

UNIVERSIDADE FEDERAL DO MARANHÃO  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE  
ELETRICIDADE  
ÁREA DE CIÊNCIA DA COMPUTAÇÃO

**MANOEL MARCONDES DE OLIVEIRA LIMA JUNIOR**

**PROPOSTA DE UM MODELO DE PREDIÇÃO DA BOLSA DE  
VALORES USANDO UMA ABORDAGEM HÍBRIDA**

São Luís  
2013

**MANOEL MARCONDES DE OLIVEIRA LIMA JUNIOR**

**PROPOSTA DE UM MODELO DE PREDIÇÃO DA BOLSA DE VALORES USANDO  
UMA ABORDAGEM HÍBRIDA**

Dissertação de Mestrado apresentada à coordenação do Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão como requisito parcial para a obtenção do título de Mestre em Engenharia de Eletricidade, na área de Ciência da Computação.

Orientador: Prof. Dr. Sofiane Labidi

São Luís  
2013

Lima Junior, Manoel Marcondes de Oliveira.

Proposta de um modelo de predição da bolsa de valores utilizando uma abordagem híbrida/ Manoel Marcondes de Oliveira Lima Junior – São Luís, 2013.

129 f.

Impresso por computador (fotocópia).

Orientador: Sofiane Labidi.

Dissertação (Mestrado) – Universidade Federal do Maranhão, Programa de Pós-Graduação em Engenharia de Eletricidade, 2013.

1. Redes Neurais - Bolsa de valores. 2. Análise de séries temporais. 3. Comitê de máquinas de aprendizado. I. Título.

CDU 004.032.26:336.761

**PROPOSTA DE UM MODELO DE PREDIÇÃO DA BOLSA DE VALORES USANDO  
UMA ABORDAGEM HÍBRIDA**

**MANOEL MARCONDES DE OLIVEIRA LIMA JUNIOR**

Dissertação de Mestrado apresentada à coordenação do Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão como requisito parcial para a obtenção do título de Mestre em Engenharia de Eletricidade, na área de Ciência da Computação.

Aprovado em / /

**BANCA EXAMINADORA**

Prof. Sofiane Labidi, Dr

(Orientador)

Prof. Zair Abdelouahab, Ph. D.

(Membro da Banca Examinadora)

Profa. Karla Donato Fook, Dra

(Membro da Banca Examinadora)

## DEDICATÓRIA

Aos meus pais,  
meu irmão  
e minha noiva

## AGRADECIMENTOS

Primeiramente a Deus, por ser meu guia em toda a vida e ter me oferecido sabedoria para fazer as melhores escolhas.

Aos meus pais Manoel Marcondes e Maria José, fonte ímpar de inspiração. Pelo amor incondicional, força e grandes ensinamentos de fundamental importância para conclusão de todos os meus projetos. Pela compreensão em todos os momentos de ausência e nas difíceis decisões que tive de fazer.

Ao meu irmão, que esteve de prontidão em momentos difíceis e sempre compartilhou seu conhecimento para me ajudar. E pela grande ajuda ao esclarecer muitas de minhas dúvidas que, com certeza, não conseguiria sozinho.

À minha Anne Carine, amiga, companheira e noiva. Um exemplo de pessoa do bem, que esteve presente quando precisei. Pelo seu carinho e compreensão nos melhores e piores momentos que tive. Pelo apoio e otimismo cético de que tudo terminaria bem e por me ajudar a mudar muitas vezes meu modo de pensar.

À minha segunda família, Antônio, Conceição, Tia Bibi, Camila e Danilo, por serem mais que um abrigo, um lugar de conforto, relaxamento e diversão.

Ao Professor Dr. Sofiane Labidi, meus sinceros agradecimentos pela oportunidade de concluir este mestrado, que mesmo não sendo seu aluno desde o início me recebeu de braços abertos e sempre ajudou em momentos de dificuldade.

A minhas revisoras Adah Matias e Jéssica Coelho que me auxiliaram em vários momentos ao longo de todo este trabalho e foram essenciais para sua conclusão.

Aos companheiros de laboratório de pesquisa (LSI), Rafael Pinheiro, Yanna, Rafael Cunha, Christian e Fernando, pelos momentos de descontração, discussões extremamente importantes para meu crescimento como pesquisador.

Aos amigos do NTI, Thiago, Ronald, Marcos, Marcio, Osvaldo, Dallyson, e muitos outros pelo exemplo de profissionalismo e por me proporcionar um ambiente agradável de trabalho.

E a todos os outros que direta ou indiretamente participaram do meu desenvolvimento pessoal e auxiliaram na conclusão deste trabalho.

*“A maioria das pessoas se interessa por ações quando todo mundo está interessado.  
O momento de interessar-se é quando ninguém mais se interessa.  
Não se ganha dinheiro comprando o que é popular.”  
(Warren Buffett)*

## RESUMO

O mercado acionário é uma importante forma para captação de recursos para investidores. Entretanto, é um mercado altamente complexo. Os investidores têm como objetivo atingir o máximo de lucro. Assim, a compra ou venda de ações deve ser feita no momento certo. Para atingir esse objetivo, técnicas de predição podem ser aplicadas em mercado acionário a fim de antever seu comportamento. Para predição em mercado acionário, existem várias técnicas que apresentam resultados promissores, mas cada uma tem suas vantagens e limitações particulares. No entanto, essas técnicas exigem um grande número de variáveis e uma arquitetura complexa. O presente trabalho descreve uma proposta de um modelo de predição objetivando elevar as chances de rentabilidade em investimentos com ações. O modelo proposto utiliza uma abordagem híbrida por meio da definição de um comitê de máquinas de aprendizado que combinam vantagens das três técnicas (estatística, rede neural e indicadores técnicos) a fim de superar as suas respectivas limitações. A abordagem proposta é analisada através de uma comparação com seus previsores individuais sob a perspectiva de duas métricas, o que demonstrou ganhos significativos em termos de rentabilidade e precisão quando utilizado um comitê com abordagem de predição envolvendo pesos nas soluções dos previsores individuais.

**Palavras-chave:** Bolsa de Valores, Análise de Séries Temporais, Redes Neurais, Comitê de máquinas de aprendizado, Predição Financeira.

## ABSTRACT

The stock market is a highly complex market and an important way to raise funds for investors. Investors aim to achieve maximum profit. Thus, the purchase or sale of shares must be made on time. To achieve this goal, prediction techniques can be applied to the stock market in order to predict their behavior. There are several techniques that show promising results for prediction on the stock market, but each one has particular advantages and limitations. Thus, these techniques require a large number of variables and a complex architecture. This paper describes a proposal of a prediction approach model designed to raise the return's chances on investments in shares. It uses a hybrid approach based on the definition from a committee of learning machines which combines the advantages of three techniques (statistical, neural network and technical indicators) in order to overcome their respective limitations. The proposed approach is examined through a comparison with their individual forecasters according to the perspective of two metrics, which ones demonstrate significant gains in terms of profitability and precision when used with a committee prediction approach involving weights in the solution of individual forecasters.

**Keywords:** Stock Market, Time Series Analysis, Neural Networks, Machine Learning Committee, Finantial prediction.

## LISTA DE FIGURAS

Figura 1	- Ciclo de ondas de Elliot.....	24
Figura 2	- Gráfico de evolução de preços.....	25
Figura 3	- Gráfico de barras.....	26
Figura 4	- Gráfico de <i>candlesticks</i> .....	26
Figura 5	- Estrutura e tipos de <i>candlestick</i> .....	27
Figura 6	- Representação gráfica da média móvel de 5 dias.....	29
Figura 7	- Representação gráfica de uma série temporal.....	32
Figura 8	- Estrutura básica de um neurônio artificial.....	43
Figura 9	- Arquitetura geral de uma RNA.....	45
Figura 10	- Passos do algoritmo de treinamento supervisionado.....	47
Figura 11	- Superfície multidimensional da função de erro.....	49
Figura 12	- Funcionamento básico da composição.....	55
Figura 13	- Resultado da utilização de ensemble comparada a previsores isolados	61
Figura 14	- Informações detalhadas do padrão “ <i>Matching Low+</i> ”.....	68
Figura 15	- Representação gráfica do candlestick <i>Matching Low +</i> .....	69
Figura 16	- Informações obtidas do <i>YahooFinance</i> .....	77
Figura 17	- Representação gráfica do padrão <i>candlestick hammer</i> .....	81
Figura 18	- Fluxo de funcionamento do componente estatístico.....	81
Figura 19	- Fluxo de funcionamento do componente neural .....	84
Figura 20	- Fluxo de funcionamento do componente técnico.....	86

## LISTA DE TABELAS

Tabela 1	- Descrição dos parâmetros da URL padronizada do YahooFinance.....	76
Tabela 2	- Valores estimados para combinação de saídas dos previsor técnico.....	87
Tabela 3	- Experimentos realizados.....	92
Tabela 4	- Eficácia dos previsores individuais aplicados a PETR4.....	95
Tabela 5	- Eficácia do <i>ensemble</i> aplicado a PETR4.....	96
Tabela 6	- Eficiência dos previsores individuais aplicados a PETR4.....	96
Tabela 7	- Eficiência do <i>ensemble</i> aplicado a PETR4.....	97
Tabela 8	- Eficácia dos previsores individuais aplicados a AMIL.....	98
Tabela 9	- Eficácia do <i>ensemble</i> aplicado a AMIL.....	98
Tabela 10	- Eficiência da abordagem aplicada a AMIL3.....	99
Tabela 11	- Eficiência do <i>ensemble</i> aplicada a AMIL3.....	99

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	13
<b>1.1</b>	<b>Problemática.....</b>	14
<b>1.2</b>	<b>Objetivos.....</b>	16
1.2.1	Objetivo Geral.....	16
1.2.2	Objetivos Específicos.....	16
<b>1.3</b>	<b>Justificativa e relevância.....</b>	16
<b>2</b>	<b>BOLSA DE VALORES.....</b>	18
<b>2.1</b>	<b>Ações.....</b>	19
<b>2.2</b>	<b>Análises de Mercado.....</b>	20
<b>2.3</b>	<b>Análise Técnica.....</b>	22
<b>2.4</b>	<b>Séries Temporais.....</b>	30
<b>3</b>	<b>PREDIÇÃO DE SÉRIES TEMPORAIS.....</b>	34
<b>3.1</b>	<b>Definições.....</b>	35
<b>3.2</b>	<b>Indicadores Técnicos.....</b>	36
3.2.1	Relative Strength Index (RSI).....	37
3.2.2	Average Directional Index (ADX).....	38
3.2.3	Bollinger Bands (BBS).....	39
<b>3.3</b>	<b>Redes Neurais Artificiais.....</b>	40
3.3.1	Histórico.....	41
3.3.2	Estrutura.....	42
3.3.3	Arquitetura e Aprendizado.....	44
3.3.4	Predição utilizando Redes Neurais.....	51
<b>3.4</b>	<b>Comitê de Aprendizado.....</b>	53
3.4.1	Historico.....	56
3.4.2	Previsores.....	56
3.4.3	Arquitetura.....	57
3.4.4	Métodos de Combinação.....	58
3.4.5	Predição usando ensenbles.....	59
<b>3.5</b>	<b>Considerações Finais.....</b>	61
<b>4</b>	<b>TRABALHOS RELACIONADOS.....</b>	64
<b>4.1</b>	<b>Médias movies em Predição de Séries Temporais.....</b>	64

<b>4.2</b>	<b>Predição de séries temporais baseado em redes neurais artificiais e transformada de Wavelet.....</b>	65
<b>4.3</b>	<b>Predição de estimativa de rentabilidade do mercado acionário usando candlesticks.....</b>	67
<b>4.4</b>	<b>Comitê de Máquinas em Predição de Séries Temporais.....</b>	69
<b>4.5</b>	<b>Discussão.....</b>	70
<b>5</b>	<b>MODELO HÍBRIDO DE PREDIÇÃO.....</b>	74
<b>5.1</b>	<b>Metodologia.....</b>	75
<b>5.2</b>	<b>Solução Proposta.....</b>	78
5.2.1	Previsor Estatístico.....	80
5.2.2	Previsor Neural.....	83
5.2.3	Previsor Técnico.....	85
<b>5.3</b>	<b>Experimentos e Simulações.....</b>	88
5.3.1	Cenário.....	88
5.3.2	Métricas.....	90
5.3.3	Experimentos.....	91
<b>6</b>	<b>RESULTADOS E DISCUSSÃO.....</b>	94
<b>6.1</b>	<b>Considerações Iniciais.....</b>	94
<b>6.2</b>	<b>Análise dos resultados.....</b>	99
<b>7</b>	<b>CONCLUSÃO.....</b>	104
<b>7.1</b>	<b>Trabalhos Futuros.....</b>	105
<b>7.2</b>	<b>Publicações.....</b>	106
	<b>REFERÊNCIAS.....</b>	107
	<b>APÊNDICES.....</b>	113
	<b>ANEXOS.....</b>	121

## 1 INTRODUÇÃO

O Mercado de Capitais vem adquirindo cada vez mais um papel fundamental no processo de desenvolvimento da economia mundial. Sendo considerado um ambiente de realização de transações dinâmicas, proporcionando liquidez aos títulos financeiros emitidos pelas empresas. É também uma importante fonte de recursos financeiros para as empresas, fornecendo a elas capital para fazer frente aos seus investimentos, viabilizando seu processo de capitalização. (CAVALCANTE, 2005)

Ao utilizar o Mercado de Capitais através de um intermediário como, por exemplo, a Bolsa de Valores, um investidor está adquirindo valores mobiliários (produtos e/ou serviços). A Bolsa de Valores é uma oportunidade bastante atrativa para um investidor obter patrimônios de modo a fornecer recursos para o crescimento de uma empresa. A captação de recursos através do mercado acionário faculta às empresas um aumento de sua produção e geração de novos empregos. (MARANGONI, 2010)

O mercado acionário é um lugar onde ocorre a compra e a venda de ações. Ao adquirir ações de uma determinada empresa, um investidor se torna proprietário de uma parte da empresa conforme o número de ações detidas por ele. O seu objetivo, certamente, é atingir o máximo de lucro. Assim, a compra ou venda de ações deve ser feita no momento certo. Quando realiza uma operação de venda ou compra, um investidor deve levar em consideração uma ampla quantidade de informações, tais como: volume de transações financeiras e principais índices do mercado e sua tendência atual. Neste cenário, é de grande valia o desenvolvimento de um mecanismo de apoio à tomada de decisão, indicando o melhor momento de realizar a venda ou compra de ações ou até aguardar um momento melhor a fim de alcançar melhores resultados. Para atingir esse objetivo, técnicas de predição podem ser aplicadas em mercado acionário para antever seu comportamento.

Segundo Rocha e Macedo (2011), a utilização da predição no mercado acionário pode tornar investimentos em ações algo muito rentável. Para que isto ocorra, existem equações matemáticas que permitem antecipar possíveis mudanças em valores futuros de ações; há modelos baseados no comportamento do mercado ao longo de um período de tempo; e, por último, existem algoritmos de identificação de padrões a partir de informações do mercado. Desta maneira, é possível direcionar as aplicações financeiras conforme as indicações de mudanças do comportamento do mercado através de análises dos resultados das técnicas de predição. A aplicação de modelos computacionais baseada em técnicas de Inteligência Artificial (IA) em mercado acionário tem sido objeto de estudo recorrente na

comunidade científica que almeja prover modelos ou ambientes computacionais que explorem informações relativas ao mercado, provendo técnicas com finalidades distintas para simulações de pregão de bolsa de valores, sistemas autônomos de investimento, predição do mercado, entre outros.

O presente trabalho descreve um novo modelo de predição objetivando elevar as chances de rentabilidade em investimentos que utilizando ações na Bolsa de Valores. Este modelo utiliza uma abordagem híbrida que combina vantagens e funcionalidades apesentadas por três técnicas, além de possuir um objetivo secundário de superar as respectivas destas três técnicas. Para tanto, é desenvolvido um comitê de máquinas de aprendizagem para combinar as técnicas usadas, que são: estatística, rede neural e indicadores técnicos. Estas três técnicas são muito conhecidas na literatura por apresentarem resultados promissores, mas cada uma com suas vantagens e desvantagens particulares.

No estudo, os seguintes questionamentos são levantados: como aprimorar a predição em mercado acionário usando técnicas da IA, nomeadamente em relação às redes neurais; quais são as variáveis que devem ser consideradas; e como obter melhores resultados combinando as vantagens das técnicas estatísticas, redes neurais e indicadores técnicos financeiros.

## 1.1 Problemática

Segundo Guarnieri (2006), nas últimas décadas a tecnologia da informação contribuiu de forma significativa para facilitar e maximizar os lucros em investimentos na bolsa de valores auxiliando investidores no processo de tomada de decisão, por facilitar complexos cálculos com base estatística e matemática. Entretanto, técnicas computacionais vão muito além de auxiliar cálculos complicados. É possível modelar o comportamento dos preços na bolsa de valores e predizer valores futuros a partir de valores passados.

Conforme Costa e Marcondes (2012), o volume financeiro negociado na Bolsa de Valores cresce a cada ano e, de acordo com um relatório divulgado pela BM&FBOVESPA, no fim de 2012 este volume de negociações ultrapassou uma marca histórica encerrando com mais de um trilhão de reais. Isto ocorre, principalmente, devido ao crescimento do mercado nacional, à redução da taxa de juros e à valorização da moeda, o que motiva novas empresas e investidores a buscarem na bolsa uma forma de crescimento pela arrecadação de capital e

aumento de patrimônio. Anualmente, a Bovespa divulga um relatório com informações sobre o estado do mercado acionário nacional, informando o volume de negociações anual, número de negociações anual, rentabilidade dos acionistas, fatia do mercado ocupada por investidores estrangeiros, dentre outras.

De acordo com Souza (2012), a formação dos preços de cada ativo negociado em bolsa de valores ocorre de acordo com a demanda de interesse pelo ativo. Esta especificação possui um comportamento bem dinâmico, sendo que a oscilação do preço é diretamente influenciada pela lei da oferta e procura pelo ativo. A natureza dinâmica apresentada pelos preços chama a atenção de pequenos investidores a grandes e sólidas instituições bancárias, motivados pela grande oportunidade de investimento em ativos da bolsa, quando realizados da maneira correta ao identificar o melhor momento. Diante deste contexto de negociações financeiras, o momento certo de realizar as operações, seja ao comprar ou vender ativos, torna-se fator decisivo para alcançar o sucesso no mercado acionário, ou seja, está diretamente ligado aos lucros ou prejuízos dos investidores.

Durante investimentos no mercado acionário, o objetivo é sempre comprar quando o preço estiver o mais barato (valor mais baixo) realizável e vender quando o preço estiver bem mais caro (atingir seu valor mais alto). Sendo possível antecipar o comportamento de ativos financeiros é possível efetuar estas operações nos momentos citados potencializando os lucros alcançados pelos investidores e reduzindo as chances de prejuízo. Esta antecipação de comportamento é chamada de predição, conforme Rocha e Macedo (2011) mencionaram em seu trabalho, a predição no mercado acionário pode aumentar a rentabilidade de investimentos em ações.

A previsão ou predição pode ser entendida como uma forma de antecipação ou expectativa de algum acontecimento, não uma certeza. Mesmo que não seja uma garantia de ocorrência, é possível se preparar para eventos que seriam completamente inesperados, possibilitando a capacidade de planejar ações diante de cenários futuros. A existência de previsões aplicadas em várias áreas mostra o quanto é interessante sua utilização, como por exemplo, a tão popular previsão do tempo, ou outras não tão conhecidas como a previsão da vazão de um rio muito utilizada em projetos de usinas hidrelétricas. (SOUZA, 2012)

## 1.2 Objetivos

### 1.2.1 Objetivo Geral

Desenvolver uma abordagem híbrida de predição de comportamento da bolsa de valores utilizando um modelo baseado na técnica de comitê de máquinas de aprendizado do tipo *ensemble*.

### 1.2.2 Objetivos Específicos

Com o objetivo de alcançar o objetivo geral pretendido, buscar-se-á atingir os seguintes objetivos específicos.

- A construção de um dataset a partir do *YahooFinance*;
- A preparação do dataset através de etapas de pré-processamentos;
- A execução dos previsores;
- A combinação e geração do resultado da previsão a ser apresentado.

## 1.3 Justificativa e relevância

Como a predição aplicada à bolsa de valores é uma área que provoca muito interesse pelas perspectivas financeiras que são geradas, muitos modelos de predição foram propostos dentro do contexto da análise técnica, que corresponde a um segmento de análise do mercado acionário que utiliza equações matemáticas e gráficos como principal instrumento para realizar a predição. Dentre as soluções voltadas para este tipo de análise, destacam-se a utilização de padrões gráficos e de *candlestick* em cotações e os indicadores técnicos de mercado. Neste cenário, a tecnologia da informação surge como um facilitador, reduzindo a complexidade de utilização destes modelos através de ferramentas computacionais. Uma das formas de realizar isto é através da inteligência artificial que permite a construção de modelos autônomos para auxiliar investidores no processo de tomada de decisão em aplicações financeiras do mercado acionário.

Observa-se na literatura que técnicas de reconhecimento e classificação de padrões de *candlestick* aplicados em cotações, assim como, as técnicas de redes neurais usadas na definição de preços futuros de ações e, finalmente, as técnicas de indicadores técnicos para determinar a tendência do mercado possuem características inerentes que as tornam complementares e adequadas para o desenvolvimento de um modelo combinado para realização de predição aplicado à bolsa de valores. E que a partir de comitês de máquina de aprendizado é possível combinar diferentes técnicas de predição em uma única soluções no mínimo tão boa quanto a melhor de suas soluções individuais.

## 2 BOLSA DE VALORES

Conforme Pinheiro (2005), a Bolsa de Valores é um local que oferece, de forma transparente, condições e sistemas necessários para a realização de negociações de compra e venda de títulos e valores mobiliários. Além disto, tem como responsabilidades orientar e fiscalizar os serviços prestados por intermediários e facilitar a divulgação de informações sobre as empresas e negócios que são realizados sob seu controle. Os pregões diários representam um mercado contínuo e propiciam liquidez às aplicações que são negociadas.

Segundo Prochnow (2013), no Brasil, a bolsa de valores é representada pela BM&FBOVESPA, popularmente conhecida por Bovespa, que é o resultado da fusão entre dois grandes mercados brasileiros, a Bolsa de Valores de São Paulo (Bovespa) e a Bolsa de Mercados & Futuros (BM&F), sendo a única bolsa de valores, mercadorias e futuros atualmente em operação no Brasil. Esta fusão ocorreu em meados de 2008 e gerou a maior bolsa da América Latina, sendo considerada uma das maiores do mundo. Segundo o relatório anual publicado pela BM&FBOVESPA no fim de 2012, o volume financeiro deste ano alcançou um valor histórico de R\$ 1,78 trilhão, assim como número de negociações. Outro dado interessante do relatório mostra a importância e o crescimento da Bovespa, informando que as empresas pagaram um total de R\$ 45, 82 bilhões aos seus acionistas.

É importante situar dentro deste contexto de investimentos dois fatores: onde a Bovespa se encaixa e como ela é acessada. Hierarquicamente, a Bolsa de Valores é um dos elementos do mercado acionário, que por sua vez está inserido no mercado de capitais. Esta estrutura fornece suporte à captação de recursos financeiros através da distribuição e circulação de valores, além de realizar a própria regulamentação deste processo.

Vidotto, Migliato e Zambon (2009), afirmam que o mercado de capitais representa uma das formas de negociação, com o principal propósito de proporcionar liquidez<sup>1</sup> aos títulos emitidos pelas empresas e viabilizar seu processo de capitalização. Corroborando com Pinheiro (2005), o mercado de capitais é um conjunto de operações de transferência de recursos financeiros, realizadas entre agentes poupadore e investidores por meio de intermediários financeiros. Neste mercado encontram-se as instituições financeiras como a bolsa de valores, bancos e companhias corretoras de investimento.

---

<sup>1</sup> Em termos financeiros, liquidez significa a facilidade que um ativo pode ser convertido em moeda de troca, já no mercado de capitais este conceito está diretamente ligado ao interesse dos investidores em negociar um ativo, podendo variar dentro de diversos fatores como o tipo de investimento, risco agregado, perspectiva de lucro ou do segmento do mercado que o ativo pertence.

As corretoras são instituições autorizadas para representar os investidores, seus clientes, diretamente na bolsa de valores. De acordo com Costa (2012), as corretoras possuem como função atuar como um intermediador entre as transações de compra e venda de ativos por parte de seus clientes. Além desta função, também compete às corretoras proporcionar um ambiente transparente e seguro para que os investidores realizem seus negócios com os ativos de interesse, com uma ressalva que somente através das corretoras os investidores possuem acesso aos sistemas de negociação da BM&FBOVESPA.

O mercado acionário, ou de ações, é um segmento do Mercado de Capitais que tem como principal função promover a negociação de ações, compreendida em dois momentos distintos: a colocação primária de ações novas emitidas pelas empresas e a negociação secundária de ações que já foram colocadas em circulação previamente. A negociação primária permite a capitalização das empresas por disponibilizar parte de seu capital para negociações em Bolsa, já a secundária gera uma capitalização para os investidores que realizarem negociações lucrativas. A seguir a definição de ação é feita juntamente com alguns conceitos importantes relacionados a ações.

## **2.1 Ações**

Uma ação representa um título negociável que é a menor parte em que se divide o capital de uma empresa disponibilizado para negociação em bolsa de valores. Segundo Hilgert (2011), um investidor que a adquire transforma-se imediatamente em um sócio da empresa, ou seja, de um negócio, passando a correr os riscos que geram lucros e prejuízos como qualquer outro empresário. O autor ainda complementa sobre a diferença fundamental existente entre a entrada e saída de uma sociedade de maneira administrativa na empresa e via ações, enquanto a primeira requer um processo burocrático intenso e desmotivante, a segunda é feita eletronicamente em alguns segundos.

As ações são classificadas das seguintes formas: tipo, registro, mercado que é oferecida, risco e liquidez. É importante destacar a liquidez como uma importante classificação que pode proporcionar altas rentabilidades quando utilizadas em investimentos bem planejados, tanto a longo quanto a curto prazo. Segundo Rojo e Sousa (2012), ações são agrupadas em duas categorias quanto a sua liquidez:

- Primeira linha (blue chips ): ações com grande volume negociado e grande número de compradores e vendedores, ou seja, possuem alta liquidez e movimentam milhares de negócios diariamente;
- segunda linha (small caps): ações menos negociadas e consequentemente com menor liquidez, possuem alta volatilidade e alto risco agregado devido a sensibilidade a acontecimentos em seu portfólio.

A principal diferença entre investimentos nessas duas categorias está na amplitude dos rendimentos. As ações de primeira linha não tem sua liquidez afetada facilmente. Já as de segunda linha, possuem uma alta probabilidade de impacto em seus resultados, agregando um maior risco ao investimento e consequentemente um maior lucro ou prejuízo. Estas oportunidades de rentabilidade ocorrem pela oscilação de preços que no caso das *small caps* ocorrem com uma maior frequência, e requer uma análise mais minuciosa pelo investidor, pois o risco é mais eminente. (LUEDERS, 2008)

Hilgert (2011) afirma que, diante do cenário econômico atual uma empresa não pode depender exclusivamente dos seus próprios recursos para financiar seu crescimento, visto que o mercado atual envolve tecnologia, intensas competições e globalização. Para contornar as adversidades modernas estas empresas podem recorrer a uma alternativa chamada de abertura de capital que permite a capitalização da empresa através da aquisição de novos sócios que compram uma fração do capital da empresa emitido publicamente através de ações. Esta opção é muito vantajosa para empresas em pleno crescimento, pois há necessidade de recursos para financiar seus projetos de expansão.

Para realizar investimentos de forma consciente no mercado acionário é necessário analisá-lo cuidadosamente para entender seu comportamento. Na subseção a seguir é feita uma descrição detalhada das formas de análise do mercado acionário.

## 2.2 Análises de Mercado

No mercado acionário, existem dois importantes tipos de análises usadas pelos investidores, que são: a análise fundamentalista e a análise técnica. Segundo Souza (2012), estas duas formas de análise representam duas correntes de pensamento distintas, comumente citadas como escola técnica e fundamentalista, Paula e Triches (2010), acrescentam que a utilização de cada uma é passível de discussões, por exemplo, quando utilizada para formar

uma carteira de investimento<sup>2</sup> mais rentável. Deste modo, as pesquisas feitas nessas duas áreas concentram-se em aprimorar as estratégias que são usadas.

A escola fundamentalista se preocupa com o estudo dos negócios financeiros da empresa visando entender melhor a natureza das características operacionais da empresa, e com base nas demonstrações financeiras, pode-se estimar o valor futuro de uma ação (REIS, 2007). Assim, a escola fundamentalista se apoia na análise contábil das corporações como o balanço patrimonial e capacidade gerencial para determinar o preço das ações, influenciado pelas avaliações de perspectivas de rendimentos da empresa.

Já a escola técnica, que possui grande influência das Teorias de Dow e de Elliot, é baseada na evolução dos preços e do volume de transações que envolvem uma ação ao longo do tempo. A avaliação oferecida por esta técnica ocorre através de gráficos, equações ou modelos matemáticos e estatísticos sobre os preços, sem se preocupar com os fundamentos empresariais ou contábeis que geraram as oscilações nos preços. (PAULA; TRICHES, 2010)

As duas formas de análise são suportadas por um grupo de metodologias e ferramentas elaboradas para fornecer suporte à decisão para investidores no mercado acionário, através da solução de um problema comum, antecipar o comportamento futuro do mercado diante de informações do passado e do presente. Guarnieri (2006), comenta que baseado em informações ou conclusões destas duas escolas, o preço do ativo pode ser previsto, assim como a sua direção da oscilação.

No trabalho que está sendo proposto, a análise técnica é de fundamental importância, por vários motivos: 1) utiliza dados passados, diferentemente da análise fundamentalista que usa informações contábeis; 2) possui aplicabilidade em curto prazo, sendo utilizada por muitos investidores no mundo todo; 3) não existe nenhuma restrição ao uso de seus conceitos sobre alguma ação específica, mesmo não conhecendo características individuais dela; e, 4) fornece pontos bem definidos de entrada e saída, o que facilita sua utilização.

Na subseção a seguir serão explicados conceitos importantes da análise técnica como as teorias que a fundamentam, bem como as principais ferramentas utilizadas, iniciando por um histórico de onde ela surgiu.

---

<sup>2</sup> Carteiras de investimento é o nome dado a um grupo de ações que são utilizadas por apenas um cliente ou grupo de clients em um investimento único.

