

EDGARD DEVANIR AMOROSO

**UM ESTUDO DE CASO SOBRE MINERAÇÃO DE DADOS COMO INSTRUMENTO
DE APRENDIZADO PARA O INVESTIDOR DO MERCADO DE AÇÕES**

Dissertação apresentada ao Programa de Pós-Graduação “*Stricto-Sensu*” em Gestão do Conhecimento e Tecnologia da Informação da Universidade Católica de Brasília, como requisito para a obtenção do Título de Mestre.

Orientador: Prof. Dr. Hércules Antonio do Prado

Co-orientador: Prof. Dr. Edilson Ferneda

Brasília
2009

A524e Amoroso, Edgard Devanir.

Um estudo de caso sobre mineração de dados como instrumento de
aprendizado para o investidor do mercado de ações / Edgar Devanir
Amoroso. – 2009.

121 f. ; il. ; 30 cm

Dissertação (mestrado) – Universidade Católica de Brasília, 2009.

Orientação: Francisco Hércules Antonio do Prado

Co-orientação: Edilson Ferneda

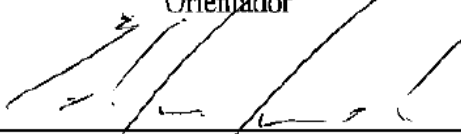
1. Gestão do conhecimento. 2. Investimento – Bolsa de valores.
3. Mineração de dados I. Prado, Francisco Hércules Antonio do, orient.
II. Ferneda, Edilson. III. Título.

CDU 005.94:33

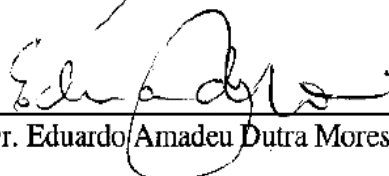
Dissertação de autoria de Edgar Devanir Amoroso, intitulada “Um Estudo de Caso sobre Mineração de Dados como Instrumento de Aprendizado para o Investidor do Mercado de Ações”, requisito parcial para obtenção do grau de Mestre em Gestão do Conhecimento e da Tecnologia da Informação, defendida e aprovada, em 26 de fevereiro de 2009, pela banca examinadora constituída por:



Prof. Dr. Hércules Antonio de Prado
Orientador



Prof. Dr. Edilson Farneda



Prof. Dr. Eduardo Amadeu Dutra Moresi



Profa. Dra. Lourdes Mattos Brasil

Brasília

2009

Dedico o presente estudo a todos os
professores que participaram, direta ou
indiretamente deste, apoiando-o e
subsidiando-o tecnicamente.
A todos minha gratidão.

AGRADECIMENTOS

Aos Prof. Hércules Antonio Prado e Prof. Edilson Ferneda, que generosamente dedicaram tempo e sabedoria na condução e orientação dos trabalhos. Sempre de forma paciente, amável e acolhedora, ensinaram-me a superar os obstáculos e não desanimar.

À Prof^a. Dra. Lourdes Mattos Brasil pelo apoio e pronta colaboração nos momentos que tive as minhas dificuldades e dúvidas.

A todos os professores do Mestrado em Gestão do Conhecimento e da Tecnologia da Informação com os quais pude conviver e que muito contribuíram com seus ensinamentos e experiência.

Aos meus familiares pelo apoio incondicional e constante incentivo, compreendendo minhas ausências e o cansaço, que, por vezes, me abateram e me tornaram impaciente.

À minha querida e amada esposa Sônia que, além de exemplo, muitas vezes deixou seus afazeres cotidianos, para me auxiliar a enfrentar os desafios.

A todos os professores participantes da pesquisa, que se dispuseram, graciosamente, mesmo que de forma anônima, a contribuir com este estudo. Não fosse sua disponibilidade teria sido árduo terminá-lo.

E, sobretudo, a Deus, que inegavelmente esteve ao meu lado, guiando meus passos, iluminando meu caminho e dando-me forças para não desanimar. Com sua ajuda pude crer que tudo é possível quando se tem fé.

RESUMO

Amoroso, Edgard Devanir. **Um estudo de caso sobre mineração de dados como instrumento de aprendizado para o investidor do mercado de ações**. 2009. 121 f. Dissertação (Mestrado) - Programa de Pós-Graduação "*Stricto-Sensu*" em Gestão do Conhecimento e Tecnologia da Informação, Universidade Católica de Brasília, Brasília, 2009.

O processo de decisão sobre investimentos no mercado de ações envolve a utilização de informações oriundas de diversas fontes: a experiência própria do investidor, meios de comunicação em geral e resultados de estudos baseados em modelos. Nesta última fonte, usualmente, são utilizados dois modelos: o fundamentalista e o grafista, o primeiro baseado nos princípios de governança da organização em que se quer investir, e o segundo tomando como base os limites mínimo e máximo dos preços das cotações das ações, tratados segundo um modelo estatístico. O foco desta pesquisa é o estudo de modelos de Mineração de Dados como instrumento de aprendizado ao investidor, complementando as fontes já utilizadas. Foi estudado um modelo de séries temporais baseado em redes neurais artificiais, utilizando o método CRISP-DM para desenvolvimento da aplicação. Foram utilizados dados correspondentes a dez anos de cotações de ações de seis empresas, negociadas na Bovespa. O estudo evidenciou que, mais do que a predição de valores precisos para as cotações, tal modelo serve como importante fonte de aprendizado para os investidores.

Palavras-chave: Gestão do Conhecimento. Aprendizagem Corporativa. Modelos Preditivos. Mineração de Dados. Mercado de Ações.

ABSTRACT

The decision process on investment in the stock market involves the use of information from several sources: the experience of the investor, media in general and results of studies based on models. This last source, usually used two models: the fundamentalist and the graphic, the first based on the principles of governance of the organization in which you want to invest, and the second on the basis the minimum and maximum limits of price quotations of the shares, treated according a statistical model. The focus of this research is the study of models of Data Mining as instruments of learning to the investor, complementing the sources already used. We studied a model of time series based on artificial neural networks, using the CRISP-DM to the development application. We used data corresponding to ten years of stock prices of six companies, traded on Bovespa. The study showed that more than the prediction of precise values for the price, this model serves as an important source of learning for investors.

Keywords: Knowledge Management. Corporative Learning. Data Mining. Stock Market. Predictive Models.

LISTA DE QUADROS

Quadro 1 - Os mercados do Dinheiro.....	26
Quadro 2 - Registro - 01 - a Cotações dos papéis por dia	44
Quadro 3 - Detalhes do Registro 00 – Header	45
Quadro 4 - Detalhes do Registro 99 – Trailer.....	45
Quadro 5 - Empresas componentes da carteira de ações considerada.....	51
Quadro 6 - Campos selecionados para a realização dos trabalhos	52
Quadro 7 - Empresas e tipos de ações	53

LISTA DE FIGURAS

Figura 1 - Espiral do Conhecimento, segundo Nonaka e Takeuchi.....	24
Figura 2 - Representação simplificada de um neurônio.	34
Figura 3 - Modelo matemático de um neurônio.	34
Figura 4 - Exemplo de uma RNA.....	35
Figura 5 - Exemplos de séries temporais.	38
Figura 6 - Tendência de uma série.....	39
Figura 7 - Fases da metodologia CRISP-DM	39
Figura 8 - Ferramenta Easynn – Seleccionando registros para <i>querying</i>	57
Figura 9 - Ferramenta Easynn - RNA.....	58
Figura 10 - Ferramenta MatLab – Criação RNA.....	61
Figura 11 - Compreensão do Negócio	82
Figura 12 - Compreensão dos Dados	84
Figura 13 - Preparação dos Dados	86
Figura 14 - Modelagem	87
Figura 15 - Universo de Técnicas.....	87
Figura 16 - Avaliação	88
Figura 17 - Desenvolvimento.....	90
Figura 18 - Rede 1	116
Figura 19 - Rede 2	116
Figura 20 - Rede 3	117
Figura 21 - Rede 4	117
Figura 22 - Rede 5	118
Figura 23 - Rede 6	118
Figura 24 - Rede 7	119
Figura 25 - Rede 8	119
Figura 26 - Rede 9	120
Figura 27 - Rede 10	120
Figura 28 - Rede 11	121
Figura 29 - Rede 12	121

LISTA DE TABELAS

Tabela 1 -	Tabela com tipos de investidores da Bovespa	29
Tabela 2 -	Ações do tipo ON e PN movimentadas no período considerado	49
Tabela 3 -	Total de cotações referentes da carteira de ações considerada	51
Tabela 4 -	Total de cotações de ações ON referentes a ações da carteira considerada.....	53
Tabela 5 –	Treinamento para geração de teste de desempenho.....	63
Tabela 6 -	Campos de entrada e saída das RNA.....	64
Tabela 7 -	Redes 1 a 6 com predição de 30 dias corridos.	65
Tabela 8 -	Redes 7 a 12 com predição de 15 dias corridos.	65
Tabela 9 -	Redes de 1 a 6 e a importância de seus campos de entrada no aprendizado	66
Tabela 10 -	Redes de 7 a 12 e a importância de seus campos de entrada no aprendizado.	67
Tabela 11 -	Variação das predições de cotações da empresa Telesp	70
Tabela 12 -	Variação das predições de cotações da empresa Petrobrás	70
Tabela 13 -	Variação das predições de cotações da empresa Embraer	71
Tabela 14 -	Variação das predições de cotações da empresa Bradesco.....	71
Tabela 15 -	Variação das predições de cotações da empresa Banco do Brasil.....	72
Tabela 16 –	Variação das predições de cotações da empresa Vale do Rio Doce...	72
Tabela 17 -	Tabela com dados da empresa Telesp – período de 30 dias.....	96
Tabela 18 -	Tabela com dados da empresa Telesp – período de 15 dias.....	97
Tabela 19 –	Tabela com dados da empresa Petrobrás – período de 30 dias.....	98
Tabela 20 –	Tabela com dados da empresa Petrobrás – período de 15 dias.....	99
Tabela 21 –	Tabela com dados da empresa Embraer – período 30 dias.....	100
Tabela 22 –	Tabela com dados da empresa Embraer – período 15 dias.....	101
Tabela 23 –	Tabela com dados da empresa Bradesco – período de 30 dias.	102
Tabela 24 –	Tabela com dados da empresa Bradesco – período de 15 dias.	103
Tabela 25 –	Tabela com dados da empresa Banco do Brasil – período 30 dias. ..	104
Tabela 26 –	Tabela com dados da empresa Banco do Brasil – período 15 dias. ..	105
Tabela 27 –	Tabela com dados da empresa Vale do Rio Doce – predição de 30 dias.....	106
Tabela 28 –	Tabela com dados da empresa Vale do Rio Doce – predição de 15 dias.....	107
Tabela 29 –	Tabela com dados para validação – Janeiro 2008.....	108

Tabela 30 – Tabela com dados para validação – Fevereiro 2008.....	109
Tabela 31 – Tabela com dados para validação – Março 2008.....	110
Tabela 32 – Tabela com dados para validação – Abril 2008	111
Tabela 33 – Tabela com dados para validação – Maio 2008.....	112
Tabela 34 – Tabela com dados para validação – Junho 2008.....	113
Tabela 35 – Tabela com dados para validação – Julho 2008	114
Tabela 36 – Tabela com dados para validação – Agosto 2008	115

LISTA DE SIGLAS

BBM	– Bolsa Brasileira de Mercadorias
BM&F	– Bolsa de Mercadorias e Futuros
BOVESBA	– Bolsa de Valores da Bahia
BOVESPA	– Bolsa de Valores de São Paulo
BOVMESB	– Bolsa de Valores de Minas Gerais, Espírito Santo e Brasília
BVRJ	– Bolsa de Valores do Rio de Janeiro
CETIP	– Câmara de Custódia e Liquidação
CNBV	– Comissão Nacional de Bolsas de Valores
CRISP-DM	– CRoss Industry Standard Process of Data Mining
DCBD	– Descoberta de Conhecimentos em Banco de Dados
GC	– Gestão do Conhecimento
KDD	– Knowledge Discovery Data
MatLab	– Matrix Laboratory
ON	– Ordinárias Nominativas
PN	– Preferenciais Nominativas
RAM	– Randon Acess Memory
RNA	– Redes Neurais Artificiais
TWI	– Taiwan Weighted Index (Índice da bolsa de valores tailandesa)
WEKA	– Waikato Environment Knowledge

SUMÁRIO

1 INTRODUÇÃO.....	15
1.1 CONTEXTUALIZAÇÃO	15
1.2 REVISÃO DE LITERATURA	15
1.3 JUSTIFICATIVA.....	20
1.4 FORMULAÇÃO DO PROBLEMA	21
1.5 OBJETIVOS	21
1.5.1 Objetivo geral.....	21
1.5.2 Objetivos específicos.....	21
1.6 ORGANIZAÇÃO DA DISSERTAÇÃO.....	22
2 REFERENCIAL TEÓRICO	23
2.1 GESTÃO DO CONHECIMENTO.....	23
2.2 MERCADO DE CAPITALIS	25
2.3 MERCADO DE AÇÕES.....	26
2.3.1 Considerações iniciais.....	26
2.3.2 A BOVESPA	28
2.3.3 Ações.....	29
2.3.4 Tipos de ações.....	30
2.3.5 Análise de preço de ações	30
2.4 DESCOBERTA DO CONHECIMENTO EM BANCO DE DADOS.....	31
2.5 REDES NEURAIS ARTIFICIAIS.....	33
2.6 SÉRIES TEMPORAIS	37
2.7 CRISP-DM.....	39
2.8 APRENDIZAGEM POR SIMULAÇÃO	41
3 MATERIAL E MÉTODOS.....	43
3.1 CLASSIFICAÇÃO DA PESQUISA.....	43
3.2 PRESSUPOSTOS	43
3.3 COLETA E ANÁLISE DOS DADOS	43
3.4 DELIMITAÇÃO DO ESTUDO	45
4 ESTUDO DE CASO.....	46
4.1 FASE 1: COMPREENSÃO DO NEGÓCIO.....	46
4.1.1 Determinação dos objetivos do negócio.....	46
4.1.2 Avaliação da situação	47
4.1.3 Determinação das metas de mineração de dados.....	47

4.1.4	Produção do plano de projeto.....	48
4.2	FASE 2: COMPREENSÃO DOS DADOS.....	48
4.2.1	Coleta dos dados iniciais.....	48
4.2.2	Descrição dos dados	49
4.2.3	Exploração dos dados	49
4.2.4	Verificação da qualidade dos dados.....	50
4.3	FASE 3: PREPARAÇÃO DOS DADOS	50
4.3.1	Seleção dos dados	50
4.3.2	Limpeza dos dados	51
4.3.3	Construção dos dados.....	52
4.3.4	Integração dos dados	54
4.3.5	Formatação dos dados	55
4.4	FASE 4: MODELAGEM	55
4.4.1	Seleção da técnica de modelagem	55
4.4.2	Geração de teste de desempenho	62
4.4.3	Construção do modelo	63
4.4.4	Avaliação do modelo.....	64
4.5	FASE 5: AVALIAÇÃO.....	69
4.5.1	Avaliação dos resultados	69
4.5.2	Revisão do processo	74
4.5.3	Determinação dos próximos passos	75
4.6	FASE 6: IMPLANTAÇÃO.....	75
5	CONCLUSÃO E TRABALHOS FUTUROS.....	76
	REFERÊNCIAS.....	78
	APÊNDICE A – METODOLOGIA CRISP-DM.....	82
	APÊNDICE B – TABELAS DOS ARQUIVOS DE COTAÇÕES	91
	APÊNDICE C – RESUMO DAS PREDIÇÕES POR EMPRESA.....	95
	APÊNDICE D – RESULTADOS DA VALIDAÇÃO E TESTE DAS RNA	108
	APÊNDICE E – DESENHO DAS REDES NEURAIS ARTIFICIAIS.....	116

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

Numa economia globalizada, impulsionada por notícias oriundas de múltiplas fontes, às vezes positivas, às vezes negativas, a aplicação de recursos financeiros é um verdadeiro quebra-cabeças para os investidores. Todo investidor, no momento de efetivá-las, pensa em três aspectos básicos: retorno, prazo e segurança do meio em que está aplicando. Nesse contexto, a rentabilidade está diretamente relacionada ao grau de risco que a aplicação representa. Quanto maior a probabilidade de gerar grandes rendimentos num curto, médio ou longo prazo, maior o risco que o investidor está correndo.

Existem tipos de aplicações com rentabilidade menor e com baixíssimo risco envolvido. É o caso de aplicações em caderneta de poupança ou fundos de investimentos lastreados através de títulos do governo federal. Outras, entretanto, têm alto risco envolvido, tais como fundos de investimentos lastreados por ações de empresas, comercializadas através da bolsa de valores ou as próprias ações, compradas diretamente através de consórcio de investidores ou mesmo por investidores individuais. Ao investidor, cabe a definição do grau de risco que quer correr, levando em consideração a rentabilidade pretendida para o seu capital.

1.2 REVISÃO DE LITERATURA

Tendo em vista a volatilidade e a insegurança verificada no mercado financeiro, incluindo o de ações, tem ocorrido grande interesse por parte de pesquisadores interessados no comportamento desse mercado e de seus atores. Nesse sentido, diversos trabalhos vêm sendo direcionados à identificação de fatores que possam auxiliar os especialistas na compreensão desse mercado. Entre as diversas abordagens para tal problema está a utilização de técnicas de Inteligência Artificial.

Mueller (1996) propõe um modelo preditivo baseado em redes neurais artificiais como forma de antecipação do comportamento de séries de preços de ações a partir da identificação de regularidades no comportamento dessas cotações no mercado. No texto, foram conceituados temas tais como: redes neurais artificiais, tópicos

ligados a ações e suas implicações nas organizações, tipos de análise das cotações de ações, séries temporais e apresentadas as conclusões obtidas com o trabalho após um ciclo de 20.000 iterações, comparando valores previstos e valores reais através de vários gráficos. Dentre as propostas de trabalhos futuros apresentados pelo autor, destacam-se:

- A especificação de alguma metodologia que oriente a realização da previsão dos preços das ações para períodos superiores ao instante seguinte de tempo, procurando identificar as tendências gerais e os pontos de inflexão mais relevantes na série de cotações das ações. Através da identificação de pontos de inflexão, o investidor poderá programar o momento ideal para comprar e vender ações;
- Uma análise do retorno sobre o retorno em investimentos em ações, considerando o preço das mesmas, observando as sazonalidades envolvendo o *dia da semana*, o *mês do ano* e outras anomalias que possam ser relevantes;
- A avaliação do desempenho dos modelos conexionistas na previsão dos preços das ações considerando, além da série de preços analisada, outras séries temporais que tenham influência significativa no mercado de ações.

Fang e Xu (2002) investiga a previsibilidade de lucros combinando análise técnica e previsões em série temporais convencionais. Enquanto explora componentes previsíveis, como funções de preços passados ou lucros, regras técnicas e série de tempo, observa a existência de diferentes aspectos de previsibilidade de mercado. Identifica períodos que tendem a lucros ou a prejuízos. O autor aplica sua pesquisa ao índice *Dow Jones* combinando regras técnicas comerciais e previsões de série temporais. Para o autor, a previsibilidade pode ser explicada em grande parte por meio de correlações entre lucros não triviais, ou seja, não previsíveis. Prejuízos não podem ser justificados somente pelos erros de previsão que ocorrem na comercialização não síncrona, ou seja, em tempos e períodos diferentes e não contíguos.

Westerhoff (2003) estuda a efetividade de limites de preços em mercados especulativos. O autor propõe um modelo estocástico não linear que estima a confiabilidade dos aplicadores em análise técnica e fundamental para determinar as aplicações. A dinâmica do modelo simula fatos como o aparecimento de bolhas e excesso de volatilidade. A partir deste modelo, busca provar que os limites de preço têm po-

tencial para reduzir volatilidade e desvios dos fundamentos do mercado acionário. Para o autor, quanto mais os aplicadores se prestam ao comportamento de extrapolar tendências, melhores são os resultados alcançados pela adoção dos limites de preço. Nesse sentido, os limites funcionam como balizadores para os aplicadores na comercialização de ações.

Leigh et al. (2004) utiliza a Descoberta de Conhecimento em Banco de Dados (DCBD) em um estudo de caso sobre as cotações da Bolsa de Valores de Nova Iorque e apresenta um modelo de análise sobre variações significativas nestas cotações. O trabalho se serve de cotações no período compreendido entre 1981 e 1999. Utiliza Engenharia do Conhecimento, *Data Mining* ou Mineração de Dados (MD), Séries Temporais e análise de comportamento financeiro, tentando identificar padrões de comportamento para justificar a razão de variações significativas.

Vasco (2004) analisa a construção de modelos de perfil de clientes em uma instituição financeira, com o objetivo de obtenção de regras que sejam úteis para avaliação de risco de crédito, utilizando o processo de DCBD e técnicas de MD. Para a geração dos modelos de perfil dos clientes, utiliza a metodologia CRISP-DM (*CRoss Industry Standard Process of Data Mining*) para organizar e executar as diversas atividades previstas e buscar obter os resultados com qualidade. Os resultados alcançados permitiram uma caracterização dos clientes com base no seu histórico de comportamento, além dos seus dados cadastrais.

Bohm e Wenzelburger (2004), servindo-se de técnicas estocásticas, investiga o desempenho de uma carteira de ações e sua diversidade de convicções comparando-a com diversas carteiras e fatores que influenciam os diferentes desempenhos identificados entre elas.

Paulos (2004) relata uma experiência vivenciada com aplicações realizada com ações da empresa *WorldCom*, estudando as variações do mercado, baseado na análise gráfica, apresentando erros na análise de risco e uma crença não justificada na reversão de tendências de retorno dos valores das cotações a patamares compatíveis com os valores aplicados. Apresenta recomendações tais como a diversificação das aplicações em uma carteira selecionada de ações. Argumenta ainda que os investidores, quando envolvidos com a aplicação de grandes volumes financeiros, devem estar atentos ao que ocorre nas bolsas do mundo inteiro, pois existe uma grande chance de que o que está acontecendo no mercado acionário em outros países venha influenciar positiva ou negativamente os demais mercados.

Paulos (2004) cita que a utilização da análise fundamentalista geralmente se situa no lado dos números, parecendo um tanto conflitante com a ética predominante no mercado de ações, baseada na esperança, nos sonhos, na visão e num certo romantismo com tonalidades monetárias. Salienta ainda que “os fundamentos representam para os investimentos o mesmo que (de maneira estereotipada) o casamento representa para o romance, ou o que os legumes representam nas refeições – apesar de saudáveis, nem sempre são empolgantes”. O conhecimento do tema é fundamental e importante para qualquer investidor que queira entrar nesse mercado.

Marshall et al. (2006) apresenta os resultados de uma pesquisa sobre análise utilizando a técnica de *Candlestick*, da variação no comportamento de preços de mercado. Essa técnica surgiu no Japão no século XVIII para tratar oscilações na cotação do mercado do arroz, nas antigas bolsas de arroz de Osaka. O termo *Candlestick* (candelabro, em inglês) se deve ao fato dos elementos gráficos utilizados na representação dos preços praticados pelo mercado lembrarem velas, distribuídas sobre a área do gráfico. A análise se faz através da identificação de "figuras" formadas pelos "candles" (velas) em determinado ponto da tendência de mercado. A partir destas variações, identificam-se tendências de pressão de venda ou de compra do produto.

Andersen et al. (2005) descrevem um mecanismo que opera com sistemas complexos adaptativos e lida com a previsibilidade diária de cotações com o apoio de agentes inteligentes, transientemente desacoplado da história passada. Os autores demonstram e testam suas teorias por meio de séries temporais para predição a partir de séries reais observadas no mercado financeiro.

Segundo Fu et al. (2006), a análise técnica, ou grafista, é uma das principais ferramentas para a investigação do comportamento do mercado de ações. Nesse contexto, os autores consideram necessária a identificação de padrões técnicos em relação ao comportamento do preço das ações através de gráficos. Ao identificar pontos perceptualmente importantes (PPI) a partir do espaço temporal, séries temporais e esquemas podem ser comparados. Três tipos de distâncias para identificação dos PPI são comparadas neste trabalho: euclidiana, perpendicular e vertical. Após a identificação dos PPI, os autores introduzem uma abordagem baseada em esquemas e outra baseada em regras para casamento de padrões.

Wang e Chan (2007) buscam identificar tendências do índice *Nasdaq* em relação ao índice tailandês TWI (*Taiwan Weighted Index*). A partir deste estudo, elabo-

ram um modelo baseado em padrões de reconhecimento dessas tendências e realizam diversas experiências para validar o modelo proposto.

Gestel et al. (2006) estuda a equidade de predição no índice de preço histórico do setor químico europeu utilizando um modelo linear de correção de erros baseado em redes bayesianas. A partir da análise dos resultados do modelo linear, esse modelo é aplicado com sucesso fora da amostra inicial de predição.

Gradojevic (2006) desenvolveu uma nova abordagem para gerar estratégias na negociação do mercado cambial, baseada em previsões a partir de um modelo híbrido entre lógica *fuzzy* e RNA, gerando previsões de taxas de câmbio com um dia de antecedência. Isso permitiu a investigação da evolução de um modelo dinâmico *neuro-fuzzy*.

Chiarella, He e Hommes (2007) analisam estatisticamente o mercado de ações visando explicar vários fenômenos de preço, incluindo bolhas de mercado, variações abruptas e resistências de preços. Propõem um modelo dinâmico para a procura de ativos negociados no mercado financeiro, combinando análises fundamentalista e grafista.

Oliveira Jr. (2007) compara modelos lineares e redes neurais artificiais como tecnologias geradoras de previsões de valores financeiros. Para o autor, a previsão é um meio importante de informação para subsidiar a tomada de decisões, visando atingir os objetivos pretendidos. Essa meta é fundamental, principalmente nas organizações, devido à necessidade de tomada de decisões relevantes e confiáveis. Seu modelo pretende reduzir os riscos de liquidez da instituição financeira, a melhorar os lucros pela antecipação de operações de compra e venda de títulos no mercado interbancário.

Nesta seção foram apresentados alguns trabalhos relativos à busca de modelos preditivos para o mercado financeiro, visando identificar contribuições relevantes para o problema abordado nesta dissertação. Vários trabalhos são apresentados a cada ano, entretanto, nenhum dos modelos apresentados conseguiram apresentar uma fórmula definitiva para enriquecer o conhecimento do acionista direcionando a melhor hora de comprar ou vender ações.

1.3 JUSTIFICATIVA

Segundo a Bolsa de Valores de São Paulo (BOVESPA, 2008), o processo de globalização resultou num intenso intercâmbio financeiro entre os países, tendo o mercado acionário adquirido importante papel no cenário financeiro mundial. Nesse sentido, os países em desenvolvimento procuram abrir suas economias para receber investimentos externos. Quanto mais desenvolvidas e globalizadas encontram-se as suas economias, mais ativos são seus mercados de capitais. O desenvolvimento de um país é fortemente influenciado pelo fluxo desse tipo de capital, uma das principais fontes para novos investimentos e a consequente criação de novos empregos, contribuindo para o progresso do país.

Uma vez que capitais estrangeiros estão disponíveis para serem aplicados em qualquer lugar do planeta, diversas formas de aplicações encontram-se disponíveis, tais como: poupança, investimentos imobiliários, mercado de ouro, mercado de moedas, mercado de capitais, entre outros. Alguns tipos de aplicação têm rendimento limitado, tornando-se mais atrativos para investidores que não estão dispostos a correr riscos de perdas financeiras. Outras aplicações com rendimentos mais atrativos oferecem um risco maior, podendo significar perdas no capital investido.

Existem casos em que o investidor dispõe de todos os seus bens, transformando-os em recursos disponíveis para aplicação em Bolsa de Valores. A existência de parâmetros históricos e de um modelo para a detecção de padrões de comportamento de rentabilidade de ações da Bolsa de Valores poderia permitir ao investidor aplicar recursos com maior segurança.

Como visto na seção anterior, há diversos trabalhos voltados para subsidiar o investidor do mercado de ações. Entretanto, um modelo com alta confiabilidade que garanta ganhos efetivos ainda está para ser criado. Por outro lado, observa-se que muito do aprendizado que um *trader*¹ acumula vem da prática vivenciada no dia-a-dia. Este aprendizado poderia ser beneficiado pela adoção de instrumentos de modelagem como o apresentado neste trabalho. Ainda que o *trader* não chegue a criar um modelo efetivo, o esforço para se construir um modelo de trabalho lhe permitirá conhecer cada vez mais a dinâmica do mercado.

¹ Investidor experiente no mercado de ações.

1.4 FORMULAÇÃO DO PROBLEMA

Um acionista, no momento de decidir o melhor momento em comprar ou vender ações no mercado, lança mão de diversos instrumentos, tais como as análises fundamentalista ou grafista. Agregado a esses tipos de análise, conta-se com uma gama significativa de notícias nacionais e internacionais vinculadas às empresas que compõem esse mercado. Como verificado anteriormente, muitos trabalhos tentam agregar novos métodos ou ferramentas com potencial de auxiliar o *trader*. O foco principal desse trabalho é agregar à plataforma de um investidor a prática de modelagem baseada em MD.

Nesse sentido, foi feito um estudo preliminar (AMOROSO, 2008) no qual foram analisados dados históricos de cotações de ações com o uso de séries temporais e observado um comportamento satisfatório em alguns dos modelos gerados. A partir da abordagem realizada, o presente trabalho busca identificar oportunidades de geração de conhecimento ao acionista a partir da predição de valores das cotações, minimizando riscos na hora de comprar ou vender ações.

Durante a elaboração desta dissertação, procurou-se evidenciar a viabilidade de adoção de ferramentas de modelagem devido à usabilidade alcançada pelas ferramentas atuais.

1.5 OBJETIVOS

1.5.1 Objetivo geral

O objetivo deste trabalho é a realização de experimentos de modelagem de séries temporais de preços de ações, visando mostrar a viabilidade deste instrumento como fonte de aprendizado para um investidor do mercado de ações.

1.5.2 Objetivos específicos

Os objetivos específicos deste trabalho são os seguintes:

- Estudar o problema de previsão de cotação de preços no âmbito do mercado de ações;
- Estudar modelos de predição em séries temporais com Redes Neurais Artificiais e selecionar um que seja adequado para o problema em questão;
- Avaliar o modelo escolhido por meio de um Estudo de Caso;
- Evidenciar o potencial do modelo como fonte de aprendizado sobre a dinâmica do mercado de ações.

1.6 ORGANIZAÇÃO DA DISSERTAÇÃO

O presente trabalho está estruturado da seguinte maneira. No Capítulo 1, é feita a apresentação do tema do trabalho, a revisão da literatura, a justificativa, a formulação do problema e os objetivos a serem alcançados com a realização do presente trabalho. No Capítulo 2, apresentam-se os conceitos extraídos da literatura a respeito da importância da gestão do conhecimento, mercado de capitais, descoberta de conhecimento em banco de dados, redes neurais artificiais, séries temporais e metodologia CRISP-DM. No Capítulo 3, é feita uma exposição dos procedimentos, materiais e métodos utilizados para a elaboração do trabalho, apresentando a classificação da pesquisa, suposições a respeito do trabalho, informações sobre a coleta e análise dos dados e a delimitação do estudo realizado. No Capítulo 4, apresenta-se a aplicação do modelo no estudo de caso proposto, conforme disposto na metodologia CRISP-DM. No Capítulo 5, são descritos os resultados obtidos em relação aos resultados esperados. No Capítulo 6, apresentam-se as conclusões obtidas com o trabalho realizado, considerações e proposta de trabalhos futuros.

2 REFERENCIAL TEÓRICO

Neste capítulo, são apresentadas as bases conceituais sobre os principais temas que permeiam o desenvolvimento do trabalho proposto.

2.1 GESTÃO DO CONHECIMENTO

Segundo Drucker (1993), “os grandes ganhos de produtividade, daqui para frente, advirão das melhorias na gestão do conhecimento”. Embora o contexto da Gestão do Conhecimento (GC) vivenciado naquela época estivesse direcionado prioritariamente para a mudança na forma de gestão nas empresas, já se pregava sua utilização como forma de obtenção, entre diversas outras possibilidades, de melhores resultados financeiros (CAVALCANTI, 2000).

A GC vem se tornando uma preocupação cada vez maior no dia-a-dia de todos. Num ambiente onde as mudanças são constantes, e a competitividade, para as organizações, é uma palavra que pode significar sua sobrevivência.

Segundo Nonaka e Takeuchi (1997), o conhecimento pode ser definido como *explícito* ou *tácito*. Conhecimento explícito é aquele que indivíduos conseguem comunicar verbalmente ou de forma estruturada, a partir de suas experiências vivenciadas. Já conhecimento tácito é aquele que o indivíduo possui a partir das experiências vivenciadas, entretanto a mesma é pessoal.

Conhecimento tácito é um termo cunhado por Polanyi (1983) e significa aquele conhecimento que possibilita a uma pessoa resolver problemas, mas que é de difícil comunicação. Segundo o autor, pode-se saber mais do que se é capaz de expressar e nem tudo o que as pessoas sabem pode ser codificado em documentos e ferramentas, já que as pessoas não sabem ao certo o que sabem.

Nonaka e Takeuchi (1997), ao estenderem para o campo da prática os fundamentos filosóficos do conhecimento tácito, propuseram um modelo de criação do conhecimento nas organizações estruturado em duas dimensões: ontológica e epistemológica. Na dimensão ontológica, a criação do conhecimento é entendida como um processo que amplia, em termos organizacionais, o conhecimento criado pelos indivíduos como parte da rede de conhecimentos da organização. Já na dimensão epistemológica, o conhecimento humano é expandido através da interação social

entre conhecimento tácito e conhecimento explícito. Esta interação ficou conhecida como Modelo SECI (Socialização, Externalização, Compartilhamento e Internalização), também chamada de Espiral do Conhecimento, como mostrado na Figura 1.



Figura 1 - Espiral do Conhecimento, segundo Nonaka e Takeuchi (1997, p. 80)

Krogh (2001, p. 69) cunhou o termo “especialista solícito” e o definiu como “o indivíduo que atinge determinado nível de excelência pessoal, envolvendo conhecimentos tanto tácitos como explícitos, e se considera responsável pelo compartilhamento do processo”. A criação de conhecimento em uma comunidade ou em uma organização que convive com diferentes tipos de conhecimento, se concretiza através da ajuda e compartilhamento entre seus membros.

A colaboração, por sua vez, é um elemento crítico para o sucesso das estratégias organizacionais. A colaboração deve existir entre as estruturas funcionais, entre os níveis hierárquicos e entre as fronteiras da organização (CROSS, 2004). Quanto maior a colaboração, mais estreito o relacionamento e mais próxima estará a organização de seus objetivos. Como o trabalho acontece através de redes colaborativas informais que transpassam os limites físicos e hierárquicos, os gestores têm um forte desafio que é o de proporcionar um ambiente onde a colaboração seja mais efetiva.

Nas redes informais e organizacionais, a troca de informação é intensa e constante. Tratam-se de ambientes onde experiências, aprendizados e dúvidas, entre outros aspectos pessoais e profissionais, são compartilhados através da sociali-

zação dos conhecimentos tácitos (NONAKA; TAKEUSHI, 1997, p. 69-71). As redes formais ou informais são, nesse contexto, na maioria das vezes, fontes de criação e propagação de conhecimento (ALLEE, 2003, p. 78-84).

O conceito de GC parte da premissa de que todo o conhecimento dos processos de negócio existente “na cabeça de seus profissionais”, ou “no coração de seus processos” fazem parte do conhecimento organizacional de propriedade da empresa (PARADIGMA, 2008). Nesse contexto, caso os processos e os produtos oriundos dos mesmos não estejam documentados, a empresa estará permanentemente refém das pessoas que detêm esse conhecimento tácito.

Muito se fala a respeito da importância do conhecimento para atuação no mundo dos negócios, principalmente quando se trata de investimentos financeiros. Vivencia-se um cenário de rara complexidade, tanto no mundo corporativo como na sociedade em geral. É nesse contexto que a GC se tornou um importante recurso estratégico. É preciso lidar com informações sobre o ambiente em que se está inserido o negócio específico, com informações oriundas das demonstrações contábeis das empresas, com a evolução de seus negócios, tanto no presente como nas prospecções de novos produtos e serviços.

Para tanto, deve-se considerar a forma da organização tratar interna e externamente a GC organizacional como um dos tópicos que envolvem a governança corporativa, pois esse conhecimento é um fator fundamental no momento de estabelecer o seu valor, por exemplo, em termos de sua cotação no mercado de ações. A capacitação de seus profissionais é uma arma fundamental para o fomento dos negócios organizacionais, principalmente quando a organização atua no mercado de capitais.

2.2 MERCADO DE CAPITAIS

As economias dos países são compostas de diversos segmentos, estando, entre eles, o mercado do dinheiro. O Quadro 1 ilustra esses segmentos.

