

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

FÁBIO DAROS DE FREITAS

**MODELO DE SELEÇÃO DE CARTEIRAS BASEADO EM
ERROS DE PREDIÇÃO**

**VITÓRIA, E.S.
2008**

FÁBIO DAROS DE FREITAS

**MODELO DE SELEÇÃO DE CARTEIRAS BASEADO EM
ERROS DE PREDIÇÃO**

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Doutor em Engenharia Elétrica, na área de concentração em Automação.

Orientador: Prof. Dr. Alberto Ferreira De Souza

Vitória, E.S.
2008

FÁBIO DAROS DE FREITAS

**MODELO DE SELEÇÃO DE CARTEIRAS BASEADO EM
ERROS DE PREDIÇÃO**

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Doutor em Engenharia Elétrica, na área de concentração em Automação.

Aprovada em 18 de Dezembro de 2008.

COMISSÃO EXAMINADORA

Prof. Dr. Alberto Ferreira De Souza
Universidade Federal do Espírito Santo — UFES
Orientador

Prof. Dr. Ailson Rosetti de Almeida
Universidade Federal do Espírito Santo — UFES

Prof. Dr. Evandro Ottoni Teatini Salles
Universidade Federal do Espírito Santo — UFES

Profa. Dra. Eliana Zandonade
Universidade Federal do Espírito Santo — UFES

Prof. Dr. Felipe Maia Galvão França
Universidade Federal do Rio de Janeiro — UFRJ

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

Freitas, Fábio Daros de, 1968–

F866m Modelo de seleção de carteiras baseado em erros de predição /
Fábio Daros de Freitas. — 2008.
xix, 143 f. : il.

Orientador: Alberto Ferreira de Souza.
Tese (doutorado) — Universidade Federal do Espírito Santo, Centro
Tecnológico.

1. Redes neurais (Computação). 2. Teoria da previsão. 3. Análise de
séries temporais. 4. Mercado financeiro. 5. Carteiras (Finanças) — Admi-
nistração. 6. Investimentos. I. Souza, Alberto Ferreira de. II. Universidade
Federal do Espírito Santo. Centro Tecnológico. III. Título.

CDU: 621.3

*We each pay a fabulous price
For our visions of paradise
But a spirit with a vision is a dream
With a mission*

Neil Peart

Agradecimentos

Concluir uma longa jornada sem contar com a ajuda de outros é algo bastante improvável. Ao longo da minha jornada na elaboração desta Tese de Doutorado acumulei débitos de gratidão com aqueles que me apoiaram e colaboraram para a conclusão deste trabalho. Rendo aqui meus sinceros agradecimentos:

A Deus, a fonte de tudo.

A Elianete, meu amor e minha companheira, e aos nossos filhos, Jéssica, Gabriel e Rafael, por me acompanharem nessa jornada, tolerando minhas ausências, compartilhando as dores e delícias do caminho, e alegrando a vida — devo isso a vocês.

Aos meus Pais, Paulo e Nilma, pelo alento e exemplo; aos meus irmãos, cunhadas e cunhado, Paulo e Zelma, Christian e Ana Kátia, e Lúcia e Alexandre, pelo amor, amizade e companheirismo de sempre; e, claro, aos meus sobrinhos Paulo Neto e Maria Clara, por sempre semearem alegria.

Ao Prof. Alberto Ferreira De Souza, orientador deste trabalho e amigo de muitas outras jornadas, pela sua dedicação, entusiasmo e otimismo sempre inabaláveis.

Ao Prof. Ailson Rosetti de Almeida, orientador na primeira fase deste trabalho, pelo apoio e confiança sempre depositados.

Ao Prof. Francisco José Negreiros Gomes (*in memorian*), orientador da minha dissertação de Mestrado e saudoso amigo, pelo início de tudo, ao me apresentar o frutífero problema da Seleção de Carteiras.

Ao Prof. Elias de Oliveira, pelas proveitosas discussões e preciosos aconselhamentos durante a fase final deste trabalho.

A Laura Gadelha Xavier, delegada da Delegacia da Receita Federal do Brasil em Vitória E.S. (DRF/VIT/ES), pelo apoio durante este trabalho.

A Mário Pedrosa Bassul, chefe do Serviço de Tecnologia e Segurança da Informação da DRF/VIT/ES, pelo apoio incondicional, companheirismo e sábios conselhos.

Aos colegas do Centro de Microinformática da DRF/VIT/ES, Antenor César Dalvi, Christian Daros de Freitas, Ricardo Caetano, e Sérgio Nogueira, companheiros do maravilhoso mundo da programação de computadores, por toda ajuda e apoio que me deram durante este trabalho.

A Wanda Testoni Gomes, chefe do Serviço de Programação e Logística da DRF/VIT/ES, e aos colegas Rosária, Marcelo, Chiquinho, Calebe e Carmem, pelo companheirismo e apoio durante o desenvolvimento deste trabalho.

A José Augusto Teixeira, chefe da Divisão de Tecnologia e Segurança da Informação da Superintendência Regional da Receita Federal do Brasil na 7ª Região Fiscal — Rio de Janeiro e Espírito Santo — (DITEC/SRRF/7ªRF), e a Álvaro Milhomem, Luiz Roberto Lobianco e demais colegas da DITEC/SRRF/7ªRF, pelo incondicional apoio durante este trabalho.

A Regina Henriques, Maria Célia, Denise e Hélio, colegas da Receita Federal do Brasil e da Subcomissão Técnica para a CNAE-Subclasses, pelo incentivo e apoio durante o desenvolvimento deste trabalho.

A Júlio, Joeldo, e demais colegas da então Equipe de Desenvolvimento dos Programas do IRPF da Receita Federal do Brasil, pelas sempre provocantes e enriquecedoras discussões.

Aos colegas do Laboratório de Computação de Alto Desempenho (LCAD) do DI/UFES pelo apoio e colaboração na manutenção do nosso “propulsor” — o cluster Enterprise.

Aos amigos da vida, Abner, André, Dilcemar, Henilton, e Humberto, e aos amigos do *almoço tecnológico*, pelo incentivo e inúmeras discussões sobre este trabalho.

Por último, mas não por menos, a todos que colaboraram de alguma forma para conclusão deste trabalho, especialmente a Eric Raymond e Linus Torvalds.

Conteúdo

1	Introdução	20
1.1	Motivações	22
1.2	Objetivos	24
1.3	Contribuições	25
1.4	Organização desta Tese	26
2	Trabalhos Correlatos	28
2.1	Predição de Ações e <i>Trading Systems</i>	30
2.2	Predição de Ações na Formação de Carteiras	32
3	Noções Fundamentais	35
3.1	Mercado de Ações	35
3.2	Retorno de Investimentos	38
3.3	Predição de Retornos de Ações com Redes Neurais	41
3.3.1	Redes Neurais Artificiais	41
3.3.2	Rede Neural Auto-regressiva	52
3.4	O Problema da Seleção de Carteiras	56
3.4.1	O Modelo Média-Variância para Seleção de Carteiras	56
3.4.2	A Fronteira Eficiente	59
3.4.3	Diversificação Eficiente e Número de Ações	61

4	Seleção de Carteiras com Predição de Retornos	62
4.1	Rede Neural Auto-regressiva com Referência Móvel	62
4.2	Retorno Predito e Risco das Ações	65
4.3	Retorno Predito e Risco da Carteira	66
4.4	Modelo de Otimização	68
4.5	O Modelo Média-Variância como um Caso Particular do Modelo de Seleção de Carteiras Baseado em Erros de Predição	70
5	Métodos	72
5.1	Estratégia de Investimento	73
5.2	Dados	74
5.3	Arcabouço Computacional	79
5.3.1	Ambiente de Seleção de Carteiras e Simulação de Investimento . .	79
5.3.2	Preditores Neurais Auto-regressivos com Referência Móvel	81
5.4	Métricas	85
5.4.1	Medidas de Desempenho dos Preditores	85
5.4.2	Medidas de Desempenho dos Modelos de Seleção de Carteiras . . .	87
5.4.3	Sentido da Melhoria das Métricas Utilizadas	88
6	Experimentos	90
6.1	Investimento com Dados Artificiais	90
6.2	Investimento com Dados Reais	94
6.2.1	Predição de Retornos	94
6.2.2	Fronteiras Eficientes	98
6.2.3	Simulação de Investimento de Baixo Risco	101
6.2.4	Simulação de Investimento de Risco Moderado	108
6.2.5	Simulação de Investimento de Alto Risco	115
6.3	Sumário dos Experimentos	121

7	Discussão	123
7.1	Previsibilidade dos Mercados	123
7.2	Fricções do Mercado	124
7.2.1	Custos Operacionais	125
7.2.2	Indivisibilidade dos Lotes-padrão das Ações	126
7.2.3	Modelo Estendido	126
7.3	Análise Crítica deste Trabalho de Pesquisa	127
8	Conclusão	130
8.1	Sumário	130
8.2	Conclusões	131
8.3	Trabalhos Futuros	131
	Referências Bibliográficas	133

Lista de Tabelas

5.1	Relação das 52 ações utilizadas nos experimentos.	75
5.2	Sumário das 52 séries de 413 retornos semanais.	77
5.3	Sumário das 52 séries de 142 retornos semanais.	78
5.4	Sentido da melhoria das métricas de avaliação de desempenho.	88
6.1	Simulação de investimento com dados artificiais — desempenho semanal. .	93
6.2	Sumário do desempenho dos 52 preditores dos retornos semanais das ações.	95
6.3	Sumário dos índices de Normalidade dos erros de predição e dos retornos.	98
6.4	Simulação de investimento de baixo risco — desempenho semanal.	106
6.5	Simulação de investimento de baixo risco — intervalos de confiança dos retornos semanais médios no trimestre.	107
6.6	Simulação de investimento de risco moderado — desempenho semanal. . .	113
6.7	Simulação de investimento de risco moderado — intervalos de confiança dos retornos semanais médios no trimestre	114
6.8	Simulação de investimento de alto risco — desempenho semanal.	119
6.9	Simulação de investimento de alto risco — intervalos de confiança dos re- tornos semanais médios no trimestre.	120

Lista de Figuras

3.1	Exemplos de séries históricas de preços de ações do mercado brasileiro. . .	39
3.2	Exemplos de séries históricas de retornos de ações do mercado brasileiro. .	40
3.3	Modelo do neurônio artificial.	42
3.4	Funções de ativação.	45
3.5	Rede neural <i>feedforward</i> totalmente conectada com uma camada.	46
3.6	Rede neural <i>feedforward</i> totalmente conectada com duas camadas.	46
3.7	Rede Neural Auto-regressiva com uma camada.	53
3.8	Rede Neural Auto-regressiva com duas camadas.	55
3.9	Exemplo de fronteira eficiente.	60
4.1	Aplicação do método auto-regressivo com referência móvel no preditor RNAR-RM (4,1).	64
4.2	Exemplo de preditor RNAR-RM (4,1).	65
4.3	Modelo distribucional dos retornos das ações para o modelo de seleção de carteiras baseado em erros de predição e para o modelo média-variância. .	69
5.1	Janela temporal deslizando para estimação dos parâmetros dos modelos. . .	79
5.2	Preditor RNAR-RM (4,1) empregado nos experimentos.	82
5.3	Janela temporal deslizando para treinamento e predição de retornos.	85
6.1	Simulação de investimento com dados artificiais.	91

6.2	Índices de Normalidade dos erros de predição e dos retornos.	97
6.3	Fronteiras eficientes e intervalos de confiança.	99
6.4	Fronteiras eficientes e coeficientes de variação.	100
6.5	Simulação de investimento de baixo risco — retornos acumulados semanais.	102
6.6	Simulação de investimento de baixo risco — retornos acumulados semanais e intervalos de confiança.	104
6.7	Simulação de investimento de baixo risco — intervalos de confiança dos retornos semanais médios no trimestre.	107
6.8	Exemplo de fronteira eficiente com ativo sem risco.	109
6.9	Simulação de investimento de risco moderado — retornos acumulados se- manais.	110
6.10	Simulação de investimento de risco moderado — retornos acumulados se- manais e intervalos de confiança.	111
6.11	Simulação de investimento de risco moderado — intervalos de confiança dos retornos semanais médios no trimestre.	114
6.12	Simulação de investimento de alto risco — retornos acumulados semanais.	116
6.13	Simulação de investimento de alto risco — retornos acumulados semanais e intervalos de confiança.	117
6.14	Simulação de investimento de alto risco — intervalos de confiança dos re- tornos semanais médios no trimestre.	120

Lista de Símbolos e Abreviações

$C()$	Função de custo financeiro do investimento.
I	Conjunto de informações acerca dos custos operacionais.
M	Número de ações da carteira.
\mathbb{N}	Conjunto dos números naturais.
N	Número de observações da série de retornos da ação.
n	Número de observações da série de erros de predição da ação.
P	Preço da ação.
p	Ordem da auto-regressão.
\mathbf{P}'	Série histórica de preços da ação.
Q	Número de lotes-padrão da ação.
r	Retorno da ação.
\mathbf{r}'	Série histórica de retornos da ação.
\bar{r}	Retorno esperado da ação no modelo média-variância.
v, σ^2	Risco da ação no modelo média-variância.
r_p	Retorno da carteira.
\bar{r}_p	Retorno esperado da carteira no modelo média-variância.
V, σ_p^2	Risco da carteira no modelo média-variância.
γ	Risco interativo entre duas ações no modelo média-variância.
\hat{r}	Retorno predito da ação.

\hat{r}^L	Retorno predito líquido da ação.
ε	Erro de predição da ação.
ε'	Série histórica dos erros de predição da ação.
$\hat{v}, \sigma_\varepsilon^2$	Risco da ação no modelo de seleção de carteiras baseado em erros de predição.
\hat{r}_p	Retorno predito da carteira no modelo de seleção de carteiras baseado em erros de predição.
$\hat{V}, \hat{\sigma}_p^2$	Risco da carteira no modelo de seleção de carteiras baseado em erros de predição.
$\hat{\gamma}$	Risco interativo entre duas ações no modelo de seleção de carteiras baseado em erros de predição.
Γ	Matriz de covariâncias.
R_d	Retorno de carteira desejado.
R_{rf}	Retorno do ativo sem risco.
$\mathcal{S}_T()$	Sistema de predição de séries históricas.
W	Capital.
X	Participação da ação na carteira.
z	Referência móvel.
$AR(p)$	Modelo auto-regressivo de ordem p .
ARE	Retorno acumulado.
H_0	Hipótese nula.
H_R	Taxa de acerto dos retornos.
H_{R-}	Taxa de acerto dos retornos negativos.
H_{R+}	Taxa de acerto dos retornos positivos.
CI	Intervalo de confiança.
IN	Índice de Normalidade.

$MAPE$	Erro percentual absoluto médio.
ME	Erro médio.
PCM	Taxa de mudança da carteira.
$RMSE$	Raiz do erro quadrático médio.
$RNAR(p)$	Rede Neural Auto-regressiva de ordem p .
$RNAR-RM(p,k)$	Rede Neural Auto-regressiva com Referência Móvel de ordem p, k .
TI	Índice de troca da carteira.
V	Coefficiente de variação.

Resumo

Este trabalho apresenta um novo modelo de seleção de carteiras baseado em erros de predição que captura oportunidades de investimento no curto prazo. Nós utilizamos preditores neurais auto-regressivos com referência móvel para prever os retornos futuros das ações, e uma medida de risco baseada nos seus erros de predição foi derivada de forma a manter a mesma fundamentação estatística do modelo média-variância. Os efeitos da diversificação eficiente se aplicam através da seleção de preditores com perfis de erros de predição baixos e complementares. Um grande conjunto de experimentos com dados reais do mercado de ações brasileiro foi conduzido para avaliar o modelo de seleção de carteiras baseado em erros de predição, o qual contou com o exame da Normalidade dos erros de predição. Nossos resultados principais mostraram que é possível obter erros de predição Normais a partir de séries de retornos não Normais e que o modelo de seleção de carteiras baseado em erros de predição capturou corretamente oportunidades de curto prazo, desempenhando melhor que o modelo média-variância e superando o índice de mercado.

Abstract

This work presents a new prediction errors-based portfolio optimization model that captures short-term investment opportunities. We used autoregressive moving references neural network predictors to predict the stock's returns and derived a risk measure based on the predictor's errors of prediction that maintains the same statistical foundation of the mean-variance model. The efficient diversification effects hold by selecting predictors with low and complimentary error profiles. A large set of experiments with real data from the Brazilian stock market was employed to evaluate our portfolio optimization model, which included the examination of the Normality of the errors of prediction. Our main results showed that it is possible to obtain Normal prediction errors with non-Normal series of stock returns, and that the prediction errors-based portfolio optimization model better captured the short term opportunities, outperforming the mean-variance model and beating the market index.

Publicações Derivadas desta Tese

- 1 FREITAS, F. D. et al. Portfolio selection with predicted returns using neural networks. In: *IASTED International Conference on Artificial Intelligence and Applications*. Marbella, Spain: [s.n.], 2001. p. 99–103.
- 2 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. Autoregressive neural network predictors in the brazilian stock markets. In: *VII Simpósio Brasileiro de Automação Inteligente (SBAI)/II IEEE Latin American Robotics Symposium (IEEE-LARS)*. São Luis, Brasil: [s.n.], 2005. p. 1–8. ISBN 85-85048-55-7.
- 3 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. Avaliação de preditores neurais auto-regressivos no mercado de ações. In: *VII Congresso Brasileiro de Redes Neurais (CBRN)*. Natal, Brasil: [s.n.], 2005. p. 1–6. ISSN 1808-8589.
- 4 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. A prediction-based portfolio optimization model. In: *5th International Symposium On Robotics and Automation — ISRA 2006*. Hidalgo, Mexico: [s.n.], 2006. p. 520–525. ISBN 970-769-070-4.
- 5 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. Um modelo de otimização de carteiras baseado em predição. In: *XXXVIII Simpósio Brasileiro de Pesquisa Operacional (SBPO)*. Goiânia-GO, Brasil: [s.n.], 2006. p. 1–8.
- 6 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. Prediction-based portfolio optimization model using neural networks. *Neurocomputing*, 2009. (in-press).

Capítulo 1

Introdução

A seleção de investimentos é um problema fundamental da área de finanças e de grande relevância tanto na academia quanto na indústria, consistindo basicamente na escolha de oportunidades de investimento que otimizem as preferências do investidor. Para tal, um projeto de investimento é definido a partir da escolha de: (i) um *horizonte de investimento*, que é o período de tempo no qual o investimento será realizado, isto é, o período que se inicia no aporte inicial de recursos e termina na aferição do resultado final; (ii) um conjunto de oportunidades de investimento, constituído pelos *ativos alvos* nos quais o investimento será realizado ao longo do seu respectivo horizonte; e (iii) uma função de preferência, ou *função utilidade* do investidor, que represente o nível de satisfação relativa do investidor com o resultado do investimento [1, 2]. O resultado do investimento tipicamente é medido como a variação da riqueza alcançada no final do horizonte de investimento, denominado *retorno do investimento*, ou simplesmente *retorno*.

O planejamento do investimento consiste na formulação e solução de um problema de otimização que maximize a expectativa de retorno do investimento no final do horizonte, ou *retorno esperado*, considerando os retornos esperados de cada oportunidade de investimento. Contudo, este retorno esperado pode não se realizar. Neste caso, o retorno do investimento observado no final do seu horizonte, ou *retorno realizado*, pode ser inferior ou superior ao retorno esperado, produzindo perdas ou excesso de ganhos, respectivamente. Portanto, cada oportunidade de investimento apresenta um grau de incerteza associado à realização do seu retorno esperado. Quando essa incerteza associada ao retorno esperado pode ser mensurada por meio do cálculo da probabilidade de ocorrência de cada um dos possíveis resultados, ela é denominada *risco*, e também deve ser considerada no problema de otimização formulado durante o planejamento do investimento [1, 2, 3, 4].

O planejamento do investimento e o problema de otimização associado podem ser tratados pelo ramo da *Análise de Investimentos* denominado *Teoria de Carteiras*¹. Neste escopo, os *títulos financeiros* são um conjunto particular de ativos alvos de especial interesse. Estes ativos são contratos legais que representam o direito ao recebimento de vantagens futuras sob um conjunto de condições determinadas, e são negociados em algum tipo de mercado organizado. Certificados de Depósito Bancário, Letras do Tesouro Nacional, caderneta de poupança, ações, debêntures e bônus são alguns exemplos de títulos financeiros, e podem ser classificados como *ativos financeiros de risco*, quando o risco é significativamente diferente de zero, ou *ativos financeiros sem risco*, quando o risco é igual ou muito próximo a zero. Ações e debêntures são um exemplo de ativos financeiros de risco, enquanto os Certificados de Depósito Bancário, Letras do Tesouro Nacional e caderneta de poupança são exemplos de ativos financeiros sem risco².

A função utilidade do investidor deve representar, em termos de preferência, a sua atitude em relação ao retorno e risco do investimento. Em 1952, Harry Max Markowitz pioneiramente formulou as propriedades de uma função utilidade que descreve a atitude do investidor racional³ diante do risco [6], e estabeleceu a sua relação com o retorno e risco de carteiras através do seu celebrado modelo média-variância [7, 8]. Muito embora os benefícios da estratégia da diversificação de investimentos em carteiras já fossem amplamente aceitos e estabelecidos à época, pela primeira vez um arcabouço matemático-estatístico era apresentado propondo uma modelagem quantitativa para o tratamento simultâneo do problema de diversificação e da maximização do valor esperado da função utilidade do investidor, ou *utilidade esperada* [9]. Desta forma, o modelo de Markowitz tornou-se um marco na *Moderna Teoria de Carteiras*, e a sua fundamentação estatística e formulação matemática, aliadas à utilidade quadrática, produziram o conceito fundamental da *fronteira eficiente* [7]. Neste modelo, o risco total do investimento em uma carteira composta por várias ações é minimizado através da seleção ótima de ações com baixo risco conjunto, o que produz um mecanismo de compensação de perdas conhecido como *diversificação eficiente*.

O processo de otimização de carteiras consiste em encontrar, num grande universo de ações, as participações (isto é, as proporções individuais) de cada ação que minimizam o risco da carteira para um determinado retorno de carteira desejado, ou, no problema dual, que maximizam o retorno da carteira para um determinado risco desejado. Variando-se o retorno desejado para a carteira (ou o risco no problema dual) ao longo de todos os valores

¹Uma carteira é uma combinação linear de ativos cuja soma dos pesos X_i de cada ativo i , ou participações, é unitária, ou seja, $\sum X_i = 1$.

²Quando não caracterizado de outro modo, utilizaremos o termo *ativo* como sinônimo para os ativos financeiros de risco ou, mais especificamente, para as ações negociadas em bolsa de valores.

³O *investidor racional* é um modelo de comportamento diante da incerteza que: (a) associa um número, a sua *utilidade*, a cada resultado possível do seu investimento, e (b) quando exposto à escolha entre alternativas incertas, escolhe aquela com o maior valor esperado da sua *utilidade* [5, pp. 208]. Notadamente, este comportamento adere à preferência de mais riqueza a menos riqueza, e a aversão ao risco.

possíveis no universo de ações, encontramos um subconjunto das carteiras possíveis (conjunto de oportunidades) que contém apenas carteiras ótimas. Estas carteiras delineiam um *locus* característico no espaço bidimensional risco-retorno, que é a *fronteira eficiente*. As carteiras da fronteira eficiente são denominadas *carteiras eficientes* — não existem outras carteiras no conjunto de oportunidades que apresentem menor risco para o mesmo nível de retorno, ou maior retorno para o mesmo nível de risco [5].

O modelo média-variância supõe que as séries de retornos⁴ de cada ação seguem uma distribuição Normal, e utiliza a média destas séries como o retorno esperado da ação, a variância da série como medida do risco individual da ação e a covariância entre os retornos das ações como medida do risco conjunto de cada par de ações. Após o modelo média-variância, surgiram vários modelos de seleção de carteiras que utilizam suas premissas básicas (de que as séries de retornos são Normais e que os seus momentos em relação à média podem ser utilizados como medida de retorno esperado e risco) [3, pp. 219–252]. Em todos estes modelos, conhecidos atualmente como *modelos clássicos*, o retorno esperado da carteira é dado pela combinação linear das participações de cada ação na carteira e seus retornos esperados (seus retornos médios). O risco da carteira, por sua vez, tem definições variadas, mas está sempre relacionado aos momentos em relação à média da combinação linear das participações e das séries de retornos das ações [3].

1.1 Motivações

Muito embora os modelos clássicos de seleção de carteiras gozem de ampla aceitação e prestígio na indústria, as suas premissas básicas de que (i) as séries de retornos são Normais, (ii) a variância destas séries pode ser utilizada como medida de risco e (iii) a média das séries pode ser utilizada como predição para o retorno futuro, têm sido desafiadas de diversas formas.

As séries históricas dos retornos freqüentemente se distanciam da Normalidade, apresentando distribuições com significativos graus de assimetria e curtose [10, 11], o que torna a variância (ou o desvio padrão) das séries de retornos uma medida inadequada do risco das ações [4, pp. 156]. Além disso, a previsibilidade dos mercados de ações ainda é uma questão aberta na teoria das finanças; a *Hipótese do Mercado Eficiente*, arcabouço teórico que tem guiado esta discussão, ainda se encontra sob revisão e testes empíricos [12, 13, 14, 15]. Isso porque a eficiência dos mercados implica em um modelo de *passeio aleatório*⁵ para o movi-

⁴A diferença entre o preço da ação no tempo t e o preço no tempo $t - 1$, dividida pelo preço no tempo $t - 1$ (ver Eq. 3.2, pág. 38).

⁵O modelo de *passeio aleatório* implica em incrementos temporais aleatórios, independentes e não correlacionados para as séries de preços das ações.

mento das séries de preços das ações, onde quaisquer novas informações são prontamente assimiladas pelo mercado, que imediatamente promove a precificação das ações e não deixa margem para a exploração de vantagens preditivas detectáveis nas séries de preços das ações. Desta forma, num mercado eficiente, a melhor predição para o preço futuro de uma ação seria o seu último preço.

Porém, a presença de padrões previsíveis, tais como correlações seriais⁶ e padrões que surgem dos efeitos de calendário, de pessimismo da mídia e até mesmo de resultados de competições esportivas, tem sido verificada em testes com dados reais do mercado [15, 16, 17, 18, 19]. No caso particular do investimento de curto prazo, a utilização dos retornos médios como predição para os retornos futuros, como nos modelos clássicos, exerce um efeito de filtragem passa-baixas no comportamento dinâmico dos mercados, produzindo estimativas imprecisas dos retornos futuros de curto prazo, o que eleva o risco das estratégias de investimento baseadas nestas estimativas.

As deficiências dos modelos clássicos, em particular quanto à incapacidade de tirarem proveito de oportunidades de curto prazo, nos motivaram a investigar novos modelos de seleção de carteiras capazes de explorar oportunidades preditivas detectáveis nas séries de retornos de ações.

A predição de séries históricas, ou *séries temporais*, tem sido tradicionalmente tratada pelos modelos lineares da *Análise de Séries Temporais* [20, 21, 22]. Entretanto, as últimas duas décadas testemunharam o surgimento de vários métodos de aprendizagem de máquina para a predição de séries temporais, tais como as redes neurais artificiais, os algoritmos genéticos, as máquinas de vetor de suporte (*support vector machines*) e os conjuntos nebulosos (*fuzzy sets*), entre outros [23, 24, 25, 26, 27, 28, 29, 30, 31].

Neste trabalho de pesquisa, nós investigamos o uso de redes neurais na predição do retorno futuro de ações e o uso dos retornos preditos e dos erros de predição na formulação de modelos de seleção de carteiras baseados em predição de retornos. A capacidade de produzir mapeamentos não-lineares e a natureza não assistida da estimação dos parâmetros das redes neurais são características vantajosas para a sua aplicação na predição dos retornos futuros de ações [32, 33, 34, 35, 36, 37, 38, 39], tornando-as adequadas para aplicações em larga escala. Essas características nos motivaram a utilizar preditores baseados em redes neurais neste trabalho. É importante mencionar que outros métodos de aprendizado de máquina também podem ser apropriados para a predição do retorno futuro de ações.

⁶Correlação de uma variável com ela própria em sucessivos intervalos amostrais, também denominada *autocorrelação*.

1.2 Objetivos

Em trabalhos anteriores [33, 34], nós investigamos a Normalidade dos erros de predição de preditores neurais auto-regressivos utilizados na predição dos retornos semanais de ações do mercado brasileiro. Nossos resultados mostraram maior evidência de Normalidade nas séries de erros de predição do que nas séries de retornos. Assim, em trabalhos subsequentes, nós exploramos a Normalidade destes erros de predição na construção de um modelo de seleção de carteiras que utiliza a predição de retornos num arcabouço Normal multivariado [35, 36, 40].

A predição de séries de retornos de ações é reconhecida como uma tarefa difícil [41, 42]. Preditores individuais (de uma ação) apresentam desempenho modesto, tipicamente exibindo erros de predição significativos [33, 34, 42]. Contudo, preditores podem ser combinados de forma a explorar as complementaridades no comportamento dos seus erros de predição, possivelmente levando a bons resultados para as combinações lineares destas predições. Uma combinação linear deste tipo, com soma dos pesos igual a um, pode representar o retorno predito de uma carteira num modelo de seleção de carteiras baseado em predição. Este contexto determinou o principal objetivo deste trabalho de pesquisa, que é investigar a seguinte hipótese:

Um modelo de seleção de carteiras com função de risco baseada nos erros de predição de preditores neurais auto-regressivos pode ser formulado para aderir ao arcabouço Normal multivariado e obter retorno superior ao modelo média-variância para os mesmos níveis de risco.

Com este objetivo, formulamos o *modelo de seleção de carteiras baseado em erros de predição* — um modelo de seleção de carteiras de investimento que explora predições de retornos de ações por meio de um arcabouço de diversificação eficiente Normal multivariado — e comparamos experimentalmente o desempenho deste modelo com o do modelo média-variância de Harry Markowitz. Nossos resultados mostraram que o modelo de seleção de carteiras baseado em erros de predição obtém retornos significativamente superiores aos do modelo média-variância para níveis equivalentes de risco. Além disso, para as carteiras ótimas de maior retorno esperado, os resultados do nosso modelo superaram em muito os resultados do modelo média-variância e do índice IBOVESPA — em uma simulação de investimento com as ações que integraram o índice IBOVESPA no período de 142 semanas entre 5-Jan-2005 e 19-Set-2007, o modelo de seleção de carteiras baseado em erros de predição alcançou um retorno acumulado de 389%, enquanto que o modelo média-variância alcançou 99% e o índice IBOVESPA atingiu 219% no período.

A formulação do modelo de seleção de carteiras baseado em erros de predição usa as variâncias e covariâncias das séries de erros de predição dos retornos futuros das ações para obter a variância da combinação linear das participações e das séries de erros de predição das ações como medida de risco da carteira. Assim, a Normalidade dos erros de predição é fundamental para a viabilidade do modelo. Por essa razão, avaliamos a Normalidade dos erros de predição dos preditores empregados no modelo de seleção de carteiras baseado em erros de predição, e também dos retornos das ações utilizadas no modelo média-variância. O teste de aderência do *Qui-quadrado* no nível de significância de 1% foi utilizado para calcular o percentual das séries que tiveram a Normalidade aceita em cada uma das 142 semanas, e a média alcançada pelas séries dos erros de predição foi superior à média das séries de retornos, segundo o *teste-t* no nível de significância de 5%. Este resultado indica que as séries dos erros de predição foram mais Normais que as séries dos retornos das ações.

1.3 Contribuições

A principal contribuição desta tese foi a criação de um modelo de seleção de carteiras que explora um conjunto de predições de retornos futuros de ações por meio de um arcabouço de diversificação eficiente Normal multivariado. Este modelo viabiliza estratégias de diversificação eficiente para o problema de investimento no curto prazo. A comparação do desempenho do nosso modelo com o desempenho do modelo média-variância e do índice IBOVESPA integra esta contribuição.

As principais vantagens desta abordagem em relação aos métodos de referência são:

- A utilização de um estimador do retorno futuro das ações baseado na predição de séries temporais permite capturar a dinâmica dos movimentos do mercado, contornando assim as características de filtragem observadas quando da utilização da média das séries de retornos. Com isso, torna-se possível derivar estratégias mais adequadas ao investimento no curto prazo.
- A utilização de preditores baseados em redes neurais artificiais não requer a identificação e estimação de parâmetros como nos métodos da Análise de Séries Temporais, permitindo a exploração de um grande número de oportunidades de investimento de forma automatizada e não assistida.
- O modelo privilegia uma métrica simples e intuitiva para avaliar os preditores — o desvio padrão (ou a variância) dos erros de predição.

Algumas contribuições adicionais deste trabalho são:

- O desenvolvimento de medidas de risco e de um modelo de otimização de carteiras baseados na predição de retornos que dependem apenas dos retornos preditos e das variâncias e covariâncias dos erros de predição. Graças a isso, o arcabouço de otimização de carteiras desenvolvido independe do método de predição utilizado, podendo ser aplicado virtualmente a qualquer método de predição que forneça os seus parâmetros de modelo (retornos preditos, variâncias e covariâncias dos erros de predição).
- A proposta de um novo preditor de séries temporais baseado na *rede neural auto-regressiva com referência móvel*. Este método busca emular o comportamento intuitivo que tipicamente empregamos para analisar visualmente o gráfico de uma série histórica com a finalidade de extrapolar seu próximo valor, qual seja, concentrar nossa atenção visual nos últimos pontos do gráfico, criando uma janela virtual que delimita uma região capaz de oferecer uma imagem com quantidade de informação visual suficiente para a extrapolação do próximo ponto do gráfico.
- O arcabouço proposto pode ser utilizado em aplicações do tipo *combinação de preditores (pooling of forecasts)* [43], onde várias predições de uma mesma série histórica, tipicamente obtidas através de métodos distintos, são combinadas com o objetivo de se obter uma predição mais precisa para o valor futuro da série. Desta forma, o modelo de seleção de carteiras baseado em erros de predição pode minimizar a variância dos erros de predição neste tipo de aplicação. De fato, o espectro de aplicações deste arcabouço não se restringe apenas aos problemas de economia e finanças, encontrando campo de utilização em diversos problemas de engenharia, especialmente da engenharia elétrica, nas aplicações em controle e comunicações, por exemplo.

1.4 Organização desta Tese

Após esta introdução, o Capítulo 2 apresenta uma revisão da literatura nas áreas correlatas a este trabalho, destacando desenvolvimentos recentes na predição de séries financeiras e na incorporação de métodos preditivos na formação de carteiras. Em seguida, o Capítulo 3 apresenta as noções fundamentais relacionadas aos temas abordados nos demais capítulos. Primeiro, é apresentada uma breve introdução ao mercado de ações e ao problema do investimento, e, na sequência, é apresentada a abordagem padrão para a construção de preditores de séries temporais baseados em redes neurais. O capítulo é encerrado com a apresentação do modelo média-variância para a seleção de carteiras.

O Capítulo 4 apresenta o preditor baseado na rede neural auto-regressiva com referência móvel e formula o modelo de seleção de carteiras baseado em erros de predição. Este modelo, bem como suas hipóteses subjacentes, medidas de retorno e risco, e modelo de otimização, são apresentados nesse capítulo. Uma comparação qualitativa do modelo de seleção de carteiras baseado em erros de predição frente ao modelo média-variância e uma análise do modelo média-variância como um caso particular do modelo de seleção de carteiras baseado em erros de predição encerram o capítulo.

O Capítulo 5 apresenta os métodos utilizados nas avaliações experimentais realizadas neste trabalho, destacando a estratégia de investimento implementada, os dados utilizados e as medidas de desempenho empregadas.

O Capítulo 6 apresenta os experimentos empregados na avaliação de desempenho do modelo de seleção de carteiras baseado em erros de predição frente ao modelo média-variância e ao referencial do mercado — o índice IBOVESPA. No Capítulo 7, discutimos este trabalho à luz da previsibilidade dos mercados e melhorias a serem incorporadas no nosso modelo para acomodar fricções do mercado e apresentamos uma análise crítica deste trabalho de pesquisa. Por fim, no Capítulo 8, apresentamos nossas conclusões e perspectivas de trabalhos futuros.

Capítulo 2

Trabalhos Correlatos

A *Moderna Teoria de Carteiras* oferece um conjunto de métodos quantitativos através dos quais o investidor racional pode otimizar a diversificação dos seus investimentos [3]. Sua origem se confunde com a proposição do modelo média-variância por Harry Markowitz em 1952 [7]; porém, à medida que o modelo média-variância ganhava espaço na teoria das finanças, sedimentando o arcabouço da diversificação eficiente, algumas das suas premissas fundamentais e aspectos de aplicabilidade também passavam a ser desafiados.

A explicação dos movimentos dos preços das ações, por exemplo, recebeu grande atenção a partir da *Hipótese do Mercado Eficiente* [12, 13, 44, 45, 46], que assume um modelo de *passeio aleatório* (*random walk*)¹ para a seqüência das mudanças nos preços das ações, contrastando com a hipótese clássica dos preços seguirem um *Movimento Browniano*² e suas mudanças seguirem uma distribuição Normal [47]. A Normalidade das séries dos retornos das ações, premissa fundamental do modelo média-variância, também passa a ser questionada através da proposição de outros modelos distribucionais para os retornos das ações [11, 44, 48, 49].

¹O modelo de *passeio aleatório*, ou *passeio casual*, para uma variável aleatória X descreve o seu movimento no tempo t como $X_t = a_1 + \dots + a_t$, ou, $X_t = X_{t-1} + a_t$, onde os valores de $\{a_t, t \geq 1\}$ formam uma seqüência aleatória independente e identicamente distribuída (i.i.d.). O valor de X_t tende a oscilar no entorno da sua média $\mu_X = t\mu_a$ com variância igual $\sigma_X^2 = t\sigma_a^2$ [22]. Ou seja, segundo esse modelo, o preço da ação no tempo t é dado pelo preço da ação no tempo $t - 1$ somado a um termo aleatório. A explicação do movimento dos preços das ações pelo modelo de passeio aleatório é bastante considerada na literatura, apesar do modelo admitir uma probabilidade maior que zero de que o preço da ação assuma um valor negativo, o que, evidentemente, não encontra abrigo na realidade.

²O *Movimento Browniano Padrão* é um processo contínuo $\{W(t), 0 \leq t \leq 1\}$, tal que: (a) $W(0) = 0$ e (b) para quaisquer instantes $0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq 1$, as variáveis aleatórias $W(t_2) - W(t_1), W(t_3) - W(t_2), \dots, W(t_k) - W(t_{k-1})$ são independentes, com $W(s) - W(t)$ seguindo uma distribuição Normal com média zero e variância $s - t$; e (c) as trajetórias de $W(t)$ são contínuas e com probabilidade igual a um. Como conseqüência, o processo tem incrementos independentes e estacionários, com covariância $\gamma(s, t) = \min(s, t)$ [21].

A notória correção do modelo média-variância ao capturar as relações de interdependência de cada ativo com todos os demais, tem como consequência um custo computacional quadrático em relação ao número de ativos (ver Seção 3.4, pág. 56). Isto foi um dos principais impedimentos para a sua pronta adoção em larga escala pelo mercado³, e teve como consequência o surgimento de aproximações e simplificações do modelo original, bem como o desenvolvimento de métodos computacionais mais eficientes [50, 51, 52, 53, 54, 55, 56].

A incorporação de métodos preditivos na seleção de carteiras tem sido vislumbrada desde a proposição original de Markowitz [7], quando ele cita a possibilidade da obtenção de valores esperados e covariâncias dos retornos das ações “razoáveis”, por meio da utilização de métodos estatísticos em uma primeira aproximação, seguida pelo julgamento de especialistas para ajustes finais nos valores obtidos. Assim, o conjunto eficiente de carteiras média-variância poderia ser encontrado a partir destes conjuntos revisados dos valores esperados e covariâncias dos retornos das ações. Ainda, Markowitz sugere a possibilidade do cálculo dos valores esperados e covariâncias dos retornos das ações a partir de segmentos das séries históricas em apenas um determinado período passado⁴, indicando a possibilidade de se utilizar outros estimadores que não os momentos das distribuições para o retorno esperado e risco das ações. Contudo, o foco desta busca tradicionalmente se manteve atrelado à difícil tarefa da obtenção das distribuições *a priori* das séries de retornos [11, 57].

Recentemente, o significativo aumento do desempenho dos computadores e o surgimento de paradigmas de computação paralela de baixo custo [58] estabeleceram um ambiente propício para se reexaminar a aplicação do arcabouço média-variância em problemas do porte demandado pelo mercado. Mais especificamente, este ambiente traz novas possibilidades em relação à obtenção de melhores estimativas para o retorno esperado e risco das ações e, conseqüentemente, para o retorno esperado e risco das carteiras. Este cenário sugere não apenas a retomada da aplicação do modelo média-variância, mas, principalmente, a exploração de novos paradigmas baseados na sua modelagem original. Nesse sentido, a popularização das técnicas de inteligência artificial e inteligência computacional permite ultrapassar limitações e explorar novos modelos preditivos neste problema [41].

³Ainda na década de 1950, apesar já haver métodos eficientes para a resolução do problema de programação quadrática produzido pelo modelo média-variância, como o método da *Linha Crítica* [8] por exemplo, a dificuldade de acesso a recursos computacionais adequados restringia a aplicação do modelo a problemas modestos. Mesmo na década de 1960, o alto custo dos computadores ainda era um fator limitante. Sharpe [50] devotou espaço significativo da apresentação do seu *Modelo Diagonal* na análise do seu custo computacional. Ele destacou que, numa máquina IBM 7090, o problema quadrático do modelo média-variância era limitado a 249 ações e demorava 33 minutos para ser computado para uma carteira com 100 ações, enquanto que o *Modelo Diagonal* era limitado a 2.000 ações e computado em apenas 30 segundos para as mesmas 100 ações — o aluguel de um IBM 709, que era uma máquina da geração anterior ao IBM 7090, custava cerca de US\$ 300 por hora —; portanto, a análise do custo computacional dos modelos era um elemento a ser considerado no projeto de investimento.