## 2.3 Análise Técnica

A origem da escola técnica tem sua marca no Japão, segundo Noronha (1995), por volta de 1700 quando as famílias japonesas negociavam a produção de arroz. Com o objetivo de atender a demanda crescente que envolvia as negociações do arroz, em Osaka, foi criada a chamada Bolsa de Osaka que possuía o papel de centralizar as negociações e precificações desta mercadoria. As primeiras notícias que se tem conhecimento sobre o uso de técnicas voltadas para predição, projeção de preços futuros, são da família Homma que utilizava gráficos para monitorar os preços nesta bolsa, como discorre Guarnieri (2006b). No início do século XIX, Charles Dow escreveu diversos artigos sobre o comportamento do mercado acionário para o *Wall Street Journal*. Seu trabalho foi direcionado para análise sobre a oscilação dos preços dos ativos na bolsa de Nova York, chegando à conclusão que este movimento obedecia a uma tendência. Além disto, Charles Dow e Edward Jones criaram um índice que era capaz de refletir o comportamento médio do mercado através da média de evolução de preços na Bolsa de Nova York originando o famoso índice Dow Jones Industrial Average. De posse de um índice de mercado e da tendência dos ativos os investidores poderiam nortear suas decisões de compra e venda de ativos e consequentemente atingir lucros maiores. Em 1932, Robert Rhea apresentou uma obra chamada *Dow Theory* que foi um aperfeiçoamento da obra *The Stock Market Barometer* de William Hamilton (1921) quando os editoriais publicados por Dow foram organizado e complementados na forma dos princípios conhecidos atualmente como Teoria Dow. (GUARNIERI, 2006a)

Segundo Noronha (1995), a Teoria Dow baseia-se em seis princípios, que são:

- Princípio 1 (Os índices descontam tudo): os índices refletem todos os fatos e ocorrências, pois sua formação se refere à atividade combinada de milhares de investidores. Segundo o Banco Central, os índices de preços são números que agregam e representam os preços de uma determinada cesta de produtos e sua oscilação mede a variação média dos preços dos produtos que o formam, fazendo uma analogia ao mercado acionário os produtos seriam as ações;
- Princípio 2 (o mercado apresenta três tendências): O segundo aborda a movimentação dos preços em tendências, com três movimentos diferenciados pelo prazo, sendo a primária com duração de um ano, a secundária com duração de alguns meses e a terciária de algumas semanas;

- Princípio 3 (a tendência primária possui três fases): a tendência primária é dividida em três fases tanto para mercados de alta quanto de baixa. Nos mercados de alta a primeira fase é chamada de acumulação momento pelo qual os investidores mais bem informados entram no mercado, seguida da segunda fase nomeada de alta sensível, onde os ditos seguidores dos primeiros investidores também entram no mercado, e finalmente a terceira fase chamada de euforia quando os investidores bem informados saem do mercado. Após estas fases o período de alta encerra e ocorre uma inversão na tendência sendo iniciada a baixa. De modo análogo ao período de alta, a baixa é marcada pelas fases distribuição, pânico e baixa lenta;
- Princípio 4 (as reversões precisam ser confirmadas): O quarto princípio reforça a necessidade de índices ou indicadores de mercado para confirmarem as reversões de tendência, que são o período de interseção entre duas tendências em sentidos opostos, uma de alta e outra de baixa em qualquer ordem;
- Princípio 5 (o volume acompanha a tendência): O quinto princípio informa que o volume de negociações que ocorreram no mercado acompanha a tendência, seja de alta ou de baixa;
- Princípio 6 (a tendência perdura até ser confrontada): a extinção de uma tendência ocorre apenas quando confrontada por uma oposta, a tendência perdura até uma confirmação contrária;

Ainda no século XIX, Ralph Elliott estudando o índice Dow Jones, identificou um padrão recorrente no comportamento da oscilação deste índice formando ondas repetidas de avanço e declínio. Seus estudos culminaram em uma publicação chamada “Princípio das Ondas”. O trabalho de Elliott representa um complemento aos estudos de Dow, por idealizar a existência de tendências, onde afirma que a aparente desorganização do mercado acionário não existia, mas sim a ocorrência de um padrão de precificação contínuo e recorrente através da formação de ondas. A base desta teoria encontra-se em conceitos distintos entre a filosofia e a matemática. Sua base filosófica é fundamentada na busca em outras ciências pela explicação dos ciclos no mercado, e durante suas pesquisas Elliott chegou à conclusão que tudo na natureza possui ciclos e como o mercado é movimentado por investidores, que são influenciados por leis da natureza, então justifica o mercado se mover em ciclos. (GUARNIERI, 2006a)

A fundamentação matemática desta teoria fez-se necessária quando Elliot precisou dar exatidão ao padrão identificado pela sua base filosófica. Aprofundando suas pesquisas Elliot encontrou em estudos do matemático Leonardo Fibonacci as explicações que procurava. Fibonacci em uma publicação “Liber Abaci” responde a uma questão de reprodução de coelhos ao longo de um ano através da série matemática chamada de Sequência de Fibonacci. Elliot percebeu que a razão de 0,618 entre os números da sequência se aplica perfeitamente a sua teoria, auxiliando na projeção dos movimentos de avanço e declínio das ondas (Guarnieri, 2006b). Também são utilizados valores percentuais, no caso 62% e seu complemento de 38%, e na prática alguns analistas de mercado ainda utilizam o valor de 50%.

Na teoria das ondas de Elliot, um ciclo completo é formado por oito ondas, cinco durante a tendência de alta e três na de baixa, dividindo o ciclo em dois grupos ou ciclos menores, o de alta e o de baixa, como pode ser observado na Figura 1. Observa-se que no ciclo de alta existem três ondas de crescimento chamadas de propulsoras e representadas na figura 1 pelos números 1, 3 e 5, ainda no ciclo de alta as ondas de número 2 e 4 são chamadas de corretivas valores. Observa-se que o ciclo de baixa existem duas ondas retrativas (a e c) divididas por uma onda propulsora (b). As ondas propulsoras impulsionam o mercado para valores mais elevados e as corretivas ajustam os preços em um valor mais estável durante o crescimento, já a onda retrativa é oposta à propulsora do outro ciclo.

Figura 1 pelos números 1, 3 e 5,

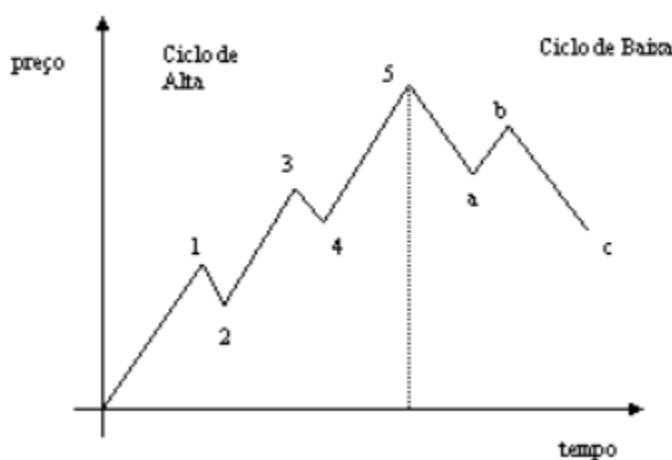


Figura 1 - Ciclo de ondas de Elliot.  
Fonte - Guarnieri, 2006

A análise técnica, ou também chamada de grafista pela intensa utilização de gráficos para auxiliar ou justificar suas análises, possui como foco a evolução dos preços das

ações ao longo do tempo conforme a utilização de duas principais ferramentas: os gráficos e os indicadores técnicos. Esta análise utiliza dados referentes aos preços das ações ocorridos no passado para gerar sinalizações de compra e venda de ações, além de aplicar estas cotações no cálculo e representação das ferramentas citadas.

## Representação gráfica

A utilização dos gráficos concentra-se mais na parte visual da análise técnica e, desta maneira, utiliza gráficos como ferramenta para determinar os melhores momento e preço para comprar e vender ações. Basicamente, as previsões são realizadas através da identificação de elementos no gráfico de cotações que indicam a continuação ou inversão de uma tendência. Os gráficos de preços são divididos, de acordo com sua forma de representação, em três categorias: o de evolução de preços, o de barras e o de *candlesticks*. É possível afirmar que estas categorias, respectivamente citadas, evoluem no quesito representação.

O gráfico de evolução de preços que é apresentado como um simples gráfico que usa apenas um único valor referente ao preço do ativo, normalmente utiliza o preço de fechamento. A Figura 2 ilustra a representação nomeada por evolução dos preços, este nome se refere exatamente ao que o gráfico informa que é apenas a evolução dos preços. Esta forma de representação não oferece muitos benefícios para os analistas de mercado por utilizar somente um dos quatro valores que a ação possui por período, que são a abertura, fechamento, máximo e mínimo. Para uma rápida visualização de como está a oscilação de preços da ação dentro de um período este gráfico é bem útil, mas planejar estratégias de investimento com apenas um dos valores da ação pode ser pouco efetiva e possivelmente bem arriscada.



O gráfico de barras foi uma tentativa ocidental para representar melhor a evolução dos preços adicionando outros valores como abertura, fechamento, máximo e mínimo. A representação gráfica em barras, ilustrada pela Figura 3, fornece um grande benefício ao utilizar a barra que possui todos os valores da cotação da ação, quando comparada a anterior que usava apenas um dos valores (preço de fechamento), o que restringia muito as possíveis análises sobre seus valores. Esta exibição é mais rica em detalhes e permite ao analista de mercado oferecer uma quantidade maior de sugestões baseadas nas informações deste gráfico. Uma das poucas desvantagens aparentes desta representação é que não é possível, rapidamente, visualizar se o período informado por cada barra encerrou com lucro ou prejuízo, pois não é muito aparente esta informação.

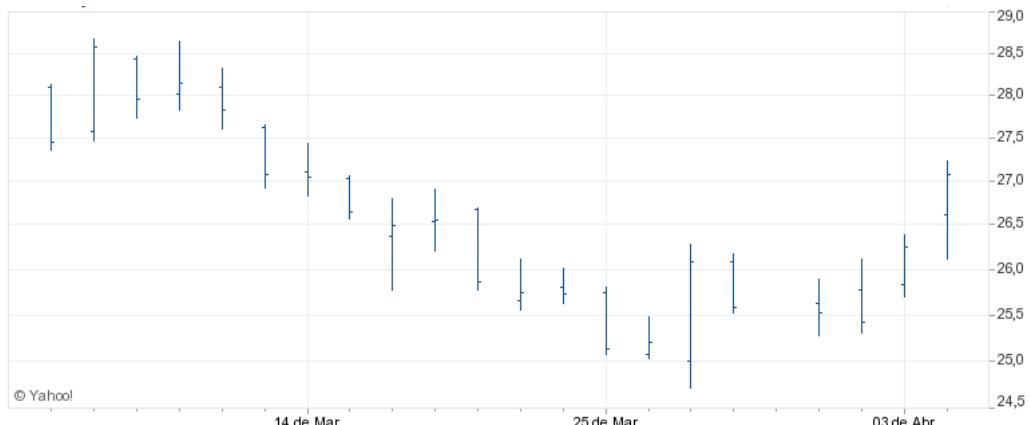


Figura 3 - Gráfico de barras

Já o gráfico de *candlestick*, de *candle* ou candelabro japonês, é uma representação oriental que possui a melhor representação visual da oscilação de preços das ações, pois envolve até informações visuais se o preço encerrou maior ou menor que iniciou, informando lucros ou prejuízos por período, que nesta representação significa por *candlestick*.

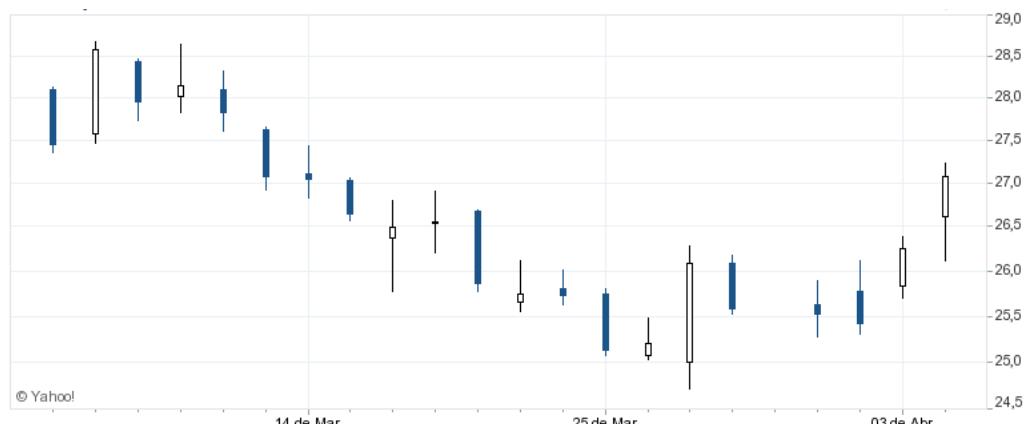


Figura 4 - Gráfico de candlesticks

A última forma de representação gráfica fornece uma maior quantidade de informação no gráfico através do conjunto de figuras individuais *candlestick*, ilustrados na Figura 4. Os benefícios desta representação são aparentes ao observar o gráfico. De imediato, duas características chamam a atenção e proporcionam análises mais poderosas que as fornecidas pelos outros, por exemplo: utilização de cores para diferenciar períodos positivos de negativos e a formação de figuras para cada período. Atualmente, analistas de mercado, denominados grafistas, usam esta representação para fornecer sugestões para seus clientes. Além disto, os *candlesticks* acarretam em outro grande benefício com a possibilidade de utilizar estas figuras como padrões recorrentes no mercado e analisar seus acontecimentos para gerar previsões futuras.

Nos gráficos de *candlestick* os padrões são, na maioria das vezes, representados individualmente através de *candlesticks*, mas também podem ser representados por um grupo de *candlesticks*. Um *candlestick* representa uma figura individual do gráfico de *candles* sendo formado por quatro valores: o preço de abertura, fechamento, máximo e o mínimo. (DEBASTIANI, 2007). Cada *candlestick* representa a variação da cotação do ativo para um intervalo de data, podendo ser de minutos, horas, dias, semanas e assim por diante. Ou seja, um *candlestick* diário apresenta os preços diários referente aos quatro valores, um candlestick semanal os preços referentes a semana. A Figura 5 exibe a estrutura básica de um *candlestick*.

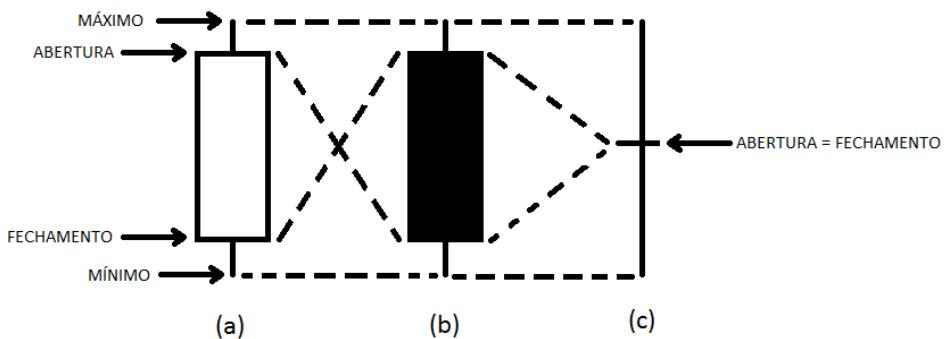


Figura 5 - Estrutura e tipos de *candlestick*.

O espaço definido entre o preço de abertura e fechamento representa o corpo do *candlestick*. Por trás deste corpo existe a linha vertical que constitui o intervalo de valores entre o preço máximo e mínimo do *candlestick* que quando são maiores que o corpo formam as conhecidas sombras superiores e inferiores. Quando o valor de fechamento é superior ao de abertura será formado um *candle* branco ou de cor clara informando lucro ou valores positivos (a), mas quando o valor de fechamento for inferior ao de abertura o corpo do *candle*

será preto ou alguma cor escura, indicando prejuízo ou valores negativos (b), e em um terceiro e último caso o valor de abertura pode ser igual ao de fechamento formando um *candle* sem corpo (c). Algumas literaturas apresentam as cores referentes ao direcionamento do padrão de outras formas, como verde para *candle* positivo e vermelho para negativo, o importante é apenas diferenciá-los. Neste trabalho será adotado o padrão branco e preto, respectivamente para encerramento em alta e em baixa.

Alguns padrões são representados apenas por *candlesticks* individuais, mas são ricos em informações úteis para antecipar o comportamento futuro da ação, além de permitir sugestões para tomada de decisão. A outra forma de ocorrência de padrões *candlestick* ocorre através da formação de grupos de *candles*, que por sua vez fornecem valores igualmente importantes no que diz respeito à predição (alguns padrões de *candle* utilizados neste trabalho encontram-se no Anexo A). A capacidade de combinação entre *candles* para formar novos padrões gera uma infinidade de possibilidades de formação de padrões. Já existem vários catalogados com algumas informações associadas aos seus acontecimentos, mas é notável que todas as possibilidades de padrões ainda não foram descobertas e é possível a existência de algum padrão importante ainda não conhecido. Acredita-se que os padrões recorrentes seguem uma ordem, não assumindo comportamento aleatório, como explica Paula (2010). Além disto, estes padrões não são simples manifestações de dados econômicos, mas também representam a combinação entre emoção e lógica dos agentes econômicos que atuam neste mercado.

## **Indicadores técnicos**

Um indicador técnico na maioria das vezes é definido apenas por uma equação matemática que utiliza como dados de entrada os valores da cotação de uma ação, normalmente apenas o preço de fechamento. No início, os indicadores possuíam função complementar a análise do mercado e não apresentavam peso ou confiança para ser fator decisivo nesta análise. Em contrapartida, hoje em dia este instrumento é considerado parte importante do processo de tomada de decisão dos investidores, sendo peça fundamental para análise do comportamento futuro do ativo. (NORONHA, 1995)

Os indicadores técnicos oferecem benefícios para a escola técnica principalmente com o surgimento da tecnologia da informação que facilitou a incorporação de cálculos matemáticos e estatísticos de maneira mais ágil para atingir resultados que auxiliassem o processo de tomada de decisão de investidores.

Independente do processo de elaboração do indicador, seja ele envolvendo simples ou complexas equações matemáticas ou um grande número de variáveis, não é prejudicial utilizar um simples procedimento de cálculos, como o de médias aritméticas, para analisar o mercado (PAULA; TRINCHES, 2010). As médias móveis representam uma forma muito eficaz de acompanhar a tendência do mercado e consistem em médias sequenciais (móveis) de um conjunto (janela de tempo) de preços da ação. Este indicador é eficaz para acompanhar, com um pequeno atraso, a tendência que os preços seguem e, para realizar seu cálculo é necessário definir prioritariamente o seu período da aplicação ou janela de tempo. Esta janela informa quantos valores farão parte do cálculo da média e à medida que novos valores são inseridos, os mais antigos deixam de participar do cálculo, por isso o nome média móvel.

A Figura 6 ilustra a representação de uma média móvel de cinco dias sobre um gráfico de *candles*. É importante perceber que a janela temporal de cálculo da média informa o tamanho dos segmentos consecutivos da cotação que irão ser submetidos incrementalmente ao cálculo da média móvel, e o nome móvel significa o movimento que este segmento faz no decorrer do conjunto de valores até atingir seu fim.

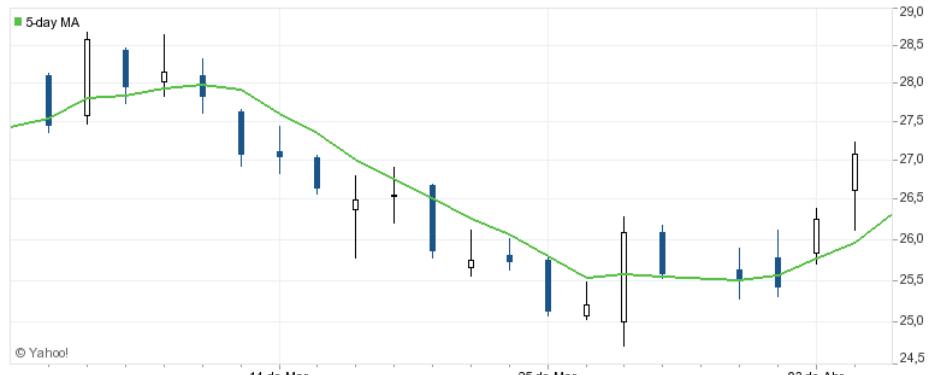


Figura 6 - Representação gráfica da média móvel de 5 dias

A forma mais simples de calcular este indicador é feita através de uma simples média aritmética, também conhecida por média móvel simples - MMS (*Simple Moving Average* - SMA), como ilustrado pela Figura 6 uma média móvel de cinco dias  $MMS_5$  e representada pela Equação 2.1, onde para cada valor  $MMS_N$  será calculada a média de valores  $X_i$  nos últimos  $N$  dias.

$$MMS_n = \frac{\sum_{i=1}^n x_i}{n} \quad (2.1)$$

A média móvel é um indicador bem flexível, o que permite a criação de inúmeras variações a partir de pequenas alterações em sua equação base. O que é fundamental para seu cálculo é apenas a janela temporal que desliza ao longo dos valores que são passados para sua equação. Uma variação muito utilizado destas médias está presente no trabalho de Guarnieri (2006) denominada média móvel exponencial - MME (ou *Exponential Mean Average – EMA*), definida de acordo com a Equação 2.2, onde o cálculo da média de hoje ( $MME_N$ ) está diretamente influenciado pelo valor de ontem ( $MME_{N-1}$ ) e a constante  $k$  é definida pelo valor de  $\frac{2}{N+1}$ .

$$MME_N = MME_{N-1} + k * (X_i - MME_{N-1}) \quad (2.2)$$

Ao contrário da média móvel simples, a exponencial atribui maior importância aos valores mais recentes, proporcionando uma suavização nos resultados e por consequência reduz a quantidade de sinais de compra e venda gerados. Os valores utilizados pelas médias móveis supracitadas são formados a partir de uma sequência ordenada de valores, no caso das cotações diárias uma sequência de preços de fechamento por dia. Isto caracteriza, de maneira superficial, uma série temporal que será explicada de maneira mais detalhada na subseção a seguir.

## 2.4 Séries Temporais

De acordo com Villanueva (2006), o estudo de séries temporais principalmente almejando a predição, ganhou destaque nos últimos anos. Segundo Soares (2008), uma série temporal é definida por um conjunto de observações sequenciais. Observar um determinando fenômeno ao longo do tempo define informalmente uma série temporal, então, dados coletados ao longo do tempo podem ser representados através de uma série temporal (MALETZKE, 2009).

Uma série temporal  $Z$  de tamanho  $n$  é uma coleção ordenada de valores  $Z_t$ , com  $t$  variando de 1 a  $n$  tal como a Equação 2.3.

$$Z = (Z_t; Z_1, Z_2, \dots, Z_n) \quad (2.3)$$

Onde:

- $Z_t$  corresponde ao valor da série temporal no instante  $t$ ;
- $t$  representa o período de observação assumindo valores de **1 a n**.

As observações  $Z_t$  forma uma série temporal  $z_1, z_2, \dots, z_n$ , definida por valores contínuos em um período ininterrupto ou em intervalos constantes. Desta maneira, a natureza das observações sobre este período classifica a série temporal em dois tipos contínua ou discreta. Uma série é dita contínua quando as observações são realizadas ao longo do tempo, ininterruptamente, as discretas são observações realizadas em intervalos de tempo especificados, espaçados igualmente. Um exemplo simples que facilita a compreensão destes dois tipos de séries utiliza como base a medida de horas em uma série contínua que possui observações no período de uma hora, já na série discreta seus valores seriam observados a cada hora.

É possível efetuar um processo chamado de discretizar uma série contínua, ou simplesmente obter uma série discreta a partir de uma contínua. Para que isto aconteça é necessário realizar um mapeamento entre os valores contínuos em um número reduzido de intervalos através de um processo de transformação dos dados. É importante destacar que existe diferença entre natureza da observação e os valores dos dados observados, o primeiro pode ser descrito de maneira contínua enquanto seus dados podem ser discretos. (MALETZKE, 2009)

Além da definição formal, uma série temporal pode ser expressa de forma gráfica, como ilustra a Figura 7. A representação gráfica de uma série temporal permite a realização de algumas análises prévias sobre os dados as quais são demasiadamente complexas diante de sua representação a partir de uma equação, como observado visualmente as ondas de Elliot, comentada anteriormente.

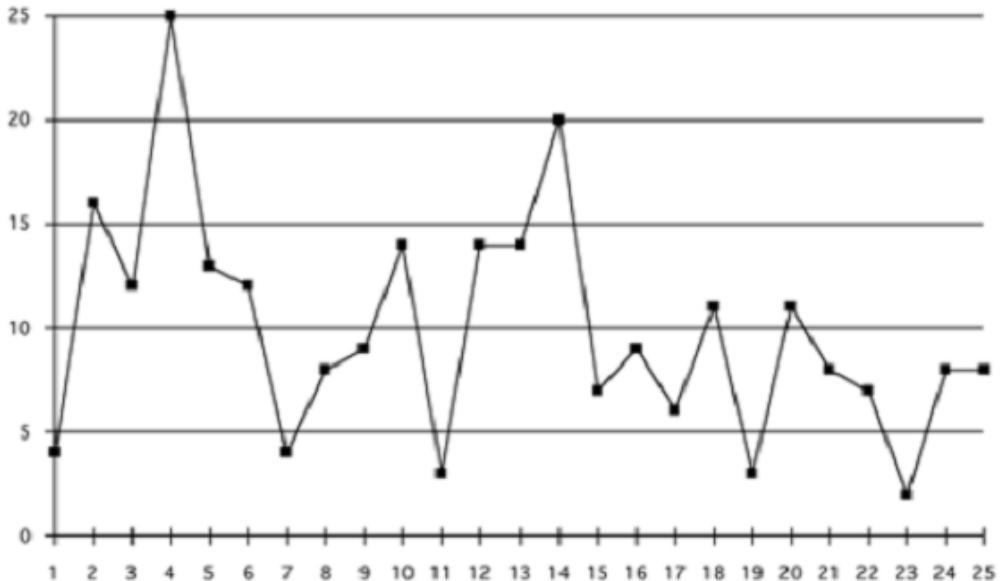


Figura 7- Representação gráfica de uma série temporal

A partir da curva de valores apresentada pelo gráfico de uma série temporal, algumas características são observadas. Alguns autores apontam para quatro elementos formadores de uma série temporal:

Segundo Souza (2008, p. 13), são eles:

- Tendência: sentido de deslocamento da série em alguma direção;
- Ciclo: movimento ondulatório repetido que pode vir a ser periódico;
- Sazonalidade: movimento de curta duração que representa mudanças eventuais;
- Ruído aleatório: variabilidade intrínseca aos dados que não pode ser modelada.

A tendência é o elemento mais importante para a abordagem proposta dentro do contexto financeiro de bolsa de valores, e compreende um movimento unidirecional, crescente ou decrescente, que uma série temporal assume por um longo período de tempo. É discutível o que seria este longo período, e acima de tudo o período das observações influencia a magnitude deste período, sendo necessário considerar o número de observações disponível. Por exemplo, se o período total de observações é de 5 anos e a série apresenta um movimento unidirecional por 1 ano consecutivo, é possível afirmar que neste período a série apresenta uma tendência. O valor definido para “longo” no mercado financeiro está comumente relacionado a investimentos com o prazo maior que um ano, neste caso cotações também são representações de séries temporais e apresentam tendências neste prazo, do mesmo modo existem tendências de médio e curto prazo no mercado acionário. Assim como as cotações, os indicadores igualmente representam exemplos de séries temporais apresentando tendências e todas as outras características.

Independente de suas características ou classificações, o estudo realizado sobre séries temporais possui vários objetivos distintos, em algumas situações o foco da análise é a estrutura da série, ou seu comportamento. Existem casos até que o principal interesse está voltado para relação entre duas séries temporais diferentes ou segmentos de uma mesma série (SOUZA, 2008). De modo geral, os principais objetivos relacionados ao estudo de séries temporais são listados a seguir:

- Descrição: propriedades das séries, padrão de tendência, alterações estruturais;
- Explicação: o comportamento de uma série explica outra;
- Controle de processos: controle estatístico de qualidade, definição de portfólio;
- Predição: antecipar valores futuros baseado em valores do passado.

A aplicação de séries temporais ocorre em vários domínios e a condição da relação temporal existente entre os dados é um elemento muito importante em diversos domínios, o que pode resultar em uma análise mais completa dos dados. Dentro do contexto da predição, é possível afirmar que a mesma só ocorre devido a dependência entre os valores que compõe a série. A elaboração de um modelo de predição é focada exatamente nesta relação entre os elementos do passado para, baseado nisto prever valores futuros. (VILLANUEVA, 2006)

Para exemplificar como a predição pode ser útil e aplicada em vários cenários, segue as seguintes situações descritas por Maletzke (2009). A identificação a partir perfil de vendas de um produto em relação a outro de acordo com um período do ano, e neste período intensificar campanhas diante desta previsão de demanda. A depreciação de peças de máquinas pode ser disposta em uma série temporal com o objetivo de maximizar o funcionamento do equipamento como um todo em detrimento da previsão de manutenção de suas partes. Na área médica, a partir de um eletrocardiograma de um paciente, é possível identificar anormalidades que representem patologias futuras não perceptíveis pelo olhar dos médicos, auxiliando neste diagnóstico. Na área financeira, mais importante para este trabalho, busca-se prever comportamentos futuros de certas observações como a de ações cotadas na bolsa. Acredita-se que esta tarefa de predizer o mercado financeiro seja altamente complexa, pois muitos defendem que o comportamento deste mercado possui um alto nível de não-determinismo.

### **3 PREDIÇÃO DE SÉRIES TEMPORAIS**

O início de abordagens de predição data de antes de 1920, mas somente com a criação de Yule, em 1927, que um modelo foi proposto utilizando a dependência entre a predição e valores antigos. Com isto, acreditava-se que o comportamento da série dependia de modelos lineares, por mais de uma década. Entretanto, foi percebido ao longo do tempo que modelos lineares não eram suficientes para representar bem séries temporais, principalmente pela característica não-linear que boa parte das séries apresentavam em suas tendências.

Com o advento do computador na década de 50, ocorreu uma revolução nas abordagens de predição propostas até então, devido a grande ajuda proporcionada pela tecnologia da informação. Nesta época vários métodos foram propostos e novos estudos foram realizados nesta área de predição. Por volta de 1976, foi elaborada a técnica Box & Jenkins, que representa uma metodologia completa de predição dividida em várias fases, iniciando na identificação de um modelo que mais se ajusta aos dados, seguida da estimação dos parâmetros e finalizando com a verificação da validade do modelo executado. Este modelo possui algumas restrições, as quais podem ser consideradas desvantagens, como um tamanho mínimo para amostra e a interferência de um analista na fase de identificação do modelo. (PASQUOTTO, 2010)

Após a elaboração do Box & Jenkins, diversas técnicas vêm sendo projetadas, elaboradas e testadas com um objetivo em foco: produzir um modelo de predição o mais preciso possível. Diante disto, os esforços foram direcionados para automatização dos modelos através de abordagens baseadas, por exemplo, na área de Inteligência Artificial (IA), pois o desejo por modelos autônomos (inteligentes) é tão forte quanto atingir bons níveis de precisão para os valores previstos. Neste contexto existem modelos inspirados em comportamentos da natureza, como o apresentado por Cortez e Neves (2002), que discorre sobre as Redes Neurais Artificiais (RNA) que imitam o cérebro humano e também Lima (2004) que apoiados pela ideia da existência de alguma incerteza na escolha de um modelo melhor, a opção de combiná-los poderá reduzir a instabilidade da predição, melhorando sua precisão, com a comprovação atingida em seus experimentos informando melhores taxas de convergência para combinação de modelos.

Segundo Costa (2008), o desenvolvimento de técnicas de predição possuiu um avanço considerável nos últimos anos, fundamentado em áreas de conhecimento como: matemática, estatística e computação. Pesquisadores vêm desenvolvendo abordagens, técnicas

e algoritmos fazendo uso de previsões em séries temporais, tais como Marques e Gomes (2009), Oliveira (2007), Parracho e Pattern (2010) e Faria (2008a). Tais trabalhos demonstram a aplicabilidade das técnicas de previsão na antecipação do comportamento do mercado acionário. Existem muitos métodos para realizar a previsão de séries temporais, como observado por Aikes Junior *et al* (2012), desde modelos estatísticos bem complexos a outros simples e intuitivos, cada um apresentando suas próprias capacidades e limitações. Ainda assim, uma mesma série pode ser analisada e prevista por muitos modelos diferentes.

A análise de séries temporais possibilita a previsão de valores futuros, o que permite oferecer suporte à tomada de decisão em várias áreas como: Produção, Recursos Humanos, Marketing e Financeira. Tais técnicas utilizam suporte computacional a fim de definir as projeções e tendências do futuro. Porém, como advertido por Souza (2006), é necessário tomar cuidado com o uso da previsão para apoiar decisões importantes, considerando que as mesmas estão sujeitas a erros. Os principais conceitos que envolvem previsão de séries temporais são abordados na subseção seguinte.

