

UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA E MODELAGEM
QUANTITATIVA

**NÚMERO DE NEGÓCIOS E VOLUME FINANCEIRO
NA PREVISIBILIDADE DO IBOVESPA:
UMA APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS**

MONOGRAFIA DE ESPECIALIZAÇÃO

Paulo Sergio Ceretta

Santa Maria, RS, Brasil
2008

**NÚMERO DE NEGÓCIOS E VOLUME FINANCEIRO
NA PREVISIBILIDADE DO IBOVESPA:
UMA APLICAÇÃO DE REDES NEURAS ARTIFICIAIS**

por

Paulo Sergio Ceretta

Monografia apresentada ao Curso de Especialização do Programa de Pós-Graduação em Estatística e Modelagem Quantitativa, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Especialista em Estatística e Modelagem Quantitativa**

Orientador: Prof. Dr. Ivanor Müller

Santa Maria, RS, Brasil

2008

**Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Programa de Pós-Graduação em Estatística e Modelagem
Quantitativa**

A Comissão Examinadora, abaixo assinada,
aprova a Monografia de Especialização

**NÚMERO DE NEGÓCIOS E VOLUME FINANCEIRO NA
PREVISIBILIDADE DO IBOVESPA: UMA APLICAÇÃO DE REDES
NEURAS ARTIFICIAIS**

elaborada por

Paulo Sergio Ceretta

**como requisito parcial para obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa**

COMISSÃO EXAMINADORA:

Ivanor Müller, Dr.
(Presidente/Orientador)

Luis Felipe Dias Lopes, Dr. (UFSM)

Fernando Hepp Pulgati, Dr. (UFSM)

Santa Maria, Outubro de 2008.

AGRADECIMENTOS

Utilizo este espaço para transmitir meus sinceros agradecimentos ao meu orientador, Ivanor Müller, pelo conhecimento, transparência, dedicação e acompanhamento. Agradeço também a todos os professores do Programa de Pós-Graduação em Estatística e Modelagem Quantitativa pelo empenho e competência na condução do ensino; a secretária Sandra, pela simpatia, simplicidade e disponibilidade para ajudar alunos e professores; aos colegas, que de uma forma ou de outra contribuíram significativamente para meu processo de aprendizado; à minha família, pelo afeto, confiança e apoio, meu muito obrigado. Agradecimento especial a Graciele, pelo acompanhamento ao longo do curso. Por fim, agradeço à Universidade Federal de Santa Maria, por disponibilizar aos seus alunos um ensino gratuito e de qualidade.

RESUMO

Monografia de Especialização
Programa de Pós-Graduação em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria, RS, Brasil

NÚMERO DE NEGÓCIOS E VOLUME FINANCEIRO NA PREVISIBILIDADE DO IBOVESPA: UMA APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS

Autor: Paulo Sergio Ceretta
Orientador: Dr. Ivanor Müller
Data e Local de Defesa: Santa Maria, Dezembro de 2008.

Este estudo tem por objetivo investigar a utilização de informações sobre o número de negócios e o volume financeiro negociado como variáveis explanatórias que melhoram o desempenho na previsibilidade do retorno do Ibovespa. São utilizadas informações diárias da carteira teórica do Ibovespa no período compreendido entre 23/01/2006 até 18/04/2008. São aplicados três modelos de previsão baseados em Redes Neurais Artificiais com duas camadas intermediárias e funções de ativação sigmóides. Na fase de treinamento, as RNAs são ajustadas automaticamente seguindo um critério de erro pré-definido tendo como base uma amostra de 758 observações. Na fase de previsão, as RNAs são acompanhadas 50 passos a frente para verificação de desempenho. Os resultados obtidos permitem afirmar que a inclusão de informações sobre número de negócios e volume de negócio melhora o ajuste das RNAs na fase de treinamento, mas a melhoria não se mantém na fase de previsão.

Palavras-chave: Mercado de ações; Eficiência de mercado; Redes neurais artificiais.

ABSTRACT

Monografia de Especialização
Programa de Pós-Graduação em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria, RS, Brasil

NUMBER OF BUSINESS AND FINANCIAL VOLUME IN PREDICTABILITY OF IBOVESPA: AN APPLICATION OF ARTIFICIAL NEURAL NETWORKS

Author: Paulo Sergio Ceretta

Advisor: Dr. Ivanor Muller

Date and Place of Defense: Santa Maria, December, 2008.

This study focused on the use of information on the number of business and financial volume traded as explanatory variables that improve the performance in the predictability of Ibovespa return index. They are used daily information from the Ibovespa theoretical portfolio during the period from 23/01/2006 until 18/04/2008. Three models are applied to forecast based on Artificial Neural Networks with two layers of features and intermediate sigmoid activation. At training, the ANNs are adjusted automatically according to a criterion of error pre-defined based on a sample of 758 points. At the time of the forecast, the tracks ANNs are 50 steps ahead for verification of performance. The results show that the inclusion of information on number of business and volume of business improves the fit of ANNs in the phase of training, but the improvement was not maintained during the forecast.

Key-words: Stock markets; Market efficiency; Artificial neural networks.

LISTA DE ILUSTRAÇÕES

FIGURA 1 –	Séries ruído branco	23
FIGURA 2 –	Representação do neurônio humano	26
FIGURA 3 –	Representação do neurônio artificial	26
FIGURA 4 –	RNA <i>feedforward</i> (2:2:1)	29
FIGURA 5 –	RNA <i>feedforward</i> (2:2:2:1)	29
FIGURA 6 –	RNA recorrente (2:2:1)	30
FIGURA 7 –	Série temporal dos preços da carteira teórica da Bovespa	32
FIGURA 8 –	Série temporal do número de negócios	33
FIGURA 9 –	Série temporal do volume financeiro	33
FIGURA 10 –	Estrutura das RNAs a serem aplicadas	35
FIGURA 11 –	Estrutura da RNA com 5 variáveis de entrada	40
FIGURA 12 –	Treinamento da RNA(5:3:3:1)	41
FIGURA 13 –	Estrutura da RNA com 10 variáveis de entrada (10:3:3:1)	43
FIGURA 14 –	Treinamento da RNA (10:3:3:1)	45
FIGURA 15 –	Estrutura da RNA com 15 variáveis de entrada (15:3:3:1)	47
FIGURA 16 –	Treinamento da RNA (15:3:3:1)	48

LISTA DE TABELAS

TABELA 1 – Estatísticas descritivas das variáveis	38
TABELA 2 – Pesos sinápticos da RNA (5:3:3:1)	39
TABELA 3 – Resultados da RNA (5:3:3:1)	40
TABELA 4 – Pesos sinápticos da RNA (10:3:3:1)	42
TABELA 5 – Resultados da RNA (10:3:3:1)	44
TABELA 6 – Pesos sinápticos da RNA (15:3:3:1)	46
TABELA 7 – Resultados da RNA (15:3:3:1)	48

LISTA DE ANEXOS

ANEXO A - Programação do modelo RNA com cinco variáveis de entrada	55
ANEXO B - Programação do modelo RNA com dez variáveis de entrada	57
ANEXO C - Programação do modelo RNA com quinze variáveis de entrada	59

SUMÁRIO

1 INTRODUÇÃO	11
1.1 Preliminares	11
1.2 Objetivos	13
1.2.1 Objetivo Geral	13
1.2.2 Objetivos Específicos	13
1.3 Justificativa	13
2 PREVISIBILIDADE, SERIES TEMPORAIS E REDES NEURAIIS	15
2.1 Previsibilidade no mercado financeiro	15
2.2 Séries temporais	16
2.2.1 Estacionariedade	17
2.2.1.1 Coeficiente de correlação	19
2.2.1.2 Função de autocorrelação	19
2.2.1.3 Teste de Ljung e Box	21
2.2.2 Ruído branco	22
2.3 Redes neurais artificiais	24
2.3.1 Arquitetura de rede	28
2.3.1.1 Redes neurais alimentadas adiante com camada única	28
2.3.1.2 Redes neurais alimentadas adiante com múltiplas camadas	29
2.3.1.3 Redes recorrentes	30
3 ASPECTOS METODOLÓGICOS	32
3.1 Amostra e dados	32
3.2 Modelagem	34
3.3 Treinamento e teste	36
3.4 <i>Software</i>	36

	12
4 RESULTADOS OBTIDOS	38
4.1 Modelo com informações sobre os retornos	39
4.2 Modelo com informações sobre os retornos e número de negócios	42
4.3 Modelo com informações sobre os retornos, número de negócios e volume Financeiro	45
5 CONCLUSÃO	50
REFERÊNCIAS BIBLIOGRÁFICAS	52
ANEXOS	55

1 INTRODUÇÃO

No estudo de séries temporais financeiras, diversos aspectos podem ser analisados, mas a questão central é a possibilidade de adquirir conhecimento sobre padrões que se desenvolvem na passagem do tempo.

Esses padrões são identificados através de ajustes de modelos econométricos aos eventos ao longo do tempo. Posteriormente, os modelos econométricos ajustados são utilizados para realizar uma estimativa de resultados que provavelmente deverão ocorrer no futuro. De forma geral, as estimativas dos modelos econométricos tem por objetivo extrapolar o domínio temporal conhecido.

1.1 Preliminares

Em Finanças, o conhecimento prévio do comportamento de valores em uma série temporal é algo que torna possível a formulação de estratégias operacionais mais consistentes e coerentes ao risco do negócio.

O reconhecimento prévio da existência de padrões no comportamento da volatilidade de um ativo, por exemplo, possibilita que se estabeleça um contrato de *hedging* impedindo que a empresa ou investidor fique exposto a elevada instabilidade do mercado. Uma consequência direta dessa proteção é impedir perdas de riqueza devido à oscilação excessiva de preços.

Uma prática comum, nos estudos de séries temporais financeiras, tem sido a aplicação de modelos estatísticos de regressão. Os modelos de previsão de séries temporais mais amplamente utilizados são os autoregressivos (AR), médias móveis (MA) e autoregressivos com médias móveis (ARMA).

Redes Neurais Artificiais (RNAs), do inglês *Artificial Neural Networks* (ANNs) tem recebido atenção crescente, tanto no campo teórico como no prático, ver, por exemplo, Kuan e White (1994) e Qi (1996). Avanços no poder de processamento de computadores associado

ao desenvolvimento de uma ampla variedade de *softwares* de aplicação especializada contribuem de maneira decisiva para o aprimoramento das RNAs.

Redes Neurais Artificiais têm sido aplicadas na solução de problemas complexos que anteriormente eram enfrentados com abordagens mais simples e menos abrangentes.

Como resultado desse cenário tem-se, hoje, uma ampla variedade de formulações baseadas em RNAs aplicada em diversas áreas. O sucesso dessas aplicações tem levado a um grande crescimento de literatura disponível na área de administração, como exemplo pode-se citar: Franses e Dijk (2003) onde são apresentados modelos de séries temporais não lineares em finanças; Zhang (2004) que se constitui como uma das principais referências em aplicações de Redes Neurais Artificiais em previsão de negócios; McNelis (2005) onde é focado Redes Neurais Artificiais em finanças e Kamruzzaman e Sarker (2006) que apresentam diversas aplicações de Redes Neurais Artificiais em finanças e manufatura.

A essência desta investigação repousa sobre o aspecto da estimação de valores futuros da série de retornos do Ibovespa com base em informações passadas. A dependência de uma variável endógena no tempo t em relação a outras variáveis exógenas no tempo $t-i$ ($i=1, \dots, T$) é chamado de dependência nas defasagens em tempo discreto.

A utilização de informações defasadas para a obtenção de estimativas futuras é bastante óbvio, visto que se está procurando antever o comportamento de uma variável no tempo t com base em informações já ocorridas, no tempo $t-i$. A parte da questão cronológica, existem alguns aspectos relevantes que sustentam a idéia de observar o passado para estimar o futuro, são motivos psicológicos, motivos associados a incerteza e motivos institucionais, dentre outros.

Os motivos associados a aspectos psicológicos são baseados no fato de que as pessoas não conseguem alterar instantaneamente seus padrões de comportamento. Por exemplo, pessoas que se tornam milionárias de forma rápida, continuam mantendo diversos padrões de comportamento, pois não é possível assimilar um novo padrão de comportamento dado que esse novo padrão não era conhecido, também é difícil abandonar estilos de decisão desenvolvidos e solidificados ao longo de muito tempo.

A incerteza é outro aspecto que pode interferir na decisão no tempo t . Por exemplo, o aumento de preço de um ativo, do ponto de vista de um investidor pode ser uma incerteza, pois este não consegue assimilar de imediato se a elevação do preço será considerada normal pelo mercado ou tenderá a reverter nos próximos dias.

Por fim, existem aspectos institucionais que impedem um perfeito compasso dos eventos. Por exemplo, após a queda do preço de um ativo, ele se torna uma opção muito atraente no mercado, porém quando os investidores decidem por sua aquisição precisam se desfazer de parte de seus investimentos atuais. O processo de conversão de parte dos investimentos atuais em recursos de liquidez imediata para aplicação no ativo subavaliado, geralmente requer um período de um ou dois dias. Nesse período os recursos ficam “presos”, mas ao serem liberados seguirão um comportamento ocorrido em $t-i$.

1.2 Objetivos

O propósito da aplicação de RNAs em séries temporais financeiras advém da necessidade de se verificar se a estrutura mais flexível das redes neurais possibilita que se tenha uma estimativa mais consistente para estimativas futuras das variações de preço do índice da BOVESPA.

Embora muitas tenham sido as pesquisas desenvolvidas sobre a possibilidade de previsão do comportamento das variações de preço da carteira do Ibovespa, pouco tem sido o consenso sobre como e quais as variáveis a serem utilizadas nesse processo. Dentro do contexto do mercado financeiro, este estudo tem por objetivo analisar a utilidade da inclusão de informações sobre número de negócios e volume financeiro na previsão das variações do Ibovespa através das Redes Neurais Artificiais.

