

**UNIVERSIDADE FEDERAL DE ALFENAS
INSTITUTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

José Alexandre Macedo

**MÁQUINA DE COMITÊ PARA A PREVISÃO DO PREÇO
MÉDIO DIÁRIO DAS AÇÕES DA BOLSA DE VALORES DO
ESTADO DE SÃO PAULO**

Alfenas, 01 de Julho de 2010.

UNIVERSIDADE FEDERAL DE ALFENAS
INSTITUTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

**MÁQUINA DE COMITÊ PARA A PREVISÃO DO PREÇO
MÉDIO DIÁRIO DAS AÇÕES DA BOLSA DE VALORES DO
ESTADO DE SÃO PAULO**

José Alexandre Macedo

Monografia apresentada ao Curso de Bacharelado em
Ciência da Computação da Universidade Federal de
Alfenas como requisito parcial para obtenção do Título de
Bacharel em Ciência da Computação.

Orientador: Prof. Humberto César Brandão de Oliveira

Alfenas, 01 de Julho de 2010.

José Alexandre Macedo

**MÁQUINA DE COMITÊ PARA A PREVISÃO DO PREÇO
MÉDIO DIÁRIO DAS AÇÕES DA BOLSA DE VALORES DO
ESTADO DE SÃO PAULO**

A Banca examinadora abaixo-assinada aprova a monografia apresentada como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação pela Universidade Federal de Alfenas.

Prof. Dr. Marcos de Mendonça Passini

Universidade Federal de Alfenas

Prof. Dr. Marcelo Lacerda Rezende

Universidade Federal de Alfenas

Prof. Humberto César Brandão de Oliveira

Universidade Federal de Alfenas

Alfenas, 01 de Julho de 2010.

Dedico este trabalho a minha mãe,
A meu pai,
A minha irmã e
A Lorena.

AGRADECIMENTO

Agradeço a Deus, em primeiro lugar, por sempre estar direcionando e abençoando a minha vida, me colocando todo tempo nas horas certas nos lugares certos.

Agradeço a minha mãe Angélica, meu pai José Reinaldo e minha irmã Bianca por todo incentivo e motivação durante esta jornada longa que está apenas no começo.

Agradeço a Lorena, por ser essa pessoa tão especial, carinhosa e amorosa que torna minha vida mais feliz desde o começo desta Faculdade.

Agradeço ao meu orientador Humberto, pela sua contribuição com uma pequena parcela de seu enorme conhecimento compartilhado comigo durante este curso.

Agradeço a minha tia Adriana, por ajudar a tornar possível a conclusão deste curso.

Agradeço aos meus companheiros de república Danilo, Luís Theodoro e Neubio, que foram muito mais do que companheiros nestes anos de Faculdade e tornaram-se verdadeiros amigos que espero levar por toda vida.

Agradeço a todos os demais, que de alguma forma contribuíram para o meu crescimento, minha educação e para a realização deste trabalho.

"O mistério não é um muro onde a inteligência esbarra, mas um oceano onde ela
mergulha."

Gustav Thibon

RESUMO

O mercado de ações brasileiro, representado pela BOVESPA, é uma fonte de renda variável que está em crescimento no Brasil onde, com cautela e estratégias corretas, é possível obter bons lucros. Entretanto, sua natureza complexa e a falta de conhecimento podem gerar prejuízos irre recuperáveis. Com isso, estratégias que auxiliem as negociações realizando previsões dos preços das ações podem ajudar as pessoas que possuem dificuldade na análise de mercado, além de possibilitar lucros maiores com a utilização de algoritmos de Inteligência Artificial que usam dados passados do mercado acionário para prever os preços futuros das ações. Esta monografia tem como objetivo verificar se a utilização de máquinas de comitê para a previsão do preço médio diário das ações, que combinam os resultados de diversos modelos de previsão, geram resultados melhores ou no mínimo tão bons quanto os resultados dos melhores previsores. Além disso, esta monografia define um sistema *online* para apoio à decisão ao investidor que pode ser utilizado com qualquer modelo ou técnica computacional de previsão para apresentar aos investidores sugestões de compra e venda diárias de ação e para tornar possível o acompanhamento dos lucros dos papéis que o investidor tenha em sua carteira.

Palavras-Chave: máquina de comitê, mercado de ações, inteligência artificial, sistema *online*.

ABSTRACT

The Brazilian stock market, represented by BOVESPA, is a variable income source which is increasing in Brazil, allowing high profits, since one has caution and right strategies. However, its complex nature and the lack of knowledge can lead to irretrievable economic losses. Then, strategies which help negotiations by predicting stock prices can be useful to people who has difficulty in market analysis, besides allowing higher profits by using Artificial Intelligence algorithms which use past data from the stock market to predict future stock prices. The objective of this monograph is to verify whether the utilization of committee machines to predict the average daily market prices, which combine the results of several prediction models, generates better (or at least as good as) results when compared to those obtained by the best predictors. Moreover, this monograph defines an online decision support system for investors, which can be used with any prediction model or computer technique to present suggestions of daily stock buying and selling suggestions to the investors, and to allow them to follow profits of the papers in their stock book.

Keywords: comittee machine, stock market, artificial intelligence, online system.

LISTA DE FIGURAS

FIGURA 1 - HISTÓRICO DO ÍNDICE BOVESPA	33
FIGURA 2 - EXEMPLO DE GRÁFICO <i>CANDLESTICK</i>	36
FIGURA 3 - REPRESENTAÇÃO ATRAVÉS DE ÁRVORE SINTÁTICA DA FUNÇÃO $\text{SEN}(X/Y - 8)$	44
FIGURA 4 - ESTRUTURA DO ALGORITMO DE PROGRAMAÇÃO GENÉTICA	45
FIGURA 5 - MUTAÇÃO EM UMA SUB-ÁRVORE.....	48
FIGURA 6 - CRUZAMENTO DE DOIS INDIVÍDUOS.....	48
FIGURA 7 - NEURÔNIO ARTIFICIAL.....	49
FIGURA 8 - EXEMPLO DE REDE COM MÚLTIPLAS CAMADAS	50
FIGURA 9 - ESTRUTURA BÁSICA DA MÁQUINA DE COMITÊ.....	52
FIGURA 10 - FLUXOGRAMA DE EXECUÇÃO DA METODOLOGIA.....	58
FIGURA 11 - SÉRIES DOS VALORES MÉDIOS DAS AÇÕES.....	60
FIGURA 12 - ESQUEMA DA MÁQUINA DE COMITÊ	62
FIGURA 13 - FLUXOGRAMA DA EXECUÇÃO DOS ALGORITMOS PARA CADA PAPEL.....	64
FIGURA 14 - GRÁFICO DOS RESULTADOS PARA O PAPEL BBDC4.....	69
FIGURA 15 - GRÁFICO DOS RESULTADOS PARA O PAPEL ITUB4	70
FIGURA 16 - GRÁFICO DOS RESULTADOS PARA O PAPEL PETR4.....	71
FIGURA 17 - GRÁFICO DOS RESULTADOS PARA O PAPEL USIM5	72
FIGURA 18 - GRÁFICO DOS RESULTADOS PARA O PAPEL VALE5.....	73
FIGURA 19 - TELA DE CADASTRO DO INVESTIDOR	81
FIGURA 20 - CADASTRO DE NOVOS PAPÉIS.....	81
FIGURA 21 - TELA DE GERENCIAMENTO DA CARTEIRA	82
FIGURA 22 - TELA DE AGENDAMENTO DE ANÁLISE	83
FIGURA 23 - TELA DE VISUALIZAÇÃO DAS ANÁLISES AGENDADAS.....	83
FIGURA 24 - TELA DE VISUALIZAÇÃO DAS ANÁLISES AGENDADAS.....	84
FIGURA 25 - TELA DE VISUALIZAÇÃO DAS ANÁLISES PASSADAS.....	85
FIGURA 26 - TELA DE ANÁLISE DOS ÚLTIMOS 30 DIAS.....	85
FIGURA 27 - TELA DE SIMULAÇÕES DE LUCRO.....	86
FIGURA 28 - TELA DE CONFIGURAÇÕES	87
FIGURA 29 - TELA PARA CONTATO.....	88
FIGURA 30 - DIAGRAMA DE CLASSES DO SISTEMA.....	90
FIGURA 31 - DIAGRAMA DE ENTIDADE E RELACIONAMENTO DO SISTEMA	91

LISTA DE TABELAS

TABELA 1- CUSTOS PARA TAXA DE CORRETAGEM VARIÁVEL	38
TABELA 2 - EMOLUMENTO COBRADO DE OPERAÇÕES NORMAIS	38
TABELA 3 - EMOLUMENTO COBRADO DE OPERAÇÕES <i>DAY TRADE</i>	38
TABELA 4 - INFORMAÇÕES ESTATÍSTICAS DO CONJUNTO DE DADOS DE 2009	60
TABELA 5 - INFORMAÇÕES DOS RESULTADOS PARA BRADESCO	69
TABELA 6 - INFORMAÇÕES DOS RESULTADOS PARA ITAU-UNIBANCO	71
TABELA 7 - INFORMAÇÕES DOS RESULTADOS PARA PETROBRAS	72
TABELA 8 - INFORMAÇÕES DOS RESULTADOS PARA USIMINAS.....	73
TABELA 9 - INFORMAÇÕES DOS RESULTADOS PARA VALE DO RIO DOCE.....	74
TABELA 10 - TESTE DE TUKEY ITUB4	75
TABELA 11 - TESTE DE TUKEY BBDC4.....	75
TABELA 12 - TESTE DE TUKEY PETR4	76
TABELA 13 - TESTE DE TUKEY USIM5	76
TABELA 14 - TESTE DE TUKEY VALE5.....	76

LISTA DE ABREVIACÕES

BOVESPA	Bolsa de Valores do Estado de São Paulo
CE	Computação Evolucionária
PG	Programação Genética
RL	Regressão Linear
RNA	Redes Neurais Artificiais
MLP	<i>Multi-Layer Perceptron</i>
MVS	Máquinas de Vetor de Suporte
EMQ	Erro Médio Quadrático
SSE	Soma dos Erros Quadráticos

SUMÁRIO

1 INTRODUÇÃO	25
1.1 JUSTIFICATIVA E MOTIVAÇÃO	26
1.2 PROBLEMATIZAÇÃO.....	27
1.3 HIPÓTESES.....	27
1.4 OBJETIVOS.....	28
1.4.1 Gerais	28
1.4.2 Específicos	28
1.5 ORGANIZAÇÃO DA MONOGRAFIA.....	29
2 MERCADO DE AÇÕES.....	31
2.1 INTRODUÇÃO AO MERCADO.....	31
2.2 OPERAÇÕES.....	33
2.3 MODELOS DE ANÁLISE DE MERCADO.....	34
2.3.1 Análise Fundamentalista.....	35
2.3.2 Análise Técnica	35
2.3.2.1 Indicadores Financeiros	36
2.4 CUSTOS OPERACIONAIS.....	37
2.5 SUMÁRIO DO CAPÍTULO	39
3 REFERENCIAL TEÓRICO.....	41
3.1 PREVISÃO DE SÉRIES TEMPORAIS	41
3.1.1 Fatores que influenciam as séries temporais	42
3.1.1.1 Sazonalidades.....	42
3.1.1.1 Tendências	42
3.1.1.1 Ciclos	43
3.1.1.1 Aleatoriedade.....	43
3.1.2 Métodos para Previsão de Séries Temporais	43
3.1.2.1 Programação Genética.....	44
3.1.2.1 Criação da População Inicial.....	46
3.1.2.1 Avaliação dos Indivíduos	46
3.1.2.1 Métodos de Seleção.....	47
3.1.2.1 Operadores Genéticos	47
3.1.2.1 Redes Neurais Artificiais	49
3.1.2.1 Regressão Linear	51
3.2 MÁQUINAS DE COMITÊ	52
3.3 REVISÃO DE LITERATURA	53
3.4 SUMÁRIO DO CAPÍTULO	55
4 PROJETO.....	57
4.1 PROJETO	57
4.2 METODOLOGIA	58
4.2.1 Análise de Dados.....	58
4.2.1.1 Definição dos dados utilizados	58
4.2.1.1 Conjunto de Dados	59
4.2.2 Normalização dos Dados	61
4.2.3 Aquisição dos Modelos de Previsão	61

4.2.4 Métricas de criação das Máquinas de comitê.....	61
4.2.5 Tipos de Resultados Gerados.....	63
4.2.6 Execução dos Algoritmos	63
4.2.7 Testes Estatísticos	64
4.3 SUMÁRIO DO CAPÍTULO	65
5 RESULTADOS	67
5.1 CONFIGURAÇÕES DAS INSTÂNCIAS DOS MODELOS DE PREVISÃO.....	67
5.1.1 Configurações do Algoritmo de PG.....	67
5.1.2 Configurações do Algoritmo de RNA	68
5.2 ANÁLISE DOS RESULTADOS	68
5.3 RESULTADOS DOS TESTES ESTATÍSTICOS.....	74
6 SISTEMA ONLINE DE APOIO A DECISÃO AO INVESTIDOR.....	79
6.1 OBJETIVOS DO SISTEMA.....	79
6.2 DESCRIÇÃO DOS USUÁRIOS	80
6.3 FUNCIONALIDADES.....	80
6.3.1 Funcionalidades do Investidor	80
6.3.1.1 Cadastro.....	80
6.3.1.1 Gerenciamento da Carteira	81
6.3.1.1 Gerenciamento de Análises Futuras	82
6.3.1.1 Visualização das previsões do dia atual.....	84
6.3.1.1 Visualização do Histórico de previsões.....	84
6.3.1.1 Análise dos últimos 30 dias.....	85
6.3.1.1 Simulações.....	86
6.3.1.1 Configurações.....	86
6.3.1.1 Contato	87
6.3.2 Funcionalidades do Administrador	88
6.3.2.1 Gerenciar Usuários.....	88
6.3.2.1 Rentabilidade dos Papéis	89
6.3.2.1 Melhores Investidores.....	89
6.3.2.1 Acessos ao Sistema	89
6.4 DOCUMENTAÇÃO DE CLASSES	90
6.5 DOCUMENTAÇÃO DO BANCO DE DADOS	91
6.6 FUNCIONAMENTO DO SISTEMA.....	92
6.7 DISCUSSÃO FINAL DO CAPÍTULO	94
7 CONCLUSÕES	95
7.1 CONCLUSÕES DOS EXPERIMENTOS	95
7.2 CONCLUSÕES DO SISTEMA <i>ONLINE</i>	96
8 REFERÊNCIAS BIBLIOGRÁFICAS	97

1

Introdução

Este capítulo apresenta uma introdução ao mercado de ações e aos objetivos desta monografia. Na Seção 1.1 é apresentada a Justificativa que levou a escrita deste trabalho. Na Seção 1.2 é descrita a problematização que se deseja responder, em seguida na Seção 1.3 a hipótese é definida. Na Seção 1.4 os objetivos gerais e específicos são apresentados e por fim na Seção 1.5 a organização da monografia é apresentada.

O mercado de ações é uma fonte de renda variável de natureza instável e complexa, que é influenciada por inúmeros fatores sejam estes, políticos, econômicos, sociológicos ou psicológicos. Estes fatores levam os investidores a estarem sempre procurando estratégias ou mecanismos para reduzir as perdas financeiras. A previsão dos valores das ações deste mercado é uma das estratégias para reduzir estas perdas e auxiliar os investidores na tomada de decisão de compra e venda destas ações.

Entre as estratégias de análise para previsões mais comuns, estão, a fundamentalista e a técnica (Piazza, 2009). A estratégia fundamentalista se baseia no contexto socioeconômico das empresas onde os dados sobre as empresas são obtidos das mais variadas fontes, a fim de tentar determinar o valor da ação desta no mercado. Enquanto a análise técnica utiliza como base apenas as informações históricas das ações para tentar prever o valor futuro destas. Contudo, a utilização de técnicas computacionais pode auxiliar na análise dessas informações de forma rápida, precisa e inteligente. A partir disso, muitas abordagens que utilizam técnicas de Inteligência Artificial vêm sendo desenvolvidas para auxiliar na previsão das séries temporais do mercado de ações para que seja possível indicar o valor futuro das ações neste mercado. Uma das estratégias utilizadas é a combinação de vários modelos de previsão chamada de máquina de comitê, como foi apresentado por Castro; Braga & Andrade (2005). A estratégia é também

conhecida por *ensemble*. Com esta técnica é possível aproveitar o bom desempenho de cada predictor para certos conjuntos de dados, sendo possível assim, melhorar os resultados finais.

O mercado de ações possui várias informações que são passíveis de extrapolação, como por exemplo, o valor do índice BOVESPA, o valor de abertura, o valor de fechamento e o valor médio das ações, sendo este último utilizado neste trabalho. O valor médio diário das ações diz respeito à média aritmética entre o valor máximo e mínimo que o preço de uma ação atinge ao longo do dia. A sua previsão pode auxiliar o investidor a definir o valor de compra e venda para as ações que deseja negociar. Com isso o desenvolvimento de previsores capazes de calcular o valor médio das ações com uma baixa taxa de erro pode facilitar a negociação das ações sendo usado como um indicador que auxilia no processo de tomada de decisão.

1.1 Justificativa e Motivação

Os sete primeiros meses do ano de 2009 foram marcados pela maior valorização do mundo, tratando-se de mercados acionários, alcançada pela BOVESPA, permitindo a seus investidores atingirem ganhos de até 87% no período (Exame, 2009). Esta informação mostra que o Brasil está economicamente mais estável, pois, teve grande facilidade para sair da crise econômica mundial de 2008, iniciada pela falência do banco de investimentos norte americano Lehman Brothers, que causou grande instabilidade financeira e devido a qual muitos países ainda não se recuperaram.

Ainda em 2009, a BOVESPA se tornou a quarta maior do mundo em valor de mercado e a 12ª somando o valor das ações de todas as companhias negociadas, tendo um valor total avaliado em 1 trilhão de dólares (Exame, 2009). Este *status* reforça a posição do Brasil como grande polo financeiro da América Latina e um dos principais dos países emergentes, refletindo assim o novo patamar econômico do país.

Estas informações são importantes, pois, mostram que a BOVESPA está se tornando uma opção atrativa de investimentos para os brasileiros que estão acostumados a utilizar a renda fixa. Com isso, novas técnicas e ferramentas que auxiliem na tomada de decisão da compra e venda de ações podem facilitar as negociações para o investidor, levando em conta ainda o potencial número de novos investidores que estão surgindo devido a estabilidade econômica do país.

1.2 Problematização

A fim de auxiliar os investidores financeiros na tomada de decisão da compra e venda de ações, surge à necessidade de responder a seguinte pergunta: Será que a utilização de máquinas de comitê para previsão dos preços médios diários das ações gera erros menores ou iguais aos menores erros gerados por diversos previsores isoladamente?

1.3 Hipóteses

A hipótese H_0 foi definida de forma que “as máquinas de comitê garantem erros menores ou iguais aos menores erros de vários previsores isolados do preço médio diário das ações”, enquanto a hipótese H_1 define que “as máquinas de comitê não garante erros menores ou iguais aos menores erros de vários previsores isolados do preço médio diário das ações”.

1.4 Objetivos

1.4.1 Gerais

Utilizar máquinas de comitê através dos modelos de previsão Programação Genética, Redes Neurais Artificiais e Regressão Linear para verificar se os resultados obtidos são melhores ou iguais aos melhores resultados dos modelos separadamente. Além disso, com base nos resultados alcançados, disponibilizar um sistema *online* de apoio à decisão ao investidor na Bolsa de Valores.

