

UNIVERSIDADE DO VALE DO RIO DOS SINOS — UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA
NÍVEL MESTRADO

PAULO A. SCHMIDT

PREVISÃO DE PREÇOS DE AÇÕES NO PERÍODO INTRADIÁRIO POR MEIO DE
FOCUSED TIME LAGGED FEEDFORWARD NETWORKS

SÃO LEOPOLDO
2015

Paulo A. Schmidt

PREVISÃO DE PREÇOS DE AÇÕES NO PERÍODO INTRADIÁRIO POR MEIO DE
FOCUSED TIME LAGGED FEEDFORWARD NETWORKS

Dissertação apresentada como requisito parcial
para a obtenção do título de Mestre pelo
Programa de Pós-Graduação em Computação
Aplicada da Universidade do Vale do Rio dos
Sinos — UNISINOS

Orientador:
Prof. Dr. João F. Valiati

São Leopoldo
2015

AGRADECIMENTOS

A realização do trabalho aqui apresentado exigiu muito esforço, dedicação e horas dispendidas. Algumas pessoas foram fundamentais ao longo desta caminhada, tornando-a, de alguma forma, mais branda e interessante.

Agradeço, inicialmente, ao professor Dr. João F. Valiati por toda orientação, motivação, questionamentos e às diversas reuniões dedicadas a mim e a esta dissertação, fundamentais para minha evolução pessoal dentro da vida acadêmica.

A todos os professores do Programa Interdisciplinar de Pós-Graduação em Computação Aplicada da UNISINOS, que contribuíram para a minha formação, com seus ensinamentos, conselhos e experiência.

Aos colegas de trabalho e amigos, alguns deles já tendo trilhado este mesmo caminho, pelo apoio, incentivo e interesse no meu crescimento pessoal e profissional.

À UNISINOS, pelo auxílio financeiro através de bolsa de estudos, e por toda infraestrutura disponibilizada, garantindo um curso eficiente e agradável.

Aos meus pais, responsáveis por me incentivarem a sempre buscar nos estudos o conhecimento necessário para atingir metas mais altas.

E em especial agradeço minha noiva, Cláudia, pelo amor, paciência, compreensão, e pelas palavras motivacionais que me impulsionaram ao longo de todo o período do Mestrado.

“Os que se encantam com a prática sem a ciência são como os timoneiros que entram no navio sem timão nem bússola, nunca tendo certeza do seu destino”.
(Leonardo da Vinci)

RESUMO

A previsão de preços de ações é um assunto de grande interesse tanto por parte de agentes de mercado quanto da comunidade científica e acadêmica. Ao mesmo tempo, o problema é considerado como um dos mais desafiadores no tratamento de séries temporais, dada sua natureza altamente dinâmica. Uma ampla gama de estudos propõe-se a abordar o tema. Alguns com resultados bastante promissores fazem uso de Redes Neurais Artificiais (RNAs) do tipo *Focused Time Lagged FeedForward Network* (FTLFN), as quais apresentam mecanismos de memória capazes de detectar padrões temporais. Em muitos casos, no entanto, as capacidades da rede neural não são devidamente exploradas, limitando-se a testes com um conjunto mínimo de parâmetros. Além disso, a maioria dos estudos de previsões de preços de ações possui como foco períodos de baixa frequência, como dias ou meses. Contudo, devido à facilidade de acesso à informação nas últimas décadas e à automatização das negociações em bolsas de valores, estas são realizadas cada vez mais sob horizontes de curto prazo, como horas, minutos ou segundos. Existe, portanto, a necessidade de se expandir o conhecimento em relação a previsões dentro deste cenário. Neste sentido, este trabalho tem como objetivo uma investigação das reais potencialidades de previsão das FTLFNs sobre preços de ações no período intradiário. Sua memória de curto prazo e tamanho de camada oculta são explorados de forma ampla e aprofundada, através dos quais se buscou identificar o impacto das diferentes configurações nos resultados de acurácia dentro do contexto considerado. Na tentativa de oferecer suporte a melhores previsões, analisa-se também a influência de indicadores da Análise Técnica sobre o modelo. De forma mais geral, procura-se ampliar o entendimento a respeito tanto das capacidades de previsão das redes do tipo FTLFN como de sua empregabilidade em séries temporais financeiras intradiárias, ainda pouco exploradas na literatura. Os resultados obtidos mostram que, assim como investidores humanos, também as FTLFNs são capazes de se beneficiar enormemente de padrões formados pelos históricos dos sinais de entrada, a fim de prover previsões de maior qualidade dentro do contexto proposto neste trabalho. O mesmo não pode ser afirmado a respeito dos indicadores da Análise Técnica escolhidos, uma vez que em sua grande maioria aumentam os erros de previsão. As evidências apresentadas baseiam-se em experimentações sobre diferentes conjuntos de sinais, oferecendo robustez às conclusões alcançadas e permitindo que a metodologia e os resultados sirvam como base para futuras pesquisas relacionadas a previsões dentro de cenários de alta frequência.

Palavras-chave: Redes Neurais Artificiais. FTLFN. Intradiário. Análise Técnica. Bolsa de valores.

ABSTRACT

Stock price prediction is a subject of great interest for both market agents and scientific and academic community. At the same time, this problem is considered to be one of the most challenging in time series forecasting, due to its highly dynamic nature. A large amount of researches have proposed to address the issue. Some of them, with very promising results, adopt the Focused Time Lagged FeedForward Network (FTLFN), a type of Artificial Neural Network (ANN) that offers memory mechanisms capable of detecting temporal patterns. In many cases, however, the neural network's capacities are not properly explored, being limited to tests with a minimum set of parameters. Besides, most of the studies on stock price prediction focus on low-frequency periods, such as days or months. On the other hand, due to the ease of access to information in the last decades and the automation of trades in stock market, these are getting more oftenly executed over short-term horizons, like hours, minutes or seconds. Therefore, there is a need to expand the knowledge related to forecasts in this scenario. With that in mind, this research has the objective of investigating the FTLFN's potential on stock price forecasting over the intraday period. Its short-term memory and hidden layer size are widely and deeply explored, so the impact of different configurations on the accuracy results could be measured. Also, Technical Analysis indicators are built and utilized as input signals to the network, with their possible contributions to stock prediction being verified. From a general perspective, the work proposes the extension of the understanding regarding the FTLFN's forecasting capabilities, as well as its use with intraday financial time series, which still require further exploration in literature. The obtained results show that, as human investors do, also FTLFNs are capable of taking enormous advantage from input signals' history on providing better prediction quality within the proposed context. The same cannot be said for the supporting Technical Analysis indicators chosen, since they mostly increase forecasting errors. Evidences are presented based on the experimentation over several sets, bringing robustness to the conclusions and allowing the methodology and the results to serve as base for future researches related to predictions on high-frequency trading scenarios.

Keywords: Artificial Neural Networks. FTLFN. Intraday. Technical Analysis. Stock market.

LISTA DE FIGURAS

Figura 1:	Gráfico de barras	26
Figura 2:	Média Móvel	28
Figura 3:	Bandas de Bollinger	29
Figura 4:	Índice de Força Relativa	30
Figura 5:	Modelo matemático de um neurônio	32
Figura 6:	Representação simplificada de uma rede neural artificial	34
Figura 7:	<i>Focused Time Lagged FeedForward Networks</i>	38
Figura 8:	Fluxo das etapas da metodologia	49
Figura 9:	Divisão dos dados por janela deslizante	55
Figura 10:	Variação no erro de acordo com o tamanho da linha de atraso - Petrobras	61
Figura 11:	Variação no erro de acordo com o tamanho da linha de atraso - Vale	62
Figura 12:	Variação no erro de acordo com o número de neurônios na camada oculta - Petrobras	63
Figura 13:	Variação no erro de acordo com o número de neurônios na camada oculta - Vale	64
Figura 14:	Influência de Indicadores da Análise Técnica - Petrobras, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta	65
Figura 15:	Influência de Indicadores da Análise Técnica - Vale, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta	66
Figura 16:	Degradação no desempenho de previsão da FTLEFN ao longo de uma semana de dados - Petrobras, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta	67
Figura 17:	Degradação no desempenho de previsão da FTLEFN ao longo de uma semana de dados - Vale, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta	68
Figura 18:	Degradação no desempenho de previsão ao longo de duas e três semanas - Petrobras, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta	69
Figura 19:	Degradação no desempenho de previsão ao longo de duas e três semanas - Vale, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta	70

LISTA DE TABELAS

Tabela 1:	MAPE médio sobre dados transformados	42
Tabela 2:	Comparação de previsões diárias e intradiárias	44
Tabela 3:	Acurácia de classificação	46
Tabela 4:	Formato de retorno dos dados pelo <i>TradeZone Data Center</i>	53
Tabela 5:	Definições sobre a FTLFN	58
Tabela 6:	Conjuntos de sinais de entrada utilizados nas experimentações	58
Tabela 7:	MAPE de cada conjunto de entrada, por número de atrasos, representados como a média dos valores obtidos para os testes com 10, 20, 30, 50 e 80 neurônios na camada oculta - Petrobras	85
Tabela 8:	MAPE de cada conjunto de entrada, por número de atrasos, representados como a média dos valores obtidos para os testes com 10, 20, 30, 50 e 80 neurônios na camada oculta - Vale	85
Tabela 9:	MAPE de cada configuração de número de atrasos e neurônios na camada oculta, representados como a média entre os 11 conjuntos de sinais de entrada - Petrobras	87
Tabela 10:	MAPE de cada configuração de número de atrasos e neurônios na camada oculta, representados como a média entre os 11 conjuntos de sinais de entrada - Vale	87
Tabela 11:	MAPE de cada conjunto de entrada, por configuração de indicadores da Análise Técnica utilizados como sinais auxiliares - 15 atrasos, 30 neurônios na camada oculta - Petrobras	89
Tabela 12:	MAPE de cada conjunto de entrada, por configuração de indicadores da Análise Técnica utilizados como sinais auxiliares - 15 atrasos, 30 neurônios na camada oculta - Vale	89
Tabela 13:	MAPE médio de cada período da primeira semana de teste, demonstrando a degradação no desempenho da FTLFN - 15 atrasos, 30 neurônios na camada oculta - Petrobras	91
Tabela 14:	MAPE médio de cada período da primeira semana de teste, demonstrando a degradação no desempenho da FTLFN - 15 atrasos, 30 neurônios na camada oculta - Vale	92
Tabela 15:	MAPE médio de cada conjunto de entrada, considerando uma, duas e três semanas de teste, sem re-treino da FTLFN - 15 atrasos, 30 neurônios na camada oculta - Petrobras	93
Tabela 16:	MAPE médio de cada conjunto de entrada, considerando uma, duas e três semanas de teste, sem re-treino da FTLFN - 15 atrasos, 30 neurônios na camada oculta - Vale	93

LISTA DE SIGLAS

ARIMA	<i>Autoregressive Integrated Moving Average</i>
BACEN	Banco Central
BB	Bandas de Bollinger
CBLC	Companhia Brasileira de Liquidação e Custódia
CDB	Certificado de Depósito Bancário
CVM	Comissão de Valores Mobiliários
FFNN	<i>FeedForward Neural Networks</i>
FTDNN	<i>Focused Time-Delay Neural Network</i>
FTLFN	<i>Focused Time Lagged FeedForward Network</i>
GARCH	<i>Generalized Autoregressive Conditional Heteroskedasticity</i>
HTTP	<i>Hypertext Transfer Protocol</i>
IFR	Índice de Força Relativa
LPA	Lucro por ação
MAPE	<i>Mean Absolute Percentage Error</i>
MLP	<i>Multilayer Perceptron</i>
MM	Média Móvel
NARX	<i>Nonlinear Autoregressive Network with Exogenous Inputs</i>
RMSE	<i>Root Mean Squared Error</i>
RNAs	Redes Neurais Artificiais
ROE	<i>Return over equity</i>
SVM	<i>Support Vector Machines</i>
TDNN	<i>Time-Delay Neural Networks</i>

SUMÁRIO

1 INTRODUÇÃO	19
1.1 Motivação	20
1.2 Objetivo	22
1.3 Organização do Trabalho	22
2 FUNDAMENTAÇÃO TEÓRICA	23
2.1 O Mercado de Capitais e a Bolsa de Valores	23
2.2 Análises do Mercado	25
2.2.1 Análise Fundamentalista	25
2.2.2 Análise Técnica	25
2.3 Previsão de Séries Temporais	30
2.3.1 Discretização dos dados	31
2.4 Redes Neurais Artificiais	31
2.4.1 O Neurônio Artificial	32
2.4.2 Arquitetura da Rede Neural Artificial	33
2.4.3 Aprendizado	34
2.4.4 Regularização Bayesiana	36
2.4.5 Focused Time Lagged FeedForward Networks	37
2.5 Considerações	39
3 TRABALHOS RELACIONADOS	41
3.1 FTLFN sobre séries temporais financeiras	41
3.2 Previsões no período intradiário através de redes neurais com atrasos no tempo	43
3.3 Indicadores da Análise Técnica como suporte a previsões de alta frequência	44
3.4 Discussão	46
4 ABORDAGEM EXPERIMENTAL	49
4.1 Metodologia	49
4.2 Seleção de dados	52
4.3 Pré-processamento	53
4.3.1 Limpeza da base	53
4.3.2 Sinais auxiliares da Análise Técnica	54
4.3.3 Particionamento dos dados	54
4.4 Modelagem Neural	56
4.5 Experimentações planejadas	57
4.6 Resultados	59
4.6.1 Resultados das experimentações sobre a memória de linha de atraso	60
4.6.2 Resultados das experimentações sobre o número de neurônios	62
4.6.3 Resultados das experimentações sobre indicadores da Análise Técnica	65
4.6.4 Resultados para experimentações sobre a degradação da rede	67
4.7 Discussão	70
5 CONCLUSÃO	75
REFERÊNCIAS	79

APÊNDICE A	TABELAS DE RESULTADOS DOS EXPERIMENTOS SOBRE ATRASOS	85
APÊNDICE B	TABELAS DE RESULTADOS DOS EXPERIMENTOS SOBRE NEURÔNIOS	87
APÊNDICE C	TABELAS DE RESULTADOS DOS EXPERIMENTOS SOBRE ANÁLISE TÉCNICA	89
APÊNDICE D	TABELAS DE RESULTADOS DOS EXPERIMENTOS SOBRE DEGRADAÇÃO	91

1 INTRODUÇÃO

A bolsa de valores é um dos indicadores primários da força econômica de um país. Por estar sujeita às leis da oferta e da procura, se os preços estão em alta significa que a demanda por ações é maior do que sua oferta. Usualmente, este cenário indica que as empresas estão conseguindo gerar retornos consistentes com seus investimentos e distribuindo parte de seus lucros aos acionistas. Do contrário, se as cotações apresentam-se em queda, tem-se um cenário de oferta maior que demanda pelos papéis das empresas. Isto pode indicar uma dificuldade por parte delas de executarem seus projetos, como consequência de inúmeros fatores, incluindo má administração governamental do país como um todo.

Além disso, o mercado acionário é um importante facilitador e promotor da formação de capital, levando ao crescimento econômico através de investimentos reais (LAU; DEMIR; BILGIN, 2013). Empresas abrem capital em busca de recursos para fomentar seu desenvolvimento, e investidores adquirem ações esperando retornos favoráveis das respectivas companhias. O grande viés deste tipo de investimento, porém, é o risco, uma vez que não existem garantias de rendimentos mínimos e há possibilidade, inclusive, de perda total do capital investido. Por esse motivo, o estudo da previsão de preços de ações se justifica, elevando a expectativa de retornos positivos, diminuindo riscos e consequentemente reforçando a ideia do mercado acionário como uma boa opção de geração de valor.

Segundo a Hipótese de Mercados Eficientes (FAMA, 1965), os preços de ações e índices no mercado acionário já refletem toda nova informação que pudesse afetá-los. Portanto, qualquer tentativa de previsão destes preços seria uma tarefa fútil. Ainda, de acordo com o estudo, as cotações seguem um padrão de comportamento aleatório, não podendo ser previstas com mais de 50% de acurácia (BOLLEN; MAO; ZENG, 2011). Por outro lado, diversos estudos questionam a hipótese de Fama e mostram que, de fato, é possível prever os preços com algum grau de precisão, não sendo estes, portanto, aleatórios. Como exemplo, Atsalakis e Valavanis (2009) enumeram mais de 100 artigos publicados, onde são aplicadas diferentes técnicas computacionais para previsão no mercado financeiro, a grande maioria deles apresentando resultados que refutam a Hipótese de Mercados Eficientes.

No entanto, a previsão dos preços de ações, assim como suas movimentações, têm sido tratados como alguns dos mais desafiadores cenários entre aqueles relacionados à previsão de séries temporais. Isso se deve ao fato de que o mercado acionário é essencialmente dinâmico, não linear e não paramétrico. Além disso, é afetado por muitos fatores macro e microeconômicos, como eventos políticos, fundamentos de empresas, condições econômicas gerais, expectativas de investidores, escolhas de investidores institucionais, movimentos de outros mercados, psicologia dos investidores, etc (TAN; QUEK; NG, 2007).

Entre as técnicas popularmente utilizadas por agentes de mercado, na tentativa de identificar movimentos futuros, estão a Análise Técnica e Análise Fundamentalista. A Análise Técnica estuda dados históricos de ações através de gráficos e indicadores. São observados topos e

fundos, tendências e padrões formados pelo comportamento dos preços ao longo do tempo. Como suporte para a tomada de decisão de compra ou venda do ativo são utilizados indicadores técnicos calculados a partir de preços e volumes, que modelam algum aspecto do movimento das cotações. Médias móveis e osciladores são exemplos de indicadores empregados regularmente por analistas (ACHELIS, 2013). Já a Análise Fundamentalista preocupa-se com as condições econômicas das empresas, na tentativa de determinar seu valor de mercado intrínseco. É um método de pesquisa que estuda informações financeiras, provendo meios de comparação entre os desempenhos de diferentes empresas e fornecendo uma perspectiva sobre o potencial preço futuro de um ativo. Os indicadores resultantes desta análise são chamados de múltiplos de mercado, sendo o P/L (razão entre o preço e o lucro) um dos mais utilizados para se determinar quanto atrativo o preço da ação se apresenta (THOMSETT, 1998).

Existe, ainda, um grande número de estudos que se utiliza de modelos de regressão na tentativa de prever séries temporais, aplicando-os a preços futuros de ações. Incluem modelos lineares, como o alisamento exponencial de Holt e ARIMA (*Autoregressive Integrated Moving Average*), e não lineares, como GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*) e Redes Neurais Artificiais (RNAs). Estas últimas têm ganhado excelente reputação como aproximadores não lineares. São altamente tolerantes à imprecisão e exibem bons resultados mesmo em ambientes onde os dados apresentam ruídos. Tais características tornam as RNAs interessantes candidatas a serem aplicadas a problemas de domínio financeiro, especialmente em bolsas de valores, onde os dados exibem altos índices de ruído e incerteza (VANSTONE; FINNIE, 2009).

1.1 Motivação

Nas últimas décadas, uma grande quantidade de trabalhos tem sido publicada sobre a aplicação de redes neurais a problemas relacionados ao mercado financeiro, como a seleção de ativos para otimização de portfólios, a previsão de preços futuros e a influência de notícias sobre as ações. Dentre aqueles estudos que focam na previsão do preço de determinado ativo, grande parte deles tem como objetivo prever o fechamento do dia seguinte. Outros focam no longo prazo, prevendo meses à frente. Poucos, porém, abordam previsões dentro do intervalo intradiário.

Variações de curto prazo são de difícil previsibilidade, mas frequentemente são a origem de ganhos e perdas em negociações de ações. Os motivos para o uso de operações sobre este horizonte de tempo incluem necessidade de liquidez, rebalanceamento de portfólios, antecipação de mudanças em leis de impostos, mas principalmente a possibilidade de lucros rápidos e em maior escala, uma vez que para *day trades* há a opção de alta alavancagem do capital investido. *Day traders* são investidores que se posicionam em determinado ativo e liquidam tal posição em um mesmo dia. Correspondem a uma grande parcela dos atuantes do mercado (SITTE; SITTE, 2000). Jordan e Diltz (2003) mostram, no entanto, que existem duas vezes mais *day traders*

que sofrem prejuízos do que aqueles que obtêm lucros, e apenas 20% deles conseguem acumular ganhos substanciais. Um dos motivos para tal estatística é o fato de que estes investidores baseiam suas decisões em interpretações muito subjetivas sobre as informações e os preços, ao mesmo tempo em que sofrem forte influência de fatores psicológicos (KUO; LIN, 2013). Portanto qualquer solução que melhore a previsão de preços no curto prazo, diminuindo riscos e eliminando a subjetividade, é de grande utilidade neste tipo de situação.

Surgidas na última década, as negociações de alta frequência, ou *High-Frequency Trading* (HFT), também são capazes de se beneficiar enormemente de previsões no período intradiário. Neste tipo de negociação são realizadas operações de compra e venda de ativos em espaços muito curtos de tempo, como minutos ou segundos, através de algoritmos, sem qualquer intervenção humana. Estudos sugerem que entre 50 e 70% de todo volume de ações negociado nos EUA, nos últimos 3 anos, tem origem em HFT (SILVA et al., 2014). Esta estratégia também é utilizada frequentemente por formadores de mercado, caracterizados por empresas ou indivíduos comprometidos em comprar e vender ativos de forma regular e contínua. Sua importância está em prover liquidez aos papéis, especialmente aqueles menos negociados, e descobrir faixas de preços de maior interesse público (MADUREIRA; UNDERWOOD, 2008). A partir destas observações, conclui-se que apesar de ainda receberem pouca atenção em pesquisas acadêmicas, previsões intradiárias de preços podem ser de alta relevância para os mercados atuais.

Até hoje, diversos foram os modelos de RNAs utilizados na previsão de preços de mercado, destacando-se as *FeedForward Neural Networks* (FFNN), as *Time Delay Neural Networks* (TDNN), as *Support Vector Machines* (SVMs), os sistemas *Fuzzy*, entre outros, como pesquisado em Atsalakis e Valavanis (2009). Outro tipo importante de rede neural neste contexto são as *Focused Time Lagged FeedForward Networks* (FTLFN). Sua principal característica é a memória de curto prazo, implementada na camada de entrada através de atrasos no sinal fornecido. Tal funcionalidade permite à rede realizar processamentos temporais e reconhecer sequências de eventos, baseando-se em históricos recentes de valores e possíveis dependências entre estes dados (HAYKIN, 2001).

Diferentes autores já demonstraram que preços passados podem ser utilizados com sucesso para prever retornos futuros, supondo-se determinado grau de confiança (LO; MACKINLAY, 1988). Outros estudaram a importância da consideração de atrasos na auto-correlação entre retornos no mercado financeiro (DING; GRANGER; ENGLE, 1993). Portanto a memória de curto prazo das FTLFNs as tornam candidatas interessantes para o tratamento de previsão de preços. Apesar de já terem sido empregadas para este propósito, conforme literatura relacionada, por vezes suas potencialidades são exploradas de forma bastante vaga. De acordo com Braga, Carvalho e Ludermir (2014), a capacidade das FTLFNs de aproximar valores futuros é altamente dependente de ajustes adequados de seus diferentes parâmetros, os quais se apresentam em grande número. Desta forma, são necessárias diversas experimentações a fim de se determinar a configuração que habilita a rede a realizar as melhores previsões. Poucos são os estudos que tratam o problema através desta perspectiva. Em grande parte, apenas um conjunto

mínimo de parametrizações é adotado para testes.

1.2 Objetivo

O objetivo deste trabalho é investigar a efetividade do uso de redes neurais artificiais do tipo FTLFN na previsão de preços de ações de alta liquidez, sobre o intervalo intradiário. Neste estudo são consideradas diferentes configurações de atrasos para os sinais de entrada, variações no tamanho da camada oculta da rede, e uso conjunto de indicadores da Análise Técnica, no sentido de fornecer indicativos de quando seu uso é recomendado. Para tanto, são definidos os seguintes objetivos específicos:

- Realização de experimentos exploratórios que captem a influência do tamanho da memória de linha de atraso sobre o desempenho de previsão da rede FTLFN;
- Realização de experimentos comparativos entre redes FTLFNs com diferentes números de neurônios em sua camada oculta, com o intuito de detectar configurações que melhor permitem à rede generalizar o conhecimento adquirido;
- Identificação do impacto na minimização do erro de previsão das redes neurais, pelo uso de sinais adicionais, gerados a partir de preços intradiários, baseados em indicadores da Análise Técnica;
- Produção de uma análise crítica do estudo, apontando suas contribuições para o campo de previsão de valores por meio de redes neurais FTLFN, assim como novas possíveis explorações.

1.3 Organização do Trabalho

O presente trabalho apresenta, no capítulo 2, conceitos relativos ao mercado de ações e à Análise Técnica, assim como o funcionamento das redes neurais, com foco nas FTLFNs. O capítulo 3 relata trabalhos relacionados a este estudo, abrangendo as diferentes fases da previsão de preços por meio de redes neurais. No capítulo 4 é apresentada a metodologia utilizada na obtenção e preparação dos dados, assim como a modelagem neural empregada. Ainda, neste mesmo capítulo, são apresentados detalhes das experimentações realizadas, juntamente com os respectivos resultados obtidos. O capítulo 5 encerra o estudo, apresentando conclusões acerca do trabalho desenvolvido, além de contribuições e possíveis extensões.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são descritos os principais conceitos necessários para a compreensão do domínio deste trabalho e das técnicas empregadas. Inicialmente são apresentados conceitos básicos sobre o Mercado de Capitais e Bolsas de Valores, com ênfase no mercado brasileiro. Após, discute-se as formas mais comuns de análise de empresas e ações adotadas por investidores e agentes de mercado. Em seguida, é realizada uma breve revisão sobre técnicas de previsão de séries temporais, no contexto da Bolsa de Valores. Descreve-se, ainda, como são discretizadas as séries temporais para uso no mercado financeiro. Por fim, as redes neurais artificiais são detalhadas, com ênfase no modelo específico utilizado neste estudo, a FTLFN.

2.1 O Mercado de Capitais e a Bolsa de Valores

Segundo Pinheiro (2012), o Mercado de Capitais é um conjunto de instituições financeiras, como a Bolsa de Valores, os bancos e as companhias corretoras de investimento, que negociam com títulos e valores mobiliários, objetivando a transferência de recursos dos agentes compradores para os agentes vendedores. Tem como proposta, ainda, viabilizar a capitalização das empresas e dar liquidez aos títulos emitidos por elas.