⁴Este método de cálculo das médias, conhecido como *média móvel*, é bastante utilizado como ferramenta de predição na *Análise Técnica* de ações [2].

A seguir, serão apresentados alguns trabalhos relacionados à utilização de técnicas de inteligência artificial e inteligência computacional na seleção de investimentos, subdivididos em trabalhos sobre predição de ações e *trading systems*, e trabalhos sobre predição de ações na formação de carteiras.

2.1 Predição de Ações e *Trading Systems*

Chen et al. [59] propuseram uma arquitetura do tipo *Flexible Neural Tree* (FNT) para predição de séries temporais. A FNT implementa um modelo evolutivo de rede neural *feed-forward* multicamadas derivando um conjunto de operadores adequados para esta finalidade. Em um trabalho subsequente [60], os autores utilizaram várias FNTs para prever conjuntamente os índices de ações NASDAQ-100 e S&P CNX NIFTY do mercado americano. O procedimento evolutivo implementado selecionou a topologia e as entradas e atrasos temporais (*time lags*) mais importantes para a rede neural, usando como base as cotações diárias do preço de abertura, de fechamento e máximo diário. Liu, Ng e Quek [61] apresentaram um método baseado no aprendizado por reforço (*Reinforcement Learning*) para encontrar simultaneamente o número de entradas e seus atrasos temporais para um preditor neural. O desvio-padrão discretizado dos dados foi utilizado como *estado*, uma medida baseada no inverso da raiz do erro quadrático médio (erro RMSE) foi utilizada com *recompensa*, e as probabilidades de se escolher a dimensão e o atraso baseado no *estado* corrente foram utilizadas como *política*. O método foi avaliado na predição de dois modelos caóticos padrão e também na predição da série dos preços diários da ação da General Motors, para a qual as dimensões e atrasos obtidos apresentaram um menor erro RMSE comparado com aqueles obtidos pelos outros métodos de comparação. Contudo, os preditores de Chen et al. e Liu, Ng e Quek não foram utilizados para a construção de estratégias automatizadas de investimento. Entretanto, nos últimos anos, bastante atenção tem sido devotada à utilização de redes neurais artificiais na construção de sistemas que implementam estratégias automatizadas de investimento, denominados *trading systems*.

Tipicamente, os *trading systems* baseados em redes neurais empregam preditores neurais para prever os preços das ações, seus retornos ou outras medidas relacionadas, e utilizam estas predições para gerar sinais que instruem a compra, venda ou retenção das ações. Lee et al. [62] propuseram uma abordagem do tipo *dividir para conquistar* que utilizou múltiplos agentes neurais *Q-Learning* na construção de um *trading system*. O problema do investimento foi dividido em dois subproblemas: o da escolha do momento em que se deve comprar ou vender os ativos (*timing*) e o da obtenção do melhor preço de compra e venda dos ativos (*precificação*). Quatro agentes especializados na geração dos sinais e das ordens de compra e venda para cada respectivo subproblema foram distintamente empregados para esta final-

idade. O sistema foi testado empiricamente no investimento no índice Coreano KOSPI, e apresentou resultados superiores aos dos arcabouços alternativos investigados pelos autores.

Thawornwong e Enke [63] utilizaram uma técnica de *mineração de dados* para selecionar variáveis econômicas e financeiras com capacidade preditiva. As variáveis selecionadas foram utilizadas na construção de um preditor neural para os sinais dos retornos de ações. Estes sinais preditos foram utilizados para produzir uma estratégia de negociação mensal com a finalidade de decidir quando investir totalmente na carteira do índice S&P 500 ou na T-Bill (Letras do Tesouro Americano — ativo sem risco). Os resultados mostraram que os retornos mensais obtidos foram sempre superiores aos dos demais métodos de modelagem comparados.

Moody e Saffell [30] propuseram um método de *reforço direto*, denominado *recurrent reinforcement learning* (RLL), para o desenvolvimento de um *trading system*. O RLL não necessita aprender uma *função de valor*, e as estratégias de negociação são aprendidas diretamente a partir dos dados de entrada, dispensando a necessidade da predição de valores intermediários. Os autores empregaram o RLL em negociações com moedas e na alocação de ativos entre o índice S&P 500 e o T-Bill, alcançando resultados superiores aos dos sistemas que utilizaram o método *Q-Learning*.

Pantazopoulos et al. [64] apresentaram dois *trading systems* baseados em preditores *neurofuzzy*. Primeiro, uma rede neural recorrente foi utilizada para implementar a predição multivariada do preço da ação da IBM. As entradas da rede foram formadas por um conjunto de índices do mercado e cada neurônio da camada de saída forneceu uma predição do preço futuro da ação baseada em um destes índices, enquanto que os erros acumulados da rede neural forneceram uma medida de confiabilidade associada a esta predição. As predições individuais obtidas foram combinadas para produzir uma única predição do preço da ação, e simulações com uma estratégia de reinvestimento total baseada nestas predições, conduzidas durante um período de seis anos, alcançaram um retorno anualizado de 20,9%. Em seguida, empregando subconjuntos dos dados de treinamento que precederam um aumento, sustentação ou decréscimo da volatilidade, os autores treinaram três redes neurais para produzir predições *fuzzy* da tendência de curto prazo do aumento, sustentação ou decréscimo da volatilidade. Uma estratégia de negociação (compra ou venda segundo a volatilidade predita) para opções no índice S&P 500 durante 10.000 dias foi derivada, e enquanto uma carteira implementada segundo o índice S&P 500 e iniciada com \$1.000 alcançou um valor final de \$3.862, o investimento baseado no *trading system* proposto resultou num valor final de \$625.120, a partir do mesmo valor inicial e no mesmo período, porém exibiu uma volatilidade superior.

Hellström [65] criou um *ranking* normalizado para um conjunto de ações a partir dos seus retornos individuais. Os *ranks* diários das ações foram preditos por um modelo linear, e

foi alcançado 63% de acerto na predição de qual ação supera a média das ações no *ranking*. O autor construiu uma estratégia de investimento no mercado sueco de ações utilizando esses *ranks* preditos, e essa estratégia superou significativamente o desempenho do índice de mercado Swedish General. Num trabalho subsequente [66], o autor otimizou um conjunto de *indicadores técnicos*⁵ através de uma modelagem baseada em uma *janela deslizante* temporal, que produziu parâmetros diferentes para a otimização das regras de negociação em cada janela de tempo. A otimização das regras de negociação contou com uma restrição que rejeitou soluções associadas a poucos exemplos, e isso produziu um efeito de regularização que aumentou a generalização das regras de negociação obtidas. Seus resultados na geração de estratégias de investimento com 32 ações do mercado sueco alcançaram uma taxa de acerto dos sinais de negociação entre 59% e 64%, enquanto as estratégias *naive* comparadas atingiram menos que 53%.

A abordagem dos *trading systems* tem produzido resultados bastante atraentes em termos de retorno do investimento. Os *trading systems* tipicamente são concebidos para maximizar o retorno do investimento através da alternância de posição em um único ativo — notadamente aquele com a maior expectativa de retorno — ou entre um ativo de risco e um ativo de renda fixa. Entretanto, os maiores retornos de um ativo de risco são esperados nos seus períodos de maior volatilidade, o que implica em um maior risco associado, e apesar do investidor racional buscar o lucro, ele também é averso ao prejuízo [6, 7] (ver também [2, pp. 595]). Isto torna imperativa a adoção de mecanismos de controle de risco nas estratégias de investimento, como forma de satisfazer as necessidades do investidor racional.

2.2 Predição de Ações na Formação de Carteiras

Hellström [67] apresentou um algoritmo para a remoção de *outliers*⁶ que produziu uma redução de 8,9% no erro RMSE conjunto para as predições da matriz de covariâncias dos retornos de 24 ativos do mercado sueco de ações e de 3,5% para a matriz de covariâncias dos retornos de 29 ativos do índice americano Dow Jones. Em seguida, o autor otimizou carteiras com base nestas predições e verificou que a remoção dos *outliers* não afetou significativamente o cômputo das carteiras ótimas.

Lazo, Vellasco e Pacheco [68] propuseram um sistema híbrido genético-neural e estatístico para a otimização de carteiras. Um algoritmo genético foi utilizado para otimizar

⁵Indicadores técnicos são ferramentas e heurísticas utilizadas pela *Análise Técnica* de ações que buscam detectar padrões recorrentes nas séries de preços do mercado e antecipar orientações de compra e venda com base nestas predições [2].

⁶*Outliers* são valores atípicos presentes nas séries históricas, normalmente gerados por eventos espúrios ou erros de amostragem.

o modelo média-variância com um conjunto de 137 ações e as 13 ações com as maiores participações foram selecionadas para investimento. Em seguida, foi construído um modelo GARCH para a previsão da volatilidade (variância) das ações selecionadas e uma rede neural foi empregada para prever o retorno semanal futuro destas ações, utilizando seus retornos passados e as previsões da volatilidade produzidas pelo modelo GARCH. Um segundo algoritmo genético utilizou os retornos preditos, as previsões da volatilidade e as covariâncias das séries de retornos para selecionar carteiras com máximo retorno ou mínimo risco. Simulações de investimento foram realizadas durante 49 semanas entre fevereiro de 1998 e março de 1999, e o sistema proposto pelos autores apresentou desempenho superior ao do índice IBOVESPA — enquanto o IBOVESPA obteve um retorno médio semanal de $-0,221\%$ e uma variância de $59,395\%$, a carteira que maximizou o retorno alcançou um retorno médio semanal de $5,227\%$ e uma variância de $142,391\%$, e a carteira que minimizou o risco alcançou um retorno médio semanal de $0,540\%$ e uma variância de $12,209\%$.

Leung, Daouk e Chen [69] propuseram uma abordagem multiobjetivo que empregou programação por metas e conceitos da análise média-variância-assimetria para otimizar carteiras formadas com quatro previsões distintas de um único ativo, obtidas com o método de alisamento exponencial adaptativo, com duas variações do modelo ARIMA e com um preditor baseado em redes neurais. Combinando os excessos de retorno médios de cada método de previsão e um conjunto de preferências do investidor em relação a média, variância e assimetria, o modelo atuou como um combinador de preditores que obteve uma previsão de “consenso” ótima que se beneficiou da redução da variância (risco) oferecida pelo modelo de diversificação empregado. Seus experimentos de investimento nos índices S&P 500, FTSE 100 e NIKKEI 225 alcançaram resultados superiores aos métodos que utilizaram um modelo de previsão único e aos métodos padrão para a combinação de preditores que foram avaliados no estudo.

Hung, Cheung e Xu [70] estenderam o *trading system* ASLD [71] adicionando um mecanismo de controle de risco baseado em redes neurais. Durante o treinamento, quando os preços futuros são conhecidos, o sistema ASLD estendido (EASLD) examina um conjunto de N ativos e utiliza as suas séries de preços conhecidos para gerar instruções de compra, venda ou retenção das suas posições nestes ativos. Os ativos que recebem um sinal de compra têm suas participações ótimas calculadas com o modelo média-variância e com um modelo desenvolvido pelos autores, que combina o *Sharpe ratio* [72] e o *risco downside* [73] — esse modelo também utiliza os retornos médios como retornos esperados das ações. Em seguida, N redes neurais, uma para cada ativo, são treinadas para produzir diretamente estas participações ótimas ou, quando a compra de posições no respectivo ativo não for indicada, produzir os sinais de venda ou retenção. Durante o teste, quando os preços futuros são desconhecidos, as saídas destas redes neurais produzem os sinais de compra, venda ou retenção de participações eficientes nos ativos — ou seja, é esperado que as redes neurais tenham aprendido

a operação completa do sistema de negociação com controle de risco (instruções de negociação e otimização de carteiras). Os experimentos apresentados pelos autores simularam o investimento em uma carteira de seis índices de mercado globais — S&P 500 (USA), Hang Seng Index (Hong Kong), Shanghai SE Composite (China), NIKKEI 255 (Japão), CAC 40 (França) e Australia SE (Austrália) — durante 365 dias e mostraram os benefícios do mecanismo de controle de risco, alcançando retornos superiores aos do ASLD com somente 20% do seu risco.

O nosso modelo de seleção de carteiras baseado em erros de predição contempla mecanismos de controle de risco e também pode ser utilizado como um *trading system*, conforme mostraremos na Seção 6.2.5.

Capítulo 3

Noções Fundamentais

3.1 Mercado de Ações¹

As empresas têm um papel fundamental na organização socioeconômica do mundo moderno. São agentes geradores de emprego e desenvolvimento, e tipicamente os maiores responsáveis pela produção de bens e serviços [74].

A empresa, “organização particular, governamental, ou de economia mista, que produz e/ou oferece bens e serviços, com vista, em geral, à obtenção de lucros”², é uma entidade organizada que congrega capital e trabalho, constituída basicamente na forma de sociedades. Dentre as diversas formas societárias estabelecidas pelo Novo Código Civil brasileiro, instituído pela Lei nº 10.406/02 [75], as mais comuns são as *sociedades limitadas* e as *sociedades anônimas*.

As sociedades limitadas são constituídas através de contrato entre os sócios, denominado *contrato social*, que estabelece as participações individuais no capital social e no resultado da empresa. Destacam-se ainda o caráter nominativo das cotas e a responsabilidade de cada sócio restrita ao valor das suas cotas. Contudo, todos os sócios respondem solidariamente pela integralização do capital social. Esta última característica faz com que estas sociedades tenham dificuldades no aporte vultoso de recursos para empreendimentos de risco.

Nesse contexto, as sociedades anônimas, ou *companhias*, são uma alternativa societária que favorece o aporte vultoso de recursos para empreendimentos de risco. Nesta modalidade societária, o capital divide-se em ações e a responsabilidade dos acionistas (sócios) é proporcional à quantidade de ações (cotas) de sua propriedade.

¹Esta seção é uma versão atualizada de Freitas [32, Seção 2.1]

²Novo Dicionário Aurélio da Língua Portuguesa, 2ª edição revista e ampliada.

Ações são valores mobiliários³ emitidos pelas companhias, que representam as menores parcelas nas quais se divide o capital da companhia [76]. Em última instância, as ações são títulos de propriedade das empresas constituídas como sociedades anônimas, negociados tanto no *mercado primário* quanto no *mercado secundário*.

As negociações no mercado primário se dão na oferta pública de ações (*initial public offering* — IPO), quando a companhia emite uma quantidade de ações e a receita obtida através desta *subscrição* é revertida para a companhia. No mercado secundário, as ações são negociadas nas Bolsas de Valores ou mercados de balcão, com operações realizadas sob as formas de negociação a vista, a termo, a futuro ou opções⁴.

Quanto à espécie, as ações são classificadas como:

- Ações ordinárias — conferem direito a voto nas assembléias gerais das empresas, com cada ação equivalendo a um voto, e proporcionam participação nos resultados da empresa.
- Ações preferenciais — possuem prioridade na distribuição de dividendos e no reembolso de capital no caso da liquidação da companhia, podem ter direito a voto em ocasiões especiais (tipicamente após três exercícios consecutivos sem distribuição), desde que previsto no Estatuto Social da empresa.

Quanto à forma de circulação, as ações podem ser:

- Ao portador — sem um titular especificado na emissão, o proprietário é o dono do título.
- Nominativas — são emitidas através de certificados em nome de um titular. A venda deve ser registrada na empresa que a emitiu.

³Valor mobiliário é um título de investimento que a sociedade anônima emite para a obtenção de recursos. Além das ações, há diversos valores mobiliários derivados, tais como as debêntures, os bônus de subscrição e os certificados de emissão de garantia.

⁴Operação a vista é a compra ou venda de uma determinada quantidade de ações com pagamento realizado pelo comprador no ato do fechamento do negócio e entrega das ações objeto da transação realizada pelo vendedor no terceiro dia útil após a realização do negócio, período denominado “D+3”. Operações a termo são realizadas através de contratos para compra ou a venda de uma determinada quantidade de ações a um preço fixado, com liquidação em um prazo determinado pelos contratantes de no mínimo 16 e no máximo 999 dias corridos. Operações no mercado futuro são realizadas através de contratos padronizados de compra ou venda de ações a um preço acordado entre as partes e com liquidação em uma data futura específica, previamente autorizada — um contrato futuro é, em sua essência, um contrato a termo padronizado. Operações no mercado de opções são realizadas através da negociação de direitos de compra ou venda de um lote de ações, com preços e prazos de exercício pré-estabelecidos contratualmente. O titular de uma opção de compra pode exercê-la até a data de vencimento ou revendê-la ao mercado, enquanto que o titular de uma opção de venda pode exercê-la somente na data de vencimento revendê-la ao mercado [77].

- Nominativas endossáveis — são ações nominativas na emissão mas que podem ser transformadas em ações ao portador nas demais transferências através de endosso do título.
- Escriturais — são ações nominativas que não têm emissão de certificado e são mantidas numa conta de depósito do seu titular. O controle é realizado por instituições financeiras autorizadas, em um tipo de conta corrente na qual os valores são lançados a débito ou a crédito dos titulares, não havendo portanto movimentação física de documentos.

Uma vez subscritas, as ações tornam-se bens negociáveis cujos preços variam ao longo do tempo, constituindo assim instrumentos de investimento que permitem a realização de lucro ou prejuízo em transações de compra e venda. As principais entidades mediadoras destas transações são as *corretoras de valores*, que se congregam em outras entidades, as *Bolsas de Valores*.

As Bolsas de Valores são associações civis sem fins lucrativos cujos sócios são as corretoras de valores. As transações de compra e venda são realizadas dentro do recinto das Bolsas de Valores pelos funcionários ou sistemas das corretoras durante um expediente denominado *pregão*, que pode ser viva-voz ou eletrônico, ou após o encerramento do pregão principal, no serviço eletrônico denominado *after-market*.

Dentre as principais atividades de competência das Bolsas de Valores, destacam-se [74]:

- Manter local ou sistema adequado à realização de operações de compra e venda de títulos e valores mobiliários em mercado livre e aberto, especialmente organizado e fiscalizado pelas corretoras-membro e pelas autoridades reguladoras do mercado.
- Dotar o referido local ou sistema de todos meios necessários à pronta e eficiente realização e visibilidade das operações.
- Divulgar as operações realizadas com rapidez e riqueza de detalhes.

A principal Bolsa de Valores brasileira é a Bolsa de Valores de São Paulo — BOVESPA [78]⁵.

⁵Em 2008, a Bolsa de Mercadorias & Futuros (BM&F) foi integrada à Bolsa de Valores de São Paulo (BOVESPA), formando a BM&FBOVESPA S.A — Bolsa de Valores, Mercadorias e Futuros.

3.2 Retorno de Investimentos

Um projeto de investimento, na sua forma mais simples, é concebido a partir de um período de tempo denominado *horizonte de investimento*, onde o capital inicial W_i é investido no seu início e o capital final W_f é recuperado no seu final. O capital final obtido é, obviamente, uma medida útil acerca do desempenho deste investimento, mas torna-se sem utilidade na comparação entre investimentos cujos capitais iniciais aportados tenham sido diferentes. Neste caso, uma medida relativa que seja independente do capital inicial deve ser utilizada. Esta medida é denominada *retorno do investimento*, ou simplesmente *retorno*, e se propõe a quantificar a variação do capital ao longo do horizonte de investimento. O *retorno líquido simples* [22] de uma ação, ou simplesmente *retorno da ação*, é definido como a seguir.

O retorno de L períodos da ação no tempo t , r_{tL} , é definido como a diferença entre o preço da ação no tempo t e o preço da ação no tempo $t - L$, dividido pelo preço da ação no tempo $t - L$, conforme mostrado na Eq. 3.1:

$$r_{tL} = \frac{P_t - P_{t-L}}{P_{t-L}}, \quad t \geq L, \quad (3.1)$$

onde P_t e P_{t-L} são os preços da ação nos tempos t e $t - L$, respectivamente ⁶.

Neste trabalho, estaremos particularmente interessados em modelos e análises para retornos de um período ($L = 1$), desta forma, a Eq. 3.1 é reescrita como:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}, \quad t \geq 1, \quad (3.2)$$

onde r_t é o retorno de um período da ação no tempo t , e P_t e P_{t-1} são os preços da ação nos tempos t e $t - 1$, respectivamente.

As medidas de retorno definidas nas Eqs. 3.1 e 3.2 são relativas e adimensionais. Por essa razão, é comum encontrá-las definidas na literatura como valores percentuais, onde são multiplicadas por um fator de 100.

Os investimentos no mercado de ações são tipicamente executados de forma consecutiva no tempo ou têm seus retornos intermediários observados em intervalos regulares de

⁶A Eq. 3.1 é uma expressão simplificada para o retorno da ação, que desconsidera os pagamentos em moeda ocorridos no período, como os dividendos por exemplo, bem como outros ajustes nos preços das ações, como os eventos de *desdobramento* (*split*) e *agrupamento* (*reverse split*). Os pagamentos em moeda ocorridos no período são somados ao numerador da Eq. 3.1 [2], enquanto os ajustes nos preços são tratados pela normalização dos preços pelos fatores de ajuste nos respectivos períodos [76]. As séries históricas utilizadas neste trabalho tiveram seus preços devidamente ajustados pelas sequências de pagamentos e eventos ocorridos com cada ação.

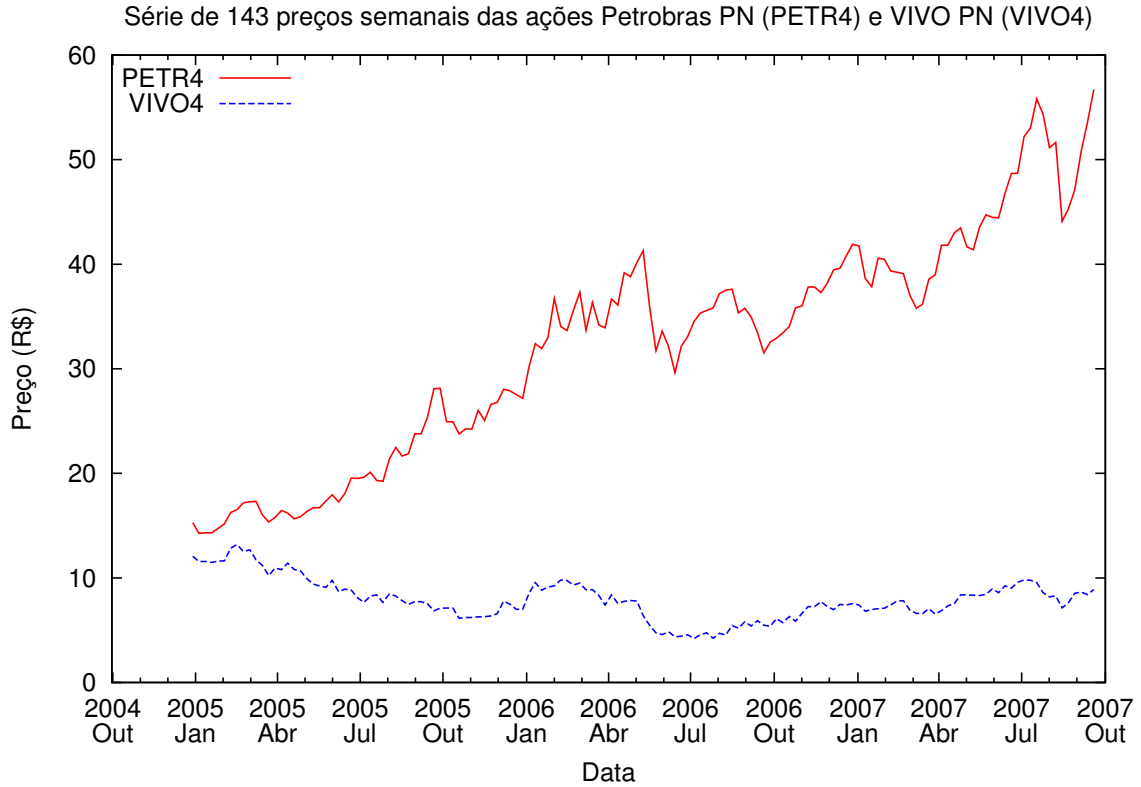


Figura 3.1: Exemplos de séries históricas de preços de ações do mercado brasileiro. São mostradas as séries de 143 preços semanais de ações de duas grandes empresas do mercado brasileiro dos setores de petróleo (PETROBRAS — PETR4) e telefonia (Vivo — VIVO4), observados às quartas-feiras entre 29-Dez-2004 e 19-Set-2007. Os perfis distintos de valorização (PETR4) e desvalorização (VIVO4) no período são evidenciados.

tempo dentro do horizonte de investimento. As séries de observações dos retornos realizados são uma importante fonte de informação acerca do desempenho futuro destas ações — seu *retorno esperado*. Neste trabalho, nós utilizamos esta fonte de informação para obter as previsões dos retornos futuros das ações como a seguir.

As séries dos $N + 1$ preços e dos N retornos de uma ação⁷, observados sequencialmente e regularmente no tempo, formam as *séries históricas*, ou *séries temporais* [20, 22], \mathbf{P}' e \mathbf{r}' , respectivamente, que são definidas como:

$$\mathbf{P}' = (P_0, P_1, \dots, P_N), \text{ e} \quad (3.3)$$

$$\mathbf{r}' = (r_1, r_2, \dots, r_N). \quad (3.4)$$

As Figuras 3.1 e 3.2 mostram dois exemplos típicos de séries de preços e de retornos do

⁷A série de retornos de uma ação é derivada da sua série de preços conforme a Eq. 3.2, e, portanto, tem um elemento a menos.

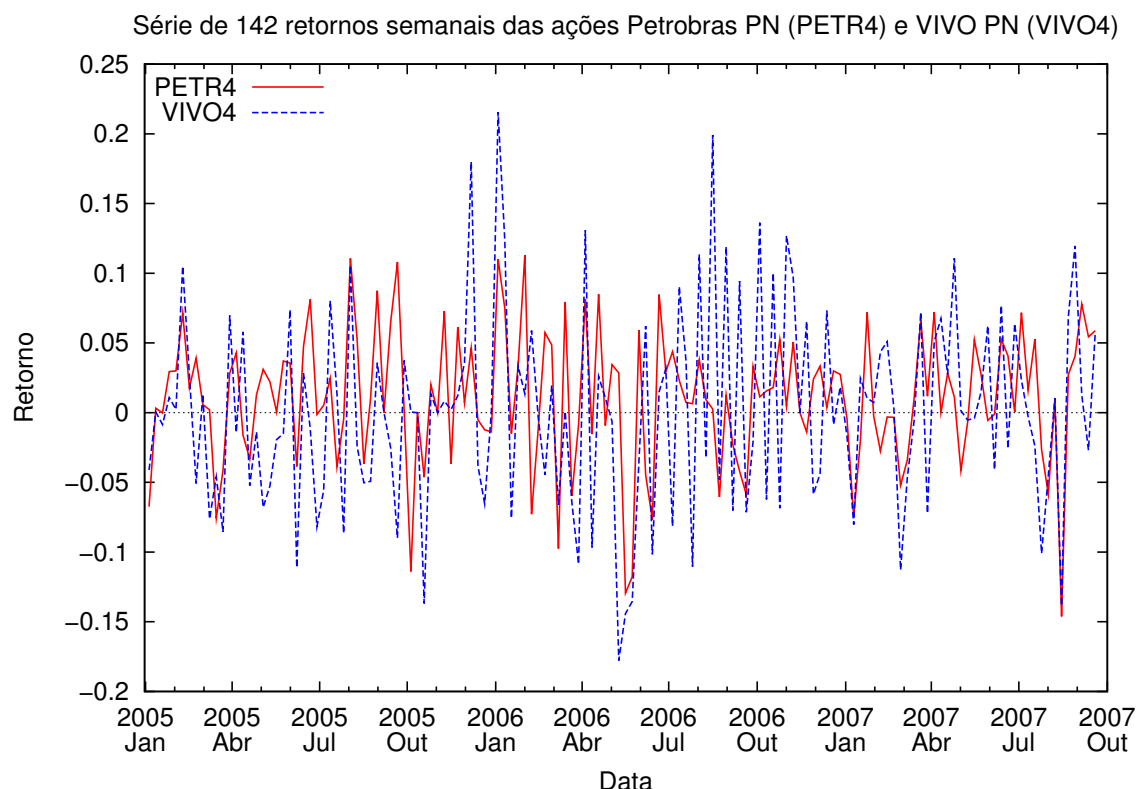


Figura 3.2: Exemplos de séries históricas de retornos de ações do mercado brasileiro. São mostradas as séries de 142 retornos semanais de ações de duas grandes empresas do mercado brasileiro dos setores de petróleo (PETROBRAS — PETR4) e telefonia (Vivo — VIVO4). As séries foram calculadas entre 05-Jan-2005 e 19-Set-2007 a partir dos preços das ações observados às quartas-feiras. Os perfis distintos de oscilação são evidenciados, onde a ação VIVO4 apresenta uma maior oscilação que a ação PETR4.

mercado de ações brasileiro. Nas figuras são mostradas ações de duas grandes empresas do mercado, PETR4 (Petróleo Brasileiro S.A. — PETROBRAS PN) e VIVO4 (Vivo Participações S.A. PN), ambas listadas na Bolsa de Valores de São Paulo — BOVESPA, a maior bolsa brasileira. A Fig. 3.1 mostra as séries de 143 preços semanais das ações PETR4 e VIVO4, observados entre 29-Dez-2004 e 19-Set-2007, onde podemos notar os perfis distintos de valorização da ação PETR4 e de desvalorização da ação VIVO4; enquanto que a Fig. 3.2 mostra as séries dos 142 retornos calculados para as duas ações entre 05-Jan-2005 e 19-Set-2007, onde notamos que a ação VIVO4 apresenta uma oscilação superior à ação PETR4.

3.3 Predição de Retornos de Ações com Redes Neurais

A abordagem tradicional para o problema da predição de retornos de ações é oferecida pela *Análise de Séries Temporais*, que deriva diversas famílias de modelos estocásticos para a previsão de séries históricas [20, 21, 22, 79, 80]. Contudo, por tratarem de modelos paramétricos, a natureza assistida e iterativa do processo de identificação e estimação de parâmetros, e, principalmente, da seleção de modelos [81, 82], traz dificuldades práticas em aplicações com um grande número de séries históricas. As estruturas adaptativas fornecidas pelas redes neurais artificiais, por outro lado, são uma alternativa de modelagem não paramétrica bastante adequada para estes tipos de aplicações [32, 33, 34, 41, 83].

3.3.1 Redes Neurais Artificiais⁸

As redes neurais artificiais são um paradigma de computação que se inspira no modelo computacional do cérebro humano [87]. A partir da interconexão massiva de unidades elementares de processamento, denominadas *neurônios*, segundo uma determinada arquitetura, o cérebro humano torna-se um complexo computador não-linear e paralelo [88]. Esta máquina implementa um sistema de processamento de informação no qual seus componentes continuamente processam as informações recebidas nas suas entradas, transformando-as em novas informações nas suas saídas.

As conexões entre os neurônios, ou entre estruturas sensoriais e neurônios, são, de forma simplificada, realizadas por circuitos denominados *sinapses*, que controlam o fluxo de informações através de mecanismos de excitação e inibição. Desta forma, a função de transformação implementada nos componentes do cérebro é fundamentalmente determinada pelos padrões de interconexão dos neurônios, pelos seus níveis de ativação e pelos estados (excitação ou inibição) das suas sinapses. Uma das características fundamentais do cérebro humano é a modificação destes padrões sinápticos de forma adaptativa em resposta às mudanças no ambiente, através de mecanismos denominados de *aprendizado*.

As redes neurais artificiais são construídas a partir de modelos matemáticos de componentes do cérebro humano, e conseguem reproduzir, em determinados níveis, diversas propriedades do cérebro humano. Dentre as propriedades mais importantes exibidas pelas redes neurais artificiais, destacam-se:

⁸Diversos métodos de classificação têm recebido o nome de *redes neurais*, tais como as redes neurais sem peso (*Weightless Neural Network* — WNN) [84] e as redes neurais probabilísticas (*Probabilistic Neural Network* — PNN) [85]. Esta seção trata apenas dos modelos clássicos de redes neurais, com neurônios artificiais derivados do modelo de McCulloch e Pitts [86].

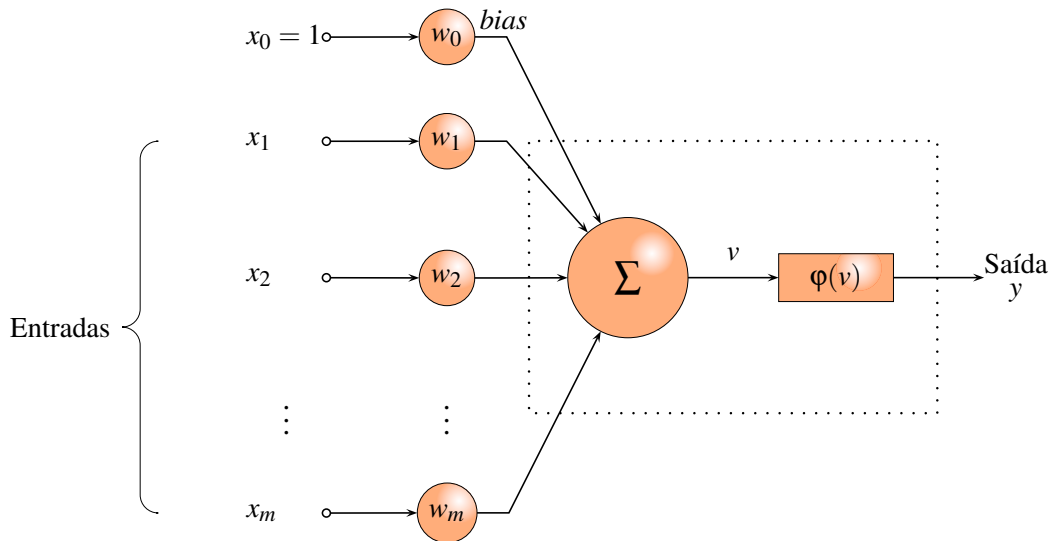


Figura 3.3: Modelo do neurônio artificial. O neurônio artificial produz um campo local induzido, v , que é resultado do produto escalar do seu vetor de entradas, \mathbf{x} , e do seu vetor de pesos sinápticos, \mathbf{w} . A saída, ou ativação, do neurônio, y , é então obtida pela aplicação do seu campo local induzido a uma função de ativação $\phi()$.

- **Aprendizado** — através da apresentação sistemática de exemplos de entradas e saídas, a rede neural consegue, por meio de um algoritmo de treinamento adequado, ajustar sua representação interna de conhecimento de forma a mapear as relações entre as entradas e as saídas apresentadas.
- **Não-linearidade** — as redes neurais podem aproximar mapeamentos de entrada e saída com não-linearidade virtualmente de qualquer ordem, propriedade especialmente importante quando o fenômeno a ser aprendido é de natureza inerentemente não-linear [88, pp. 208–209].
- **Plasticidade** — uma vez treinada para representar as entradas e saídas em um determinado ambiente, a representação do conhecimento da rede neural pode ser adaptada a partir de treinamento adicional para incorporar mudanças no ambiente. Esta propriedade é determinante na representação de fenômenos não-estacionários.
- **Generalização** — a partir do treinamento adequado para a representação de um conjunto de entradas e saídas, a rede neural pode produzir saídas razoáveis para entradas desconhecidas, isto é, distintas daquelas entradas que foram apresentadas durante o treinamento.

Neurônio Artificial

Um modelo típico do neurônio artificial com m entradas é mostrado na Fig. 3.3, onde os valores x_1, x_2, \dots, x_m são aplicados a cada entrada do neurônio e os valores w_1, w_2, \dots, w_m formam os *pesos sinápticos*, ou simplesmente *pesos*, associados a cada uma destas entradas. Cada peso multiplica o valor da sua respectiva entrada, fornecendo um modelo matemático do estado da sua sinapse (excitação ou inibição). O neurônio, desta forma, implementa um combinador linear que computa o produto escalar entre o vetor de entradas $\mathbf{x} = (x_1, x_2, \dots, x_m)$ e o vetor de pesos $\mathbf{w} = (w_1, w_2, \dots, w_m)$. O peso w_0 aplicado à entrada fixa $x_0 = 1$ é denominado *bias*, e é tratado como um parâmetro externo ao neurônio artificial, com a finalidade específica de controlar o nível da saída v do combinador linear. A saída do combinador linear, denominada *campo local induzido*, é dada pela Eq. 3.5:

$$\begin{aligned} v &= \sum_{i=1}^m x_i w_i + w_0 \\ &= \sum_{i=0}^m x_i w_i, \end{aligned} \quad (3.5)$$

onde v é a saída do combinador linear, x_1, \dots, x_m são as entradas do neurônio, w_1, \dots, w_m são os pesos sinápticos, x_0 tem valor fixo igual a um e w_0 é o *bias* do neurônio.

A saída, ou *ativação*, do neurônio, y , é obtida aplicando-se a *função de ativação* $\varphi()$ à saída do combinador linear, v , conforme mostrado na Fig. 3.3. Desta forma, temos que:

$$y = \varphi(v), \quad (3.6)$$

onde v é a saída do combinador linear do neurônio artificial, calculada pela Eq. 3.5, e $\varphi()$ é a função de ativação do neurônio artificial.

A função de ativação é o componente que possibilita que a saída do neurônio tenha uma relação não-linear com as suas entradas, e, conseqüentemente, que o neurônio consiga representar mapeamentos de entrada e saída mais complexos do que as combinações lineares das suas entradas.

A escolha da função de ativação do neurônio artificial é uma decisão de projeto que depende de vários aspectos, tais como as características do problema a ser aprendido pela rede neural, o algoritmo de treinamento que será utilizado e aspectos de implementação e desempenho.

Diversas funções de ativação são descritas na literatura [87, 88], dentre as mais utilizadas temos:

- *Função Linear* — produz uma ativação linear em v , com inclinação dada pelo parâmetro a e nível dado pelo parâmetro b , conforme a Eq. 3.7:

$$\varphi(v) = av + b. \quad (3.7)$$

A Fig. 3.4a mostra um exemplo da função linear para $a = 0,5$ e $b = 0,5$.

- *Função Linear por Partes* — intercala uma região linear em v , com inclinação a e nível b , situada entre os limiares v_0 e v_1 , com duas regiões constantes em 0 e 1, respectivamente, conforme a Eq. 3.8:

$$\varphi(v) = \begin{cases} 0, & v \leq v_0; \\ av + b, & v_0 < v < v_1; \\ 1, & v \geq v_1. \end{cases} \quad (3.8)$$

A Fig. 3.4b mostra um exemplo da função linear por partes para $a = 0,5$, $b = 0,5$, $v_0 = -1$, e $v_1 = 1$.

- *Função Limiar* — também conhecida como *função degrau unitário*, ou *função de Heaviside*, assume valor zero para v negativo e valor unitário nos outros casos, conforme a Eq. 3.9:

$$\varphi(v) = \begin{cases} 0, & v < 0; \\ 1, & v \geq 0. \end{cases} \quad (3.9)$$

A Fig. 3.4c mostra a função limiar, que é a função de ativação utilizada no neurônio de McCulloch e Pitts [86].

- *Função Sigmóide* — as sigmóides são uma classe de funções estritamente crescentes com formato de “S”, que apresentam transições suaves entre regiões próximas da linearidade e regiões fortemente não-lineares. A *função logística* é uma sigmóide bastante utilizada, sendo definida conforme a Eq. 3.10, onde o parâmetro a fornece a sua inclinação, o parâmetro b o seu nível e o parâmetro c a sua amplitude:

$$\varphi(v) = \frac{c}{1 + \exp(-av)} + b. \quad (3.10)$$

A Fig. 3.4d mostra exemplos da função logística com $c = 1$, $b = 0$ e inclinações $a = 0,1, 0,2, 0,5$, e $1,0$. Como pode ser observado na figura, à medida que o valor

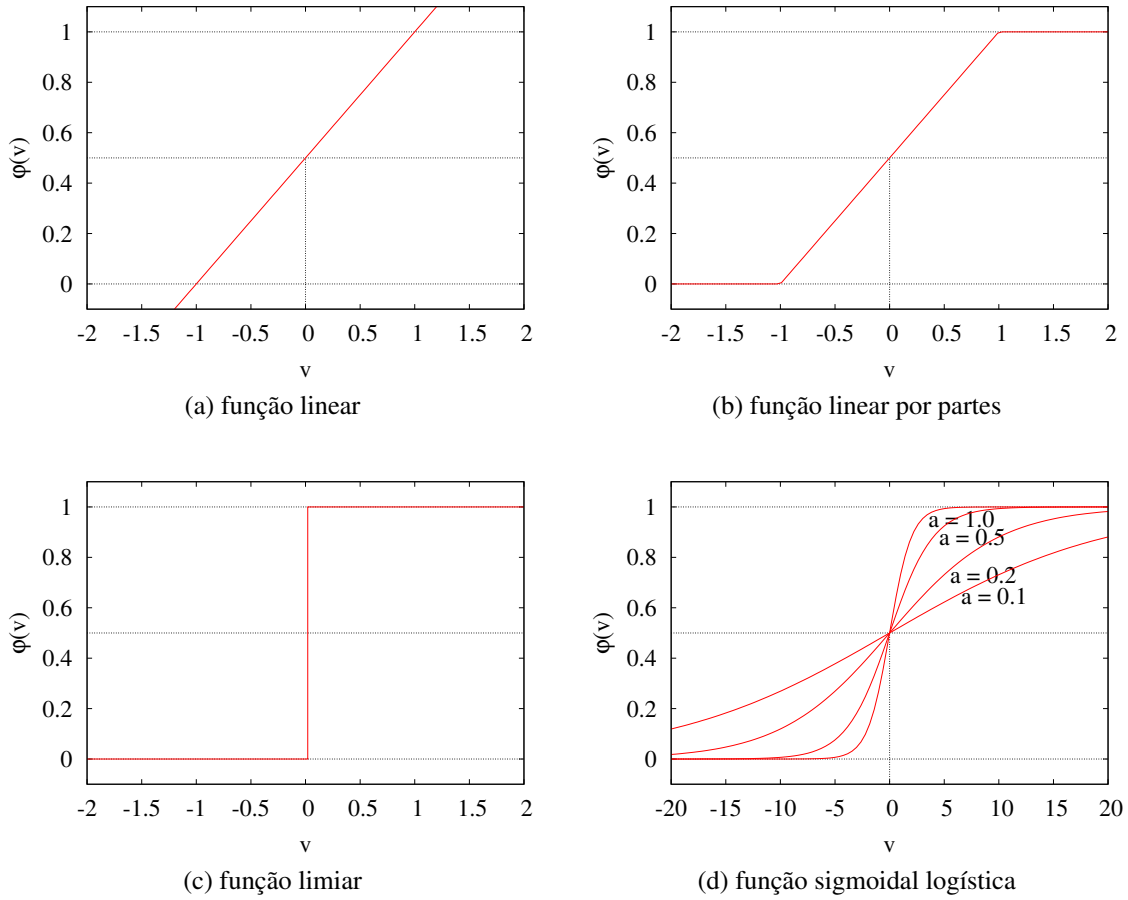


Figura 3.4: Exemplos de algumas funções de ativação: (a) função linear, (b) função linear por partes, (c) função limiar e (d) função sigmoideal logística.

de a diminui, o formato da função logística aproxima-se do formato da função linear; e à medida que o valor de a aumenta, seu formato aproxima-se do formato da função limiar. A função logística é diferenciável em todo seu domínio, o que é uma propriedade necessária nos diversos algoritmos de treinamento baseados em gradiente local, tal como o algoritmo *back-propagation* [88, 89, 90]. Estas características fazem da função logística uma escolha padrão para a função de ativação nos neurônios artificiais.