1.2.1 Objetivo Geral

Verificar se a inclusão de informações sobre o número de negócios e o volume financeiro é capaz de possibilitar melhorias nos modelos de estimação do futuro comportamento das variações do índice Ibovespa, além das estimativas fornecidas pela análise padrão que se utiliza apenas das informações das próprias variações de preço em períodos de tempo defasados.

1.2.2 Objetivos Específicos

De forma mais restrita, este estudo apresenta os seguintes objetivos específicos:

- ✓ Verificar a adequação do uso de informações sobre o comportamento passado das variações de preço do Ibovespa para estimativas de valores futuros através da aplicação da metodologia das redes neurais artificiais;
- ✓ Verificar a adequação do uso de informações adicionais referentes ao número de negócios e ao volume financeiro, além das informações sobre variações de preços passados nas estimativas de valores futuros do Ibovespa.

1.3 Justificativa

Mcnelis (2005) destaca que a principal motivação para o uso da metodologia de redes neurais artificiais na modelagem das variações do Ibovespa advém de que em mercados financeiros, variações lentas seguidas de bruscas oscilações nos preços são comuns de ocorrer. Principalmente, em períodos de instabilidade econômica. Nesse ambiente, modelos não lineares de maior flexibilidade pode ser uma opção mais eficiente na representação da série de dados.

Esse capítulo procurou destacar a relevância do conhecimento prévio das variações de preços no mercado de uma forma geral, o surgimento das RNAs como uma metodologia alternativa de previsão e os objetivos do estudo. Após essa breve introdução, o restante da monografia está assim estruturado: na segunda seção é apresentada uma sucinta revisão sobre previsibilidade no mercado financeiro, na terceira seção são descritos aspectos sobre RNAs e suas semelhanças com o cérebro humano, na quarta seção são detalhados os aspectos metodológicos e, na quinta seção são apresentados os resultados obtidos. Por último é realizada a conclusão do estudo.

2 PREVISIBILIDADE, SÉRIES TEMPORAIS E REDES NEURAS

2.1 Previsibilidade no mercado financeiro

A possibilidade de realizar previsões com elevada taxa de sucesso é um tópico que instiga a mente de muitos estudiosos em Finanças. Inicialmente, Bachelier (1900) propôs o modelo de caminho aleatório para explicar como evoluía a oscilação dos preços no mercado. Posteriormente, Kendall (1953) ganhou notoriedade sobre o tema ao investigar o comportamento das variações dos preços de ações e mercadorias na Bolsa de Valores de Londres. A conclusão foi de que as variações dos preços eram erráticas, ou seja, não havia nenhuma informação nos retornos passados que ajudasse a prever o seu comportamento futuro. Dessa forma, os retornos ocorriam de maneira completamente aleatória.

Procurando sintetizar a idéia sobre o comportamento das flutuações dos preços, Samuelson (1965) desenvolveu a idéia de que em um mercado informacionalmente eficiente, os preços mudam, mas são imprevisíveis. A idéia básica é de que os preços incorporam as expectativas e informações de todos os participantes do mercado. Sendo a informação disponível a todos, variações de preços para os próximos períodos serão reflexos de informações associadas a esses períodos. À medida que as informações chegam de forma aleatória ao mercado, as variações de preços deveram seguir o mesmo caminho.

A hipótese de que as variações dos preços se comportam de maneira aleatória é fortemente questionada em 1988 com a publicação de três artigos clássicos em Finanças, são dos estudos de Fama e French (1988), Lo e MacKinlay (1988) e de Poterba e Summer (1988). Fama e French (1988) analisaram os retornos mensais no mercado norte-americano e evidenciaram a presença de autocorrelação serial negativa, afirmando que de 25% a 40% das variações de preços poderiam ser previsíveis no longo prazo, a partir de retornos passados.

Lo e McKinlay (1988), também encontraram evidências de autocorrelação positiva para retornos semanais em *portfolios* formados por ações de pequenas empresas no mercado norte-americano. Poterba e Summers (1988), por sua vez, utilizando o teste de quociente de variância, encontraram evidências de autocorrelação positiva no curto prazo e autocorrelação

negativa no longo prazo, tanto no mercado norte-americano quanto em outros mercados desenvolvidos.

Posteriormente, diversos estudos empregando uma ampla variedade de testes apresentaram resultados conflitantes, uns aceitando e outros rejeitando a hipótese de aleatoriedade nas variações de preço no mercado financeiro. A divergência, na maioria dos casos, é devido às características específicas de cada estudo, por exemplo, período da amostra mais longo ou mais curto, a situação dominante do mercado financeiro no momento do estudo era de recessão, crescimento ordenado ou euforia. O tipo de teste aplicado, também, é um fator que contribui muito para aceitar ou rejeitar a presença de comportamento aleatório nos dados (variações imprevisíveis).

No Brasil, Ceretta e Costa Jr. (2002) apresentaram uma revisão, síntese e novos resultados sobre o caminho aleatório nos mercados de capitais da América Latina no período de 1990 até 1999. Nesse estudo, os autores analisaram os retornos semanais de índices de mercado da Argentina, Brasil, Chile, Colômbia, México e Venezuela. Os resultados indicaram um comportamento distinto para os diversos mercados. Os mercados da Argentina e do Brasil se mostraram perfeitamente de acordo com a hipótese do caminho aleatório, os mercados do Chile e da Colômbia rejeitaram a aleatoriedade e, para os mercados do México e da Venezuela não foi possível fazer uma afirmação conclusiva sobre a aleatoriedade ou não das variações de preços.

No contexto de previsibilidade no mercado financeiro, mais recentemente, Oliveira, Montini e Bergmann (2007) empregaram comparativamente modelos autoregressivos integrados de médias móveis com heteroscedasticidade condicional autoregressiva generalizada – ARIMA-GARCH e Redes Neurais Artificiais. Nesse estudo foram analisadas séries de preços de ações de diversos setores. A partir dos resultados obtidos, os autores concluíram que nenhuma das duas modelagens se mostrou dominante, em alguns casos o modelo ARIMA-GARCH apresentou melhor desempenho e em outros, as Redes Neurais Artificiais apresentaram melhor desempenho. Na conclusão, os autores sugeriram que as duas técnicas sejam utilizadas de forma complementar par melhorar a capacidade de decisão do analista.

2.2 Series temporais

Séries temporais univariadas são classificadas como modelos específicos onde o foco é a estimação de valores futuros de uma série utilizando-se somente informações contidas em seus próprios valores passados e de um componente de erro. Brooks (2008) enfatiza que modelos de séries temporais são considerados desprovidos de uma construção teórica fundamentada, ou seja, são utilizados para capturar aspectos empíricos relevantes que podem ser oriundos da influência de muitos eventos estruturais não observados ou de difícil mensuração.

Uma série temporal pode ser entendida como um conjunto de valores que estão dispostos em ordem cronológica, assumindo-se que, na maioria dos casos, o tempo é uma variável discreta. Kirchgassner e Wolters (2008) afirmam que as séries temporais são consideradas muito importantes desde os primórdios da ciência natural, afirmando que astrônomos da Babilônia se utilizavam de séries temporais do posicionamento relativo de estrelas e planetas para antever eventos astronômicos.

2.2.1 Estacionariedade

Uma série temporal é parte de um processo observado na seqüência do tempo, $t=1, \dots, T$. Para identificar a dependência no tempo, será utilizado o subscrito t para identificar uma observação individual e T para identificar o número de observações do processo observado.

A fundamentação básica para estudo de séries temporais é a estacionariedade. Uma série $\{r_t\}$ é dita ser fracamente estacionária (estacionária de segunda ordem, estacionária em covariância) se a média de r_t e a covariância entre r_t e r_{t-k} são invariantes no tempo. Se a série $\{r_t\}$ é fracamente estacionária, tem-se que:

$$\text{Média} = E(r_t) = E(r_{t-k}) = \mu, \quad [1]$$

$$\text{Variância} = E[(r_t - \mu)^2] = E[(r_{t-k} - \mu)^2] = \sigma^2, \quad [2]$$

$$\text{Covariância} = E[(r_t - \mu)(r_{t-k} - \mu)] = E[(r_{t-j} - \mu)(r_{t-j-k} - \mu)] = \gamma_k, \text{ ou } [3]$$

$$\text{Autocovariância} = \text{Cov}(r_t, r_{t-k}) = \text{Cov}(r_{t-j}, r_{t-j-k}) = \gamma_k. \quad [4]$$

Conseqüentemente, a covariância entre as observações na série $\{r_t\}$ é um função que depende da distância intervalar entre t e k e não do tempo t ou k em que ocorrem. Em termos práticos, se uma série de retornos $\{r_t | t=1, \dots, T\}$ é fracamente estacionária se o gráfico dos retornos ao longo de T identificar valores flutuantes com variações constantes ao redor de um nível fixo. A estacionariedade fraca torna possível que se realize inferências sobre valores futuros dos retornos.

Na condição de estacionariedade fraca, assume-se implicitamente que os dois primeiros momentos, média e variância de $\{r_t\}$ são finitos (não explosivos, flutuantes ao redor de um valor esperado). Assim, se a série temporal $\{r_t\}$ tem distribuição normal, então satisfaz a condição de estacionariedade fraca, mas a recíproca não é verdadeira. É importante destacar que γ_k é comumente referenciado como autocovariância de r_t na defasagem k , portanto:

$$\text{Para } k = 0 \text{ tem-se que } \gamma_{k=0} = E[(r_t - \mu)(r_{t-0} - \mu)] = E[(r_t - \mu)^2] = \text{Cov}(r_t, r_t) = \text{Var}(r_t) \quad [5]$$

$$\text{Para } k = -1 \text{ tem-se que } \gamma_{k=-1} = E[(r_t - \mu)(r_{t-1} - \mu)] = \text{Cov}(r_t, r_{t-1}) \quad [6]$$

$$\text{Para } k = +1 \text{ tem-se que } \gamma_{k=+1} = E[(r_t - \mu)(r_{t+1} - \mu)] = \text{Cov}(r_t, r_{t+1}). \quad [7]$$

Dado a definição de que a covariância depende apenas do espaçamento entre as defasagens pode-se afirmar que $\gamma_{k=-1} = \gamma_{k=+1}$, pois o espaçamento é o mesmo, ou seja, $|k| = 1$.

2.2.1.1 Coeficiente de correlação

O coeficiente de correlação entre duas variáveis X e Y mede a dependência linear entre ambas e pode assumir valores de $-1 \leq \rho \leq 1$, tal que $\rho_{x,y} = \rho_{y,x}$. O coeficiente de correlação é definido como:

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}, \quad [8]$$

$$\rho_{x,y} = \frac{E[(X - \mu_x)(Y - \mu_y)]}{\sqrt{E(X - \mu_x)^2 E(Y - \mu_y)^2}}, \quad [8.1]$$

onde μ_x e μ_y são as médias de X e Y , respectivamente. Se X e Y são independentes $\rho_{x,y} = 0$, mas se $\rho_{x,y} = 0$ não se pode afirmar que exista ausência de dependência entre X e Y . Quando a análise é realizada para dados de uma amostra $\{(x_t, y_t)\}$, $t=1, \dots, T$, o coeficiente de correlação é definido por:

$$\hat{\rho}_{x,y} = \frac{\sum_{t=1}^T (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^T (x_t - \bar{x})^2 \sum_{t=1}^T (y_t - \bar{y})^2}}, \quad [9]$$

onde \bar{x} e \bar{y} são as médias amostrais de X e Y , dadas por $\bar{x} = \sum_{t=1}^T x_t$ e $\bar{y} = \sum_{t=1}^T y_t$, respectivamente.

2.2.1.2 Função de autocorrelação

Autocorrelação é uma extensão do conceito de correlação, nesse caso, aplicado sobre as defasagens da variável analisada. Assim, o coeficiente de correlação entre r_t e r_{t-l} é chamado de autocorrelação da primeira defasagem. Na forma geral, o coeficiente de autocorrelação é obtido por:

$$\rho_k = \frac{Cov(r_t, r_{t-k})}{\sqrt{Var(r_t)Var(r_{t-k})}}, \quad [10]$$

$$\rho_k = \frac{Cov(r_t, r_{t-k})}{Var(r_t)}, \quad [10.1]$$

$$\rho_k = \frac{\gamma_k}{\gamma_0}. \quad [10.2]$$

Observe que, por definição, quando $k=0$, $\rho_0=1$, pois $Cov(r_t, r_{t-0})/Var(r_t) = Var(r_t)/Var(r_t)$. Uma série r_t , fracamente estacionária, será não correlacionada serialmente se, e somente se, $\rho_k = 0$ para todo $k>0$. Para dados de uma amostra $\{r_t\}$, $t=1, \dots, T$, sendo \bar{x} média aritmética de $\{r_t\}$, obtêm-se o coeficiente de autocorrelação amostral relativo a $t-1$ por:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^T (r_{t+1} - \bar{r})(r_{t-k} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2}, \quad [11]$$

$$\hat{\rho}_1 = \frac{\sum_{t=2}^T (r_t - \bar{r})(r_{t-1} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2}. \quad [11.1]$$

Dispondo diversos valores de $\hat{\rho}_k$, com $(0 \leq k < T - 1)$, na forma gráfica, tem-se o chamado correlograma amostral ou *autocorrelation function* – ACF. Mills (1999) destaca que a ACF é de grande importância na modelagem das dependências e juntamente com a média e a variância descrevem a evolução de r_t . A ACF evidencia como r_t está relacionado (no tempo) com seus valores anteriores, construindo uma idéia de extensão e grau de memória do processo.

2.2.1.3 Teste de Ljung e Box

Após a obtenção de diversos valores de coeficientes de autocorrelação é, geralmente, preferível verificar a significância conjunta desses coeficientes. Para tal, aplica-se a estatística *portmanteau* (geral) proposta por Box e Pierce (1970) que é dada por:

$$Q^*_{(m)} = T \sum_{k=1}^m \hat{\rho}_k . \quad [12]$$

As hipóteses verificadas pela estatística $Q^*_{(m)}$ são:

$$H_0: \hat{\rho}_1 = \dots = \hat{\rho}_m = 0 ,$$

$$H_1: \hat{\rho}_k \neq 0 \text{ para pelo menos algum } k \in \{1, \dots, m\} .$$

Para amostras grandes, $Q^*_{(m)}$ segue uma distribuição qui-quadrado com m graus de liberdade. Se o valor calculado exceder o valor crítico de $Q^*_{(m)}$ na distribuição χ^2 ao nível específico de significância, rejeita-se a hipótese nula de que todos os coeficientes são iguais a zero.

