

**UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS**

Vitor Nazareth de Souza

**REDES NEURAIAS ARTIFICIAIS APLICADAS A
PREVISÃO DE CURTO PRAZO DE SÉRIES
TEMPORAIS FINANCEIRAS**

São Carlos

2018

Vitor Nazareth de Souza

**REDES NEURAIS ARTIFICIAIS APLICADAS A
PREVISÃO DE CURTO PRAZO DE SÉRIES
TEMPORAIS FINANCEIRAS**

Monografia apresentada ao Curso de Curso de Engenharia Elétrica com Ênfase em Sistemas de Energia e Automação, da Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Engenheiro Eletricista.

Orientador: Prof. Dr. Takaaki Ohishi

São Carlos

2018

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da
EESC/USP com os dados inseridos pelo(a) autor(a).

N845r Nazareth de Souza, Vitor
Redes Neurais Artificiais Aplicadas a Previsão de
Curto Prazo de Séries Temporais Financeiras / Vitor
Nazareth de Souza; orientador Takaaki Ohishi. São
Carlos, 2018.

Monografia (Graduação em Engenharia Elétrica com
ênfase em Sistemas de Energia e Automação) -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2018.

1. Redes Neurais Artificiais. 2. Wavelets. 3.
Predição de Séries Temporais Financeiras. 4. Gestão de
Risco. I. Título.

FOLHA DE APROVAÇÃO

Nome: Vitor Nazareth de Souza

Título: “Redes neurais artificiais aplicadas a previsão de curto prazo de séries temporais financeiras”

Trabalho de Conclusão de Curso defendido e aprovado

em 30 / 11 / 2018,

com NOTA 10,0 (Dez , zero), pela Comissão Julgadora:

Prof. Dr. Takaaki Ohishi - Orientador - UNICAMP

Prof. Associado Rogério Andrade Flauzino - SEL/EESC/USP

Prof. Dr. Danilo Hernane Spatti - SSC/ICMC/USP

Coordenador da CoC-Engenharia Elétrica - EESC/USP:

Prof. Associado Rogério Andrade Flauzino

*Este trabalho é dedicado aos meus pais e irmã,
por todo apoio, carinho e por nunca deixarem de acreditar no meu potencial.
Hoje sou apenas reflexo disso tudo. Muito obrigado!*

Agradecimentos

Agradeço primeiramente à minha família: Benedito, Roseli e Nathália. Muito obrigado por toda motivação, suporte, paciência e por estarem presentes em todas as minhas realizações.

À meu orientador, Prof. Dr. Takaaki Ohishi, pelas instruções, conhecimento e dicas para realização deste trabalho.

Aos amigos Rafael Taranto Polizel, Rodolfo Brigato Ferreira e Weber Daas Santos pelos anos de convívio, estudos e alegrias.

Aos amigos Fabrício Tietz, Gabriel Ramadan, Lucas Gianotti, João Victor Barbosa, Rodolfo Takeshi e Gustavo Silva pelas conversas, estudos e companheirismo.

À Universidade de São Paulo, por toda estrutura, recursos e suporte para o desenvolvimento do engenheiro.

*“The true sign of intelligence is
not knowledge but imagination.”*

Albert Einstein

Resumo

SOUZA, V. N. **REDES NEURAIS ARTIFICIAIS APLICADAS A PREVISÃO DE CURTO PRAZO DE SÉRIES TEMPORAIS FINANCEIRAS**. 2018. 78p. Monografia (Trabalho de Conclusão de Curso) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2018.

O presente trabalho trata de uma abordagem para predição de séries temporais ruidosas, com modelo desconhecido e que apresentam grandes não-linearidades. O modelo em questão combina uma rede neural de múltiplas camadas com o filtro de *wavelet*. A justificativa para a junção dessas abordagens reside no fato possuírem características complementares para o tratamento de séries temporais financeiras. As redes neurais conseguem aproximar qualquer função de modelo gerador da série, sendo capazes de tratar grandes não-linearidades, aprendendo a partir de amostras dos próprios dados. Os filtros de volatilidade, como o de *wavelet*, são responsáveis por atenuar os efeitos do ruído, possibilitando que a rede abstraia o máximo de informação relevante das séries financeiras. A previsão de séries temporais é de grande importância, uma vez que podem servir como base para a tomada de decisão e planejamento estratégico. Quando se trata de séries temporais financeiras, sua previsão possibilita uma melhora na gestão de riscos em posições de investimento no mercado financeiro.

Palavras-chave: Redes Neurais Artificiais; *Wavelets*; Predição de Séries Temporais Financeiras; Gestão de Risco.

Abstract

SOUZA, V. N. . 2018. 78p. Monografia (Trabalho de Conclusão de Curso) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2018.

The present work deals with an approach to predict noisy time series with an unknown model and with large nonlinearities. The model in question combines a multilayer neural network with the *wavelet* filter. The justification for joining these approaches lies in the fact that they have complementary characteristics for the treatment of financial time series. Neural networks can approximate any function of the generator model of the series, being able to treat large non-linearities, learning from samples of the data itself. Volatility filters, such as *textit* wavelet, are responsible for attenuating the effects of noise, allowing the network to abstract the maximum information relevant to the financial series. The forecast of time series is of great importance, since it can serve as a basis for strategic and planning decisions. When it comes to financial time series, its forecast allows for an improvement in the management of risks in financial market investment positions.

Keywords: Artificial neural networks; *Wavelets*; Prediction of Time Series Financial; Risk management.

Lista de figuras

| | |
|---|----|
| Figura 1 – Fluxograma das Etapas de Pesquisa | 24 |
| Figura 2 – Modelo Matemático do Neurônio | 32 |
| Figura 3 – Rede Neural de Camada Única | 33 |
| Figura 4 – Rede Neural de Multicamada | 33 |
| Figura 5 – Modelo de um Sistema Dinâmico | 37 |
| Figura 6 – Árvore de decomposição em dois níveis de uma série temporal via decomposição por <i>wavelets</i> | 39 |
| Figura 7 – Validação Cruzada | 47 |
| Figura 8 – Fluxograma para a previsão com filtro de <i>wavelets</i> e redes neurais recorrentes | 49 |
| Figura 9 – Formas de Onda do Log Retorno da ação ELET3 - com e sem filtro . . | 52 |
| Figura 10 – Previsão da Série ELET3 | 54 |
| Figura 11 – Formas de Onda do Log Retorno da ação CPFE3 - com e sem filtro . . | 56 |
| Figura 12 – Previsão da Série CPFE3 | 57 |
| Figura 13 – Formas de Onda do Log Retorno da ação CMIG4 - com e sem filtro . . | 59 |
| Figura 14 – Previsão da Série CMIG4 | 60 |
| Figura 15 – Formas de Onda do Log Retorno da ação CPLE6 - com e sem filtro . . | 62 |
| Figura 16 – Previsão da Série CPLE6 | 63 |
| Figura 17 – Histograma do Erro de Treinamento com filtro de <i>wavelet</i> - ELET3. . . | 71 |
| Figura 18 – Histograma do Erro de Treinamento com filtro de <i>wavelet</i> - CPFE3. . . | 73 |
| Figura 19 – Histograma do Erro de Treinamento com filtro de <i>wavelet</i> - CMIG4. . . | 75 |
| Figura 20 – Histograma do Erro de Treinamento com filtro de <i>wavelet</i> - CPLE6. . . | 77 |

Lista de tabelas

| | |
|--|----|
| Tabela 1 – Topologias das RNAs | 46 |
| Tabela 2 – Melhores Topologias das RNAs para Previsão sem Filtro da Série ELET3 | 51 |
| Tabela 3 – Previsão sem Filtro da Série ELET3 | 52 |
| Tabela 4 – Melhores Topologias das RNAs para Previsão da Série ELET3 | 53 |
| Tabela 5 – Previsão da Série ELET3 | 53 |
| Tabela 6 – Melhores Topologias das RNAs para Previsão sem Filtro da Série CPFE3 | 55 |
| Tabela 7 – Previsão sem Filtro da Série CPFE3 | 55 |
| Tabela 8 – Melhores Topologias das RNAs para Previsão da Série CPFE3 | 56 |
| Tabela 9 – Previsão da Série CPFE3 | 57 |
| Tabela 10 – Melhores Topologias das RNAs para Previsão sem Filtro da Série CMIG4 | 58 |
| Tabela 11 – Previsão sem Filtro da Série CMIG4 | 58 |
| Tabela 12 – Melhores Topologias das RNAs para Previsão da Série CMIG4 | 59 |
| Tabela 13 – Previsão da Série CMIG4 | 60 |
| Tabela 14 – Melhores Topologias das RNAs para Previsão sem Filtro da Série CPLE6 | 61 |
| Tabela 15 – Previsão sem Filtro da Série CPLE6 | 61 |
| Tabela 16 – Melhores Topologias das RNAs para Previsão da Série CPLE6 | 62 |
| Tabela 17 – Previsão da Série CPLE6 | 63 |

Sumário

| | | |
|------------|---|-----------|
| 1 | INTRODUÇÃO | 21 |
| 1.1 | Justificativa e Motivação | 22 |
| 1.2 | Objetivos | 22 |
| 1.3 | Estrutura do Trabalho | 23 |
| 2 | SÉRIES TEMPORAIS | 25 |
| 2.1 | Introdução à Séries Temporais | 25 |
| 2.1.1 | Séries Temporais Financeiras | 25 |
| 2.1.2 | Análise das Séries Temporais Financeiras | 26 |
| 2.2 | Modelos para Predição de Séries Temporais | 27 |
| 2.2.1 | Modelos Lineares para Predição de Séries Temporais | 27 |
| 2.2.2 | Modelos para Predição de Séries Temporais Financeiras | 28 |
| 2.2.2.1 | Predição de Séries Temporais Financeiras com Redes Neurais | 28 |
| 3 | REDES NEURAS ARTIFICIAIS | 31 |
| 3.1 | Conceitos | 31 |
| 3.1.1 | Introdução às Redes Neurais | 31 |
| 3.1.2 | Arquiteturas das Redes Neurais | 32 |
| 3.1.3 | Importância das Redes Neurais Artificiais | 34 |
| 3.1.4 | Aprendizado | 35 |
| 3.1.4.1 | Aprendizado Supervisionado | 35 |
| 3.1.4.2 | Aprendizado Não Supervisionado | 36 |
| 4 | FILTROS | 37 |
| 4.1 | Análise Comparativa entre as funções de Fourier e de <i>Wavelets</i> | 37 |
| 4.2 | Filtro de <i>Wavelets</i> | 38 |
| 4.2.1 | Modelagem do Filtro de <i>Wavelets</i> | 40 |
| 5 | METODOLOGIA | 43 |
| 5.1 | Determinação das Variáveis de Entrada | 43 |
| 5.2 | Tratamento das Variáveis do Modelo e Normalização dos Dados | 44 |
| 5.3 | Redes Neurais Recorrentes | 45 |
| 5.4 | Algoritmo de Treinamento | 46 |
| 5.4.1 | Validação Cruzada | 46 |
| 5.4.2 | Estatísticas de Erro | 47 |
| 5.4.3 | Decomposição do Esquema Operacional | 48 |

| | | |
|------------|--|-----------|
| 6 | RESULTADOS | 51 |
| 6.1 | Análise da Série ELET3 | 51 |
| 6.1.1 | Previsão com redes neurais recorrentes | 51 |
| 6.1.2 | Previsão com filtro de <i>waveles</i> e redes neurais recorrentes | 52 |
| 6.2 | Análise da Série CPFE3 | 54 |
| 6.2.1 | Previsão com redes neurais recorrentes sem filtro de wavelets | 55 |
| 6.2.2 | Previsão com filtro de <i>waveles</i> e redes neurais recorrentes | 55 |
| 6.3 | Análise da Série CMIG4 | 58 |
| 6.3.1 | Previsão com redes neurais recorrentes sem filtro de wavelets | 58 |
| 6.3.2 | Previsão com filtro de <i>waveles</i> e redes neurais recorrentes | 58 |
| 6.4 | Análise da Série CPLE6 | 60 |
| 6.4.1 | Previsão com redes neurais recorrentes sem filtro de wavelets | 60 |
| 6.4.2 | Previsão com filtro de <i>waveles</i> e redes neurais recorrentes | 61 |
| 6.5 | Considerações Parciais | 63 |
| 7 | CONCLUSÃO | 65 |
| | REFERÊNCIAS | 67 |
| | APÊNDICES | 69 |
| | APÊNDICE A – HISTOGRAMA DOS ERROS DE TREINAMENTO DAS REDES NEURAI NA PRESENÇA DO FILTRO DE <i>WAVELET</i> - ELET3 | 71 |
| | APÊNDICE B – HISTOGRAMA DOS ERROS DE TREINAMENTO DAS REDES NEURAI NA PRESENÇA DO FILTRO DE <i>WAVELET</i> - CPFE3 | 73 |
| | APÊNDICE C – HISTOGRAMA DOS ERROS DE TREINAMENTO DAS REDES NEURAI NA PRESENÇA DO FILTRO DE <i>WAVELET</i> - CMIG4 | 75 |
| | APÊNDICE D – HISTOGRAMA DOS ERROS DE TREINAMENTO DAS REDES NEURAI NA PRESENÇA DO FILTRO DE <i>WAVELET</i> - CPLE6 | 77 |

1 Introdução

A alta complexidade dos fenômenos do mundo real reflete-se por comportamentos caóticos não lineares (SHUDHEER, 2015), tornando sua modelagem excessivamente árdua. A maior parte da literatura relacionada a técnicas de modelagem clássicas, como os modelos estatísticos ARIMA (Auto-regressivo Integrado de Médias Móveis) e o GARCH (Auto-regressivo com Heterocedasticidade Condicional Generalizada) têm produzido previsões insatisfatórias (OLIVEIRA, 2007; ARAUJO, 2016), sendo que suas desvantagens podem ser justificadas pelo elevado grau de abstração em relação a estrutura do modelo real (SCHOPF, 2007). Em outras palavras, os modelos estatísticos tradicionais não conseguem descrever as equações que governam o comportamento das variáveis de entrada de um modelo real que se quer prever.

Nesse contexto, as Redes Neurais Artificiais (RNA) surgem como uma alternativa atrativa para a tarefa de predição de séries temporais, uma vez que possuem grande poder computacional devido à sua estrutura maciçamente paralela e distribuída e de sua capacidade de aprender para generalizar (HAYKIN, 2001). As redes neurais têm sido utilizadas com sucesso em uma vasta gama de aplicações de sistemas de modelagem financeira, mostrando superioridade sobre os modelos estatísticos ARIMA e GARCH em muitos casos (MEDEIROS; RECH, 2006; SOMERS; SCARBOROUGH, 2006; OLIVEIRA, 2007). Segundo os autores (MORETTIN; TOLOI, 2004; LIMA, 2011; OLIVEIRA, 2007), as condições atuais de um sistema determinam em algum grau o futuro de tal forma que podem haver muitas iterações e complexas relações entre as variáveis envolvidas, como é o caso dos mercados financeiros.

O uso de redes neurais perante sistemas complexos, é motivado pelo fato de sua estrutura ser não linear e pela sua capacidade de estender o número de argumentos de entrada. A não linearidade de uma rede neural é de um tipo especial, pois está presente em cada neurônio. Ainda, outra vantagem do uso das redes neurais consiste na utilização do aprendizado supervisionado, o qual permite o ajuste das conexões sinápticas a partir de um conjunto de dados. Esse ajuste é feito de maneira gradativa, semelhante à inferência estatística não-paramétrica (SCHOPF, 2007), não realizando suposições sobre o processo gerador de dados (BREIMAN, 2003; ARAUJO, 2016).

Segundo (HAYKIN, 2001), mais uma conjuntura corrobora o uso de redes neurais: a adaptabilidade. Para processos não estacionários, as redes neurais podem ser treinadas visando a sua adaptação defrente as modificações que ocorrem no ambiente, abstraindo (aproximando) as equações geradoras de uma série temporal. Ainda, (WAN; R.MERWE, 2001) defendem que uma rede neural com um número arbitrário de neurônios e com duas ou mais camadas pode aproximar qualquer função uniformemente contínua.

As redes neurais costumam apresentar bom desempenho para a predição de séries temporais não lineares e desconhecidas. Todavia, a predição de séries temporais ruidosas dificulta a performance das redes. Nesse cenário, as redes neurais apresentam dificuldades em identificar o que é a série original e o que é ruído. O uso da decomposição de séries temporais via wavelets, visando a análise de séries temporais, proporciona uma alternativa para redução de ruídos. wavelets são funções que consistem em fracionar a série temporal original em duas subséries, uma relativa às altas frequências e outra às baixas frequências, com o intuito de amenizar os efeitos do ruído nas previsões (GENÇAY; SELÇUK; WHITCHER, 2002).

1.1 Justificativa e Motivação

A fundamentação teórica que norteia esse trabalho a fim de se realizar a predição de séries financeiras é embasada nos estudos de (MANDELROT; HUDSON, 2004), onde os autores afirmam que é possível identificar padrões de comportamento em uma série de dados e conseqüentemente realizar predições. Ao comprovar esse fato, os autores mostram que várias séries financeiras possuem a propriedade de memória. Dessa forma, a partir de um conjunto de observações passadas do comportamento de uma série temporal, podem-se fazer previsões sobre o seu comportamento futuro e verificar o quão precisa essas previsões podem ser (LIMA, 2011).

A motivação do trabalho é estruturada na necessidade do gerenciamento de risco nas estratégias de investimentos, onde faz-se necessária projeções financeiras relativas a volatilidade de mercado, principalmente as de curto prazo, para tomadas de decisão e planejamento de suas posições no mercado a vista (por exemplo, ações) e futuros.

Em suma, pode-se dizer que em toda atividade financeira, principalmente na gestão de riscos, há a necessidade de se prever a volatilidade dos ativos, com o propósito de se verificar a possibilidade de perda futura (JORION, 2007).

1.2 Objetivos

O objetivo deste trabalho de conclusão de curso é a obtenção de um modelo para predição de séries temporais financeiras através do uso combinado do filtro de wavelet e de redes neurais artificiais recorrentes multicamadas, onde o desempenho do modelo não deverá ser somente mensurado pela diferença (erro) entre o valor real do ativo financeiro e o valor previsto pela rede, mas também pela capacidade de prever a tendência dos movimentos diários da série temporal, isto é, apresentar a capacidade de estimar os instantes de inversão de movimento da série.