Segmentos	Prazos	Finalidade	Tipo de intermediação
Monetário	Curtíssimo, curto	Controle da liquidez monetária da economia, suprimentos momentâneos de caixa	Bancário e não bancário
Crédito	Curto e médio	Financiamento do consumo e capital de giro das empresas	Bancário e não bancário
Câmbio	À vista, curto	Conversão de valores em moedas estrangeiras e nacional	Bancário e auxiliar (corretoras)
Capitais	Média, longo	Financiamento de investimentos, de giros e especiais (habitação)	Não bancário

Quadro 1 - Os mercados do Dinheiro (CAVALCANTE, 2005)

Segundo Pinheiro (2002), o mercado de capitais pode ser definido como sendo um conjunto de instituições que negociam com títulos e valores mobiliários, objetivando a canalização dos recursos dos agentes compradores para os agentes vendedores. Ou seja, o mercado de capitais representa um sistema de distribuição de valores mobiliários que tem o propósito de viabilizar a capitalização das empresas e dar liquidez aos títulos emitidos por elas. O surgimento do mercado de capitais ocorreu a partir do momento que o mercado de crédito deixou de atender às necessidades da atividade produtiva, no sentido de suprir as necessidades de capital das empresas com prazos, custos e exigibilidades adequados.

2.3 MERCADO DE AÇÕES

A seguir, alguns conceitos básicos sobre o mercado de ações são apresentados.

2.3.1 Considerações iniciais

Segundo Cavalcante (2005), a necessidade de estreitar o relacionamento entre as principais praças de mercado brasileiras e a procura por meios que facilitem a realização dos negócios entre as corretoras das diversas bolsas de valores e destas entre si levou à criação da Comissão Nacional de Bolsas de Valores – CNBV, hoje

denominada CNB, associação civil sem fins lucrativos responsável pelos interesses e pela integração de suas associadas, as bolsas.

Os objetivos estatutários da CNB são (BOVESPA, 2008):

- Estimular a colaboração entre as associadas;
- Defender o interesse das associadas perante as autoridades;
- Propor medidas ou modificações na legislação;
- Zelar pela manutenção de elevados padrões éticos de conduta no relacionamento entre as associadas;
- Atuar como porta voz com os meios de comunicação ou veículos próprios, das reivindicações, propostas e opiniões do conjunto das associadas;
- Indicar membros para integrar órgãos públicos e privados em que tal representação seja exigida ou solicitada;
- Estudar os atos do Conselho Monetário Nacional, da Comissão de Valores Mobiliários e do Banco Central do Brasil, expedindo às associadas esclarecimentos e recomendações que facilitem a sua aplicação;
- Organizar e promover cursos, seminários, congressos e conferências, bem como a realização de estudos e pesquisas sobre assuntos de interesse das associadas;
- Representar as associadas em congressos, comissões, plenários e assemelhados, nacionais e internacionais, que tratem de temas relevantes para os mercados operados pelas mesmas;
- Editar, publicar, distribuir ou veicular material educativo ou promocional relacionado aos mercados operados pelas associadas;
- Manter e ampliar as relações e o intercâmbio internacional permanente com bolsas de outros países.

As bolsas associadas à CNB são:

- Bolsa de Valores de São Paulo – Bovespa
- Bolsa de Mercadorias e Futuros – BM&F
- Bolsa de Valores do Rio de Janeiro – BVRJ
- Bolsa de Valores de Minas, Espírito Santo e Brasília – BOVMESB
- Bolsa de Valores da Bahia – Sergipe – Alagoas – BOVESBA
- Bolsa Brasileira de Mercadorias – BBM

- Câmara de Custódia e Liquidação – CETIP

2.3.2 A BOVESPA

A Bovespa é o maior centro de negociações com ações na América Latina. Ela desenvolveu um cadastro de empresas cotadas nos seus pregões a fim de facilitar aos usuários a consulta das diversas informações sobre determinada empresa listada. Esse programa visa atrair o investidor interessado no mercado acionário e recebeu o nome de “Bovespa vai até você” (BOVESPA, 2008).

Segundo Lassalvia (2005), a busca de melhores práticas de governança corporativa levou a Bovespa a reunir uma série de condutas a serem seguidas pelas empresas (opção), controladores e administradores. Tais condutas podem levar as empresas a serem distinguidas no mercado de capitais, de acordo com os seus compromissos assumidos junto ao mercado, podendo estar classificadas no *Novo Mercado*, no qual as empresas são classificadas segundo dois níveis de governança corporativa. Por definição da Bovespa (2008), esse Novo Mercado é regido por um conjunto de normas de conduta para as empresas, seus administradores e controladores. A adesão a tais práticas é um dos fatores que distingue as companhias segundo o nível de comprometimento assumido.

As companhias habilitadas no nível 1 devem se comprometer, principalmente, com melhorias na prestação de informações ao mercado e com a dispersão acionária. Já as empresas habilitadas no nível 2, adotam um conjunto bem mais amplo de práticas de governança e de direitos adicionais para os acionistas minoritários.

Os mercados mundiais movimentam bilhões de dólares diariamente. Uma parcela significativa desses negócios é realizada nas bolsas de valores ao redor do mundo. As bolsas são instrumentos fundamentais para propiciar liquidez e segurança aos investidores de ações e, conseqüentemente, facilitar a captação de recursos pelas empresas, permitindo, em última análise, o desenvolvimento econômico dos países.

Eis alguns números da Bovespa em relação à quantidade de investidores, número de clubes de investimento, volume financeiro de transações entre outros:

- Existem cerca de 2.372 clubes de investimentos cadastrados até o final do mês de março de 2008;

- Somente em 2008, foram criados 265 novos clubes de investidores;
- O patrimônio líquido totalizou R\$14,94 bilhões em fevereiro de 2008;
- O número de cotistas alcançou 154.625 em fevereiro de 2008;
- A Companhia Brasileira de Liquidação e Custódia (CBLC) revelou que o volume financeiro das operações com empréstimos de ações alcançou R\$ 30,1 bilhões em março de 2008;
- Foram realizadas 49.989 operações no mês de março de 2008;
- O volume financeiro negociado em março de 2008 foi de R\$ 111,32 bilhões;
- 31,24% no número de negócios da Bovespa foram realizados através de *Home Broker*, instrumento que permite a negociação de ações via Internet.

Os investidores estrangeiros continuam liderando a movimentação financeira. Os percentuais por área em relação ao volume total negociado segundo a Bovespa são apresentados na Tabela 1.

Tabela 1 - Tabela com tipos de investidores da Bovespa

Tipo de investidor	Percentual
1. Investidores estrangeiros	36,31%
2. Investidores institucionais	27,07%
3. Pessoas físicas	26,94%
4. Instituições financeiras	7,41%
5. Empresas	2,18%
6. Outros	0,09%
Total	100,00%

2.3.3 Ações

Quando uma empresa precisa de capital externo à organização para a realização de investimentos, oferece ações ao mercado através das bolsas de valores. Uma ação é uma pequena porção da sociedade de uma empresa. A partir do momento que uma ação é negociada, o portador de uma ou mais ações torna-se proprietário daquela porção da empresa. Numa definição mais formal (CAVALCANTE,

2008), pode-se dizer que “ações são títulos nominativos negociáveis que, representam, para quem as possui, uma fração do capital social de uma empresa”.

2.3.4 Tipos de ações

As ações podem ser:

- Ordinárias Nominativas (ON), que concedem ao seu proprietário o direito de voto nas assembleias de acionistas da empresa (BOVESPA, 2008);
- Preferenciais Nominativas (PN), que concedem ao seu proprietário a preferência no recebimento dos resultados ou no reembolso do capital em caso de liquidação da empresa. Esse tipo de ação não dá ao seu proprietário direito a voto nas assembleias de acionistas da empresa ou restringem esse voto, que é declarado expressamente no estatuto social da empresa (BOVESPA, 2008).

As ações preferenciais podem ainda ser diferenciadas por classes: A, B, C ou alguma outra letra que apareça após o "PN". As características de cada classe são estabelecidas no estatuto social da empresa emissora da ação. Essas diferenças variam de empresa para empresa, não sendo possível uma definição geral dessas classes de ações (BOVESPA, 2008).

2.3.5 Análise de preço de ações

Um investidor, no momento da aplicação em ações, deve analisar como se encontra o mercado naquele momento, atentando principalmente para o índice de fechamento de um determinado pregão. Bolsa “em alta”, “estável” e “em baixa” significam respectivamente que, depois do fechamento do dia, o índice geral de fechamento de determinado pregão foi superior, permaneceu nos mesmos patamares ou foi inferior ao índice de fechamento do pregão do dia anterior (BOVESPA, 2008).

Para uma análise mais refinada do comportamento do mercado, existem dois tipos de abordagens normalmente utilizadas: a análise fundamentalista e a análise técnica (BOVESPA, 2008).

A análise fundamentalista considera a situação financeira, econômica e mercadológica de uma empresa, com suas expectativas e projeções de mercado. Para tanto, é necessária uma análise criteriosa da situação financeira da organização por meio de seu último balanço publicado ou de informações veiculadas sobre a mesma nos principais meios de comunicação. Essa análise é normalmente utilizada para definir o valor de mercado de uma empresa e comparar sua cotação com o atual mercado em termos de quantidade de ações que a empresa possui e o seu preço de cotação vigente (PINHEIRO, 2002).

A análise técnica, também conhecida como Análise Gráfica, é uma ferramenta utilizada tanto por investidores profissionais (também conhecidos como *traders*) quanto por amadores para o estudo de ações individuais. Baseia-se na ideia de que os preços dos ativos negociáveis se movem de acordo com padrões repetitivos e identificáveis. Ela registra em gráficos as atividades de preços e volumes e inferem, a partir de sua história gráfica, as prováveis tendências de preços desses ativos no mercado (PINHEIRO, 2002).

Essa abordagem é fundamentada pela Teoria de Dow (MATSURA, 2007). De acordo com essa teoria, os preços dos ativos refletem a dinâmica do mercado em relação a todas as informações relevantes a respeito da mesma.

Nesse contexto, a teoria apresenta quatro abordagens:

- O preço desconta tudo – toda influência externa gera impactos positivos ou negativos em relação ao preço dos ativos;
- O preço tem tendência – analisando o preço na linha do tempo, observa-se que tendências se repetem;
- Os padrões se repetem – existem padrões que se repetem na linha do tempo da mesma forma que o preço tem tendência;
- A história se repete – fatores extraordinários e sazonais se repetem.

Entre os tipos de gráficos que são adotados nesse tipo de análise estão os gráficos de barras e *Candlestick*.

2.4 DESCOBERTA DO CONHECIMENTO EM BANCO DE DADOS

Fayyad et al. (1996) apresentam os fundamentos da DCBD. Este trabalho representa um importante marco para os estudiosos sobre o tema e foi resultado da

compilação de vários artigos descritos por uma série de *Workshops* que foram realizados entre os anos de 1989 e 1994, que abordavam processos, modelos de classificação e “clusterização” e perspectivas estatísticas.

A Descoberta de Conhecimento em Banco de Dados (DCBD) é um processo iterativo e incremental para a extração de padrões novos, úteis e previamente desconhecidos a partir de um conjunto relevante de dados (FAYYAD et al., 1996). DCBD envolve, de modo geral três fases: pré-processamento, modelagem e pós-processamento. No pré-processamento, um conjunto de dados de interesse é tratado visando torná-lo o mais representativo do fenômeno estudado. A fase de modelagem (a MD), tem como finalidade a obtenção de padrões no conjunto de dados tratado. Por fim, no pós-processamento os padrões são interpretados com vistas à sua utilidade para a resolução do problema que justificou todo o ciclo de DCBD.

Inmom et al. (2001) afirmam que em MD, o explorador tem liberdade para escolher e tratar tanto de suas fontes internas de dados quanto das externas. Entretanto, ele deve se suplementar com mais dados para uma análise mais profunda. O grau de verdade das informações depende das suposições que foram adotadas e até que ponto as mesmas são verdadeiras. Analisando esses conceitos sob esse ângulo, é importantíssima a escolha dos dados a serem tratados para a descoberta de padrões de comportamento e conhecimento.

Segundo Amaral (2001), DCBD é um processo não trivial de identificação válida dos padrões de dados. É a descoberta de novos conhecimentos, seja padrões, tendências, associações, probabilidades ou fatos, que não são óbvios ou de fácil identificação. Esses padrões podem ser utilizados para a geração de uma ação prática transformando-se em uma vantagem competitiva. Padrões são reconhecidos quando podem ser transformados em informações identificáveis pelos usuários.

A seguir, são apresentados alguns casos de sucesso relacionados ao tema deste trabalho:

- *American Express*: utilização de marketing direcionado registrou um aumento percentual de 10 a 15 por cento no uso de cartões de crédito (FAYYAD, 1996);
- *Dun & Bradstreet*: utilização de um grande *data warehouse*, denominado Scantrack, com dados de vendas em vários supermercados espalhados pelos E.U.A. As vendas são registradas e armazenadas num computador

central e servem para direcionar a reposição de estoques, em épocas específicas do ano com foco em um público específico (CARVALHO, 2001);

- *Keystone Financial*: através de *marketing* foi selecionado aos resultados de um modelo de DCBD, que gerou 1,6 milhões de dólares em novos rendimentos somente sobre três promoções (GRAETTINGER, 1999).

A seguir, algumas áreas de aplicação de utilização da DCBD:

- *Marketing*: verificação de preferências e comportamento de clientes para a utilização de marketing direto;
- Medicina: através do comportamento de pacientes é possível identificar terapias de sucesso e pesquisar padrões de doenças novas;
- Instituições Financeiras: através de análise de informações é possível identificar oportunidades de vendas de produtos, identificar possíveis comportamentos em fraudes, identificar clientes preferenciais e/ou inadimplentes, correlações entre indicadores financeiros;
- Seguradoras: verificação de padrões de incidência de sinistros, identificação de fraudes em procedimentos médicos, verificação de distorções em procedimentos médicos, identificação de padrões de comportamento de clientes perigosos e de comportamentos fraudulentos;
- Transportadoras: identificação de padrões de comportamento de carga e rotas mais utilizadas em determinados períodos.

2.5 REDES NEURAIS ARTIFICIAIS

O cérebro é composto por alguns bilhões de neurônios. Neurônios são células formadas por *dendritos*, *corpo* e *axônio* (FERNEDA, 2006), como mostrado de forma simplificada na Figura 2. Estímulos gerados por neurônios vizinhos são captados pelos dendritos e transmitidos ao *corpo* do neurônio. Quando tais estímulos atingem certo limiar, o corpo envia novo impulso, transmitido pelo axônio às células vizinhas por meio de *sinapses*. Este processo se repete por várias camadas de neurônios. Assim, informações são processadas de forma que o cérebro gere reações físico-químicas.

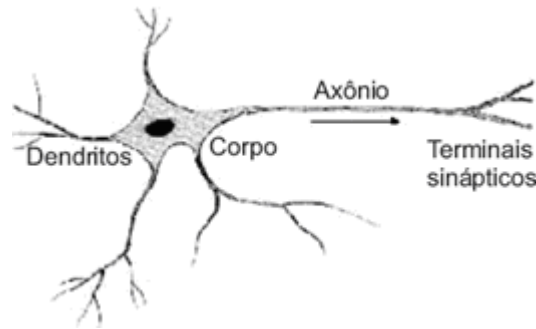


Figura 2 - Representação simplificada de um neurônio

A habilidade dos animais realizarem funções complexas como a aprendizagem é consequência direta do processamento paralelo e distribuído da rede de neurônios do cérebro. Novos conhecimentos ou experiências podem levar a rearranjos das redes de neurônios, reforçando ou inibindo algumas sinapses (HAYKIN, 2001, p. 32-36).

McCulloch e Pitts (1943) foram pioneiros na busca por modelos computacionais que simulassem o funcionamento das células do cérebro. Os resultados animadores levaram pesquisadores a proporem métodos inovadores de aprendizagem para as Redes Neurais Artificiais (RNA), como o *perceptron* de Rosenblatt (1958), que finalmente Minsky e Pappert (1969) mostraram ter sérias limitações. Porém, durante os anos 80, avanços tecnológicos e metodológicos fazem reavivar o entusiasmo pelo tema. A Figura 3 é uma simplificação do modelo de um neurônio artificial (HAYKIN, 2001, p. 36).

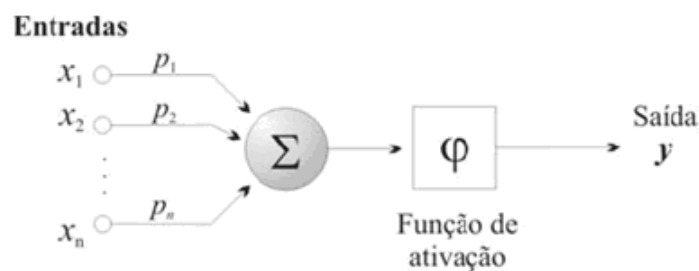


Figura 3 - Modelo matemático de um neurônio (FERNEDA, 2006)

Este modelo é composto por três elementos básicos: (i) um conjunto de conexões de entrada (x_1, x_2, \dots, x_n) associadas a pesos (p_1, p_2, \dots, p_n), (ii) um acumulador da soma ponderada ($\sum_{i=1}^n x_i p_i$) dos sinais de entrada e (iii) uma função de ativação (ϕ)

que limita o intervalo permissível de amplitude do sinal de saída (γ) a um valor fixo. Os pesos simulam o comportamento das conexões entre os neurônios. Pesos com valores positivos representam conexões excitatórias enquanto os valores negativos representam conexões inibitórias. O valor resultante da soma ponderada das entradas é avaliado pela *função de ativação*, que define a saída do neurônio.

Uma combinação de diversos neurônios resulta numa RNA, modelo que busca simular o processamento de informação dos cérebros. Seus componentes, os neurônios, são unidades de processamento simples que se unem por meio de conexões sinápticas. De uma forma simplificada, uma RNA pode ser vista como um grafo onde os nós representam os neurônios e os arcos as sinapses, como exemplificado na Figura 4.

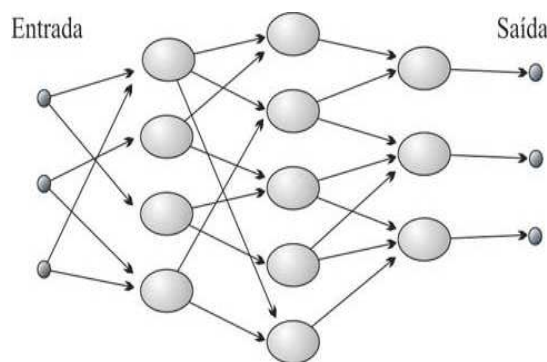


Figura 4 - Exemplo de uma RNA

As RNA podem ser classificadas segundo sua arquitetura e sua forma de ajustar os pesos associados às conexões durante o processo de ajuste dos pesos. A arquitetura de uma RNA restringe o tipo de problema no qual a rede poderá ser utilizada. Ela é definida pelo número de camadas (*camada única* ou *múltiplas camadas*), pelo número de nós em cada camada, pelo tipo de conexão entre os nós (*feedforward* ou *feedback*) e por sua topologia (HAYKIN, 2001, p. 46-49).

Uma das propriedades das RNA é a de se servir de estratégias de ajuste dos pesos de cada uma de suas conexões, processo esse chamado de *aprendizagem*. Assim, a partir de exemplos e de contra-exemplos, uma RNA *aprende* e faz inferências sobre o que foi aprendido, possibilitando a melhoria gradativa do seu desempenho. (BRAGA; CARVALHO; LUDERMIR, 2000)

O aprendizado das RNA se dá de forma *supervisionada* ou *não-supervisionada*. No primeiro caso, é apresentado à RNA um conjunto de padrões de entrada e

seus padrões de saída correspondentes. Ou seja, é necessário um conhecimento prévio do comportamento que se deseja ou se espera da rede. Compara-se a resposta fornecida pela RNA com a resposta esperada. Em função do erro verificado, são feitos ajustes buscando a melhoria da suas futuras respostas. (BRAGA; CARVALHO; LUDERMIR, 2000)

No aprendizado não-supervisionado, somente os padrões de entrada estão disponíveis para a RNA. A rede as processa e, em função das regularidades encontradas, tenta progressivamente estabelecer representações internas para codificar características e classificá-las automaticamente. Este tipo de aprendizado só é possível quando existe redundância nos dados de entrada, para que se consiga encontrar padrões em tais dados. (BRAGA; CARVALHO; LUDERMIR, 2000)

Denomina-se ciclo ou época uma apresentação de todos os n pares (entrada e saída) do conjunto de treinamento ao processo de aprendizado. A correção dos pesos num ciclo pode ser executada de dois modos (CARVALHO, 2004):

- *Modo padrão*, onde a correção dos pesos acontece a cada apresentação à rede de um exemplo do conjunto de treinamento. Cada correção de pesos baseia-se somente no erro do exemplo apresentado naquela iteração. Assim, em cada ciclo ocorrem n correções.
- *Modo batch*, onde apenas uma correção é feita por ciclo. Todos os exemplos do conjunto de treinamento são apresentados à rede, seu erro médio é calculado e a partir deste erro fazem-se as correções dos pesos.

As RNA podem ser aplicadas na solução de uma grande quantidade de problemas complexos. Citamos abaixo alguns exemplos:

- A sua aplicação em software de reconhecimento de voz, que necessitam aprender a conhecer a voz de determinadas pessoas;
- Na geração de séries temporais;
- Em software de OCR, que é responsável pela tradução de uma imagem para texto onde ele necessita aprender a reconhecer caracteres a partir de uma imagem;
- Em aplicações mais complexas como usinas, mercado financeiro, etc.

Um modelo de RNA bastante difundido é a rede *Multilayer Perceptron* (MLP), utilizada na resolução de problemas de diversas áreas do conhecimento humano, tais como a engenharia, a medicina e a linguística, em função de sua facilidade e

simplicidade de implementação. Braga, Ludemir e Carvalho (2001, p. 72) exemplificam algumas destas implementações: reconhecimento de caracteres, previsão do comportamento de ações na bolsa, verificação de assinaturas, segurança em transações com cartões de crédito e diagnóstico médico.

As redes *Perceptron*, de apenas uma camada, conseguem resolver apenas problemas linearmente separáveis. Já uma rede MLP com uma camada intermediária é suficiente para aproximar qualquer função contínua, ao passo que duas camadas intermediárias são suficientes para aproximar qualquer função matemática, contínua ou não (CYBENKO, apud BRAGA; LUDEMIR; CARVALHO, 2001, p. 49).

Para Braga, Ludemir e Carvalho (2001, p. 49), uma grande dificuldade surge com o treinamento das RNA, pois pode haver convergência para um mínimo local em função da distribuição dos dados. Para minimizar este problema, a rede MLP deve ser implementada com uma camada intermediária formada por um conjunto de redes *Perceptron* para cada grupo de entradas linearmente separáveis. Outra solução é treinar toda a rede de uma única vez. Todavia, neste caso, surge um outro problema: como treinar os neurônios da camada intermediária?

Rumelhart (apud BRAGA; LUDEMIR; CARVALHO, 2001, p. 50) propôs na década de 80 um método para resolução deste problema baseado em gradiente descendente, onde a função de ativação precisa ser contínua e diferenciável, para que o gradiente possa ser calculado e direcionar o ajuste dos pesos e, ainda ser, de preferência não decrescente. Esta função deve informar os erros da camada de resposta às camadas anteriores. A alternativa encontrada foi a utilização da função de ativação do tipo sigmoidal.

2.6 SÉRIES TEMPORAIS

Segundo Morettin (2006), uma série temporal é qualquer conjunto de observações ordenadas no tempo. São citados como exemplos de séries temporais: valores diários de poluição na cidade de São Paulo, valores mensais de temperatura na cidade de Cananeia-SP, índices diários da Bolsa de Valores de São Paulo, precipitação atmosférica anual na cidade de Fortaleza, número médio anual de manchas solares e registro de marés no porto de Santos.

Nos exemplos acima citados, temos séries temporais contínuas e discretas (Figura 5). As contínuas são medidas analógicas no tempo e as discretas são observações tomadas em intervalo de tempo específicos. Séries temporais contínuas podem ser discretizadas pela amostragem de valores em intervalos de tempos usualmente iguais.

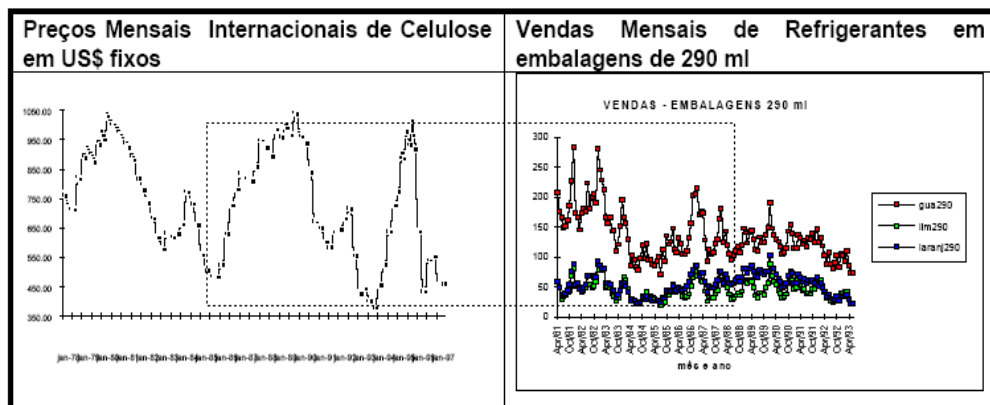


Figura 5 - Exemplos de séries temporais contínuas e discretas (BARROS, 2008)

Segundo Morettin (2006), há basicamente dois enfoques usados na análise de séries temporais. No primeiro, a análise é feita no domínio temporal e os modelos propostos são modelos paramétricos (com um número finito de parâmetros). No segundo, a análise é conduzida no domínio de frequências e os modelos são não paramétricos.

A maneira tradicional de se analisar uma série temporal é através de sua decomposição nos componentes de tendência, ciclo e sazonalidade. A tendência de uma série indica o seu comportamento de “longo prazo”, isto é, se ela sobe, desce ou permanece estável, e qual a velocidade dessas mudanças. Nos casos mais comuns, considera-se a tendência constante, linear ou quadrática, conforme Figura 6. (BARROS, 2008).

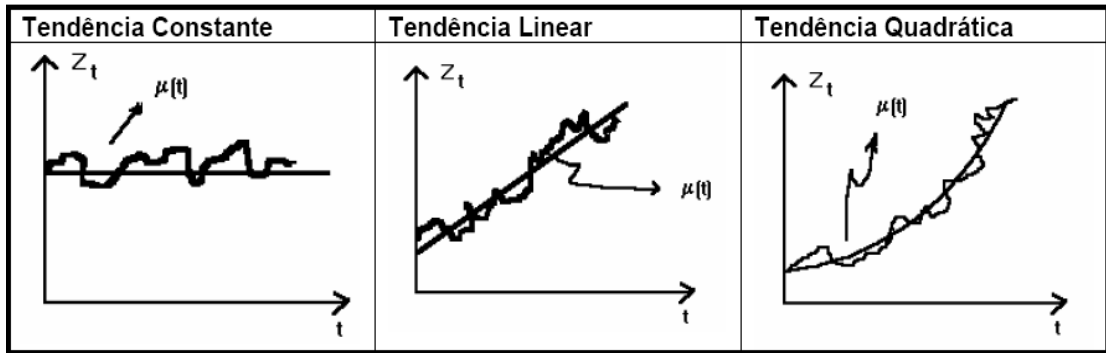


Figura 6 - Tendência de uma série

2.7 CRISP-DM

CRISP-DM (*C*Ross-*I*ndustry *S*tandard *P*rocess for *D*ata *M*ining) é um modelo de processo que fornece um roteiro para a condução de um projeto de DCBD. A metodologia CRISP-DM (2000) possui um conjunto de fases que devem ser cumpridas para cobrir todo o ciclo de vida de um projeto de MD, como mostrado na Figura 7. Nessa figura, as setas indicam as dependências mais frequentes e importantes entre as fases, enquanto o círculo externo simboliza o ciclo normal da DCBD.

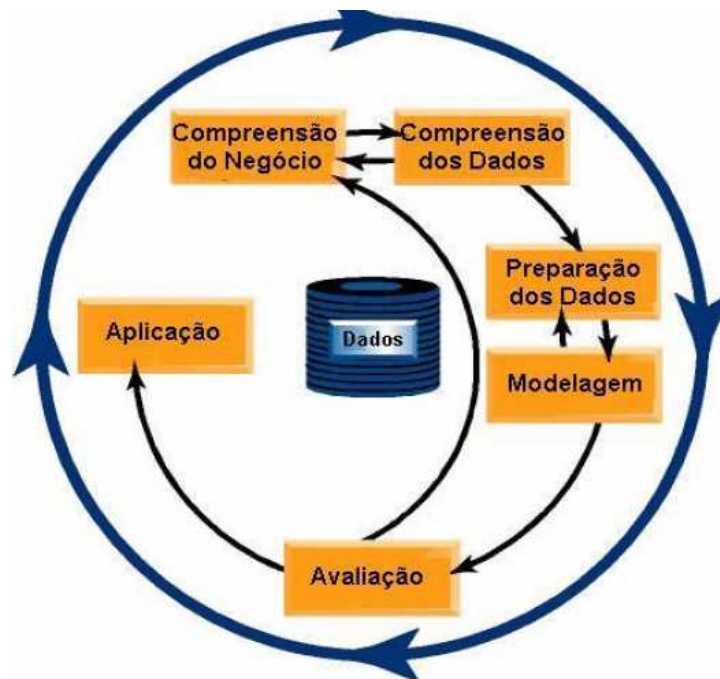


Figura 7 - Fases da metodologia CRISP-DM (CRISP-DM, 2008)

As fases do ciclo de vida do projeto, conforme disposto na figura 7 são as seguintes:

- *Compreensão do negócio.* Nesta fase é que são determinados os objetivos, as necessidades e exigências do cliente a partir da perspectiva do negócio tratado. Nesta fase é muito importante identificar fatores que possam influenciar os resultados de KDD. As tarefas realizadas nesta fase são: (i) determinação dos objetivos do negócio, (ii) avaliação da situação, (iii) determinação das metas de MD e (iv) produção do plano de projeto.
- *Compreensão dos dados.* Envolve identificar quais os dados serão importantes para o estudo que será realizado e uma primeira avaliação do conteúdo dessas informações. As tarefas realizadas nesta fase são: (i) coleta dos dados iniciais, (ii) descrição dos dados, (iii) exploração dos dados e (iv) verificação da qualidade dos dados.
- *Preparação dos dados.* Nesta fase encontra-se o ponto mais importante de todo o processo. É onde será gerado o produto de entrada para a realização do trabalho de MD. É onde será criado e validado o modelo proposto. As tarefas realizadas nesta fase são: (i) seleção dos dados, (ii) limpeza dos dados, (iii) construção dos dados, (iv) integração dos dados e (v) formatação dos dados.
- *Modelagem.* Trata-se do ponto onde são selecionadas e aplicadas as técnicas de MD mais adequadas à situação em estudo, levando em consideração os objetivos pretendidos. No caso de seleção de mais de uma técnica, as mesmas devem ser trabalhadas em separado. As tarefas realizadas nesta fase são: (i) seleção da técnica de modelagem, (ii) geração de teste de desempenho, (iii) construção do modelo e (iv) avaliação do modelo.
- *Avaliação.* Consiste basicamente em verificar se os modelos escolhidos atingiram os objetivos selecionados, revisando os passos seguidos no modelo. As tarefas realizadas nesta fase são: (i) avaliação dos resultados, (ii) revisão do processo e (iii) determinação dos próximos passos.
- *Aplicação.* Compara os resultados obtidos com os objetivos traçados no início do projeto. As tarefas realizadas nesta fase são: (i) planejamento do desenvolvimento, (ii) planejamento do monitoramento e a manutenção, (iii) produção do relatório final e (iv) revisão do projeto.

O detalhamento das tarefas que compõem cada uma das fases da metodologia CRISP-DM, é apresentado no Apêndice A.

2.8 APRENDIZAGEM POR SIMULAÇÃO

A utilização de recursos computacionais como instrumento de aprendizagem é cada vez mais freqüente e, apesar de real, muitas vezes não é perceptível por quem a utiliza.

Ao analisar tudo aquilo que, em nossa forma de pensar, depende da oralidade, da escrita e da impressão, descobriremos que apreendemos o conhecimento por simulação, típico da cultura informática, com os critérios e os reflexos mentais ligados às tecnologias intelectuais anteriores. (LEVY, 2007, Contra-capá).

Levy (2007) afirma ainda que o conhecimento por simulação é uma das novas ferramentas do saber que a ecologia cognitiva informatizada transporta. O autor salienta que as tecnologias intelectuais situam-se fora dos sujeitos cognitivos, como um computador sobre uma mesa ou um livro na mão de seu leitor. Mas elas estão também entre os sujeitos com códigos compartilhados, textos que circulam ao nosso redor, programas que copiamos, imagens que imprimimos e transmitimos por via internet. Ao conectar os sujeitos, as técnicas de comunicação e de representação estruturam uma rede cognitiva coletiva que contribuem para determinação de suas propriedades. As tecnologias intelectuais estão representadas pelos sujeitos por meio da imaginação e da aprendizagem. Mesmo com as mãos vazias e sem se mexer, pensa-se por meio de formas escritas mentais, métodos, regras, compassos, quadros, grafos, oposições lógicas, cantigas algorítmicas, modos de representação e de visualização.

Segundo Levy (2007), a interiorização das tecnologias intelectuais pode ser muito forte, de forma intuitiva, quase um reflexo, como pode ser o conhecimento utilizado em uma linguagem natural utilizada pelo ser humano. O processo de metamorfose sociotécnica era lento na maior parte da sociedade do passado. Entretanto, o seu ritmo acelerou-se nas últimas décadas, primeiro no ocidente, espalhando-se por todo o planeta. A utilização dos recursos informacionais deixou de ser um luxo e passou a ser uma necessidade para quase toda a comunidade. Na maioria da ve-

zes, quando não se conta com a Internet é que se observa e se pensa em como seria difícil se ela não existisse.

Para Langenohl e Schmidt-Beck (2004), enquanto a memória semântica é constituída de conhecimentos no sentido estrito do termo, conhecimentos adquiridos pelos processos de aprendizagem cognitiva, a memória episódica retém representações de eventos vivenciados na vida cotidiana pelo assunto em questão. Nesse sentido, sistemas computacionais que simulem a aquisição de conhecimento (como as Redes Neurais Artificiais, por exemplo) podem representar um importante instrumento de aprendizagem cognitiva, ampliando o conhecimento ao nível da memória semântica.

Langenohl e Schmidt-Beck (2004) ponderam ainda que o processo de aprendizagem pela experiência parece levar à necessidade de se ignorar parte dos conhecimentos adquiridos no ensino formal. O profissional está aprendendo algo de novo todos os dias, atribuindo a si qualidade, forçando-se a não desistir frente a dificuldades e ampliando sua capacidade de aprendizagem a partir das experiências e simulações vivenciadas. Isso reforça a importância do aprendizado por simulação discutida por Levy (2007).

Baseando-se nos conceitos apresentados pelos autores acima citados, a simulação e a vivência de um determinado assunto é fator determinante de agregação de conhecimento ao profissional, principalmente quando vinculado a um assunto tão complexo quanto mercado de ações.

3 MATERIAL E MÉTODOS

3.1 CLASSIFICAÇÃO DA PESQUISA

Segundo Moresi (2004), uma pesquisa pode ser classificada segundo sua natureza, sua abordagem, seus fins e seus meios. Esta pesquisa é de natureza experimental, por buscar um modelo que permita uma melhoria no aprendizado do investidor do mercado de ações. Sua abordagem é quantitativa, por lidar com valores de cotações de ações em um período de 10 anos. Quanto ao meio, é do tipo laboratorial, pois lida com dados históricos e uma ferramenta de MD.

3.2 PRESSUPOSTOS

Este trabalho pressupõe o emprego de fundamentos de apoio à predição de cotações de ações, considerando seus dados históricos, favorecendo o aprendizado sobre o investimento em ações.

3.3 COLETA E ANÁLISE DOS DADOS

Para a realização deste trabalho, foi realizada uma busca em fontes de valores históricos de cotações de ações. Privilegiou-se a Bovespa como fonte principal (BOVESPA, 2008), de onde foram obtidas cotações históricas no período compreendido entre janeiro de 1986 a setembro de 2008.

O arquivo de séries históricas é composto por três tipos de registros: (i) registro de cotações dos papéis por dia (Quadro 2), (ii) Registro *header* (Quadro 3) e (iii) Registro *trailer* (Quadro 4).