### 3.1 Definições

A previsão de séries temporais, de acordo com Souza (2006), é compreendida como o ato de antecipar valores futuros com base no conhecimento de valores passados ou existentes até o atual momento. A quantidade de valores futuros é denominada neste contexto por horizonte ou janela de previsão.

De maneira mais formal, a previsão de séries temporais pode ser definida da seguinte forma: seja uma série temporal  $Z = (Z_t; t = 1, 2, \dots, n)$ , sendo  $t$  o período de tempo,  $S_t$  as variáveis observadas e o horizonte de previsão definido por  $h$ . O valor futuro desta série é determinado após seu período de observação, logo, este valor corresponde a  $Z_{n+h}$ . Com a condição mínima de existência da previsão estabelecida sobre a restrição de  $h \geq 1$ .

De acordo com Souza (2008), as abordagens previsão de séries são organizadas nas seguintes categorias:

- Técnicas descritivas: a representação gráfica dos dados é fundamental para revelar padrões comportamentais importantes como tendências e ciclos;
- Técnicas lineares: utiliza modelos probabilísticos e métodos não paramétricos como alisamento exponencial;

- Técnicas não-lineares: englobam modelos não-determinísticos, com comportamento relativamente aleatório.

Devido ao comportamento não-linear apresentado por séries temporais financeiras, de acordo com Veiga (2012), acredita-se que os modelos mais adequados para predição destas séries sejam também os não-lineares, por conseguirem mapear adequadamente as características apresentadas pelas séries. Entretanto, o custo computacional para sua representação é numericamente maior que os lineares, além da dificuldade de formulação matemática devido ao ambiente possuir comportamento aleatório. O modelo mais utilizado desta categoria são as redes neurais artificiais (RNA), e como afirma Cupertino (2009), elas apresentam um grande diferencial pois na maioria dos casos uma RNA é um sistema adaptável que muda sua estrutura de acordo com informações externas ou internas a rede. Villanueva (2006), acrescenta que por serem aproximadores universais, a RNA é adequada e amplamente utilizada para predizer séries temporais.

Por outro lado, pesquisadores sugerem a utilização de abordagens híbridas de predição para melhorar os resultados apresentados por redes neurais, isto ocorre através da composição de múltiplos previsores em uma solução unificada. Como observado nos trabalhos de Lima (2004) e Nikoo et al (2007), ao aplicarem abordagens de combinação como modelo de predição de séries temporais.

Um modelo simples e bem eficaz de predizer o comportamento de uma série temporal é representado pelas médias móveis, como descrito por Guarnieri (2006), Oliveira (2007) e Faria (2008a). Estes trabalhos demonstram a aplicação financeira das médias móveis no domínio do mercado acionário, e diante das complexas oscilações de preços das ações esta técnica antecipa expectativas de preços futuros (crescimentos e quedas no mercado) permite acompanhar os valores com um pequeno atraso. Caso seja utilizada como principal método de predição para investimentos pode ser considerada uma desvantagem, por outro lado, do ponto de vista de uma ferramenta auxiliar de investimentos é cogitado como limitação.

As principais abordagens de predição utilizadas como base para o modelo proposto neste trabalho serão descritas detalhadamente nas subseções a seguir.

### **3.2 Indicadores Técnicos**

Os indicadores econômicos são bem conhecidos como técnica para predição financeira, especialmente os indicadores técnicos representantes da análise técnica do

mercado acionário, considerados pioneiros no segmento de predição. Segundo Costa (2008), indicadores técnicos são utilizados, principalmente, para informar a tendência atual do mercado e momentos de reversão, o que permite sugerir quando comprar e vender ações almejando o máximo lucro nos pontos de reversão da tendência. A ocorrência de tendências é um dos grandes atrativos para predição de séries temporais financeiras, como relatado por Souza (2006), quando ocorrem tanto com inclinações positivas quanto negativas.

Noronha (1995), acrescenta que no início os indicadores possuíam função complementar a análise do mercado e não apresentavam muita influência como fator decisivo na análise, em contrapartida, hoje em dia este instrumento é considerado como fundamental no processo de tomada de decisão dos investidores na análise do comportamento futuro das ações.

Dentre os vários indicadores técnicos empregados para antecipar reversões, identificar e confirmar tendências, destacam-se: o *Relative Strength Index* (RSI), o *Average Directional Index* (ADX) e as *Bollinger Bands* (BBS), estes indicadores serão comentados separadamente a seguir

### 3.2.1 Relative Strength Index (RSI)

O RSI foi criado por Welles Wilder em 1978 para medir a força e a velocidade do movimento atual dos preços comparando a média dos preços que fecharam em alta pelos que fecharam em baixa em uma determinada janela de tempo. Seus valores variam entre 0 e 100 eliminando distorções na evolução dos preços através da suavização de movimentos bruscos de alta ou baixa, como definido pela Equação 3.1, onde  $\sum\varphi$  é a média dos preços que fecharam em alta e  $\sum\omega$  é a média dos preços que fecharam em baixa. De acordo com Costa (2008), é comum utilizar a janela de tempo de nove (9) ou catorze (14) dias para calcular estas duas médias respectivamente, sendo muito importante perceber que quanto menor a janela de tempo mais sensível será o indicador a grandes oscilações e maior será sua amplitude nos resultados.

$$RSI = 100 - \frac{100}{1 + \frac{\sum\varphi}{\sum\omega}} \quad (3.1)$$

A aplicação deste indicador depende exclusivamente de dois valores limites para seu resultado, escolhidos em alguns casos por 30 e 70, como descrito por Costa (2008), ou

como Bruni (2009), define de maneira mais rigorosa como 20 e 80. Estes dois valores definem dois momentos específicos onde serão aplicadas estratégias de investimento sugeridas pelo indicador. Para o primeiro caso, quando o resultado está inferior a 30 a ação é considerada como sobre vendida<sup>3</sup>, portanto quando o RSI ultrapassar o valor 30 será informado um sinal de compra naquele instante. Por outro lado, quando este mesmo indicador apresenta resultados superiores a 70 a ação é considerada sobre comprada<sup>4</sup>, de maneira análoga ao sobre vendido, após isto será iniciada uma reversão informando declínio de valores e quando o resultado do indicador cair para valores inferiores a 70 será informado um sinal de venda.

### 3.2.2 Average Directional Index (ADX)

O ADX, também criado por Welles Wilder em 1978 é classificado como um direcionador que identifica a intensidade de uma tendência e sua direção através da diferença entre os valores mínimos a máximos analisados em uma janela de tempo (VORA, 2010). O seu valor flutua entre 0 e 100, onde valores acima de 60 são muito raros e os valores limites empregados no indicador são de 20 para o mais baixo e 40 o mais alto. Valores inferiores a 20 indicam uma tendência fraca e maiores que 40 uma forte tendência, não informando a direção da tendência apenas sua intensidade.

O ADX é derivado dos indicadores Indicador Direcional Positivo ( $DI^+$ ) e Negativo ( $DI^-$ ), também desenvolvidos por Wilder, os quais formam a base de cálculo do ADX, como ilustra a Equação 3.2. As informações completas sobre as equações dos indicadores direcionais ( $DI^+$  e  $DI^-$ ) encontram-se no Anexo B.

$$ADX = 100 * MME_n \left( \frac{|DI^+ - DI^-|}{DI^+ + DI^-} \right) \quad (3.2)$$

O valor calculado pelo  $DI^+$  mede a força dos movimentos altistas<sup>5</sup>, já o  $DI^-$  a força dos movimentos baixistas<sup>6</sup>, em um período normalmente utilizado de 14 dias, que pode ser alterado de acordo com a preferência do analista. O cálculo destes dois indicadores

<sup>3</sup> Em um momento sobre vendido muitos investidores já foram tentados a vender suas ações e está próximo de atingir o valor mais baixo após um intenso ou prolongado movimento de desvalorização

<sup>4</sup> O momento sobre comprado significa que vários investidores já foram tentados a comprar uma ação após um longo movimento de valorização e a ação está próxima a alcançar seu valor mais alto

<sup>5</sup> Entende-se por movimento altista aquele que estimula uma tendência de alta.

<sup>6</sup> Análogo ao movimento altista, o baixista provoca uma tendência de baixa.

depende de uma  $MME_{14}$  de outro indicador, denominado *True Range*, conforme descrito no Anexo B. Em suma o ADX apenas mede a intensidade da tendência, assim as possíveis interpretações (operações financeiras) são formadas pelo  $DI^+$  e  $DI^-$ , quando suas linhas se cruzam. Um sinal de compra ocorre quando a linha do  $DI^+$  ultrapassa os valores da linha do  $DI^-$ , caso contrário é gerado um sinal de venda quando a linha do  $DI^+$  cruza a do  $DI^-$  para cima. Neste caso o ADX é utilizado para atribuir um grau de confiança nestas operações, quanto mais alto o valor do ADX maior a confiabilidade destas operações.

### 3.2.3 Bollinger Bands (BBs)

As BBs, criadas por John Bollinger nos anos 80, determinam zonas de *sobrecompra* e *sobrevenda* da cotação atual frente aos valores mais recentes sinalizando a continuidade ou interrupção da tendência atual, além de repercutir mudanças na volatilidade do mercado. São definidas por três linhas (bandas): i) a banda central (Banda média) calculada através da média móvel dos últimos  $n$  dia  $MMS_n$ , a banda superior que é definida a partir da soma da banda média com o desvio-padrão, e iii) a banda inferior que é obtida pela diferença entre a banda média e o desvio-padrão. O cálculo das bandas superior e inferior é definido pela Equação 3.3, ou seja, em função do desvio padrão relativo a media móvel ou banda média, onde são normalmente utilizados os valores de  $D = 2$  e  $n = 5$ . Isto informa que estas duas bandas circundam a média em  $D$  vezes o seu desvio padrão para mais ou para menos. (KANNAN; SEKAR; SATHIK, 2010)

$$BBs = BM_{ni} \pm D * \sqrt{\frac{\sum (x_i - BM_{ni})^2}{n}} \quad (3.3)$$

Quando os valores das bandas estão muito afastados denota uma elevada volatilidade das cotações, já quando estão mais próximos significa que existe uma baixa volatilidade. As bandas de Bollinger fornecem algumas interpretações importantes, além de sinais para operar, segundo Kannan (2010), descritas abaixo:

- Oscilação do preço: o valor da cotação tende a ficar entre as bandas;
- Preço alvo: um movimento que inicia-se em uma banda e alcança a outra banda;
- Estreitamento: aponta uma fase de consolidação da cotação.

- Sinalização de compra: quando o preço estiver abaixo da banda inferior ele tende a crescer;
- Sinalização de venda: quando o preço estiver acima da banda superior ele tende a cair.

Além disso, esta técnica pode validar sinais de outros indicadores como, por exemplo, o RSI ou o ADX.

### **3.3 Redes Neurais Artificiais**

A tecnologia das Redes Neurais Artificiais (RNAs) é uma tecnologia que tem origem em disciplinas, como a matemática, a estatística, a física, a engenharia, a neurociência e a ciência da computação. As áreas de suas aplicações são o reconhecimento de padrões, processamento de sinais e controle e análise de séries temporais que têm como virtude a capacidade de aprender a partir de um conjunto de dados de entrada. (HAYKIN, 2001) Consiste em um modo de abordar a solução de problemas simulando comportamentos inteligentes através dos circuitos cerebrais. Da mesma forma, estes circuitos neurais artificiais poderão se auto-organizar, quando apresentado a ambientes diversos, criando suas próprias representações internas. (BARRETO, 2002)

Segundo Cupertino (2009), uma RNA é um modelo matemático não-linear usado para encontrar relacionamentos complexos entre a entrada e saída de dados. Isto possibilita a capacidade de aprender e agir diante de situações adversas, bem como adquirir conhecimento por meio da experiência e observação.

Existem dificuldades inerentes em tentar imitar o sistema nervoso humano, enquanto este é formado por bilhões de neurônios e uma estrutura altamente complexa de conexões, o modelo matemático de uma RNA é formado por centenas, no máximo milhares, de elementos de processamento interconectados. Estes elementos são conhecidos como neurônios artificiais e são responsáveis por efetuar operações simples. As RNAs se diferenciam, principalmente, por três conceitos básicos: o modelo do neurônio artificial, sua estrutura de conexão e a regra de aprendizado.

Soluções baseadas em RNAs apresentam robustez e flexibilidade para representar modelos não-lineares, destacando uma vantagem muito importante para sua utilização no contexto de séries temporais (BALESTRASSI; SAMOHYL; MEURER, 1998). Para realizar

uma comparação entre soluções desta natureza e outras propostas Villanueva (2006), apresentou uma tabela comparativa com 22 pesquisas que utilizaram RNAs e alguma abordagem de outra categoria, por exemplo modelos baseados em regressão. Entre estas metodologias, as RNAs mostraram superioridade em 15 destes trabalhos, sendo inferiores em apenas 4. Entretanto também foi percebido que em alguns casos as séries temporais permitiram algum tipo de pré-processamento. Com isto, novas comparações foram realizadas utilizando este pré-processamento em todos os casos, e novamente as RNAs mantiveram superioridade.

As principais diferenças entre as várias RNAs existentes concentram-se em três conceitos básicos: neurônio artificial, arquitetura e processo de aprendizagem. Antes de conhecer estes conceitos será feito um breve histórico sobre redes neurais artificiais.

### 3.3.1 Histórico

Conforme Cupertino (2009), as primeiras pesquisas sobre redes neurais datam dos anos 40 em uma publicação nomeada por “A logical Calculus of the Ideas Immanent in Nervous Activity” que relatava uma analogia feita entre células nervosas e o processo eletrônico, realizada por Warren McCulloch e Walter Pitts, na Universidade de Illinois. Na mesma década, o biólogo e psicólogo Donald Hebb escreveu um livro chamado “The Organization of Behavior”, em 1949, que não possuía ideias pioneiras, mas descrevia o comportamento quantitativo das sinapses e o processo de treinamento humano através de um princípio de aprendizado em sistemas nervosos complexos.

O mesmo autor supracitado afirma que, durante a década seguinte ocorreram muitos acontecimentos nesta área de pesquisa pela empolgação de novas descobertas. O primeiro simulador de uma cadeia neural chamado SNARC foi construído em 1951 por Marvin Minsky, co-fundador do Laboratório de Inteligência Artificial do MIT. Apesar de nunca ter executado alguma aplicação interessante, o SNARC podia ajustar seus pesos sinápticos automaticamente e servia de inspiração para ideias nesta área. A primeira Conferência Internacional de Inteligência Artificial ocorreu em 1956, onde foi apresentado pelo pesquisador da IBM Nathaniel Rochester um modelo de rede neural artificial que consistia em uma simulação de centenas de neurônios interconectados para avaliar suas respostas a estímulos ambientais. Em 1959 foi criada uma rede de múltiplos neurônios com um modelo matemático de sinapses humanas baseada nas pesquisas de McCulloch. Seu

criador, Frank Rosenblatt, batizou sua rede como perceptron e pelas inúmeras contribuições técnicas e complexas pesquisas é considerado por muitos como fundador da neurocomputação.

Afirma ainda Cupertino (2009), que as pesquisas e o investimento financeiro sobre redes neurais reduziram um pouco com a publicação de uma obra chamada “Perceptron” por Minsky e Seymour Papert. Este documento publicado no fim da década de 50 contrariava o modelo proposto por Rosenblatt afirmando que sua abordagem não era promissora. Anos após foram prometidos computadores com poder de raciocínio igual ou superior ao cérebro humano, comprometendo a credibilidade de futuros estudos nesta área, devido a estas previsões exageradas.

Os anos 80 foram marcados pelo renascimento das pesquisas na área de neurociência, John Hopfield, físico e biólogo, ao se interessar pela área publicou diversos artigos em 1982, além de criticar as pesquisas de Minsky e Papert realizadas na década de 50. E por volta de 1986, o livro “Parallel Distributed Processing: Explorations in the Microstructure of Cognition” em seus dois volumes de autoria dos professores David Rumelhart e James McClelland, apresentaram um modelo matemático e computacional que proporciona o treinamento supervisionado dos neurônios artificiais. Com isto surgia a ideia do algoritmo backpropagation. (CUPERTINO, 2009)

A primeira Conferência de Redes Neurais aconteceu em 1987, quando foram formados os vários grupos desta área, como a Sociedade Internacional (International Neural Network Society – INNS) E após estes acontecimentos, muitas instituições formaram institutos de pesquisa e programas de educação em neurocomputação. (CUPERTINO, 2009)

### 3.3.2 Modelo de neurônio artificial

De acordo com Cupertino (2009) e Braga (2000), a unidade elementar de uma RNA é o neurônio artificial, responsável pelo processamento dos dados submetidos à rede, correspondente ao ilustrado pela Figura 8. Assim como um neurônio biológico, o artificial possui um conjunto de entradas (dendritos) e uma saída (axônio). As informações presentes na entrada são ponderadas pelos pesos sinápticos de forma semelhante às sinapses nervosas, a qual determina o quanto à entrada influenciou o neurônio. Isto gera um grau de estimulação do neurônio determinado pela função de ativação, que determina o cálculo da sua saída. As

entradas do neurônio são fatores importantes para definir o grau de excitação ou inibição do neurônio, sendo responsável por transmitir ou não uma informação para sua saída.

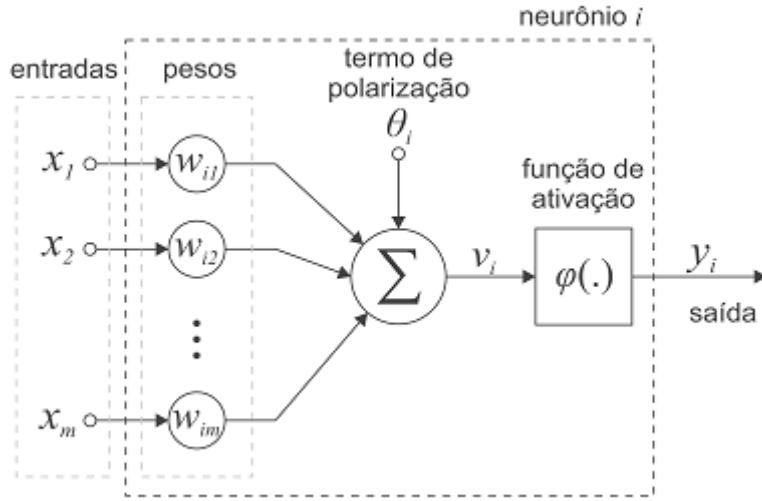


Figura 8 - Estrutura básica de um neurônio artificial.  
Fonte – Cupertino, 2009.

A representação matemática do funcionamento de um neurônio artificial é formalmente definida pela Equação 3.4.

$$Y_i = f\left(\sum_{i=0}^n w_i x_i\right) \quad (3.4)$$

Onde:

- $x_i$  é o valor da  $i$ -ésima entrada;
- $w_i$  é o valor correspondente ao peso da entrada  $x_i$ ;
- $n$  é o número de entradas e consequentemente pesos
- $f$  é a função de ativação utilizada;
- $Y_i$  é o estado de ativação ou saída do neurônio.

Como observado na Equação EEE, o resultado do somatório das  $n$  entradas  $x_i$  ponderadas com seus pesos respectivos  $w_i$  representa a ativação interna e será utilizado para calcular a função de ativação  $f$ . Já a função de ativação de um neurônio artificial define como ele responde aos estímulos de sua entrada pois calcula o resultado final do neurônio utilizando

a própria entrada. Dentre as funções de ativação conhecidas a mais usada é a sigmoidal. A função sigmoide, ou logística, é uma função contínua que tem a forma de um “S” e faz uma transição gradual entre seus dois extremos, variando entre 0 e 1 (PASQUOTTO, 2010). Como ilustrado de forma matemática pela Equação 3.5.

$$\varphi(v) = \frac{1}{(1 + e^{-v})} \quad (3.5)$$

Pasquotto (2010), para calcular a função sigmoide é necessário usar a constante de Napier ( $e=2,7182$ ). A função de ativação dos neurônios é responsável pela variação imposta à informação passada na saída. As informações recebidas após este processamento interno podem ser majoradas ou minoradas. É possível utilizar qualquer equação para representar a função de ativação do neurônio, exceto por alguns algoritmos de treinamento que exigem algumas características como a diferenciabilidade<sup>7</sup> para cálculos de gradiente, por exemplo.

As funções da família sigmoide definem um grupo de funções de ativação mais comumente utilizadas. O gráfico da função sigmoide possui a forma visual da letra “S”, com valores estritamente crescentes e apresenta uma estrutura com intervalos lineares e não lineares, o que possibilita o mapeamento de dados também com este comportamento. Seus valores de saída estão compreendidos no intervalo [0,1], mas quando é desejável valores negativos, por exemplo, no intervalo [-1,1], é utilizada a tangente hiperbólica que também pertence à família sigmoide. (BRAGA, 2000)

A ligação entre dois ou mais neurônios é feita através da conexão entre a saída do anterior com a entrada do posterior. Isto ocorre, pois o cálculo resultante de um neurônio serve como informação de entrada para outro, definindo um fluxo de informação até a saída final da rede. Estas conexões definem a topologia da rede que é considerado um detalhe importante na construção de RNAs. (BARRETO, 2002)

### 3.3.3 Arquitetura e aprendizado

Cupertino (2009), afirma que a organização geral de uma RNA é disposta através de uma estrutura em camadas, onde cada camada é formada por um conjunto de neurônios que possuem conexões com neurônios da camada seguinte, definindo uma comunicação entre

---

<sup>7</sup> Quando é possível calcular as derivadas parciais

camadas, como ilustrado pela Figura 9. A primeira camada da rede é chamada de camada de entrada, é canal de informação por onde os dados entram na rede, ou em outros termos, é o local por onde a rede recebe os estímulos externos. Em sua estrutura mais simples, a entrada da rede é seguida pela camada de saída que informa o resultado do processamento da entrada calculado ao longo das conexões existentes entre os nós da entrada e da saída, estas camadas delimitam o escopo de funcionamento da rede. Uma estrutura um pouco mais complexa envolve camadas intermediárias, também denominadas camadas ocultas, sendo que todos os nós que não pertencem as duas camadas limítrofes encontram-se nas camadas intermediárias. Os pesos que são definidos entre os nós armazenam o conhecimento adquirido pela rede e definem o nível atingido pela sua capacidade de aprendizado. (BRAGA, 2000)

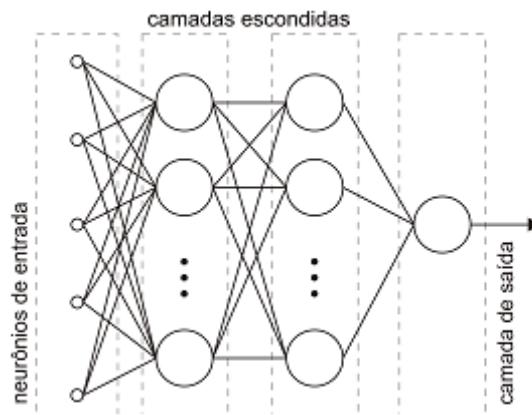


Figura 9 - Arquitetura geral de uma RNA.

Fonte – Cupertino, 2009

A topologia apresentada pela RNA pode possuir duas classificações, recorrentes ou não recorrentes. A Figura 9 ilustra uma topologia não recorrente, pois existem apenas conexões entre neurônios de camadas posteriores e não existem conexões entre neurônios da mesma camada. O padrão de conexões mais comum em RNAs é o feedforward quando é estabelecido um fluxo unidirecional no sentido entrada-saída sem a existência de realimentação como nas redes recorrentes.

Os resultados calculados pela RNA são influenciados por todos os elementos presentes em sua arquitetura, a quantidade de neurônios, número de camadas e os valores dos pesos sinápticos podem interferir no seu desempenho e precisão. Não existe na literatura uma configuração ideal, ou um meio de chegar a uma estrutura ideal para rede. Diante do conjunto finito de dados submetido à entrada existem infinitos mapeamentos que irão corresponder bem a esta entrada dada como amostra. Portanto, a arquitetura ideal para representar bem uma das possíveis soluções para o problema que a rede for aplicada deve ser a que apresentar o

melhor resultado. Acredita-se que o método de tentativa e erro é o mais adequado para definir uma arquitetura inicial da rede, diante dos parâmetros que precisam ser ajustados antes de sua utilização na prática, apenas aproximando do ideal para resolver o problema. (COSTA, 2008)

Segundo Cupertino (2009), o treinamento de uma RNA é realizado através de aplicação dos valores de entrada de maneira sequencial à medida que os pesos da rede são ajustados de acordo com as regras de treinamento adotadas. O objetivo do treinamento das RNAs é fazer com que um conjunto de entradas produza de maneira correspondente um grupo de valores de saída. Ao longo do treinamento os pesos vão gradualmente convergindo para valores que colaborem com a geração da saída desejada a partir da entrada oferecida. Isto ocorre de tal maneira que após uma entrada gerar uma saída correspondente e for submetida uma nova entrada semelhante sua saída deve ser também semelhante à primeira.

Barreto (2002), cita a necessidade de utilizar alvos para as saídas desejadas de cada entrada durante o treinamento o divide em duas categorias: supervisionado e não supervisionado. Ambos aproveitam de um conjunto de dados de treinamento, entretanto o primeiro requer também os valores com as respectivas saídas para as entradas, formando o chamado par de treinamento. Atualmente, existe uma grande variedade de algoritmos de treinamento de RNAs, para qualquer uma das classificações, sendo que o mais conhecido é o *backpropagation* (ou algoritmo por retropropagação). Que recebe este nome devido a suas duas fases de funcionamento, a propagação das informações de entrada no sentido normal da rede e a retropropagação para ajustes de peso no sentido contrário.

Para realização de um treinamento supervisionado é necessário definir o critério de parada, de modo geral existem duas formas de interromper o treinamento de uma RNA: função de erro e número de épocas. Uma função de erro, como o MSE ilustrado pela Equação 3.6, é utilizada para medir a diferença entre o que a rede fornece como saída e a saída pretendida. Enquanto a entrada não gerar a saída desejada, o valor do erro é utilizado para ajustar os pesos da rede colaborando para que isto ocorra. Este processo ocorre repetidas vezes até que seja atingido um valor pré-determinado para o erro. Uma segunda forma de parada para o treinamento ocorre com o alcance de outro valor definido a priori, chamado número de épocas que corresponde à apresentação completa do conjunto de treinamento à rede. Portanto, o treinamento de uma RNA irá encerrar assim que um limiar de erro for alcançado ou ultrapassar o número de épocas estabelecido. (BARRETO, 2002)

$$MSE = \frac{\sum_{n=1}^N [y(n) - \delta(n)]^2}{N} \quad (3.6)$$

O *Mean Squared Error* (MSE) é uma forma bem comum de medir o desempenho da rede, que por sua vez avalia o aprendizado da rede diante dos valores submetidos durante seu treinamento comparados com o que foi calculado. O MSE é um estimador de adequação ao modelo, tendendo a zero quando a rede atinge o maior valor de adequação possível. Para casos em que o MSE apresenta valores inferiores a 0.01, é considerado que a rede atingiu um patamar ideal para aplicação de predição de séries temporais em casos reais (SOUZA, 2008). A Figura 10 ilustra de maneira geral como funciona um treinamento supervisionado utilizando uma medida de erro como critério de parada.

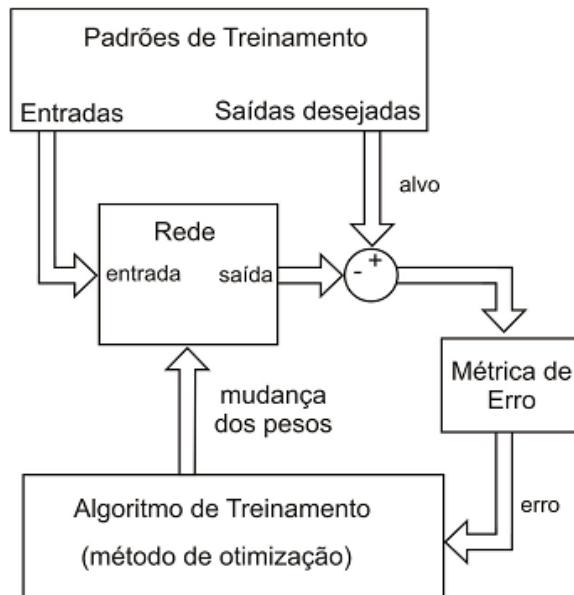


Figura 10 - Passos do algoritmo de treinamento supervisionado.  
Fonte – Cupertino, 2009

Quando se aborda a questão de treinamento e aprendizado em RNAs, vários fatores precisam ser levados em consideração. Durante o projeto arquitetural de uma RNA é possível que ocorram alguns problemas que merecem destaque. É importante elaborar cuidadosamente a arquitetura da rede para evitar um problema chamado de **erro de aproximação**. Segundo Villanueva (2006), isto ocorre quando a rede não oferece a flexibilidade necessária para modelar o problema, se tratando de dependências estruturais. Dependendo do tipo de problema a ser aproximado é notória a necessidade diferenciada do número de neurônios em qualquer camada, sobretudo na camada oculta que irá diferenciar as soluções projetadas.

Os valores utilizados no treinamento da rede possuem papel fundamental para o sucesso na fase de treinamento. Cada função de ativação utilizada na RNA possui uma faixa de valores de acordo com suas equações, por exemplo a sigmoide varia no intervalo [0,1]. Deste modo, os dados submetidos à rede durante o treinamento precisam estar nesta faixa para que a rede consiga generalizar o problema de maneira adequada.

A falta de utilização de critérios de parada para o treinamento é outro acontecimento que pode ocasionar em problemas muito sérios para utilização posterior da rede. É provável que durante o treinamento chegue um momento em que a rede aparenta estar representando bem o problema devido a redução da taxa de erro, mas sua capacidade de generalização começa a piorar. A generalização é a capacidade da rede de responder adequadamente a novos dados fora do conjunto de treinamento, quando este valor diminui é considerado que a rede se especializou no conjunto de testes ou ocorreu um sobreajuste. De modo contrário se o treinamento for interrompido antes da rede atingir um nível satisfatório de aproximação ao problema é dito que ocorreu um sub-ajuste. Quando ocorre algum destes problemas ocorre ao longo do treinamento é considerado o acontecimento de um **erro de estimativa**.

Outro fator gerador de problemas durante o treinamento da rede e a taxa de aprendizado utilizada. Esta taxa presente nos algoritmos de treinamento informa em qual proporção os pesos serão alterados, com valores no intervalo [0,1]. Para uma baixa taxa de aprendizado é capaz da RNA levar um tempo muito alto para finalizar o treinamento, já valores altos podem incapacitar o treinamento de atingir o nível de aprendizado desejado. (VILLANUEVA, 2006)

Kohavi (1995), afirma que para evitar alguns problemas citados existe uma etapa fundamental no treinamento de RNAs chamada de validação dos resultados solucionando problemas de estimativa, e uma técnica muito comum utilizada nesta fase é conhecida por validação cruzada. Nesta técnica o conjunto de dados total que são submetidos à rede é dividido em três grupos: treinamento, validação e teste. É adotado que todos os grupos representam bem o problema abordado. Nesta técnica, o conjunto de treinamento é responsável por ajustar os pesos da rede, e a avaliação do erro é realizada através do segundo conjunto de dados (validação). Quando o erro calculado neste segundo grupo apresentar algum crescimento o treinamento é encerrado antes que ocorra o *overfitting*. O último grupo de dados é utilizado para simular o comportamento da rede em um ambiente real, espera-se que um bom desempenho com a validação ocasione o mesmo com o teste.

A arquitetura mais comum de RNAs é chamada de *Multi-Layer Perceptron* ou simplesmente MLP. Sua estrutura é definida por uma rede multicamada formada por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. E as conexões existentes entre os neurônios são do tipo *feedforward* (PASQUOTTO, 2008). Esta rede tem sido muito utilizada para resolver problemas complexos de várias naturezas através de um treinamento supervisionado guiado pelo algoritmo de treinamento *backpropagation*.