Através da Bolsa de Valores são realizadas negociações de compra e venda de títulos e valores mobiliários. Ela fornece suporte à captação de recursos financeiros através da distribuição e circulação de valores, e realiza a regulamentação deste processo. Além disso, 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. No Brasil, a companhia responsável por gerenciar tais atividades é a BM&F Bovespa, cujo volume financeiro total em 2013 atingiu a marca recorde de R\$ 1,83 trilhão, enquanto que a média diária de volume financeiro alcançou R\$ 7,41 bilhões, valor também inédito na história (BM&F BOVESPA, 2013). Tais números mostram o tamanho e a importância deste mercado, na forma como se encontra hoje. A Bovespa proporciona um ambiente virtual, onde é possível realizar operações em diversos tipos de mercado, sendo os mais importantes os mercados à vista, a termo, futuro e de opções. O mercado à vista é o mais comumente utilizado pelo investidor pessoa física. Nele são realizadas negociações de compra e venda de valores mobiliários em tempo real, com prazo de liquidação física e financeira regulamentados pela CBLC (Companhia Brasileira de Liquidação e Custódia). Tais títulos são emitidos por empresas de capital aberto e recebem o nome de "ações" (PINHEIRO, 2012).

Quando empresas sentem a necessidade de financiamento para execução de seus projetos e atividades, uma das alternativas é a abertura de capital na Bolsa de Valores, emitindo ações. Uma ação corresponde à representação da menor fração do capital social de uma empresa. O comprador deste título torna-se sócio da companhia, possuindo poderes e direitos sobre a mesma, limitados de acordo com o tipo e quantidade da ação adquirida (CAVALCANTE; MI-

SUMI; RUDGE, 2009). Além disso, o comprador passa a receber dividendos, ou seja, parte dos lucros auferidos pela empresa, enquanto for proprietário de alguma ação correspondente. O preço de uma ação é determinado de acordo com a lei da oferta e da demanda, ou seja, quanto menor a oferta e maior a demanda, maiores são os preços. Já quando existe uma demanda menor que a oferta, as ações tendem a ficar mais baratas. Os preços estão em constante movimento, e o investidor não possui a informação *a priori* do rendimento relativo aos seus investimentos. Daí sua nomenclatura de renda variável, diferentemente da renda fixa (fundos de renda fixa, cadernetas de poupança, CDBs e títulos públicos), onde os investimentos têm a rentabilidade atrelada de alguma forma a um índice ou indicador razoavelmente previsível e, teoricamente, incorrem em menor risco (HISSA, 2008).

Atualmente são negociadas na Bovespa ações de cerca de 350 companhias. Aquelas que atendem determinados critérios de negociabilidade são inclusas no Índice Bovespa, ou Ibovespa. Este indicador foi criado com o objetivo de refletir o desempenho médio das cotações dos ativos de maior representatividade do mercado de ações brasileiro, com seus integrantes respondendo por mais de 80% do número de negócios e do volume financeiro no mercado à vista. São, portanto, ativos de alta liquidez, ou seja, passíveis de serem comprados e vendidos com facilidade. Esta característica permite que investidores adquiram e desfaçam-se de ações rapidamente, sendo fundamental em negociações de curto prazo (THOMSETT, 1998). O valor do Índice Bovespa é calculado em tempo real e corresponde, em moeda corrente, ao de uma carteira teórica de ações constituída em 02/01/1968 a partir de uma aplicação hipotética, sendo reajustada regularmente de acordo com os critérios descritos em sua metodologia e com a distribuição de proventos pelas empresas emissoras (BM&F BOVESPA, 2013). Exemplos de empresas brasileiras listadas no Ibovespa são: Petrobras, Vale, Ambev, Bradesco, Itaú, etc.

Investidores não possuem acesso direto aos sistemas de negociação da BM&F Bovespa para efetuarem suas transações de compra e venda de valores. Eles precisam, para isso, contratar os serviços de corretoras de valores, instituições com constituição autorizada pelo Banco Central (BACEN) e exercício autorizado pela Comissão de Valores Mobiliários (CVM). Uma das principais funções de uma corretora é atuar como intermediadora das transações de compra e venda de títulos mobiliários por parte dos seus clientes. Igualmente importante, ela deve proporcionar um ambiente transparente e seguro, adequado à realização de negócios com valores mobiliários. Através de seus assessores, as corretoras ainda auxiliam o investidor a escolher as ações de acordo com seu objetivo financeiro, definem o perfil deste investidor, informam clientes sobre novo produtos, entre outras atribuições (COSTA; MARCONDES, 2012).

Tanto investidores quanto assessores de corretoras que sugerem ativos para seus clientes precisam, antes de mais nada, estudar e analisar o mercado e as empresas com as quais estão dispostos a negociar na Bolsa de Valores. As cotações das ações sofrem influência de uma infinidade de fatores, muitos deles até mesmo imprevisíveis, como: taxa de juros, inflação, guerras, epidemias, cenários macro e micro econômicos, taxa de desemprego. Portanto, a tarefa de se determinar qual a melhor empresa para se investir, assim como o melhor momento para

tal, é de grande complexidade e requer análises cuidadosas, na tentativa de se prever um preço ou comportamento futuro de determinada ação.

2.2 Análises do Mercado

Duas estratégias econômicas bastante utilizadas no mercado são a Análise Fundamentalista e a Análise Técnica. Ambas englobam ferramentas e metodologias próprias que fornecem suporte à tomada de decisão relativa à compra ou venda de ativos em bolsas de valores. Adotar uma ou outra estratégia depende do objetivo do investidor e, apesar dos pensamentos distintos das duas correntes, não é incomum o uso conjunto delas, haja visto que tanto a Análise Técnica quanto a fundamentalista oferecem informações importantes sobre as ações.

Nas subseções a seguir são explicados conceitos relativos a ambas abordagens, em especial à Análise Técnica, empregada com mais frequência por investidores com interesse em retornos no curto prazo.

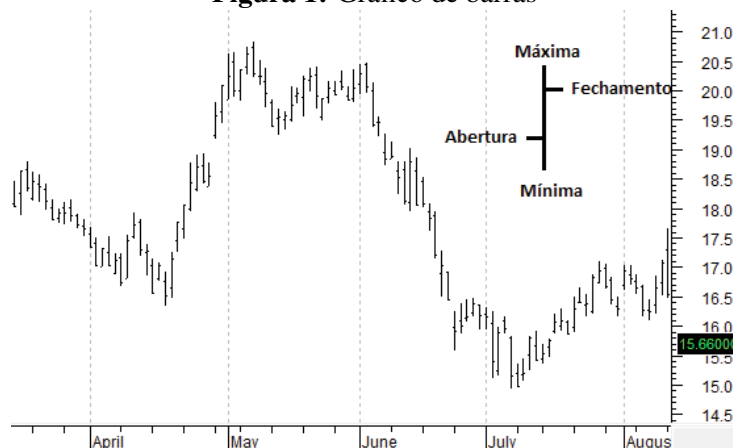
2.2.1 Análise Fundamentalista

Thomsett (1998) define a Análise Fundamentalista com um método de pesquisa que estuda informações financeiras básicas, com o intuito de prever lucros, ofertas e demandas, força da indústria, habilidades gerenciais, e outros assuntos intrínsecos que afetam o valor de mercado e o potencial de crescimento de uma ação. Seu principal benefício é prover meios para comparações entre companhias, facilitando o processo de escolha quando se deseja investir. Indicadores fundamentalistas clássicos, utilizados por analistas de mercado, são o LPA (lucro por ação), o P/L (preço da ação em relação ao lucro por ação), o ROE (retorno financeiro sobre o patrimônio líquido), o *Dividend Yield* (razão entre o dividendo pago por ação e o preço da ação), entre outros.

Apesar de sua enorme utilidade na decisão de compra e venda de determinada ação, a Análise Fundamentalista é direcionada a prospectos de longo prazo e àqueles que se engajam na tentativa de prever preços e movimentos meses ou anos no futuro (THOMSETT, 1998). O presente trabalho, ao contrário, preocupa-se com modificações mais imediatas no mercado, não fazendo uso, portanto, da Análise Fundamentalista. A prática de mercado, para este caso, é o uso da Análise Técnica.

2.2.2 Análise Técnica

Segundo Murphy (1999), a Análise Técnica é o estudo dos movimentos do mercado, primariamente, através do uso de gráficos, com o propósito de prever futuras tendências de preços. Este método está baseado em três premissas principais: o mercado desconta tudo, os preços se movem em tendências e a história se repete. Dizer que o mercado desconta tudo significa

Figura 1: Gráfico de barras

Fonte: Elaborado pelo autor desta dissertação.

afirmar que todos os fatores que possam afetar o preço de uma ação são, de fato, refletidos neste preço e, portanto, nada além do estudo de seus movimentos é necessário. A identificação de tendências, por sua vez, é o princípio básico que leva ao uso de gráficos. Existe um corolário a essa premissa que diz que a probabilidade de uma tendência continuar seu movimento é maior do que a probabilidade de sua reversão. É bastante comum a investidores, portanto, a adoção da estratégia de *trend-following*, seguindo uma tendência até o surgimento de sinais de reversão. Já a premissa de que a história se repete está relacionada com o estudo da psicologia humana. Segundo pesquisadores da área, padrões identificados em gráficos de preços passados revelam a psicologia de mercados em alta ou em baixa e tendem a apresentar comportamentos muito semelhantes no futuro. Ou seja, a previsão de movimentos se baseia no estudo do passado.

Gráficos fazem parte da fundamentação da Análise Técnica. Eles mostram o histórico de preços de uma determinada ação, durante um intervalo de tempo específico. Diversos são os tipos de gráficos utilizados por analistas técnicos, como o de linha e o de *candlesticks*. Segundo Murphy (1999), no entanto, o mais popular é o gráfico de barras. É assim chamado porque cada período do tempo que se está analisando é representado por uma barra vertical, com duas pequenas marcas horizontais, como pode ser visualizado na Figura 1. Nesta imagem específica, cada barra representa os preços de abertura e fechamento de um dia, assim como os preços máximo e mínimo atingidos no decorrer do mesmo (STEVENS, 2002).

Além de gráficos, a Análise Técnica faz uso extensivo de modelos matemáticos que exibem diversos aspectos das atividades dos preços, como momentos promissores de compra e venda ou taxas de valorização e desvalorização da ação. Tais modelos são denominados indicadores técnicos (STEVENS, 2002). Quando do início de sua concepção, estas ferramentas possuíam função apenas complementar à análise de mercado, e não apresentavam peso considerável ou confiança para serem fatores decisivos de análise. Hoje em dia, no entanto, são consideradas parte importante do processo de tomada de decisão dos investidores, sendo peça fundamental para análise do comportamento futuro do ativo (ACHELIS, 2013).

Devido a sua natureza matemática e precisa, indicadores técnicos tendem a ser utilizados de forma mecânica, sem a devida interpretação do cenário e do momento em que o mercado está inserido. Neste contexto, é importante o uso de mais de um indicador técnico, além do estudo dos padrões dos gráficos, para que haja uma ponderação em relação à confirmação de sinais de compra e venda de ativos (STEVENS, 2002). Atualmente existem centenas de indicadores técnicos, e a escolha de quais utilizar recai, na maioria das vezes, sobre a preferência do investidor. Entre os mais populares estão as Médias Móveis (MM), as Bandas de Bollinger (BB) e o Índice de Força Relativa (IFR) (MURPHY, 1999).

2.2.2.1 Médias Móveis

As Médias Móveis são utilizadas em larga escala, e são a base para diversos outros indicadores. Representam valores médios extraídos de um corpo de dados sequenciais em uma determinada janela de tempo. Em uma média de 10 períodos, por exemplo, o valor em um momento específico indica o preço médio dos últimos 10 períodos, incluindo o corrente, sendo 10 o tamanho da janela de tempo considerada. Diz-se que ela é móvel porque, no período seguinte, a cotação mais antiga é eliminada do cálculo, enquanto que o novo preço é incluso. A média resultante se move relativamente para o próximo período (NORONHA, 1995). As MMs pertencem à categoria dos rastreadores de tendência, ou seja, são eficazes para acompanhar e evidenciar a direção seguida pelos preços, com um pequeno atraso.

Variações para o cálculo das MMs são comuns. A Média Móvel Simples reflete a média aritmética dos preços considerados, e sua função é expressa pela seguinte fórmula:

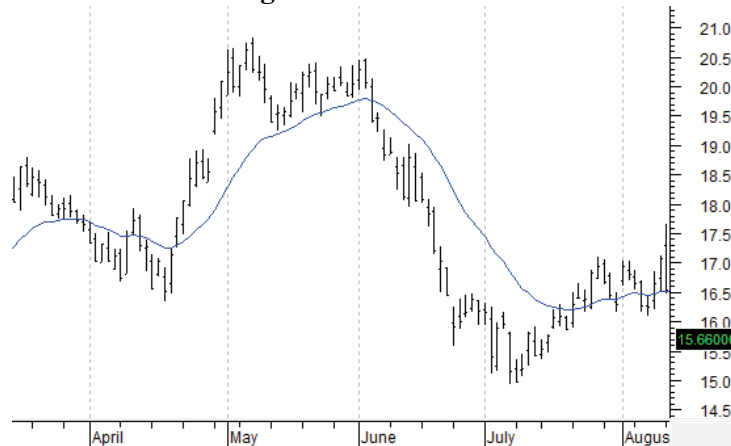
$$MMSt = \frac{1}{n} \sum_{t=1}^n p_t \quad (2.1)$$

onde n é o tamanho da janela considerada para o cálculo da média, e p é o valor do período t a ser calculado. Já a Média Móvel Exponencial atribui peso maior a cotações mais recentes, representando uma média ponderada de observações passadas.

Diversas pesquisas relativas ao uso de redes neurais aplicadas à previsão de preços utilizam as Médias Móveis. Alguns exemplos são Chang et al. (2011), Armano, Marchesi e Murru (2005) e Yang, Wu e Lin (2012). A Figura 2 exibe uma Média Móvel Simples de 20 períodos sobre um gráfico de barras. É possível notar que sua direção acompanha o movimento dos preços, com relativo atraso.

2.2.2.2 Bandas de Bollinger

As Bandas de Bollinger foram concebidas por John Bollinger nos anos 80, partindo da ideia já existente de bandas de tendências que englobam a maior parte dos preços sendo executados sobre uma ação. Sua estrutura forma envelopes que identificam momentos em que o ativo

Figura 2: Média Móvel

Fonte: Elaborado pelo autor desta dissertação.

possa estar caro ou barato, de forma relativa. Bollinger modificou de forma sutil a técnica, agregando a ela conhecimento sobre a volatilidade corrente através do uso de desvios padrões (BOLLINGER, 2001). As Bandas de Bollinger fornecem, portanto, duas informações importantes: momentos em que os preços estão atingindo níveis extremos relativamente a períodos anteriores, e o nível de dispersão desses preços. Quando as bandas superior e inferior se encontram muito próximas uma da outra, significa que a volatilidade das cotações está baixa. Quando estão afastadas, mostra uma volatilidade elevada. Quando o preço atinge a banda superior, há uma indicação de preços bastante elevados quando comparados ao seu histórico definido pelo parâmetro da média móvel; já quando atinge a banda inferior, pode-se concluir o oposto (ACHELIS, 2013).

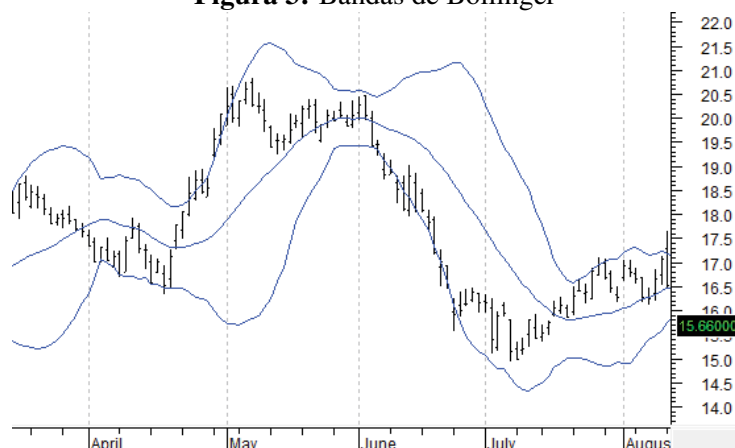
As Bandas de Bollinger consistem de uma média móvel aritmética com período parametrizável; uma banda superior, calculada pela soma de um número de desvios padrões à média móvel; e uma banda inferior, calculada pela subtração do mesmo número de desvios padrões. Na Figura 3 é possível visualizar um gráfico de barras com uma média móvel central de 20 períodos, e as bandas superior e inferior situando-se a 2 desvios padrões da mesma.

Assim como as médias móveis, as Bandas de Bollinger fazem parte da classe de rastreadores de tendência. O cálculo das bandas é definido pela seguinte equação:

$$BBs_i = MMS_i \pm d \sqrt{\frac{\sum_{j=i-n}^i (p_j - MMS_i)^2}{n}} \quad (2.2)$$

onde MMS é a Média Móvel Simples no período i , calculada a partir da parametrização do número de períodos n (tamanho da janela), d é o parâmetro indicativo do número de desvios padrões e p é o preço do ativo. O cálculo da raiz quadrada refere-se ao desvio padrão para o período considerado.

Exemplos de pesquisas que abordam o uso de Bandas de Bollinger na previsão de preços são Gomide e Milidiú (2010) e Oliveira, Nobre e Zárte (2013).

Figura 3: Bandas de Bollinger

Fonte: Elaborado pelo autor desta dissertação.

2.2.2.3 Índice de Força Relativa

O Índice de Força Relativa pertence à classe dos osciladores, indicadores que medem o momento do mercado, ou a taxa de mudança do preço. Ascendências e declínios abruptos de seus valores indicam um forte momento do mercado. Além disso, sinalizam pontos de esgotamento deste momento ao atingirem regiões extremas. A função do IFR é a de comparar a magnitude dos ganhos recentes com a das perdas recentes. Para isso, calcula a razão entre a média dos preços de fechamento dos últimos períodos em que houve valorização do ativo, e a média daqueles em que houve desvalorização. Os valores do IFR são normalizados entre 0 e 100, e quando alcançam determinados patamares costumam, segundo analistas técnicos, indicar sinais importantes. Como prática comum, valores acima de 70 são considerados mercados com excesso de demanda, podendo levar a correções de preços para baixo. Um índice abaixo de 30, por sua vez, indica excesso de oferta, sugerindo uma valorização da ação nos próximos períodos (STEVENSON, 2002).

O IFR toma como parâmetro apenas o número de períodos a serem considerados para o cálculo, e sua equação é definida como segue:

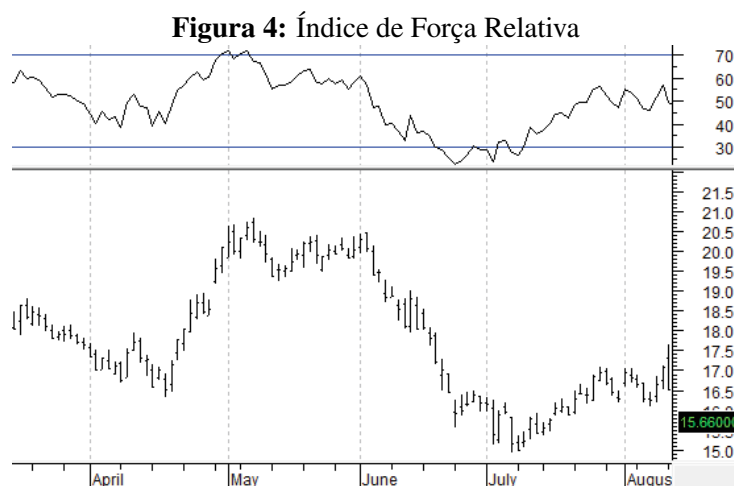
$$IFR_i = 100 - \frac{100}{1 + \frac{mediaGanhos}{mediaPerdas}} \quad (2.3)$$

onde

$mediaGanhos = (\text{total de ganhos durante os } n \text{ instantes da janela de tempo})/n$ e

$mediaPerdas = (\text{total de perdas durante os } n \text{ instantes da janela de tempo})/n$.

A Figura 4 mostra, em sua parte superior, o Índice de Força Relativa calculado para o gráfico de barras correspondente. Notam-se linhas horizontais nos níveis de 30 e 70 pontos, indicando as zonas extremas. É possível reparar que após o IFR atingir estes limites houve correções nos preços, revelando uma situação em que o indicador mostrou-se útil em sinalizar excessos de



Fonte: Elaborado pelo autor desta dissertação.

valorização e desvalorização da ação. Entre outros, os estudos de Lam (2004) e Ticknor (2013) utilizam-se do IFR na tentativa de melhorar a acurácia de suas redes neurais.

2.3 Previsão de Séries Temporais

Uma série temporal é uma sequência de observações de um dado fenômeno, o qual é observado em um espaço contínuo ou discreto do tempo (ARAÚJO, 2010). Contínuo quando as observações são realizadas de forma ininterrupta ao longo do tempo, como um impulso elétrico. Discreto quando intervalos de tempo ocorrem entre as observações. Segundo Morettin e Toloí (1987), a análise de séries temporais pode ter diversos objetivos: investigar o mecanismo gerador da série, descrever seu comportamento, procurar periodicidades relevantes nos dados, ou fazer previsões de valores futuros.

O objetivo dos modelos de previsão aplicados a uma determinada série temporal é prover um mecanismo que permita, com relativa acurácia, a previsão de valores futuros dessa série, dados valores passados da mesma. Tais mecanismos tentarão identificar certos padrões regulares no conjunto de dados, criando um modelo capaz de gerar os próximos padrões temporais (ARAÚJO, 2010). No escopo de séries financeiras, modelos de regressão, como ARIMA, são utilizados com frequência. No entanto, para sistemas não lineares que apresentam alto ruído e dinamicidade em seu comportamento, como é o caso de séries de preços de bolsas de valores, técnicas da Inteligência Artificial têm se mostrado mais adequadas, expondo larga habilidade de modelagem (NOURANI; KISI; KOMASI, 2011). Uma das ferramentas mais promissoras neste campo são as redes neurais artificiais, capazes de lidar com a não linearidade em modelos, além de serem tolerantes a ruídos nos dados. Seu funcionamento é detalhado em seções posteriores deste trabalho.

2.3.1 Discretização dos dados

No mercado de ações, as séries temporais são formadas pelo histórico de preços nos quais alguma transação foi realizada. Cada uma dessas transações dá origem a um *tick*, o qual armazena informações pertinentes ao negócio. Mercados de alta liquidez geram milhares de *ticks* a cada dia e, pela sua natureza, tais dados são irregularmente espaçados no tempo. A maioria dos estudos na literatura financeira, no entanto, lida com séries temporais homogêneas, ou seja, cujos dados são igualmente distribuídos no tempo. Isso ocorre, principalmente, devido a maior facilidade com que essa organização da informação é tratada pelo aparato estatístico já existente (GENÇAY et al., 2001).

No meio financeiro, a homogeneização destes dados se dá pelo agrupamento dos *ticks* na periodicidade desejada, de acordo com o nível de detalhe que se quer extrair da informação. Ticknor (2013) e Chang et al. (2011), por exemplo, abordam previsões diárias de preços. Neste caso, os dados são discretizados de forma que cada *tick* represente toda informação referente à movimentação de preços dentro de um dia de negociação. Martinez et al. (2009), por sua vez, adotam a divisão de dados em intervalos fixos de 15 minutos. Quando analisados desta maneira, ou seja, sobre uma frequência maior que um dia, o período é denominado intradiário. A exploração de espaços curtos de tempo permite que as implicações de novas informações e seus respectivos efeitos sobre os preços dos ativos sejam capturados o mais rápido possível, permitindo também uma reação aos mesmos de forma mais eficiente (KIM; RYU, 2014).

2.4 Redes Neurais Artificiais

Segundo Haykin (2001), uma rede neural artificial (RNA) é um processador paralelo distribuído, composto por unidades de processamento simples, com propensão natural para armazenar conhecimento e torná-lo disponível para uso. As unidades de processamento são denominadas neurônios, sendo as forças entre suas conexões chamadas de pesos sinápticos. O conhecimento é adquirido do ambiente ao qual a rede é exposta, ou seja, dos dados a ela submetidos. Através de um processo de aprendizagem, os pesos sinápticos da rede são modificados de maneira ordenada, de forma a se alcançar determinado objetivo proposto.

O poder computacional das redes neurais deriva, em parte, da sua capacidade de generalizar o conhecimento adquirido. A generalização refere-se à produção de saídas razoáveis para entradas não encontradas durante a fase de aprendizagem da rede neural. Esta característica, juntamente com sua estrutura distribuída paralela, permite às RNAs resolverem problemas complexos e de larga escala.

As grandes áreas de aplicação das redes neurais são o reconhecimento e classificação de padrões, o processamento de sinais, e o controle e análise de séries temporais. Ainda, de acordo com Kosko (1992), elas podem ser programadas ou treinadas para resolver problemas de otimização combinatória, filtrar ruídos em dados de medição e controlar problemas mal definidos.

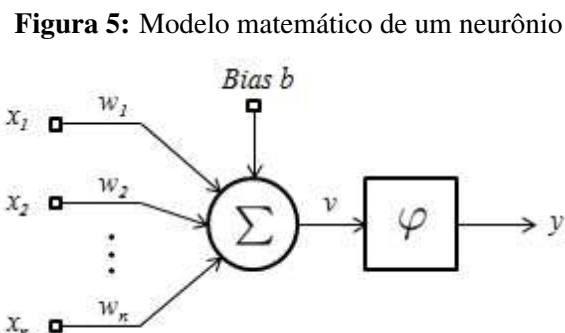
Resumidamente, as RNAs são capazes de estimar amostras de funções quando não se conhecem suas formas.

Redes neurais possuem propriedades interessantes que as permitem resolver problemas específicos. Neurônios podem ser caracterizados a fim de tratar funções não lineares, caso o mecanismo responsável pelo sinal de entrada também apresente esta natureza. Mapeamentos entre entradas e saídas podem ser construídos através do processo denominado aprendizado supervisionado, onde são apresentadas as saídas desejadas para um conjunto de entradas, permitindo à rede ajustar suas respostas na tentativa de diminuir o erro de previsão. Além disso, RNAs possuem capacidade inerente de adaptação, uma vez que podem ser facilmente retreinadas para lidar com pequenas modificações nas condições do ambiente em que opera (HAYKIN, 2001).

2.4.1 O Neurônio Artificial

O neurônio artificial é uma unidade de processamento de informação com papel fundamental para a operacionalização da rede neural. A Figura 5 apresenta a estrutura tradicional de um neurônio artificial. Cada entrada x_i corresponde a uma característica da amostra de dados sendo apresentada à rede. Os pesos w_i associados a cada entrada determinam a força do sinal de disparo da conexão, ou seja, a representatividade de cada atributo para a rede. Os valores de entrada x_i são multiplicados por seus pesos correspondentes w_i e agregados sob um combinador linear. Nesta etapa também há a inclusão de um sinal de *bias*, capaz de aumentar ou diminuir o valor líquido v de saída do acumulador (HAYKIN, 2001).