Outro objetivo mais específico do trabalho é o de verificar se o uso do filtro de wavelet para a redução do ruído da série temporal original resulta em uma melhoria na

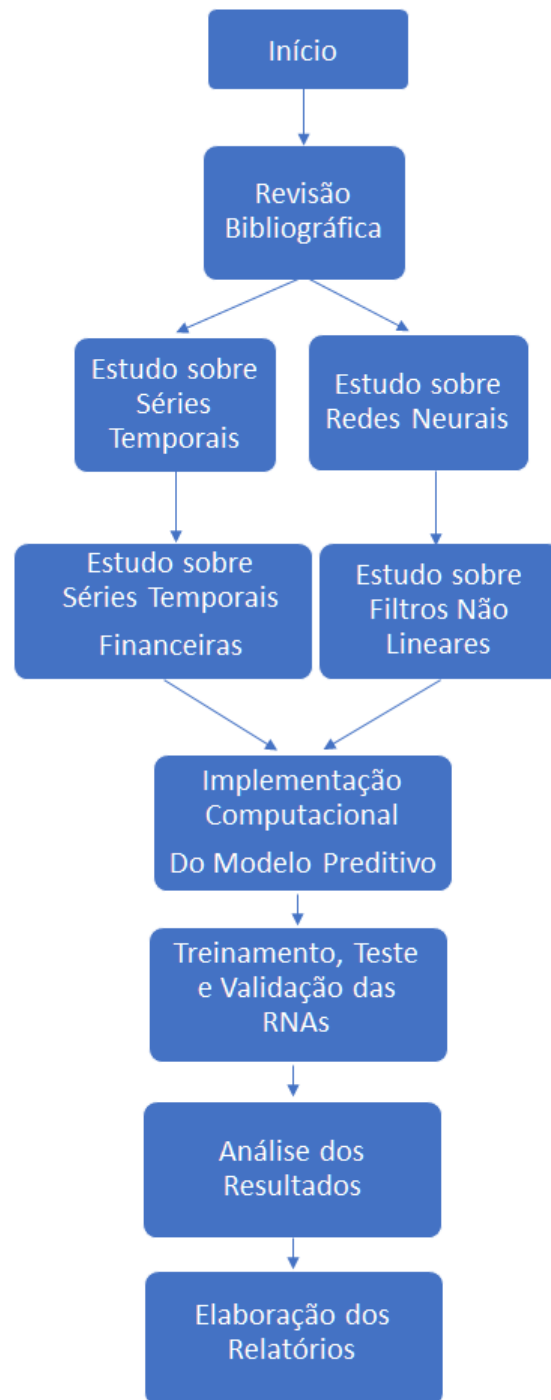
qualidade preditiva quando comparado a modelos de redes neurais

1.3 Estrutura do Trabalho

A revisão bibliográfica, a respeito de séries temporais e de redes neurais artificiais, são abordados nos Capítulo 2 e Capítulo 3. Neles, procurou-se abstrair conhecimentos teóricos e metodologias propostas por diversos autores, a fim de que este trabalho de conclusão de curso tivesse alguma sustentação acadêmica para constatar sua viabilidade. Neste instante, avaliou-se a necessidade da "limpar" as séries temporais financeiras, com a intenção de melhorar a qualidade da previsão das redes. A base literária presente no Capítulo 4 propõe o uso de uma ferramenta que seja capaz de reduzir ruídos de uma série financeira, mas com preservação das suas características gerais de tendência, ciclos e sazonalidades. No Capítulo 5, são apresentadas as metodologias utilizadas para a implementação do estudo, bem como foram feitos os tratamentos dos dados. Os resultados obtidos são exibidos no Capítulo 6, e em seguida, no Capítulo 7, são desenvolvidas as conclusões, as quais são feitas após análises descritivas dos resultados.

O fluxograma descritivo das etapas empregadas nesse trabalho pode ser representado pela Figura 1:

Figura 1: Fluxograma das Etapas de Pesquisa



Fonte: Autoria própria.

2 Séries Temporais

Essa seção será destinada a apresentar os conceitos e os aspectos teóricos e práticos que envolvem a análise e previsão de séries temporais com o uso de redes neurais.

2.1 Introdução à Séries Temporais

De uma forma geral, uma Série Temporal (ST) é definida como qualquer conjunto de dados estocásticos ordenados obtidos de intervalos regulares no tempo por um período específico de observações (ST_1), que apresentam dependência serial (dependência entre instantes de tempo). Enquadra-se nessa categoria uma vasta quantia de fenômenos de natureza física, biológica, econômica, entre outros; e pode-se representá-lo matematicamente por: $y_t = y_1, y_2, \dots, y_n$.

A curva representativa do movimento obtido pelo processo estocástico é chamada de trajetória. (MORETTIN; TOLOI, 2004) definem ST como parte de uma trajetória dentre muitas outras que poderiam ter sido observadas de um processo estocástico (este, por sua vez, é definido como processos cuja evolução no tempo é regida por leis probabilísticas).

Em sequência, será apresentada uma descrição sobre as séries temporais oriundas de dados do mercado financeiro, uma vez que a cerne deste trabalho são as séries temporais financeiras (não serão tratadas aqui as demais séries).

2.1.1 Séries Temporais Financeiras

Esta seção é destinada a apresentar alguns conceitos, preceitos, bem como concepções teóricas e práticas que circundam as séries temporais financeiras, embasados nos trabalhos realizados por (ENDERS, 2004), (MORETTIN; TOLOI, 2004) e (LIMA, 2011).

Segundo (ENDERS, 2004), as séries temporais financeiras não são serialmente correlacionadas, mas sim dependentes, dado que as mesmas diferem das demais séries temporais por apresentarem peculiaridades próprias de seus elementos. Ainda, o autor diz que as séries financeiras apresentam em seus log retornos a presença de clusters de volatilidade, os quais podem ser definidos de diferentes formas, mas não são notadamente observáveis. A definição desses grupos de volatilidade auxilia no processo de modelagem dos fenômenos que geram as séries temporais, visando uma melhor previsão.

Segundo (LIMA, 2011), a variância de um log retorno para um dado período de tempo depende dos log retornos passados, além de outros elementos até então desconhecidos, de modo que sua variância condicional não coincide com a variância total da série. A variância de um log retorno é de fundamental importância em finanças, pois têm em sua essência a medida do risco relacionada às variações da série temporal ao longo do tempo

(sua aplicação é demasiada comum no âmbito das teorias de finanças, como em modelos de gestão de riscos, precificação de ativos, marcação a mercado, entre outros).

(MORETTIN; TOLOI, 2004) justifica o uso dos log retornos em modelos financeiros pelo fato de apresentarem propriedades estatísticas como estacionalidade e ergodicidade. Essa última, permite medir e prever valores futuros de uma série temporal através da estimação estatística da interpretação dos dados passados.

2.1.2 Análise das Séries Temporais Financeiras

O estudo de séries financeiras é estruturada por meio dos cálculos dos retornos, obtidos a partir da série de preços dos ativos financeiros. (TSAY, 2005) considera os preços dos ativos financeiros como sendo variáveis aleatórias, uma vez que podem admitir função de distribuição própria, dada à possibilidade teórica dos preços poderem alcançar valores infinitos e não poderem ser negativos.

Sendo o preço de um ativo em um certo instante t dado por P_t , e em um instante anterior $t - 1$ por P_{t-1} , pode-se definir o retorno, discreto ou contínuo (log retorno), de um ativo no período $t - 1$ a t a partir das seguintes equações:

Retorno Discreto

$$R(t) = \frac{P_t}{P_{t-1}} - 1 \quad (2.1)$$

Retorno Contínuo

$$R(t) = \ln \left(\frac{P_t}{P_{t-1}} \right) \quad (2.2)$$

A interpretação dada pelo desvio padrão do retorno é o que possibilita a diferenciação para o uso do retorno discreto ou do contínuo. O desvio padrão do retorno discreto é conhecido como risco e o do retorno contínuo como volatilidade.

Frente às incertezas no mercado financeiro, tanto nacional quanto internacional, causadas por cenários políticos, crises econômicas, oscilações de moedas, entre outros (tais incertezas de mercado, aliadas à fatores especulativos, podem gerar grandes oscilações na série temporal dos log retornos, culminando nos clusters de volatilidade que podem ser verificados ao observar a evolução do comportamento dos dados), (SABINO E BRESSAN, 2009) defendem que a modelagem da variabilidade dos ativos financeiros proporciona a previsão dos valores futuros dos ativos financeiros, gerando uma métrica útil de risco que poderá ser assumida pelos investidores no mercado financeiro.

A partir do Acordo de Basiléia II, modelos voltados a previsão de volatilidade passaram a integrar o cotidiano de instituições financeiras (em um cenário global), com enfoque no gerenciamento de riscos financeiros. Ainda, (ANDERSEN; BOLLERSLEV;

CHRISTOFFERSEN, 2006) propõe que a volatilidade tem ganhado espaço em pesquisas que envolvem econometria de séries temporais e previsão.

2.2 Modelos para Predição de Séries Temporais

Esta seção discorrerá sobre os principais modelos adotados na literatura relacionados à previsão de séries temporais, dando enfoque nas técnicas adotadas para o desenvolvimento deste trabalho. A previsão de séries temporais é de grande importância, uma vez que podem servir como base para a tomada de decisão e planejamento gerencial e/ou estratégico.

Os modelos de previsão de séries temporais podem ser classificados em dois grandes grupos: lineares e não-lineares. Em geral, os modelos lineares são tratados pela literatura de séries temporais clássicas, as quais supõe que as séries são estacionárias (se a série permanece ao redor de uma média constante, refletindo um equilíbrio estável) (YAN, 2012) ou que podem ser transformadas em séries estacionárias através de alguma transformação simples, como a diferenciação (CASTRO, 2016).

Entretanto, séries reais não exibem características de um processo linear, onde a alta complexidade dos fenômenos do mundo real reflete-se por comportamentos caóticos não lineares (CASTRO, 2016).

2.2.1 Modelos Lineares para Predição de Séries Temporais

Aqui, serão apresentados, brevemente, os principais métodos para previsão linear de séries temporais. Em geral, a maior parte dos modelos apresentados são constituídos de filtros com combinações de médias simples e ponderadas.

- **Média Móvel Simples:** Essa técnica consiste em calcular a média aritmética das k observações mais recentes da série, onde a estimativa não leva em consideração as observações mais antigas, fazendo a tendência variar suavemente com o tempo e esquecendo o passado mais distante. O termo média móvel dá-se pela substituição da observação mais antiga pela mais recente, a cada passo de tempo;
- **Alisamento Exponencial Simples:** Esse modelo representa uma média ponderada que atribui pesos maiores às observações mais recentes da série;
- **Alisamento Exponencial Linear de Brown:** Consiste em calcular um segundo valor exponencialmente alisado e diferentemente dos modelos de Médias Móveis Simples e Alisamento Exponencial Simples, aplicam-se também à modelos que não são localmente constantes;
- **Alisamento Exponencial Quadrático de Brown:** Semelhante ao modelo anterior, mas com a tendência na forma quadrática, ou seja, necessita-se de um terceiro alisamento;

- **Modelos de Auto-Regressão:** Supõem que os valores da série sejam linearmente relacionados com seus próprios valores defasados com a média do ruído igual a zero;
- **Modelos ARIMA:** Constituem o principal exemplo de métodos lineares. O modelo ARIMA (*Autoregressive Integrates Moving Averages*) é uma combinação de três componentes, interpretados como filtros: o componente auto-regressivo, o componente de médias móveis e o filtro de integração. Podem ser configurados para realizar a previsão de até três casos de modelos de séries temporais: processos lineares estacionários, processos lineares não-estacionários e processos de memória longa.

2.2.2 Modelos para Predição de Séries Temporais Financeiras

Essa seção irá dar enfoque apenas aos modelos de redes neurais, uma vez que o presente trabalho tem como objetivo desenvolver um modelo de rede neural para previsão do retorno logarítmico do preço de ações, mas também fará uma breve citação sobre os principais modelos utilizados na literatura à respeito das previsões de séries temporais financeiras.

O modelos de predição para séries temporais já retificados pela literatura mostram que para séries financeiras os melhores são os não-lineares. Para (TSAY, 2005), modelos TAR (Limiar Auto-Regressivo) que incluem não linearidades na esperança, são paradigmas de volatilidade estocásticas que modelam a variância através de um processo não observado. Segundo (ENDERS, 2004), os modelos da família ARCH e GARCH são os mais comuns a serem aplicados na modelagem da variância condicionada. Ainda, existem os modelos de Redes Neurais Artificiais, que segundo (HAYKIN, 2001), são modelos de processamentos paralelos distribuídos, formados por unidades de ajustes simples, que têm capacidade natural de armazenar conhecimento experimental e torná-lo disponível para o uso.

2.2.2.1 Predição de Séries Temporais Financeiras com Redes Neurais

Nas últimas décadas, algoritmos de aprendizado de máquinas (machine learning) têm sido largamente aplicados como alternativa aos métodos estatísticos clássicos de previsão (KANTZ; SCHREIBER, 2004). Em suma, algoritmos de machine learning assumem que o mecanismo de geração dos dados é complexo e desconhecido, utilizando os dados de entradas para prever a saída, realizando poucas suposições sobre o processo gerador dos dados (CASTRO, 2016).

As redes neurais artificiais são ferramentas extremamente adaptáveis em um ambiente dinâmico, detendo a capacidade de aprender rapidamente padrões complexos e tendências nos dados, moldando-se às mudanças. A ideia do uso das redes neurais está fundamentada em como projetar um modelo e de tratá-lo como um sistema dinâmico, onde, a partir desse modelo, ambiciona-se estimar um valor futuro com base no conhecimento de

um conjunto informacional de valores passados, o qual contém informação suficiente para se prever a evolução futura do sistema.

Segundo (HAYKIN, 2001), pode-se citar algumas características das redes neurais que justificam e fundamentam seu uso em problemas que envolvem a predição de séries temporais:

- Possibilidade de considerar o perfil não-linear do processo gerador do sistema;
- Sem necessidade de deter profundo conhecimento estatístico sobre o contexto no qual a rede está inserida;
- Capacidade de aproximar qualquer função de natureza contínua;
- Capacidade de aprendizado e de generalização, apresentando bom desempenho na resposta a dados que não faziam parte do conjunto de treinamento.

Em suma, as redes neurais sobressaem-se aos demais modelos por aprenderem padrões subjacentes nos dados, culminando em um melhor desempenho do que os métodos estatísticos tradicionais quando o processo gerador dos dados é desconhecido, não-linear ou não estacionário (CASTRO, 2001).

O cálculo do próximo instante da série temporal é dado por:

$$y(t) = \sum_{n=1}^T a(n)y(t-n) + e(t) = \hat{y}(t) + e(t) \quad (2.3)$$

Sendo:

- $y(t)$ o valor a ser calculado;
- T o número de termos anteriores a t ;
- $y(t-n)$ são os T valores anteriores da série;
- $a(n)$ é o peso de cada observação passada;
- $e(t)$ é o erro de cada previsão.

3 Redes Neurais Artificiais

A tecnologia das Redes Neurais Artificiais (RNA's) constituem modelos computacionais que ambicionam solucionar problemas de reconhecimento de padrões que geralmente são baseados em um conjunto de informações previamente conhecido. Nesse capítulo, será abordado os conceitos e estruturas das RNA's, os processos de aprendizagem das redes e as aplicações em diferentes áreas.

3.1 Conceitos

Tal seção será destinada a discorrer sobre os conceitos fundamentais das RNA's, visando uma melhor compreensão das bases e do funcionamento das mesmas.

3.1.1 Introdução às Redes Neurais

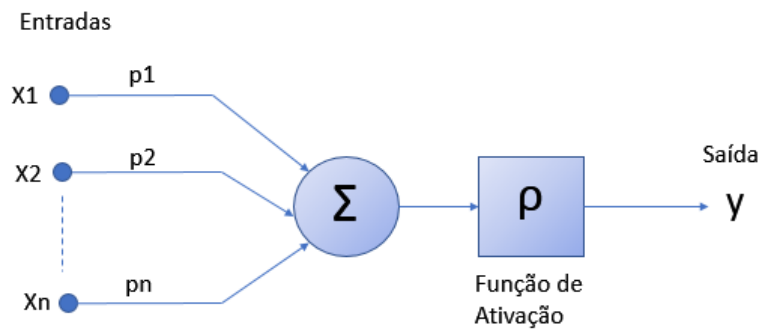
Como dito acima, as RNA's visam resolver problemas de reconhecimento de padrões que geralmente são baseados em um conjunto de informações previamente conhecido. Normalmente tais conjuntos de dados são divididos em grupo de treinamento e um grupo de teste. Os estudos de redes neurais tem sido instigado pelas diferenças apresentadas entre o funcionamento do cérebro humano e o de um computador tradicional, ou seja, são estimulados pela busca da compreensão das capacidades da natureza humana, as quais possibilitam que as pessoas realizem computações (como o reconhecimento de padrões e controle sensório-motor) de modo mais rápido e preciso do que os computadores tradicionais.

As redes neurais objetivam, em sua maioria, solucionar problemas de inteligência artificial, modelando sistemas através de circuitos (conexões) que possibilitam a simulação de um sistema nervoso humano, compreendendo a faculdade que o mesmo tem de aprender e agir diante de diferentes cenários apresentados, incluindo o fato de obter conhecimento através da experiência e da observação. (HAYKIN, 2001) define uma rede neural como um processador maciçamente distribuído e paralelo, constituído de unidades simples de processamento, com tendência natural de transformar conhecimento experimental em conhecimento pronto para uso.

O conhecimento das redes neurais é obtido a partir do seu processo de aprendizagem e pela existência das forças de conexão entre os neurônios, chamadas de pesos sinápticos, usadas para armazenar o conhecimento adquirido. O algoritmo de aprendizagem das redes neurais nada mais é do que um processo no qual os pesos sinápticos da rede são ajustados de maneira ordenada, através de um treinamento para que a mesma trabalhe reconhecendo padrões à forma na qual foi projetada. A seguir, a Figura 2 exibe a forma na qual cada

entrada dos neurônios artificiais é modelada através dos pesos sinápticos P_1, P_2, \dots, P_n . Para cada entrada X_i do neurônio atribui-se uma ponderação P_i , de forma que todas as entradas são multiplicadas por algum fator pré-determinado pelo neurônio. A saída final do neurônio artificial (Y) resulta da aplicação de alguma função de ativação sobre a saída intermediária do corpo celular.

Figura 2: Modelo Matemático do Neurônio

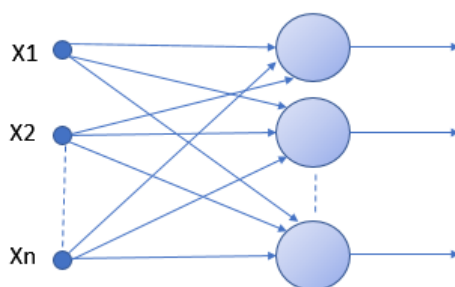


Fonte: Autoria própria.

3.1.2 Arquiteturas das Redes Neurais

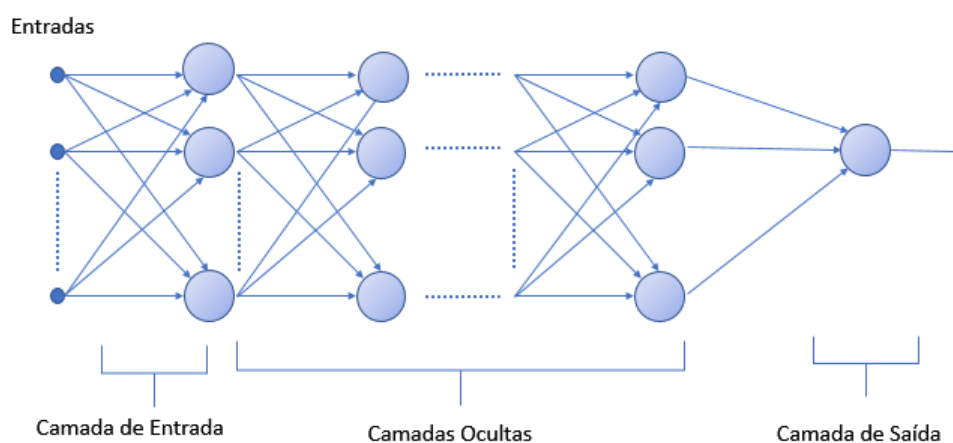
As redes neurais podem apresentar diferentes formas de agrupamento dos neurônios e de arquitetura. Uma rede neural pode ter uma ou várias camadas. As que possuem apenas uma camada são as redes que apresentam um nó entre a entrada e a saída da rede, como pode ser exemplificado na Figura 3. Tais redes são indicadas para a solução de problemas linearmente separáveis. As redes com mais de uma camada entre as camadas de entrada e saída são chamadas de redes multicamadas. As RNA's multicamadas possuem as chamadas camadas escondidas (hidden), que também são chamadas de intermediárias ou ocultas (vide Figura 4). Nas camadas ocultas, os sinais são passados para os neurônios segundo às funções de transferência que cada neurônio possui.

Figura 3: Rede Neural de Camada Única



Fonte: Autoria própria.

Figura 4: Rede Neural de Multicamada



Fonte: Autoria própria.

As redes neurais também podem ser classificadas quanto à sua conectividade, elas podem ser totalmente conectadas ou parcialmente conectadas. No caso de uma rede totalmente conectada, cada neurônio possuirá ligações com todos os neurônios da camada posterior. Caso a rede seja parcialmente conectada, apenas algumas dessas conexões não existirão. Nesse estudo, apenas serão utilizadas redes totalmente conectadas.

Por fim, as redes neurais são agrupadas em relação ao modo que os sinais se propagam dentro da rede. Nesse sentido, as redes neurais podem ser classificadas em redes progressivas ou recorrentes.