Topologias de Redes Neurais

As redes neurais são construídas a partir da interconexão de neurônios artificiais segundo uma organização topológica denominada *arquitetura*, ou *topologia*, da rede. Tipicamente, os neurônios são organizados em camadas, que recebem como entrada as sinapses da camada

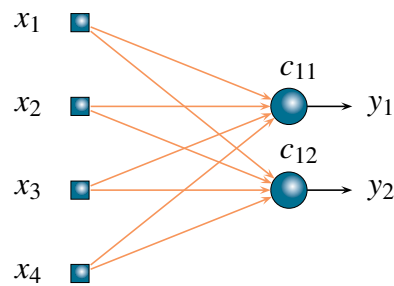


Figura 3.5: Exemplo de rede neural *feedforward* totalmente conectada com uma camada.

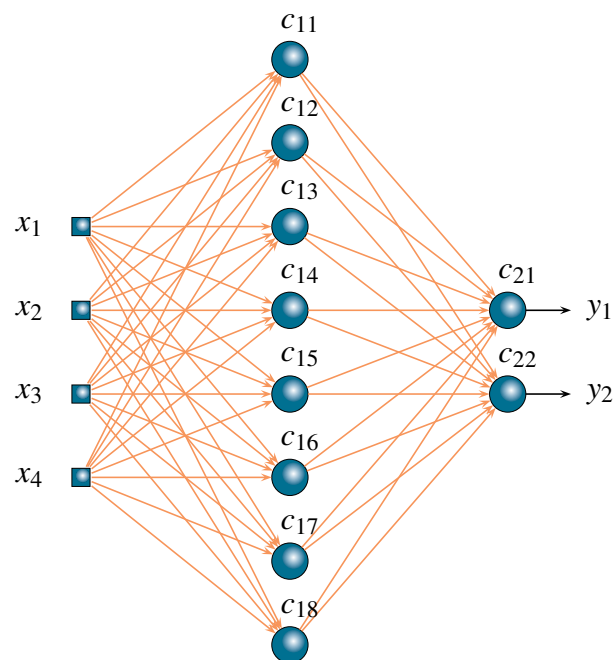


Figura 3.6: Exemplo de rede neural *feedforward* totalmente conectada com duas camadas.

anterior e projetam suas saídas na camada posterior. As entradas da rede se conectam aos neurônios da sua primeira camada e as saídas da rede são obtidas nos neurônios da camada de saída. As camadas intermediárias entre as entradas e a camada de saída recebem o nome de camadas *ocultas*.

Em relação à orientação das interconexões dos neurônios, a rede neural pode ser do tipo *com alimentação direta*, ou *feedforward*, quando as saídas dos neurônios de uma camada se conectam somente às entradas dos neurônios da camada posterior, formando um grafo acíclico direcionado; ou do tipo *recorrente*, quando há caminhos de retro-alimentação (*feed-back*) entre as camadas [88].

A Fig. 3.5 mostra um exemplo elementar de uma rede neural *feedforward* mono-camada com topologia 4:2 — quatro entradas diretamente conectadas à camada de saída com dois neurônios. Nessa rede, as entradas x_1 , x_2 , x_3 e x_4 são conectadas diretamente à camada de saída formada pelos neurônios c_{11} e c_{12} , que produzem as saídas da rede, y_1 e y_2 . A Fig. 3.6 mostra uma extensão da rede neural da figura anterior para uma topologia multicamada 4:8:2, através da inserção de uma camada oculta com oito neurônios (c_{11} a c_{18}). Estes dois exemplos apresentam redes *feedforward totalmente conectadas*, onde cada entrada se conecta a todos os neurônios da primeira camada, cada neurônio da primeira camada se conecta a todos os neurônios da segunda camada, e assim sucessivamente. Muito embora as redes totalmente conectadas sejam padrão na literatura e nas aplicações, a escolha do arranjo de interconexão dos neurônios da rede também é um aspecto de topologia a ser considerado.

A escolha da topologia de rede neural é uma decisão de projeto norteadas, entre outros fatores, pela complexidade do problema a ser aprendido. Minsky e Papert [91] demonstraram que as topologias mono-camada conseguem aprender apenas os problemas linearmente separáveis, que são aqueles cuja superfície de decisão é um hiperplano [88, pp. 138], e, portanto, os problemas mais complexos devem ser atacados através de topologias multicamadas. A topologia da rede também impõe restrições quanto à escolha do algoritmo de treinamento, sendo que para as redes *feedforward* multicamadas, a escolha padrão é o algoritmo *back-propagation*.

Algoritmo *Back-propagation*

Para que uma rede neural possa ser utilizada em uma tarefa específica, como na predição de séries históricas por exemplo, ela necessita adquirir previamente o conhecimento necessário para desempenhar essa tarefa de forma satisfatória. Enquanto a quantidade de neurônios, o número de camadas e as funções de ativação escolhidas são parâmetros fixos no modelo da rede neural, os pesos sinápticos são seus parâmetros ajustáveis. O estado do conjunto de pesos representa o conhecimento adquirido durante a fase de aquisição de conhecimento, denominada *treinamento*, ou *aprendizado*, que precede a fase de utilização, ou *operação*, da rede neural.

O treinamento *supervisionado* consiste na utilização sistemática de um conjunto de exemplos de entradas e saídas, denominado *conjunto de treinamento*, num processo iterativo de ajuste dos pesos da rede. Estes ajustes adaptam a representação interna de conhecimento da rede, fazendo com que ela aprenda a representar as relações existentes entre as entradas e suas respectivas saídas. Cada elemento do conjunto de treinamento deve conter uma entrada da rede neural e a sua respectiva saída desejada, formando um *par de treinamento* $\langle \mathbf{x}, \mathbf{d} \rangle$,

onde \mathbf{x} é o vetor de entradas da rede e \mathbf{d} é o vetor de saídas desejadas. Por exemplo, para as redes neurais das Figuras 3.5 e 3.6, temos os vetores $\mathbf{x} = (x_1, x_2, x_3, x_4)$ e $\mathbf{d} = (y_1, y_2)$.

Desta forma, o erro entre a saída desejada e a saída produzida pela rede neural pode ser calculado para cada par de treinamento. Rumelhart, Hinton e Williams [89] derivaram um método eficiente para utilizar este erro como fonte de informação para o ajuste dos pesos de uma rede neural multicamadas: o algoritmo *back-propagation* propaga o erro presente na saída da rede no sentido inverso das suas camadas — da saída para a entrada; calculando, em cada camada, os ajustes dos pesos que corrigem a saída da rede de forma a minimizar o seu erro.

Sejam o conjunto de treinamento $\mathbb{T} = \{\langle \mathbf{x}, \mathbf{d} \rangle_1, \dots, \langle \mathbf{x}, \mathbf{d} \rangle_{N_T}\}$, onde $\mathbf{x} = (x_1, \dots, x_{M_0})$ e $\mathbf{d} = (d_1, \dots, d_{M_L})$ são os vetores de entradas e de saídas desejadas; o vetor de saídas $\mathbf{y} = (y_1, \dots, y_{M_L})$; e o vetor de erros $\mathbf{e} = (e_1, \dots, e_{M_L})$, respectivamente, para uma rede neural de L camadas, com M_0 entradas (camada 0) e M_L neurônios na camada de saída (camada L). Para cada par de treinamento $\langle \mathbf{x}, \mathbf{d} \rangle_n$, representamos seus vetores \mathbf{x}_n e \mathbf{d}_n , bem como os vetores de saída e de erro associados, respectivamente, como $\mathbf{x}(n)$, $\mathbf{d}(n)$, $\mathbf{y}(n)$ e $\mathbf{e}(n)$; e calculamos o erro produzido em cada neurônio j da camada de saída da rede neural como:

$$e_j^{(L)}(n) = d_j^{(L)}(n) - y_j^{(L)}(n), \quad (3.11)$$

onde o índice sobrescrito (\cdot) indica em que camada da rede neural a grandeza é observada, e, nesse caso, $e_j^{(L)}(n)$ é o erro no neurônio j da camada de saída da rede (camada L) para o par de treinamento n ; e $d_j^{(L)}(n)$ e $y_j^{(L)}(n)$ são, respectivamente, a saída desejada e a saída produzida no neurônio de saída j para o par de treinamento n .

A energia do erro total na camada de saída da rede neural para o par de treinamento n é dada por:

$$\mathcal{E}(n) = \frac{1}{2} \sum_{j=1}^{M_L} (e_j^{(L)}(n))^2, \quad (3.12)$$

onde $\mathcal{E}(n)$ é a energia total dos erros $e_j^{(L)}(n)$ observados em cada um dos M_L neurônios de saída para o par de treinamento n .

Assim, a energia média do erro total na camada de saída da rede neural, para todo o conjunto de treinamento \mathbb{T} , é dada por:

$$\bar{\mathcal{E}} = \frac{1}{N_T} \sum_{n=1}^{N_T} \mathcal{E}(n), \quad (3.13)$$

onde $\bar{\mathcal{E}}$ é a energia média do erro total na camada de saída da rede neural, N_T é o número de exemplos do conjunto de treinamento e $\mathcal{E}(n)$ é a energia do erro total na camada de saída

para o par de treinamento n . Os valores de $\mathcal{E}(n)$ e, conseqüentemente, de $\bar{\mathcal{E}}$, dependem dos pesos aprendidos pela rede neural, e a energia média $\bar{\mathcal{E}}$ é utilizada como uma medida de desempenho do treinamento, representando uma *função de custo* a ser minimizada.

Os passos do algoritmo *back-propagation* são descritos a seguir.

1. *Inicialização* — o algoritmo *back-propagation* calcula ajustes incrementais nos pesos da rede que minimizam a energia média do erro total na sua saída. Portanto, o estado inicial dos pesos tem impacto determinante no desempenho do algoritmo — por exemplo, no caso de muitos neurônios serem inicializados com a ativação saturada, serão produzidos gradientes locais muito pequenos, que impedirão a efetiva correção dos valores dos pesos, prendendo a busca no espaço de pesos em um mínimo local com alta energia.

Assim, na ausência de informações *a priori*, os valores iniciais dos pesos da rede devem ser obtidos a partir de uma distribuição uniforme com média zero e variância escolhida de forma que o desvio padrão dos campos locais induzidos dos neurônios estejam na transição entre as regiões lineares e não-lineares das funções de ativação.

Após a inicialização, a energia média $\bar{\mathcal{E}}$ é minimizada através de um procedimento iterativo com dois passos, que são executados para cada um dos N_T pares de treinamento, e descritos como:

2. *Passo direto* — o vetor $\mathbf{x}(n)$ é inserido na entrada da rede, produzindo as saídas $y_j^{(1)}(n)$ dos neurônios da primeira camada oculta (camada 1) conforme as Eqs. 3.5 e 3.6. Estas saídas alimentam as entradas dos neurônios da próxima camada oculta (camada 2), cujas saídas $y_j^{(2)}(n)$ são produzidas fazendo $x_i = y_j^{(1)}(n)$ nas Eqs. 3.5 e 3.6. Este processo se repete, camada a camada, no sentido da camada de saída da rede neural (camada L), até que o vetor de saída $\mathbf{y}^{(L)}(n)$ seja obtido e o vetor de erros $\mathbf{e}^{(L)}(n)$ possa ser calculado conforme a Eq. 3.11. Neste passo, os pesos da rede não são alterados.
3. *Passo reverso* — o vetor de erros $\mathbf{e}^{(L)}(n)$ é utilizado para ajustar os pesos da camada de saída da rede. Em seguida, estes pesos da camada de saída são utilizados para ajustar os pesos da camada anterior, e este processo se repete, camada a camada, no sentido das entradas da rede, até que os pesos da primeira camada oculta da rede tenham sido ajustados. O método padrão para o cálculo dos ajustes dos pesos é a *regra delta generalizada*, proposta por Rumelhart, Hinton e Williams [89], e descrita como a seguir [88, 92].

Seja $w_{ji}^{(l)}$ o peso que conecta a saída de cada neurônio i da camada $l - 1$ à entrada do neurônio j da camada l . Então, para cada par de treinamento n , o valor de $w_{ji}^{(l)}(n)$ que

minimiza a energia média do erro total na saída da rede é dado por:

$$w_{ji}^{(l)}(n) = w_{ji}^{(l)}(n-1) + \Delta w_{ji}^{(l)}(n), \quad (3.14)$$

onde $w_{ji}^{(l)}(n)$ é o valor ajustado do peso que conecta o neurônio i da camada $l-1$ ao neurônio j da camada l para o par de treinamento n ; $w_{ji}^{(l)}(n-1)$ é o valor anterior deste peso, obtido da mesma forma para o par de treinamento $n-1$; e $\Delta w_{ji}^{(l)}(n)$ é o ajuste atual deste peso para o par de treinamento n , calculado pela *regra delta generalizada*, como:

$$\Delta w_{ji}^{(l)}(n) = \alpha \delta_j^{(l)}(n) y_i^{(l-1)}(n) + \eta \Delta w_{ji}^{(l)}(n-1), \quad (3.15)$$

onde α é a *taxa de aprendizado*, constante que determina a intensidade do ajustes dos pesos a cada iteração; $\delta_j^{(l)}(n)$ é o gradiente local do erro para o neurônio j da camada l , que determina a direção da busca no espaço de pesos (ver Eq. 3.16 a seguir); $y_i^{(l-1)}(n)$ é a saída do neurônio i da camada $l-1$, que se conecta ao neurônio j da camada l ($y_i^{(0)}(n)$ corresponde à entrada da rede $x_i(n)$); η é a *inércia*, ou *momento*, constante que tem efeito estabilizador na busca, aumentando os incrementos dos pesos quando o erro converge, e diminuindo os incrementos quando o erro diverge [88, pp. 170–171]; $\Delta w_{ji}^{(l)}(n-1)$ é o ajuste anterior do peso, obtido da mesma forma para o par de treinamento $n-1$.

O gradiente local do erro para o neurônio j , $\delta_j^{(l)}(n)$, é obtido de forma distinta para os neurônios da camada de saída e para os neurônios das demais camadas da rede, sendo calculado como:

$$\delta_j^{(l)}(n) = \begin{cases} \phi_j'(v_j^{(L)}(n)) e_j^{(L)}(n), & \text{na camada de saída;} \\ \phi_j'(v_j^{(l)}(n)) \sum_{k=1}^{M_{l+1}} \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n), & \text{nas demais camadas;} \end{cases} \quad (3.16)$$

onde para a camada de saída (camada L), o gradiente local do erro no neurônio j para o par de treinamento n , $\delta_j^{(L)}(n)$, é calculado pelo produto da primeira derivada da função de ativação do neurônio j aplicada ao seu campo local induzido, $\phi_j'(v_j^{(L)}(n))$, pelo erro na saída da rede, $e_j^{(L)}(n)$. Para as demais camadas da rede neural ($l < L$), o gradiente local do erro no neurônio j para o par de treinamento n , $\delta_j^{(l)}(n)$, é calculado pelo produto da primeira derivada da função de ativação do neurônio j aplicada ao seu campo local induzido, $\phi_j'(v_j^{(l)}(n))$, pelo somatório do produto dos gradientes dos M_{l+1} neurônios da camada $l+1$ pelos pesos $w_{kj}^{(l+1)}(n)$ que conectam o neurônio j da camada l ao neurônio k da camada $l+1$.

Conforme mostrado na Eq. 3.16, os pesos da camada de saída da rede são ajustados diretamente em função dos erros na saída da rede, enquanto os pesos das demais camadas da rede são ajustados a partir dos valores dos pesos e respectivos gradientes dos erros nos neurônios da sua camada posterior.

Os passos direto e reverso do algoritmo *back-propagation* são executados para todos N_T pares do conjunto de treinamento \mathbb{T} . Cada iteração em um dos N_T pares de treinamento recebe o nome de *época*, e a rede neural é sucessivamente treinada até que um critério de parada seja atingido.

Como não há prova de convergência para algoritmo *back-propagation*, também não há um critério de parada único e bem definido para o treinamento [88, pp. 173]. Tipicamente, são utilizados critérios de parada baseados na estabilização da convergência da energia média do erro total na saída da rede, ou da convergência dos gradientes locais, para o conjunto de treinamento; ou ainda, são utilizados critérios que estabelecem um número máximo de épocas para o treinamento.

O controle da convergência do treinamento, através da escolha de um critério de parada adequado, é um elemento crucial para o sucesso da aplicação das redes neurais em problemas práticos do mundo real. Uma rede neural treinada de forma insuficiente não consegue extrair uma representação adequada do mapeamento de entradas e saídas contido no conjunto de treinamento. Como resultado, a rede acaba por implementar um modelo muito elementar para o mapeamento desejado, ocorrendo um subajuste do modelo (*underfitting*). Neste caso, diz-se que o modelo apresenta uma alta *polarização* (*bias*). Por outro lado, se a rede neural for excessivamente treinada com os exemplos do conjunto de treinamento, o modelo aprendido será altamente especializado nos mínimos detalhes do mapeamento de entradas e saídas contido no conjunto de treinamento (inclusive no ruído), ocorrendo um sobreajuste do modelo (*overfitting*). Neste outro caso, diz-se que o modelo apresenta uma alta *variância* e a rede neural apresentará uma baixa generalização, resultando num desempenho insuficiente para entradas distintas daquelas do conjunto de treinamento.

Um controle de treinamento adequado deve buscar um compromisso ótimo entre a polarização e a variância da rede (*bias-variance tradeoff*) para o problema de interesse [88]. Uma técnica bastante utilizada para atacar este problema é a *validação cruzada* (*cross-validation* — CV) [88, pp. 213], que consiste em separar uma parcela dos dados disponíveis para o treinamento da rede e formar um *conjunto de validação*, $\mathbb{V} = \{\langle \mathbf{x}, \mathbf{d} \rangle_1, \dots, \langle \mathbf{x}, \mathbf{d} \rangle_{N_V}\}$, com pares que não serão utilizados para ajustar os pesos da rede neural. Assim, após treinar a rede neural da maneira tradicional durante um determinado número de épocas, o conjunto de validação, \mathbb{V} , é apresentado para a rede (sem ajustar os seus pesos) e o seu erro total é calculado conforme as Eqs. 3.12 e 3.13. Em seguida, o treinamento é retomado com o conjunto

de treinamento, \mathbb{T} , para mais um determinado número de épocas, e assim sucessivamente. A evolução do erro total do conjunto de validação, \mathbb{V} , calculado desta forma, pode ser utilizada como critério de parada. Por exemplo, o método de *parada antecipada* (*early stopping*) [88, pp. 215] interrompe o treinamento quando o erro total do conjunto de validação começa a aumentar⁹.

Após ter sido treinada, a rede está pronta para ser testada com dados distintos daqueles utilizados nos conjuntos de treinamento e validação, com os quais é formado o *conjunto de teste*, ou *conjunto de predição*, $\mathbb{P} = \{\langle \mathbf{x}, \mathbf{d} \rangle_1, \dots, \langle \mathbf{x}, \mathbf{d} \rangle_{N_p}\}$. O desempenho da rede neural no conjunto \mathbb{P} pode então ser mensurado a partir do cálculo do seu erro total, conforme as Eqs. 3.12 e 3.13, e utilizado como parâmetro de avaliação quanto ao desempenho esperado desta rede neural na sua aplicação prática.

3.3.2 Rede Neural Auto-regressiva

A predição baseada em séries históricas para o retorno futuro de uma ação pode ser definida como o processo no qual se utiliza, no tempo t , os elementos da série histórica de retornos, $\mathbf{r}' = (r_{t-(N-1)}, \dots, r_{t-1}, r_t)$, para obter uma estimativa \hat{r}_{t+l} do retorno futuro r_{t+l} , onde $l \geq 1$. O valor de l afeta diretamente a escolha do método de predição utilizado. Para $l = 1$, ou seja, a predição para um período futuro, a escolha padrão são os métodos auto-regressivos [20, 21, 22, 93], enquanto que para $l > 1$, ou seja, a predição para múltiplos períodos futuros, técnicas como o *método das Diferenças Temporais* [83, 94] são utilizadas; para este último caso, também podem ser utilizadas l predições recursivas para um período futuro [20]. Este trabalho investiga somente o problema da predição para um período futuro.

Assim, a predição do retorno futuro de ações pode ser obtida através de um sistema de predição de séries históricas, \mathcal{S}_P , definido como:

$$\hat{r}_{t+1} = \mathcal{S}_P(\mathbf{r}'), \quad (3.17)$$

onde \hat{r}_{t+1} é uma estimativa para o valor de r_{t+1} , obtida no tempo t a partir das informações disponíveis na sua série histórica, \mathbf{r}' .

Uma característica intrínseca da maioria das séries históricas formadas a partir da observação de fenômenos do mundo real é a existência de dependência temporal entre seus valores. Isto é, seus valores não são independentes e apresentam função de autocorrelação com

⁹ Algumas variações da validação cruzada são apresentadas na literatura, como a *validação cruzada múltipla* (*K-fold cross-validation*), que utiliza múltiplos conjuntos de validação; e o método *deixe um de fora* (*leave-one-out cross-validation*), que utiliza um conjunto de validação com apenas um exemplo, sendo adequado para o caso da escassez de dados de treinamento [88, pp. 217].

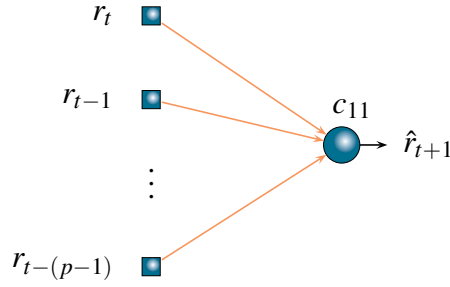


Figura 3.7: Rede Neural Auto-regressiva com uma camada.

coeficientes $\rho_k \neq 0, k \geq 1$ [20]. Ou seja, existe uma função que relaciona cada valor destas séries históricas com seus respectivos valores defasados no tempo. Desta forma, é esperado que o sistema de predição \mathcal{S}_p seja capaz de se beneficiar da existência desta característica estrutural das séries históricas para produzir predições úteis para os seus valores futuros.

O *modelo auto-regressivo* de ordem p , $AR(p)$ [20], é um modelo linear estacionário amplamente utilizado no arcabouço da Análise de Séries Temporais, e define um processo estocástico na forma:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \cdots + \phi_p \tilde{z}_{t-p} + a_t, \quad (3.18)$$

com

$$\tilde{z}_t = z_t - \mu, \quad (3.19)$$

onde \tilde{z}_t é o desvio do processo, no tempo t , em relação à sua média, μ ; $\phi_1, \phi_2, \dots, \phi_p$ são os pesos de cada valor defasado $\tilde{z}_{t-1}, \tilde{z}_{t-2}, \dots, \tilde{z}_{t-p}$ na regressão, respectivamente; e a_t é um *ruído branco*¹⁰. A ordem da regressão, p , é um parâmetro fixo do modelo, e os seus $p+2$ parâmetros variáveis, $\mu, \phi_1, \phi_2, \dots, \phi_p$, e σ_a^2 , são estimados a partir das séries históricas de interesse. Diversos métodos para a estimação desses parâmetros são descritos na literatura [20, 21, 22, 93], sendo tipicamente compostos por procedimentos assistidos e iterativos. Assim, a modelagem auto-regressiva produz um sistema de predição na forma:

$$\hat{r}_{t+1} = \mathcal{S}_p(r_{t-(p-1)}, \dots, r_{t-1}, r_t). \quad (3.20)$$

A estrutura do modelo $AR(p)$ da Eq. 3.18 assemelha-se à do modelo do neurônio artificial da Fig. 3.3, com entradas $x_1 = \tilde{z}_{t-1}, x_2 = \tilde{z}_{t-2}, \dots, x_m = \tilde{z}_{t-p}$ e função de ativação linear com $a = 1$ e $b = 0$, conforme as Eqs. 3.5, 3.6 e 3.7.

¹⁰O processo estocástico denominado *ruído branco* corresponde a uma série estacionária formada por variáveis aleatórias a_t não correlacionadas, isto é, com covariância $Cov(a_t, a_s) = 0$ para $t \neq s$, tipicamente com média $\mu_a = 0$ e variância σ_a^2 finita [20, 21].

White [23] propôs a implementação de um modelo $AR(p)$ para a predição de séries financeiras através de uma rede neural como a da Fig. 3.7, com p entradas $x_1 = r_t$, $x_2 = r_{t-1}$, \dots , $x_p = r_{t-(p-1)}$ e um neurônio linear com saída igual a \hat{r}_{t+1} , para o caso da predição do retorno futuro de ações. Este modelo de predição de séries históricas baseado em redes neurais é denominado *rede neural auto-regressiva* de ordem p , $RNAR(p)$, e implementa o seguinte modelo de regressão linear:

$$\hat{r}_{t+1} = w_{11,0} + w_{11,1} r_t + w_{11,2} r_{t-1} + \dots + w_{11,p} r_{t-(p-1)}, \quad (3.21)$$

onde \hat{r}_{t+1} é a predição para o valor de r_{t+1} ; $w_{11,1}, \dots, w_{11,p}$ são os pesos dos p valores defasados da série na regressão (os pesos que conectam o neurônio c_{11} às entradas $r_t, \dots, r_{t-(p-1)}$), respectivamente; e $w_{11,0}$ é uma constante de nível (*bias* do neurônio c_{11}). O modelo de rede neural apresentado na Fig. 3.7 também é denominado *Filtro Neural Focado* (*Focused Neuronal Filter*) [88, pp. 644].

Contudo, muitas séries históricas do mundo real, especialmente as séries financeiras, se afastam da linearidade e da estacionariedade, freqüentemente mudando seu regime estrutural ao longo do tempo e exibindo memória de longa duração [49, 95]. Portanto, estas séries não são adequadamente representadas por modelos lineares como os das Eqs. 3.18, 3.19 e 3.21, demandando modelos não-lineares mais complexos para esta finalidade.

A rede neural auto-regressiva é uma estrutura capaz de responder a esta demanda, oferecendo modelos auto-regressivos não-lineares virtualmente de qualquer ordem [88, pp. 208–209] e, desta forma, oferecendo vantagens frente aos modelos oferecidos pela Análise de Séries Temporais. Por exemplo, através da utilização de uma função de ativação não-linear, a rede neural auto-regressiva da Fig. 3.7 pode implementar um modelo auto-regressivo não-linear da forma:

$$\hat{r}_{t+1} = \varphi(w_{11,0} + w_{11,1} r_t + w_{11,2} r_{t-1} + \dots + w_{11,p} r_{t-(p-1)}), \quad (3.22)$$

onde \hat{r}_{t+1} é a predição para o valor de r_{t+1} ; $\varphi()$ é uma função não-linear diferenciável; $w_{11,1}, \dots, w_{11,p}$ são os pesos dos p valores defasados da série na regressão (os pesos que conectam o neurônio c_{11} às entradas $r_t, \dots, r_{t-(p-1)}$), respectivamente; e $w_{11,0}$ é uma constante de nível (*bias* do neurônio c_{11}).

Modelos não-lineares mais complexos podem ser obtidos através da inserção de camadas ocultas na rede neural auto-regressiva. Por exemplo, a Fig. 3.8 mostra uma rede neural auto-

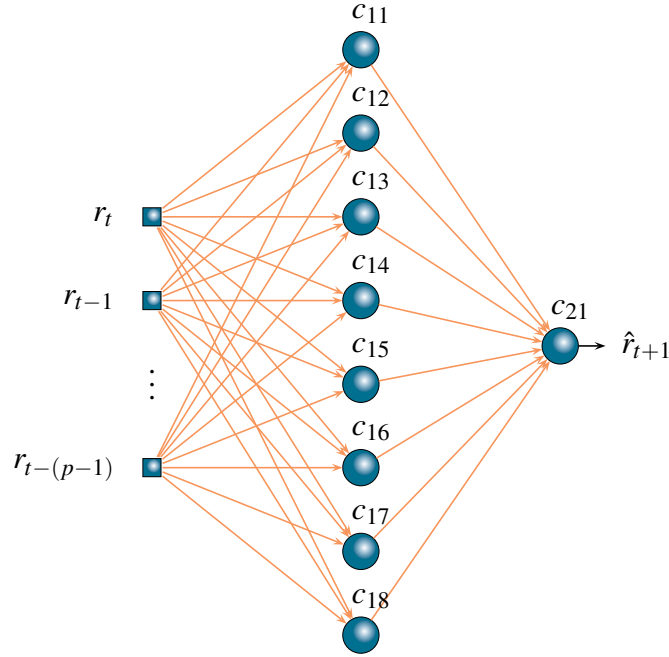


Figura 3.8: Rede Neural Auto-regressiva com duas camadas.

regressiva com uma camada oculta com oito neurônios, que implementa o seguinte modelo:

$$\begin{aligned} \hat{r}_{t+1} = & \varphi(w_{21,0} + w_{21,11} \varphi(w_{11,0} + w_{11,1} r_t + w_{11,2} r_{t-1} + \cdots + w_{11,p} r_{t-(p-1)}) \\ & + w_{21,12} \varphi(w_{12,0} + w_{12,1} r_t + w_{12,2} r_{t-1} + \cdots + w_{12,p} r_{t-(p-1)}) \\ & + w_{21,13} \varphi(w_{13,0} + w_{13,1} r_t + w_{13,2} r_{t-1} + \cdots + w_{13,p} r_{t-(p-1)}) \\ & + w_{21,14} \varphi(w_{14,0} + w_{14,1} r_t + w_{14,2} r_{t-1} + \cdots + w_{14,p} r_{t-(p-1)}) \\ & + w_{21,15} \varphi(w_{15,0} + w_{15,1} r_t + w_{15,2} r_{t-1} + \cdots + w_{15,p} r_{t-(p-1)}) \\ & + w_{21,16} \varphi(w_{16,0} + w_{16,1} r_t + w_{16,2} r_{t-1} + \cdots + w_{16,p} r_{t-(p-1)}) \\ & + w_{21,17} \varphi(w_{17,0} + w_{17,1} r_t + w_{17,2} r_{t-1} + \cdots + w_{17,p} r_{t-(p-1)}) \\ & + w_{21,18} \varphi(w_{18,0} + w_{18,1} r_t + w_{18,2} r_{t-1} + \cdots + w_{18,p} r_{t-(p-1)})), \end{aligned} \quad (3.23)$$

onde \hat{r}_{t+1} é a predição para o valor de r_{t+1} ; $\varphi()$ é uma função não-linear diferenciável; $w_{1X,1}, \dots, w_{1X,p}$, com $X = 1, \dots, 8$, são os pesos dos p valores defasados da série na regressão (os pesos que conectam o neurônio c_{1X} às entradas $r_t, \dots, r_{t-(p-1)}$), respectivamente; $w_{1X,0}$ são constantes parciais de nível (*bias* dos neurônios c_{1X}); $w_{21,11}, \dots, w_{21,18}$ são os pesos que conectam o neurônio de saída c_{21} aos neurônios c_{1X} da camada 1; e $w_{21,0}$ é o *bias* do neurônio c_{21} .

Desta forma, após ter sido adequadamente treinado com os pares entrada-saída do seu conjunto de treinamento, o preditor baseado na rede neural auto-regressiva implementa um

modelo de regressão não-linear múltipla para a série de retornos [88, pp. 635-660], com parâmetros estimados pelo algoritmo de treinamento utilizado. O modelo de rede neural apresentado na Fig. 3.8 também é denominado *Rede Neural Feedforward Focada com Atrasos Temporais* (*Focused Time Lagged Feedforward Network* — TLFN) [88, pp. 644].

3.4 O Problema da Seleção de Carteiras

A proposição do modelo média-variância por Harry Markowitz [7] em 1952 é um marco na *Moderna Teoria de Carteiras*. Com este trabalho, Markowitz foi laureado com o prêmio Nobel de Economia em 1990 “pelo trabalho pioneiro na teoria econômica das finanças” [96]. No encerramento da sua conferência de premiação, ele observa este aspecto histórico destacando [97]:

Finalizando, eu gostaria de adicionar um comentário a respeito da teoria das carteiras como parte da microeconomia da atitude sob incerteza. Nem sempre ela foi considerada assim. Por exemplo, quando defendi a minha tese como estudante do Departamento de Economia da Universidade de Chicago, o Professor Milton Friedman¹¹ argumentou que a teoria das carteiras não era Economia, e que eles não poderiam me conceder um título de Ph.D. em Economia com uma tese que não era em Economia. Eu assumi que ele estava sendo somente parcialmente sério, visto que eles me concederam o título sem longo debate. Sobre os méritos dos seus argumentos, neste ponto eu estou bastante disposto a admitir: na época em que eu defendi minha tese, a teoria das carteiras não era parte da Economia. Mas hoje ela é.

Apesar de já haver esforços para a solução simultânea dos problemas da escolha sob incerteza e da diversificação naquela época [98], era a primeira vez que um arcabouço quantitativo que maximiza uma função utilidade era proposto neste sentido [9].

3.4.1 O Modelo Média-Variância para Seleção de Carteiras

Uma carteira é um conjunto de M ações, cada qual com um respectivo peso, ou *participação*, associado. Cada participação X_i , $i = 1, \dots, M$, com $0 \leq X_i \leq 1$ e $\sum X_i = 1$, representa a fração do valor da carteira investido na ação i . O retorno realizado da carteira no tempo t ,

¹¹Prêmio Nobel de Economia em 1976.

r_{pt} , por sua vez, é dado pela combinação linear das participações e dos retornos individuais das suas ações componentes, ou seja:

$$r_{pt} = \sum_{i=1}^M X_i r_{it}, \quad (3.24)$$

onde M é o número de ações na carteira, X_i é a participação da ação i na carteira e r_{it} é o retorno realizado desta ação no tempo t .

A medida do retorno esperado de uma ação, segundo a proposição original de Markowitz, é a média aritmética da série de retornos desta ação, definida como:

$$\bar{r} = \frac{1}{N} \sum_{t=1}^N r_t, \quad (3.25)$$

onde \bar{r} é o retorno esperado da ação para o tempo $t = N + 1$, N é o tamanho da série histórica e r_t é o retorno realizado da ação no tempo t .

De forma similar, o retorno esperado da carteira para o tempo $t = N + 1$, \bar{r}_p , é definido pela combinação linear das participações e dos retornos esperados das suas ações:

$$\bar{r}_p = \sum_{i=1}^M X_i \bar{r}_i, \quad (3.26)$$

onde M é o número de ações na carteira, X_i é a participação da ação i na carteira e \bar{r}_i o retorno esperado desta ação.

O risco da ação no modelo média-variância é definido como a variância da sua série de retornos:

$$v = \sigma^2 = \frac{1}{N-1} \sum_{t=1}^N (r_t - \bar{r})^2, \quad (3.27)$$

onde v é o risco associado ao retorno esperado da ação para o tempo $t = N + 1$, N é o tamanho da série histórica, r_t é o retorno realizado da ação no tempo t e \bar{r} é o seu retorno esperado para o tempo $t = N + 1$. A medida do risco da ação proposta por Markowitz busca refletir a incerteza da realização do retorno esperado (Eq. 3.25) — a ocorrência do retorno esperado (valor médio) seria menos provável em séries com maior dispersão (maior risco) do que em séries com menor dispersão (menor risco).

O risco da carteira proposto por Markowitz é a variância da combinação linear das participações das ações na carteira e das suas séries de retornos, definida em termos da covariância

entre cada par de ações como:

$$V = \sigma_p^2 = \sum_{i=1}^M \sum_{j=1}^M X_i X_j \gamma_{ij}, \quad (3.28)$$

onde V é o risco total da carteira, que é igual à variância σ_p^2 da combinação linear das participações e das séries de retornos, M é o número de ações na carteira, X_i e X_j são as participações individuais das ações i e j na carteira, respectivamente, e γ_{ij} é a covariância dos retornos do par da ações i e j , que é definida como:

$$\gamma_{ij} = \frac{1}{N-1} \sum_{t=1}^N (r_{it} - \bar{r}_i)(r_{jt} - \bar{r}_j), \quad (3.29)$$

onde N é o tamanho das séries históricas, r_{it} e r_{jt} são os retornos realizados das ações i e j no tempo t , e \bar{r}_i e \bar{r}_j são os retornos esperados das ações i e j , respectivamente, para o tempo $t = N + 1$.

A Eq. 3.28 pode ser reescrita como:

$$\begin{aligned} V = \sigma_p^2 &= \sum_{i=1}^M X_i^2 \sigma_i^2 + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \gamma_{ij} \\ &= \sum_{i=1}^M X_i^2 v_i + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \gamma_{ij}, \end{aligned} \quad (3.30)$$

onde V é o risco total da carteira, M é o número de ações na carteira, X_i e X_j são as participações individuais das ações i e j na carteira, respectivamente, v_i (σ_i^2) é o risco individual (variância da série de retornos) da ação i e γ_{ij} é a covariância das séries de retornos do par de ações i e j .

A Eq. 3.30 evidencia a composição do risco total da carteira, que é composto pela contribuição dos riscos individuais de cada uma das M ações, representada pelo primeiro somatório, e pela contribuição dos riscos combinados de cada par de ações, representada pelo duplo somatório. A covariância dos pares de ações é a medida do *risco interativo*, ou *risco conjunto*, entre duas ações, e o seu efeito no risco da carteira V é crucial para a obtenção de carteiras com risco inferior à soma ponderada dos riscos individuais das suas ações (primeiro somatório da Eq. 3.30) — a contribuição fundamental do modelo média-variância de Markowitz foi a formalização da possibilidade de minimização do risco da carteira por meio da minimização deste risco interativo [7].

O modelo média-variância formula que o risco da carteira, V , para um retorno de carteira desejado, R_d , pode ser minimizado por meio da resolução do problema de minimização quadrático definido como:

Minimize

$$V = \sum_{i=1}^M X_i^2 v_i + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \gamma_{ij}, \quad (3.31)$$

Sujeito a

$$\sum_{i=1}^M X_i \bar{r}_i = R_d, \quad (3.32)$$

$$\sum_{i=1}^M X_i = 1, \quad (3.33)$$

$$X_i \geq 0, \quad i = 1, \dots, M. \quad (3.34)$$

A Eq. 3.31 é a função objetivo a ser minimizada, o risco da carteira média-variância, V , a Eq. 3.32 é a restrição que garante o retorno de carteira desejado, R_d , a Eq. 3.33 é a restrição que garante a integralidade da carteira com a alocação total dos recursos disponíveis e a Eq. 3.34 restringe o modelo para participações positivas apenas.

3.4.2 A Fronteira Eficiente

Um conjunto de carteiras de risco mínimo pode ser obtido por meio da resolução do problema de minimização definido no modelo média-variância (Eqs. 3.31, 3.32, 3.33 e 3.34) para vários valores de retorno de carteira desejado, R_d . Estas carteiras são denominadas *carteiras eficientes*, e a estratégia de diversificação de investimentos baseada neste arcabouço é denominada *diversificação eficiente*. Cada carteira eficiente exibe a propriedade particular de não haver, segundo as definições do modelo, outra carteira no conjunto de oportunidades que apresente um menor risco para um mesmo retorno de carteira, ou, no problema dual, um maior retorno para um mesmo nível de risco.