1.4.2 Específicos

- Estudar o mercado financeiro e suas formas de análise;
- Estudar Programação Genética, Redes Neurais Artificiais e Regressão Linear;
- Desenvolver os algoritmos para os modelos definidos;
- Gerar máquinas de comitê com os resultados dos algoritmos;
- Definir e executar testes estatísticos com os resultados obtidos;
- Analisar os requisitos do sistema de apoio à decisão do investidor;
- Escolher e estudar tecnologias a serem utilizadas pelo sistema;
- Implementar e divulgar o sistema *online*;
- Publicar artigos com os resultados obtidos.

1.5 Organização da Monografia

Esta monografia está organizada da seguinte forma: no Capítulo 2 é realizada uma introdução ao mercado de ações, apresentando suas características, os tipos de operações permitidas, além da explicação dos modelos de análise, Fundamentalista e Técnica, terminando o capítulo explicando os custos para se operar neste mercado. O Capítulo 3 apresenta o referencial teórico deste trabalho, separado em três partes, onde na primeira é apresentado o referencial teórico da previsão de séries temporais, englobando os modelos utilizados neste trabalho, na segunda parte, é apresentado o referencial teórico das máquinas de comitê e por fim são apresentados os trabalhos encontrados na literatura referentes a previsão de séries temporais no mercado de ações. O Capítulo 4 descreve a metodologia utilizada no desenvolvimento desta monografia, contemplando a análise dos dados, a normalização escolhida, a construção dos modelos de previsão e as configurações das máquinas de comitê. Em seguida é descrito como os resultados são gerados, como os algoritmos são executados e por fim qual o teste estatístico escolhido. O Capítulo 5 apresenta os resultados dos experimentos realizados e a análise do teste estatístico usado. O Capítulo 6 descreve o sistema *online* de apoio a decisão ao investidor desenvolvido com a intenção de ser um sistema independente do modelo ou técnica de previsão utilizado. O Capítulo 7 mostra a conclusão de acordo com os resultados atingidos e com o sistema *online* desenvolvido.

2

Mercado de Ações

Este capítulo descreve o Mercado de Ações e algumas de suas peculiaridades. Na Seção 2.1 é feita uma introdução ao Mercado esclarecendo o que são ações e alguns termos comuns a este ambiente. A Seção 2.2 mostra os tipos de operações possíveis para a negociação de ações. Os modelos de análise do Mercado de Ações, Fundamentalista e Técnica, são descritos na Seção 2.3 e os custos operacionais para as operações de compra e venda são apresentados na Seção 2.4.

2.1 Introdução ao Mercado

O mercado de ações é uma importante opção de investimento para pessoas e instituições interessadas em formas de renda variáveis. Baseia-se na negociação de valores mobiliários com o objetivo de proporcionar liquidez aos títulos emitidos por empresas, sendo uma forma de captação de recursos que permite o desenvolvimento destas, contribuindo assim, para o progresso do país e na geração de novos empregos (Mcapitais, 2008).

No Brasil, o mercado de ações é representado pela Bolsa de Valores do Estado de São Paulo (BOVESPA), responsável por proporcionar todas as condições e sistemas indispensáveis para as negociações de compra e venda de valores mobiliários de forma transparente. Uma das importâncias das bolsas de valores advém do seu “objetivo de capitalismo moderno: o estímulo à poupança do grande público e ao investimento em empresas em expansão, que, diante deste apoio, poderão assegurar as condições para seu desenvolvimento.” (Mcapitais, 2008, p. 14).

Os valores mobiliários negociados no mercado de ações são comumente chamados de ações, as quais se podem definir como sendo “um título negociável que representa uma fração mínima do capital social de uma empresa de capital aberto, ou seja, de uma empresa do tipo S.A. (Sociedade Anônima)” (Piazza, 2009,

p. 21). Ao comprar ações o investidor está adquirindo uma pequena parcela da empresa, tornando-se assim um sócio desta.

As ações podem ser divididas em dois grupos, as preferenciais, com rendimentos periódicos, emitidas por empresas do tipo, públicas, compradoras em transações de fusão ou sofrendo perdas e precisando de financiamento adicional. Estas ações não dão direito a voto e quem as possui tem preferência nas distribuições de lucros em caso de falência, e as ações ordinárias, onde apesar de não possuir garantia do recebimento do fluxo de dinheiro, quem as possui são os verdadeiros proprietários da empresa, pois estas ações dão o direito a voto, além disso, seus donos estão assegurados de que não poderão perder mais do que investiram na empresa (Gitman, 2006).

A compra e venda de ações é realizada atualmente através do *home broker*, um serviço semelhante ao *home banking*, através do qual é possível executar diversas operações na Bolsa de Valores de forma prática e instantânea. O seu surgimento possibilitou que os pequenos investidores se tornassem mais presentes neste mercado de renda variável, um dos motivos para isso é a dispensa da necessidade de intermediadores para o fechamento das negociações (Piazza, 2009), pois estes pequenos investidores ficavam em segundo plano perto de bancos e instituições que devido aos grandes volumes negociados sempre tinham a preferência do corretor que só atendia um cliente por vez ao telefone.

Com o passar do tempo surgiu a necessidade de representar o comportamento do mercado em determinados períodos de tempo. Para isso, alguns investidores criaram os índices, que determinam o desempenho médio dos preços de uma carteira hipotética. A principal informação destes índices é a lucratividade, baseada nas variações ocorridas entre um período e outro. Os índices refletem o comportamento dos investimentos em ações e conseqüentemente as tendências gerais da economia (Assaf Neto, 2006).

Um investidor, para conhecer o desempenho médio das ações mais estáveis, seguras e de alta liquidez (*blue chips*) do mercado de ações brasileiro, utiliza o Índice Bovespa (Ibovespa). Sua importância advém do fato de que este retrata o comportamento dos principais papéis negociados na BOVESPA mostrando, assim, o perfil dos negócios realizados e também de sua tradição. Este índice manteve a

integridade de sua série histórica e não sofreu modificações metodológicas desde sua implementação em 1968 (Ibove, 1999).

A Figura 1 exemplifica um gráfico com as cotações diárias do Ibovespa no período de Junho de 2009 a Junho de 2010.



Figura 1 - Histórico do Índice Bovespa
Fonte: Fyadoo (2010)

2.2 Operações

As negociações no mercado de ações, de acordo com Assaf Neto (2006), podem ocorrer de três formas distintas:

- A vista: Os títulos vendidos são entregues no segundo dia útil e o pagamento e recebimento entregues no terceiro dia útil com a liquidação financeira realizada pela *Cia. Brasileira de Liquidação e Custódia*;
- A termo: Um contrato determina a quantidade das ações negociadas, prazo de liquidação para uma data futura e um preço de acordo comum para as partes. Apesar de possuir uma data para liquidação uma operação a termo

pode ser liquidada antes de seu vencimento e qualquer papel da BOVESPA pode ser negociado nesta modalidade;

- A opção: Neste caso são negociados os direitos de compra e venda de ações com um prazo e preço estabelecidos em contrato, esses direitos são chamados de prêmios e o pagamento destes pode ser negociado se o investidor não quiser exercer o direito até a data de vencimento da opção (Assaf Neto, 2006).

Outra forma de se negociar neste mercado é através de operações *day trade*, que segundo Giampietro *et al.* (2007), significa a compra e venda de uma mesma quantidade de ações de uma determinada empresa, realizada no mesmo dia pelo mesmo investidor, através da mesma corretora e liquidada pelo mesmo agente de compensação. “As operações *day trade* têm intuito especulativo, ou seja, as pessoas compram e vendem a volatilidade do papel no mesmo dia com o objetivo de ganhar uma ou várias vezes em um mesmo dia” (Focaccia *et al.*, 2002, p. 58).

O investidor pode também realizar operações de venda sem possuir os papéis, o que é chamado de negociação a descoberto ou venda a descoberto. Nesta negociação a corretora obtém um empréstimo das ações com outro investido. Entretanto é necessária a transação de um valor mínimo determinado pela corretora, assim como a compra da mesma quantidade de ações vendida até o final do dia em questão (Assaf Neto, 2006).

2.3 Modelos de Análise de Mercado

A principal tarefa do investidor concentra-se nas análises dos papéis para que estes lhe tragam o retorno esperado de acordo com o capital aplicado. Para isso, os modelos de análise procuram projetar o comportamento futuro dos papéis permitindo a formulação de previsões com base em diversas informações.

Existem duas abordagens básicas para a análise de ações: a análise fundamentalista e a análise técnica, apresentadas a seguir.

2.3.1 Análise Fundamentalista

A análise fundamentalista busca se basear nos fundamentos da empresa, como, lucro líquido, grau de endividamento, patrimônio líquido, pagamento de dividendos e em notícias como fusões ou venda da empresa, para verificar a saúde financeira desta e assim enxergar um futuro promissor ou sinais de desvalorização das ações, utilizando essas informações como base para projetar funções de lucro tomando assim a melhor decisão de compra ou venda com uma visão projetiva e não histórica (Piazza, 2009).

De acordo com Penman (2004), o analista deve possuir uma boa compreensão do negócio e de como as declarações financeiras medem o sucesso em curso da empresa. Com isso, ele deve ser capaz de converter todo este entendimento em uma avaliação única da empresa. Este processo abrange o uso de projeções futuras e torna a análise fundamentalista essencialmente subjetiva, o que acaba exigindo do investidor determinada experiência técnica e formações especializadas para obtenção de estimativas precisas e confiáveis.

Os investidores fundamentalistas podem ser caracterizados, segundo Lima *et al.* (2009), como um investidor que se apoia no fundamento de que o risco pode ser reduzido a partir da diversificação e que o mercado não remunerará o risco não sistemático com as características individuais do negócio. Com isso, o investidor mantém uma carteira diversificada e aguarda informações a respeito do risco sistemático das ações para negociar.

2.3.2 Análise Técnica

A análise técnica, conhecida também por análise gráfica, não está relacionada com a situação clínica da empresa e nem com notícias divulgadas sobre esta. Ela se baseia na projeção do comportamento dos preços das ações utilizando padrões passados do mercado (Piazza, 2009).

Esta análise considera para o estudo do comportamento das ações as informações sobre parâmetros de oferta e procura e a evolução histórica das cotações dos papéis, com essas informações o investidor elabora previsões sobre o desempenho futuro das ações (Assaf Neto, 2006).

Segundo Assaf Neto (2006, p. 193), “as variações nos preços das ações guardam uma relação entre si, descrevendo uma tendência de mercado. Dessa forma, pelas movimentações de preços passadas é possível explicar suas evoluções futuras.”.

A principal ferramenta utilizada por esta análise são os gráficos que descrevem oscilações e tendências dos preços das ações, permitindo, juntamente com os indicadores financeiros, a identificação das melhores oportunidades de transações.

A Figura 2 apresenta um exemplo de um gráfico de *candlestick* comumente utilizado por investidores. Sua vantagem está na quantidade de informações representadas em cada barra, sendo estas, o valor máximo, mínimo, de abertura e de fechamento de certo período de tempo.



Figura 2 – Exemplo de gráfico *candlestick*
Fonte: Fyadoo (2010)

2.3.2.1 Indicadores Financeiros

Os indicadores são utilizados para avaliar as condições reais do mercado, mostrando um possível sentido para este. Existem três grupos de indicadores: o grupo de acompanhamento de tendências, que ajuda o investidor a se manter na tendência de alta ou de baixa sendo os indicadores mais comuns as médias móveis, o MACD e os sistemas direcionais; o grupo de osciladores, que ajudam a localizar pontos de inflexão responsáveis por mostrar os momentos em que o mercado está

em baixa e preste a subir ou em alta e preste a cair, neste grupo os principais indicadores são os índices de força relativa, o estocástico e a taxa de mudança; e o grupo de indicadores mistos que mostram o animo do mercado, apontando se este está aumentando ou diminuindo, os indicadores em destaque deste grupo são o *Bullish Consensus*, *Commitments of Traders* e *New High New Low Index* (Elder, 2002, p. 94).

A utilização dos indicadores média móvel, índice de força relativa e *bollinger bands* está presente neste trabalho, sendo empregados como dados de entrada dos algoritmos como pode ser visto na Seção 4.2. Em seguida estes indicadores são descritos, segundo Elder (2004).

- **Média Móvel:** Este indicador mostra o valor médio das ações em um determinado período de tempo. Por exemplo, uma média móvel de 5 dias mostra o preço médio dos últimos 5 dias.
- **Índice de Força Relativa:** Este indicador mede a força de qualquer veículo de negociação através do monitoramento nas mudanças dos preços. O valor deste indicador varia entre 0 e 100, onde ao atingir valores pertos de 0 e perto de 100, identifica-se sinais de mudanças nas direções dos preços.
- **Bollinger Bands:** O indicador *bollinger bands* é representado por 3 linhas, chamadas de bandas. A banda central é uma média móvel das cotações da ação nos últimos n dias. As bandas, superior e inferior são calculadas em função do desvio padrão relativo à banda central e possuem a característica de se afastar durante os períodos de volatilidade e de se aproximar durante os períodos mais calmos.

2.4 Custos Operacionais

As operações de compra e venda no mercado de ações, enviadas à BOVESPA e executadas, possuem algumas taxas referentes aos custos operacionais.

A primeira taxa a ser analisada é chamada corretagem, ela é cobrada pela corretora por cada operação realizada, podendo ser um valor fixo independente do valor da operação ou variável de acordo com a Tabela 1 baseada em um modelo

padrão da BOVESPA. A escolha ideal de corretagem a se pagar vai depender das estratégias de investimento utilizadas.

Tabela 1- Custos para taxa de corretagem variável

<i>Valor Negociado</i>	<i>Percentual</i>	<i>Parte Fixa</i>
Até 135,07	-	R\$2,70
135,08 a 498,62	2,0%	-
498,63 a 1514,69	1,5%	R\$2,49
1514,70 a 3029,39	1,0%	R\$10,06
Acima de 3029,39	0,5%	R\$25,21

Outra taxa cobrada tanto na operação de compra quanto de venda é o emolumento que representa uma pequena porcentagem do capital aplicado. A porcentagem cobrada varia de acordo com o tipo de operação realizada, para operações normais e *day trade* as Tabela 2 e Tabela 3, respectivamente, representam a porcentagem de emolumento cobrado.

Tabela 2 - Emolumento cobrado de operações normais

<i>Emolumentos – Operações Normais</i>	
Taxa de Negociação (BOVESPA)	0,027%
Taxa de Liquidação (CBLC)	0,008%
Total	0,035%

Tabela 3 - Emolumento cobrado de operações *day trade*

<i>Emolumentos – Operações Day Trade</i>	
Taxa de Negociação (BOVESPA)	0,019%
Taxa de Liquidação (CBLC)	0,006%
Total	0,025%

Os últimos gastos que os investidores terão são referentes ao pagamento do imposto de renda que representa uma parcela de 15% do lucro líquido obtido, sendo cobrado mensalmente para as vendas de ações que ultrapassem R\$ 20.000,00. Entretanto, se o imposto a ser pago for inferior a R\$ 10,00 torna-se desnecessário também o recolhimento. A corretora ainda deve reter 0,005% do total de vendas que ultrapassem R\$ 20.000,00 por questão de controle da Receita Federal. As porcentagens cobradas mudam no caso de operações *day trade*, onde não existe

mais isenção de impostos para operações inferiores a R\$ 20.000,00 e com um valor de tributação de 20% calculado sobre o ganho líquido da operação (Piazza, 2009).

Estas taxas foram detalhadas, pois são utilizadas nos cálculos de lucro do simulador de previsões encontrado no sistema *online* desenvolvido neste trabalho.

2.5 Sumário do Capítulo

O Capítulo de Mercado de Ações contextualizou o leitor com relação aos termos, operações e modelos de análise comuns ao mercado. É possível verificar que operar no mercado de ações, além de necessitar dos conhecimentos básicos apresentados neste capítulo, depende da escolha e utilização de estratégias que auxiliem a definir os melhores momentos de compra e venda das ações. A utilização de técnicas computacionais através da previsão de séries temporais, apresentada no próximo capítulo, remove do investidor a necessidade de ter que se aprofundar nas técnicas e indicadores existentes para os modelos de análise.

3

Referencial Teórico

Este capítulo apresenta o Referencial Teórico partindo da Previsão de Séries Temporais na Seção 3.1, apresentando as características e os fatores que as influenciam, seguido dos métodos utilizados para previsão, destacando os métodos utilizados neste trabalho. Na Seção 3.2 é apresentada a máquina de comitê para a combinação de vários modelos de previsão. Por fim, na Seção 3.3, é descrito uma revisão de literatura dos trabalhos relacionados.

3.1 Previsão de Séries Temporais

A fim de prever o preço médio diário das ações negociadas no mercado financeiro, surge a necessidade do estudo da previsão de séries temporais para compreender suas características e métodos existentes, tornando possível a escolha de quais serão utilizados nas previsões financeiras.

De acordo com Cortez (2002), a vontade de compreender o passado e prever o futuro, leva o homem a procurar leis que expliquem o comportamento de fenômenos e acontecimentos. Na ausência de regras que definam o comportamento de um sistema, procura-se determinar o seu comportamento futuro a partir de observações do passado.

A sequência de fenômenos ou acontecimentos pode ser descrita através de uma série temporal, que é definida como sendo uma coleção de observações feitas sequencialmente ao longo do tempo. A característica mais importante destes tipos de dados é a dependência entre as observações vizinhas (Ehlers, 2007).

Segundo Morettin & Tolo (1981), os objetivos de se analisar uma série temporal são os seguintes:

- a) Descrição: propriedades da série como, por exemplo, o padrão de tendência, a existência de alterações estruturais entre outros;

- b) Explicação: construir modelos que permitam explicar o comportamento da série no período observado;
- c) Controle de Processos: por exemplo, controle estatístico de qualidade;
- d) Previsão: prever valores futuros com base em valores passados.

3.1.1 Fatores que influenciam as séries temporais

3.1.1.1 Sazonalidades

A sazonalidade é uma característica de muitas séries temporais que exibem um comportamento que tende a se repetir de tempos em tempos (Ehlers, 2007). Como exemplo, espera-se que as vendas mensais de chocolate tenham um pico de alta durante a época da Páscoa e talvez um pico de baixa durante o mês de dezembro devido ao calor e a venda de outros produtos. Provavelmente este padrão se repetirá ao longo de vários anos.

3.1.1.1 Tendências

Outra característica das séries temporais são as tendências de crescimento (ou decréscimo) que podem possuir diversos padrões. Segundo Ehlers (2007), as tendências possuem os seguintes tipos de variação:

- Linear: Como exemplo, a cada ano o aumento esperado de lucro bruto nas vendas de ovos de páscoa de uma empresa é de R\$ 500.000,00.
- Exponencial: Como exemplo, a cada ano o lucro bruto na venda de ovos de páscoa aumenta a um fator 1,7.
- Amortecido: Como exemplo, as vendas dos ovos de chocolate branco têm um aumento esperado de 85% sobre o ano anterior. Se o aumento esperado for de R\$ 1.000.000,00 no primeiro ano, no segundo ano será de R\$ 1.850.000,00, no terceiro ano será de R\$ 3.422.500,00 e assim por diante.

3.1.1.1 Ciclos

Algumas séries temporais possuem a capacidade de apresentar oscilações periódicas nos dados, como exemplo as séries construídas com base em dados que descrevem a procura de produtos agrícolas, a venda de automóveis ou até a evolução de indicadores macro econômicos (Cortez, 1997) .

3.1.1.1 Aleatoriedade

A aleatoriedade é o componente das séries que não é previsível, ou seja, a parte que não pode ser explicada. Devido a existência deste componente as previsões nunca são 100% corretas (Cortez, 1997). Apesar disso, é possível obter boas previsões sobre determinadas séries temporais com o auxílio de métodos para a previsão de séries temporais.

3.1.2 Métodos para Previsão de Séries Temporais

Os estudos envolvendo as séries temporais têm como objetivo a realização de previsões com base em informações passadas de fenômenos ou acontecimentos. Estas previsões são alcançadas com o uso de modelos estatísticos ou computacionais.