Nome do Campo/Descrição	Conteúdo	Tipo e Tamanho	Posição Inicial	Posição Final
TIPREG-Tipo de Registro	Fixo "01"	N(02)	01	02
Data do pregão	Formato "AAAAMMDD"	X(08)	03	10
CODBDI – Código BDI	Ver Tabela apêndice B	X(02)	11	12
CODNEG – Código de negociação do papel		X(12)	13	24
TPMERC- Tipo de mercado	Ver Tabela apêndice B	N(03)	25	27
NOMRES- Nome resumido da empresa emissora do papel		X(12)	28	39
ESPECI-Especificação do papel	Ver Tabela apêndice B	X(10)	40	49
PRAZOT- Prazo em dias do mercado a termo		X(03)	50	52
MODREF- Moeda de referência	Moeda usada pregão	X(04)	53	56
PREABE – Preço de abertura do papel – mercado do pregão		N11V99	57	69
PREMAX – Preço máximo do papel – mercado do pregão		N11V99	70	82
PREMIN – Preço mínimo do papel – mercado do pregão		N11V99	83	95
PREMED – Preço médio do papel – mercado do pregão		N11V99	96	108
PREULT – Preço último negócio do papel – mercado do pregão		N11V99	109	121
PREOFC – Preço da melhor oferta de compra do papel – mercado do pregão		N11V99	122	134
PREOFV – Preço da melhor oferta de venda do papel – mercado do pregão		N11V99	135	147
TOTNEG – NEG – Número de negócios efetuados com o papel – mercado do pregão		N(05)	148	152
QUATOT – Quantidade total de títulos negociados nesta papel – mercado		N(18)	153	170
VOLTOT – Volume total de títulos negociados nesta papel – mercado		N11V99	171	188
PREEXE – Preço de exercício para o mercado de opções ou valor do contrato para o mercado de opções ou termo secundário		N11V99	189	201
INDOPC – Indicador de correção de preços de exercícios ou valor do contrato para o mercado de opções ou termo secundário	Ver Tabela apêndice B	N(01)	202	202
DATVEN – Data do vencimento para o mercado de opções ou termo secundário	Formato "AAAAMMDD"	N(08)	203	210
FATCOT – Fator de cotação do papel	"1" -cotação.unitária "1000" – lote 1000 p/lote	N(07)	211	217
PTOEXE – Preço de exercício em pontos para opções referenciadas em dólar ou valor de contrato em pontos para termo secundário		N07V06	218	230
CODISI – Código do papel no sistema ISIN ou código interno do papel	A partir de 15.05.1995. Ver apêndice B	X(12)	231	242
DISMES - Número de negociação		N(03)	243	245

Quadro 2 - Registro - 01 - a Cotações dos papéis por dia (BOVESPA, 2008)

Nome do Campo/Descrição	Conteúdo	Tipo e Tamanho	Posição Inicial	Posição Final
Tipo de Registro	Fixo "00"	N(02)	01	02
Nome do Arquivo	Fixo "COTAHIST.AAAA"	X(13)	03	15
Código da Origem	Fixo "BOVESPA"	X(08)	16	23
Data da Geração do Arquivo	Formato "AAAAMMDD"	N(08)	24	31
reserva	Preencher com brancos	X(214)	32	245

Quadro 3 - Detalhes do Registro 00 – Header (BOVESPA, 2008)

Nome do Campo/Descrição	Conteúdo	Tipo e Tamanho	Posição Inicial	Posição Final
Tipo de Registro	Fixo "99"	N(02)	01	02
Nome do Arquivo	Fixo "COTAHIST.AAAA"	X(13)	03	15
Código da Origem	Fixo "BOVESPA"	X(08)	16	23
Data da Geração do Arquivo	Formato "AAAAMMDD"	N(08)	24	31
Total de registros	Incluir também os registros header e trailer	N(11)	32	42
Reserva	Preencher com brancos	X(214)	43	245

Quadro 4 - Detalhes do Registro 99 – Trailer (BOVESPA, 2008)

Foi utilizada a metodologia CRISP-DM para a construção do modelo, por abranger todas as etapas de DCBD.

3.4 DELIMITAÇÃO DO ESTUDO

O estudo foi desenvolvido a partir de dados referentes à série histórica de 10 anos (janeiro de 1998 a setembro de 2008) de um conjunto de ações de primeira linha, as chamadas *blue chips*.

4 ESTUDO DE CASO

Neste capítulo é descrito o desenvolvimento do modelo de predição, com base na metodologia CRISP-DM. A apresentação segue a estrutura de fases e nomenclatura dessa metodologia.

4.1 FASE 1: COMPREENSÃO DO NEGÓCIO

4.1.1 Determinação dos objetivos do negócio

Os mercados financeiros mundiais vêm apresentando uma grande evolução nos últimos anos, tornando-se muito mais competitivos, complexos e globais. Em uma economia globalizada, onde cada vez mais os mercados estão dependentes de capitais internacionais para impulsionar a sua economia, o mercado brasileiro também está se adequando cada vez mais a esta realidade. Com a estabilidade econômica ocorrida nos últimos anos, o mercado de títulos brasileiros tornou-se atrativo e seguro para investidores estrangeiros, sejam eles públicos ou privados. Em uma carteira de investidores é comum se encontrar ações de empresas públicas ou privadas.

O foco deste trabalho é o mercado de ações, mais especificamente considera-se uma carteira de ações consideradas de primeira linha. A antecipação de valores de cotações de ações dessa carteira significará mais uma importante fonte de conhecimento e aprendizado para o investidor.

Uma ação representa a menor parcela em que se divide um capital social de uma empresa. O investidor em ações é um co-proprietário da sociedade anônima da qual é acionista, participando de seus resultados através de dividendos distribuídos aos acionistas por ocasião periódica do fechamento de balanços. As ações são conversíveis em moeda corrente a qualquer tempo, através de negociação em bolsas de valores ou mercado de balcão devidamente organizado.

4.1.2 Avaliação da situação

A aplicação no mercado de ações é uma operação de alto risco, devido à sua imprevisibilidade, posto que esse mercado pode ser influenciado positiva ou negativamente por fatores internos ou externos. Esses fatores podem ser, por exemplo, políticos, climáticos ou mesmo o humor dos investidores ao comprar ou vender ações. Um exemplo real que o mercado americano tem vivenciado recentemente é o problema com as empresas financeiras que aplicaram seus recursos em operações com hipotecas imobiliárias. A inadimplência desse tipo de operação agregada à baixa dos valores dos imóveis ocasionou a quebra de alguns bancos do setor. Esses acontecimentos têm influenciado significativamente o mercado de compra e venda de ações. Nesse mercado uma margem variável de acerto na predição de valores das cotações de ações em torno de 80% seria de grande valia, tendo em vista os altos valores negociados.

Segundo Piazza (2008), os riscos em aplicações no mercado de ações podem ser reduzidos caso o investidor observe fatores tais como:

- Não operar contra a tendência de mercado;
- Não investir em apenas uma empresa;
- Procurar diversificar seu capital em dois ou mais papéis, montando uma carteira (conjunto de ativos que um investidor possui em seu poder em um determinado momento);
- Ter uma estratégia bem definida e segui-la;
- Saber o momento de vender ou comprar.

4.1.3 Determinação das metas de mineração de dados

O objetivo deste trabalho é a proposição de um modelo preditivo do valor de cotação de ações de uma carteira pré-definida, visando apoiar a decisão sobre investimentos em ações.

4.1.4 Produção do plano de projeto

O plano de projeto deverá ser composto dos seguintes passos:

- Buscar valores das cotações de ações da BOVESPA;
- Tratar as cotações obtidas;
- Definir a carteira;
- Definir a periodicidade para aplicação da estratégia;
- Escolher a ferramenta de MD;
- Definir um modelo para tratamento dos dados na ferramenta de MD;
- Preparar os dados para aplicação em uma ferramenta de predição de valores;
- Treinar a RNA para um ano de amostra.
- Aplicar os dados na ferramenta selecionada com uma série de 10 anos;
- Analisar os resultados comparando-os com as cotações reais;
- Fornecer o modelo criado para investidores que atuam no mercado de ações.

4.2 FASE 2: COMPREENSÃO DOS DADOS

4.2.1 Coleta dos dados iniciais

Os dados a serem utilizados no estudo de caso foram obtidos no *site* da Bolsa de Valores de São Paulo². Os mesmos são de uso público e estão disponíveis a todos os interessados no mercado de ações, bastando se cadastrar e realizar o *download* das informações.

Os dados disponíveis são compostos dos seguintes arquivos:

- Arquivo contendo o leiaute dos arquivos de cotações, com o nome de *SeriesHistoricas_Layout.pdf*;
- Arquivos contendo as cotações mensais das ações, com o nome de *CO-TAHIST_Mxxaaaa.pdf*, onde *xx* representa o mês e *aaaa* representa o ano das cotações;

² www.bovespa.com.br

- Arquivos contendo as cotações diárias das ações, com o nome de CO-TAHIST_Axxxx, onde o xxxx representa o ano das cotações. O arquivo contém as cotações diárias de todos os dias úteis de um determinado ano;

4.2.2 Descrição dos dados

A descrição dos arquivos obtidos junto à Bovespa estão descritos no capítulo Materiais e Métodos e o detalhamento das tabelas dos campos que compõem esses arquivos está no Apêndice B.

4.2.3 Exploração dos dados

Os dados disponibilizados pela Bovespa referem-se a *(i)* uma série histórica mensal que se inicia em janeiro de 2007 e *(ii)* uma série histórica diária que se inicia em 01 de janeiro de 1986. Neste trabalho, foi utilizada a série histórica diária referente ao período de 01 de janeiro de 1998 e 31 de agosto de 2008. O volume de dados envolvidos nessa série corresponde a um total de 1.903.767 observações. Nesses dados, verificou-se a predominância das ações dos tipos ON e PN, com suas variações. A Tabela 2 reflete os totais dessas ações em relação ao total de observações consideradas.

Apesar de se verificar um número maior de ações do tipo PN, essas estão diluídas em 31 variações, enquanto que o tipo ON só apresenta 4 variações, como se pode verificar na Tabela 3 do Apêndice B.

Tabela 2 - Ações do tipo ON e PN movimentadas no período considerado

Ano	Volume total	Tipos de ações mais movimentadas					
		ON		PN		Outras	
		Total	%	Total	%	Total	%
1998	104.972	30.874	29,4	70.852	67,5	3.246	3,1
1999	129.843	41.335	31,8	83.245	64,1	5.263	4,1
2000	153.684	44.006	28,6	101.806	66,2	7.872	5,1
2001	141.850	38.646	27,2	96.506	68,0	6.698	4,7
2002	129.975	36.123	27,8	88.004	67,7	5.848	4,5
2003	148.114	40.630	27,4	101.530	68,5	5.954	4,0
2004	166.939	45.789	27,4	114.932	68,8	6.218	3,7
2005	181.410	49.759	27,4	123.428	68,0	8.223	4,5
2006	198.363	65.645	33,1	121.724	61,4	10.994	5,5
2007	280.257	112.892	40,3	149.218	53,2	18.147	6,5
2008*	268.360	113.923	42,5	136.478	50,9	17.959	6,7
Total	1.903.767	619.622		1.187.723		96.422	

* Até agosto.

4.2.4 Verificação da qualidade dos dados

Quanto à qualidade dos dados selecionados, foi verificado que não existem cotações para todos os dias úteis do ano, variando em função da ação da empresa selecionada ter sido ou não comercializada naquele dia. Foi verificado ainda que as demais informações, constantes dos diversos campos existentes nos arquivos são de excelente qualidade e não foram verificadas inconsistências de nenhuma natureza nos dados selecionados.

4.3 FASE 3: PREPARAÇÃO DOS DADOS

4.3.1 Seleção dos dados

A seleção inicial considerou apenas dois tipos específicos de ações: (i) PN, ações ao portador que não dá direito a voto nas assembleias de acionistas, e (ii) ON, ações ordinárias nominativas que dá direito a voto nas assembleias de acionistas.

Dentre as empresas constantes das cotações apresentadas, foi definida uma carteira de ações formada pelas empresas constantes no Quadro 6. Para isso, foi utilizado o campo NOMRES (ver Quadro 5), que representa a identificação da empresa emissora do papel. Dessa forma, o número de cotações ON e PN dessa carteira dentro do período considerado é mostrado na Tabela 3. Por conveniência no tratamento dos dados, foram gerados arquivos por empresa e por ano.

#	Nome da empresa	Atividade	Ramo
(1)	Bradesco	Financeira	Privado
(2)	Banco do Brasil	Financeira	Público
(3)	Embraer	Aviação	Privado
(4)	Petrobras	Petroquímico	Público
(5)	Telesp	Telefonia	Privado
(6)	Vale do Rio Doce	Mineração	Privado

Quadro 5 - Empresas componentes da carteira de ações considerada

Tabela 3 - Total de cotações referentes da carteira de ações considerada

Ano	Total de ações ON ou PN	Total de cotações das empresas da carteira considerada*							
		(1)	(2)	(3)	(4)	(5)	(6)	Total	%
1998	101.726	1.134	2.420	612	966	1.029	823	6.984	6.984
1999	124.580	1.087	2.447	826	973	1.044	904	7.281	7.281
2000	145.812	1.107	2.474	992	993	1.010	908	7.484	7.484
2001	135.152	1.127	2.166	983	984	984	904	7.148	7.148
2002	124.127	996	1.808	1.023	996	1.003	989	6.815	6.815
2003	142.160	1.195	1.494	1.000	1.001	1.000	1.000	6.690	6.690
2004	160.721	1.553	1.604	996	996	996	996	7.141	7.141
2005	173.187	1.141	1.491	993	996	996	996	6.613	6.613
2006	187.369	1.094	1.418	699	985	984	984	6.164	6.164
2007	262.110	2.481	2.319	1.058	3.917	1.215	3.712	14.702	14.702
2008*	250.401	2.097	1.811	917	3.708	837	3.752	13.122	13.122
Total	1.807.345	15.012	21.452	10.099	16.515	11.098	15.968	90.144	

* (1) Bradesco, (2) Banco do Brasil, (3) Embraer, (4) Petrobras, (5) Telesp, (6) Vale do Rio Doce.

4.3.2 Limpeza dos dados

Por ser uma base de dados gerada pela própria Bovespa respeitando critérios de integridade, consistência e confiabilidade, não foram identificados ruídos ou incorreções na mesma.

4.3.3 Construção dos dados

Em função dos objetivos deste trabalho, foram selecionados, dentre os campos existentes, os constantes no Quadro 6.

Nome do campo na tabela	Conteúdo do campo
DATA-AAAA	Ano da cotação
DATA-MM	Mês da cotação
DATA-DD	Dia da cotação
NOMRES	Nome resumido da empresa
ESPECI	Tipo de ação
PREABE	Preço de abertura do papel no pregão
PREMAX	Preço máximo do papel no pregão
PREMIN	Preço mínimo do papel no pregão
PREMED	Preço médio do papel no pregão
PREULT	Preço do último negócio do papel no pregão
PREOFC	Preço da melhor oferta de compra do papel
PREOFV	Preço da melhor oferta de venda do papel
FATCOT	Fator de cotação do preço
TOTNEG	Número de negócios efetuados com o papel

Quadro 6 - Campos selecionados para a realização dos trabalhos

A partir de uma análise das cotações existentes para os tipos de ação ON e PN, verificou-se que (i) não existiam cotações para todos os dias úteis do período considerado e (ii) as ações do tipo ON possuíam maior quantidade de cotações diárias. Isso torna as ações do tipo ON constituem um conjunto mais homogêneo, o que garante a construção de um modelo de melhor qualidade. Assim, foram selecionadas as ações do tipo ON, conforme mostrado no Quadro 7.

Empresas	Histórico de negociação
Banco do Brasil	[1998-2004]: AÇÕES ON * [2004-2006]: AÇÕES ON [2006-2008]: AÇÕES ON NM
Bradesco	[1998-2001]: AÇÕES ON * [2001-2004]: AÇÕES ON * E N1 [2004-2008]: AÇÕES ON NM
Petrobrás	[1998-2000]: AÇÕES ON [2000-2008]: AÇÕES ON *
Telesp	[1998-2005]: AÇÕES ON * [2001-2002]: AÇÕES ON *ANT [2005-2008]: AÇÕES ON
Embraer	[1998-2000]: AÇÕES ON [2000-2008]: AÇÕES ON *
Vale do Rio Doce	[1998-2003]: AÇÕES ON [2003-2008]: AÇÕES ON N1

Quadro 7 - Empresas e tipos de ações

Após a seleção do tipo ON, foi detectada a existência de mais de uma cotação para o mesmo dia, para os tipos de ações selecionados. Como a técnica utilizada para a geração da série temporal exige que se tenha apenas uma cotação diária, foi necessário um critério para selecionar a cotação que desse maior cobertura temporal. Para tanto, utilizou-se o campo CODBDI (ver Quadro 2 e Apêndice B), que identifica o tipo de negociação realizada na bolsa. Dentre os tipos de negociação existentes (lote padrão, leilão, permuta e fracionário), foi selecionado, segundo o critério definido, o tipo *lote padrão*. Na Tabela 4 é apresentada a quantidade de registros por empresa.

Durante a montagem dos dados contendo as informações que eram necessárias para se trabalhar na ferramenta de predição de valores, verificou-se a falta de cotação para alguns dias úteis dos anos selecionados. Nesse caso, foi incluído um novo registro preenchido com os valores das cotações verificadas no último dia útil disponível. Isso reflete uma prática do mercado, que considera última cotação como fator básico para a determinação do preço do dia. As cotações de ações de todas as empresas foram tratadas com esse procedimento, visando garantir a completude das informações componentes do arquivo da carteira selecionada e que serão tratadas posteriormente. Assim, ao arquivo com as cotações da carteira que contava, inicialmente, com 12.951 registros, foram incluídos 2.871 novos registros, o que representa um total de 2.637 cotações diárias para cada empresa.

Tabela 4 - Total de cotações de ações ON referentes a ações da carteira considerada

Ano	Total de cotações das empresas da carteira considerada						Total
	(1)	(2)	(3)	(4)	(5)	(6)	
1998	125	225	62	239	230	158	1.039
1999	127	229	203	237	192	225	1.213
2000	138	201	201	216	224	238	1.218
2001	144	160	216	213	246	228	1.207
2002	152	232	211	233	238	241	1.307
2003	142	233	233	233	222	226	1.289
2004	136	211	214	217	226	225	1.229
2005	139	229	217	220	217	233	1.255
2006	143	218	222	222	222	222	1.249
2007	138	151	213	213	213	221	1.149
2008*	93	118	148	139	141	157	796
Total	1.477	2.207	2.140	2.382	2.371	2.374	12.951

* (1) Bradesco, (2) Banco do Brasil, (3) Embraer, (4) Petrobras, (5) Telesp, (6) Vale do Rio Doce.

Foi analisada também a ocorrência de desmembramentos ou consolidações de lotes de ações. Por exemplo, se a cotação de mercado das ações da Petrobrás fosse R\$ 200,00 e ocorresse o seu desmembramento, de tal forma que o lote original fosse composto de 100 ações e o lote futuro fosse composto de 10 ações, o seu valor de mercado poderia ter se alterado de R\$ 200,00 para R\$ 20,00. Esta é uma prática comum no mercado de ações que se reflete na composição do preço das ações e pode afetar negativamente a construção do modelo preditivo. O arquivo foi integralmente inspecionado para detectar esse tipo de ocorrência. Para isso, foi utilizado o campo referente ao Fator de Cotação de Preço (campo FATCOT do Quadro 2), que identifica a quantidade de ações por lote comercializado. Para a carteira de ações considerada, não foi constatada nenhum desmembramento ou consolidação de lotes de ações.

4.3.4 Integração dos dados

Esta etapa não aplica a este trabalho, uma vez que os dados utilizados vieram de uma única fonte de informação na forma de uma base de dados já integrada.

4.3.5 Formatação dos dados

Visando adequar a formatação dos dados para a construção do modelo, os dados de cada ano foram organizados em uma única série anual, na qual o primeiro dia corresponde ao primeiro dia útil do ano e o último dia corresponde ao último dia útil do ano. Os dias das semanas foram tomados como grupos de cinco dias úteis e os dos meses levando em consideração o próprio calendário. Esta opção foi adotada como forma de padronização da série. Nos seguintes campos foi registrada a padronização adotada:

- DATA-UTIL-SEMANA representa o dia útil da semana. O arquivo foi iniciado como se o primeiro dia útil de 1998 fosse o primeiro dia útil da semana.
- DIA-UTIL-MÊS representa o dia útil do mês, considerando a partir do primeiro dia útil do mês. Dependendo do mês e da quantidade de feriados, pode ocorrer variações entre 18 e 23 dias úteis.
- DIA-UTIL-ANO representa o dia útil do ano, considerando a partir do primeiro dia útil do ano. Dependendo da quantidade de feriados no ano, a quantidade de dias úteis pode variar de 245 a 250.

4.4 FASE 4: MODELAGEM

4.4.1 Seleção da técnica de modelagem

Uma vez preparados os dados para a geração do modelo e, considerando que a técnica definida a ser utilizada são as RNA, partiu-se para a escolha da ferramenta mais adequada para a realização dos trabalhos. Nesta seção discutem-se três opções de ferramentas populares para a construção de modelos de séries temporais: EasyNN, MatLab e Weka.

EasyNN (2008) é uma ferramenta que implementa uma RNA do tipo *multilayer perceptron*, algoritmo de treinamento com retropropagação e função de ativação logística. Possui como características principais (i) opção para otimizar as taxas de aprendizado e de momento com base em uma estrutura de envoltório (Kohavi, 1995), (ii) opção para redução das taxas de aprendizado e de momento durante o





treinamento por meio de intervenções diretas durante o processo de aprendizado e (iii) cálculo automático do número de neurônios das camadas intermediárias. O EasyNN oferece os seguintes critérios de parada:

- Quando a média quadrática dos erros calculados sobre o conjunto de treinamento for menor que um valor pré-determinado,
- Quando todos os valores dos erros forem menores que um valor determinado, em um determinado ciclo de treinamento, em um tempo decorrido determinado e
- Quando ocorre aumento da taxa de erros nos dados de validação da RNA.

A ferramenta possui três modos de operação: treinamento, validação e teste. A carga dos dados de treinamento é relativamente simples, permitindo arquivos no formato .txt ou .csv. A primeira linha do arquivo deve conter o nome dos campos sendo que, para cada campo constante da tabela deve ser selecionado o seu tipo e especificado se o campo é de entrada, saída ou excluído no tratamento da RNA.

Uma vez carregado o arquivo, o passo seguinte será a definição dos campos que irão receber os valores estimados para as variáveis de saída, conforme apresentado na Figura 8. As linhas que receberão os valores previstos pela rede após o treinamento da RNA, são devidamente selecionadas através da opção de *validating*.

O próximo passo é a parametrização da RNA. Entre outros parâmetros, a ferramenta permite que se especifique (i) o tempo para a geração de cada RNA quando se optar pela otimização de algum parâmetro, (ii) parada por limite de erros e (iii) parada por número máximo de ciclos. Uma vez parametrizada a rede, o sistema exibe a configuração completa e habilita o modo de treinamento.

Para iniciar o treinamento da RNA, deve-se selecionar a opção *Action – Start Learning*, do menu principal. Com isso, o sistema disponibiliza uma tela com os parâmetros de treinamento e o sistema inicia o seu aprendizado. A RNA construída pode ser visualizada a partir do acionamento de um ícone específico localizado na barra perpendicular direita do aplicativo [], conforme figura 9. É possível também visualizar, mesmo durante o treinamento: (i) a evolução da taxa de erro [], (ii) a importância de cada um dos campos de entrada no treinamento da rede [], (iii) o estado da rede, ou seja, os valores associados a cada um dos nós da rede [] e (iv) as

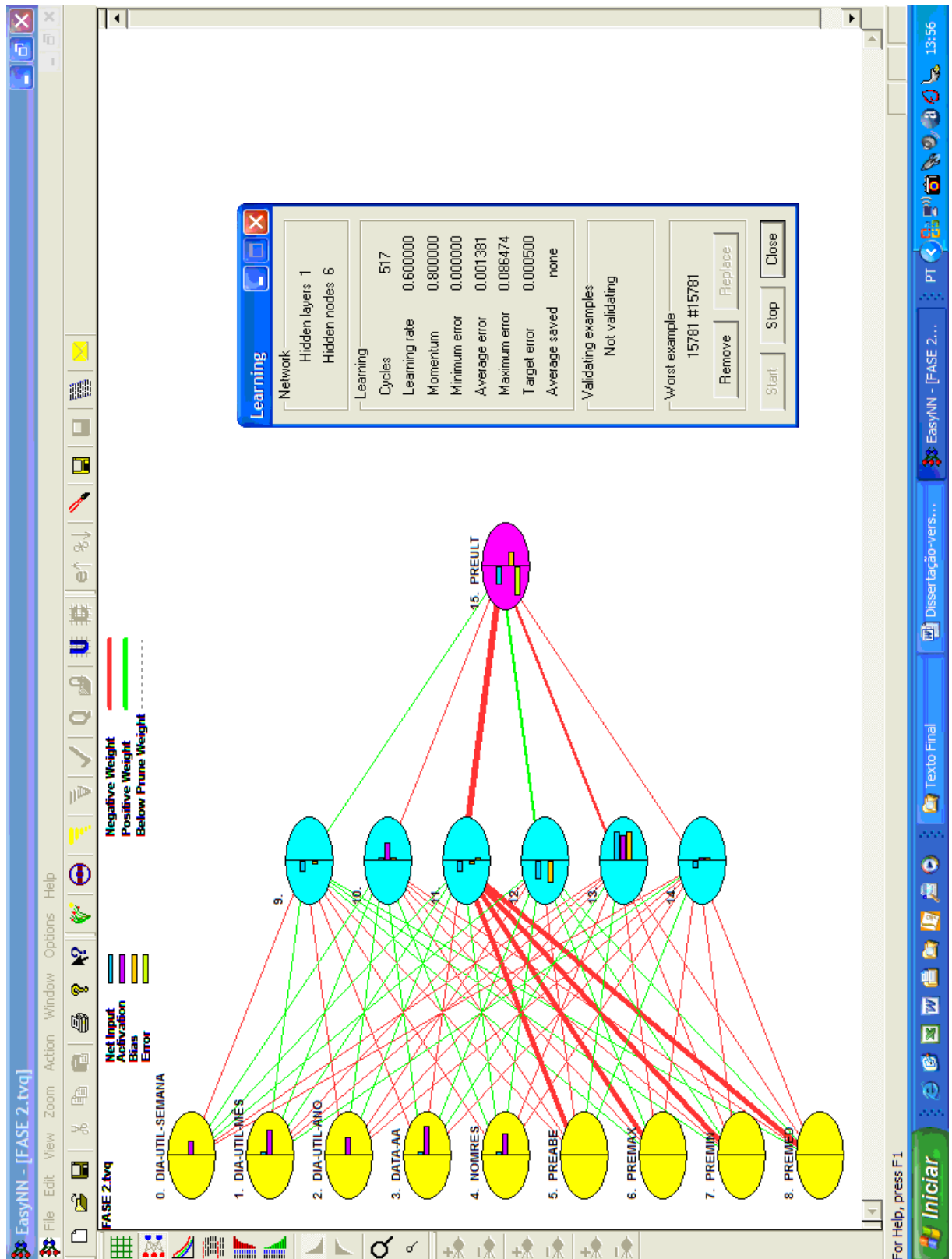


Figura 9 - Ferramenta Easynn - RNA

Ao final do treinamento da RNA para a obtenção dos valores previstos para as cotações, é possível visualizar os valores previstos na sequência dos valores de treinamento³.

É preciso atentar para um pequeno problema na ferramenta. Foi verificado que a existência de um espaço em branco logo após o último registro do arquivo pode ocasionar erro na carga dos dados.

O Weka (*Waikato Environment Knowledge Analysis*) (WAIKATO, 2007) é um produto de uso liberado, onde é possível realizar alterações em sua programação em função das necessidades detectadas. Seu manuseio, porém, não é intuitivo, tendo como fator negativo a sua navegabilidade e usabilidade, levando em consideração a possível utilização por parte de usuários que não estão inseridos na área de TI.

A ferramenta oferece quatro modos de operação:

- *Simple CLI* apresenta uma janela para executar os algoritmos através de linha de comando. Isto porque a primeira versão do Weka não dispunha de uma interface gráfica e as classes de seus pacotes eram executadas por linhas de comando.
- Explorer oferece o módulo gráfico para execução dos algoritmos, com as seguintes abas:
- *Preprocess* permite as funções de abrir arquivo, abrir URL, abrir um banco de dados, editar o banco de dados aberto e salvar arquivos de dados. Oferece ainda alguns métodos para filtragem de dados e a visualização da distribuição das variáveis a seleção de variáveis a serem utilizadas no treinamento.
- *Classify* é destinada a configurar e executar tarefas de classificação e regressão, como a escolha de um dos algoritmos de classificação oferecidos, definição de opções de testes (conjunto próprio de treinamento, conjunto independente de treinamento, validação cruzada ou percentual de dados a ser utilizada no teste, entre outras funções) e execução dos algoritmos com visualização dos resultados.

³ Isso é feito selecionando a linha definida para predição e pressionando a tecla <enter> no campo de saída.

- *Cluster* é voltada para a análise de agrupamentos, com a escolha entre um conjunto de algoritmos com essa finalidade, parametrização e execução dos mesmos, com visualização dos resultados.
- *Associate* permite a escolha, configuração e execução de algoritmos para a geração de regras de associação.
- *Select attributes* oferece diversos algoritmos para a seleção de atributos, tais como busca exaustiva e algoritmos genéticos.
- *Visualise* permite a visualização dos dados em diversos tipos de gráficos bidimensionais.
- *Experimenter* permite definir experimentos mais complexos, visando executar um ou mais algoritmos sobre um ou vários conjuntos de dados de entrada, e comparar estatisticamente seus resultados.
- *KnowledgeFlow* permite executar as mesmas opções do *Explorer*, com uma configuração gráfica inspirada em ferramentas do tipo *data-flow* para selecionar componentes e conectá-los em projetos de MD. Por exemplo, poderia-se construir um projeto com as seguintes etapas: carga dos dados, aplicação de algoritmos de tratamento e análise, escolha do tipo de avaliação desejada, visualização dos resultados.

MatLab (*Matrix Laboratory*) é um programa interativo para cálculos numéricos e geração de gráficos científicos. Seu ponto forte está na manipulação e cálculos matriciais, por exemplo, resolução de sistemas lineares, cálculo de autovalores e autovetores e fatoração de matrizes. Além disso, muitas funções especializadas já estão internamente implementadas. Em muitos casos, não há a necessidade de se escrever muito código adicional. Outros dois pontos fortes da ferramenta é a possibilidade de criação e manipulação de gráficos científicos e a possibilidade de extensão por meios de pacotes comerciais ou escritos pelo próprio usuário (TONINI, 2002).

MatLab oferece diversas opções de manipulação. Para a criação de um modelo, é possível ao usuário escolher o tipo de modelo desejado em função do tipo ou das características do assunto tratado. Por exemplo, a interface para a criação, parametrização, visualização e execução de uma RNA é habilitada pelo comando *nn-toll*. A Figura 10 apresenta essa interface.

As ferramentas foram analisadas em termos de tempo de aprendizagem e dificuldade de operação. Verificou-se que a ferramenta EasyNN atende os requisitos

necessários para a realização dos trabalhos, levando-se em conta a sua facilidade de operação e interação.

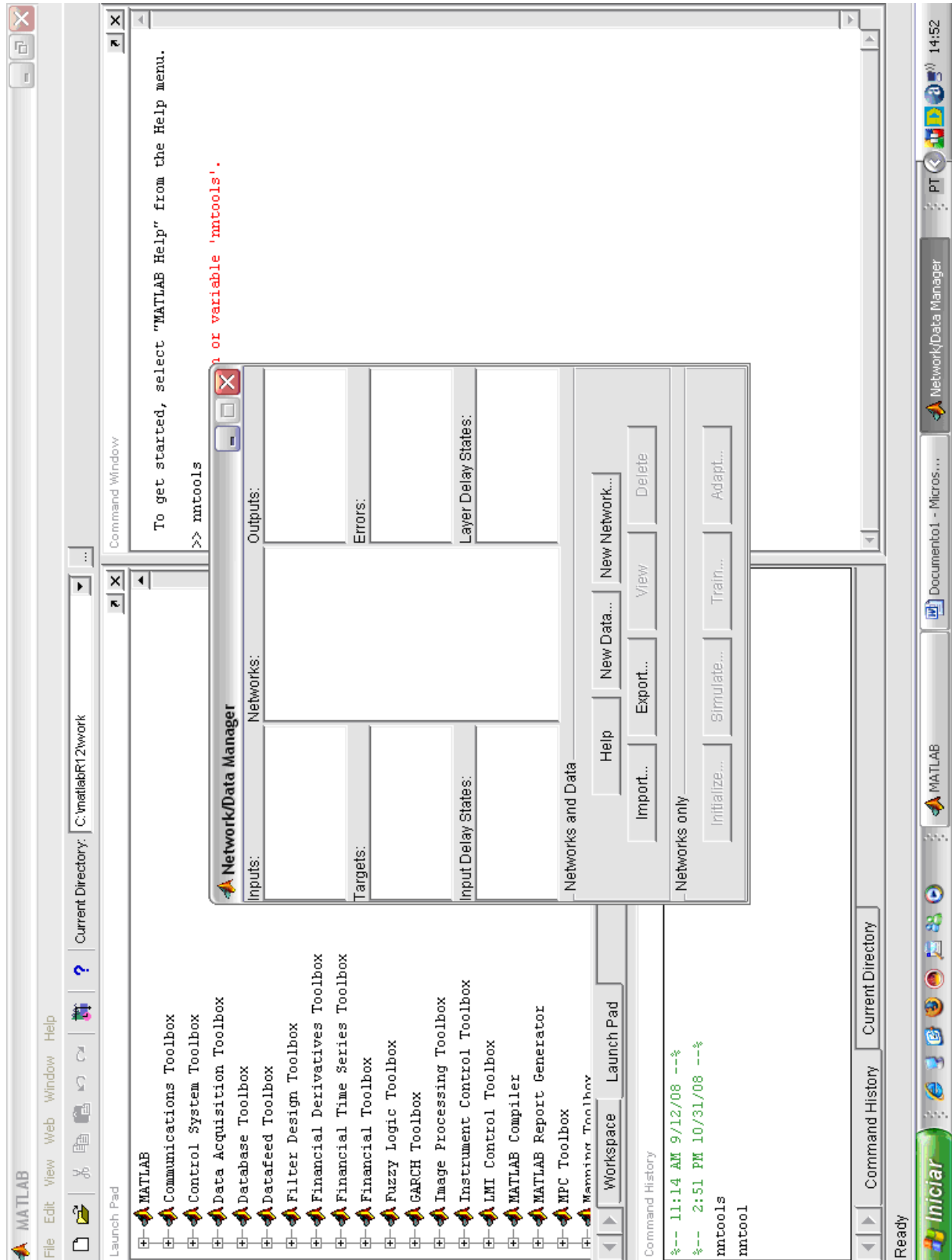


Figura 10 - Ferramenta MatLab – Criação RNA

Cada uma das ferramentas analisadas possuem o seu diferencial específico. A escolha recaiu sobre o EasyNN pelo domínio dessa ferramenta pelo autor.

O *hardware* para o desenvolvimento dos trabalhos foi um *notebook* com processador AMD Sempron “Móbile 3000”, frequência de processamento de 1,8 gigahertz, com 1,12 gigabytes de *Random Access Memory* (RAM).

Para a redundância de eventos de falhas ou sobrecarga de processamento no processo de treinamento das RNA foi utilizado um *desktop* Pentium IV, frequência de processamento de 2,8 GHz, com 1 GB de memória RAM. Ambos os hardwares possuem memória cache de 1 MB, que possibilita uma maior velocidade de processamento dos dados.

Servindo como equipamento complementar que, foi disponibilizado ainda um notebook com processador AMD Turion X2 *Dual-Core*, frequência de processamento de 2 GHz, com 4 GB de *Random Access Memory* (RAM), disco rígido de 250 GB e sistema operacional *Windows Vista Home Premium* 64 bits.

4.4.2 Geração de teste de desempenho

Inicialmente, foram gerados oito conjuntos de dados para a construção da série histórica considerando-se o seguinte esquema: (i) treinamento com dados de janeiro de 1998 a dezembro de 2007 e teste com dados de janeiro de 2008; (ii) treinamento com dados de janeiro de 1998 a janeiro de 2008 e teste com dados de fevereiro de 2008; (iii) treinamento com dados de janeiro de 1998 a fevereiro de 2008 e teste com dados de março de 2008; (iv) treinamento com dados de janeiro de 1998 a março de 2008 e teste com dados de abril de 2008; (v) treinamento com dados de janeiro de 1998 a abril de 2008 e teste com dados de maio de 2008; (vi) treinamento com dados de janeiro de 1998 a maio de 2008 e teste com dados de junho de 2008; (vii) treinamento com dados de janeiro de 1998 a junho de 2008 e teste com dados de julho de 2008; e (viii) treinamento com dados de janeiro de 1998 a julho de 2008 e teste com dados de agosto de 2008.

Foram realizados oito ensaios com estes conjuntos de dados, selecionando-se um mes da série histórica para a validação do modelo. Os oito conjuntos utilizados para a validação encontram-se no Apêndice D.

As variações dos resultados do treinamento das oito redes, comparados com valores reais são apresentados na Tabela 5.

Tabela 5 – Resultado do treinamento para geração de teste de desempenho

Mês	Menor	Maior
jan/08	-1,6801	5,5667
fev/08	-7,0250	0,9243
mar/08	-4,8912	1,1304
abr/08	-2,3508	5,2291
mai/08	-3,1226	2,1954
jun/08	-2,4531	2,7270
jul/08	-3,3942	6,3448
ago/08	-3,2075	3,6251
Média	-3,5156	3,4679

4.4.3 Construção do modelo

A visualização das RNA geradas através da ferramenta utilizada encontra-se disponível do Apêndice E.