Conforme Werbos (1974), o algoritmo de treinamento *backpropagation* é dito como mais utilizado, tanto pelo histórico de aplicação quanto pelas diversas abordagens estendidas de sua abordagem. Seu funcionamento é fundamentado sobre dois passos, o primeiro no sentido direto da rede e o segundo no sentido reverso. No passo direto a saída é gerada através dos dados de entrada e caso não compreenda a saída desejada será calculado um erro. Nesta etapa os pesos serão ajustados camada a camada no sentido reverso, de acordo com o erro calculado na etapa anterior. O fim do treinamento ocorre dependendo do critério de parada escolhido.

Conforme Pasquotto (2010), algoritmos semelhantes ao *backpropagation* utilizam como parâmetro para atualização dos pesos o sentido oposto ao do gradiente da função de erro, que por sua vez é uma generalização da regra delta que define o ajuste nos pesos proporcionalmente ao erro calculado. Devido ao cálculo do gradiente, os pesos irão mudar no sentido da descida mais acentuada da superfície multidimensional da função de erro, caracterizando uma redução gradual do erro. Esta heurística, chamada de gradiente descendente, é ilustrada pela Figura 11 e pode ser explicada a partir da seguinte analogia: uma pessoa encontra-se em um terreno montanhoso e possui o objetivo de ir ao ponto mais baixo deste terreno. Neste contexto, o terreno simboliza a função de erro com picos e vales, a busca pelo ponto de menor elevação movendo-se na direção de maior inclinação de descida representa o gradiente descendente.

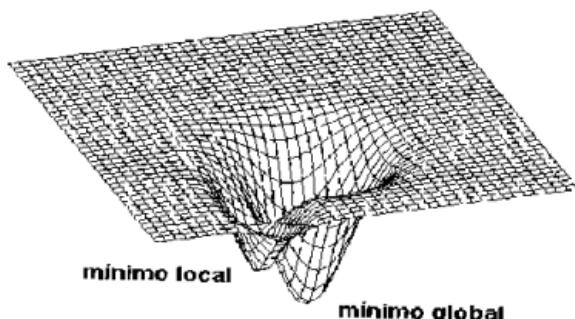


Figura 11- Superfície multidimensional da função de erro.  
Fonte - Pasquotto (2010).

Afirma Pasquotto (2010), que um problema clássico, e muito comentado pela literatura, associado ao *backpropagation* ocorre devido à utilização do gradiente descendente para cálculo da função de erro, como comentado por Existe um risco nesta abordagem de encontrar um vale que possua um valor mínimo comparado aos vales ao seu redor, entretanto pode não ser um valor mínimo global para toda a superfície de erro ou absoluto para função, caracterizando a ocorrência do problema chamado de mínimo local. Na prática, as superfícies de erro que serão formadas pelo gradiente descendente são extremamente convolutas e não convexas, podendo gerar soluções “não ótimas” através do encontro de mínimos locais.

## **Propagação Resiliente**

Como a construção de uma rede neural precisa ser nivelada ao longo de vários testes é fundamental que mais de um algoritmo de treinamento seja utilizado durante as simulações de configuração para definir o que possua a melhor precisão. Existem vários algoritmos de treinamento para RNAs *feedforward*, como o tradicional *backpropagation*, o *Resilient Propagation*, o *Levenberg Marquardt* e muitos outros, sendo que boa parte destes algoritmos derivam do *backpropagation* ou da ideia de retropropagação, destacando o *backpropagation* uma grande inspiração para treinamento de RNAs.

Soares (2008), evidencia-se o *Resilient Propagation* (RPROP), que foi criado como uma alternativa mais eficiente que o *backpropagation* para o treinamento de redes do tipo MLP, evitando problemas relacionados aos mínimos locais, sendo sugerido por alguns autores como uma ótima alternativa para predição de séries temporais. O algoritmo de treinamento RPROP é uma melhoria sobre a técnica de aprendizagem por retropropagação de erro. O RPROP foi proposto por Riedmiller e Braun (1993), para treinamento de RNAs do tipo MLP com uma alteração na atualização dos pesos sinápticos que são calculados baseados em informações de gradientes locais, isto requer que cada peso possua seu valor próprio de atualização para determinar o tamanho da alteração que o peso irá receber.

Assim, Souza (2008), acrescenta que o valor de do peso evolui durante todo o processo de aprendizado, baseando-se no sinal local da função de erro, seguindo as seguintes regras ilustradas abaixo na Equação 3.7.

$$\Delta_{ij}(n+1) = \begin{cases} n^+ \Delta_{ij}(n), \text{ se } \frac{\partial E}{\partial w_{ij}} n \frac{\partial E}{\partial_{ij}}(n+1) > 0 \\ n^- \Delta_{ij}(n), \text{ se } \frac{\partial E}{\partial w_{ij}} n \frac{\partial E}{\partial_{ij}}(n+1) < 0 \\ \Delta_{ij}(n), \text{ caso contrário} \end{cases} \quad (3.7)$$

Um detalhe crucial sobre o funcionamento do RPROP está relacionado ao fator  $n$ , que pode ser positivo ou negativo. Isto implica diretamente no cálculo estabelecido sobre os pesos sinápticos, pois diferente do *backpropagation* que depende exclusivamente da magnitude da variação da função de erro, o RPROP depende dos termos deste gradiente, com a atualização dos pesos dependendo apenas da variação dos sinal das derivadas parciais apresentada na Equação 3.7. Como pode ser observado na mesma equação o valor do ajuste depende apenas do sinal da derivada parcial, quando o sinal altera significa que a taxa de ajuste possivelmente foi exagerada e o fator precisa ser decrementado pois existe uma possível ocorrência de um mínimo local da função de erro, precisando reduzir o valores individuais de atualização  $\Delta_{ij}$  proporcional ao fator  $n^-$ . Para casos de manutenção do sinal o fator de atualização é aumentado, pela conclusão que o peso relacionado ocorreu na direção que minimiza a função de erro E, mantendo a convergência a um mínimo global através do aumento do fator  $n^+$  implicando no aumento dos valores individuais  $\Delta_{ij}$ . (PASQUOTTO, 2010)

### 3.3.4 Predição utilizando redes neurais

Braga (2000) destaca que a abordagem de RNAs é muito atraente no contexto de predição de séries temporais por ser um modelo não-paramétrica. Isto é, não há necessidade de conhecer detalhes do problema para que seja utilizada, basta realizar um mapeamento dos valores da série entre a entrada e saída da rede para que esta seja capaz de estimar a predição de novos valores. Braga acrescenta que diante de abordagens tradicionais, as RNAs possuem outro fator favorável relacionado a não haver limitação para o número de variáveis de entrada, diferente de abordagens clássicas como modelos auto-regressivos. Souza (2008) afirma que as RNAs destacam-se como modelos adequados para representar o comportamento não-linear apresentado pelas séries temporais, e realizar predição sobre seus valores futuros. Isto torna as RNAs modelos adequados para reproduzir fenômenos com comportamento não-linear tal

como séries temporais, conforme comentado por Souza (2008). Ou seja, não é necessário conhecer previamente detalhes sobre o problema, apenas utilizar dados passados para predizer os futuros. Portanto, ao utilizar redes neurais não é necessário ter previamente informações específicas sobre a série, como a frequência de valores, ocorrência de ciclos ou sazonalidade, mas apenas os dados que a compõem utilizados durante seu treinamento.

Destacam-se duas formas de aplicar RNAs para realizar predição de séries temporais de cotações: preços futuros de ações ou a tendência representada pela oscilação da série. É desejável saber precisamente qual o valor que a ação irá apresentar em um período futuro, pois trata-se de uma forma de investimento extremamente rentável e igualmente difícil de se prever devido o comportamento aleatório que se encontra o mercado acionário. Por outro lado, estimar a direção que movimento do mercado irá apresentar fornece informações importantes, sabendo a tendência de preços é possível realizar investimentos lucrativos. Diante destas duas formas de predição discutidas muitos autores propuseram soluções utilizando RNAs como foco principal para predizer o mercado acionário, como apresentado por Sousa (2008), Cupertino (2009), Pasquotto (2010). Apesar das vantagens e desvantagens apresentadas por estes trabalhos, é importante perceber a importância que foi dada a esta abordagem de predição do mercado acionário, destacando as RNAs como técnica muito empregada para objetivo.

Embora ao usar RNAs não seja necessário possuir conhecimentos específicos sobre o problemas, pois sua flexibilidade se adequa a complexidade solução, é necessário projetar sua arquitetura para que a solução seja viabilizada, que não é uma tarefa nada fácil. Isto ocorre, normalmente, através do processo de tentativa e erro, ajustando os parâmetros da rede na direção que a mesma forneça melhores valores de saída. Mas no tocante a predição, é fundamental que dois elementos de sua arquitetura sejam especificamente definidos: a entrada da rede e a sua saída. Para cada um dos elementos é necessário definir o número de neurônios, padrão de conexão e função de ativação. Como são duas camadas diferentes é possível que todos estes valores sejam diferentes entre si.

A entrada da rede pode ser estabelecida apenas com os preços de fechamento presentes na cotação, por exemplo: a partir da definição de um atraso de dez preços correspondente aos últimos dez valores da série temporal é formada a entrada de uma RNA. Igualmente como a utilização única de preços de fechamento, também pode ser definida através dos quatro valores (abertura, fechamento, máximo e mínimo) da cotação formando grupos de *candlesticks* e neste caso o número de neurônios teria que ser quatro vezes o número de padrões de *candlesticks* para suportar seus valores. Saindo do âmbito de preços

puramente extraídos da cotação, outras informações como resultados do cálculo de indicadores técnicos da mesma forma podem ser utilizados como entrada da rede. É evidente o número de possibilidades que podem definir a entrada da rede, além das combinações entre estas sugestões discutidas. Mas, para efeito de satisfazer as duas formas comuns de predição da curva de cotações citadas anteriormente, os valores de fechamento são suficientes, tanto para predição de preços quanto de tendência.

Igualmente desafiador como projetar a entrada da rede encontra-se sua saída, que possui uma estrutura simples e representa a predição propriamente dita. O número de neurônios nesta camada constituem o grau de predição que pode ser alcançado, quanto mais neurônios maior será o alcance da predição. Associando esta camada aos conceitos de predição, a saída corresponde à janela de predição. A saída, normalmente, é projetada de acordo com sua entrada, pois representa a conclusão do processamento de valores que foram adicionados na rede pela camada de entrada. Utilizando a mesma analogia feita com a entrada para definição de uma camada simplificada para predição, para estabelecer uma saída com apenas um neurônio representando o preço do dia seguinte satisfaz as duas formas comuns de predição de preço e tendência.

### **3.4 Comitê de Aprendizado**

Alternativas para combinação de métodos de predição vêm sendo pesquisadas. Há a abordagem de combinação pela criação de comitês de máquinas de aprendizado. Comitê de máquinas de aprendizado é uma abordagem que trata da fusão do conhecimento adquirido por especialistas (métodos de previsão ou simplesmente previsores) para atingir uma decisão global que visa ser superior àquela alcançável por qualquer um deles atuando isoladamente (HAYKIN, 2001).

Comitê de máquina, ou comitê de aprendizado, é uma forma de aprendizado de máquina<sup>8</sup> que busca superar o desempenho obtido por uma abordagem que utiliza uma máquina de aprendizado operando isoladamente, sua ideia é baseada no princípio “dividir-para-conquistar” combinando diferentes abordagens como componentes em uma única solução (LIMA, 2004). E como esta abordagem apresenta uma solução composta é de indispensável que sua proposta de utilização envolva alguma comparação experimental com

---

<sup>8</sup> Segmento da Inteligência Artificial dedicado ao desenvolvimento de algoritmos e técnicas que permitam o computador aprender ou aperfeiçoar seu desempenho em alguma tarefa

as soluções individuais que a compõe, mostrando as vantagens e desvantagens de cada uma e fundamentando a escolha.

Historicamente a ideia de comitê de máquina surgiu por volta da década de 90 quando os pesquisadores estavam em busca de conceitos que superassem as dificuldades enfrentadas por abordagens neurais após um longo tempo de estudo sobre RNAs que ocorria desde o final dos anos 50. Neste contexto os comitês tentam agregar o conhecimento adquirido pelos seus componentes de modo que sua solução combinada apresente desempenho superior à obtida por qualquer um dos seus componentes atuando de maneira isolada. (VILLANUEVA, 2006)

Algumas características inerentes a séries temporais, financeiras ou não, como quantidade finita de amostras para treinamento, ruído e até *outliers*<sup>9</sup> dificultam a construção de um modelo que atinja a máxima capacidade de generalização. Realizando uma comparação direta com RNAs que são consideradas um dos melhores previsores de séries temporais, os comitês apresentam certa superioridade diante alguns problemas inerentes a estas redes, como: a convergência do algoritmo de treinamento para um mínimo local, riscos associados à sobreajustes durante treinamentos e a própria complexidade do problema a ser solucionado, que pode ser simplificada ou dividida entre os elementos do comitê. O sucesso de redes neurais é inteiramente dependente da superação de erros de estimativa e de aproximação, justificando a busca por alternativas que contornem esta limitação. Ao comparar comitês com abordagens individuais como RNAs, apontado por Villanueva (2006), é constatado uma melhor capacidade de generalização e tolerância a ruídos nos dados. Além disto, cada componente do comitê não precisa apresentar flexibilidade compatível com a complexidade total do problema, uma vez que não são considerados isoladamente, mas sim em comitê.

A elaboração e construção da arquitetura de um comitê depende de como o problema será solucionado pelos seus componentes, se parcialmente por cada componente ou de maneira independente. Contudo, esta definição gera uma classificação particular para os comitês, quando um comitê delibera a responsabilidade de solucionar um problema de maneira integral para cada um de seus componentes este recebe a classificação de estático, de maneira oposta o comitê dinâmico possui componentes que resolvem parte da tarefa principal e a solução de um único componente não comprehende a solução para o problema. O comitê estático é também conhecido como *ensemble* (palavra francesa que significa junto) e sua arquitetura geral pode ser observada na Figura 12.

---

<sup>9</sup> É considerada uma observação que possui um valor julgado discrepante por estar bem longe dos demais valores da série, em estatística os *outliers* implicam em prejuízos na interpretação dos resultados.

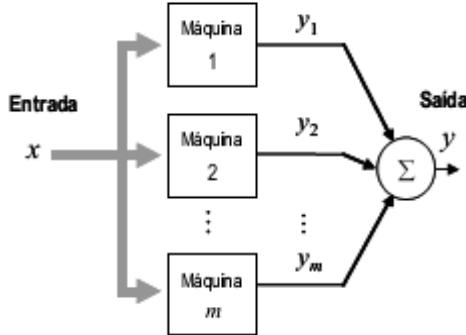


Figura 12 - Funcionamento básico da composição

Como observado na arquitetura de um *ensemble*, vários componentes irão trabalhar em paralelo para resolver uma tarefa e apresentar como resultado uma solução candidata para o problema a partir de uma mesma entrada de dado que é fornecida igualmente para todos os componentes. Por mais que as soluções individuais forneçam um resultado completo para o problema, não são necessariamente isoladas umas das outras. Um requisito para o sucesso na utilização de *ensemble* está relacionado a como os componentes irão generalizar suas soluções para o problema, o que não faz sentido adotar os mesmos procedimentos e hipóteses para modelar cada previsor, pois é objetivado utilizar componentes diferentes. Sendo assim, é importante que o processo de construção dos componentes seja independente um do outro, utilize abordagens e treinamentos diferentes, pois é de grande importância a combinação de componentes que possuam erros descorrelacionados através da diversidade de sua elaboração.

Existe uma grande quantidade de fatores retratados por Villanueva (2006), que motivam a utilização de ensemble, como fatores estatísticos envolvendo solução e fracasso. Para uma abordagem que envolve combinação falhe é necessário que mais da metade de seus componentes também falhe. Do ponto de vista do volume de dados, soluções individuais podem não modelar de maneira total um problema com grandes volumes gerando um erro de estimativa que pode ser contornado a partir do particionamento dos dados entre componentes de um comitê, ou não generalizar a solução quando este volume for relativamente pequeno também contornado através do uso de vários componentes. A complexidade do problema pode acarretar dois grandes contratemplos como erros de aproximação ou custos computacionais altíssimos quando tratados por componentes individuais.

### 3.4.1 Histórico

Um breve histórico sobre ensembles inicia em 1990 com Hansen e Salamon ao demonstrarem que o desempenho de generalização de uma RNA pode ser melhorado utilizando um *ensemble* de RNAs configuradas com o mesmo número de neurônios. No mesmo ano, Schapire provou que um classificador forte no sentido de aprendizado provavelmente aproximadamente (do inglês *probably approximately correct*) correto pode ser gerado pela combinação de classificadores fracos através do algoritmo de *boosting*.

Woods et al., (1997), afirmam que estes trabalhos iniciais incentivaram a comunidade que expandiu esta nova abordagem através de inúmeras propostas envolvendo comitês. Corroborando com os autores acima, Xu et al (1992), acrescentam que na década de 90, muitos trabalhos foram publicados como combinação de vários classificadores e seleção dinâmica de classificadores. Na década seguinte foi proposto por Kuncheva (2001), o trabalho “fusão de classificação” e também pelo mesmo autor em 2004 o “ensemble de classificadores”.

A principal diferença entre estas abordagens está no processo utilizado para gerar os componentes individuais ou na estratégia empregada na combinação ou seleção dos componentes.

### 3.4.2 Previsores

Conforme Villanueva (2006) e Lima (2004), dentro do contexto da predição os componentes que formam o comitê também são conhecidos como previsores. Sendo estes independentes entre si, haja vista que cada um recebe os dados passados pelo comitê de forma individual e após isto calculam separadamente suas saídas correspondentes e a retornam para o comitê que efetuar a combinação com os outros previsores e gera uma resposta final unificada. Como cada previsor funciona de maneira autônoma do ponto de vista da interferência do comitê, os previsores podem ser completamente diferentes entre si, tanto sua estrutura quanto seu comportamento, sendo desejável que exista estas diferenças nos previsores proporcionando chances de maximizar a capacidade de generalização do comitê.

Os mesmos autores supracitado afirmam que do ponto de vista estrutural, cada previsor se encaixa no comitê como componentes de software são projetados. O comitê oferece uma interface de entrada de dados que o previsor irá receber e informa a saída necessária que o previsor precisa gerar. Como esta entrada irá gerar uma saída é responsabilidade do previsor executar. Diante disto, é preciso que cada previsor adapte ou transforme os valores da entrada que o comitê irá informar de acordo com sua necessidade.

Outro aspecto importante sobre a independência entre os previsores está relacionado ao ciclo de vida ou tempo de duração para cada processamento individual encerrar, que eventualmente serão desiguais por causa das diferentes abordagens utilizadas em cada previsor. Este acontecimento não pode prejudicar o funcionamento do comitê que precisa estar preparado para aguardar ou não todos os previsores concluírem seus cálculos para iniciar o processo de combinação dos resultados. Como os previsores de um ensemble produzem soluções completas, não precisam iniciar e encerrar juntos para que o comitê atinja seu resultado final, dependendo da estratégia adotada pela combinação mesmo que um ou mais previsores falhem (não alcancem a solução desejada) é possível que comitê seja capaz de gerar a solução final a partir da combinação de parte dos previsores. (VILLANUEVA, 2006; LIMA, 2004)

### 3.4.3 Arquitetura

A arquitetura de um comitê, assim como em RNAs, envolve tanto a parte estrutural como estratégica responsável pelo seu funcionamento, como: a construção dos previsores envolvidos neste processo, além de aspectos relacionados com a padronização de dados de entrada e métodos de combinação dos resultados .

Para estabelecer uma arquitetura de um comitê é necessário seguir algumas fases bem estruturadas como um projeto, como sugeridas por Villanueva (2006). A etapa inicial compreende a geração dos componentes, e neste ponto é fundamental a preocupação com os erros de treinamento e ausência de relação entre os perfis de erros dos previsores. Em casos de *ensembles* homogêneos, os que utilizam previsores da mesma natureza como redes neurais, é recomendado que o processo de construção inserisse alguma forma diversidade através de diferentes arquiteturas, algoritmos de aprendizagem ou inicialização aleatória de pesos. *Ensembles* heterogêneos promovem uma diversificação maior por apresentar abordagens

utilizando técnicas com pequenas diferenças ou totalmente diferentes. Outra forma de possibilitar a diversidade nesta solução é realizada através dos dados que irão formar o conjunto de treinamento dos previsores que podem ou não ser os mesmos para cada um, existe uma técnica de geração de conjunto de treinamentos distintos a partir de uma mesma fonte muito utilizada chamada de *bagging*. Por outro lado, a escolha deste grupo particionado de treinamento pode interferir no desempenho dos preditores, outra ferramenta chamada *boosting* é capaz de transformar previsores fracos em fortes através da reorganização do conjunto de treinamentos.

Na segunda etapa de construção da arquitetura será efetuada a seleção dos componentes que irão ser utilizados para construir a solução final. Isto é utilizado quando ocorre a necessidade de um refinamento no ensemble com o objetivo de melhorar seu desempenho global de generalização, pois é possível que certos previsores possuam taxas de generalização bem inferiores à da maioria, degradando a generalização do comitê em detrimento de pequenos problemas individuais. Este processo ocorre quando os componentes do comitê são gerados no processo anterior e a solução para o problema exige uma taxa de generalização superior a apresentada por alguns integrantes do comitê, ocorrendo uma exclusão dos resultados submetidos à próxima etapa, de composição da solução final.

A última etapa da arquitetura apresentada é chamada de composição, ocorrendo a partir deste ponto a junção das soluções individuais para construção de uma informação única que representa o resultado final do comitê. A forma como a composição ocorre no comitê depende exclusivamente de como o problema foi modelado e qual a solução desejada. Dentro do contexto da predição de valores, como é o caso de séries temporais, é possível utilizar composições de várias maneiras, desde as mais simples médias, até suas derivações com uso de ajustes em coeficientes de ponderação ou até utilizar complexas equações não lineares. (VILLANUEVA, 2006)

### 3.4.4 Métodos de Combinação

A estratégia mais simples de combinação dentro dos comitês de aprendizado ocorre por meio da média aritmética entre os resultados dos previsores, o que permite uma estabilização em um valor médio de seus resultados. Uma pequena melhoria nesta abordagem, discutida por Ribeiro (2009) é oferecida basicamente pela atribuição de pesos nas saídas dos

previsores, gerando um somatório ponderado para cálculo da saída. A representação matemática desta estratégia de combinação é ilustrada na Equação 3.8 a seguir:

$$f_c = \sum_{i=1}^n w_i f_i \quad (3.8)$$

Esta equação define a uma abordagem de composição, onde  $f_c$  é a saída combinada de  $n$  métodos envolvidos no comitê, que por sua vez possui um grupo de previsores com saídas individuais  $f_i$  associadas aos pesos  $w_i$ . O uso de pesos relacionados aos resultados dos previsores elenca por meio de um nível de importância cada um, em termos percentuais a soma de todos os pesos deve totalizar 100% (RIBEIRO, GOLDSHIMIDT; CHOREN, 2009). Estas duas variações citadas apresentam características lineares e também são chamadas de combinação linear.

Conforme Villanueva (2006) e Lima (2004), além desta forma de combinação envolvendo equações é possível realizar a combinação de várias maneiras. Uma forma bem simples de efetuar a combinação é realizada através do método “Votação”, onde a solução final é apoiada pela maioria dos previsores, comumente utilizado em problemas de classificação esta combinação resulta na classe majoritária classificada pelos previsores. Apesar de não requerer altos custos computacionais esta abordagem pode reduzir o desempenho do ensemble caso os previsores possuam desempenho global ruim. Uma variação mais rigorosa da votação também em abordagens de classificação é chamada de “Unanimidade”, a qual exige uma concordância total entre os resultados de todos os previsores para que o resultado desta combinação resulte nesta classe unâime. Já uma abordagem mais moderada ocorre com a “Pluralidade limitada” a combinação resulta em uma classe quando o número de componentes que a classificaram é maior que o número de componentes das outras classes. É possível também que um dos previsores aprenda a combinar a saída dos outros, onde este previsor torna-se o método de combinação na chamada “Generalização Empilhada”.

### 3.4.5 Predição usando ensembles

Para utilizar um ensemble como previsor de séries temporais é necessário projetar seus componentes para este fim e definir a abordagem de combinação que será utilizada. As

etapas de geração e seleção dos previsores pode ser omitida, mas seus erros podem reduzir severamente o desempenho do comitê. Entretanto, caso exista garantia que os previsores apresentem diversificação, devido a suas estruturas ou processo de construção, e seus erros individuais estão dentro de um patamar aceitável é possível utilizar o comitê para predição sem estas duas etapas.

A utilização mais comum apresentada na literatura de comitês utiliza RNAs como previsores do comitê e através das ferramentas disponíveis em abordagens de combinação propõe melhorias no desempenho da RNA, quando esta realiza a mesma tarefa que faria de maneira isolada. Um trabalho realizado desta maneira foi feito por Kim et al (2010) com seu *ensemble* homogêneo composto de previsores tipo RNA para prever a falência de empresas coreanas baseado em alguns índices disponíveis de relatórios financeiros destas empresas. Por outro lado, existem os *ensembles* heterogêneos que sugerem previsores de abordagens diferentes, não apenas arquiteturas diferentes de RNA. O trabalho de Chintra et al (2010), usou um *ensemble* heterogêneo como classificador de predição de séries temporais, seus três previsores eram de abordagens totalmente diferentes: RNA com função de ativação RBF, *K-Near Neighbor* (KNN ou técnica do vizinho mais próximo)<sup>10</sup> e *Self Organization Map*<sup>11</sup> (SOM ou mapas auto-organizáveis).

Os resultados obtidos pela predição de um único previsor na forma de uma RNA MLP foram comparados a várias abordagens ensemble por Villanueva (2006, p. 68), com o objetivo de avaliar qual a melhor abordagem para predição. Um total de 30 simulações foram realizadas com 5 séries temporais. A RNA foi configurada com 30 neurônios ocultos e 1000 épocas de treinamento com validação cruzada. Os *ensembles* possuíam 50 previsores do tipo RNA MLP com os mesmos parâmetros de configuração do previsor individual, diferenciando apenas pelas estratégias de construção, seleção e composição presentes apenas no *ensemble*. Após as simulações foi constatado que em todos os casos o *ensemble* apresentou melhores resultados que a única MLP. Vale ressaltar que a etapa de geração não gerou em todos os casos previsores adequados para utilização e o que manteve a solução com bons resultados foi a etapa de seleção dos previsores.

Ainda no trabalho de Villanueva (2006), é apresentado graficamente como se comportou a abordagem de comitê diante de soluções individuais para predição de uma série temporal, como ilustrado na Figura 13. É possível perceber que os previsores individuais se

---

<sup>10</sup> KNN é uma técnica que baseia-se na ideia que as propriedades de um ponto são semelhantes a dos pontos de sua vizinhança

<sup>11</sup> SOM, também conhecida como Rede de Kohonen, é uma rede neural de duas camadas com a primeira estruturada em forma de matriz e a segunda completamente conectada à primeira

aproximaram bem da solução ideal para o problema, mas a combinação realizada pelo *ensemble* melhorou os resultados da predição atingindo uma generalização ainda maior.

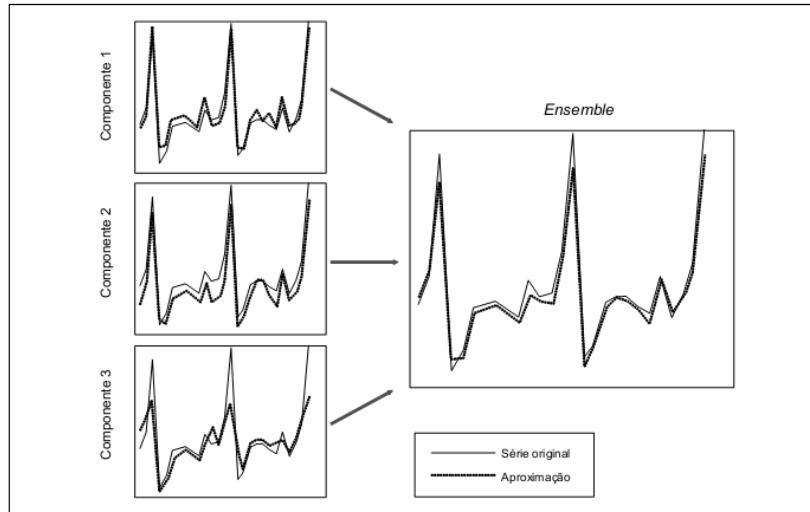


Figura 13 - Resultado da utilização de ensemble comparada a previsões isoladas

### 3.5 Considerações Finais

Existem diversas maneiras de antecipar o comportamento de séries temporais financeiras. De maneira geral, a utilização de indicadores técnicos pode ocorrer isoladamente com o objetivo de informar sinais de compra e venda, mas geralmente são combinados entre si no intuito de alcançar melhores resultados, também porque em alguns casos seus resultados não geram conclusões seguras sobre como operar. ou em algumas circunstâncias o investidor exige uma garantia maior de confirmação para sua aplicação financeira, sendo necessário um indicador complementar. A aplicação de indicadores técnicos como instrumento de análise é bem flexível e o utilizador desta ferramenta possui ao seu dispor a organização personalizada de sua estratégia de investimento através de ajustes nas configurações dos indicadores.

Já a utilização de redes neurais neste contexto de predição oferece grandes benefícios relacionados à aproximação de funções não-lineares, e vantagens como independência de conhecimento prévio sobre modelo. Além disto, já existe um grande período de estudos nesta área de neurocomputação e uma vasta quantidade de trabalhos propostos com abordagens das mais diversas. Entretanto, principalmente se tratando do mercado financeiro que é um ambiente considerado caótico por depender de inúmeros fatores intangíveis, é necessário que a RNA apresente desempenho e precisão bem aperfeiçoados, não

apenas para oferecer confiabilidade em seus resultados e utilização, mas por se tratar de uma aplicação direcionada para investimentos, que envolvem lucros e prejuízos.

O ponto de partida da utilização de RNAs carece que uma metodologia formal, que a literatura não oferece mesmo após tantos anos de estudo, tornando o início de seu projeto um passo certamente obscuro. Isto ocorre tanto pela perspectiva de projeto e estruturação da arquitetura da rede, quanto pela definição do processo de aprendizagem que envolve o algoritmo de treinamento e seus parâmetros. Apesar de alguns autores sugerirem alguns algoritmos de treinamento para cada contexto de utilização das RNAs, não existe um paralelo formal entre este contexto e o algoritmo correspondente, implicando em uma grande quantidade de testes e simulações para que seja alcançada a configuração ideal necessária para atingir com êxito a predição.

Por último, a utilização de RNAs como previsor de séries temporais, particularmente no mercado acionário, apresenta bons resultados diante de várias abordagens propostas na literatura, destacando-a como uma alternativa viável para antecipar o comportamento do mercado acionário. Os resultados apresentados por redes aplicadas neste contexto permitem uma boa margem de interpretações que podem direcionar estratégias de investimento bem rentáveis. Uma forma de acrescentar às contribuições já oferecidas pela RNA pode ser feita através da combinação de sua utilização com outra metodologia com o mesmo objetivo de predição, por exemplo indicadores técnicos usados para aplicações financeiras.

A outra abordagem apresentada foram os comitês, com um destaque aos *ensembles* como previsor de séries temporais por apresentar perspectivas promissoras devido às limitações superadas de tecnologias como RNAs e a expectativa de atingir resultados melhores que abordagens individuais ou até seus próprios previsores executados isoladamente. O uso de *ensembles* homogêneos já possui comprovação na literatura que atinge bons resultados, principalmente na otimização de RNAs.

No caso do mercado acionário, existem diversos métodos para definir estratégias financeiras para investimento em ações como o reconhecimento de padrões, indicadores técnicos e índices fundamentalistas, preços de commodities e fatores intangíveis como políticas internacionais. Portanto, um *ensemble* heterogêneo utilizado neste contexto poderia combinar algumas destas abordagens em uma solução única, propondo melhores resultados que suas utilizações individuais, como por exemplo utilizar como previsor do mercado acionário informações técnicas e fundamentalistas, combinando as duas escolas que dividem opiniões de muitos investidores deste mercado.