- **Redes Neurais Progressivas:** todos os neurônios de uma camada são conectados a todos os neurônios da camada seguinte, não havendo laços de realimentação. O fluxo de informação segue apenas um sentido: da camada de entrada até a camada de saída (feed-forward);

- **Redes Neurais Recorrentes:** apresentam uma ou mais conexões entre neurônios que constituem um ciclo. Os ciclos são responsáveis por armazenar e transmitir as experiências de um neurônio para outro, formulando uma memória interna que auxilia no aprendizado de dados sequenciais (HAYKYN). Os ciclos de retroalimentação podem ser utilizados em qualquer lugar da rede neural e em qualquer direção.

3.1.3 Importância das Redes Neurais Artificiais

Segundo (HAYKIN, 2001), dada a capacidade de generalização das informações das redes neurais, o que possibilita o cálculo de saídas adequadas para entradas que não estavam presentes no arquivo de treinamento, as RNAs podem ser empregadas em problemas que envolvem:

- **Reconhecimento de Padrões:** também conhecido como classificação, essa tarefa é um processo em que um conjunto de entradas é referida a uma classe entre um conjunto definido de classes. Seu emprego envolve análise de crédito; classificação de riscos de doenças a partir de indicadores; identificação de graus e classificações de tumores;
- **Controle de Processos:** em casos onde o modelo de um processo a se controlar não é disponível, as redes neurais são incorporadas em um controlador;
- **Aproximação de funções:** uma RNA multicamada treinada via algoritmo de retro-propagação pode ser usada como um aproximador de funções de caráter geral (SCHOPF, 2007; CYBENKO, 1989);
- **Monitoramento:** tarefa possibilitada graças a capacidade de resposta, praticamente imediata, de uma rede direta empregada em cenários nos quais a rapidez de resposta é um fator primordial (por exemplo, a detecção de vibrações em um reator nuclear);
- **Processamento Temporal:** para que as RNAs consigam trabalhar com informações em sequencia temporal, é necessário que haja uma representação explícita ou implícita no tratamento de dados pela rede. A representação explícita do tempo modifica os dados para dar a idéia de tempo, isso é, o tempo está presente na própria estrutura de dados. A representação explícita do tempo é usada em redes alimentadas adiante atrasadas no tempo.

Já na representação implícita do tempo, os dados entram normalmente nas redes, em sequencia temporal (sem repetição). A cada instante n , entram na rede apenas os dados relativos a esse instante. Toda a estrutura de atrasos temporais é feita internamente pela rede através de sua arquitetura. Para realizar a representação temporal internamente, existem a Redes de Jordan, a Redes de Elman (ELMAN, 1990) e a Rede de Elman Estendida.

Na Rede de Jordan, a realimentação ocorre apenas no sentido dos nós de saída para os nós de entrada, enquanto que na Rede de Elman a realimentação sai apenas dos nós das camadas escondidas. Na Rede de Elman Estendida, a realimentação sai dos nós da camada escondida e também dos nós de saída.

O processamento temporal permite a previsão de séries temporais, filtragem de ruído, controle adaptativo e também a identificação de sistemas.

3.1.4 Aprendizado

As RNAs são capazes de aprender por exemplos, determinando a intensidade de conexões entre os neurônios pertencentes à rede. O algoritmo de aprendizado é caracterizado como um conjunto de procedimentos definidos para ajustar os parâmetros livres (pesos sinápticos) de uma RNA, a fim de que a mesma possa aprender uma determinada função. Para o aprendizado das redes neurais, foram propostos diversos métodos de treinamento, sendo estes subdivididos em dois grandes grupos: o aprendizado supervisionado e o não supervisionado. O uso de uma RNA na resolução de um determinado problema, passa em um primeiro instante por um processo de aprendizagem, onde a rede ambiciona extrair "conhecimentos" fundamentais de padrões de informações apresentados a ela, a fim de modelar uma interpretação própria. Para os dois modelos de aprendizado, supervisionado e o não supervisionado, existem vantagens e desvantagens, as quais serão exibidas a seguir.

3.1.4.1 Aprendizado Supervisionado

O aprendizado supervisionado é realizado a partir de amostras de entradas e saídas previamente conhecidas do sistema, no qual a saída atual da rede neural é comparada com a saída desejada. Os pesos sinápticos terão seus valores iniciais setados aleatoriamente, e serão ajustados via algoritmo de aprendizagem pela rede em um processo iterativo.

O método de aprendizagem procura minimizar o fluxo corrente de erros de todos os elementos em processamento, uma vez que o ajuste do peso sináptico depende do valor esperado e do sinal atual de saída da rede. Tal redução global de erros trabalha modificando continuamente os pesos até que a rede alcance uma certa precisão satisfatória. Neste tipo de aprendizagem, as RNAs necessitam ser treinadas antes de serem utilizadas.

O treinamento é realizado apresentando os sinais de entrada e saída à rede (conjunto de treinamento) e é considerado completo quando a rede neural atinge determinado nível de performance, isto é, quando a rede atinge uma precisão estatística conforme as produções de saída necessárias para uma dada sequência de entradas. Os dados de treinamento precisam ser suficientemente grande para comportar informações relevantes para que a rede aprenda os padrões de entrada.

O conjunto de testes é destinado a analisar o que a rede pode ser capaz de realizar com os dados que ainda não foram apresentados à mesma. Caso o resultado de saída do

sistema não for suficientemente satisfatório para este novo conjunto de dados, pressupõe-se que o treinamento da rede não foi adequado. Tal avaliação é crítica para confirmar que a rede não memorizou um determinado conjunto de dados, mas sim "aprendeu" os padrões envolvidos na aplicação, ou seja, para assegurar se a rede deteve a capacidade de generalização.

3.1.4.2 Aprendizado Não Supervisionado

A rede neural aprende diretamente das características intrínsecas dos dados, sem que haja a necessidade de um conjunto de treinamento. Os ajustes dos pesos sinápticos dão-se através do monitoramento do desempenho interno das redes, analisando as tendências dos sinais de entradas e por conseguinte adaptar-se às necessidades da rede.

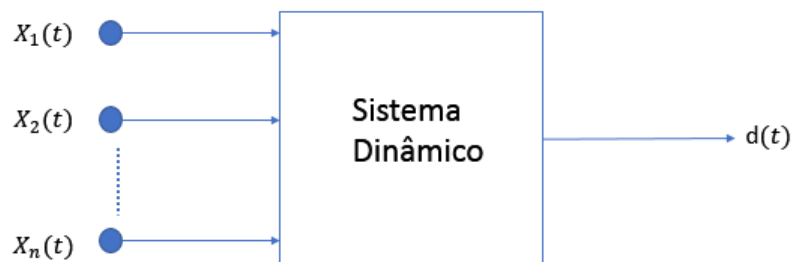
O método de aprendizado não supervisionado mais usual é a Regra de Aprendizagem Competitiva, onde ocorre uma competição entre os neurônios. Ao passo que a competição pelo aprendizado ocorre efetivamente, somente os pesos pertencentes à unidade de processamento "vencedora" (neurônio melhor responder às características dos dados de entrada) serão ajustados.

4 Filtros

(MORETTIN; TOLOI, 2004) definem filtros como um instrumento que é responsável pela passagem de componentes com frequências em uma determinada faixa de frequência. Em outras palavras, o filtro é uma técnica empregada para "limpar" as séries temporais, removendo/tratando componentes ruidosos.

(HAYKIN, 2001) defende que o processo de filtragem envolve o cálculo de dois tipos de sinais: uma saída, representada por $y(t)$, gerada em resposta aos n elementos do vetor de estímulo $x(t)$, e um sinal de erro, representado por $e(t)$, o qual é mensurado pela comparação da saída $y(t)$ com a saída correspondente $d(t)$, produzida pelo sistema agindo como uma resposta desejada ou sinal alvo. Tal processo pode ser visto na Figura 5, onde há a representação da filtragem como um sistema dinâmico:

Figura 5: Modelo de um Sistema Dinâmico



Fonte: Adaptado de (HAYKIN, 2001).

(MORETTIN; TOLOI, 2004) designam o uso do processo de filtragem ao campo da análise espectral, visando decompor a série temporal em seus componentes senoidais com coeficientes aleatórios não correlacionados. Para (LIMA, 2011), o principal fator que justifica o uso desse tipo de análise no tratamento das séries temporais reside no fato de que o espectro fornece uma descrição bastante simples do efeito de uma transformação linear de um processo estacionário.

4.1 Análise Comparativa entre as funções de Fourier e de Wavelets

Em diversas circunstâncias as informações essenciais para entendimento de um processo só pode ser vista no domínio da frequência e não do tempo, ou ainda, a informação mais importante está "escondida" em determinadas frequências do sinal original. A representação do sinal no domínio da frequência é obtido aplicando a transformada de Fourier (TF) à série original, a qual é expressa no domínio do tempo. Tem-se como resultado um

conjunto de frequências que caracterizam o sinal original. Todavia, tal transformação só pode ser aplicada a sinais estacionários. Uma vez que aplicada a sinais não estacionários, a TF trará apenas as informações de quais componentes de frequência que contém o sinal. A TF apenas indicará o conteúdo espectral do sinal, mas não fornecerá o instante ou intervalo de tempo em que essas componentes espectrais aparecem.

Em uma tentativa de contornar este problema, surgiu a Transformada por Janelas de Fourier (TPJF), que é uma generalização da TF, onde sua aplicação visava obter informação do sinal tanto no domínio do tempo quanto no da frequência. Entretanto, sua metodologia apresenta um problema: a escolha da largura da janela de amostragem.

Como majoritariamente as séries financeiras apresentam comportamento não estacionário, tais como mudanças nas tendências e quebres estruturais, as funções de Fourier seriam ineficientes para capturar os eventos. Assim, a Transformada de *Wavelets* surge como uma importante ferramenta para análise de séries não estacionárias, dado que elas são funções dependentes de ambos os domínios da frequência e do tempo.

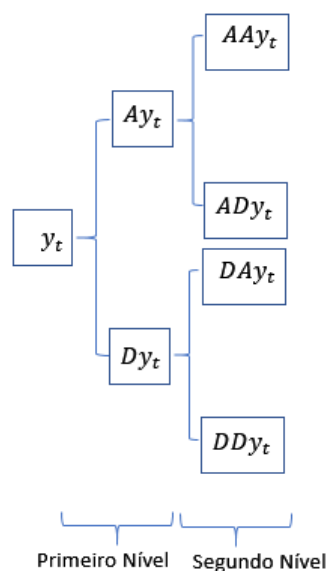
Os métodos de filtragem para séries temporais financeiras já revisados pela literatura indicam o uso de dois filtros: o de *wavelets* e o de Kalman. Isso, pois ambos apresentam melhores desempenho perante a sinais não-estacionários. O presente trabalho faz uso apenas dos filtros de *wavelets*, por isso, na seção a seguir segue uma descrição, sobre alguns aspectos básicos, somente para esse tipo de filtro.

4.2 Filtro de *Wavelets*

A filtragem por *wavelets* tem por objetivo realizar a separação dos dados da série temporal original em duas outras sub-séries por seus componentes de frequência. Segundo (MISITI, 2007), as *wavelets* são funções matemáticas que ampliam os intervalos de dados, possibilitando que cada componente seja alocado em sua respectiva escala. Para (POLIKAR, 2010), este processo de filtragem, para alocar cada componente da série temporal na sua devida escala, nada mais é do que a identificação dos coeficientes correspondentes a cada escala (alta ou baixa frequência), originando as novas sub-séries.

(LIMA, 2011) cita o fato de que esse tipo de filtragem pode ser iterativo. Utilizando-se uma banda de filtros que aplicadas sucessivamente irão decompor o sinal sempre em duas novas faixas de frequência, uma de alta frequência e uma segunda de baixa, criando uma árvore de decomposição com 2^n caminhos diferentes para a decomposição da série. O objetivo da utilização dos filtros de *wavelets* é o de tentar suavizar os efeitos das variações no mercado financeiro através da criação das sub-séries de alta e de baixa frequência. A árvore de decomposição pode ser vista abaixo, na Figura 6:

Figura 6: Árvore de decomposição em dois níveis de uma série temporal via decomposição por *wavelets*



Fonte: Adaptado de (MISITI, 2007).

As *wavelets* são funções dependentes de ambos os domínios da frequência e do tempo, sendo que quando da frequência, é via diferimento da sua forma de onda; e quando do tempo, é através da translação da onda (PERCIVAL; WALDEN, 2006). Isso constitui-se como uma vantagem para o tratamento local de dados, quando comparado as funções de Fourier, pois essas não se ajustam à análise de dados descontínuos, como é o caso das séries financeiras. O uso das *wavelets* é válido graças a sua capacidade de decompor uma série temporal em escalas referentes tanto ao domínio da frequência quanto ao domínio do tempo (GENÇAY; SELÇUK; WHITCHER, 2002).

A análise de *wavelets* é definida pela aplicação da transformada de *wavelet* com diferentes valores para os parâmetros de dilatação e translação, representando a decomposição da série original em seus respectivos componentes localizados no tempo e na frequência para esses parâmetros. A transformada inversa consiste em aplicar os filtros inversos no sinal decomposto, objetivando a reconstrução do sinal original (juntando novamente as duas bandas de frequência).

O uso das *wavelets* tem como objetivo a redução do ruído, processo conhecido como denoising; o qual realiza a escolha de quais coeficientes devem ser mantidos nas sub-séries para preservar a informação e consistência dos dados da série financeira original. Ou seja, a análise de *wavelets* promove o cancelamento daqueles componentes que provocam ruído na série no domínio do tempo, proporcionando que a série fique mais suavizada.

4.2.1 Modelagem do Filtro de *Wavelets*

As *wavelets* podem alterar a forma e comprimento de onda, uma vez que sua formulações matemáticas se expressão através de dilatações ou compressões e translações a partir de uma forma de onda original, conhecida como ondaleta mãe (PERCIVAL; WALDEN, 2006; LIMA, 2011). Nesse processo, o uso das *wavelets* não rompe nenhum valor da série original, pois ela é capaz de notar os pontos de máxima e mínima oscilação.

A interpretação do uso das funções de *wavelets* neste trabalho, pode ser descrita por: obter-se duas sub-séries, uma de alta e outra de baixa frequência, para suavizar os efeitos das variações no mercado financeiro sobre uma série financeira.

A seguir, será abordada a modelagem matemática das funções de *wavelets*, com base nos estudos e menções dos autores (MISITI, 2007; SCHOPF, 2007; CARVALHO, 2013).

Para (MISITI, 2007), para que uma função Ψ_t possa ser considerada uma *wavelets*, ela deve atender as seguintes propriedades:

- A)

$$\int_{-\infty}^{\infty} \Psi(t) dt = 0 \quad (4.1)$$

- B)

$$\int_{-\infty}^{\infty} |\Psi(t)|^2 dt < \infty \quad (4.2)$$

Ou seja, pertencem ao conjunto das funções quadrado integráveis, representadas por $L^2(IR)$. (MORETTIN, 1999) define o produto interno de duas funções $f(t)$ e $g(t)$ em $L^2[a, b]^2$ como sendo:

$$\langle f(t), g(t) \rangle = \int_{-a}^b f(t) \cdot * g(t) dt \quad (4.3)$$

O autor determina a transformada de *wavele* (WT) como sendo o produto interno do sinal de teste com as funções base $\Psi(t)(a, b)(t)$:

$$WT_X^\psi(a, b) = \Psi_X^\psi(a, b) = \int y(t) \cdot \psi_{a,b}^* dt \quad (4.4)$$

sendo

- $\psi_{a,b} = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right)$
- $y(t)$ é a série financeira

Para (MORETTIN, 1999; SCHOPF, 2007), o objetivo é gerar um espaço de estado a partir da função Ψ , que é obtida através de dilatações do parâmetro a ou compressões e translações definidos pelo parâmetro b :

$$\Psi_{a,b}(x) = |a|^{-\frac{1}{2}} \Psi\left(\frac{x-b}{a}\right), a, b \in \mathfrak{R}, a \neq 0 \quad (4.5)$$

Segundo (SCHOPF, 2007), uma análise de *wavelets* é definida pela aplicação da transformada de *wavelets* com diversos valores para os parâmetros de dilatação e translação, podendo representar a decomposição da série original em seus respectivos componentes localizados no tempo e na frequência para esses parâmetros. A forma de onda Ψ é conhecida como ondaleta mãe (*mother wavelet*). (MORETTIN, 1999) sugere que os parâmetros a e b assumam os seguintes valores especiais: $a = 2^{-j}$, para a dilatação binária e $b = k2^{-j}$, para a translação diática. Resultando em:

$$\Psi_{j,k}(x) = 2^{\left(\frac{j}{2}\right)} \Psi(2^j x - k), j, k \in Z \quad (4.6)$$

Também, pode-se gerar uma ondaleta pai (Φ), a partir da função escala, o qual é solução da equação:

$$\Phi(x) = \sqrt{2} \sum_k^{\infty} l_k \Phi(2x - k) \quad (4.7)$$

sendo $l_k = \int_{-\infty}^{\infty} \Phi(t) \Phi(2t - k) dt$, chegamos em:

$$\Phi_{j,k}(x) = 2^{\left(\frac{j}{2}\right)} \Phi(2^j x - k), j, k \in Z \quad (4.8)$$

Em outras palavras, a análise de wavelet é feita pela aplicação sucessiva da transformada de wavelet, variando-se os valores dos parâmetros a e b . (GENÇAY; SELÇUK; WHITCHER, 2002) afirma que uma série financeira pode ser decomposta por uma sequência de projeções de *wavelets* pai e mãe, tendo como base as funções Φ e Ψ , de acordo com as Equação 4.6 e Equação 4.8.

Para (SCHOPF, 2007), a *wavelet* mãe funciona como uma janela móvel de abertura finita que percorre a série temporal, onde a abstração dos pontos de alta e baixa frequência se dá pela translação e dilatação da forma da *wavelet*. Tal observação pode ser descrita por:

$$y_t = \sum_k a_{j,k} \Phi_{j,k}(t) + \sum_k d_{j,k} \Psi_{j,k}(t) + \sum_k d_{j-1,k} \Psi_{j-1,k}(t) + \dots + \sum_k d_{1,k} \Psi_{1,k}(t) \quad (4.9)$$

, sendo:

- j o número de componentes;

- k o número que varia de 1 até o número de coeficiente

Os coeficientes $a_{j,k}, d_{j,k}, \dots, d_{1,k}$ são os coeficientes das transformadas de *wavelets*, dadas pela Equação 4.10 e pela Equação 4.11, conhecidas respectivamente por parte de aproximação e parte de detalhe.

$$a_{a,k} = \int \Phi_{j,k}(t)y_t dt \quad (4.10)$$

$$d_{j,k} = \int \Psi_{j,k}(t)y_t dt \quad (4.11)$$

5 Metodologia

Aqui, serão apresentados os principais procedimentos desenvolvidos no trabalho, tais como: métodos empregados, procedimentos de análise, previsão e avaliação dos resultados e as ferramentas utilizadas como filtros de volatilidade.

Como já citado na Capítulo 1, o objetivo do presente estudo é o de estruturar conceitos a fim de compreender o uso de filtros de volatilidade juntamente com redes neurais, com o intuito de reduzir os ruídos e melhorar a qualidade da previsão. Para tal, a primeira etapa do estudo foi focada em revisões bibliográficas para servirem de apoio teórico e para direcionarem o trabalho de modo a adotar ferramentas que melhor se adequassem às séries temporais financeiras. Nesse momento, foram estudados livros, artigos e produções acadêmicas nacionais e internacionais (como as referências bibliográficas abordavam diferentes tipos de séries temporais em diferentes campos de estudo, deu-se foco aquelas que eram direcionadas ao estudo de séries financeiras, já que pretendia-se melhor compreender a forma como esse campo de estudo foi abordado e explorado).

