência artificial*. São Paulo: Campus, 2004.
- OLAH, C. Understanding lstm networks. 2015. Disponível em: <<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>>.
- OLIVEIRA, F.A.de et al. The use of artificial neural networks in the analysis and prediction of stock prices. In: *IEEE International Conference on Systems, Man and Cybernetics*. 2011. p. 2151–2155.
- OLIVEIRA, N.; CORTEZ, P.; AREAL, N. Some experiments on modeling stock market behavior using investor sentiment analysis and posting volume from Twitter. In: *Proceedings of the 3rd International Conference on Web Intelligence, Mining and Semantics - WIMS '13*. New York, New York, USA: ACM Press, 2013. p. 1. ISBN 9781450318501. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2479787.2479811>>

OLIVEIRA, M. A. DE. Aplicação de Redes Neurais Artificiais na Análise de Séries Temporais Econômico-Financeiras. Tese (Doutorado). USP - Universidade de São Paulo, Faculdade de Economia, Administração e Contabilidade, 2007.

PYTHON. Python. Disponível em: <<http://www.python.org/>>. Acesso em: 03 fev. 2018.

POLIKAR, R. Ensemble based systems in decision making. IEEE Circuits and systems magazine, v. 6, n. 3, p. 21–45, 2006.

RESENDE, F. S. Ciência de dados aplicada à BM&Fbovespa. Curitiba: Appris, 2016.

ROCA, R. L. R. DE S.; MÓL, A. L. R. Previsão do Índice Bovespa por Meio de Redes Neurais Artificiais. Business and Management Review, v. 4, n. 10, p. 377–390, 2015.

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.

SAAD, E. W.; PROKHOROV, D. V.; WUNSCH, D. C. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. Neural Networks, IEEE Transactionson, IEEE, v. 9, n. 6, p. 1456–1470, 1998.

SEOKHO, K.; CHO, S.; KANG, P. Multi-class classification via heterogeneous ensemble of one-class classifiers. Engineering Applications of Artificial Intelligence, v. 43, p. 35–43, 2015.

SILVA, E.J. Modelagem e aplicação de técnicas de aprendizado de máquina para negociação em alta frequência em bolsa de valores. Dissertação (Mestrado) - Curso de Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

SILVA, I. N. DA; SAPATTI, D. H.; FLAUZINO, R. A. Redes Neurais Artificiais: Para Engenharia e Ciências Aplicadas. São Paulo: Artliber, 2010.

SHARKEY, A.J.C. Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems. Springer Verlag London Limited, London, UK, 1999.

STEFKA, D.; HOLENA, M. Dynamic classifier aggregation using interaction-sensitive fuzzy measures. Fuzzy Sets and Systems, v. 270, p. 25–52, 2015.

SUGIYAMA, M. Chapter 1 - Statistical Machine Learning. In: SUGIYAMA, M. (Ed.). Introduction to Statistical Machine Learning. Boston: Morgan Kaufmann, 2016. p. 3–8. ISBN 978-0-12-802121-7. Disponível em: <<http://www.sciencedirect.com/science/article/pii/B9780128021217000121>>

TAPIA-ROSETO, A. et al. Fusion of preferences from different perspectives in a decision-making context. Information Fusion, v. 29, p. 120–131, 2016.

UOL. UOL Economia Cotações. Disponível em: <<https://economia.uol.com.br/cotacoes>>. Acesso em: 4 jan. 2018.

VASCONCELOS, B. F. B. Poder preditivo de métodos de machine learning com processos de seleção de variáveis: uma aplicação às projeções de produto de países. Tese (Doutorado). UNB - Universidade de Brasília, Faculdade de Economia, Administração e Contabilidade, 2017.

WEIGEND, A.; GERSHEFELD, N. Time Series Prediction: Forecasting the Future and Understanding the Past. Reading: Addison-Wesley, 1994.

WEST, D.; DELLANA, S.; QIAN, J. Neural network ensemble strategies for financial decision applications. Computers & operations research, Elsevier, v. 32, n. 10, p. 2543–2559, 2005.

WU, Z.-H. Z. J.-X.; CHEN, Y. J. S.-F. Genetic algorithm based selective neural network ensemble. In: CITESEER. International Joint Conference on Artificial Intelligence, 17. Seattle, WA, 2001. v. 2, p. 797–802.

YAHOO. Yahoo Finanças. Disponível em: <<https://br.financas.yahoo.com>>. Acesso em: 4 jan. 2018.

ZANG, G.P. Avoiding pitfalls in neural network research. *IEEE Transactions on Systems, Man and Cybernetics, Part C*, p. 3-16, 2007.

ZHOU, Z.-H.; WU, J.; TANG, W. Ensembling neural networks: many could be better than all. *Artificial intelligence*, Elsevier, v. 137, n. 1, p. 239–263, 2002.