É evidente que a utilização de comitês, ou propriamente de ensembles, não irá produzir necessariamente taxas de acerto excessivas. Acredita-se que a avaliação de uma abordagem de combinação apresente valores no mínimo iguais ao atingido pelo melhor de seus previsores, mas é desejado pelo menos um ganho de desempenho frente a este previsor.

Por fim, a utilização de *ensembles* como previsor de séries temporais apresenta um conjunto de benefícios interessantes como modelagem de problemas complexos através da divisão do problema entre os previsores, redução de custo computacional promovida por esta divisão e melhoria de generalização a partir da combinação da solução de vários previsores. Estes fatores o tornam uma ótima oportunidade de pesquisa e aplicação no contexto de investimentos baseado em previsão de séries temporais.

## **4 TRABALHOS RELACIONADOS**

Na literatura, vários modelos de predição do mercado acionário têm sido propostos para permitir ou facilitar a predição de comportamentos apresentados pelo mercado acionário, indicando, principalmente, a oscilação dos preços das ações. A seguir serão descritos alguns trabalhos que possuem abordagens, experimentos ou resultados importantes para compreensão da problemática de predição aplicada a séries temporais financeiras, sobretudo no mercado acionário.

### **4.1 Médias Móveis em predição de Séries Temporais**

Guarnieri (2006a), destacou a importância da utilização das médias móveis como uma ferramenta fundamental para predição de séries temporais, em bolsas de valores. Apesar desta técnica tanto conceitos quanto a forma de aplicação bem simples, seus resultados são promissores.

Em seu trabalho, o autor mostrou que quando as médias móveis são calculadas de forma ponderada melhora os resultados da predição de séries temporais , tanto para critérios de eficácia quanto de eficiência. Guarnieri (2006a), definiu eficiência como a capacidade de atingir sinais corretos, ou seja, comprar quando o preço estiver baixo e vender quando estiver alto. Já a eficácia diz respeito a rentabilidade que pode ser obtida a partir dos valores previstos, o quão lucrativa é considerada a metodologia.

A estratégia de operação utilizada por Guarnieri foi o cruzamento da média móvel com a curva de cotações da ação, sinalizando compra quando a média cruza para cima a curva de cotação e venda quando o cruzamento ocorre no sentido contrário. E um diferencial importante apresentado por este trabalho foi a informação do cálculo de corretagem sobre as operações permitindo uma análise mais refinada dos impactos nos resultados com e sem corretagem. E foi observado nos resultados apresentados pelo autor que a inclusão da corretagem prejudicou várias abordagens pela grande quantidade de operações que técnicas baseadas em média móvel proporcionam. Mas, ainda assim mais de 60% das abordagens mantiveram retornos positivos pela capacidade preditiva.

Seus experimentos foram realizados a partir de um conjunto de médias móveis de 13, 21, 34, 55, 89, 144 e 233 utilizando três formas diferentes de cálculo: média móvel simples (MMS), média móvel ponderada (MMP) e média móvel exponencial (MME). Com isto são totalizados 42 experimentos, pois incluem os cálculos com e sem corretagem. A estratégia que utilizou a média móvel exponencial de 13 dias (MME13) foi a mais eficiente tanto para casos sem corretagem quanto para sua inclusão. Do mesmo modo, com a inclusão da corretagem a estratégia mais eficaz também foi a MME13. Quando a corretagem foi inserida um número maior de estratégias obteve prejuízo e a mais eficaz foi a MMP34, diferente da exponencial que vinha aparecendo com melhor nos outros quesitos.

Além disto, os experimentos foram feitos através de um comparativo entre estratégias de médias móveis e a estratégia *buy-and-hold* que consiste em comprar e permanecer por um bom tempo sem realizar a operação de venda, muito comum em operações de longo prazo. Observou-se que o ganho de desempenho conferido pelas MME e MMP sobre a MMS estão relacionados a suas equações que possibilitam um maior grau de importância para os valores mais recentes, o que foi eliminado ao adicionar a corretagem.

Por fim o autor agrupou os dados em uma tabela que apresenta informações sobre a média de rentabilidade total das três médias. A MMP apresentou a maior rentabilidade média e a menor variabilidade nos retornos comparada aos outros dois tipos de médias móveis. Então, mesmo que algumas MMEs possuam alto grau de eficiência e eficácia, esta técnica possui uma presença mais significativa de risco que a MMP. Ademais, o autor completa que esta ferramenta apresentou retorno positivo acima de 66% dos casos avaliados e é uma ferramenta de fácil utilização pelo público em geral pela simplicidade de seu cálculo e entendimento.

## **4.2 Predição de séries temporais baseado em Redes Neurais Artificiais e Transformada de Wavelet**

Soares (2008), definiu um modelo de predição de séries temporais econômicas fundamentado em Redes Neurais Artificiais. Seu modelo é utiliza de informações advindas da análise técnica e fundamentalista, por meio de dois módulos em seu modelo. O módulo técnico faz uso de um histórico de cotações formando uma série temporal econômica e o fundamentalista utiliza variáveis econômicas que impacta direta ou indiretamente a série

temporal. O mesmo autor, utiliza duas fontes de dados em seu trabalho com informações de janeiro de 1999 a agosto de 2006, uma para cada módulo, sendo a Bovespa fornecedora de dados para o técnico e a Fundação Getúlio Vargas a do fundamentalista. Após a aquisição dos dados cada módulo segue seu fluxo independente um do outro até terminar seus processamentos individuais e repassar as informações para a RNA.

O técnico inicia extraiendo série temporal a partir preços que formam a cotação diária. Em seguida, dois procedimentos são utilizados para facilitar a predição da RNA melhorando seu processo de aprendizado, primeiramente a série temporal será submetida a transformada de wavelet para reduzir as altas frequências presentes em seus valores, e logo em seguida, serão selecionados alguns padrões de treinamento a partir do algoritmo de projeções sucessivas (SOARES, 2008). Já o módulo fundamentalista objetiva explicar o comportamento da série a partir do ponto de vista de dados financeiros relacionados com a mesma. Foram escolhidas variáveis macroeconômicas que estão relacionadas ao desempenho geral da economia tanto no mercado nacional quanto internacional como a cotação do dólar, e variáveis financeiras que informam o desempenho individual da empresa como por exemplo seu lucro líquido. Destes dois grupos de variáveis serão selecionadas as que estiverem estatisticamente correlacionadas com a série através do coeficiente de correlação de Spearman (SOARES, 2008).

Os módulos técnico e fundamentalista fornecem dados para treinamento de uma RNA de Elman<sup>12</sup> que realiza a predição. A estrutura da rede é determinada pelos modelos, de acordo com a quantidade de padrões selecionados pelo técnico e variáveis financeiras pelo fundamentalista. E após o treinamento da rede foram utilizados dados de agosto a novembro de 2006 para validar a predição. O autor menciona como ressalva que durante o projeto da rede, que envolve muito do método tentativa-erro, foi observado que dentre os principais algoritmos de treinamento disponíveis o *Resilient Propagation* obteve os melhores resultados para questão velocidade de treinamento quanto possuir menor chance de atingir de mínimos locais.

Para realização dos experimentos foram selecionadas 15 séries temporais de empresas diferentes cotadas na bolsa. Os resultados do modelo proposto pelo autor foram comparados com outras duas técnicas clássicas, uma utiliza auto-regressão em um modelo ARIMA e a outra usa uma RNA apenas com dados técnicos. E seus resultados foram analisados a partir da medida de erro *Root Mean Square Error of Prediction – RMSEP*.

---

<sup>12</sup> RNA que apresenta comportamento recorrente, ou seja, as saídas dos neurônios não são necessariamente ligadas na entrada dos neurônios da camada seguinte, podendo ocorrer ligações entre neurônios da mesma camada ou de camadas anteriores.

O autor supracitado, constatou a partir dos resultados que a adição de informações fundamentalistas melhoraram a predição do modelo quando comparado aos resultados das outras técnicas, atingindo uma melhora de 47% em relação a RNA e até de 76% comparada ao ARIMA. Porém, alguns experimentos não selecionaram nenhuma variável fundamentalista para treinamento da rede, nestes casos, o modelo proposto foi ligeiramente melhor que o ARIMA e em alguns casos foi pior que a RNA.

#### **4.3 Predição de estimativa de rentabilidade do mercado acionário usando candlesticks**

Morris (2006), descreve em seu livro uma abordagem estatística de investimento no mercado acionário baseada em estimativas de padrões *candlestick*. Estes padrões são descritos de acordo com algumas informações como frequência de acontecimento e perspectiva de lucros ou prejuízos. Os dados de rentabilidade calculados estão organizados em um intervalo de sete dias posteriores ao acontecimento do padrão, com informações percentuais para cada um destes dias. Foi utilizado mais de 14 milhões de dias para gerar estas informações estatísticas, acredita-se que isto fortalece as informações adquiridas a partir desta metodologia.

O resultado de seu trabalho foi uma grande base de informações, nomeada de catálogo de padrões, em que cada elemento (*candlestick*) do catálogo possui um grupo de informações estatísticas sobre sua ocorrência, bem como exigências para sua utilização. Os dados de cada elemento podem ser observados na Figura , que ilustra informações como a frequência que o padrão apresentou ao longo do *dataset*, os percentuais de rendimento durante os próximos sete dias após sua ocorrência e requisitos para sua utilização como tendência e confirmação.

<b>Nome do Padrão:</b>	Matching Low +							<b>Tipo:</b> R+					
<b>Nome Japonês:</b>	Niten zoko/kenuki												
<b>Tendência Requerida:</b>	Sim			<b>Confirmação:</b> Não									
<b>Frequência (MDaysBP):</b>	590												
Estatísticas do padrão de 7.275 Ações Ordinárias, com mais de 14,6 milhões de dias de Dados													
<b>Intervalo (dias)</b>	1	2	3	4	5	6	7						
<b>% Ganhos</b>	69	64	62	61	60	59	59						
<b>Média % de Ganho</b>	3,63	4,71	5,42	5,98	6,64	6,98	7,37						
<b>% Perdas</b>	31	36	38	39	40	41	41						
<b>Média % de Perdas</b>	-2,60	-3,42	-3,92	-4,39	-4,75	-5,13	-5,48						
<b>Lucro líquido</b>	1,23	1,43	1,55	1,65	4,75	5,13	1,82						

Figura 14 - Informações detalhadas do padrão “Matching Low” + adaptado de (MORRIS, 2006)

A confirmação que um padrão requer informa, de modo geral, o grau de confiança que possui, apresentando valores como não requerida, requerida ou sugerida. Estes valores enumeram sua confiabilidade, pois um padrão que não precisa de uma confirmação pode ser utilizado com maior segurança que um que necessita de alguma forma de confirmação posterior sua identificação. Outro aspecto fundamental que está presente nestes detalhes supracitados é o tipo da tendência que o padrão ocorre, representado pelo sinal após seu nome identificador, que no caso da figura anterior Matching Low + indica tendência altista, podendo variar entre + (*bullish*) e – (*bearish*). Por último, é possível notar na Figura na parte direita superior o tipo do padrão, apresentando valores de continuação (C) ou reversão (R), com possibilidades de reversão bullish (R+) ou bearish (R-).

Como o foco deste trabalho é a utilização de padrões de *candlestick*, Morris também sugere formas de identificá-los através de um conjunto de regras que envolve a tendência vigente e o intervalo ocupado pelo corpo e sombra do padrão e uma figura exemplificativa. Para exemplificar como funciona este processo, o padrão chamado “Matching Low +” possui como regra de identificação os seguintes quesitos: tendência de baixa, identificação de dois *candles* com o primeiro sendo um longo dia de baixa e o segundo também de baixa com o preço de fechamento coincidente com o primeiro dia, como ilustrado pela Figura . O dia longo é considerado por morris como um *candle* com um corpo que esteja percentualmente acima da média dos últimos *candles*.

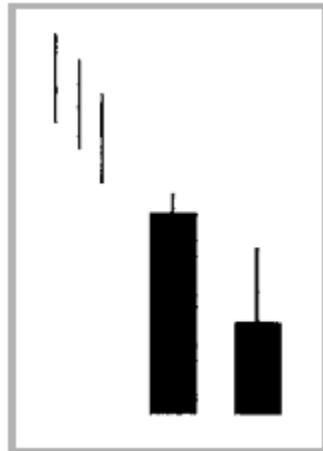


Figura 15 - Representação gráfica do *candlestick Matching Low+*

Uma informação importante comentada por Morris (2006) em seu trabalho é que ninguém deve usar sua abordagem estatística de investimento baseada em *candlesticks* irrestritamente, por mais que exista uma forte segurança por trás de seus dados estatísticos em uma grande base de dados formada por um conjunto de mais de 14 milhões de dias analisados, em alguns casos ocorre a necessidade de complementar alguma exigência do padrão como uma tendência ou confirmação e em outros casos os *candlesticks* se enquadraram muito melhor como ferramenta complementar a outra da própria escola técnica.

A característica “tendência requerida”, na Figura 14, ilustra uma pequena limitação no trabalho de Morris, por dois motivos. Primeiro, existem padrões que apresentam as mesmas características quando identificados, mas possuem duas classificações, uma para tendência de alta e outra para tendência de baixa. Segundo, é necessária a utilização de uma abordagem complementar que forneça a informação sobre a tendência atual do mercado para classificar os padrões dependentes da tendência.

#### **4.4 Comitê de Máquinas para predizer séries temporais**

Os autores Kim e Kang (2010), propuseram um modelo de predição baseado em um *ensemble* heterogêneo denominado PAPEM (Pattern Prediction Ensemble Model) com principal objetivo de melhorar o desempenho da predição de séries temporais através de uma estrutura multinível de *ensemble*. Este modelo híbrido é baseado em três tipos de métodos

comuns no domínio da predição: Redes Neurais com função Base-Radial (RBF), o método de K-vizinhos próximos (KNN) e os Mapas Auto-Organizáveis (SOM).

O modelo proposto pelos autores divide a realização da predição em dois níveis que funcionam da seguinte forma: o primeiro nível de predição é formado por modelos elaborados pelas três abordagens isoladas e o seu resultado é utilizado para treinar o *ensemble* que se encontra no segundo nível e possui sua estrutura definida a partir de diferentes configurações destas mesmas abordagens.

Para avaliar o modelo proposto em seu trabalho os autores utilizaram três datasets com séries temporais de diferentes categorias. O *dataset* chamado *Mackey-Glass* é formado a partir de uma equação diferencial não-linear, outro chamado de *Sunspot* corresponde a dados de manchas solares adquiridos a partir de um instituto de pesquisa chamado SIDC (Solar Influences Data Center) e o último *dataset* possui dados de cotações formadas pelo índice IBEX. Segundo o autor as séries temporais *Mackey-Glass* e *Sunspot* que compõe dois *datasets* são largamente utilizadas para comparar diferentes modelos de predição, por isso um *dataset* foi criado a partir de uma série temporal financeira para verificar a generalização e a precisão de predição calculada.

A avaliação das simulações foram feitas através de três medidas de erro utilizada para avaliar a precisão da predição de valores o MSE, MAPE e RMSE. Estas medidas de erro calculam diferenças entre o valor previsto e o valor real. A avaliação realizada sobre os resultados mostrou que a abordagem ensemble proposta possui melhor desempenho que seus componentes individuais.

Como resultado deste trabalho a abordagem proposta mostrou superioridade em todos os aspectos analisados quando comparada a métodos independentes (RBF, KNN e SOM) e até mesmo a ensembles homogêneos utilizando estes três métodos. O autor finaliza com uma conclusão bem ousada que devido ao comportamento muito bom do PAPEM é possível utilizá-lo para predição de padrões em qualquer tipo de dados de séries temporais não lineares.

## 4.5 Discussão

A leitura destes trabalhos motivou questionamentos importantes sobre a predição de séries temporais, de tal forma que formasse a base para o modelo híbrido de predição. A

principal questão é qual a melhor abordagem para predição de séries temporais financeiras. Todos os trabalhos comentados trouxeram benefícios e apresentaram diversas contribuições.

As médias móveis utilizadas por Guarnieri (2006a), são um ótimo exemplo de aplicabilidade de um simples e efetivo modelo de predição. Este trabalho desperta o interesse para usar indicadores técnicos, visto que o mais simples deles alcançou bons resultados, então o uso de outros mais elaborados e adequados para outras situações em que as médias móveis não se comportem bem obteriam melhores resultados. Isto foi evidenciado por Petry (2010), ao utilizar as Bandas de Bollinger e sugerir o uso combinado com o Índice de Força Relativa (IRF), por Vidotto (2009) que utilizou o MACD como principal ferramenta para sugestão de operações sobre ações de cinco empresas escolhidas de forma aleatória em um espaço amostral de noventa empresas aplicado sobre as cotações de 2006 atingindo uma rentabilidade média de 26% dos investimentos, ou com Petrokas, Famá e Modro (2010), que comparou investimentos diferentes de acordo com sua rentabilidade entre os anos de 2008 e 2009. O objeto de estudo foi analisar os rendimentos da renda fixa representada pela caderneta de poupança contra a renda variável representada pela análise técnica. As rentabilidades obtidas pela análise técnica foram bem superiores e eram obtidas através de três *setups*, que são roteiros de atividades ou estratégias que buscam indicar os melhores momentos para efetuar operações de compra e venda. Estes três *setups* são configurados através de indicadores técnicos como médias móveis, o IRF ou simplesmente pela identificação de pontos máximos e mínimos em tendências, que o autor chamou de ponto contínuo.

A utilização de abordagens computacionais de predição com indicadores técnicos destaca-se pelo uso de técnicas bem simples a complexas equações que representam o cálculo dos indicadores. Novas estratégias a partir de indicadores técnicos surgem a todo momento no mercado acionário, destacando esta ferramenta como uma ótima escolha para predição de séries temporais financeiras.

As RNAs são aplicadas com os mais diversos objetivos de predição, por exemplo: antecipar preços futuros, tendência, reversões, ou o importante índice IBOVESPA<sup>13</sup>. Isto pode ser observado nos trabalhos de Soares (2008), Marangoni (2010), Faria (2008) e Bauer (2009). Soares (2008), destaca uma importância muito grande ao pré-processamento dos dados submetidos a rede cujo objetivo é facilitar seu treinamento e aprendizagem, também relatado por Costa (2008), a importância do pré-processamento para o sucesso da RNA. A existência de inúmeras abordagens utilizando RNAs ocorre porque ao alterar o algoritmo de treinamento é configurada uma nova solução para o problema, ilustrado nos testes de Souza

---

<sup>13</sup> Principal índice da Bovespa, calculado a partir das ações de maior volume de negociação.

(2012), que demonstrou o potencial de uma simples MLP comparada a outras arquiteturas bem mais complexas. Gouveia (2009), demonstrou que o *backpropagation* é inferior ao *Resilient Propagation* (RPROP) por este segundo convergir mais rapidamente e ser menos suscetível a problemas clássicos de algoritmos baseados no gradiente descendente como os mínimos locais. Já Souza (2008), usou uma RNA para prever preços de *commodities* e durante sua discussão não recomendou o uso desta mesma rede em outro domínio, pois fatores influenciadores que moldaram sua rede no domínio de *commodities* podem não estar presentes em outros mercados que formaram séries temporais mais específicas, ou em séries não-financeiras, podendo ocasionar prejuízos para utilização da RNA nestes domínios fora de seu projeto inicial de aplicação.

Diante das considerações acerca RNAs a utilização do algoritmo de treinamento *Resiliente Propagation* (RPROP) como melhor alternativa para predição no contexto de séries temporais merece destaque pelos vários autores que comprovaram a superioridade deste algoritmo quando comparado com tradicional *backpropagation*, sendo apontado como um substituto melhor que seu antecessor. Além de evidências importantes sobre uma configuração simples de MLP apresentar resultados tão bons quanto arquiteturas mais bem elaboradas.

A utilização de padrões gráficos em séries temporais reúne uma grande quantidade de técnicas distintas devido a existência de várias formas e tipos de padrões que ocorrem nas cotações das ações. A própria curva formada pela oscilação dos preços forma padrões gráficos, porém os mais conhecidos são os *candlesticks*, que fornecem inúmeras possibilidades de estratégias de investimento relacionadas a sua ocorrência histórica. Fato evidenciado por como apresentado por Morris (2006), ao definir estimativas financeiras associadas ao acontecimento de *candlesticks*, além de definir regras para auxiliar o reconhecimento e a classificação destes padrões. Uma ferramenta que efetue a identificação destes tipo de padrão facilitado pela sugestão de regras de identificação e consiga associar aos percentuais de rendimento que cada um possui oferece boas perspectivas e contribuições para compor um previsor aplicado a investimentos na bolsa de valores.

Diante das abordagens promissoras de predição que foram citadas e aprofundando no tema de comitê de máquinas de aprendizado, surge a ideia central do modelo híbrido de predição. Através da utilização de comitês do tipo *ensemble* como peça fundamental para realização da predição é possível mesclar múltiplas abordagens distintas em uma única metodologia híbrida auxiliados pelo processo de combinação do *ensemble*. Um benefício

importante deste tipo de abordagem é a garantia que o seu resultado final é no mínimo igual ao melhor de seus previsores individuais, quando não é superior a todos os previsores usados.

A utilização de modelos preditivos de forma híbrida ou através de combinação é uma tendência, visto que os autores observaram ganhos significativos com a utilização de abordagens que combinam modelos preditores, quando utilizados em iguais condições, a abordagens com um único modelo preditor, como constatado por Villanueva (2006). Isto ocorre tanto para ensembles homogêneos utilizados por Chitra *et al* (2010), quanto para heterogêneos usados por Kim e Kang (2010). Do mesmo modo Ribeiro et al (2009), ao estudar diferentes formas de predição advindas da matemática ou baseados em técnicas de inteligência artificial percebeu que a utilização de sistemas híbridos, ou comitês, primam pelo aproveitamento individual de diferentes previsores para solucionar um problema comum. Nesta mesma linha de pesquisa, Macedo (2010), comprovou estatisticamente a hipótese que a utilização de comitês geram resultados melhores ou no mínimo tão bons quanto o melhor resultado dos modelos executados isoladamente, por outro lado o mesmo autor reforça que isto não permite afirmar que seu comitê funcionará de maneira semelhante para todas as ações da bolsa.

Ao fim da análise destes trabalhos relacionados e outros presentes na literatura tornou-se muito interessante a elaboração de uma abordagem híbrida de predição utilizando um grupo heterogêneo de técnicas derivadas de áreas de conhecimento diferentes como reconhecimento de padrões, economia e matemática.

## 5 MODELO HÍBRIDO DE PREDIÇÃO

O modelo de predição proposto combina as seguintes técnicas: i) mapeamento estatístico de rendimentos baseado em padrões *candlestick*; ii) redes neurais artificiais (RNA) para a previsão de preços futuros; e iii) utilização de indicadores técnicos para a definição da tendência do mercado. A ocorrência de *candlesticks* está associada a valores percentuais de lucro ou prejuízo para um intervalo de dias futuros, uma RNA permite antecipar o valor do próximo dia e calcular uma estimativa de crescimento ou queda dos preços da ação baseado no valor mais atual, e indicadores técnicos possibilitam a confirmação de tendências e pontos de reversão ao longo das cotações financeiras. Com essas técnicas é possível definir um modelo híbrido para predição do comportamento do mercado acionário que alcance resultados mais promissores do que modelos que apenas baseiam-se na utilização única de uma das técnicas supracitadas como o caso dos trabalhos de Guarnieri (2006a), Soares (2008), Morris (2006) e Kim e Kang (2010) discutidos no capítulo 4 como trabalhos relacionados.

Villanueva (2006), evidenciou que é muito pouco provável que apenas uma técnica seja capaz de acertar eficientemente a antecipação do comportamento do mercado, visto que essas técnicas não são completas e independentes uma das outras. Segundo Ribeiro, Goldshmidt e Choren (2009), a combinação de mais de uma técnica pode possibilitar melhores resultados, uma vez que as limitações de um ou mais previsores podem ser superadas pelos benefícios dos outros. Recentes trabalhos comprovam a eficiência de abordagens baseadas em combinação na área financeira, Kim e Kang (2010), utilizou um ensemble homogêneo composto por RNAs para prever a falência de empresas baseado em indicadores financeiros. Chitra *et al* (2010), usou um ensemble heterogêneo baseado em tecnologias de RNAs e mapas auto organizáveis para predição de séries temporais e também aplicou a séries temporais financeiras. Ambos os trabalhos de Kim e Kang (2010) e Chitra (2009) realizaram experimentos com o ensemble e com os previsores individualmente para demonstrar que o ensemble apresenta como benefício a melhoria da predição.

Construir um previsor aplicado ao mercado acionário utilizando dados de *candlesticks* acrescenta o conhecimento de padrões gráficos em séries temporais financeiras, que segundo teorias como a de Elliot e de Dow a aparente desorganização do mercado acionário não existe, mas sim um conjunto de padrões recorrentes que se repetem frequentemente nas cotações. Já as RNAs, por serem consideradas aproximadores universais, representam bem o suposto comportamento aleatório de séries temporais financeiras prevendo

valores futuros com alta precisão. O uso de indicadores técnicos como modelo de predição acontece profissionalmente por corretores que analisam gráficos de ações, pois esta ferramenta fornece resultados bem efetivos para antecipação de comportamentos oscilatórios na curva de cotações, indicando operações de compra e venda baseadas em informações sobre a continuação da tendência ou momentos de reversão.

Portanto, acredita-se que as técnicas de mapeamento de padrões de *candlestick*, de RNAs e de indicadores técnicos possuem características que as tornam complementares e adequadas para o desenvolvimento de um modelo híbrido. Desta forma, este modelo pode determinar a tendência do comportamento oscilatório do mercado acionário. E para o desenvolvimento de um modelo com estas técnicas é necessário a utilização da abordagem de comitê de máquinas de aprendizado, e como os previsores individuais já estão definidos segundo as três abordagens citadas o processo de construção da arquitetura do comitê é reduzido, mesmo presumindo a existência de fases como geração e seleção de previsores. Algumas abordagens utilizam comitês do tipo *ensembles* homogêneos com todos os previsores construídos a partir de RNAs como o caso de Kim e Kang (2010), já outras abordagens utilizam *ensembles* heterogêneos com previsores utilizando diferentes abordagens computacionais semelhante ao que é almejado pelo modelo proposto.

A abordagem proposta não usa apenas uma combinação de metodologias diferentes de predição, mas abordagens com áreas de conhecimento distintas (economia, estatística e inteligência artificial) que podem oferecer benefícios muito superiores quando comparada a *ensembles* homogêneos ou porventura um heterogêneo que utilize abordagens computacionais baseadas na mesma área de conhecimento, por exemplo estatística

## 5.1 Metodologia

A construção de um modelo híbrido de predição financeira requer uma estruturação criteriosa desde aquisição dos dados das cotações ao resultado final obtido, o qual provém da combinação de resultados individuais dos previsores que formam o *ensemble* heterogêneo. As seguintes etapas foram definidas para o modelo de predição que está sendo proposto, sendo que uma ou mais podem ser executadas simultaneamente:

- Definir um dataset a partir de informações de cotações, indicadores técnicos, séries temporais e etc;

- Carregar os valores das cotações através de plataforma *web* e mantê-la localmente para processamento do *dataset*;
- Criar série temporal a partir dos valores contidos no *dataset*;
- Calcular indicadores técnicos a partir de série temporal;
- Projetar rede neural utilizando série temporal financeira como entrada;
- Desenvolver reconhecimento de padrões em candlesticks;
- Definir estratégias de investimento baseado nos resultados dos previsores;
- Elaborar a solução híbrida de predição combinando resultados de vários previsores.

Inicialmente, um dataset local é construído a partir dos dados de cotações que são obtidos através da plataforma do *YahooFinance*, conforme processo de aquisição de cotações utilizado por Gupta, Sundaram e Shamsi (2006). Esta plataforma disponibiliza informações financeiras que podem ser requisitadas via URL<sup>14</sup>. Em uma requisição pode ser passado os seguintes parâmetros(que são localizados na URL após o sinal de interrogação): código da ação, mês inicial, dia inicial, ano inicial, dia final, mês final, ano final e agrupador de valores. Na Tabela 1, são mostrados cada parâmetro e a sua respectiva descrição. como pode ser observado a seguir um exemplo de aquisição das cotações de ações da Petrobras entre os anos de 2008 a 2012.

Tabela 1 - Descrição dos parâmetros da URL padronizada do *YahooFinance*

Parâmetro	Descrição
s	Código da ação
a	Mês inicial
b	Dia inicial
c	Ano inicial
d	Mês final
e	Dia final
f	Ano final
g	Tipo de agrupamento

Para exemplificar a utilização dos parâmetros listados na Tabela 1 será utilizada a seguinte requisição ao *YahooFinance*:

- <http://ichart.finance.yahoo.com/table.csv?s=PETR4.SA&a=00&b=1&c=2008&d=00&e=31&f=2012&g=d>.

---

<sup>14</sup> Do inglês Uniform Resource Locator que significa Localizador Padrão de Recursos, é o endereço de um recurso disponível na rede, por exemplo na internet.

Observa-se como valor do primeiro parâmetro o identificador PETR4.SA representando a Petrobras, seguido dos parâmetros que definem o intervalo da cotação (a, b, c, d, e, f) com o duas ressalvas: o mês aparece antes do dia na sequência de data e os meses iniciam em 0 para janeiro e encerram em 11 para dezembro. O último parâmetro chamado de tipo de agrupamento e definido pelo identificador “g” pode assumir valores como “d” quando a cotação for apresentada com valores diários, ou podendo receber o valor “m” para distribuição da cotação de forma mensal.

Além desses parâmetros, existem outros que permitem formatar os dados ou calcular algumas outras informações úteis para análise do mercado financeiro, mas foram omitidos por saírem do foco do modelo de predição desejado. A Figura 16 ilustra o resultado de uma requisição de dados de uma determinada ação de acordo com o seguinte formato: data, valor da abertura, máximo, mínimo, fechamento e volume.

29/01/2013	19,35	19,37	18,84	19,11	27.097.900	VOLUME
28/01/2013	19,73	19,81	19,27	19,37	10.122.600	
24/01/2013	19,37	19,75	19,37	19,61	15.012.000	
23/01/2013	19,42	19,63	19,23	19,55	17.200.800	
22/01/2013	19,42	19,61	19,23	19,50	31.535.100	
21/01/2013	19,57	19,6	19,27	19,39	18.086.200	
18/01/2013	19,8	19,89	19,54	19,56	16.913.500	
17/01/2013	19,86	19,93	19,6	19,71	19.719.600	
16/01/2013	19,89	19,89	19,6	19,84	16.797.000	
15/01/2013	20,01	20,24	19,69	19,82	29.633.900	
14/01/2013	20,01	20,24	19,63	19,72	28.502.400	
11/01/2013	19,85	20,04	19,7	19,88	18.223.600	
10/01/2013	19,77	20,05	19,54	19,93	37.536.200	
09/01/2013	19,64	19,87	19,46	19,68	24.361.100	
08/01/2013	20,11	20,23	19,46	19,5	29.091.300	
07/01/2013	20,48	20,67	19,95	20,08	28.069.600	

Figura 16 - Informações obtidas do *YahooFinance*

Como observador na Figura 16, cada linha da cotação contém informação referente a uma data, estruturada conforme uma forma de agrupamento de valores (no caso desta figura é apresentado o agrupamento diário), em seguida contém quatro valores que formam um *candlestick*, e por fim o volume de transações financeiras.