Foram geradas doze RNA, combinando as variáveis de entrada, como apresentado na Tabela 6. Dessas RNA, seis (R_1 - R_6) referem-se à predição do mês de agosto de 2008 a partir do treinamento com o restante dos dados (de janeiro de 1998 a julho de 2008) e seis (R_7 - R_{12}) referem-se à predição da segunda quinzena do mês de agosto de 2008 a partir do treinamento com o restante dos dados (de janeiro de 1998 até a primeira quinzena de agosto de 2008).

A escolha dos períodos de uma quinzena e um mês para predição foi baseada nos estudos de Mueller (1996), segundo o qual períodos de predições com prazos muito longos tendem a gerar altos índices de erros.

No Apêndice C são apresentadas as 12 RNA com o desenho dos neurônios de entrada, saída e processamento geradas na ferramenta de mineração.

Tabela 6 - Campos de entrada e saída das RNA

Campos de entrada	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	R ₇	R ₈	R ₉	R ₁₀	R ₁₁	R ₁₂
dia-util-semana	x	x	x	x	x	x	x	x	x	x	x	x
dia-util-mês	x	x	x	x	x	x	x	x	x	x	x	x
dia-util-ano	x	x	x	x	x	x	x	x	x	x	x	x
data-aa	x	x	x	x	x	x	x	x	x	x	x	x
data-mm	x	x				x			x	x	x	
data-dd	x	x				x			x	x	x	
nomres	x	x	x	x	x	x	x	x	x	x	x	x
preabe	x		x				x		x			
premax	x		x				x		x			
premin	x		x				x		x			
premed	x		x				x		x			
Campos de saída												
preabe					x	x					x	x
premax					x	x					x	x
premin					x	x					x	x
premed					x	x					x	x
Preult	x	x	x	x	x	x	x	x	x	x	x	x

4.4.4 Avaliação do modelo

Os resultados finais após a impostação dos dados são apresentados nas Tabelas 7 e 8. Alguns desses parâmetros foram utilizados para a criação e treinamento das 12 RNA na ferramenta de MD e seus respectivos parâmetros de configuração.

Os valores adotados para os parâmetros listados nas Tabelas 7 e 8 foram os seguintes:

- Ciclos. Assim como no trabalho do Mueller (1996), foram previstos para todas as RNA um aprendizado de 20.000 ciclos.
- *Learning rate*, taxa de aprendizado que regula a velocidade em que os dados são ajustados. Essa taxa foi escolhida automaticamente pela ferramenta adotada.
- *Momentum*, taxa de aceleração do aprendizagem em função da variação de erro. Foi escolhida automaticamente pela ferramenta.
- *Minimum error* representa o erro mínimo apresentado pela RNA.
- *Maximum error* representa o erro máximo apresentado pela RNA.

- *Average error* representa o erro médio apresentado pela RNA.
- *Target error* representa o erro mínimo configurado para o aprendizado da RNA. Essa taxa é um dos fatores que definem a parada do aprendizado quando o mesmo é atingido. Durante o período de aprendizado da RNA, ocorreram algumas paradas por erro antes da finalização pelos 20.000 ciclos previstos, necessitando a reconfiguração desse parâmetro, até a configuração final constante das Tabelas 7 e 8.
- *Total de registros* representa o total de registros que foram inseridos para a criação da RNA.

Tabela 7 - Redes 1 a 6 com predição de 30 dias corridos

Características	Números do treinamento					
	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
Ciclos	20000	20000	20000	20000	20000	20000
Learning rate	1,0	1,0	1,0	0,6	1,0	0,8
Momentum	0,8	0,8	0,8	0,8	0,7	0,6
Minimum error	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001
Average error	0,0009	0,0106	0,0009	0,0106	0,0165	0,0124
Maximum error	0,0237	0,8105	0,0163	0,8370	0,7861	0,7662
Target error	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005
Total de registros	15820	15820	15820	15820	15820	15820
Registros utilizados treinamento	15674	15665	15678	15696	15696	15625

Tabela 8 - Redes 7 a 12 com predição de 15 dias corridos

Características	Números do treinamento					
	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
Ciclos	20000	20000	20000	20000	20000	20000
Learning rate	1,0	0,6	1,0	1,0	0,8	1,0
Momentum	0,8	0,8	0,8	0,8	0,8	0,6
Minimum error	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
Average error	0,0009	0,0137	0,0009	0,0140	0,0113	0,0144
Maximum error	0,0201	0,8472	0,0177	0,7715	0,7741	0,8045
Target error	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005
Total de registros	15820	15820	15820	15820	15820	15820
Registros utilizados treinamento	15735	15756	15736	15729	15729	15715

- *Registros utilizados no treinamento* representa a quantidade de registros que foram utilizados na ferramenta de MD para a geração da predição de valores. Foi especificada para a ferramenta a exclusão no treinamento dos registros que estavam apresentando uma alta taxa de erro no aprendizado. A identificação e exclusão destes registros pode ser realizada através da opção de visualização dos detalhes do aprendizado.

Outro fator a ser considerado em relação ao aprendizado das 12 RNA é a importância que cada campo de entrada teve no momento do aprendizado da RNA. A seguir são apresentadas as Tabelas 9 e 10 com estes valores.

Com relação à importância dos dados para as RNA para a predição de um mês (ver Tabela 9), observou-se o seguinte:

R₁: A variação entre o menor e o maior grau de participação foi de 48,8932, onde o maior grau de participação no aprendizado foi de 49,0412 para o campo de Preço Máximo de Cotação na data e o de menor participação foi o Dia Útil da Semana com 0,1480. Essa RNA foi construída com 11 campos de entrada.

R₂: A variação entre o menor e o maior grau de participação foi de 202,8779, campo de Identificação da Empresa (NOMRES) participou com 203,7627 e o de menos importância foi o Dia Útil da Semana com 0,8848. Essa RNA foi construída com 7 campos de entrada.

R₃: A variação entre o menor e o maior grau de participação foi de 50,7509, onde o maior grau de participação no aprendizado foi de 52,4357 para o campo de Preço Máximo de Cotação na data e o de menor participação foi o Dia Útil da Semana com 0,1437. Essa RNA foi construída com 9 campos de entrada.

R₄: A variação entre o menor e o maior grau de participação foi de 176,2002, onde o maior grau de participação no aprendizado foi de 176,4192 para o campo de Identificação da Empresa (NOMRES) e o de menor participação foi o Dia Útil da Semana com 0,2190. Essa RNA foi construída com 5 campos de entrada.

R₅: A variação entre o menor e o maior grau de participação foi de 780,2013, onde o maior grau de participação no aprendizado foi de 784,8896 para o campo de Identificação da Empresa (NOMRES) e o de

menor participação foi o Dia Útil da Semana com 4,6883. Essa RNA foi construída com 5 campos de entrada.

R₆: A variação entre o menor e o maior grau de participação foi de 693,2676, onde o maior grau de participação no aprendizado foi de 694,6056 para o campo de Identificação da Empresa (NOMRES) e o de menor participação foi o Dia Útil da Semana com 1,3380. Essa RNA foi construída com campos de entrada.

Tabela 9 - Redes de 1 a 6 e a importância de seus campos de entrada no aprendizado

Entradas do treinamento	REDES – Predição 30 dias					
	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
DIA-UTIL-SEMANA	0,1480	0,8848	0,1437	0,2190	4,6883	1,3380
DIA-UTIL-MÊS	0,5427	7,6678	1,5682	0,4265	7,2488	13,8500
DIA-UTIL-ANO	3,0330	102,4784	1,6848	4,9634	35,1004	167,0981
DATA-AA	7,4516	63,6688	2,7435	49,7723	346,9562	167,0981
DATA-MM	2,9360	88,8688				119,1765
DATA-DD	0,7262	4,2783				8,5576
NOMRES	2,9520	203,7627	0,7983	176,4192	784,8896	694,6056
PREABE	35,0167		39,4738			
PREMAX	49,0412		52,4357			
PREMIN	37,6812		41,5317			
PREMED	44,3291		39,3941			

Tabela 10 - Redes de 7 a 12 e a importância de seus campos de entrada no aprendizado.

Entradas do treinamento	REDES – Predição 15 dias					
	R ₇	R ₈	R ₉	R ₁₀	R ₁₁	R ₁₂
DIA-UTIL-SEMANA	0,2249	0,3841	0,2225	0,5770	4,2704	2,1042
DIA-UTIL-MÊS	0,1069	0,9821	1,0813	1,9418	50,7324	2,0972
DIA-UTIL-ANO	0,5485	8,1745	3,6612	40,3258	104,1739	45,7625
DATA-AA	4,3322	39,9833	6,6180	107,0983	413,0861	234,9099
DATA-MM			3,2505	31,5814	86,8125	
DATA-DD			1,0444	3,9319	46,6823	
NOMRES	0,2245	218,74000	3,0333	298,7131	1089,7238	799,5406
PREABE	37,9261		40,9396			
PREMAX	51,4195		52,8178			
PREMIN	37,6812		41,5317			
PREMED	44,3291		39,3941			

No caso das RNA para predição de uma quinzena (Tabela 10), as observações sobre a importância dos dados são:

R₇: A variação entre o menor e o maior grau de participação foi de 51,1946, onde o maior grau de participação no aprendizado foi de 51,4195 para o campo de Preço Máximo de Cotação na data e o de menor participação foi o Dia Útil da Semana com 0,2249. Essa RNA foi construída com 9 campos de entrada.

R₈: A variação entre o menor e o maior grau de participação foi de 218,3559, campo de Identificação da Empresa (NOMRES) participou com 218,7400 e o de menos importância foi o Dia Útil da Semana com 0,3841. Essa RNA foi construída com 5 campos de entrada.

R₉: A variação entre o menor e o maior grau de participação foi de 52,5953, onde o maior grau de participação no aprendizado foi de 52,8178 para o campo de Preço Máximo de Cotação na data e o de menor participação foi o Dia Útil da Semana com 0,2225. Essa RNA foi construída com campos de entrada.

R₁₀: A variação entre o menor e o maior grau de participação foi de 298,1360, onde o maior grau de participação no aprendizado foi de 298,713 para o campo de Identificação da Empresa (NOMRES) e o de menor participação foi o Dia Útil da Semana com 0,5770. Essa RNA foi construída com 5 campos de entrada.

R₁₁: A variação entre o menor e o maior grau de participação foi de 1085,4496, onde o maior grau de participação no aprendizado foi de 1.089,7200 para o campo de Identificação da Empresa (NOMRES) e o de menor participação foi o Dia Útil da Semana com 4,2704. Essa RNA foi construída com 7 campos de entrada.

R₁₂: A variação entre o menor e o maior grau de participação foi de 797,4358, onde o maior grau de participação no aprendizado foi de 799,540 para o campo de Identificação da Empresa (NOMRES) e o de menor participação foi o Dia Útil da Semana com 2,1042. Essa RNA foi construída com 5 campos de entrada.

Analisando o resultado das 12 redes foi observado que o campo de menor importância no treinamento em todas elas foi o campo de Dia Útil da Semana, o que condiz com a limitação introduzida na modelagem da série histórica (ver seção

4.3.5). Quando o campo Preço Máximo de Cotação participa do treinamento, ele é o mais importante para o aprendizado da rede. Onde não existem os campos de entrada de valores de cotação, o campo mais importante é o identificador da empresa da cotação.

4.5 FASE 5: AVALIAÇÃO

Nesta fase serão apresentadas tabelas que foram obtidas a partir do treinamento das RNA criadas e em seguida serão tecidos comentários a respeito desses dados.

4.5.1 Avaliação dos resultados

O Apêndice C apresenta 12 tabelas, por empresa, com as 12 RNA criadas na ferramenta de MD e que contém as previsões.

As Tabelas 11 a 16 representam as variações entre os valores das previsões e os valores efetivos das cotações de cada uma das empresas escolhidas neste trabalho, respectivamente: Telesp, Petrobrás, Embraer, Bradesco, Banco do Brasil e Vale do Rio Doce. O campo Menor % de variação indica que o valor real de cotação foi maior que o valor identificado pela rede. Já o campo Maior % de variação indica que o valor identificado pela rede foi maior que o valor real de negociação da ação da referida empresa. Os campos variações a menor e a menor identifica, dentro do período preditivo definido a quantidade, a quantidade de dias de cotações identificadas pela rede superior ou inferior aos valores efetivamente negociados.

Tabela 11 - Variação das predições de cotações da empresa Telesp

	Menor % variação	Maior % Variação	Variações a menor	Variações a maior
Rede 1	-1,04	+1,56	12	9
Rede 2	-11,45	-8,08	21	0
Rede 3	-0,68	+1,49	9	12
Rede 4	-5,92	-9,33	21	0
Rede 5 (*)	-17,88	-14,90	21	0
Rede 6	-24,82	-97,93	21	0
Rede 7	-0,72	+0,66	6	4
Rede 8	-16,3	-19,03	10	0
Rede 9	-0,13	+1,06	3	7
Rede 10	-9,12	-12,05	10	0
Rede 11	-13,49	-16,28	0	0
Rede 12(**)	-6,31	-15,69	10	0

(*) – mesmo valor de predição

(**) – o percentual de variação de erro de acerto foi crescente.

Tabela 12 - Variação das predições de cotações da empresa Petrobrás

	Menor % Variação	Maior % Variação	Variações a menor	Variações a maior
Rede 1	-0,79	+1,84	5	16
Rede 2	-7,0	7,15	12	9
Rede 3	-0,32	2,06	1	20
Rede 4	-11,85	-20,81	21	0
Rede 5	+2,72	+15,09	0	21
Rede 6	+10,60	+23,75	0	21
Rede 7	-0,63	+1,65	4	6
Rede 8	++7,86	+18,95	0	10
Rede 9	-0,90	+1,59	4	6
Rede 10	-1,22	+10,28	4	6
Rede 11	+2,48	+13,06	0	10
Rede 12(*)	+0,40	-9,74	9	1

(*) – mesmo valor de predição

Tabela 13 - Variação das predições de cotações da empresa Embraer

	Menor % Variação	Maior % Variação	Variações a menor	Variações a maior
Rede 1	-2,35	+4,58	13	8
Rede 2(**) (*)	0,0	+55,83	0	21
Rede 3	-2,35	+4,06	15	6
Rede 4(**)	+3,98	+18,76	0	21
Rede 5	-8,10	+49,58	7	14
Rede 6	+14,43	-26,47	15	6
Rede 7	+0,76	-2,29	7	3
Rede 8	+3,09	+54,35	0	10
Rede 9	-1,41	+1,59	4	6
Rede 10	+55,54	+70,86	0	10
Rede 11 (*)	+40,93	+49,20	0	10
Rede 12(*)	-5,87	-23,74	10	0

(*) – praticamente o mesmo valor de predição

(**) – valor de predição com variação percentual acerto decrescente

Tabela 14 - Variação das predições de cotações da empresa Bradesco

	Menor % Variação	Maior % Variação	Variações a menor	Variações a maior
Rede 1	-0,34	-1,74	21	0
Rede 2	-19,92	-32,12	21	0
Rede 3	+0,62	-0,72	5	16
Rede 4 (*)	-9,64	-20,10	21	0
Rede 5	-80,99	-90,40	21	0
Rede 6 (**)	-15,00	-24,13	21	0
Rede 7 (**)	+0,12	-1,33	9	1
Rede 8	-10,87	-14,79	10	0
Rede 9	-0,37	-1,59	10	0
Rede 10	-11,49	-14,82	10	0
Rede 11	-28,29	-31,68	10	0
Rede 12	-29,76	-35,38	10	0

(*) – praticamente o mesmo valor de predição

(**) – valor de predição com variação percentual acerto decrescente

Tabela 15 - Variação das predições de cotações da empresa Banco do Brasil

	Menor % Variação	Maior % Variação	Variações a menor	Variações a maior
Rede 1	-1,36	2,52	4	17
Rede 2	+26,12	+52,45	0	21
Rede 3	-0,75	+2,84	4	17
Rede 4 (*)	-1,49	+12,41	3	18
Rede 5	+10,64	+26,23	0	21
Rede 6 (*)	+4,95	-8,06	18	3
Rede 7	+0,22	-1,12	3	7
Rede 8 (*)	-2,98	+5,38	3	7
Rede 9	+0,09	-1,53	5	5
Rede 10	-5,17	-33,35	10	0
Rede 11 (**)	-4,31	-63,47	10	0
Rede 12	+68,95	+105,04	0	10

(*) – praticamente o mesmo valor de predição

(**) – predição de cotação destoante das demais apresentadas

Tabela 16 – Variação das predições de cotações da empresa Vale do Rio Doce

	Menor % Variação	Maior % Variação	Variações a menor	Variações a maior
Rede 1	+1,34	-1,86	9	12
Rede 2 (*)	+3,29	+14,30	0	21
Rede 3	+1,07	-1,47	7	14
Rede 4	+9,30	+35,05	0	21
Rede 5 (*)	+14,49	+27,30	0	21
Rede 6	+31,23	+57,54	0	21
Rede 7	+0,64	-1,74	4	6
Rede 8	+1,0	+9,68	0	10
Rede 9	-1,97	+0,54	4	6
Rede 10 (**)	+4,29	+13,17	0	10
Rede 11 (*)	+0,88	+10,15	0	10
Rede 12(*)	-1,15	+7,93	2	8

(*) – praticamente o mesmo valor de predição

(**) – valor de predição com variação percentual acerto decrescente

Considerações sobre as predições apresentadas

Para analisar qual a melhor predição da RNA para cada uma das empresas tratada temos que analisar dois tipos de abordagens. No caso de compra de ações, se a predição for superior ao valor real de compra, considera-se como ideal que o capital a ser investido será menor do que o valor previsto. Já no caso de venda de

ações, caso a predição for inferior ao valor real de mercado, indica que o valor de capital a ser recuperado nesse caso, será superior ao valor previsto.

Em todas as RNA existe uma variação entre valores abaixo e acima do valor real de mercado previsto para os dias trabalhados.

Analisando as Tabelas 11 a 16, verifica-se que as RNA R_1 , R_3 , R_7 e R_9 foram as que apresentaram o valor das cotações mais próximos se comparados os valores de predição e real de cotação. As demais RNA apresentaram uma variação muito grande ou para maior ou para menor.

Para a RNA R_7 , conforme consta da Tabela 14 com cotações do Bradesco, observou-se que o valor de predição para 30 dias apresentou praticamente o mesmo valor.

Analisando os números finais apresentados pela ferramenta de MD, foram exatamente as RNA R_1 , R_3 , R_7 e R_9 as que apresentaram a menor taxa de erro de aprendizagem, 0,0234 (R_1), 0,0163 (R_3), 0,0201 (R_9) e 0,0177 (R_{11}).

Confirmando o observado por Mueller (1996), constatou-se ainda que quanto menor o período de predição de valores, a tendência de acerto dos valores para um período menor de tempo é maior, inclusive a taxa de erro de aprendizagem tende a ser menor. Esse fato pode ser verificado nas Tabelas de 11 a 16, onde as 6 primeiras RNA trazem a predição para 30 dias corridos e as 6 últimas apresentam predição para 15 dias corridos.

Analisando, ainda, os campos de entrada e saída, conforme disposto na Tabela 6, verifica-se que as RNA que apresentaram melhor resultado foram aquelas que tiveram a maior quantidade de variáveis de entrada, como é o caso das RNA R_1 e R_9 , que tiveram 11 campos de entrada e 1 campo de saída, e as RNA R_3 e R_7 , que contaram com 9 campos de entrada e 1 campo de saída.

Após a análise dos resultados obtidos, foi elaborada uma nova experiência selecionando uma das empresas, no caso a Petrobrás e a introdução de 2 novos campos de saída: Percentual de Rendimento Diário e Percentual de Rendimento Mensal. O campo Percentual de Rendimento Diário levou em consideração o valor da cotação de fechamento do dia anterior e o valor da cotação de fechamento do dia atual. O campo Percentual de Rendimento Mensal levou em consideração o acumulado mensal do campo Percentual de Rendimento Diário.

Os campos que participaram da RNA estão presentes no Quadro 8.

Campos de entrada	Descrição
DIA-UTIL-MÊS	Dia útil do mês
DIA-UTIL-ANO	Dia útil do ano
DATA-AA	Ano da cotação
DATA-MM	Mês da cotação
DATA-DD	Dia da cotação
PREABE	Preço de abertura da cotação na data
PREMAX	Preço máximo de cotação na data
PREMIN	Preço mínimo de cotação na data
PREMED	Preço médio da cotação na data
Campo de saída	Descrição
PERC_DIA	Rendimento da cotação no dia
PERC_MES	Rendimento da cotação no mês

Quadro 8 – Campos que fizeram parte do treinamento

No modelo apresentado a RNA não conseguiu prever percentuais de rendimento com uma taxa de acerto aceitável, apresentando um valor de rendimento único como predição, totalmente desconexo com os valores reais apresentados pelas cotações.

4.5.2 Revisão do processo

É recomendável que, dentro da disponibilidade de tempo para a realização dos trabalhos, seja realizada uma revisão nas demais redes que apresentaram uma alta taxa de erro de aprendizagem e verificar o comportamento das mesmas com cotações de outras empresas que não constaram da carteira selecionadas.

Observou-se ainda que o aumento da quantidade de campos de saída de 1 para 5 não produziu efeitos positivos, apesar de existir RNA semelhante em termos de campos de entrada. Os campos inseridos não conseguiram melhorar a taxa de aprendizagem das respectivas RNA. Possivelmente, isto poderia ser melhorado com uma parametrização mais fina que, entretanto, não é possível na ferramenta adotada.

4.5.3 Determinação dos próximos passos

Considerando que os objetivos traçados de obtenção de um modelo que possa subsidiar o aprendizado em aplicações no mercado de ações foram alcançados com o presente trabalho, entretanto o mesmo necessita evoluir nos seguintes pontos:

- Aplicação do modelo em outras ferramentas de MD;
- Aplicação de cotações de outras ações que possam compor a carteira definida nesse trabalho;
- Inclusão de outros campos tais como cotação de moedas; taxa de inflação ou outros indicadores econômicos e verificar a influência dos mesmos no resultado obtidos pela RNA;
- Evolução RNA que não apresentaram resultado satisfatório, mudando a forma de trabalho.

4.6 FASE 6: IMPLANTAÇÃO

A implantação do modelo em situações reais de aprendizado pode ser melhor explorada considerando-se as observações dispostas nas seções 4.5.3.

5 CONCLUSÃO E TRABALHOS FUTUROS

O mercado acionário é movido por lucros financeiros dos investidores. Esse mercado é composto de um conjunto complexo de fatores que influenciam a evolução dos valores das cotações diárias. Em algumas situações, podem ser identificadas empresas que estão enfrentando problemas de mercado, influenciando o valor de suas cotações, podendo gerar perdas financeiras ao investidor, e consequentemente levar o mesmo a procurar novos tipos de aplicações ou a buscar ações de outras empresas, visando gerar ganhos financeiros.

A atuação do analista de mercado neste contexto necessita de um conhecimento que transcende aquele adquirido por meio do ensino formal. O aprendizado formal pode e deve ser complementado com o aprendizado cognitivo vivenciado pelo analista no seu dia-a-dia ou por meio de simulações sobre o tema (LÉVY, 2007) realizadas com o auxílio de ferramentas tecnológicas. Nesse sentido, o presente trabalho permitiu uma ampla exploração de conceitos sobre o mercado de ações e a expansão do conhecimento do especialista envolvido.

A busca por trabalhos relacionados ao tema evidenciou a existência de várias linhas de pesquisa relacionadas. Tais linhas incluem outros aspectos importantes que, em um esforço de extensão do modelo criado, podem levar ainda a um aprofundamento do aprendizado.

As linhas de pesquisa mencionadas propõem diversos métodos que podem ser adotados para a análise dos preços de cotação das ações, mas vale ressaltar que, muito mais importante do que a forma de análise da variação de preços, está um fator difícil de ser mensurado: o humor e a vontade do investidor em sair ou permanecer no mercado. Esse fator dificilmente poderá ser representado em qualquer modelo que for criado para previsão da cotação de ações no mercado. O que se pode afirmar é uma tendência no mercado mundial terá uma alta probabilidade de influenciar o mercado nacional, exceto se fatores internos neutralizarem essa força internacional.

Considerando a experiência adquirida neste estudo, sugerem-se como trabalhos futuros os seguintes tópicos:

- Realizar um experimento controlado visando mensurar o aprendizado com a aplicação do processo por *traders* e identificar a sua efetividade e possíveis melhorias;

- Ampliar o aprendizado por simulação pela adoção de outras técnicas ou a modificação da atual para incluir:
 - A análise de sensibilidade do modelo a partir da inserção de dados sobre: (i) moedas estrangeiras, (ii) indicadores econômicos, (iii) fatores políticos e (iv) fatores climáticos;
 - Utilização de valores de rendimentos de ações como variável de saída;
 - Utilização do modelo em outras carteiras de ações;
 - A predição baseada em janelas de tempo variáveis ao invés de um período fixo;
 - A modificação da série histórica para alinhar os dias da semana de modo a capturar a sazonalidade dos preços no período.
- Desenvolver modelos para outros tipos de aplicações, como: (i) comercialização de produtos através da Bolsa de Mercadorias e Futuros e (ii) compra e venda de moeda estrangeira, podendo trabalhar com uma carteira de moedas.

REFERÊNCIAS

- ALLEE, V. **The future of knowledge**: increasing prosperity through value networks. Burlington: Butterworth-Heinemann, 2003.
- AMARAL, F. C. N. **Data Mining**: técnicas e aplicações para o marketing direto. São Paulo: Berkeley, 2001.
- AMOROSO, E. D.; PRADO, H. A.; BRASIL, L. M.; ALVARENGA, R.; FERNEDA, E.; . Previsão de séries temporais utilizando redes neurais artificiais em cotação de ações da bolsa de valores. In: CONTECSI - CONGRESSO INTERNACIONAL DE GESTÃO DE TECNOLOGIA E SISTEMAS DE INFORMAÇÃO 5., São Paulo. **Anais...** São Paulo: FEA/USP, 2008.
- ANDERSEN, J. V.; SORNETTE, D. A mechanism for pockets of predictability in complex adaptive systems. **Europhys. Lett.**, v. 70, n. 5, p. 697-703, 2005.
- BARROS, M. Introdução às séries temporais e ao modelo ARIMA. In: _____. **Processos estocásticos**. cap. 5. Disponível em: <<http://www.mbarros.com/id1.html>>. Acesso em: 29 ago. 2008.
- BOHM, V.; WENZELBURGER, J. On the performance of efficient portfolios. **J. Econ. Dyn. Control.**, v. 29, n. 4, p. 721-740, 2005.
- BOLSA DE VALORES DE SÃO PAULO. Disponível em: <<http://www.bovespa.com.br>>. Acesso em: 04 abr. 2008.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais**: teoria e aplicações. Rio de Janeiro: LTC, 2000.
- CARVALHO, A. P. L. F. **Redes neurais artificiais** – Introdução - 2004. Disponível em: <<http://www.icmc.usp.br/~andre/body.htm>>. Acesso em: 03 fev. 2008.
- CAVALCANTE, F.; MISUMI, J. Y.; RUDGE, L. F. **Mercado de capitais**: o que é e como funciona. Rio de Janeiro: Elsevier, 2005.
- CAVALCANTI, M.; GOMES, E. A nova riqueza das organizações: os capitais do conhecimento. **Revista TN Petróleo**, v. 3, n. 16, p. 1-10, 2000.
- CHIARELLA, C.; HE, X.; HOMMES, C. H. **A dynamic analysis of moving average rules**. Tinbergen Institute. Discussion Paper No. TI 05-057/1. May, 2005.
- CRISP-DM. **CRISP-DM 1.0**:- Step-by-step data mining guide. Disponível em <<http://www.crisp-dm.org>>. Acesso em: 11 nov. 2008.
- CROSS, R.; PARKER, A. **The hidden of social networks**: understanding how work really gets done in organizations. Boston, Massachusetts: Harvard Business School Press, 2004.
- DRUCKER, P. **Post-capitalist society**. Oxford: Butterworth-Heinemann, 1993.

EASY-NNA PLUS. **Version 10.0d**: Neural Planner Software, Jul. 2008. Disponível em: <<http://www.easynn.com/>>. Acesso em: 12 set. 2008.

EASY-NNB PLUS. **EasyNN-plus help**: the user interface manual. Neural Planner Software, Jul. 2008. Disponível em: <<http://www.easynn.com/>>. Acesso em: 14 fev. 2009.

FANG, Y.; XU, D. The predictability of asset returns: an approach combining technical analysis and time series forecasts. **Int. J. Forecast.**, v. 19, n. 3, p. 369-385, 2003.

FAYYAD, U. M.; PIATETSKY-SHAPIO, G.; SMYTH, P. **From data mining to knowledge discovery**: an overview. Menlo Park, CA: MIT Press, 1996.

FERNEDA, E. Redes neurais e sua aplicação em sistemas de recuperação de informação. **Ciênc. Inf.**, v. 35, n. 1, p. 25-30, jan./abr. 2006.

FU, T.-C.; CHUNG, F.-L.; LUK, R.; NG, C.-M. Stock time series pattern matching: template-based vs. rule-based approaches. **Eng. Appl. Artif. Intell.**, v. 20, n. 3, p. 347-364, 2007.

GESTEL, T. Van; Espinoza, M.; Baesens, B.; Suykens, J. A. K.; Brasseur, C.; De Moor, B. A Bayesian nonlinear support vector machine error correction model. **J. Forecast.**, v. 25, n. 2, p. 77-100, 2006.

GRADOJEVIC, N. Non-linear, hybrid exchange rate modeling and trading profitability in the foreign exchange market. **J. Econ. Dyn. Control.**, v. 31, n. 2, p. 557-574, 2007.

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

INMON, W. H.; TERDEMAN, R. H.; IMHOFF, C. **Data Warehousing**: como transformar informações em oportunidades de negócios. São Paulo: Berkeley, 2001.

KOHAVI, R. **Wrappers for performance enhancement and oblivious decision graphs**. 1995. 302 f. Dissertation (Doctor of Philosophy) – Department Computer Science, Stanford University, 1995.

KROGH, Georg Von. Solicitude nas Organizações. In: _____. **Facilitando a criação de conhecimento**: reinventando a empresa com o poder da inovação contínua. Rio de Janeiro: Campus, 2001. cap. 3.

LANGENOHL, A; SCHMIDT-BECK, K. **The First Stock**: How Financial Professional View their Careers. Burlington: Butterworth-Heinemann, 2003.

LASSALVIA, F. A. **O novo mercado de capitais da Bovespa**. 2005. Monografia (MBA Executivo em Finanças) – University Giessen, Germany, 2004.

LEIGH, W.; MODANI, N.; Hightower, R. A computational implementation of stock charting: abrupt volume increase as signal for movement in New York Stock Exchange composite index. **Decis. Support Syst.**, v. 37, n. 4, 2004.

LÉVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática. Rio de Janeiro: Editora 34, 2008.

MARSHALL, B. R.; YOUNG, M. R.; ROSE, L. C. Candlestick technical trading strategies: Can they create value for investors? **J. Bank. Financ.**, v. 30, n. 8, p. 2303-2323, 2006.

MATSURA, E. **Comprar ou Vender? Como investir na bolsa utilizando análise gráfica**. 6.ed. São Paulo: Saraiva, 2007.

McCULLOCH, W. S.; PITTS, W. H. A logical calculus of the ideas immanent in nervous activity. **Bull. Math. Biophys.**, n. 5, p.115-133, 1943.

MINSKY, M. L.; PAPPERT, S. **Perceptron**: an introduction to computational geometry. Cambridge: MIT Press, 1969.

MORESI, E. A. D. **Metodologia da pesquisa**. Brasília: Universidade Católica de Brasília, 2004. p. 125. Obra não-publicada.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: Edgard Blucher, 2006.

MUELLER, A. **Uma aplicação de redes neurais artificiais na previsão do mercado acionário**. 1996. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Florianópolis, 1996.

NONAKA, I.; TAKEUCHI, H. **Criação de conhecimento na empresa**. Rio de Janeiro: Campus, 1997.

OLIVEIRA JR, Waldomiro. **Estudo comparativo entre modelos lineares e redes neurais artificiais como tecnologias geradoras de expectativas e previsões de valores financeiros**. 2007. Dissertação (Mestrado em Gestão do Conhecimento e da Tecnologia da Informação) – Universidade Católica de Brasília, Brasília, 2007.

PARADIGMA. A gestão do conhecimento na prática. Disponível em: <<http://www.paradigma.com.br/gestao-do-conhecimento-na-pratica/view>>. Acesso em: 18 ago. 2008. (Também publicado na **Revista HSM Management**, n. 42, 2004).

PAULOS, J. A. **A lógica do mercado de ações**: uma análise prática do mercado de ações. Tradução de Afonso Celso da Cunha Serra. Rio de Janeiro: Elsevier, 2004.

PIAZZA, M. C. P. **Bem vindo à bolsa de valores**. 7. ed. São Paulo: Novo Conceito, 2008.

PINHEIRO, J. L. **Mercado de capitais**: fundamentos e técnicas. 2. ed. São Paulo: Atlas, 2002.

POLANYI, M. **The Tacit Dimension**. Gloucester: Peter Smith, 1983, reimpressão da edição de Doubleday & Company, Inc. 1966.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and retrieval in the brain. **Psychol. Rev.**, v. 65, p. 386-408, 1958.

RUSSELL, S. J. **Inteligência artificial**. Tradução da segunda edição. Rio de Janeiro: Elsevier, 2004.

THE UNIVERSITY OF WAIKATO. **Weka 3**: Data Mining Software in Java. 1999. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 12 set. 2008.

TONINI, A. M.; SCHETINO, D. N. **MATLAB para Engenharia**. Belo Horizonte: Centro Universitário de Belo Horizonte, 2002. Disponível em: <academcos.cefetmg.br/admin/downloads/2101/ApostilaMatLab_UNI.pdf>. Acesso em: 12 set. 2008.

WANG, J.; CHAN, S. Stock market trading rule discovery using pattern recognition and technical analysis. **Expert Syst. Appl.**, v. 33, n. 2, p. 304-315, 2007.

WESTERHOFF, F. Speculative markets and the effectiveness of price limits. **J. Econ. Dyn. Control.**, v. 28, n. 3, p. 493-508, 2003.

APÊNDICE A – Metodologia CRISP-DM

A metodologia CRISP-DM é constituída de seis fases: (1) Compreensão do Negócio; (2) Compreensão dos Dados; (3) Preparação dos Dados; (4) Modelagem; (5) Avaliação; e (6) Aplicação.

Neste apêndice, cada uma destas fases será explicada detalhadamente para uma melhor compreensão da maneira como o trabalho de DCBD foi conduzido, visto que esta foi a metodologia de DCBD utilizada para a criação do modelo para trabalhar as cotações de ações da Bolsa de Valores de São Paulo.

Fase 1: Compreensão do Negócio

Nesta fase é que são determinados os objetivos, as necessidades e exigências do cliente a partir da perspectiva do negócio tratado. Nesta fase é muito importante identificar fatores que possam influenciar os resultados de KDD. As atividades desta fase estão representadas na figura 11.

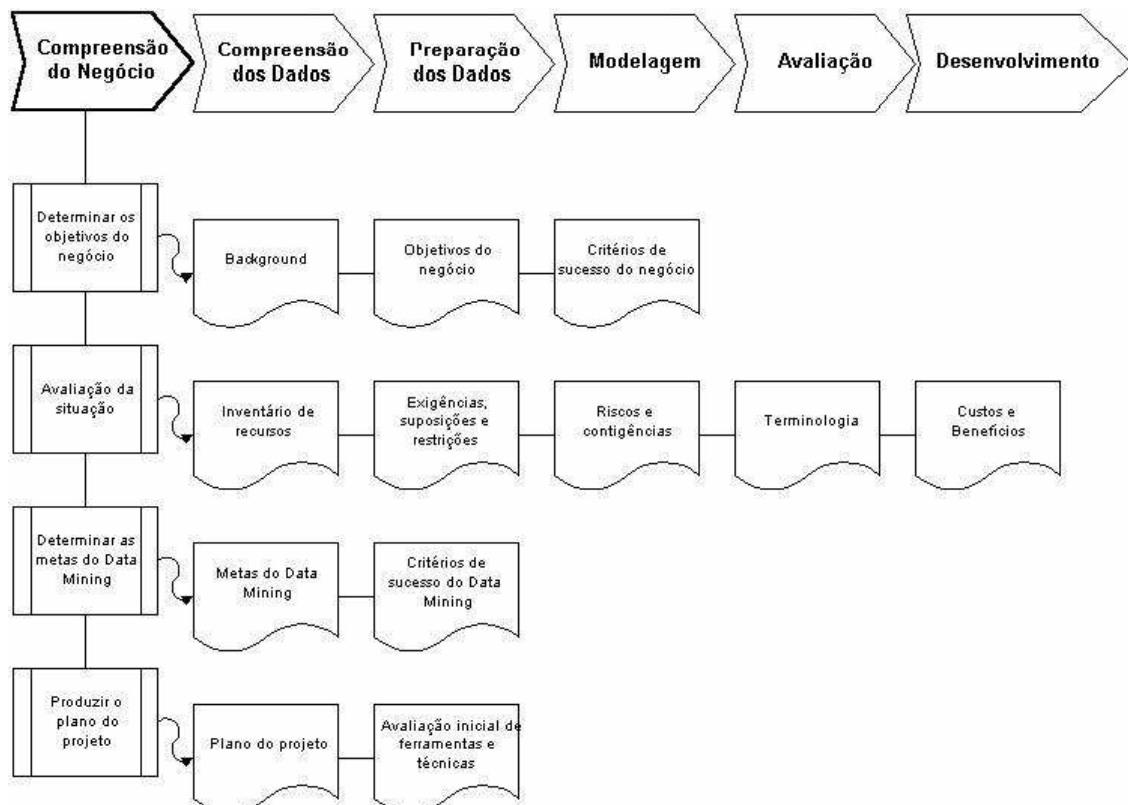


Figura 11 - Compreensão do Negócio (CRISP-DM [CRISP, 2000])

a) Determinação dos Objetivos do Negócio

O primeiro objetivo do analista de dados é entender completamente, de uma perspectiva empresarial, o que o cliente realmente quer realizar. A meta do analista é descobrir fatores importantes, no princípio, que podem influenciar o resultado do projeto. Uma possível consequência de negligenciar este passo é gastar um grande esforço produzindo respostas certas para as perguntas erradas. A informação é registrada e é conhecida a sua situação no negócio.