5.1 Determinação das Variáveis de Entrada

Para a determinação das variáveis de entrada do modelo, foram levadas em consideração variáveis da conjuntura da macroeconomia, como as taxas de câmbio e índices da *BM&FBOVESPA* (tais variáveis, em um primeiro momento, são consideradas variáveis candidatas a entrada do modelo). O uso dos índices pode ser justificado pelo fato de serem indicadores de desempenho de um conjunto de ações, ou seja, mostram a valorização de um determinado grupo de papéis ao longo do tempo. Os preços das ações podem variar por fatores relacionados à empresa ou por fatores externos, como o crescimento do país, do nível de emprego e da taxa de juros. Assim, as ações de um índice podem apresentar um comportamento diferente no mesmo período, podendo ocorrer valorização ou desvalorização.

Foram considerados três grandes grupos de índices da *BM&FBOVESPA*:

- **Índices Amplos:** contempla os Índices Bovespa (Ibovespa), Brasil 50 (IBrX 50), Brasil 100 (IBrX 100) e o Brasil Amplo (IBrA);
- **Índices Setoriais:** abrange os Índices Energia Elétrica (IEE), Industrial (INDX), Consumo (ICON), Imobiliário (IMOB), Financeiro (IFNC), Materiais Básicos (IMAT) e Utilidade Pública (UTIL).
- **Índices de Segmento:** é composto pelos Índices MidLarge Cap (MLCX), Small Cap (SMLL), Valor *BM&FBOVESPA* (IVBX 2) e pelo Dividendos *BM&FBOVESPA*

(IDIV);

Para as taxas de câmbio, foram levadas em consideração as seguintes moedas:

- USD/BRL: Dólar americano por Real Brasileiro;
- EUR/BRL: Euro por Real Brasileiro;
- JPY/BRL: Iene Japonês por Real Brasileiro;
- GBP/BRL: Libra Esterlina por Real Brasileiro;
- AUD/BRL: Dólar Australiano por Real Brasileiro;
- CHF/BRL: Franco Suiço por Real Brasileiro;
- CAD/BRL: Dólar Canadense por Real Brasileiro;
- CNY/BRL: Yuan Chinês por Real Brasileiro.

Em posse de todas essas séries temporais candidatas a variáveis de entrada do modelo, o passo seguinte foi aplicar uma função de correlação de Pearson no maior intervalo de tempo possível (uma vez que a data de criação de alguns dos índices *BM&FBOVESPA* são mais recentes do que as das ações estudadas). Foram escolhidas como entradas do modelo as variáveis que apresentaram correlação acima de 0,7 (correlação forte).

Para efeito de exemplificação, a ação da empresa Companhia Energética de Minas Gerais (CMIG4) será tomada como *Benchmark*. Nesse caso, as variáveis que passaram a integrar os dados de entrada do modelo foram: o Índice Energia Elétrica, e as taxas de câmbio USD/BRL, EUR/BRL, JPY/BRL e CNY/BRL. Além dessas séries temporais, também foram utilizadas os dados de preços diários das ações, tais como o valor de fechamento, abertura, máxima e mínima.

5.2 Tratamento das Variáveis do Modelo e Normalização dos Dados

Como dito na subseção 2.1.2, o primeiro procedimento a ser feito para o tratamento das séries financeiras é a formação da série dos log retornos através da Equação 2.1.2. Em posse de uma série temporal descrita por $y_t = (y_1, y_2, \dots, y_n)$, os retornos logarítmicos são dados pela Equação 2.1.2, $R_t = \ln \left(\frac{y_t}{y_{t-1}} \right)$.

Os dados de entrada da RNA necessitam, na maioria dos casos, de normalização. Esse processo é aconselhável uma vez que coloca o intervalo de dados dentro dos limites que facilitam a tarefa da rede para realizar a minimização do erro da saída. Então, é aplicada

a normalização proposta por (AZOFF, 1994), a qual normaliza a amostra total da série em um intervalo $[0,1]$, dada pela expressão:

$$Z_t = \frac{R_t - \min(R_t)}{\max(R_t) - \min(R_t)} \quad (5.1)$$

É fácil ver que, após previsto Z_{t+1} , para reconstruir seu valor no nível da série original, basta aplicar a Equação 5.4.2:

$$y_{t+1} = e^{(Z_{t+1}[\max(R_t) - \min(R_t)] + \min(R_t)) + \ln(y_t)} \quad (5.2)$$

5.3 Redes Neurais Recorrentes

A partir dos conceitos explicados em subseção 3.1.2, justifica-se nesse trabalho o uso das Redes Neurais Recorrentes com representação implícita do tempo (Elman, redes com memórias de longo prazo).

Em relação à topologia, para cada horizonte de tempo no qual se almejava realizar uma previsão, foram declaradas 15 topologias de redes distintas, variando-se o número de camadas escondidas e o número de neurônios cada camada. Essa metodologia foi adotada pois o número de camadas escondidas e o número de neurônios em cada uma das camadas escondidas determinam a complexidade do modelo da RNA e como não há regra para determinar tais parâmetros, não sabemos a princípio quais especificações seriam as melhores para influir na relação entrada-saída de modo eficaz. Ainda, a metodologia é estruturada no fato de que as séries financeiras apresentam em seu processo gerador características tidas como aleatórias, ou seja, buscou-se encontrar a RNA mais apta a extrair as estatísticas de ordem superior do processo que rege o comportamento das variáveis de entrada, sendo capaz de evitar um mínimo local.

Além disso, se usarmos poucos neurônios na camada escondida, a rede neural é incapaz de modelar dados mais complexos, resultando em uma baixa capacidade de generalização (underfitting). Por outro lado, se muitos neurônios forem utilizados, o treinamento da rede irá se tornar excessivamente longo e a rede pode sofrer problemas de sobre-ajuste (overfitting), resultando na perda da capacidade preditiva da rede, uma vez que observa-se pequenos desvios de previsão para os dados usados na fase de treino, mas grandes desvios quando novos dados de entrada são utilizados.

A seguir, na Tabela 1 podem ser vistas as diferentes topologias definidas no trabalho:

Tabela 1: Topologias das RNAs

| Topologia | Número de Camadas | Neurônios em Cada Camada |
|-----------|-------------------|--------------------------|
| 1 | 3 | [7 5 1] |
| 2 | 3 | [3 2 1] |
| 3 | 3 | [3 9 1] |
| 4 | 3 | [2 3 1] |
| 5 | 4 | [1 2 3 1] |
| 6 | 4 | [1 6 3 1] |
| 7 | 4 | [7 5 3 1] |
| 8 | 4 | [3 2 4 1] |
| 9 | 4 | [1 5 7 1] |
| 10 | 4 | [4 2 3 1] |
| 11 | 4 | [3 2 10 1] |
| 12 | 5 | [5 2 3 2 1] |
| 13 | 5 | [7 5 3 2 1] |
| 14 | 5 | [5 2 4 3 1] |
| 15 | 5 | [3 2 1 2 1] |

Fonte: Autoria própria.

5.4 Algoritmo de Treinamento

Conforme dito na subseção 3.1.4.1, este trabalho é instigado a adotar o método de aprendizado supervisionado, onde, resumidamente, o aprendizado é baseado em um conjunto de exemplos de entradas-saídas que é apresentada a rede. A partir da entrada, a rede inicializa seu processamento e a saída obtida é comparada com a saída esperada. Se esses valores não forem iguais, então é aplicado um método para ajuste dos pesos sinápticos, a fim de chegar-se a um erro mínimo ou a um erro aceitável.

O algoritmo de aprendizado utilizado foi o *backpropagation*. Nele, o ajuste dos pesos sinápticos ocorre pela execução de um processo de otimização realizado em duas fases: *forward* e *backward*. Na fase *forward*, os dados de entrada alimentam a rede e são propagados para frente até que as saídas dos nós da última camada sejam obtidas. Já na fase do *backward*, o erro entre a resposta gerada pela rede e a saída real é utilizado para ajustar os pesos sinápticos das conexões da rede (o sinal de erro é propagado na direção da camada de entrada e o gradiente decrescente é usado para ajustar os parâmetros da rede). Neste trabalho, foi utilizado o algoritmo de Levenberg-Marquardt, que é uma otimização do do gradiente decrescente.

5.4.1 Validação Cruzada

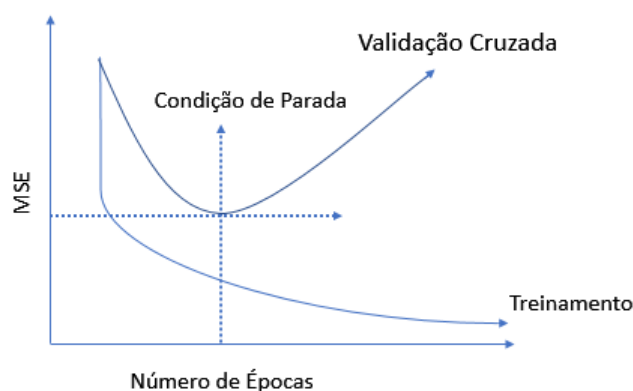
Segundo (HAYKIN, 2001), um dos entraves do uso das RNAs reside na dificuldade da identificação do melhor ponto de parada de treinamento, pois o erro treinamento é inicializado com um alto valor, mas decresce rapidamente e continua diminuindo lentamente,

tendendo a atingir um mínimo local na superfície do erro.

Assim, para identificar um ponto de parada do aprendizado de modo a se obter uma melhor generalização da rede, pode-se aplicar a técnica da regra de parada com base na validação cruzada (MATHIAS; DIAS, 2007).

A validação cruzada consiste em medir o desempenho da rede de épocas em épocas, utilizando um conjunto de dados diferentes dos usados para estimar os parâmetros durante o treinamento. Conforme o treinamento se segue, os resultados obtidos na validação cruzada tendem a cair e a ideia é de que quando estes resultados comecem a subir o treinamento seja finalizado, conforme ilustrado pela Figura 7.

Figura 7: Validação Cruzada



Fonte: Adaptado de (HAYKIN, 2001).

Esse processo é importante pois garante que não haja excesso de treinamento da rede neural, evitando problemas relacionados a overfitting. Durante a implementação computacional, foi adotado que as épocas das redes neurais tivessem início em 20 e fossem acrescidas de 20 unidades até chegarem ao limiar de 1000 épocas.

Ainda, durante o processo de treinamento, mais especificamente na etapa de teste dos dados das redes neurais, avaliou-se o desempenho de quais topologias eram mais adequadas para a previsão de k passos à frente. Para a etapa de validação do modelo, Capítulo 6, utilizou-se a topologia das redes que mais vezes foram acertivas durante a fase de teste.

5.4.2 Estatísticas de Erro

Segundo (AHLBURG, 1992), a escolha da medida de análise dos erros de previsão deve levar em consideração o tamanho da amostra a ser prevista. Ainda segundo o autor, para pequenos passos a frente na previsão, deve-se utilizar o MAPE (*Mean Absolute*

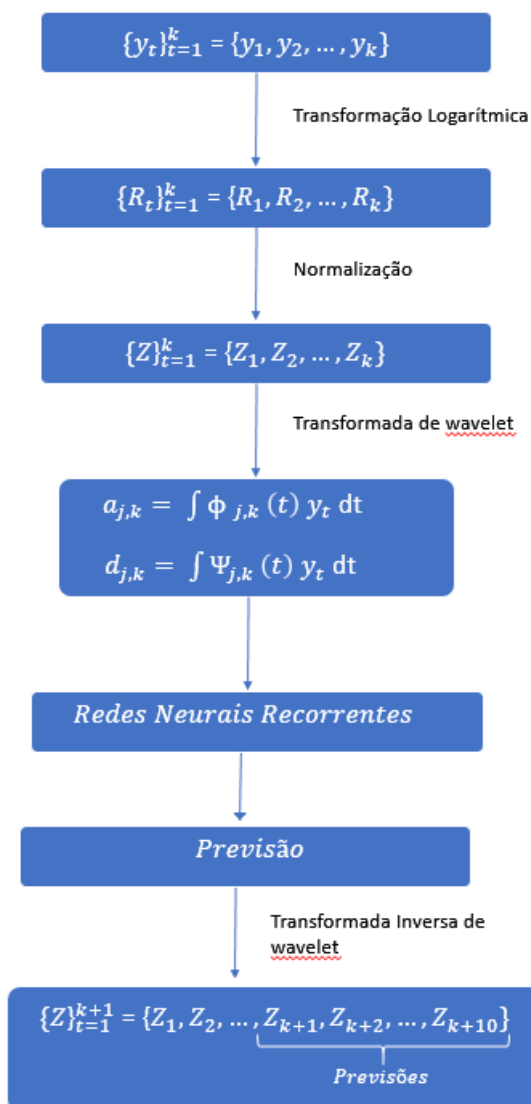
Percentage Error). Como neste estudo faz-se previsões para $k=10$ passos a frente, a medida de análise do erro será o MAPE (mean absolute percentage error):

$$MAPE_h = \frac{\sum_{i=1}^h \left[\frac{\epsilon_i}{y_i} \right]}{h} \cdot 100\% \quad (5.3)$$

5.4.3 Decomposição do Esquema Operacional

Toda a implementação computacional foi feita no *Matlab*. O fluxograma descritivo, implementado nesse trabalho, para a previsão de séries temporais com filtro de *wavelets* e redes neurais, pode ser vista no esquema operacional, representado pela Figura 8. Em relação aos filtros de *wavelets*, (SCHOPF, 2007) indica o uso da *wavelet* de Haar (ou "daubesch1"), pois ela apresenta melhor desempenho para séries financeiras com agrupamentos de volatilidade.

Após a transformada inversa de wavelet, deverá ser feita a desnormalização dos dados previstos pela rede, para reconstruir o valor do sinal no nível da série original, através da Equação 5.4.2.

Figura 8: Fluxograma para a previsão com filtro de *wavelets* e redes neurais recorrentes

Fonte: Autoria Própria

6 Resultados

Neste capítulo, serão exibidos as simulações computacionais como forma de avaliação e validação do modelo proposto. Para tal, serão avaliados o desempenho do modelo frente a quatro grupos de ações: *CMIG4*, *CPLE6*, *ELET3* e *CPFE3*.

6.1 Análise da Série ELET3

6.1.1 Previsão com redes neurais recorrentes

Nesta etapa, não fez-se uso do filtro não linear. A rede neural foi treinada sendo "alimentada" apenas por sinais do retorno logarítmico. A Tabela 2 apresenta as topologias, definidas no seção 5.3, que foram mais eficientes para cada passo k na fase de treinamento, juntamente com o erro MAPE:

Tabela 2: Melhores Topologias das RNAs para Previsão sem Filtro da Série ELET3

| k passos à diante | Topologia | MAPE |
|---------------------|-----------|-------|
| 1 | 3 | 4.35% |
| 2 | 7 | 4.37% |
| 3 | 2 | 4.3% |
| 4 | 3 | 4.41% |
| 5 | 5 | 2.47% |
| 6 | 5 | 4.31% |
| 7 | 10 | 4.39% |
| 8 | 6 | 4.29% |
| 9 | 3 | 4.40% |
| 10 | 2 | 4.43% |

Fonte: Autoria própria.

Após o treinamento, os dados de mercado do dia 05/09/2018 foram apresentado às redes, visando que as mesmas realizassem previsões para um horizonte de dez dias. A Tabela 3 abaixo fornece os dados previstos, trazidos ao nível real da série, o erro contido em cada previsão e uma métrica de avaliação relativa a capacidade de "sensibilidade preditiva" das redes.

Tabela 3: Previsão sem Filtro da Série ELET3

| Data a ser Prevista | Cotação Real | Valor Previsto | Erro | Predição da Tendência de Movimentos |
|---------------------|--------------|----------------|---------|-------------------------------------|
| 06/09/2018 | 16.05 | 16.44 | -2.39% | Não |
| 10/09/2018 | 15.89 | 16.44 | -2.31% | Não |
| 11/09/2018 | 14.81 | 16.45 | 9.95% | Não |
| 12/09/2018 | 15.51 | 16.43 | 5.60% | Não |
| 13/09/2018 | 15.24 | 17.01 | -10.38% | Não |
| 14/09/2018 | 14.71 | 16.42 | -10.39% | Não |
| 17/09/2018 | 15.31 | 16.43 | 6.83% | Não |
| 18/09/2018 | 15.97 | 16.43 | -2.87% | Não |
| 19/09/2018 | 15.22 | 16.40 | -7.20% | Não |
| 20/09/2018 | 15.38 | 16.92 | -9.12% | Não |

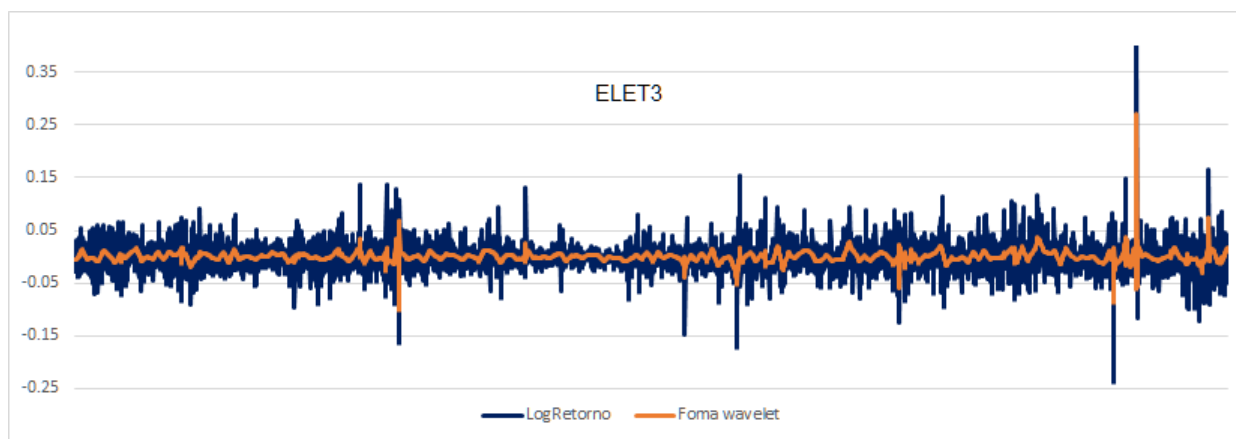
Fonte: Autoria própria.

Nota-se que os valores oriundos das redes neurais indicam que elas não apresentaram uma boa capacidade preditiva, uma vez que seus valores em 80% dos casos ficaram "estacionados", praticamente constantes. Além disso, o erro da previsão apresentou-se elevado, com uma média próxima aos 10.40%.

6.1.2 Previsão com filtro de *waveles* e redes neurais recorrentes

Ao aplicar o filtro de *wavelet* sobre a série do log retorno, nota-se o a remoção dos componentes ruidosos, conforme a Figura 9:

Figura 9: Formas de Onda do Log Retorno da ação ELET3 - com e sem filtro



Fonte: Autoria Própria.

Após o treinamento das redes, a Tabela 4 exibe qual a topologia de rede neural apresentou melhor desempenho para a previsão de k passos à frente, sendo k variando uma até dez unidades, conforme proposto na subseção 5.4.1. Nesta tabela, também é possível analisar o erro MAPE relacionado a cada k e a cada topologia.

Tabela 4: Melhores Topologias das RNAs para Previsão da Série ELET3

| k passos à diante | Topologia | MAPE |
|---------------------|-----------|-------|
| 1 | 5 | 0.64% |
| 2 | 9 | 0.80% |
| 3 | 12 | 1.17% |
| 4 | 6 | 1.07% |
| 5 | 6 | 1.13% |
| 6 | 8 | 1.17% |
| 7 | 9 | 1.17% |
| 8 | 1 | 1.14% |
| 9 | 8 | 1.23% |
| 10 | 6 | 1.26% |

Fonte: Autoria própria.

Com o modelo já treinado, informou-se os dados de mercado, conforme dito em seção 5.1, do dia 05/09/2018 para a previsão dos preços dos próximos dez dias. À estes novos valores, aplicou-se a transformada inversa de wavelet, conforme indicado na subseção 4.2.1 e ilustrado na Figura 8 e em seguida utilizou-se a Equação 5.4.2, a fim de reconstruir o valor previsto ao nível da série original.

A Tabela 5 traz os valores previstos e os valores reais da cotação da ação, bem como o erro percentual da previsão e um campo destinado a verificar se o modelo proposto foi capaz de detectar as inversões de movimento da série.