Posteriormente, após a etapa de construção do dataset, ocorre o de pré-processamento destes dados a fim de facilitar e possibilitar o processo de predição. Esta etapa consiste da adequação a previsores que possuem um conjunto de requisitos que precisam ser atendidos durante o pré-processamento. Por exemplo, um previsor formado por RNAs tem um aumento de desempenho significativo quando os dados de entrada são normalizados antes da utilização, já outro que precise de uma entrada na forma de séries temporais precisa reduzir as

informações do *dataset*. No pré-processamento definido são realizadas as seguintes tarefas: i) extração das séries temporais com o preço de fechamento da ação por data especificada, ii) uma transformação com a finalidade de limpeza ou simplificação dos dados através da normalização e iii) identificação de tendências curtíssimas no período de até quatro dias do *dataset*. Uma vez que o *dataset* é composto de um grande grupo de informações, parte destas não são relevantes para qualquer previsor construído, por exemplo: um previsor que usa uma série temporal só precisa de duas informações de cada linha dos dados obtidos que formaram o *dataset*, já um previsor que utilize reconhecimento de padrões de *candlestick* precisaria de todas as informações do *dataset*.

Em seguida a etapa de pré-processamento, ocorre a execução dos previsores estatístico, neural e técnico. O estatístico tem como finalidade a geração da saída com o valor estimado de lucro que cada *candlestick* oferece. Já o neural fornece também uma estimativa, porém de crescimento ou queda no preço da ação. E por último, o técnico através do cálculo de indicadores técnicos informar momentos de reversão na tendência atual da cotação. Estas técnicas funcionam de maneira independente e ao fim de seus processamentos têm suas saídas combinadas em um único resultado.

Por ultimo, a geração do resultado final a partir dos três resultados gerados anteriormente ocorre através da ponderação destes resultados de acordo com critérios de importância pelo qual cada previsor recebe um peso. Essas duas últimas etapas, de execução dos previsores e combinação dos resultado, serão detalhadamente abordadas a seguir, na próxima subseção.

## 5.2 Solução Proposta

O modelo de predição proposto consiste essencialmente de três previsores que compõe um comitê de máquinas de aprendizado. O previsor estatístico tem como função realizar estimativas de rendimento através da técnica de padrões de *candlestick*. Já o previsor neural é o responsável por calcular o preço de fechamento do dia seguinte, além do percentual de crescimento ou queda de preços através da técnica de rede neural artificial. Por último, o previsor técnico tem como funcionalidade informar a propensão da cotação da ação iniciar uma tendência de alta ou de baixa nos seus preços através da técnica de indicadores técnicos.

É importante destacar que cada previsor tem como saída um resultado individual e que, posteriormente, estes resultados gerados são combinados. Um dos métodos mais utilizados para essa tarefa é a ponderação dos resultados de acordo com critérios de importância, onde os previsores recebem um peso de acordo com seu grau de relevância para solução final e sua saída é ponderada de acordo o valor deste peso (RIBEIRO *et al*, 2009). No modelo proposto, a combinação dos resultados dos previsores ocorre através da Equação 4.1.

$$f_c = \sum_{i=1}^M w_i f_i \quad (3.9)$$

A Equação 4.1 acima define como os valores de cada resultado são combinados, onde o valor de saída  $f_c$  é calculado através de um somatório ponderado dos resultados dos  $M$  previsores. Seja  $f_i$  o resultado individual de cada previsor e  $w_i$  o peso que define o grau de importância da solução que o previsor apresenta para o *ensemble*.

O comitê desenvolvido possui a estrutura de um *ensemble* (como explicado na subseção 3.5) onde cada previsor opera isoladamente, possui abordagens distintas e é definido por etapas sequenciais. A linha de execução geral de um previsor passa por três etapas: 1) recebimento da entrada, 2) processamento de predição e 3) geração da saída. Não há exigência quanto a ordem para execução dos previsores individualmente do comitê, contanto que cada um realize as operações de acordo com esta linha de execução definida o comitê irá funcionar adequadamente e a combinação irá ocorrer após o cálculo de todos os três previsores que o compõe. Mas, existe uma condição quanto ao resultado calculado, que precisa estar definido sob o mesmo domínio, como explicado, todos são valores de expectativas de rentabilidade da ação. Seus resultados individuais representam uma solução definitiva para o problema, mas sob a perspectiva de uma abordagem específica e o comitê apenas combina estas soluções particulares em uma mais geral.

O resultado final desta abordagem de predição pode ser interpretado de várias maneiras, como por exemplo: informar em termos percentuais qual a tendência que o ativo irá apresentar para os próximos dias, ou informar qual o preço da ação para o dia ou período seguinte, ou é possível refinar a saída dos previsores e informar de maneira simples se é um bom momento para comprar ou vender o ativo. As subsecções seguintes discutem e explicam detalhadamente o funcionamento de cada previsor e a realização dos passos da execução cada um.

### 5.2.1 Previsor Estatístico

Um previsor estatístico segue uma abordagem voltada para o reconhecimento de padrões que pode antecipar oscilações do mercado financeiro. Esta antecipação (predição) pode ser alcançada usando padrões de *candlestick* e uma base de dados estatísticos. O previsor estatístico que está sendo proposto baseia-se na abordagem de Morris (2006).

Um padrão gráfico de *candlestick* possui uma forma estrutural específica que o define, podendo ser descrita através de regras com suas características como amplitude entre abertura e fechamento, existência de sombras inferiores ou superiores, dentre outros. Tomando como exemplo o padrão chamado *hammer* (martelo), explicado por Debastiani (2007), sua estrutura é formada por um único *candlestick* com um pequeno corpo sobre uma grande sombra formando uma figura semelhante a um martelo, o qual até o identifica como a tradução de seu nome *hammer*. Portanto, para identificar um *hammer* é necessário que duas regras sejam satisfeitas para cada tipo:

- O pequeno corpo está na extremidade superior do intervalo de negociações;
- A cor do *candlestick* não importa (preto para encerramento em queda e branco para encerramento em alta);
- A sombra inferior precisa ser pelo menos duas vezes maior que o corpo;
- Não deve haver sombras superiores.

A representação gráfica da ocorrência deste padrão é ilustrada pela Figura 17, e as classificações altista e baixista estão relacionadas ao direcionamento do padrão. Um padrão altista ocorre em mercados crescentes e indica que o padrão fechou com lucro, já os baixistas acontecem na queda e informam prejuízos. Quando existem duas formas de classificar o padrão cada uma possui um valor diferente para o percentual de ganho e perda associado ao padrão, sendo assim, é necessário identificar qual a tendência atual que o mercado segue antes de efetuar a classificação do padrão, para evitar predições errôneas.

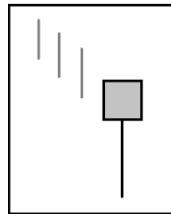


Figura 17 - Representação gráfica do padrão *candlestick hammer*

A ocorrência de cada padrão conhecido pelo previsor gera uma expectativa de rentabilidade, seja ela de lucros ou prejuízos. Portanto, cada padrão está associado a dois grupos de informações, primeiro as regras de identificação que definem suas ocorrências e em segundo lugar os valores que informam a rentabilidade quando cada padrão ocorre. A Figura 18 ilustra o fluxo de funcionamento do previsor estatístico que está sendo proposto.



Figura 18 - Fluxo de funcionamento do componente estatístico

O início da execução deste previsor requer um pré-processamento que extraia informações do *dataset* e as converta em *candlesticks*, no caso dos dados já informados na Figura é excluído apenas o volume de transações para execução do primeiro passo do fluxo. A entrada de dados deste previsor contém informação de *candlesticks* referentes aos últimos quatro dias de cotação do ativo, construídos a partir de informações pré-processadas contidas no *dataset*. Esta entrada possui o valor fixo de quatro dias de cotação pois o maior padrão de *candlestick* que é identificado pelas regras possui quatro *candlesticks*.

Em seguida, com base nestas informações da entrada, a identificação do padrão é realizada conforme o grupo de regras definidas, que estabelece as características de vários padrões conforme descrito no Anexo A deste trabalho. O passo seguinte fornece a tendência atual que o mercado segue, pois em alguns casos o padrão exige que o mercado esteja em uma tendência com direção especificada. Tendo a estrutura do padrão e a tendência o próximo passo é iniciado para realizar a classificação deste padrão, também acontecendo só em alguns casos quando uma mesma estrutura de um padrão ocorre em mais de uma tendência proporcionando rentabilidades diferente para cada uma.

A última etapa calcula a saída do previsor baseado em dois valores apresentados pelo padrão de *candlestick*, o percentual de ganho e de perda. A fim de tirar uma estimativa da rentabilidade oferecida através destes dois valores, é utilizada a Equação 4.2 que calcula uma proporção entre os possíveis lucros e prejuízos. Existem outras informações que a ocorrência dos padrões informam como frequência de acontecimento, média de ganhos e perdas, mas não há necessidade de utilizar mais do que a equação precisa para calcular a saída do previsor. Além do mais, a inclusão de mais informações talvez desgastasse o desempenho do previsor ou incluísse informações desnecessárias ou redundantes.

A saída calculada não utiliza todas as informações fornecidas pelo estudo de padrões realizado por Morris (2006). Primeiramente, por há necessidade de informar alguns valores como a frequência que este padrão teve nos experimentos, ou informações similares como o percentual de ganho e a média percentual de ganho, selecionando apenas os dois necessários para aferir o lucro proporcionado pelo padrão. Uma grande quantidade de dados pode interferir diretamente no desempenho do previsor, motivando também a utilização apenas de dados necessários para predição para não prejudicar este componente que é utilizado para investimentos financeiros.

$$\text{Lucro Estimado} = \% \text{ ganho} / \% \text{ perda} \quad (4.2)$$

A Equação 4.2 representa a proposta de construção do previsor estatístico, calculando o seu resultado final apenas com as seguintes informações: o percentual de ganho e de perda que cada padrão identificado pode gerar. Estes dois valores informam a expectativa de lucro para investidores que iniciem suas aplicações a partir da ocorrência de um padrão identificado baseando-se em previsões para o dia seguinte. Como só é calculado o lucro

estimado para padrões que o previsor conhece, não resulta em um grande número de operações de compra e venda sugeridas por interpretações deste previsor.

### 5.2.2 Previsor neural

Um previsor neural tem em sua estrutura, como um dos principais elementos, uma rede neural artificial que possui como objetivo antecipar o dia seguinte de um ativo através do preço de fechamento da ação. A RNA proposta é definida através da arquitetura MLP *feedforward* e utiliza o algoritmo de treinamento RPROP, o qual foi selecionado devido às vantagens oferecidas para o domínio do mercado financeiro e também relacionadas à predição de séries temporais, na predição de preços. Apesar de um previsor neural permitir mais de uma forma de predição como, por exemplo, o reconhecimento de padrões nas cotações ou na classificação destes padrões (tarefa realizada pelo previsor estatístico descrito anteriormente), o que está sendo proposto é capaz de predizer o preço de fechamento do dia seguinte.

A MLP é definida através da camada de entrada, saída e oculta. A camada de entrada foi idealizada com 10 nós que correspondem respectivamente aos 10 últimos valores dos preços de fechamento da ação. Já a camada de saída da MLP apresenta um nó que tem como valor futuro do 11º preço de fechamento do , o dia seguinte. Por último, a camada oculta define como a entrada vai gerar a predição através do processamento destes dados pelas conexões ponderadas entre a entrada e a saída, sendo utilizada a função de ativação sigmoide por padrão em todas as camadas. Como não existe na literatura uma forma de construir uma RNA ideal, tampouco uma estrutura padronizada e adequada para aplicar na predição, foram usadas duas heurísticas para definir a estrutura da rede relativa, principalmente, a sua camada oculta. Primeiramente foram geradas aleatoriamente várias configurações variando apenas o número de neurônios na sua camada oculta (de 10 a 30 neurônios), mantendo a entrada com 10 neurônios e saída com um neurônio. Após isto os pesos foram atribuídos aleatoriamente e as redes foram executadas até atingir o valor da medida de erro MSE inferior a 0,01 com o algoritmo *backpropagation*. As cinco melhores configurações que apresentaram o melhor resultado, ou seja, atingiu uma generalização no menor tempo (número de épocas) de execução foram selecionadas como candidatas para representar o previsor neural. De acordo com o levantamento bibliográfico realizado para projetar este previsor este algoritmo de treinamento não é o ideal para utilizar em predição de

séries temporais, então em um segundo momento foram realizadas simulações com as cinco melhores redes aplicadas a algoritmos baseados em variações do backpropagation com a mesma condição de idealidade de atingir o valor do MSE inferior a 0,01. Ao fim destes dois períodos de simulação a configuração ideal da arquitetura da rede ficou com 20 neurônios utilizando o algoritmo de treinamento *Resilient Propagation*.

Para realizar o treinamento da rede deste previsor é necessário separar uma parte do dataset para este treinamento dividindo-o em dois grupos: um ano de cotação para treinamento e o restante para realizar a predição. O algoritmo de treinamento utilizado é o *Resilient Propagation* e o treinamento é executado até a rede atingir um erro  $MSE < 0,01$ . A Figura 19 ilustra as etapas do funcionamento do previsor neural, que são: i) extração de informações do *dataset*, ii) pré-processamento e treinamento da rede, iii) execução do previsor e iv) cálculo do resultado final.

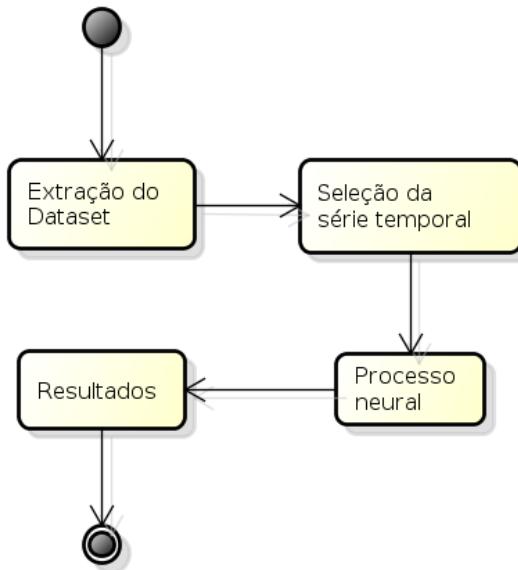


Figura 19 - Fluxo de funcionamento do componente neural

Inicialmente o *dataset* precisa ser processado para extrair e transformar suas informações em uma série temporal para ser utilizada pela rede. Nesta primeira etapa é criada uma série temporal com os preços de fechamento de todo o *dataset* para utilização do previsor.

Após a criação da série temporal utilizada pelo previsor é iniciada uma próxima etapa que seleciona os valores desta série para formar o conjunto de dado utilizado para treinamento da rede e o restante para aplicação da predição, além de normalizar os dados desta série para serem utilizados pela rede. A realização da predição exige que seja utilizada a

mesma quantidade de valores de entrada usada durante o treinamento da rede, portanto 10 preços de fechamentos referentes aos 10 dias anteriores ao previsto. O processamento deste previsor segue para outra etapa submetendo os valores necessários da entrada para rede que calcula o valor resultante da predição que é representado pelo preço do dia seguinte.

Por último, será calculado o resultado final do previsor baseado na saída que a rede calcula e o valor atual que a ação possui, como observado na Equação 4.3. Como a rede possui uma janela temporal de um dia, que significa predizer o preço de fechamento de um dia futuro, esta saída não pode ser utilizada diretamente como resultado do previsor, pois o comitê irá combinar expectativas de lucro ou prejuízo, então a partir desta equação é possível calcular uma estimativa do lucro esperado para o dia seguinte.

$$\text{Lucro estimado} = \frac{\text{Valor previsto}}{\text{Valor atual}} \quad (4.3)$$

Diferente do previsor estatístico que informa estimativas de lucro apenas quando identifica alguma ocorrência de padrão conhecido, o previsor neural irá indicar estimativas de lucro para todos os dias previstos pela RNA, tanto provocando lucro quanto prejuízo. Isto gera uma maior quantidade de operações sugeridas pelo resultado final do previsor neural comparada ao estatístico, porém as chances de predizer informações errôneas também é maior. Desta maneira, um terceiro previsor, o técnico, pode ser utilizado para complementar os dois descritos anteriormente.

### 5.2.3 Previsor técnico

Um previsor técnico tem a responsabilidade de obter a tendência atual do mercado e, assim, antecipar o melhor momento para iniciar ou encerrar um investimento. Este previsor utiliza informações de três indicadores técnicos clássicos: o RSI que informa momentos de reversão na tendência de preços, o ADX que confirma a existência de uma tendência e sua direção e, por fim, as BBs que permitem mensurar a volatilidade do mercado além de definir uma área de valores para o preço do ativo e também informando momentos de reversão para a tendência atual. Para obter o valor de cada indicador que forma este previsor é necessário calcular suas equações explicadas na subseção 2.5 que discute sobre indicadores da análise técnica.

A Figura ilustra as etapas do funcionamento do previsor técnico, que são: i) a extração de informações do *dataset*, ii) execução dos indicadores técnicos em paralelo e iii) cálculo do resultado final do previsor.

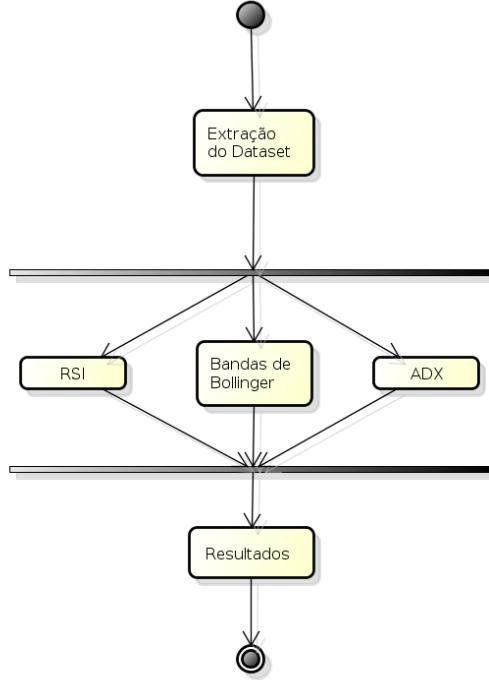


Figura 20 - Fluxo de funcionamento do componente técnico

O funcionamento do previsor técnico inicia semelhante ao previsor neural, realizando inicialmente a extração de uma série temporal do *dataset* com os preços de fechamento da ação para calcular os indicadores técnicos.

Após a série temporal ser construída o fluxo irá se subdividir na iniciando uma próxima etapa que envolve o cálculo de cada indicador técnico separadamente formando soluções parciais deste previsor. Após calcular o indicador, seu resultado é interpretado de acordo com as especificações estabelecidas por cada um que foram discutidas na subseção 2.5, por exemplo: se o ADX ultrapassar o valor 40 e o  $DI^+$  ultrapassar o valor do  $DI^-$  é informado um crescimento para ação proporcionando lucro para investimentos iniciados na data atual. Os resultados calculados pelos indicadores possibilitam três valores: 1 para tendência alta, -1 para tendência baixa e 0 para indeterminações.

A última etapa inicia quando os resultados que foram interpretados dos indicadores são submetidos a Equação 4.4 que pondera suas saídas, com pesos igualmente distribuídos para cada indicador resultando na saída final deste previsor técnico. Assim sendo, cada indicador recebe o peso equivalente a um terço do total, atingindo com isto o valor

máximo de 33% da solução do previsor. Devido a utilização de percentuais de 33% para cada indicador, mesmo que todos informarem chances de crescimento o resultado final deste previsor será de 99%. De maneira simples, quando um indicador confirmar uma tendência de alta o resultado é de 0,33, para dois indicadores 0,66 e três indicadores 0,99. A Tabela 1 mostra as combinações dos possíveis resultados calculados pelos indicadores, com a proporção de um terço igual para cada um

$$\text{Lucro estimado} = \frac{\text{ADX} * 33 + \text{RSI} * 33 + \text{BB} * 33}{100} \quad (4.4)$$

Tabela 2 - Valores estimados para combinação de saídas dos previsor técnico.

<b>Classificação</b>	<b>ADX</b>	<b>RSI</b>	<b>BB</b>	<b>Colunas2</b>
Alta	1	1	1	99%
Alta	1	1	0	66%
Alta	1	0	0	33%
Indefinido	0	0	0	0%
Alta	1	1	-1	33%
Indefinido	1	-1	0	0%
Baixa	-1	0	0	-33%
Baixa	-1	-1	0	-66%
Baixa	-1	-1	-1	-99%

A utilização de uma equação ponderada para calcular o resultado final deste previsor proporciona uma flexibilidade interessante no que diz respeito a alteração dos pesos associados aos indicadores, que pode ocorrer conforme algum critério específico, por exemplo: de acordo com a rentabilidade individual proporcionada durante os experimentos, sendo que o mais rentável fica com uma maior porcentagem da formação da solução final. E como observado na Tabela 2 o grau de certeza sobre uma tendência de alta e possíveis lucros nos investimentos depende exclusivamente das confirmações de todos os indicadores. Por fim, de acordo com a independência entre os previsores, o previsor estatístico pode utilizar os resultados calculados por este previsor (técnico) para identificar a tendência atual e operar sem a necessidade de identificar a tendência a partir de qualquer outro método não tão adequado quanto indicadores técnicos.

### 5.3 Experimentos e simulações

Para exemplificar e avaliar o modelo híbrido proposto, foram definidos experimentos com o comitê de aprendizado e com seus os três previsores separadamente. Após a realização dos experimentos, uma análise comparativa é feita entre os resultados alcançados. O código fonte que estrutura o funcionamento do comitê implementado como *ensemble* está descrito no Apêndice A.

É importante realizar simulações separadamente com o comitê e seus previsores para provar a premissa básica de comitê de máquinas, a qual afirma que ocorre a combinação das vantagens de previsores individuais e a superação de suas desvantagens. Com a utilização de experimentos separados para o comitê e os previsores é possível analisar mais detalhadamente o comportamento de cada um e realmente identificar suas características mais particulares diante de métricas específicas para este domínio.

Para efetuar a avaliação do modelo proposto foi definido um cenário que usa informações reais do mercado acionário (obtidos a partir de requisições à plataforma *YahooFinance*) e também métricas que foram selecionadas a partir de trabalhos semelhantes, alguns comentados na seção de trabalhos relacionados, que realizaram avaliações de abordagens de predição do mercado acionário ou dentro deste contexto de investimento em bolsa de valores.

Nas próximas subseções, o cenário de uma aplicação é descrito de acordo com seus parâmetros como o valor inicial do investimento, os prazos que delimitam as cotações e a quantidade de ações envolvidas. Em seguida, as métricas usadas são explicadas conforme os motivos que levaram à sua escolha e seus resultados esperados. Por último, os experimentos são relatados através da execução do cenário aplicado às métricas permitirão a avaliação dos resultados.

### 5.3.1 Cenário

O cenário do mercado acionário elaborado leva em consideração o investimento financeiro, o prazo de aplicação e a ação negociada. Estes três parâmetros possuem o objetivo de representar uma aplicação corriqueira que é realizada na bolsa de valores por um investidor comum. Não serão discutidos detalhes mais complexos de investimento, sendo restrito a estes três parâmetros qualquer variação que possa ocorrer no cenário utilizado. Além disto, a escolha destes parâmetros simplifica o cenário por restringir o grau de variação de suas características, mas é possível aplicar outros parâmetros e consequentemente adicionar complexidade ao cenário de investimento.

É fundamental a determinação de um capital inicial para ser utilizado ao longo do investimento, definindo o valor para o primeiro parâmetro. Este valor aumenta e diminui de acordo com as operações de compra e venda das ações e o seu preço no momento da operação. Em uma aplicação real, para cada operação realizada é cobrada uma taxa de corretagem que, de maneira elementar, é um encargo financeiro referente a taxas de administração do banco, corretora ou corretor autônomo responsável por este capital investido.

O prazo de investimentos em ações variam de minutos a anos, o intervalo escolhido para este cenário, que é utilizado como segundo parâmetro, foi de quatro anos consecutivos de cotação com início em janeiro de 2009 encerrando em dezembro de 2012. Acredita-se que com este prazo seja possível analisar o comportamento da abordagem híbrida e seus previsores diante de várias situações de investimento que ocorreram ao longo deste período, são mais de 900 dias de cotações reais. Também é esperado que os dados utilizados sejam variados para envolver diferentes situações de investimento.

Finalizando os parâmetros, o terceiro é definido a partir da escolha das ações, que não precisa obedecer a rigorosos critérios econômicos sobre sua situação atual, contanto que exista a disponibilidade de acesso a suas cotações para construção do *dataset*. Mas é fundamental a utilização de pelo menos duas ações diferentes para que seja analisado o comportamento da abordagem de maneira comparativa. E para realizar uma melhor análise comparativa é fundamental a escolha de ações com características diferentes, como por exemplo diferentes grupos de classificação, pois não adiantaria analisar empresas do mesmo ramo e que talvez passem por momentos econômicos semelhantes. Para evitar que os experimentos sejam prejudicados pela escolha das ações, foi adotado o critério de seleção em duas categorias diferentes, uma ação da *blue ship* e outra *small cap*.

Com estes três parâmetros é definido um cenário de aplicação financeira pela qual a predição pode ser aplicada. Pois, além de características, o cenário possui regras que guiam sua utilização como as estratégias de investimento utilizadas nos experimentos: como os previsores indicam operações, e qual a forma de combinação o comitê utiliza. Como o resultado final tanto dos previsores quanto do comitê são estimativas de crescimento ou queda do mercado, entende-se que crescimento gera lucro e queda causa prejuízo, assim que se informado um crescimento de qualquer intensidade é sugerida uma operação de compra, e para queda é sugerida uma operação de venda. Existem momentos que os previsores não conseguem efetuar sugestões gerando uma indeterminação e ausência de qualquer tipo de operação. Esta mesma ideia ocorre com os indicadores que formam o previsor técnico e também com o comitê que na confirmação de pelo menos um dos componentes são sugeridas operações.

Quando o capital estiver investido no último dia do experimento este será retirado da aplicação e seu rendimento calculado de acordo com o preço de fechamento presente no último dia do experimento. Isto pode gerar valores tanto maiores quanto menores que o histórico de cotações, mas é realizado apenas para informar o valor final do investimento em caixa e padronizar os resultados.

### 5.3.2 Métricas

Os resultados dos experimentos realizados são avaliados através das seguintes métricas: a eficácia e a eficiência. A primeira métrica está relacionada ao rendimento alcançado durante a utilização da abordagem de predição no cenário descrito, e significa eficácia financeira. Com esta métrica, o capital inicial definido para utilização no cenário é analisado de acordo com as alterações ocorridas ao longo do período de investimento. Portanto, serão medidos valores como: o máximo atingido, o mínimo, a média, maior perda, maior ganho e o valor final do rendimento. Estes valores informam a eficácia da abordagem, e justifica sua utilização do ponto de vista financeiro caso proporcione lucro ou não para os investidores que a utilizarem.

Já a segunda métrica proporciona um aspecto importante relacionado ao quanto a predição é precisa. Este parâmetro mensura a proporção existente entre a previsão e o valor real, com relação a informações discordantes ou exatas do que foi previsto. Deste modo, a

métrica calcula uma importante proporção que permite identificar a ocorrência de falsos positivos nas operações sugeridas. Os falsos positivos acontecem quando um valor é apontado como correto, mas na verdade não é, no caso da predição quando aponta um crescimento do valor atual e ocorre uma queda, e o caso contrário. Estas falsas indicações possuem grande importância para análise final da solução, pois indica o quanto foi sugerido de maneira errônea. Sendo que quando uma falsa indicação acontece eventualmente será ocasionado um prejuízo nos investimentos que se basearam nesta indicação.

Conforme Guarnieri (2006), estas duas métricas selecionadas para analisar os resultados também foram usadas nas avaliações realizadas. Através dessas métricas, é possível mensurar o quanto uma abordagem oferece de rentabilidade para um investidor e a proporção de operações que foram geradas de maneira correta, o código fonte responsável pelo cálculo destas métricas encontra-se no Apêndice B. Com isto, são fornecidas informações necessárias para uma análise completa da aplicação de estratégias de investimentos em bolsa de valores, ao serem aplicadas em todos os experimentos, tanto aos previsores individualmente quanto no comitê. Assim, é possível verificar se o principal benefício oferecido por modelos baseado em comitê é atingido, de ser no mínimo tão bom quanto o melhor de seus previsores individuais.

A simulação da abordagem proposta em um ambiente real é fundamental para avaliar seu comportamento diante de diversas situações, normalmente realizada através de uma simulação histórica ou *backtesting*. Esta simulação é realizada com dados passados com o objetivo de reproduzir situações o mais próximas possível de um cenário real (KABASHIMA, 2005). A aplicação destas métricas durante todo o *backtesting* ocorre sobre os resultados de uma abordagem de investimento (previsores individuais ou em comitê) aplicada a uma ação. Para efeito de análise cada par (abordagem, ação) resulta em um experimento distinto e representa uma solução para o problema, e precisa ser executado e analisado separadamente para envolver o máximo de situações que forneçam informação para comparação de resultados.

### 5.3.3 Experimentos

O experimento realizado consiste do investimento em duas ações que pertencem as categorias *blue ship* e *small caps*, com os respectivos códigos: PETR4 e AMIL3. Ocorrendo durante o prazo que compreende entre 01/01/2009 à 31/12/2012, com um investimento financeiro inicial de R\$ 10.000,00 onde são realizadas diversas operações de

compra e venda de ações com o objetivo de capitalizar o investidor, sendo possível ocorrer prejuízos para estratégias mal elaboradas e lucros para estratégias mais refinadas. Uma estratégia mal elaborada pode ser caracterizada pela fraca fundamentação teórica que acarreta na escolha errônea de regras para definir os momentos que irão ocorrer as operações de compra e venda, e analogamente boas estratégias são, normalmente, bem fundamentadas sob conceitos econômicos sólidos definidos por autores e profissionais renomados.

Para cada ação os experimentos são divididos em duas etapas de execução: previsores individuais e abordagem combinada. Inicialmente serão executados os três previsores e os indicadores técnicos de maneira individual e isolada, em um segundo momento a abordagem *ensemble* será executada utilizando dois métodos de combinação diferentes(voto majoritário igualitário e ponderação de previsores), totalizando oito experimentos para cada um dos dois ativos selecionados, como ilustrado pela Tabela 3.

Tabela 3 - Experimentos realizados

Etapa	Previsor	Legenda das tabelas
Previsores individuais	Indicador ADX	ADX
	Indicador RSI	RSI
	Indicador BB	BB
	Previsor Técnico	Técnico
	Previsor Neural	Neural
	Previsor Estatístico	Estatístico
Abordagem combinada	Ensemble majoritário igualitário	Ensemble voto
	Ensemble com ponderação de previsores	Ensemble peso

O comitê foi dividido entre dois grupos segundo suas abordagem de combinação com o intuito de avaliar seus resultados a partir de dois pontos de vista distintos, primeiramente quando da atribuição de pesos aos seus previsores que ocorre após executar cada previsor individualmente e associar pesos de acordo com seus resultados analisados sob alguma métrica, e a segunda abordagem utiliza o voto majoritário igualitário que resulta em uma combinação de seus resultados sem a utilização do cálculo ponderado em cada um dos previsores.

A execução dos experimentos, para cada ação, ocorre de acordo com o seguinte roteiro: primeiramente é realizada a simulação de investimentos da ação, em seguida os dados correspondentes as operações e ao capital financeiro são coletados, e por último, as métricas serão aplicadas para calcular um conjunto de dados estatísticos sobre rentabilidade e precisão de todas as abordagens. Já o tempo de execução dos experimentos não é levado em

consideração, pois está fora do escopo do projeto. Porém, é importante destacar que a velocidade alcançada por cada resultado da predição é um fator que pode ser um diferencial para ferramentas utilizadas no domínio da bolsa de valores. E não houve um estudo aprofundado para definir qual seria o intervalo de tempo para avaliar bem a abordagem, mas foi escolhido sob a restrição de ser após o ano de 2008, pois historicamente foi um período problemático para a bolsa de valores por representar o início de uma crise mundial que representando um fator intangível difícil de ser previsto, então não é interessante usá-lo nestes experimentos.