Posteriormente, com objetivo de aumentar o poder do teste associado à estatística $Q^*_{(m)}$ frente a pequenas amostras, Ljung e Box (1978) propuseram alterações na formulação de Box e Pierce (1970). Agora o teste é obtido por:

$$Q_{(m)} = T(T+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{T-k} , \quad [13]$$

$$Q_{(m)} = T(T+2) \sum_{k=1}^m (T-k)^{-1} \hat{\rho}_k^2 \sim \chi^2_{\alpha, m} . \quad [13.1]$$

Para dados amostrais $Q_{(m)}$ segue uma distribuição qui-quadrado com m graus de liberdade. As hipóteses do teste são as mesmas do teste original de Box e Pierce (1970). H_0 será rejeitada se $Q_{(m)} > \chi^2_{\alpha, m}$, onde α representa o nível de significância e m os graus de

liberdade, ou se *p-value* for menor ou igual ao nível de significância α (geralmente de 1% a 5%).

Analisando-se [13 e 13.1], constata-se que, assintoticamente (quando o tamanho da amostra tende ao infinito), os termos $(T + 2)$ e $(T - k)$ tendem a se cancelar e a estatística de Ljung-Box converge para Box-Pierce.

Em termos práticos, pode-se perceber que o resultado do teste é afetado pela escolha do valor de m . Na literatura de Finanças há sugestões para que se utilizem valores pequenos para m , ou seja, $m \approx \ln T$. Com esse artifício, consegue-se um aumento no poder do teste.

2.2.2 Ruído branco

A terminologia ruído branco (*white noise*), derivada de sinal branco, é muito utilizada na literatura de Finanças. Lathi (2007) destaca que um sinal é um conjunto de dados ou informação com certas propriedades. Sinais podem ser modificados e modelados a fim de se extrair informações adicionais.

O adjetivo branco é utilizado para fazer referência a óptica, onde a cor branca tem um espectro mais largo e constante se comparado com o espectro da luz colorida. Isso se deve ao fato da luz branca ser obtida por meio da combinação simultânea de diversas frequências cromáticas. Por exemplo, ao girar rapidamente um disco colorido sobre um eixo fixo, as diversas cores convergem para o branco impossibilitando que se identifiquem quais as reais cores presentes no disco. Dessa forma, o termo sinal (conjunto de dados) passa a ser chamado de ruído (conjunto de dados que não permitem informações adicionais), um chiado indecifrável.

Embora não seja objetivo deste estudo, é interessante destacar que existem diversos tipos de ruídos brancos. Por exemplo, algumas dessas seqüências são: i) ruído branco uniforme, oscilação segundo uma distribuição uniforme $[-2,5; \dots; 2,5]$ não correlacionada em torno de zero; ii) ruído branco gaussiano, oscilação segundo uma distribuição gaussiana [média = 0 e desvio padrão = 1] não correlacionada em torno de zero e, iii) ruído branco *Cauchy*, oscilação segundo uma distribuição *Cauchy* (0, 1) não correlacionada em torno de zero.

A Figura 1 ilustra graficamente essas três formas distintas de ruído branco.

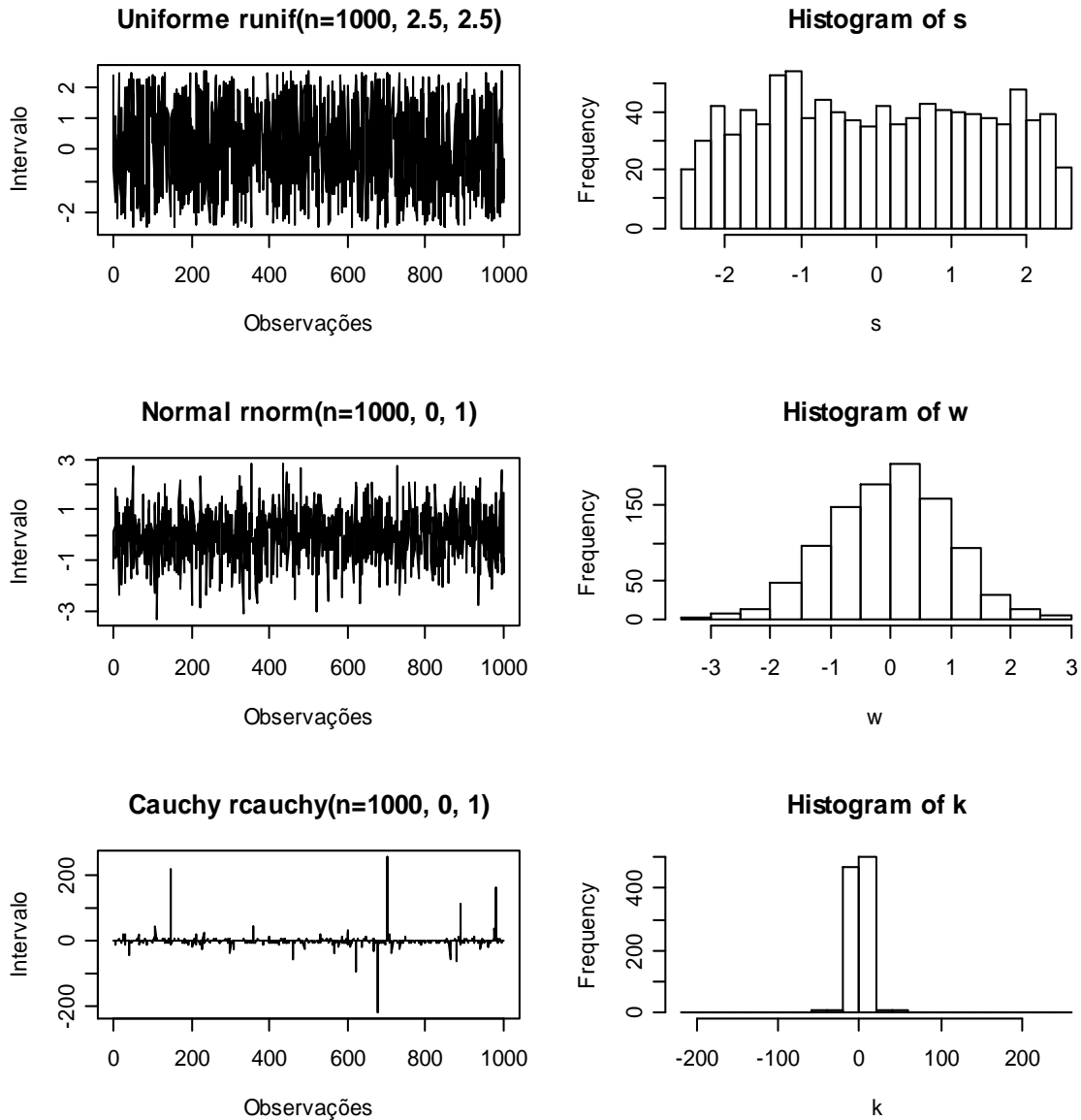


FIGURA 1 – Séries ruído branco.

Na primeira linha da Figura 1 tem-se uma série ruído branco uniforme e ao lado o seu histograma de freqüência; na segunda linha, uma série ruído branco gaussiano e seu respectivo histograma; na terceira linha, uma série *cauchy* e seu histograma.

Em termos formais, uma série de retornos r_t ($t = -\infty, \dots, +\infty$) é chamada de ruído branco se a seqüência for constituída de variáveis aleatórias. Nessa definição, tem-se que:

$$E(r_t) = \mu \text{ para } \forall t; \quad [14]$$

$$E[(r_t - \mu)^2] = \sigma^2 \text{ para } \forall t; \quad [14.1]$$

$$E(r_t r_{t-k}) = \begin{cases} \sigma^2 & \text{se } k = 0 \\ 0 & \text{para } \forall t \text{ quando } k \neq 0. \end{cases} \quad [14.2]$$

Então, se uma série é ruído branco deverá ter média e variância constante, com autocovariância igual a zero, exceto na defasagem zero, ou seja, as observações não são correlacionadas com os outros valores da seqüência.

Brooks (2008) enfatiza que adicionando a restrição de que r_t é distribuído normalmente, então o coeficiente de autocorrelação amostral deve repousar entre os limites da distribuição normal $\hat{\rho} \sim N(0, 1/T)$. Onde T identifica o tamanho da amostra. Com base nessa definição, pode-se conduzir um teste de significância para o coeficiente de autocorrelação construindo-se uma região de não-rejeição (intervalo de confiança) para qualquer coeficiente estimado, determinando se ele é estatisticamente diferente de zero. Por exemplo, para 95%, a região de não rejeição é dada por $\pm 1,96(1/\sqrt{T})$.

Mais precisamente, uma série ruído branco deverá ter a *autocorrelation function* – ACF for igual à zero. Em termos práticos, se todas as ACFs forem consideradas zero, então a série é ruído branco. Para verificar tal situação, aplica-se o teste de Ljung e Box (1978).

2.3 Redes neurais artificiais

O desenvolvimento de trabalhos sobre Redes Neurais Artificiais (RNAs) tem como marco inicial os estudos de McCulloch e Pitts (1943). Psicólogo e neurofisiologista, McCulloch desenvolveu diversos estudos procurando representar e modelar eventos no sistema nervoso e em parceria com o matemático Pitts apresentou ensaios a respeito de redes lógicas de neurônios artificiais chamados de neurônio.

O foco inicial desses estudos estava direcionado na descrição de um modelo artificial de neurônio, ou seja, uma simplificação sobre o neurônio biológico com base no conhecimento da época sobre o mesmo. Posteriormente, Hebb (1949) demonstrou como o processo de aprendizagem ocorre em redes artificiais através da variação de pesos de entrada do neurônio.

Desde os pioneiros estudos de McCulloch e Pitts (1943) até meados de 1980 houve importantes avanços em RNAs, porém grande parte do desenvolvimento foi de lento crescimento. A lenta evolução foi devido a maior exigência de tempo para a solução de problemas não lineares.

Os problemas não lineares eram considerados difíceis de aprender. Após meados de 1980 ocorreu um rápido crescimento no interesse pelas RNAs, muito provavelmente estimulado pelo avanço da tecnologia de microcomputadores.

Para um melhor entendimento das RNAs deve-se fazer um paralelo com o processamento que ocorre no cérebro humano. O cérebro humano contém aproximadamente 100 bilhões de neurônios e 100 trilhões de sinapses.

Conforme a Figura 2, neurônio é a célula do sistema nervoso responsável pela condução do impulso nervoso (sinal) para outros nervos ou células. Um neurônio recebe um sinal através dos seus dendritos que são numerosos prolongamentos especializados na recepção de estímulos nervosos. Após o recebimento do sinal, o mesmo é repassado a outra célula através do axônio.

Os neurônios recebem de forma contínua impulsos de neurônios próximos ou de outras células através das sinapses nervosas localizadas nas extremidades dos dendritos. O contato físico entre dendritos não ocorre, o espaço entre eles é chamado de fenda sináptica. Dos axônios são liberados os neurotransmissores que atravessam a fenda sináptica e estimulam receptores nos dendritos, ocorrendo, dessa forma, o impulso nervoso de um neurônio para outro.

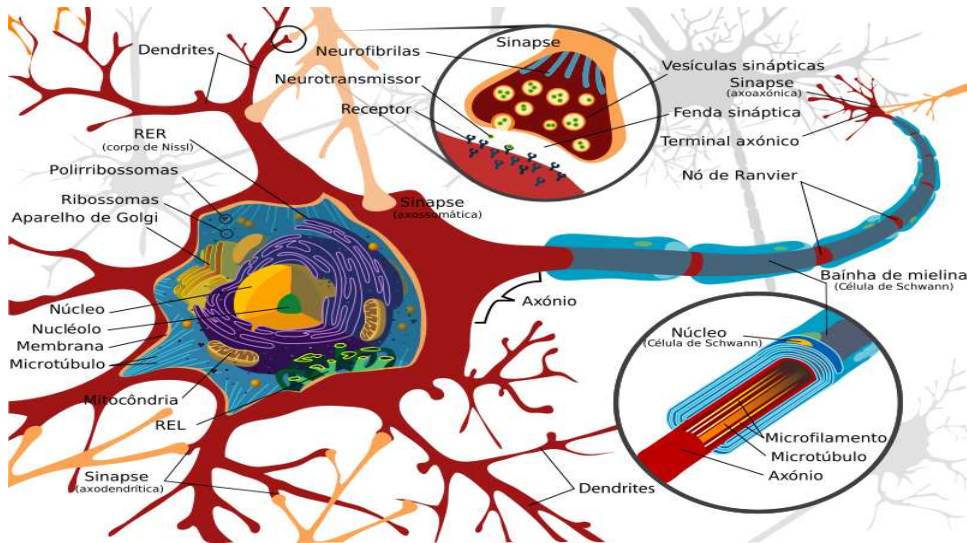


FIGURA 2 – Representação do neurônio humano.

Fonte: Imagem de domínio público.

http://commons.wikimedia.org/wiki/Image:Complete_neuron_cell_diagram_pt.svg

Haykin (2001) descreve que, de forma análoga ao neurônio humano, o modelo de neurônio artificial é composto por três elementos principais: i) elos, ii) somador e, iii) função de ativação.

Os elos conduzem os sinais de entrada, cada sinal x_j ($j = 1, \dots, n$) é ponderado com peso sináptico específico w_j . Como pode ser observado na Figura 3, o conjunto de sinais de entrada ponderados por pesos sinápticos é combinado e, dependendo das ações inibidoras ou excitatórias de cada sinapse, poderá disparar uma saída de sinal. O sinal de saída do neurônio é limitado a valores finitos por uma função de ativação linear ou não linear.