A função de ativação φ avalia o valor v e é responsável por limitar, ou não, seu intervalo de amplitude. Segundo Freeman e Skapura (1991), a escolha da função de ativação depende de como se deseja representar a saída do neurônio. Na literatura relacionada ao uso de redes neurais como previsoras de séries temporais é comum o uso das funções lineares e sigmoidais. A função linear, como o próprio nome diz, mapeia linearmente valores de entrada em valores de saída e é apresentada conforme a equação abaixo:



Fonte: Elaborado pelo autor desta dissertação, com base em Haykin (2001).

$$\varphi(v) = \alpha v \quad (2.4)$$

onde α é um número real que define a saída linear para o valor de entrada v . Geralmente, para problemas de aproximação de funções utiliza-se a função de ativação linear nos neurônios da camada de saída da rede neural.

Já as funções sigmoidais utilizadas nas camadas intermediárias das redes neurais são as responsáveis por capacitá-las a resolverem problemas não lineares no espaço de entrada (BRAGA; CARVALHO; LUDERMIR, 2014). Elas podem ser encontradas de diversas formas, sendo a tangente hiperbólica utilizada com frequência. Tal função limita a saída a valores entre -1 e 1, permitindo que a média entre todos os neurônios seja próxima a zero. De acordo com Haykin (2001), esta característica traduz-se em uma convergência mais rápida de aprendizado em redes com grande número de conexões sinápticas e que empregam o algoritmo de retro-propagação. A função tangente hiperbólica é definida como segue:

$$\varphi(v) = \frac{2}{1 + e^{-2v}} - 1 \quad (2.5)$$

onde o parâmetro v representa o valor de entrada da função. Quando a função é associada a um neurônio, v origina-se do resultado do acumulador descrito anteriormente.

Finalmente, a representação matemática do funcionamento de um neurônio artificial pode ser formalmente definida pela seguinte equação:

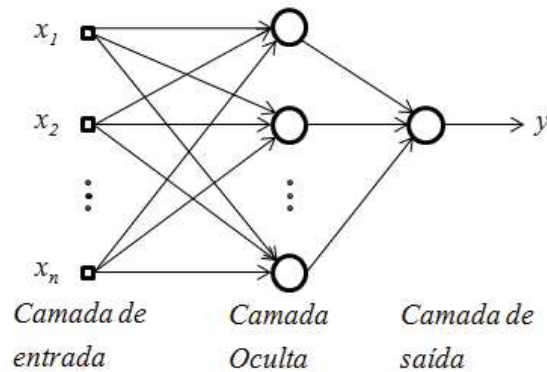
$$y = \varphi\left(\sum_{i=1}^n x_i w_i + b\right) \quad (2.6)$$

onde y representa o valor de saída do neurônio, φ a função de ativação, n define o número de entradas do neurônio em questão, w_i é o peso associado ao valor de entrada x_i , e b é o *bias*.

2.4.2 Arquitetura da Rede Neural Artificial

O neurônio artificial é a unidade elementar de uma rede neural artificial (RNA). Neurônios podem ser organizados em camadas, de forma a se configurar arquiteturas específicas. A arquitetura classifica a rede neural e restringe o tipo de problema no qual ela poderá ser utilizada. É definida pelo número de camadas, pelo número de neurônios em cada camada e pelo tipo de conexão entre os nós, as quais remetem à sua topologia (BRAGA; CARVALHO; LUDERMIR, 2014). A Figura 6 apresenta uma representação simplificada de uma RNA.

Quanto ao número de camadas, as RNAs podem ser classificadas como de camada única ou de múltiplas camadas. As redes de camada única são aquelas onde existe apenas um nó entre qualquer entrada e qualquer saída da rede. Suas limitações estão no fato de que só conseguem trabalhar com a classificação de padrões linearmente separáveis (MITCHELL, 1997), portanto de pouca utilidade a problemas reais, de maior complexidade. As redes mais comumente utilizadas são as de múltiplas camadas, também conhecidas como *Multilayer Networks*,

Figura 6: Representação simplificada de uma rede neural artificial

Fonte: Elaborado pelo autor desta dissertação.

ou *Multilayer Perceptron* (MLP). Estas apresentam uma camada de entrada, uma ou mais camadas ditas ocultas, e uma camada de saída. A camada de entrada armazena um número suficiente de neurônios representativos das diferentes variáveis, relacionadas ao domínio ou aplicação, a serem utilizadas pela rede, e tais neurônios apenas armazenam e propagam os estímulos externos recebidos. Já as camadas ocultas permitem que a rede aprenda tarefas complexas, extraindo características mais significativas a partir dos padrões de entradas. O tamanho desta camada é dependente do problema que se está tratando, e experimentações geralmente podem ajudar a decidir o número de neurônios que apresentam os melhores resultados. Em relação à camada de saída, o número de neurônios normalmente também é definido de acordo com o problema. Para aqueles de classificação, por exemplo, cada neurônio representa uma classe. Já para problemas de previsão de valores, um único neurônio na camada de saída pode ser suficiente, ao expor o valor previsto.

Quando a RNA não apresenta ciclos entre as conexões de seus neurônios, ela é classificada como do tipo *feedforward*. Caso a saída de algum neurônio na i -ésima camada seja usada como entrada para neurônios em camadas de índice menor ou igual a i , a rede é denominada de cíclica, recorrente, ou também com *feedback*. Tais tipos de conexão são especialmente úteis para a representação de tempo de forma implícita pela rede neural, quando valores previstos são entradas importantes para previsões seguintes (HAYKIN, 2001).

2.4.3 Aprendizado

Redes neurais são sistemas dinâmicos capazes de serem treinados para estimar funções de entrada e saída. Dados amostrados dão forma e programam sua evolução no tempo. Diferentemente de estimadores estatísticos, as RNAs estimam uma função sem um modelo matemático de dependência das saídas em relação às entradas. Elas aprendem a partir de experiências com dados numéricos e, por vezes, linguísticos (KOSKO, 1992). Os métodos de aprendizado das redes neurais são classificados em dois paradigmas principais: supervisionado e não supervisionado.

O aprendizado não supervisionado trata de problemas de agrupamento e redução de dimensionalidade, utilizando-se apenas dos dados fornecidos como entrada da rede. Já no aprendizado supervisionado, os dados de entrada a serem utilizados pelo algoritmo para aprendizagem são apresentados juntamente com seus valores esperados de saída. É o caso, por exemplo, dos problemas de classificação, como reconhecimento de dígitos, e dos problemas de regressão, como a previsão de preços no mercado de ações (BISHOP, 2007).

A técnica de aprendizado supervisionado mais comumente utilizada com redes do tipo MLP é o algoritmo de retro-propagação, ou *backpropagation*, cujo funcionamento é descrito a seguir. Um padrão de entrada é apresentado à primeira camada da RNA, que repassa o sinal aos neurônios da camada seguinte por meio das conexões sinápticas, já computando seus respectivos pesos e funções de ativação. O processo é repetido por todas as camadas, até que a saída desse ciclo seja gerada. Esta fase é chamada de propagação. Na camada de saída, o valor resultante em cada um de seus neurônios é comparado ao resultado esperado através de uma função custo. Uma das mais utilizadas é o erro médio quadrático (*mean squared error*, MSE), cuja fórmula é apresentada a seguir:

$$\varepsilon = \frac{1}{2N} \sum_{n=1}^N \sum_{j \in C} (d_{j,n} - y_{j,n})^2 \quad (2.7)$$

onde N é o número de amostras para as quais o erro está sendo calculado, C representa o conjunto de neurônios j na camada de saída, $d_{j,n}$ é o valor esperado do j -ésimo neurônio para a n -ésima amostra apresentada à rede, e $y_{j,n}$ é o valor obtido pela rede do j -ésimo neurônio para a n -ésima amostra. O objetivo do processo de aprendizagem das redes neurais é ajustar os parâmetros da rede de forma que este erro seja minimizado (HAYKIN, 2001).

Para a minimização do erro, geralmente adota-se o método de otimização *Gradiente Descendente*, que segundo Freeman e Skapura (1991) procura modificar os pesos sinápticos na direção da maior taxa de decaimento da função erro. Este resultado é alcançado calculando-se recursivamente os termos dos gradientes locais que representam a participação de cada neurônio no erro da rede, da seguinte forma:

$$\delta_j^{(l)} = \begin{cases} \varepsilon_j^{(L)}(n) \varphi_j'(v_j^{(L)}(n)) & \text{para o neurônio } j \text{ da camada de saída } L \\ \varphi_j'(v_j^{(l)}(n)) \sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n) & \text{para o neurônio } j \text{ da camada oculta } l \end{cases}$$

onde $\delta_j^{(l)}$ representa o gradiente local para o neurônio j da camada l , ε é o erro calculado sobre o j -ésimo neurônio da camada de saída L , φ' é o resultado da derivada da função de ativação, tomando como entrada o valor v , w_{kj} é o peso da conexão sináptica entre os neurônios j e k , e n é o número da amostra apresentada à rede.

Com o gradiente local calculado, inicia-se a fase de retro-propagação, onde o sinal do erro é transmitido de volta à camada imediatamente anterior. No entanto, cada neurônio dessa camada recebe apenas uma porção do erro total gerado, baseado na sua contribuição para o valor de

saída original. O mesmo acontece para todas as camadas anteriores, até a camada de entrada. Com base nos erros recebidos, os pesos das conexões são atualizados, convergindo a rede para um estado que reconheça todos os padrões de treinamento a ela apresentados, caracterizando o aprendizado da mesma. O cálculo dos pesos atualizados nesta etapa é realizado através da seguinte formulação:

$$w_{j,i}^{(l)}(n+1) = w_{j,i}^{(l)}(n) + \eta \delta_j^{(l)}(n) y_i^{(l-1)}(n) \quad (2.8)$$

onde $w_{j,i}^{(l)}(n)$ é o valor atual do peso entre os neurônios i e j do conector de camadas l , $\delta_j^{(l)}$ é o gradiente local e $y_i^{(l-1)}$ é o valor de saída do neurônio i do conector de camadas $l-1$. η refere-se à taxa de aprendizado a qual a rede é submetida, e segundo Haykin (2001), precisa ser selecionada de forma cuidadosa, uma vez que determina a estabilidade e a convergência do algoritmo.

O processo iterativo e de aprendizado é caracterizado pela repetição do ciclo de propagação e retro-propagação até que determinado critério de parada seja atingido. Alguns dos mais utilizados são o número de épocas ou iterações do algoritmo, e o atingimento de um erro máximo ou uma taxa mínima de mudança do gradiente. Após a finalização do treinamento, é esperado que a RNA responda com uma saída ativa, caso novas entradas contenham padrões que remetam a características aprendidas dos indivíduos utilizados no treinamento. Do contrário, caso não existam padrões parecidos, as saídas devem inibidas (FREEMAN; SKAPURA, 1991).

2.4.4 Regularização Bayesiana

Um problema muito comum com algoritmos de aprendizagem supervisionada por meio de retro-propagação e Gradiente Descendente é seu potencial de *overfitting* (FORESEE; HAGAN, 1997), quando até mesmo ruídos são considerados na formulação da função subjacente. Conforme aumenta o tamanho da rede, através da adição de novos neurônios nas camadas ocultas, maior também é a chance de ocorrência deste tipo de situação, a qual leva a uma perda de generalização por parte da rede, quando treinada excessivamente. Uma forma de evitar este problema é o uso de um conjunto de validação, através do qual pode-se determinar um ponto de parada do treinamento quando se detecta o início da ocorrência de *overfitting* (TICKNOR, 2013).

Já uma alternativa que não exige o passo adicional de validação é a introdução da Regularização Bayesiana ao algoritmo de treinamento. De acordo com Burden e Winkler (2009), em redes neurais capacitadas com esta técnica, seu treinamento conduz à descoberta e evidência de ligações entre neurônios efetivamente relevantes para a formação do conhecimento, sendo que ligações desnecessárias têm seus pesos conduzidos a zero. Geralmente, este número efetivo de conexões é consideravelmente menor que em uma rede neural padrão totalmente conectada. Ainda, segundo Foresee e Hagan (1997), a ideia geral da regularização é manter pesos pequenos para as redes neurais, reduzindo o potencial de *overfitting*, de forma que elas possam produzir

saídas mais suaves, aumentando sua capacidade de generalização.

Em geral, o passo de treinamento das RNAs foca na redução do erro médio quadrático entre as saídas do modelo e seus respectivos valores esperados, através da equação 2.7. A Regularização Bayesiana adiciona um termo extra a essa equação:

$$F = \beta\varepsilon + \alpha\varepsilon_w \quad (2.9)$$

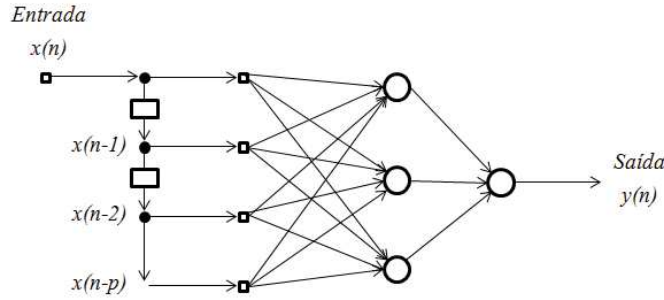
onde ε_w representa a soma dos quadrados dos pesos da rede, e α e β são parâmetros da função objetivo F . O tamanho relativo de cada um desses parâmetros determina a ênfase do treinamento. Se $\alpha \ll \beta$, o algoritmo terá como foco a diminuição do erro médio quadrático. Já se $\alpha \gg \beta$, o treinamento dará ênfase à redução dos pesos, produzindo saídas de forma mais generalizada (FORESEE; HAGAN, 1997). A otimização dos parâmetros de regularização α e β requer a resolução da matriz Hessiana de $F(w)$, onde w é o vetor de pesos da rede, no ponto mínimo w^{MP} , para a qual Foresee e Hagan (1997) propuseram uma aproximação de Gauss-Newton. Esta técnica reduz a chance de se alcançar um mínimo local, aumentando a capacidade de generalização da rede.

Redes neurais com Regularização Bayesiana foram estudadas por Ticknor (2013) para prever preços no mercado financeiro, apresentando resultados bastante interessantes. Na pesquisa, o modelo proposto com esta técnica mostra desempenhos tão bons quanto modelos mais avançados, sem a necessidade de pré-processamento de dados, testes de sazonalidade ou análises de ciclos. Também Ardalani-Farsa e Zolfaghari (2010) empregam a técnica com sucesso na previsão de séries temporais caóticas.

2.4.5 Focused Time Lagged FeedForward Networks

O tratamento do tempo, seja ele contínuo ou discreto, é de fundamental importância no processo de aprendizagem de diferentes tarefas cognitivas, como a visão, a fala, o processamento de sinais e o controle motor (HAYKIN, 2001). Também na previsão de preços de ações, o tempo tem papel central. Como visto em seções anteriores, segundo a Análise Técnica, padrões de movimento de preços identificados no passado tendem a apresentar comportamento semelhante no futuro. É interessante, portanto, adaptar as redes neurais de forma que sejam capacitadas a captar comportamentos de séries temporais e, conseqüentemente, estejam aptas a solucionar problemas desta natureza.

Uma das formas de se alcançar esse objetivo é através da representação implícita do tempo na forma de sinal de entrada de uma rede neural, inserindo a estrutura temporal deste sinal na estrutura espacial da rede. Além disso, para captar o conceito de tempo a rede neural deve possuir um certo tipo de memória. Uma forma de introduzir esta ideia à rede é através da substituição dos neurônios na camada de entrada por uma estrutura chamada de *memória de linha de atraso derivada*. Através dela o sinal de entrada é decomposto em atrasos, os quais formam, por sua vez, novas entradas para a rede. RNAs deste tipo são denominadas *Focused Time Lagged Feed-*

Figura 7: *Focused Time Lagged FeedForward Networks*

Fonte: Elaborado pelo autor desta dissertação, com base em Haykin (2001).

Forward Networks. O termo *focused* advém do fato de que a estrutura interna de memória está focada apenas na camada de entrada da rede, ao contrário de outros modelos onde os atrasos aparecem também em camadas intermediárias (PRÍNCIPE; EULIANO; LEFEBVRE, 2000).

Além da memória, a qual mantém informações relevantes sobre o passado, outra estrutura importante para que a rede neural consiga prever padrões temporais é um associador capaz de relacionar os diferentes passos no tempo. Nas FTLFNs, este associador é uma rede MLP estática, exibida na Figura 7, juntamente com a estrutura de memória (COULIBALY; EVORA, 2007).

Dado um sinal de entrada consistindo do valor presente $x(n)$ e de p valores passados $x(n-1), \dots, x(n-p)$ armazenados na memória de linha de atraso, os pesos da rede neural são ajustados para minimizar o erro médio quadrado entre a saída da rede, $y(n)$, e a resposta desejada (HAYKIN, 2001). A saída desta rede, para uma rede hipotética de camada oculta única, é calculada da seguinte maneira:

$$y(n) = \sum_{j=1}^{m_l} w_j y_j(n) + b_0 \quad (2.10a)$$

$$y_j(n) = \varphi\left(\sum_{l=0}^p w_j(l)x(n-l) + b_j\right) \quad (2.10b)$$

onde o conjunto $\{w_j\}_{j=1}^{m_l}$ representa os pesos sinápticos do neurônio de saída, m_l é o tamanho da camada oculta, b é o valor de bias, $y_j(n)$ é a saída do neurônio j , p é o número de atrasos considerados para a memória, x é o valor do sinal, e φ é a função de ativação.

O tamanho da memória, ou seja, o número de atrasos necessários para descrever as características do padrão de entrada, precisa ser determinado de acordo com o problema que se está tratando, através de diferentes experimentações, assim como acontece com outros parâmetros das MLPs. Não existe uma quantidade fixa, adaptada a toda e qualquer situação. Pelo fato da memória se encontrar apenas na camada de entrada da rede neural, é possível utilizar o algoritmo de *backpropagation* padrão para treiná-la (HAYKIN, 2001).

O uso de FTLFNs é bastante difundido na literatura, apesar de não haver consenso quanto

a esta nomenclatura. Costa e Crepaldi (2014) utilizam FTLFN, como no presente trabalho. Coulibaly e Evora (2007) adotam a terminologia TLFN (*Time-Lagged FeedForward Network*), e Ouazzane et al. (2012) usam FTDNN (*Focused Time-Delay Neural Network*). Existem, também, diversos trabalhos onde este tipo de rede neural não recebe qualquer nomenclatura em particular, apesar da especificação da rede ser compatível com as FTLFNs, como é o caso em Geva e Zahavi (2014).

2.5 Considerações

Neste capítulo foram apresentados conceitos sobre o Mercado de Capitais e a Bolsa de Valores que colaboram para a compreensão do problema sendo abordado no trabalho proposto. Descreveram-se os tipos de análises de mercado, com ênfase em três indicadores da Análise Técnica a serem utilizados nos experimentos: Médias Móveis, Índice de Força Relativa e Bandas de Bollinger.

Foram apresentadas características de séries temporais, assim como modelos utilizados historicamente na previsão das mesmas. Com foco maior no objetivo principal deste trabalho, relatou-se a forma como elas são discretizadas no meio financeiro, de acordo com o nível de detalhe que se quer extrair da informação.

As redes neurais artificiais foram apresentadas de forma teórica através dos principais conceitos e formulações matemáticas envolvidas, detalhando-se sua estrutura e processo de aprendizado. Por fim, foram analisadas particularidades da FTLFN, rede temporal utilizada neste trabalho.

A revisão realizada tem como objetivo permitir um melhor entendimento dos métodos e técnicas aplicados no presente estudo e nos trabalhos relacionados, a serem apresentados no próximo capítulo.

3 TRABALHOS RELACIONADOS

A gama de estudos envolvendo a aplicação de redes neurais artificiais a problemas de previsões de séries temporais é ampla e diversificada. Neste capítulo são apresentados quatro trabalhos recentes sobre o assunto, cujas abordagens e características se assemelham de alguma forma a esta dissertação.

Costa (2012) investigou em sua dissertação o desempenho de redes neurais temporais do tipo FTLFN na previsão de séries financeiras, focando-as sobre ativos do mercado acionário brasileiro. Os resultados foram comparados com aqueles produzidos por outro modelo de rede, as NARX (*Nonlinear Autoregressive Network with Exogenous Inputs*), cuja principal diferença se encontra na recorrência do sinal de saída. Como extensão deste trabalho, Costa e Crepaldi (2014) aprofundaram a discussão sobre o tema, expondo novas potencialidades das redes FTLFN. No estudo, desenvolveram um método capaz de reverter de forma otimizada dados transformados pelo método Box-Cox, tirando proveito da memória de curto prazo deste tipo de rede neural.

O trabalho de Gomide e Milidiú (2010) apresentou modelos de redes neurais com atrasos, caracterizando-se em FTLFNs, capazes de prever preços em operações específicas do mercado, destacando a efetividade de se introduzir dados intradiários no processo. De forma análoga, o estudo de Silva et al. (2014) demonstrou especial interesse na aplicação de redes neurais artificiais sobre dados do mercado de alta frequência, buscando identificar oscilações positivas nos preços em intervalos curtos de tempo. Na tentativa de melhorar as capacidade preditivas, em ambos os trabalhos foram utilizados atributos gerados a partir de indicadores da Análise Técnica.

Ao final do capítulo é realizada uma breve discussão sobre os trabalhos apresentados. São pontuadas as características mais interessantes em cada um deles, a serem utilizadas também nesta dissertação. Da mesma forma, são debatidas possíveis melhorias nos estudos.

3.1 FTLFN sobre séries temporais financeiras

O trabalho de Costa (2012) possui como foco a comparação entre os desempenhos de dois tipos de redes neurais temporais, FTLFN e NARX, na previsão de séries temporais financeiras. As NARX são redes neurais dinâmicas e recorrentes, cujo vetor de sinal aplicado à camada de entrada da RNA consiste de uma janela de dados constituída dos seguintes componentes: valores presentes e passados da entrada exógena, originária de fora da rede, e valores atrasados do sinal de saída, sobre os quais é feita a regressão da saída do modelo (HAYKIN, 2001). Redes recorrentes são arquitetadas de forma a realizar previsões de múltiplos passos no futuro, através do uso do sinal de saída como entrada da rede no passo seguinte de previsão. Este processo de realimentação é chamado de *feedback* e supostamente aumenta a capacidade de mapeamento não linear da rede. As FTLFNs, por sua vez, também utilizam-se de atrasos na camada de

entrada, porém não fazem uso da recorrência. É importante observar que, neste estudo, as FTLFNs foram chamadas de FTDNNs.

A partir de critérios definidos previamente, como volume e frequência de negociação, os movimentos dos preços diários das ações do Banco do Brasil (BBAS3) e do Bradesco (BBDC3 e BBDC4) foram selecionados como séries temporais a serem previstas, entre os anos de 2009 e 2012. O intervalo corresponde a 735 observações para cada série. Além das experimentações com preços, também foram feitas análises com transformações dos dados, sendo elas as diferenças de primeira ordem, os retornos logarítmicos e a transformação Box-Cox. O autor argumenta que esta abordagem se justifica pela busca de estacionariedade e gaussianidade aproximada das séries temporais. Segundo alguns estudos, como Heravi, Osborn e Birchenhall (2004) e Nelson et al. (1999), tais características podem ajudar as redes na melhora de suas previsões.

Após testes com diferentes topologias sobre séries temporais diversas, definiram-se configurações a serem adotadas pelas redes FTLFN e NARX na previsão dos dados. Como o objetivo foi de comparar seus desempenhos, o autor procurou adotar parâmetros semelhantes entre ambas arquiteturas. Ficou estipulado o uso de duas camadas ocultas, com 10 neurônios em cada uma delas. Para ativação dos neurônios nesta camada, foi utilizada a função tangente hiperbólica. Já para a camada de saída, adotou-se a função de transferência linear. A rede NARX diferenciou-se por receber em sua camada de entrada um *feedback* com dois períodos de atraso.

O treinamento das redes ocorreu com 80% dos dados disponíveis, sendo o conjunto de teste formado pelos 20% restantes. Foram realizadas previsões para cinco modalidades de preços de cada série: abertura, fechamento, máximo, mínimo e preço médio para cada dia de negociação. Os resultados mostraram que, de forma geral, tanto a rede FTLFN quanto a NARX são eficientes no uso do fator temporal das séries para prever valores futuros. No experimento sobre dados originais (sem transformação), as NARX apresentaram MAPE (*Mean Absolute Percentage Error*) de 1,731, mostrando relativa superioridade em relação à FTLFN, cujo resultado foi de 2,472. Ainda assim, ambas apresentaram erros médios de previsão menores que o modelo estatístico ARMA-GARCH, largamente utilizado como referência em análises e previsões de séries temporais financeiras. Já nos testes de previsão sobre os dados transformados, os erros médios apresentados pela FTLFN foram levemente inferiores aos da NARX, conforme a Tabela 1.

Tabela 1: MAPE médio sobre dados transformados

Rede	Diferenças	Retornos	Box-Cox
NARX	1,512	1,353	1,900
FTLFN	1,327	1,208	1,857

Fonte: Elaborado pelo autor desta dissertação, com base em Costa (2012).

O artigo de Costa e Crepaldi (2014) complementa o trabalho de dissertação descrito acima. Nele, deseja-se prover um método para reverter dados transformados via Box-Cox ao seu formato original, através de redes neurais FTLFN. Segundo os autores, quando esta reversão é

realizada por métodos estatísticos sobre dados que não possuem características gaussianas, o processo está sujeito a um determinado viés, levando a erros. Empregando-se a FTLFN, que não requer qualquer suposição sobre a distribuição dos dados transformados, buscou-se eliminar este viés. Outra justificativa apresentada pelos autores para o uso deste tipo de rede neural é que ela mantém uma memória de curto prazo, levando em consideração possíveis dependências dos dados a valores passados.

A topologia utilizada para a rede foi a mesma adotada na dissertação, assim como o número de atrasos do sinal de entrada, fixados em 8. Os dados de entrada da FTLFN consistiram de valores transformados por Box-Cox, tendo como alvo a previsão de valores originais, sem qualquer tipo de transformação. Foram utilizadas 40 séries temporais com diferentes características de tamanho, frequência, estacionariedade e normalidade, incluindo preços diários de ações e índices da bolsa de valores. Em todas elas, independente de suas particularidades, os resultados mostraram que a reversão da transformação dos dados por FTLFN apresentou MAPE menor em comparação com sua reversão algébrica. Confirmou-se, então, a capacidade deste tipo de rede neural temporal em aproximar valores desconhecidos, quaisquer que sejam as distribuições do dados.