Além do objetivo empresarial primário, há tipicamente outras perguntas empresariais relacionadas que o cliente gostaria de conhecer. Por exemplo, a meta empresarial primária po-

deria ser a manutenção de clientes, predizendo quando eles estão propensos a irem para um concorrente.

O critério para um resultado de sucesso é descrito para o projeto do ponto de vista empresarial. Isto pode ser bastante específico e possível de ser medido objetivamente, como redução de clientes em certo nível ou no geral e subjetivamente como "procurar estreitar os laços no relacionamento com os clientes". No caso anterior deveria ser indicada uma pessoa.

b) Avaliação da Situação

Esta tarefa envolve encontrar dados mais detalhados sobre os recursos, limitações, suposições e outros fatores que deveriam ser considerados quando determinando o objetivo da análise de dados e o plano de projeto. As tarefas anteriores devem ser rapidamente obtidas para que se chegue ao objetivo principal.

Os recursos disponíveis para o projeto são listados, e incluem: pessoal (negócio peritos, peritos de dados, apoio técnico, pessoal de MD), dados, recursos computacionais (plataformas de hardware) e software (ferramentas de MD ou outro software pertinente).

As exigências do projeto, como o tempo de conclusão, são relacionadas, inclusive as suposições feitas pelo projeto. Estas podem ser suposições sobre os dados que podem ser conferidas durante a mineração dos dados, mas também pode incluir suposições não confirmadas do negócio no qual o projeto está inserido. É particularmente importante listar as limitações que possam influir na validade dos resultados. Estas podem ser limitações na disponibilidade de recursos, mas também podem incluir limitações tecnológicas como o tamanho dos dados para torná-los fáceis de modelar.

Os riscos ou eventos que poderiam acontecer, são dimensionados, analisando fatos que possam causar demora ou falhas no projeto. São estabelecidos planos de contingência correspondentes para os riscos dimensionados.

Construir uma análise de custo-benefício para o projeto, que compare os custos do projeto com o benefício potencial para o negócio se tudo ocorrer como o esperado. A comparação deverá ser tão específica quanto possível.

c) Determinação das Metas de Mineração de Dados

Uma meta empresarial declara os objetivos na terminologia empresarial. Uma meta de MD declara objetivos do projeto em termos técnicos. Por exemplo, a meta empresarial poderia ser "catalogar o aumento das vendas a clientes existentes". Uma meta da MD poderia ser "predizer quanto um cliente irá comprar, dado as suas compras durante os últimos três anos, informação demográfica (idade, salário, cidade, etc.) e o preço do produto."

Os resultados desejados do projeto são descritos e possibilitam a realização dos objetivos do negócio.

d) Produção do Plano de Projeto

Um plano de projeto é produzido e visa alcançar os objetivos de MD e assim os objetivos empresariais. O plano deverá conter um conjunto antecipado de passos a serem executados durante o resto do projeto, inclusive uma seleção inicial de ferramentas e técnicas.

As fases a serem executadas no projeto são listadas, junto com a duração, recursos requeridos, entradas, resultados e dependências. Declarar explicitamente as amplas repetições nos processos de MD, como repetições da modelagem e fases de avaliação. Como parte do plano de projeto, é também importante analisar dependências entre tempo e riscos.

O plano de projeto deve ser considerado como um documento dinâmico que deve permitir ao final de cada fase uma revisão de progresso e uma atualização do plano de projeto.

Ao término da primeira fase, o projeto executa também uma avaliação inicial de ferramentas e técnicas. Aqui, ocorre a seleção de uma ferramenta de MD que suporte vários métodos

para estágios diferentes do processo, por exemplo. É importante avaliar ferramentas e técnicas cedo no processo, dado que a seleção de ferramentas e técnicas possivelmente influencia o projeto inteiro.

Fase 2: Compreensão dos Dados

A obtenção inicial dos dados inclui a carga bem como a sua compreensão. Envolve o conhecimento da definição de cada um dos atributos, seu formato, sua amplitude e seus principais conceitos. As atividades que compõem esta fase estão representadas na figura 12.

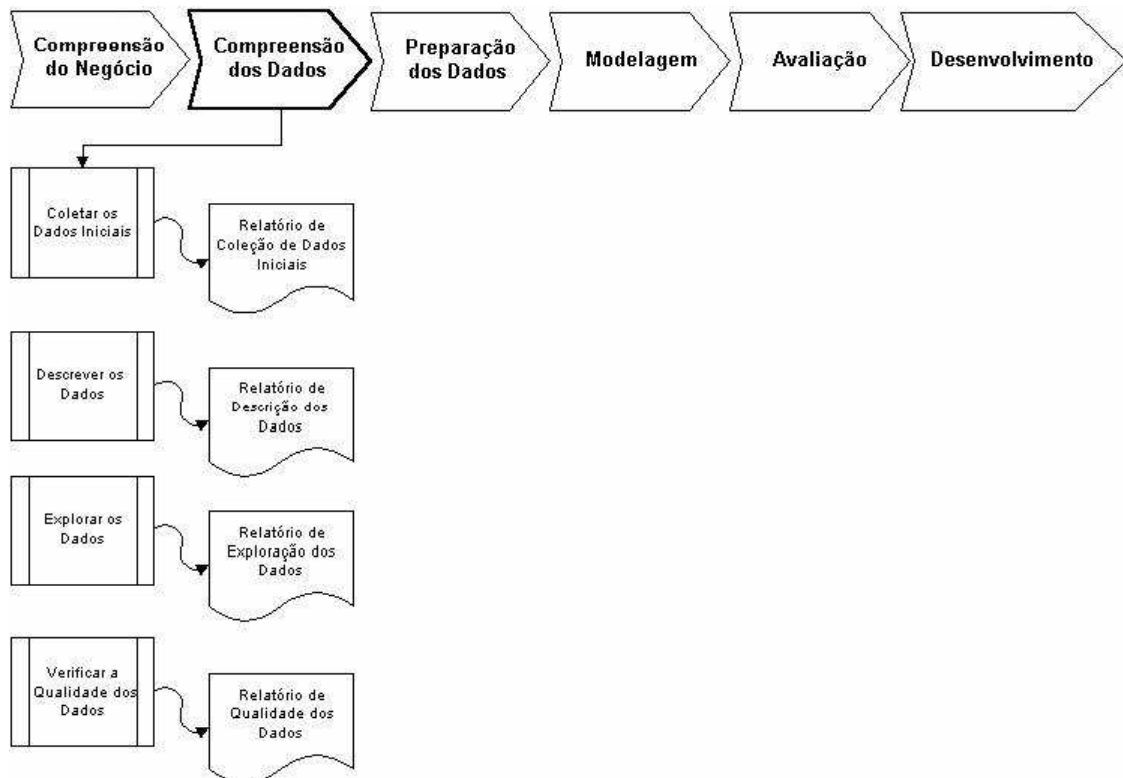


Figura 12 - Compreensão dos Dados (CRISP-DM [CRISP, 2000])

a) Coleta dos Dados Iniciais

O conjunto de dados é adquirido, juntamente com as suas localizações no projeto, os métodos usados para adquiri-los e qualquer problema encontrado. Problemas de registros encontrados e qualquer solução alcançada ajudam com em futuras reproduções deste projeto ou com a execução de futuros projetos similares.

b) Descrição dos Dados

Examinar as propriedades "bruta" e "superficial" dos dados adquiridos e elaborar um relatório dos resultados.

Descrever os dados que foram adquiridos, incluindo: o formato dos dados, a quantidade de dados, por exemplo, número de registros e campos, em cada tabela, as identidades dos campos e qualquer outra característica superficial dos dados que foram descobertos. Os dados adquiridos satisfazem as exigências pertinentes (relevantes)?

c) Exploração dos Dados

A exploração dos dados deve tomar como base as perguntas que se busca responder com a MD, visualizando, examinando e relatando os resultados. Estas incluem: distribuição de atributos chaves, por exemplo, o atributo designado de uma tarefa de predição; relações

entre pares ou números pequenos de atributos; resultados de agregações simples; propriedades de sub-populações significativa; simples análises estatísticas. Estas análises podem dirigir diretamente os objetivos de MD; elas também podem contribuir para refinar a descrição ou qualidade dos dados, introduzindo a transformação ou preparando outros dados necessários para análises adicionais.

A descrição dos resultados desta tarefa inclui os primeiros achados ou hipóteses iniciais e o impacto delas no restante do projeto. Se apropriado, poderão ser incluídos gráficos e diagramas que indiquem características dos dados ou levem a subconjuntos de dados interessantes para um exame adicional.

d) Verificação da Qualidade dos Dados

Poderá ser examinada a qualidade dos dados, dirigindo perguntas como: o dado é completo (cobre todos os casos desejados)? Ele é correto ou contém erros e se há erros como comumente eles são? Há perdas de valores nos dados? Nesse caso como são eles representados, onde eles acontecem e como comumente eles são?

Listar os resultados da verificação de qualidade dos dados; se existem problemas de qualidade, possíveis soluções. Soluções para problemas de qualidade de dados geralmente dependem fortemente dos dados e conhecimento do negócio.

Fase 3: Preparação dos Dados

Nesta fase encontra-se o ponto mais importante de todo o processo. É onde será gerado o produto de entrada para a realização do trabalho de MD. É onde será criado e validado o modelo proposto. As atividades que compõem esta fase estão representadas na figura 13.

a) Seleção dos Dados

Esta tarefa envolve a decisão dos dados a serem usados para a análise. Os critérios incluem a relevância dos objetivos da MD, qualidade e confiança técnicas tais como limites no volume dos dados ou nos tipos dos dados. Note que a seleção dos dados cobre a seleção dos atributos (colunas) como também a seleção dos registros (fileiras) em uma tabela.

Esta tarefa provê a fundamentação utilizada para se definir os dados a serem incluídos ou excluídos e as razões para tais decisões.

b) Limpeza dos Dados

Aumente a qualidade dos dados para o nível requerido pelas técnicas da análise selecionadas. Isto pode envolver a seleção de subconjuntos limpos dos dados, a inserção de opções apropriadas ou de técnicas de mais âmbito tais como a estimativa de dados faltantes.

O relatório da limpeza de dados descreve que decisões e ações foram tomadas para se dirigir os problemas da qualidade dos dados relatados durante a tarefa de Verificação da Qualidade dos Dados na fase Entendendo dos dados.

c) Construção dos Dados

Esta tarefa inclui a preparação de dados construídos, como a produção de atributos derivados, de registros inteiramente novos ou valores transformados para atributos existentes.

Atributos derivados são os atributos novos que são construídos de um ou mais atributos existentes no mesmo registro. Descrevem a criação de registros completamente novos.

d) Integração dos Dados

Estes são métodos onde a informação é combinada de várias tabelas ou registros para criar novos valores ou registros.

Tabelas Combinadas consistem em juntar duas ou mais tabelas que têm informações diferentes a respeito dos mesmos objetos.

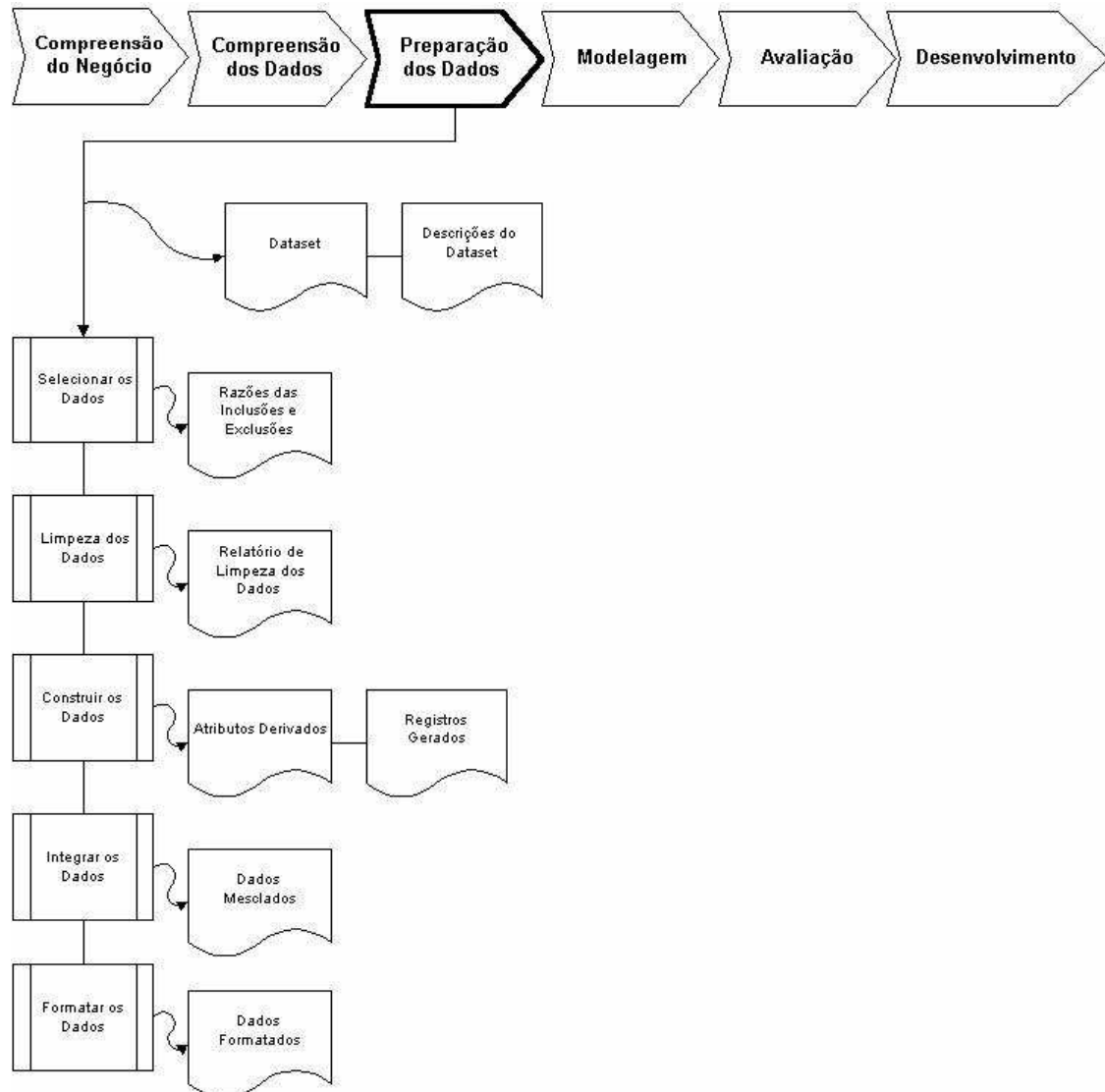


Figura 13 - Preparação dos Dados (CRISP-DM [CRISP, 2000])

e) Formatação dos Dados

Transformações de formatos se referem primeiramente a transformações sintáticas feitas nos dados que não mudam seu significado, mas podem ser requisitadas pela Ferramenta de Modelagem. Algumas ferramentas têm requisitos na ordem dos atributos tais como o primeiro campo ser um único identificador para cada registro ou o último campo ser a classe a ser prevista.

Fase 4: Modelagem

Esta fase trata basicamente da seleção e aplicação das técnicas de MD mais adequadas à situação em estudo, levando em consideração os objetivos pretendidos. No caso de seleção de mais de uma técnica, as mesmas devem ser trabalhadas em separado. As atividades que compõem esta fase estão representadas na figura 14.

a) Seleção da Técnica de Modelagem

A primeira etapa da modelagem é a seleção de uma técnica de modelagem que deve ser usada. Se uma ferramenta de modelagem já foi selecionada, esta tarefa consistirá especificamente na definição da técnica de modelagem e a sua documentação, conforme figura 15.

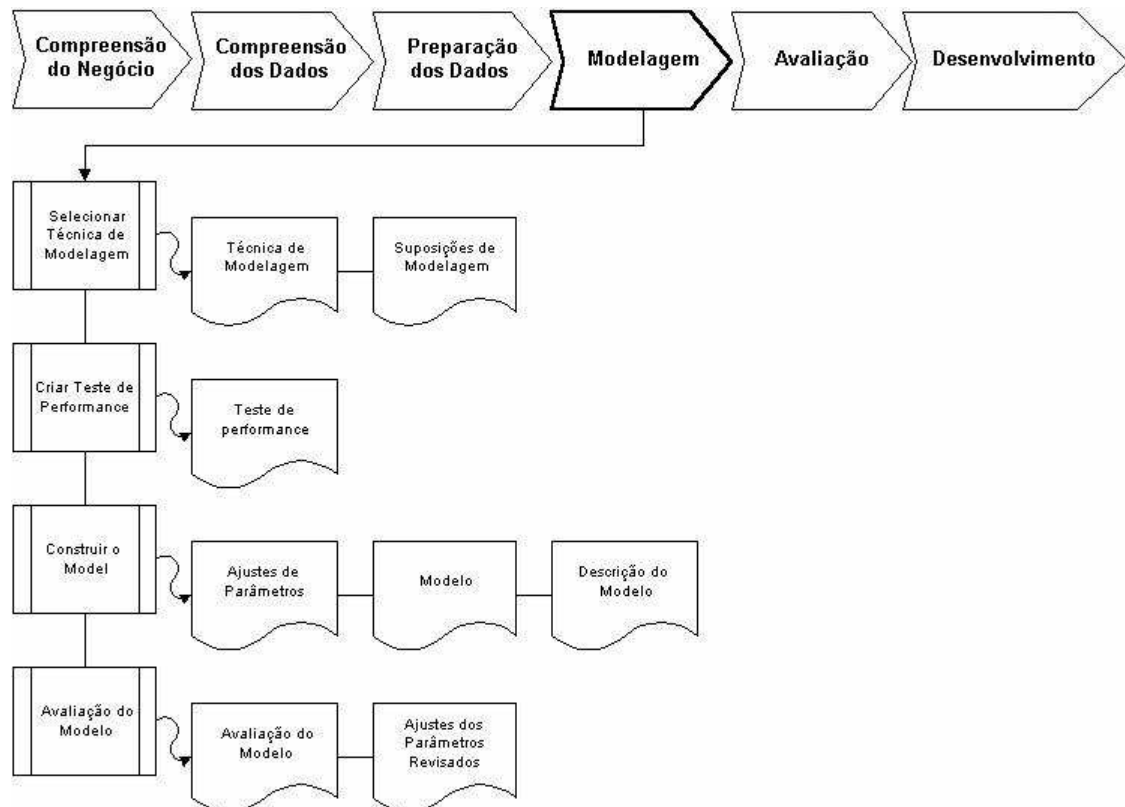


Figura 14 - Modelagem (CRISP-DM [CRISP, 2000])

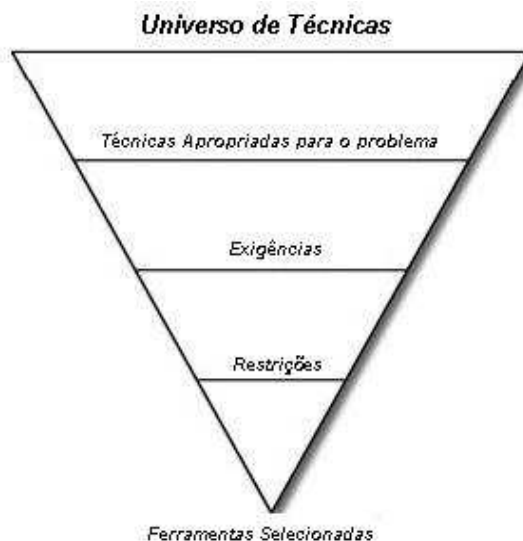


Figura 15 - Universo de Técnicas (CRISP-DM [CRISP, 2000])

b) Geração do Teste de Performance

Antes de realmente construir um modelo, será necessário gerar procedimentos ou mecanismos para testar a qualidade e validade do modelo.

Descrever o plano pretendido para treinamento, teste e validação do modelo. Um componente primário do plano é decidir como dividir o conjunto de dados disponível para o treino dos dados, teste dos dados e validação dos conjuntos de dados.

c) Construção do Modelo

Consiste em executar a ferramenta de modelagem no conjunto de dados preparados para

criar um ou mais modelos.

Com qualquer ferramenta de modelagem, geralmente tem-se um largo número de parâmetros que podem ser ajustados. Estes parâmetros devem ter seus valores escolhidos, juntamente com a razão para a escolha dos ajustes de parâmetros.

Descrever o modelo resultante, relatar a interpretação do modelo e documentar qualquer dificuldade encontrada juntamente com seu significado.

d) Avaliação do Modelo

O engenheiro da MD poderá interpretar o modelo de acordo com o domínio do seu conhecimento, o critério de sucesso da MD e o teste de *design* desejado. Esta tarefa interfere na fase de validação subsequente, onde o engenheiro da MD deve julgar o sucesso da modelagem e, juntamente com o analista de negócios e especialista no domínio, com quem deverá discutir os resultados da MD no contexto do negócio. Esta tarefa somente considera modelos onde a fase de validação também leva em conta todos os outros resultados que foram produzidos no decorrer do processo. O engenheiro da MD tenta classificar os modelos. Os modelos são acessados de acordo com os critérios de validação.

Resumir os resultados desta tarefa, listar qualidades dos modelos gerados e classificar suas qualidades em relação de uma com a outra.

De acordo com o acesso do modelo, revisar os parâmetros e ajustá-los para a próxima execução na tarefa de construção do modelo. Documentar todas as revisões e avaliações.

Fase 5: Avaliação

Esta fase consiste basicamente em verificar se os modelos escolhidos atingiram os objetivos selecionados, revisando os passos seguidos no modelo. As tarefas pertinentes a esta fase do processo estão representadas na Figura 16.

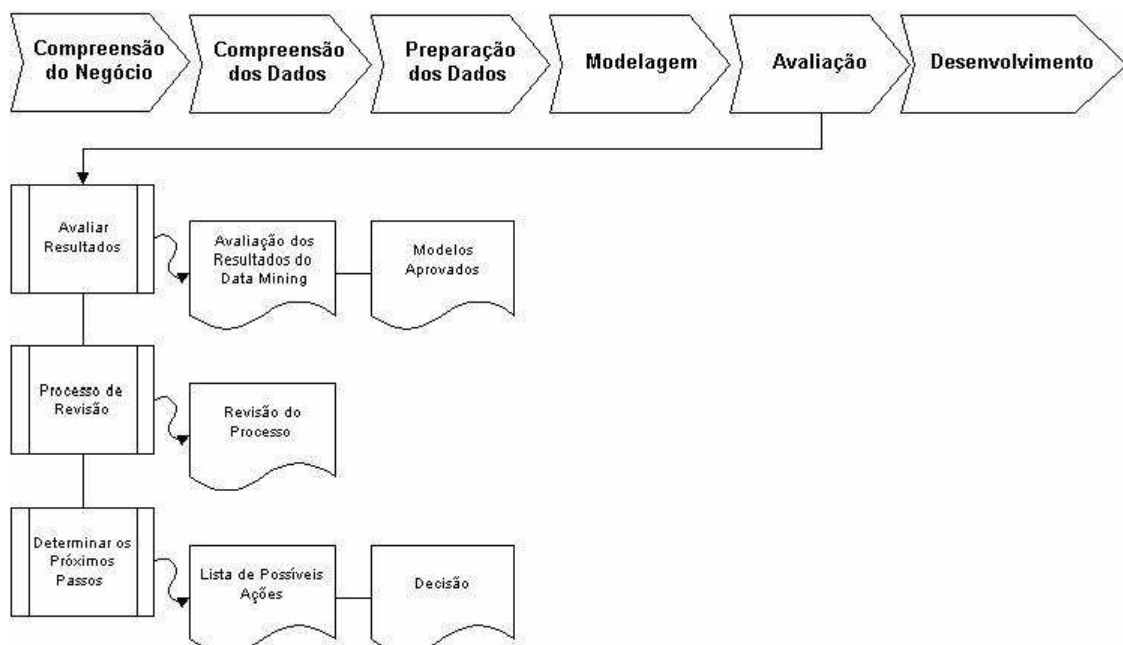


Figura 16 - Avaliação (CRISP-DM [CRISP, 2000])

a) Avaliação dos Resultados

Validações anteriores lidaram com fatores tais como precisão da aplicação de um modelo. Esse passo avalia o grau no qual o modelo vai ao encontro dos objetivos do negócio e tenta determinar se há alguma deficiência no modelo. A validação também avalia outros resulta-

dos gerados na MD. Os resultados dessa etapa cobrem modelos que são necessariamente relacionados com os objetivos de negócios originais e todos os achados que não são necessariamente relacionados com o objetivo, mas podem descobrir desafios adicionais, informação ou dicas para direções futuras.

Avaliação dos resultados da MD a respeito do sucesso dos critérios de negócio.

Resumir a avaliação dos resultados em termos dos critérios do negócio incluindo uma declaração final se o projeto já alcançou os objetivos iniciais do negócio.

Após a avaliação do modelo, se o modelo alcançar sucesso nos critérios do negócio, ele se tornará um modelo aprovado.

b) Revisão do Processo

Neste ponto, o modelo aparenta ser satisfatório. Agora é recomendada a realização de uma revisão do processo de MD para ver se nenhuma tarefa ou fator importante ficou de fora.

Resumir o processo de revisão e destacar as atividades que foram perdidas ou que deverão ser repetidas.

c) Determinação dos Próximos Passos

De acordo com os resultados da avaliação e o Processo de Revisão, o analista poderá decidir como proceder neste estágio. A decisão poderá ser a de finalizar ou preparar um novo projeto de MD.

Declarar os próximos passos potenciais juntamente com as razões pró e contra cada opção.

Fase 6: Aplicação

Nesta fase é realizada a comparação dos resultados obtidos com os objetivos traçados no início do projeto. As atividades pertinentes a esta fase do processo estão representadas na figura 17.

a) Planejamento da Aplicação

Planejar o desenvolvimento e procurar utilizar os resultados da MD no negócio, tendo em conta as avaliações dos resultados, gerando uma estratégia de desenvolvimento.

Resumir a estratégia de desenvolvimento incluindo passos necessários e como implementá-los.

b) Planejamento do Monitoramento e Manutenção

A elaboração de um plano de monitoramento e manutenção visa evitar que resultados desatualizados continuem a ser utilizados. Dependendo do estudo e dos dados, poderá ser necessária uma atualização periódica dos modelos.

Resumir a estratégia de monitoramento e manutenção incluindo passos necessários a serem executados e como implementá-los.

c) Produção do Relatório Final

No final do projeto, o líder do projeto e sua equipe redigem um relatório. Dependendo do método de Desenvolvimento, este relatório pode ser somente um resumo ou pode ser a apresentação final dos resultados da MD. O relatório escrito poderá incluir todos os resumos anteriores e organizar os resultados.

Frequentemente haverá reuniões na conclusão do projeto onde os resultados são verbalmente apresentados ao cliente.

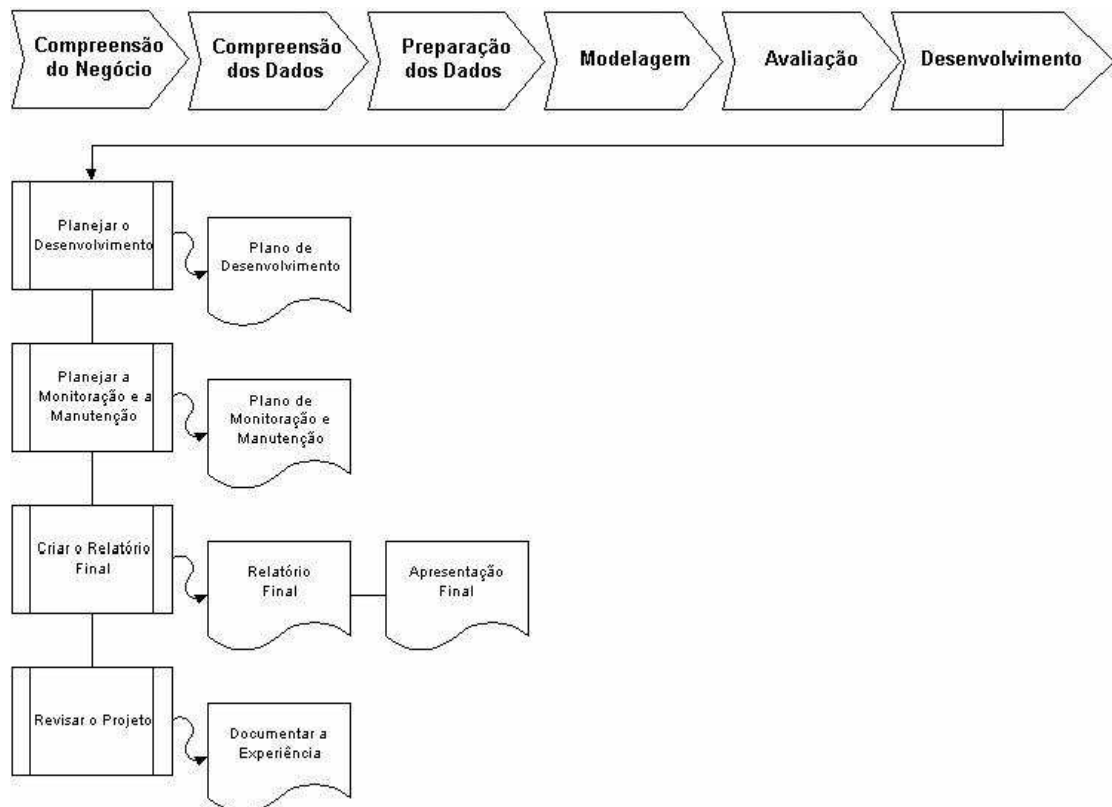


Figura 17 - Desenvolvimento (CRISP-DM [CRISP, 2000])

d) Revisão do Projeto

Avaliar o que deu certo, o que correu errado e o que poderá ser melhorado.

APÊNDICE B – Tabelas dos arquivos de cotações

1. TABELA DE CODBDI - relação dos valores para códigos de BDI

- 02 LOTE PADRÃO
- 06 CONCORDATÁRIAS
- 10 DIREITOS E RECIBOS
- 12 FUNDOS IMOBILIÁRIOS
- 14 CERTIFIC. INVESTIMENTO / DEBÊNTURES / TÍTULOS DIVIDA PÚBLICA
- 18 OBRIGAÇÕES
- 22 BÔNUS (PRIVADOS)
- 26 APÓLICES / BÔNUS / TÍTULOS PÚBLICOS
- 32 EXERCÍCIO DE OPÇÕES DE COMPRA DE ÍNDICE
- 33 EXERCÍCIO DE OPÇÕES DE VENDA DE ÍNDICE
- 38 EXERCÍCIO DE OPÇÕES DE COMPRA
- 42 EXERCÍCIO DE OPÇÕES DE VENDA
- 46 LEILÃO DE TÍTULOS NÃO COTADOS
- 48 LEILÃO DE PRIVATIZAÇÃO
- 50 LEILÃO
- 51 LEILÃO FINOR
- 52 LEILÃO FINAM
- 53 LEILÃO Fiset
- 54 LEILÃO DE AÇÕES EM MORA
- 56 VENDAS POR ALVARÁ JUDICIAL
- 58 OUTROS
- 60 PERMUTA POR AÇÕES
- 61 META
- 62 TERMO
- 66 DEBÊNTURES COM DATA DE VENCIMENTO ATÉ 3 ANOS
- 68 DEBÊNTURES COM DATA DE VENCIMENTO MAIOR QUE 3 ANOS
- 70 FUTURO COM MOVIMENTAÇÃO CONTÍNUA
- 71 FUTURO COM RETENÇÃO DE GANHO
- 74 OPÇÕES DE COMPRA DE ÍNDICES
- 75 OPÇÕES DE VENDA DE ÍNDICES
- 78 OPÇÕES DE COMPRA
- 82 OPÇÕES DE VENDA
- 83 DEBÊNTURES E NOTAS PROMISSÓRIAS
- 96 FRACIONÁRIO
- 99 TOTAL GERAL

2. TABELA DE ESPECI - relação de valores para as posições 5, 6, 7

- ON AÇÕES ORDINÁRIAS NOMINATIVAS
- PNA AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE A
- PNB AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE B
- PNC AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE C
- PND AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE D
- PNE AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE
- PNF AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE F
- PNG AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE G
- PNH AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE H
- PN AÇÕES PREFERENCIAIS NOMINATIVAS
- PNV AÇÕES PREFERENCIAIS NOMINATIVAS COM DIREITO A VOTO
- OR AÇÕES ORDINÁRIAS NOMINATIVAS RESGATÁVEIS
- PRA AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE A RESGATÁVEIS

- PRB AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE B RESGATÁVEIS
- PRC AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE C RESGATÁVEIS
- PRD AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE D RESGATÁVEIS
- PRE AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE E RESGATÁVEIS
- PRF AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE F RESGATÁVEIS
- PRG AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE G RESGATÁVEIS
- PRH AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE H RESGATÁVEIS
- PNR AÇÕES PREFERENCIAIS NOMINATIVAS RESGATÁVEIS
- PRV AÇÕES PREFERENCIAIS NOMINATIVAS COM DIREITO A VOTO RESG
- ON P AÇÕES ORDINÁRIAS NOMINATIVAS COM DIREITOS DIFERENCIADOS
- PNA P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE A C/ DIREITOS DIFER
- PNB P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE B C/ DIREITOS DIFER
- PNC P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE C C/ DIREITOS DIFER
- PND P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE D C/ DIREITOS DIFER
- PNE P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE E C/ DIREITOS DIFER
- PNF P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE F C/ DIREITOS DIFER
- PNG P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE G C/ DIREITOS DIFER
- PNH P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE H C/ DIREITOS DIFER
- PN P AÇÕES PREFERENCIAIS NOMINATIVAS COM DIREITOS DIFERENCIADOS
- PNV P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE V C/ DIREITOS DIFER
- ON P AÇÕES ORDINÁRIAS NOMINATIVAS COM DIREITOS DIFERENCIADOS
- BDR BDR
- UNT CERTIFICADO DE DEPOSITO DE AÇÕES - MISCELÂNEA
- CDA CERTIFICADO DE DEPOSITO DE AÇÕES ORDINÁRIAS
- CPA CERTIFICADOS DE POTENCIAL ADICIONAL DE CONSTRUÇÃO E OPERAÇÃO
- RON CESTA DE AÇÕES ORDINÁRIAS NOMINATIVAS
- R CESTA DE AÇÕES NOMINATIVAS
- CI FUNDO DE INVESTIMENTO
- DIR DIREITOS DE SUBSCRIÇÃO MISCELÂNEA (BÔNUS, DEBÊNTURES, ETC)
- DIR ORD DIREITOS DE SUBSCRIÇÃO EM AÇÕES ORDINÁRIAS
- DIR P/A DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE A
- DIR P/B DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE B
- DIR P/C DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE C
- DIR P/D DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE D
- DIR P/E DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE E
- DIR P/F DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE F
- DIR P/G DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE G
- DIR P/H DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE H
- DIR PRE DIREITOS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS
- PRA REC RECIBO DE SUBSCRIÇÃO EM AÇÕES RESGATÁVEIS PREF. CLASSE A
- PRB REC RECIBO DE SUBSCRIÇÃO EM AÇÕES RESGATÁVEIS PREF. CLASSE B
- PRC REC RECIBO DE SUBSCRIÇÃO EM AÇÕES RESGATÁVEIS PREF. CLASSE C
- M1 REC RECIBO DE SUBSCRIÇÃO DE MISCELÂNEAS
- DIR PRA DIREITOS DE SUBSCRIÇÃO EM AÇÕES RESGATÁVEIS PREF. CLASSE A
- DIR PRB DIREITOS DE SUBSCRIÇÃO EM AÇÕES RESGATÁVEIS PREF. CLASSE B
- DIR PRC DIREITOS DE SUBSCRIÇÃO EM AÇÕES RESGATÁVEIS PREF. CLASSE C
- LFT LETRA FINANCEIRA DO TESOURO
- BNS ORD BÔNUS DE SUBSCRIÇÃO EM AÇÕES ORDINÁRIAS
- BNS P/A BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE A
- BNS P/B BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE B
- BNS P/C BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE C
- BNS P/D BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE D
- BNS P/E BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE E
- BNS P/F BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE F
- BNS P/G BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE G
- BNS P/H BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE H
- BNS PRE BÔNUS DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS

- PCD POSIÇÃO CONSOLIDADA DA DÍVIDA
- UP PRECATÓRIO
- REC RECIBO DE SUBSCRIÇÃO MISCELÂNEA
- ON REC RECIBO DE SUBSCRIÇÃO EM AÇÕES ORDINÁRIAS
- PNA REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE A
- PNB REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE B
- PNC REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE C
- PND REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE D
- PNE REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE E
- PNF REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE F
- PNG REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE G
- PNH REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS CLASSE H
- PN REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS
- PNV REC RECIBO DE SUBSCRIÇÃO EM AÇÕES PREFERENCIAIS C/ DIREITO VOTO
- UNT UNITS
- WRT WARRANTS DE DEBÊNTURES
- OR P AÇÕES ORDINÁRIAS NOMINATIVAS RESGATÁVEIS C/ DIREITOS DIF
- PRA P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE "A" RESG. C/ DIR.DIF
- PRB P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE "B" RESG. C/ DIR.DIF
- PRC P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE "C" RESG. C/ DIR.DIF
- PRD P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE "D" RESG. C/ DIR.DIF.
- PRE P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE "E" RESG. C/ DIR.DIF
- PRF P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE "F" RESG. C/ DIR.DIF.
- PRG P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE "G" RESG. C/ DIR.DIF
- PRH P AÇÕES PREFERENCIAIS NOMINATIVAS CLASSE "H" RESG. C/ DIR.DIF
- PR P AÇÕES PREFERENCIAIS NOMINATIVAS RESGATÁVEIS C/ DIREITOS DIF
- PRV P AÇÕES PREFERENCIAIS NOMINATIVAS RESG. C/ DIR.DIF. E DIR.VOTO

3. TABELA DE INDOPC - Relação dos valores para correção de contratos

- 1 US\$ CORREÇÃO PELA TAXA DO DÓLAR
- 2 TJLP CORREÇÃO PELA TJLP
- 3 TR CORREÇÃO PELA TR
- 4 IPCR CORREÇÃO PELO IPCR
- 5 SWA OPÇÕES DE TROCA - SWOPTIONS
- 6 ÍNDICES (PONTOS) OPÇÕES REFERENCIADAS EM PONTOS DE ÍNDICE
- 7 US\$ (PROTEGIDAS) CORREÇÃO PELA TAXA DO DÓLAR - OPÇÕES PROTEGIDAS
- 8 IGPM (PROTEGIDA) CORREÇÃO PELO IGP-M - OPÇÕES PROTEGIDAS
- 9 URV CORREÇÃO PELA URV

4. TABELA DE TPMERC - relação dos valores para tipo de mercado

- 010 VISTA
- 012 EXERCÍCIO DE OPÇÕES DE COMPRA
- 013 EXERCÍCIO DE OPÇÕES DE VENDA
- 017 LEILÃO
- 020 FRACIONÁRIO
- 030 TERMO
- 050 FUTURO COM RETENÇÃO DE GANHO
- 060 FUTURO COM MOVIMENTAÇÃO CONTÍNUA
- 070 OPÇÕES DE COMPRA
- 080 OPÇÕES DE VENDA

5. TABELA PROVISÓRIA DE CONVERSÃO DE CODINT PARA ISIN

- Estrutura do campo codint
 - CÓDIGO INTERNO DE IDENTIFICAÇÃO DE PAPEL USADO ATÉ 14/05/1995. (12 POSIÇÕES)
 - SOCEMI PIC X (04) CÓDIGO DA SOCIEDADE EMISSORA.
 - TIPTIT PIC X (02) TIPO DO TÍTULO
 - ESPANO PIC X (04) ESPÉCIE DO TÍTULO
 - DISMES PIC 9 (03) COMP-3 CORRESPONDE AO CAMPO QUE CONTINUA ABAIXO DE CODISI NA DEFINIÇÃO DO LAYOUT DO COTAHIST (DISMES)
- Estrutura do campo codisi
 - CÓDIGO ISIN DE IDENTIFICAÇÃO DE PAPEL, USADO DESDE 15/05/1995. (12 POSIÇÕES)
 - ESTE CÓDIGO NÃO PODE TER BRANCOS EM SUA ESTRUTURA.
 - 1.- PAÍS PIC X (02). FIXO BR
 - 2.- EMPRESA EMISSORA PIC X (04). CAMPO SOCEMI DO CODINT
 - 3.- TIPO DE ATIVO PIC X (03).
 - 4.- IDENTIFICADOR DE EMISSÃO DE TÍTULO PIC X (02).
 - 5.- DÍGITO DE CONTROLE PIC 9 (01).
 - SE TIPTIT = AC E SEGUNDA POSIÇÃO DO ESPANO = N, TIPO DE ATIVO = ACN
 - SE TIPTIT = AC E SEGUNDA POSIÇÃO DO ESPANO = N E TERCEIRA POSIÇÃO DO ESPANO = R, TIPO DE ATIVO = ARN
 - SE TIPTIT = AC E SOCEMI = YPFL, TIPO DE ATIVO = CDA
 - SE TIPTIT = CI E SOCEMI = UBBR, TIPO DE ATIVO = CDA
 - SE TIPTIT = IN, TIPO DE ATIVO = IND
 - SE TIPTIT = B1 OU S1, TIPO DE ATIVO = A01
 - SE TIPTIT = R1, TIPO DE ATIVO = R01
 - SE TIPTIT = BC, TIPO DE ATIVO = CPB
 - SE TIPTIT = CI E SOCEMI = BCBR, TIPO DE ATIVO = CPC
 - SE TIPTIT = CI, TIPO DE ATIVO = CTF
 - SE TIPTIT = TX, TIPO DE ATIVO = IND
 - SE PRIMEIRA POSIÇÃO DO ESPANO = O, IDENTIFICADOR DE TÍTULO = OR
 - SE PRIMEIRA POSIÇÃO DO ESPANO = P, IDENTIFICADOR DE TÍTULO = PR
 - SE TERCEIRA POSIÇÃO DO ESPANO = A A H, IDENTIFICADOR DE TÍTULO = PA A PH

APÊNDICE C – Resumo das Predições por Empresa

A seguir são apresentados os gráficos com resultados provenientes da aplicação dos dados na ferramenta de mineração EasyNN.

Tabela 17 - Tabela com dados da empresa Telesp – período de 30 dias

Data	Cotação real	Fase 1		Fase 2		Fase 3		Fase 4		Fase 5		Fase 6	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
01.08.08	3900	3946	1,18	3546	-9,08	3956	1,44	3611	-7,41	3285	-15,77	1915	-50,90
04.08.08	3885	3883	-0,05	3552	-8,57	3884	-0,03	3612	-7,03	3285	-15,44	2152	-44,61
05.08.08	3899	3917	0,46	3555	-8,82	3917	0,46	3614	-7,31	3285	-15,75	2314	-40,65
06.08.08	4000	3976	-0,60	3542	-11,45	3982	-0,45	3636	-9,10	3285	-17,88	1791	-55,23
07.08.08	3901	3962	1,56	3545	-9,13	3959	1,49	3617	-7,28	3285	-15,79	1931	-50,50
08.08.08	3860	3875	0,39	3548	-8,08	3880	0,52	3619	-6,24	3285	-14,90	1980	-47,93
11.08.08	3940	3899	-1,04	3554	-9,80	3915	-0,63	3620	-8,12	3285	-16,62	2333	-40,79
12.08.08	3950	3923	-0,68	3558	-9,92	3923	-0,68	3622	-8,30	3285	-16,84	2504	-36,61
13.08.08	3989	3963	-0,65	3544	-11,16	3969	-0,50	3624	-9,15	3285	-17,65	1948	-51,17
14.08.08	3998	4025	0,68	3547	-11,28	4034	0,90	3625	-9,33	3285	-17,83	2098	-47,52
15.08.08	4000	3992	-0,20	3550	-11,25	3992	-0,20	3627	-9,33	3285	-17,88	2257	-43,58
18.08.08	3949	3976	0,68	3556	-9,95	3993	1,11	3629	-8,10	3285	-16,81	2525	-36,06
19.08.08	3999	3981	-0,45	3560	-10,98	3993	-0,15	3630	-9,23	3285	-17,85	2706	-32,33
20.08.08	3980	3975	-0,13	3546	-10,90	3981	0,03	3632	-8,74	3285	-17,46	2116	-46,83
21.08.08	3888	3901	0,33	3549	-8,72	3908	0,51	3634	-6,53	3285	-15,51	2276	-41,46
22.08.08	3917	3917	0,00	3552	-9,32	3920	0,08	3635	-7,20	3285	-16,13	2444	-37,61
25.08.08	3889	3897	0,21	3558	-8,51	3905	0,41	3637	-6,48	3285	-15,53	2727	-29,88
26.08.08	3880	3861	-0,49	3562	-8,20	3868	-0,31	3638	-6,24	3285	-15,34	2917	-24,82
27.08.08	3870	3852	-0,47	3547	-8,35	3865	-0,13	3641	-5,92	3285	-15,12	2295	-40,70
28.08.08	3900	3887	-0,33	3551	-8,95	3901	0,03	3642	-6,62	3285	-15,77	2464	-36,82
29.08.08	3882	3902	0,52	3554	-8,45	3908	0,67	3644	-6,13	3285	-15,38	2643	-31,92

Tabela 18 - Tabela com dados da empresa Telesp – período de 15 dias

Data	Cotação real	Fase 7		Fase 8		Fase 9		Fase 10		Fase 11		Fase 12	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
18.08.08	3949	3975	0,66	3238	-18,00	3991	1,06	3518	-10,91	3348	-15,22	3700	-6,31
19.08.08	3999	3975	-0,60	3238	-19,03	3995	-0,10	3517	-12,05	3348	-16,28	3668	-8,28
20.08.08	3980	3966	-0,35	3239	-18,62	3980	0,00	3519	-11,58	3348	-15,88	3585	-9,92
21.08.08	3888	3893	0,13	3239	-16,69	3911	0,59	3518	-9,52	3348	-13,89	3556	-8,54
22.08.08	3917	3904	-0,33	3238	-17,33	3926	0,23	3517	-10,21	3348	-14,53	3528	-9,93
25.08.08	3889	3889	0,00	3238	-16,74	3908	0,49	3516	-9,59	3348	-13,91	3499	-10,03
26.08.08	3880	3852	-0,72	3238	-16,55	3875	-0,13	3515	-9,41	3348	-13,71	3471	-10,54
27.08.08	3870	3851	-0,49	3239	-16,30	3865	-0,13	3517	-9,12	3348	-13,49	3322	-14,16
28.08.08	3900	3886	-0,36	3239	-16,95	3902	0,05	3516	-9,85	3348	-14,15	3297	-15,46
29.08.08	3882	3892	0,26	3239	-16,56	3914	0,82	3515	-9,45	3348	-13,76	3273	-15,69

Tabela 19 – Tabela com dados da empresa Petrobrás – período de 30 dias

Data	Cotação real	Fase 1		Fase 2		Fase 3		Fase 4		Fase 5		Fase 6	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
01.08.08	4228	4302	1,75	3932	-7,00	4297	1,63	3348	-20,81	4445	5,13	4862	15,00
04.08.08	4011	4085	1,84	3961	-1,25	4082	1,77	3327	-17,05	4462	11,24	4852	20,97
05.08.08	3965	4000	0,88	4013	1,21	4003	0,96	3306	-16,62	4480	12,99	4847	22,24
06.08.08	4071	4093	0,54	3920	-3,71	4116	1,11	3441	-15,48	4417	8,50	4855	19,26
07.08.08	4132	4181	1,19	3979	-3,70	4191	1,43	3420	-17,23	4435	7,33	4851	17,40
08.08.08	4100	4102	0,05	4036	-1,56	4110	0,24	3399	-17,10	4452	8,59	4846	18,20
11.08.08	3986	4048	1,56	4063	1,93	4055	1,73	3377	-15,28	4470	12,14	4834	21,27
12.08.08	4020	4082	1,54	4115	2,36	4098	1,94	3356	-16,52	4487	11,62	4829	20,12
13.08.08	4125	4104	-0,51	4027	-2,38	4136	0,27	3494	-15,30	4424	7,25	4838	17,28
14.08.08	4090	4109	0,46	4086	-0,10	4112	0,54	3472	-15,11	4442	8,61	4833	18,17
15.08.08	4005	4014	0,22	4143	3,45	4023	0,45	3451	-13,83	4459	11,34	4827	20,52
18.08.08	3890	3957	1,72	4168	7,15	3970	2,06	3429	-11,85	4477	15,09	4814	23,75
19.08.08	4004	3993	-0,27	4220	5,39	4029	0,62	3408	-14,89	4495	12,26	4807	20,05
20.08.08	4200	4177	-0,55	4136	-1,52	4211	0,26	3549	-15,50	4431	5,50	4819	14,74
21.08.08	4331	4337	0,14	4195	-3,14	4360	0,67	3527	-18,56	4449	2,72	4813	11,13
22.08.08	4311	4314	0,07	4253	-1,35	4324	0,30	3505	-18,70	4466	3,60	4806	11,48
25.08.08	4165	4226	1,46	4276	2,67	4242	1,85	3483	-16,37	4484	7,66	4791	15,03
26.08.08	4201	4212	0,26	4329	3,05	4232	0,74	3461	-17,61	4502	7,16	4783	13,85
27.08.08	4320	4286	-0,79	4248	-1,67	4306	-0,32	3605	-16,55	4438	2,73	4796	11,02
28.08.08	4330	4321	-0,21	4308	-0,51	4330	0,00	3582	-17,27	4456	2,91	4789	10,60
29.08.08	4266	4316	1,17	4366	2,34	4326	1,41	3560	-16,55	4473	4,85	4782	12,10

Tabela 20 – Tabela com dados da empresa Petrobrás – período de 15 dias

Data	Cotação real	Fase 7		Fase 8		Fase 9		Fase 10		Fase 11		Fase 12	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
18.08.08	3890	3954	1,65	4627	18,95	3952	1,59	4290	10,28	4398	13,06	3909	0,49
19.08.08	4004	4007	0,07	4606	15,03	4009	0,12	4284	6,99	4365	9,02	3909	-2,37
20.08.08	4200	4193	-0,17	4695	11,79	4185	-0,36	4295	2,26	4490	6,90	3909	-6,93
21.08.08	4331	4343	0,28	4672	7,87	4337	0,14	4290	-0,95	4453	2,82	3909	-9,74
22.08.08	4311	4308	-0,07	4650	7,86	4305	-0,14	4285	-0,60	4418	2,48	3909	-9,32
25.08.08	4165	4227	1,49	4629	11,14	4221	1,34	4176	0,26	4391	5,43	3909	-6,15
26.08.08	4201	4214	0,31	4607	9,66	4214	0,31	4271	1,67	4357	3,71	3909	-6,95
27.08.08	4320	4293	-0,63	4696	8,70	4281	-0,90	4282	-0,88	4481	3,73	3909	-9,51
28.08.08	4330	4318	-0,28	4674	7,94	4310	-0,46	4277	-1,22	4445	2,66	3909	-9,72
29.08.08	4266	4313	1,10	4652	9,05	4307	0,96	4272	0,14	4410	3,38	3909	-8,37

Tabela 21 – Tabela com dados da empresa Embraer – período 30 dias

Data	Cotação real	Fase 1		Fase 2		Fase 3		Fase 4		Fase 5		Fase 6	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
01.08.08	1157	1210	4,58	1803	55,83	1204	4,06	1374	18,76	1549	33,88	1324	14,43
04.08.08	1166	1155	-0,94	1803	54,63	1148	-1,54	1377	18,10	1409	20,84	1203	3,17
05.08.08	1210	1207	-0,25	1803	49,01	1204	-0,50	1380	14,05	1282	5,95	1165	-3,72
06.08.08	1190	1202	1,01	1803	51,51	1193	0,25	1384	16,30	1780	49,58	1337	12,35
07.08.08	1146	1165	1,66	1803	57,33	1154	0,70	1387	21,03	1620	41,36	1294	12,91
08.08.08	1190	1180	-0,84	1803	51,51	1174	-1,34	1390	16,81	1474	23,87	1252	5,21
11.08.08	1205	1198	-0,58	1803	49,63	1192	-1,08	1393	15,60	1341	11,29	1138	-5,56
12.08.08	1279	1249	-2,35	1804	41,05	1249	-2,35	1396	9,15	1221	-4,53	1102	-13,84
13.08.08	1286	1286	0,00	1803	40,20	1282	-0,31	1400	8,86	1695	31,80	1264	-1,71
14.08.08	1325	1308	-1,28	1803	36,08	1304	-1,58	1403	5,89	1542	16,38	1224	-7,62
15.08.08	1335	1325	-0,75	1803	35,06	1321	-1,05	1406	5,32	1403	5,09	1185	-11,24
18.08.08	1336	1335	-0,07	1804	35,03	1334	-0,15	1409	5,46	1276	-4,49	1078	-19,31
19.08.08	1305	1319	1,07	1804	38,24	1318	1,00	1412	8,20	1162	-10,96	1044	-20,00
20.08.08	1327	1323	-0,30	1803	35,87	1320	-0,53	1416	6,71	1613	21,55	1196	-9,87
21.08.08	1310	1309	-0,08	1803	37,63	1305	-0,38	1419	8,32	1467	11,98	1158	-11,60
22.08.08	1352	1352	0,00	1352	0,00	1327	-1,85	1422	5,18	1335	-1,26	1352	0,00
25.08.08	1317	1334	1,29	1804	36,98	1334	1,29	1425	8,20	1215	-7,74	1021	-22,48
26.08.08	1345	1338	-0,52	1804	34,13	1340	-0,37	1428	6,17	1107	-17,70	989	-26,47
27.08.08	1328	1333	0,38	1803	35,77	1330	0,15	1432	7,83	1535	15,59	1132	-14,76
28.08.08	1369	1356	-0,95	1803	31,70	1357	-0,88	1435	4,82	1396	1,97	1096	-19,94
29.08.08	1383	1377	-0,43	1804	30,44	1378	-0,36	1438	3,98	1271	-8,10	1062	-23,21

Tabela 22 – Tabela com dados da empresa Embraer – período 15 dias

Data	Cotação real	Fase 7		Fase 8		Fase 9		Fase 10		Fase 11		Fase 12	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
18.08.08	1336	1326	-0,75	1557	16,54	1337	0,07	2088	56,29	1947	45,73	1106	-17,22
19.08.08	1305	1309	0,31	1730	32,57	1321	1,23	2158	65,36	1947	49,20	1176	-9,89
20.08.08	1327	1317	-0,75	1368	3,09	1324	-0,23	2064	55,54	1946	46,65	1012	-23,74
21.08.08	1310	1300	-0,76	1504	14,81	1309	-0,08	2133	62,82	1947	48,63	1051	-19,77
22.08.08	1352	1321	-2,29	1664	23,08	1333	-1,41	2205	63,09	1947	44,01	1104	-18,34
25.08.08	1317	1327	0,76	1853	40,70	1338	1,59	2224	68,87	1948	47,91	1174	-10,86
26.08.08	1345	1332	-0,97	2076	54,35	1346	0,07	2298	70,86	1949	44,91	1266	-5,87
27.08.08	1328	1328	0,00	1605	20,86	1334	0,45	2199	65,59	1947	46,61	1049	-21,01
28.08.08	1369	1352	-1,24	1780	30,02	1363	-0,44	2272	65,96	1948	42,29	1102	-19,50
29.08.08	1383	1372	-0,80	1987	43,67	1385	0,14	2347	69,70	1949	40,93	1171	-15,33

Tabela 23 – Tabela com dados da empresa Bradesco – período de 30 dias

Data	Cotação real	Fase 1		Fase 2		Fase 3		Fase 4		Fase 5		Fase 6	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
01.08.08	2930	2920	-0,34	1989	-32,12	2935	0,17	2341	-20,10	433	-85,22	2223	-24,13
04.08.08	2899	2889	-0,34	2023	-30,22	2906	0,24	2341	-19,25	486	-83,24	2221	-23,39
05.08.08	2899	2889	-0,34	2040	-29,63	2906	0,24	2341	-19,25	551	-80,99	2220	-23,42
06.08.08	2899	2887	-0,41	1976	-31,84	2904	0,17	2341	-19,25	315	-89,13	2223	-23,32
07.08.08	2899	2888	-0,38	1990	-31,36	2904	0,17	2342	-19,21	341	-88,24	2221	-23,39
08.08.08	2899	2889	-0,34	2005	-30,84	2904	0,17	2342	-19,21	373	-87,13	2220	-23,42
11.08.08	2899	2886	-0,45	2042	-29,56	2905	0,21	2342	-19,21	412	-85,79	2217	-23,53
12.08.08	2899	2887	-0,41	2060	-28,94	2905	0,21	2342	-19,21	460	-84,13	2214	-23,63
13.08.08	2899	2883	-0,55	1991	-31,32	2903	0,14	2342	-19,21	286	-90,13	2219	-23,46
14.08.08	2899	2884	-0,52	2006	-30,80	2903	0,14	2343	-19,18	305	-89,48	2217	-23,53
15.08.08	2707	2686	-0,78	2023	-25,27	2707	0,00	2343	-13,45	328	-87,88	2214	-18,21
18.08.08	2705	2665	-1,48	2062	-23,77	2686	-0,70	2343	-13,38	357	-86,80	2210	-18,30
19.08.08	2705	2666	-1,44	2082	-23,03	2686	-0,70	2343	-13,38	393	-85,47	2207	-18,41
20.08.08	2620	2591	-1,11	2008	-23,36	2618	-0,08	2343	-10,57	265	-89,89	2213	-15,53
21.08.08	2669	2645	-0,90	2024	-24,17	2671	0,07	2343	-12,21	279	-89,55	2210	-18,30
22.08.08	2680	2659	-0,78	2042	-23,81	2683	0,11	2344	-12,54	296	-88,96	2206	-17,69
25.08.08	2698	2668	-1,11	2085	-22,72	2695	-0,11	2344	-13,12	317	-88,25	2201	-18,42
26.08.08	2630	2619	-0,42	2106	-19,92	2640	0,38	2344	-10,87	343	-86,96	2195	-16,54
27.08.08	2594	2584	-0,39	2026	-21,90	2610	0,62	2344	-9,64	249	-90,40	2205	-15,00
28.08.08	2639	2593	-1,74	2043	-22,58	2620	-0,72	2344	-11,18	259	-90,19	2200	-16,64
29.08.08	2694	2673	-0,78	2002	-25,69	2704	0,37	2344	-12,99	272	-89,90	2195	-18,52

Tabela 24 – Tabela com dados da empresa Bradesco – período de 15 dias

Data	Cotação real	Fase 7		Fase 8		Fase 9		Fase 10		Fase 11		Fase 12	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
18.08.08	2705	2669	-1,33	2305	-14,79	2662	-1,59	2304	-14,82	1848	-31,68	1844	-31,83
19.08.08	2705	2669	-1,33	2303	-14,86	2664	-1,52	2290	-15,34	1870	-30,87	1918	-29,09
20.08.08	2620	2604	-0,61	2311	-11,79	2594	-0,99	2319	-11,49	1800	-31,30	1693	-35,38
21.08.08	2669	2656	-0,49	2309	-13,49	2649	-0,75	2305	-13,64	1834	-31,29	1756	-34,21
22.08.08	2680	2667	-0,49	2307	-13,92	2663	-0,63	2291	-14,51	1859	-30,63	1823	-31,98
25.08.08	2698	2677	-0,78	2305	-14,57	2673	-0,93	2268	-15,94	1870	-30,69	1895	-29,76
26.08.08	2630	2624	-0,23	2303	-12,43	2619	-0,42	2255	-14,26	1886	-28,29	1973	-24,98
27.08.08	2594	2597	0,12	2312	-10,87	2584	-0,39	2283	-11,99	1834	-29,30	1737	-33,04
28.08.08	2639	2606	-1,25	2309	-12,50	2598	-1,55	2270	-13,98	1859	-29,56	1803	-31,68
29.08.08	2694	2688	-0,22	2307	-14,37	2684	-0,37	2256	-16,26	1879	-30,25	1873	-30,48

Tabela 25 – Tabela com dados da empresa Banco do Brasil – período 30 dias

Data	Cotação real	Fase 1		Fase 2		Fase 3		Fase 4		Fase 5		Fase 6	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
01.08.08	2381	2430	2,06	3259	36,88	2434	2,23	2383	0,08	2663	11,84	2223	-6,64
04.08.08	2258	2314	2,48	3142	39,15	2319	2,70	2383	5,54	2661	17,85	2220	-1,68
05.08.08	2300	2322	0,96	3092	34,43	2333	1,43	2383	3,61	2659	15,61	2218	-3,57
06.08.08	2415	2383	-1,33	3299	36,60	2397	-0,75	2383	-1,33	2672	10,64	2225	-7,87
07.08.08	2380	2391	0,46	3245	36,34	2399	0,80	2383	0,13	2669	12,14	2224	-6,55
08.08.08	2340	2359	0,81	3192	36,41	2370	1,28	2383	1,84	2667	13,97	2223	-5,00
11.08.08	2257	2309	2,30	3079	36,42	2318	2,70	2383	5,58	2665	18,08	219	-90,30
12.08.08	2180	2235	2,52	3029	38,94	2242	2,84	2383	9,31	2662	22,11	2218	1,74
13.08.08	2120	2159	1,84	3232	52,45	2165	2,12	2383	12,41	2676	26,23	2225	4,95
14.08.08	2180	2168	-0,55	3179	45,83	2177	-0,14	2383	9,31	2674	22,66	2224	2,02
15.08.08	2229	2236	0,31	3127	40,29	2247	0,81	2383	6,91	2671	19,83	2223	-0,27
18.08.08	2229	2233	0,18	3017	35,35	2247	0,81	2383	6,91	2669	19,74	2219	-0,45
19.08.08	2229	2234	0,22	2969	33,20	2248	0,85	2384	6,95	2666	19,61	2218	-0,49
20.08.08	2229	2232	0,13	3166	42,04	2245	0,72	2384	6,95	2683	20,37	2225	-0,18
21.08.08	2229	2232	0,13	3114	39,70	2245	0,72	2384	6,95	2680	20,23	2224	-0,22
22.08.08	2229	2233	0,18	3064	37,46	2246	0,76	2384	6,95	2677	20,10	2223	-0,27
25.08.08	2229	2230	0,04	2957	32,66	2246	0,76	2384	6,95	2674	19,96	2219	-0,45
26.08.08	2229	2231	0,09	2910	30,55	2246	0,76	2384	6,95	2671	19,83	2217	-0,54
27.08.08	2420	2387	-1,36	3101	28,14	2406	-0,58	2384	-1,49	2693	11,28	2225	-8,06
28.08.08	2420	2387	-1,36	3052	26,12	2406	-0,58	2384	-1,49	2689	11,12	2224	-8,10
29.08.08	2379	2380	0,04	3003	26,23	2391	0,50	2384	0,21	2685	12,86	2223	-6,56

Tabela 26 – Tabela com dados da empresa Banco do Brasil – período 15 dias

Data	Cotação real	Fase 7		Fase 8		Fase 9		Fase 10		Fase 11		Fase 12	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
18.08.08	2229	2232	0,13	2347	5,29	2229	0,00	1876	-15,84	1995	-10,50	3832	71,92
19.08.08	2229	2232	0,13	2345	5,20	2231	0,09	1790	-19,69	1919	-13,91	3766	68,95
20.08.08	2229	2234	0,22	2349	5,38	2225	-0,18	1814	-18,62	2133	-4,31	4483	101,12
21.08.08	2229	2233	0,18	2348	5,34	2228	-0,04	1730	-22,39	2049	-8,08	4408	97,76
22.08.08	2229	2232	0,13	2347	5,29	2230	0,04	1650	-25,98	1970	-11,62	4334	94,44
25.08.08	2229	2231	0,09	2345	5,20	2229	0,00	1749	-21,53	1891	-15,16	4260	91,12
26.08.08	2229	2231	0,09	2344	5,16	2231	0,09	1669	-25,12	1822	-18,26	4187	87,84
27.08.08	2420	2394	-1,07	2348	-2,98	2383	-1,53	1691	-30,12	2019	-16,57	4962	105,04
28.08.08	2420	2393	-1,12	2347	-3,02	2386	-1,40	1613	-33,35	1941	-19,79	4883	101,78
29.08.08	2379	2378	-0,04	2346	-1,39	2373	-0,25	2256	-5,17	869	-63,47	4804	101,93

Tabela 27 – Tabela com dados da empresa Vale do Rio Doce – predição de 30 dias

Data	Cotação real	Fase 1		Fase 2		Fase 3		Fase 4		Fase 5		Fase 6	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
01.08.08	4439	4470	0,70	4585	3,29	4459	0,45	4852	9,30	5082	14,49	6338	42,78
04.08.08	4119	4174	1,34	4570	10,95	4163	1,07	4960	20,42	5082	23,38	6113	48,41
05.08.08	4163	4133	-0,72	4566	9,68	4132	-0,74	5069	21,76	5082	22,08	5900	41,72
06.08.08	4220	4221	0,02	4589	8,74	4223	0,07	4840	14,69	5082	20,43	6648	57,54
07.08.08	4240	4213	-0,64	4586	8,16	4209	-0,73	4947	16,67	5082	19,86	6419	51,39
08.08.08	4169	4176	0,17	4582	9,91	4181	0,29	5056	21,28	5082	21,90	6211	48,98
11.08.08	4015	4030	0,37	4567	13,75	4023	0,20	5166	28,67	5082	26,58	5989	49,17
12.08.08	3992	4025	0,83	4563	14,30	4033	1,03	5277	32,19	5082	27,30	5778	44,74
13.08.08	4175	4135	-0,96	4587	9,87	4165	-0,24	5043	20,79	5082	21,72	6521	56,19
14.08.08	4220	4246	0,62	4583	8,60	4255	0,83	5153	22,11	5082	20,43	6302	49,34
15.08.08	4140	4136	-0,10	4579	10,60	4139	-0,02	5264	27,15	5082	22,75	6086	47,00
18.08.08	4048	4085	0,91	4564	12,75	4088	0,99	5375	32,78	5082	25,54	5866	44,91
19.08.08	4110	4109	-0,02	4560	10,95	4136	0,63	5487	33,50	5082	23,65	5657	37,64
20.08.08	4336	4306	-0,69	4584	5,72	4327	-0,21	5250	21,08	5082	17,20	6393	47,44
21.08.08	4420	4428	0,18	4580	3,62	4446	0,59	5362	21,31	5082	14,98	6176	39,73
22.08.08	4348	4361	0,30	4577	5,27	4367	0,44	5474	25,90	5082	16,88	5962	37,12
25.08.08	4216	4249	0,78	4561	8,18	4253	0,88	5586	32,50	5082	20,54	5744	36,24
26.08.08	4220	4225	0,12	4558	8,01	4241	0,50	5699	35,05	5082	20,43	5538	31,23
27.08.08	4358	4277	-1,86	4581	5,12	4294	-1,47	5461	25,31	5082	16,61	6267	43,80
28.08.08	4380	4347	-0,75	4578	4,52	4355	-0,57	5583	27,47	5082	16,03	6052	38,17
29.08.08	4347	4346	-0,02	4578	5,31	4355	0,18	5582	28,41	5080	16,86	6050	39,18

Tabela 28 – Tabela com dados da empresa Vale do Rio Doce – predição de 15 dias

Data	Cotação real	Fase 7		Fase 8		Fase 9		Fase 10		Fase 11		Fase 12	
		Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.	Predição	% var.
18.08.08	4048	4074	0,64	4440	9,68	4070	0,54	4581	13,17	4459	10,15	4369	7,93
19.08.08	4110	4116	0,15	4440	8,03	4117	0,17	4582	11,48	4459	8,49	4369	6,30
20.08.08	4336	4312	-0,55	4465	2,98	4305	-0,71	4570	5,40	4459	2,84	4369	0,76
21.08.08	4420	4431	0,25	4464	1,00	4427	0,16	4573	3,46	4459	0,88	4369	-1,15
22.08.08	4348	4354	0,14	4464	2,67	4351	0,07	4575	5,22	4459	2,55	4369	0,48
25.08.08	4216	4241	0,59	4464	5,88	4235	0,45	4576	8,54	4459	5,76	4369	3,63
26.08.08	4220	4225	0,12	4464	5,78	4223	0,07	4577	8,46	4459	5,66	4369	3,53
27.08.08	4358	4282	-1,74	4488	2,98	4272	-1,97	4566	4,77	4459	2,32	4369	0,25
28.08.08	4380	4343	-0,84	4488	2,47	4337	-0,98	4568	4,29	4459	1,80	4369	-0,25
29.08.08	4347	4340	-0,16	4485	3,17	4335	-0,28	4565	5,01	4455	2,48	4365	0,41

APÊNDICE D – Resultados da validação e teste das RNA

Tabela 29 – Tabela com dados para validação – Janeiro 2008

EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO	EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO
1	2	44,21	46,50	5,1784	4	16	50,31	51,50	2,3642
2	2	103,87	104,31	0,4271	5	16	28,78	29,30	1,8001
3	2	20,62	20,40	-1,0592	6	16	50,96	51,41	0,8797
4	2	50,51	51,50	1,9629	1	17	41,70	41,75	0,1317
5	2	30,82	31,10	0,9244	2	17	97,07	97,00	-0,0767
6	2	57,29	57,90	1,0727	3	17	19,21	19,20	-0,0491
1	3	44,18	46,50	5,2501	4	17	50,30	51,50	2,3844
2	3	102,86	102,80	-0,0618	5	17	28,67	28,99	1,1193
3	3	20,10	20,50	1,9900	6	17	50,96	51,05	0,1736
4	3	50,48	51,50	2,0253	1	18	40,87	42,40	3,7417
5	3	30,80	31,10	0,9739	2	18	97,05	97,00	-0,0527
6	3	57,58	58,20	1,0720	3	18	19,02	19,00	-0,1024
1	4	44,15	46,50	5,3159	4	18	45,35	44,85	-1,1003
2	4	98,35	96,70	-1,6801	5	18	28,27	28,55	0,9826
3	4	20,04	20,40	1,7848	6	18	50,62	51,59	1,9137
4	4	50,45	51,50	2,0823	1	21	39,27	39,00	-0,6749
5	4	30,79	31,10	1,0188	2	21	97,02	97,00	-0,0173
6	4	55,97	56,10	0,2339	3	21	18,08	18,20	0,6416
1	7	44,15	46,50	5,3309	4	21	44,15	45,00	1,9173
2	7	94,34	93,38	-1,0206	5	21	26,73	27,09	1,3402
3	7	20,10	20,30	0,9708	6	21	46,73	46,22	-1,0961
4	7	50,44	51,50	2,0990	1	22	40,30	42,00	4,2228
5	7	30,79	31,10	1,0139	2	22	97,00	97,00	0,0011
6	7	54,07	55,10	1,9049	3	22	18,74	18,60	-0,7545
1	8	44,12	46,50	5,3931	4	22	42,65	43,00	0,8169
2	8	98,99	99,00	0,0128	5	22	27,99	28,70	2,5456
3	8	20,16	20,35	0,9606	6	22	47,06	47,60	1,1375
4	8	50,41	51,50	2,1528	1	23	41,24	43,00	4,2753
5	8	30,77	31,10	1,0564	2	23	96,99	97,00	0,0092
6	8	55,78	56,20	0,7483	3	23	18,03	18,22	1,0489
1	9	44,10	46,50	5,4494	4	23	43,36	43,00	-0,8370
2	9	98,90	100,68	1,7954	5	23	26,87	27,40	1,9882
3	9	20,50	20,85	1,6968	6	23	44,48	45,00	1,1779
4	9	50,39	51,50	2,2008	1	24	42,11	43,20	2,5799
5	9	30,76	31,10	1,0942	2	24	96,99	97,00	0,0069
6	9	54,46	55,60	2,0897	3	24	18,53	18,65	0,6584
1	10	44,08	46,50	5,4995	4	24	40,76	41,40	1,5589
2	10	98,35	98,60	0,2522	5	24	28,84	29,45	2,1146
3	10	20,78	21,00	1,0557	6	24	48,75	50,12	2,8110
4	10	50,37	51,50	2,2430	1	28	42,33	43,80	3,4721
5	10	30,75	31,10	1,1272	2	28	94,51	97,12	2,7585
6	10	55,48	56,81	2,4048	3	28	18,31	18,81	2,7078
1	11	44,06	46,50	5,5434	4	28	43,66	44,75	2,4981
2	11	97,22	97,00	-0,2292	5	28	28,56	29,10	1,8989
3	11	20,28	20,30	0,1018	6	28	48,06	50,25	4,5490
4	11	50,35	51,50	2,2791	1	29	43,82	45,40	3,6049
5	11	30,20	30,12	-0,2752	2	29	95,58	97,00	1,4815
6	11	55,54	56,10	1,0022	3	29	18,61	18,80	1,0443
1	14	44,05	46,50	5,5667	4	29	43,31	44,75	3,3288
2	14	97,20	97,00	-0,2069	5	29	29,28	29,70	1,4433
3	14	19,67	20,00	1,6560	6	29	50,05	51,08	2,0671
4	14	50,34	51,50	2,3047	1	30	43,58	45,00	3,2558
5	14	29,67	30,00	1,1255	2	30	96,92	99,45	2,6090
6	14	55,28	56,00	1,2981	3	30	18,45	18,80	1,8877
1	15	41,70	41,80	0,2411	4	30	43,71	44,53	1,8711
2	15	97,15	97,00	-0,1540	5	30	29,05	30,40	4,6481
3	15	19,41	19,59	0,9276	6	30	50,54	52,39	3,6527
4	15	50,32	51,50	2,3376	1	31	43,42	43,50	0,1914
5	15	29,26	29,31	0,1586	2	31	93,72	96,40	2,8553
6	15	53,54	53,90	0,6659	3	31	18,49	19,05	3,0099
1	16	40,94	43,00	5,0311	4	31	42,75	44,05	3,0354
2	16	97,11	97,00	-0,1105	5	31	28,81	29,25	1,5185
3	16	19,05	19,30	1,2932	6	31	50,54	51,83	2,5448

LEGENDA: (1) TELES P; (2) PETROBRAS; (3) EMBRAER; (4) BRADESCO; (5) BRASIL; (6) VALE.