O *locus* bidimensional no espaço risco-retorno onde localizam-se todas as carteiras eficientes possíveis recebe o nome de *fronteira eficiente*. A Fig. 3.9 mostra um exemplo de fronteira eficiente obtida a partir de séries de retornos do mercado de ações brasileiro, onde temos os riscos das carteiras (variâncias) nas abscissas e seus retornos esperados nas ordenadas. Também é usual que a fronteira eficiente seja representada com os eixos invertidos em relação aos eixos mostrados na Fig. 3.9, ou seja, os retornos esperados nas abscissas e os riscos nas ordenadas; ou, ainda, com a utilização do risco como desvio-padrão ao invés

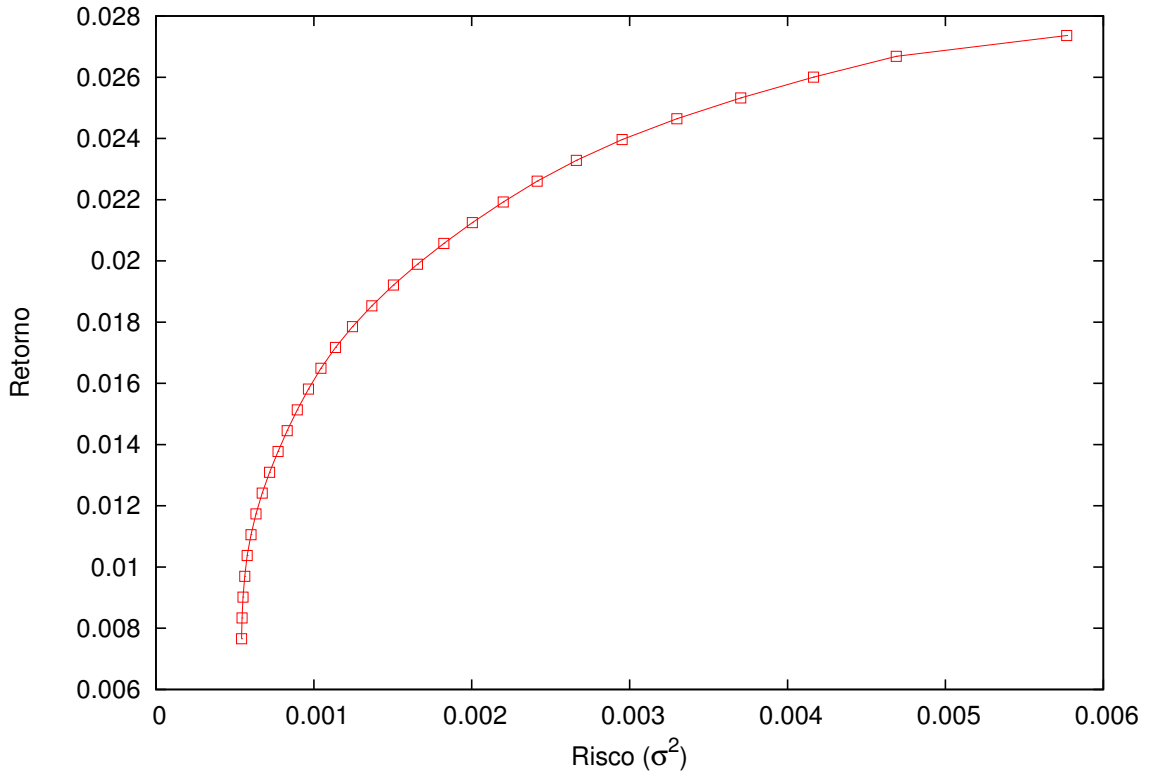


Figura 3.9: Exemplo de fronteira eficiente do modelo média-variância construída com séries de retornos semanais do mercado de ações brasileiro, a partir de 30 carteiras eficientes obtidas com a variação do retorno de carteira desejado, R_d , na região de factibilidade do modelo. Nas abscissas temos os riscos (variâncias) das carteiras e nas ordenadas seus retornos esperados. O formato da fronteira eficiente é determinado unicamente pelos retornos esperados, variâncias e covariâncias dos retornos do seu conjunto de ações.

da variância. Este último caso é adequado para a demonstração de algumas propriedades geométricas da fronteira eficiente e do modelo média-variância [4].

Cada conjunto de M ações determina uma fronteira eficiente, a qual depende dos retornos esperados (Eq. 3.25), das participações ótimas (Eq. 3.31), das variâncias das séries de retornos de cada ação (Eq. 3.27) e das covariâncias de cada par de séries de retornos das ações (Eq. 3.29). O conjunto de variâncias e covariâncias utilizado pelo modelo é computacionalmente representado como uma *matriz de covariâncias*, ou *matriz de variâncias e covariâncias*, que é uma matriz quadrada com dimensão $M \times M$ representada como:

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1M} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2M} \\ \vdots & \vdots & & \vdots \\ \gamma_{M1} & \gamma_{M2} & \cdots & \gamma_{MM} \end{bmatrix}, \quad (3.35)$$

ou

$$\mathbf{\Gamma} = \begin{bmatrix} \sigma_1^2 & \gamma_{12} & \cdots & \gamma_{1M} \\ \gamma_{21} & \sigma_2^2 & \cdots & \gamma_{2M} \\ \vdots & \vdots & & \vdots \\ \gamma_{M1} & \gamma_{M2} & \cdots & \sigma_M^2 \end{bmatrix}. \quad (3.36)$$

Os termos γ_{ij} são as covariâncias das séries de retornos das ações i e j , sendo que os termos γ_{ii} são iguais às variâncias σ_i^2 . Como a matriz de covariâncias é simétrica, o seu número de elementos distintos é dado por $(M^2 + M)/2$, o que faz com que o custo computacional e de armazenamento do modelo seja quadrático em relação ao tamanho do universo de ações, M . A matriz de covariâncias precisa ser *positiva semidefinida*¹²[8], caso contrário o modelo poderá produzir carteiras com risco negativo¹³, o que é implausível, uma vez que a variância é o segundo momento central e este é obrigatoriamente positivo ou igual a zero [99].

3.4.3 Diversificação Eficiente e Número de Ações

O efeito da diversificação na minimização do risco da carteira pode ser demonstrado para o caso particular de uma carteira uniforme [3]. Neste caso, fixando-se $X_i = X_j = 1/M$ na Eq. 3.30, rearranjando e fatorando os termos, obtemos:

$$\begin{aligned} V = \sigma_p^2 &= \sum_{i=1}^M \left(\frac{1}{M} \right)^2 \sigma_i^2 + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{M} \frac{1}{M} \gamma_{ij} \\ &= \frac{1}{M} \left(\frac{1}{M} \sum_{i=1}^M \sigma_i^2 \right) + \frac{M(M-1)}{M^2} \left(\frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \gamma_{ij} \right) \\ &= \frac{1}{M} \bar{\sigma}^2 + \left(1 - \frac{1}{M} \right) \bar{\gamma}_{ij}, \end{aligned} \quad (3.37)$$

onde $\bar{\sigma}^2$ e $\bar{\gamma}_{ij}$ são a variância média e a covariância média dos retornos, respectivamente. Quando o número M de ações na carteira aumenta, a Eq. 3.37 é aproximada por:

$$V \simeq \bar{\gamma}_{ij}, \quad (3.38)$$

onde a contribuição do risco individual de cada ação no risco da carteira é eliminada e, desta forma, o risco da carteira é dado apenas pela covariância média dos retornos das suas ações.

¹²Uma matriz quadrada real \mathbf{H} é positiva semidefinida se $\mathbf{bHb}^T \geq 0$ para todos os vetores \mathbf{b} .

¹³Como a variância da carteira é dada por $\sigma_p^2 = \mathbf{X}\mathbf{\Gamma}\mathbf{X}^T$, a condição de que $\mathbf{\Gamma}$ seja positiva semidefinida garante que a variância da carteira não seja negativa.

Capítulo 4

Seleção de Carteiras com Predição de Retornos

Desde a proposição original do modelo média-variância, aspectos como a sua factibilidade computacional, simplificações e extensões do modelo e o desenvolvimento de novas medidas de risco têm sido foco de intensa pesquisa. Estes esforços contribuíram fortemente para estabelecer a teoria de carteiras, sedimentando a sua prática na indústria e produzindo resultados bastante significativos [50, 52, 53, 54, 56, 73, 100, 101, 102]. Contudo, o valor do retorno médio como estimativa de retorno futuro tende a ser verificado somente no longo prazo, e, portanto, torna-se uma estimativa inadequada às estratégias de investimento de curto prazo, como a gerência ativa de carteiras por exemplo. A utilização de métodos de predição superiores na obtenção das estimativas de retorno futuro, aliados a medidas de risco adequadas, pode fornecer novos arcabouços preditivos para seleção de carteiras, mais adequados e alinhados às demandas das estratégias ativas de curto prazo. Neste capítulo, apresentaremos uma adaptação do modelo média-variância para utilização de retornos preditos. Este novo modelo de seleção de carteiras explora aspectos preditivos das séries de retornos ao mesmo tempo em que mantém aderência ao arcabouço Normal multivariado. Neste modelo, a variância dos erros de predição produzidos por preditores baseados em *redes neurais auto-regressivas com referência móvel* (RNAR-RM), apresentadas a seguir, é utilizada como medida de risco individual das ações.

4.1 Rede Neural Auto-regressiva com Referência Móvel

Neste trabalho, nós propomos um novo preditor neural auto-regressivo inspirado no comportamento intuitivo que tipicamente empregamos para analisar visualmente o gráfico

de uma série histórica com a finalidade de extrapolar seu próximo valor. Tendemos a concentrar nossa atenção visual nos últimos pontos do gráfico, criando uma janela virtual que delimita uma região capaz de oferecer uma imagem com quantidade de informação visual suficiente para a extrapolação do próximo ponto do gráfico. Neste método de predição, tipicamente utilizamos algum ponto dentro da região analisada como uma referência a partir da qual o valor do próximo ponto é estimado. Nós imitamos esse método intuitivo no nosso método auto-regressivo de predição por meio da subtração do valor de um dos pontos, usado como referência, dos pontos apresentados como entrada da rede neural do preditor.

Nós denominamos este preditor de *Rede Neural Auto-regressiva com Referência Móvel* — RNAR-RM (p, k) [33, 34] —, onde p é a ordem da regressão e k é o retardo do ponto de referência na linha do tempo, contado a partir de p .

Desta forma, o modelo RNAR-RM (p, k) implementa o seguinte sistema de predição:

$$\widehat{r_{t+1} - z} = S_p(r_{t-(p-1)} - z, \dots, r_{t-1} - z, r_t - z), \quad (4.1)$$

onde $\widehat{r_{t+1} - z}$ é uma estimativa para o valor de $r_{t+1} - z$, obtida no tempo t a partir das informações disponíveis na série histórica $\mathbf{r}' = (r_{t-(N-1)}, \dots, r_{t-1}, r_t)$ de tamanho N ($N \geq p + k$), p é a ordem da regressão (Seção 3.3.2) e z é a *referência móvel* dada por:

$$z = r_{t-(p-1)-k}. \quad (4.2)$$

Após o treinamento, a predição \hat{r}_{t+1} para o valor de r_{t+1} da série histórica é obtida partir da predição de $\widehat{r_{t+1} - z}$ utilizando:

$$\hat{r}_{t+1} = \widehat{r_{t+1} - z} + z. \quad (4.3)$$

A Fig. 4.1 mostra um exemplo de aplicação do método auto-regressivo com referência móvel com $p = 4$ e $k = 1$ para uma série de preços P' . No tempo t , a nova referência para os valores é determinada pelo valor de P_{t-4} , de forma que os valores de entrada da rede tornam-se $P_{t-3} - P_{t-4}$, $P_{t-2} - P_{t-4}$, $P_{t-1} - P_{t-4}$ e $P_t - P_{t-4}$. A Fig. 4.2 (pág. 65) mostra um exemplo de uma RNAR-RM (4,1) para a predição de séries de retornos.

Com a RNAR-RM, padrões de janelas de referência móvel repetidos podem ser codificados nos pesos da rede neural apenas uma única vez, reduzindo o número de padrões que a rede precisa aprender. Também, os valores de entrada da rede podem ser menores que os valores das séries (pois são diferenças entre valores próximos), o que reduz a possibilidade de saturação das saídas dos neurônios, aumentando a faixa dinâmica da rede neural e podendo dispensar algumas necessidades de pré-processamento, tais como a normalização e a

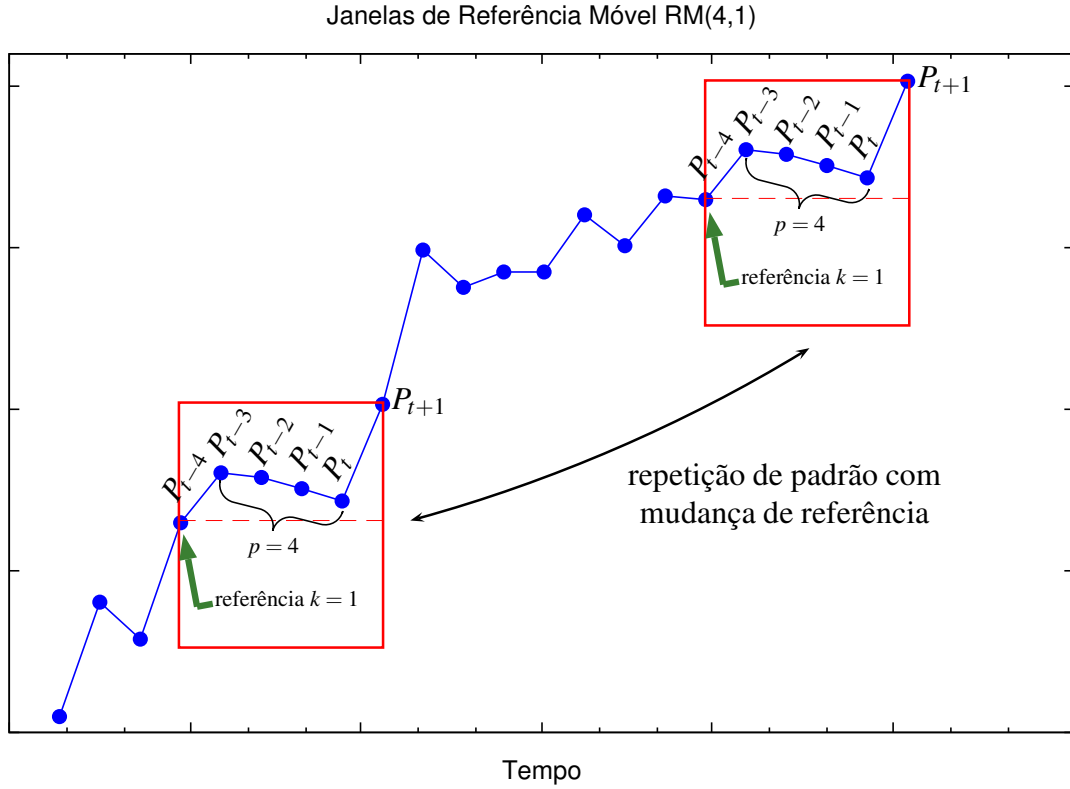


Figura 4.1: Aplicação do método auto-regressivo com referência móvel no preditor RNAR-RM (4,1). No tempo t , a referência é dada pelo valor de P_{t-4} , e as entradas da rede são $P_{t-3} - P_{t-4}$, $P_{t-2} - P_{t-4}$, $P_{t-1} - P_{t-4}$ e $P_t - P_{t-4}$. Os padrões das janelas de referência móvel repetidos podem ser codificados nos pesos da rede neural apenas uma única vez.

eliminação de tendências. Como resultado, a chance de se obter valores de pesos menores no término do treinamento aumenta, o que contribui para a regularização da rede neural [88], melhorando a sua capacidade de generalização e aumentando a sua capacidade de representação, o que permite uma melhor modelagem de séries históricas mais complexas.

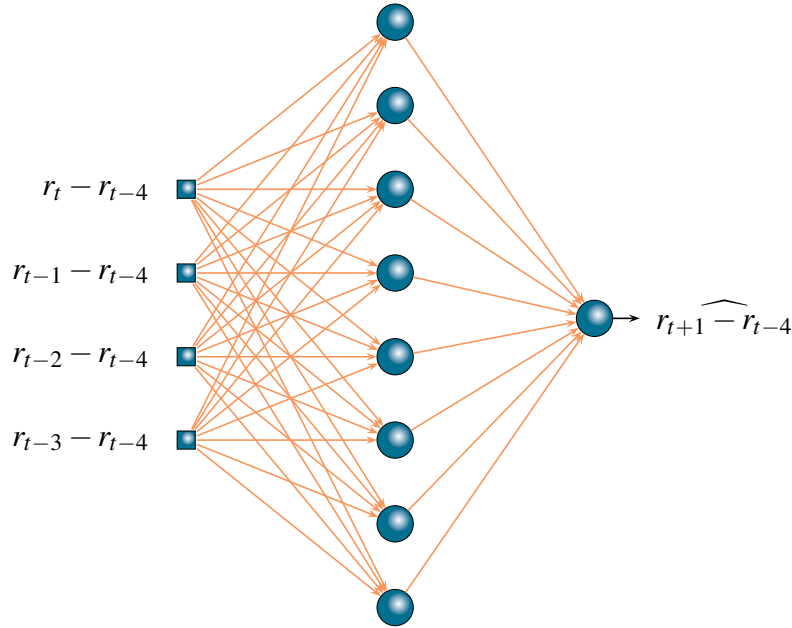


Figura 4.2: Exemplo de preditor RNAR-RM (4,1).

4.2 Retorno Predito e Risco das Ações

O *modelo de seleção de carteiras baseado em erros de predição*, que é proposto neste trabalho, utiliza o retorno predito \hat{r}_{t+1} , obtido no tempo t , como retorno esperado de cada ação no tempo $t + 1$. A medida de risco proposta no modelo, por sua vez, é baseada nos erros de predição e será apresentada a seguir.

Seja a relação entre o retorno realizado e o retorno predito da ação no tempo t dada por:

$$r_t = \hat{r}_t + \varepsilon_t, \quad (4.4)$$

onde r_t é o retorno realizado da ação no tempo t , \hat{r}_t é o retorno predito da ação no tempo t , obtido anteriormente no tempo $t - 1$, e ε_t é o erro de predição observado no tempo t .

Rearranjando os termos, podemos definir o erro de predição como:

$$\varepsilon_t = r_t - \hat{r}_t. \quad (4.5)$$

Considerando um conjunto de n predições realizadas até o tempo t , a série histórica dos n erros de predição é, então, dada por:

$$\boldsymbol{\varepsilon}' = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n). \quad (4.6)$$

A medida de risco individual das ações proposta no modelo de seleção de carteiras baseado em erros de predição é a variância da série de erros de predição (ver Eq. 4.8), que reflete a incerteza da realização do retorno predito. Ou seja, quanto maiores os valores absolutos dos erros de predição, maior a variância da sua série e maior a diferença esperada entre o retorno realizado e o retorno predito (maior a variância, maior o risco). Por outro lado, quanto menores os valores absolutos dos erros de predição, menor a variância da sua série e menor a diferença esperada entre o retorno realizado e o retorno predito (menor a variância, menor o risco). Para um preditor não tendencioso, a série dos erros de predição (Eq. 4.6) deve ser independente e identicamente distribuída (i.i.d.)¹, com média e variância dadas por:

$$\begin{aligned}\bar{\varepsilon} = \mu_{\varepsilon} &= \frac{1}{n} \sum_{t=1}^n \varepsilon_t \\ &= 0,\end{aligned}\tag{4.7}$$

e

$$\begin{aligned}\hat{v} = \sigma_{\varepsilon}^2 &= \frac{1}{n-1} \sum_{t=1}^n (\varepsilon_t - \bar{\varepsilon})^2 \\ &= \frac{1}{n-1} \sum_{t=1}^n \varepsilon_t^2,\end{aligned}\tag{4.8}$$

onde ε_t é o erro de predição no tempo t , $\bar{\varepsilon}$ é o erro de predição médio, \hat{v} é o risco da ação, que é igual à variância dos erros de predição, σ_{ε}^2 , e n é o número de observações da série de erros de predição.

4.3 Retorno Predito e Risco da Carteira

De forma similar aos modelos clássicos de seleção de carteiras, o retorno predito da carteira no modelo de seleção de carteiras baseado em erros de predição, \hat{r}_p , é a combinação linear das participações das ações na carteira e seus respectivos retornos preditos, dada por:

$$\hat{r}_p = \sum_{i=1}^M X_i \hat{r}_i,\tag{4.9}$$

onde M é o número de ações na carteira, X_i é a participação da ação i na carteira e \hat{r}_i é o retorno predito desta ação.

¹Uma série \mathbf{s} é independente e identicamente distribuída (i.i.d.) quando for estacionária e formada por valores não correlacionados. Ou seja, quando mantiver o mesmo valor esperado $E(\mathbf{s}) = \mu$ e variância $Var(\mathbf{s}) = \sigma^2$ para todos os tempos t , e a sua função de auto-covariância $Cov(s_t, s_{t+\tau})$ for igual a σ^2 para $\tau = 0$ e nula para qualquer $\tau \neq 0$ [22].

Assumindo que os erros de predição da Eq. 4.6 seguem uma distribuição Normal, podemos utilizar a modelagem Normal multivariada para o risco da carteira no modelo de seleção de carteiras baseado em erros de predição. Desta forma, o risco da carteira preditiva é dado pela variância da distribuição Normal conjunta da combinação linear das participações das ações na carteira e das suas séries de erros de predição, definida em termos das variâncias e covariâncias das séries de erros de predição como:

$$\hat{V} = \hat{\sigma}_p^2 = \sum_{i=1}^M X_i^2 \sigma_{\varepsilon_i}^2 + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \gamma_{\varepsilon_{ij}}, \quad (4.10)$$

onde \hat{V} é o risco total da carteira, que é igual à variância da combinação linear das participações das ações na carteira e das séries de erros de predição das ações, $\hat{\sigma}_p^2$, M é o número de ações na carteira, X_i e X_j são as participações individuais das ações i e j na carteira, respectivamente, e $\gamma_{\varepsilon_{ij}}$ é a covariância dos erros de predição das ações i e j , dada por:

$$\hat{\gamma}_{ij} = \gamma_{\varepsilon_{ij}} = \frac{1}{n-1} \sum_{t=1}^n \varepsilon_{it} \varepsilon_{jt}, \quad (4.11)$$

onde n é o tamanho das séries históricas de erros de predição, e ε_{it} e ε_{jt} são os erros de predição das ações i e j , respectivamente, observados no tempo t .

Na Eq. 4.10, o primeiro somatório representa a contribuição dos riscos associados às predições individuais de cada uma das M ações e o segundo grupo de somatórios representa a contribuição dos riscos combinados associados aos erros de predição dos pares de ações, ou *risco interativo dos retornos preditos*.

4.4 Modelo de Otimização

Com suas medidas de retorno e risco definidas, podemos formular o modelo de seleção de carteiras baseado em erros de predição como:

Minimize

$$\hat{V} = \sum_{i=1}^M X_i^2 \hat{v}_i + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \hat{\gamma}_{ij}, \quad (4.12)$$

Sujeito a

$$\sum_{i=1}^M X_i \hat{r}_i = R_d, \quad (4.13)$$

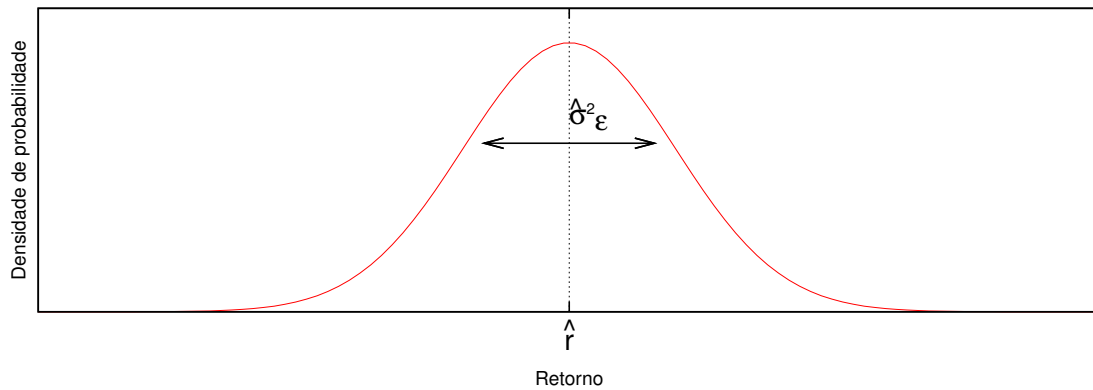
$$\sum_{i=1}^M X_i = 1, \quad (4.14)$$

$$X_i \geq 0, \quad i = 1, \dots, M. \quad (4.15)$$

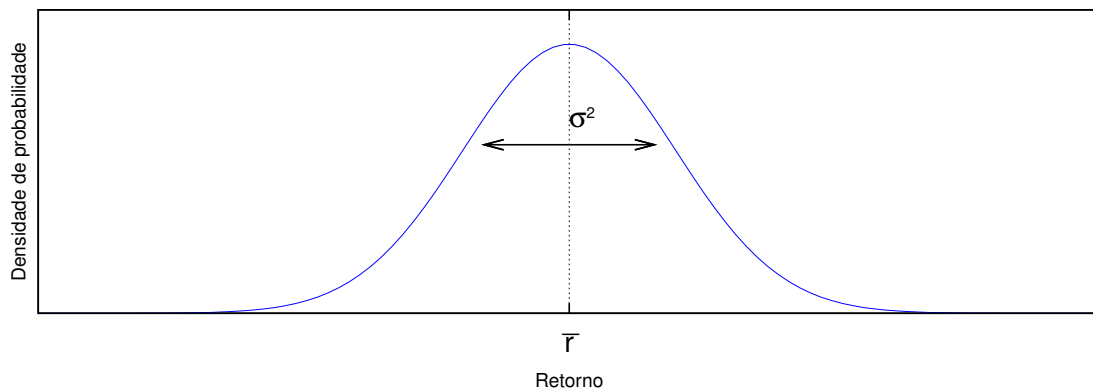
A Eq. 4.12 é a função objetivo a ser minimizada, o risco da carteira baseada em erros de predição, \hat{V} , a Eq. 4.13 é a restrição que garante o retorno de carteira desejado, R_d , a Eq. 4.14 é a restrição que garante a integralidade da carteira com a alocação total dos recursos disponíveis e a Eq. 4.15 restringe o modelo para participações positivas apenas.

Apesar de também empregar a modelagem Normal multivariada, o modelo de seleção de carteiras baseado em erros de predição difere fundamentalmente do modelo média-variância porque: (i) no modelo de seleção de carteiras baseado em erros de predição o retorno esperado de cada ação é o seu retorno predito, e não a média de sua série de retornos como no modelo média-variância; (ii) no modelo de seleção de carteiras baseado em erros de predição o retorno de uma ação é modelado como uma variável aleatória Normal com média zero (o erro de predição, ver Eqs. 4.5 e 4.7) somada ao valor de saída de um preditor (o retorno predito) — ver Fig. 4.3a —, enquanto que no modelo média-variância o retorno de uma ação é modelado como uma variável aleatória Normal — ver Fig. 4.3b —; e (iii) no modelo de seleção de carteiras baseado em erros de predição o risco individual de cada ação e o risco interativo entre os pares de ações são obtidos a partir da variância e covariância das séries de erros de predição (Eqs. 4.8 e 4.11), e não pela variância e covariância das séries de retornos como no modelo média-variância.

Predizer corretamente as séries de retornos do mercado de ações é reconhecidamente uma tarefa difícil [42]. Individualmente, os preditores de retornos de ações não alcançam um desempenho excepcional, tipicamente exibindo erros de predição significativos [33, 34]. Contudo, esses preditores podem ser sistematicamente combinados de forma a explorar a



(a) Distribuição dos retornos das ações para o modelo de seleção de carteiras baseado em erros de predição



(b) Distribuição dos retornos das ações para o modelo média-variância

Figura 4.3: Modelo distribucional dos retornos das ações para (a) o modelo de seleção de carteiras baseado em erros de predição e para (b) o modelo média-variância. No modelo de seleção de carteiras baseado em erros de predição o retorno de uma ação é modelado como uma variável aleatória Normal com média zero (os erros de predição) somada ao valor de saída do seu preditor (o retorno predito), enquanto que o modelo média-variância utiliza uma variável aleatória Normal com a média e variância da série de retornos da ação.

complementaridade dos seus desempenhos individuais a partir dos perfis dos seus erros de predição, produzindo boas predições (isto é, predições com baixa variância nos erros) para o seu agregado — o retorno da carteira baseada em erros de predição. O modelo de seleção de carteiras baseado em erros de predição explora esta possibilidade. Ele é baseado nas premissas das Eqs. 4.7, 4.8 e 4.11, e na Normalidade dos erros de predição, a qual foi verificada experimentalmente com os resultados que serão apresentados na Seção 6.2.1.

4.5 O Modelo Média-Variância como um Caso Particular do Modelo de Seleção de Carteiras Baseado em Erros de Predição

A proposição e o desenvolvimento do modelo de seleção de carteiras baseado em erros de predição foram fortemente orientados no sentido da utilização das séries de erros de predição dos preditores auto-regressivos no fornecimento de uma medida de risco aderente ao arcabouço Normal multivariado. Uma motivação particular da busca desta aderência ao modelo Normal multivariado foi a utilização do sedimentado arcabouço da programação quadrática, que dispõe de algoritmos e implementações computacionais bastante eficientes.

Apesar das diferenças entre o modelo de seleção de carteiras baseado em erros de predição e o modelo média-variância, que foram destacadas na Seção 4.4, nós verificamos analiticamente que: utilizando o retorno médio como o retorno predito do modelo de seleção de carteiras baseado em erros de predição, este se equipara ao modelo média-variância e, portanto, o modelo média-variância pode visto como um caso particular do modelo de seleção de carteiras baseado em erros de predição, conforme será demonstrado a seguir.

Seja $\hat{r}_t = \bar{r}$, então o erro de predição, ε_t (ver Eq. 4.5), pode ser aqui calculado como:

$$\varepsilon_t = r_t - \bar{r}. \quad (4.16)$$

Para $n \sim N$, segue que $\bar{\varepsilon} = 0$ (ver Eq. 4.7), pois:

$$\begin{aligned} \bar{\varepsilon} &= \frac{1}{n} \sum_{t=1}^n r_t - \bar{r} \\ &= \frac{1}{n} \sum_{t=1}^n r_t - \frac{1}{n} \sum_{t=1}^n \bar{r} \\ &= 0. \end{aligned} \quad (4.17)$$

Também, temos que $\sigma_{\varepsilon}^2 = \sigma^2$, pois:

$$\begin{aligned} \hat{v} = \sigma_{\varepsilon}^2 &= \frac{1}{n-1} \sum_{t=1}^n \varepsilon_t^2 \\ &= \frac{1}{n-1} \sum_{t=1}^n (r_t - \bar{r})^2 \\ &= \sigma^2. \end{aligned} \quad (4.18)$$

Finalizando, temos que $\hat{\gamma}_{ij} = \gamma_{ij}$, pois:

$$\begin{aligned}\hat{\gamma}_{ij} &= \gamma_{\varepsilon_{ij}} = \frac{1}{n-1} \sum_{t=1}^n \varepsilon_{it} \varepsilon_{jt} \\ &= \frac{1}{n-1} \sum_{t=1}^n (r_{it} - \bar{r}_i)(r_{jt} - \bar{r}_j) \\ &= \gamma_{ij}.\end{aligned}\tag{4.19}$$

Em todas as equações acima, n é o número de observações das séries históricas, ε_t é o erro de predição no tempo t , aqui definido conforme a Eq. 4.16, r_t é o retorno da ação no tempo t e \bar{r} é o retorno médio da ação.

Desta forma, uma vez que neste caso particular do modelo de seleção de carteiras baseado em erros de predição com $\hat{r} = \bar{r}$ as hipóteses das Eqs. 4.7, 4.8 e 4.11 são satisfeitas pelas Eqs. 4.17, 4.18 e 4.19, respectivamente, e as variâncias e covariâncias dos erros de predição são igualadas às variâncias e covariâncias dos retornos ($\hat{v} = \sigma^2$ e $\hat{\gamma}_{ij} = \gamma_{ij}$) como no modelo média-variância, os dois modelos se equivalem. Portanto, podemos tratar o modelo de seleção de carteiras baseado em erros de predição como “um novo olhar para um velho problema”, que engloba o modelo média-variância como um caso particular.

Capítulo 5

Métodos

Um grande número de experimentos com dados artificiais e reais foi conduzido para avaliar o modelo de seleção de carteiras baseado em erros de predição apresentado no Capítulo 4. Nossos experimentos foram planejados para comparar o desempenho do modelo de seleção de carteiras baseado em erros de predição com o do modelo média-variância e com o do índice de mercado IBOVESPA da Bolsa de Valores de São Paulo (BM&FBOVESPA).

Índices de mercado são carteiras de ações específicas selecionadas pelas bolsas de valores que têm por objetivo representar o mercado com um todo, tornando-se um referencial de desempenho para aqueles que participam deste mercado¹. O método de seleção das participações das ações no índice tipicamente leva em consideração diversos aspectos do desempenho das ações nos pregões passados (a metodologia de seleção do IBOVESPA é descrita em [103]). Devido ao seu caráter representativo e à ampla aceitação da dificuldade de superar seu desempenho em termos de retorno e risco [104, 105], os índices de mercado são freqüentemente utilizados como figura de mérito na avaliação de desempenho das estratégias de seleção de carteiras e gerência de fundos de investimento. Mais especificamente, os índices de mercado são referenciais importantes na avaliação da habilidade dos métodos de seleção e gerência de carteiras de produzir carteiras que antecipem os movimentos do mercado, denominada de *timing* do mercado [3, pp. 651–654]. Portanto, a comparação do desempenho de carteiras com o desempenho dos índices de mercado é um importante parâmetro de avaliação de modelos de seleção de carteiras.

Nos experimentos que serão apresentados mais adiante no Capítulo 6, nós utilizamos

¹Outras instituições, como instituições financeiras e de pesquisa e análise de mercado, também mantêm índices importantes, tais como o índice FGV 100 da Fundação Getúlio Vargas, o índice S&P 500 da agência de classificação de risco americana *Standard and Poor's* e o índice de índices CIJ mantido pela consultoria americana *Ibbotson Associates*. Visões segmentadas do mercado são obtidas através dos *índices setoriais*, que são concebidos de forma similar aos índices de mercado, porém, considerando apenas grupos de ações pertencentes a determinados setores da economia.

os métodos descritos neste capítulo para selecionar carteiras eficientes e avaliar seus desempenhos. As carteiras foram selecionadas em cada tempo t utilizando apenas os dados históricos disponíveis até esse instante, portanto sem conhecimento prévio de dados nos tempos posteriores a t ; em seguida, nós simulamos o investimento nessas carteiras e contabilizamos seus retornos no tempo $t + 1$. Este procedimento foi repetido para todo o horizonte de investimento e, ao final deste, nós comparamos o desempenho histórico das carteiras baseadas em erros de predição com o dos métodos de referência utilizados neste trabalho — o modelo média-variância e o índice IBOVESPA. Este capítulo apresenta a metodologia utilizada na execução dos experimentos e as métricas utilizadas na avaliação dos seus resultados.

5.1 Estratégia de Investimento

O modelo de seleção de carteiras baseado em erros de predição e o modelo média-variância pertencem a uma classe de estratégias de investimento denominada *abordagens de período único*, ou *modelos estáticos*. Estes modelos otimizam carteiras para apenas um período de tempo, usando os dados históricos disponíveis e os retornos esperados para o próximo período de tempo [4, pp. 139]. A estratégia de investimento implementada com estes modelos, e também empregada neste trabalho, consiste em:

- No tempo t :
 1. Obter uma fronteira eficiente para os dados disponíveis;
 2. Selecionar a carteira eficiente desejada;
 3. Investir todo o capital disponível, W_t , na carteira eficiente selecionada, investindo em cada ação i uma parcela do capital proporcional à participação da ação na carteira, $X_i W_t$.
- No tempo $t + 1$:
 4. Vender toda a carteira adquirida no tempo t e reinvestir todo o capital obtido, W_{t+1} , em uma nova carteira segundo este procedimento.

Nos experimentos realizados, que serão apresentados no Capítulo 6, consideramos as mesmas premissas subjacentes tradicionais da literatura [2, 3, 4]:

- As ações são perfeitamente divisíveis, de forma que é possível comprar as ações de qualquer carteira selecionada nas proporções das suas participações.

- Não existem fricções (custos de transação, taxação, comissões, etc...) quando da troca de posição em qualquer ação da carteira.
- Existe perfeita liquidez em todas as ações envolvidas.
- É possível comprar qualquer ação no tempo t pelo seu preço de fechamento e vendê-la no tempo $t + 1$ também pelo seu preço de fechamento.

5.2 Dados

Um aspecto de investigação primordial deste trabalho é a análise da capacidade preditiva (*timing*) do modelo de seleção de carteiras baseado em erros de predição, que foi avaliada a partir da comparação do desempenho do modelo com o do índice IBOVESPA. Para tal, restringimos o universo de ações dos experimentos apenas às ações que compuseram o IBOVESPA durante o período de simulação de investimento dos experimentos. Embora este não seja um procedimento usual na literatura², optamos por ele para obtermos uma comparação mais estreita, que evidenciasse o desempenho do modelo em termos de *timing* e diversificação através do isolamento de outras ações que pudessem produzir perdas ou ganhos excessivos não captados pelo IBOVESPA. Desta forma, nós listamos as 82 ações que participaram do índice IBOVESPA entre Janeiro de 2005 e Setembro de 2007 e selecionamos um subconjunto de 52 ações com séries históricas de preços com número de observações suficiente para o treinamento das redes neurais e o cálculo dos parâmetros necessários à otimização das carteiras, que para os nossos métodos compreendeu o período de 27-Out-1999 a 19-Set-2007³.

A Tabela 5.1 mostra a relação das 52 ações utilizadas nos nossos experimentos, destacando os períodos de presença de cada ação no IBOVESPA entre 2005 e 2007⁴. Conforme a Tabela 5.1 mostra, nem todas as 52 ações utilizadas participaram do índice IBOVESPA em todos os quadrimestres. O número de ações que participaram do IBOVESPA no período variou entre 40 e 49, do total de 52 ações disponíveis.

Nós computamos os retornos semanais para cada uma das 52 ações da Tabela 5.1 conforme a Eq. 3.2, utilizando as séries históricas dos seus preços de fechamento observados às

²O padrão da literatura é utilizar o índice de mercado no *benchmarking* de estratégias de investimento em quaisquer ações [3, pp. 652].

³Desde o final dos anos 1990, o mercado de ações brasileiro tem experimentado um processo de amadurecimento e crescimento, com aumento contínuo no número de aberturas de capital das empresas (*initial public offering* — IPO). Algumas destas novas ações foram assimiladas e valorizadas pelo mercado, passando a integrar o índice IBOVESPA na primeira metade dos anos 2000. Desta forma, 30 das 82 ações que participaram do índice IBOVESPA no período de simulação de investimento dos nossos experimentos, de 2005 a 2007, não dispunham de séries históricas com observações em todo o período demandado pelos nossos métodos, de 1999 a 2007, e, portanto, não foram incluídas nas nossas análises.

⁴O IBOVESPA é rebalanceado ao final de cada quadrimestre [103].

Tabela 5.1: Relação das 52 ações utilizadas nos experimentos e seus períodos de presença no IBOVESPA entre 2005 e 2007.