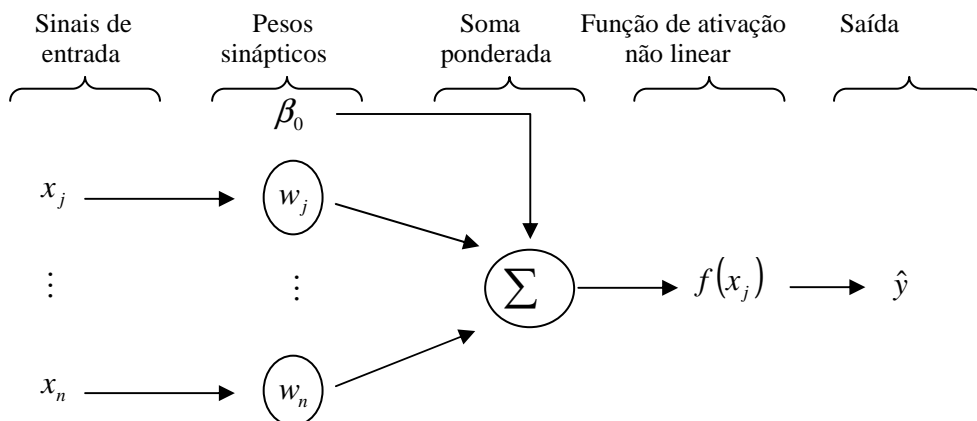


FIGURA 3 – Representação do neurônio artificial.

Fonte: Adaptado de Haykin (2001).

Analisando-se a Figura 3, pode-se constatar que uma RNA é um conjunto de neurônios artificiais. Um neurônio artificial é uma forma simplificada de representação matemática de um neurônio biológico e, por sua vez, neurônio biológico é entendido como unidades elementares de processamento de informações (sinais) em qualquer organismo vivo. Sem nenhuma pretensão de validade neurofisiológica, pode-se descrever o funcionamento do neurônio artificial da seguinte maneira:

- 1° Neurônios são unidades elementares na qual ocorre o processamento da informação;
- 2° A entrada de informação ocorre através de sinais que são transmitidos até os neurônios por meio de conexões, as quais têm a propriedade de ponderar os sinais ao transmiti-los;
- 3° Em cada neurônio ocorre o processamento interno onde atua uma função de ativação (geralmente não linear) ponderando e transformando os sinais em saída.

Haykin (2001) enfatiza que em cada neurônio artificial, representado na Figura 3, está presente uma constante (β_0) que tem por objetivo aumentar (se de sinal positivo) ou diminuir (se de sinal negativo) o valor da função de ativação. Essa constante, geralmente, é referenciada como *bias*. Na RNA a função de ativação $f(x_j)$ é a responsável pelo processamento de informações e pela geração de uma saída do neurônio. As RNAs, geralmente se utilizam de funções de ativação não lineares. As funções de ativação não lineares distribuem a não linearidade em paralelo e se comunicam por meio de conexões sinápticas. As funções de ativação mais utilizadas são as do tipo sigmóide.

A função sigmóide mais amplamente utilizada como função de ativação em RNAs é a função logística $f = 1/[1 + \text{Exp}(\psi)]$ sendo $\psi = \beta_0 + \sum_{i=1}^n x_i w_i$, onde x_i são as variáveis de entrada e w_i as ponderações da função de ativação com $i = 1, \dots, n$. A saída gerada pela função de ativação vai oscilar de forma contínua no intervalo entre 0 e 1. Sendo desejável que a função de ativação estenda a saída entre valores no intervalo de -1 a 1, deve-se utilizar a função tangente hiperbólica definida por $f = \tanh(x_i)$.

Haykin (2001) enfatiza que uma RNA é um sistema de processamento de informação em paralelo e, semelhante ao cérebro humano, trabalha com processos não lineares de

natureza complexa. A partir da forma como a RNA é estruturada e operacionalizada pode-se destacar dois aspectos básicos de semelhança ao cérebro humano: i) A RNA adquire conhecimento num exaustivo processo de treinamento-aprendizagem através de análises de variáveis em seu ambiente e, ii) O conhecimento adquirido fica armazenado nos neurônios artificiais na forma de pesos sinápticos que são formados do relacionamento entre os neurônios e as informações de entrada.

Braga, Carvalho e Ludermir (2007) destacam que uma das características mais importantes das RNAs é a sua capacidade de aprender por meio de exemplos. Nesse contexto, tem-se que o conhecimento é adquirido através do relacionamento entre neurônios que, paralelamente, ajustam à importância de certas informações através de conexões sinápticas. O ajuste das conexões ocorre em vista a melhorar o desempenho da rede conforme algum critério pré-estabelecido. A raiz do erro quadrático médio (REQM) fornecido pela RNA no processo de treinamento frente a um conjunto de dados é um exemplo de um critério de desempenho a ser melhorado, gerando, com isso, aprendizagem.

2.3.1 Arquitetura de rede

A arquitetura da rede neural artificial diz respeito a forma como são estruturados os neurônios nas camadas intermediárias, ligando os valores de entrada (x) com os valores de saída (y). Haykin (2001) classifica as arquiteturas de rede em três diferentes tipos: i) redes alimentadas adiante com camada única, ii) rede alimentadas adiante com múltiplas camadas e, iii) redes recorrentes.

2.3.1.1 Redes neurais alimentadas adiante com camada única

Um exemplo da arquitetura de uma RNA alimentada adiante de camada única é apresentado na Figura 4, onde são ilustrados dois neurônios na camada intermediária unindo os valores de entrada aos valores de saída.

O modelo representado pela Figura 4 é dado pela formulação [15].

$$f(x_t, \theta) = \beta_0 + \beta_1 g(\alpha_{10} + \alpha_{11}x_{1t} + \alpha_{12}x_{2t}) + \beta_2 g(\alpha_{20} + \alpha_{21}x_{1t} + \alpha_{22}x_{2t}) + \varepsilon_t. \quad [15]$$

Em [15], x_t é o vetor representativo dos valores de entrada x_{1t}, x_{2t} ($t = 1, \dots, n$); $\alpha_{10}, \alpha_{11}, \alpha_{12}, \alpha_{20}, \alpha_{21}, \alpha_{22}$ são os coeficientes da primeira e da segunda função de ativação, $\beta_0, \beta_1, \beta_2$ são o intercepto e os coeficientes de inclinação.

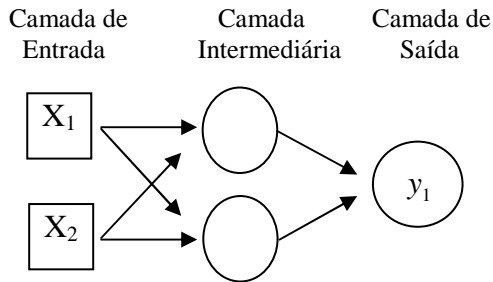


FIGURA 4 – RNA *feedforward* (2:2:1).

Este exemplo demonstra como uma simples RNA *feedforward* pode ser vista como um flexível modelo de regressão não linear, facilmente estimado com ferramentas de otimização padrão.

2.3.1.2 Redes neurais alimentadas adiante com múltiplas camadas

Como alternativa para modelagem de problemas mais complicados, tem-se as RNAs *feedforward* com duas ou mais camadas intermediárias. Haykin (2001) descreve que o fato de adicionar mais camadas ocultas, torna a rede capaz de extrair estatísticas de ordem mais elevada. A Figura 5 apresenta uma RNA *feedforward* com duas camadas intermediárias.

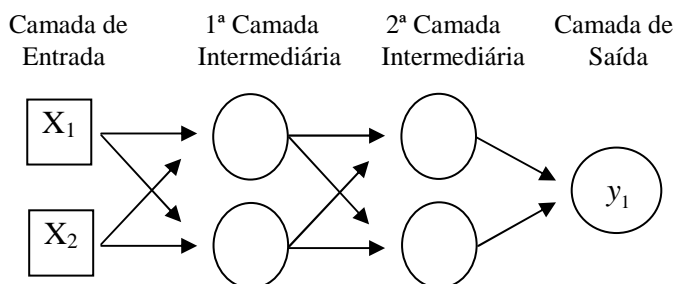


FIGURA 5 – RNA *feedforward* (2:2:2:1).

O modelo de RNA alimentada adiante com múltiplas camadas intermediárias é dado pelas formulações [16, 17].

$$y_t = s\left(\beta_0 + \sum_{i=1}^q \beta_i h_{i,t}\right) + \varepsilon_t, \quad [16]$$

$$h_{i,t} = g\left(\alpha_{i0} + \sum_{j=1}^n \alpha_{ij} x_{i,j}\right), \quad [17]$$

onde s e g são conhecidos como funções de ativação e ε_t é o termo de erro com $(0, \sigma^2)$, θ representa os parâmetros a serem estimados seguindo um critério de ajuste especificado.

2.3.1.3 Redes recorrentes

Gençay, Selçuk e Whicher (2002) destacam que a literatura de RNAs tem sido dominada pelos modelos *feedforward*. Uma rede recorrente se distingue de uma rede *feedforward* pela presença de no mínimo uma conexão de *feedback*. Uma típica representação de rede recorrente é ilustrada na Figura 6.

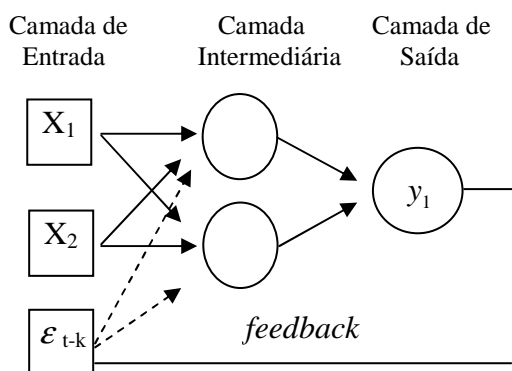


FIGURA 6 – RNA recorrente (2:2:1).

O modelo recorrente é descrito pelas formulações [18, 19].

$$y_t = s \left(\beta_0 + \sum_{i=1}^q \beta_i h_{i,t} \right) + \varepsilon_t, \quad [18]$$

$$h_{i,t} = g \left(\alpha_{i0} + \sum_{j=1}^n \alpha_{ij} x_{i,j} + \sum_{p=1}^k \delta_{i,p} \varepsilon_{t-p} \right). \quad [19]$$

Onde ε_{t-p} é o componente de *feedback* do tempo t-p. Agora, a melhoria temporal do output pode ser expressa por $y_t = f(x_{1,t}, x_{2,t}, \varepsilon_{t-p}, \theta) + \varepsilon_t$. Modelos de RNAs recorrentes também podem ser vistos como modelos não lineares autoregressivos de médias móveis – NARMA(p, q).

O presente capítulo focou sobre os aspectos teóricos básicos relativos à previsibilidade no mercado financeiro, séries temporais e redes neurais. A seguir será realizado um detalhamento dos aspectos metodológicos que nortearam o estudo.

3 ASPECTOS METODOLÓGICOS

3.1 Amostra e dados

Os dados considerados para o desenvolvimento desse estudo abrangem o período de 23/01/2006 até 18/04/2008. Os valores são relativos à frequência temporal diária e estão em moeda local. Os preços considerados referente à carteira teórica da Bovespa são relativos ao fechamento do dia. Para o número de negócios e volume financeiro, foi considerada a totalização diária.