Tabela 30 – Tabela com dados para validação – Fevereiro 2008

EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO	EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO
1	1	45,11	44,00	-2,4601	4	18	43,98	43,00	-2,2180
2	1	98,79	98,10	-0,6943	5	18	30,29	29,60	-2,2802
3	1	20,02	19,60	-2,1220	6	18	58,56	57,59	-1,6540
4	1	43,79	43,00	-1,8107	1	19	46,58	46,10	-1,0260
5	1	30,01	29,00	-3,3700	2	19	104,68	101,90	-2,6578
6	1	55,05	54,90	-0,2647	3	19	19,76	18,90	-4,3316
1	6	44,54	43,70	-1,8965	4	19	44,96	44,29	-1,4871
2	6	96,81	94,24	-2,6539	5	19	31,00	30,05	-3,0648
3	6	19,46	18,86	-3,0829	6	19	59,50	57,91	-2,6659
4	6	44,02	43,00	-2,3159	1	20	46,76	47,00	0,5096
5	6	29,34	28,35	-3,3699	2	20	104,64	103,05	-1,5191
6	6	54,80	53,11	-3,0752	3	20	19,63	19,40	-1,1957
1	7	44,49	43,90	-1,3153	4	20	44,68	43,89	-1,7766
2	7	99,48	97,36	-2,1283	5	20	31,24	30,89	-1,1049
3	7	19,08	18,60	-2,4959	6	20	60,13	60,00	-0,2228
4	7	44,00	43,00	-2,2777	1	21	46,98	46,25	-1,5561
5	7	28,90	28,00	-3,1228	2	21	103,46	101,46	-1,9358
6	7	53,35	52,40	-1,7879	3	21	20,16	19,40	-3,7885
1	8	44,54	43,89	-1,4680	4	21	44,90	44,10	-1,7753
2	8	99,94	97,64	-2,3013	5	21	32,36	31,70	-2,0496
3	8	19,22	18,79	-2,2164	6	21	60,68	59,80	-1,4548
4	8	43,99	43,00	-2,2400	1	22	46,86	46,10	-1,6127
5	8	28,93	28,00	-3,1986	2	22	102,85	101,16	-1,6442
6	8	54,33	53,50	-1,5259	3	22	19,92	19,49	-2,1769
1	11	44,71	44,50	-0,4747	4	22	45,04	44,18	-1,9063
2	11	101,89	99,52	-2,3246	5	22	31,91	31,15	-2,3704
3	11	19,92	19,50	-2,1052	6	22	60,14	60,70	0,9243
4	11	44,01	43,00	-2,3021	1	25	46,22	45,60	-1,3417
5	11	29,20	28,25	-3,2416	2	25	102,74	101,99	-0,7332
6	11	55,87	55,50	-0,6675	3	25	20,16	19,80	-1,7681
1	12	46,39	45,38	-2,1862	4	25	45,37	45,10	-0,6012
2	12	103,13	99,62	-3,4050	5	25	31,24	30,00	-3,9745
3	12	20,65	20,03	-3,0029	6	25	61,30	60,70	-0,9748
4	12	44,00	43,00	-2,2649	1	26	45,78	45,45	-0,7242
5	12	30,40	28,75	-5,4234	2	26	103,41	102,21	-1,1558
6	12	57,88	55,91	-3,4095	3	26	20,33	19,93	-1,9578
1	13	46,22	44,70	-3,2846	4	26	46,50	45,95	-1,1876
2	13	103,22	101,60	-1,5736	5	26	29,54	28,51	-3,4863
3	13	20,62	19,80	-3,9633	6	26	61,45	60,70	-1,2181
4	13	43,98	43,00	-2,2282	1	27	45,30	44,45	-1,8814
5	13	30,12	28,76	-4,5003	2	27	105,31	103,89	-1,3511
6	13	57,08	56,79	-0,5035	3	27	20,55	19,80	-3,6710
1	14	45,77	45,50	-0,5970	4	27	48,11	47,31	-1,6534
2	14	103,13	100,93	-2,1366	5	27	29,53	28,51	-3,4514
3	14	20,28	19,73	-2,7071	6	27	61,75	60,51	-2,0090
4	14	43,96	43,00	-2,1921	1	28	44,64	43,49	-2,5790
5	14	29,88	28,25	-5,4683	2	28	105,11	104,05	-1,0080
6	14	57,36	56,31	-1,8390	3	28	19,85	19,25	-3,0313
1	15	46,41	45,97	-0,9508	4	28	49,18	48,20	-2,0012
2	15	101,11	99,88	-1,2117	5	28	29,52	28,51	-3,4167
3	15	19,78	19,20	-2,9469	6	28	63,09	62,50	-0,9387
4	15	43,95	43,00	-2,1565	1	29	44,65	44,50	-0,3256
5	15	29,53	28,55	-3,3070	2	29	101,26	97,70	-3,5177
6	15	55,52	55,22	-0,5321	3	29	19,47	18,90	-2,9345
1	18	46,94	46,10	-1,7967	4	29	50,10	49,80	-0,6001
2	18	102,54	101,86	-0,6599	5	29	29,51	28,51	-3,3824
3	18	19,85	19,10	-3,7841	6	29	63,09	58,66	-7,0250

LEGENDA: (1) TELES P; (2) PETROBRAS; (3) EMBRAER; (4) BRADESCO; (5) BRASIL; (6) VALE.

Tabela 31 – Tabela com dados para validação – Março 2008

EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO	EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO
1	3	45,00	44,50	-1,1209	1	17	41,32	39,98	-3,2437
2	3	99,63	98,20	-1,4355	2	17	89,74	89,00	-0,8244
3	3	19,26	18,40	-4,4558	3	17	18,09	17,33	-4,2160
4	3	48,94	47,50	-2,9507	4	17	46,20	45,70	-1,0728
5	3	29,47	28,51	-3,2437	5	17	26,50	25,20	-4,8912
6	3	59,02	58,30	-1,2211	6	17	56,29	55,33	-1,7121
1	4	43,98	42,71	-2,8985	1	18	41,70	41,38	-0,7555
2	4	96,16	94,30	-1,9335	2	18	90,60	90,01	-0,6497
3	4	19,00	18,30	-3,6667	3	18	18,09	17,33	-4,2030
4	4	49,20	49,00	-0,4099	4	18	44,73	44,84	0,2392
5	4	29,45	28,51	-3,1897	5	18	26,45	25,75	-2,6586
6	4	58,18	57,06	-1,9292	6	18	56,99	57,08	0,1561
1	5	45,78	45,49	-0,6375	1	19	40,81	39,10	-4,1966
2	5	98,15	97,00	-1,1753	2	19	86,92	83,80	-3,5882
3	5	18,92	18,30	-3,2700	3	19	18,09	17,33	-4,1943
4	5	49,17	49,00	-0,3469	4	19	46,03	45,60	-0,9298
5	5	29,43	28,51	-3,1421	5	19	26,24	25,25	-3,7658
6	5	59,52	58,77	-1,2581	6	19	55,25	53,16	-3,7827
1	6	44,10	42,85	-2,8361	1	20	41,10	39,90	-2,9086
2	6	97,91	95,38	-2,5875	2	20	83,44	83,25	-0,2278
3	6	18,74	18,07	-3,5752	3	20	18,09	17,33	-4,1897
4	6	49,14	49,00	-0,2913	4	20	44,79	42,79	-4,4648
5	6	29,42	28,51	-3,1005	5	20	25,61	24,90	-2,7860
6	6	58,76	57,23	-2,6025	6	20	53,60	52,73	-1,6300
1	7	43,75	42,45	-2,9778	1	24	41,84	41,05	-1,8807
2	7	95,58	94,08	-1,5649	2	24	86,75	84,68	-2,3819
3	7	18,56	17,72	-4,5156	3	24	18,10	17,33	-4,2365
4	7	49,12	49,00	-0,2430	4	24	43,42	42,95	-1,0766
5	7	29,41	28,51	-3,0647	5	24	25,62	24,90	-2,8254
6	7	57,22	56,02	-2,1055	6	24	56,19	55,45	-1,3207
1	10	42,91	41,90	-2,3444	1	25	41,35	40,40	-2,3007
2	10	94,77	92,58	-2,3085	2	25	89,40	89,57	0,1952
3	10	18,32	17,80	-2,8349	3	25	18,10	17,33	-4,2339
4	10	49,12	49,00	-0,2446	4	25	44,54	43,50	-2,3386
5	10	27,51	26,50	-3,6735	5	25	25,62	24,90	-2,8277
6	10	55,27	54,50	-1,3895	6	25	57,67	57,02	-1,1231
1	11	42,90	41,60	-3,0395	1	26	40,73	39,50	-3,0301
2	11	96,31	96,20	-0,1139	2	26	90,92	90,50	-0,4667
3	11	18,39	17,80	-3,1899	3	26	18,10	17,33	-4,2353
4	11	49,10	49,00	-0,2037	4	26	45,13	44,72	-0,9178
5	11	27,36	26,59	-2,7990	5	26	25,63	24,90	-2,8341
6	11	56,59	56,74	0,2583	6	26	60,18	59,51	-1,1209
1	12	42,34	41,50	-1,9729	1	27	40,74	39,50	-3,0348
2	12	97,05	95,90	-1,1871	2	27	90,23	88,35	-2,0816
3	12	18,46	17,80	-3,5941	3	27	18,10	17,33	-4,2405
4	12	49,08	49,00	-0,1695	4	27	44,74	43,80	-2,1052
5	12	26,95	25,77	-4,3818	5	27	25,63	24,90	-2,8445
6	12	58,61	57,30	-2,2312	6	27	60,78	59,75	-1,6927
1	13	41,75	41,49	-0,6308	1	28	40,74	39,50	-3,0444
2	13	94,22	94,52	0,3160	2	28	89,40	88,14	-1,4053
3	13	18,13	17,50	-3,4560	3	28	18,10	17,33	-4,2495
4	13	49,07	49,00	-0,1419	4	28	43,85	43,00	-1,9381
5	13	26,24	25,80	-1,6878	5	28	25,63	24,90	-2,8590
6	13	57,82	57,60	-0,3882	6	28	60,16	59,80	-0,5999
1	14	41,19	40,50	-1,6866	1	31	40,74	39,50	-3,0552
2	14	94,30	93,99	-0,3241	2	31	89,88	89,67	-0,2376
3	14	18,62	18,05	-3,0864	3	31	17,83	17,47	-2,0186
4	14	49,06	49,00	-0,1206	4	31	42,83	42,30	-1,2356
5	14	27,25	26,50	-2,7456	5	31	25,64	24,90	-2,8858
6	14	58,24	58,90	1,1304	6	31	60,16	60,50	0,5637

LEGENDA: (1) TELES P; (2) PETROBRAS; (3) EMBRAER; (4) BRADESCO; (5) BRASIL; (6) VALE.

Tabela 32 – Tabela com dados para validação – Abril 2008

EMPR	DIA	PREDIÇÃO	REAL	VARIACÃO	EMPR	DIA	PREDIÇÃO	REAL	VARIACÃO
1	1	39,93	39,50	-1,0836	4	15	43,41	43,80	0,9001
2	1	89,72	91,31	1,7749	5	15	22,84	22,77	-0,3025
3	1	17,66	17,45	-1,2164	6	15	60,11	61,21	1,8329
4	1	42,03	42,70	1,5822	1	16	38,33	38,17	-0,4156
5	1	25,00	24,90	-0,4061	2	16	95,45	94,88	-0,5991
6	1	60,97	61,75	1,2755	3	16	17,56	17,31	-1,4206
1	2	39,91	39,50	-1,0174	4	16	43,41	43,80	0,9080
2	2	93,65	93,50	-0,1613	5	16	24,01	24,35	1,4010
3	2	17,66	17,31	-2,0022	6	16	60,11	61,21	1,8348
4	2	43,50	43,80	0,6863	1	17	38,81	38,80	-0,0320
5	2	24,99	24,90	-0,3551	2	17	102,72	103,60	0,8608
6	2	60,79	61,10	0,5136	3	17	17,57	17,30	-1,5312
1	3	39,88	39,50	-0,9599	4	17	43,41	43,80	0,9091
2	3	94,90	94,86	-0,0385	5	17	25,03	25,35	1,2806
3	3	17,51	17,32	-1,0847	6	17	60,11	61,21	1,8290
4	3	43,48	43,80	0,7381	1	18	38,51	38,50	-0,0311
5	3	25,13	24,90	-0,9316	2	18	103,63	104,40	0,7390
6	3	61,52	61,87	0,5629	3	18	17,69	17,60	-0,5242
1	4	39,86	39,50	-0,9109	4	18	30,75	30,80	0,1670
2	4	95,64	94,88	-0,7895	5	18	25,93	25,40	-2,0359
3	4	17,71	17,70	-0,0780	6	18	60,12	61,21	1,8157
4	4	43,46	43,80	0,7818	1	22	38,65	38,62	-0,0885
5	4	24,90	24,66	-0,9524	2	22	105,91	105,61	-0,2807
6	4	62,06	62,51	0,7199	3	22	17,74	17,32	-2,3508
1	7	39,87	39,50	-0,9256	4	22	31,11	31,00	-0,3445
2	7	95,63	94,88	-0,7806	5	22	25,19	25,10	-0,3489
3	7	18,10	17,71	-2,1347	6	22	60,09	61,21	1,8580
4	7	43,47	43,80	0,7693	1	23	38,60	38,55	-0,1268
5	7	25,25	24,95	-1,1933	2	23	104,36	104,21	-0,1457
6	7	63,73	62,41	-2,0778	3	23	17,38	17,25	-0,7709
1	8	39,41	39,59	0,4476	4	23	30,95	30,77	-0,5735
2	8	95,57	94,88	-0,7255	5	23	24,89	24,55	-1,3786
3	8	18,52	18,50	-0,1053	6	23	60,10	61,21	1,8481
4	8	43,45	43,80	0,8093	1	24	38,49	38,70	0,5478
5	8	25,01	25,10	0,3641	2	24	102,31	102,00	-0,3034
6	8	62,88	62,70	-0,2909	3	24	17,49	17,31	-1,0118
1	9	38,79	39,09	0,7620	4	24	30,59	30,85	0,8654
2	9	95,53	94,88	-0,6841	5	24	25,03	25,32	1,1654
3	9	18,47	18,05	-2,2872	6	24	62,88	62,25	-1,0026
4	9	43,43	43,80	0,8417	1	25	38,53	38,66	0,3290
5	9	24,64	24,35	-1,1850	2	25	102,55	103,00	0,4356
6	9	61,30	61,60	0,4884	3	25	17,49	17,35	-0,7995
1	10	38,95	39,39	1,1423	4	25	30,79	31,00	0,6778
2	10	95,51	94,88	-0,6560	5	25	26,37	26,60	0,8755
3	10	18,49	18,32	-0,9445	6	25	62,20	63,00	1,2879
4	10	43,42	43,80	0,8666	1	28	37,83	38,00	0,4384
5	10	23,82	23,56	-1,0853	2	28	102,47	103,00	0,5132
6	10	60,13	61,21	1,7990	3	28	17,63	17,52	-0,6245
1	11	38,70	38,98	0,7150	4	28	31,29	31,38	0,2933
2	11	95,49	94,88	-0,6410	5	28	27,19	26,75	-1,6138
3	11	18,29	18,01	-1,5283	6	28	63,74	63,86	0,1845
4	11	43,42	43,80	0,8843	1	29	37,30	37,80	1,3358
5	11	23,66	23,40	-1,0952	2	29	102,53	103,00	0,4577
6	11	60,12	61,21	1,8119	3	29	17,75	17,42	-1,8386
1	14	39,26	38,40	-2,1996	4	29	31,48	31,40	-0,2671
2	14	95,46	94,88	-0,6067	5	29	26,42	26,36	-0,2162
3	14	17,74	17,42	-1,7860	6	29	62,38	62,30	-0,1254
4	14	43,42	43,80	0,8854	1	30	37,33	38,00	1,7906
5	14	22,93	22,62	-1,3357	2	30	102,60	103,00	0,3898
6	14	60,11	61,21	1,8232	3	30	17,47	17,35	-0,7076
1	15	38,32	38,85	1,3841	4	30	30,51	30,74	0,7545
2	15	95,45	94,88	-0,5966	5	30	28,42	28,90	1,6896
3	15	17,61	17,30	-1,7871	6	30	62,38	65,64	5,2291

LEGENDA: (1) TELES P; (2) PETROBRAS; (3) EMBRAER; (4) BRADESCO; (5) BRASIL; (6) VALE.

Tabela 33 – Tabela com dados para validação – Maio 2008

EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO	EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO
1	2	40,16	39,94	-0,5365	1	16	41,83	41,65	-0,4223
2	2	103,33	103,00	-0,3237	2	16	57,34	57,90	0,9679
3	2	17,70	17,15	-3,1226	3	16	17,39	17,28	-0,6077
4	2	32,48	32,89	1,2721	4	16	33,23	33,25	0,0637
5	2	30,41	30,43	0,0804	5	16	29,07	28,93	-0,4965
6	2	65,73	65,40	-0,4950	6	16	71,41	72,09	0,9529
1	5	41,65	41,20	-1,0834	1	19	41,40	41,95	1,3298
2	5	103,32	103,00	-0,3050	2	19	59,57	59,69	0,2001
3	5	17,48	17,19	-1,6778	3	19	17,41	17,31	-0,5634
4	5	34,88	34,49	-1,1260	4	19	33,82	33,69	-0,3800
5	5	29,95	30,00	0,1766	5	19	29,07	28,93	-0,4912
6	5	66,82	66,60	-0,3320	6	19	71,53	71,50	-0,0481
1	6	41,23	40,62	-1,4884	1	20	41,36	41,30	-0,1547
2	6	103,21	103,00	-0,2027	2	20	60,89	61,50	1,0092
3	6	17,49	17,40	-0,5129	3	20	17,33	17,18	-0,8625
4	6	34,86	34,49	-1,0573	4	20	33,74	34,00	0,7792
5	6	29,28	28,90	-1,2937	5	20	29,07	28,93	-0,4676
6	6	67,69	67,82	0,1989	6	20	70,50	71,50	1,4186
1	7	40,63	40,06	-1,3935	1	21	41,36	41,30	-0,1350
2	7	103,12	103,00	-0,1134	2	21	63,22	62,30	-1,4611
3	7	17,51	17,21	-1,6929	3	21	17,04	16,65	-2,2617
4	7	34,84	34,49	-0,9950	4	21	33,56	33,85	0,8731
5	7	28,53	27,99	-1,8902	5	21	29,06	28,93	-0,4492
6	7	66,92	65,95	-1,4547	6	21	69,85	69,20	-0,9255
1	8	40,02	39,84	-0,4551	1	23	41,35	41,30	-0,1205
2	8	103,04	103,00	-0,0369	2	23	59,83	59,95	0,2056
3	8	17,35	17,15	-1,1401	3	23	15,90	15,60	-1,8789
4	8	34,82	34,49	-0,9391	4	23	33,13	33,10	-0,1031
5	8	28,78	28,75	-0,0930	5	23	29,06	28,93	-0,4376
6	8	67,25	66,74	-0,7545	6	23	67,28	68,25	1,4458
1	9	39,90	39,71	-0,4655	1	26	41,34	41,30	-0,1079
2	9	54,09	54,74	1,2004	2	26	61,25	61,70	0,7333
3	9	17,23	17,12	-0,6536	3	26	15,83	15,72	-0,7030
4	9	34,80	34,49	-0,8894	4	26	32,98	32,98	0,0008
5	9	28,51	28,55	0,1366	5	26	29,05	28,93	-0,4300
6	9	66,36	66,41	0,0737	6	26	67,90	67,50	-0,5820
1	12	40,09	40,00	-0,2197	1	27	41,34	41,30	-0,0942
2	12	54,60	54,92	0,5839	2	27	59,15	59,20	0,0816
3	12	17,29	17,28	-0,0589	3	27	15,42	15,29	-0,8135
4	12	34,80	34,49	-0,8880	4	27	32,61	32,68	0,2137
5	12	28,71	28,99	0,9792	5	27	29,05	28,93	-0,4164
6	12	66,74	67,31	0,8492	6	27	65,73	66,19	0,6982
1	13	39,69	39,50	-0,4900	1	28	41,34	41,30	-0,0865
2	13	56,00	56,24	0,4250	2	28	59,69	60,49	1,3402
3	13	17,26	17,13	-0,7566	3	28	15,11	14,96	-0,9870
4	13	34,78	34,49	-0,8428	4	28	32,71	33,00	0,8725
5	13	28,85	28,40	-1,5577	5	28	30,56	30,85	0,9460
6	13	67,32	67,50	0,2696	6	28	66,21	67,66	2,1954
1	14	40,20	40,00	-0,5009	1	29	41,34	41,30	-0,0848
2	14	55,84	55,36	-0,8600	2	29	59,34	58,01	-2,2493
3	14	17,43	17,10	-1,8771	3	29	15,04	14,98	-0,4011
4	14	34,77	34,49	-0,8035	4	29	33,82	34,15	0,9857
5	14	29,09	28,93	-0,5561	5	29	31,12	31,01	-0,3540
6	14	68,42	67,26	-1,6942	6	29	65,44	64,45	-1,5060
1	15	41,30	41,50	0,4740	1	30	41,34	41,30	-0,0891
2	15	55,72	56,40	1,2288	2	30	58,03	57,40	-1,0922
3	15	17,33	17,30	-0,1703	3	30	15,39	15,29	-0,6336
4	15	34,76	34,49	-0,7701	4	30	34,29	34,00	-0,8362
5	15	29,08	28,93	-0,5236	5	30	32,33	32,54	0,6388
6	15	69,03	69,60	0,8246	6	30	64,45	65,51	1,6447

LEGENDA: (1) TELES P; (2) PETROBRAS; (3) EMBRAER; (4) BRADESCO; (5) BRASIL; (6) VALE.

Tabela 34 – Tabela com dados para validação – Junho 2008

EMPR	DIA	PREDIÇÃO	REAL	VARIACÃO	EMPR	DIA	PREDIÇÃO	REAL	VARIACÃO
1	2	41,30	41,30	0,0000	4	16	31,40	31,59	0,6095
2	2	58,44	58,44	0,0000	5	16	29,06	29,40	1,1754
3	2	14,53	14,53	0,0000	6	16	56,39	56,80	0,7230
4	2	34,20	34,20	0,0000	1	17	39,95	39,60	-0,8698
5	2	31,67	31,67	0,0000	2	17	56,04	56,02	-0,0418
6	2	65,29	65,29	0,0000	3	17	12,43	12,37	-0,5213
1	3	40,82	40,59	-0,5524	4	17	31,37	31,64	0,8484
2	3	56,27	55,49	-1,3782	5	17	29,63	29,75	0,4132
3	3	14,28	14,20	-0,5873	6	17	58,42	58,48	0,0952
4	3	33,54	33,79	0,7337	1	18	39,64	39,30	-0,8636
5	3	30,82	30,50	-1,0240	2	18	55,34	55,30	-0,0698
6	3	62,85	62,96	0,1734	3	18	11,89	11,88	-0,0720
1	4	40,94	41,00	0,1467	4	18	32,00	32,15	0,4587
2	4	53,62	53,50	-0,2176	5	18	28,76	28,50	-0,8999
3	4	14,19	14,25	0,4238	6	18	58,16	58,17	0,0151
4	4	33,53	33,79	0,7629	1	19	39,58	39,70	0,3117
5	4	30,57	30,21	-1,1732	2	19	53,69	53,39	-0,5562
6	4	61,14	60,74	-0,6618	3	19	11,70	11,96	2,1980
1	5	41,88	41,80	-0,1918	4	19	31,30	31,30	-0,0138
2	5	55,53	56,14	1,1008	5	19	28,42	28,50	0,2767
3	5	14,17	14,10	-0,5055	6	19	57,55	58,05	0,8713
4	5	33,52	33,79	0,7940	1	20	40,29	39,30	-2,4531
5	5	31,29	31,80	1,6142	2	20	52,97	52,40	-1,0702
6	5	62,73	62,85	0,1863	3	20	11,94	12,27	2,7270
1	6	41,28	41,25	-0,0805	4	20	30,96	31,29	1,0790
2	6	56,86	56,15	-1,2474	5	20	27,88	27,41	-1,6778
3	6	13,81	13,60	-1,5097	6	20	56,09	55,44	-1,1595
4	6	33,52	33,79	0,7911	1	23	39,75	39,94	0,4837
5	6	30,83	30,93	0,3127	2	23	53,61	53,41	-0,3646
6	6	61,58	61,79	0,3346	3	23	11,95	12,02	0,5883
1	9	41,10	41,40	0,7299	4	23	30,11	30,39	0,9360
2	9	55,85	55,96	0,2028	5	23	27,20	27,04	-0,5736
3	9	13,19	13,20	0,0677	6	23	56,44	56,48	0,0659
4	9	33,47	33,79	0,9648	1	24	39,66	39,79	0,3387
5	9	30,21	30,45	0,8090	2	24	53,81	53,20	-1,1314
6	9	59,47	59,70	0,3840	3	24	11,98	12,02	0,3264
1	10	40,89	40,26	-1,5371	4	24	30,08	29,85	-0,7489
2	10	54,78	54,30	-0,8709	5	24	27,23	27,04	-0,6846
3	10	12,88	13,21	2,6002	6	24	57,49	57,49	-0,0069
4	10	33,51	33,79	0,8288	1	25	40,06	39,90	-0,3964
5	10	30,03	30,30	0,8934	2	25	53,95	54,69	1,3762
6	10	57,89	58,40	0,8856	3	25	11,98	12,02	0,3595
1	11	40,45	40,64	0,4676	4	25	29,65	29,88	0,7927
2	11	55,05	54,40	-1,1883	5	25	27,22	27,04	-0,6653
3	11	13,20	13,05	-1,1426	6	25	58,54	58,68	0,2340
4	11	33,50	33,79	0,8559	1	26	39,97	40,00	0,0647
5	11	29,81	29,70	-0,3834	2	26	54,79	54,80	0,0096
6	11	56,89	56,59	-0,5347	3	26	11,97	12,02	0,3951
1	12	40,52	40,45	-0,1772	4	26	30,18	30,22	0,1323
2	12	54,81	54,54	-0,4861	5	26	27,22	27,04	-0,6444
3	12	12,93	12,96	0,1943	6	26	56,88	56,60	-0,4983
4	12	33,49	33,79	0,8849	1	27	40,59	40,20	-0,9592
5	12	30,06	29,61	-1,4974	2	27	54,98	55,00	0,0428
6	12	56,84	56,69	-0,2658	3	27	11,97	12,02	0,4330
1	13	39,80	39,40	-0,9995	4	27	29,23	29,00	-0,7853
2	13	55,78	55,61	-0,3090	5	27	27,21	27,04	-0,6218
3	13	12,78	12,88	0,8032	6	27	56,50	56,23	-0,4818
4	13	33,48	33,79	0,9157	1	30	40,63	40,50	-0,3164
5	13	29,25	28,80	-1,5386	2	30	56,64	56,45	-0,3426
6	13	56,56	56,76	0,3531	3	30	11,95	12,02	0,6113
1	16	39,72	39,55	-0,4243	4	30	28,84	29,10	0,9099
2	16	55,93	55,79	-0,2493	5	30	27,16	27,04	-0,4508
3	16	12,62	12,68	0,4890	6	30	56,50	57,40	1,5889

LEGENDA: (1) TELES P; (2) PETROBRAS; (3) EMBRAER; (4) BRADESCO; (5) BRASIL; (6) VALE.

Tabela 35 – Tabela com dados para validação – Julho 2008

EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO	EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO
1	1	39,59	39,99	1,0153	1	17	38,65	38,49	-0,4165
2	1	56,60	56,30	-0,5351	2	17	46,41	46,00	-0,8896
3	1	12,03	12,02	-0,0477	3	17	11,71	11,63	-0,7061
4	1	28,96	29,10	0,4839	4	17	28,37	29,23	3,0165
5	1	27,38	27,04	-1,2525	5	17	25,47	24,61	-3,3942
6	1	54,74	55,01	0,4912	6	17	46,12	46,54	0,9073
1	2	39,24	39,80	1,4310	1	18	38,86	38,99	0,3349
2	2	55,29	54,00	-2,3373	2	18	46,00	46,01	0,0148
3	2	12,02	12,02	-0,0286	3	18	11,33	11,30	-0,2539
4	2	28,96	29,10	0,4961	4	18	29,09	29,30	0,7086
5	2	27,38	27,04	-1,2407	5	18	25,27	24,95	-1,2858
6	2	53,12	52,19	-1,7448	6	18	46,25	46,01	-0,5180
1	3	39,16	39,39	0,5860	1	21	39,19	39,47	0,7094
2	3	52,77	51,85	-1,7440	2	21	46,80	47,03	0,4981
3	3	12,02	12,02	-0,0097	3	21	11,50	11,43	-0,5763
4	3	28,95	29,10	0,5079	4	21	29,69	29,89	0,6662
5	3	27,38	27,04	-1,2292	5	21	25,30	25,02	-1,1022
6	3	51,27	50,95	-0,6277	6	21	46,90	47,70	1,7129
1	4	38,42	38,70	0,7203	1	22	38,80	39,40	1,5527
2	4	52,67	52,60	-0,1325	2	22	45,62	45,52	-0,2236
3	4	9,43	9,64	2,2721	3	22	11,60	12,10	4,3482
4	4	28,95	29,10	0,5195	4	22	30,04	29,92	-0,3929
5	4	23,91	23,80	-0,4634	5	22	24,30	24,00	-1,2375
6	4	51,38	51,50	0,2343	6	22	45,91	45,91	0,0089
1	7	38,30	38,20	-0,2643	1	23	39,56	39,58	0,0409
2	7	52,09	51,26	-1,5918	2	23	44,43	44,01	-0,9423
3	7	9,80	10,10	3,0215	3	23	12,08	12,21	1,0479
4	7	28,95	29,10	0,5333	4	23	29,44	29,89	1,5294
5	7	23,76	23,07	-2,9107	5	23	24,56	24,80	0,9618
6	7	52,36	51,56	-1,5355	6	23	46,25	45,87	-0,8185
1	8	37,49	37,70	0,5666	1	24	38,76	38,90	0,3733
2	8	49,51	49,73	0,4384	2	24	42,74	41,87	-2,0421
3	8	10,63	11,16	4,9524	3	24	11,79	11,91	1,0201
4	8	28,94	29,10	0,5443	4	24	29,70	29,99	0,9632
5	8	23,67	24,25	2,4693	5	24	24,72	24,73	0,0297
6	8	50,37	51,00	1,2509	6	24	44,39	43,73	-1,4863
1	10	38,22	37,40	-2,1394	1	25	38,46	39,00	1,3911
2	10	48,67	49,03	0,7333	2	25	41,49	41,71	0,5184
3	10	10,93	10,90	-0,2909	3	25	11,65	11,85	1,7009
4	10	28,94	29,10	0,5560	4	25	29,40	29,50	0,3273
5	10	24,26	24,20	-0,2656	5	25	24,52	24,25	-1,0833
6	10	50,56	51,00	0,8742	6	25	42,37	43,50	2,6772
1	11	37,43	37,79	0,9545	1	28	38,73	38,82	0,2398
2	11	49,95	49,46	-0,9726	2	28	42,82	42,08	-1,7191
3	11	10,55	10,84	2,7821	3	28	11,92	11,72	-1,7092
4	11	28,94	29,10	0,5664	4	28	28,84	29,00	0,5693
5	11	24,10	23,80	-1,2407	5	28	23,70	23,07	-2,6683
6	11	50,52	50,17	-0,6896	6	28	44,17	43,99	-0,4004
1	14	38,82	39,49	1,7286	1	29	38,90	39,00	0,2452
2	14	49,57	49,60	0,0642	2	29	42,15	42,54	0,9135
3	14	10,68	10,65	-0,3020	3	29	11,88	12,22	2,8730
4	14	28,93	29,10	0,5783	4	29	28,70	28,49	-0,7487
5	14	24,34	24,15	-0,7772	5	29	23,89	24,40	2,1380
6	14	50,92	51,00	0,1517	6	29	44,52	45,05	1,1816
1	15	38,87	39,40	1,3565	1	30	38,94	38,81	-0,3328
2	15	48,37	48,79	0,8753	2	30	43,79	44,61	1,8697
3	15	10,38	10,90	5,0169	3	30	12,22	12,28	0,4524
4	15	28,93	29,10	0,5881	4	30	28,59	28,95	1,2699
5	15	23,97	24,40	1,7741	5	30	25,02	25,28	1,0288
6	15	49,37	49,91	1,0847	6	30	47,18	48,10	1,9404
1	16	39,04	39,10	0,1632	1	31	38,85	38,83	-0,0557
2	16	47,60	48,03	0,8940	2	31	43,92	43,94	0,0384
3	16	11,27	11,99	6,3448	3	31	11,97	11,95	-0,1531
4	16	27,11	27,39	1,0179	4	31	29,39	29,35	-0,1338
5	16	25,28	25,36	0,3177	5	31	25,37	25,10	-1,0661
6	16	48,72	49,20	0,9850	6	31	47,18	47,34	0,3297

LEGENDA: (1) TELES P; (2) PETROBRAS; (3) EMBRAER; (4) BRADESCO; (5) BRASIL; (6) VALE.

Tabela 36 – Tabela com dados para validação – Agosto 2008

EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO	EMPR	DIA	PREDIÇÃO	REAL	VARIAÇÃO
1	1	39,40	39,00	-1,0215	4	15	26,89	27,07	0,6796
2	1	42,85	42,28	-1,3269	5	15	22,28	22,29	0,0261
3	1	11,95	11,57	-3,2075	6	15	41,15	41,4	0,6001
4	1	29,24	29,3	0,2206	1	18	39,67	39,49	-0,4601
5	1	24,23	23,81	-1,7275	2	18	39,47	38,9	-1,4496
6	1	44,47	44,39	-0,1826	3	18	13,19	13,36	1,2614
1	4	38,70	38,85	0,3841	4	18	26,67	27,05	1,4158
2	4	40,72	40,11	-1,4981	5	18	22,28	22,29	0,0573
3	4	11,39	11,66	2,4106	6	18	40,66	40,48	-0,4328
4	4	28,94	28,99	0,1587	1	19	39,66	39,99	0,8275
5	4	23,08	22,58	-2,1615	2	19	40,03	40,04	0,0282
6	4	41,52	41,19	-0,7995	3	19	13,02	13,05	0,2163
1	5	39,01	38,99	-0,0582	4	19	26,67	27,05	1,4393
2	5	39,89	39,65	-0,6071	5	19	22,27	22,29	0,0822
3	5	11,92	12,1	1,5042	6	19	41,10	41,1	-0,0096
4	5	28,93	28,99	0,2149	1	20	39,57	39,8	0,5731
5	5	23,19	23	-0,8011	2	20	41,85	42	0,3499
6	5	41,16	41,63	1,1320	3	20	13,08	13,27	1,4229
1	6	39,64	40,00	0,9072	4	20	26,00	26,20	0,7818
2	6	40,99	40,71	-0,6759	5	20	22,27	22,29	0,1029
3	6	11,85	11,9	0,4191	6	20	43,02	43,36	0,7993
4	6	28,91	28,99	0,2660	1	21	38,84	38,88	0,0931
5	6	23,84	24,15	1,2867	2	21	43,34	43,31	-0,0635
6	6	42,06	42,2	0,3324	3	21	12,92	13,1	1,4149
1	7	39,43	39,01	-1,0601	4	21	26,51	26,69	0,6603
2	7	41,72	41,32	-0,9700	5	21	22,26	22,29	0,1194
3	7	11,46	11,46	0,0162	6	21	44,20	44,2	-0,0038
4	7	28,90	28,99	0,3122	1	22	38,96	39,17	0,5467
5	7	23,86	23,8	-0,2363	2	22	42,99	43,11	0,2800
6	7	41,92	42,4	1,1387	3	22	13,13	13,52	2,9885
1	8	38,60	38,60	-0,0046	4	22	26,63	26,80	0,6212
2	8	40,90	41	0,2470	5	22	22,26	22,29	0,1316
3	8	11,63	11,9	2,2931	6	22	43,44	43,48	0,0957
4	8	28,89	28,99	0,3535	1	25	38,79	38,89	0,2623
5	8	23,54	23,4	-0,5879	2	25	42,15	41,65	-1,1928
6	8	41,60	41,69	0,2255	3	25	13,18	13,17	-0,0897
1	11	38,95	39,40	1,1658	4	25	26,74	26,98	0,8957
2	11	40,36	39,86	-1,2377	5	25	22,25	22,29	0,1696
3	11	11,80	12,05	2,0974	6	25	42,29	42,16	-0,3099
4	11	28,88	28,99	0,3716	1	26	38,43	38,8	0,9621
5	11	23,03	22,57	-1,9892	2	26	42,06	42,01	-0,1112
6	11	40,06	40,15	0,2320	3	26	13,23	13,45	1,6584
1	12	39,02	39,50	1,2404	4	26	26,20	26,30	0,3857
2	12	40,74	40,2	-1,3348	5	26	22,25	22,29	0,1798
3	12	12,34	12,79	3,6251	6	26	42,14	42,2	0,1342
4	12	28,87	28,99	0,4100	1	27	38,42	38,7	0,7283
5	12	22,26	21,8	-2,0500	2	27	42,81	43,2	0,9105
6	12	40,11	39,92	-0,4711	3	27	13,17	13,28	0,7998
1	13	39,48	39,89	1,0404	4	27	25,91	25,94	0,1170
2	13	41,15	41,25	0,2503	5	27	23,86	24,2	1,4073
3	13	12,72	12,86	1,1378	6	27	42,72	43,58	2,0207
4	13	28,86	28,99	0,4437	1	28	38,77	39	0,5886
5	13	21,51	21,2	-1,4220	2	28	43,08	43,3	0,5091
6	13	41,42	41,75	0,7899	3	28	13,43	13,69	1,9673
1	14	40,11	39,98	-0,3224	4	28	26,01	26,39	1,4532
2	14	40,91	40,9	-0,0144	5	28	23,86	24,2	1,4072
3	14	12,92	13,25	2,5410	6	27	42,72	43,8	2,5357
4	14	28,85	28,99	0,4727	1	29	38,85	38,82	-0,0803
5	14	21,60	21,8	0,9147	2	29	43,03	42,66	-0,8640
6	14	42,31	42,2	-0,2632	3	29	13,63	13,83	1,4757
1	15	39,68	40,00	0,7949	4	29	26,84	26,94	0,3779
2	15	40,00	40,05	0,1232	5	29	23,73	23,79	0,2516
3	15	13,08	13,35	2,0513	6	29	43,34	43,47	0,2903

LEGENDA: (1) TELES P; (2) PETROBRAS; (3) EMBRAER; (4) BRADESCO; (5) BRASIL; (6) VALE.