Os modelos estatísticos se baseiam na utilização de regressores lineares, médias móveis, alisamentos exponenciais Cf. Faria *et al.*, 2008a, p. 6, modelos auto-regressivos e na combinações destes, como é o caso do ARMA que utiliza médias moveis e modelos auto-regressivos.

Os modelos computacionais aplicam técnicas de inteligência artificial para prever os acontecimentos e fenômenos. Entre os modelos utilizados estão os algoritmos evolucionários, as redes neurais artificiais, as máquinas de vetor de suporte (MVS), a lógica *fuzzy* e também a combinação destes modelos. Além disso, é possível realizar a combinação dos resultados obtidos com vários destes modelos, técnica chamada máquina de comitê.

A utilização dos modelos estatísticos e computacionais busca encontrar padrões nas séries temporais do passado, para que seja possível a extrapolação

para o futuro. No caso deste trabalho, são analisadas as séries temporais dos preços médios diários dos papéis no mercado financeiro utilizando as máquinas de comitê para a combinação dos resultados gerados por instâncias dos modelos: programação genética, redes neurais artificiais e regressão linear. Espera-se que essa combinação encontre resultados no mínimo tão bons quanto os melhores resultados encontrados individualmente por cada instância.

3.1.2.1 Programação Genética

A Programação Genética idealizada por John Koza (Koza, 1989) pertence a uma das classes da Computação Evolucionária (CE). A CE é uma área da Inteligência Artificial responsável por englobar um conjunto de métodos computacionais inspirados na Teoria da Evolução das Espécies de Charles Darwin (Darwin, 1859) para a solução de problemas. Segundo sua teoria, na natureza sobrevivem os indivíduos que possuem maior capacidade de se adaptarem ao meio ambiente. Suas características genéticas são repassadas para as gerações seguintes e melhoradas. Assim a nova geração será composta por indivíduos com material genético melhor do que os da população anterior.

A classe da Programação Genética na CE é caracterizada pela estrutura de dados utilizada pelos indivíduos, baseada em árvores sintáticas. As árvores sintáticas são combinações de conjuntos de Funções (F) e Terminais (T), de acordo com o domínio do problema. A Figura 3 mostra um exemplo de árvore sintática.

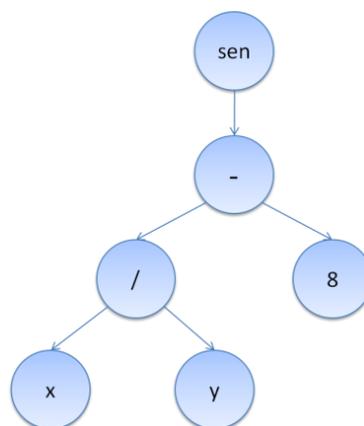


Figura 3 - Representação através de árvore sintática da função $\text{sen}(x/y - 8)$

O Algoritmo Evolutivo proposto por John Koza é composto por uma população de indivíduos capazes de evoluírem de modo a gerar novas populações

constituídas por indivíduos melhores, utilizando para isso operadores específicos para árvores sintáticas de cruzamento e mutação. O processo é conduzido por uma função de aptidão (*fitness*) que mede o quanto o indivíduo está próximo da solução do problema. Indivíduos que possuem maior capacidade de adaptação têm melhores chances de sobreviver. A Programação Genética vem sendo aplicada em diversas áreas do conhecimento, tais como biotecnologia (Koza, 1999), engenharia elétrica (Koza, 2004), análises financeiras (Kuo & Chen, 2003), classificação de padrões (Teredesai & Govindaraju, 2004) e *data mining* (Bojarczuk; Lopes & Freitas, 2004).

O fluxograma de execução da programação genética é mostrado na Figura 4.

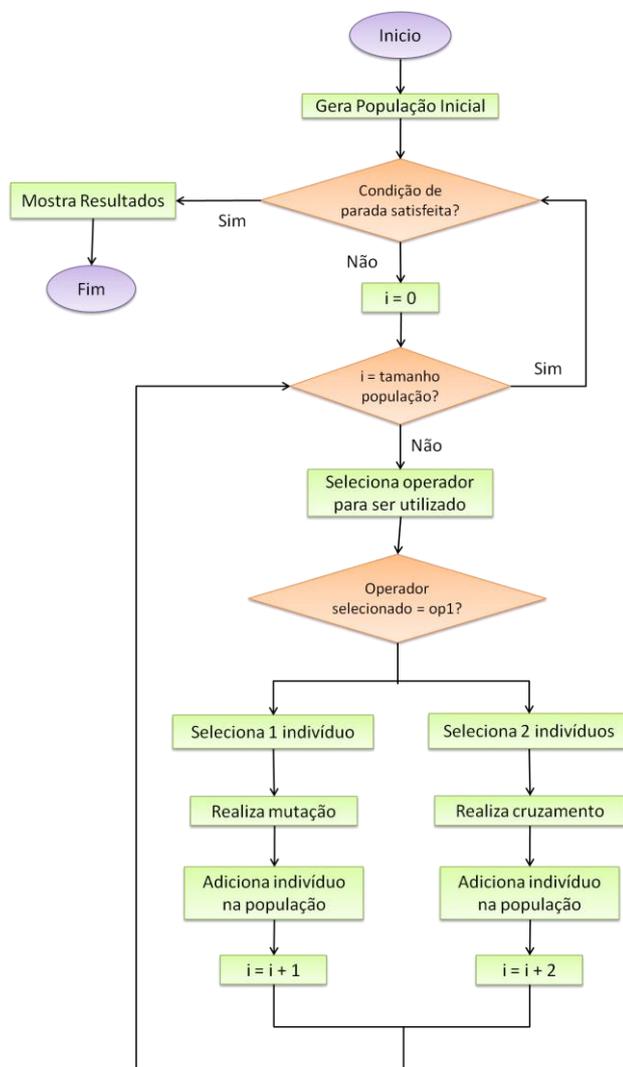


Figura 4 - Estrutura do Algoritmo de Programação Genética

3.1.2.1 Criação da População Inicial

O processo de criação da população inicial é realizado por meio de métodos que definem a altura máxima com que os indivíduos (árvores sintáticas) serão criados. Os métodos mais comuns, segundo Eiben & Smith (2003), são:

- Método *Full*: Este algoritmo define um parâmetro D_{max} que representa a altura máxima das árvores a serem criadas, desta forma todas as árvores iniciais devem obrigatoriamente possuir esta altura.
- Método *Grow*: Este algoritmo também define um parâmetro D_{max} e neste caso as árvores iniciais devem possuir uma altura entre 1 e D_{max} .
- Método *Half-and-half*: Este método cria 50% das árvores utilizando o método *Grow* e 50% utilizando o método *Full*.

3.1.2.1 Avaliação dos Indivíduos

Para avaliar o quão bom é um indivíduo, é necessário associar a cada um o valor de adaptabilidade (*fitness*) correspondente. Uma métrica utilizada para avaliar árvores sintáticas que representam expressões matemáticas é a Soma dos Erros Quadráticos (SSE) muito utilizada na arquitetura de Redes Neurais Artificiais Cf. Haykin, 1998. A equação 1 descreve o cálculo do SSE para definição do *fitness*.

$$SSE = \sum_{i=1}^n (\text{esperado}(i) - \text{obtido}(i))^2 \quad (1)$$

Onde:

esperado: representa o valor real esperado como resposta para o dia i .

obtido: representa o resultado previsto baseado na árvore sintática para o dia i .

n: total de dias para cálculo do SSE

Utilizando esta equação, quanto menor o valor do SSE, mais adaptada é a solução encontrada.

3.1.2.1 Métodos de Seleção

Para selecionar quais indivíduos da população farão parte de uma nova geração e quais deles sofrerão alterações através dos operadores genéticos, é necessário que se tenha um critério de seleção que garanta uma boa escolha.

Existem diferentes operadores de seleção e a decisão de qual destes operadores será utilizado é uma tarefa importante durante a construção do algoritmo. Alguns dos métodos explicados por Souza (2006) são descritos a seguir.

- **Torneio:** Certo número de indivíduos, que é o tamanho do torneio, é selecionado aleatoriamente, e é realizada uma competição seletiva. Os indivíduos com melhores *fitness* durante cada competição permanecem enquanto os com piores *fitness* vão sendo descartados;
- **Ranqueamento:** Neste método os indivíduos são ordenados de forma crescente de acordo com seu *fitness*, então, para cada indivíduo é atribuído uma numeração de acordo com sua posição no *ranking*. Quanto melhor o *ranking* do indivíduo, melhor sua aptidão em relação aos demais indivíduos da população e, portanto, melhores são suas chances de ser sorteado.

3.1.2.1 Operadores Genéticos

As alterações realizadas nos indivíduos para a geração de novos indivíduos são feitas por meio dos operadores genéticos. Estes operadores atuam através de mutações e cruzamentos, tornando possível o surgimento de novos indivíduos com *fitness* melhores que os das gerações anteriores.

- **Mutação:** A mutação é uma alteração aleatória realizada na estrutura dos indivíduos da população. Existem varias técnicas de mutação disponíveis na literatura, sendo a mutação de sub-árvores uma das mais conhecidas. Nesta, seleciona-se um individuo presente na população e é escolhido um nó seu aleatoriamente. Este nó passa a ser a raiz de uma nova sub-árvore criada aleatoriamente que substituirá o ramo antigo (Eiben & Smith, 2003). Na Figura 5 é possível ver um exemplo desta mutação;

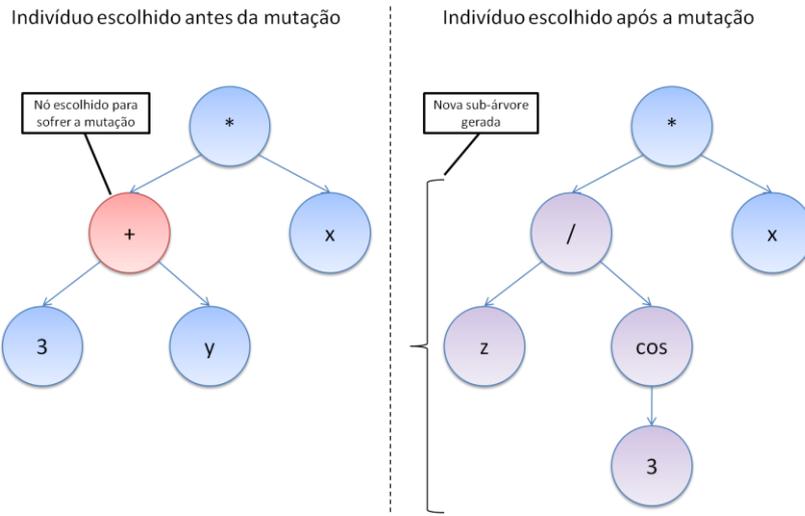


Figura 5 - Mutação em uma sub-árvore

- Cruzamento:** No cruzamento, selecionam-se dois indivíduos chamados de pais e em seguida são escolhidos um nó de cada um para trocar entre eles as sub-árvores destes nós, formando assim dois indivíduos filhos com algumas características de cada pai (Eiben & Smith, 2003). A Figura 6 exemplifica o funcionamento de um cruzamento.

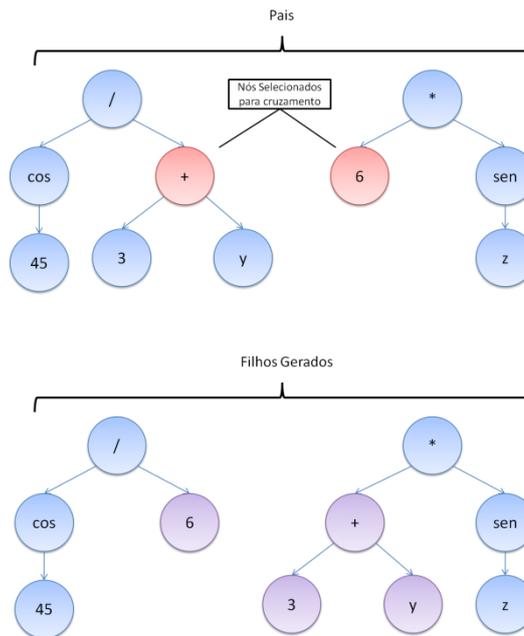


Figura 6 - Cruzamento de dois indivíduos

3.1.2.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) buscam simular o comportamento de aprendizado do cérebro humano através de modelos matemáticos, com o objetivo de poder tomar decisões com base em aprendizados passados e sem a interferência humana (Silva Filho & Abe, 2001). Entretanto, a complexidade da tarefa não permitiu, por enquanto, que fossem encontrados modelos adequados que representem de forma satisfatória a capacidade de aprendizado possível do cérebro. Contudo, com a evolução dos estudos, as RNAs se tornaram uma boa opção para realizar previsão de séries temporais e classificação de padrões.

O aprendizado das RNAs está ligado à adaptação de seus parâmetros internos de acordo com entradas externas em um processo gradativo para melhora de seu desempenho (Rezende, 2005).

Segundo Rezende (2005), a estrutura das RNAs é composta por neurônios artificiais que funcionam como elementos processadores, organizados de maneira paralela e que se comunicam para realizar o aprendizado. Um neurônio artificial possui n entradas com um peso w referente a cada uma, onde a soma das n entradas ponderadas pelos pesos w geram a saída linear u que é processada por uma função de ativação (linear ou não-linear) $f(u)$ gerando por fim a saída do neurônio como exemplificado na Figura 7.

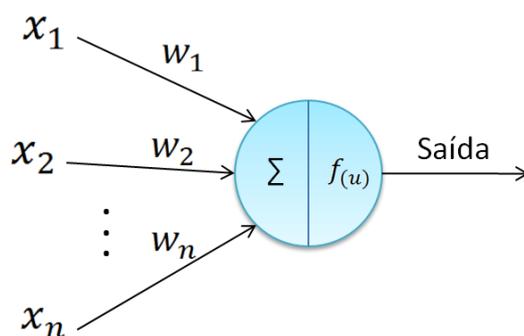


Figura 7 - Neurônio Artificial

A arquitetura (estrutura) das RNAs representa características das relações entre os neurônios artificiais, principalmente com relação ao número de camadas de neurônios. Segundo Haykin (1998, p. 46) “a maneira como os neurônios de uma rede neural estão estruturados está intimamente ligada com o algoritmo de

aprendizagem usado para treinar a rede”, ou seja, a escolha da arquitetura esta totalmente ligada à escolha dos algoritmos de aprendizagem utilizados.

É possível identificar inicialmente três classes de arquitetura de redes essencialmente diferentes (Haykin, 1998):

- **Camada Única:** Neste modelo de arquitetura existe apenas uma camada de neurônios artificiais que recebem os dados de entrada e retornam as saídas. Devido a isso, está rede é sempre do tipo alimentada a diante ou acíclica. Uma das características desta rede, descoberta por Misnky & Papert (1969), é que ela é capaz de resolver somente problemas linearmente separáveis;
- **Múltiplas Camadas:** Este tipo de arquitetura se caracteriza pela presença de uma ou mais camadas ocultas, onde a função dos neurônios destas camadas é intervir de forma útil entre os dados de entrada e a camada de saída. Estas camadas ocultas permitem à rede extrair estatísticas diferenciadas que são muito úteis quando o número de entradas é grande. A Figura 8 mostra um exemplo da Arquitetura de Múltiplas Camadas;

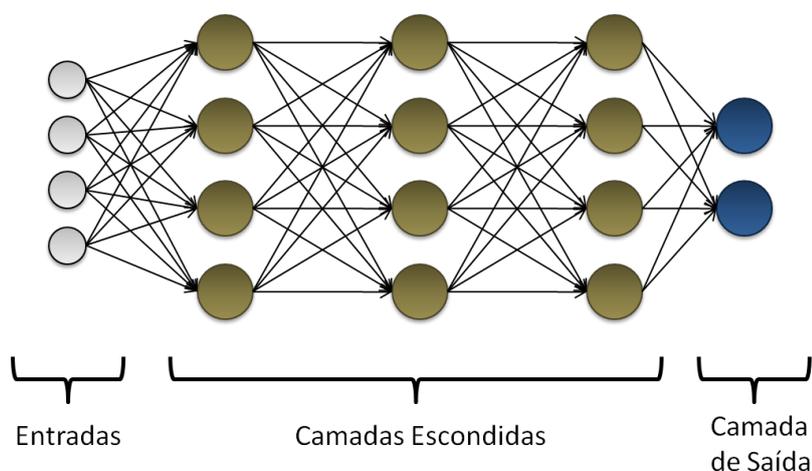


Figura 8 - Exemplo de Rede com Múltiplas Camadas

- **Recorrentes:** Este modelo se diferencia por ter pelo menos um laço de realimentação, que diz respeito à alimentação da rede com dados já processados por ela. A realimentação gera um grande impacto na capacidade de aprendizagem e no desempenho da rede (Haykin, 1998).

Neste trabalho as RNAs são utilizadas como um dos modelos para a previsão das séries temporais do preço médio das ações do mercado financeiro.

Chakraborty *et al.* (1992) mostrou a existência de modelos eficientes de RNAs que garantem bons resultados para a resolução de problemas com previsão de séries temporais.

3.1.2.1 Regressão Linear

Segundo Box; Jenkins & Reinsel (1994), a análise de regressão é uma técnica de modelagem utilizada para encontrar a relação entre uma variável dependente (resposta) contínua e uma ou mais variáveis independentes (preditoras). Além disso, a regressão é utilizada para prever novas respostas para a variável dependente utilizando como base variáveis independentes desconhecidas durante o processo de regressão.

No método da regressão linear múltipla, são permitidas duas ou mais variáveis independentes para a localização da variável dependente. O modelo da equação de regressão utilizando este método é expresso de acordo com a equação 2.

$$\gamma = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + \varepsilon \quad (2)$$

Onde:

- γ é a variável dependente.
- cada x_i é uma variável independente.
- cada β_i é um parâmetro da regressão que diz respeito à x_i .
- ε é o erro aleatório.

A principal ideia em se adicionarem diversas variáveis independentes está em poder encontrar uma estimação mais precisa. Porém, é importante haver um equilíbrio na quantidade de variáveis independentes de acordo com a capacidade preditiva do modelo para que não haja um sobreajustamento dos dados com o excesso destas variáveis.

Este trabalho utiliza a regressão linear múltipla como uma das técnicas para encontrar previsões do preço médio das ações do mercado financeiro.

3.2 Máquinas de Comitê

Com o objetivo de utilizar várias RNAs configuradas de forma diferente para certo problema de regressão, Hansen & Salamon (1990) definiram as máquinas de comitê, tais como sendo a combinação, com estratégias de decisões apropriadas, de vários modelos de previsões, a fim de encontrar um resultado equivalente ou superior ao melhor resultado descoberto.

Segundo Leone (2006) utilizam-se de máquinas de comitê com o objetivo de explorar o bom comportamento local de cada previsor que tendem a ter um desempenho melhor para determinados conjuntos de dados, aumentando com isso a precisão e confiabilidade das previsões.

A Figura 9 mostra o funcionamento de uma máquina de comitê utilizando vários previsores, onde cada um dos previsores recebe os dados de entrada necessários do problema e retorna seu resultado que é usado como entrada para a máquina de comitê a qual fará uma combinação destes para retornar apenas uma saída.