Já o intervalo de tempo foi definido em quatro anos para avaliar a abordagem proposta, E não houve um estudo aprofundado para definir qual seria o intervalo de tempo para avaliar bem a abordagem, apenas foi escolhido sob a restrição de ser após o ano de 2008 que historicamente foi um período problemático para a bolsa de valores por representar o início de uma crise mundial, e representa um fator intangível deste mercado que é difícil de ser previsto. Então, não é interessante usar este ano nos experimentos para não gerar resultados complexos e discordantes para análise.

As duas que possuem suas ações utilizadas nos experimentos são a Petrobras e o grupo AMIL. A Petrobras é uma empresa que atua no segmento de energia, prioritariamente nas áreas de exploração, produção, refino, comercialização e transporte de petróleo e seus derivados. Para a BOVESPA, a Petrobras possui como atividade principal o petróleo, gás e energia, sendo classificada também no setor de biocombustível. O código de negociação utilizado para ativos desta empresa é o PETR4, que diferencia apenas por seu sufixo “.SA” por exigência da plataforma *YahooFinance* para extrair os valores de suas cotações e formar o dataset. A AMIL Assistência Médica Internacional é uma operadora de saúde do Brasil que ocupa posição entre as cinco maiores empresas de assistência médica privada do país. A AMIL é classificada pela BOVESPA como uma prestadora de serviços médicos hospitalares, análises e diagnósticos e seu código de negociação é o AMIL3 que representa as ações ordinárias desta empresa e sua aquisição via *YahooFinance* funciona de maneira semelhante a outra ação, adicionando o “.SA” a fim.

## 6 RESULTADOS E DISCUSSÕES

Os resultados apresentados ao longo deste capítulo são organizados em dois grupos, um para cada ação escolhida para os experimentos primeiramente os resultados da PETR4 seguidos da AMIL3. Inicialmente serão analisados sob o ponto de vista da eficácia de cada abordagem (previsores individuais e abordagem combinada) e depois da eficiência aplicada do mesmo modo. Com uma pequena diferença para os resultados da eficiência que foram divididos em uma única tabela com seu lado esquerdo apresentando os dados referentes a operações realizadas pelos previsores e o lado direito da tabela com as informações totais de sugestões, incluindo sugestões sobrepostas, por exemplo: quando dez sugestões de compra ocorrem sequencialmente, uma única operação é contabilizada do lado esquerdo, mas do lado direito são contabilizadas todas. Quanto a operações corretas são as que proporcionam lucros entre uma compra e uma venda, já as sugestões corretas são as que se mantém na mesma tendência, como sugestões de compra na tendência de baixa e de venda na de alta. Operações neutras são consideradas que não foram nem compra nem venda, caracterizando um momento de indeterminação no mercado.

Após isto, na seção seguinte, são realizadas algumas análises sobre como ocorreram os experimentos e os resultados alcançados, efetuando um paralelo entre a aplicação de previsores individuais e o ensemble projetado como abordagem híbrida de predição aplicado ao mercado acionário.

### 6.1 Considerações iniciais

#### Ações da PETR4 (Petrobras)

Analizando primeiramente a eficácia presente nos resultados dos experimentos, pode-se afirmar que ocorreu como esperado de ativos pertencentes a categoria das *blue ships*, pois a variação proporcionada pela rentabilidade do investimento durante os experimentos não alcançou valores muito altos para lucros, muito menos para prejuízos representando bem o comportamento estável de uma ação de primeira linha. O lucro máximo atingido foi de R\$

16531,47 e o maior prejuízo finalizou com R\$ 7599,47. Considerando o investimento inicial de R\$ 10.000,00 e o período operacional de quatro anos contínuos, estes valores representam um resultado muito bom para rentabilidade, pois o crescimento no capital inicial acima de 65% e uma queda máxima de apenas de 25% proporcionam grandes chances para investidores neste mercado sem agregar muito risco. Os resultados referentes aos previsores individuais, seja por indicadores técnicos unitários ou pelos previsores descritos na seção 5.2, são apresentados na Tabela 4 com os respectivos valores financeiros referentes a análise da primeira métrica, a eficácia.

Tabela 4 - Eficácia dos previsores individuais aplicados a PETR4

Previsor	Capital Final	Máximo	Mínimo
RSI	R\$ 8626,15	R\$ 13434,38	R\$ 8626,15
ADX	R\$ 9196,82	R\$ 12255,22	R\$ 7937,40
BB	R\$ 9288,26*	R\$ 13552,04	R\$ 10302,74
Previsor Estatístico	R\$ 7316,75*	R\$ 15400,76	R\$ 7599,47
Previsor Neural	R\$ 8348,11	R\$ 11221,34	R\$ 8056,34
Previsor Técnico	R\$ 8506,50	R\$ 13694,40	R\$ 8450,71

\* Ocorrência de venda forçada ao fim da simulação

Com relação aos indicadores técnicos observados na Tabela 4 o lucro acima de 20% atingido pelas BBs e sua ausência de prejuízos, apesar de não atingir valores tão altos, merecem destaque entre os indicadores técnicos, comparado aos mais de 20% de prejuízo atingido pelo ADX em seu valor mínimo. Porém, como explicado na subseção 3.2, existe uma grande flexibilidade no cálculo de indicadores técnicos, então é importante ressaltar que foi utilizado nos experimentos padrão que a literatura indica para gerar estes resultados.

Já os resultados apresentados pelos previsores foram um pouco mais lucrativos, ou eficazes, que os indicadores individualmente. O lucro superior a 50% atingido pelo previsor estatístico em seu valor máximo chama a atenção pela ótima oportunidade de lucro e em contrapartida seu prejuízo foi o maior desta primeira etapa, chegando a perdas próximas de 25% do seu capital inicial. O previsor técnico obteve o valor mais seguro entre os previsores que apresentaram algum prejuízo, com resultado aproximadamente igual ao valor médio atingido por todos os seus indicadores individuais (R\$ 8955,43), com cerca de 20% em seu valor mínimo. A seguir, Tabela 5 informa os valores obtidos na segunda etapa envolvendo as duas abordagens de combinação do ensemble.

Tabela 5 - Eficácia do *ensemble* aplicado a PETR4

Valores	Capital final	Máximo	Mínimo
<i>Ensemble</i> voto	R\$ 8215,91	R\$ 15595,25	R\$ 8215,91
<i>Ensemble</i> peso	R\$ 9552,21	R\$ 16422,72	R\$ 9218,31

\* Ocorrência de venda forçada ao fim da simulação

Como teorizado, o comitê apresentou o melhor resultado quando comparado a todas as abordagens já explicadas, se tratando de valores altos para lucro e baixos para prejuízo. O experimento de combinação que foi realizado utilizando pesos encerrou com prejuízo inferior a 5% (valor mínimo) e lucro máximo ultrapassando os 60% (valor máximo), comparando estes valores ao quantitativo patrimonial investido no início dos experimentos são valores muito bons pois agrega um pequeno risco de prejuízo e uma grande possibilidade de lucro. Estes dois valores são os melhores resultados de todas as abordagens (previsores) utilizadas no experimento, tanto o maior lucro quanto o menor prejuízo. Outra informação que merece destaque foi a melhora significativa do comitê quando utilizada a abordagem de combinação através de ponderação dos resultados dos previsores, sendo que os pesos foram calculados de acordo com os valores máximos obtidos por cada previsor individualmente.

Após os dados relacionados a eficácia das abordagens serem descritos, as mesmas, são observadas sob a perspectiva da segunda métrica utilizada para análise da abordagens, a eficiência apresentada pela predição. Seguindo o mesmo roteiro da análise da eficácia, primeiramente são estruturados os dados referentes a primeira etapa de simulação correspondente aos previsores individuais e os indicadores técnicos, como observado na Tabela 6 que está dividida entre operações realizadas e sugestões fornecidas.

Tabela 6 - Eficiência dos previsores individuais aplicados a PETR4

Previsores	Operações					Sugestões				
	Correta	%Correta	Falsa	%Falso	Total	Correto	%Correto	Errado	%Falso	Neutro
RSI	1	20%	4	80%	10	22	56%	17	44%	906
ADX	9	47%	10	53%	38	25	53%	22	47%	898
BB	3	50%	3	50%	13	30	42%	42	58%	873
Técnico	16	57%	12	43%	56	64	49%	66	51%	815
Estatístico	2	29%	5	71%	15	66	55%	55	45%	824
Neural	76	48%	82	52%	316	342	50%	342	50%	261

Nota-se que para operações realizadas (lado esquerdo da tabela), alguns previsores como o estatístico e o indicador técnico RSI obtiveram resultados muito baixos mesmo com a pequena quantidade de operações realizadas, denotando uma opção muito pouco eficiente

para predição no mercado acionário. Diferente da RNA que ficou próxima de 50% e possuiu a maior quantidade de operações dos previsores individuais com cerca de 150 operações. Apenas o previsor técnico superou a quantidade de operações corretas sobre as erradas, mas não forneceu uma quantidade muito alta de operações, com menos de 30 operações. Com relação as sugestões (lado direito da tabela), todas as abordagens ficaram próximas de 50% de operações corretas destacando-se o indicador RSI e o previsor estatístico com os melhores resultados. E um destaque ainda maior às redes neurais que permitiram uma grande quantidade de sugestões e mantiveram seu percentual de acerto na mesma faixa dos demais.

A seguir são apresentados na Tabela 7 os resultados relacionados a segunda etapa de experimentos com a eficiência de utilização do *ensemble* a partir de duas abordagens de combinação.

Tabela 7 - Eficiência do *ensemble* aplicado a PETR4

Eficiência	Operações					Sugestões				
	Correta	%Correta	Falsa	%Falso	Total	Correto	%Correto	Errado	%Falso	Neutro
Previsores										
Voto	76	47%	87	53%	325	460	49%	472	51%	13
Peso	83	49%	86	51%	337	459	49%	473	51%	13

Nos resultados do ensemble é visto que a quantidade de operações realizadas manteve um percentual próximo de 50%, assim como a de sugestões oferecidas, ficando na média das abordagens individuais. Porém, a quantidade de operações tanto realizadas quanto sugeridas aumentou com a combinação, o que permite uma exploração maior das cotações da ação utilizadas, justificando os maiores lucros atingidos pelo comitê analisados pela outra métrica (eficácia).

### Ações da AMIL3 (AMIL)

Os experimentos esclareceram que ações de segunda linha, realmente possuem um comportamento muito diferente da outra categoria, com oscilações mais acentuadas. E um fator que chamou a atenção por alcançar valores muito altos, foi o lucro máximo utilizando indicadores técnicos que permitiram atingir capital máximo de R\$ 33.092,30.

Os resultados referentes a primeira etapa da avaliação da eficácia são informados na Tabela 8, que apresenta os valores financeiros atingidos pelos indicadores técnicos individualmente e os previsores do comitê.

Tabela 8 - Eficácia dos previsores individuais aplicados a AMIL3

	Capital Final	Máximo	Mínimo
RSI	R\$ 19.783,62	R\$ 18.661,37	R\$ 16.651,69
ADX	R\$ 33.146,02	R\$ 33.092,30	R\$ 9.720,28
BB	R\$ 25.679,79	R\$ 25.679,79	R\$ 13.416,67
Estatístico	R\$ 24040,03*	R\$ 15.977,06	R\$ 9.324,32
Neural	R\$ 11.794,15	R\$ 11.988,76	R\$ 8.329,01
Técnico	R\$ 26.973,46	R\$ 26.955,98	R\$ 9.720,28

\* ocorrência de venda forçada ao fim da simulação

Como observado, o ADX apresentou a maior diferença entre lucro e prejuízo, com seu valor máximo superior a 230% e o pequeno prejuízo inferior a 3%. Exceto pelo RSI os outros indicadores superaram 100% de lucro máximo ao longo dos experimentos, demonstrando uma excelente forma de operar com *small caps* a partir apenas do uso de indicadores técnicos. Com relação aos previsores, o que teve o pior resultado foi o neural, com o lucro máximo próximo de 20% e prejuízos que alcançaram a marca de 15%. Já o estatístico e técnico alcançaram lucros superiores a 50% e prejuízos inferiores a 10%, com valores impressionantes para o lucro atingido pelo previsor técnico que superou a média de lucros (R\$ 25811,15) obtida pelos seus previsores (indicadores) que o formam. A combinação dos previsores anteriormente expostos forma o resultado final da abordagem utilizando *ensemble*, como exposto no Tabela 9.

Tabela 9 - Eficácia do *ensemble* aplicado a AMIL3

	Capital Final	Máximo	Mínimo
<i>Ensemble</i> voto	R\$ 22.443,22*	R\$ 22950,44	R\$ 9789,70
<i>Ensemble</i> peso	R\$ 26794,95*	R\$ 26768,88	R\$ 9789,70

Da mesma forma como em ações de primeira linha, a utilização do comitê apresenta resultados tão bons quanto de suas melhores abordagens individuais, com um pequeno detalhe do previsor técnico que atingiu um valor ligeiramente superior. Com isto, o comitê também ultrapassou a margem de lucro de 160% do que foi investido inicialmente e seus prejuízos foram mínimos, inferiores a 10%.

A segunda métrica utilizada retrata informações bem diferentes da outra categoria de ação, onde apenas um previsor superou a quantidade média de operações corretas. A Tabela 10, ilustra os resultados calculados para análise de eficiência dos previsores individuais.

Tabela 10 - Eficiência da abordagem aplicada a AMIL3

AMIL	Operações					Sugestões				
	Previsores	Correta	%Correta	Falsa	%Falso	Total	Correto	%Correto	Errado	%Falso
RSI	2	67%	1	33%	7	31	61%	20	39%	960
ADX	12	67%	6	33%	37	18	41%	26	59%	967
BB	5	83%	1	17%	12	35	52%	32	48%	944
Técnico	19	70%	8	30%	55	73	51%	69	49%	869
Neural	84	48%	92	52%	353	367	52%	338	48%	306
Estatístico	12	67%	6	33%	37	80	49%	83	51%	848

Quanto a eficiência, muitos dos previsores superaram a proporção existente entre operações corretas pelas erradas. A não ser o neural que obteve um percentual de operações corretas ligeiramente inferior ao das erradas, porém acumulou a maior quantidade de operações realizadas pelos previsores maior que todos os outros previsores juntos. O mesmo ocorreu para as sugestões, assim como as ações de primeira linha, boa parte dos previsores não gerou tantas sugestões como o neural. Por fim, a Tabela 11 apresenta os resultados da segunda etapa dos experimentos com dados relativos a eficiência do ensemble utilizando as duas abordagens de combinação.

Tabela 11- Eficiência do *ensemble* aplicada a AMIL3

AMIL	Operações					Sugestões				
	Previsores	Correta	%Correta	Falsa	%Falso	Total	Correto	%Correto	Errado	%Falso
Voto	88	50%	88	50%	352	476	51%	449	49%	86
Peso	93	51%	90	49%	366	473	51%	452	49%	86

Ambos os parâmetros, quantidade de operações e sugestões, obtiveram percentuais corretos superiores aos errados. Mas o importante, foi atingir este valor positivo com uma quantidade de operações e sugestão de operações acima de qualquer outras abordagem de previsores individuais citada anteriormente.

## 6.2 Análise dos resultados

Conforme os resultados anteriormente obtidos através das técnicas (ADX, RSI, BBs, previsor técnico, previsor neural, previsor estatístico, comitê por votação e comitê ponderado) aplicadas aos dois *datasets* (formados pelas ações de primeira e segunda linha) constatou que uma mesma abordagem aplicada aos dois tipos de ações, por exemplo o previsor neural aplicado na predição da PETR4 gera resultados muito divergentes quando aplicado à AMIL3, diante das duas métricas (eficácia financeira e eficiência de precisão). Portanto, é possível modelar uma solução adequada e ajustada para algum domínio deste mercado, como por exemplo investimentos em ações de primeira linha.

Os previsores implementados apresentaram crescimento no capital inicial através das operações sugeridas, indicando boas oportunidades de investimento . Em todas as experiências houveram crescimento, mesmo que mínimo no valor do capital inicial dos experimentos, como o caso do previsor neural que obteve um pequeno crescimento de 10% apresentado na Tabela 4.

Ainda que na maioria dos casos tenha ocorrido prejuízos (valor mínimo nas tabelas da seção 5.3.3) em algum momento do período de experimento foram atingidos resultados avaliativos satisfatórios, pois o objetivo é avaliar cada técnica até onde alcançarem valores máximos e mínimos. Isto é, a ocorrência de prejuízos é importante para ilustrar as possibilidades de perdas que cada técnica possui, mas é comum em qualquer tipo de investimento a configuração de uma forma de interrupção das operações para evitar prejuízos (operação de *stop*), por exemplo: se os prejuízos ultrapassarem 10% encerre as operações e resgate o valor.

Cada técnica empregada representa uma solução para o problema de predição envolvendo os *datasets* criados. Visto que são previsores de um *ensemble* que exige a solução completa para o problema por cada elemento que o compõe. A partir destas técnicas é possível extrair estratégias de investimento individual de acordo com interpretações dos seus resultados:

- De maneira geral, os previsores calculam valores estimados de crescimento (lucro) e de acordo com este parecer pode ser estabelecido níveis de risco associado à estimativa de crescimento ou queda da ação prevista;
- Associando os níveis de risco e a estimativa calculada é definido o perfil de investidores, que podem ser desde os conservadores (os que exigem estimativas altas de crescimento para realizar seus investimentos) até os mais agressivos (que são mais corajosos e investem a partir de pequenas oportunidades e quase nenhuma confirmação de lucro);

- A diferença entre o lucro máximo e mínimo obtido durante as simulações estabelece o grau de confiança naquela abordagem, visto que, quanto maior a diferença entre o lucro máximo e mínimo mais arriscada é sua utilização, mas quando esta diferença é menor sua utilização é mais segura.
- O resultado do previsor estatístico que foi elaborado a partir de uma grande base de dados de cotações também informa a intensidade deste movimento de crescimento ou de queda, a partir dos padrões *candlestick*;
- A RNA é desenvolvida calcula o preço do dia seguinte, o qual pode ser usado para identificar uma tendência de curtíssimo prazo entre o preço atual e o previsto ou de curso prazo entre o primeiro preço de sua entrada (1º dia) e o preço previsto (11º dia), mesmo que estas tendências sejam meramente interpretações da predição;
- O previsor técnico oferece uma correta identificação da tendência e pontos de reversão através de conceitos econômicos, podendo ser utilizada para efetuar investimentos com maior amplitude (diferença entre o preço na operação compra comparado quando ocorreu a venda).

Para definir as operações durante os experimentos a seguinte estratégia foi adotada: a indicação de uma operação de compra é gerada quando ocorre um sinal de crescimento dos preços, de venda quando for de queda. Para calcular o resultado dos previsores durante as simulações existem as seguintes regras:

- Resultado maior que zero gera operação de compra;
- Resultado menor que zero ocasiona em operação de venda;
- Resultado zerado resulta em uma indeterminação, sem gerar operação.

Definir uma estratégia com base em qualquer sinal de crescimento ou queda acaba gerando problemas para os previsores, além das próprias limitações e desvantagens relacionadas a sua abordagem. Dentre as principais desvantagens inerentes aos próprios previsores, têm-se:

- Nem sempre o mercado apresenta padrões recorrentes, principalmente de *candlesticks*, e consequentemente não são sugeridas operações de investimento, isto prejudica o funcionamento do previsor estatístico que depende da ocorrência de padrões no mercado;
- O previsor neural é o principal culpado pelo número excessivo de operações geradas, devido a estratégia de geração de operações baseada em qualquer

crescimento ou queda, reduzindo a eficiência da abordagem híbrida durante os experimentos;

- A forma de combinação dos resultados dos indicadores para formar o previsor técnico, utilizando combinação ponderada igualitária, reduziu seu desempenho para as duas ações usadas nos experimentos.

Apesar das limitações supracitadas o comitê desenvolvido apresentou resultados esperados de uma abordagem de combinação através de *ensemble*, ao atingir resultados pelo menos tão bons quanto seu melhor previsor diante de ação tanto de primeira quanto de segunda linha. Isto demonstra que a abordagem híbrida obteve **eficácia** por gerar rendimentos financeiros altos em relação ao valor inicial do experimento. Por outro lado, a combinação dos resultados dos previsores ocasionou em uma quantidade excessiva de operações desnecessárias conhecidas como falso-positivo, o que diminuiu a expectativa de rentabilidade em comparação aos previsores individuais e causou o baixo valor de **eficiência** informado pela segunda métrica. Portanto, diante das duas métricas utilizadas para avaliar a abordagem híbrida de predição é possível afirmar:

- É muito eficaz por proporcionar alta rentabilidade, quando comparada aos resultados individuais;
- É pouco eficiente porque grande parte das operações geradas são falso-positivos.

Os resultados alcançados pelo comitê analisados pela perspectiva de cada métrica confirma a hipótese inicial que as abordagens dos previsores utilizados são complementares. Isto ocorre pois o comitê conseguiu combinar os resultados parciais de diferentes metodologias de investimento e manteve os resultados superiores ao de seus previsores, mantendo um alto nível de eficácia financeira. Já a precisão dos resultados ficou um pouco comprometida pela quantidade alta de operações geradas pelo previsor neural.

A utilização do previsor neural requer um *dataset* que inclua os dados referentes aos treinamentos . Para efeito de comparação, nos experimentos, foi usado o ano de 2008. No entanto, os valores de cotações desse ano, representam um período complicado para investimentos em ações devido à crise financeira mundial ocorrida. Assim, os experimentos reduziram o desempenho de todas as abordagens exceto as redes neurais. Anos de crise são muito importantes, pois encontram-se no grupo de fatores intangíveis que podem atrapalhar ou impossibilitar estratégias de predição, e isto prova que as RNAs são realmente aproximadores universais, apenas com a ressalva de ser treinada com um conjunto de dados representativo e diversificado.

Também observou-se que a crise de 2008 influenciou negativamente o resultado do comitê desenvolvido, uma vez que os resultados ruins de seus previsores foram a causa do resultado ruim do *ensemble*. É importante ressaltar que quando um dos previsores possui um resultado muito ruim em relação aos demais, as informações passadas para o comitê via combinação de resultados dos previsores pode acabar prejudicando o resultado final da combinação, sendo necessária a utilização da etapa de seleção de previsores para eliminar este tipo de problema.

Durante os experimentos notou-se que as RNAs obtiveram resultados ruins, ao contrário do que a literatura comenta. Então, após analisar o problema foi percebido que sua abordagem ficou em desvantagem por necessitar de um treinamento prévio para ser utilizada o que retirou um dos anos de cotação do seu *dataset*. E para avaliar seu comportamento durante todo o período de investimentos seu *dataset* foi aumentado em um ano para realização dos treinamentos com as cotações do ano de 2008. Com esta outra simulação também sendo executada foi permitido à RNA operar durante os quatro anos consecutivos de investimento, e com o novo prazo o desempenho do previsor neural cresceu de maneira impressionante, encerrando o experimento com lucro de aproximadamente 50% e prejuízos mínimos inferiores a 10% para a ação de primeira linha e superando os 200% de lucro máximo, também melhorando seu prejuízo para menos que 10% com a ação de segunda linha. O acréscimo do ano de 2008 tornou a RNA a abordagem que gera o menor prejuízo e um dos maiores lucros até este momento. Após esta melhora substancial da rede neural novos experimentos foram realizados diminuindo em um ano o prazo inicial, mas todos os previsores sofreram uma queda assustadora em seus rendimentos, chegando a prejuízos superiores a 50% e tornando inviável do ponto de vista financeiro a utilização desta abordagem na prática.

No contexto do comitê, foi notado que a alteração da estratégia de combinação é uma opção importante permitindo uma flexibilidade no resultado final apresentado. E a consideração de valores ponderados de acordo com a rentabilidade de cada previsor alcançou valores superiores quando comparados a utilização de voto majoritário com as duas ações. Sendo assim, dentro do domínio de predição, o comitê projetado neste trabalho apresentou valores muito bons tanto para possibilidade de lucro quanto prejuízo, potencializando o lucro e minimizando as chances de prejuízo, ao combinar vários previsores multidisciplinares em uma abordagem híbrida de investimento.

## 7 CONCLUSÃO

Este trabalho apresenta um modelo híbrido para o problema de predição no mercado acionário. Este mercado apresenta-se como um ambiente bem complexo devido aos inúmeros fatores e variáveis que o influenciam. Neste contexto, o modelo híbrido permite combinar as vantagens das técnicas de estimativas de padrões *candlestick*, antecipação de preços futuros a partir de redes neurais e definição da tendência atual com indicadores técnicos inseridos em três previsores de um comitê de aprendizado. Além disto, um benefício fundamental é a possibilidade de superar as limitações de cada técnica quando estas são utilizadas de maneira isolada por intermédio da combinação de seus resultados pelo comitê.

Os experimentos que foram realizados de forma individual, apresentaram as seguintes desvantagens: geração de operações desnecessárias pela grande quantidade de operações e sugestões, ou pela ausência de operação devido a momentos de indeterminação apresentada pelo previsor. Quando isto ocorre o comitê refina os resultados estabilizando a quantidade de operações e reduzindo os períodos inoperantes através de sua estratégia de combinação das operações sugeridas e realizadas pelos previsores.

Ao avaliar todos os experimentos realizados sob a perspectiva das métricas eficácia e eficiência, foi percebido que a abordagem proposta fornece boas oportunidades de investimento no mercado acionário utilizando a predição de séries temporais. Além disto, foi observado o comportamento diferente da abordagem diante de ações de primeira e segunda linha, confirmando o que a literatura informa sobre a estabilidade de ações de primeira e a alta rentabilidade proporcionada pelas oscilações das de segunda linha.

Outro marco importante deste trabalho foi verificar o comportamento das RNAs no período de 2008 em que todos os previsores, inclusive o comitê, apresentaram resultados muito insatisfatórios, enquanto a rede se comportou muito bem. Este acontecimento reforça a utilização de RNAs como aproximador universal, inclusive no âmbito de séries temporais e ressalta a ocorrida crise mundial neste mesmo ano. Por outro lado, é necessário aplicar uma estratégia mais branda para geração de suas operações, diferente da sinalização a partir de crescimento ou queda do mercado para sugerir compra ou venda de ações, pois gerou uma

quantidade desnecessária de operações que acabou prejudicando a rede e o comitê na avaliação por uma das métricas.

## 7.1 Trabalhos Futuros

Ao longo do desenvolvimento do trabalho e diante dos resultados alcançados, algumas ideias de trabalhos futuros surgiram para acrescentar contribuições à utilização de comitês de máquinas no âmbito da predição de séries temporais financeiras. Primeiramente, é importante realizar os mesmos experimentos de maneira mais numerosa, pois os realizados neste trabalho tiveram caráter comparativo, com o objetivo de analisar a abordagem híbrida proposta diante de dois pontos de vista (ações de primeira e segunda linha). Para que isto ocorra o *dataset* deve ser expandido para um número maior de anos, segundo a literatura o processo de *backtesting* pode ocorrer com os valores de cotação desde o ano que a ação entrou na bolsa até o momento atual.

Outra adição importante aos experimentos pode ser feita com um acréscimo no número de ações avaliadas, podendo manter o caráter comparativo através da inclusão de novas ações direcionadas por critérios de análise, por exemplo: incluir ações de empresas do ramo hospitalar, ou *small caps* mais recentes e mais antigas, ou a ação mais recente na bolsa e a mais antiga, dentre outros. O importante neste novo grupo de experimentos é aumentar o número de ações avaliadas, mas manter critérios que possam ser avaliadas de maneira comparativa.

Com relação as avaliações, é possível complementar o trabalho através da inclusão de mais critérios de avaliação além das métricas usadas, eficácia e eficiência, sendo uma boa oportunidade para iniciar a utilização de indicativos de desempenho pois operações financeiras são sensíveis ao tempo e não foi priorizada a avaliação do desempenho do modelo híbrido proposto. Sobre tempo utilizados nos *candlesticks* e séries temporais, é importante diminuir sua granularidade que atualmente define as séries e o *dataset* em dias, com a predição direcionada para o dia seguinte. Sendo assim, ao adicionando este critério desempenho temporal é interessante refletir sobre a alteração da predição para horas ou minutos, reduzindo o valor atual em um dia.

Como foi comprovada a utilização dos três previsores em um *ensemble* também é interessante comparar este *ensemble* com variações em sua estrutura, a partir da alteração da disposição dos seus previsores ou da metodologia de combinação. Com isto, poderia se pensar em acrescentar outros previsores ou variar os previsores existentes como por exemplo: a

criação de um novo previsor neural utilizando outra abordagem de RNAs e algoritmo de treinamento, ou outro previsor técnico com novos indicadores ou combinação dos existentes. Existe uma grande quantidade de possibilidades para variação da configuração da RNA utilizada no previsor neural, o que permite validar o modelo proposto sob diversas perspectivas não abordadas neste trabalho e, por exemplo, melhorar seu desempenho.

Por fim, os resultados alcançados pelo *ensemble* possuem grande importância para consolidar o modelo de predição proposto como uma importante ferramenta de predição por demonstrarem o bom funcionamento de uma abordagem híbrida, além de incitar a escrita de um artigo científico destacando uma grande oportunidade de atingir uma segunda publicação (a primeira está descrita na próxima seção) relacionada a predição no mercado acionário enfatizando os resultados altançados pelo *ensembles*.

## 7.2 Publicações

A partir do modelo de predição baseado em uma abordagem híbrida proposto neste trabalho foi elaborado um artigo descrevendo detalhadamente a metodologia de predição utilizando o *ensemble* com os três previsores (estatístico, neural e técnico) publicado no 10<sup>th</sup> CONTECSI – International Conference on Information Systems and Technology Management, com título de “A hybrid model for prediction the behavior of stock market” (Um modelo híbrido de predição do comportamento da bolsa de valores).

## REFERÊNCIAS

AIKES JUNIOR, J. et al. "Estudo da Influência de diversas medidas de similaridade na previsão de séries temporais utilizando o algoritmo KNN-TSP.".In: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL, 12, 2012. *Anais...* Curitiba: BRASICS, 2012.

BALESTRASSI, P. P.; SAMOHYL, R. W.; MEURER, R. Métodos de previsão usando a série econométrica de Nelson Plosser: um estudo comparativo. In: Encontro Nacional de Engenharia de Produção, 18, 1998. Niterói. *Anais...*Universidade Federal Fluminense, 1998.

BAUER, D. P. **Aplicação de redes neurais artificiais na determinação do preço de ações.** Centro Universitário Feevale, Instituto de Ciências. Novo Hamburgo, RS: Instituto de Ciências e Tecnologia, 2008.

BARRETO, J. M. "**Introdução às redes neurais artificiais.**" **Laboratório de Conexionismo e Ciências Cognitivas.** Departamento de Informática e de Estatística. Florianópolis: UFSC, 2002.

BM&FBOVESPA. Introdução ao mercado de capitais, 2010. Disponível em: em: <http://www.bmfbovespa.com.br/pt-br/a-bmfbovespa/download/merccap.pdf>. Acesso em: 28 jan. 2013.

BRAGA, C. V. **Rede Neural e Regressão Linear: comparativo entre as técnicas aplicadas a um caso prático na receita federal.** Dissertação (Mestrado Profissionalizante em Administração) - Pós-Graduação e Pesquisa em Administração e Economia. Rio de Janeiro: Faculdades IBMEC, 2010.

BRUNI, A. L. et al. "Análise técnica de ações: um estudo de caso da empresa Bombril SA." **Revista de Contabilidade da UFBA** v.3, n.3, ago., 2008, p. 66-80, 2010.

CAVALCANTE, F. **Mercado de Capitais, o que é e como funciona.** In: \_\_\_\_\_. 7 ed.Rio de Janeiro: Editora Campus, 2005.

CORTEZ, P; NEVES, J. C. F. M. **Modelos inspirados na natureza para a previsão de séries temporais.** Braga, Portugal: Universidade do Minho, 2002.

**COSTA, V. T. Análise de pré-procesamento no desempenho de previsão de Séries Temporais Financeiras.** Trabalho de Conclusão de Curso ( Graduação em Ciência da Computação) - Universidade Federal de Pernambuco, Recife, 2008.