Código	Empresa/Ação	Tipo	Presença no IBOVESPA								
			2005			2006			2007		
			Jan/ Abr	Mai/ Ago	Set/ Dez	Jan/ Abr	Mai/ Ago	Set/ Dez	Jan/ Abr	Mai/ Ago	Set/ Dez
ACES4	ACESITA S.A.	PN	✓	✓	✓	✓	✓	✓	✓	✓	✓
AMBV4	CIA BEBIDAS DAS AMERICAS - AMBEV	PN	✓	✓	✓	✓	✓	✓	✓	✓	✓
ARCE3	ARCELOR BRASIL S.A.	ON				✓	✓	✓	✓	✓	
ARCZ6	ARACRUZ CELULOSE S.A.	PNB N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
BBAS3	BCO BRASIL S.A.	ON NM	✓	✓	✓	✓	✓	✓	✓	✓	✓
BBDC4	BCO BRADESCO S.A.	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
BRKM5	BRASKEM S.A.	PNA N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
BRT04	BRASIL TELECOM S.A.	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
BRTP3	BRASIL TELECOM PARTICIPACOES S.A.	ON N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
BRTP4	BRASIL TELECOM PARTICIPACOES S.A.	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
CGAS5	CIA GAS DE SAO PAULO - COMGAS	PNA*	✓	✓	✓	✓	✓	✓	✓	✓	✓
CLSC6	CENTRAIS ELET DE SANTA CATARINA S.A.	PNB N2	✓	✓	✓	✓	✓	✓	✓	✓	✓
CMIG3	CIA ENERGETICA DE MINAS GERAIS - CEMIG	ON N1	✓	✓	✓	✓	✓	✓			
CMIG4	CIA ENERGETICA DE MINAS GERAIS - CEMIG	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
CPLE6	CIA PARANAENSE DE ENERGIA - COPEL	PNB	✓	✓	✓	✓	✓	✓	✓	✓	✓
CRUZ3	SOUZA CRUZ S.A.	ON ED	✓	✓	✓	✓	✓	✓	✓	✓	✓
CSNA3	CIA SIDERURGICA NACIONAL	ON	✓	✓	✓	✓	✓	✓	✓	✓	✓
DURA4	DURATEX S.A.	PN N1									✓
EBTP4	EMBRATEL PARTICIPACOES S.A.	PN *	✓	✓	✓	✓	✓	✓			
ELET3	CENTRAIS ELET BRAS S.A. - ELETROBRAS	ON N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
ELET6	CENTRAIS ELET BRAS S.A. - ELETROBRAS	PNB N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
ELPL5	ELETROPAULO METROP. ELET. SAO PAULO S.A.	PNA* N2						✓			
EMBR3	EMBRAER-EMPRESA BRAS DE AERONAUTICA S.A.	ON NM	✓	✓	✓	✓	✓	✓	✓	✓	✓
GGBR4	GERDAU S.A.	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
GOAU4	METALURGICA GERDAU S.A.	PN N1		✓	✓	✓	✓	✓	✓	✓	✓
ITAU4	BCO ITAU HOLDING FINANCEIRA S.A.	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
ITSA4	ITAUSA INVESTIMENTOS ITAU S.A.	PN ED N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
KLBN4	KLABIN S.A.	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
LAME4	LOJAS AMERICANAS S.A.	PN									✓
LIGT3	LIGHT S.A.	ON * NM					✓	✓	✓	✓	✓
NETC4	NET SERVICOS DE COMUNICACAO S.A.	PN N2		✓	✓	✓	✓	✓	✓	✓	✓
PCAR4	CIA BRASILEIRA DE DISTRIBUICAO	PN N1						✓	✓	✓	✓
PETR3	PETROLEO BRASILEIRO S.A. PETROBRAS	ON	✓	✓	✓	✓	✓	✓	✓	✓	✓
PETR4	PETROLEO BRASILEIRO S.A. PETROBRAS	PN	✓	✓	✓	✓	✓	✓	✓	✓	✓
PRGA3	PERDIGAO S.A.	ON NM					✓	✓	✓	✓	✓
PTIP4	CIA BRASILEIRA DE PETROLEO IPIRANGA	PN	✓	✓	✓	✓	✓	✓	✓	✓	✓
SBSP3	CIA SANEAMENTO BASICO EST SAO PAULO	ON NM	✓	✓	✓	✓	✓	✓	✓	✓	✓
SDIA4	SADIA S.A.	PN N1			✓	✓	✓	✓	✓	✓	✓
TBLE3	TRACTEBEL ENERGIA S.A.	ON NM	✓	✓							
TCSL3	TIM PARTICIPACOES S.A.	ON	✓	✓	✓	✓	✓	✓	✓	✓	✓
TCSL4	TIM PARTICIPACOES S.A.	PN	✓	✓	✓	✓	✓	✓	✓	✓	✓
TLPP4	TELEM DE SAO PAULO S.A. - TELES	PN	✓	✓	✓	✓	✓	✓	✓	✓	✓
TMCP4	TELEMIG CELULAR PARTICIPACOES S.A.	PN	✓	✓	✓	✓	✓	✓	✓	✓	✓
TNLP3	TELE NORTE LESTE PARTICIPACOES S.A.	ON	✓	✓	✓	✓	✓	✓	✓	✓	✓
TNLP4	TELE NORTE LESTE PARTICIPACOES S.A.	PN	✓	✓	✓	✓	✓	✓	✓	✓	✓
TRPL4	CIA TRANSMISSAO ENERGIA ELET PAULISTA	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
UBBR11	UNIBANCO UNIAO DE BCOS BRASILEIROS S.A.	UNT N1		✓	✓	✓	✓	✓	✓	✓	✓
USIM5	USINAS SID DE MINAS GERAIS S.A.-USIMINAS	PNA	✓	✓	✓	✓	✓	✓	✓	✓	✓
VALE3	CIA VALE DO RIO DOCE	ON N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
VALE5	CIA VALE DO RIO DOCE	PNA N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
VCPA4	VOTORANTIM CELULOSE E PAPEL S.A.	PN N1	✓	✓	✓	✓	✓	✓	✓	✓	✓
VIVO4	VIVO PARTICIPACOES S.A.	PN					✓	✓	✓	✓	✓

quartas-feiras no período de 27-Out-1999 a 19-Set-2007. Foram obtidas 52 séries de retornos com 413 observações cada, que formaram o conjunto fundamental de dados para os nossos experimentos. A amostragem no meio da semana foi escolhida para evitar os efeitos do iní-

cio e do fim da semana nos preços das ações [3, 15]. Os dados faltantes (preços das ações) foram preenchidos com os valores passados mais próximos disponíveis nas séries históricas, não foram utilizados mecanismos de detecção e remoção de *outliers* e, devido à estabilidade econômica no período dos experimentos, optamos por não deflacionar nem indexar as séries de preços.

A Tabela 5.2 mostra o sumário das 52 séries de 413 retornos semanais entre 27-Out-1999 e 19-Set-2007, e é organizada como a seguir. No topo da tabela, é apresentado o sumário dos retornos das 52 séries, mostrando a média, o valor mínimo (mín.) e o valor máximo (máx.) do número total de observações, do retorno médio (\bar{r}), da variância (σ^2) e do desvio padrão (σ) das séries de retornos. Em seguida, é apresentado o mesmo sumário para os retornos positivos e para os retornos negativos. Por fim, é apresentado o sumário das variâncias e covariâncias dos retornos das 52 séries, mostrando o número total de observações e a média, o valor mínimo (mín.) e o valor máximo (máx.) das variâncias (σ^2) e das covariâncias (γ). Conforme a Tabela 5.2 mostra, os retornos semanais médios das 52 séries no período obtiveram uma média de 0,00641 e a média das variâncias das 52 séries ficou em 0,00367. As séries apresentaram 210 retornos positivos e 180 retornos negativos, em média, no período. Os retornos positivos totalizaram 51% dos retornos ($210/413 = 0,51$) e os negativos 44% ($180/413 = 0,44$), alcançando uma média de 0,0491 e $-0,0423$, respectivamente, com média das variâncias de 0,00213 para os retornos positivos e 0,00129 para os retornos negativos. Portanto, a magnitude dos retornos positivos foi ligeiramente superior e os retornos positivos tiveram uma variância maior que a dos os retornos negativos no período.

Os experimentos de simulação de investimento foram conduzidos para o modelo de seleção de carteiras baseado em erros de predição e para o modelo média-variância durante 142 semanas entre 5-Jan-2005 e 19-Set-2007, abrangendo um período de aproximadamente três anos. A Tabela 5.3 (pág. 78) mostra o sumário das 52 séries de 142 retornos semanais das ações utilizadas neste período, e está organizada da mesma forma que a Tabela 5.2. Conforme a Tabela 5.3 mostra, os retornos semanais médios das 52 séries no período obtiveram uma média de 0,00939 e a média das variâncias das 52 séries ficou em 0,00355. As séries apresentaram 80 retornos positivos e 57 retornos negativos, em média, no período. Os retornos positivos totalizaram 56% dos retornos ($80/142 = 0,56$) e os negativos 39% ($57/142 = 0,39$), alcançando uma média de 0,0478 e $-0,0431$, respectivamente, com média das variâncias de 0,00156 para os retornos positivos e 0,00166 para os retornos negativos. Portanto, a magnitude dos retornos positivos foi ligeiramente superior e os retornos positivos e negativos tiveram uma variância muito próxima.

Nos experimentos de simulação de investimento, nós formamos a primeira carteira de cada modelo em 29-Dez-2004 utilizando a estratégia de investimento descrita na Seção 5.1

Tabela 5.2: Sumário das 52 séries de 413 retornos semanais das ações no período total abrangido pelo conjunto fundamental de dados utilizados nos experimentos. Os retornos foram observados às quartas-feiras entre 27-Out-1999 e 19-Set-2007.

Sumário dos Retornos			
	média	mín.	máx.
Número de observações	413	413	413
\bar{r}	0,00641	-0,00141	0,0139
σ^2	0,00367	0,00183	0,00983
σ	0,0594	0,0428	0,0991
Sumário dos Retornos Positivos			
	média	mín.	máx.
Número de observações	210	185	232
\bar{r}	0,0491	0,0329	0,0737
σ^2	0,00213	0,000868	0,00636
σ	0,0447	0,0295	0,0798
Sumário dos Retornos Negativos			
	média	mín.	máx.
Número de observações	180	149	209
\bar{r}	-0,0423	-0,0756	-0,0286
σ^2	0,00129	0,000664	0,00443
σ	0,0352	0,0258	0,0666
Sumário das Variâncias e Covariâncias dos Retornos			
Número de observações	52 variâncias e 2652 covariâncias		
	média	mín.	máx.
σ^2	0,00367	0,00183	0,00983
γ	0,00115	0,000050	0,00416

a partir de uma janela temporal com 104 retornos semanais passados entre 8-Jan-2003 e 29-Dez-2004. Estes retornos passados foram utilizados para a estimação dos parâmetros dos modelos, isto é, o cálculo dos erros de predição, suas variâncias e covariâncias para o modelo de seleção de carteiras baseado em erros de predição, e o cálculo dos retornos médios, variâncias e covariâncias dos retornos para o modelo média-variância. Um investimento de \$1 (uma unidade monetária) foi então realizado em cada uma destas primeiras carteiras (uma para cada modelo) em 29-Dez-2004 e seus retornos foram aferidos em 5-Jan-2005, quando uma nova carteira foi formada e todo o capital obtido foi reinvestido da mesma forma. Nós repetimos este procedimento para as 141 semanas restantes do total de 142 semanas, deslizando a janela temporal de 104 retornos uma semana por vez. As últimas car-

Tabela 5.3: Sumário das 52 séries de 142 retornos semanais das ações no período de simulação de investimento dos nossos experimentos. Os retornos foram observados às quartas-feiras entre 5-Jan-2005 e 19-Set-2007.

Sumário dos Retornos			
	média	mín.	máx.
Número de observações	142	142	142
\bar{r}	0,00939	0,000296	0,0156
σ^2	0,00355	0,00163	0,00579
σ	0,0590	0,0404	0,0761
Sumário dos Retornos Positivos			
	média	mín.	máx.
Número de observações	80	69	91
\bar{r}	0,0478	0,0308	0,0623
σ^2	0,00156	0,000626	0,00316
σ	0,0389	0,0250	0,0562
Sumário dos Retornos Negativos			
	média	mín.	máx.
Número de observações	57	47	69
\bar{r}	-0,0431	-0,0598	-0,0319
σ^2	0,00166	0,000826	0,00336
σ	0,0402	0,0287	0,0580
Sumário das Variâncias e Covariâncias dos Retornos			
Número de observações	52 variâncias e 2652 covariâncias		
	média	mín.	máx.
σ^2	0,00355	0,00163	0,00579
γ	0,00183	0,000663	0,00458

teiras foram selecionadas em 12-Set-2007 e seus retornos foram aferidos em 19-Set-2007, encerrando assim o período de simulação de investimento dos experimentos.

A Fig. 5.1 mostra um diagrama da janela temporal deslizante utilizada para a estimação dos parâmetros dos modelos durante as 142 semanas dos experimentos de simulação de investimento.

Os dados artificiais utilizados nos experimentos foram gerados de forma a manter correspondência com os mesmos períodos utilizados para os dados reais, conforme descrito acima.

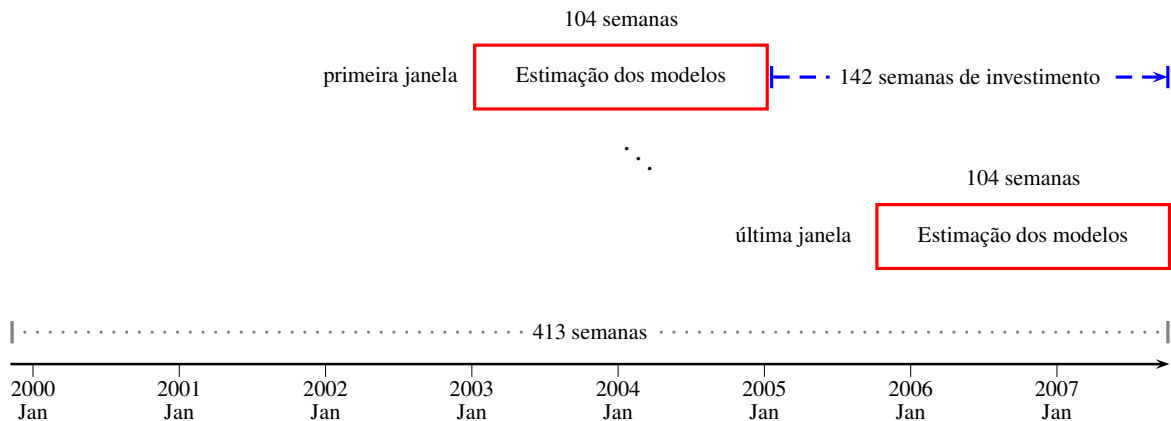


Figura 5.1: Janela temporal deslizante de 104 semanas para a estimação dos parâmetros dos modelos, utilizada durante as 142 semanas dos experimentos de simulação de investimento.

5.3 Arcabouço Computacional

A avaliação experimental do desempenho do modelo de seleção de carteiras baseado em erros de predição frente ao modelo média-variância e ao IBOVESPA, que será apresentada mais adiante no Capítulo 6, foi conduzida através de simulações de investimento obtidas com uma implementação computacional dos modelos de seleção de carteiras e da estratégia de investimento descrita na Seção 5.1. Um conjunto de aplicativos foi desenvolvido para esta finalidade. Nesta seção, apresentaremos brevemente os principais artefatos de software utilizados.

5.3.1 Ambiente de Seleção de Carteiras e Simulação de Investimento

O ambiente *Portfolio Workbench* (PWB) foi concebido como uma ferramenta computacional de suporte à pesquisa, desenvolvimento e avaliação de modelos de seleção de carteiras. A otimização das carteiras do modelo de seleção de carteiras baseado em erros de predição e do modelo média-variância, a seleção das carteiras nas fronteiras eficientes e as diversas simulações de investimento foram inteiramente conduzidas neste ambiente. O PWB é um conjunto de aplicativos desenvolvido em linguagem C para o sistema operacional Linux, que utiliza a biblioteca de interface gráfica com usuário XForms [106], o otimizador LindoAPI [107], a biblioteca *Gnu Scientific Library* (GSL) [108] e o ambiente de plotagem interativo *GnuPlot* [109], entre outras bibliotecas e utilitários padrão providos pelas distribuições tradicionais do sistema operacional Linux.

Os principais aplicativos do PWB são brevemente descritos a seguir:

- *Data Source* (PWB-DS) — implementa o principal banco de dados das séries históricas de cotações das ações (preço, volume, número de negócios, etc.) utilizado pelo ambiente PWB. O aplicativo conta com um ambiente de importação de dados extensível, onde, a partir de uma interface padrão, *drivers* para quaisquer fontes de dados e formatos de arquivos podem ser utilizados. A seleção das séries históricas de interesse é realizada através de um ambiente elementar de processamento de consultas, tendo como parâmetros de seleção as características das séries, tais como tamanho, data inicial e data final, por exemplo. O banco de dados de séries históricas foi implementado no formato texto (*ASCII*), e o modelo de dados foi mapeado na própria estrutura hierárquica do sistema de arquivos do sistema operacional. Estas escolhas foram motivadas pelas vantagens oferecidas por esta abordagem em termos de desempenho, especialmente em relação à implementação tradicional com banco de dados relacional [110]. O acesso aos dados (séries históricas e metadados) é realizado através de uma interface de programação (*Application Programming Interface* — API) única para o ambiente PWB.
- *Project Manager* (PWB-PM) — é o aplicativo que implementa as ferramentas fundamentais para seleção de carteiras e simulação de investimentos. No PWB-PM, um projeto de investimento é construído a partir da seleção do universo de ações de interesse através do PWB-DS e da definição dos seus conjuntos de dados (*data sets*) para seleção das carteiras (isto é, os dados que serão utilizados na otimização das carteiras) e para a simulação de investimento. Nesta etapa, diversos tratamentos para as séries históricas das cotações (preços) podem ser definidos: o intervalo temporal (datas) de interesse, a série de referência para indexação (por exemplo, para deflacionar ou indexar as séries), o método de preenchimento de dados faltantes e ajustes de posição e escala. A partir das séries de preços devidamente tratadas, são calculadas as séries de retornos e de retornos acumulados das ações.

Em seguida, um conjunto de carteiras eficientes pode ser obtido a partir da escolha do modelo de otimização de carteiras desejado, do seu *data set* de estimação de parâmetros e das suas restrições adicionais (como a faixa de retornos de carteira desejados, por exemplo, para o cálculo da fronteira eficiente do modelo). As matrizes de variâncias e covariâncias são calculadas e testadas quanto a sua forma positiva semidefinida, as carteiras são, então, otimizadas para os parâmetros e restrições escolhidos e sua fronteira eficiente é plotada — permitindo a inspeção das todas carteiras eficientes do modelo. A partir da fronteira eficiente e de um *data set* de simulação de investimento, as séries históricas com os resultados da simulação do investimento em todas as carteiras eficientes são produzidas (no período definido pelo *data set* de simulação)

e diversos gráficos e análises comparativas para as carteiras eficientes selecionadas podem ser construídos de forma interativa.

- *Time Series Prediction Tool* (PWB-PRED) — é uma ferramenta multimodelos para a predição de séries temporais. O aplicativo implementa a geração automática dos vetores de dados para modelos auto-regressivos e produz predições recursivas para múltiplos períodos a partir de predições para um período futuro da série histórica, recalculando automaticamente todos os vetores de dados num modelo de janela temporal deslizante. A cada nova predição, são calculadas diversas medidas de erro de predição que são armazenadas para avaliação e comparação de desempenho de preditores. Para o caso de modelos de preditores com aprendizado (por exemplo, redes neurais), a ferramenta automaticamente gera os conjuntos de treinamento, validação e teste para cada predição a partir dos vetores de dados e parâmetros escolhidos, e calcula e armazena os erros de convergência de treinamento e validação para análise e sintonia de modelos. Alguns métodos de controle de treinamento e *overfitting* são implementados na ferramenta e aplicados automaticamente a cada predição conforme uma agenda determinada.

O modelo da rede neural auto-regressiva com referência móvel foi implementado no PWB-PRED com o pacote simulador de redes neurais Aspirin/MIGRAINES v7.0 [92]. Os métodos e demais parâmetros selecionados para os preditores utilizados neste trabalho serão apresentados na Seção 5.3.2 a seguir.

5.3.2 Preditores Neurais Auto-regressivos com Referência Móvel

Nós utilizamos preditores neurais auto-regressivos com referência móvel com regressão de quarta ordem ($p = 4$) e retardo da referência de um ponto ($k = 1$), RNAR-RM (4,1), para prever os retornos das ações para um período futuro. Estes preditores foram implementados com uma rede neural *feedforward* totalmente conectada, com duas camadas ocultas, função de ativação sigmoideal no intervalo $[-1, 1]$, e topologia 4:16:4:1 (quatro neurônios de entrada, 16 neurônios na primeira camada oculta, quatro neurônios na segunda camada oculta e um neurônio de saída). Esta topologia, incluindo a ordem da regressão $p = 4$, foi determinada empiricamente⁵. White [23] propôs a utilização de $p = 5$ para a predição de retornos diários de ações e, a partir de uma avaliação empírica do efeito da variação de p (entre dois e 16) no erro de convergência do treinamento e validação na predição de séries de preços semanais de ações, nós observamos que $p = 4$ minimizava o erro *MAPE* (ver Eq. 5.3 na Seção 5.4.1) dos conjuntos de treinamento e validação [32, 111]. Nós utilizamos $p = 4$ na

⁵As séries de retornos de ações contêm dados de alta frequência, de forma que redes menores, com menos graus de liberdade, são mais propensas a apresentar um subajuste (*underfitting*) do modelo.

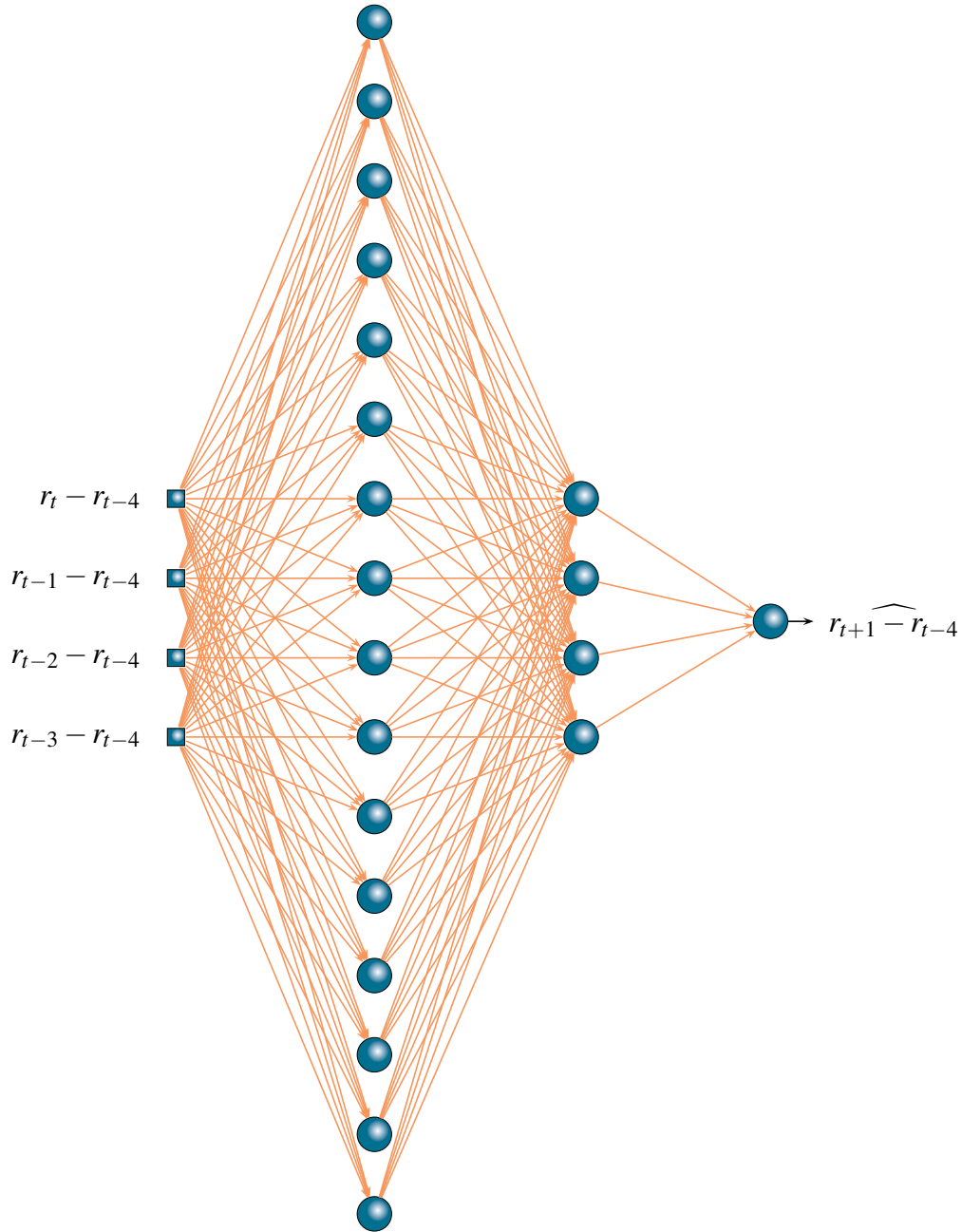


Figura 5.2: Preditor RNAR-RM (4,1) com topologia 4:16:4:1 empregado nas 12.792 predições dos retornos das 52 ações utilizadas nos nossos experimentos.

predição dos retornos semanais — a utilização da informação semanal passada numa janela de um mês (quatro semanas) mostrou-se suficiente para minimizar o erro de convergência para a predição da próxima semana. A Fig. 5.2 mostra o diagrama do preditor RNAR-RM (4,1) com topologia 4:16:4:1 utilizado nos nossos experimentos.

Nós construímos um preditor RNAR-RM (4,1) para cada uma das 52 ações utilizadas nos nossos experimentos. O treinamento e a predição de um retorno futuro com cada rede

neural utilizou uma janela temporal deslizante com 168 retornos semanais dentre os 413 disponíveis. Estas janelas foram deslocadas por 245 semanas ($413 - 168 = 245$), uma semana por vez, e isto nos permitiu obter 246 predições do retorno futuro de cada uma das 52 ações (a predição da primeira janela mais as 245 predições das demais) entre 8-Jan-2003 e 12-Set-2007. Desta forma, nossos experimentos contaram com 12.792 sessões de treinamento ($246 \text{ ciclos de treino e teste} \times 52 \text{ ações} = 12.792$). Cada sessão de treinamento foi conduzida durante 200.000 épocas utilizando o algoritmo *back-propagation* com taxa de aprendizado de 0,009 e inércia de 0,95 (ver Capítulo 3). Os retornos de cada conjunto de treinamento, validação e teste foram normalizados no intervalo $[-0.8, 0.8]$ para evitar a saturação dos neurônios da rede. Estes parâmetros de treinamento e pré-processamento foram determinados empiricamente de forma a balancear o compromisso entre a qualidade da convergência e o tempo computacional gasto em cada sessão de treinamento⁶ [32, 33, 34, 111].

Cada janela de 168 retornos forneceu os retornos semanais necessários para o treinamento da rede neural (163 pares de entradas e saídas) e para a predição do retorno futuro (um par entrada-saída) — os quatro primeiros retornos da janela formaram as quatro entradas do primeiro par de treinamento do preditor RNAR-RM (4,1), e, desta forma, cada janela de 168 retornos forneceu 164 pares de entradas e saídas ($168 - 4 = 164$). Para reduzir o *overfitting*, nós empregamos uma variação do método da validação cruzada (ver Capítulo 3), que será descrita na sequência, na qual inicialmente dividimos a janela de 163 pares de treinamento para formar um conjunto de treinamento com os primeiros 156 pares de entradas e saídas e um conjunto de validação com os últimos sete pares. O conjunto de treinamento foi utilizado para ajustar os pesos da rede neural, enquanto o conjunto de validação foi utilizado para avaliar e selecionar, a cada 1.000 épocas de treinamento (um bloco de treinamento), os pesos que alcançaram a menor raiz do erro quadrático médio de predição (*RMSE* — ver Eq. 5.2 na Seção 5.4.1). Os tamanhos dos conjuntos de treinamento e validação foram obtidos com as heurísticas apresentadas por Haykin [88, pp. 217].

As aplicações de processamento temporal, especialmente a predição de séries temporais não estacionárias, oferecem desafios para o controle do *overfitting* através dos métodos tradicionais de validação cruzada com utilização de conjuntos de treinamento e validação. Estas séries históricas mudam de regime⁷ ao longo do tempo, de forma que os seus últimos valores refletem o regime atual, o qual tem chance de persistir no próximo período futuro. Desta forma, os dados mais recentes necessitam ser incluídos no conjunto de treinamento para que

⁶Redes neurais do porte da topologia utilizada, 4:16:4:1, com 153 parâmetros livres ($((4 + 1) \times 16 + (16 + 1) \times 4 + (4 + 1) \times 1 = 153)$), alcançam bons resultados em termos de convergência de treinamento e generalização [112]. Porém, estas redes requerem um treinamento mais suave, com taxas de aprendizado menores, e, conseqüentemente, durante um número maior de épocas, para evitar o sobreajuste (*overfitting*) do modelo e a saturação das saídas dos neurônios.

⁷O regime de uma série histórica no tempo t é correspondente à sua distribuição nesse tempo, determinando características tais com a sua média, variância e tendência. Séries não estacionárias mudam a sua distribuição, e conseqüentemente, o seu regime ao longo do tempo [113].

a rede possa codificar as informações mais recentes do regime atual nos seus pesos. Por outro lado, esses dados mais recentes também devem ser incluídos no conjunto de validação para aumentar a generalização da rede para o regime atual através do controle adequado do *overfitting*. Não existem métodos padrão estabelecidos para o tratamento desta situação. Alternativas como o deslocamento sistemático ou aleatório do conjunto de validação para outras regiões no tempo se aplicam [64, 114], porém, apesar de contornarem o problema da codificação da informação recente nos pesos da rede, esses métodos aumentam a chance de se validar a generalização da rede com dados de um regime da série histórica que não mais corresponde ao regime atual. Nós atacamos este problema através de um refinamento do nosso método de controle de *overfitting*, que foi inspirado na *validação cruzada não-linear* (*Non-linear Cross Validation* — NCV) proposta por Moody [114].

No nosso procedimento de controle de *overfitting*, nós primeiramente obtemos um conjunto de pesos da rede neural, \mathbf{w}_k , no final do bloco de treinamento k (com $1 \leq k \leq 200$, para um total de 200.000 épocas de treinamento com cada bloco abrangendo 1.000 épocas). Em seguida, antes de iniciar o próximo bloco de treinamento, nós perturbamos este conjunto de pesos, \mathbf{w}_k , através de um treinamento adicional (ajuste adicional dos pesos) com o conjunto de validação durante um pequeno número de épocas (nos nossos experimentos utilizamos dez épocas), assumindo que esta perturbação ajudará a encontrar um novo conjunto de pesos, \mathbf{w}'_k , mais adequado para a predição do próximo valor da série histórica. Então, nós utilizamos a rede neural com os pesos \mathbf{w}'_k para predizer os retornos do conjunto de validação e examinamos o seu erro de predição RMSE. Se o conjunto de pesos \mathbf{w}'_k apresentar o melhor desempenho de todo o treinamento (menor erro de predição RMSE), ele é preservado. Nós, então, prosseguimos com o próximo bloco de treinamento procedendo da mesma forma. Este procedimento foi repetido para as 200.000 épocas do treinamento, no final das quais nós obtivemos o melhor conjunto de pesos do treinamento (em termos de erro de predição RMSE do conjunto de validação), o qual foi utilizado para predizer o conjunto de teste da janela de treinamento corrente.

O procedimento de treinamento e teste (predição) descrito acima foi repetido para todas as 246 predições, avançando a janela deslizante de 168 semanas uma semana por vez. A Fig. 5.3 mostra o diagrama da janela temporal deslizante utilizada para o treinamento e predição dos 246 retornos utilizados durante as 142 semanas dos experimentos de simulação de investimento. Nós computamos as 12.792 sessões de treinamento utilizando o Cluster Enterprise com 64 nós ATHLON XP 1800 do Laboratório de Computação de Alto Desempenho (LCAD) do Departamento de Informática da Universidade Federal do Espírito Santo (<http://www.lcad.inf.ufes.br>).

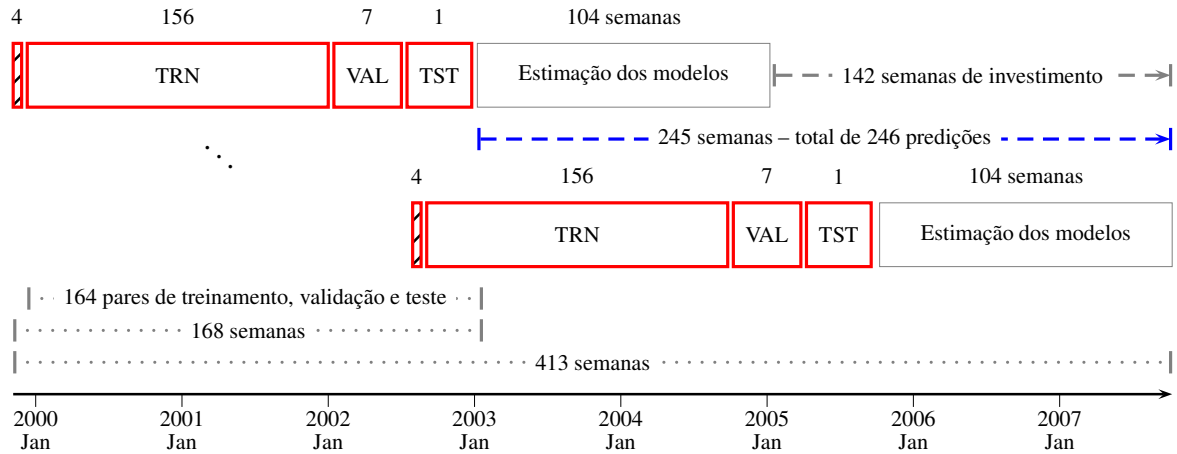


Figura 5.3: Janela temporal deslizante de 168 semanas para treinamento e predição dos 246 retornos utilizados durante as 142 semanas dos experimentos de simulação de investimento.

5.4 Métricas

Nesta seção apresentaremos as métricas que foram utilizadas na avaliação de desempenho dos preditores e dos modelos de seleção de carteiras nos experimentos que serão apresentados no Capítulo 6.

5.4.1 Medidas de Desempenho dos Preditores

O desempenho dos preditores foi avaliado a partir das suas séries de erros de predição (ver Eqs. 4.5 e 4.6), utilizando as medidas de *erro médio*, *raiz do erro quadrático médio*, *erro percentual absoluto médio* e as *taxas de acerto*, que são apresentadas a seguir.

O erro médio (*mean error* — ME) é a média das diferenças entre o retorno realizado e o retorno predito, definida como:

$$ME = \frac{1}{n} \sum_{t=1}^n r_t - \hat{r}_t, \quad (5.1)$$

onde n é o tamanho das séries históricas, e r_t e \hat{r}_t são o retorno realizado e o retorno predito, respectivamente, no tempo t . O ME é utilizado para avaliar a premissa da Eq. 4.7 ($\bar{\epsilon} = 0$), que é fundamental na formulação do modelo de seleção de carteiras baseado em erros de predição apresentado no Capítulo 4.

A raiz do erro quadrático médio (*root mean square error* — RMSE) é uma medida padrão para avaliar as diferenças entre duas séries temporais, e é definida em termos das

diferenças entre o retorno realizado e o retorno predito como:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (r_t - \hat{r}_t)^2}. \quad (5.2)$$

O RMSE, conforme definido acima, pode ser interpretado como o desvio padrão dos erros de predição ($\epsilon_t = r_t - \hat{r}_t$, ver Eq. 4.5) em relação a uma média zero — desvios $\epsilon_t - 0$, ou seja, $(r_t - \hat{r}_t) - 0$; que mostra a distância destes erros de predição em relação à situação ideal de erro de predição médio igual a zero (ver Eq. 4.7). O RMSE tem baixa proteção contra os *outliers*, boa sensibilidade para pequenas mudanças nos dados e não exhibe as assimetrias dos dados [115].

O erro percentual absoluto médio (*mean absolute percentage error* — MAPE) é definido como:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{r_t - \hat{r}_t}{r_t} \right|, \quad (5.3)$$

onde o operador $|\cdot|$ retorna o valor absoluto do seu argumento. O MAPE é amplamente utilizado como figura de mérito do desempenho de preditores na literatura, e é uma medida adimensional, com boa sensibilidade para pequenas mudanças nos dados, que não exhibe as assimetrias nos dados e apresenta muito baixa proteção contra os *outliers*.

As taxas de acerto H_R , H_{R+} , e H_{R-} , mensuram a percentagem das predições nas quais os sinais de r e \hat{r} coincidem: H_R é a percentagem dos retornos preditos que apresentaram o mesmo sinal que os retornos realizados e ambos foram diferentes de zero, H_{R+} é a percentagem para os retornos preditos positivos e H_{R-} é a percentagem para os retornos preditos negativos. Estas medidas são adequadas para a avaliação de preditores como geradores de sinais de negociação [42], e podem ser definidas como:

$$H_R = \frac{Count_{t=1}^n(r_t \hat{r}_t > 0)}{Count_{t=1}^n(r_t \hat{r}_t \neq 0)}, \quad (5.4)$$

$$H_{R+} = \frac{Count_{t=1}^n(r_t > 0 \text{ AND } \hat{r}_t > 0)}{Count_{t=1}^n(\hat{r}_t > 0)}, \quad (5.5)$$

$$H_{R-} = \frac{Count_{t=1}^n(r_t < 0 \text{ AND } \hat{r}_t < 0)}{Count_{t=1}^n(\hat{r}_t < 0)}, \quad (5.6)$$

onde a função $Count_1^n(\cdot)$ retorna o número de ocorrências do seu argumento.

5.4.2 Medidas de Desempenho dos Modelos de Seleção de Carteiras

A avaliação de desempenho das carteiras de ações foi baseada nas medidas de retorno e risco apresentadas nos Capítulos 3 e 4, e nas medidas de *retorno acumulado*, *taxa de mudança da carteira* e *índice de troca*, apresentadas a seguir.

O retorno acumulado (*accumulated return* — *ARE*) no tempo t é definido como:

$$\begin{aligned} ARE_t &= \prod_{i=1}^t (1 + r_i) \\ &= (1 + r_t) ARE_{t-1}, \quad t > 1, \end{aligned} \quad (5.7)$$

onde r_i é o retorno aritmético no período i , conforme definido na Eq. 3.2. Esta é uma medida de desempenho padrão na comparação do resultado de investimentos e relaciona o capital obtido no tempo t , W_t , com o capital inicial, W_0 , como:

$$W_t = W_0 ARE_t. \quad (5.8)$$

Todos os experimentos de simulação de investimento deste trabalho utilizaram um capital inicial $W_0 = \$1$ (uma unidade monetária).

A taxa de mudança da carteira (*portfolio change measure* — *PCM*) [116, 117] é a covariância entre os retornos das ações da carteira e as diferenças das suas respectivas participações em períodos sucessivos no tempo, definida como:

$$PCM = \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{M^{\{t,t-1\}}} r_{it} (X_{it} - X_{it-1}), \quad (5.9)$$

onde T é o número de observações do período de interesse, $M^{\{t,t-1\}}$ é o número total de ações pertencentes às carteiras dos tempos $t-1$ e t , r_{it} é o retorno da ação i no tempo t e X_{it} e X_{it-1} são as participações da ação i nas carteiras dos tempos t e $t-1$, respectivamente. Uma ação contribui positivamente para o *PCM* da carteira quando ela apresenta um retorno positivo no tempo t (ver Eq. 3.2) e sua participação é aumentada entre os tempos $t-1$ e t ($X_{it} > X_{it-1}$), ou, quando ela apresenta um retorno negativo no tempo t e sua participação é diminuída entre os tempos $t-1$ e t ($X_{it} < X_{it-1}$). O *PCM* é um indicador do alinhamento da estratégia de rebalanceamento das carteiras com os movimentos do mercado, e, portanto, serve como uma medida do *timing* (capacidade preditiva) da estratégia de investimento.

Neste trabalho, nós propomos uma nova medida que fornece o percentual do capital investido que é exposto a fricções durante a troca de participações das ações da carteira em

Tabela 5.4: Sentido da melhoria das métricas utilizadas na avaliação de desempenho dos preditores e dos modelos de seleção de carteiras.

Medidas de Desempenho dos Preditores		
Métrica	Descrição	Sentido da Melhoria
ME	erro médio	$ \downarrow $
$RMSE$	raiz do erro quadrático médio	\downarrow
$MAPE$	erro percentual absoluto médio	\downarrow
H_R	taxa de acerto total	\uparrow
H_{R+}	taxa de acerto positivo	\uparrow
H_{R-}	taxa de acerto negativo	\uparrow
Medidas de Desempenho dos Modelos de Seleção de Carteiras		
Métrica	Descrição	Sentido da Melhoria
ARE	retorno acumulado	\uparrow
PCM	taxa de mudança da carteira	\uparrow
TI	índice de troca	\downarrow

períodos sucessivos no tempo⁸. Desta forma, o índice de troca (*turnover index* — TI) [40] é definido como:

$$TI_t = \sum_{i=0}^{M^{\{t,t-1\}}} |X_{it} - X_{it-1}|, \quad (5.10)$$

onde $M^{\{t,t-1\}}$ é o número total de ações pertencentes às carteiras dos tempos t e $t-1$, X_{it} e X_{it-1} são as participações da ação i nas carteiras dos tempos t e $t-1$, respectivamente, e $|\cdot|$ retorna o valor absoluto do seu argumento. Os valores do TI estão no intervalo fechado $[0, 2]$, e seus valores extremos correspondem a manter exatamente a mesma carteira entre os tempos $t-1$ e t (uma exposição às fricções de 0% do capital), ou, vender todas as ações da carteira do tempo $t-1$ e comprar um conjunto de ações inteiramente diferente no tempo t (uma exposição às fricções de 200% do capital)⁹, respectivamente.

5.4.3 Sentido da Melhoria das Métricas Utilizadas

A Tabela 5.4 mostra um sumário do sentido da melhoria das métricas apresentadas na seção anterior, onde o símbolo \uparrow significa que quanto maior o valor da métrica, melhor; o

⁸Embora a estratégia de investimento descrita na Seção 5.1 assuma a ausência de fricções em favor da factibilidade computacional dos modelos de otimização propostos nesse trabalho (ver uma discussão adicional na Seção 7.2), o endereçamento das fricções é um elemento importante para a análise de viabilidade das estratégias de investimento produzidas pelos modelos como alternativas reais de investimento.

⁹Tipicamente, os custos de transação e outras taxas incidem tanto na compra quanto na venda de ações.

símbolo \Downarrow significa que quanto menor o valor da métrica, melhor; e o símbolo $|\Downarrow|$ significa que quanto menor o valor absoluto da métrica, melhor. Conforme a Tabela 5.4 mostra, enquanto as menores medidas de erro são as melhores (com exceção do *ME*, cujo menor valor absoluto é o melhor), as maiores taxas de acerto é que são as melhores.