3.2 Previsões no período intradiário através de redes neurais com atrasos no tempo

Em Gomide e Milidiú (2010) foram apresentados modelos de redes neurais artificiais para previsões em bolsas de valores, fazendo uso de dados dos períodos diários e intradiários. Como forma de suporte a esta tarefa e melhoria na acurácia de predição foram construídos indicadores da Análise Técnica a partir de dados históricos, utilizados como vetores de entrada da rede juntamente com os preços. Segundo os autores, estas ferramentas são muito informativas para negociações de ações.

No entanto, ao invés de prever preços, diretamente, os autores focaram nas operações denominadas *Pairs Trading* (VIDYAMURTHY, 2004). São basicamente transações entre dois ativos que mostram correlação nos padrões de variação de seus preços. Nesta estratégia, busca-se o ganho explorando-se o *spread* entre os preços, ou seja, a razão entre eles. Quando dois ativos correlacionados mantêm variações percentuais similares, seu *spread* não sofre grandes alterações. Porém, quando este valor apresenta mudanças maiores surge a oportunidade de, simultaneamente, comprar um dos ativos e vender o outro. O objetivo da construção das redes neurais, neste trabalho, foi o de identificar *spreads* máximos e mínimos entre um par de ações, dentro de um dia.

Duas abordagens foram exploradas para o problema. Na primeira, somente dados diários passados foram utilizados na tentativa de se prever os *spreads* para o dia seguinte. Já no segundo experimento, os dados passados foram complementados com valores do período intradiário conhecidos até o momento, focando a diminuição do erro de previsão para o mesmo dia. Neste

caso, foram realizadas previsões de hora em hora, totalizando 7 (das 10:00 às 17:00) no decorrer do dia.

Os pares de ações escolhidos para observações dos *spreads* foram Petrobras ON (PETR3) e Petrobras PN (PETR4), Usiminas PN (USIM5) e Gerdau PN (GGBR4), e Bradesco PN (BBDC4) e Bradespar PN (BRAP4). Foram obtidos dados históricos referentes aos preços de abertura, máximo, mínimo e de fechamento destes papéis para o período compreendido entre 18/12/2008 e 14/05/2010, correspondendo a 343 dias de negociações. Os indicadores da Análise Técnica utilizados foram: Média Móvel Simples, Média Móvel Exponencial, Índice de Força Relativa e Bandas de Bollinger.

A rede neural adotada para os experimentos foi uma MLP de três camadas, treinada com o algoritmo de *backpropagation*. É importante notar que apesar dos autores não empregarem qualquer nomenclatura específica para a rede, foram utilizados atrasos no sinal de entrada da mesma, configurando uma rede temporal do tipo FTLFN. Testes foram realizados para determinação da janela de entrada (número de atrasos) e do conjunto de atributos que produzissem as maiores acurácias de previsão. Verificaram-se os melhores resultados com o uso de 5 atrasos sobre o sinal de entrada dos preços diários máximo e mínimo, juntamente com os dados gerados pelos indicadores da Análise Técnica Média Móvel Exponencial e Bandas de Bollinger. A RNA utilizada no período intradiário diferenciou-se por receber, além destes dados, os valores dos *spreads* máximo e mínimo atingidos no dia corrente até o momento da previsão.

Para apresentação dos resultados, os autores empregaram a métrica MAPE, assim como no trabalho de Costa (2012). Quando comparadas com outras técnicas, as FTLFNs mostraram desempenho bastante satisfatório na previsão dos *spreads* diários, de acordo com a estratégia definida inicialmente. Concluiu-se, ainda, que a inclusão de atributos relativos ao período intradiário levou a um impacto muito positivo nas acurácias de previsão. A Tabela 2 apresenta os percentuais MAPE alcançados para cada par de ações, comparando as duas periodicidades abordadas.

Tabela 2: Comparação de previsões diárias e intradiárias

	Diário		Intradiário	
	Min MAPE	Max MAPE	Min MAPE	Max MAPE
PETR4 x PETR3	0,76	1,05	0,53	0,83
USIM5 x GGBR4	1,06	1,14	0,93	1,38
BBDC4 x BRAP4	3,41	4,05	3,26	3,41

Fonte: Elaborado pelo autor desta dissertação, com base em Gomide e Milidiú (2010).

3.3 Indicadores da Análise Técnica como suporte a previsões de alta frequência

O trabalho de Silva et al. (2014) teve como objetivo identificar oscilações positivas de preços dentro de intervalos de 5, 10 e 15 minutos, a fim de obter lucros com operações sobre ativos

financeiros dentro do período de um dia de negociação, também chamadas de *day trade*. Para tanto, empregou redes neurais artificiais, fazendo uso de indicadores da Análise Técnica como auxiliares preditivos. Segundo os autores, a abordagem sobre o intradiário é importante para a inserção mais frequente de ordens de negociações, trazendo liquidez ao mercado e contribuindo com o campo de *trades* de alta frequência. Além disso, argumentou-se a favor do trabalho como uma forma útil de suporte ao processo de formação de mercado.

No contexto desta pesquisa, as redes neurais não possuem função de previsão numérica, sendo no entanto utilizadas como classificadores. Segundo Haykin (2001), neste tipo de tarefa o objetivo é relacionar cada vetor de entrada a uma categoria específica dentro de um universo finito discreto. Poucas modificações são necessárias na arquitetura da rede. A principal diferença está na função de ativação dos neurônios de saída, que neste caso costumam ser não lineares.

Na pesquisa de Silva et al. (2014) foi empregada uma RNA do tipo MLP treinada com algoritmo de *backpropagation*. Sua topologia apresenta três camadas, sendo o número de neurônios na camada oculta dependente da quantidade de atributos de entrada da rede, de acordo com o teste sendo realizado. Já a camada de saída foi configurada com apenas um neurônio, capaz de identificar a existência ou não de uma oscilação positiva no preço do ativo.

Os dados utilizados neste artigo referem-se às empresas Banco Bradesco (BBDC4), Banco do Brasil (BBAS3), Usiminas (USIM5), Petrobras (PETR4) e Vale (VALE5), relativos aos preços de abertura, máximo, mínimo e de fechamento de cada uma dessas ações, nos intervalos de 5, 10 e 15 minutos. Estes três diferentes períodos foram estudados como forma de se avaliar a influência da granularidade temporal na previsão de flutuações de preços de um ativo. O conjunto de dados compreende 26 dias de negociações entre novembro e dezembro de 2013.

O estudo de Silva et al. (2014) prestou elevada importância ao uso de indicadores da Análise Técnica como suporte para previsões de preços. Argumentou-se que indicadores com alto grau de correlação em relação aos movimentos de ativos financeiros tendem a gerar bons modelos preditivos. Os autores optaram por utilizar 10 deles para geração de valores de entrada para a RNA, entre os quais o Índice de Força Relativa, as Médias Móveis e as Bandas de Bollinger.

Como forma de divisão dos dados, foram adotados valores do início de cada dia até as 14:00 como base de treinamento, e o restante para testes. Os atributos de entrada da RNA foram compostos pelos preços no momento imediatamente anterior à previsão, juntamente com os valores de cada indicador da Análise Técnica no mesmo instante de tempo. O vetor de saída foi constituído por valores binários, onde 1 indica que o preço do ativo sofrerá uma valorização de R\$ 0,04 (quatro centavos de real) dentro do intervalo desejado, e 0 sinaliza o oposto. Apesar deste valor fixo representar diferentes grandezas percentuais para cada ação, ele foi escolhido propositalmente para demonstrar o bom desempenho das redes neurais mesmo quando a expectativa de oscilação varia em função do preço do papel. Além do mais, os quatro centavos são suficientes para cobrir os custos das operações de alta frequência.

Nos resultados foram apresentadas as acurácias de classificação, comparando as diferentes

ações e periodicidades adotadas. Foram verificadas taxas de classificações corretas de até 83%, como foi o caso para os papéis da Usiminas. Em relação ao tempo utilizado para previsão foram observados os melhores resultados dentro da janela de 5 minutos em todos os casos. Na Tabela 3 é possível verificar os percentuais de acertos para as duas ações com melhores desempenhos, segmentados conforme o período de teste.

Tabela 3: Acurácia de classificação

Ativo	Período (minutos)	Acurácia
BBAS3	15	63,88%
	10	65,83%
	05	77,58%
USIM5	15	70,57%
	10	76,41%
	05	83,86%

O estudo mostrou que o uso combinado de preços e indicadores da Análise Técnica conseguem capturar de forma eficiente tendências de mercado, gerando resultados bastante satisfatórios especialmente em espaços curtos de tempo. A estratégia utilizada proporciona aos formadores de mercado a possibilidade de inserir grandes quantidades de ordens de compra e venda, uma vez que elas são acionadas com mais frequência, aumentando a liquidez do mercado.

3.4 Discussão

Como visto em três dos quatro trabalhos correlatos acima descritos, as FTLFNs apresentam-se como ferramentas de grande potencial preditivo quando aplicadas sobre séries temporais financeiras. Costa (2012) mostrou que as FTLFNs conseguem prever preços de ações apresentando erros menores que modelos estatísticos conhecidos, como ARMA-GARCH. Quando os dados foram transformados, a capacidade preditiva da rede aumentou ainda mais, mostrando resultados levemente superiores também às NARX. Já Costa e Crepaldi (2014) empregaram com sucesso as FTLFNs na reversão de transformações de dados pela técnica Box-Cox. Segundo os autores, a memória de curto prazo oferecida por este tipo de rede neural foi crucial para que sua aplicabilidade mostrasse resultados mais otimizados em relação às reversões algébricas. E Gomide e Milidiú (2010) demonstraram um bom desempenho das FTLFNs na previsão de preços diários e intradiários, quando são agregados indicadores da Análise Técnica ao modelo. Os resultados promissores apresentados pelos estudos foram de grande motivação para o desenvolvimento da presente dissertação, uma vez que no trabalho proposto também há a intenção de se adotar redes do tipo FTLFN como previsoras de preços de ações.

Apesar das FTLFNs serem empregadas com sucesso sobre séries temporais financeiras, há uma carência na literatura por estudos que a examinem mais detalhadamente na previsão de preços de ativos no período intradiário. Uma extensa revisão dos estudos publicados nesta área

evidenciou a dificuldade em se encontrar tais pesquisas. Os trabalhos aqui relatados são algumas das poucas exceções com relativo relacionamento ao tema. Gomide e Milidiú (2010) abordaram a previsão de preços por FTLFNs, porém o fizeram com foco maior sobre os resultados possíveis a serem alcançados para uma estratégia específica de negociação de mercado. Silva et al. (2014) também trouxeram questões importantes sobre o período intradiário, expondo a contribuição positiva da inclusão de dados deste tipo ao modelo, aumentando sua acurácia. Sua rede neural, no entanto, foi adaptada à classificação de padrões, e não à aproximação de valores, como nos outros trabalhos correlatos. Em nenhum deles, portanto, há uma discussão e exploração mais abrangente, capaz de ser reutilizada em diferentes contextos. O estudo aqui proposto pretende ampliar a literatura sobre previsões no período intradiário, dada a sua importância no contexto atual de mercado. Diferentemente de Gomide e Milidiú (2010), no entanto, não procura identificar possíveis retornos sobre estratégias específicas. Ao contrário, busca descobrir potencialidades do uso das redes FTLFN sobre dados intradiários de maneira geral, permitindo que sejam utilizadas em situações diversas.

A principal característica das FTLFNs é sua memória de curto prazo implementada na camada de entrada, capaz de prover subsídios concretos para previsões de séries temporais, baseado em históricos de valores. Ainda assim, muitos estudos que adotam este tipo de rede neural, incluindo os aqui expostos, por vezes exploram de forma vaga esta potencialidade. Como exemplo, no trabalho de Costa (2012) o número de atrasos no sinal de entrada foi determinado de forma bastante genérica, tomando uma média dos desempenhos de dois tipos de redes diferentes. Costa e Crepaldi (2014) decidiram pelo uso de 8 atrasos, sem maiores explicações em relação ao embasamento para tal decisão. Neste caso, não é possível determinar se os resultados seriam impactados de forma significativa ao se empregar mais ou menos atrasos. Em Gomide e Milidiú (2010) foram testados apenas 5 e 10 atrasos, e os resultados apresentados unicamente para a melhor configuração. A evolução das previsões pelo uso de diferentes tamanhos de memória não é considerado, portanto o estudo torna-se muito particular para o cenário proposto, e as conclusões advindas das experimentações podem não servir facilmente como base para pesquisas mais abrangentes. Outros trabalhos utilizam métodos variados para a escolha do melhor número de atrasos, como em Neto, Calvalcanti e Ren (2009), onde a análise de correlação é empregada. Esta é uma abordagem válida, mas não provê uma visão geral do comportamento da rede quando são explorados atrasos em torno e além daquele selecionado.

Observou-se também limitações na exploração do número de neurônios na camada oculta, por parte dos trabalhos correlatos. Gomide e Milidiú (2010) adotam uma configuração fixa, baseada em uma média geométrica entre os tamanhos das camadas de entrada e de saída. Silva et al. (2014) também fixaram o número de neurônios, de acordo com a quantidade de atributos e classes com as quais a rede neural lida. Já Costa (2012) e Costa e Crepaldi (2014) citaram a realização de testes de performance para determinação do número de neurônios nas camadas ocultas, porém omitiram a evolução dos resultados que levaram à escolha final. Além disso, nenhum dos estudos possibilita que se tire conclusões a respeito do relacionamento entre o

número de neurônios na camada oculta e o tamanho da memória de linha de atraso escolhida.

Os trabalhos de Gomide e Milidiú (2010) e Silva et al. (2014) fizeram uso de indicadores da Análise Técnica como suporte às previsões por redes neurais, e apresentaram resultados aparentemente promissores. No entanto, não há evidências da contribuição de cada indicador para a melhora do desempenho da rede. De fato, segundo Setiono e Liu (1997), alguns atributos podem até mesmo prejudicar a acurácia de predição. Gomide e Milidiú (2010) apenas concluíram que o uso conjunto da Média Móvel Exponencial e das Bandas de Bollinger levaram aos menores erros de previsão, mas seus impactos relativos foram suprimidos do artigo. E Silva et al. (2014) fizeram uso de 10 diferentes tipos de indicadores, porém não foi possível compreender a importância de cada um deles separadamente. Portanto, para ambos os trabalhos, torna-se difícil medir o peso da Análise Técnica sobre os resultados gerais, e se os mesmos mostrariam piora significativa caso algum dos indicadores não fosse utilizado. O presente estudo procura diferenciar-se dos aqui relatados, realizando uma abordagem comparativa, com a inclusão gradativa de indicadores ao modelo, buscando evidenciar de forma mais precisa a influência do uso destas técnicas.

Apesar de todos os trabalhos correlatos apresentados estudarem a previsão de preços de ações, em nenhum deles houve uma preocupação em relação à dinamicidade e constantes mudanças nas condições de mercado. Devido a estas características, caso a rede neural não seja submetida a re-treinamentos periódicos, seu desempenho tende a se degradar com o tempo (KAASTRA; BOYD, 1996). As arquiteturas e parametrizações escolhidas nos estudos, portanto, podem não ser de grande valia para futuros trabalhos, mas sim as respectivas metodologias adotadas. Esta dissertação, no entanto, procurou dar mais ênfase a tal questão, tentando clarificar a importância desta abordagem sobre pesquisas relacionadas a séries temporais financeiras.

4 ABORDAGEM EXPERIMENTAL

Neste trabalho é proposta a avaliação da capacidade das redes neurais artificiais do tipo FTLFN de prever séries temporais financeiras, mais especificamente preços de ações. Procura-se diferenciar de outros estudos ao investigar sua aplicabilidade especificamente no período intradiário, o qual possui características próprias, incomuns a períodos diários ou mensais, adotados pela maioria das pesquisas nesta área, conforme literatura já referenciada anteriormente. Além disso, são examinadas alterações no desempenho da rede quando diferentes tamanhos de memórias de linhas de atraso são aplicados aos sinais de entrada, assim como quando se varia o número de neurônios na camada oculta da rede.

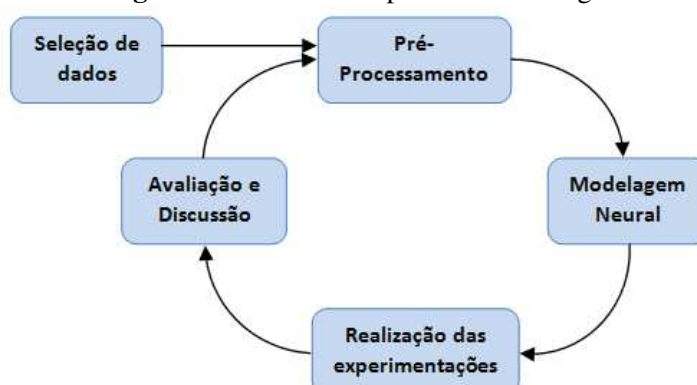
Sendo assim, este capítulo descreve a metodologia empregada na elaboração da pesquisa. São detalhadas as diversas tarefas envolvidas na construção de modelos para o problema proposto, que requereu uma estruturação criteriosa, abrangendo desde a coleta dos dados até a aplicação de métricas para avaliação dos resultados.

Após a metodologia, são detalhadas de forma técnica as principais decisões em termos de seleção dos dados, pré-processamento sobre os mesmos e configurações da rede neural utilizada. Já a seção seguinte a estas relata os experimentos realizados e os respectivos resultados obtidos. As experimentações foram planejadas de maneira que as descobertas ampliem o conhecimento a respeito das FTLFNs quando expostas a diferentes sinais de entrada, configurações da rede e indicadores da Análise Técnica, podendo servir como base para futuras pesquisas, ao mesmo tempo em que expande a literatura em torno do assunto. Por fim, é apresentada uma avaliação e discussão geral dos resultados, buscando relacioná-los com situações práticas de mercado.

4.1 Metodologia

A Figura 8 ilustra o fluxo de etapas da metodologia aplicada neste trabalho. Foram consideradas cinco tarefas principais, com escopos bem definidos sobre cada uma delas.

Figura 8: Fluxo das etapas da metodologia



Fonte: Elaborado pelo autor desta dissertação.

A primeira etapa refere-se à seleção dos dados a serem utilizados ao longo do trabalho. Para a escolha dos mesmos, foram levados em consideração aspectos relativos ao cenário sendo abordado, de maneira que os tipos e formatos dos dados refletissem aqueles adotados na realidade dos mercados. Ainda relacionado a este quesito, também pensou-se nas características de negociação das ações, como liquidez e volume, as quais deveriam viabilizar uma futura aplicação do método aqui apresentado em situações reais. Além disso, decidiu-se pelo estudo de ações de empresas que sejam foco comum de outros trabalhos, permitindo, assim, a complementação entre eles ou até mesmo comparações de resultados das diferentes abordagens.

Decididas as ações com as quais se trabalhar, de acordo com os critérios acima mencionados, a próxima tarefa referiu-se à coleta dos dados. Esta fase foi de extrema importância para o estudo aqui exposto, uma vez que quebras no relacionamento temporal das séries empregadas seriam prejudiciais para quaisquer conclusões a respeito do uso de históricos de preços pelas FTLFNs. Por este propósito, decidiu-se pela coleta de dados a partir de ferramentas e serviços de um *vendor*¹ Bovespa, ou seja, uma companhia autorizada pela BM&F Bovespa a comercializar e distribuir informações e notícias geradas pelos mercados, proporcionando maior confiabilidade ao processo como um todo.

Com os históricos de preços, volumes e quantidades em mãos, foram aplicadas atividades de pré-processamento, com o objetivo de consolidar a base de dados a ser utilizada em todas as experimentações. Foram excluídos valores e preços fora do horário normal de negociação, assim como atributos redundantes advindos da ferramenta de captura de dados. Também foi necessário o ajuste de valores negativos referentes à quantidade de ações negociadas no início do dia, resultante de um cálculo genérico para o restante dos valores em diferentes horários.

Outra tarefa importante dentro da etapa de pré-processamento foi a criação dos novos atributos representativos dos indicadores da Análise Técnica utilizados como entradas extras para a rede neural. A partir dos históricos de preços das ações capturados na fase anterior, geraram-se os atributos correspondentes às Médias Móveis, ao Índice de Força Relativa e às Bandas de Bollinger. Com a disponibilidade de todos os dados a serem usados nos experimentos, prosseguiu-se à sua segmentação e organização em diferentes conjuntos de treinamento e teste, necessários para o aprendizado e generalização do conhecimento adquirido por parte das redes neurais. Além da divisão em conjuntos, também aplicou-se a divisão dos dados sequenciais em períodos de tempo limitados. Esta abordagem é chamada de *janelas deslizantes*, e se mostra interessante na medida em que se propõe a evitar a degradação da rede conforme ela é utilizada ao longo do tempo (KAASTRA; BOYD, 1996). Na literatura relacionada à previsão de séries de preços, é comum o uso desta técnica, como visto em Freitas, Souza e Almeida (2009) e Liang et al. (2013).

A fase de modelagem neural da metodologia corresponde à escolha da implementação da rede neural do tipo FTLFN a ser empregada nas experimentações, assim como à decisão de seus parâmetros. Para este trabalho, optou-se pela FTLFN disponível na *Neural Network Toolbox* do

¹<http://www.bmfbovespa.com.br/pt-br/servicos/servicos-de-informacao/sinal-de-informacoes/vendors.aspx>

ambiente de programação funcional MATLAB². Apesar da documentação oficial do software (BEALE; HAGAN; DEMUTH, 2013) apresentar a rede sob o nome de *Focused Time-Delay Neural Networks* (FTDNN), ela apresenta as mesmas características das FTLFNs presentes na literatura correspondente. A ferramenta oferece uma estrutura bastante completa para a criação e simulação de redes neurais, permitindo um alto grau de liberdade em suas configurações, facilitando testes ao mesmo tempo em que garante robustez ao processo. As parametrizações das redes ocorreram de acordo com os objetivos definidos para o estudo. Tentando-se identificar o potencial de previsão de preços das FTLFNs para o período intradiário, estipularam-se diferentes números de neurônios a serem testados na camada oculta, o número de atrasos considerados para os sinais de entrada, assim como as variáveis a serem utilizadas para suporte a previsões mais precisas.

Na etapa de realização de experimentações foram executados os *scripts* desenvolvidos no MATLAB, com o uso da FTLFN e suas respectivas parametrizações, conforme definido previamente. As capacidades da rede neural foram exploradas baseando-se nos atributos coletados e criados para as ações selecionadas. À medida em que os experimentos eram finalizados, seus resultados eram agregados de forma a permitir a construção de tabelas e gráficos que facilitassem a avaliação. Devido à grande variedade de testes planejados, esta etapa foi a que exigiu maior demanda em termos de tempo e poder computacional, ao longo de todo o trabalho.

Após a finalização de cada experimento, foi possível avaliar e interpretar os padrões de comportamento da FTLFN quando sujeita à previsão de preços de ações no período intradiário, de acordo com as diferentes configurações de parâmetros e sinais de entrada adotados. Os resultados obtidos, por vezes, motivaram a busca de refinamentos, exigindo nova execução do processo, a partir da fase de pré-processamento, configurando ciclos na metodologia, conforme ilustrado na Figura 8. Após a identificação de um modelo satisfatório na etapa de avaliação foi importante rever os diferentes passos a partir do pré-processamento, certificando-se de que cada um deles atendeu devidamente ao objetivo exigido. Portanto, também neste caso existiu o ciclo disposto na metodologia.

Alguns estudos relacionados a previsões de preços em bolsas de valores utilizam retornos percentuais, ao longo do tempo, para medir o desempenho de redes neurais, como é o caso em Chavarnakul e Enke (2009). Apesar de se apresentar como uma boa métrica para comparação entre diferentes pesquisas, ela é altamente dependente da modelagem conjunta de um *trading system*. Nestes sistemas é necessário definir pontos de compra e venda de ações, de acordo com alguma estratégia, além de considerar custos envolvidos em transações, entre outras características de negociações no mercado. O problema com esta abordagem é que, em certas instâncias, o desempenho da RNA é menos relevante do que a estratégia e as escolhas envolvendo o *trading system*.

Como o objetivo deste estudo é explorar as capacidades de aproximação de valores, por parte da FTLFN, optou-se pelo emprego do MAPE (*Mean Absolute Percentage Error*) para

²Versão 8.4.0.150421 (R2014b), 64bits

apresentação dos resultados. Esta métrica permite determinar a robustez dos modelos de previsão (TICKNOR, 2013), possui sensibilidade a pequenas mudanças nos dados, baixa proteção a valores discrepantes e não demonstra assimetria dos dados (FREITAS; SOUZA; ALMEIDA, 2009). Além disso, é amplamente utilizado em estudos deste tipo, como é o caso em Chang e Liu (2008) e Hassan, Nath e Kirley (2007). Ao longo desta dissertação, frequentemente utiliza-se o termo acurácia ao se analisar os resultados obtidos. É importante salientar que está se referindo, nestes casos, aos valores MAPE correspondentes, e não à métrica acurácia comumente utilizada para determinação do desempenho de algoritmos de classificação, por exemplo.

Ao final do trabalho são discutidos os resultados alcançados, buscando-se definir a contribuição do estudo para a ampliação do conhecimento em relação ao tópico de previsões no período intradiário, através de redes neurais do tipo FTLFN. São sugeridas extensões e modificações para pesquisas futuras, permitindo a continuação do trabalho e possíveis melhorias.

4.2 Seleção de dados

Sendo o objetivo da rede neural, neste trabalho, realizar previsões em intervalos de 15 minutos, é importante que os ativos estudados apresentem características de alta liquidez, permitindo que sejam negociados em diferentes níveis de preços e de forma ágil a qualquer momento do dia. Para tanto, optou-se por adotar as ações de Petrobras (PETR4) e Vale (VALE5) para os experimentos. Ambas fazem parte do Índice Bovespa, indicador que inclui os ativos de maior negociabilidade e representatividade do mercado de ações brasileiro. Além disso, também compõe o Índice Brasil 50 (IBrX), que mede o retorno das 50 ações mais negociadas na BM&F Bovespa em termos de liquidez. Outro motivo para a escolha destes dois ativos em particular é que suas ações valorizaram-se de maneiras bastante distintas nos últimos anos, permitindo que o desempenho da rede neural seja apresentado sobre dois cenários com diferentes realidades. Diversos estudos fazem uso conjunto destas mesmas ações, como Alvim e Milidiú (2013), Martinez et al. (2009) e Nizer e Nievola (2012), tornando-as candidatas ideais para complementações entre diferentes pesquisas. Em relação aos atributos gerados a partir da Análise Técnica, foram selecionadas as Médias Móveis, as Bandas de Bollinger e o Índice de Força Relativa. O emprego destes indicadores, tanto de forma isolada quanto em conjunto, é bastante comum na literatura relacionada, como visto em Chang et al. (2011) e Teixeira e Oliveira (2010).