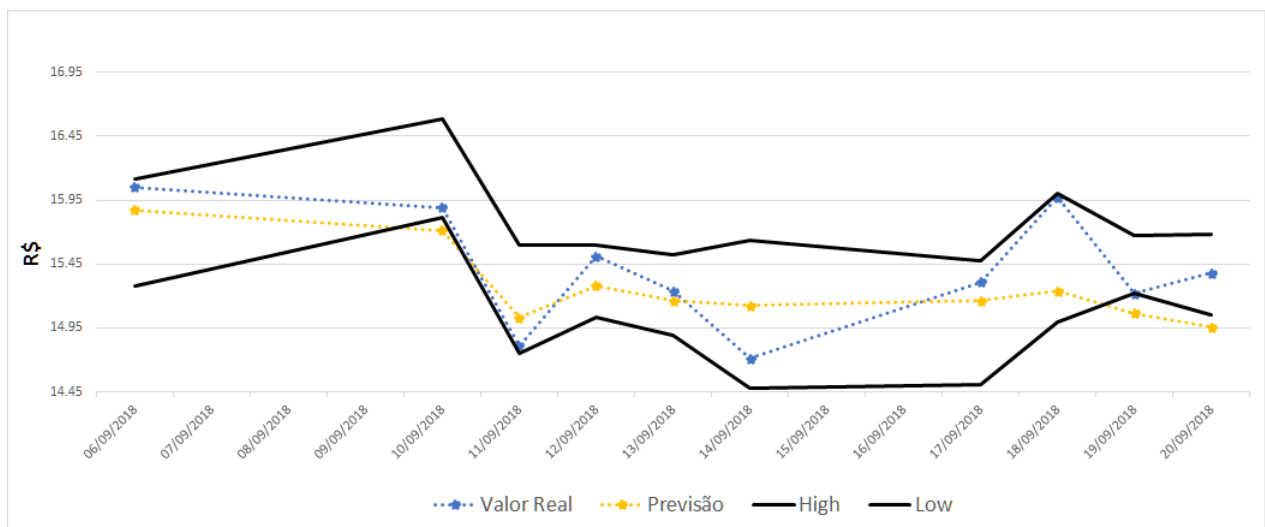
Tabela 5: Previsão da Série ELET3

| Data a ser Prevista | Cotação Real | Valor Previsto | Erro | Predição da Tendência de Movimentos |
|---------------------|--------------|----------------|--------|-------------------------------------|
| 06/09/2018 | 16.05 | 15.88 | 1.10% | Sim |
| 10/09/2018 | 15.89 | 15.71 | 1.13% | Sim |
| 11/09/2018 | 14.81 | 15.03 | -1.47% | Sim |
| 12/09/2018 | 15.51 | 15.28 | 1.5% | Sim |
| 13/09/2018 | 15.24 | 15.16 | 0.49% | Sim |
| 14/09/2018 | 14.71 | 15.13 | -2.76% | Sim |
| 17/09/2018 | 15.31 | 15.16 | 0.98% | Sim |
| 18/09/2018 | 15.97 | 15.24 | 4.8% | Sim |
| 19/09/2018 | 15.22 | 15.07 | 1.02% | Sim |
| 20/09/2018 | 15.38 | 14.96 | 2.84% | Não |

Fonte: Autoria própria.

Para melhor visualização dos dados obtidos, a Figura 10 traz um gráfico comparativo entre os valores gerados pelo modelo e os valores reais do fechamento para as ações. Nele, também há a informação das cotações de máxima e mínima e do dia, parâmetros os quais foram escolhidos para servirem como um intervalo de confiança (não é esperado que o modelo de filtros de *wavelets* e de redes neurais apresentem muitas previsões fora desse intervalo).

Figura 10: Previsão da Série ELET3



Fonte: Autoria Própria.

Pode-se considerar satisfatório desempenho apresentado pelo modelo proposto perante à série temporal dos preços da ação ELET3, pois em 90% dos casos a previsão das tendências de movimento foram feitas com êxito e em 70% os dados estiveram dentro do intervalo de confiança.

6.2 Análise da Série CPFE3

Dado que a metodologia empregada para essa série financeira foi idêntica à apresentada na seção anterior e também é idêntica às demais, a partir daqui serão apresentados os resultados finais de modo mais direto.

6.2.1 Previsão com redes neurais recorrentes sem filtro de wavelets

Dos dados de treinamento, tem-se:

Tabela 6: Melhores Topologias das RNAs para Previsão sem Filtro da Série CPFE3

| k passos à diante | Topologia | MAPE |
|---------------------|-----------|-------|
| 1 | 5 | 3.28% |
| 2 | 6 | 3.28% |
| 3 | 6 | 3.29% |
| 4 | 5 | 3.28% |
| 5 | 7 | 3.25% |
| 6 | 2 | 3.27% |
| 7 | 13 | 3.29% |
| 8 | 11 | 3.29% |
| 9 | 5 | 3.30% |
| 10 | 8 | 3.33% |

Fonte: Autoria própria.

Alimentando as redes com as cotações do dia 05/09/2018, tem-se:

Tabela 7: Previsão sem Filtro da Série CPFE3

| Data a ser Prevista | Cotação Real | Valor Previsto | Erro | Predição da Tendência de Movimentos |
|---------------------|--------------|----------------|-------|-------------------------------------|
| 06/09/2018 | 23.23 | 23.48 | 3.34% | Não |
| 10/09/2018 | 23.31 | 23.38 | 4.14% | Não |
| 11/09/2018 | 23.20 | 23.38 | 3.65% | Não |
| 12/09/2018 | 23.49 | 23.37 | 5.02% | Não |
| 13/09/2018 | 24.28 | 24.37 | 8.55% | Não |
| 14/09/2018 | 24.08 | 23.39 | 7.55% | Não |
| 17/09/2018 | 23.98 | 23.44 | 6.85% | Não |
| 18/09/2018 | 23.72 | 23.38 | 5.98% | Não |
| 19/09/2018 | 23.64 | 23.37 | 5.65% | Não |
| 20/09/2018 | 23.85 | 23.38 | 6.56% | Não |

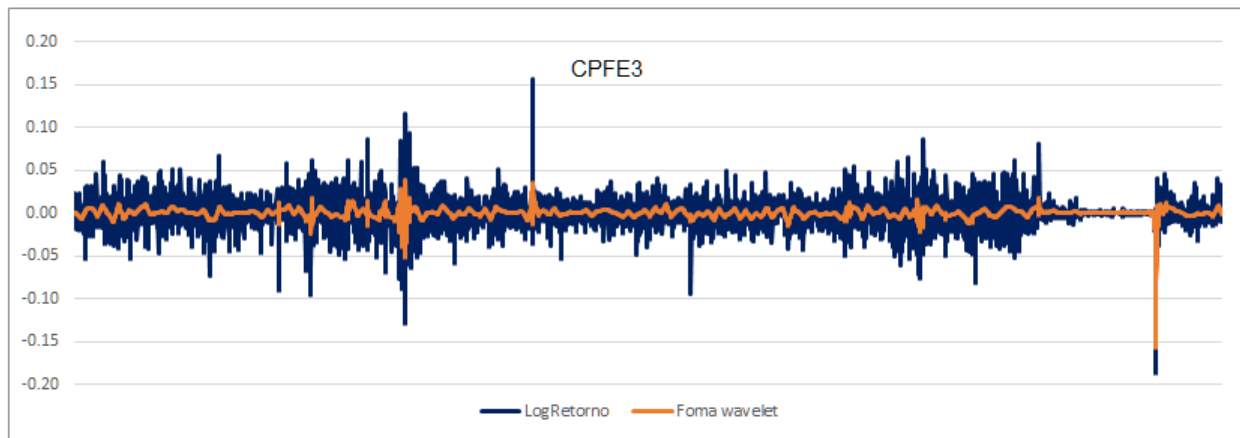
Fonte: Autoria própria.

Observa-se que os valores obtidos das redes neurais não apresentaram uma boa capacidade preditiva, uma vez que sua amostra ficou quase constante, não apresentando sensibilidade ao movimento.

6.2.2 Previsão com filtro de *waveles* e redes neurais recorrentes

O uso do filtro de *wavelet* sobre a série do log retorno "limpa" a série, conforme a Figura 11:

Figura 11: Formas de Onda do Log Retorno da ação CPFE3 - com e sem filtro



Fonte: Autoria Própria.

As topologias mais adequadas para a série CPFE3 de acordo com cada passo de previsão, podem ser vistas na Tabela 12:

Tabela 8: Melhores Topologias das RNAs para Previsão da Série CPFE3

| k passos à diante | Topologia | MAPE |
|---------------------|-----------|-------|
| 1 | 5 | 0.39% |
| 2 | 6 | 0.41% |
| 3 | 12 | 0.56% |
| 4 | 4 | 0.54% |
| 5 | 8 | 0.57% |
| 6 | 5 | 0.64% |
| 7 | 8 | 0.62% |
| 8 | 2 | 0.70% |
| 9 | 6 | 0.64% |
| 10 | 5 | 0.68% |

Fonte: Autoria própria.

Utilizando-se dos mesmos atributos citados no subseção 6.1.2, a Tabela 13 traz as informações à despeito do desempenho preditivo:

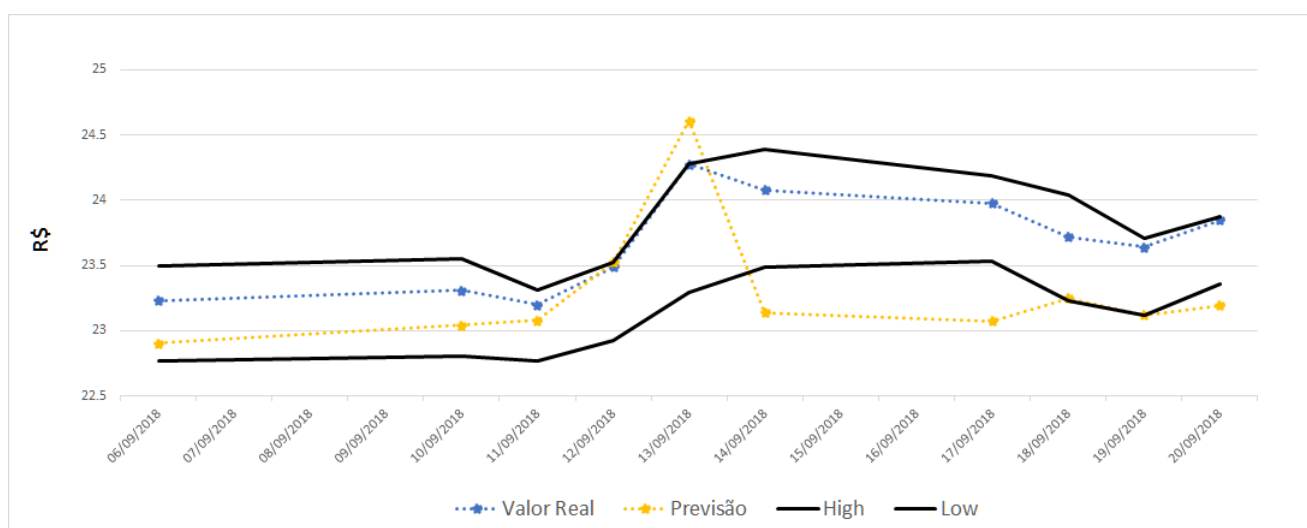
Tabela 9: Previsão da Série CPFE3

| Data a ser Prevista | Cotação Real | Valor Previsto | Erro | Predição da Tendência de Movimentos |
|---------------------|--------------|----------------|--------|-------------------------------------|
| 06/09/2018 | 23.23 | 23.90 | 1.43% | Sim |
| 10/09/2018 | 23.31 | 23.04 | 1.16% | Sim |
| 11/09/2018 | 23.20 | 23.08 | 0.51% | Não |
| 12/09/2018 | 23.49 | 23.53 | -0.16% | Sim |
| 13/09/2018 | 24.28 | 24.61 | -1.34% | Sim |
| 14/09/2018 | 24.08 | 23.14 | 4.07% | Sim |
| 17/09/2018 | 23.98 | 23.08 | 3.91% | Sim |
| 18/09/2018 | 23.72 | 23.25 | 2.02% | Não |
| 19/09/2018 | 23.64 | 23.12 | 2.25% | Sim |
| 20/09/2018 | 23.85 | 23.20 | 2.82% | Sim |

Fonte: Autoria própria.

Graficamente, o comportamento dos valores previstos pelo modelo podem ser vistos abaixo, na Figura 12.

Figura 12: Previsão da Série CPFE3



Fonte: Autoria Própria.

Aqui, também é considerado satisfatório desempenho apresentado pelo modelo proposto defronte à série temporal dos preços da ação CPFE3, uma vez que em 80% dos casos a previsão das tendências de movimento foram feitas com êxito e em 60% das vezes os dados estiveram dentro do intervalo de confiança.

6.3 Análise da Série CMIG4

6.3.1 Previsão com redes neurais recorrentes sem filtro de wavelets

A partir dados de treinamento, tem-se:

Tabela 10: Melhores Topologias das RNAs para Previsão sem Filtro da Série CMIG4

| k passos à diante | Topologia | MAPE |
|---------------------|-----------|-------|
| 1 | 5 | 2.23% |
| 2 | 6 | 2.12% |
| 3 | 6 | 2.53% |
| 4 | 5 | 2.67% |
| 5 | 7 | 2.88% |
| 6 | 2 | 3.24% |
| 7 | 13 | 4.48% |
| 8 | 11 | 3.87% |
| 9 | 5 | 4.10% |
| 10 | 8 | 4.27% |

Fonte: Autoria própria.

Alimentando as redes com as cotações do dia 05/09/2018, tem-se:

Tabela 11: Previsão sem Filtro da Série CMIG4

| Data a ser Prevista | Cotação Real | Valor Previsto | Erro | Predição da Tendência de Movimentos |
|---------------------|--------------|----------------|--------|-------------------------------------|
| 06/09/2018 | 7.25 | 7.29 | 0.61% | Sim |
| 10/09/2018 | 7.24 | 7.28 | 0.61% | Sim |
| 11/09/2018 | 7.02 | 6.27 | -3.50% | Sim |
| 12/09/2018 | 7.04 | 6.31 | -3.66% | Sim |
| 13/09/2018 | 6.81 | 6.30 | -6.67% | Não |
| 14/09/2018 | 6.80 | 6.30 | -6.85% | Não |
| 17/09/2018 | 6.96 | 7.31 | -4.73% | Não |
| 18/09/2018 | 7.29 | 7.31 | -0.32% | Não |
| 19/09/2018 | 7.15 | 7.31 | -2.17% | Não |
| 20/09/2018 | 7.17 | 7.32 | -2.01% | Não |

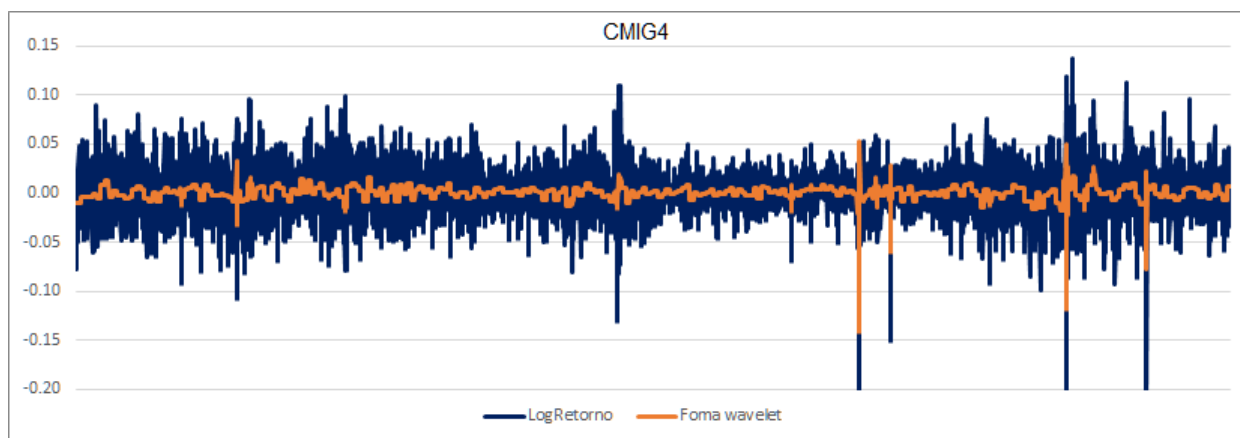
Fonte: Autoria própria.

Mais uma vez, nota-se que os valores obtidos das redes neurais não apresentaram uma boa capacidade preditiva, já que sua previsão permanece quase constante, não apresentando sensibilidade ao movimento.

6.3.2 Previsão com filtro de *wavelets* e redes neurais recorrentes

O uso do filtro de *wavelet* sobre a série do log retorno "limpa" a série, conforme a Figura 13:

Figura 13: Formas de Onda do Log Retorno da ação CMIG4 - com e sem filtro



Fonte: Autoria Própria.

As topologias mais adequadas para a série CMIG4 de acordo com cada passo de previsão, com seus respectivos erro MAPE, podem ser vistas na Tabela 12:

Tabela 12: Melhores Topologias das RNAs para Previsão da Série CMIG4

| k passos à diante | Topologia | MAPE |
|---------------------|-----------|-------|
| 1 | 4 | 0.66% |
| 2 | 2 | 0.67% |
| 3 | 14 | 0.67% |
| 4 | 12 | 0.65% |
| 5 | 14 | 0.66% |
| 6 | 6 | 0.66% |
| 7 | 11 | 0.72% |
| 8 | 10 | 0.74% |
| 9 | 5 | 0.78% |
| 10 | 9 | 0.80% |

Fonte: Autoria própria.

Utilizando-se dos mesmos atributos citados no subseção 6.1.2, a Tabela 13 traz as informações à despeito do desempenho preditivo:

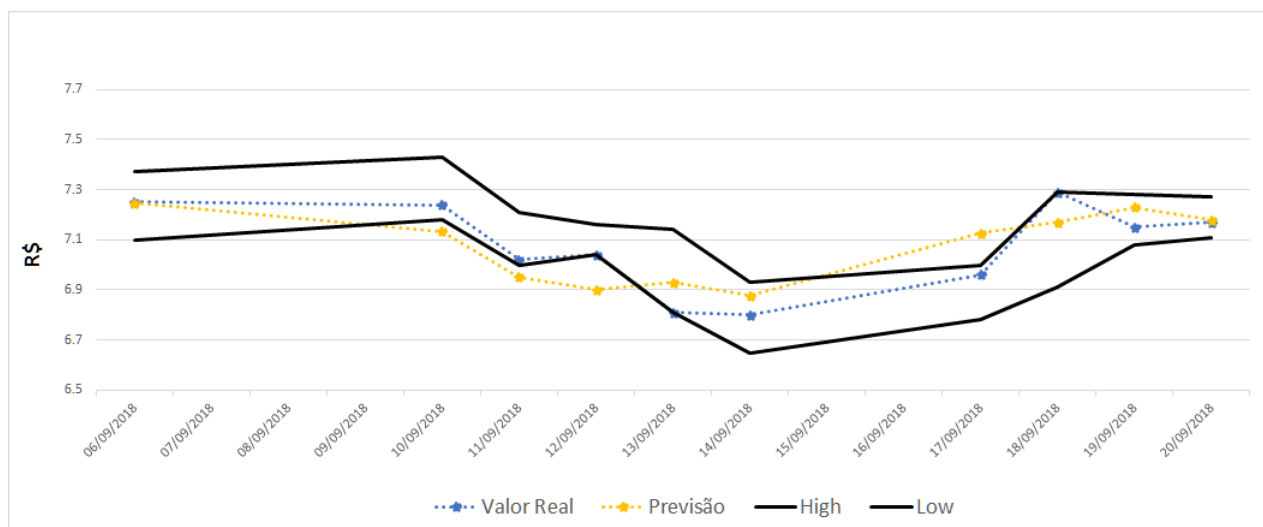
Tabela 13: Previsão da Série CMIG4

| Data a ser Prevista | Cotação Real | Valor Previsto | Erro | Predição da Tendência de Movimentos |
|---------------------|--------------|----------------|--------|-------------------------------------|
| 06/09/2018 | 7.25 | 7.24 | 0.07% | Sim |
| 10/09/2018 | 7.24 | 7.14 | 1.49% | Sim |
| 11/09/2018 | 7.02 | 6.95 | 1.07% | Sim |
| 12/09/2018 | 7.04 | 6.90 | 2.03% | Não |
| 13/09/2018 | 6.81 | 6.93 | -1.73% | Não |
| 14/09/2018 | 6.80 | 6.87 | -1.12% | Sim |
| 17/09/2018 | 6.96 | 7.13 | -2.34% | Sim |
| 18/09/2018 | 7.29 | 7.17 | 1.68% | Sim |
| 19/09/2018 | 7.15 | 7.23 | -1.11% | Não |
| 20/09/2018 | 7.17 | 7.18 | -0.14% | Não |

Fonte: Autoria própria.