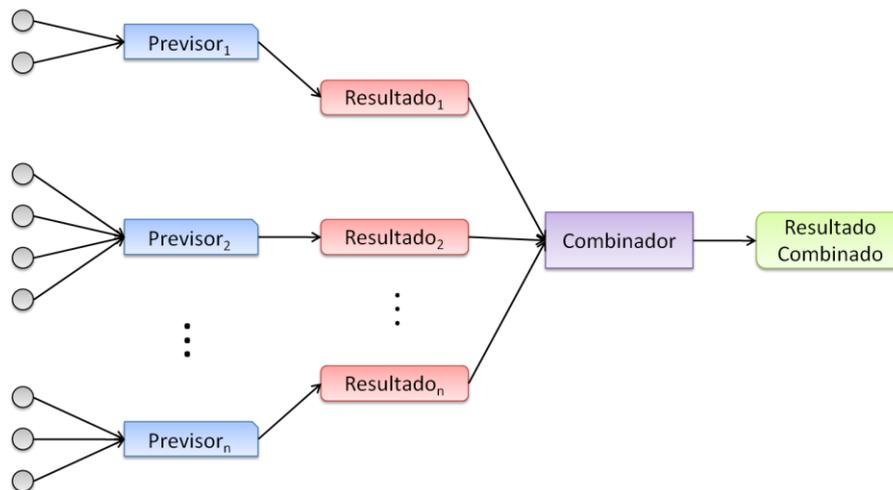


Figura 9 - Estrutura Básica da Máquina de Comitê

A combinação dos previsores pode ocorrer de várias formas segundo Leone (2006), através de votação múltipla ou votação majoritária para classificação de padrões e através de médias aritméticas ou médias ponderadas para problemas de

regressão, ou ainda utilizando-se otimização linear, pesos dinâmicos ou modelos de previsão.

Neste trabalho serão utilizados como entradas para as máquinas de comitê três modelos de previsões: RNA MLP, Programação Genética e Regressão Linear com o objetivo de melhorar os resultados da previsão do valor médio diário das ações, já que, a decisão coletiva produzida por um conjunto de previsores é provavelmente melhor do que as decisões tomadas pelos previsores individualmente (Hansen & Salamon, 1990).

3.3 Revisão de Literatura

O estudo da previsão das séries temporais do mercado financeiro é facilmente encontrado na literatura, porém, não existe uma base de dados única utilizada pelos pesquisadores, o que dificulta a comparação entre os trabalhos existentes. Devido a isso serão apresentados, a seguir, os últimos resultados obtidos, entretanto sem a comparação entre as estratégias propostas.

Campos *et al.* (2004) com a utilização de RNAs MLP obteve previsões para o índice BOVESPA com um erro médio relativo de 5,06% durante o período de treinamento e de 5,53% durante o período de testes, com um desvio padrão de 1,21% e 1,92% respectivamente, mostrando assim que é possível prever o valor de um indicador com a utilização de Redes Neurais Artificiais MLP.

Cabral (2005) comparou o potencial de duas técnicas para previsão de séries temporais, sendo uma as máquinas de vetor de suporte (MVS) e outra as Redes Neurais Artificiais com topologias MLP e Elman. Em seus experimentos com as séries do mercado de ações da BOVESPA, Cabral (2005) utilizou como método de avaliação o cálculo do erro médio quadrático (EMQ). O menor EMQ obtido utilizando-se RNA Elman foi de 0,0013 enquanto para a MLP foi possível alcançar EMQ de 0,0017, porém o melhor resultado obtido foi na utilização de MVS onde conseguiu um EMQ de 0,00003.

Utilizando como método a lógica *fuzzy*, Souto-Maior (2007), previu o valor do Índice Bovespa utilizando como base de teste dados de Janeiro de 2001 a

Fevereiro de 2005 e obteve rendimentos de 41,48% até 77,80% variando-se a taxa de erro pré determinada.

Faria *et al.* (2008b), também obteve sucesso na previsão do índice Bovespa utilizando Redes Neurais Artificiais, atingindo erros médios em torno de 2%. Através da técnica de janelamento obteve um erro durante o treinamento de 10^{-3} , suficiente para previsões razoáveis do índice BOVESPA.

Martinez *et al.* (2009), teve como objetivo a previsão dos valores máximos e mínimos das ações da Bovespa, utilizando como método computacional as Redes Neurais Artificiais e como método de comparação os resultados de Connor & Madden (2006) *apud* Martinez et al. (2009, p. 5), onde o melhor resultado obtido apresentou um erro percentual médio absoluto em torno de 50% menor que os melhores resultados de Connor & Madden (2006).

Yen *et al.* (2007), fez uma comparação para o mercado financeiro de Taiwan de previsões utilizando Programação Genética e Redes Neurais Artificiais e obteve melhores resultados em todas as previsões realizadas com a utilização das RNAs. Mostrando que soube encontrar uma boa configuração para sua RNA.

Castro; Braga & Andrade (2005), aplicaram um modelo de máquina de comitê utilizando os resultados de três RNAs com arquitetura MLP para previsão da série de índices de fechamento diário da Bolsa de Valores de São Paulo. Utilizou como entrada para as RNAs os índices de fechamento diário das bolsas internacionais Tóquio, Londres, Nova Iorque e Frankfurt, o fechamento da cotação diária do Dólar x Real e o valor do fechamento do volume negociado diariamente. Para o cálculo dos resultados Castro; Braga & Andrade (2005) utilizou a raiz dos erros quadráticos médios, obtendo para cada RNA respectivamente os erros 70,40, 68,36 e 114,35 e para a máquina de comitê que utilizou estas RNAs o erro de 37,30 mostrando assim que o modelo combinado apresentou uma grande melhora nos resultados.

Mello (2004) teve como objetivo o desenvolvimento de uma RNA *feedforward* com algoritmo de treinamento *backpropagation* para a previsão dos preços futuros de ações das empresas Eletrobrás, Petrobrás e Vale do Rio Doce. Foram utilizadas 16 variáveis como entrada entre preços de abertura, máximos e mínimos, fechamentos, médias móveis, índices, taxas de poupança e taxas de dólar. Os resultados encontrados foram obtidos com o cálculo do erro quadrático médio,

para a Eletrobrás os erros durante a fase de teste variaram entre 0,0068 e 0,02, para a Petrobrás variaram entre 0,0103 e 0,0144 e para a Vale do Rio Doce variaram entre 0,0140 e 0,0583, mostrando que obteve sucesso nos seus objetivos.

Pontes (2008) utilizou sistemas multiagentes inteligentes com um total de 146 agentes inteligentes, tendo como dados de entrada índices de força relativa com períodos de 7, 14, 21 e 28 dias, tendo como entrada dados referente ao ano de 2008 e obtendo, com isso, uma média de acerto de 78,75% com a previsão feita pelo agente que utilizou Máquina de Vetor de Suporte.

3.4 Sumário do Capítulo

Este capítulo apresentou o Referencial Teórico em que este trabalho se embasou, descrevendo a previsão de séries temporais e as técnicas utilizadas neste trabalho para a previsão do valor médio diário das ações de algumas empresas brasileiras. Além disso, foram descritos os trabalhos relacionados que também buscavam realizar previsões no mercado financeiro, mostrando que previsões nesta área da economia são comumente realizadas, apesar de que não foram encontrados trabalhos relacionados que realizem a previsão do valor médio diário das ações como ocorre nesta monografia. O próximo capítulo utiliza todo o embasamento teórico para a construção do projeto desta monografia.

4

Projeto

Este capítulo apresenta a metodologia utilizada para responder a hipótese H_0 . Na seção 4.2.1 é realizado uma análise dos dados utilizados. A seção 4.2.2 apresenta como foi definida a normalização dos dados. A seção 4.2.3 mostra como os modelos de previsão utilizados foram construídos. As métricas para combinação dos modelos através das máquinas de comitê são descritas na seção 4.2.4. Em seguida a forma como os resultados são gerados é descrita na seção 4.2.5. Na seção 4.2.6 é apresentado como os experimentos foram realizados e na seção 4.2.7 o teste estatístico usado para verificar a hipótese H_0 é apresentado.

4.1 Projeto

A fim de encontrar melhores resultados na busca por previsões do preço médio diário das ações torna-se importante descobrir se a utilização das máquinas de comitê para a combinação de modelos de previsores trará resultados melhores ou no mínimo tão bons quanto o melhor resultado descoberto isoladamente. Sendo assim, a hipótese H_0 foi definida de forma que “as máquinas de comitê garantem erros menores ou iguais aos menores erros de previsores isolados do preço médio diário das ações”, enquanto a hipótese H_1 ficou definida como “as máquinas de comitê não garantem erros menores ou iguais aos menores erros de previsores isolados do preço médio diário das ações”.

A seção 4.2, a seguir, descreve todo processo necessário para alcançar a hipótese H_0 . O processo inclui a organização, obtenção e o tratamento dos dados, a construção dos modelos de previsão, a definição de métricas de combinação de resultados para as máquinas de comitê, a escolha dos tipos de resultados gerados, a execução dos algoritmos e os testes de hipótese aplicados.

4.2 Metodologia

A metodologia deste trabalho envolve uma pesquisa do tipo experimental que passa pelos passos descritos no fluxograma da Figura 10 para responder a hipótese H_0 definida. Todos os passos do fluxograma são detalhadamente descritos em seguida.

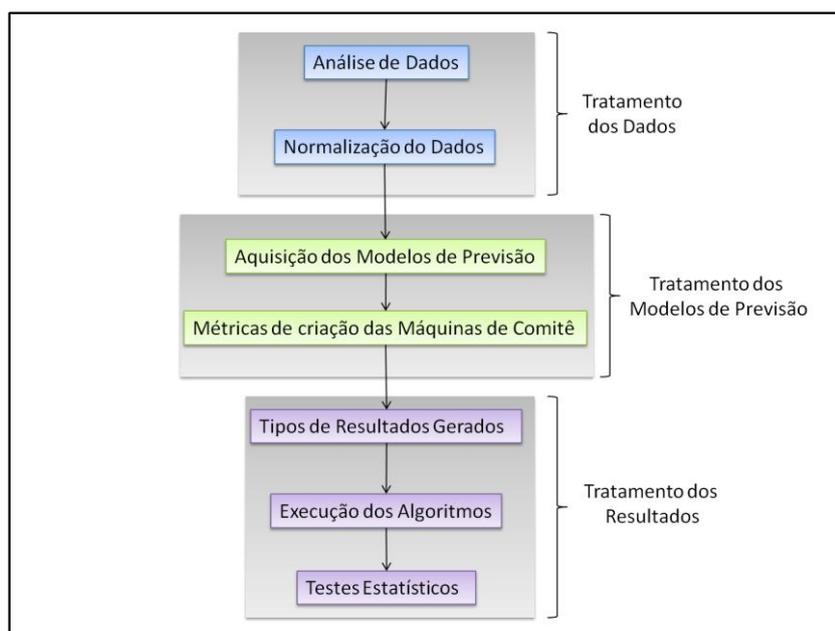


Figura 10 - Fluxograma de Execução da Metodologia

4.2.1 Análise de Dados

4.2.1.1 Definição dos dados utilizados

Os dados necessários para este trabalho são informações passadas do mercado de ações que possam influenciar o direcionamento futuro dos preços das ações. Para chegar aos dados que serão utilizados pelos modelos é necessário passar por um processo de obtenção, seguido pelo pré-processamento de alguns dados primários.

O processo de obtenção é feito através da base de dados financeiros do Yahoo (Fyahoo, 2010), que disponibiliza gratuitamente os dados primários necessários, sendo estes, as informações diárias do valor de abertura, valor máximo, valor mínimo, valor de fechamento e volume referentes a papéis de

empresas da BOVESPA. A obtenção dos dados foi automatizada utilizando, para isso, um endereço *web* disponibilizado pela empresa Yahoo Inc. que de acordo com informações como nome do papel, data inicial e data final ele retorna um arquivo no formato *comma-separated values* (csv) com os dados relacionados. Esta automatização visa facilitar os testes e experimentos necessários.

O segundo passo para obtenção dos dados de entrada utilizados pelos modelos de previsão é a realização de um pré-processamento dos dados primários. O pré-processamento gera uma base de dados de entrada com as seguintes informações para cada dia: valor da abertura, valor de fechamento do dia anterior, preço médio da ação do dia anterior e os indicadores financeiros, média móvel do dia anterior, índice de força relativa do dia anterior e o *bollinger bands* do dia anterior. Cada dia possui como saída o valor médio da ação para este, que é representado pela média entre o valor máximo e o valor mínimo da ação no dia.

O cálculo dos indicadores financeiros utilizados neste trabalho é realizado com o auxílio da Biblioteca de Análises Técnica para Java (Tail) (Santos, 2008), uma biblioteca *open-source* que abstrai os componentes básicos da Análise Técnica e além de facilitar os cálculos de indicadores, fornece ferramentas para manipular, criar e avaliar estratégias de compra e venda.

4.2.1.1 Conjunto de Dados

Os papéis escolhidos para realização dos experimentos são alguns dos que compõe o índice BOVESPA que representa os principais papéis negociados na BOVESPA. Foram escolhidos os papéis das empresas Vale do Rio Doce (VALE5), Petrobrás (PETR4), Usiminas (USIM5), Bradesco (BBDC4) e Itaú-Unibanco (ITUB4). Serão realizados experimentos para todos os dias do ano de 2009 com a intenção de prever o valor médio diário de cada um dos papéis escolhidos.

A Tabela 4 apresenta as informações estatísticas dos dados referentes às séries do valor médio diário dos papéis utilizados para experimentos no período do ano de 2009.

Tabela 4 - Informações estatísticas do Conjunto de Dados de 2009

<i>Papel</i>	<i>Máxima</i>	<i>Mínima</i>	<i>Média</i>	<i>Desv. Pad.</i>	<i>Coef. Var.(%)</i>
VALE5	43,21	25,22	33,19	5,24	15,78
PETR4	39,75	23,31	31,83	4,15	13,02
USIM5	53,53	23,21	39,30	8,76	22,28
BBDC4	38,40	19,81	29,09	5,35	18,40
ITUB4	40,24	20,89	31,23	5,00	16,02

Os papéis apresentam desvios padrão semelhantes, exceto para o papel da Usiminas que além do desvio padrão maior, possui uma porcentagem do coeficiente de variação bem maior que os demais papéis o que implica numa grande variação dos valores médios diários da ação, podendo afetar o potencial das previsões para este papel.

O conjunto de dados que será utilizado para realização das previsões pode ser visualizado na Figura 11, onde para cada papel são apresentadas os valores médios diários. Pelo gráfico é possível observar uma variação relativamente maior nos valores dos preços médios da Usiminas, informação apresentada anteriormente através do coeficiente de variação destes dados e confirmada com a visualização do gráfico.

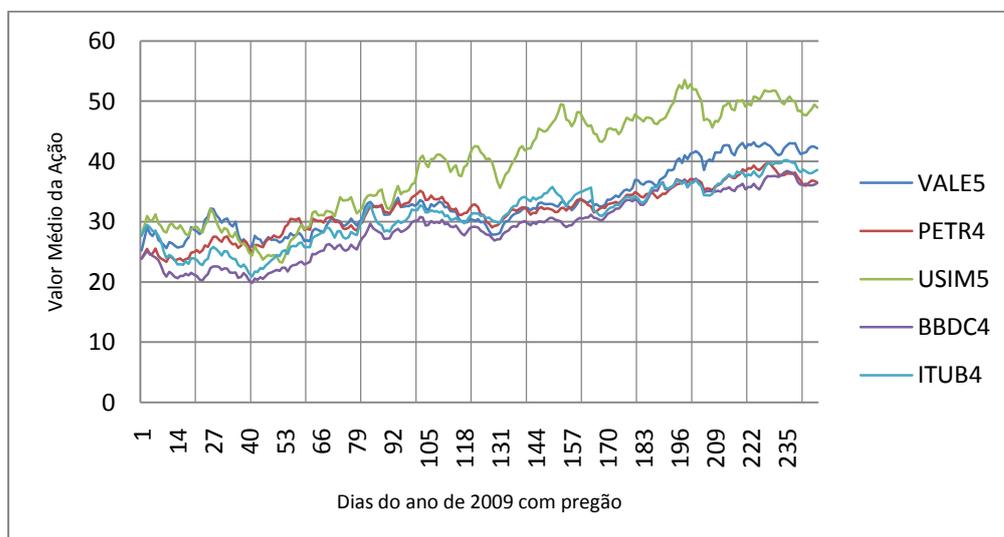


Figura 11 - Séries dos valores médios das ações

4.2.2 Normalização dos Dados

A normalização dos dados é uma tarefa que tem como objetivo transformar os valores dos dados de entrada e saída em valores dentro de uma faixa estabelecida, mantendo a proporção entre estes dados, visando com isso melhorar as previsões, diminuindo os grandes desvios que podem existir entre os dados.

Neste trabalho o modelo que necessita da normalização dos dados é a RNA. Seus dados de entrada e saída são normalizados de acordo com a equação 3 apresentada em seguida, de forma que x_{in} representa o padrão x_i normalizado, $\max(X)$ representa o maior valor do conjunto de entradas e saídas e k representa o total de entradas mais saídas para um dia.

$$x_{in} = \frac{x_i}{\max(X)} \text{ para } i = 1 \dots k \quad (3)$$

4.2.3 Aquisição dos Modelos de Previsão

Para realização deste trabalho foi necessário a utilização de algoritmos responsáveis por gerar instâncias dos modelos de previsão escolhidos. Destes algoritmos, o criado para o modelo de programação genética foi desenvolvido pelo autor durante a execução deste trabalho. O algoritmo do modelo de redes neurais artificiais foi criado utilizando o *framework* NeuralNetwork disponível em <http://bcc.unifal-mg.edu.br/~ricardo/modelos/neural20.zip> e o algoritmo do modelo de regressão linear foi criado utilizando a biblioteca LinearRegression disponível em <http://bcc.unifal-mg.edu.br/~ricardo/modelos/linearregression.zip>.

4.2.4 Métricas de criação das Máquinas de comitê

As máquinas de comitê utilizadas neste trabalho buscam combinar os resultados de instâncias dos modelos de previsão, programação genética, redes neurais artificiais e regressão linear a fim de otimizar os resultados encontrados por cada um isoladamente. As combinações de resultados são realizadas utilizando métricas que combinam, de acordo com a mesma data, os valores encontrados por dois ou três previsores retornando uma nova previsão para a data. A Figura 12 exemplifica a

combinação dos resultados dos três modelos implementados, utilizando uma métrica qualquer pela máquina de comitê.

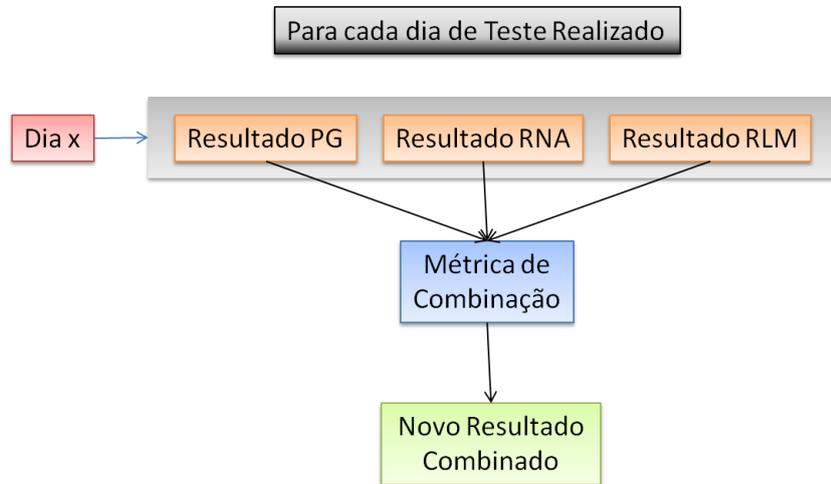


Figura 12 - Esquema da máquina de comitê

Neste trabalho foram realizadas as seguintes combinações de modelos, PG com RNA, PG com RL, RNA com RL e PG com RNA com RL. A métrica escolhida para combinação foi a média ponderada que apresentou a melhor redução de erros em relação a média aritmética, média harmônica e média quadrática. A equação 4 apresenta a forma como o cálculo da média ponderada é realizado.