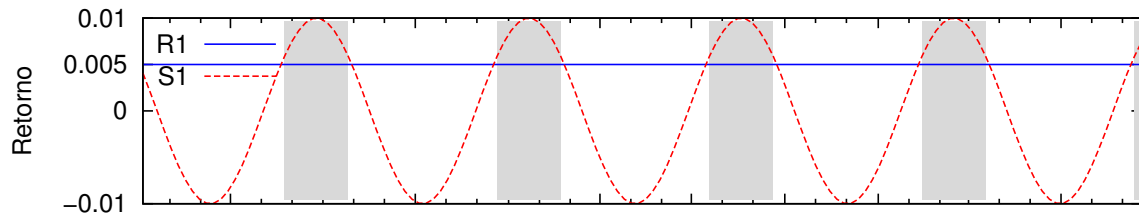
Capítulo 6

Experimentos

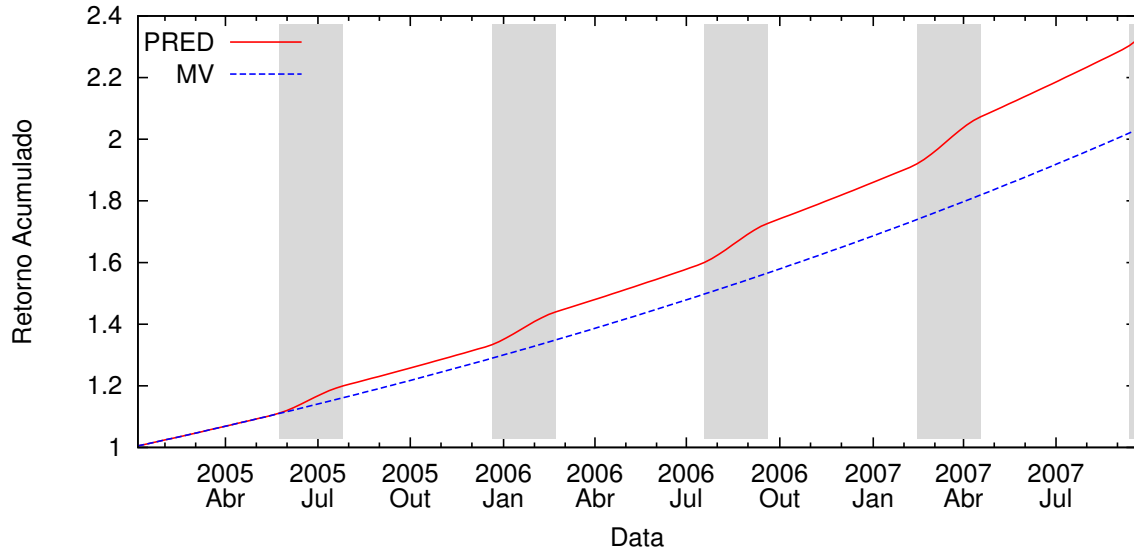
Este capítulo apresenta o conjunto de experimentos que nós utilizamos para avaliar o desempenho do modelo de seleção de carteiras baseado em erros de predição frente ao modelo média-variância e ao índice IBOVESPA da Bolsa de Valores de São Paulo. Nossas avaliações incluíram um experimento controlado de simulação de investimento com dados artificiais, onde demonstramos a capacidade do modelo de seleção de carteiras baseado em erros de predição de explorar oportunidades de curto prazo não capturadas pelo modelo média-variância; e um conjunto de experimentos com dados reais do mercado de ações brasileiro, onde avaliamos o desempenho dos preditores neurais auto-regressivos com referência móvel, a Normalidade dos seus erros de predição e o desempenho dos modelos de seleção de carteiras no investimento no curto prazo. Todos os experimentos desse capítulo foram conduzidos e avaliados utilizando os métodos apresentados no Capítulo 5.

6.1 Investimento com Dados Artificiais

O objetivo do experimento de simulação de investimento com dados artificiais foi demonstrar como o modelo de seleção de carteiras baseado em erros de predição se beneficia de oportunidades preditivas de curto prazo que não são visíveis ao modelo média-variância. Para isso, nós conduzimos um experimento de simulação de investimento utilizando duas ações artificiais, R1 e S1, cujas séries de retornos são mostradas na Fig. 6.1a. A ação R1 tem um retorno semanal constante de 0,005, enquanto que a ação S1 tem retornos semanais senoidais variando entre $-0,01$ e $0,01$ com média igual a zero. O experimento consistiu em implementar a estratégia de investimento da Seção 5.1 para investir nas carteiras que maximizassem o retorno esperado durante o período de 142 semanas definido na Seção 5.2.



(a) Séries de retornos das ações artificiais R1 e S1



(b) Retornos acumulados do investimento nas ações artificiais R1 e R2

Figura 6.1: Simulação de investimento com dados artificiais. A partir das (a) séries de retornos para as ações artificiais R1 e S1, foram obtidos os (b) retornos acumulados para o modelo de seleção de carteiras baseado em erros de predição e para o modelo média-variância. O modelo de seleção de carteiras baseado em erros de predição (PRED) sistematicamente alternou suas posições entre as ações R1 e S1, utilizando a última quando seu retorno predito era superior (regiões sombreadas do gráfico), enquanto que o modelo média-variância (MV) somente utilizou a ação R1, porque o retorno médio de S1 é zero.

A Fig. 6.1b mostra o desempenho do modelo de seleção de carteiras baseado em erros de predição (PRED) e do modelo média-variância (MV) neste experimento, apresentando seus retornos acumulados (ver Eq. 5.7) durante o período de simulação de investimento — conforme visto na Seção 5.2, as séries históricas artificiais foram geradas de forma a corresponder com os períodos das séries reais, de forma que o período de simulação de investimento deste experimento utilizou as últimas 142 semanas (de 5-Jan-2005 a 19-Set-2007) das 413 semanas disponíveis no nosso conjunto de dados artificiais. Os preditores neurais aprenderam a prever as duas séries da Fig. 6.1a com erro igual a zero, permitindo que o modelo de seleção de carteiras baseado em erros de predição acumulasse retornos superiores através da alternância de participações entre R1 e S1, utilizado aquela ação com o maior retorno predito (a curva PRED na Fig. 6.1b); enquanto que o modelo média-variância utilizou

somente a ação R1, porque ela sempre apresentou um retorno médio superior ao de S1 (a curva MV na Fig. 6.1b).

A Tabela 6.1 sumariza os resultados obtidos com todas as métricas empregadas na avaliação dos modelos (ver Seção 5.4.2). Conforme mostrado na Tabela 6.1, o modelo de seleção de carteiras baseado em erros de predição apresentou um retorno acumulado (*ARE*) final de 2,321, valor 14,32% maior que os 2,0303 alcançados pelo modelo média-variância. O modelo de seleção de carteiras baseado em erros de predição acumulou retornos nas taxas semanais de 0,005 (ação R1) e 0,01 (ação S2 — áreas sombreadas da Fig. 6.1), selecionando sempre o maior retorno predito a cada semana e alcançando um retorno semanal médio de 0,00595, enquanto o modelo média-variância acumulou retornos somente na taxa de 0,005 (ação R1).

O modelo de seleção de carteiras baseado em erros de predição apresentou um *TI* médio de 0,127, devido às nove trocas de posição entre R1 e S1 nas 142 semanas ($(9 \times 2)/142 = 0,127$), e um *PCM* positivo pequeno, que contabilizou apenas os nove períodos em que o modelo corretamente trocou de posições entre R1 e S1, escolhendo sempre aquela ação com o maior retorno predito. O modelo média-variância, por sua vez, apresentou um *TI* zero e um *PCM* zero, porque ele somente utilizou a ação R1 em todas as semanas, não efetuando troca de posições. Os dois modelos produziram carteiras com somente uma ação em todas as semanas.

Tabela 6.1: Simulação de investimento com dados artificiais. Sumário do desempenho semanal do modelo de seleção de carteiras baseado em erros de predição e do modelo média-variância.

Carteiras Baseadas em Erros de Predição selecionadas no máximo \hat{r}_p				Carteiras Média-Variância selecionadas no máximo \bar{r}_p			
retorno acumulado		2,321		retorno acumulado		2,0303	
retorno semanal <i>TI</i> <i>PCM</i>	média	σ^2	σ	média	σ^2	σ	
	0,00595	$2,822 \cdot 10^{-06}$	0,00168	0,005	0	0	
	0,127	0,241	0,491	0	0	0	
				<i>PCM</i>			
média		mín.	máx.	média		mín.	máx.
1		1	1	1		1	1
número de ações				número de ações			

6.2 Investimento com Dados Reais

Esta seção apresenta os experimentos de simulação de investimento com as 52 ações do mercado brasileiro listadas na Seção 5.2 durante o período de 142 semanas entre 5-Jan-2005 e 19-Set-2007. Nestes experimentos, nós implementamos a estratégia de investimento da Seção 5.1 e avaliamos o desempenho do modelo de seleção de carteiras baseado em erros de predição em uma situação real de investimentos (mantidas as premissas subjacentes tradicionais da Seção 5.1), comparando os seus resultados com os do modelo média-variância e com o desempenho do índice IBOVESPA no período.

Primeiro, nós apresentaremos uma avaliação do desempenho dos preditores utilizados nos experimentos — os preditores RNAR-RM (4,1) (ver Seção 4.1) —, a partir da análise do conjunto de predições obtido. Em seguida, algumas das fronteiras eficientes obtidas pelos modelos serão mostradas e analisadas, e, na sequência, será apresentada uma avaliação do desempenho do modelo de seleção de carteiras baseado em erros de predição na simulação de investimento em três diferentes níveis de risco ao longo das fronteiras eficientes: baixo risco, risco moderado e alto risco.

6.2.1 Predição de Retornos

Um conjunto de 12.792 predições foi formado com as 246 predições obtidas entre 8-Jan-2003 e 12-Set-2007 para cada uma das 52 ações de interesse ($246 \times 52 = 12.792$, ver Seção 5.3.2). Estas predições foram utilizadas na simulação de investimento com o modelo de seleção de carteiras baseado em erros de predição — as 246 predições de cada ação abrangeram desde as 104 predições da primeira janela deslizante para estimação dos parâmetros do modelo até as 142 predições do período de simulação do investimento (ver Fig. 5.3, pág. 85). Os 52 preditores para os retornos futuros das ações (um preditor para cada ação) foram implementados com redes neurais RNAR-RM (4,1) (ver Seção 4.1), cujo treinamento e predições foram obtidos com os dados, topologia, janelas deslizantes e métodos descritos na Seção 5.3.2.

Desempenho dos Preditores

A Tabela 6.2 sumariza o desempenho dos 52 preditores RNAR-RM (4,1). Ela mostra a média, a variância (σ^2), o desvio padrão (σ), o valor mínimo (mín.) e o valor máximo (máx.) de cada uma das métricas da Seção 5.4.1, calculadas a partir dos resultados das 246 predições produzidas por cada um dos 52 preditores (total de 12.792 predições). Conforme mostrado

Tabela 6.2: Sumário do desempenho dos 52 preditores empregados na predição dos retornos semanais das 52 ações que participaram do índice do IBOVESPA durante o período de simulação de investimento. No total foram obtidas 12.792 predições.

Métrica	média	σ^2	σ	mín.	máx.
ME	-0,000578	0,0000110	0,00324	-0,00869	0,00716
$RMSE$	0,0725	0,000254	0,0159	0,050	0,130
$MAPE$	9,398	163,103	12,771	3,070	86,80
H_R	0,513	0,00125	0,0354	0,448	0,591
H_{R+}	0,610	0,00301	0,0549	0,50	0,696
H_{R-}	0,383	0,00330	0,0575	0,280	0,549

na Tabela 6.2, o ME médio no conjunto dos 52 preditores foi muito próximo a zero e com um baixo desvio padrão. Nós aplicamos o teste t bilateral de uma amostra [118, pp. 244] para o ME da Tabela 6.2 no nível de significância de 5% ($H_0 : ME \text{ médio} = 0$), e encontramos um p -valor de 0,2042¹, que indica que o ME médio de -0,000578 é significativamente igual a zero, portanto, validando a premissa da Eq. 4.7 ($\bar{\epsilon} = 0$) do modelo de seleção de carteiras baseado em erros de predição.

O $RMSE$, o $MAPE$ e o H_R alcançaram valores típicos para estas aplicações [42, 61]. A taxa de acerto H_{R+} média de 61% mostrou que os preditores alcançaram um desempenho médio 11% acima do acaso (50%) na predição dos retornos positivos do mercado, chegando a um máximo de 69% de acerto, enquanto que a predição dos retornos negativos alcançou uma taxa de acerto H_{R-} média de 38%. Uma vez que a energia do erro minimizada pelo algoritmo *back-propagation* é simétrica (isto é, erros positivos e negativos contribuem da mesma forma, ver Eq. 3.12 da Seção 3.3.1), essa assimetria verificada nas taxas de acerto segmentadas (H_{R+} e H_{R-}) pode ser explicada pela melhor generalização dos preditores para os retornos positivos, devido ao maior número de exemplares de retornos positivos nos conjuntos de treinamento dos preditores — dos 413 retornos utilizados no período total dos experimentos, 210 (51%) foram positivos e 180 (44%) foram negativos (ver Tabela 5.2, pág. 77); e dos 142 retornos utilizados no período de simulação de investimento dos experimentos, 80 (56%) foram positivos e 57 (39%) foram negativos (ver Tabela 5.3, pág. 78).

Com exceção do ME e do H_{R+} , o desempenho dos preditores foi modesto. Com um ME próximo de zero ($\bar{\epsilon} \sim 0$), o $RMSE$ pode ser visto como o desvio padrão dos erros de predição, dado pela raiz quadrada da variância dos erros de predição, $\sqrt{\sigma_{\epsilon}^2}$ (ver Eqs. 4.5, 4.7, 4.8 e 5.2). Desta forma, comparando o desvio padrão dos erros de predição com o desvio padrão dos retornos das ações, verificamos que o $RMSE$ médio de 0,0725 foi um pouco superior ao

¹Um p -valor maior que o valor do nível de significância do teste indica que a hipótese nula (H_0) não pode ser rejeitada no nível de significância do teste. No caso contrário, rejeita-se H_0 neste nível de significância. O p -valor é o menor nível de significância no qual H_0 pode ser rejeitada [118, pp. 225].

desvio padrão médio de 0,0590 dos retornos das ações (ver Tabela 5.3, pág. 78), mostrando que, apesar dos preditores não oferecerem um desempenho superior em termos de desvio padrão (risco), o modelo de seleção de carteiras baseado em erros de predição foi capaz de se valer dos resultados coletivos das predições, especialmente dos H_{R+} , para alcançar desempenhos de investimento superiores, conforme será mostrado nas Seções 6.2.3, 6.2.4 e 6.2.5.

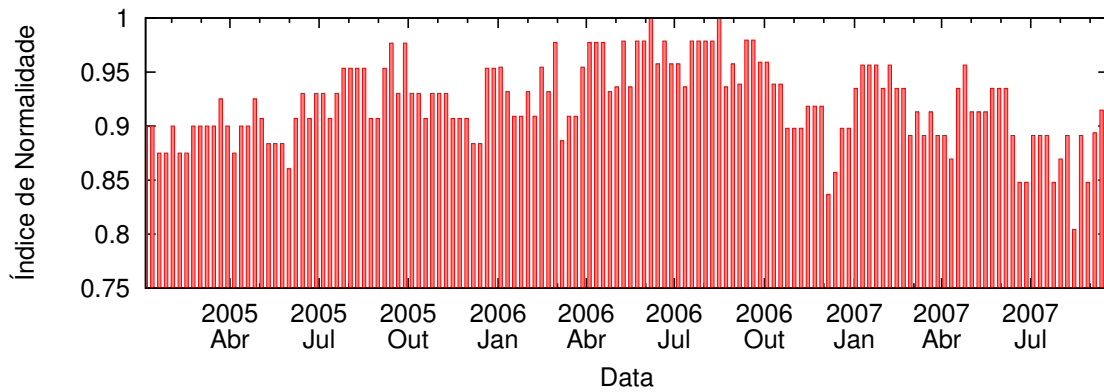
Normalidade dos Erros de predição e dos Retornos

Os experimentos de simulação de investimento foram conduzidos durante 142 semanas entre 5-Jan-2005 e 19-Set-2007 implementando a estratégia de investimento da Seção 5.1. Nós formamos a primeira carteira de cada modelo em 29-Dez-2004, apurando o seu retorno em 5-Jan-2005, e a última carteira em 12-Set-2007, apurando o seu retorno em 19-Set-2007, totalizando assim 142 carteiras para cada modelo. A cada semana, nós computamos a fronteira eficiente do modelo de seleção de carteiras baseado em erros de predição utilizando os retornos preditos e as variâncias e covariâncias da janela temporal deslizante de 104 erros de predição das 52 ações, e computamos a fronteira eficiente do modelo média-variância utilizando os retornos médios e as variâncias e covariâncias da janela temporal deslizante de 104 retornos das 52 ações (ver Seção 5.2 e Fig. 5.1). Cada posição da janela no tempo produziu 52 séries de 104 erros de predição e 52 séries de 104 retornos (uma série de erros de predição e uma série de retornos para cada ação). A janela inicial foi deslocada 141 vezes, avançando uma semana por vez, totalizando 142 posições (uma para cada semana).

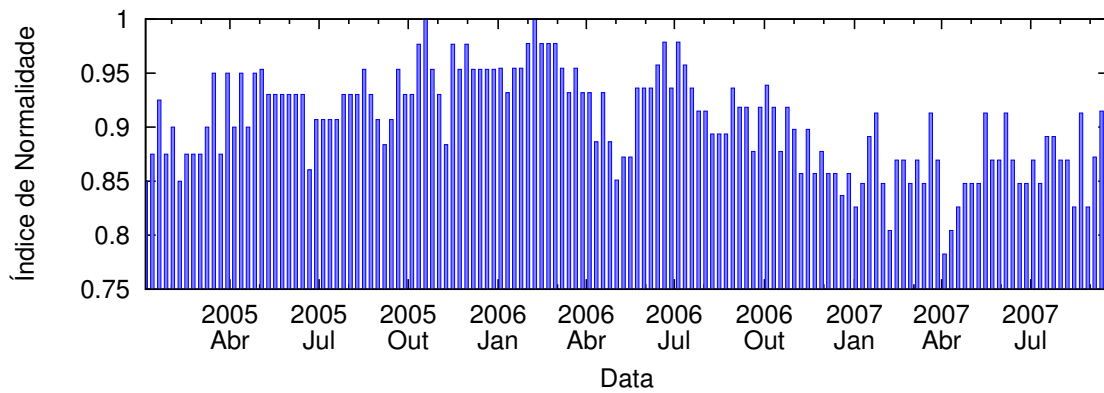
Nós examinamos a Normalidade dessas 142 séries de 104 erros de predição e 142 séries de 104 retornos de cada uma das 52 ações. No total, foram 14.768 séries examinadas ($284 \times 52 = 14.768$), sendo 7.384 séries de erros de predição e 7.384 séries de retornos. Para isso, a cada semana nós medimos as percentagens das 52 séries de erros de predição e das 52 séries de retornos que tiveram a Normalidade aceita (isto é, não rejeitada) de acordo com o *teste de aderência do Qui-quadrado* (*Chi-square test for goodness of fit — χ^2*) [118, pp. 262] no nível de significância de 1%. Nós denominamos esta medida de *índice de Normalidade (IN)*, definida como:

$$IN_t = \frac{\text{número de séries com Normalidade aceita no tempo } t}{\text{número total de séries no tempo } t}. \quad (6.1)$$

Nem todas as 52 ações participaram de todas as 142 fronteiras eficientes semanais de cada modelo durante os experimentos, conforme mostrado na Tabela 5.1 (pág. 75). Apenas as ações que participavam do índice IBOVESPA naquela semana foram utilizadas nas otimizações das carteiras. O número de ações utilizadas em cada semana variou de entre 40 e 49 dentre as 52 ações disponíveis. Conforme apresentado na Seção 5.2, nós adotamos este



(a) Índices de Normalidade para as séries de erros de predição



(b) Índices de Normalidade para as séries de retornos

Figura 6.2: Índices de Normalidade (IN) das (a) 142 séries de 104 erros de predição e (b) 142 séries de 104 de retornos que foram utilizadas para a estimação dos parâmetros do modelo de seleção de carteiras baseado em erros de predição e do modelo média-variância, respectivamente, a cada semana do período de investimento.

procedimento para permitir uma comparação mais precisa do desempenho dos modelos de seleção de carteiras analisados com o desempenho do índice IBOVESPA.

Os IN calculados em cada uma das 142 semanas de simulação de investimento são mostrados na Fig. 6.2. Conforme a Fig. 6.2 mostra, os IN das séries de erros de predição foram superiores e variaram menos que os IN das séries de retornos.

A Tabela 6.3 sumariza os IN das 142 séries de erros de predição e das 142 séries de retornos, mostrando a média, a variância (σ^2), o desvio padrão (σ), o valor mínimo (mín.) e o valor máximo (máx.) dos IN semanais. Conforme a Tabela 6.3 mostra, a média dos IN dos erros de predição ficou em aproximadamente 92%, com uma variância de 0,00134, enquanto que a média dos IN dos retornos ficou em aproximadamente 90%, com uma variância de 0,00198. Nós aplicamos o teste t pareado das diferenças das médias dos IN da Tabela 6.3 no nível de significância de 5% ($H_0 : IN \text{ médio dos erros de predição} \leq IN \text{ médio dos retornos}$),

Tabela 6.3: Sumário dos índices de Normalidade (IN) para as 142 séries de erros de predição e para as 142 séries de retornos utilizadas na estimação dos parâmetros dos modelos de seleção de carteiras durante as 142 semanas do período de simulação de investimento.

	média	σ^2	σ	mín.	máx.
IN dos erros de predição	0,922	0,001 34	0,036 6	0,804	1
IN dos retornos	0,905	0,001 98	0,044 5	0,783	1

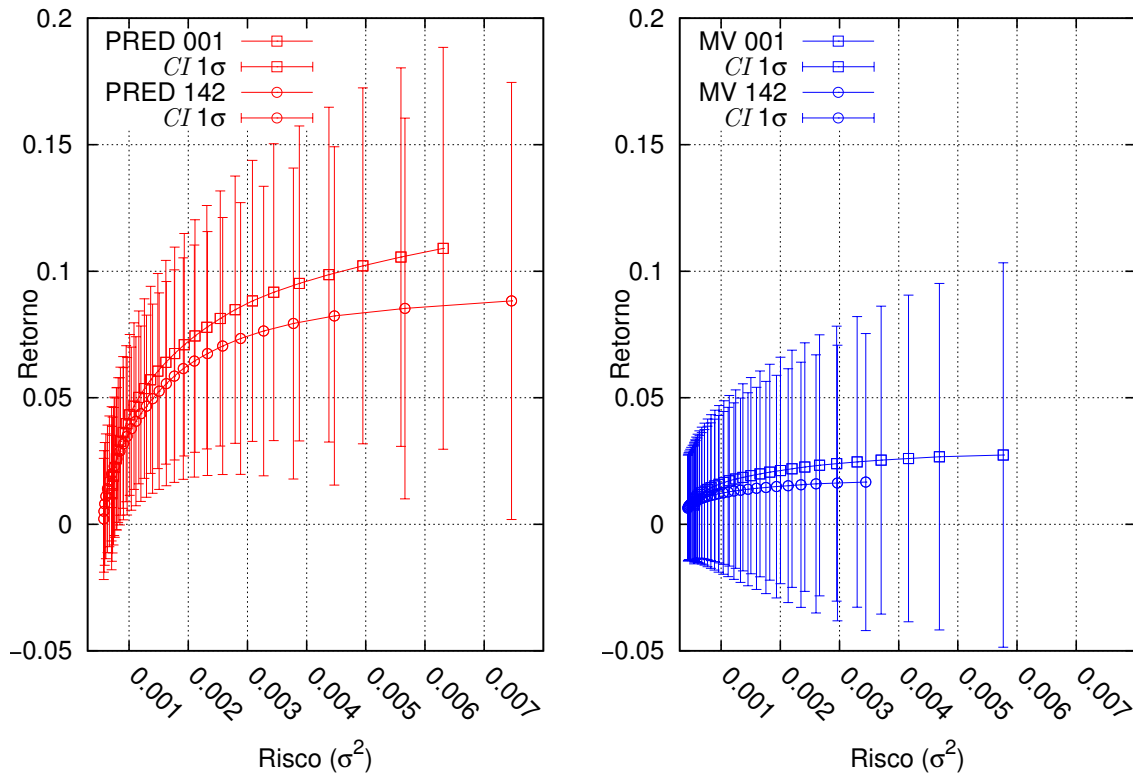
e encontramos um p -valor de $3,512 \cdot 10^{-05}$. Este resultado indica que o IN médio dos erros de predição de 0,922 é significativamente maior que o IN médio dos retornos de 0,905, e, portanto, que as séries dos erros de predição foram mais Normais que as séries de retornos no período considerado.

6.2.2 Fronteiras Eficientes

O primeiro passo da estratégia de investimento empregada nos nossos experimentos de simulação de investimento consistiu em obter uma fronteira eficiente para cada modelo de seleção de carteiras (ver Seção 5.1). Conforme apresentado na Seção 3.4.2, a fronteira eficiente é formada pelo conjunto de carteiras de risco mínimo para vários valores de retorno de carteira desejado. Nos experimentos que serão apresentados mais adiante nas Seções 6.2.3, 6.2.4, e 6.2.5, nós variamos o retorno de carteira desejado, R_d , dentre 30 valores na região de factibilidade de cada modelo. Desta forma, nós obtivemos fronteiras eficientes com 30 carteiras para cada uma das 142 semanas do período de simulação de investimento.

Os limiares do intervalo de factibilidade dos retornos das fronteiras eficientes (o retorno mínimo e o retorno máximo das suas ordenadas — ver Fig. 3.9, pág. 60), que são necessários para a parametrização do modelo de seleção de carteiras baseado em erros de predição e do modelo média-variância para R_d (Eqs. 4.13 e 3.32), foram obtidos através da resolução de dois casos particulares dos modelos, conforme indicado por Elton et al. [3]. O retorno mínimo do intervalo foi o retorno da carteira de mínimo risco com R_d irrestrito (problema primal, minimizando as Eqs. 4.12 e 3.31 sem as restrições das Eqs. 4.13 e 3.32, respectivamente, para os dois modelos); e o retorno máximo foi o retorno da carteira de máximo retorno (problema dual, maximizando as Eqs. 4.9 e 3.26 sem as restrições de risco das Eqs. 4.12 e 3.31, respectivamente, para os dois modelos). Desta forma, um total de 9.088 problemas de otimização foi resolvido nos experimentos $((32 + 32) \times 142 = 9.088)$.

Cada fronteira eficiente do modelo de seleção de carteiras baseado em erros de predição foi computada utilizando os retornos preditos e as variâncias e covariâncias das 52 séries de 104 erros de predição das ações, e cada fronteira eficiente do modelo média-variância



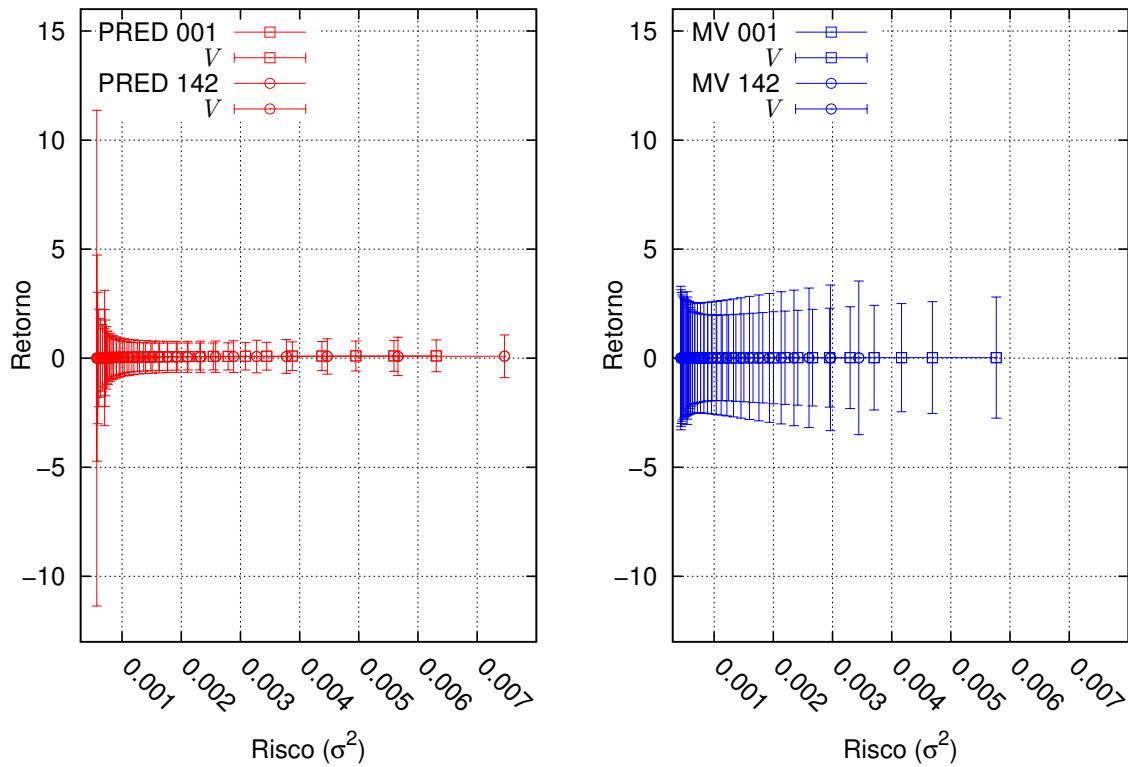
(a) Fronteiras eficientes e CI 1 σ do modelo de seleção de carteiras baseado em erros de predição

(b) Fronteiras eficientes e CI 1 σ do modelo média-variância

Figura 6.3: Duas fronteiras eficientes calculadas nas semanas $t = 1$ e 142 e seus intervalos de confiança de 68, 27% (CI 1 σ): (a) PRED001 e PRED142 para o modelo de seleção de carteiras baseado em erros de predição e (b) MV001 e MV142 para o modelo média-variância. O modelo de seleção de carteiras baseado em erros de predição obteve retornos esperados superiores para os mesmos níveis de risco e apresentou uma menor probabilidade de obter retornos negativos que o modelo média-variância.

foi computada utilizando os retornos médios e as variâncias e covariâncias das 52 séries de 104 retornos das ações. Estas séries foram obtidas através da janela temporal deslizante de estimação de parâmetros dos modelos descrita na Seção 5.2 e ilustrada na Fig. 5.1 (pág. 79). A primeira fronteira eficiente foi computada em 29-Dez-2004 e a última foi computada em 12-Set-2007. No total, foram computadas 142 fronteiras eficientes para cada modelo.

A Fig. 6.3a mostra duas das 142 fronteiras eficientes do modelo de seleção de carteiras baseado em erros de predição, PRED001 e PRED142, obtidas na primeira e na última semana do período de simulação de investimento, $t = 1$ e 142, respectivamente; e a Fig. 6.3b mostra o mesmo para o modelo média-variância, com as fronteiras eficientes MV001 e MV142 obtidas da mesma forma. As carteiras eficientes são plotadas nos pontos da Fig. 6.3 (quadrados e círculos) junto com barras verticais com comprimento de 1 σ (um desvio padrão), calculadas



(a) Fronteiras eficientes e V do modelo de seleção de carteiras baseado em erros de predição

(b) Fronteiras eficientes e V do modelo média-variância

Figura 6.4: Duas fronteiras eficientes calculadas nas semanas $t = 1$ e 142 e seus coeficientes de variação (V): (a) PRED001 e PRED142 para o modelo de seleção de carteiras baseado em erros de predição e (b) MV001 e MV142 para o modelo média-variância. O modelo de seleção de carteiras baseado em erros de predição obteve valores de V menores que a metade daqueles do modelo média-variância para a maioria das suas carteiras.

com a raiz quadrada do risco da carteira (variância). Cada barra de 1σ demarca a região onde o retorno da carteira eficiente é esperado com probabilidade 68,27% de acordo com a distribuição Normal, e constitui um *intervalo de confiança* (*confidence-interval* — *CI*) de 68,27%, ou *CI* 1σ [118, pp. 202]. Conforme mostrado nas curvas e nos respectivos *CI* da Fig. 6.3, a maioria das carteiras baseadas em erros de predição produziu retornos esperados maiores que as carteiras média-variância para os mesmos níveis de risco, e com uma menor probabilidade de obter retornos negativos (houve menos $\hat{r}_p - \hat{\sigma}_p$ negativos que $\bar{r}_p - \sigma_p$ negativos, mostrado nas barras verticais dos gráficos).

A Fig. 6.4 mostra as mesmas fronteiras eficientes da Fig. 6.3 plotadas com seus *coeficientes de variação* (*coefficient of variation* — V) [118, pp. 93] dos retornos esperados das carteiras, $V = \hat{\sigma}_p / \hat{r}_p$ e $V = \sigma_p / \bar{r}_p$, para o modelo de seleção de carteiras baseado em erros de predição e para o modelo média-variância, respectivamente. Conforme mostrado na Fig. 6.4,

o modelo de seleção de carteiras baseado em erros de predição (Fig. 6.4a) obteve valores de V menores que a metade dos valores daqueles do modelo média-variância (Fig. 6.4b) para a maioria das carteiras. As duas primeiras carteiras do modelo de seleção de carteiras baseado em erros de predição tiveram retornos preditos muito próximo de zero (ver Fig. 6.3a), e, neste caso, a medida do coeficiente de variação é inadequada para comparação dessas carteiras, conforme comentado por Spiegel e Stephens [118, pp. 94].

No nosso conhecimento, esta é a primeira vez na literatura que as fronteiras eficientes são analisadas com o auxílio dos intervalos de confiança dos retornos esperados e dos coeficientes de variação dos riscos das carteiras eficientes.

6.2.3 Simulação de Investimento de Baixo Risco

Este experimento avalia o desempenho do modelo de seleção de carteiras baseado em erros de predição no investimento de baixo risco. Para isso, nós utilizamos as fronteiras eficientes obtidas na Seção 6.2.2 para implementar o segundo passo da estratégia de investimento da Seção 5.1, selecionando, a cada semana, carteiras com retornos semanais esperados \hat{r}_p e \bar{r}_p iguais a 0,0135, respectivamente, para o modelo de seleção de carteiras baseado em erros de predição e para o modelo média-variância. Estes retornos esperados correspondem aos níveis de risco do início das fronteiras eficientes dos modelos (ver Fig. 6.3, pág. 99). A primeira carteira de cada modelo foi selecionada em 29-Dez-2004 e a última foi selecionada em 12-Set-2007. Os resultados do investimento nas carteiras selecionadas foram observados no decorrer das 142 semanas do período de simulação de investimento, entre 5-Jan-2005 e 19-Set-2007.

A Fig. 6.5 mostra os retornos acumulados semanais (ARE , ver Eq. 5.7) obtidos com o modelo de seleção de carteiras baseado em erros de predição (PRED), com o modelo média-variância (MV) e com o índice IBOVESPA (IBOV). Conforme a Fig. 6.5 mostra, o modelo de seleção de carteiras baseado em erros de predição superou o modelo média-variância no final das 142 semanas e acompanhou o índice IBOVESPA de forma mais precisa que o modelo média-variância. O modelo média-variância oscilou em torno do índice IBOVESPA de forma mais intensa que o modelo de seleção de carteiras baseado em erros de predição, apresentando um desempenho inferior ao do índice no final do período.

A Fig. 6.6 (pág. 104) apresenta uma avaliação dos retornos acumulados semanais dos modelos de seleção de carteiras no contexto do arcabouço Normal, mostrando as curvas dos retornos acumulados *realizados* e dos retornos acumulados *esperados* do modelo de seleção de carteiras baseado em erros de predição (Fig. 6.6a) e do modelo média-variância (Fig. 6.6b). A curva dos retornos acumulados realizados é obtida da forma usual (ver

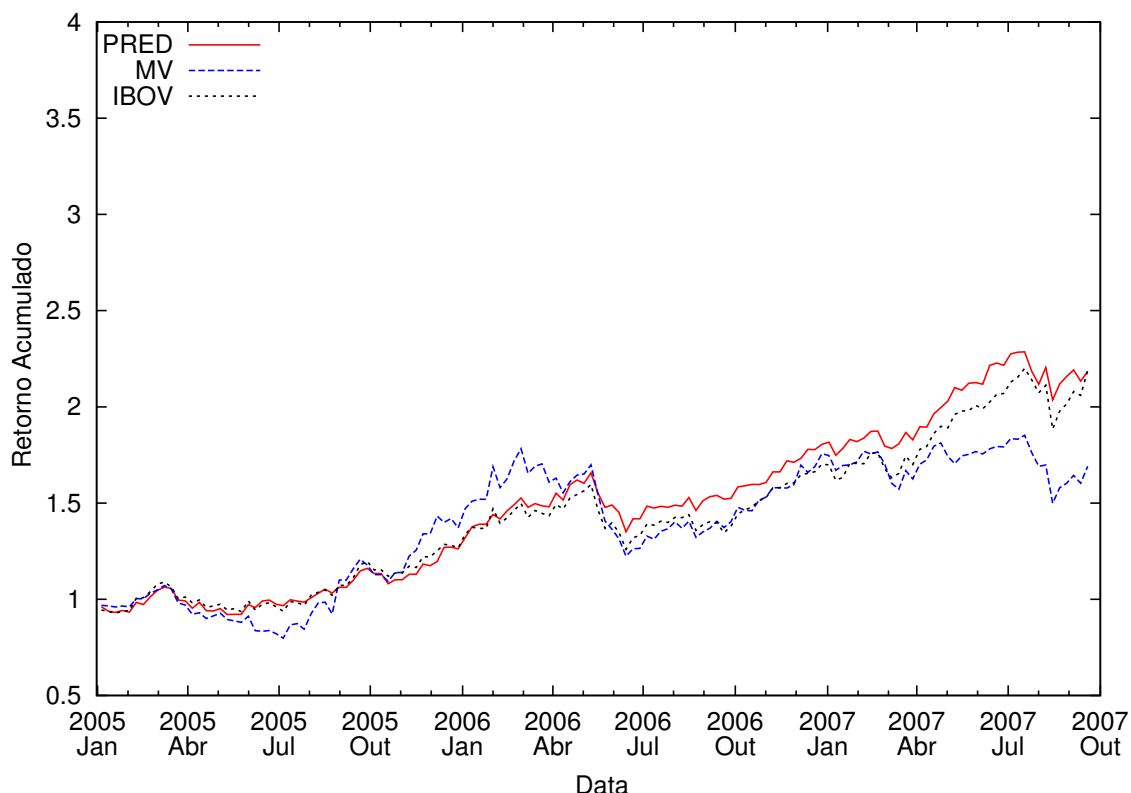


Figura 6.5: Simulação de investimento de baixo risco. Retornos acumulados semanais das 142 carteiras do modelo de seleção de carteiras baseado em erros de predição (PRED) e do modelo média-variância (MV), e do índice IBOVESPA (IBOV). As carteiras foram selecionadas com \hat{r}_p e \bar{r}_p iguais a 0,0135, respectivamente, para os modelos. O modelo de seleção de carteiras baseado em erros de predição superou o modelo média-variância no final das 142 semanas e acompanhou o IBOVESPA de forma mais precisa.

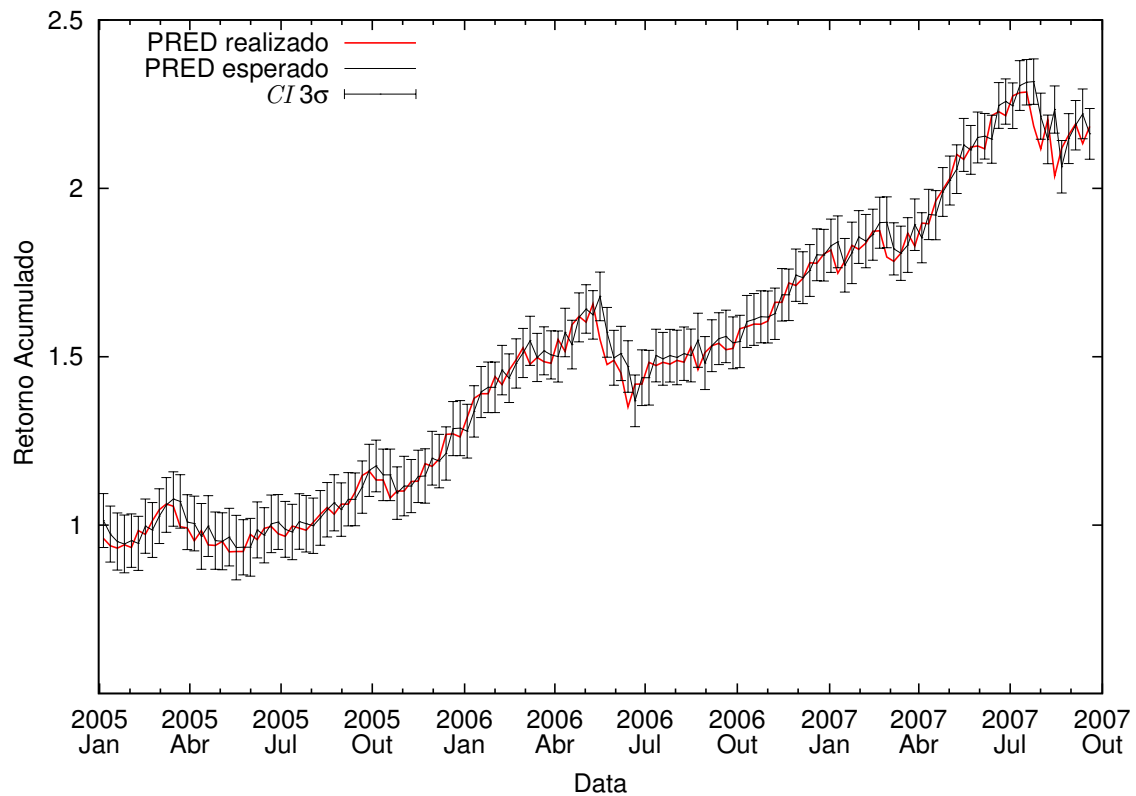
Eq. 5.7), enquanto que a curva dos retornos acumulados esperados é obtida conforme a seguir. Em cada tempo t , o retorno esperado da carteira selecionada no tempo $t - 1$ é acumulado ao retorno acumulado realizado do modelo no tempo $t - 1$ de forma similar à Eq. 5.7, e então é plotado como o retorno acumulado esperado para o tempo t . Em seguida, nós plotamos uma barra vertical com comprimento de 3σ (três desvios padrão, isto é, três vezes a raiz quadrada da variância, ou risco, da carteira selecionada no tempo $t - 1$) centrada nesse retorno acumulado esperado para o tempo t . Estas barras de 3σ correspondem aos intervalos de confiança de 99,73% ($CI\ 3\sigma$), onde o retorno acumulado da carteira no tempo t é esperado com uma probabilidade de 99,73% de acordo com a distribuição Normal [118, pp. 202]. Conforme a Fig. 6.6 (pág. 104) mostra, os dois modelos apresentaram um comportamento muito próximo do esperado no arcabouço Normal durante as 142 semanas de simulação de investimento.