Para a seleção dos dados referentes às cotações intradiárias de PETR4 e VALE5 utilizou-se a ferramenta *TradeZone Data Center* (TZDC), desenvolvida pelo *vendor IT Evolution*³. Através de requisições HTTP (*Hypertext Transfer Protocol*) a um servidor local, instalado pela ferramenta, é possível obter históricos de cotações de qualquer ação negociada na Bovespa, na periodicidade desejada. Os dados disponibilizados são ajustados de forma automática a cada distribuição de provento, desdobramento e agrupamento de ações, ou qualquer outro fator que afete os preços. Isso facilita seu tratamento, no sentido em que não há a necessidade de ajustes

³<http://www.itevolution.com.br/>

Tabela 4: Formato de retorno dos dados pelo *TradeZone Data Center*

Data	Hora	Abert.	Máx.	Mín.	Fech.	Qtd.	Vol.
02/01/2014	10:15	16,11	16,16	16,05	16,05	635	6.791.410
02/01/2014	10:30	16,05	16,07	15,99	16,01	806	9.589.700
02/01/2014	10:45	16,01	16,07	16,00	16,04	542	4.288.391
02/01/2014	11:00	16,04	16,06	15,94	15,95	884	9.423.199

Fonte: Elaborado pelo autor desta dissertação.

manuais dessa natureza, ao mesmo tempo em que os mantém constantemente atualizados.

Por meio de um programa de desenvolvimento próprio do autor, na linguagem *C#*, capaz de se comunicar com o servidor TZDC, foram coletadas cotações de PETR4 e VALE5, discretizadas a cada 15 minutos, para o período compreendido entre 1º de janeiro de 2014 e 17 de outubro de 2014. O montante dos dados correspondeu a cerca de 6.000 registros para cada ativo. A Tabela 4 apresenta um exemplo do formato dos registros retornados pela ferramenta *TradeZone*, conforme requerido pelo programa elaborado. Os campos *Abert.*, *Máx.*, *Mín.*, *Fech.*, *Qtd.* e *Vol.* representam, respectivamente, os preços de abertura, preço máximo, preço mínimo, preço de fechamento, quantidade e volume negociados para o intervalo correspondente do registro.

O conjunto de dados, após extraído, foi armazenado temporariamente em um arquivo local, sendo posteriormente transferido para o ambiente MATLAB através de sua ferramenta de importação. A partir deste momento, toda tarefa relativa ao trabalho proposto foi executada neste *software*. Sua escolha deu-se pela oferta de diferentes pacotes de funcionalidades necessárias ao desenvolvimento desta pesquisa, assim como por sua robustez e aceitação no meio acadêmico e científico.

4.3 Pré-processamento

O histórico de preços, volumes e número de negócios adquiridos do *vendor* Bovespa inclui apenas as informações básicas a serem utilizadas pelas redes neurais. Além disso, dados irrelevantes e capazes de alterar negativamente os resultados também estão presentes. Desta forma, foram necessárias etapas de pré-processamento, a fim de preparar os dados de acordo com os requisitos exigidos pelas experimentações planejadas.

4.3.1 Limpeza da base

A base de dados importada para o MATLAB inclui preços referentes ao período de *after-market*, das 17:30 às 18:00, ou seja, após o horário normal de negociação no mercado à vista, que ocorre das 10:00 às 17:00 (não há qualquer tipo de negociação entre as 17:00 e as 17:30). Durante o *after-market*, o número de negócios e a liquidez dos ativos tornam-se bastante reduzidos, e a variação percentual máxima do preço das ações é limitada a 2% em relação a seus preços de fechamento (BM&F BOVESPA, 2009). Assim, como forma de não se trabalhar com

informações de menor relevância, focando apenas sobre movimentações com maior liquidez, decidiu-se neste estudo eliminar da base de dados observações referentes ao período de *after-market*. Ao final da tarefa, a base resultante apresentou um total 5.515 registros para cada ação, e passou a ser utilizada para todo o restante das atividades do estudo.

4.3.2 Sinais auxiliares da Análise Técnica

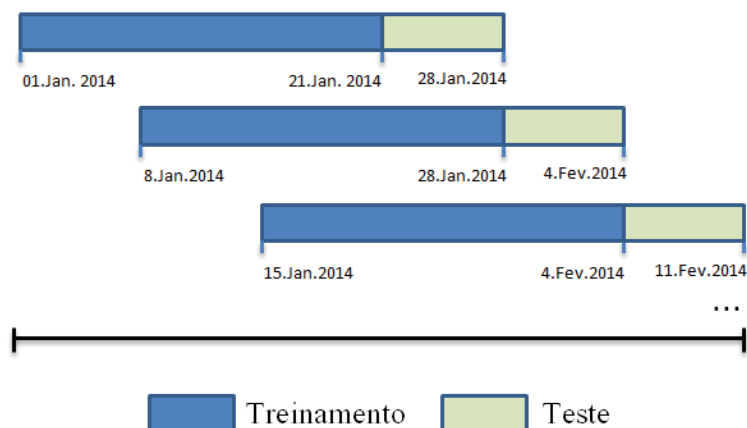
O próximo passo dentro da etapa de pré-processamento foi a criação dos novos atributos relativos aos indicadores da Análise Técnica. Os valores de fechamento de cada período foram utilizados como base para esta tarefa, sendo submetidos a funções do pacote de ferramentas *Financial Toolbox*. Com a função *tsmovavg* criaram-se as Médias Móveis; *rsindex* deu origem à série do Índice de Força Relativa; e a função *bollinger* gerou os dados para as Bandas de Bollinger.

A parametrização dos indicadores técnicos está relacionada ao período que se deseja analisar, assim como à velocidade de reação desejada, conforme os preços dos ativos sofrem alterações (MURPHY, 1999). Ainda assim, não existem valores pré-determinados que produzam melhores ou piores resultados, independente da empresa que se está negociando e do momento. A escolha dos parâmetros, portanto, recai sobre a preferência de cada investidor, os quais muitas vezes os decidem baseados em testes retroativos de desempenho.

Para esta pesquisa, optou-se por empregar parâmetros de acordo com o utilizado comumente em situações de análises reais de mercado, como definido em (STEVENS, 2002). Segundo o autor, para análises de curto prazo, tanto a Média Móvel quanto o Índice de Força Relativa fornecem bons resultados quando parametrizados para 9 períodos. Já a média móvel das Bandas de Bollinger é padronizada para 20 períodos, com suas bandas superior e inferior sendo calculadas sobre 2 desvios padrões, independente do prazo sendo analisado. Foram estas, portanto, as configurações utilizadas para o cálculo dos indicadores adotados neste trabalho. Além disso, os indicadores de Análise Técnica foram agregados de forma gradativa aos demais sinais de entrada. Por exemplo, primeiramente apenas a série de preços de fechamento foi utilizada como sinal de entrada da rede neural. Em um segundo experimento, além desta também utilizou-se a série referente à Média Móvel como sinal de entrada. Com isso, procurou-se avaliar a contribuição de cada atributo, de forma isolada e em conjunto, para a acurácia de previsão. A seção de experimentações planejadas detalha as diferentes combinações e parametrizações adotadas em cada uma delas.

4.3.3 Particionamento dos dados

Na tentativa de simular a previsão de preços durante um espaço maior de tempo, assim como evitar a degradação na acurácia da rede neural devido a mudanças nas condições do mercado acionário, foi utilizado nesta pesquisa a técnica de *janelas deslizantes*, ainda na fase de

Figura 9: Divisão dos dados por janela deslizante

Fonte: Elaborado pelo autor desta dissertação.

pré-processamento. Este método consiste em particionar os dados em séries de conjuntos sobrepostos ao longo do tempo, no formato treino-validação-teste (KAASTRA; BOYD, 1996). No trabalho proposto, no entanto, devido à introdução da Regularização Bayesiana ao algoritmo de treinamento, não foi necessário o uso de um conjunto de validação, sendo os dados divididos unicamente em treino e teste. Tomou-se para cada agrupamento do tipo treino-teste o montante de dados compreendendo 4 semanas de cotações intradiárias, sendo 3 delas reservadas para treino e a última para teste, representando respectivamente 75% e 25% do total de dados, conforme o estudo de Geva e Zahavi (2014). O período de 4 semanas equivale, em média, a 560 observações, sendo cada uma delas uma discretização de 15 minutos no tempo. Os conjuntos foram formados por blocos de dados contínuos, tentando-se evitar a quebra do entrelaçamento temporal entre os preços.

Uma vez que a FTLFN tenha sido treinada com dados das 3 semanas e testada sobre a semana seguinte deste primeiro agrupamento, o resultado do teste é armazenado, e a rede é submetida a um novo agrupamento treino-teste. Nesta composição, a primeira semana do conjunto de treinamento anterior é descartada, e o conjunto de teste anterior passa a fazer parte dos dados de treino do novo conjunto. Já a etapa de testes deste novo ciclo ocorre sobre a semana seguinte. Esta segmentação é ilustrada na Figura 9, e foi seguida até que o conjunto de teste coincidissem com a última fração dos dados disponíveis. Sua implementação foi realizada previamente ao início das experimentações, diretamente no MATLAB, deixando os conjuntos preparados para uso na fase seguinte.

Dividindo-se os dados desta maneira, entre 1º de janeiro de 2014 e 17 de outubro de 2014, foram gerados 35 agrupamentos treino-teste diferentes, onde cada conjunto de treino englobou uma média de 395 registros, e cada conjunto de teste agregou cerca de 132 registros.

A fim de confirmar que a afirmação em Kaastra e Boyd (1996) a respeito da degradação da rede também é válida para o cenário abordado neste trabalho, foi realizado um experimento que demonstra graficamente a evolução do erro de previsão ao longo de uma semana, para

cada intervalo de 15 minutos, sem que houvesse novos treinamentos da rede durante o período. Ainda, foram simuladas situações de aplicação das FTLFNs para previsões durante duas e três semanas à frente, sem re-treinos. Os resultados são apresentados na seção correspondente.

4.4 Modelagem Neural

A implementação da FTLFN oferecida pelo ambiente MATLAB, com nome técnico de *timedelaynet*, provê grande flexibilidade para parametrizações, além de funções automatizadas para a correta manipulação dos sinais de entrada de acordo com o tamanho da memória de linha de atraso especificado, apresentando-se como ideal para as experimentações do trabalho proposto. É importante que os dados fornecidos à rede sejam corretamente dispostos, a fim de representar fielmente a dinâmica temporal presente nas situações reais de mercado. Para tanto, a *toolbox* de redes neurais do MATLAB fornece a função *preparets*, capaz de simplificar a tarefa de reformatação das séries temporais de entrada e de saída esperadas. Ela realiza de forma automática o deslocamento do número de passos necessários dentro de cada série, criando estruturas preenchidas adequadamente com os estados dos atrasos, conforme o parâmetro de tamanho da memória de linha de atraso configurado para a rede (BEALE; HAGAN; DEMUTH, 2013). Antes de qualquer ciclo de treinamento da rede neural utilizada nas experimentações deste trabalho, a função *preparets* foi executada sobre os dados, preparando devidamente as estruturas para processamento.

Uma importante tarefa de pré-processamento para dados utilizados por redes neurais é a sua normalização, impedindo que valores grandes de entrada não sobreponham valores menores durante o treinamento, o que diminuiria sua importância para previsões e aumentaria o erro da rede (TICKNOR, 2013). Para esta atividade, utilizou-se a função *mapminmax* do MATLAB, executada automaticamente sobre todas as entradas da rede neural, quando configurada desta maneira. Ela transforma os dados de entrada da rede, garantindo que todos os valores se encontrem dentro do intervalo $[-1,1]$ (BEALE; HAGAN; DEMUTH, 2013). O processo inverso, ou seja, a desnormalização, é aplicada sobre os dados de saída da RNA, de forma que reflitam os valores em seus dimensionamentos originais.

Para os diferentes testes, a FTLFN foi configurada com três camadas, sendo uma de entrada, uma oculta e uma de saída. Esta escolha fundamentou-se no fato de que duas camadas ocultas raramente melhoram o modelo, podendo inclusive introduzir um risco maior de convergência a mínimos locais (CYBENKO, 1989), e também pelo uso comum desta mesma abordagem em trabalhos correlatos, como Silva et al. (2014) para o período intradiário, e Oliveira, Nobre e Zárte (2013) com foco na previsão diária. O número de neurônios na camada de entrada é dependente dos sinais de entrada e do número de atrasos utilizados sobre os mesmos, os quais variaram em diferentes experimentações. A camada de saída foi configurada com apenas um neurônio, referente à cotação a ser prevista. Já para a camada oculta foram testadas diferentes quantidades de neurônios, uma vez que sua definição depende altamente do problema que se

está tratando (AZOFF, 1994). Optou-se, então, pela verificação do comportamento da rede com 10, 20, 30, 50 e 80 neurônios nesta camada, buscando padrões de melhorias ou pioras nos resultados de acordo com cada configuração. O uso além dos limites de 10 e 80 não foi experimentado, porém os resultados apresentados neste trabalho proveem uma visão geral da evolução da rede sob as quantidades de neurônios testadas, e a necessidade de talvez se expandir o experimento para diferentes abordagens.

As funções de ativação dos neurônios das camadas ocultas foram definidas como do tipo sigmóide tangente hiperbólica. Já para o neurônio de saída foi configurada a função linear. Tais escolhas baseiam-se na necessidade de tratar a não-linearidade das séries temporais em questão, assim como fornecer valores previstos por meio de regressão. Além do mais, a literatura relacionada apresenta em grande parte as mesmas configurações, como visto em Ticknor (2013) e Sitte e Sitte (2000). Para o aprendizado da rede, empregou-se o algoritmo de otimização gradiente descendente com Regularização Bayesiana, implementado no MATLAB através da função de treinamento *trainbr*.

Um dos principais aspectos da FTLFN explorados neste estudo foi o impacto do tamanho da memória de linha de atraso sobre a acurácia de previsão. Para tanto, experimentações foram realizadas com 3, 5, 7, 9, 11, 13, 15, 20 e 25 atrasos sobre todos os sinais de entrada apresentados à rede. Como exemplo, se em determinado experimento são fornecidos como entrada para a rede os sinais de preço de fechamento, de máxima e de mínima, e é determinada uma memória de 3 atrasos, significa que está se provendo à rede 3 atrasos sobre cada um dos 3 sinais, totalizando 9 variáveis de entrada para processamento pela FLTFN. O objetivo desta abordagem foi de encontrar indicativos da importância do uso de padrões passados nos preços intradiários para prover melhores resultados de previsões. O experimento não busca determinar o desempenho específico de cada número de atrasos, mas sim entender o comportamento geral da rede pela adoção de diferentes valores para este parâmetro. Desta forma, escolheram-se valores espaçados entre si, compreendendo um intervalo significativo a partir do qual fosse possível concluir a respeito do funcionamento da FTLFN sob este aspecto.

A tabela 5 resume as principais definições envolvendo a FTLFN para realização das experimentações.

4.5 Experimentações planejadas

Como visto em seções anteriores, este trabalho propõe explorar as capacidades das FLTFNs de prever preços de ações no período intradiário. Para tanto, uma ampla e extensa gama de experimentações foram planejadas, permitindo que conclusões concretas fossem tiradas a partir dos resultados.

No mercado financeiro é possível fazer uso de variadas informações sobre os preços e volumes negociados dos ativos, na tentativa de prever comportamentos futuros. Também para o presente trabalho adotou-se a mesma ideia, razão pela qual foram capturados históricos de diver-

Tabela 5: Definições sobre a FTLFN

Parâmetro	Valor
Funções de pré-processamento dos dados	<i>preparets e mapminmax</i>
Número de camadas	3 (1 de entrada, 1 oculta, 1 de saída)
Número de neurônios na camada oculta	10, 20, 30, 50, 80
Número de neurônios na camada de saída	1
Função de ativação na camada oculta	Sigmóide tangente hiperbólica
Função de ativação na camada de saída	Linear
Algoritmo de treinamento	Regularização Bayesiana
Tamanho da memória de linha de atraso	3, 5, 7, 9, 11, 13, 15, 20, 25

Fonte: Elaborado pelo autor desta dissertação.

os atributos para Petrobras e Vale. Não há, porém, um consenso sobre o tipo de informação que produz melhores ou piores resultados. Por este motivo, neste estudo decidiu-se pela aplicação de diferentes conjuntos de sinais como entradas para as múltiplas configurações de redes neurais experimentadas. Identificou-se duas vantagens principais para esta abordagem. Primeiro, o resultado final da experimentação com uma determinada configuração pode ser apresentado como uma média entre os valores obtidos individualmente, para cada um dos conjuntos. Este método traz robustez às conclusões apresentadas, uma vez que diminui a dependência das mesmas a um conjunto específico de dados fornecido à rede, e provê uma noção do comportamento geral desta rede em relação a suas parametrizações. A segunda vantagem é que permite identificar, entre todos os conjuntos testados, aqueles que geram os melhores resultados ao longo do tempo, tornando-se os principais candidatos para refinamentos em trabalhos futuros.

Como elaborado anteriormente, através da importação dos dados de Petrobras e Vale, teve-se acesso aos seus históricos de preços de abertura, fechamento, máximo, mínimo, quantidade e volume negociados em intervalos de 15 minutos. A partir destes, construíram-se 11 conjuntos de sinais de entrada, sobre os quais todas as experimentações planejadas neste estudo foram executadas. A Tabela 6 aponta estas 11 combinações.

Tabela 6: Conjuntos de sinais de entrada utilizados nas experimentações

Nº do Conjunto	Composição do conjunto de sinais de entrada
1	Fechamento
2	Fechamento + Máxima + Mínima
3	Fechamento + Abertura
4	Fechamento + Máxima + Mínima + Abertura
5	Fechamento + Quantidade
6	Fechamento + Abertura + Quantidade
7	Fechamento + Máxima + Mínima + Abertura + Quantidade
8	Fechamento + Volume
9	Fechamento + Máxima + Mínima + Abertura + Volume
10	Fechamento + Quantidade + Volume
11	Fechamento + Máxima + Mínima + Abertura + Quantidade + Volume

Os experimentos desta pesquisa foram divididos de forma a se atingir três objetivos principais. Inicialmente, a intenção foi identificar o comportamento da rede FTLFN quando exposta a atrasos de diferentes magnitudes sobre os sinais de entrada. O segundo objetivo foi o de verificar o impacto sobre a acurácia de previsão da rede ao se variar o número de neurônios em sua camada oculta. Finalmente, foram realizadas análises para se compreender a influência dos sinais da Análise Técnica como forma de suporte a previsões de preços mais precisas. Experimentos extras foram executados a fim de se detectar a degradação na capacidade de previsão intradiária da FTLFN ao longo do tempo, identificando-se possíveis necessidades de adaptações nos tamanhos dos conjuntos de treino e teste da rede neural.

Dado que redes neurais com arquiteturas idênticas podem apresentar diferentes resultados de previsão, uma vez que seus pesos são inicializados de forma aleatória em cada execução (HAYKIN, 2001), neste trabalho foram realizadas 5 execuções de cada experimento, extraíndo-se a média aritmética entre eles como resultado final a ser apresentado. A estratégia procura suavizar o viés ao qual as redes neurais estariam sujeitas, caso apenas uma execução fosse realizada sob um cenário onde os pesos iniciais da rede não contribuíssem para uma previsão de preço satisfatória, de acordo com o esperado para a média.

Como definido na seção referente à metodologia, a métrica empregada neste trabalho para exposição dos resultados foi o MAPE. Seu cálculo é elaborado da seguinte forma:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{r_t - p_t}{r_t} \right| \quad (4.1)$$

onde n é o tamanho do período definido para cálculo, em número de passos da série temporal, e r_t e p_t são os retornos esperado e previsto no tempo t , respectivamente. Considerando-se o particionamento dos dados através de janelas deslizantes adotados neste trabalho, o MAPE foi calculado em momentos distintos ao longo do processo de treinamento e teste das redes FTLFN. Dentro de cada agrupamento, o MAPE é calculado sobre a semana utilizada para teste, após a rede ter sido treinada com os dados relativos às 3 semanas anteriores. Depois da obtenção do resultado individual para cada um dos 35 agrupamentos, é calculada a média aritmética entre eles, fornecendo o valor representante do erro percentual absoluto médio para todo o período compreendido. A mesma abordagem é adotada usualmente na literatura, como em Chen, Leung e Daouk (2003), Liang et al. (2013) e Freitas, Souza e Almeida (2009).

4.6 Resultados

Diferentes experimentações foram realizadas a fim de se explorar de forma bastante abrangente o potencial de previsão das redes FTLFN, sob o cenário de previsão de preços de ações no período intradiário, mais especificamente em intervalos de 15 minutos. No entanto, como destacado previamente, os resultados referem-se não somente a previsões dentro de um único dia, mas sim à aplicação de uma mesma rede ao longo do período de uma semana, após trei-

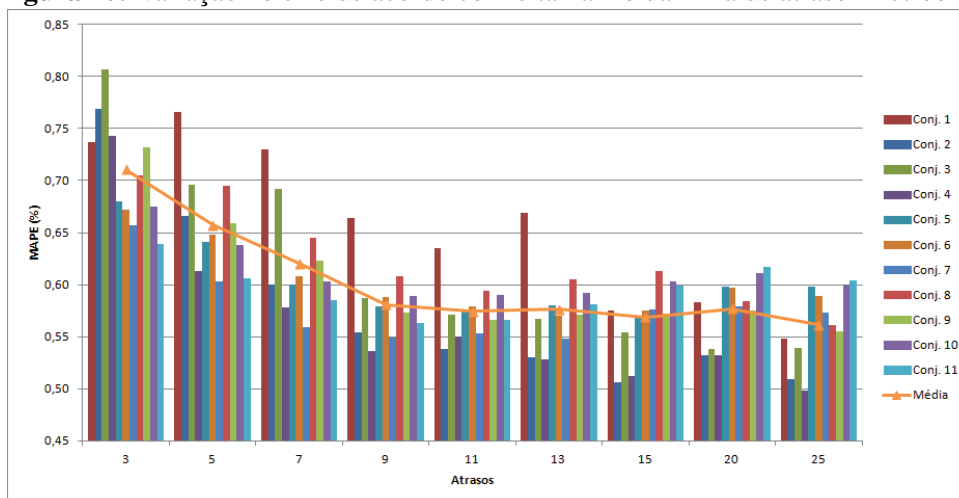
namento com dados relativos às três semanas anteriores, sendo o procedimento repetido para toda a série temporal disponível. A subseção 4.6.4 inclui uma exceção a esta abordagem, uma vez que experimenta-se também o uso da rede, sem re-treinos, em períodos maiores do que uma semana. Esta seção descreve cada um dos testes, destacando sua relevância e principais descobertas.

4.6.1 Resultados das experimentações sobre a memória de linha de atraso

Como mencionado anteriormente, experimentações foram realizadas com diversos valores para a memória de linha de atraso, na tentativa de identificar padrões de mudanças nas previsões das FTLFNs no período intradiário. Cada valor de atraso foi testado sobre os 11 conjuntos de sinais de entrada. Em cada experimentação, o atraso sendo testado foi aplicado sobre todos os sinais pertencentes ao conjunto corrente. Por exemplo, no teste com 3 atrasos sobre o conjunto 8, este número de atrasos foi considerado tanto para o sinal de Fechamento quanto para o de Volume, totalizando 6 valores de entrada fornecidos à rede neural. Além disso, sujeitou-se cada combinação de atraso e conjunto a 5 diferentes arquiteturas de FTLFN, com números distintos de neurônios na camada oculta: 10, 20, 30, 50 e 80. Os resultados dispostos nas Figuras 10 e 11 referem-se às médias dos valores MAPE calculados para Petrobras e Vale, respectivamente, para as 5 redes. Tal abordagem teve como objetivo prover conhecimento a respeito do comportamento geral da FTLFN à medida em que um número crescente de atrasos é adicionado aos sinais de entrada, ao mesmo tempo em que diminui a dependência dos resultados a um tamanho específico de camada oculta. O apêndice A contém as tabelas referentes aos resultados de Petrobras e Vale, a partir das quais construíram-se as figuras.

A Figura 10 evidencia, para o caso de previsões de preços intradiários da Petrobras, a capacidade das FTLFNs de se beneficiarem significativamente de sua principal característica: o uso de uma memória de linha de atraso na camada de entrada da rede, que a permite capturar padrões de comportamento ao longo do tempo. Considerando-se apenas 3 atrasos para cada sinal de entrada, a rede foi capaz de prever preços com um MAPE médio de 0,7105. No entanto, ao se explorar mais profundamente o uso de históricos maiores, considerando-se 25 passos no passado, foi possível melhorar a acurácia de previsão em 20,97%, levando o MAPE a 0,5615.

Os maiores ganhos, em termos percentuais, foram vistos entre 3 e 9 atrasos. Isto significa que o fornecimento de um histórico de preços e volumes até 2 horas e 15 minutos (9 períodos de 15 minutos) imediatamente anteriores ao momento de análise se apresentou como de grande utilidade para que a FTLFN conseguisse construir um conhecimento a respeito dos padrões de comportamento intradiários para a ação da Petrobras, podendo generalizar tal conhecimento e inferir preços futuros com base em novos padrões adquiridos. Mudanças menores no erro médio foram observadas entre 9 e 25 atrasos, com variações negativas e positivas à medida em que dados de prazos mais longos foram sendo adicionados ao modelo, indicando que preços e volumes anteriores a 2 horas e 15 minutos não se mostraram igualmente relevantes para a

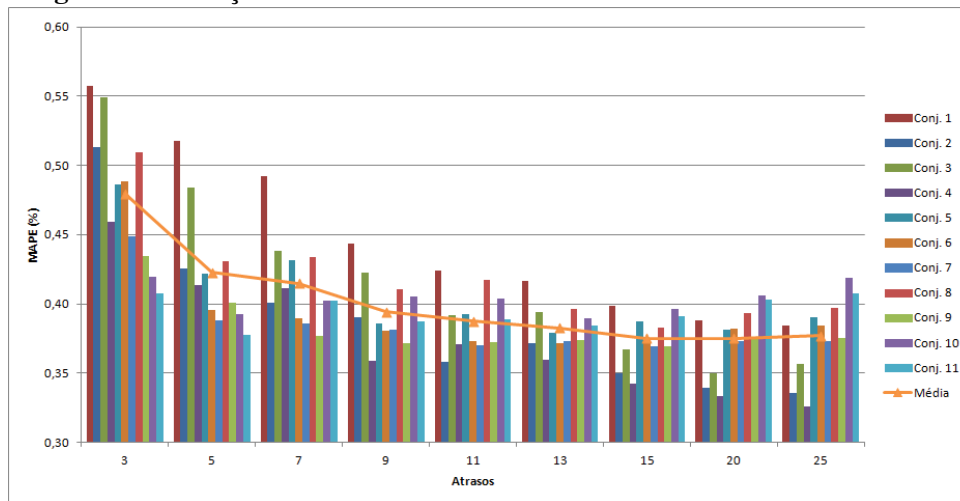
Figura 10: Variação no erro de acordo com o tamanho da linha de atraso - Petrobras

Fonte: Elaborado pelo autor desta dissertação.

melhoria de previsões. Ainda assim, o experimento com 25 atrasos, correspondentes a 6 horas e 15 minutos, foi o que mostrou o menor MAPE médio dentro do contexto proposto neste trabalho. Este fato é um indicativo da possibilidade da FTLFN não ter atingido seu limite de processamento, e ser capaz de identificar padrões de movimentos de preços e volumes mais duradouros.