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \quad (4)$$

Onde:

- \bar{x} é o novo resultado para um dia de experimento;
- n é o número de previsores utilizados;
- x_i é o resultado encontrado pelo previsor i ;
- w_i é o peso dado ao resultado do previsor i .

O peso atribuído ao resultado de cada previsor é calculado de acordo com o desempenho médio dos previsores nas análises passadas, ou seja, os melhores previsores terão uma ponderação maior para o cálculo da média.

4.2.5 Tipos de Resultados Gerados

Os resultados das previsões do preço médio diário das ações são apresentados através do erro relativo entre o valor real do preço médio e o valor previsto. O erro relativo é apresentado na equação 5 em seguida.

$$ER_{xi} = \frac{EA_{xi}}{x_i} \text{ para } i = 1 \dots k \quad (5)$$

Onde:

- ER_{xi} = erro relativo do dia i ;
- $EA_{xi} = |x_i - \bar{x}_i|$ = erro absoluto;
- x_i = valor real do preço médio no dia i ;
- \bar{x}_i = valor previsto do preço médio no dia i ;
- k = total de dias com previsões

Cada previsão realizada produz uma lista com uma data, o valor real e o valor previsto para cada dia de previsão requisitado. Estas listas são utilizadas para calcular os erros relativos encontrados em cada dia, servindo de base para a realização dos experimentos necessários na comprovação ou não da hipótese H_0 .

4.2.6 Execução dos Algoritmos

A execução dos algoritmos consiste de repetidas execuções de instâncias dos modelos de previsões utilizados e das máquinas de comitê definidas, com o objetivo de gerar dados para formulação do teste estatístico que visa responder a veracidade da hipótese H_0 . A Figura 13 apresenta o fluxograma dos passos necessários para a execução dos experimentos utilizando os dados de todos os dias do ano de 2009.

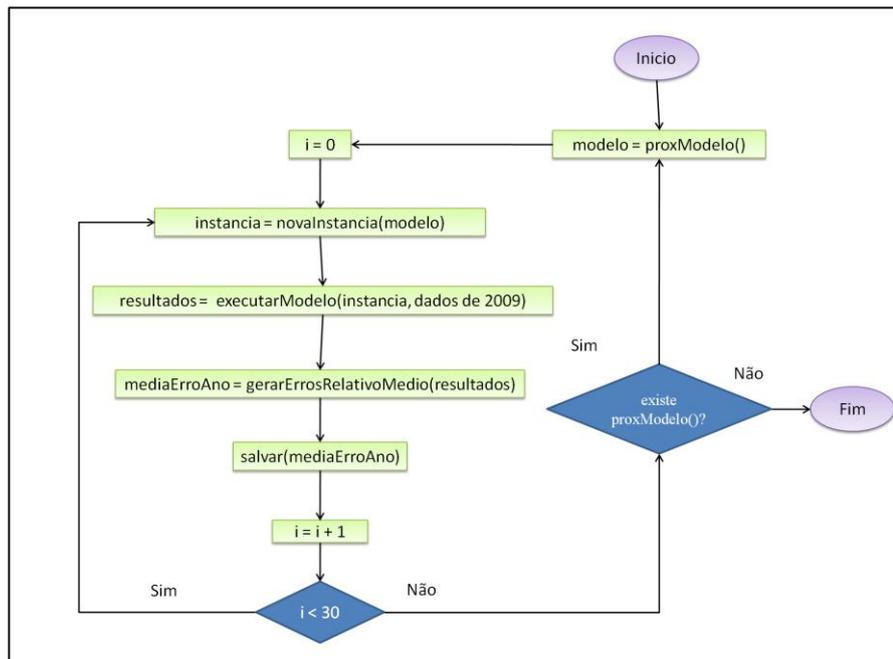


Figura 13 - Fluxograma da execução dos algoritmos para cada papel

O fluxograma da Figura 13 apresenta o experimento que é realizado com cada papel da BOVESPA utilizado neste trabalho. Nele, cada modelo e técnica utilizados, que incluem programação genética, redes neurais artificiais, regressão linear, máquinas de comitê com média aritmética, máquinas de comitê com média ponderada, máquinas de comitê com média harmônica e máquinas de comitê com média quadrática são executadas 30 vezes, gerando para cada execução uma nova instância do modelo ou técnica, onde o erro relativo médio dos resultados gerados é salvo para utilização durante o teste estatístico.

4.2.7 Testes Estatísticos

A comprovação da hipótese H_0 depende da realização de um teste estatístico que além de mostrar que a máquina de comitê é superior ou não a utilização dos modelos isoladamente, mostre também se existem diferenças estatisticamente significantes a ponto de agrupar separadamente os tratamentos. Estes tratamentos são todos os resultados gerados separados por modelos e técnicas de combinação utilizados.

O teste utilizado neste trabalho para a comparação de médias e comprovação da hipótese é o teste de Tukey (Tukey, 1977), que permite encontrar contrastes entre os tratamentos, tornando possível separá-los em grupos que possuem equivalências nos resultados.

Antes do teste de Tukey foi realizada a análise de variância. Esta análise permite constatar se existem diferenças significativas entre os tratamentos em questão. Porém, mesmo existindo diferenças significativas, é possível que não se encontrem diferenças significativas entre contrastes de médias, por este motivo que o teste de Tukey foi aplicado em seguida.

A análise de variância e o teste de Tukey deste trabalho foram realizados utilizando um software gratuito e open-source de estatística chamado R, desenvolvido na Universidade de Auckland na Nova Zelândia e de grande aceitação em Universidades e Empresas.

4.3 Sumário do Capítulo

O capítulo do Projeto apresentou as informações dos dados utilizados e as condições necessárias para a realização dos experimentos, detalhando também a forma de execução utilizada nestes experimentos, além do teste estatístico com os resultados obtidos. A partir do próximo capítulo serão apresentados os resultados desta monografia, começando pelos resultados obtidos com a realização dos experimentos e a análise do teste estatístico, seguido do sistema *online* para apoio a decisão do investidor.

5

Resultados

Este capítulo apresenta os resultados encontrados através da realização dos experimentos. Na Seção 5.1 são apresentadas as configurações utilizadas pelos modelos de previsão PG e RNA para a geração dos resultados. A Seção 5.2 apresenta a análise dos resultados obtidos e em seguida na Seção 5.3 são apresentados os resultados obtidos pelo teste estatístico.

5.1 Configurações das Instâncias dos Modelos de Previsão

Os modelos de previsão, PG e RNA, possuem alguns parâmetros que precisam ser configurados para a realização dos experimentos. A escolha destes parâmetros afeta diretamente os resultados das previsões realizadas. Devido a isto, eles foram cuidadosamente escolhidos em testes anteriores a execução dos experimentos. Em seguida serão apresentadas as configurações utilizadas neste trabalho para os modelos de previsão PG e RNA.

5.1.1 Configurações do Algoritmo de PG

O algoritmo de programação genética foi elaborado criando a população inicial através do método *half-and-half*, detalhado na Seção 3.1.2.1, com 50 indivíduos na população, sendo que, a altura máxima dos indivíduos (árvores sintáticas) é 10. Em cada geração 20 indivíduos sofrem mutação, escolhidos de forma aleatória e 40 recombinações são realizadas. A seleção dos indivíduos para uma nova geração foi feita utilizando o Torneio Binário, no qual são escolhidos dois indivíduos aleatoriamente e o que possui melhor *fitness* é transferido para a nova geração. O torneio se repete até o número de indivíduos da nova população ser igual a 50. O cálculo do *fitness* de cada indivíduo foi gerado através do inverso da soma dos

erros quadráticos de acordo com a equação 6. Uma instância do algoritmo termina de ser executada quando atinge o número de 200 gerações.

$$fitness = \frac{1}{(\sum_{i=1}^n (esperado(i) - obtido(i))^2)} \quad (6)$$

Onde:

- **esperado**: representa o valor real esperado como resposta para o dia i .
- **obtido**: representa o resultado previsto baseado na árvore sintática para o dia i .
- **n**: total de dias para calculo do SSE

5.1.2 Configurações do Algoritmo de RNA

Para os Algoritmos de RNA serem executados é necessário a definição das seguintes configurações: número de camadas, número de neurônios nas camadas, taxa de aprendizagem, termo de momentum e as funções de ativação. As configurações utilizadas neste trabalho foram definidas com base em testes com diferentes combinações de valores empregados. As configurações escolhidas foram: 2 camadas intermediárias com 6 neurônios artificiais cada uma, uma taxa de aprendizagem de 0.8 por camada, um termo de momentum de 0.6 para a primeira camada, 0.6 e 0.5 para as camadas intermediárias e 0.4 para a camada de saída. Para todas as camadas foram utilizados parâmetros lineares como função de ativação.

5.2 Análise dos Resultados

Os resultados gerados pelos experimentos são apresentados em seguida de acordo com cada papel escolhido. Cada gráfico mostra o erro relativo médio dos resultados dos modelos de previsão e das máquinas de comitê de acordo com as previsões realizadas para todo o ano de 2009.

A Figura 14 apresenta o gráfico boxplot dos resultados das previsões para o Bradesco.

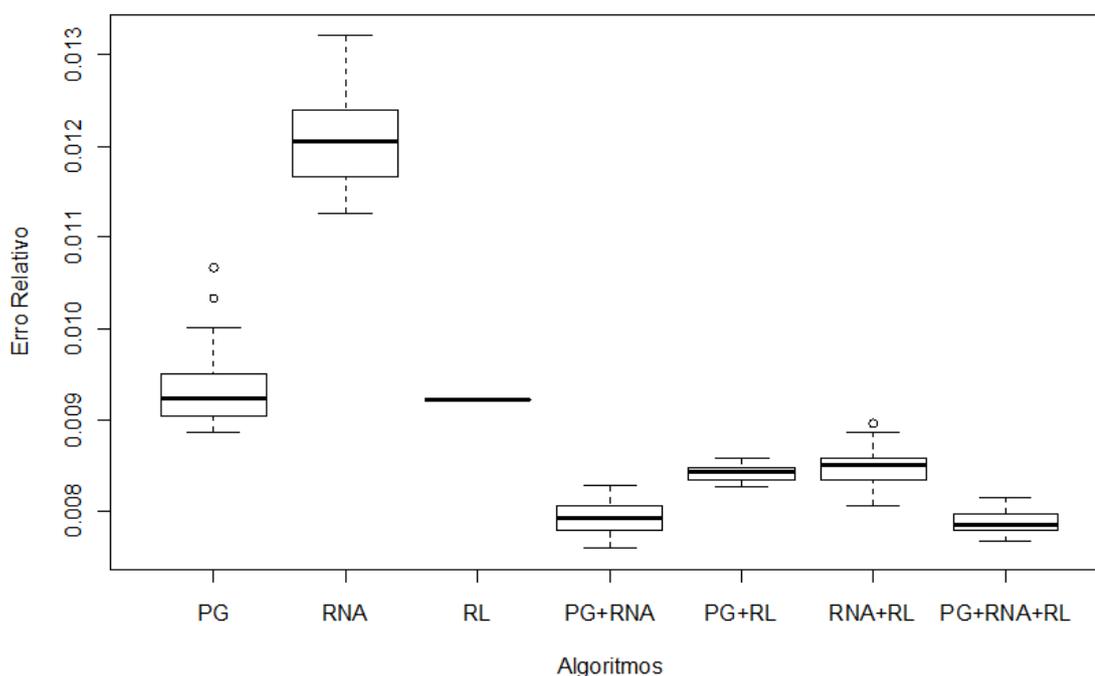


Figura 14 - Gráfico dos resultados para o papel BBDC4

O gráfico mostra que os maiores erros foram encontrados com o modelo de previsão RNA. No caso da RL, como todos os resultados são idênticos por este ser um modelo determinístico, seu boxplot é representado por uma linha. Os melhores resultados foram obtidos com a combinação dos resultados dos modelos. O modelo de previsão PG e a combinação RNA+RL geraram *outliers*, o que mostra que alguns resultados das distribuições estavam muito diferentes dos demais. Além disso, é possível observar, através da simetria dos quartis, que a melhor distribuição dos resultados está na combinação PG+RNA.

A Tabela 5 apresenta mais algumas informações referentes aos resultados para as previsões do valor médio diário das ações do Bradesco.

Tabela 5 - Informações dos resultados para Bradesco

	PG	RNA	RL	PG+RNA	PG+RL	RNA+RL	PG+RNA+RL
Média	0,94%	1,21%	0,92%	0,79%	0,84%	0,85%	0,79%
Coef. Var.	4,60%	4,00%	0,00%	2,37%	1,07%	2,43%	1,61%
Desvio Pad.	0,04%	0,05%	0,00%	0,02%	0,01%	0,02%	0,01%

Os resultados para o Bradesco apresentaram em média o menor erro relativo com a combinação dos modelos PG+RNA+RL obtendo uma média de

0,79% de erro. Porém, apenas com os resultados do teste estatístico realizado é possível dizer se existem modelos equivalentes estatisticamente ou não. Os resultados que mais variaram foram os do modelo PG com um coeficiente de variação de 4,6%. O maior desvio padrão foi de 0,05% gerado pelo modelo RNA. As técnicas de combinação, de forma geral, apresentaram um baixo desvio padrão e um baixo coeficiente de variação.

A Figura 15 apresenta o gráfico com os resultados do experimento para o preço médio diário das ações da empresa Itaú-Unibanco.

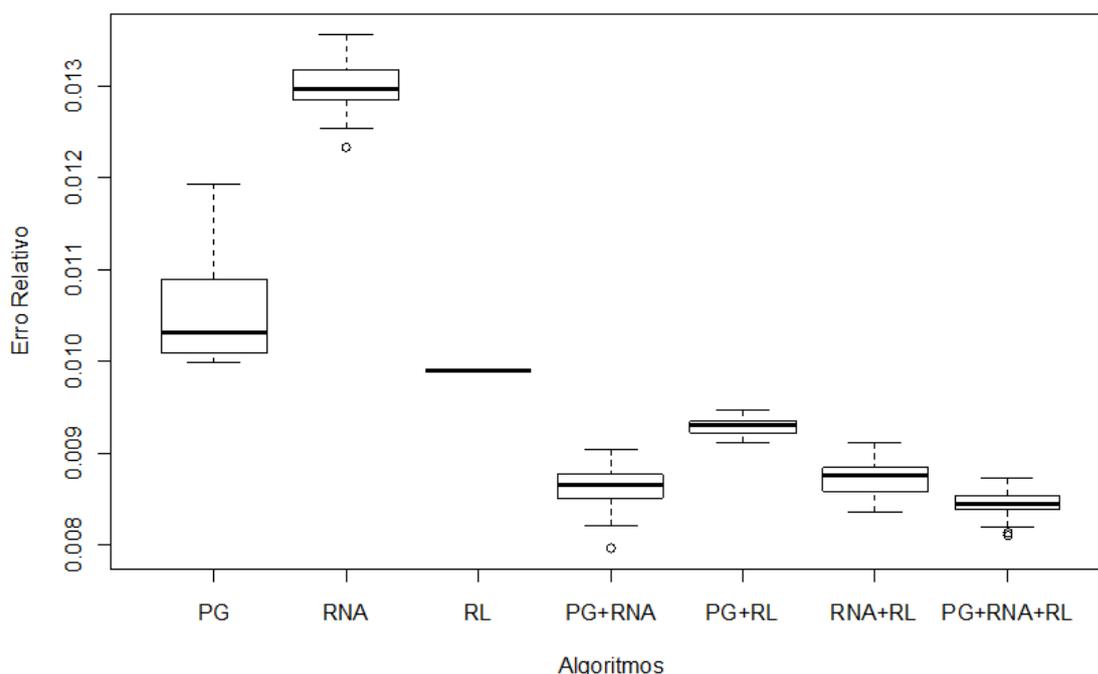


Figura 15 - Gráfico dos resultados para o papel ITUB4

O gráfico boxplot do Itaú-Unibanco mostra, também, os maiores erros relativos encontrados com o modelo de previsão RNA. É possível notar alguns *outliers* para o modelo RNA e para as combinações PG com RNA e PG com RNA com RL, o que mostra que apresentaram alguns dos resultados muito diferentes dos demais. Os resultados são mais semelhantes e com poucas variações para a combinação PG com RL.

A Tabela 6 apresenta outras informações dos resultados para o valor médio diário das ações do Itaú-Unibanco.

Tabela 6 - Informações dos resultados para Itau-Unibanco

	PG	RNA	RL	PG+RNA	PG+RL	RNA+RL	PG+RNA+RL
Média	1,05%	1,30%	0,99%	0,86%	0,93%	0,87%	0,84%
Coef. Var.	5,18%	2,25%	0,00%	2,59%	0,94%	1,95%	1,70%
Desvio Pad.	0,05%	0,03%	0,00%	0,02%	0,01%	0,02%	0,01%

Na média o menor erro relativo encontrado foi 0,8% para a combinação PG com RNA com RL. O maior erro foi de 1,3% para o modelo de previsão RNA. Os resultados que mais variaram foram os do modelo PG com um coeficiente de variação de 5,1%, assim como também foi o modelo PG que obteve o maior desvio padrão. Assim como os resultados do Bradesco, as melhores médias e as menores variações de resultados são provenientes das combinações dos modelos. Nos resultados do teste estatístico é possível verificar quais modelos e combinações para as ações do Itaú-Unibanco são equivalentes.

A Figura 16 apresenta o gráfico para os resultados encontrados para a previsão do valor médio diário das ações da empresa Petrobrás. O modelo PG apresentou uma melhor distribuição dos resultados, possível de se identificar nos tamanhos semelhantes dos quartis. O modelo PG apresentou 2 *outliers*, o que mostra a existência de resultados diferentes dos demais.

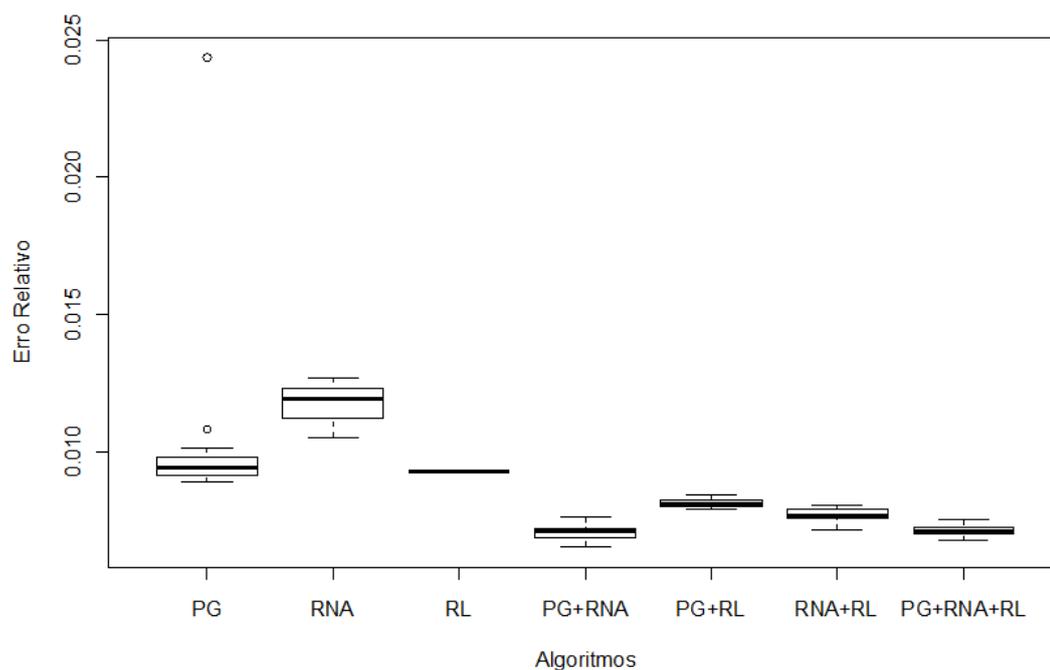


Figura 16 - Gráfico dos resultados para o papel PETR4

A Tabela 7 apresenta outras informações relevantes sobre os resultados para as previsões do valor médio diário das ações da Petrobrás. O menor erro relativo na média foi obtido com a combinação PG com RNA com erro de 0,7%. O maior erro relativo, assim como os demais, foi obtido com o modelo de previsão RNA com erro de 1,2%. O maior coeficiente de variação e o maior desvio padrão foram do modelo de previsão PG com os valores 28,2% e 0,3% respectivamente.