APÊNDICE E – Desenho das redes neurais artificiais

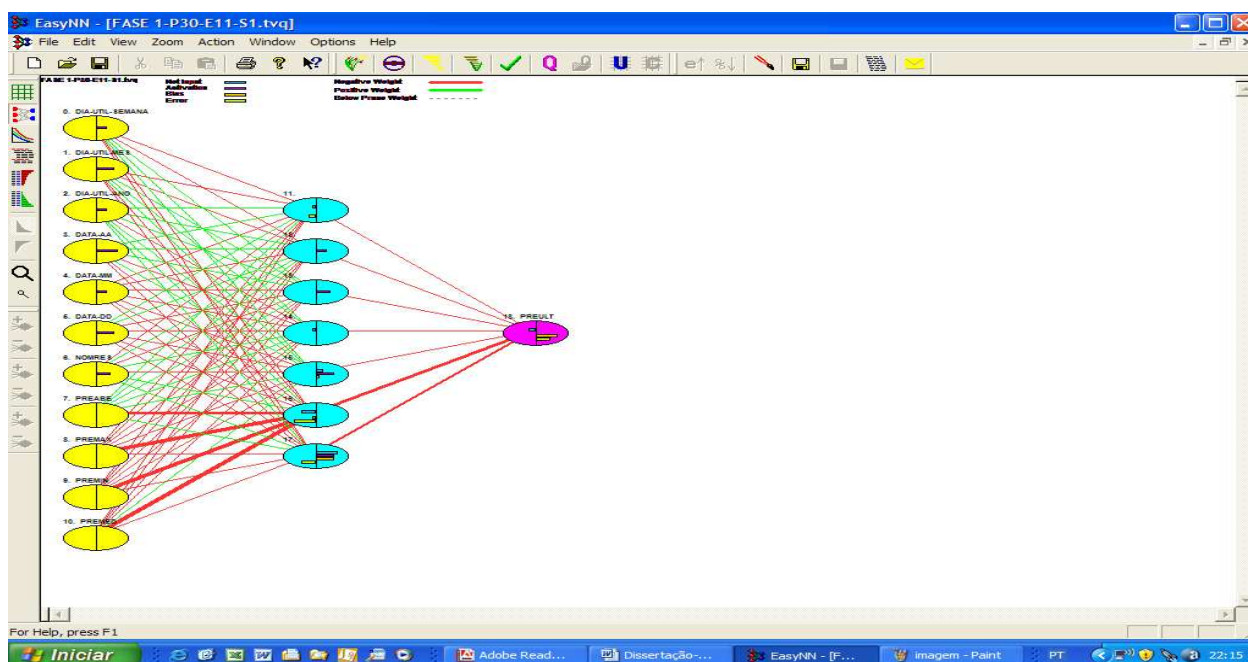


Figura 18 - Rede 1

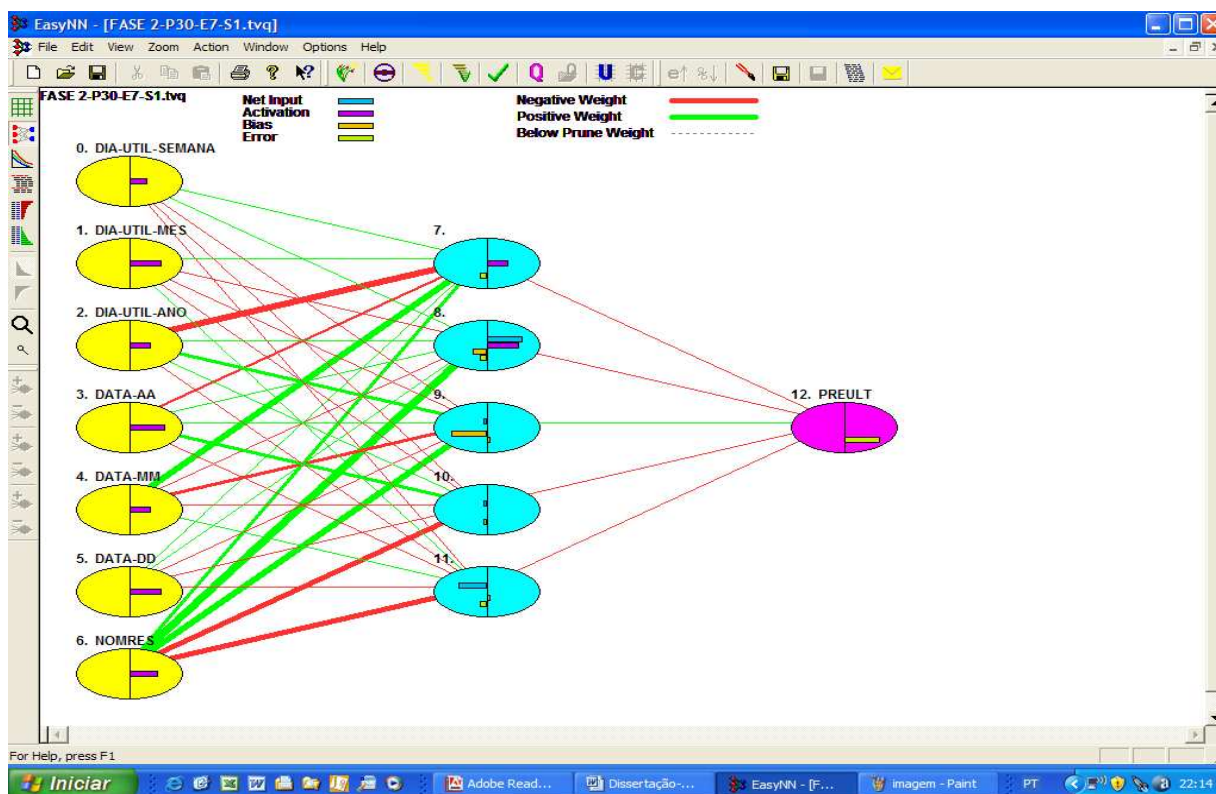


Figura 19 - Rede 2

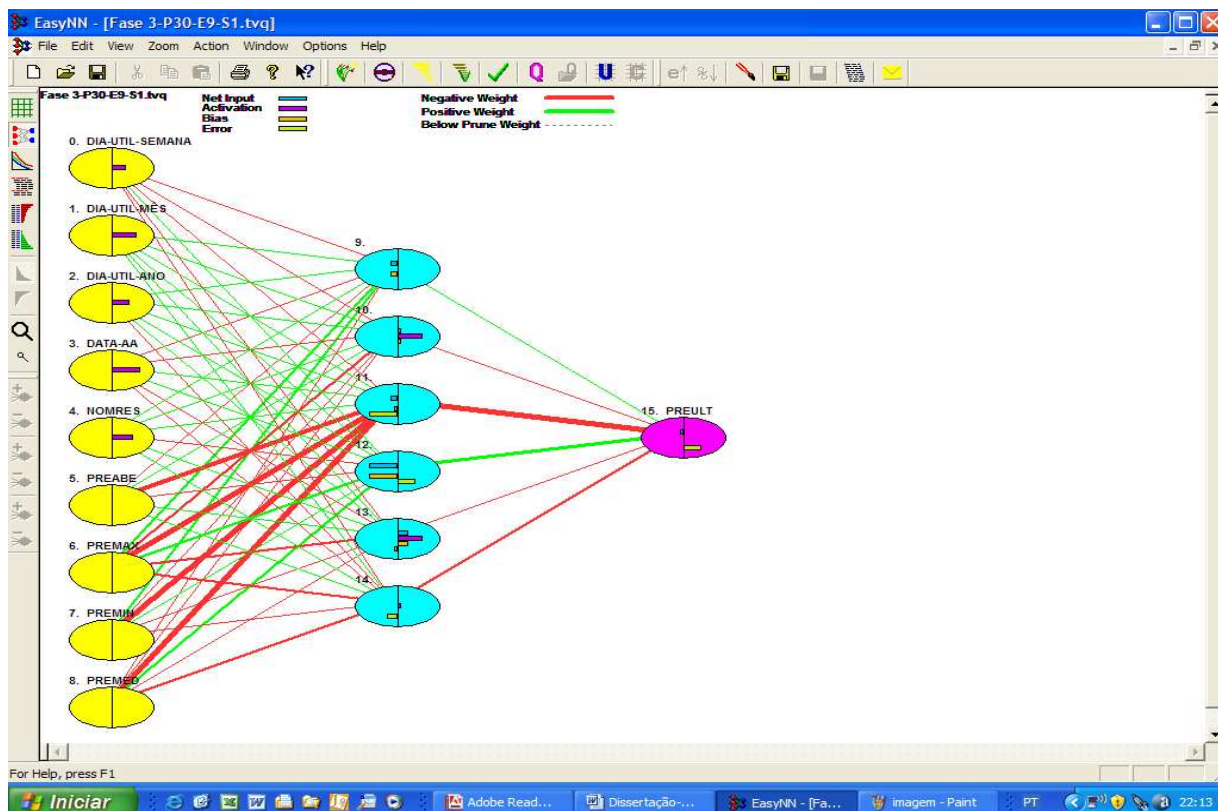


Figura 20 - Rede 3

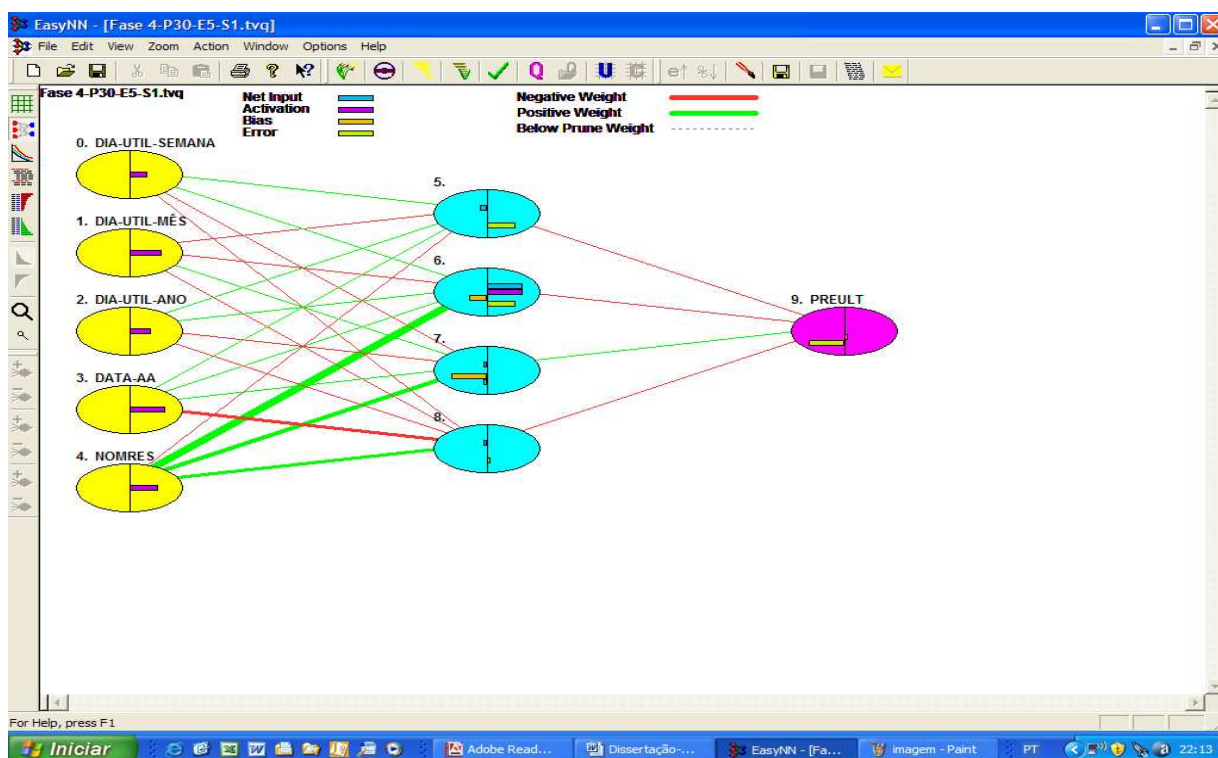


Figura 21 - Rede 4

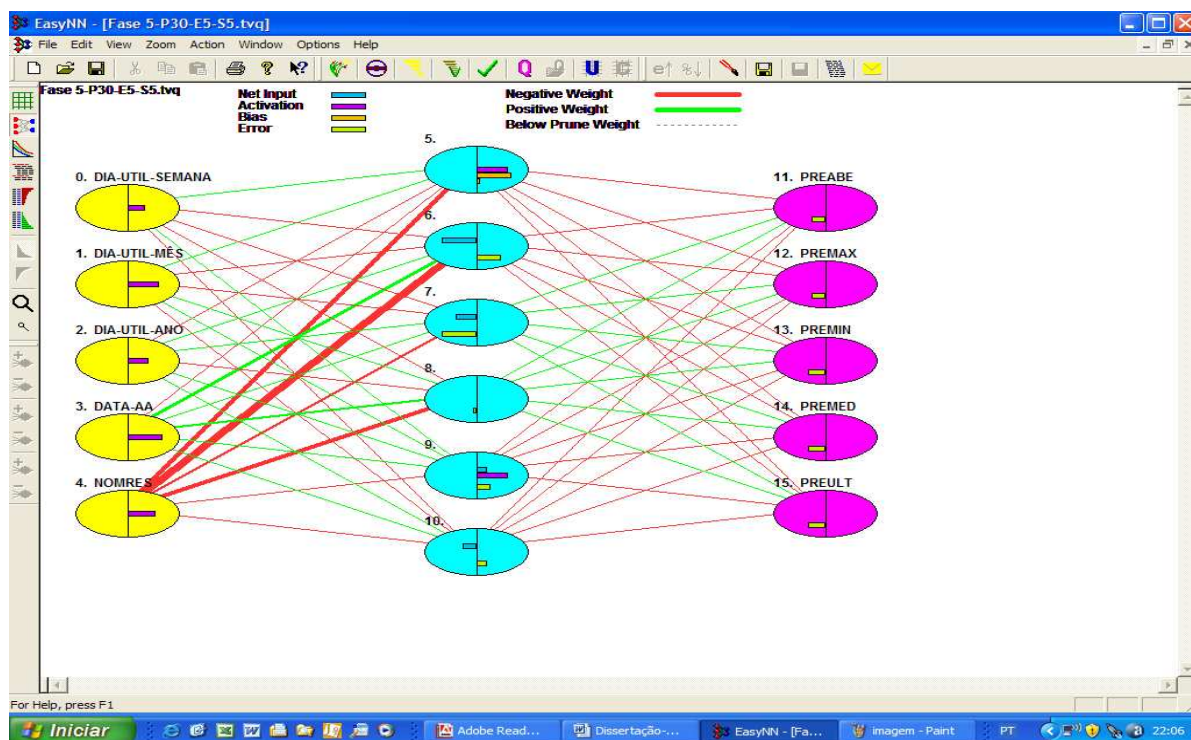


Figura 22 - Rede 5

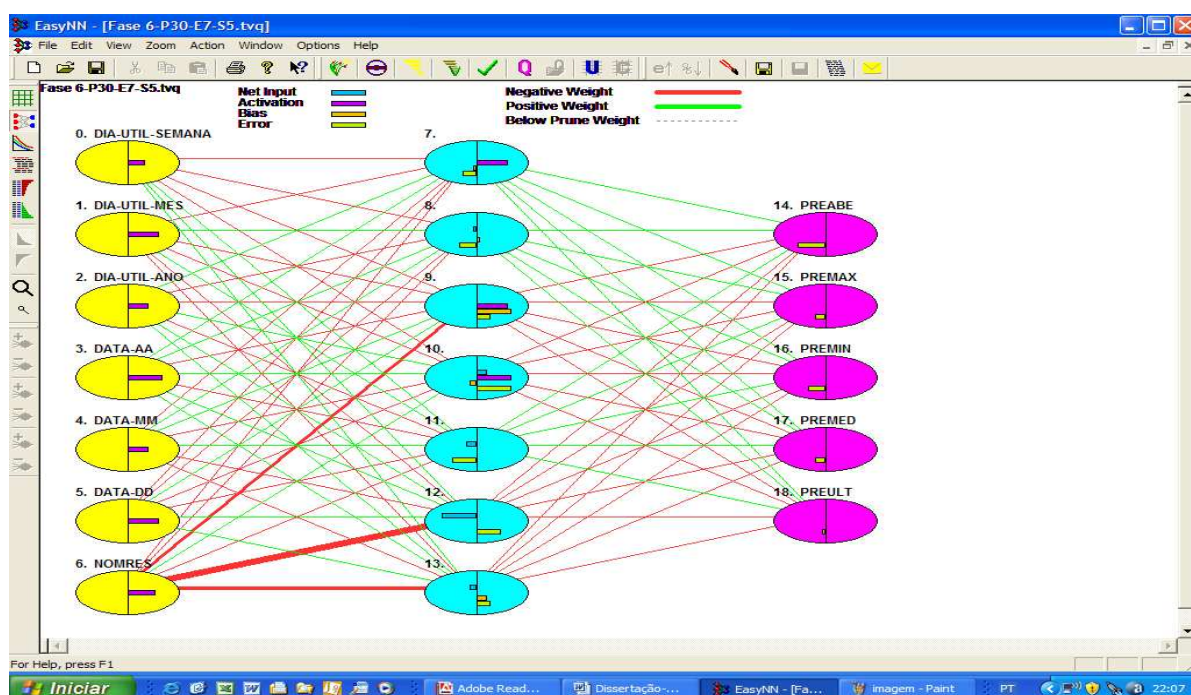


Figura 23 - Rede 6

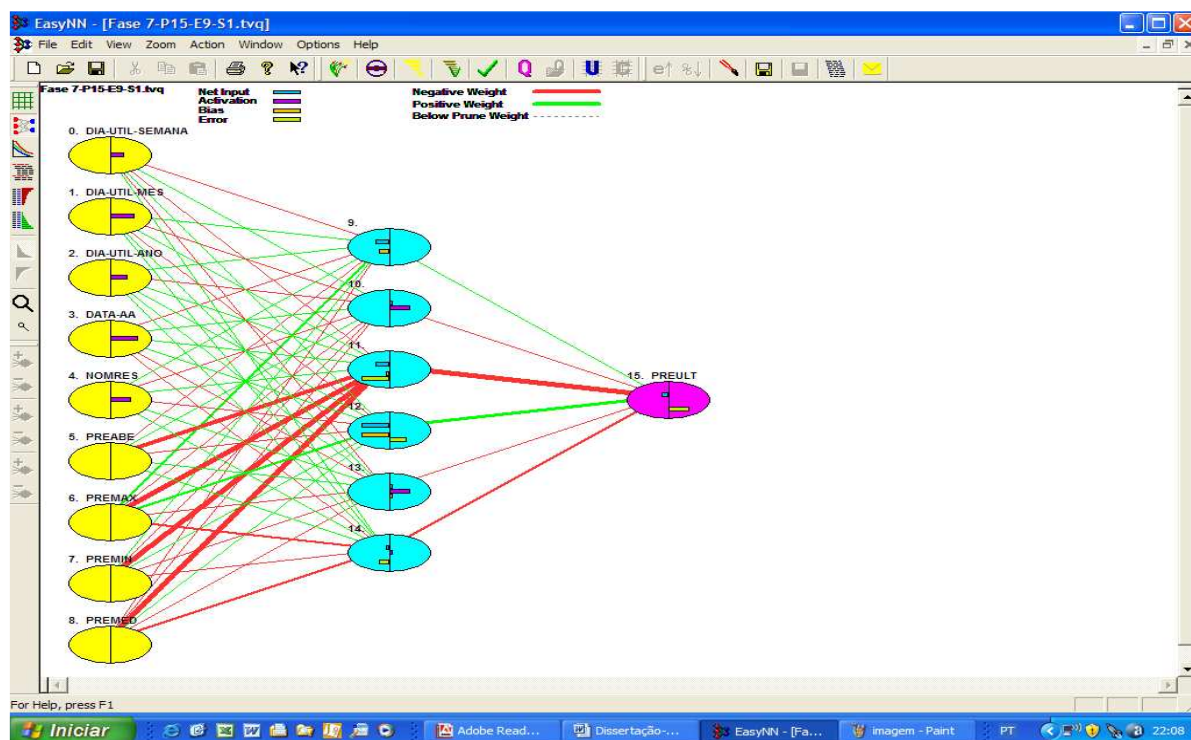


Figura 24 - Rede 7

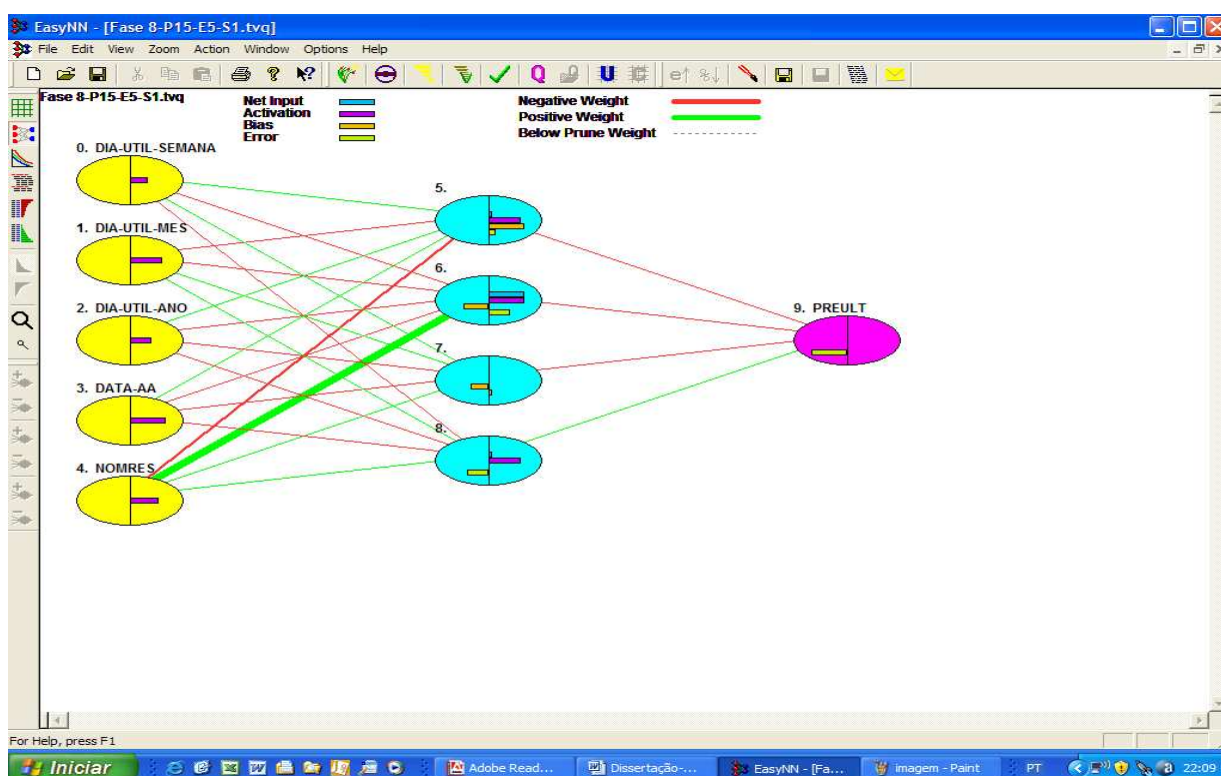


Figura 25 - Rede 8

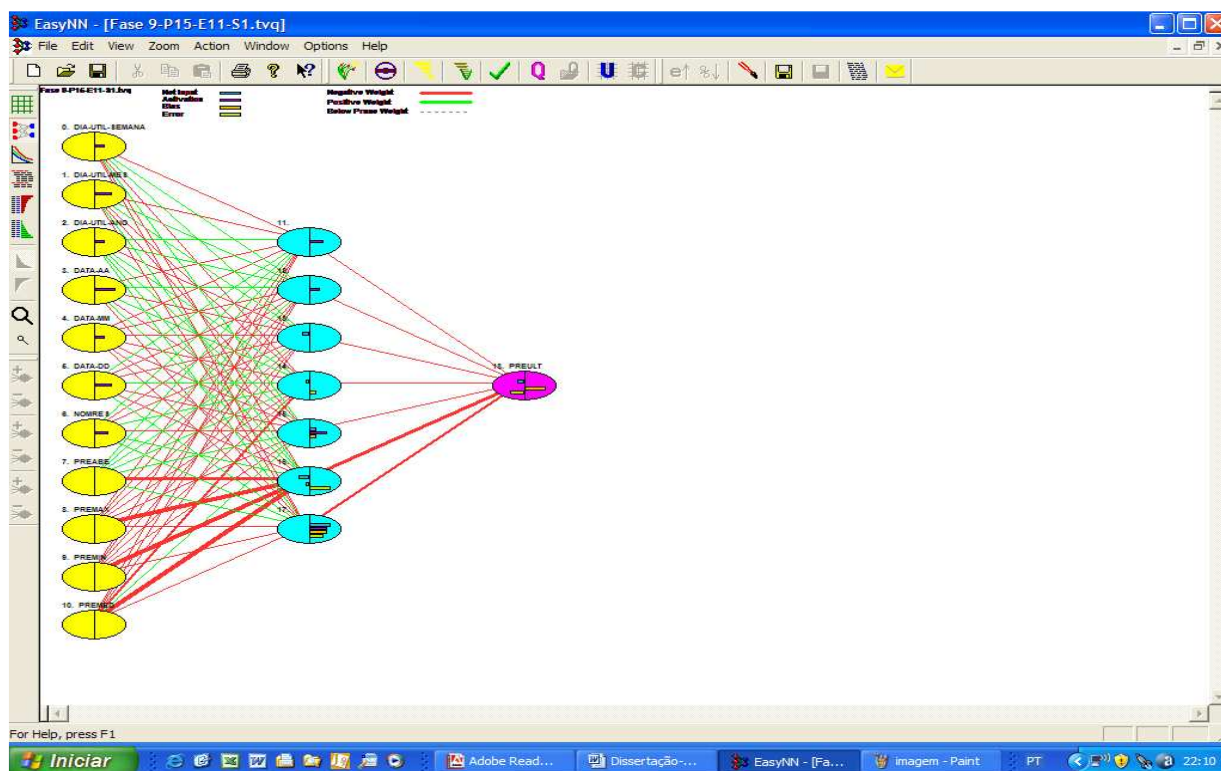


Figura 26 - Rede 9

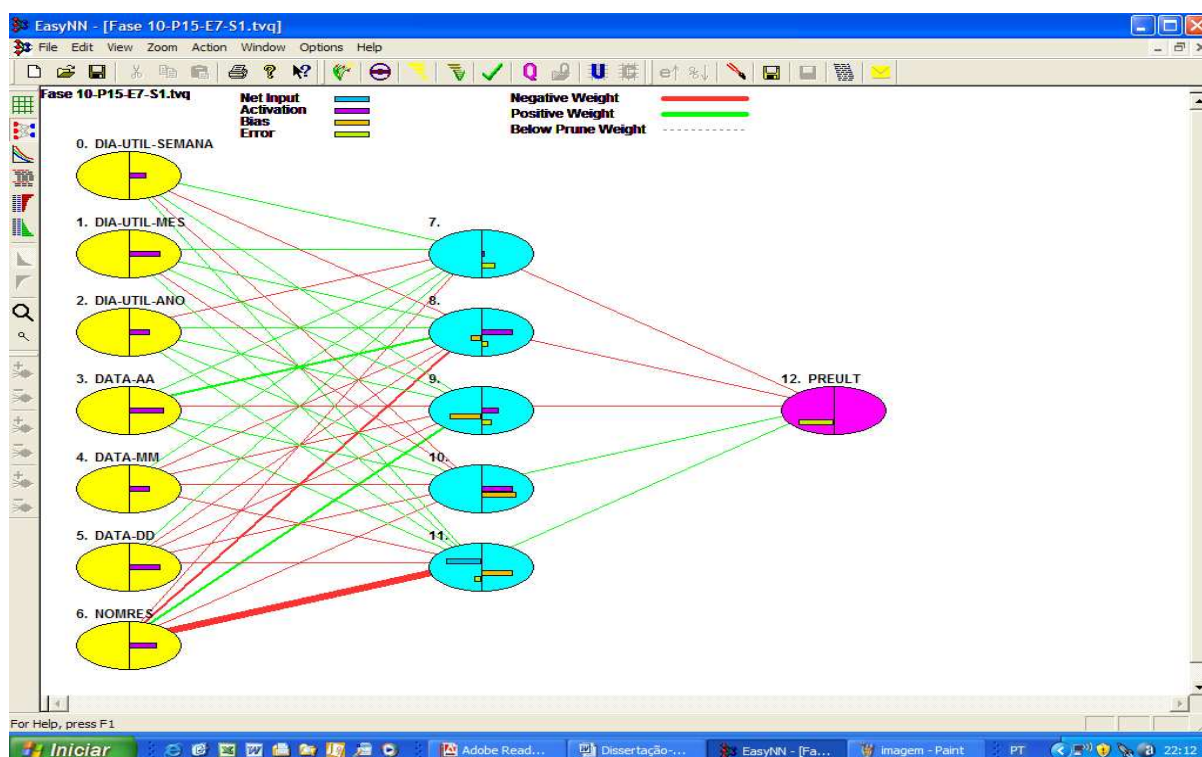


Figura 27 - Rede 10

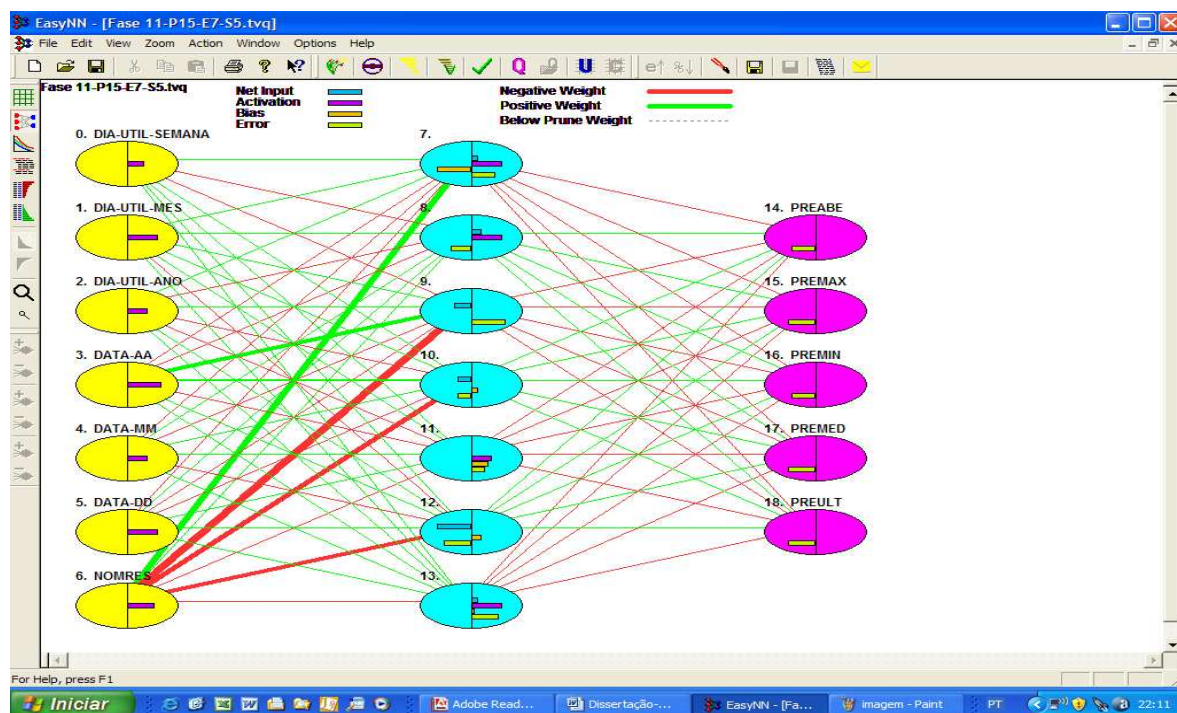


Figura 28 - Rede 11

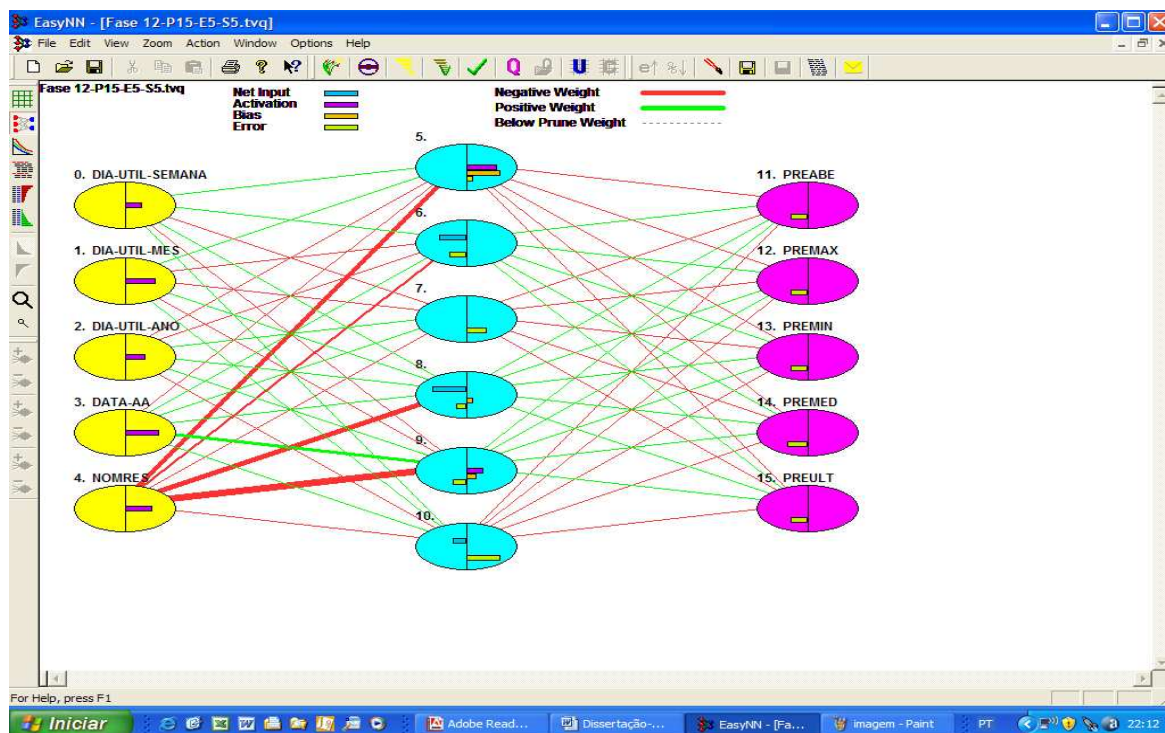


Figura 29 - Rede 12