Apesar de, até o momento, ter-se discutido resultados referentes a erros calculados como médias entre 11 diferentes conjuntos de sinais de entrada da rede neural, alguns conjuntos destacam-se por apresentarem desempenhos consistente e consideravelmente melhores que a média. Este é o caso dos conjuntos 2 e 4 para os experimentos com Petrobras. A partir do uso de 11 atrasos, ambos apresentaram os menores valores MAPE em relação às combinações restantes, até o teste com 25 passos passados. O conjunto 2 apresentou seu menor MAPE, de 0,5062, com 15 atrasos, representando um ganho de 10,99% sobre o valor médio de 0,5687 para este mesmo tamanho de memória. Já o conjunto 4 alcançou o erro mínimo quando foram considerados 25 atrasos, chegando a um MAPE de 0,4977, 11,35% menor que a média de 0,5615. Através desta análise é possível concluir que o fornecimento à FTLFN unicamente dos preços de fechamento, máximo, mínimo, e opcionalmente abertura, discretizados em intervalos de 15 minutos, permite que a mesma realize previsões no intervalo intradiário de forma mais precisa do que através do emprego do restante dos agrupamentos experimentados.

O comportamento geral da FTLFN foi muito similar quando experimentada com a Vale, como visto na Figura 11, com grandes melhorias na acurácia de previsão entre 3 e 9 atrasos. O MAPE para a primeira configuração foi de 0,4794, decaindo a 0,3941 para 9 atrasos, representando um ganho de 17,79%. O incremento no histórico fornecido à rede, entre 9 e 20 períodos, ainda permitiu a diminuição consecutiva do erro em todos os casos, apesar de em menor escala do que o observado até então. Com 20 atrasos, equivalentes a 5 horas de negociação, a FTLFN atingiu seu melhor resultado médio de previsão intradiária para a Vale, apresentando um MAPE

Figura 11: Variação no erro de acordo com o tamanho da linha de atraso - Vale

Fonte: Elaborado pelo autor desta dissertação.

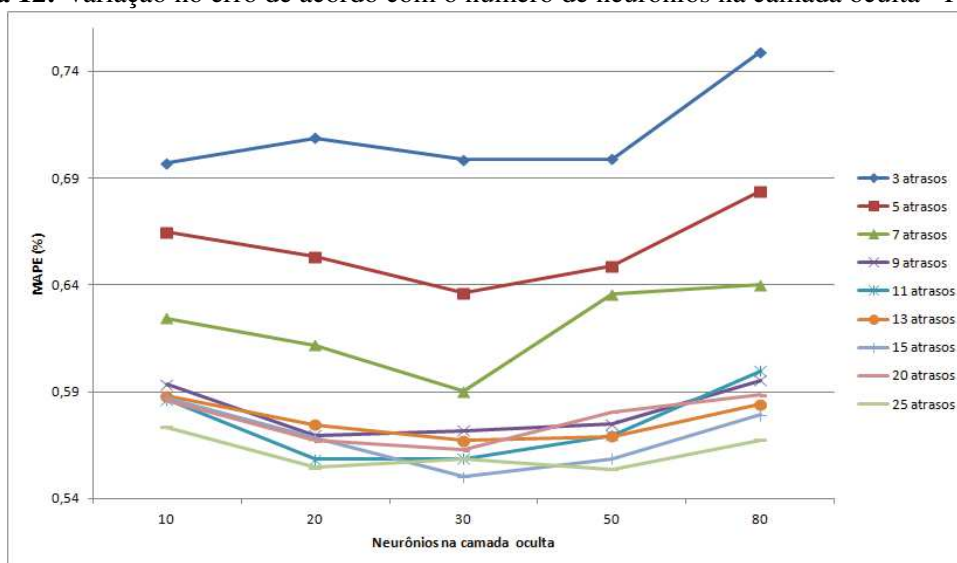
de 0,3749. Este valor representa um aumento de 21,80% na precisão da rede, em comparação com o teste onde 3 atrasos foram utilizados. Contrário ao que ocorreu com a Petrobras, porém, um histórico de 25 passos no passado não ajudou a reduzir ainda mais o erro. De fato, este tamanho de memória de linha de atraso elevou o erro médio em 0,59%, mostrando um MAPE de 0,3771. No entanto, isto não é um indicativo de que a FTLFN tenha alcançado seu desempenho máximo para a Vale. Experimentações futuras, que estendam o tamanho da memória de linha de atraso, podem encontrar resultados de previsão ainda melhores.

Assim como nos experimentos com a Petrobras, também no caso da Vale há destaque para os conjuntos 2 e 4, os quais apresentaram erros médios de previsão significativamente menores do que o restante dos agrupamentos, à medida em que históricos mais longos foram considerados sobre os sinais de entrada. A partir de 15 atrasos, ambos apresentaram consistentemente os menores valores MAPE, alcançando a aproximação máxima entre preços previstos e esperados quando fornecidos 25 atrasos a seus sinais de entrada. Para o conjunto 2, este valor correspondeu a um MAPE de 0,3359, um ganho de 10,93% sobre o erro médio para o emprego deste tamanho de memória de linha de atraso. O conjunto 4, por sua vez, apresentou um MAPE ainda menor, de 0,3260, 13,54% abaixo do erro médio.

4.6.2 Resultados das experimentações sobre o número de neurônios

Com o intuito de se determinar configurações de redes capazes de melhor processar os sinais de entrada e fornecer previsões de preços de forma mais precisa, foram realizados testes com 10, 20, 30, 50 e 80 neurônios na camada oculta da FTLFN. A subseção anterior exibiu os erros médios para os 11 conjuntos, calculados a partir dos valores obtidos individualmente sobre cada uma dessas diferentes configurações de rede. Já a presente subseção discretiza os resultados pelo número de neurônios utilizados na camada oculta da FTLFN, apresentando-os

Figura 12: Variação no erro de acordo com o número de neurônios na camada oculta - Petrobras



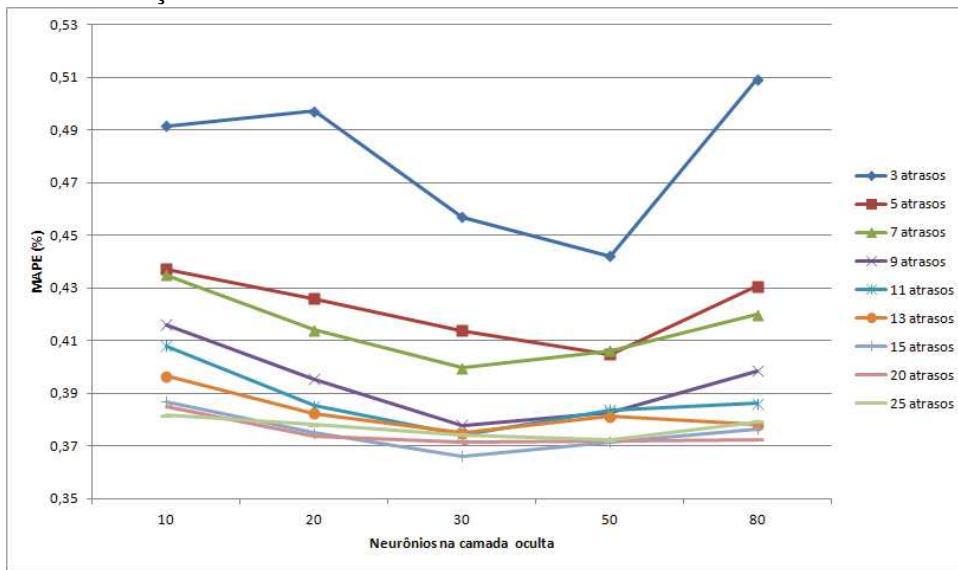
Fonte: Elaborado pelo autor desta dissertação.

através de gráficos onde cada ponto representa a média aritmética dos valores MAPE calculados individualmente para os 11 conjuntos de sinais de entrada, considerando-se o tamanho da camada oculta definida no eixo horizontal. Pretendeu-se, assim, identificar as parametrizações que conduziram efetivamente aos menores erros de previsão. O apêndice B detalha os resultados dispostos nos gráficos de Petrobras e Vale, através de tabelas.

O gráfico da Figura 12 provê uma forma de se determinar o número de neurônios que permitiu à rede fornecer os menores erros de previsão para a Petrobras. Com exceção de quando 3 atrasos são utilizados na memória de linha de atraso, todas as outras configurações têm seus erros médios diminuídos com o uso de 20 neurônios na camada oculta da FTLFN, quando comparados com 10 neurônios. O MAPE médio para esta rede com camada oculta maior é de 0,5962, representando um ganho de 2,44% sobre a de 10 neurônios, cujo MAPE é de 0,6111. Cinco configurações de atrasos (5, 7, 13, 15 e 20) mostram resultados ainda melhores com a adição de mais 10 neurônios à camada oculta, totalizando 30. Este tamanho também foi o que apresentou o menor MAPE médio dentre todos os experimentados, com um valor de 0,5882. Somente quando 25 atrasos foram empregados sobre os sinais de entrada, pode-se obter vantagens pelo uso de 50 neurônios. As redes sob as demais configurações de memória de linha de atraso tiveram aumento em seus erros médios sob esta circunstância. Já a empregabilidade de uma camada oculta com 80 neurônios foi responsável pelo maior MAPE médio do experimento, de 0,6208, sendo seu uso menos interessante até mesmo do que camadas ocultas com 10 neurônios.

Analisando-se os experimentos de forma individual, é possível identificar que a combinação de um tamanho de memória de 15 atrasos e camada oculta de 30 neurônios resultou no menor MAPE entre todos os experimentos realizados para a Petrobras, com valor de 0,5502.

Também é interessante notar que as FTLFNs cujos sinais de entrada foram sujeitos a memó-

Figura 13: Variação no erro de acordo com o número de neurônios na camada oculta - Vale

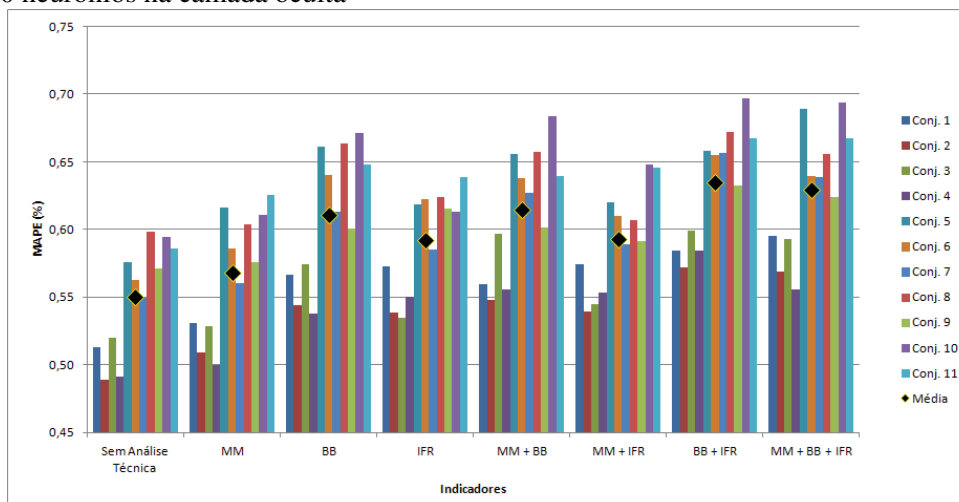
Fonte: Elaborado pelo autor desta dissertação.

rias de linha de atraso mais curtas apresentaram impactos maiores nos resultados conforme o número de neurônios em suas camadas ocultas foram sendo modificados. Este comportamento pode ser observado através do cálculo do desvio padrão do MAPE para cada configuração de atrasos. Para memórias de tamanho 3, 5 e 7, seus respectivos desvios padrões são 0,0220, 0,0180 e 0,0201. Já a partir de 13 atrasos, os desvios padrões se mantêm sempre abaixo de 0,0150, atingindo o valor mínimo de 0,0086 com o número máximo de atrasos experimentado.

Diferentemente do que aconteceu com Petrobras, onde os erros mínimos de previsão foram alcançados pelo uso de 20 ou 30 neurônios na camada oculta da FTLFN, os experimentos para a Vale exigiram 30 ou 50 para atingirem tal condição. A Figura 13 demonstra o comportamento geral obtido através das experimentações. Adotando-se 10 neurônios, observou-se um MAPE médio de 0,4153. Com 20, apenas a rede submetida a sinais com 3 atrasos não apresentou diminuição no erro de previsão, o que coincide com o padrão percebido para Petrobras. Nesta configuração, o MAPE médio mostrou o valor de 0,4030, um ganho de 2,96% sobre a parametrização anterior. Já o emprego de 30 neurônios na camada oculta foi benéfico a todos os tamanhos de históricos aplicados sobre os sinais de entrada, apresentando um MAPE médio de 0,3899. Da mesma forma como para a Petrobras, esta foi a definição de camada oculta que melhor permitiu à FTLFN aproximar valores previstos e esperados para a Vale. Ainda assim, em 3 experimentações (atrasos de 3, 5 e 25) o uso de 50 neurônios melhorou ainda mais os resultados. Esta configuração se apresentou como a segunda melhor, com um MAPE médio de 0,3907. Quanto ao uso de 80 neurônios na camada oculta, o mesmo foi benéfico unicamente para o experimento com 13 atrasos sobre os sinais de entrada, mas ainda assim este não foi o melhor resultado para este tamanho de memória. Para as configurações de atrasos restantes, o número maior de neurônios ocultos aumentou os erros médios de previsão.

De forma análoga ao ocorrido com a Petrobras, também para a Vale o melhor resultado

Figura 14: Influência de Indicadores da Análise Técnica - Petrobras, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta



Fonte: Elaborado pelo autor desta dissertação.

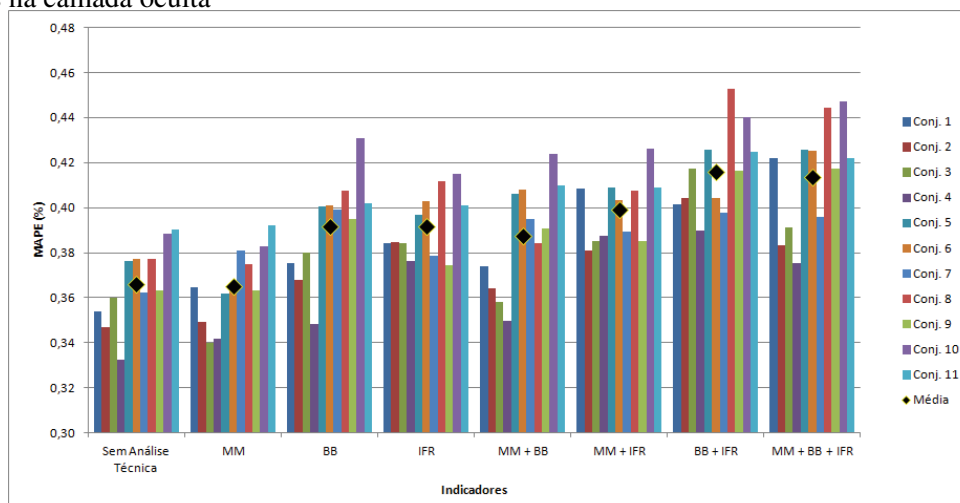
individual foi atingido através do uso de uma memória de linha de atraso de 15 passos e de uma camada oculta da FTLFN composta por 30 neurônios, mostrando um MAPE de 0,3662.

O comportamento referente à maior estabilização de resultados ao se variar o número de neurônios na camada oculta, pelo uso de uma memória de linha de atraso maior, também se mostrou presente nos experimentos com a Vale. Para 3 atrasos, o desvio padrão observado entre os valores medidos para 10, 20, 30, 50 e 80 neurônios foi de 0,0285. Entre 5 e 11 atrasos, este valor encontrou-se entre 0,0124 e 0,0149. E a partir de uma memória de 13 passos no passado, o desvio padrão diminuiu constantemente, variando de 0,0082 até o valor mínimo encontrado no experimento, de 0,0037 para 25 atrasos.

4.6.3 Resultados das experimentações sobre indicadores da Análise Técnica

O objetivo deste experimento foi verificar o impacto nas previsões por parte das FTLFNs, no período intradiário, ao se introduzir indicadores da Análise Técnica (e combinações dos mesmos) ao modelo. Eles são fornecidos à rede no mesmo formato que os sinais de preços e volumes. Por exemplo, no caso do experimento onde foram empregados 3 atrasos sobre o conjunto 8, agregando-se como suporte a Média Móvel, supriu-se a camada de entrada da FTLFN com 9 valores: 3 referentes aos atrasos do sinal de Fechamento, 3 do Volume, e 3 do indicador Média Móvel. Dado que, no presente momento do estudo, os resultados para os testes com atrasos e neurônios na camada oculta já se encontravam disponíveis, as configurações que forneceram as melhores aproximações médias entre valores esperados e previstos foram utilizadas como base para a inclusão dos indicadores. Tais parâmetros correspondem a 15 atrasos e 30 neurônios para ambas Petrobras e Vale. Esta abordagem permite que se identifique se sinais adicionais são capazes de melhorar ainda mais as previsões, além dos resultados otimizados

Figura 15: Influência de Indicadores da Análise Técnica - Vale, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta



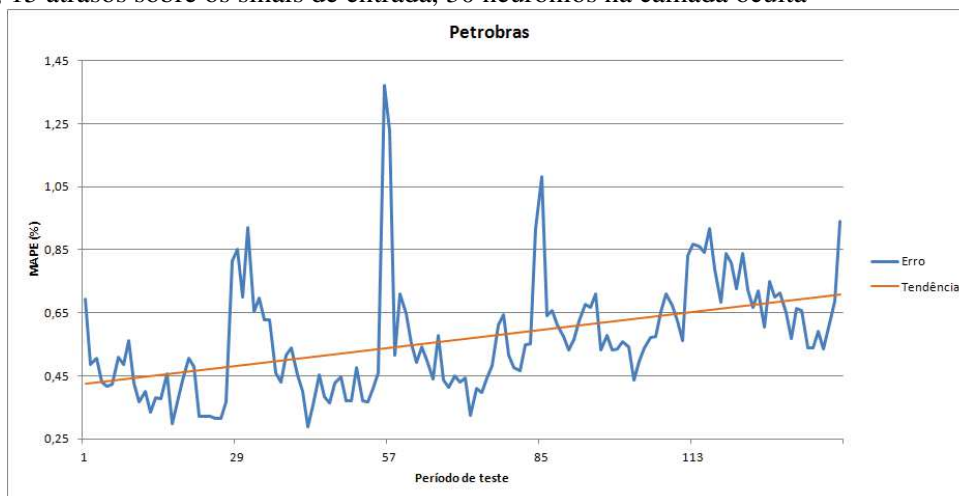
Fonte: Elaborado pelo autor desta dissertação.

encontrados dentro do contexto proposto neste trabalho. O apêndice C expõe tabelas com os valores MAPE advindos das experimentações com a Análise Técnica.

A Figura 14 mostra os resultados referentes à Petrobras. Fica claro que, na média, os indicadores da Análise Técnica selecionados não proveram meios para melhor aproximar previsões de preços no período intradiário. De fato, todos contribuíram para o aumento do erro, sendo eles empregados isoladamente, em pares, ou os três de forma conjunta. Ao se agregar apenas um indicador ao modelo, juntamente com os sinais de preço, volume e quantidade, o menor valor MAPE, de 0,5677, se deu para a Média Móvel, significando um aumento no erro de previsão de 3,18% quando comparado ao modelo sem indicadores da Análise Técnica, com MAPE de 0,5502. Já as Bandas de Bollinger e o IFR pioram o resultado inicial em 11,04% e 7,60%, respectivamente, com MAPE de 0,6109 e 0,5920. Considerando indicadores em pares, o menor erro de previsão alcançado foi através da Média Móvel utilizada em conjunto com o IFR, com um MAPE de 0,5929, uma perda de 7,77% sobre o modelo sem indicadores. Os pares remanescentes, por sua vez, apresentaram resultados piores, sendo aquele formado por Bandas de Bollinger e IFR o que apresentou o maior erro entre todos os experimentos com Análise Técnica para Petrobras, com MAPE de 0,6344, 15,30% acima do modelo base. A agregação dos três indicadores ao modelo, de forma paralela, foi responsável por um MAPE de 0,6291.

Para a Vale, como demonstrado na Figura 15, a adição da Média Móvel ao modelo melhora levemente o resultado médio, diminuindo o MAPE em 0,29%, de 0,3662 para 0,3651, do modelo sem Análise Técnica. As Bandas de Bollinger e o IFR, no entanto, apresentaram aumento no erro da ordem de 6,92% e 7,00%, respectivamente, com MAPEs de 0,3915 e 0,3918. O uso conjunto destes dois indicadores apresentou a pior acurácia de previsão no experimento com a Vale, com um MAPE de 0,4158, refletindo em uma perda de 13,55% quando comparado ao modelo base. O melhor par, unindo Bandas de Bollinger e Média Móvel, mostrou um MAPE

Figura 16: Degradação no desempenho de previsão da FTLFN ao longo de uma semana de dados - Petrobras, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta



Fonte: Elaborado pelo autor desta dissertação.

de 0,3876. Agregando-se os três indicadores ao mesmo tempo, atingiu-se um MAPE de 0,4135, muito próximo ao pior resultado atingido dentro do contexto proposto.

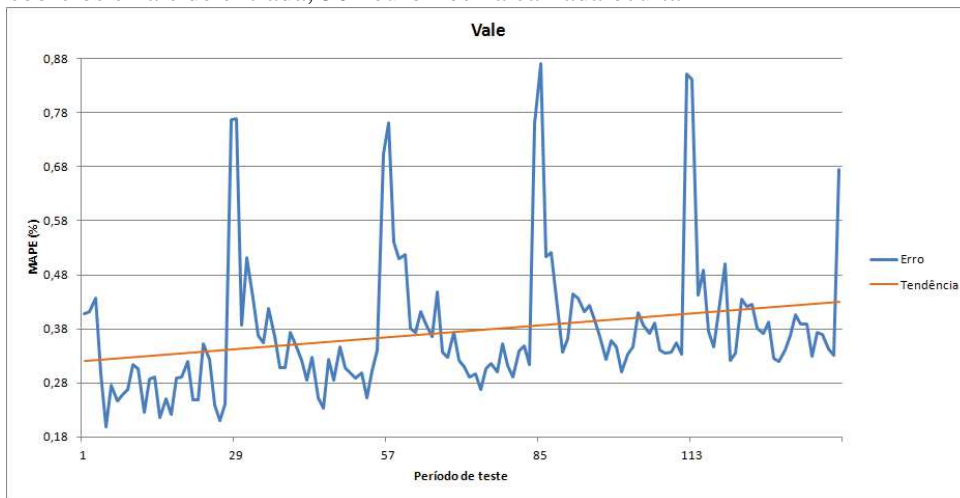
4.6.4 Resultados para experimentações sobre a degradação da rede

O uso de *janelas deslizantes* neste trabalho, como especificado anteriormente, deveu-se à sua proposta de evitar a degradação na acurácia de previsão da rede, conforme ela é utilizada ao longo do tempo, sem um treinamento periódico com novos dados (KAASTRA; BOYD, 1996).

Como forma de se visualizar o nível de tal degradação para redes do tipo FTLFN, quando expostas a cenários intradiários de previsão, foi realizada uma experimentação adicional na presente dissertação. Baseado na literatura correspondente, espera-se uma diminuição no desempenho da rede ao longo do tempo, independente da parametrização e do modelo adotados. Portanto, para este experimento tomou-se como base as configurações que apresentaram os melhores resultados nos testes anteriores, sendo a rede FTLFN parametrizada com 30 neurônios na camada oculta, e a memória de linha de atraso com 15 períodos.

Da mesma forma como realizado nas experimentações anteriores, treinou-se a rede com dados relativos a 3 semanas de negociações de preços, e testou-se o desempenho da mesma sobre dados da semana seguinte. Os conjuntos de teste são compostos, individualmente, por 140 períodos referentes a 15 minutos cada. Armazenou-se o erro verificado para cada um destes períodos. O mesmo procedimento foi realizado sobre todos os 35 agrupamentos treino-teste, de forma a compreender a totalidade do histórico capturado, sendo ainda repetido 5 vezes, como explicado na seção relativa às experimentações planejadas. Após completada a execução para os 11 conjuntos de sinais de entrada, tomou-se a média dos valores MAPE observados período a período dentro das semanas de teste. Assim, foi possível construir os gráficos das Figuras

Figura 17: Degradação no desempenho de previsão da FTLFN ao longo de uma semana de dados - Vale, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta



Fonte: Elaborado pelo autor desta dissertação.

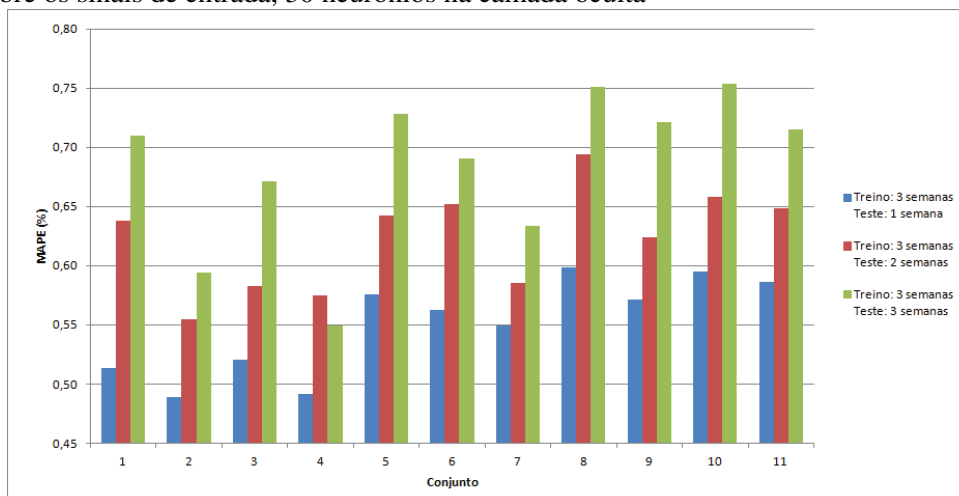
16 e 17, onde se observam as evoluções dos erros médios de previsão ao longo do intervalo para Petrobras e Vale, respectivamente. No apêndice D encontram-se as tabelas com os valores relacionados.