Tabela 7 - Informações dos resultados para Petrobras

	<i>PG</i>	<i>RNA</i>	<i>RL</i>	<i>PG+RNA</i>	<i>PG+RL</i>	<i>RNA+RL</i>	<i>PG+RNA+RL</i>
Média	0,99%	1,18%	0,93%	0,7%	0,81%	0,77%	0,71%
Coef. Var.	28,22%	5,25%	0,00%	3,67%	1,81%	3%	2,63%
Desvio Pad.	0,28%	0,06%	0,00%	0,03%	0,01%	0,02%	0,02%

A Figura 17 apresenta o gráfico com os resultados das previsões para a empresa Usiminas. Os resultados mostraram vários *outliers* para o modelo de PG, resultados com os maiores erros para o modelo RNA, assim como para os demais papéis, uma distribuição mais equivalente para os quartis da combinação RNA com RL e os menores erros, assim como para os demais papéis, para as combinações dos modelos.

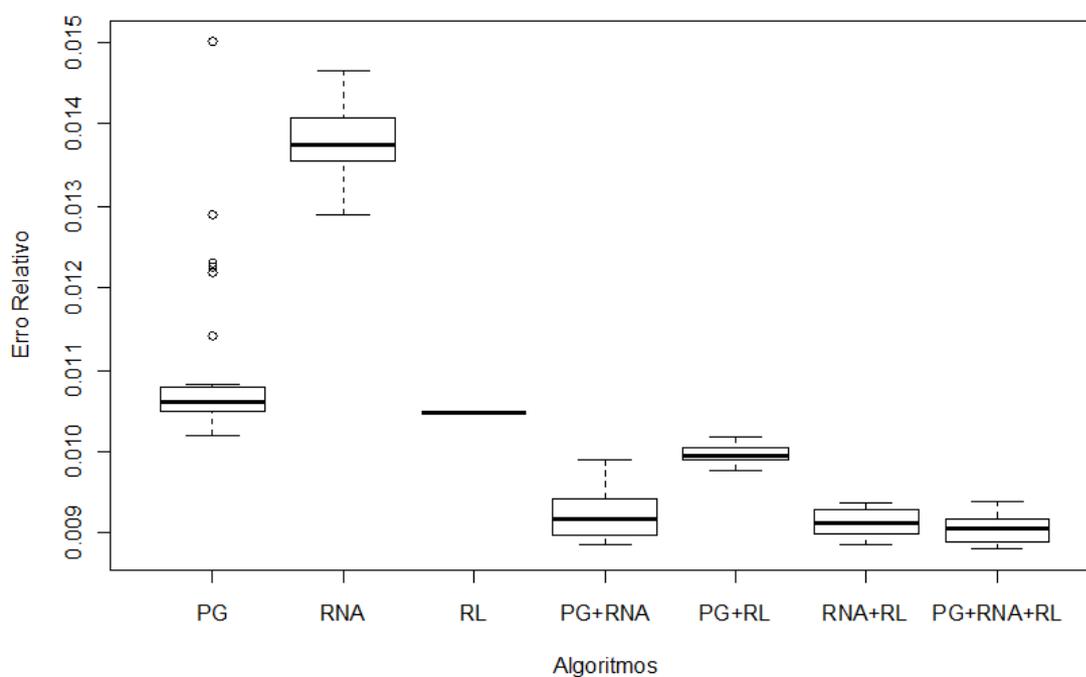


Figura 17 - Gráfico dos resultados para o papel USIM5

A Tabela 8 descreve outras informações dos resultados das previsões para a empresa Usiminas, que obteve um menor erro relativo médio com a combinação dos modelos PG, RNA e RL com erro de 0,9%. Seguindo os outros papéis, o maior erro relativo médio foi do modelo RNA com erro de 1,4%. O coeficiente de variação maior foi do modelo PG de 9,4%, assim como o maior desvio padrão de 0,1%.

Tabela 8 - Informações dos resultados para Usiminas

	PG	RNA	RL	PG+RNA	PG+RL	RNA+RL	PG+RNA+RL
Média	1,1%	1,38%	1,05%	0,92%	1%	0,91%	0,9%
Coef. Var.	9,39%	2,97%	0,00%	2,82%	1,07%	1,70%	1,88%
Desvio Pad.	0,1%	0,04%	0,00%	0,03%	0,01%	0,02%	0,02%

A Figura 18 apresenta o gráfico dos resultados das previsões realizadas para a empresa Vale do Rio Doce, onde, de forma geral não houve grandes variações nos resultados dos modelos e das combinações destes. Assim como na maioria dos papéis o modelo PG apresentou *outliers* que demonstram resultados fora do comum para este modelo. Além do modelo PG, as combinações RG com RNA e PG com RL também apresentaram alguns *outliers*. O modelo que apresentou uma melhor distribuição dos resultados foi o RNA, que obteve também os maiores erros relativo. Assim como para os demais papéis as combinações dos modelos apresentaram os menores erros relativo.

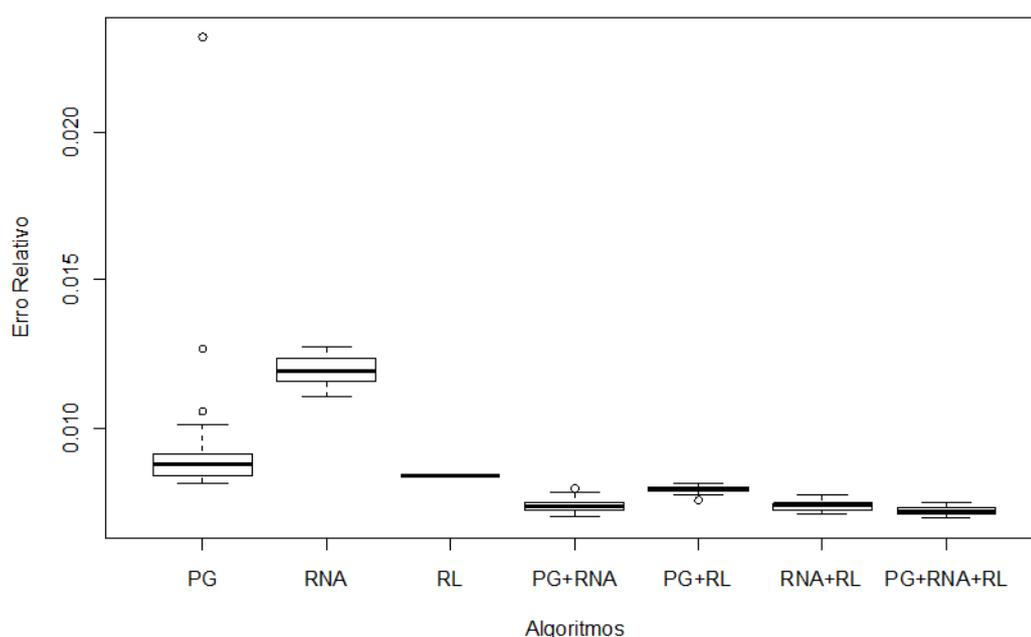


Figura 18 - Gráfico dos resultados para o papel VALE5

A Tabela 9 apresenta outras informações dos resultados para a Vale do Rio Doce, onde, o menor erro relativo foi de 0,7%, encontrado pela combinação PG, RNA e RL. Assim como todos os papéis o maior erro relativo médio foi do modelo RNA. Os resultados que apresentaram maior variação foram os do modelo PG com o coeficiente de variação de 29,3%, o que foi causado pelos *outliers* existentes. O maior desvio padrão é de 0,2% e está presente, também, no modelo PG.

Tabela 9 - Informações dos resultados para Vale do Rio Doce

	<i>PG</i>	<i>RNA</i>	<i>RL</i>	<i>PG+RNA</i>	<i>PG+RL</i>	<i>RNA+RL</i>	<i>PG+RNA+RL</i>
Média	0,95%	1,19%	0,84%	0,74%	0,80%	0,74%	0,72%
Coef. Var.	29,39%	3,91%	0,00%	2,98%	1,39%	2,18%	1,84%
Desvio Pad.	0,28%	0,05%	0,00%	0,02%	0,01%	0,02%	0,01%

Houve nos resultados algumas semelhanças interessantes. A maioria dos resultados do modelo PG tiveram *outliers*, o que mostra que pode existir algum problema no algoritmo do modelo, passível de um outro estudo. Para todos os papéis os maiores erros foram do modelo RNA. Para todos os papéis os melhores resultados estiveram entre as combinações dos modelos. Entretanto, é através do teste estatístico apresentado na Seção 5.3 que será possível confirmar a hipótese H_0 .

5.3 Resultados dos Testes Estatísticos

O teste estatístico de Tukey realizado com os resultados das previsões para cada empresa foi gerado com um nível de 5% de significância, sendo que o modelo RL não foi utilizado no teste por ser um modelo determinístico. Cada tabela possui a coluna “grupos” que mostra o agrupamento entre os tratamentos, representando a equivalência na execução dos modelos ou técnicas que estão agrupados.

A análise de variância apresentou valores inferiores a 5% para todos os papéis o que mostra que existem diferenças significativas entre os tratamentos.

A Tabela 10 apresenta os resultados do teste de Tukey para as ações do Itau-Unibanco. É possível verificar que os tratamentos 4 e 5 e os tratamentos 5 e 6 possuem equivalência, logo, estes pertencem aos mesmos grupos, isso ocorre pois

caso os algoritmos sejam executados novamente os resultados dentro dos grupos podem se inverter. Comparando-se as médias de tratamento é possível verificar que o tratamento 6 obteve a menor média com 0,84%, o que mostra que a combinação PG, RNA e RL e a combinação PG e RNA se equivalem e são as melhores para as previsões do preço médio diário das ações do Itau-Unibanco.

Tabela 10 - Teste de Tukey ITUB4

<i>Grupos</i>	<i>Tratamentos</i>	<i>Médias</i>
a	RNA(1)	1,30%
b	PG(2)	1,05%
c	PG+RL(3)	0,93%
d	RNA+RL(4)	0,87%
de	PG+RNA(5)	0,86%
e	PG+RNA+RL(6)	0,84%

A Tabela 11 apresenta o teste de Tukey para as previsões da empresa Bradesco, onde foram agrupados os tratamentos 3 e 4 e os tratamentos 5 e 6. Sendo que, comparando as médias de tratamento os melhores resultados foram do grupo com os tratamentos 5 e 6 representado pelos resultados das combinações PG, RNA e RL e PG com RNA os dois tratamentos com médias de tratamento de 0,79%.

Tabela 11 - Teste de Tukey BBDC4

<i>Grupos</i>	<i>Tratamentos</i>	<i>Médias</i>
a	RNA(1)	1,21%
b	PG(2)	0,94%
c	RNA+RL(3)	0,85%
c	PG+RL(4)	0,84%
d	PG+RNA(5)	0,79%
d	PG+RNA+RL(6)	0,79%

A Tabela 12 apresenta os resultados do teste de Tukey para as previsões da empresa Petrobras. Neste teste foi possível verificar que os tratamentos 3 e 4 se equivalem, assim como os tratamentos 4, 5 e 6 que fazem parte do grupo com os melhores resultados com todos os tratamentos sendo combinações dos modelos e médias de tratamento entre 0,7% e 0,71%.

Tabela 12 - Teste de Tukey PETR4

<i>Grupos</i>	<i>Tratamentos</i>	<i>Médias</i>
a	RNA(1)	1,18%
b	PG(2)	0,99%
c	PG+RL(3)	0,81%
cd	RNA+RL(4)	0,77%
d	PG+RNA+RL(5)	0,71%
d	PG+RNA(6)	0,7%

A Tabela 13 apresenta o teste estatístico realizado com os resultados das previsões para a Usiminas. Neste teste, os tratamentos 4, 5 e 6 passaram a pertencer ao mesmo grupo, possuindo resultados que podem inverter as ordens das menores médias dentro do grupo, porém sendo este grupo o que possui as menores médias de tratamento e sendo composto somente de combinações dos modelos de previsão.

Tabela 13 - Teste de Tukey USIM5

<i>Grupos</i>	<i>Tratamentos</i>	<i>Médias</i>
a	RNA(1)	1,38%
b	PG(2)	1,1%
c	PG+RL(3)	1%
d	PG+RNA(4)	0,92%
d	RNA+RL(5)	0,91%
d	PG+RNA+RL(6)	0,9%

A Tabela 14 apresenta o teste de Tukey para as previsões da empresa Vale do Rio Doce. Neste teste, todas as combinações foram agrupadas, possuindo equivalência nos resultados, mostrando também que a utilização da máquina de comitê apresentou melhores resultados que os modelos isoladamente.

Tabela 14 - Teste de Tukey VALE5

<i>Grupos</i>	<i>Tratamentos</i>	<i>Médias</i>
a	RNA(1)	0,0120
b	PG(2)	0,0094
c	PG+RL(3)	0,0079
c	RNA+RL(4)	0,0074
c	PG+RNA(5)	0,0074
c	PG+RNA+RL(6)	0,0072

Com os testes apresentados acredita-se que estatisticamente a hipótese H_0 é válida para os papéis estudados. Pois, em todos os casos, as combinações de modelos geradas obtiveram melhores resultados que os modelos isoladamente. Além disso, várias das combinações, dependendo do papel, possuíram equivalência nos resultados, o que mostra que de acordo com o papel é possível utilizar várias combinações e obter respostas semelhantes.

6

Sistema *Online* de Apoio a Decisão ao Investidor

Este capítulo apresenta um sistema online desenvolvido para complementar qualquer método de previsão que gere previsões para operações diárias de compra e venda. A Seção 6.1 descreve o objetivo do Sistema. A Seção 6.2 descreve os usuários permitidos ao Sistema, junto ao perfil de cada um. A Seção 6.3 descreve as funcionalidades do Sistema separadas por tipo de usuário. Em seguida nas Seções 6.4 e 6.5 são apresentados e descritos os diagramas de classe e de entidade e relacionamento do sistema, respectivamente. A Seção 6.6 descreve o funcionamento do sistema e as peculiaridades que devem ser tratadas pelos métodos de previsão que utilizarão o sistema. A Seção 6.7 finaliza o capítulo apresentando a discussão final e possíveis trabalhos futuros.

6.1 Objetivos do Sistema

A partir do momento que existem modelos de previsão responsáveis por gerar previsões dos preços futuros das ações no mercado de ações, surge a necessidade de um sistema que apresente de uma forma simples para o usuário final os resultados gerados pelas previsões. Com isso, o objetivo do sistema *online* de apoio à decisão ao investidor é fornecer uma interface entre o usuário e qualquer previsor que se deseje utilizar para disponibilização de sugestões diárias de valores para compra e venda de ações, sendo que, o sistema foca em sugestões para operações do tipo *day trading*. Entre outros objetivos estão: possibilitar ao investidor manter uma carteira de investimentos com os papéis da BOVESPA que deseja acompanhar as previsões; agendar análises futuras destes papéis definindo quais riscos entre baixo, médio e alto o investidor deseja correr; visualizar o desempenho passado dos papéis através de gráficos e tabelas; realizar simulações de lucro com a utilização do sistema.

Com as previsões sugeridas pelo sistema é necessário apenas inseri-las no livro de ofertas do *homebroker* de preferência do usuário e aguardar até o final do pregão, onde o sistema informa se a previsão sugerida obteve sucesso ou não para cada papel solicitado. Na seção 6.6 é possível acompanhar todo o fluxo de funcionamento do Sistema com os possíveis imprevistos e os tratamentos para estes.

6.2 Descrição dos Usuários

O sistema permite a interação com dois tipos de usuários: um chamado Investidor, caracterizado por clientes iniciantes ou profissionais que buscam realizar negociações utilizando as previsões do sistema, e o outro chamado Administrador, que além de poder realizar negociações pode gerenciar os usuários, acessos, papéis mais rentáveis entre outras coisas. Na Seção 6.3 são descritas todas as funcionalidades de acordo com o perfil do usuário.

6.3 Funcionalidades

6.3.1 Funcionalidades do Investidor

6.3.1.1 Cadastro

O usuário Investidor pode realizar o cadastro para ter acesso ao sistema e suas funcionalidades informando, seu nome, CPF, telefone, *email* e senha. A Figura 19 mostra a tela de cadastro com as informações necessárias.

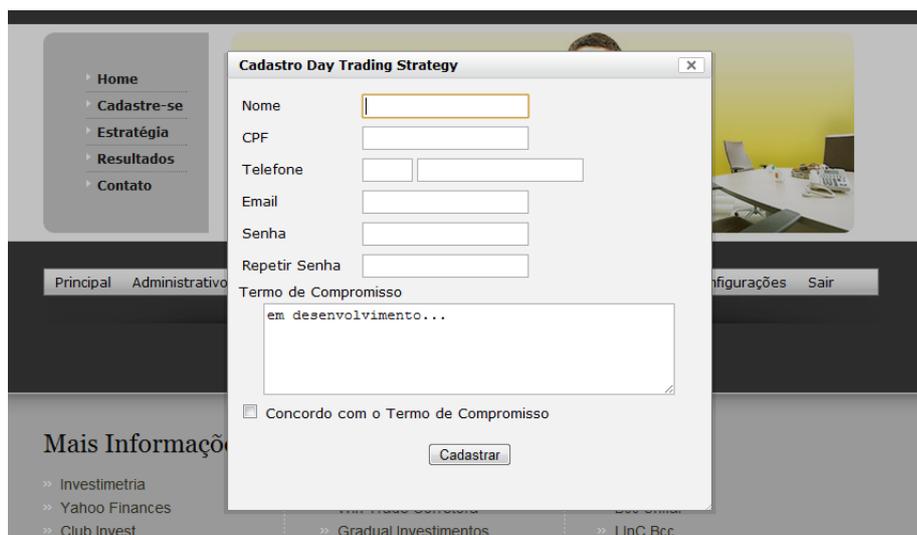


Figura 19 - Tela de Cadastro do Investidor

6.3.1.1 Gerenciamento da Carteira

O sistema disponibiliza, para cada usuário, uma carteira de investimentos onde é possível adicionar, alterar e remover papéis que poderão ser utilizados para agendamento de análises futuras, além da possibilidade de acompanhamento da lucratividade com a utilização do sistema.

Ao adicionar um novo papel em sua carteira o usuário precisa escolher os riscos que espera correr quando for agendar análises, sendo que, existem três possibilidades de risco, baixo, médio e alto, aumentando as possibilidades de grandes ganhos ou perdas de acordo com o aumento do risco. O usuário pode escolher para cada papel que adiciona mais de uma possibilidade risco. A Figura 20 apresenta a Tela de Cadastro dos Papéis na Carteira.

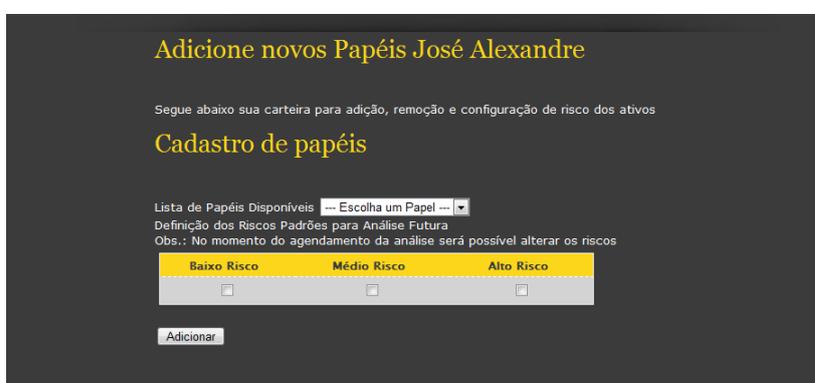


Figura 20 - Cadastro de Novos Papéis

A Figura 21 apresenta a Tela para Gerenciamento da Carteira, onde é possível alterar os riscos dos papéis ou excluí-los.