Graficamente, o comportamento dos valores previstos pelo modelo podem ser vistos abaixo, na Figura 14.

Figura 14: Previsão da Série CMIG4



Fonte: Autoria Própria.

Aqui, ainda aceita-se a performance do modelo como adequado já que em 50% dos casos a previsão das tendências de movimento foram feitas com êxito e em 60% das vezes os dados estiveram dentro do intervalo de confiança.

6.4 Análise da Série CPLE6

6.4.1 Previsão com redes neurais recorrentes sem filtro de wavelets

Dos dados de treinamento, obtém-se:

Tabela 14: Melhores Topologias das RNAs para Previsão sem Filtro da Série CPLE6

| k passos à diante | Topologia | MAPE |
|---------------------|-----------|-------|
| 1 | 10 | 3.95% |
| 2 | 12 | 4.24% |
| 3 | 5 | 4.27% |
| 4 | 6 | 4.39% |
| 5 | 2 | 4.91% |
| 6 | 7 | 5.01% |
| 7 | 6 | 5.36% |
| 8 | 9 | 5.38% |
| 9 | 8 | 6.29% |
| 10 | 10 | 6.45% |

Fonte: Autoria própria.

Alimentando as redes com as cotações do dia 05/09/2018, tem-se:

Tabela 15: Previsão sem Filtro da Série CPLE6

| Data a ser Prevista | Cotação Real | Valor Previsto | Erro | Predição da Tendência de Movimentos |
|---------------------|--------------|----------------|--------|-------------------------------------|
| 06/09/2018 | 20.69 | 22.11 | -6.46% | Não |
| 10/09/2018 | 21 | 21.93 | -4.25% | Sim |
| 11/09/2018 | 20.57 | 21.95 | -6.32% | Não |
| 12/09/2018 | 20.55 | 21.93 | -6.24% | Não |
| 13/09/2018 | 20.05 | 21.94 | -8.65% | Não |
| 14/09/2018 | 20.35 | 21.93 | -7.23% | Não |
| 17/09/2018 | 21.01 | 21.98 | -4.44% | Sim |
| 18/09/2018 | 21.31 | 22.00 | -3.14% | Sim |
| 19/09/2018 | 21.34 | 22.03 | -3.03% | Sim |
| 20/09/2018 | 21.81 | 22.03 | -1.01% | Não |

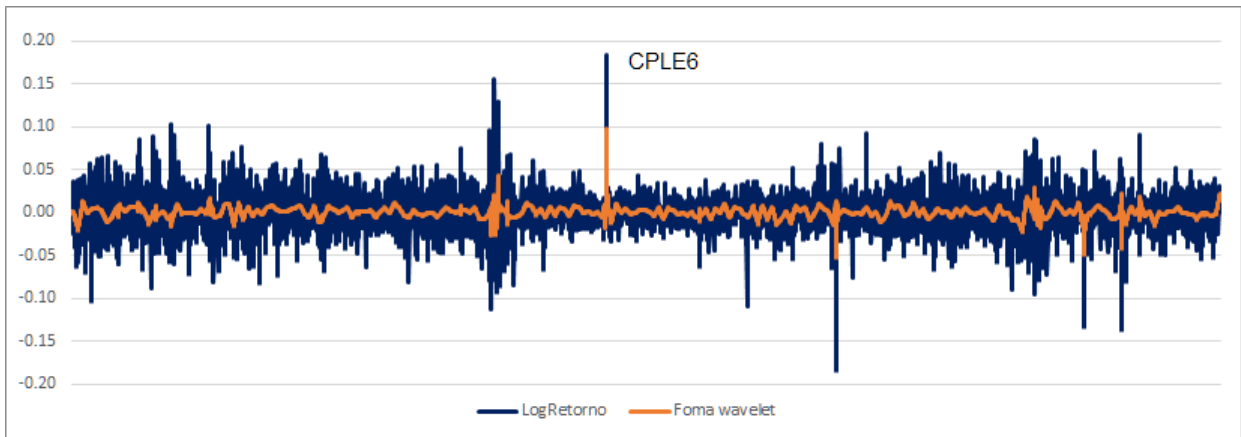
Fonte: Autoria própria.

Novamente, verifica-se que os valores obtidos das redes neurais sem os filtros de wavelets não apresentaram uma boa capacidade preditiva. Constatado-se mais uma vez que as previsões permaneceram praticamente estáticas.

6.4.2 Previsão com filtro de *wavelets* e redes neurais recorrentes

O uso do filtro de *wavelet* sobre a série do log retorno "limpa" a série, conforme a Figura 15:

Figura 15: Formas de Onda do Log Retorno da ação CPLE6 - com e sem filtro



Fonte: Autoria Própria.

As topologias mais adequadas para a série CPLE6 de acordo com cada passo de previsão, com seus respectivos erro MAPE, podem ser vistas na Tabela 12:

Tabela 16: Melhores Topologias das RNAs para Previsão da Série CPLE6

| k passos à diante | Topologia | MAPE |
|---------------------|-----------|-------|
| 1 | 13 | 0.33% |
| 2 | 11 | 0.41% |
| 3 | 5 | 0.53% |
| 4 | 13 | 0.61% |
| 5 | 9 | 0.69% |
| 6 | 13 | 0.76% |
| 7 | 11 | 0.82% |
| 8 | 12 | 0.87% |
| 9 | 5 | 0.93% |
| 10 | 9 | 0.98% |

Fonte: Autoria própria.

Utilizando-se dos mesmos atributos citados no subseção 6.1.2, a Tabela 17 traz as informações à despeito do desempenho preditivo:

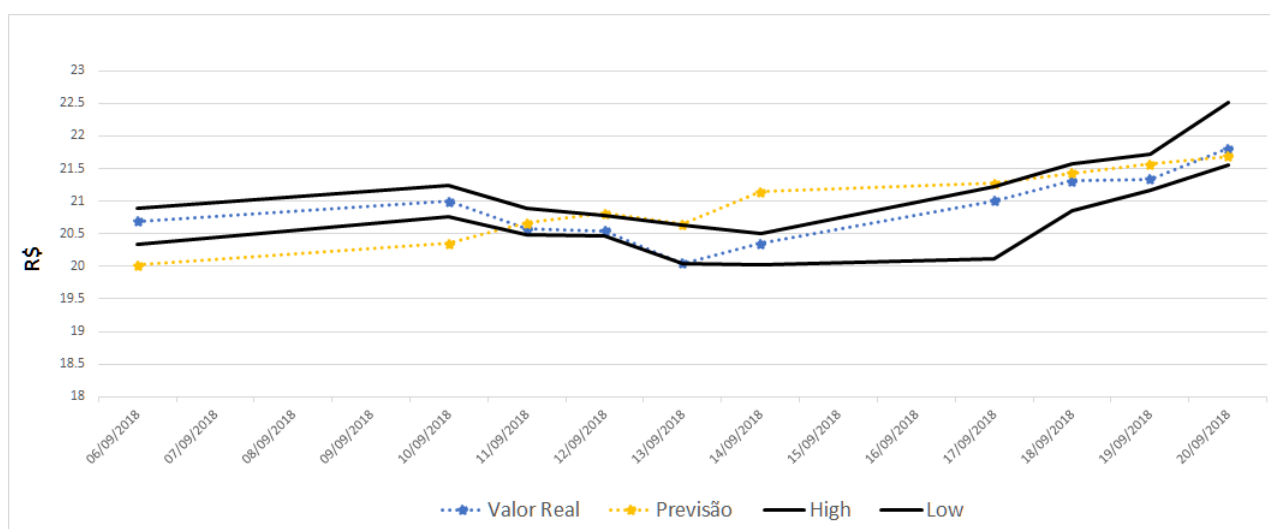
Tabela 17: Previsão da Série CPLE6

| Data a ser Prevista | Cotação Real | Valor Previsto | Erro | Predição da Tendência de Movimentos |
|---------------------|--------------|----------------|--------|-------------------------------------|
| 06/09/2018 | 20.69 | 20.37 | 0.07% | Não |
| 10/09/2018 | 21 | 20.41 | 2.89% | Sim |
| 11/09/2018 | 20.57 | 20.66 | 1.07% | Não |
| 12/09/2018 | 20.55 | 20.82 | 2.03% | Não |
| 13/09/2018 | 20.05 | 20.65 | -1.73% | Sim |
| 14/09/2018 | 20.35 | 21.14 | -1.12% | Sim |
| 17/09/2018 | 21.01 | 21.27 | -2.34% | Sim |
| 18/09/2018 | 21.31 | 21.42 | 1.68% | Sim |
| 19/09/2018 | 21.34 | 21.56 | -1.11% | Sim |
| 20/09/2018 | 21.81 | 21.69 | -0.14% | Sim |

Fonte: Autoria própria.

Graficamente, o comportamento dos valores previstos pelo modelo podem ser vistos abaixo, na Figura 16.

Figura 16: Previsão da Série CPLE6



Fonte: Autoria Própria.

Novamente, o desempenho do modelo combinado de filtro de *wavelet* e redes neurais pode ser classificada como satisfatória, já que em 70% dos casos a previsão das tendências de movimento foram feitas com êxito e em 80% das vezes os dados estiveram dentro do intervalo de confiança.

6.5 Considerações Parciais

Foi verificado que o desempenho do uso combinado de filtros de volatilidade com redes neurais produzem melhores previsões, quando comparado a modelos que não fazem

uso destes filtros.

A introdução do filtro de *wavelet* possibilitou que a rede conseguisse abstrair as informações essenciais das séries temporais originais, culminando em predições que na maioria das vezes foram capazes de detectar as inversões e as tendências de movimento da série temporal real. Tal fato foi observado em uma considerável redução do erro MAPE de treinamento e também do erro percentual das previsões obtidas. Ainda, possibilitou que o modelo estimador tivesse erro com média próxima de zero e desvio padrão pequeno (com a maioria dos testes dentro do intervalo de dois desvios), como pode ser visto no Apêndice (onde k representa o passo da previsão) .

Além disso, as topologias de redes que mostraram-se mais recorrentes, quando utilizadas em conjunto com o filtro de volatilidade, foram as que tinham 4 camadas (duas camadas escondidas, além da entrada e da saída), correspondendo a 65% dos casos. As topologias com 5 camadas, apresentaram melhor desempenho em 22,5% dos eventos e as com 3 camadas em 12,5%. Não foi observado uma tendência de topologia ideal para cada passo k da previsão.

7 Conclusão

Este trabalho de conclusão de curso teve por objetivo estudar a realização das previsões de séries temporais financeiras com o uso combinado de filtros de volatilidade e redes neurais artificiais, visando a redução do ruído nos dados e consequente melhora na previsão. A hipótese de que os filtros de volatilidade pudessem de fato resultar em uma melhora no desempenho preditivo das redes neurais foi confirmada em todos os testes realizados no Capítulo 6, onde foi feita a comparação entre as previsões de sinais relativos aos log-retornos em duas etapas: sem serem tratados e com o tratamento proposto.

Os filtros de volatilidade via funções de *wavelets* mostram-se eficientes não só por serem capazes de reduzir o erro das previsões, como evidenciado nas tabelas da Capítulo 6, através da atenuação do ruído, mas também por auxiliarem as redes neurais a abstraírem os componentes fundamentais das séries temporais, possibilitando que suas previsões apresentem boa assertividade em relação aos momentos de inversão de movimento das séries. Além disso, foi verificado que dependendo do tamanho k do passo a ser feita a previsão e do tipo da série temporal, diferentes topologias apresentam melhor capacidade de aprendizado e generalização.

Os resultados obtidos corroboram para a difusão do emprego de ferramentas neuro-estatísticas voltadas ao gerenciamento de risco e para tomadas de decisão e planejamento gerencial e/ou estratégico, uma vez que a vasta gama de modelos matemáticos e estatísticos presentes na literatura não foram capazes de explicar por sí só crises no mercado financeiro e falhas nos modelos de previsão.

Referências

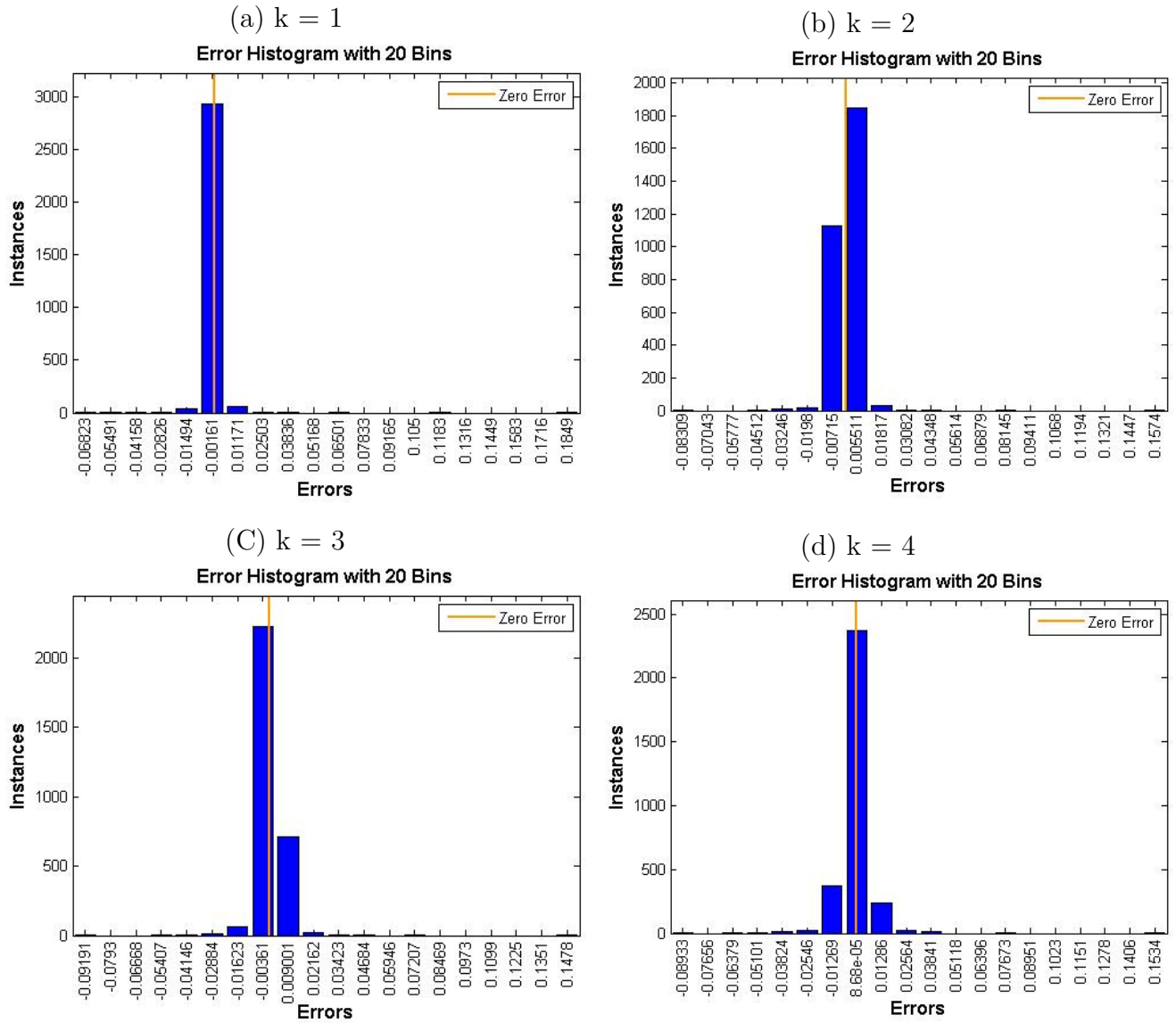
- AHLBURG, D. **Error measures and the choice of a forecast method.** [S.l.]: international Journal os Dorecasting, 1992.
- ANDERSEN, T.; BOLLERSLEV, T.; CHRISTOFFERSEN, P. **Volatility and correlation forecasting.** [S.l.: s.n.], 2006.
- ARAUJO, J. Tese de Mestrado, Universidade Pontifica Catolica do Rio de Janeiro. **Aplicação de SRV e ESN à Previsão de Series do Mercado de Seguros.** Rio de Janeiro: [s.n.], 2016.
- AZOFF, E. **Neural network time series forecasting of financial markets.** [S.l.]: John Wiley & Sons, 1994.
- BREIMAN, L. **Statistical modeling: two cultures. Quality control and applied statistics.** [S.l.: s.n.], 2003.
- CARVALHO, T. Tese de Doutorado, Universidade Federal de Minas Gerais. **Análise e Implementação de Metodos Aplicados aos Sinais Variantes no Tempo Presentes em Sistemas Eletricos de Potência.** Belo Horizonte: [s.n.], 2013.
- CASTRO, M. Tese de Doutorado, Universidade de Campinas. **Predição Não-Linear de Series Temporais Usando Redes Neurais RBF por Decomposição em Componentes Principais.** Campinas: [s.n.], 2001.
- CYBENKO, G. **Approximation by Superpositions of a Sigmoidal Function.** [S.l.: s.n.], 1989. 303-314 p. (Mathematics of Control, Signal and Sytems).
- ELMAN, J. **Finding Structure in Time. Cognitive Science.** [S.l.: s.n.], 1990.
- ENDERS, W. **Applied econometric time series.** second. [S.l.]: John Wiley & Sons, 2004.
- GENÇAY, R.; SELÇUK, F.; WHITCHER, B. **An introduction to wavelets and other filtering methods in finance and economics.** [S.l.]: New York: Academic Press, 2002.
- HAYKIN, S. **Redes Neurais: princípios e prática.** second. [S.l.]: BR: Bookman, 2001.
- JORION, P. **Value at Risk.** [S.l.]: McGraw-Hill, 2007.
- KANTZ, H.; SCHREIBER, T. **Nonlinear time series analysis.** Cambridge University: [s.n.], 2004.
- LIMA, F. Livre Docência, Universidade de São Paulo. **Modelos de Previsão de Series Temporais Financeiras com Combinação de Filtros de Kalman e Wavelets.** Ribeirão Preto: [s.n.], 2011.
- MANDELBROT, B.; HUDSON, R. **Mercados Financeiros fora de controle: a teoria dos fractais explicando o comportamento dos mercados.** [S.l.]: Elsevier, 2004.