A Tabela 6.4 (pág. 106) sumariza o desempenho semanal do modelo de seleção de carteiras baseado em erros de predição, do modelo média-variância e do índice IBOVESPA nas 142 semanas do período de simulação de investimento, e está organizada como a seguir. No topo da tabela, são apresentados os sumários para o modelo de seleção de carteiras baseado em erros de predição e o modelo média-variância, mostrando o retorno acumulado no período; a média, variância (σ^2) e desvio padrão (σ) dos retornos semanais e dos *TI*; o *PCM* no período; e a média, o valor mínimo (mín.) e o valor máximo (máx.) do número de ações das carteiras no período. Na parte inferior da tabela, é apresentado o sumário para o índice IBOVESPA, mostrando o seu retorno acumulado no período; a média, variância (σ^2) e desvio padrão (σ) dos seus retornos semanais; e a média, o valor mínimo (mín.) e o valor máximo (máx.) do número de ações que integrou o índice no período. Conforme a Tabela 6.4 mostra, o modelo de seleção de carteiras baseado em erros de predição alcançou um retorno acumulado de 2,184, que foi 29% superior ao retorno acumulado de 1,691 alcançado pelo modelo média-variância e próximo ao retorno acumulado de 2,189 obtido pelo IBOVESPA no período. O modelo de seleção de carteiras baseado em erros de predição obteve um risco (variância) *ex-post* de 0,000744, que foi 57% menor que o risco de 0,00175 obtido pelo modelo média-variância e 25% menor que o risco de 0,000998 obtido pelo IBOVESPA. Portanto, o modelo de seleção de carteiras baseado em erros de predição alcançou retornos nos mesmos níveis do mercado (índice IBOVESPA) e com um risco 25% menor.

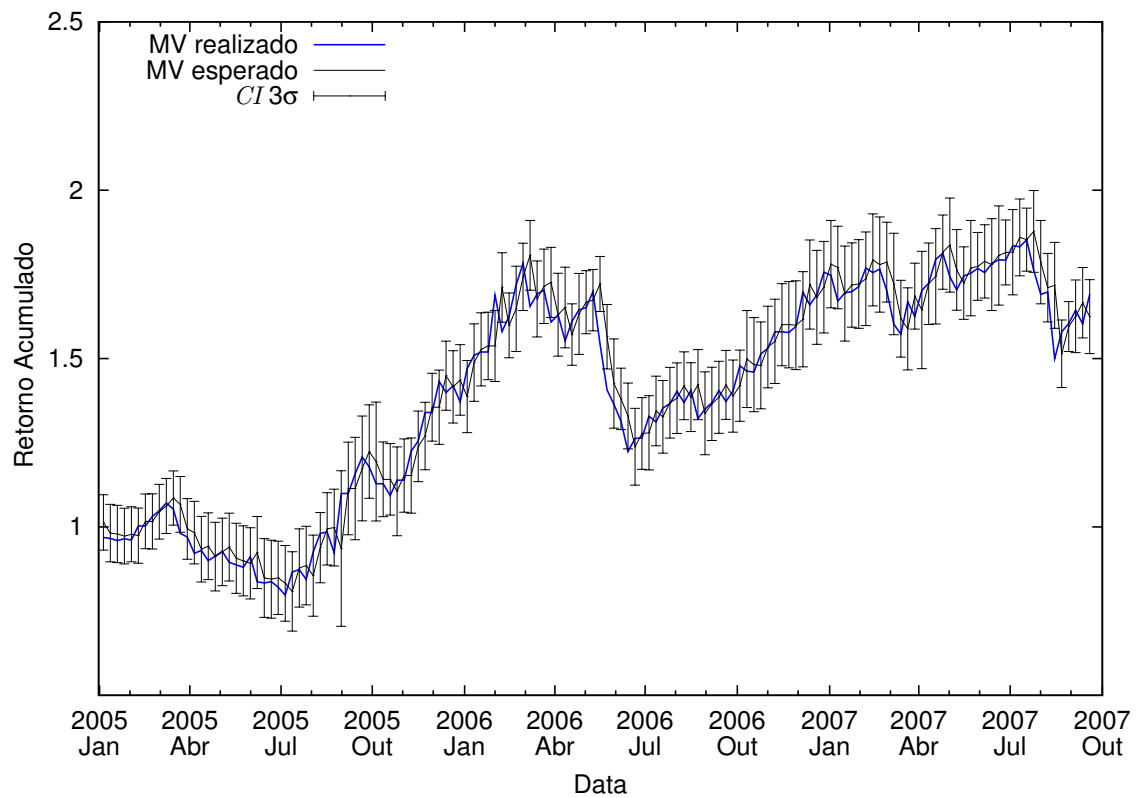
Conforme mostrado anteriormente nas Eqs. 3.37 e 3.38 (pág. 61), o valor da covariância média das ações é uma linha de base para a diversificação do risco no arcabouço média-variância. Considerando esse resultado para os riscos *ex-post*, o risco de 0,000744 alcançado pelo modelo de seleção de carteiras baseado em erros de predição foi 59% menor que a covariância média de 0,00183 apresentada pelos retornos das ações no período (ver Tabela 5.3, pág. 78), enquanto que o risco de 0,00175 alcançado pelo modelo média-variância ficou próximo ao valor da covariância média dos retornos das ações.

Os dois modelos apresentaram um *PCM* negativo e baixo, e o *TI* médio de 42% do modelo de seleção de carteiras baseado em erros de predição foi um pouco melhor que o *TI* médio de 45% do modelo média-variância.

Neste e nos demais experimentos deste capítulo, o número de ações de cada carteira considerou apenas as ações com participações X_i maiores que o valor de corte de 0,000001. Assim, o número de ações das carteiras baseadas em erros de predição variou entre 12 e 34, com uma média de 21 ações, enquanto que o número de ações das carteiras média-variância variou entre dois e 21, com uma média de sete ações. Portanto, o modelo de seleção de carteiras baseado em erros de predição diversificou melhor os investimentos, aproximando-se do referencial de 30 ações proposto por Statman [119].



(a) Retornos acumulados e $CI\ 3\sigma$ para o modelo de seleção de carteiras baseado em erros de predição



(b) Retornos acumulados e $CI\ 3\sigma$ para o modelo média-variância

Figura 6.6: Simulação de investimento de baixo risco. Retornos acumulados realizados e esperados com seus intervalos de confiança de 99,73% ($CI\ 3\sigma$): (a) modelo de seleção de carteiras baseado em erros de predição e (b) modelo média-variância.

A avaliação trimestral do desempenho de investimentos é de especial interesse na indústria. Os resultados operacionais das empresas, os valores de diversos indicadores econômicos e o resultado das contas públicas, entre outros, são divulgados trimestralmente. Nós avaliamos os resultados trimestrais obtidos pelos modelos de seleção de carteiras e pelo IBOVESPA calculando os seus retornos semanais médios no trimestre e os seus retornos acumulados trimestrais². A Tabela 6.5 (pág. 107) mostra os intervalos de confiança (*CI*) de 95% dos retornos semanais médios no trimestre para o modelo de seleção de carteiras baseado em erros de predição (PRED), para o modelo média-variância (MV) e para o índice IBOVESPA (IBOV). Conforme mostrado na Tabela 6.5, a média do retorno semanal médio no trimestre de 0,00601 obtida pelo modelo de seleção de carteiras baseado em erros de predição foi maior que a média de 0,00492 obtida pelo modelo média-variância e um pouco menor que a média de 0,00643 obtida pelo IBOVESPA. O modelo de seleção de carteiras baseado em erros de predição obteve um erro padrão de 0,00427, que foi menos que a metade do erro padrão de 0,00891 do modelo média-variância e menor que o erro padrão de 0,00509 do IBOVESPA. A Fig. 6.7 (pág. 107) mostra uma representação gráfica das médias e dos intervalos de confiança da Tabela 6.5, onde é evidenciado o posicionamento da média do modelo de seleção de carteiras baseado em erros de predição (um pouco abaixo do IBOVESPA e um pouco acima do modelo média-variância) e o seu menor erro padrão (menor risco).

O modelo de seleção de carteiras baseado em erros de predição obteve um retorno acumulado trimestral médio de 1,080 com variância de 0,00908, enquanto que o modelo média-variância obteve um retorno acumulado trimestral médio de 1,069 com variância de 0,0355. Nós aplicamos o *teste t* pareado na comparação das médias dos modelos no nível de significância de 5% (H_0 : *ARE* médio do modelo de seleção de carteiras baseado em erros de predição \leq *ARE* médio do modelo média-variância), e encontramos um *p-valor* de 0,3843. Este resultado indica que a diferença entre as médias dos retornos acumulados trimestrais dos modelos não é significativa.

As variâncias dos retornos acumulados trimestrais dos dois modelos foram testadas com o *teste F* [118, pp. 246] no nível de significância de 5% (H_0 : razão da variância do modelo de seleção de carteiras baseado em erros de predição pela variância do modelo média-variância ≥ 1), e encontramos um *p-valor* de 0,02124. Este resultado indica que a variância (risco) alcançada pelo modelo de seleção de carteiras baseado em erros de predição foi significativamente inferior à variância (risco) do modelo média-variância.

²Os retornos semanais médios no trimestre e os retornos acumulados trimestrais foram calculados partir dos retornos semanais com as Eqs. 3.25 e 5.7, respectivamente, durante 13 semanas (52 semanas no ano / 4 trimestres no ano = 13 semanas por trimestre). Nós repetimos o último retorno semanal disponível para obter 143 (142 + 1 = 143) semanas e exatamente 11 retornos trimestrais (143 semanas / 13 semanas por trimestre = 11 trimestres).

Tabela 6.4: Simulação de investimento de baixo risco. Sumário do desempenho semanal dos modelos de seleção de carteiras e do índice IBOVESPA durante as 142 semanas do período de investimento.

Carteiras Baseadas em Erros de Predição selecionadas em $\hat{r}_p = 0,0135$				Carteiras Média-Variância selecionadas em $\bar{r}_p = 0,0135$			
retorno acumulado	2,184			retorno acumulado	1,691		
	média	σ^2	σ		média	σ^2	σ
retornos semanais	0,00589	0,000744	0,0273	retornos semanais	0,00456	0,00175	0,0418
<i>TI</i>	0,421	0,0580	0,241	<i>TI</i>	0,454	0,104	0,323
<i>PCM</i>	−0,000590			<i>PCM</i>	−0,000636		
	média	mín.	máx.		média	mín.	máx.
número de ações	21	12	34	número de ações	7	2	21

Índice IBOVESPA			
retorno acumulado	2,189		
	média	σ^2	σ
retornos semanais	0,00603	0,000998	0,0316
	média	mín.	máx.
número de ações	57	53	63

Tabela 6.5: Simulação de investimento de baixo risco. Intervalos de confiança (CI) de 95% dos retornos semanais médios no trimestre durante as 142 semanas do período de investimento.

	Média	Erro padrão	Intervalo de confiança da média (95%)
PRED	0,00601	0,00427	$0,00175 \leq 0,00601 \leq 0,0103$
MV	0,00492	0,00891	$-0,00399 \leq 0,00492 \leq 0,0138$
IBOV	0,00643	0,00509	$0,00133 \leq 0,00643 \leq 0,0115$

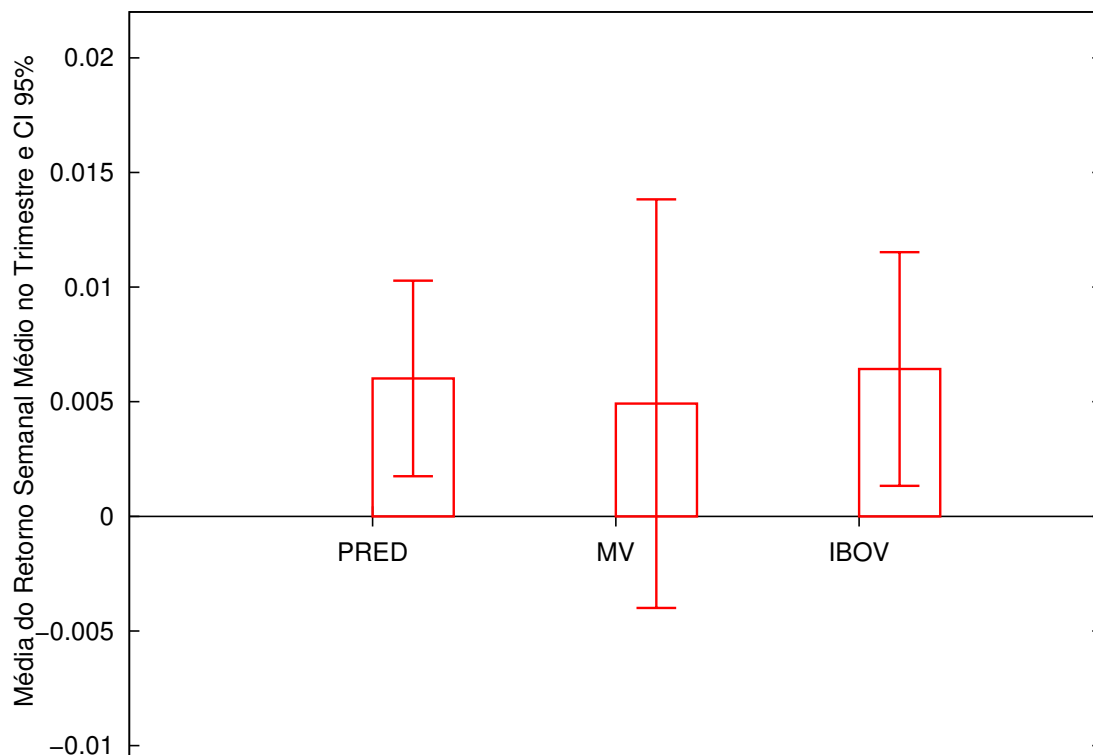


Figura 6.7: Simulação de investimento de baixo risco. Intervalos de confiança (CI) de 95% dos retornos semanais médios no trimestre para o modelo de seleção de carteiras baseado em erros de predição (PRED), para o modelo média-variância (MV) e para o índice IBOVESPA durante as 142 semanas do período de investimento. O menor risco apresentado pelo modelo de seleção de carteiras baseado em erros de predição é evidenciado pelo seu menor erro padrão da média, mostrado nas barras verticais.

6.2.4 Simulação de Investimento de Risco Moderado

Os investidores freqüentemente se deparam com o problema de diversificar seus investimentos em uma carteira de ações e um *ativo sem risco* (*risk-free asset*), ou *ativo de renda fixa*. Esses ativos são aqueles com risco (variância) igual ou muito próximo a zero, como, por exemplo, os Certificados de Depósito Bancário, as Letras do Tesouro Nacional e a caderneta de poupança. Para este caso, é demonstrado na literatura que o conjunto de soluções ótimas é formado pelas combinações da participação no ativo sem risco e da participação em uma determinada carteira da fronteira eficiente [3, 4]. Esta carteira, P , está localizada no ponto em que uma reta que cruza a ordenada dos retornos no valor do retorno esperado do ativo sem risco, R_{rf} , tangencia a fronteira eficiente, plotada com o desvio padrão das carteiras nas abscissas, conforme mostrado na Fig. 6.8. Todas as combinações da participação no ativo sem risco e da participação na carteira P são eficientes. A localização da carteira P depende do valor de R_{rf} e do formato da fronteira eficiente, este último determinado pelos retornos e riscos das suas ações. De uma forma geral, quando os valores de R_{rf} são razoáveis, a carteira P localiza-se próximo ao meio da fronteira eficiente das ações. Neste experimento, nós selecionamos as carteiras do modelo de seleção de carteiras baseado em erros de predição e do modelo média-variância segundo esse critério de diversificação, utilizando um R_{rf} semanal de 0,00114, que corresponde à taxa de juros de 6% ao ano da caderneta de poupança.

A Fig. 6.9 (pág. 110) mostra os retornos acumulados semanais obtidos com o modelo de seleção de carteiras baseado em erros de predição (PRED), com o modelo média-variância (MV) e com o índice IBOVESPA (IBOV) durante as mesmas 142 semanas do período de simulação de investimento do experimento anterior da Seção 6.2.3, e com a mesma estratégia de investimento e os mesmos dados e datas utilizados. Conforme mostrado na Fig. 6.9, o modelo de seleção de carteiras baseado em erros de predição superou o modelo média-variância e o IBOVESPA em quase todas as semanas. É importante mencionar que o modelo média-variância obteve um desempenho inferior ao do IBOVESPA.

A Fig. 6.10 (pág. 111) apresenta a mesma avaliação dos modelos de seleção de carteiras no contexto do arcabouço Normal descrita no experimento anterior da Seção 6.2.3, mostrando os retornos acumulados realizados e esperados com seus intervalos de confiança de 99,73% ($CI\ 3\sigma$) para o modelo de seleção de carteiras baseado em erros de predição (Fig. 6.10a) e para o modelo média-variância (Fig. 6.10b). Conforme mostrado na Fig. 6.10, o desempenho do modelo de seleção de carteiras baseado em erros de predição se distanciou mais do esperado no arcabouço Normal que o desempenho do modelo média-variância no conjunto das 142 semanas de simulação de investimento.

Neste experimento, nós selecionamos carteiras com retornos preditos e riscos maiores que os das carteiras selecionadas no experimento anterior da Seção 6.2.3. Por conta disso, o

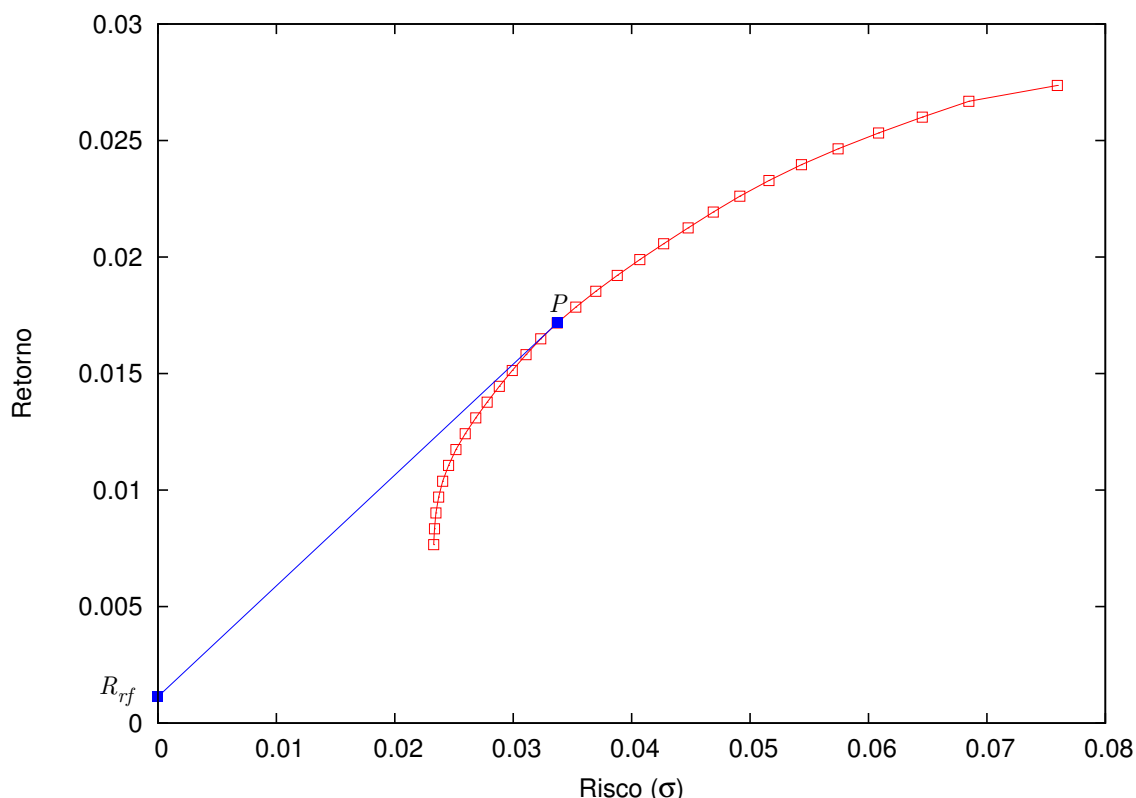


Figura 6.8: Exemplo de fronteira eficiente com ativo sem risco. Todas as combinações da participação no ativo sem risco com retorno esperado R_{rf} e da participação na carteira P , ao longo da tangente que os conecta, são eficientes.

modelo de seleção de carteiras baseado em erros de predição produziu vários retornos preditos com valores acima dos retornos realizados no mercado. Entretanto, estes retornos foram preditos na direção correta (ver o H_{R+} das predições na Tabela 6.2), causando a separação das curvas realizada e esperada observadas em diversos pontos da Fig. 6.10a, especialmente na segunda metade das 142 semanas.

A Tabela 6.6 (pág. 113) sumariza o desempenho semanal dos modelos e do índice IBOVESPA nas 142 semanas do período de simulação de investimento, e está organizada da mesma forma que a Tabela 6.4 do experimento anterior da Seção 6.2.3. Conforme a Tabela 6.6 mostra, o modelo de seleção de carteiras baseado em erros de predição alcançou um retorno acumulado de 3,0786, que foi 81% superior ao retorno acumulado de 1,700 alcançado pelo modelo média-variância e 41% superior ao retorno acumulado de 2,189 do IBOVESPA. O modelo de seleção de carteiras baseado em erros de predição obteve um risco (variância) *ex-post* de 0,00200, que foi 82% superior ao risco de 0,00110 obtido pelo modelo média-variância e o dobro do risco de 0,000998 obtido pelo IBOVESPA.

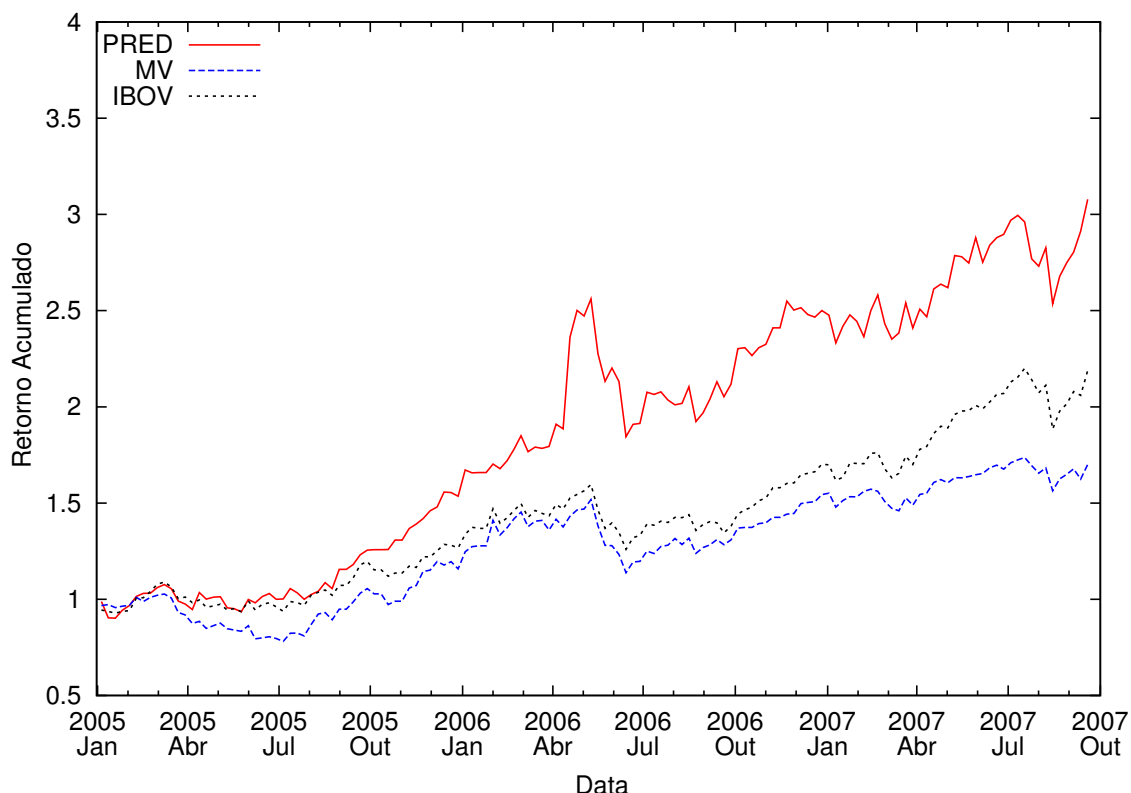


Figura 6.9: Simulação de investimento de risco moderado. Retornos acumulados semanais das 142 carteiras do modelo de seleção de carteiras baseado em erros de predição (PRED) e do modelo média-variância (MV), e do índice IBOVESPA (IBOV). As carteiras foram selecionadas no ponto de tangência das respectivas fronteiras eficientes com o ativo sem risco ($R_{rf} = 0,00114$). O modelo de seleção de carteiras baseado em erros de predição superou o modelo média-variância e o índice IBOVESPA na maioria das semanas.

O modelo de seleção de carteiras baseado em erros de predição alcançou um *PCM* de 0,00134, enquanto o modelo média-variância alcançou um *PCM* de $-0,000645$. Isto mostra que o modelo de seleção de carteiras baseado em erros de predição obteve uma melhor capacidade de antecipação dos movimentos do mercado (*timing* do mercado) que o modelo média-variância, valendo-se do desempenho do conjunto das predições dos retornos futuros das ações (ver Tabela 6.2).

O modelo de seleção de carteiras baseado em erros de predição alcançou um *TI* médio de aproximadamente 176%, enquanto o modelo média-variância alcançou aproximadamente 26%. Isto mostra que o modelo de seleção de carteiras baseado em erros de predição produziu um comportamento mais agressivo que o modelo média-variância, alterando suas participações nas ações de forma mais intensa, na tentativa de antecipar os movimentos de curto prazo do mercado.

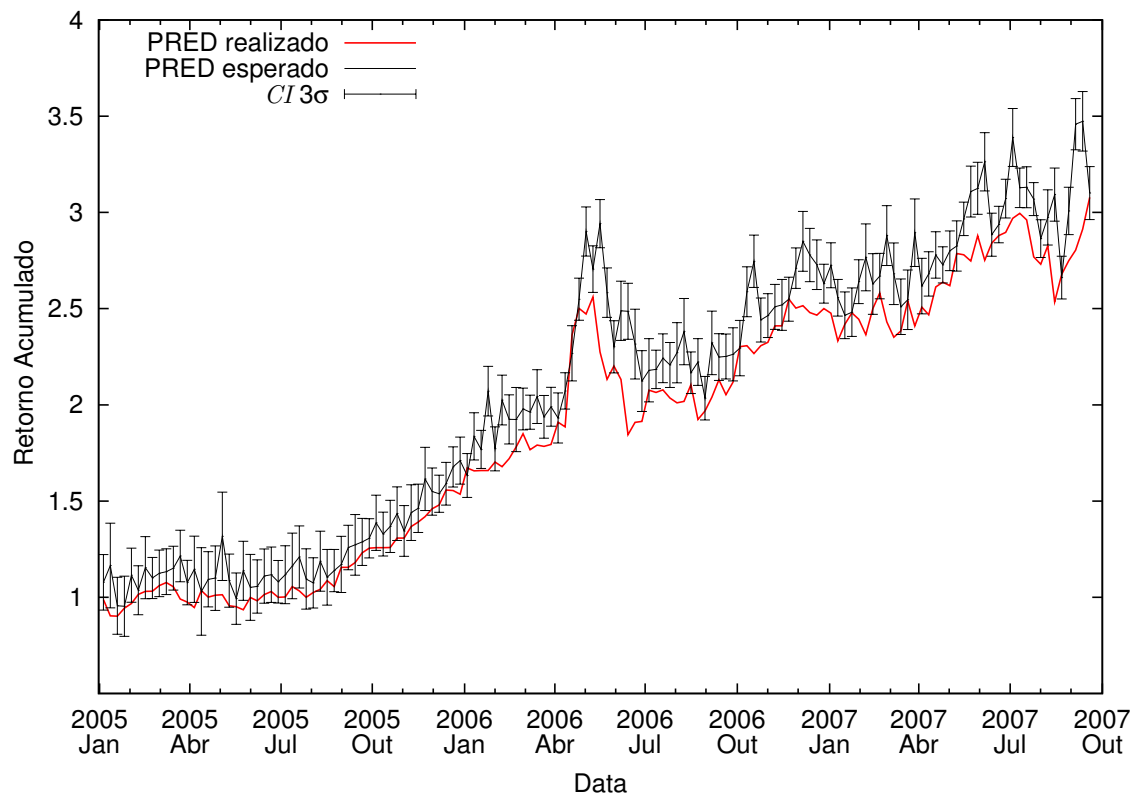
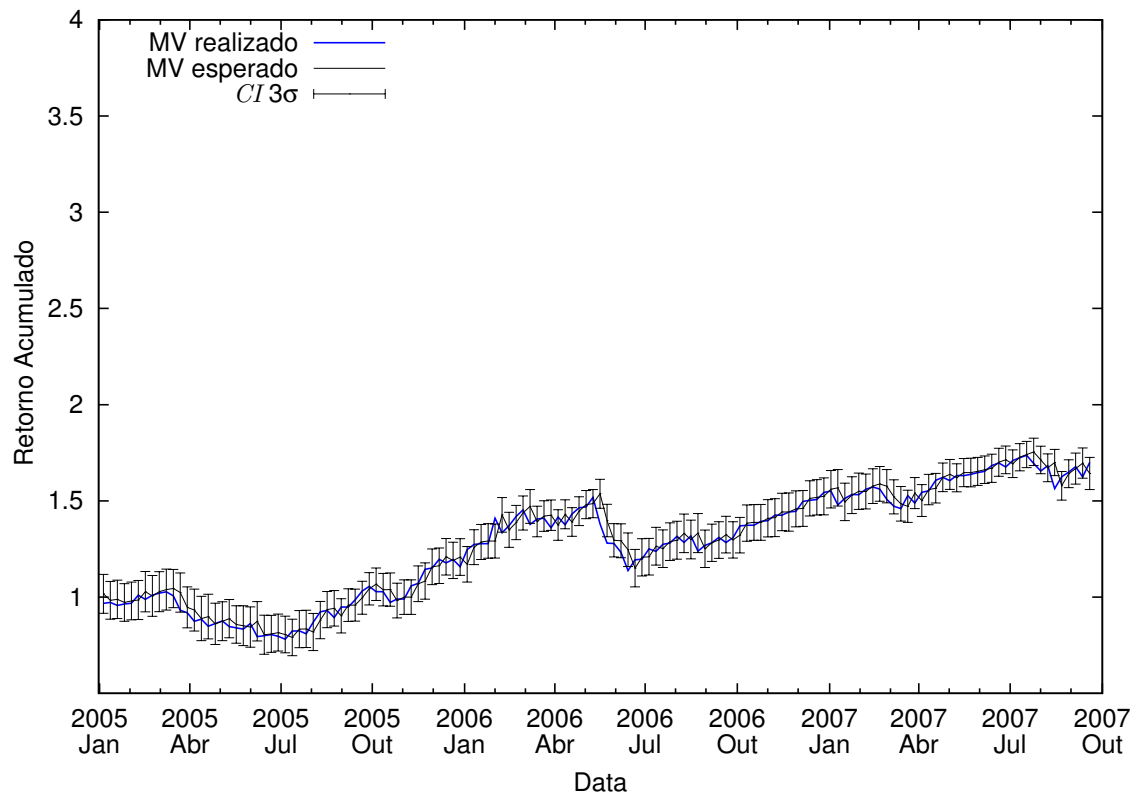
(a) Retornos acumulados e $CI\ 3\sigma$ para o modelo de seleção de carteiras baseado em erros de predição(b) Retornos acumulados e $CI\ 3\sigma$ para o modelo média-variância

Figura 6.10: Simulação de investimento de risco moderado. Retornos acumulados realizados e esperados com seus intervalos de confiança de 99,73% ($CI\ 3\sigma$): (a) modelo de seleção de carteiras baseado em erros de predição e (b) modelo média-variância.

O número de ações das carteiras baseadas em erros de predição variou entre dois e 14, com uma média de seis ações, enquanto que o número de ações das carteiras média-variância variou entre seis e 24, com uma média de 11 ações. O modelo de seleção de carteiras baseado em erros de predição apresentou uma diversificação mais modesta que o modelo média-variância, uma vez que foram selecionadas carteiras com retornos superiores devido à diferença de formato das suas fronteiras eficientes (ver Fig. 6.3, pág. 99), e um menor número de ações foi capaz de oferecer retornos esperados nesses níveis.

Nós realizamos a mesma avaliação dos resultados trimestrais obtidos pelos modelos de seleção de carteiras e pelo IBOVESPA apresentada no experimento anterior da Seção 6.2.3, calculando seus retornos semanais médios no trimestre e seus retornos acumulados trimestrais da mesma forma. A Tabela 6.7 (pág. 114) mostra os intervalos de confiança (*CI*) de 95% dos retornos semanais médios no trimestre para o modelo de seleção de carteiras baseado em erros de predição (PRED), para o modelo média-variância (MV) e para o índice IBOVESPA (IBOV). Conforme mostrado na Tabela 6.7, a média do retorno semanal médio no trimestre de 0,00926 obtida pelo modelo de seleção de carteiras baseado em erros de predição foi aproximadamente o dobro da média de 0,00459 obtida pelo modelo média-variância e 44% acima da média de 0,00643 obtida pelo IBOVESPA. O modelo de seleção de carteiras baseado em erros de predição obteve um erro padrão de 0,00435, que foi 34% menor que o erro padrão de 0,00667 do modelo média-variância e 14% menor que o erro padrão de 0,00509 do IBOVESPA. Observa-se ainda que a média de 0,00459 do retorno semanal médio no trimestre do modelo média-variância ficou abaixo do limite inferior do *CI* 95% de 0,00491 do modelo de seleção de carteiras baseado em erros de predição. A Fig. 6.11 (pág. 114) mostra uma representação gráfica das médias e dos intervalos de confiança da Tabela 6.7, onde é evidenciada superioridade da média do modelo de seleção de carteiras baseado em erros de predição e o seu menor erro padrão (menor risco).

O modelo de seleção de carteiras baseado em erros de predição obteve um retorno acumulado trimestral médio de 1,117 com variância de 0,00988, enquanto que o modelo média-variância obteve um retorno acumulado trimestral médio de 1,0620 com variância de 0,0204. Nós aplicamos o teste *t* pareado na comparação das médias dos modelos no nível de significância de 5% (H_0 : *ARE* médio do modelo de seleção de carteiras baseado em erros de predição \leq *ARE* médio do modelo média-variância) e encontramos um *p-valor* de 0,0211. Este resultado indica que o *ARE* médio do modelo de seleção de carteiras baseado em erros de predição é significativamente maior que *ARE* médio do modelo média-variância.

As variâncias dos retornos acumulados trimestrais dos dois modelos foram testadas com o teste *F* no nível de significância de 5% (H_0 : razão da variância do modelo de seleção de carteiras baseado em erros de predição pela variância do modelo média-variância = 1) e encontramos um *p-valor* de 0,2686. Este resultado indica que não há diferença significativa

Tabela 6.6: Simulação de investimento de risco moderado. Sumário do desempenho semanal dos modelos de seleção de carteiras e do índice IBOVESPA durante as 142 semanas do período de investimento.

Carteiras Baseadas em Erros de Predição				Carteiras Média-Variância			
selecionadas no ponto de tangência das fronteiras eficientes com o retorno esperado do ativo sem risco ($R_{rf} = 0,00114$)				selecionadas no ponto de tangência das fronteiras eficientes com o retorno esperado do ativo sem risco ($R_{rf} = 0,00114$)			
retorno acumulado	3,0786			retorno acumulado	1,700		
retornos semanais	média	σ^2	σ	retornos semanais	média	σ^2	σ
	0,00892	0,00200	0,0448		0,00429	0,00110	0,0332
<i>TI</i>	1,768	0,0720	0,268	<i>TI</i>	0,263	0,0157	0,125
<i>PCM</i>	0,00134			<i>PCM</i>	−0,000645		
número de ações	média	mín.	máx.	número de ações	média	mín.	máx.
	6	2	14		11	6	24

Índice IBOVESPA			
retorno acumulado	2,189		
retornos semanais	média	σ^2	σ
	0,00603	0,000998	0,0316
número de ações	média	mín.	máx.
	57	53	63

Tabela 6.7: Simulação de investimento de risco moderado. Intervalos de confiança (CI) de 95% dos retornos semanais médios no trimestre durante as 142 semanas do período de investimento.

	Média	Erro padrão	Intervalo de confiança da média (95%)
PRED	0,009 26	0,004 35	$0,00491 \leq 0,00926 \leq 0,0136$
MV	0,004 59	0,006 67	$-0,00209 \leq 0,00459 \leq 0,01131$
IBOV	0,006 43	0,005 09	$0,00133 \leq 0,00643 \leq 0,0115$

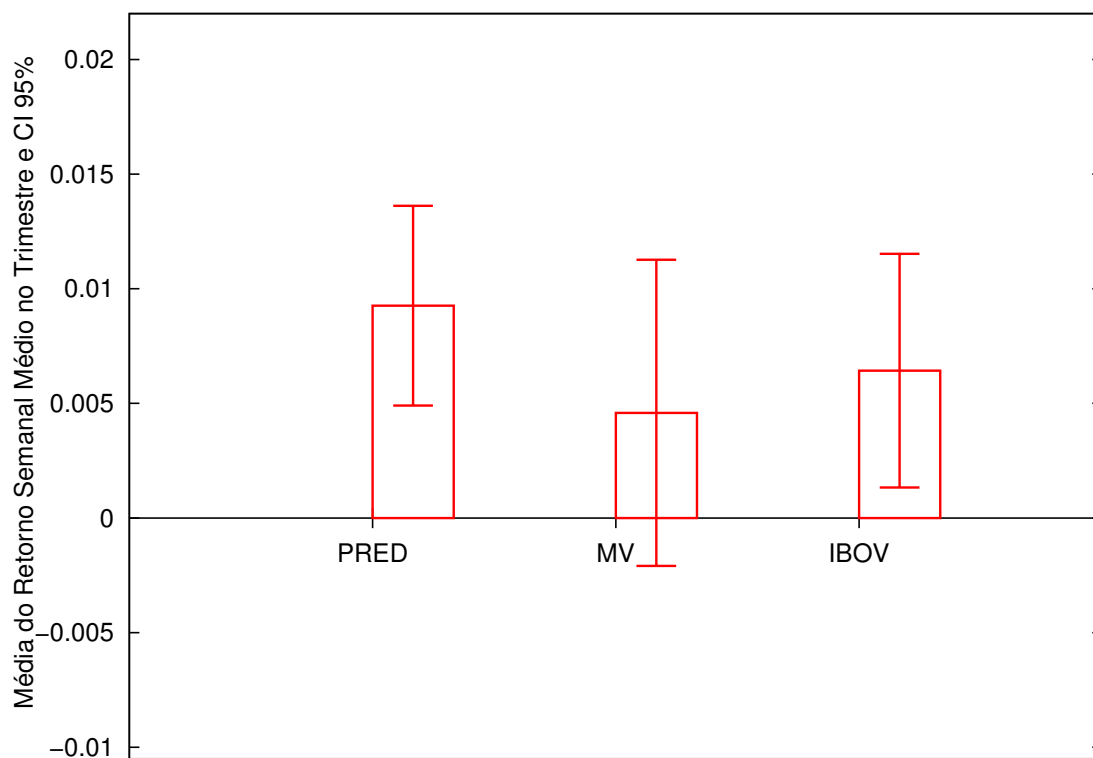


Figura 6.11: Simulação de investimento de risco moderado. Intervalos de confiança (CI) de 95% dos retornos semanais médios no trimestre para o modelo de seleção de carteiras baseado em erros de predição (PRED), para o modelo média-variância (MV) e para o índice IBOVESPA durante as 142 semanas do período de investimento. O menor erro padrão da média dos retornos trimestrais alcançados pelo modelo de seleção de carteiras baseado em erros de predição em relação ao modelo média-variância é evidenciado pelo tamanho das barras verticais PRED e MV.

entre as variâncias (riscos) alcançadas pelo modelo de seleção de carteiras baseado em erros de predição e pelo modelo média-variância.

6.2.5 Simulação de Investimento de Alto Risco

Este experimento avalia o desempenho dos modelos de seleção de carteiras na maximização dos retornos esperados, ou seja, atuando como um *trading system* tradicional. Nessa aplicação, os modelos devem alocar todo o capital disponível somente na ação com o maior retorno esperado — o retorno predito para o modelo de seleção de carteiras baseado em erros de predição, e o retorno médio para o modelo média-variância —, portanto, selecionando as carteiras na extremidade direita das suas fronteiras eficientes³.

A Fig. 6.12 mostra os retornos acumulados semanais obtidos com o modelo de seleção de carteiras baseado em erros de predição (PRED), com o modelo média-variância (MV) e com o índice IBOVESPA (IBOV) durante as mesmas 142 semanas do período de simulação de investimento dos experimentos anteriores, com a mesma estratégia de investimento e os mesmos dados e datas. Conforme mostrado na Fig. 6.12, o modelo de seleção de carteiras baseado em erros de predição fortemente superou o modelo média-variância e o IBOVESPA em quase todas as semanas do período. É importante mencionar que o modelo média-variância obteve um desempenho bastante inferior ao do IBOVESPA no período.