Figura 21 - Tela de Gerenciamento da Carteira

6.3.1.1 Gerenciamento de Análises Futuras

Uma Análise Futura é caracterizada por um período com uma data inicial, uma data final e uma lista de papéis dos quais o usuário Investidor deseja receber as previsões do Sistema.

No Gerenciamento de Análises Futuras o usuário pode agendar novas análises e visualizar as existentes. Ao agendar uma nova análise, o usuário deve informar o período, os papéis e os riscos para cada papel. Inicialmente são apresentados os riscos que foram definidos quando o usuário adicionou os papéis na carteira. A Figura 22 apresenta a tela para agendamento de análises.



Figura 22 - Tela de Agendamento de Análise

A visualização das análises agendadas apresenta para cada análise o seu *status*, podendo ser aguardando, executando e finalizado. O *status* aguardando mostra que as análises ainda não começaram a serem apresentadas ao usuário, este *status* é mantido até a data inicial ser atingida. O *status* executando mostra que a análise está em execução, ou seja, esta entre a data inicial e a data final, este *status* é mantido até a data final ser atingida, então o *status* é alterado automaticamente para finalizado. A Figura 23 apresenta a Tela de Visualização das Análises Agendadas com seus respectivos status.

Análises Agendadas

Período Inicial	Período Final	Ações	Status
2010-03-01	2010-03-01	USIM5 VALES	finalizado
2010-03-19	2010-08-31	BBDC4 GGBR4 PETR4 USIM5 VALES	executando
2010-04-16	2010-04-30	PETR4 USIM5 VALES	finalizado
2010-04-15	2010-04-22	USIM5	finalizado
2010-04-17	2010-09-24	PETR4 USIM5	executando

Agendar Nova Análise Futura

Figura 23 - Tela de Visualização das Análises Agendadas

6.3.1.1 Visualização das previsões do dia atual

Ao entrar no sistema o usuário Investidor encontra como primeira informação as previsões das análises agendadas para o dia atual. Caso não existam análises agendadas para o dia atual, a tabela de previsões aparecerá vazia.

Depois de visualizar as previsões dos papéis no sistema o usuário deve ir até seu *homebroker* e inserir no livro de ofertas as sugestões de compra e venda. Ao final do dia, caso o valor de compra ou de venda não tenha sido atingido, o usuário deve tomar as medidas descritas na seção 6.6.

A Figura 24 apresenta a Tela de Visualização das análises agendadas pelo usuário.



Data	Papel	Risco	Compra	Venda	Máximo Real	Mínimo Real	Fechamento	Retorno(%)
19/03/2010	BBDC4	médio	31.88	32.04				
19/03/2010	GGBR4	médio	27.77	27.98				
19/03/2010	PETR4	médio	37.0	37.09				
19/03/2010	USIM5	médio	54.93	56.36				
19/03/2010	VALE5	médio	47.4	47.67				

Análises Passadas

Figura 24 - Tela de Visualização das Análises Agendadas

6.3.1.1 Visualização do Histórico de previsões

As previsões passadas para a carteira do Investidor podem ser visualizadas em uma tela com todas as análises e seus respectivos retornos. Os retornos são calculados com base nas regras definidas na Seção 6.6. A Figura 25 apresenta um exemplo da tela de visualização das análises passadas.

Análises passadas para seus papéis

<< first < prev 1 2 3 4 5 6 7 8 9 10 next >>

Data	Papel	Risco	Ordem de Compa	Ordem de Venda	Fechamento	Retorno(%)
19/03/2010	BBDC4	médio	31.88	32.04	31.5	-1.19
18/03/2010	BBDC4	médio	31.58	31.84	31.71	0.41
17/03/2010	BBDC4	médio	32.15	32.34	31.75	-1.24
16/03/2010	BBDC4	médio	31.5	31.74	32.03	0.76
15/03/2010	BBDC4	médio	31.49	31.76	31.75	0.86
12/03/2010	BBDC4	médio	32.24	32.42	31.82	-1.3
11/03/2010	BBDC4	médio	31.67	31.94	32.15	0.85
10/03/2010	BBDC4	médio	31.82	32.05	31.98	0.72

<< first < prev 1 2 3 4 5 6 7 8 9 10 next >>

Figura 25 - Tela de Visualização das Análises Passadas

6.3.1.1 Análise dos últimos 30 dias

Os lucros obtidos com as sugestões do sistema, em longo prazo, podem ser visualizados de duas formas: através de um gráfico dinâmico com a lucratividade dos papéis nos últimos 30 dias e através de uma tabela com a lucratividade média dos papéis nos últimos 30 dias. Através destas informações é possível analisar os papéis que estão apresentando uma maior rentabilidade, tornando possível, assim, melhorar a tomada de decisões do Investidor com relação ao montante que deve ser investido em cada papel que ele acompanha.

A Figura 26 apresenta um exemplo da tela de Análise de Lucros para uma carteira com cinco papéis.

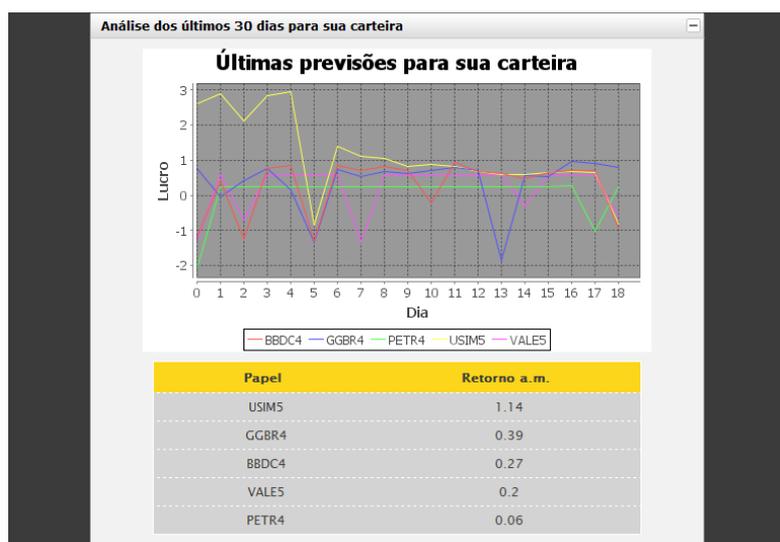


Figura 26 - Tela de Análise dos Últimos 30 dias

6.3.1.1 Simulações

Através do sistema é permitido ao usuário Investidor a realização de simulações de lucro obtidos com a utilização das previsões do sistema. O usuário informa além do montante e o papel que deseja simular o período inicial, o período final, a retirada ou não da taxa de emolumento e o tipo de corretagem que será cobrado durante a simulação, sendo que, os valores utilizados para calcular os gastos com emolumento e com corretagem variável são padrões da BOVESPA presentes na Seção 2.4.

A possibilidade de realizar simulações de lucro apresenta para o usuário uma forma de saber o quanto ele poderia estar ganhando ou perdendo em um determinado período para cada papel com a utilização do sistema, lembrando que são descontadas as taxas de corretagem e de emolumento (opcional) apresentando ao final para o usuário o lucro líquido obtido.

A Figura 27 apresenta a tela de configurações para realização de simulações de lucro.

Simule seus lucros com o Day Trading Strategy!

Configurações de Corretagem

Tipo de Corretagem

Fixa Variável

Valor 0.0

Configurações da Simulação

Valor do Investimento (R\$) 0.0

Retirar emolumento

Escolha um Papel --- Escolha um Papel ---

Início da Simulação

Fim da Simulação

Simular!

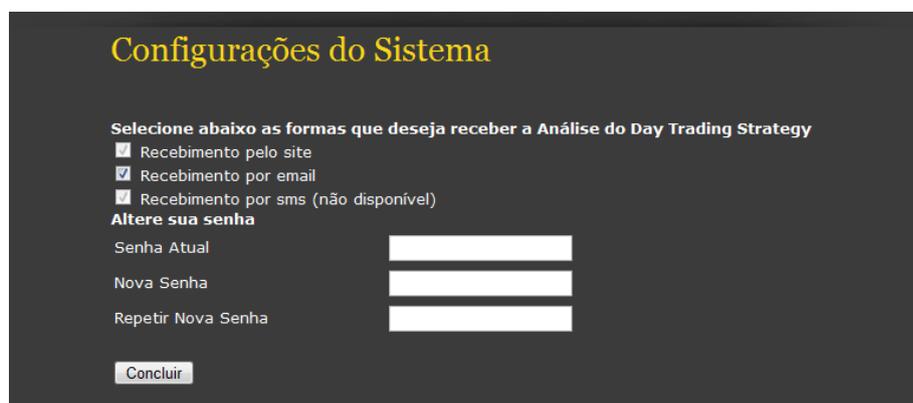
Figura 27 - Tela de Simulações de Lucro

6.3.1.1 Configurações

O usuário Investidor pode ainda configurar as formas através das quais deseja receber suas análises agendadas, estando disponível sempre o recebimento através

do site e opcionalmente o recebimento por *email*. O recebimento via sms¹ não foi desenvolvido. Ainda está disponível para o usuário nas configurações, a alteração de sua senha, digitando a senha atual uma vez e a nova senha duas vezes.

A Figura 28 apresenta a tela de configurações do usuário Investidor.



Configurações do Sistema

Selecione abaixo as formas que deseja receber a Análise do Day Trading Strategy

- Recebimento pelo site
- Recebimento por email
- Recebimento por sms (não disponível)

Altere sua senha

Senha Atual

Nova Senha

Repetir Nova Senha

Figura 28 - Tela de Configurações

As várias formas possíveis de recebimento das análises agendadas tornam o processo de visualização de previsões mais agradável e fácil para o usuário. Lembrando que o usuário deve inserir em seu *homebroker* logo no início do dia de operações, as análises agendadas, devido ao risco dos valores previstos serem atingidos logo no início do pregão, como é descrito na seção 6.6.

6.3.1.1 Contato

O usuário pode entrar em contato com os desenvolvedores do sistema para mais informações, parcerias ou sugestões. Para entrar em contato o usuário precisa informar seu nome, *email*, a informação que deseja tratar e sua mensagem. Estas informações serão enviadas para o *email* da equipe que analisa e retorna uma resposta.

¹ *short message service* - Serviço disponível em telefones celulares digitais que permite o envio de mensagens curtas entre estes equipamentos e entre outros dispositivos de mão.

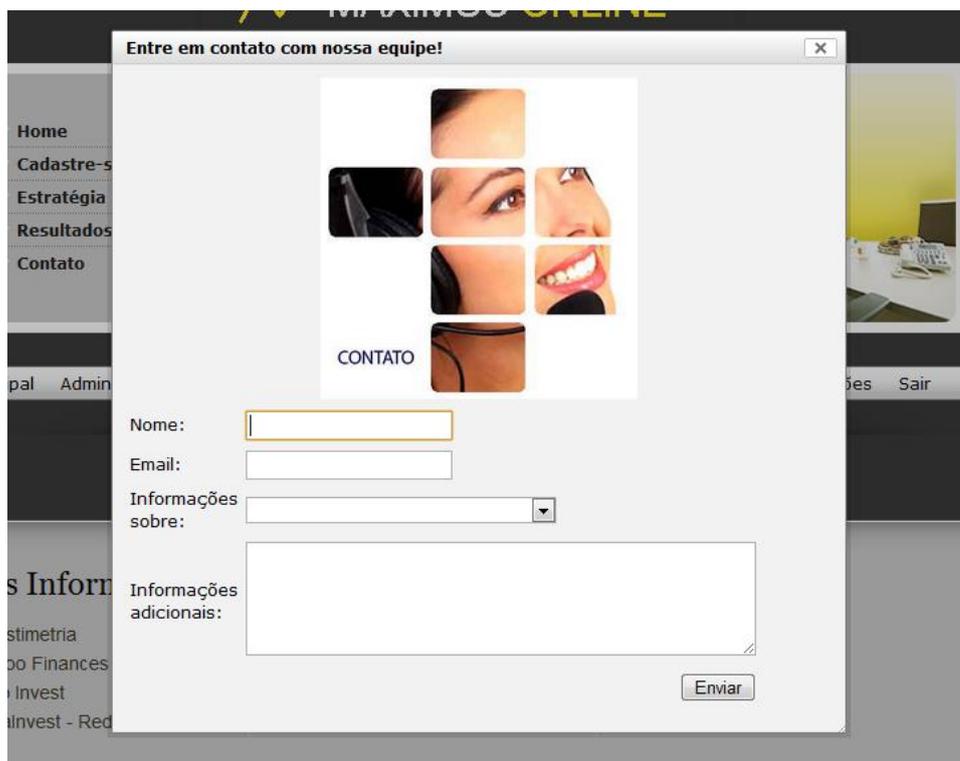


Figura 29 - Tela para Contato

6.3.2 Funcionalidades do Administrador

O usuário Administrador, além das funcionalidades básicas do usuário Investidor, possui uma visão gerencial do sistema, com permissão para funcionalidades, descritas a seguir, que mostram informações úteis para a gerência do sistema.

6.3.2.1 Gerenciar Usuários

O usuário Administrador pode atribuir permissão de administrador a usuários Investidores. Além disso, pode excluir usuários, verificar o total de usuários no sistema, verificar o usuário com maior número de acessos, verificar o usuário com maior número de agendamentos e, por fim, verificar o usuário com maior número de análises agendadas, onde é feita uma conta com a soma do total de análises de cada agendamento de acordo com o número de papéis e período escolhido.

Esta funcionalidade permite ao usuário Administrador ter um controle dos usuários cadastrado e dos usuários que mais utilizam o sistema, tornando possível o oferecimento de pacotes e promoções para estes.

6.3.2.1 Rentabilidade dos Papéis

Os papéis mais rentáveis podem ser visualizados em duas tabelas: uma com os papéis mais rentáveis do último dia de análises e outra com a média de rentabilidade dos últimos trinta dias de análise. Nos dois casos são apresentados os cinco principais papeis de cada lista. A importância destas listas advém do fato de tornar possível as sugestões de melhores carteiras para o usuário Investidor, facilitando assim a escolha de papéis para sua carteira.

6.3.2.1 Melhores Investidores

Os melhores investidores podem ser caracterizados de três formas diferentes: na primeira, é apresentada uma lista com os dez investidores que mais realizam análises, mostrando assim, os usuários que mais estão utilizando o sistema. Na segunda, é apresentada uma lista com os dez investidores com melhores carteiras no último dia de análise. Com esta lista é possível visualizar as combinações de papéis que estão gerando maiores retornos, o que possibilita no futuro sugeri-las a outros investidores. Na terceira lista são apresentados os dez investidores com melhor carteira nos últimos 30 dias, mostrando de uma forma mais estável as melhores coleções de papéis utilizadas.

6.3.2.1 Acessos ao Sistema

Nesta funcionalidade é apresentado o total de acessos ao sistema e os cinco investidores com o maior número de acessos, possibilitando assim, verificar o andamento dos números de acessos ao mesmo.

6.4 Documentação de Classes

A partir das funcionalidades do sistema apresentadas, foi elaborado, utilizando os princípios da linguagem de modelagem UML, o diagrama de classes, que segundo Bezerra (2007, p. 95), “permite compreender como o sistema está estruturado internamente para que as funcionalidades externamente visíveis sejam produzidas”. Uma versão simplificada deste diagrama contendo as classes no domínio do negócio em questão é apresentada na Figura 30, pois é mais simples e de fácil compreensão do que os diagramas de classes de especificação e de implementação.

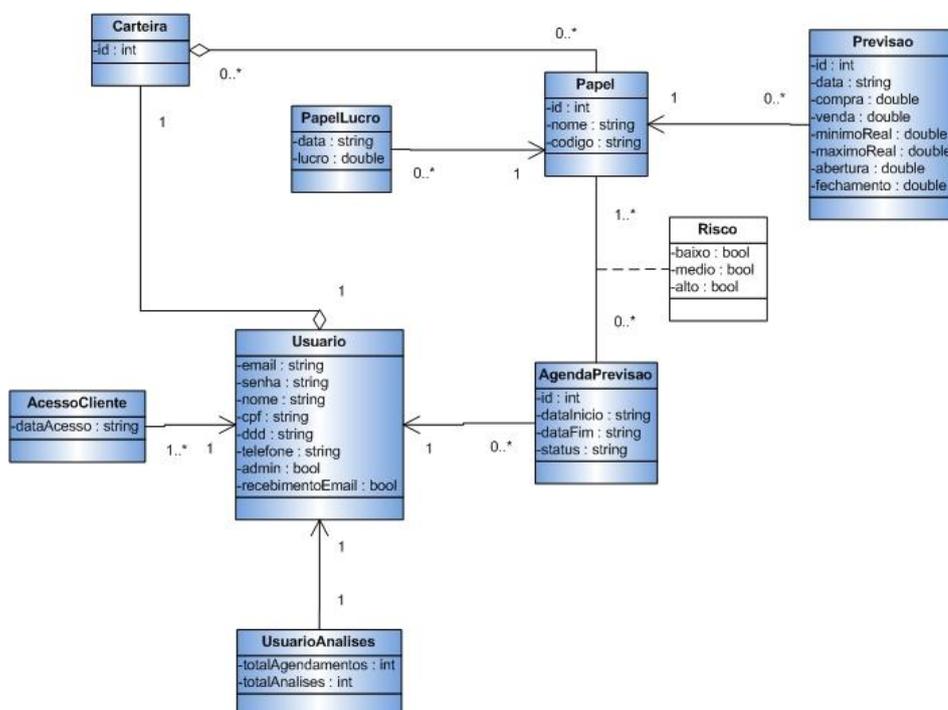


Figura 30 - Diagrama de Classes do Sistema

O diagrama de classes do domínio mostra a relação entre as classes do sistema. Nele, um usuário possui uma carteira, pode realizar de nenhum até muitos agendamentos de previsão, pode efetuar de um até muitos acessos ao sistema e alcança um total de agendamentos e um total de análises. Uma carteira no sistema é composta de nenhum até muitos papéis e um papel apresenta de nenhuma até muitas previsões e respectivamente de nenhum até muitos lucros de acordo com as previsões. Um agendamento de previsões possui de um até muitos

papeis e cada papel do agendamento de previsão possui os riscos que o Usuário deseja correr com este.

O diagrama das classes do domínio, além de auxiliar na compreensão do sistema com o relacionamento entre as classes, mostra os atributos e a multiplicidade de cada classe facilitando, assim, na construção do diagrama de entidades e relacionamento utilizado para a construção do Banco de Dados e explicado a seguir na Seção 6.5.

6.5 Documentação do Banco de Dados

O diagrama de entidade e relacionamento é a base para a construção do banco de dados do sistema e surge do mapeamento das classes definidas no diagrama de classe para o modelo relacional. A Figura 31 apresenta o diagrama de entidade e relacionamento do sistema.