Percebe-se claramente a tendência à degradação no desempenho de previsão da FTLFN à medida em que se realizam previsões para períodos mais distantes do conjunto de treinamento, sem que haja re-treino da rede com dados recentes. O aumento no erro é mais abrupto para a Petrobras, porém também para a Vale este comportamento é bastante significativo.

Os períodos de 1 a 28 equivalem aos *ticks* de 15 minutos representantes das 7 horas de negociação do primeiro dia de teste. Neste intervalo, o MAPE médio para a Petrobras foi de 0,4332. Já no segundo dia, definido entre os períodos 29 e 56, a medida de erro apresentou o valor de 0,5287, 22,04% pior do que o anterior. O terceiro e quarto dias, compreendidos entre os períodos 57 e 84, e 85 e 112, respectivamente, mostraram MAPEs médios de 0,5416 e 0,6178, confirmando o aumento gradativo do erro ao longo do tempo. Finalmente, o quinto dia dentro da semana de teste apresentou o maior MAPE médio, de 0,7159, 15,88% menos preciso que o dia anterior, e 65,24% pior que o primeiro dia testado. É interessante notar a partir do gráfico, ainda, a ocorrência de picos nos erros em regiões próximas à alternância entre os dias. O comportamento demonstra uma maior dificuldade da FTLFN em prever mudanças mais abruptas nos preços, os quais acontecem com frequência em ocasiões como a abertura e o fechamento das negociações dentro do período de um dia.

Para a Vale, o MAPE médio do primeiro dia apresentou o valor de 0,3015, seguido de 0,3614 para o dia seguinte, representando uma piora de 19,88%. No terceiro e quarto dias obtiveram-se os erros de 0,3879 e 0,4159, respectivamente, dando continuidade à degradação no desempenho da rede. O quinto dia, com MAPE de 0,4096, representou uma leve melhora de 1,50% em relação ao dia anterior, porém ainda assim 35,87% pior do que o primeiro dia de teste após

Figura 18: Degradação no desempenho de previsão ao longo de duas e três semanas - Petrobras, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta



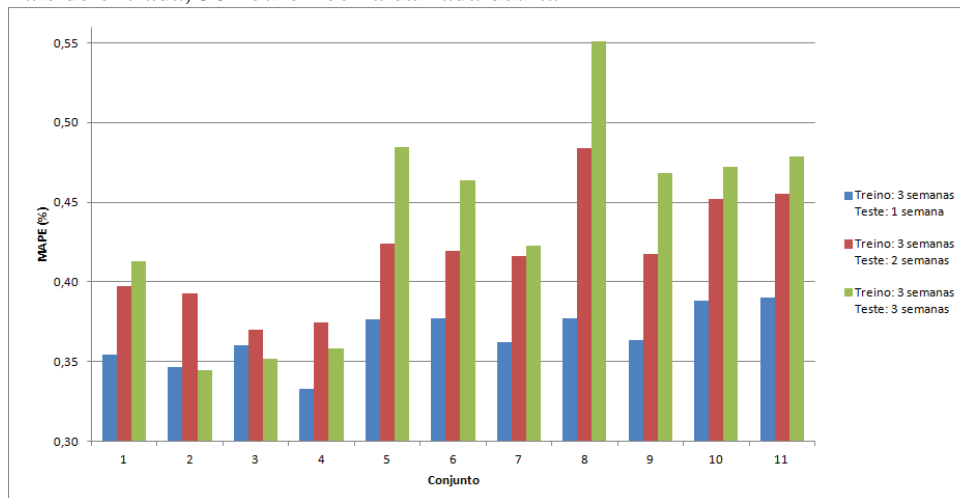
Fonte: Elaborado pelo autor desta dissertação.

o treino. Também no caso deste ativo, e ainda de forma mais evidente que para a Petrobras, é possível notar a presença de picos nos valores MAPE, próximos ao fechamento das negociações em um dia e à abertura do dia seguinte. Estas ocorrências confirmam a observação relatada anteriormente a respeito de maiores imprecisões nas previsões da FTLFN quando sujeitas a modificações bruscas nos preços.

O experimento foi estendido, com o intuito de demonstrar a degradação no desempenho caso a FTLFN recebesse treinamentos ainda menos frequentes. Foram realizadas experimentações onde a mesma rede foi utilizada para previsões durante duas e três semanas, sem novos re-treinamentos durante os períodos. A Figura 18 explicita os resultados para Petrobras, os quais também foram discretizados por conjunto de sinais de entrada. No apêndice D são apresentadas as tabelas correspondentes. De maneira geral, esta experimentação confirmou o padrão de degradação visto no teste anterior, onde quanto maior o tempo que se faz uso da rede sem novos treinamentos, pior é o seu desempenho. Em média, o MAPE para uma, duas e três semanas seguidas de teste apresentou, respectivamente, os valores de 0,5503, 0,6230 e 0,6834. Houve, portanto, um aumento no erro da ordem de 13,22% ao se estender o uso da mesma rede de uma para duas semanas. E empregando-a ainda para três semanas, a piora no MAPE foi de 9,69% em relação a duas semanas, e de 24,20% comparando-se com apenas uma semana. Observando-se os conjuntos individualmente, somente o 4 não apresentou o mesmo comportamento da média, uma vez que o uso da rede para três semanas levou a um erro médio levemente menor do que para duas. Porém, também neste caso é possível afirmar que re-treinamentos semanais da FTLFN conduziram a um desempenho melhor do que caso eles ocorressem em espaços mais distantes no tempo.

O gráfico correspondente para a Vale é apresentado na Figura 19. Também neste caso é confirmada a degradação no desempenho da FTLFN, de maneira geral, ao utilizá-la por um

Figura 19: Degradação no desempenho de previsão ao longo de duas e três semanas - Vale, 15 atrasos sobre os sinais de entrada, 30 neurônios na camada oculta



Fonte: Elaborado pelo autor desta dissertação.

tempo maior sem o emprego do re-treino. Quando, após três semanas de treino, ela foi adotada apenas para previsões dentro da primeira semana seguinte, o MAPE médio apresentado foi de 0,3662. Já quando testada para duas e três semanas, produziu erros de valores 0,4186 e 0,4371, respectivamente, o equivalente a pioras de 14,30% e 19,37% em relação ao resultado para uma semana. Houveram três exceções a este comportamento, neste experimento. Para o conjunto 4, assim como ocorrido com a Petrobras, foi observada uma diminuição no erro médio ao se utilizar a mesma rede por três semanas, comparado ao uso por duas semanas. Ainda assim, ambos os valores MAPE foram superiores ao emprego da FTLFN apenas durante a primeira semana após o treino. Já para os conjuntos 2 e 3 não foi possível confirmar a degradação no desempenho pelo uso mais prolongado da rede. Para duas semanas o erro foi maior que para uma semana, porém para três semanas as previsões médias foram as mais precisas entre as três configurações testadas. Ainda que esta diferença tenha sido pequena, as evidências mostram que, mesmo podendo-se afirmar, de forma genérica, a ocorrência da degradação no desempenho da FTLFN no período intradiário ao não se fazer uso de treinamentos constantes, certos conjuntos de sinais podem não sofrer a mesma influência negativa do tempo em suas capacidades preditivas.

4.7 Discussão

Um dos objetivos do trabalho proposto nesta dissertação foi o de entender o comportamento da FTLFN quando sujeita a históricos de dados dentro do período intradiário. As experimentações referentes à memória de linha de atraso permitem tirar conclusões a respeito desse objetivo.

Tanto para o caso da Petrobras quanto da Vale, a construção dos sinais de entrada com 3 atrasos forneceu os piores resultados médios entre os testes realizados. À medida em que histó-

ricos maiores de preços e volumes foram providos à FTLFN para a tomada de decisão, melhor foi sua capacidade de aproximar os resultados previstos daqueles que efetivamente aconteceram. Esta afirmação é verdadeira até um tamanho de memória de 11 passos no passado para Petrobras, equivalente a 2 horas e 45 minutos, e de 20 passos para a Vale, ou 5 horas de dados do mercado. A partir destes pontos, o fornecimento de maiores informações à rede não se mostrou igualmente relevante. Para a Vale, o uso de 25 atrasos aumentou o erro de previsão, enquanto que para a Petrobras houveram pequenas variações, positivas e negativas, muito próximas entre si, a partir de 13 atrasos.

No entanto, o resultado apresentado no experimento com a memória de linha de atraso para a Petrobras, onde a FTLFN apresentou seu menor erro quando foram fornecidos à ela sinais de entrada com 25 atrasos, leva à conclusão de que a rede pode não ter atingido seu limite máximo de aprendizado e capacidade de generalização. Portanto, é clara a possibilidade de se obter aproximações intradiárias entre preços esperados e previstos ainda melhores, através da identificação de padrões de movimentos de preços e volumes mais duradouros, fornecendo-se históricos maiores à rede. O mesmo pode ser dito a respeito da Vale, cujo melhor resultado, dentro dos parâmetros adotados no contexto deste trabalho, foi alcançado com o uso de 20 atrasos, muito próximo ao limite testado de 25 passos no passado. É possível que uma memória de linha de atraso maior diminua os erros de previsões intradiárias para este ativo, assim como aconteceu com a Petrobras.

É interessante reparar que os resultados apresentados nesta pesquisa demonstram a importância de se explorar a principal característica das FTLFNs, que é a sua capacidade de aprendizado através de padrões de preços e volumes passados, apresentados a ela em sua camada de entrada. No entanto, existe uma dificuldade de se encontrar na literatura trabalhos que explorem esta funcionalidade de forma extensa. Enquanto que nos experimentos relatados acima os menores erros são atingidos a partir de uma memória de linha de atraso de 20 passos, os estudos de Gomide e Milidiú (2010) e Costa (2012) não ultrapassam 10 atrasos em seus testes, enquanto que Costa e Crepaldi (2014) reportam os resultados apenas para o uso de 8 atrasos, como já discutido na seção de trabalhos correlatos. Assim como exposto nesta dissertação, independente da abordagem adotada, é importante que estudos de previsões por redes do tipo FTLFN embasem resultados em experimentações abrangentes, o que parece ter sido limitado nas pesquisas relatadas.

O teste sobre 11 diferentes conjuntos de sinais de entrada e 5 tamanhos de camada oculta forneceu robustez aos resultados, uma vez que limitou a exposição da FTLFN a especificidades de um determinado modelo. Além disso, a abordagem sobre os diversos conjuntos permitiu a identificação de alguns mais promissores, habilitando a FTLFN a apresentar melhores previsões do que outros. Estes resultados demonstram que apesar de se poder observar um padrão geral de comportamento deste tipo de rede no intervalo intradiário, a seleção cuidadosa dos sinais de entrada tem papel fundamental na busca por limites máximos de aproximação de preços que a FTLFN possa alcançar.

Comparando-se os experimentos entre Petrobras e Vale é possível inferir que a FTLFN foi capaz de interpretar de forma mais precisa os dados relativos à Vale, ou que os preços desta são mais previsíveis que os da Petrobras, uma vez que todos os resultados médios apresentados para a Vale mostraram erros menores do que aqueles relacionados à Petrobras. De forma mais enfática, o menor MAPE médio para a Petrobras ainda é maior do que o pior resultado encontrado para a Vale.

Em relação aos experimentos sobre o número de neurônios na camada oculta da FTLFN, Petrobras e Vale também apresentaram similaridades e diferenças em seus resultados. Em ambos os casos, os limites inferior e superior testados, de 10 e 80 neurônios, respectivamente, foram responsáveis pelos maiores erros médios encontrados na experimentação. Os melhores resultados para Petrobras mostraram-se, em sua maioria, quando da adoção de 20 ou 30 neurônios na camada oculta, sugerindo limites ideais para o ativo, dentro do contexto proposto neste trabalho. Por outro lado, o fato do uso de 50 neurônios apresentar o menor erro para 25 atrasos provê uma ideia de que futuras experimentações sobre esta ação, considerando estruturas de memória mais longas, possam requerer um número de neurônios superior na camada oculta da rede, a fim de fornecer previsões mais precisas. No caso da Vale, 30 e 50 neurônios foram suficientes para se alcançar os menores valores médios MAPE, dependendo do tamanho da memória de linha de atraso adotado para os sinais de entrada. Conclui-se, assim, que para se encontrar o tamanho ideal da camada oculta da FTLFN, depende-se fortemente de experimentações amplas, uma vez que se provaram distintos para dois ativos testados sob condições de parametrizações similares.

Analisando-se o mesmo experimento sob diferente aspecto, foi possível reparar um claro comportamento, tanto para Petrobras quanto para Vale, em que quanto maior o histórico de preços e volumes fornecido à rede, menores foram os impactos nos resultados ao se variar o tamanho da camada oculta da FTLFN. Ou seja, ao se adotar memórias de linha de atraso curtas, experimentações com diferentes números de neurônios tiveram maior importância, uma vez que influenciaram vigorosamente nos erros de previsão. À medida em que se forneceu sinais com atrasos mais longos, o tamanho da camada oculta provocou modificações menos expressivas nos resultados. Estas observações levam à conclusão de que provendo-se à rede padrões maiores de movimentos de preços e volumes, é possível não somente melhorar o desempenho das FTLFNs no contexto deste estudo, mas também torná-las mais estáveis em relação aos erros de aproximação.

A partir do conteúdo relatado até o momento, é possível afirmar que as explorações tanto do tamanho da memória de linha de atraso, quanto do número de neurônios na camada oculta da FTLFN, são de grande importância para uma melhor previsão de preços no período intradiário. Ambos contribuem, dentro de limitações, para melhores aproximações entre valores esperados e obtidos.

Os resultados das experimentações sobre a inclusão da Análise Técnica ao modelo preditivo, da forma como propostas, mostraram que os indicadores não foram capazes de contribuir positivamente para que as FTLFNs apresentassem melhoras em suas previsões de preços no pe-

ríodo intradiário, dentro do cenário estudado. Apenas para a Vale, a agregação da Média Móvel ao modelo trouxe diminuição do erro MAPE. Ainda assim, a magnitude do ganho foi ínfima, de baixa representatividade. De acordo com Murphy (1999), investidores costumam utilizar um indicador para confirmar a informação fornecida por outro, aumentando a confiança nos sinais de compra ou venda identificados. Contrariando esta afirmativa, o uso conjunto de múltiplos indicadores, considerando os cenários explorados nesta pesquisa, aumentou o ruído ao qual a FTLFN se sujeitou, elevando também o erro médio produzido. Dentro do contexto desta dissertação, portanto, a exploração de diferentes atrasos e tamanhos da camada oculta da FTLFN foi mais promissora que a exploração da Análise Técnica, no suporte a melhores aproximações de previsão.

Os resultados negativos obtidos, no entanto, não significam que indicadores da Análise Técnica devam ser descartados como ferramentas auxiliares a previsões de preços por FTLFNs, no período intradiário. Diferentes abordagens e a escolha entre centenas de indicadores disponíveis podem, eventualmente, servir de apoio a previsões mais precisas. Segundo Achelis (2013), alguns indicadores funcionam melhor em mercados sem tendência, enquanto outros em mercados com tendência. Portanto, pode ser interessante determinar o ferramental a ser utilizado de acordo com este critério, por exemplo.

Dado o comportamento de degradação da FTLFN apresentado para o período intradiário no experimento correspondente, com exceção de conjuntos isolados cujos comportamentos desviaram-se levemente do restante, fica clara a importância do treinamento frequente da rede, agregando-se dados recentes ao modelo de tempos em tempos. Visto que, de maneira geral, houve um aumento significativo no erro de previsão, tanto para Petrobras quanto para Vale, ao se estender o uso da rede para duas, e principalmente para três semanas, o emprego da técnica de *janelas deslizantes* mostrou-se de enorme importância para que se atingissem os resultados apresentados no restante das experimentações. Através dos testes sobre a degradação, também confirmou-se que o conhecimento advindo da literatura a respeito deste método é igualmente válido para o contexto abordado neste trabalho.

Mesmo analisando-se unicamente a primeira semana de testes após o treinamento da FTLFN, observou-se uma degradação importante no desempenho da rede entre o primeiro e o último dia. Assim, é possível que treinamentos mais periódicos, considerando-se intervalos menores que o utilizado nesta pesquisa, apresentem resultados ainda melhores de previsão intradiária.

5 CONCLUSÃO

Redes neurais do tipo FTLFN têm apresentado resultados bastante promissores ao serem adotadas como ferramentas preditivas no mercado de ações. Seu mecanismo de memória é capaz de detectar padrões temporais de preços e volumes negociados, e utilizá-los como base para a inferência de valores futuros. Grande parte dos estudos que a utilizam, no entanto, exploram de forma limitada suas reais capacidades, adotando parametrizações de maneira empírica e muitas vezes obscuras ao público acadêmico. Existe também uma divergência entre a literatura disponível relacionada a previsões de preços de ações, e a realidade do mercado, no que tange à periodicidade analisada. Enquanto que há uma facilidade maior de se encontrar pesquisas abordando a baixa frequência, como dias e meses, atualmente as negociações em bolsas de valores são realizadas cada vez mais sob horizontes de curto prazo, sejam horas, minutos ou segundos. Assim, o presente trabalho propôs-se a expandir o conhecimento relativo principalmente sobre estes dois aspectos.

Foram realizadas experimentações com o objetivo de se entender, de forma aprofundada, o impacto no comportamento das FTLFNs a partir do uso de diferentes profundidades de históricos de sinais e de variados tamanhos de camada oculta. Os dados apresentados à rede e o horizonte de previsão são relativos ao período intradiário, sendo possível medir o desempenho deste tipo de rede sob condições mais próximas ao estado atual dos mercados. Além disso, estudou-se também o papel de indicadores da Análise Técnica como ferramentas de suporte a melhores aproximações. A partir dos resultados obtidos neste trabalho, é possível listar as seguintes contribuições frente a literatura:

- Há grandes vantagens em se explorar de forma ampla o tamanho da memória de linha de atraso referente aos sinais fornecidos como entrada à FTLFN, no período intradiário, uma vez que se observou um impacto relevante no desempenho de previsão da rede pela adoção de diferentes configurações deste parâmetro.
- Os maiores ganhos nas aproximações entre valores previstos e esperados ocorreram ao se empregar até 9 atrasos, de 15 minutos cada, para ambas Petrobras e Vale. No entanto, mesmo a partir de históricos de 11 passos pode-se observar diminuição nos erros de previsão, ainda que relativamente menores.
- Dentro do contexto proposto, existe a possibilidade de que o fornecimento à rede de sinais com mais de 25 atrasos, equivalentes a históricos de 6 horas e 15 minutos, resultem em erros de previsão ainda menores do que os observados, uma vez que para a Petrobras o valor médio neste passo foi o mínimo atingido no experimento, e para a Vale o valor se mostrou muito próximo ao mínimo.
- Através dos conjuntos de sinais de entrada formados unicamente pelos preços de fechamento, máximo e mínimo, e alternativamente também pelo preço de abertura, a FTLFN

atingiu os menores erros médios de previsão, tanto para Petrobras quanto para Vale. Este é um indicativo de que são, portanto, os mais promissores em possíveis trabalhos que os utilizem para aplicações do modelo em situação de negociações reais.

- A realização de experimentos com a FTLFN, considerando diferentes conjuntos de sinais de entrada, permite uma maior confiabilidade em relação aos resultados, uma vez que estes são apresentados como valores médios, provendo uma noção geral do comportamento da rede.
- O número ideal de neurônios na camada oculta da FTLFN, a fim de alcançar resultados de previsão mais precisos, pode ser encontrado através de refinamentos das experimentações, variando de acordo com as características dos dados submetidos à rede. De maneira geral, para a base de preços e volumes utilizados neste trabalho, FTLFNs com tamanhos de camada oculta entre 20 e 50 neurônios apresentaram os menores erros médios.
- Quanto mais longos os padrões dos sinais intradiários expostos na camada de entrada, maior a estabilidade da rede frente a mudanças no número de neurônios da camada oculta. Em outras palavras, preços previstos pela rede FTLFN estão sujeitos a menores variações pelas modificações em sua camada oculta, quando são capazes de visualizar padrões de movimentos passados mais duradouros.
- A agregação, ao modelo, de sinais construídos a partir dos indicadores da Análise Técnica Média Móvel, Bandas de Bollinger e Índice de Força Relativa, não se apresentou benéfica para melhorias de previsões. Tanto o uso individual dos mesmos, quanto de forma conjunta, fez com que os erros médios aumentassem, de maneira geral.

Agregando-se os itens acima expostos, espera-se que o estudo desenvolvido nessa dissertação sirva como base para trabalhos futuros que se utilizem de redes neurais do tipo FTLFN para a previsão de preços intradiários. Neste sentido, sugere-se como possíveis extensões a esta pesquisa os seguintes aspectos:

- A exploração de diferentes configurações dos parâmetros da FTLFN, assim como de memórias de linhas de atraso mais longas, dentro do mesmo contexto proposto neste trabalho, na tentativa de se encontrar os limites a partir dos quais exista uma real degradação no desempenho de previsão.
- Alguns conjuntos específicos de sinais de entrada permitiram à rede neural mostrar um desempenho muito superior à média. Portanto, o desenvolvimento de um trabalho com foco nestes agrupamentos possui o potencial de apresentar resultados surpreendentes de previsão.

- Apenas três indicadores da Análise Técnica foram utilizados neste estudo, ao passo que centenas deles estão disponíveis. Assim, muitos ainda podem ser experimentados com variadas parametrizações, buscando-se contribuições particulares, como visto em diferentes pesquisas.
- A identificação de sazonalidades ao longo das séries temporais de preços, possibilitando que indicadores da Análise Técnica melhor adaptados a cada situação sejam devidamente adotados.
- Um estudo comparativo entre a FTLFN e outros tipos de rede neurais aptas a captarem comportamentos temporais, como as RNN (*Recurrent Neural Network*) com arquitetura LSTM (*Long Short-Term Memory*), utilizando a mesma abordagem do trabalho aqui proposto.
- Criando-se regras para a montagem e desmontagem de posições em ações, e considerando-se custos financeiros envolvidos, é possível simular a aplicação do modelo desenvolvido neste estudo a sistemas de negociação real, no período intradiário, demonstrando potenciais retornos financeiros.

REFERÊNCIAS

- ACHELIS, S. **Technical Analysis from A to Z**. 2. ed. [S.l.]: McGraw-Hill, 2013. 400 p.
- ARAÚJO, R. A. Hybrid intelligent methodology to design translation invariant morphological operators for Brazilian stock market prediction. **Neural Networks**, [S.l.], v. 23, n. 10, p. 1238–1251, 2010.
- ARDALANI-FARSA, M.; ZOLFAGHARI, S. Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks. **Neurocomputing**, [S.l.], v. 73, n. 13–15, p. 2540–2553, 2010.
- ARMANO, G.; MARCHESI, M.; MURRU, A. A hybrid genetic-neural architecture for stock indexes forecasting. **Information Sciences**, [S.l.], v. 170, n. 1, p. 3–33, 2005.
- ATSALAKIS, G. S.; VALAVANIS, K. P. Surveying stock market forecasting techniques – Part II: soft computing methods. **Expert Systems with Applications**, [S.l.], v. 36, n. 3, Part 2, p. 5932–5941, 2009.
- AZOFF, E. M. **Neural Network Time Series: forecasting of financial markets**. 1. ed. [S.l.]: Wiley, 1994. 212 p.
- BEALE, M. H.; HAGAN, M. T.; DEMUTH, H. B. **Neural Network Toolbox User's Guide**. [S.l.]: The MathWorks, Inc., 2013.
- BISHOP, C. M. **Pattern Recognition and Machine Learning**. 1. ed. [S.l.]: Springer, 2007.
- BM&F Bovespa. Disponível em:
<http://www.bmfbovespa.com.br/Pdf/cap_06_do_after_market.pdf>. Acesso em: abr. 2015.
- BM&F Bovespa. Disponível em: <<http://bmfbovespa.com.br/pt-br/noticias/2014/BMFBOVESPA-divulga-balanco-de-operacoes-de-2013-2014-01-07.aspx>>. Acesso em: out. 2014.
- BOLLEN, J.; MAO, H.; ZENG, X. Twitter mood predicts the stock market. **Journal of Computational Science**, [S.l.], v. 2, n. 1, p. 1–8, 2011.
- BOLLINGER, J. A. **Bollinger on Bollinger Bands**. 1. ed. [S.l.]: McGraw-Hill, 2001. 227 p.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: teoria e aplicações**. 2. ed. [S.l.]: LTC, 2014. 226 p.
- BURDEN, F.; WINKLER, D. Bayesian regularization of neural networks. **Methods in Molecular Biology**, [S.l.], v. 458, p. 23–42, 2009.
- CAVALCANTE, F.; MISUMI, J. Y.; RUDGE, L. F. **Mercado de Capitais**. 7. ed. [S.l.]: Elsevier, 2009. 424 p.
- CHANG, P.-C.; LIAO, T. W.; LIN, J.-J.; FAN, C.-Y. A dynamic threshold decision system for stock trading signal detection. **Applied Soft Computing**, [S.l.], v. 11, n. 5, p. 3998–4010, 2011.