- MATHIAS, I.; DIAS, A. **Modulo de Validação Cruzada para Treinamento de Redes Neurais Artificiais com Algoritmos Backpropagation e Resilient Propagation**. [S.l.: s.n.], 2007.
- MEDEIROS, M.; RECH, G. **Building Neural Network Models for Time Series: A Statistical Approach**. [S.l.]: Journal of Forecasting, 2006.
- MISITI, M. **Wavelets and their applications**. [S.l.]: John Wiley, 2007.
- MORETTIN, P. **Ondas e ondaletas**. São Paulo: Edusp, 1999.
- MORETTIN, P.; TOLOI, C. **Análise de Séries Temporais**. São Paulo: Blucher, 2004.
- OLIVEIRA, M. Tese de Doutorado, Universidade de São Paulo. **Aplicação de Redes Neurais Artificiais na Analise de Series Temporais Econômico - Financeiras**. São Paulo: [s.n.], 2007.
- PERCIVAL, D.; WALDEN, A. **Wavelets methods for time series analysis**. Cambridge: [s.n.], 2006.
- POLIKAR, R. **The wavelet tutorial**. [S.l.]: Disponível em: <http://users.rowan.edu/polikar/WTtutorial.html>, 2010.
- SCHOPF, E. Tese de Mestrado, Universidade Federal do Rio Grande do Sul. **Metodo Neuro-Estatístico para Predição de Series Temporais Ruidosas**. Porto Alegre: [s.n.], 2007.
- SOMERS, M.; SCARBOROUGH, D. **Neural Networks in Organizational Research: Applying Pattern Recognition to the Analysis of Organizational Behavior**. [S.l.: s.n.], 2006.
- TSAY, R. **Analysis of financial time series. Wiley series in probability and statistics**. [S.l.: s.n.], 2005.
- WAN, E.; R.MERWE. **Efficient Derivative-Free Kalman Filters for Online Learning**. [S.l.]: Proceedings of the European Symposium on Artificial Neural Networks, 2001.

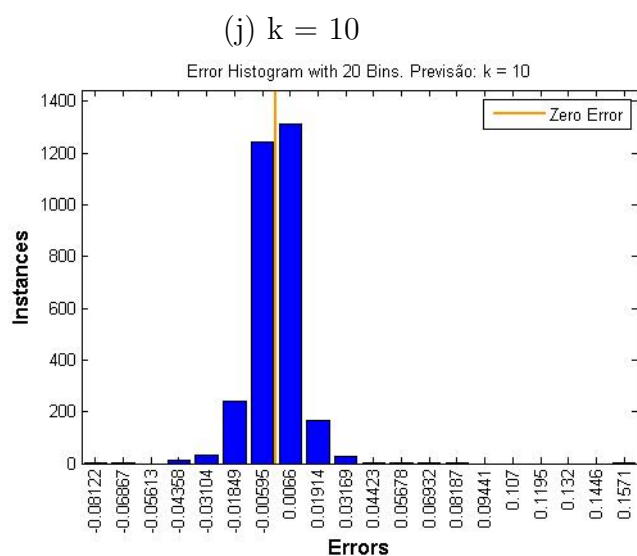
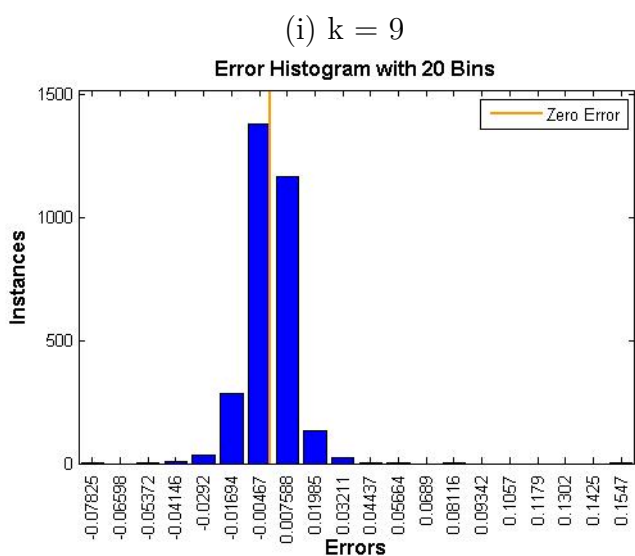
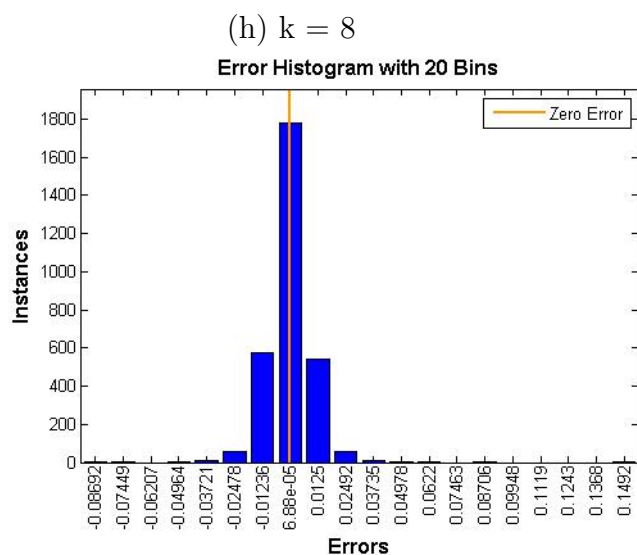
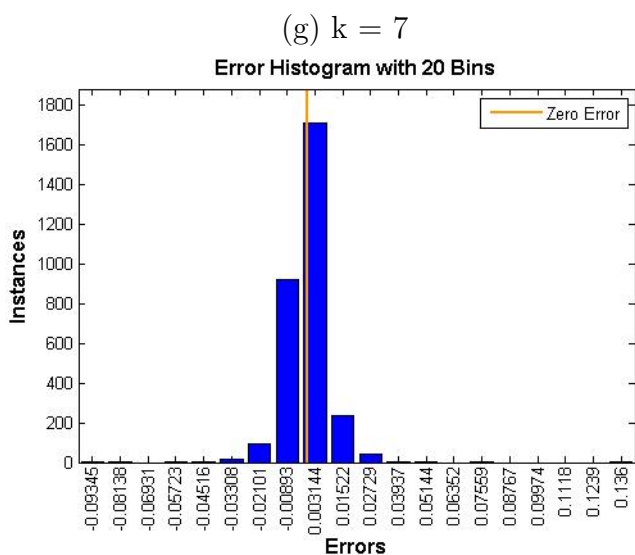
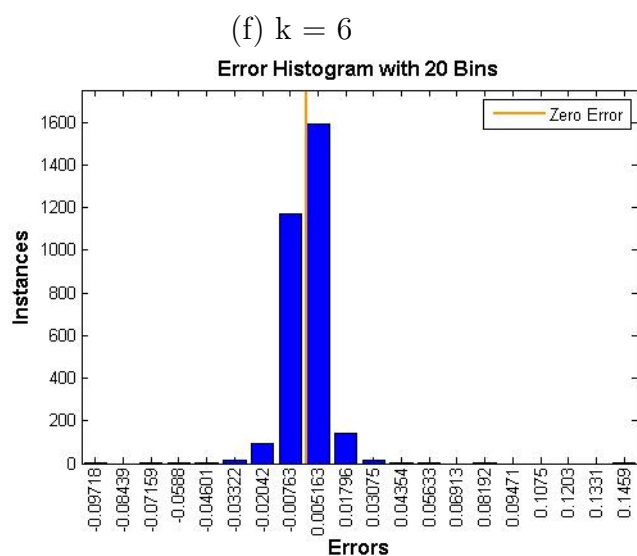
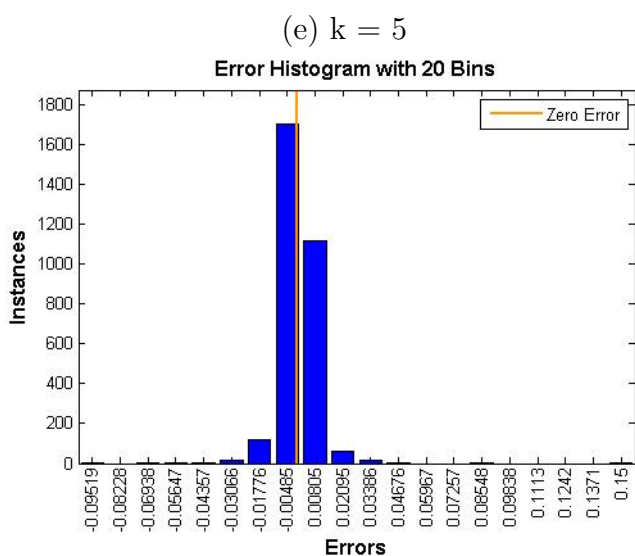
Apêndices

APÊNDICE A – Histograma dos Erros de Treinamento das Redes Neurais na Presença do Filtro de *Wavelet* - ELET3

Figura 17: Histograma do Erro de Treinamento com filtro de *wavelet* - ELET3.



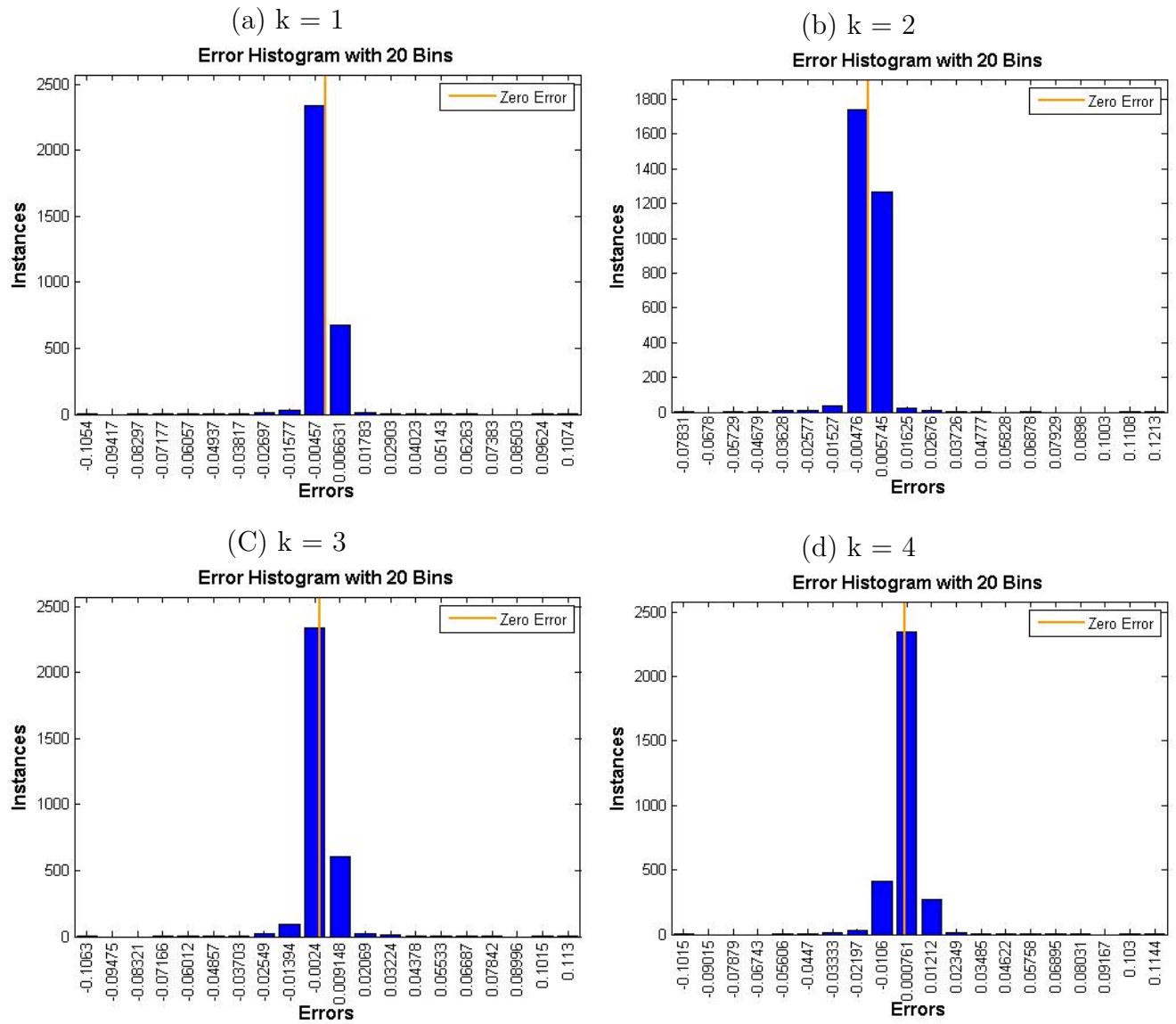
Fonte: Autoria própria.



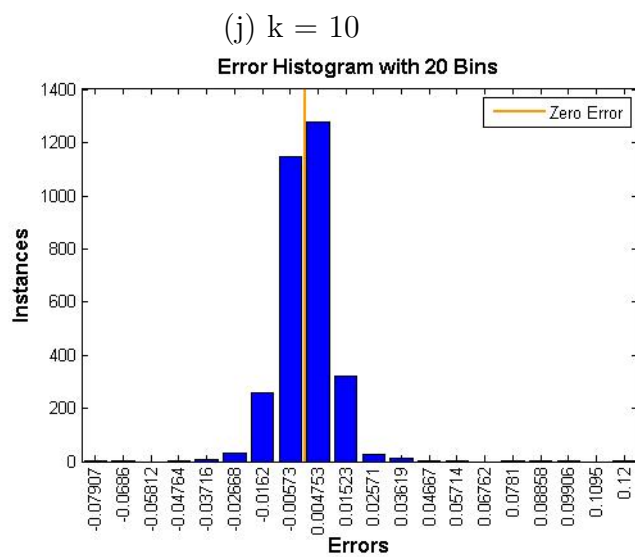
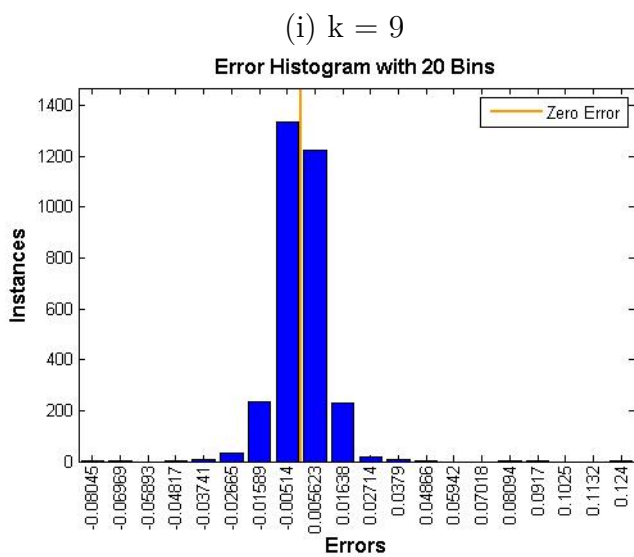
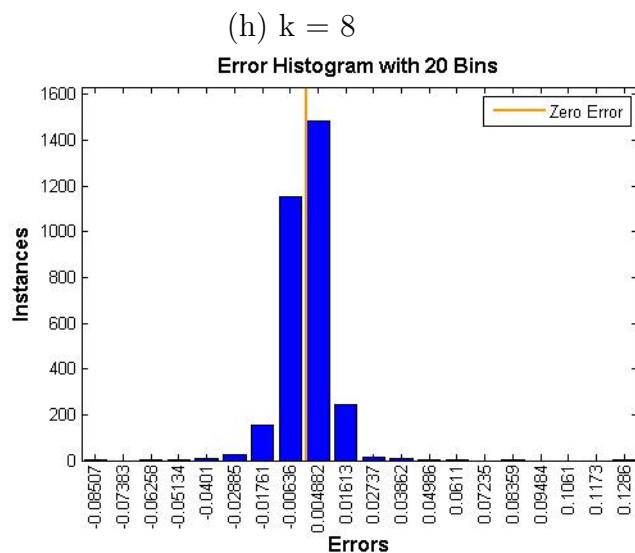
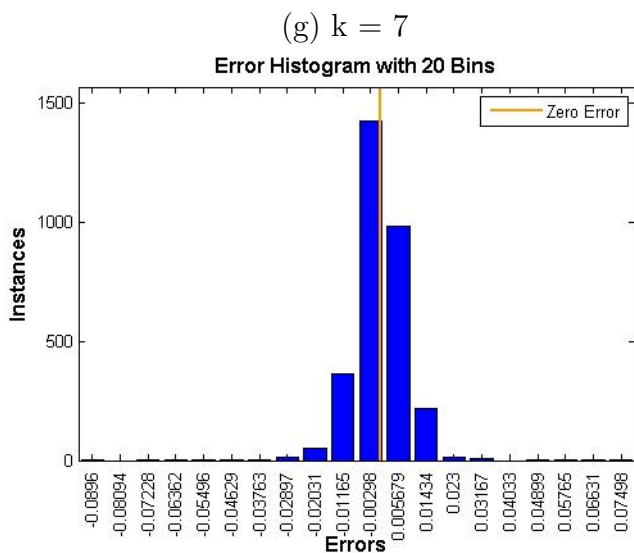
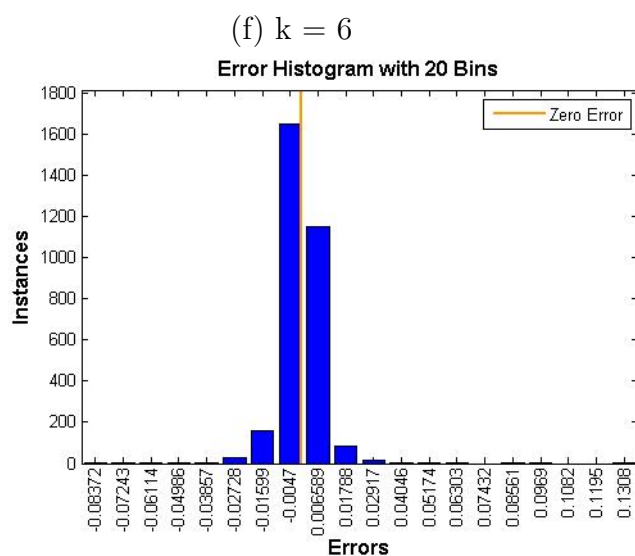
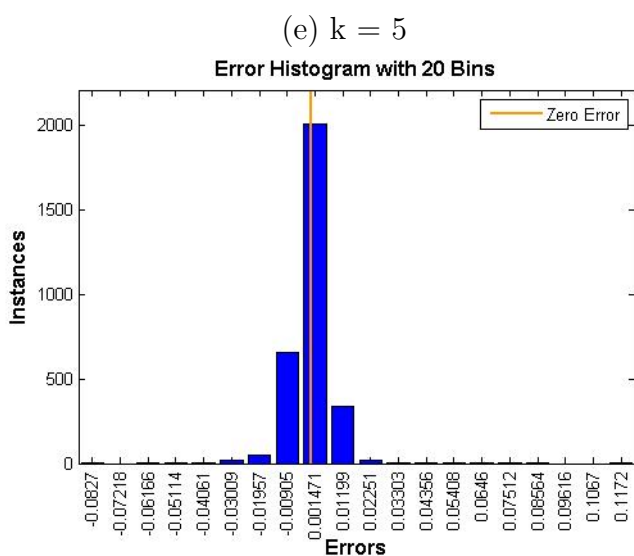
Fonte: Autoria própria.

APÊNDICE B – Histograma dos Erros de Treinamento das Redes Neurais na Presença do Filtro de *Wavelet* - CPFE3

Figura 18: Histograma do Erro de Treinamento com filtro de *wavelet* - CPFE3.