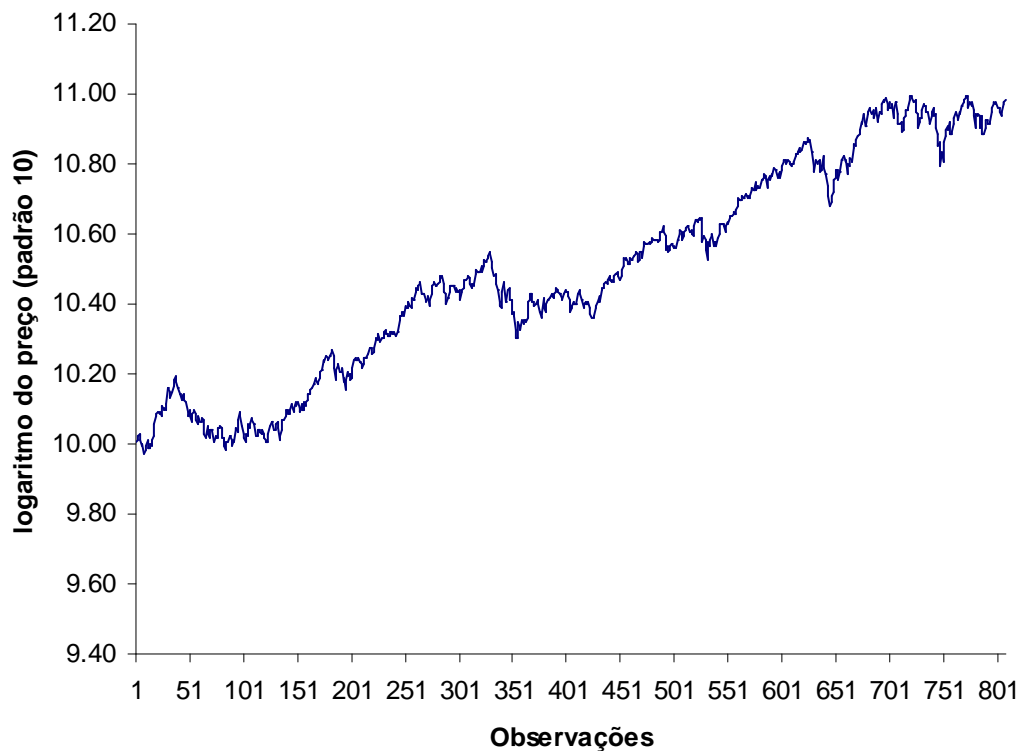


FIGURA 7 – Série temporal dos preços da carteira teórica da Bovespa

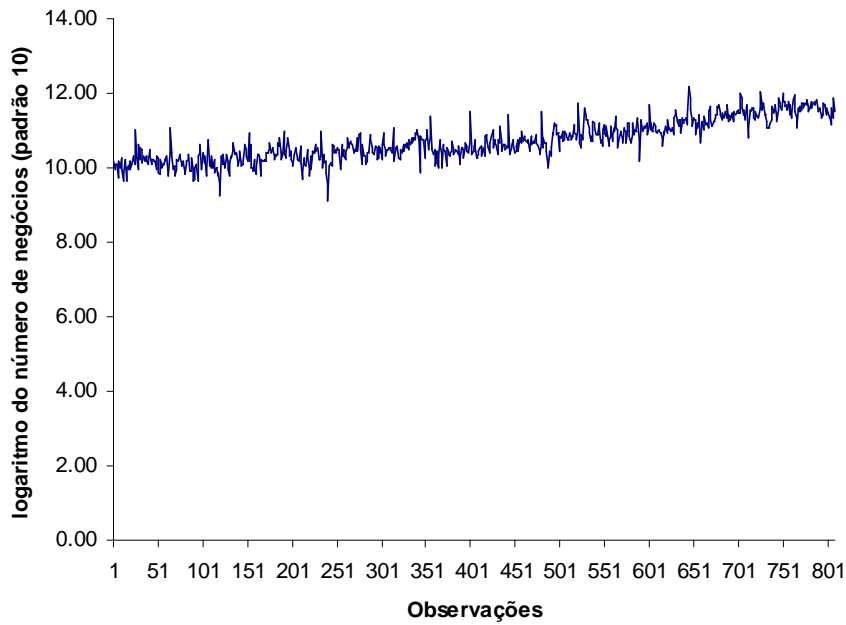


FIGURA 8 – Série temporal do número de negócios

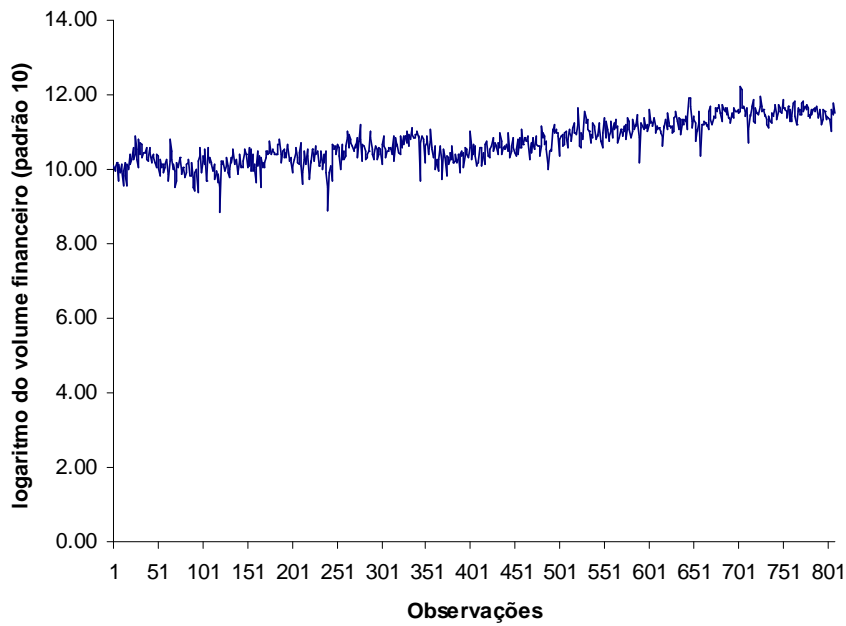


FIGURA 9 – Série temporal do volume financeiro

A taxa de retorno foi calculada pela diferença do logaritmo do preço de fechamento do índice conforme a seguinte equação:

$$r_{i,t} = \ln(P_{i,t}) - \ln(P_{i,t-1}), \quad [20]$$

onde $r_{i,t}$ é o retorno do índice i no tempo t , $P_{i,t}$ e $P_{i,t-1}$ são os preços de fechamento no tempo t e no período imediatamente anterior. Para o cálculo foram considerados os ajustes relativos a dividendos, bonificações e outros ocorridos no período. Procedimento semelhante foi adotado para a variação do número de negócio e variação do volume financeiro.

Procurando melhorar o processo de ajuste pelo método dos mínimos quadrados não linear, as variáveis (retorno, variação do número de negócios e variação no volume financeiro) foram padronizadas no intervalo $[0,1]$. A padronização não afeta as características dos dados e torna bem mais rápido a obtenção de uma solução computacional. A padronização adotada foi obtida pela aplicação da fórmula:

$$r_{i,t}^* = [r_{i,t} - \min(r_i)] / [\max(r_i) - \min(r_i)]. \quad [21]$$

onde $r_{i,t}^*$ é a variável padronizada, $\min(r_i)$ é o valor mínimo e $\max(r_i)$ é o valor máximo da série temporal das variáveis.

3.2 Modelagem

As formulações [21, 22 e 23] detalham, de maneira genérica, o modelo de RNAs utilizada para identificar a influência das variáveis números de negócios e volume financeiro na previsibilidade do retorno do Ibovespa.

$$RNA(x|\phi, \alpha, \beta) = \left[1 + \text{Exp} \left(\phi_0 + \sum_{k=1}^3 \phi_k g_k \right) \right]^{-1}; \quad [21]$$

$$g_k(f_{k,j}) = \left[1 + \text{Exp} \left(\alpha_0 + \sum_{k=1}^3 \alpha_{k,j} f_{kj} \right) \right]^{-1}, \text{ para } j = 1, 2, 3; \quad [22]$$

$$f_j(x_{ij}) = \left[1 + \text{Exp} \left(\beta_0 + \sum_{k=1}^{r,s,p} \beta_{ij} x_{ij} \right) \right]^{-1}, \text{ para } j = 1, 2, 3; \text{ e } r = 5, s = 10 \text{ e } p = 15. \quad [23]$$

No primeiro modelo de RNA, composto pela combinação das formulações [21, 22, 23], será considerado como relevantes para a previsibilidade do retorno do Ibovespa apenas suas informações defasadas em até (t-5). Nesse caso, há cinco variáveis de entrada na RNA.

Na primeira e na segunda camada intermediária serão utilizadas três funções sigmóides escolhidas de forma arbitrária.

O procedimento de duas camadas intermediárias com três funções sigmóides será utilizado nos três modelos de RNA com objetivo de facilitar a comparação de desempenho pela raiz do erro quadrático médio.

A Figura 10 ilustra as RNAs a serem aplicadas.

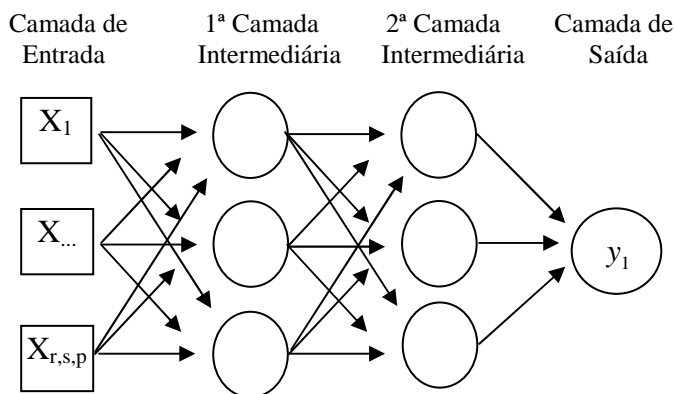


FIGURA 10 – Estrutura das RNAs a serem aplicadas.

No segundo modelo são considerados informações defasadas em até (t-5) das variações do número de negócios além das variáveis do primeiro modelo. No terceiro modelo são acrescentadas informações defasadas do volume financeiro em até (t-5) além das variáveis do segundo modelo, totalizando 15 variáveis de entrada.

3.3 Treinamento e teste

Após a padronização dos dados de entrada, o processo de aplicação da RNA envolve mais dois estágios: i) treinamento e, ii) teste. O treinamento é um procedimento no qual a rede vai aprender as relações que governam a variação dos dados de saída. No treinamento, os coeficientes (pesos) da rede são modificados de forma a minimizar a raiz do erro quadrático médio entre os valores de saída observados e os estimados dos primeiros 758 valores.

Na prática não se pode determinar se um modelo é bom ou não com base em apenas uma realização. Portanto, na fase de teste, um conjunto de dados separados, composto de 50 valores, não utilizados pela rede na fase de treinamento será considerado para avaliar o desempenho levando-se em consideração os coeficientes estimados.

Na fase de teste, a rede será avaliada com base nos valores da raiz do erro quadrático médio que são gerados do confronto entre os 50 valores realizados e suas estimativas. A raiz do erro quadrático médio é definida pela formulação [23].

$$REQM = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2} . \quad [23]$$

Em [23] T é o número total de observações, y_t ($t = 1, \dots, T$) são os valores observados e \hat{y}_t os valores estimados pela rede neural artificial. Um modelo será melhor do que outro quando o seu $REQM$ for menor.

3.4 Software

O software utilizado foi o software denominado de *R*. O software *R* é uma linguagem e ambiente para computação estatística e gráfica. *R* fornece uma ampla variedade de estatísticas tais como modelagem linear e não-linear, testes estatísticos clássicos, análise de séries temporais, *clustering*, análise multivariada, redes neurais entre outras.

O *R* é disponibilizado como software livre sob os termos da *Free Software Foundation's GNU (General Public License)* em forma de código fonte. A operacionalidade do *R* se concretiza através de pequenos pacotes. Atualmente, janeiro de 2009, há

aproximadamente 1.600 pacotes disponíveis para serem livremente utilizados no R. O pacote específico utilizado no presente estudo foi o AMORE versão 0.29. AMORE é um sistema de redes neural artificial altamente flexível. O pacote permite formar uma rede *feedforward* multicamadas com um mínimo de programação. A programação elaborado para os três modelos analisados encontra-se nos Anexos A, B e C .

4 RESULTADOS OBTIDOS

Inicialmente são apresentados, na Tabela 1, os valores das estatísticas descritivas para as três séries de dados. Na primeira coluna de valores tem-se as estatísticas relativas a variável retorno (diferença do logaritmo do preço), na segunda coluna as estatísticas da variação do número de negócio (diferença do logaritmo do número de negócios) e na última coluna as estatísticas da variação do volume financeiro (diferença do logaritmo do volume financeiro).

Tabela 1 – Estatísticas descritivas das variáveis

Estatísticas	Retorno	Negócio	Volume
Período de treinamento			
Média	0.0012	0.0019	0.0017
Mediana	0.0019	-0.0138	-0.0127
Desvio padrão	0.0166	0.2826	0.2922
Excesso de Curtose	0.9062	1.7226	1.4501
Assimetria	-0.3204	0.1465	0.1198
Mínimo	-0.0686	-0.9973	-1.2242
Máximo	0.0578	1.0854	1.2757
Observações	758	758	758
Período de teste			
Média	0.0019	0.0019	0.0044
Mediana	0.0029	-0.0023	-0.0112
Desvio padrão	0.0180	0.2123	0.2115
Excesso de Curtose	0.6792	0.3066	0.1031
Assimetria	-0.5447	0.3978	0.4197
Mínimo	-0.0514	-0.4675	-0.3931
Máximo	0.0387	0.5287	0.5809
Observações	50	50	50

Observa-se na parte superior da Tabela 1, que a série de retorno se caracteriza com excesso de curtose o que lhe confere uma aparência da curva de distribuição de frequência mais elevada e caudas mais longas que a de uma distribuição normal. A assimetria negativa caracterizada pela mediana maior que a média induz a possibilidade de uma maior ocorrência

de valores extremos negativos. De uma forma geral, para todas as variáveis, as características do período de treinamento são semelhantes aos do período de teste.

Os resultados obtidos através da aplicação dos três modelos de RNAs são apresentados a seguir. O primeiro modelo procurou o melhor ajuste às variações do Ibovespa considerando apenas seus valores defasados ($t-1, \dots, t-5$) como sendo relevantes. O segundo modelo acrescenta as informações defasadas ($t-1, \dots, t-5$) do número de negócios. Por fim, o terceiro modelo acrescenta as informações defasadas ($t-1, \dots, t-5$) do volume financeiro.

4.1 Modelo com informações sobre os retornos

Os resultados da RNA que considera apenas as informações defasadas dos retornos são apresentados na Tabela 2 e é ilustrada na Figura 11. Deve-se salientar que os resultados da Tabela 2, são relativos ao período de treinamento da RNA, ou seja, o período utilizado como amostra para que a rede aprenda os padrões e possa, posteriormente, realizar a previsão. Na fase de treinamento foram consideradas as primeiras 758 observações e, após, foram realizadas 50 previsões fora da amostra.

Tabela 2 – Pesos sinápticos da RNA (5:3:3:1)

	F1	F2	F3	G1	G2	G3	K1
Bias	-0,1099	0,0230	-0,0503				
r_{t-1}	-0,0134	-0,1340	0,0096				
r_{t-2}	-0,1448	0,2966	-0,1785				
r_{t-3}	0,1949	0,1129	0,2635				
r_{t-4}	0,1710	-0,0507	-0,2114				
r_{t-5}	0,0510	0,0913	0,1084				
Bias				-0,0175	-0,0499	0,2161	
F1				0,0007	-0,0242	-0,0472	
F2				0,1441	0,2517	0,0434	
F3				0,1442	-0,1949	-0,0739	
Bias							0,1418
G1							0,0866
G2							-0,1631
G3							0,1180

Os valores relativos a F1, F2 e F3 são da primeira camada intermediária;

Os valores relativos à G1, G2 e G3 são da segunda camada intermediária.

A Figura 11 ilustra o desempenho da RNA (5:3:3:1) onde fica evidente as ligações entre os valores de entrada (*inputs*) com a primeira camadas intermediária, essa com a segunda camada intermediária e, por fim com os valores de saída (*outputs*).