- CHANG, P.-C.; LIU, C.-H. A TSK type fuzzy rule based system for stock price prediction. **Expert Systems with Applications**, [S.l.], v. 34, n. 1, p. 135–144, 2008.
- CHAVARNAKUL, T.; ENKE, D. A hybrid stock trading system for intelligent technical analysis-based equivolume charting. **Neurocomputing**, [S.l.], v. 72, n. 16–18, p. 3517–3528, 2009.
- CHEN, A.-S.; LEUNG, M. T.; DAOUK, H. Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index. **Computers & Operations Research**, [S.l.], v. 30, n. 6, p. 901–923, 2003.
- COSTA, A. F. **Previsão de séries temporais financeiras por meio de redes neurais dinâmicas e processos de transformação de dados: uma abordagem empírico-comparativa**. 2012. 97 p. Dissertação (Mestrado em Ciência da Computação) — UNESP, São Paulo, 2012.
- COSTA, A. F.; CREPALDI, A. F. The bias in reversing the Box–Cox transformation in time series forecasting: an empirical study based on neural networks. **Neurocomputing**, [S.l.], v. 136, p. 281–288, 2014.
- COSTA, F. V.; 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, 2012, São Paulo, SP. **Anais...** [S.l.: s.n.], 2012.
- COULIBALY, P.; EVORA, N. Comparison of neural network methods for infilling missing daily weather records. **Journal of Hydrology**, [S.l.], v. 341, n. 1–2, p. 27–41, 2007.
- CYBENKO, G. Continuous valued neural networks with two hidden layers are sufficient. **Mathematics of Control, Signal and Systems**, [S.l.], v. 2, p. 304–314, 1989.
- DING, Z.; GRANGER, C. W.; ENGLE, R. F. A long memory property of stock market returns and a new model. **Journal of Empirical Finance**, [S.l.], v. 1, n. 1, p. 83–106, 1993.
- FAMA, E. F. The Behavior of Stock-Market Prices. **The Journal of Business**, [S.l.], v. 38, n. 1, p. 34–105, 1965.
- FORESEE, F.; HAGAN, M. Gauss-Newton approximation to Bayesian learning. In: INTERNATIONAL CONFERENCE ON NEURAL NETWORKS, 1997. **Anais...** [S.l.: s.n.], 1997. v. 3, p. 1930–1935.
- FREEMAN, J.; SKAPURA, D. **Neural Networks: algorithms, applications, and programming techniques**. [S.l.]: Addison-Wesley, 1991. 550 p.
- FREITAS, F. D.; SOUZA, A. F.; ALMEIDA, A. R. Prediction-based portfolio optimization model using neural networks. **Neurocomputing**, [S.l.], v. 72, n. 10–12, p. 2155–2170, 2009.
- GENÇAY, R.; DACOROGNA, M.; MULLER, U. A.; PICTET, O.; OLSEN, R. **An Introduction to High-Frequency Finance**. 1. ed. [S.l.]: Elsevier Science, 2001. 383 p.
- GEVA, T.; ZAHAVI, J. Empirical evaluation of an automated intraday stock recommendation system incorporating both market data and textual news. **Decision Support Systems**, [S.l.], v. 57, p. 212–223, 2014.
- HASSAN, M. R.; NATH, B.; KIRLEY, M. A fusion model of HMM, ANN and GA for stock market forecasting. **Expert Systems with Applications**, [S.l.], v. 33, n. 1, p. 171–180, 2007.

- HAYKIN, S. **Neural Networks: a comprehensive foundation**. 2nd. ed. NJ, USA: Prentice Hall, 2001. 842 p.
- HERAVI, S.; OSBORN, D. R.; BIRCHENHALL, C. Linear versus neural network forecasts for European industrial production series. **International Journal of Forecasting**, [S.l.], v. 20, n. 3, p. 435–446, 2004.
- HISSA, M. B. **Sobreviva na Bolsa de Valores**. 1. ed. [S.l.]: Campus Elsevier, 2008. 224 p.
- JORDAN, D.; DILTZ, D. The profitability of day traders. **Financial Analysts Journal**, [S.l.], v. 59, n. 6, p. 85–95, 2003.
- KAASTRA, I.; BOYD, M. Designing a neural network for forecasting financial and economic time series. **Neurocomputing**, [S.l.], v. 10, n. 3, p. 215–236, 1996.
- KIM, J. S.; RYU, D. Intraday price dynamics in spot and derivatives markets. **Physica A: Statistical Mechanics and its Applications**, [S.l.], v. 394, p. 247–253, 2014.
- KOSKO, B. **Neural networks and fuzzy sytems: a dynamical systems approach to machine intelligence**. 1. ed. [S.l.]: Prentice-Hall, 1992.
- KUO, W.-Y.; LIN, T.-C. Overconfident individual day traders: evidence from the Taiwan futures market. **Journal of Banking & Finance**, [S.l.], v. 37, n. 9, p. 3548–3561, 2013.
- LAM, M. Neural network techniques for financial performance prediction: integrating fundamental and technical analysis. **Decision Support Systems**, [S.l.], v. 37, n. 4, p. 567–581, 2004.
- LAU, C. K. M.; DEMIR, E.; BILGIN, M. H. Experience-based corporate corruption and stock market volatility: evidence from emerging markets. **Emerging Markets Review**, [S.l.], v. 17, p. 1–13, 2013.
- LIANG, X.; CHEN, R.-C.; HE, Y.; CHEN, Y. Associating stock prices with web financial information time series based on support vector regression. **Neurocomputing**, [S.l.], v. 115, p. 142–149, 2013.
- LO, A.; MACKINLAY, A. Stock market prices do not follow random walks: evidence from a simple specification test. **Review of Financial Studies**, [S.l.], v. 1, n. 1, p. 41–66, 1988.
- MADUREIRA, L.; UNDERWOOD, S. Information, sell-side research, and market making. **Journal of Financial Economics**, [S.l.], v. 90, n. 2, p. 105–126, 2008.
- MARTINEZ, L. C.; HORA, D. N.; PALOTTI, J. R. M.; MEIRA, W.; PAPPA, G. L. From an artificial neural network to a stock market day-trading system: a case study on the BM&F BOVESPA. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 2009. **Anais...** [S.l.: s.n.], 2009. p. 2006–2013.
- MITCHELL, T. M. **Machine Learning**. 1. ed. [S.l.]: McGraw-Hill Education, 1997. 432 p.
- MORETTIN, P. A.; TOLOI, C. M. C. **Previsão de séries temporais**. 1. ed. [S.l.]: Atual Editora, 1987. 436 p.
- MURPHY, J. J. **Technical Analysis of the Financial Markets: a comprehensive guide to trading methods and applications**. [S.l.]: New York Institute of Finance, 1999. 576 p.

NELSON, M.; HILL, T.; REMUS, W.; O'CONNOR, M. Time series forecasting using neural networks: should the data be deseasonalized first? **Journal of Forecasting**, [S.l.], v. 18, n. 5, p. 359–367, 1999.

NETO, M. C. A.; CALVALCANTI, G. D. C.; REN, T. I. Financial time series prediction using exogenous series and combined neural networks. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 2009. **Anais...** [S.l.: s.n.], 2009. p. 149–156.

NIZER, P. S. M.; NIEVOLA, J. C. Predicting published news effect in the Brazilian stock market. **Expert Systems with Applications**, [S.l.], v. 39, n. 12, p. 10674–10680, 2012.

NORONHA, M. **Análise técnica: teorias, ferramentas e estratégias**. 1. ed. Rio de Janeiro: Editec, 1995. 392 p.

NOURANI, V.; KISI, O.; KOMASI, M. Two hybrid Artificial Intelligence approaches for modeling rainfall–runoff process. **Journal of Hydrology**, [S.l.], v. 402, n. 1–2, p. 41–59, 2011.

OUAZZANE, K.; LI, J.; KAZEMIAN, H. B.; JING, Y.; BOYD, R. An Artificial Intelligence-based language modeling framework. **Expert Systems with Applications**, [S.l.], v. 39, n. 5, p. 5960–5970, 2012.

PINHEIRO, J. L. **Mercado de Capitais: fundamentos e técnicas**. 6. ed. [S.l.]: ATLAS, 2012. 560 p.

PRÍNCIPE, J. C.; EULIANO, N. R.; LEFEBVRE, W. C. **Neural and adaptive systems: fundamentals through simulations**. [S.l.]: Wiley, 2000. 672 p.

SETIONO, R.; LIU, H. Neural-network feature selector. **IEEE Transactions on Neural Networks**, [S.l.], v. 8, n. 3, p. 654–662, 1997.

SILVA, E.; CASTILHO, D.; PEREIRA, A.; BRANDAO, H. A neural network based approach to support the Market Making strategies in High-Frequency Trading. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 2014. **Anais...** [S.l.: s.n.], 2014. p. 845–852.

SITTE, R.; SITTE, J. Analysis of the predictive ability of time delay neural networks applied to the S&P 500 time series. **IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews**, [S.l.], v. 30, n. 4, 2000.

STEVENS, L. **Essential Technical Analysis: tools and techniques to spot market trends**. 1. ed. [S.l.]: John Wiley & Sons, Inc., 2002. 390 p.

TAN, T. Z.; QUEK, C.; NG, G. S. BIOLOGICAL BRAIN-INSPIRED GENETIC COMPLEMENTARY LEARNING FOR STOCK MARKET AND BANK FAILURE PREDICTION. **Computational Intelligence**, [S.l.], v. 23, n. 2, p. 236–261, 2007.

TEIXEIRA, L. A.; OLIVEIRA, A. L. I. de. A method for automatic stock trading combining technical analysis and nearest neighbor classification. **Expert Systems with Applications**, [S.l.], v. 37, n. 10, p. 6885–6890, 2010.

THOMSETT, M. C. **Mastering Fundamental Analysis: how to spot trends and pick winning stocks like the pros**. 1. ed. [S.l.]: Kaplan Business, 1998. 256 p.

TICKNOR, J. L. A Bayesian regularized artificial neural network for stock market forecasting. **Expert Systems with Applications**, [S.l.], v. 40, n. 14, p. 5501–5506, 2013.

VANSTONE, B.; FINNIE, G. An empirical methodology for developing stockmarket trading systems using artificial neural networks. **Expert Systems with Applications**, [S.l.], v. 36, n. 3, Part 2, p. 6668–6680, 2009.

VIDYAMURTHY, G. **Pairs Trading: quantitative methods and analysis**. 1. ed. [S.l.]: Wiley, 2004. 224 p.

YANG, K.; WU, M.; LIN, J. The application of fuzzy neural networks in stock price forecasting based On Genetic Algorithm discovering fuzzy rules. In: INTERNATIONAL CONFERENCE ON NATURAL COMPUTATION, 2012. **Anais...** [S.l.: s.n.], 2012. p. 470–474.

APÊNDICE A TABELAS DE RESULTADOS DOS EXPERIMENTOS SOBRE ATRASOS

Tabela 7: MAPE de cada conjunto de entrada, por número de atrasos, representados como a média dos valores obtidos para os testes com 10, 20, 30, 50 e 80 neurônios na camada oculta - Petrobras

	Número de atrasos								
	3	5	7	9	11	13	15	20	25
Conj. 1	0,7372	0,7655	0,7304	0,6643	0,6347	0,6686	0,5748	0,5834	0,5486
Conj. 2	0,7690	0,6660	0,6004	0,5546	0,5380	0,5306	0,5062	0,5318	0,5097
Conj. 3	0,8070	0,6958	0,6918	0,5870	0,5709	0,5672	0,5539	0,5377	0,5392
Conj. 4	0,7426	0,6136	0,5779	0,5357	0,5500	0,5286	0,5123	0,5326	0,4977
Conj. 5	0,6800	0,6415	0,6002	0,5790	0,5765	0,5804	0,5692	0,5978	0,5976
Conj. 6	0,6722	0,6476	0,6080	0,5878	0,5791	0,5700	0,5749	0,5968	0,5891
Conj. 7	0,6569	0,6028	0,5596	0,5487	0,5530	0,5486	0,5765	0,5795	0,5729
Conj. 8	0,7045	0,6951	0,6452	0,6078	0,5942	0,6049	0,6128	0,5838	0,5615
Conj. 9	0,7315	0,6585	0,6228	0,5733	0,5664	0,5706	0,5726	0,5754	0,5554
Conj. 10	0,6754	0,6376	0,6032	0,5896	0,5903	0,5920	0,6035	0,6109	0,6001
Conj. 11	0,6386	0,6062	0,5849	0,5636	0,5663	0,5811	0,5992	0,6167	0,6045
Média	0,7105	0,6573	0,6204	0,5810	0,5745	0,5766	0,5687	0,5769	0,5615

Tabela 8: MAPE de cada conjunto de entrada, por número de atrasos, representados como a média dos valores obtidos para os testes com 10, 20, 30, 50 e 80 neurônios na camada oculta - Vale

	Número de atrasos								
	3	5	7	9	11	13	15	20	25
Conj. 1	0,5576	0,5179	0,4920	0,4434	0,4241	0,4168	0,3983	0,3884	0,3840
Conj. 2	0,5131	0,4253	0,4008	0,3902	0,3583	0,3717	0,3501	0,3397	0,3359
Conj. 3	0,5491	0,4841	0,4384	0,4223	0,3916	0,3943	0,3674	0,3495	0,3568
Conj. 4	0,4594	0,4134	0,4113	0,3587	0,3709	0,3595	0,3424	0,3336	0,3261
Conj. 5	0,4865	0,4218	0,4315	0,3854	0,3925	0,3789	0,3871	0,3809	0,3899
Conj. 6	0,4884	0,3956	0,3895	0,3804	0,3727	0,3716	0,3739	0,3818	0,3841
Conj. 7	0,4489	0,3877	0,3856	0,3811	0,3700	0,3730	0,3691	0,3728	0,3729
Conj. 8	0,5093	0,4310	0,4335	0,4107	0,4173	0,3961	0,3827	0,3930	0,3972
Conj. 9	0,4347	0,4010	0,3771	0,3718	0,3723	0,3737	0,3693	0,3754	0,3753
Conj. 10	0,4198	0,3928	0,4024	0,4050	0,4041	0,3892	0,3963	0,4064	0,4190
Conj. 11	0,4073	0,3776	0,4023	0,3870	0,3890	0,3845	0,3913	0,4030	0,4073
Média	0,4795	0,4226	0,4150	0,3942	0,3875	0,3827	0,3753	0,3750	0,3771

APÊNDICE B TABELAS DE RESULTADOS DOS EXPERIMENTOS SOBRE NEURÔNIOS

Tabela 9: MAPE de cada configuração de número de atrasos e neurônios na camada oculta, representados como a média entre os 11 conjuntos de sinais de entrada - Petrobras

	Número de neurônios na camada oculta					Média	Desv.Pad.
	10	20	30	50	80		
3 atrasos	0,6969	0,7087	0,6988	0,6991	0,7491	0,7105	0,0220
5 atrasos	0,6647	0,6533	0,6362	0,6487	0,6839	0,6573	0,0180
7 atrasos	0,6245	0,6118	0,5902	0,6356	0,6402	0,6205	0,0201
9 atrasos	0,5935	0,5697	0,5719	0,5748	0,5955	0,5811	0,0124
11 atrasos	0,5863	0,5583	0,5586	0,5697	0,5997	0,5745	0,0181
13 atrasos	0,5880	0,5746	0,5672	0,5692	0,5841	0,5766	0,0091
15 atrasos	0,5871	0,5685	0,5503	0,5588	0,5793	0,5688	0,0149
20 atrasos	0,5861	0,5670	0,5627	0,5806	0,5886	0,5770	0,0115
25 atrasos	0,5735	0,5544	0,5585	0,5538	0,5674	0,5615	0,0086
Média	0,6111	0,5962	0,5882	0,5989	0,6208	-	-

Tabela 10: MAPE de cada configuração de número de atrasos e neurônios na camada oculta, representados como a média entre os 11 conjuntos de sinais de entrada - Vale

	Número de neurônios na camada oculta					Média	Desv.Pad.
	10	20	30	50	80		
3 atrasos	0,4916	0,4973	0,4569	0,4422	0,5094	0,4795	0,0286
5 atrasos	0,4373	0,4259	0,4139	0,4049	0,4308	0,4226	0,0131
7 atrasos	0,4350	0,4141	0,3997	0,4061	0,4198	0,4150	0,0136
9 atrasos	0,4162	0,3954	0,3780	0,3828	0,3985	0,3942	0,0150
11 atrasos	0,4081	0,3853	0,3744	0,3837	0,3862	0,3875	0,0124
13 atrasos	0,3965	0,3825	0,3749	0,3813	0,3781	0,3827	0,0083
15 atrasos	0,3869	0,3750	0,3662	0,3716	0,3765	0,3753	0,0076
20 atrasos	0,3851	0,3738	0,3714	0,3721	0,3724	0,3750	0,0057
25 atrasos	0,3816	0,3782	0,3743	0,3723	0,3793	0,3771	0,0038
Média	0,4153	0,4030	0,3899	0,3907	0,4056	-	-

APÊNDICE C TABELAS DE RESULTADOS DOS EXPERIMENTOS SOBRE ANÁLISE TÉCNICA

Tabela 11: MAPE de cada conjunto de entrada, por configuração de indicadores da Análise Técnica utilizados como sinais auxiliares - 15 atrasos, 30 neurônios na camada oculta - Petrobras

	Sem Análise Técnica	MM	BB	IFR	MM+ BB	MM+ IFR	BB+ IFR	MM+ BB+ IFR
Conj. 1	0,5132	0,5309	0,5662	0,5728	0,5597	0,5743	0,5847	0,5948
Conj. 2	0,4893	0,5091	0,5442	0,5386	0,5477	0,5396	0,5718	0,5692
Conj. 3	0,5203	0,5281	0,5746	0,5347	0,5964	0,5450	0,5991	0,5926
Conj. 4	0,4916	0,4999	0,5380	0,5499	0,5560	0,5533	0,5844	0,5554
Conj. 5	0,5760	0,6162	0,6611	0,6187	0,6561	0,6197	0,6584	0,6892
Conj. 6	0,5627	0,5856	0,6400	0,6225	0,6383	0,6100	0,6549	0,6396
Conj. 7	0,5493	0,5603	0,6132	0,5848	0,6274	0,5892	0,6568	0,6390
Conj. 8	0,5984	0,6035	0,6639	0,6236	0,6570	0,6066	0,6721	0,6560
Conj. 9	0,5714	0,5758	0,6006	0,6153	0,6018	0,5916	0,6322	0,6240
Conj. 10	0,5946	0,6105	0,6713	0,6131	0,6840	0,6483	0,6969	0,6934
Conj. 11	0,5860	0,6252	0,6478	0,6389	0,6396	0,6453	0,6677	0,6677
Média	0,5503	0,5678	0,6110	0,5921	0,6149	0,5930	0,6345	0,6292

Tabela 12: MAPE de cada conjunto de entrada, por configuração de indicadores da Análise Técnica utilizados como sinais auxiliares - 15 atrasos, 30 neurônios na camada oculta - Vale

	Sem Análise Técnica	MM	BB	IFR	MM+ BB	MM+ IFR	BB+ IFR	MM+ BB+ IFR
Conj. 1	0,3541	0,3647	0,3755	0,3843	0,3739	0,4085	0,4013	0,4220
Conj. 2	0,3468	0,3492	0,3677	0,3848	0,3642	0,3808	0,4044	0,3833
Conj. 3	0,3599	0,3401	0,3799	0,3842	0,3582	0,3854	0,4174	0,3914
Conj. 4	0,3326	0,3416	0,3482	0,3764	0,3495	0,3873	0,3897	0,3754
Conj. 5	0,3765	0,3618	0,4006	0,3970	0,4062	0,4090	0,4255	0,4255
Conj. 6	0,3772	0,3649	0,4009	0,4027	0,4079	0,4031	0,4043	0,4255
Conj. 7	0,3622	0,3811	0,3990	0,3786	0,3951	0,3894	0,3979	0,3957
Conj. 8	0,3771	0,3750	0,4076	0,4115	0,3843	0,4077	0,4528	0,4441
Conj. 9	0,3633	0,3633	0,3950	0,3743	0,3909	0,3853	0,4163	0,4172
Conj. 10	0,3882	0,3829	0,4310	0,4149	0,4238	0,4260	0,4400	0,4470
Conj. 11	0,3904	0,3923	0,4018	0,4012	0,4098	0,4090	0,4247	0,4218
Média	0,3662	0,3652	0,3916	0,3918	0,3876	0,3992	0,4158	0,4135

APÊNDICE D TABELAS DE RESULTADOS DOS EXPERIMENTOS SOBRE DEGRADAÇÃO

Tabela 13: MAPE médio de cada período da primeira semana de teste, demonstrando a degradação no desempenho da FTLFN - 15 atrasos, 30 neurônios na camada oculta - Petrobras

Dia 1		Dia 2		Dia 3		Dia 4		Dia 5	
Per.	MAPE	Per.	MAPE	Per.	MAPE	Per.	MAPE	Per.	MAPE
1	0,6925	29	0,8525	57	1,2261	85	1,0823	113	0,8678
2	0,4852	30	0,6991	58	0,5143	86	0,6408	114	0,8618
3	0,5046	31	0,9221	59	0,7099	87	0,6570	115	0,8429
4	0,4300	32	0,6533	60	0,6467	88	0,6109	116	0,9165
5	0,4180	33	0,6979	61	0,5526	89	0,5760	117	0,7904
6	0,4253	34	0,6270	62	0,4939	90	0,5324	118	0,6825
7	0,5102	35	0,6264	63	0,5435	91	0,5639	119	0,8381
8	0,4873	36	0,4607	64	0,4984	92	0,6236	120	0,8082
9	0,5635	37	0,4289	65	0,4399	93	0,6767	121	0,7249
10	0,4254	38	0,5170	66	0,5770	94	0,6676	122	0,8373
11	0,3683	39	0,5394	67	0,4367	95	0,7112	123	0,7222
12	0,4019	40	0,4541	68	0,4138	96	0,5315	124	0,6683
13	0,3352	41	0,4022	69	0,4497	97	0,5798	125	0,7194
14	0,3814	42	0,2889	70	0,4306	98	0,5331	126	0,6063
15	0,3781	43	0,3584	71	0,4423	99	0,5363	127	0,7495
16	0,4571	44	0,4546	72	0,3250	100	0,5597	128	0,6992
17	0,2972	45	0,3849	73	0,4106	101	0,5432	129	0,7141
18	0,3675	46	0,3629	74	0,3966	102	0,4384	130	0,6522
19	0,4366	47	0,4279	75	0,4442	103	0,4945	131	0,5671
20	0,5074	48	0,4481	76	0,4824	104	0,5374	132	0,6656
21	0,4791	49	0,3705	77	0,6119	105	0,5724	133	0,6585
22	0,3218	50	0,3722	78	0,6445	106	0,5756	134	0,5379
23	0,3229	51	0,4773	79	0,5163	107	0,6542	135	0,5388
24	0,3222	52	0,3697	80	0,4749	108	0,7085	136	0,5923
25	0,3143	53	0,3680	81	0,4659	109	0,6731	137	0,5370
26	0,3162	54	0,4094	82	0,5479	110	0,6248	138	0,6176
27	0,3673	55	0,4588	83	0,5522	111	0,5628	139	0,6879
28	0,8141	56	1,3717	84	0,9172	112	0,8301	140	0,9405
Média	0,4332		0,5287		0,5416		0,6178		0,7159

Tabela 14: MAPE médio de cada período da primeira semana de teste, demonstrando a degradação no desempenho da FTLFN - 15 atrasos, 30 neurônios na camada oculta - Vale

Dia 1		Dia 2		Dia 3		Dia 4		Dia 5	
Per.	MAPE	Per.	MAPE	Per.	MAPE	Per.	MAPE	Per.	MAPE
1	0,4078	29	0,7690	57	0,7617	85	0,8710	113	0,8415
2	0,4123	30	0,3870	58	0,5396	86	0,5134	114	0,4417
3	0,4372	31	0,5109	59	0,5090	87	0,5208	115	0,4883
4	0,2971	32	0,4359	60	0,5173	88	0,4364	116	0,3762
5	0,1980	33	0,3668	61	0,3804	89	0,3376	117	0,3465
6	0,2746	34	0,3550	62	0,3736	90	0,3623	118	0,4300
7	0,2466	35	0,4182	63	0,4122	91	0,4440	119	0,5006
8	0,2574	36	0,3663	64	0,3893	92	0,4375	120	0,3219
9	0,2680	37	0,3075	65	0,3656	93	0,4122	121	0,3347
10	0,3147	38	0,3088	66	0,4485	94	0,4242	122	0,4346
11	0,3065	39	0,3741	67	0,3362	95	0,3960	123	0,4210
12	0,2261	40	0,3470	68	0,3278	96	0,3667	124	0,4256
13	0,2863	41	0,3219	69	0,3727	97	0,3228	125	0,3806
14	0,2916	42	0,2859	70	0,3220	98	0,3581	126	0,3724
15	0,2162	43	0,3283	71	0,3102	99	0,3461	127	0,3936
16	0,2508	44	0,2525	72	0,2918	100	0,3015	128	0,3263
17	0,2218	45	0,2325	73	0,2965	101	0,3336	129	0,3191
18	0,2886	46	0,3244	74	0,2677	102	0,3460	130	0,3410
19	0,2908	47	0,2859	75	0,3057	103	0,4108	131	0,3679
20	0,3190	48	0,3462	76	0,3166	104	0,3873	132	0,4062
21	0,2481	49	0,3088	77	0,3010	105	0,3717	133	0,3879
22	0,2485	50	0,2992	78	0,3529	106	0,3917	134	0,3892
23	0,3517	51	0,2896	79	0,3120	107	0,3404	135	0,3292
24	0,3245	52	0,2994	80	0,2904	108	0,3348	136	0,3740
25	0,2394	53	0,2528	81	0,3390	109	0,3371	137	0,3705
26	0,2104	54	0,3027	82	0,3493	110	0,3550	138	0,3421
27	0,2410	55	0,3405	83	0,3144	111	0,3334	139	0,3315
28	0,7661	56	0,7029	84	0,7591	112	0,8519	140	0,6756
Média	0,3015		0,3614		0,3879		0,4159		0,4096

Tabela 15: MAPE médio de cada conjunto de entrada, considerando uma, duas e três semanas de teste, sem re-treino da FTLFN - 15 atrasos, 30 neurônios na camada oculta - Petrobras

Conj.	Semanas de teste		
	1	2	3
1	0,5132	0,6378	0,7097
2	0,4893	0,5543	0,5937
3	0,5203	0,5830	0,6710
4	0,4916	0,5749	0,5495
5	0,5760	0,6419	0,7285
6	0,5627	0,6519	0,6907
7	0,5493	0,5851	0,6337
8	0,5984	0,6940	0,7509
9	0,5714	0,6237	0,7213
10	0,5946	0,6582	0,7534
11	0,5860	0,6483	0,7149
Média	0,5503	0,6230	0,6834

Tabela 16: MAPE médio de cada conjunto de entrada, considerando uma, duas e três semanas de teste, sem re-treino da FTLFN - 15 atrasos, 30 neurônios na camada oculta - Vale

Conj.	Semanas de teste		
	1	2	3
1	0,3541	0,3975	0,4131
2	0,3468	0,3930	0,3446
3	0,3599	0,3702	0,3518
4	0,3326	0,3748	0,3582
5	0,3765	0,4242	0,4844
6	0,3772	0,4193	0,4636
7	0,3622	0,4161	0,4227
8	0,3771	0,4841	0,5509
9	0,3633	0,4172	0,4684
10	0,3882	0,4522	0,4721
11	0,3904	0,4555	0,4788
Média	0,3662	0,4186	0,4371