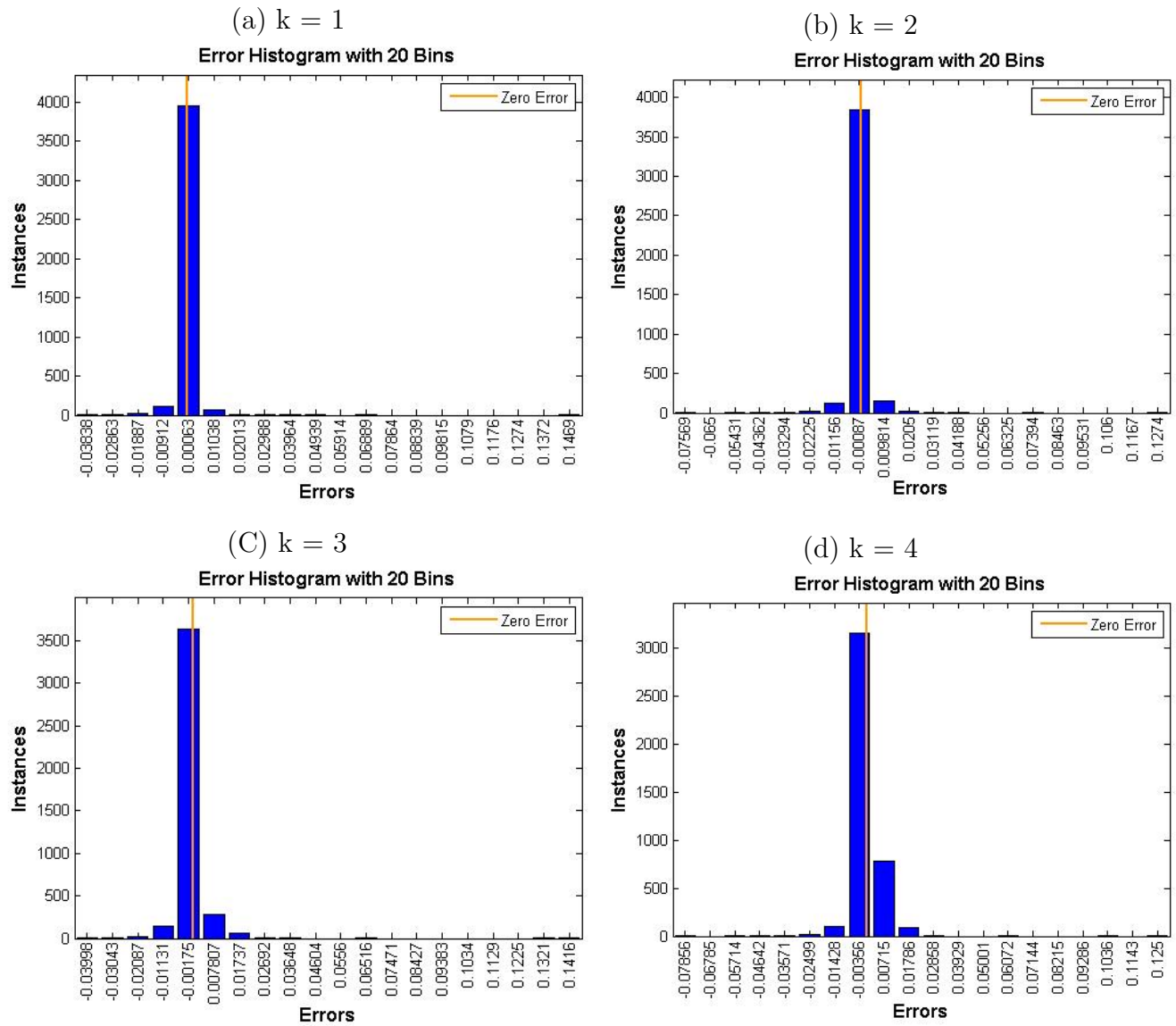
Fonte: Autoria própria.



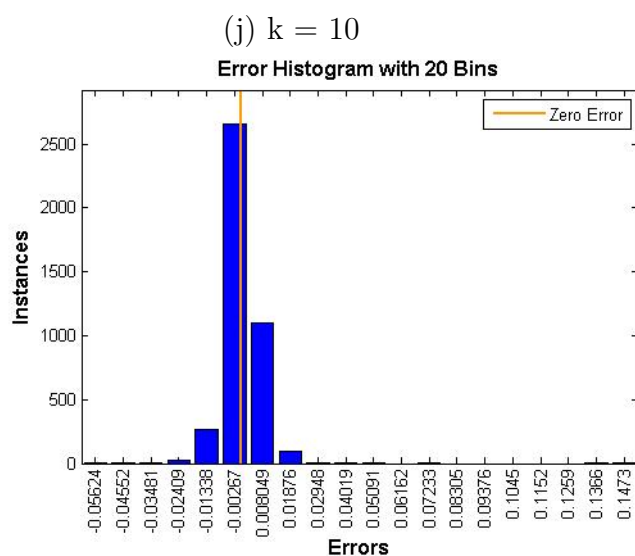
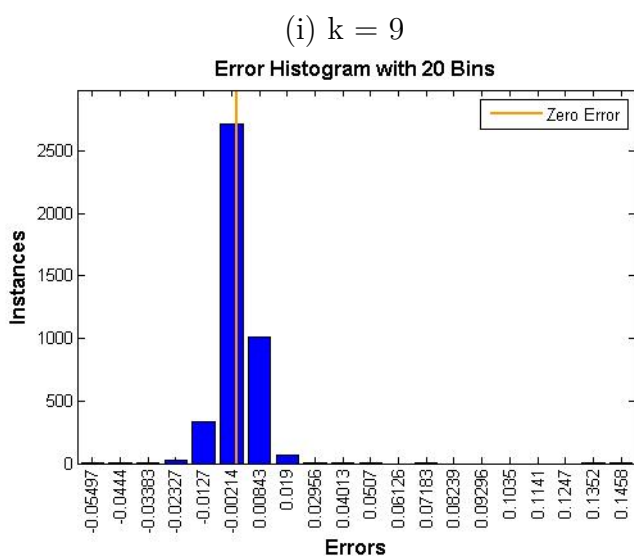
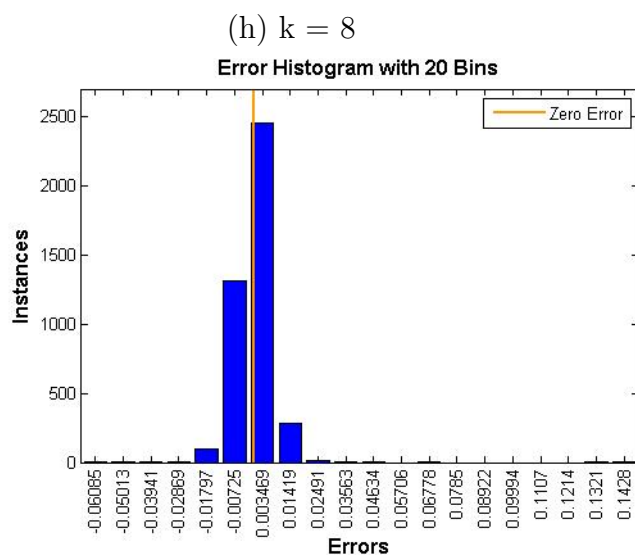
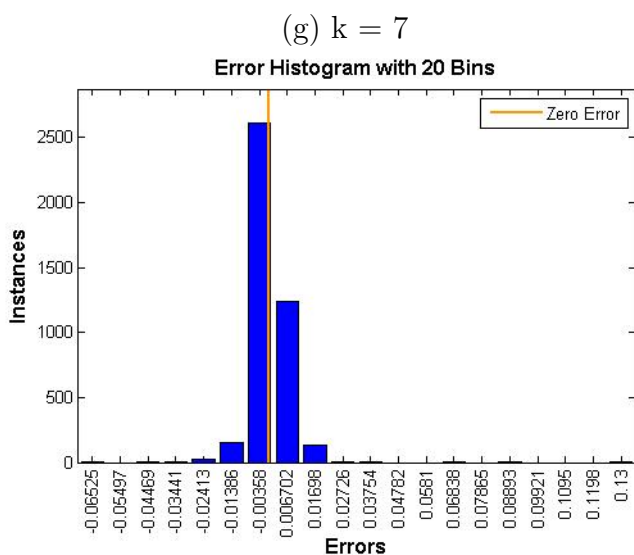
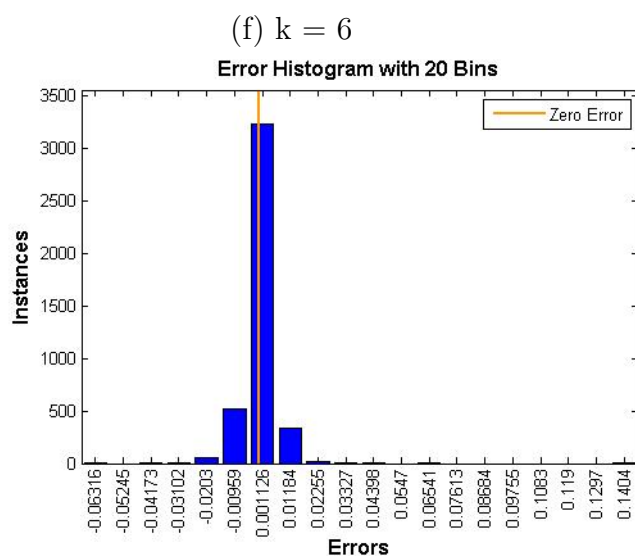
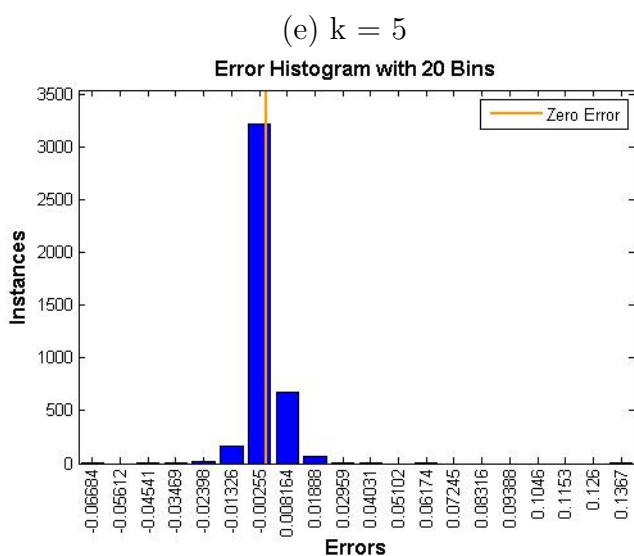
Fonte: Autoria própria.

APÊNDICE C – Histograma dos Erros de Treinamento das Redes Neurais na Presença do Filtro de *Wavelet* - CMIG4

Figura 19: Histograma do Erro de Treinamento com filtro de *wavelet* - CMIG4.



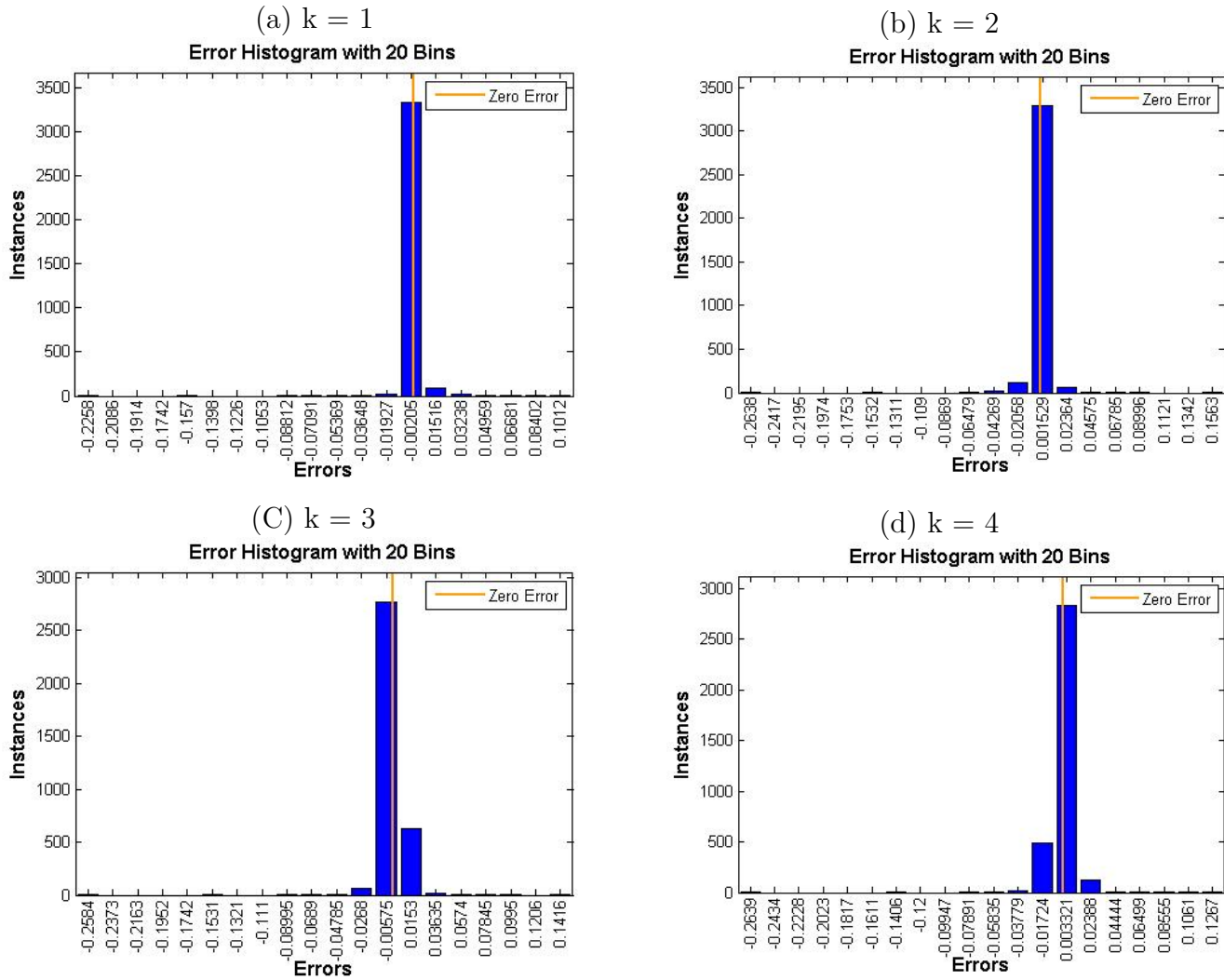
Fonte: Autoria própria.



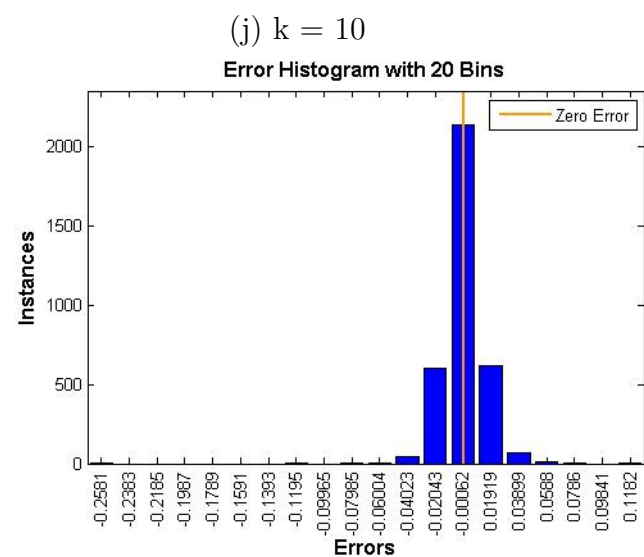
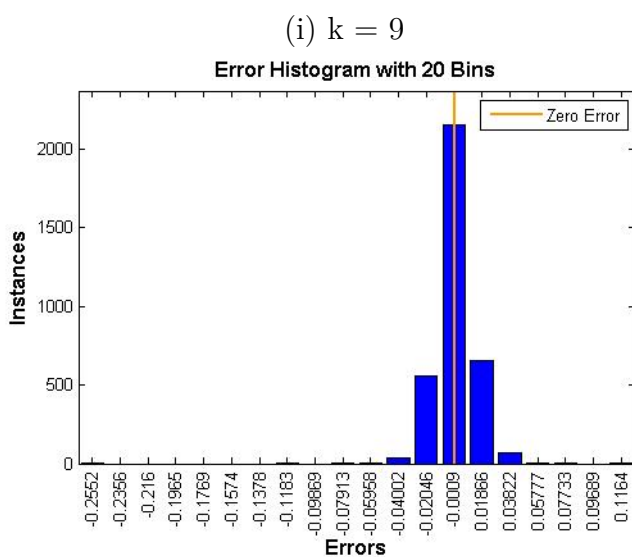
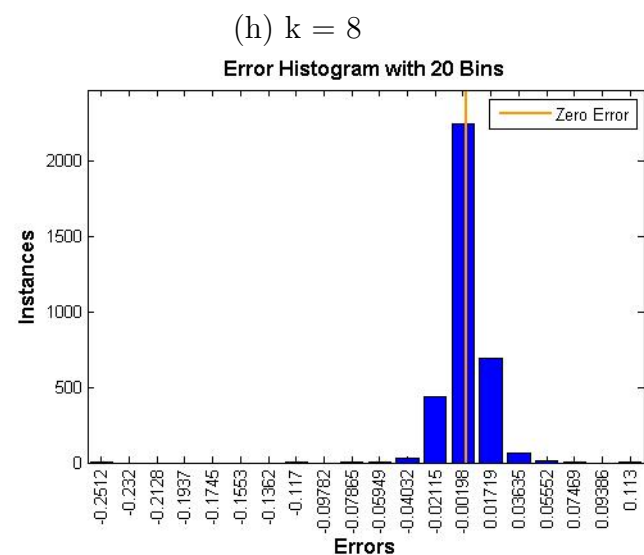
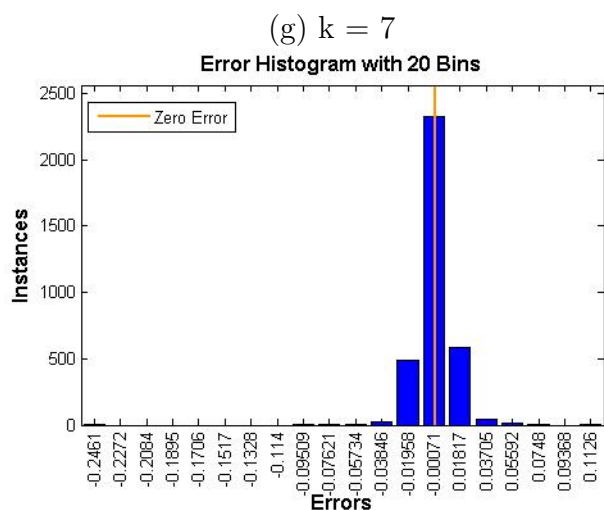
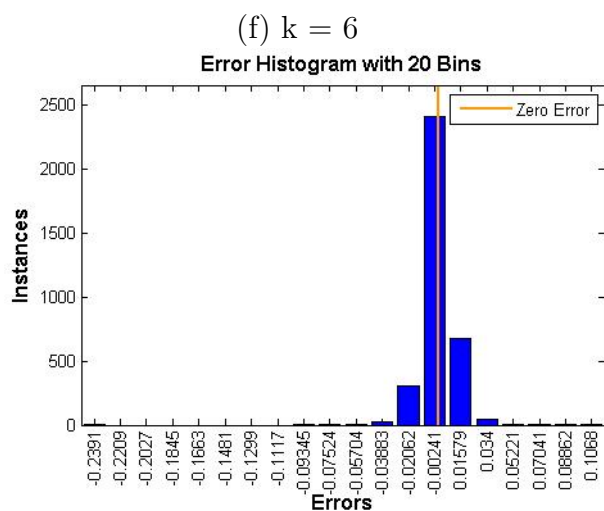
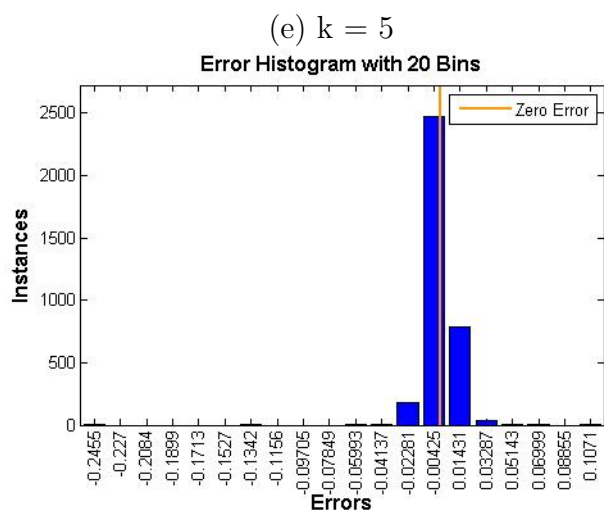
Fonte: Autoria própria.

APÊNDICE D – Histograma dos Erros de Treinamento das Redes Neurais na Presença do Filtro de *Wavelet* - CPLE6

Figura 20: Histograma do Erro de Treinamento com filtro de *wavelet* - CPLE6.



Fonte: Autoria própria.



Fonte: Autoria própria.