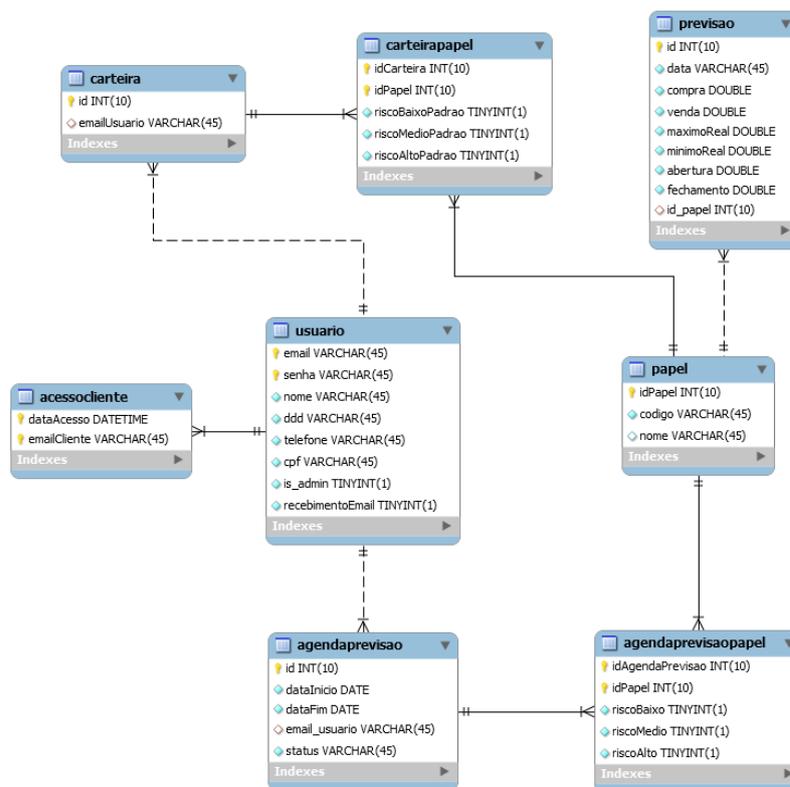


Figura 31 - Diagrama de Entidade e Relacionamento do Sistema

O diagrama de entidade e relacionamento do sistema apresenta as tabelas que são necessárias armazenar no banco de dados para persistir os dados dos usuários e as previsões realizadas por um modelo de previsão. Através do banco de dados que ocorre a integração dos modelos de previsão com o sistema, sendo, a tabela previsão a responsável pela integração, pois, é utilizada pelo modelo de previsão para persistir as previsões, sendo inserido nela os valores de compra e venda sugeridos pelo previsor, enquanto o Sistema utiliza esta tabela para apresentar as sugestões aos usuários de acordo com suas carteiras e análises agendadas.

A interação entre o modelo de previsão e o sistema online de apoio ao investidor é apresentado a seguir na Seção 6.6.

6.6 Funcionamento do Sistema

O funcionamento do sistema se baseia na interação deste com seu banco de dados, o usuário e um previsor. Entretanto, estes elementos devem respeitar a uma sequência básica diária, no sentido de possibilitar o correto funcionamento do sistema.

O primeiro acontecimento básico cabe ao previsor utilizado pelo sistema. Este previsor deve ser capaz de inserir até o momento da abertura do pregão as sugestões de compra e venda de todos os papéis da BOVESPA para o dia. Este é um acontecimento fundamental para o correto funcionamento, pois, os valores sugeridos podem ocorrer logo no início do pregão e caso o banco de dados não esteja populado com as sugestões no início, o uso destas para o dia estará comprometido.

O segundo acontecimento deve ocorrer exatamente em seguida à inserção das sugestões no banco de dados e diz respeito ao envio das sugestões para o *email* dos usuários que desejam recebê-las por *email*. Este acontecimento deve ser realizado pelo previsor para que seja imediata à geração das sugestões.

O terceiro acontecimento diz respeito ao usuário que deve visualizar logo no início do pregão as sugestões geradas para suas análises agendadas e inseri-las em seu *homebroker*. Feito isso o usuário deve aguardar até o final do pregão para saber se as ordens de compra e venda foram realizadas com sucesso, ou se houve algum imprevisto durante o dia. As possibilidades de ocorrência e as decisões que devem ser tomadas pelo usuário são apresentadas a seguir:

- **Compra seguida da venda:** Neste caso, a negociação é realizada com os preços sugeridos pelo sistema, sem nenhum imprevisto para ser tratado.
- **Venda seguida da compra:** Caso o valor sugerido para venda seja atingido antes do valor de compra surge a necessidade da realização de uma venda a descoberto, onde são vendidas ações sem possuí-las, com o comprometimento de comprá-las até o final do pregão. Após a venda a descoberto o valor de compra sugerido é atingido, fechando o pregão com sucesso.
- **Somente compra:** Quando, após todo o pregão, somente o preço de compra é atingido, o usuário deve realizar uma venda a mercado ao fim do pregão, onde a negociação é realizada com o valor que ação estiver valendo, podendo assim, fechar o dia com lucro ou prejuízo de acordo com o preço de fechamento da ação.
- **Somente venda:** Quando, após todo o pregão, somente o preço de venda é atingido, o usuário deve realizar uma compra a mercado ao fim do pregão, onde a negociação é realizada com o valor que ação estiver valendo, podendo assim, fechar o dia com lucro ou prejuízo de acordo com o preço de fechamento da ação.
- **Nem a compra nem a venda:** Neste caso, não são efetuadas negociações, logo, não existem lucros nem prejuízos.

Ao final do pregão o sistema é atualizado com as informações dos valores de máximo, mínimo e fechamento das ações, com os quais é possível gerar os cálculos que medem os lucros obtidos para cada um dos papéis de acordo com a realização ou não das operações de compra e venda.

6.7 Discussão Final do Capítulo

A estratégia *day trading* definida pelo sistema para compra e venda diária de ações possui algumas falhas que não foram tratadas neste trabalho. Em seguida são apresentados duas destas falhas que podem ser corrigidas em trabalhos futuros.

A primeira é que no período inicial dos pregões, normalmente, ocorrem grandes variações nos preços das ações, o que pode levar, de acordo com o modelo de previsão utilizado e o treinamento deste modelo, ao fato dos preços de compra e venda serem atingidos antes do usuário receber as sugestões e inseri-las no livro de ofertas. Uma das formas de diminuir este problema pode estar na utilização dos dados *intra-day*, que consistem das informações minuto a minuto dos preços e negociações das ações, para o treinamento dos modelos de previsão, buscando acostamá-los sem os dados iniciais do pregão.

A segunda falha está na não utilização da técnica de *stop móvel*, que poderia alavancar os lucros diários do usuário a partir do momento em que os valores de compra ou de venda sugeridos foram atingidos. Esta técnica pode ser considerada com a utilização da base *intra-day* pelos modelos de previsão utilizados.

Este trabalho não pode resolver estas falhas, pois não dispunha dos dados *intra-day* das ações para realizar experimentos com os modelos de previsão utilizados.

7

Conclusões

Este capítulo apresenta as conclusões desta monografia e está organizado da seguinte forma: Na Seção 7.1 são apresentadas as conclusões para os experimentos e testes estatísticos realizados e na Seção 7.2 são apresentadas as conclusões para o Sistema Online para Apoio a Decisão ao Investidor desenvolvido neste trabalho.

As conclusões deste trabalho estão divididas em conclusões dos experimentos realizados para validar estatisticamente a hipótese H_0 e conclusões do Sistema Online para Apoio a Decisão do Investidor e são apresentadas em seguida.

7.1 Conclusões dos Experimentos

A realização dos experimentos para a previsão o valor médio diário de alguns dos papéis da BOVESPA permitiu verificar que estatisticamente a hipótese H_0 é válida, ou seja, a utilização das máquinas de comitê geram resultados melhores ou no mínimo tão bons quanto o melhor resultado dos modelos executados isoladamente. Esta comprovação não permite que se possa dizer que as máquinas de comitê funcionarão bem para a previsão dos demais papéis da BOVESPA nos quais não foram realizados experimentos.

O teste de Tukey realizado, além de validar estatisticamente a hipótese H_0 , mostrou que várias das máquinas de comitê desenvolvidas se equivalem, pois, foram atribuídas a grupos iguais no teste. Porém, as máquinas que se equivalem variam de acordo com os papéis, chegando a haver papéis como o da empresa Vale do Rio Doce onde todas as máquinas de comitê eram equivalentes. Em 80% dos papéis as menores médias de tratamento encontradas foram da combinação PG, RNA e RL.

A conclusão mais importante é que a utilização das máquinas de comitê melhoraram os resultados em todos os papéis que foram realizados experimentos,

sendo que o teste de Tukey mostrou que os resultados das máquinas de comitê não são equivalentes aos resultados de nenhum dos modelos isoladamente, deixando claro que os resultados das máquinas de comitê são realmente melhores.

7.2 Conclusões do Sistema *Online*

O sistema *online* desenvolvido foi uma tentativa de apresentar de uma forma mais agradável ao usuário final as previsões para operações de compra e venda diárias realizadas por modelos e técnicas computacionais, além de tornar possível o acompanhamento das previsões dos papéis de sua carteira, mostrando sempre os lucros possíveis de serem obtidos com os modelos e técnicas de previsão utilizados.

A criação deste sistema permite que qualquer usuário com acesso a Internet possa acompanhar as sugestões de compra e venda de ações feitas por modelos computacionais, complementando assim, as máquinas de comitê criadas neste trabalho para a previsão do preço médio diário das ações.

8

Referências Bibliográficas

Assaf Neto, A. *Mercado Financeiro*. 7. ed. São Paulo: Atlas, 2006.

Bezerra, E. *Princípios De Análise E Projeto De Sistemas Com UML*. Rio de Janeiro: Elsevier, 2007.

Bojarczuk, C. C.; Lopes, H. S.; Freitas, A. A. *A constrained-syntax genetic programming system for discovering classification rules: application to medical data sets*, *Artificial Intelligence in Medicine*, v. 30, n. 1, p. 27-48, 2004.

Box, G. E. P.; Jenkins, G. M.; Reinsel, G. C. *Time Series Analysis, Forecasting and Control*. 3. ed. Englewood Cliffs: Prentice Hall, 1994.

Cabral, G. G. *Ferramenta para auxílio na previsão de séries temporais com intervalos de confiança usando Máquinas de Vetor de Suporte*. 2005. 77 f. Trabalho de Conclusão de Curso - Engenharia da Computação, Escola Politécnica de Pernambuco, Recife, 2005.

Campos, P. G.; Oliveira, E. M. J.; Ludermir, T. B. & Araujo, A. F. R. *MLP Networks for Classification and Prediction with Rule Extraction Mechanism*. In: *International Joint Conference on Neural Networks, 2004, Budapest. Proceedings of IJCNN 2004*. Los Alamitos : IEEE, 2004. p. 1387-1392.

Castro, C. L.; Braga, A. P.; Andrade, A. V. *Aplicação de um Modelo Ensemble de Redes Neurais Artificiais para Previsão de Séries Temporais não Estacionárias*. In: *Encontro Nacional de Inteligência Artificial, 5., 2005, São Leopoldo. Anais do Enia*. São Leopoldo: Sociedade Brasileira de Computação, 2005. p. 722-731.

Chakraborty K.; Mehrotra, K.; Mohan, C. K. & Ranka, S. *Forecasting the Behavior of Multivariate time Series Using Neural Networks*. 1992. *Neural Networks*, v. 5, n. 6, p. 961-970.

- Connor, N.; Madden, M. G. *A neural network approach to predicting stock exchange movements using external factors*, Knowl.-Based Syst., 19(5), 371-378, 2006. *apud* Martinez, L. C. et al. *From an Artificial Neural Network to a Stock Market Day-Trading System: A Case Study on the BM&F BOVESPA*. In: The International Joint Conference on Neural Networks, 2009, Atlanta - EUA, 2009.
- Cortez, P. A. R., *Algoritmos Genéticos e Redes Neurais na Previsão de Séries Temporais*. 1997. 105 f. Dissertação (Mestrado em Informática) – Departamento de Informática, Universidade do Minho, Braga, 1997.
- Cortez, P. A. R., *Modelos Inspirados na Natureza para a Previsão de Séries Temporais*. 2002. 188 f. Tese (Doutorado em Informática) – Departamento de Informática, Universidade do Minho, Braga, 2002.
- Darwin, C., *On the Origin of Species by Means for Natural Selection*. London: Jhon Murray, 1859.
- Ehlers, R. S. *Análise de Séries Temporais*. Desenvolvimento de material didático ou instrucional - Apostila. 4. ed. Universidade de São Paulo, São Paulo, 2007.
- Eiben, A.E.; Smith, J. E. *Introduction to Evolutionary Computing*. *Natural Computing Series*, MIT Press. Berlin: Springer, 2003.
- Elder, A. *Aprenda a Operar no Mercado de Ações*. Tradução de Afonso Celso da Cunha Serra. 7. ed. Rio de Janeiro: Elsevier, 2002.
- Elder, A. *Como Se Transformar Em Um Operador E Investidor De Sucesso*. Rio de Janeiro: Editora Campus, 2004.
- Faria, E. L.; Albuquerque M. P.; Alfolso, J. L. G.; Albuquerque, M. P. & Cavalcante, J. T. P. *Previsão de Séries Temporais utilizando Métodos Estatísticos*. CBPF e PUC RJ, Rio de Janeiro. 2008a.

Faria, E. L.; Albuquerque M. P.; Alfolso, J. L. G.; Albuquerque, M. P. & Cavalcante, J. T. P. *Previsão do Mercado de Ações Brasileiro utilizando Redes Neurais Artificiais*, CBPF e PUC RJ. Rio de Janeiro. 2008b.

Focaccia, A. M. C.; Rocha, D. F. C.; Heineck, R. H.; Cechin, S. G. & Añaña, E. S. *Mercado de Opções*. Revista do Ccei, Bagé, v. 6, n. 10, p. 56-64, ago. 2002.

Fyadoo. *Yahoo! Finance*. Califórnia. 2010. Disponível em: <<http://finance.yahoo.com/>>. Acesso em: 05 Julho 2010.

Giampietro, A. C. T.; Sarraceni, J. M.; Montanha, R. C. L.; Horita, R. Y. & Indalêncio, T. C. *As Bolsas De Valores Estão Cada Vez Mais Fazendo Parte Da Vida Dos Brasileiros?* In: Simpósio de Educação, 1, 2007, Lins-SP: Centro Universitário Católico Salesiano Auxilium: Disponível em: <<http://www.unisalesiano.edu.br/encontro2007/trabalho/aceitos/CC34169684842A.pdf>>. Acesso em 25 set. 2009.

Gitman, L. J. *Princípios De Administração Financeira Essencial*. 2. ed. Porto Alegre: Bookman, 2006.

Hansen, L. K.; Salamon, P. *Neural network ensembles*, IEEE Trans. Pattern Analysis and Machine Intelligence 12, no 10, 1990, p. 993-1001.

Haykin, S. *Neural Networks: A comprehensive Foundation*. 2. ed. Ontario: Prentice Hall, 1998.

Ibove99. *Índice Bovespa Definição e Metodologia*. São Paulo. 1999. Disponível em: <<http://www.bovespa.com.br/Pdf/Indices/IBovespa.pdf>>. Acesso em: 09 jun. 2010.

Koza, J. R.; Jones, L. W.; Keane, M. A.; Streeter, M. J. & Al-Sakran, S. H. *Towards industrial strength automated design of analog electrical circuits by means of genetic programming*, In: Genetic Programming Theory and Practice II, Kluwer, p. 13-15, 2004.

- Koza, J. R. *Genetic Programming III: Automatic Programming and Automatic Circuit Synthesis*, Morgan Kaufmann; 1st edition, 1999.
- Koza, J. R. *Hierarchical genetic algorithms operating on populations of computer programs*. Proceedings of the 11th International Joint Conference on Artificial Intelligent (IJCAI-89). Detroit, MI. p. 768-774. Morgan Kaufmann, 1989.
- Kuo, T.; Chen, S. *Genetic programming and international short-term capital flow*. Computing in Economics and Finance, v. 1, p. 11-13, 2003.
- Leone Filho, M. A. *Previsão de carga de curto prazo usando ensembles de previsores selecionados e evoluídos por Algoritmos Genéticos*. 2006. 143 f. Dissertação (Mestrado) - Universidade Estadual de Campinas, Campinas, 2006.
- Lima, V. S.; Ilha, H. F.; Scalzer, R. S. & Galdi, F. C. Análise Fundamentalista Sob A Perspectiva Do Analista De Mercado: Um Estudo De Caso Na Aes Tietê Comparando Os Modelos De Fluxo De Caixa Descontado e Aeg Ohlson (1995). In: 6º Congresso Usp Iniciação Científica Em Contabilidade, 6., 2009. São Paulo. *Anais do 6º Congresso USP de Iniciação Científica em Contabilidade*. São Paulo: Usp, 2009. p. 1-15.
- Martinez, L. C.; Hora, D. N.; Palotti, J. R. M.; Metra Junior, W. & Pappa, G. L. *From an Artificial Neural Network to a Stock Market Day-Trading System: A Case Study on the BM&F BOVESPA*. In: The International Joint Conference On Neural Networks, 2009, Atlanta - EUA, 2009.
- Mcapitais. *Mercado de Capitais Introdução*. São Paulo. 2008. Disponível em: <<http://www.bmfbovespa.com.br/Pdf/merccap.pdf>>. Acesso em: 16 Maio 2010.
- Mello, M. T. *Aplicação De Redes Neurais Artificiais No Processo De Precificação De Ações*. 2004. 49 f. Monografia (Bacharelado) - Universidade Federal de Pelotas, Pelotas, 2004.
- Misnky, M. L.; Papert, S. A. *Perceptrons: an introduction to computational geometry*. Massachussets: MIT Press, 1969.

- Morettin, P. A.; Toloi, C. M. *Modelos para Previsão de Séries Temporais*. Rio de Janeiro: IMPA, 1981.
- Penman, S. H. *Financial statement analysis and security valuation*. 2. ed. Boston: Irwin Mc Graw Hill, 2004.
- Piazza, M. C. *Bem Vindo à Bolsa de Valores*. 8. ed. rev. e ampl. São Paulo: Novo Conceito, 2009.
- Pontes, M. R. *Descoberta de sinais de compra e venda ótimos para análise técnica financeira do mercado de ações usando redes neurais artificiais*. 2008. 49 f. Monografia (Bacharelado) - Escola Politécnica de Pernambuco, Recife, 2008.
- Exame. *Revista Exame: O Brilho da Bolsa*. São Paulo: Abril, 28 Ago. 2009.
- Rezende, S. O. *Sistemas Inteligentes Fundamentos e Aplicações*. Barueri: Manole. 2005.
- Santos, M. V. A.; Takinami, A.; Goldman, A. & Fernandes, C. Tail - A Java Technical Analysis Library. *cse*, p.463-470, 2008. 11th IEEE International Conference on Computational Science and Engineering, 2008.
- Silva Filho, J. I.; Abe, J. M. *Fundamento das Redes Neurais Artificiais Paraconsistentes: Destacando Aplicações em Neurocomputação*. São Paulo: Arte e Ciência, 2001. 296 p.
- Souto-Maior, C. D. *Previsão Da Direção De Movimento De Índices De Ações Usando Um Sistema Fuzzy*. 2007. 197 f. Dissertação (Mestrado em Administração), Universidade Federal de Santa Catarina, Florianópolis, 2007.
- Souza, L. V. *Programação Genética e Combinação de Preditores para Previsão de Séries Temporais*. 2006. 154 f. Tese (Doutorado) - Universidade Federal do Paraná, Curitiba, 2006.

Teredesai, A.; Govindaraju, V. *Issues in evolving GP based classifiers for a pattern recognition task*, In: Proceedings of the 2004 IEEE Congress on Evolutionary Computation, p. 509-515, 2004.

Tukey, J. W. *Exploratory Data Analysis*. Reading, Mass.: Addison-Wesley Pub., 1977.

Yen, M.; Chou, T.; Li, H. & Ho, Y. *Using Neural Network and Genetic Programming Techniques to Forecast Inter-Commodity Spreads*. Department of Finance, Chaoyang University of Technology, Taiwan. 2007.