COSTA, F. V. da; MARCONDES, R. C. Recursos estratégicos em corretoras de valores mobiliários visando a busca de vantagem competitiva: pela abordagem da RBV. In: SEMINÁRIOS EM ADMINISTRAÇÃO – SEMEAD, 15, out. 2012.

CHITRA , et al. An Ensemble Model of Multiple Classifiers for time series prediction international. **Journal of Computer Theory and Engineering**, v. 2, n. 3, jun., 2010.

**CUPERTINO, L.F. Modelagem do módulo de Young em nanocompósitos através de Inteligência Computacional.** 2009. 82f. Dissertação (Mestrado em Engenharia Elétrica) - Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 2009.

DEBASTIANI, C. A. **Candlestick:** um método para ampliar lucros na Bolsa de Valores. São Paulo: Novatec, 2007.

FARIA, E. L. **Previsão de séries temporais utilizando métodos estatísticos.** CBPF 003 - Centro Brasileiro de Pesquisas Físicas - Novas Técnicas, PUC-RJ, Rio de Janeiro, 2008a.

\_\_\_\_\_. **Previsão do mercado de ações brasileiro utilizando redes neurais artificiais.** CBPF - Centro Brasileiro de Pesquisas Físicas - Novas Técnicas, PUC-RJ, Rio de Janeiro, 2008b.

GOUVEIA, M. C. **Estudo e implementação de um algoritmo de localização baseado na correspondência de mapas.** 82f.Dissertação (Mestrado Integrado - Engenharia Electrotécnica e de Computadores). Universidade do Porto, 2009.

GUARNIERI, O. C. **Um estudo empírico da eficiência da análise técnica como instrumento na predição do comportamento dos preços das ações:** o caso Embraer. 121f.Dissertação (Mestrado em Gestão e Desenvolvimento Regional) - Departamento de Economia, Contabilidade e Administração, Universidade de Taubaté, 2006.

\_\_\_\_\_. O impacto do custo de corretagem nas operações com utilização da média móvel na tomada de decisão em investimento: o caso da EMBRAER. In: CONGRESSO USP INICIAÇÃO CIENTÍFICA EM CONTABILIDADE, 3., 2006, São Paulo. **Anais ...São Paulo**, 2006.

GUPTA, K.; SUNDARAM, S.; SHAMSI, F. Dara Warehousing and Mining: Prediction Market Direction, 2006. Disponível em: <http://cs.nyu.edu/~sss374/datamine.pdf>. Acesso em: 25 jan. 2013.

HAYKIN, S. **Redes Neurais**: princípios e práticas, 2. ed. Porto Alegre: BOOKMAN, 2000.

HILGERT, S. P. **Educação**: aprenda a investir na bolsa de valores com ênfase em análise técnica. São Paulo: XP Educação, 2011.

KABASHIMA, E. M. "Pairs Trading: Aplicação no mercado acionário brasileiro." 56f. Dissertação (Mestrado em Finanças e Economia Empresarial). - Escola de Economia, 2005.

KANNAN, K. S.; SEKAR, P. S.; SATHIK, M. M. Financial Stock Market Forecast using Data Mining Techniques. Hong Kong: In: CONFERENCE OF ENGINEERS AND COMPUTER SCIENTIST, Proceedings..., 2010.

KIM, M.J.; KANG, D.K. Ensemble with neural networks for bankruptcy prediction, Elsevier **Expert Systems with Applications**, v. 37, p. 3373–3379, april, 2010.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE. **Anais**..., 1995. v. 14, p. 1137–1145.

KUNCHEVA, L. I. Classifier ensembles for changing environments. 5th Interantional Workshop on Multiple Classifier Systems, **Lecture Notes in Computer Science**, v. 3077, p. 1-15, 2004.

LEVINSON, M. **Guide to Financial Markets**. 4. ed. Londres: The Economist, 2005.

LUEDERS, A. **Investindo em small caps**: um roteiro completo para se tornar um investidor de sucesso. São Paulo, Campus: Elsevier, 2008.

LIMA, C. A. de M. "Comitê de Máquinas: uma abordagem unificada empregando máquinas de vetores-suporte." Tese (Doutorado em Engenharia Elétrica) – Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. Campinas, 2004. MALETZKE, A. G. "Uma metodologia para extração de conhecimento em séries temporais por meio da identificação de motifs e da extração de características." Dissertação ( Mestrado em Ciências de Computação e Matemática Computacional) – Universidade de São Paulo – ICMC-USP, 2009.

MARANGONI, P. H. **Redes neurais artificiais para previsão de séries temporais no mercado acionário.** Trabalho de Conclusão de Curso (Graduação em Ciências Econômicas) - Universidade Federal de Santa Catarina, 2010.

MARQUES, F. C.; GOMES, R. M. **Análise de séries temporais aplicadas ao mercado financeiro com o uso de algoritmos genéticos e lógica nebulosa.** Belo Horizonte: CEFET-MG, 2009.

MORRIS, G. L. **Candlestick Charting Explained** - timeless techniques for trading stocks and futures. 3. ed. Nova Iorque: McGraw-Hill, 2006.

NIKOOA, H., et al. "Using a trainable neural network ensemble for trend prediction of Tehran stock exchange." **IJCSNS**, v. 7, n.12, dec., 2007.

NORONHA, M. **Análise técnica:** teorias, ferramentas e estratégias. Rio de Janeiro: Editec, 1995.

OLIVEIRA, P. C. Séries Temporais: analisar o passado, predizer o futuro. **Documento Técnico.** Departamento de Engenharia Informática, Coimbra: Universidade de Coimbra, 2007.

PASQUOTTO, J. L. D. **Previsão de séries temporais no varejo brasileiro:** uma investigação comparativa da aplicação de redes neurais recorrentes de Elman. Dissertação. (Mestrado em Administração) - Faculdade de Economia Administração e Contabilidade da USP, 2011.

PARRACHO, P. M. **PATTERN: Identificação de padrões em mercados bolsistas.** Dissertação (Mestrado em Engenharia Informática e de Computadores) - Universidade Técnica de Lisboa - Instituto Técnico Superior, Lisboa, 2010.

PAULA, A. de; TRICHES, D. O mercado acionário e o desempenho dos ativos financeiros no Brasil com emprego da análise técnica. **Perspectiva Econômica** , v.6, n.1, p. 37- 62, jan./jun. 2010.

PETROKAS, L. A; FAMÁ, R.; MODRO, W. M. A utilização da análise técnica contribui para o aumento da riqueza dos investidores.In: SEMINÁRIOS EM ADMINISTRAÇÃO – SEMEAD, 13, 2010.

PETRY, A. C. **Protótipo para previsão do mercado de ações utilizando Bandas de Bollinger.** Trabalho de Conclusão de Curso (Graduação em Ciência da Computação) - Universidade Regional de Blumenau, 2010.

PINHEIRO, J. L. **Mercado de capitais:** fundamentos e técnicas. 3 ed. São Paulo: Atlas, 2005.

PROCHNOW, F. A. "Programação genética para predição de séries temporais aplicados a mercados financeiros." Trabalho de Conclusão (Graduação em Ciência da Computação) - Universidade Federal do Rio Grande do Sul, 2013.

REIS, C. E. dos; TRICHES, D. Seleção e composição de uma carteira de ações com base na técnica grafista. **Perspectiva Econômica**, v.3, n.1, p. 1 - 26, jan./jun.2007.

RIBEIRO, C.; GOLDSHIMIDT; R.; CHOREN, R. **Métodos para previsão de séries temporais e suas tendências de desenvolvimento.** Rio de Janeiro: Instituto Militar de Engenharia - IME, 2009.

ROCHA, H. R.; MACEDO, M. A. Previsão do Preço de Ações Usando Redes Neurais. In: CONGRESSO USP DE INICIAÇÃO CIENTÍFICA EM CONTABILIDADE, 8. São Paulo, 2011.

ROJO, C. A.; SOUSA, A. F. de. O mercado brasileiro de small caps: perspectivas de stakeholders institucionais. **Ciências Sociais em Perspectiva**, v. 11, p. 1-29, 2012.

ROQUE, R. C. **Estudo sobre a empregabilidade da previsão do índice BOVESPA usando redes neurais artificiais.** Trabalho de Conclusão de Curso (Graduação em Engenharia Eletrônica e de Computação) - Universidade Federal do Rio de Janeiro, 2009.

SOARES, A. S. **Predição de séries temporais econômicas por meio de Redes neurais artificiais e transformada de Wavelet:** combinando modelo técnico e fundamentalista. 98f Dissertação (Mestrado em Engenharia Elétrica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2008.

SOUZA, R. C. **Previsão de séries temporais utilizando rede neural treinada por filtro de kalman e evolução diferencial.** 82f Dissertação (Mestrado em Engenharia de Produção e Sistemas) - Programa de Pós-Graduação em Engenharia de Produção e Sistemas, Pontifícia Universidade Católica do Paraná, Curitiba, 2008.

SOUZA, I. M. M. **Um estudo comparativo para previsão da cotação de ações da BM&FBOVESPA utilizando redes neurais artificiais.** Trabalho de Conclusão de Curso. (Graduação em Engenharia da Computação) - Escola Politécnica - Universidade de Pernambuco, Engenharia da Computação, 2012.

SOUZA, L. V. de. **Programação genética e combinação de preditores para previsão de séries temporais.** 145f. Tese (Doutorado em Ciências) - Pós Graduação em Métodos Numéricos em Engenharia – Programação Matemática - Universidade Federal do Paraná. 2006.

VEIGA, C. P. da, et al. "Impacto financeiro dos erros na previsão empresarial: um estudo comparativo entre modelos lineares e redes neurais." **Revista Produção Online**, v. 12, n.3, p. 629-656. jul./set. 2012.

VIDOTTO, R. S.; MIGLIATO, A. L. T; ZAMBON, A. C. O Moving Average Convergence-Divergence como ferramenta para a decisão de investimentos no mercado de ações. **Revista de Administração Contemporânea**, v.13, n.2, p. 291-309, abr./jun. 2009.

VILLANUEVA, W. J. **Comitê de máquinas em predição de séries temporais.** 178f. Dissertação (Mestrado em Engenharia Elétrica) - Pós-graduação da Faculdade de Engenharia Elétrica e de Computação, Campinas: Universidade Estadual de Campinas, 2006.

VORA, M. N. "Genetic Algorithm for Trading Signal Generation." 2010 International Conference on Business and Economics. In: **INTERNATIONAL CONFERENCE ON BUSINESS AND ECONOMICS**, 2010.

XU, L.; KRZYZAK, A.; SUEN, C. Y. Methods of combining multiple classifiers and their applications to handwriting recognition. **IEEE Transactions on Systems, Man and Cybernetics**, v. 22, n. 3, p. 418-435, 1992.

WERBOS, P. J. **Beyond Regression:** new tools for prediction and analysis in the behavioral sciences. PhD thesis, Harvard University, 1974.

WOODS, K. et al. Combination of multiple classifiers using local accuracy estimates. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 19, n. 4, p. 405-410, 1997.

## APÊNDICES

## Apêndice A – CÓDIGO SIMPLIFICADO DA ABORDAGEM DO *ENSEMBLE*

```

public class Committee {

    TimeSerie pattern;
    TimeSerie neural;
    TimeSerie technical;

    public TimeSerie decision;
    public TimeSerie result;

    public int combineStrategy;
    public static double patternWeight=1;
    public static double neuralWeight=1;
    public static double technicalWeight=1;

    public String log;

    public static final int MAJORITY_VOTE=0;
    public static final int WEIGHTED_AVERAGE=1;

    public Committee(int strategy, TimeSerie p, TimeSerie n, TimeSerie
t, TimeSerie r) {
        pattern = p;
        neural = n;
        technical = t;

        result = r;

        decision = new TimeSerie();

        log = "";

        combineStrategy = strategy;

        validate();

        evaluate();
    }

    private void validate() {

    }

    private void evaluate() {
        Iterator<TimeSerieItem> itPattern = pattern.iterator();
        Iterator<TimeSerieItem> itNeural = neural.iterator();
        Iterator<TimeSerieItem> itTechnical = technical.iterator();

        Iterator<TimeSerieItem> itResult = result.iterator();

        log += "Date \t Price \t Pattern \t Neural \t Technical \t
Decision \n";

        while(itPattern.hasNext()){

            TimeSerieItem tsiPattern = itPattern.next();
            TimeSerieItem tsiNeural = itNeural.next();
            TimeSerieItem tsiTechnical = itTechnical.next();
        }
    }
}

```

```

        TimeSerieItem tsiResult = itResult.next();

        double count = 0;

        if (combineStrategy == Committee.MAJORITY_VOTE) {

            if (tsiPattern.getValue() == TraderDecisionConstants.BUY) {
                count++;
            }
            else
                if (tsiPattern.getValue() == TraderDecisionConstants.SELL) {
                    count--;
                }

            if (tsiNeural.getValue() == TraderDecisionConstants.BUY) {
                count++;
            }
            else
                if (tsiNeural.getValue() == TraderDecisionConstants.SELL) {
                    count--;
                }

            if (tsiTechnical.getValue() == TraderDecisionConstants.BUY) {
                count++;
            }
            else
                if (tsiTechnical.getValue() == TraderDecisionConstants.SELL) {
                    count--;
                }
            }

        if (combineStrategy == Committee.WEIGHTED_AVERAGE) {

            if (tsiPattern.getValue() == TraderDecisionConstants.BUY) {
                count += 1 * patternWeight;
            }
            else
                if (tsiPattern.getValue() == TraderDecisionConstants.SELL) {
                    count -= 1 * patternWeight;
                }

            if (tsiNeural.getValue() == TraderDecisionConstants.BUY) {
                count += 1 * neuralWeight;
            }
            else
                if (tsiNeural.getValue() == TraderDecisionConstants.SELL) {
                    count -= 1 * neuralWeight;
                }

            if (tsiTechnical.getValue() == TraderDecisionConstants.BUY) {
                count += 1 * technicalWeight;
            }
            else
                if (tsiTechnical.getValue() == TraderDecisionConstants.SELL) {
                    count -= 1 * technicalWeight;
                }
        }
    }
}

```

```

        }
    }

// change strategy scale
// if(count>1){
//     if(count>=0){
//         decision.addValue(tsiPattern.getDate(),
// (double)TraderDecisionConstants.BUY);
//         log += DataHelper.formataData(DataHelper.DEFAULT_PATTERN,
// tsiResult.getDate()+"\t"+TraderDecisionConstants.strDecision(tsiResult.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiPattern.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiNeural.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiTechnical.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(TraderDecisionConstants.BUY)+"\n";
//     }
//     else if(count<-1){
//         else if(count<0){
//             decision.addValue(tsiPattern.getDate(),
// (double)TraderDecisionConstants.SELL);
//             log += DataHelper.formataData(DataHelper.DEFAULT_PATTERN,
// tsiResult.getDate()+"\t"+TraderDecisionConstants.strDecision(tsiResult.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiPattern.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiNeural.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiTechnical.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(TraderDecisionConstants.SELL)+"\n";
//         }
//         else{
//             decision.addValue(tsiPattern.getDate(),
// (double)TraderDecisionConstants.HOLD);
//             log += DataHelper.formataData(DataHelper.DEFAULT_PATTERN,
// tsiResult.getDate()+"\t"+TraderDecisionConstants.strDecision(tsiResult.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiPattern.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiNeural.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(tsiTechnical.getValue().intValue())+"\t"+TraderDecisionConstants.strDecision(TraderDecisionConstants.HOLD)+"\n";
//         }
//     }
// }
}

```

## Apêndice B – CÓDIGO SIMPLIFICADO DO CÁLCULO DAS MÉTRICAS

```

public class Investment {

    public static boolean isVerbose=false;
    public static boolean isProfitability=false;

    double maxShare = 0.0d;
    double maxInvestment = 0.0d;
    Date maxShareDate = new Date();
    Date maxInvestmentDate = new Date();

    double minShare = 1000000.0d;
    double minInvestment = 1000000.0d;
    Date minShareDate = new Date();
    Date minInvestmentDate = new Date();

    //profitability
    double share = 10000;
    double investment = 0.0d;
    DecimalFormat decimal;

    //precision
    int correctOperation=0, failOperation=0, totalOperation=0;
    double lastShare=0;

    //evaluate
    int realCorrect=0,failCorrect=0,totaltotal;

    public Investment(TimeSerie tsPrice, TimeSerie tsDecision) {

        decimal = new DecimalFormat( "0.00" );

        Iterator<TimeSerieItem> itPrice = tsPrice.iterator();
        Iterator<TimeSerieItem> itDecision = tsDecision.iterator();

        double price = itPrice.next().getValue();
        TimeSerieItem tsi = itDecision.next();
        double lastDecision = tsi.getValue();

        Date lastDate = new Date();
        double lastPrice = 0.0d;

        if(lastDecision==TraderDecisionConstants.BUY) {
            investment = share/price;

            lastShare=share;

            share = 0;
            if(isVerbose)
                System.out.println(tsi.getDate()+""
                    "+price+"\t [BUY] \t s="+decimal.format(share)+"\t\t");
        }
    }
}

```

```

i)+"+decimal.format(investment));

        if(isProfitability)

    System.out.println(DataHelper.formataData(DataHelper.DEFAULT_PATTER
N, tsi.getDate()+" "+decimal.format(lastShare));
}

while(itDecision.hasNext()){
    TimeSerieItem tsiPrice = itPrice.next();
    TimeSerieItem tsiDecision = itDecision.next();

    if( (lastDecision==TraderDecisionConstants.SELL &
lastPrice < tsiPrice.getValue())
        ||
(lastDecision==TraderDecisionConstants.BUY && lastPrice >
tsiPrice.getValue())){
            realCorrect++;
        }
        else if( (lastDecision==TraderDecisionConstants.SELL
&& lastPrice > tsiPrice.getValue())
        ||
(lastDecision==TraderDecisionConstants.BUY && lastPrice <
tsiPrice.getValue())){
            failCorrect++;
        }
        else{
            totaltotal++;
        }

    lastPrice = tsiPrice.getValue();
    lastDate = tsiPrice.getDate();
    if(lastDecision==TraderDecisionConstants.SELL
&& share==0){

        share = investment *
tsiPrice.getValue();
        investment=0.0d;

        if(share>lastShare){
            correctOperation++;
        }
        else{
            failOperation++;
        }

        totalOperation++;

        if(isVerbose)

    System.out.println(DataHelper.formataData(DataHelper.DEFAULT_PATTER
N, tsiPrice.getDate()+" "+tsiPrice.getValue()+"\t [SELL] \t
s)+"+decimal.format(share)+"\t\t i)+"+decimal.format(investment));

        if(isProfitability)

    System.out.println(DataHelper.formataData(DataHelper.DEFAULT_PATTER
N, tsiPrice.getDate()+" "+decimal.format(share));

        if(share>maxShare) {
            maxShare = share;
}

```

```

maxShareDate=tsiPrice.getDate();
}

    if (share<minShare) {
        minShare = share;
    }

minShareDate=tsiPrice.getDate();
}
}

//buy
else
if (lastDecision<=TraderDecisionConstants.BUY && investment==0) {

    totalOperation++;
    lastShare = share;

    investment =
share/tsiPrice.getValue();
    share = 0.0d;
    if (isVerbose)

        System.out.println(DataHelper.formataData(DataHelper.DEFAULT_PATTER
N, tsiPrice.getDate()+" "+tsiPrice.getValue()+"\t [BUY] \t
s="+decimal.format(share)+"\t\t i="+decimal.format(investment));

    if (isProfitability)

        System.out.println(DataHelper.formataData(DataHelper.DEFAULT_PATTER
N, tsiPrice.getDate()+" "+decimal.format(lastShare));

    if (investment>maxInvestment) {
        maxInvestment = investment;
    }

    maxInvestmentDate=tsiPrice.getDate();
}

    if (investment<minInvestment) {
        minInvestment = investment;
    }

    minInvestmentDate=tsiPrice.getDate();
}

}

else {
    if (isVerbose)

        System.out.println(DataHelper.formataData(DataHelper.DEFAULT_PATTER
N, tsiPrice.getDate()+" "+tsiPrice.getValue()+"\t [HOLD] \t
s="+decimal.format(share)+"\t\t i="+decimal.format(investment));
    if (isProfitability) {
        double tempShare = lastShare
!= 0 ? lastShare : share;

        System.out.println(DataHelper.formataData(DataHelper.DEFAULT_PATTER
N, tsiPrice.getDate()+" "+decimal.format(tempShare));
    }
}

lastDecision = tsiDecision.getValue();

```

```

        }

        if(share==0.0) {
            share = investment*lastPrice;
            investment=0.0d;
            System.out.println(lastDate+" "+lastPrice+"\t
[SELL] \t s="+decimal.format(share)+"\t\t i="+decimal.format(investment));
        }
        if(isVerbose)
            System.out.println("\n\nMAX share = " + maxShare + "
in date = " +DataHelper.formataData(DataHelper.DEFAULT_PATTERN,
maxShareDate));

    }

public void statistics(){
    System.out.println("share = " + decimal.format(share));
    System.out.println("investment = " +
decimal.format(investment));

    System.out.println("Operation (c/f/t) =
"+correctOperation+"/"+failOperation+"/"+totalOperation);

    System.out.println("Operation (r/f/tt) =
"+realCorrect+"/"+failCorrect+"/"+totaltotal);

    System.out.println("MAX investment = " +
decimal.format(maxInvestment) + " in date = "
+DataHelper.formataData(DataHelper.DEFAULT_PATTERN, maxInvestmentDate));
    System.out.println("MAX share = " + decimal.format(maxShare)
+ " in date = " +DataHelper.formataData(DataHelper.DEFAULT_PATTERN,
maxShareDate));

    System.out.println("MIN investment = " +
decimal.format(minInvestment) + " in date = "
+DataHelper.formataData(DataHelper.DEFAULT_PATTERN, minInvestmentDate));
    System.out.println("MIN share = " + decimal.format(minShare)
+ " in date = " +DataHelper.formataData(DataHelper.DEFAULT_PATTERN,
minShareDate) + "\n\n");

    isVerbose = false;
}

}

```

## ANEXOS

## **Anexo A - PADRÕES CANDLESTICK E REGRAS DE IDENTIFICAÇÃO**

### ***Hammer***

Este padrão possui característica altista por se encontrar em um ponto de reversão de tendência e sua representação gráfica é ilustrada a seguir.



Regras que definem o *hammer*:

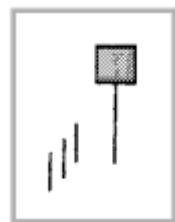
- Tendência de baixa;
- O pequeno corpo deve estar totalmente na parte superior;
- A sombra inferior deve ser de duas a três vezes o tamanho do corpo;
- Não deve haver a ocorrência de sombras superiores, apenas inferior
- A cor do padrão não influencia;

A ocorrência deste padrão antecipa os seguintes aspectos para apenas um dia seguinte:

- Tipo: reversão altista;
- Média de ganho: 3,06;
- Media de perdas: -3,25;

### ***Hanging Man***

Este padrão possui característica baixista por se encontrar em um ponto de reversão de uma tendência de alta, curiosamente o contrário do *hammer*. Sua representação gráfica é ilustrada a seguir.



Regras que definem o *hanging man*:

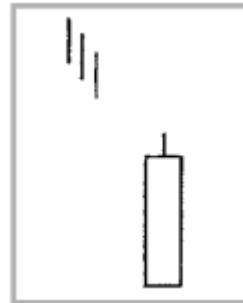
- Tendência de alta;
- O pequeno corpo deve estar totalmente na parte superior;
- A sombra inferior deve ser de duas a três vezes o tamanho do corpo;
- Não deve haver a ocorrência de sombras superiores, apenas inferior
- A cor do padrão não influencia;

A ocorrência deste padrão antecipa os seguintes aspectos para apenas um dia seguinte:

- Tipo: reversão baixista;
- Média de ganho: 3,63;
- Media de perdas: -2,81;

### **Belt Hold+**

Este padrão possui característica altista por se encontrar em um ponto de reversão de uma tendência de baixa. Sua representação gráfica é ilustrada a seguir.



Regras que definem o *belt hold+*:

- Tendência de baixa;
- O grande corpo deve estar totalmente na parte inferior;
- A sombra superior deve ser de duas a três vezes menor que o tamanho do corpo;
- Não deve haver a ocorrência de sombras inferiores, apenas superior
- O padrão deve fechar em alta;

A ocorrência deste padrão antecipa os seguintes aspectos para apenas um dia seguinte:

- Tipo: reversão altista;
- Média de ganho: 2,95;
- Media de perdas: -2,55;

### **Belt Hold-**

Este padrão possui característica baixista por se encontrar em um ponto de reversão de uma tendência de alta, curiosamente o contrário do *belt hold+*. Sua representação gráfica é ilustrada a seguir.



Regras que definem o *belt hold*-:

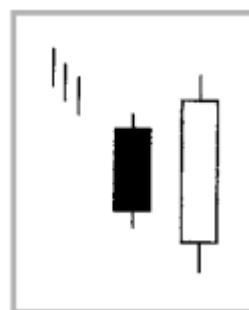
- Tendência de alta;
- O grande corpo deve estar totalmente na parte superior;
- A sombra inferior deve ser de duas a três vezes menor que o tamanho do corpo;
- Não deve haver a ocorrência de sombras superiores
- O padrão deve fechar em baixa;

A ocorrência deste padrão antecipa os seguintes aspectos para apenas um dia seguinte:

- Tipo: reversão baixista;
- Média de ganho: 2,72;
- Media de perdas: -2,69;

### *Engulfing +*

Este padrão possui característica altista por se encontrar em um ponto de reversão de uma tendência de baixa. Sua representação gráfica é ilustrada a seguir.



Regras que definem o *engulfing+*:

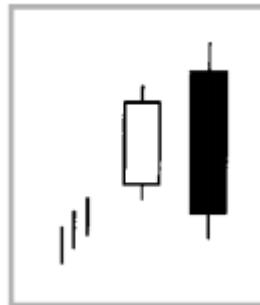
- Tendência de baixa;
- O pequeno corpo que fecha em baixa deve estar totalmente envolvido por um segundo corpo que fechou em alta;
- Sombra inferiores e superiores não importam;
- O primeiro padrão deve fechar em baixa e o segundo em alta;

A ocorrência deste padrão antecipa os seguintes aspectos para apenas um dia seguinte:

- Tipo: reversão altista;
- Média de ganho: 2,76;
- Media de perdas: -2,74;

### ***Engulfing -***

Este padrão possui característica baixista por se encontrar em um ponto de reversão de uma tendência de alta, curiosamente o contrário do *Engulfing+*. Sua representação gráfica é ilustrada a seguir.



Regras que definem o *engulfing-*:

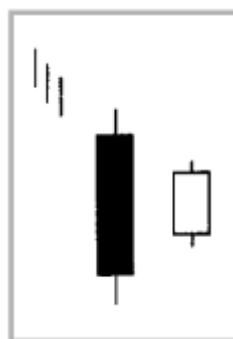
- Tendência de alta;
- O pequeno corpo que fecha em alta deve estar totalmente envolvido por um segundo corpo que fechou em baixa;
- Sombra inferiores e superiores não importam;
- O primeiro padrão deve fechar em alta e o segundo em baixa;

A ocorrência deste padrão antecipa os seguintes aspectos para apenas um dia seguinte:

- Tipo: reversão baixista;
- Média de ganho: 2,72;
- Média de perdas: -2,69;

### ***Harami+***

Este padrão possui característica altista por se encontrar em um ponto de reversão de uma tendência de baixa. Sua representação gráfica é ilustrada a seguir.



Regras que definem o *harami+*:

- Tendência de baixa;
- O grande corpo que fechou em baixa deve envolver totalmente um segundo corpo que fechou em alta;
- Sombra inferiores e superiores não importam;

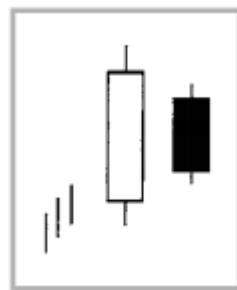
- O primeiro padrão deve fechar em baixa e o segundo em alta;

A ocorrência deste padrão antecipa os seguintes aspectos para apenas um dia seguinte:

- Tipo: reversão altista;
- Média de ganho: 2,69;
- Media de perdas: -2,46;

### ***Harami -***

Este padrão possui característica baixista por se encontrar em um ponto de reversão de uma tendência de alta, curiosamente o contrário do *Harami+*. Sua representação gráfica é ilustrada a seguir.



Regras que definem o *harami-*:

- Tendência de alta;
- O grande corpo que fecha em alta deve estar envolver totalmente um segundo corpo que fechou em baixa;
- Sombra inferiores e superiores não importam;
- O primeiro padrão deve fechar em alta e o segundo em baixa;

A ocorrência deste padrão antecipa os seguintes aspectos para apenas um dia seguinte:

- Tipo: reversão baixista;
- Média de ganho: 2,15;
- Media de perdas: -2,28;

## Anexo B - EQUAÇÕES DOS INDICADORES TÉCNICOS

### RSI

O Índice de Força Relativa mede a aceleração do movimento de preços utilizando a equação base a seguir:

$$RSI = 100 - \frac{100}{1 + \frac{\sum \varphi}{\sum \omega}}$$

Onde:

- $n$  é o número utilizado para calcular as médias, normalmente 9, 14 ou 25;
- $\sum \varphi$  é a média dos últimos  $n$  períodos que fecharam em alta;
- $\sum \omega$  é a média dos últimos  $n$  períodos que fecharam em baixa;

### ADX

O Índice de Direção do Movimento mede a força de uma tendência, sua intensidade independente da direção, através da equação base a seguir:

$$ADX = 100 * MME_n \left( \left| \frac{DI^+ - DI^-}{DI^+ + DI^-} \right| \right)$$

A composição do ADX é feita a partir de dois outros indicadores direcionais, um positivo  $DI^+$  e outro negativo  $DI^-$ , onde:

- $n$  é o número utilizado para calcular as médias, normalmente fixado em 14 para todos os cálculos do indicador;
- $MME_n$  é a média móvel exponencial

$$DI^+ = 100 * \frac{MME_n(DM^+)}{ATR}$$

$$DI^- = 100 * \frac{MME_n(DM^-)}{ATR}$$

E o cálculo individual de cada direcional é definido a partir das seguintes condições de existência:

$$DM^+ = \begin{cases} U; U > D \& U > 0 \\ 0; \text{caso contrário} \end{cases}$$

$$DM^- = \begin{cases} D; D > U \& D > 0 \\ 0; \text{caso contrário} \end{cases}$$

Com:

- U sendo o maior valor entre a diferença entre o máximo do dia seguinte pelo atual ou zero;
- D, de forma análoga, a diferença entre o menor valor do dia seguinte pelo atual ou zero;

$$U = \text{Maior} (Max_{n+1} - Max_n, 0)$$

$$D = \text{Maior} (Min_n - Min_{n+1}, 0)$$

O ATR é outro indicador que informa o grau de interesse ou desinteresse em uma ação, seja qual for sua direção, calculado a partir da seguinte equação, onde:

- O  $ATR$  é uma média móvel exponencial dos valores do  $TR$
- O  $TR$  é calculado pela maior diferença entre as três condições definidas em sua equação;

$$ATR = MME_n(TR)$$

$$TR = \text{Maior} [(Max_i - Min_i), (Max_i - X_{i-1}), (Min_i - X_{i-1})]$$

BBs

As Bandas de Bollinger definem uma zona para a variação de preços entre duas curvas calculadas através do desvio padrão de uma média móvel, denominada banda média, onde:

- $n$  é o número de preços presentes nas médias móveis deste indicador, definido em 20 atrasos;
- $D$  é o desvio padrão aplicado sobre as bandas, normalmente com valor 2;
- $BM_n$  é uma média móvel simples

$$\text{Banda média} = BM_n = MMS_n$$

$$\text{Banda superior} = BM_n + D * \sqrt{\frac{\sum(x_i - BM_n)^2}{n}}$$

$$\text{Banda inferior} = BM_n - D * \sqrt{\frac{\sum(x_i - BM_n)^2}{n}}$$