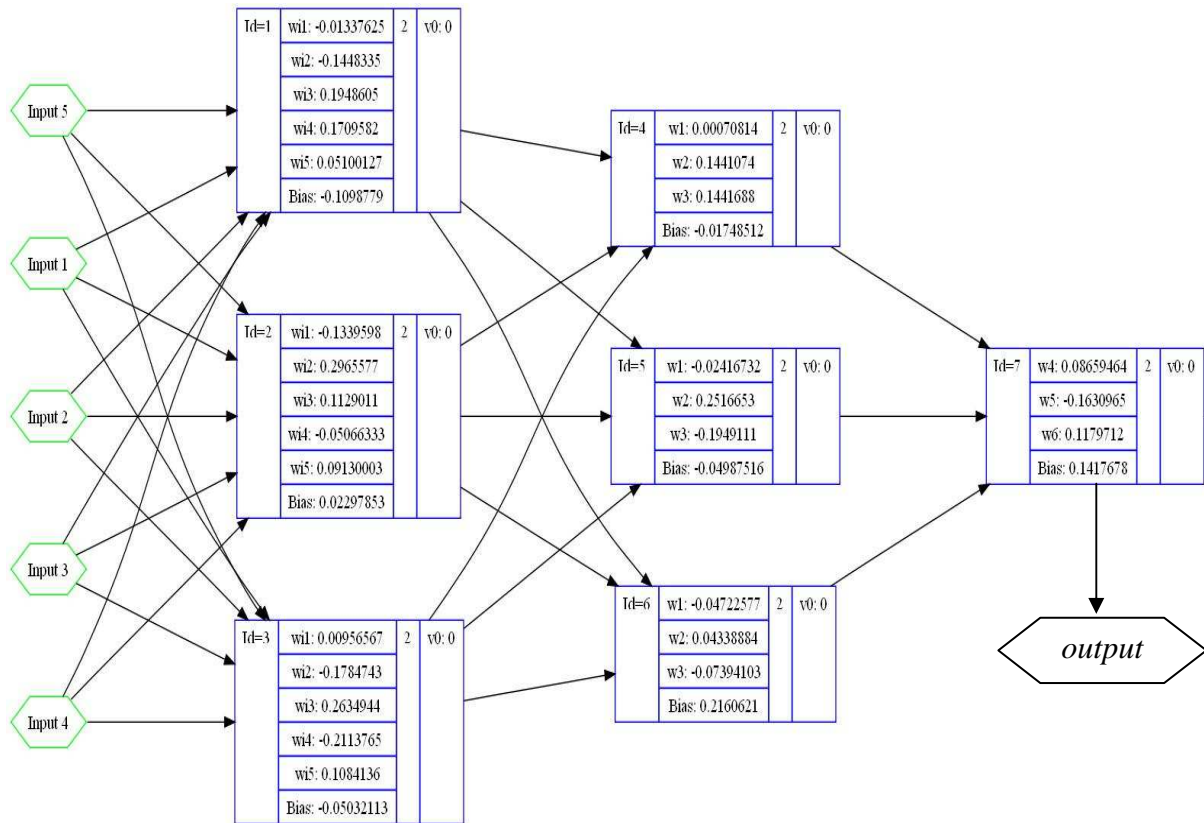


FIGURA 11 – Estrutura da RNA com 5 variáveis de entrada.

Tabela 3 – Resultados da RNA (5:3:3:1)

Estatísticas	Fase	
	Treinamento	Previsão
Erro quadrático médio (raiz)	0,1264	0,1422
Correlação linear	0,2674	0,0996
(<i>p</i> -valor)	(0,000)	(0,4914)
R^2 ajustado	0,0690	-0,0322
Teste de Ljung-Box	4,6143	
(<i>p</i> -valor)	(0,9154)	
Teste de Jarque-Bera	46,3010	
(<i>p</i> -valor)	(0,000)	

Treinamento = primeiras 758 observações;

Previsão = 50 valores estimados após o período de treinamento.

Na Tabela 3, tem-se algumas estatísticas para análise do desempenho da RNA (5:3:3:1), tanto na fase de treinamento quanto na fase de previsão. Na fase de treinamento, a raiz do erro quadrático médio foi de 0,1264 com uma correlação linear significativa de 0,2674. Esses valores garantiram um coeficiente de determinação ajustado de 6,9%. O resíduo da rede se mostra sem dependência serial, porém distinto dos padrões de uma distribuição normal.

A Figura 12 apresenta um diagnóstico visual da RNA (5:3:3:1), através da análise visual constata-se que os valores estimados do retorno do Ibovespa no período de treinamento ficam bem concentrados em torno da média, identificando poucos períodos de instabilidade. Isto, de certa forma, subestima o risco de mercado.

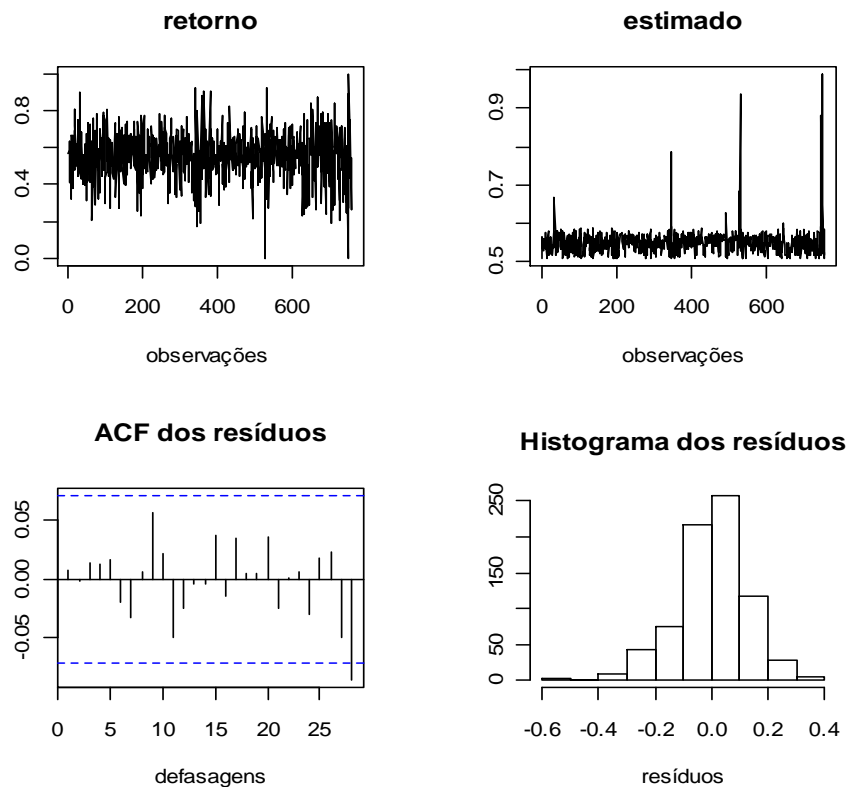


FIGURA 12 – Treinamento da RNA(5:3:3:1).

A função de autocorrelação dos resíduos – ACF e o histograma de frequência evidenciam uma série não autocorrelacionada, mas distinta dos padrões de uma distribuição normal.

4.2 Modelo com informações sobre os retornos e número de negócios

Os resultados da RNA que considera informações defasadas dos retornos e do número de negócios são apresentados na Tabela 4. Sua estrutura é ilustrada na Figura 13. Nesse modelo foram consideradas 10 variáveis de entrada com três funções sigmóides de ativação na primeira camada intermediária e três funções sigmóides de ativação na segunda camada intermediária.

Tabela 4 – Pesos sinápticos da RNA (10:3:3:1)

	F1	F2	F3	G1	G2	G3	K1
Bias	-0,0798	-0,1208	0,1359				
r_{t-1}	-0,1663	0,1310	-0,1654				
r_{t-2}	-0,1912	0,1735	0,0077				
r_{t-3}	0,2463	-0,2443	0,1031				
r_{t-4}	-0,0668	0,0469	-0,0239				
r_{t-5}	-0,2081	0,0113	-0,1320				
neg_{t-1}	-0,0880	-0,0046	-0,2108				
neg_{t-2}	-0,0497	0,1901	-0,0720				
neg_{t-3}	-0,1623	-0,1818	-0,0777				
neg_{t-4}	-0,0346	-0,0294	0,0357				
neg_{t-5}	0,1611	-0,0690	0,0496				
Bias				-0,0191	-0,1155	0,0798	
F1				-0,1238	-0,1323	0,2421	
F2				-0,1387	0,2220	-0,1153	
F3				-0,2372	-0,2258	-0,0150	
Bias							-0,1313
G1							-0,0921
G2							0,1706
G3							0,2169

Os valores relativos a F1, F2 e F3 são da primeira camada intermediária;

Os valores relativos à G1, G2 e G3 são da segunda camada intermediária.

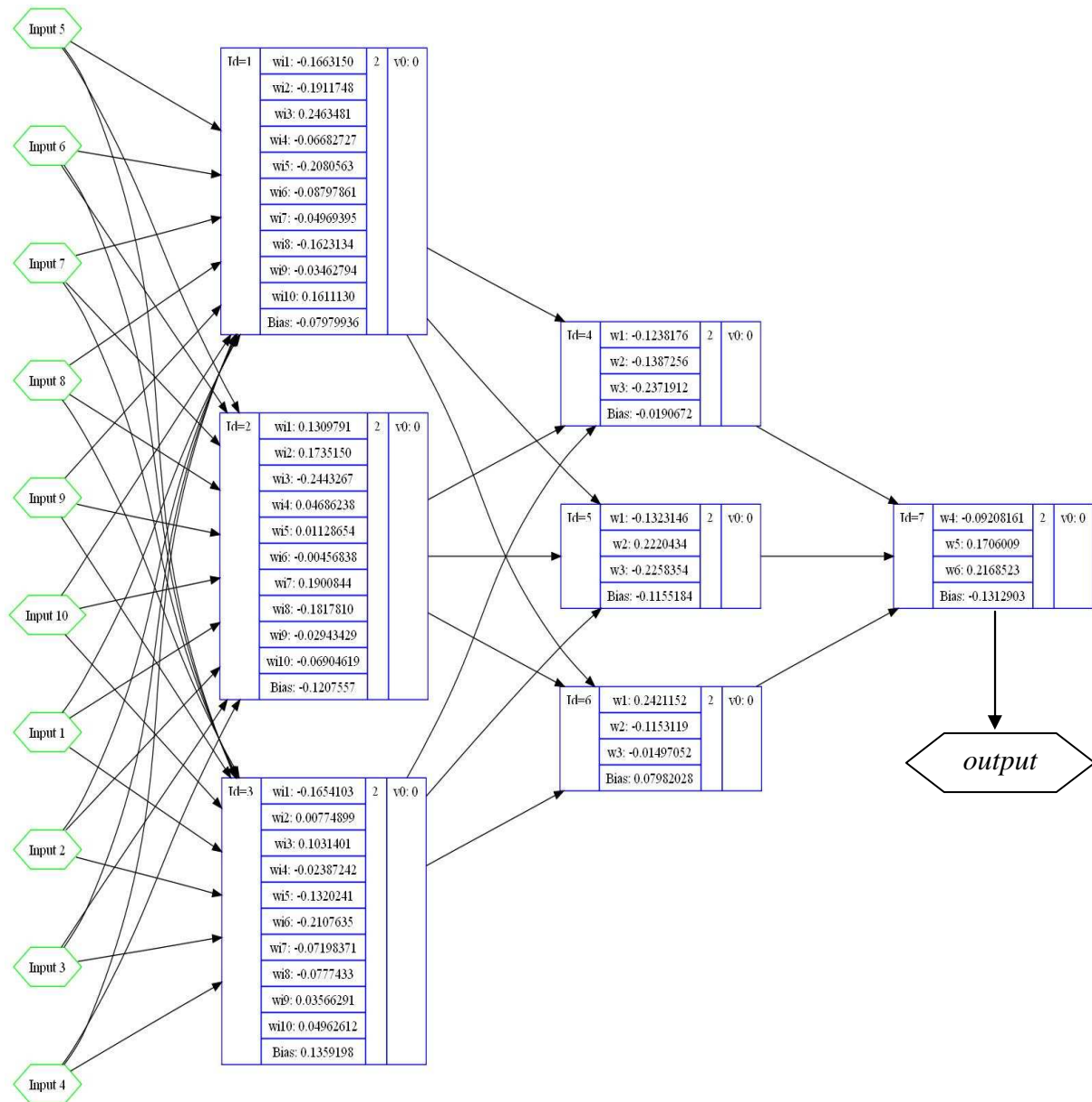


FIGURA 13 – Estrutura da RNA com 10 variáveis de entrada (10:3:3:1).

Na Tabela 5, tem-se os resultados da RNA (10:3:3:1). Na fase de treinamento o erro quadrático médio é de 0,1192 aumentando para 0,1449 na fase de previsão fora da amostra. A correlação linear de 0,4167 passa a ser negativa na ordem de -0,0791. O coeficiente de determinação, de 0,1715 na fase de treinamento passa para -0,0360 na fase de previsão.

Nesse ponto é importante destacar que a inclusão de um maior número de variáveis de entrada (informações sobre as variações do número de negócios) torna possível uma redução do critério de ajuste (raiz do erro quadrático médio) e um aumento do coeficiente de correlação linear e coeficiente de determinação ajustado em relação do modelo que utiliza apenas informação dos retornos passados.

Poder-se-ia pensar que essa melhora significa um ganho de previsibilidade, mas não o é. Fica evidente que, tanto o a raiz do erro quadrático médio quanto à correlação linear e o coeficiente de determinação não conseguem manter o bom desempenho na fase de teste (previsão). O diagnóstico visual do modelo da RNA (10:3:3:1) é ilustrado na Figura 14.

Tabela 5 – Resultados da RNA (10:3:3:1)

Estatísticas	Fase	
	Treinamento	Previsão
Erro quadrático médio (raiz)	0,1192	0,1449
Correlação linear	0,4167	0,0791
(<i>p-valor</i>)	(0,000)	(0,5849)
R ² ajustado	0,1715	-0,0360
Teste de Ljung-Box	8,0310	
(<i>p-valor</i>)	(0,6258)	
Teste de Jarque-Bera	32,2480	
(<i>p-valor</i>)	(0,000)	

Treinamento = primeiras 758 observações;

Previsão = 50 valores estimados após o período de treinamento.

Em termos ilustrativos, a Figura 14 permite uma análise visual da RNA (10:3:3:1). Pode-se constatar que a série estimada de retorno do Ibovespa apresenta-se com maior oscilação torno da média se comparado com a estimativa da RNA (5:3:3:1).

Com uma maior dispersão, tem-se uma expectativa mais realista do risco associado ao mercado financeiro. A função de autocorrelação dos resíduos (ACF) identifica um comportamento não autocorrelacionado serialmente. O histograma de frequência identifica um comportamento mais próximo de um formato de “sino”, porém ainda muito distinto dos padrões de uma curva normal.

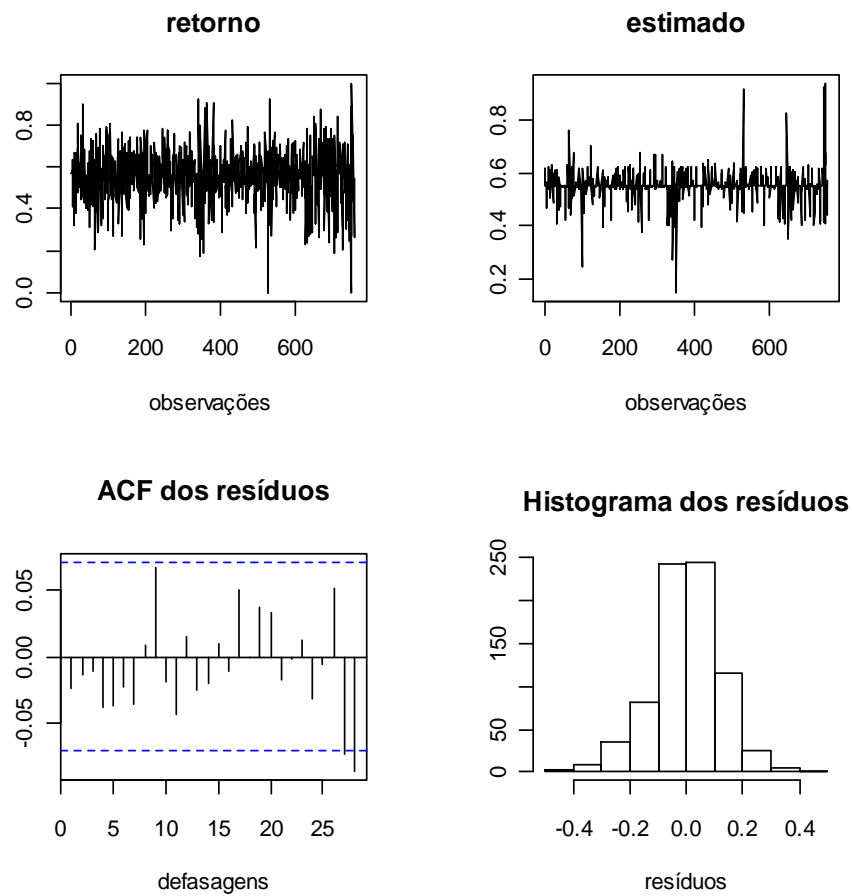


FIGURA 14 – Treinamento da RNA (10:3:3:1).

Até o momento, pode-se destacar que acrescentando informações sobre o número de negócios tem-se um ajuste mais preciso na fase de treinamento, mas que não se mantém na fase de teste.

4.3 Modelo com informações sobre os retornos, número de negócios e volume financeiro

O terceiro modelo de RNA aplicado envolve a utilização de informações defasadas do retorno, variação do número de negócios e variação do volume financeiro como variáveis explanatórias das variações do Ibovespa. Os resultados dos pesos sinápticos são apresentados na Tabela 6.

Tabela 6 – Pesos sinápticos da RNA (15:3:3:1)

	F1	F2	F3	G1	G2	G3	K1
Bias	-0,0860	-0,0963	-0,0365				
r_{t-1}	0,1001	0,1433	-0,0789				
r_{t-2}	-0,1978	-0,0921	0,1854				
r_{t-3}	0,1219	-0,0941	0,1741				
r_{t-4}	0,0544	-0,0940	-0,0706				
r_{t-5}	0,1435	-0,0551	-0,1235				
neg_{t-1}	-0,0793	-0,0561	0,0865				
neg_{t-2}	-0,0563	-0,0690	0,1398				
neg_{t-3}	0,0710	0,0360	0,0914				
neg_{t-4}	-0,2018	-0,0860	-0,1555				
neg_{t-5}	0,1833	0,1404	0,0387				
vol_{t-1}	0,1743	0,0441	0,0671				
vol_{t-2}	-0,0295	-0,0558	-0,1354				
vol_{t-3}	-0,1578	0,1277	0,2006				
vol_{t-4}	0,1847	0,1564	-0,0752				
vol_{t-5}	0,0302	0,0808	0,2000				
Bias				0,0243	-0,2155	-0,0089	
F1				-0,0092	0,0988	-0,0757	
F2				0,1521	-0,1961	-0,1492	
F3				0,1169	-0,0559	-0,1238	
Bias							0,1136
G1							-0,0767
G2							-0,0700
G3							-0,0095

A estrutura da RNA estimada com 15 variáveis de entrada, três neurônios na primeira camada intermediária e outros três na segunda camada intermediária é ilustrada na Figura 15.

Na Tabela 7, tem-se os resultados da RNA (15:3:3:1) ajustada. A fase de treinamento da RNA que considerou 15 variáveis de entrada possibilitou a obtenção de valores mais ajustados do que os modelos anteriores. A raiz o erro quadrático médio foi reduzido para 0,1139. A correlação linear aumentou para 0,4992 e o coeficiente de determinação ajustado indica que o modelo treinado prediz 24,30% das variações do Ibovespa.

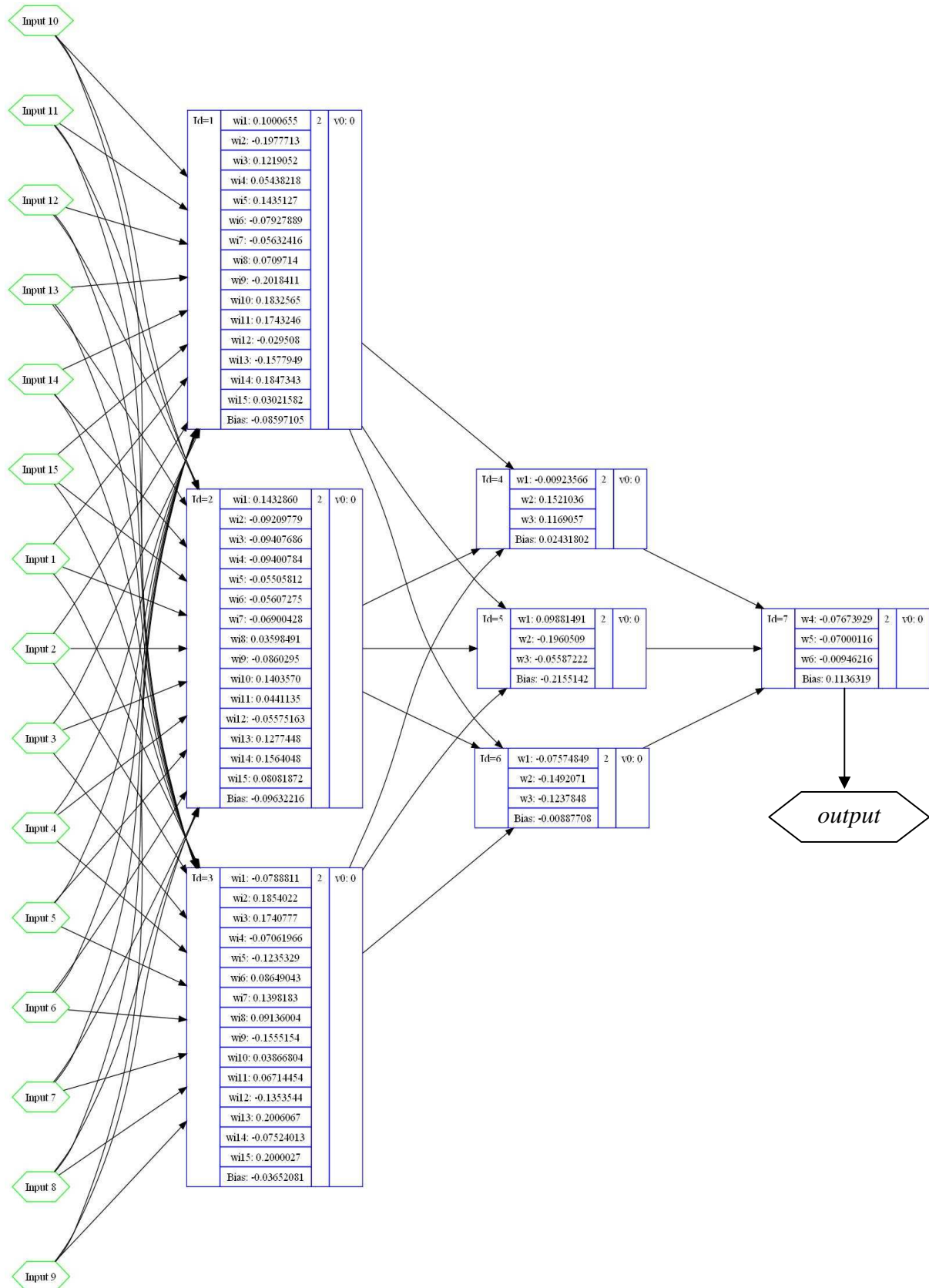


FIGURA 15 – Estrutura da RNA com 15 variáveis de entrada (15:3:3:1).

Tabela 7 – Resultados da RNA(15:3:3:1)

Estatísticas	Fase	
	Treinamento	Previsão
Erro quadrático médio (raiz)	0,1139	0,1432
Correlação linear	0,4992	0,1660
(<i>p</i> -valor)	(0,000)	(0,2475)
R ² ajustado	0,2430	-0,0136
Teste de Ljung-Box	10,4159	
(<i>p</i> -valor)	(0,4048)	
Teste de Jarque-Bera	50,0790	
(<i>p</i> -valor)	(0,000)	

Treinamento = primeiras 758 observações;

Previsão = 50 valores estimados após o período de treinamento.

Em termos de mercado financeiro, os valores obtidos pela RNA (15:3:3:1) podem ser considerados de boa performance. Porém, o modelo ajustado não consegue manter sua qualidade na fase de previsão. Dessa forma, parece que a boa qualidade do ajuste do modelo não é garantia de igual qualidade na fase de previsão. A Figura 16 permite um diagnóstico visual desse ajuste.

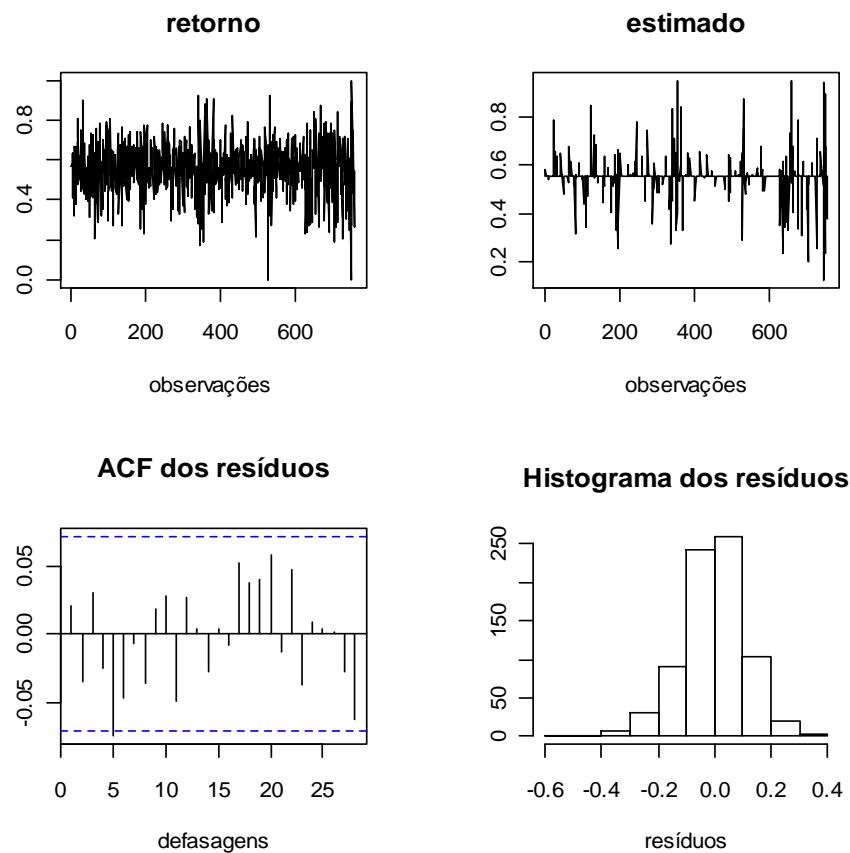


FIGURA 16 – Treinamento da RNA (15:3:3:1).

Observando-se a série estimada pela RNA (15:3:3:1) pode-se perceber que ela apresenta uma oscilação maior em torno da média do que as estimativas dos modelos anteriores. Dessa forma tem-se uma noção ainda mais realista do risco de mercado. Como nos modelos anteriores, a função de autocorrelação dos resíduos e seu histograma, também identificam resíduos não autocorrelacionados e distribuídos de forma distinta de uma distribuição normal.

Analisando os resultados da aplicação dos três modelos de Redes Neurais Artificiais, pode-se constatar que com o aumento do número de variáveis de entrada, ou seja, considerando dados sobre o número de negócios além dos dados sobre as variações de preços, é possível melhorar o ajuste no período de treinamento.

Esse ajuste fica ainda melhor com acréscimo de dados sobre o volume financeiro, a raiz do erro quadrático médio diminui e a correlação linear aumenta. Porém, a melhoria na fase de treinamento não se mantém na fase de previsão, indicando, dessa forma, que dados sobre o número de negócios e volume financeiro não são relevantes para a previsão das variações de preço do Ibovespa.

5 CONCLUSÃO

Este estudo foi desenvolvido com o objetivo de verificar a utilidade de se considerar informações sobre o número de negócios e o volume financeiro negociado em bolsa como variáveis explanatórias das variações do Ibovespa. O que se está procurando é tentar melhorar um modelo de previsão que se utiliza apenas dos dados das variações de preços passados para fornecer uma estimativa de variações de preço no futuro. A investigação é baseada na aplicação da metodologia de Redes Neurais Artificiais com duas camadas intermediárias e com funções de ativação sigmóides.

São construídos três modelos de redes neurais artificiais. Os modelos são estruturados com duas camadas intermediárias. No primeiro modelo são utilizados apenas os dados sobre as variações nos preços dos últimos cinco dias. No segundo modelo são acrescentadas ao primeiro modelo dados dos últimos cinco dias do número de negócios realizados. No terceiro modelo são acrescentados ao segundo modelo dados dos últimos cinco dias do volume financeiro movimentado pelos negócios na bolsa.

Após o treinamento dos três modelos, procedeu-se o cálculo da raiz do erro quadrático médio, coeficiente de determinação ajustado e grau de correlação. Também foram calculados o teste de Ljung-Box e Jarque-Bera a fim de verificar se os modelos foram ajustados adequadamente. Todos os testes aplicados identificaram que os resíduos dos modelos não se apresentaram como uma distribuição normal, mas não apresentaram dependência, assim podem ser considerado como sendo ruído branco.

Conclui-se que o terceiro modelo, a RNA que considerou informações sobre o número de títulos e o volume financeiro negociado na bolsa de valores além de informações sobre as variações de preço, fornece um ajuste melhor do que o modelo que considera apenas informações de retornos passados. Essa afirmação é válida para o período de treinamento.

Porém, na fase de previsão a RNA que melhor se ajustou a série do Ibovespa na fase de treinamento não conseguiu manter as suas propriedades, oferecendo uma melhor estimativa, ou seja, as propriedades de ajustes se deterioraram no período fora da amostra.

Dessa forma, pode-se afirmar que as informações defasadas sobre número de títulos e volume financeiro negociado em até cinco dias não são úteis para previsão de futuro comportamento das oscilações do mercado financeiro.

Esses resultados vêm corroborar com a hipótese de eficiência na forma fraca do mercado financeiro, ou seja, informações passadas sobre preços são inúteis para modelos de previsibilidade com objetivos de ganhos anormais, visto que os preços já refletem totalmente as informações disponíveis.

REFERÊNCIAS BIBLIOGRÁFICAS

BACHELIER, L. Theory of speculation (1900). In COOTNER, P. The random character of stock market prices. Cambridge: MIT Press, pag. 17-78, 1964.

BOX, G. E. P.; PIERCE, D. A. Distributions of residual autocorrelations in autoregressive integrated moving average models. *Journal of the American Statistics Association*, v. 65, p. 1509-1526, 1970.

BRAGA, A. P.; CARVALHO, A. P. L.; LUDERMIR, T. B. Redes neurais artificiais. Rio de Janeiro: LTC, 2007.

BROOKS, C. *Introductory econometrics for finance*. Cambridge, 2008

CERETTA, P. S.; COSTA, JR., N. C. A. Caminho aleatório: revisão, síntese e novos resultados nos mercados de capitais da América Latina. *Revista de Administração (RA-USP)*, Vol. 37, n. 3, 2002.

FAMA, E. F.; FRENCH, K. R. Permanent and temporary components of stock prices. *Journal of Political Economics*, Vol, 96, n. 2, pag. 246-273, 1988.

FRANSES, P. H.; DIJK, D. V. *Non-linear time series models in empirical finance*. New York: Cambridge University Press, 2003.

GENÇAY, B.; SELÇUK, F.; WHITCHER, B. An introduction to wavelets and other filtering methods in finance and economics. California: Elsevier, 2002.

HAYKIN, S. *Redes neurais: princípios e práticas*. Porto Alegre: Bookman, 2001.

HEBB, D. O. *The organization of behavior*. Wiley, 1949.

KAMRUZZAMAN, J.; BEGG, R.; SARKER, R. *Artificial neural networks in finance and*

manufacturing. London: Idea Group Publishing, 2006.

KENDALL, L. Theory analysis of economic time series – Part I: Prices (1953). In COOTNER, P. *The random character of stock market prices*. Cambridge: MIT Press, pag. 85-99, 1964.

KIRCHGASSNER, G.; WOLTERS, J. *Introduction to modern time series analysis*. Springer, 2008.

KUAN, C. M.; WHITE, H. Artificial networks: an econometric perspective (with discussion). *Econometric Reviews*, Vol.13, p. 1-143, 1994.

LATHI, B. P. *Sinais e sistemas lineares*. Porto Alegre: Bookman, 2007.

LJUNG, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. *Biometrika*, v. 65(2), p. 297-303, 1978.

LO, A. W.; MACKINLAY, c. Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies*, Vol. 1, n. 1, pag. 41-46, 1988.

MCCULLOCH, W.S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysycs*, v. 5, p. 115-133, 1943.

McNELIS, P. D. *Neural networks in finance*. California: Elsevier Academic Press, 2005.

MILLS, T. C. *The econometric modeling of financial time series*. New york: Cambridge University Press, 1999.

OLIVEIRA, M. A.; MONTINI, A. A.; BERGMANN, D. R. *Previsão de retornos de ações de empresas dos setores financeiro, alimentos, industrial e de serviços por meio de Redes Neurais e modelos ARIMA-GARCH*. Anais do Enanpad 2007. XXX Encontro da ANPAD. Rio de Janeiro, 2007.

POTERBA, J.; SUMMER, L. Mean reversion in stock prices: evidences and implications. *Journal of Financial Economics*, Vol. 22, pag. 27-59, 1988.

QI, M. *Financial applications of artificial neural networks*. In MADDALA, G. S.; RAO, C. D. Handbook of statistics. Vol. 14: Statistical methods in finance. Amsterdam: Elsevier

Science, p. 529-552, 1996.

SAMUELSON, P. A. Proof that property anticipated prices fluctuate randomly. *Industrial Management Review*, Vol. 6, pag. 41-49, 1965.

ZHANG, G. P. *Neural networks in business forecasting*. London: Idea Group Publishing, 2004.

ANEXO A

Programação do modelo RNA com cinco variáveis de entrada


```

setwd("F:/Monografia")
da=read.table("todosMONO.txt")
daf=read.table("todosforaMONO.txt")
library(AMORE)
target =da [,1]          # y treinamento
targetf=daf[,1]         # y fora da amostra
imput <- matrix(c(da[,2],da[,3],da[,4],da[,5],da[,6]),758, 5) # Imput treino
imputf <- matrix(c(daf[,2],daf[,3],daf[,4],daf[,5],daf[,6]),50, 5) # Imput fora da amostra
net <-
  newff(n.neurons=c(5,3,3,1),
        learning.rate.global=1e-2, # Modelo
        momentum.global=0.5,
        error.criterium="LMS",
        Stao=NA,
        hidden.layer="sigmoid",
        output.layer="sigmoid",
        method="ADAPTgdwm")
result <-
  train(net, imput, target, error.criterium="LMS",
        report=TRUE, show.step=1000, n.shows=1000)
treino <- sim(result$net, imput)
graphviz.MLPnet(net, 2)
eqm1=(mean((target-treino)^2))
reqm1=reqm1^0.5
fora <- sim(result$net, imputf)
eqm2=(mean((targetf-fora)^2))
reqm2=reqm2^0.5
reqm2
cor.test(treino,target)
cor.test(fora,targetf)
r2ajustado=1-(((758-1)/(758-3))*(1-cor(treino, target)^2))
r2ajustado
r2ajustadof=1-(((50-1)/(50-3))*(1-cor(for, targetf)^2))
r2ajustadof
Box.test((target-treino), lag=10, type="Ljung")
jbTest(target-treino)
par(mfrow=c(2,2))
plot.ts(target, main="retorno", xlab="observações",ylab="")
plot.ts(treino, main="estimado", xlab="observações", ylab="")
acf(target-treino, main="ACF dos resíduos", xlab="defasagens",ylab="")
hist(target-treino,main="Histograma dos resíduos",xlab="resíduos", ylab=" ")
net$neurons[[1]]$weight # Pesos
net$neurons[[2]]$weight
net$neurons[[3]]$weight
net$neurons[[4]]$weight
net$neurons[[5]]$weight
net$neurons[[6]]$weight
net$neurons[[7]]$weight
net$neurons[[1]]$bias # Constante
net$neurons[[2]]$bias
net$neurons[[3]]$bias
net$neurons[[4]]$bias
net$neurons[[5]]$bias
net$neurons[[6]]$bias
net$neurons[[7]]$bias

```

ANEXO B

Programação do modelo RNA com dez variáveis de entrada

```

setwd("F:/Monografia")
da=read.table("todosMONO.txt")
daf=read.table("todosforaMONO.txt")
library(AMORE)
target =da [,1]          # y treinamento
targetf=daf[,1]         # y fora da amostra
imput <- matrix(c(da[,2],da[,3],da[,4],da[,5],da[,6],
                 da[,7],da[,8],da[,9],da[,10],da[,11]),758, 10) # Imput treino

imputf <- matrix(c(daf[,2],daf[,3],daf[,4],daf[,5],daf[,6],
                 daf[,7],daf[,8],daf[,9],daf[,10],daf[,11]),50, 10) # Imput fora da amostra
net <-
  newff(n.neurons=c(10,3,3,1),
        learning.rate.global=1e-2, # Modelo
        momentum.global=0.5,
        error.criterium="LMS",
        Stao=NA,
        hidden.layer="sigmoid",
        output.layer="sigmoid",
        method="ADAPTgdwm")
result <-
  train(net, imput, target, error.criterium="LMS",
        report=TRUE, show.step=1000, n.shows=1000)
treino <- sim(result$net, imput)
graphviz.MLPnet(net, 2)
eqm1=(mean((target-treino)^2))
reqm1=eqm1^0.5
fora <- sim(result$net, imputf)
eqm2=(mean((targetf-fora)^2))
reqm2=eqm2^0.5
reqm2
cor.test(treino,target)
cor.test(fora,targetf)
r2ajustado=1-(((758-1)/(758-3))*(1-cor(treino, target)^2))
r2ajustado
r2ajustadof=1-(((50-1)/(50-3))*(1-cor(for, targetf)^2))
r2ajustadof
Box.test((target-treino), lag=10, type="Ljung")
jbTest(target-treino)
par(mfrow=c(2,2))
plot.ts(target, main="retorno", xlab="observações",ylab="")
plot.ts(treino, main="estimado", xlab="observações", ylab="")
acf(target-treino, main="ACF dos resíduos", xlab="defasagens",ylab="")
hist(target-treino,main="Histograma dos resíduos",xlab="resíduos", ylab=" ")
net$neurons[[1]]$weight # Pesos
net$neurons[[2]]$weight
net$neurons[[3]]$weight
net$neurons[[4]]$weight
net$neurons[[5]]$weight
net$neurons[[6]]$weight
net$neurons[[7]]$weight
net$neurons[[1]]$bias # Constante
net$neurons[[2]]$bias
net$neurons[[3]]$bias
net$neurons[[4]]$bias
net$neurons[[5]]$bias
net$neurons[[6]]$bias
net$neurons[[7]]$bias

```

ANEXO C

Programação do modelo RNA com quinze variáveis de entrada

```

setwd("F:/Monografia")
da=read.table("todosMONO.txt")
daf=read.table("todosforaMONO.txt")
library(AMORE)
target =da [,1]          # y treinamento
targetf=daf[,1]         # y fora da amostra
imput <- matrix(c(da[,2],da[,3],da[,4],da[,5],da[,6],
                 da[,7],da[,8],da[,9],da[,10],da[,11],
                 da[,12],da[,13],da[,14],da[,15],da[,16]),758, 15) # Imput treino

imputf <- matrix(c(daf[,2],daf[,3],daf[,4],daf[,5],daf[,6],
                 daf[,7],daf[,8],daf[,9],daf[,10],daf[,11],
                 daf[,12],daf[,13],daf[,14],daf[,15],daf[,16]),50, 15) # Imput fora da amostra
net <-
  newff(n.neurons=c(15,3,3,1),
        learning.rate.global=1e-2, # Modelo
        momentum.global=0.5,
        error.criterium="LMS",
        Stao=NA,
        hidden.layer="sigmoid",
        output.layer="sigmoid",
        method="ADAPTgdwm")
result <-
  train(net, imput, target, error.criterium="LMS",
        report=TRUE, show.step=1000, n.shows=1000)
treino <- sim(result$net, imput)
graphviz.MLPnet(net, 2)
eqm1=(mean((target-treino)^2))
reqm1=reqm1^0.5
reqm1
fora <- sim(result$net, imputf)
eqm2=(mean((targetf-fora)^2))
reqm2=reqm2^0.5
reqm2
cor.test(treino,target)
cor.test(fora,targetf)
r2ajustado=1-(((758-1)/(758-3))*(1-cor(treino, target)^2))
r2ajustado
r2ajustadof=1-(((50-1)/(50-3))*(1-cor(for, targetf)^2))
r2ajustadof
Box.test((target-treino), lag=10, type="Ljung")
jbTest(target-treino)
par(mfrow=c(2,2))
plot.ts(target, main="retorno", xlab="observações",ylab="")
plot.ts(treino, main="estimado", xlab="observações", ylab="")
acf(target-treino, main="ACF dos resíduos", xlab="defasagens",ylab="")
hist(target-treino,main="Histograma dos resíduos",xlab="resíduos", ylab=" ")
net$neurons[[1]]$weight # Pesos
net$neurons[[2]]$weight
net$neurons[[3]]$weight
net$neurons[[4]]$weight
net$neurons[[5]]$weight
net$neurons[[6]]$weight
net$neurons[[7]]$weight
net$neurons[[1]]$bias # Constante
net$neurons[[2]]$bias
net$neurons[[3]]$bias
net$neurons[[4]]$bias
net$neurons[[5]]$bias
net$neurons[[6]]$bias
net$neurons[[7]]$bias

```