A Fig. 6.13 apresenta a mesma avaliação dos modelos de seleção de carteiras no contexto do arcabouço Normal apresentada nos experimentos anteriores, mostrando os retornos acumulados realizados e esperados com seus intervalos de confiança de 99,73% ($CI\ 3\sigma$), para o modelo de seleção de carteiras baseado em erros de predição (Fig. 6.13a) e para o modelo média-variância (Fig. 6.13b). Conforme mostrado na Fig. 6.13, o desempenho do modelo de seleção de carteiras baseado em erros de predição se distanciou mais do esperado no arcabouço Normal que o modelo média-variância no conjunto das 142 semanas de simulação de investimento. Esse comportamento foi similar ao do experimento anterior da Seção 6.2.4, e foi causado pelos mesmos motivos observados naquela seção.

A Tabela 6.8 (pág. 119) sumariza o desempenho semanal do modelo de seleção de carteiras baseado em erros de predição, do modelo média-variância e do índice IBOVESPA nas 142 semanas do período de simulação de investimento, e está organizada da mesma forma que as Tabelas 6.4 e 6.6 dos experimentos anteriores. Conforme a Tabela 6.8 mostra, o modelo de seleção de carteiras baseado em erros de predição alcançou um retorno acumulado de 3,894, que foi 292% superior ao retorno acumulado de 0,993 alcançado pelo modelo média-variância e 78% superior ao retorno acumulado de 2,189 alcançado pelo IBOVESPA. O modelo de seleção de carteiras baseado em erros de predição obteve um risco (variância) *ex-post* de 0,00383, que foi 21% acima do risco de 0,00316 obtido pelo modelo média-variância e 284% acima do risco de 0,000998 obtido pelo IBOVESPA.

³No caso de haver mais de uma ação com os mesmos valores de retorno esperado e risco, quaisquer combinações destas ações também maximizam o retorno esperado. Contudo, pelo *princípio da parcimônia* (*Occam's razor*), o sistema de otimização deve escolher apenas uma das ações neste caso.

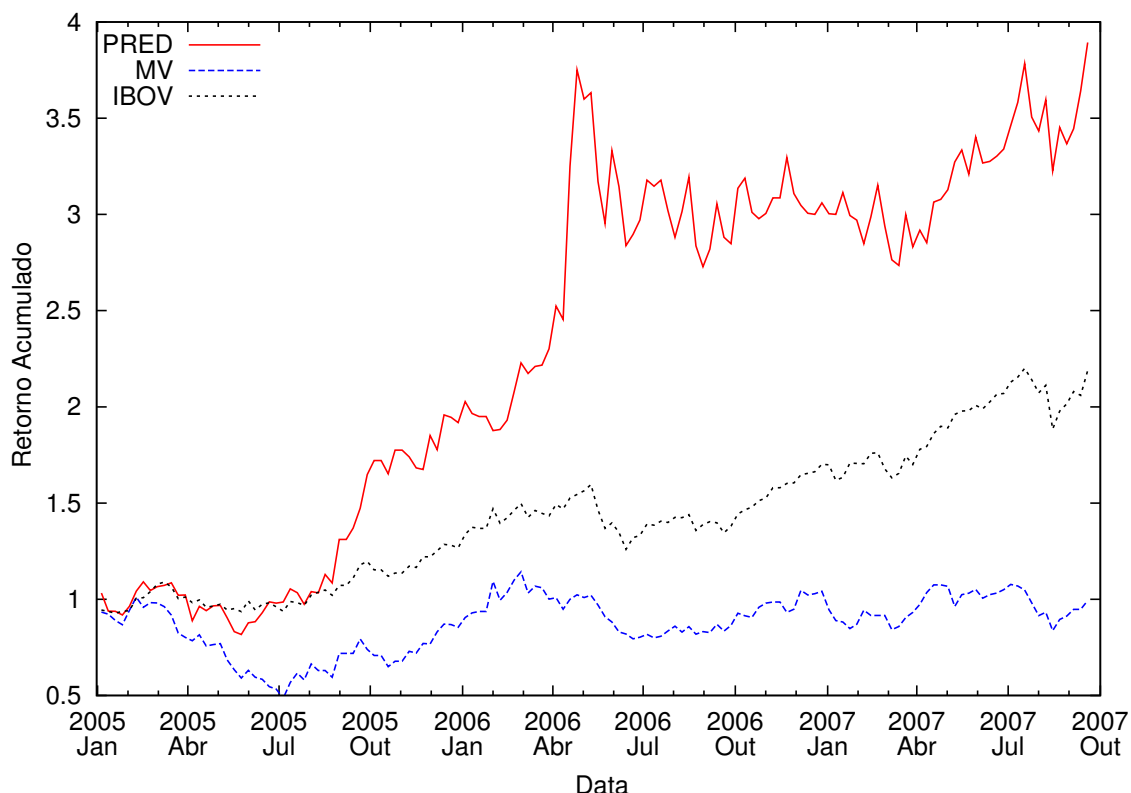


Figura 6.12: Simulação de investimento de alto risco. Retornos acumulados semanais das 142 carteiras do modelo de seleção de carteiras baseado em erros de previsão (PRED) e do modelo média-variância (MV), e do índice IBOVESPA (IBOV). As carteiras foram selecionadas no máximo \hat{r}_p para o modelo de seleção de carteiras baseado em erros de previsão e no máximo \bar{r}_p para o modelo média-variância. O modelo de seleção de carteiras baseado em erros de previsão fortemente superou o modelo média-variância e o índice IBOVESPA em quase todas as semanas.

O modelo de seleção de carteiras baseado em erros de previsão alcançou um *PCM* de 0,00601, enquanto o modelo média-variância alcançou um *PCM* de 0,00284. Isto mostra que o modelo de seleção de carteiras baseado em erros de previsão obteve um *timing* do mercado cerca de duas vezes melhor que o do modelo média-variância, valendo-se do desempenho do conjunto das previsões do retorno futuro das ações (ver Tabela 6.2).

O modelo de seleção de carteiras baseado em erros de previsão alcançou um *TI* médio de aproximadamente 190%, enquanto que o modelo média-variância alcançou aproximadamente 70%. Isto mostra que o modelo de seleção de carteiras baseado em erros de previsão produziu um comportamento extremamente mais agressivo que o modelo média-variância, trocando a ação da sua carteira em quase todas as semanas. Os dois modelos utilizaram somente uma ação nas suas carteiras em todas as 142 semanas.

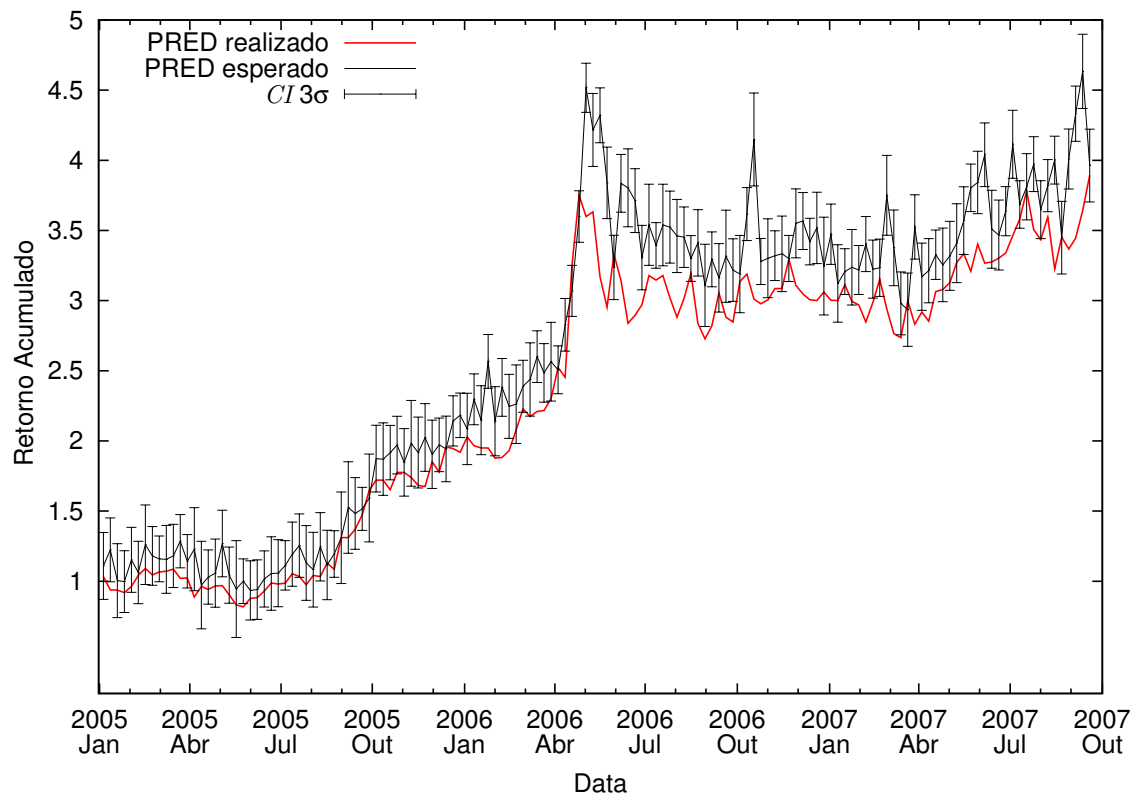
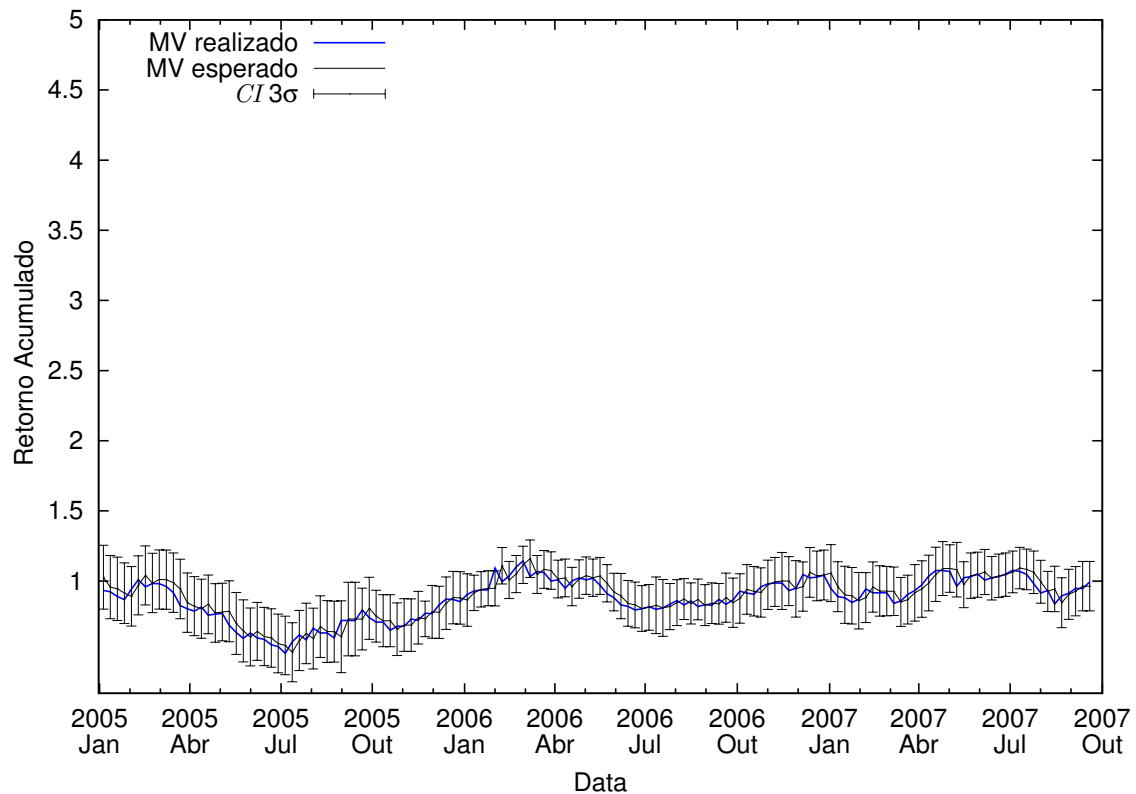
(a) Retornos acumulados e $CI\ 3\sigma$ para o modelo de seleção de carteiras baseado em erros de predição(b) Retornos acumulados e $CI\ 3\sigma$ para o modelo média-variância

Figura 6.13: Simulação de investimento de alto risco. Retornos acumulados realizados e esperados com seus intervalos de confiança de 99,73% ($CI\ 3\sigma$): (a) modelo de seleção de carteiras baseado em erros de predição e (b) modelo média-variância.

Nós realizamos a mesma avaliação dos resultados trimestrais obtidos pelos modelos de seleção de carteiras e pelo IBOVESPA apresentada nos experimentos anteriores, calculando seus retornos semanais médios no trimestre e seus retornos acumulados trimestrais da mesma forma. A Tabela 6.9 (pág. 120) mostra os intervalos de confiança (*CI*) de 95% dos retornos semanais médios no trimestre para o modelo de seleção de carteiras baseado em erros de predição (PRED), para o modelo média-variância (MV) e para o índice IBOVESPA (IBOV). Conforme mostrado na Tabela 6.9, a média do retorno semanal médio no trimestre de 0,0118 obtida pelo modelo de seleção de carteiras baseado em erros de predição foi 548% acima da média de 0,00182 obtida pelo modelo média-variância e 83% acima da média de 0,00643 obtida pelo IBOVESPA. O modelo de seleção de carteiras baseado em erros de predição obteve um erro padrão de 0,00906, que foi 19% menor que o erro padrão de 0,0112 do modelo média-variância e 35% menor que o erro padrão de 0,00509 do IBOVESPA. Observa-se ainda que a média de 0,00182 do retorno semanal médio no trimestre do modelo média-variância ficou abaixo do limite inferior do *CI* 95% de 0,00278 do modelo de seleção de carteiras baseado em erros de predição. A Fig. 6.14 (pág. 120) mostra uma representação gráfica das médias e dos intervalos de confiança da Tabela 6.9, onde é evidenciada superioridade da média do modelo de seleção de carteiras baseado em erros de predição.

O modelo de seleção de carteiras baseado em erros de predição obteve um retorno acumulado trimestral médio de 1,155 com variância de 0,0460, enquanto que o modelo média-variância obteve um retorno acumulado trimestral médio de 1,0250 com variância de 0,0462. Nós aplicamos o *teste t* pareado na comparação das médias dos modelos no nível de significância de 5% (H_0 : *ARE* médio do modelo de seleção de carteiras baseado em erros de predição \leq *ARE* médio do modelo média-variância) e encontramos um *p-valor* de 0,0256. Este resultado indica que o *ARE* médio do modelo de seleção de carteiras baseado em erros de predição é significativamente maior que *ARE* médio do modelo média-variância.

As variâncias dos retornos acumulados trimestrais dos dois modelos foram testadas com o *teste F* no nível de significância de 5% (H_0 : razão da variância do modelo de seleção de carteiras baseado em erros de predição pela variância do modelo média-variância = 1) e encontramos um *p-valor* de 0,995. Este resultado indica que não há diferença significativa entre as variâncias (riscos) alcançadas pelo modelo de seleção de carteiras baseado em erros de predição e pelo modelo média-variância.

Tabela 6.8: Simulação de investimento de alto risco. Sumário do desempenho semanal dos modelos de seleção de carteiras e do índice IBOVESPA durante as 142 semanas do período de investimento.

Carteiras Baseadas em Erros de Predição selecionadas no máximo \hat{r}_p				Carteiras Média-Variância selecionadas no máximo \bar{r}_p			
retorno acumulado		3,894		retorno acumulado		0,993	
retornos semanais <i>TI</i>	média	σ^2		retornos semanais <i>TI</i>	média	σ^2	
	0,0114 1,901	0,00383 0,1901	0,0619 0,436		0,00149 0,454	0,00316 0,707	0,0562 0,841
<i>PCM</i>		0,00601		<i>PCM</i>		0,00284	
número de ações	média	mín.		número de ações	média	mín.	
	1	1	1		1	1	1

Índice IBOVESPA			
retorno acumulado		2,189	
retornos semanais	média	σ^2	
	0,00603	0,000998	
número de ações	média	mín.	
	57	53	63

Tabela 6.9: Simulação de investimento de alto risco. Intervalos de confiança (CI) de 95% dos retornos semanais médios no trimestre durante as 142 semanas do período de investimento.

	Média	Erro padrão	Intervalo de confiança da média (95%)
PRED	0,0118	0,00906	$0,00278 \leq 0,01180 \leq 0,0209$
MV	0,00182	0,0112	$-0,00937 \leq 0,00182 \leq 0,0130$
IBOV	0,00643	0,00509	$0,00133 \leq 0,00643 \leq 0,0115$

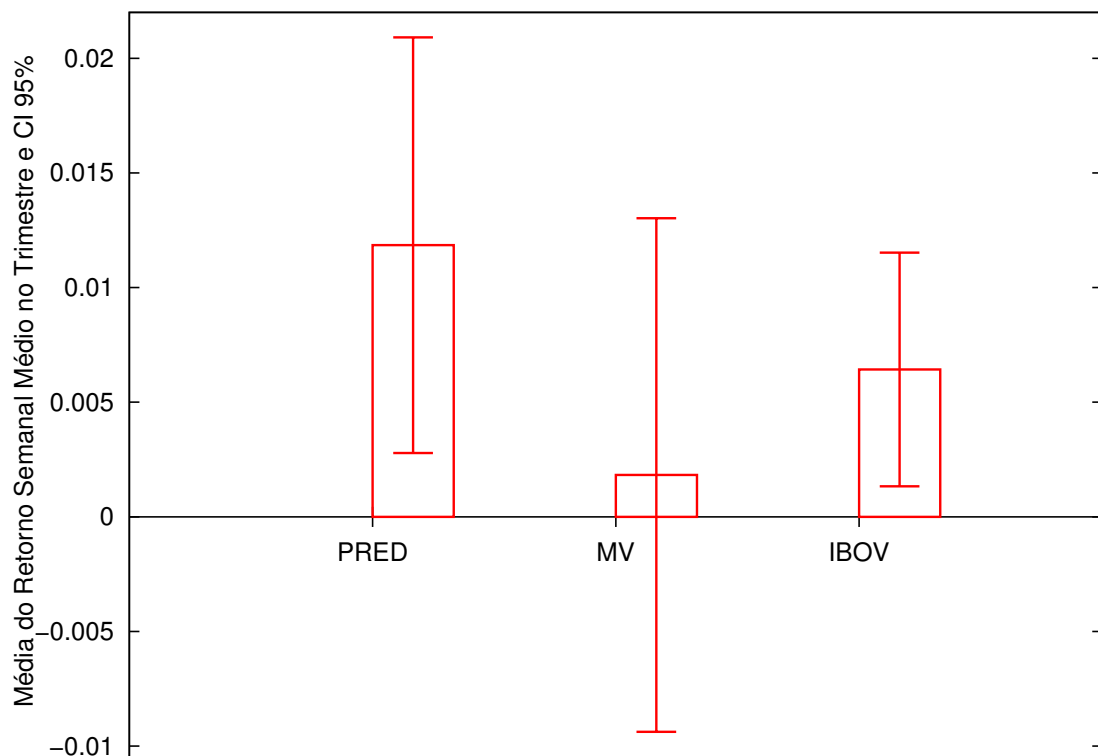


Figura 6.14: Simulação de investimento de alto risco. Intervalos de confiança (CI) de 95% dos retornos semanais médios no trimestre para o modelo de seleção de carteiras baseado em erros de predição (PRED), para o modelo média-variância (MV) e para o índice IBOVESPA durante as 142 semanas do período de investimento. O menor erro padrão da média dos retornos trimestrais alcançados pelo modelo de seleção de carteiras baseado em erros de predição em relação ao modelo média-variância é evidenciado pelo tamanho das barras verticais PRED e MV.

6.3 Sumário dos Experimentos

Os experimentos com dados artificiais da Seção 6.1 demonstraram como o modelo de seleção de carteiras baseado em erros de predição tira proveito dos padrões de previsibilidade das séries de retornos das ações.

Nos experimentos com dados reais, a avaliação do desempenho dos preditores apresentada na Seção 6.2.1 validou a hipótese subjacente da Eq. 4.7 ($\bar{\epsilon} = 0$) do modelo de seleção de carteiras baseado em erros de predição e mostrou que 61% das predições dos retornos positivos foram preditos com o sinal correto. As fronteiras eficientes calculadas na Seção 6.2.2 mostraram que o modelo de seleção de carteiras baseado em erros de predição produziu um menor risco associado ao retorno esperado que o modelo média-variância. Os experimentos de investimento de baixo risco da Seção 6.2.3 mostraram que o modelo de seleção de carteiras baseado em erros de predição superou o desempenho semanal do modelo média-variância com um risco 57% menor e acompanhou o desempenho semanal do índice IBOVESPA com um risco 25% menor. Os experimentos de investimento de risco moderado da Seção 6.2.4 mostraram que o modelo de seleção de carteiras baseado em erros de predição superou o desempenho trimestral do modelo média-variância e obteve um retorno semanal acumulado 81% superior ao do modelo média-variância e 41% superior ao do índice IBOVESPA. Os experimentos de investimento de alto risco da Seção 6.2.5 mostraram que o modelo de seleção de carteiras baseado em erros de predição superou o desempenho trimestral do modelo média-variância e obteve um retorno semanal acumulado 292% superior ao do modelo média-variância e 78% superior ao do índice IBOVESPA.

O principal argumento que levou à hipótese investigada nesta Tese (ver Seção 1.2) é que apesar dos preditores individuais (de uma ação) apresentarem desempenho modesto, eles podem ser combinados de forma a explorar as complementaridades no comportamento dos seus erros de predição, levando a bons resultados para as combinações lineares destas predições. Nós verificamos a validade deste argumento através da comparação dos *RMSE* (ver Seção 5.2) das diferenças entre o retorno realizado e o retorno predito das carteiras baseadas em erros de predição do experimento de baixo risco da Seção 6.2.3 e do experimento de alto risco da Seção 6.2.5. No experimento de alto risco, foram selecionadas carteiras com apenas uma ação em todas as 142 semanas. Desta forma, o desempenho das carteiras deste experimento foi exposto unicamente aos desempenhos individuais dos preditores, uma vez que não houve combinação de preditores neste experimento, e o *RMSE* obtido para as 142 semanas de simulação de investimento foi de 0,0805. Já no experimento de baixo risco, as carteiras foram selecionadas combinando 21 preditores em média (número de ações na carteira) durante as 142 semanas. Desta forma, o desempenho das carteiras deste experimento foi exposto ao desempenho da combinação dos preditores, e o *RMSE* obtido para as

142 semanas de simulação de investimento foi de 0,0176, portanto, 78% menor que o *RMSE* dos preditores individuais do experimento de alto risco.

O *RMSE* equivale à raiz quadrada do *risco da predição* de Moody [114]. Tomando então os quadrados dos *RMSE* observados acima (equivalentes às variâncias dos erros de predição, ou o risco no nosso modelo de seleção de carteiras baseado em erros de predição), temos uma redução de 95% do risco da predição devido ao efeito da exploração das complementaridades dos erros de predição, verificando desta forma o argumento principal da hipótese central desta Tese.

Capítulo 7

Discussão

A seleção de carteiras de investimento ainda é uma questão aberta na teoria das finanças. Temas como o comportamento dos investidores, a utilidade esperada e arcabouços probabilísticos, modelos dinâmicos, modelagem do risco e distribuições de retornos, fricções do mercado, modelos preditivos e modelos robustos, apenas para citar alguns, ainda são foco de intensa atividade de pesquisa [3, 9, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130]. Neste capítulo, discutiremos este trabalho de pesquisa à luz da previsibilidade dos mercados e das melhorias a serem incorporadas ao modelo de seleção de carteiras baseado em erros de predição para acomodar as fricções do mercado, bem como apresentaremos uma análise crítica deste trabalho.

7.1 Previsibilidade dos Mercados

O interesse na exploração de modelos de predição na seleção de carteiras não é recente, porém, foi tradicionalmente focado na difícil tarefa de determinar as distribuições *a priori* das séries de retornos das ações [57]. A predição do retorno futuro de ações é reconhecida como uma tarefa difícil [33, 41, 42] e, nos anos recentes, técnicas de inteligência artificial e aprendizado de máquina também têm sido aplicadas na predição de outras medidas, tais como sinais de compra e venda de ações [30, 70] e posição em *ranking* de ações [65]. Os métodos de inteligência computacional são mais adequados às aplicações que examinam um grande número de ações do que a determinação das distribuições das séries de retornos ou a identificação e estimação de modelos estocásticos para predição de séries temporais.

O desempenho de preditores baseados em redes neurais artificiais na predição dos retornos futuros de ações depende de vários aspectos, tais como a topologia da rede, o método

de treinamento e o tratamento dos conjuntos de dados disponíveis [114]. A idéia central deste trabalho é que, apesar dos preditores do retorno futuro de ações não alcançarem bons desempenhos individualmente, eles podem ser combinados de forma a produzir resultados superiores quando utilizados no investimento em várias ações. Os resultados experimentais mostrados no Capítulo 6 são uma evidência da validade dessa hipótese.

Neste trabalho, nós utilizamos arquiteturas padrão de preditores neurais auto-regressivos com referência móvel, com topologia e parâmetros de treinamento fixos para todas as ações (ver Seções 4.1 e 5.3.2). As séries de retornos de ações são não-estacionárias e não-lineares [41], apresentando características individuais bastante distintas (ver Fig. 3.2), e, portanto, devem ser modeladas individualmente de forma a otimizar o modelo auto-regressivo implementado para cada ação (ver Fig. 3.8 e Eq. 3.23). Contudo, apesar da utilização de topologia e parâmetros de treinamento fixos nos preditores prejudicar a otimalidade das predições, nós optamos por esta estratégia para investigar as hipóteses da Seção 1.2, na mesma linha de investigação das aplicações de *combinação de preditores* [43, 69].

7.2 Fricções do Mercado

Os modelos de seleção de carteiras apresentados neste trabalho e suas aplicações na simulação de investimentos tiveram como base as premissas padrão da literatura que foram apresentadas na Seção 5.1 [2, 3, 4], especialmente considerando que:

- As ações são perfeitamente divisíveis, de forma que é possível comprar as ações de qualquer carteira selecionada nas proporções das suas participações.
- Não existem fricções (custos de transação, taxação, comissões, etc.) quando da troca de posição em qualquer ação da carteira.

Contudo, os problemas reais de investimento, especialmente aqueles de interesse do investidor corporativo, desafiam estas premissas clássicas de diversas formas: (i) a indivisibilidade dos lotes-padrão das ações pode dificultar a implementação das soluções ótimas dos modelos, prejudicando a otimalidade da estratégia de investimento (a participação X_i é um número real positivo, e a fração do capital total W a ser investido na ação i , $X_i W$, pode não ser um múltiplo inteiro do preço do lote-padrão da ação)¹; (ii) os custos de transação e outras taxas e encargos podem consumir os retornos realizados; e (iii) podem ser demandadas restrições

¹Muito embora seja possível comprar apenas uma ação no *mercado fracionário*, essas dificuldades se aplicam mesmo nessa situação, em função do capital disponível e do valor do preço das ações em si.

estratégicas que tenham por objetivo estabelecer cotas de participação das ações nas carteiras, como, por exemplo, no controle acionário de companhias e no cumprimento de metas de participação setorial. Nos últimos anos, aspectos desta natureza passaram a ser tratados de forma conjunta no corpo de conhecimento das *fricções do mercado* (*market frictions*) e vistos não somente como restrições operacionais, mas também como oportunidades de lucro [131, 132].

7.2.1 Custos Operacionais

Diversos custos financeiros incidem sobre o investimento em ações, tais como os custos operacionais cobrados pelas bolsas de valores, as taxas de corretagem cobradas pelas corretoras que implementam as ordens de compra e venda das ações e os impostos. O modelo de cobrança destes custos depende da regulamentação do mercado financeiro e bolsas de valores e também das políticas tributárias vigentes. No mercado brasileiro, estes custos tipicamente são percentuais cobrados sobre o volume financeiro (isto é, a soma das compras e vendas realizadas no dia), como no caso das taxas operacionais e de corretagem, ou sobre o lucro realizado, como no caso dos impostos e das taxas de *performance* cobradas pelas corretoras. As tabelas das alíquotas aplicadas tipicamente são regressivas com percentuais variando em função do volume, como no caso dos custos de transação e taxas de corretagem [133], e, mais recentemente, em função do horizonte do investimento, como no caso de alguns impostos [134]. Adicionalmente, algumas destas taxas têm níveis de limiar (onde há um valor mínimo cobrado) ou valores fixos sobre as transações, e, ainda, alguns impostos incluem incentivos e regimes de compensação de prejuízo.

A diversidade metodológica do cálculo dos custos operacionais acima apresentados torna a sua modelagem exata de difícil tratamento, principalmente devido à sua não-linearidade e dependência dos preços das ações e do horizonte de investimento. Entretanto, o tratamento dos custos operacionais no modelo de seleção de carteiras baseado em erros de predição pode ser obtido de forma aproximada. Para isso, podemos definir o *retorno predito líquido* da ação i descontando o seu custo operacional do seu retorno predito como:

$$\hat{r}_i^L = \hat{r}_i - \frac{C(X_i W, I)}{X_i W}, \quad (7.1)$$

onde \hat{r}_i^L é o retorno predito líquido da ação i , \hat{r}_i é o retorno predito da ação i , $X_i W$ é o volume financeiro a ser investido na ação i , ou seja, a participação X_i do capital total W , e $C(X_i W, I)$ é a função de custo financeiro do investimento na ação i , que depende do volume financeiro, $X_i W$, e do conjunto de informações acerca dos custos operacionais, I . Desta forma, o modelo de seleção de carteiras baseado em erros de predição pode incorporar o tratamento dos custos

operacionais através da utilização do retorno predito líquido, \hat{r}_i^L , como o retorno esperado da ação i .

7.2.2 Indivisibilidade dos Lotes-padrão das Ações

Conforme apresentado na Seção 3.1, as ações tipicamente são comercializadas em lotes-padrão que agrupam, na maioria das vezes, 1.000 ações. Uma vez que a participação X_i da ação i na carteira é um número real positivo, a fração do capital total W a ser investido na ação i , $X_i W$, pode não ser um múltiplo inteiro do preço do lote-padrão da ação. Com isso, a factibilidade das soluções ótimas do modelo de seleção de carteiras baseado em erros de predição pode ser comprometida, prejudicando a otimalidade da estratégia de investimento. Este efeito tende a ser minimizado com o aumento do capital disponível para o investimento, W , porém, ele pode ser bastante significativo no caso contrário, especialmente quando os lotes-padrão das ações de interesse atingem valores muito elevados.

Um melhoramento do modelo de seleção de carteiras baseado em erros de predição deve considerar apenas quantidades inteiras, Q_i , dos lotes-padrão da ação i nas suas soluções. Assim, uma vez que Q_i lotes-padrão da ação i com preço P_i podem ser comprados com o volume financeiro $X_i W$, temos que:

$$Q_i = X_i \frac{W}{P_i}, \quad (7.2)$$

onde Q_i é o número de lotes-padrão da ação i a ser comprado, X_i é a participação da ação i na carteira, P_i é o preço do lote-padrão da ação i e W é o capital total disponível.

Desta forma, o tratamento da indivisibilidade dos lotes-padrão das ações pode ser incorporado ao modelo de seleção de carteiras baseado em erros de predição através da inserção da seguinte restrição adicional no modelo da Seção 4.4:

$$X_i \frac{W}{P_i} \in \mathbb{N}, \quad i = 1, \dots, M. \quad (7.3)$$

Conforme as Eqs. 7.2 e 7.3 mostram, restringir as soluções do modelo de otimização da Seção 4.4 somente àquelas que tornam o lado esquerdo da Eq. 7.3 um inteiro positivo garante que $X_i W$ seja um múltiplo inteiro do preço do lote-padrão da ação.

7.2.3 Modelo Estendido

Inserindo os tratamentos das Seções 7.2.1 e 7.2.2 acima no modelo de seleção de carteiras baseado em erros de predição da Seção 4.4, podemos formular um modelo de otimização

inteiro-quadrático misto com restrições não-lineares da forma:

Minimize

$$\hat{V} = \sum_{i=1}^M X_i^2 \hat{v}_i + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M X_i X_j \hat{v}_{ij}, \quad (7.4)$$

Sujeito a

$$\sum_{i=1}^M \left(X_i \hat{r}_i - \frac{C(X_i W, I)}{W} \right) = R_d, \quad (7.5)$$

$$\sum_{i=1}^M X_i = 1, \quad (7.6)$$

$$X_i \frac{W}{P_i} \in \mathbb{N}, \quad i = 1, \dots, M. \quad (7.7)$$

$$X_i \geq 0, \quad i = 1, \dots, M, \quad (7.8)$$

onde a Eq. 7.4 é a função objetivo a ser minimizada, ou seja, o risco da carteira baseada em erros de predição, \hat{V} , a Eq. 7.5 é a restrição que garante o retorno líquido de carteira desejado, R_d , considerando a função de custo financeiro do investimento na ação i , $C(X_i W, I)$, a Eq. 7.6 é a restrição que garante a integralidade da carteira com a alocação total dos recursos disponíveis, a Eq. 7.7 é a restrição que garante soluções com $X_i W$ sendo um múltiplo inteiro do preço do lote-padrão da ação i , P_i , e a Eq. 7.8 restringe o modelo para participações positivas apenas.

A resolução deste modelo de otimização inteiro-quadrático misto com restrições não-lineares é intratável pelos métodos de otimização tradicionais. Portanto, este problema deve ser atacado através da investigação de novas técnicas baseadas em metaheurísticas como *algoritmos genéticos*, *simulated annealing* e *particle swarm optimization* [135].

7.3 Análise Crítica deste Trabalho de Pesquisa

O ponto focal deste trabalho é a idéia de que apesar dos preditores não apresentarem um bom desempenho individual na predição dos retornos futuros de um conjunto de ações, eles podem ser sistematicamente combinados de forma a explorar a complementaridade dos seus desempenhos individuais, produzindo boas predições para o seu agregado — o retorno da carteira baseada em erros de predição. No modelo de seleção de carteiras baseado em erros de predição, o desempenho individual dos preditores é mensurado de forma similar ao *risco da predição* proposto por Moody [114], que, sob a hipótese da Eq. 4.7 do modelo ($\bar{\epsilon} = 0$), equivale à variância dos erros de predição da Eq. 4.8. O modelo explora a complementari-

dade dos desempenhos individuais dos preditores a partir das covariâncias dos seus erros de predição, atribuindo participações maiores aos preditores com menor variância dos erros de predição e aos pares de preditores com covariância negativa (ver Eq. 4.12).

Um objetivo adicional na proposição do modelo de seleção de carteiras baseado em erros de predição foi manter a aderência ao arcabouço Normal multivariado e ao modelo de programação quadrática, uma vez estes são os métodos mais praticados na indústria [122]. A verificação experimental da hipótese dos $\bar{\epsilon} = 0$ da Eq. 4.7 (ver Seção 6.2.1) e da hipótese da Normalidade dos erros de predição (ver Seção 6.2.1) contribuiu para este objetivo.

Através da exploração de aspectos preditivos das séries de retornos das ações, o modelo de seleção de carteiras baseado em erros de predição pode ser mais adequado ao investimento de curto prazo e à gerência ativa de carteiras do que o modelo média-variância. Os resultados experimentais apresentados no Capítulo 6 são evidência da validade desta hipótese. Contudo, vale comentar alguns aspectos em que este trabalho pode ser expandido e melhorado, no sentido da ampliação das situações de comparação e da melhoria dos métodos e resultados apresentados nos Capítulos 5 e 6, conforme a seguir.

A utilização de preditores neurais auto-regressivos com as mesmas topologias e parâmetros de treinamento para todas as séries de retornos das ações foi aderente ao objetivo da verificação das hipóteses da Seção 1.2. Porém, a avaliação do modelo de seleção de carteiras baseado em erros de predição com resultados de predição superiores, obtidos por exemplo através da otimização dos preditores para cada ação analisada, pode permitir a comparação do ganho de desempenho relativo oferecido pelo modelo em função de níveis de erros de predição distintos, e possibilitar a avaliação do grau de dependência do modelo com a qualidade dos preditores.

A restrição do universo de ações para somente aquelas participantes do índice IBOVESPA no período de simulação de investimento teve o objetivo de propiciar um “jogo justo” na avaliação do desempenho do modelo de seleção de carteiras baseado em erros de predição em relação ao índice. Isto é, verificar se o modelo de seleção de carteiras baseado em erros de predição seria capaz de superar o desempenho do IBOVESPA, em termos de retorno e risco, otimizando as participações nas mesmas ações do índice segundo seus próprios critérios e de forma isolada das demais oportunidades de investimento do mercado. Os resultados experimentais das Seções 6.2.3, 6.2.4 e 6.2.5 verificaram esta capacidade. Contudo, a utilização do modelo com um universo de ações expandido (ou até mesmo excludente daquele do IBOVESPA) pode permitir a verificação do comportamento dos ganhos de desempenho obtidos, em termos de retorno e risco, em função da busca num espaço ampliado de ações e, conseqüentemente, poderíamos quantificar o ganho exploratório do modelo em função do universo de ações.

A Seção 7.2 deste capítulo discutiu o tratamento das fricções do mercado no modelo de seleção de carteiras baseado em erros de predição, apresentando um modelo estendido para esta finalidade. A inclusão deste tratamento no modelo pode permitir a quantificação e a comparação dos seus resultados em diversas situações de relaxamento das premissas subjacentes da Seção 5.1. Avaliações dessa natureza são uma contribuição bastante significativa para o estudo da aplicação do modelo de seleção de carteiras baseado em erros de predição em situações reais de investimento.

O Capítulo 2 apresentou a revisão de literatura desta Tese, descrevendo trabalhos correlatos aos temas deste trabalho de pesquisa. O modelo de seleção de carteiras baseado em erros de predição é aqui comparado àqueles trabalhos apresentados que exploram predição e diversificação como estratégia para o investimento no curto prazo como a seguir.

O sistema proposto por Lazo, Vellasco e Pacheco [68] utiliza predição de retornos e diversificação eficiente, porém, as suas medidas de risco foram derivadas das séries de retornos somente (na estimação dos parâmetros do modelo GARCH e na obtenção das covariâncias dos retornos das ações), sem levar em consideração os desempenhos dos preditores de cada ação, e, portanto, não incorporaram essa informação na medida de risco das ações. O modelo de seleção de carteiras baseado em erros de predição trata este problema derivando suas medidas de risco a partir do desempenho individual e coletivo dos preditores dos retornos futuros das ações, utilizando as variâncias e covariâncias das suas séries de erros de predição.

O *trading system* EASLD proposto por Hung, Cheung e Xu [70] difere fundamentalmente do modelo de seleção de carteiras baseado em erros de predição de duas formas: (i) o EASLD utiliza redes neurais na predição dos sinais de compra e venda das ações (que incluem as participações das ações na carteira) baseado nos preços e sinais de compra e venda passados, enquanto que o modelo de seleção de carteiras baseado em erros de predição utiliza as redes neurais para prever os retornos futuros das ações e empregá-los (juntamente com seus riscos associados) na otimização das carteiras; (ii) o EASLD emprega a diversificação para resolver as deficiências dos *trading systems* relacionadas à utilização de um único ativo e à ausência de controle de risco, comentadas no início da Seção 2.2. Isto é feito treinando as redes neurais diretamente nas carteiras ótimas obtidas com o modelo média-variância e com um modelo desenvolvido pelos autores, que combina o *Sharpe ratio* [4] e o *risco downside* [73] — este modelo também utiliza os retornos médios como retornos esperados das ações. Porém, não é garantido que as carteiras preditivas resultantes sejam ótimas, uma vez que as suas participações são computadas diretamente pelas redes neurais. O modelo de seleção de carteiras baseado em erros de predição, por sua vez, garante a otimalidade das suas carteiras preditivas selecionando sempre carteiras eficientes em todos os períodos do investimento.

Capítulo 8

Conclusão

8.1 Sumário

Neste trabalho, nós propusemos, implementamos e avaliamos um novo preditor para o retorno futuro de ações denominado *rede neural auto-regressiva com referência móvel* (RNAR-RM) e um novo modelo de otimização de carteiras de ações denominado *modelo de seleção de carteiras baseado em erros de predição*.

Para cada predição, o preditor RNAR-RM utiliza variáveis de regressão que são as diferenças entre os valores da série histórica de retornos e um determinado retorno passado, que é utilizado como referência para a predição do retorno futuro. A Normalidade dos erros de predição produzidos pelos preditores RNAR-RM foi verificada em 92% das ações utilizadas, dentre as quais 90% apresentaram séries de retornos Normais. Isto mostrou que é possível obter erros de predição Normais a partir de séries de retornos não-Normais, e suportou o desenvolvimento do nosso modelo de seleção de carteiras baseado em erros de predição dentro do arcabouço Normal multivariado.

Simulações de investimento com o modelo de seleção de carteiras baseado em erros de predição alcançaram um retorno 292% superior ao retorno do modelo média-variância de Markowitz, com variância apenas 21% superior. Além disso, com carteiras selecionadas especificamente para oferecer baixos níveis de risco, o nosso modelo conseguiu acompanhar o desempenho do índice de mercado IBOVESPA com uma variância (risco) 25% menor.

8.2 Conclusões

A seleção de investimentos é um problema fundamental da área de finanças e de grande relevância tanto na academia quanto na indústria. Os modelos clássicos de seleção de carteiras, que utilizam os retornos médios das ações como estimador do retorno esperado, são incapazes de tirar proveito das oportunidades de investimento de curto prazo através da detecção de padrões de previsibilidade nas séries de retornos das ações. Isto abre terreno para a proposição de novos modelos preditivos para a seleção de carteiras que aliem métodos preditivos à diversificação eficiente.

Nesse contexto, nós formulamos a hipótese de que é possível obter retornos superiores aos do modelo média-variância através de um modelo de seleção de carteiras com função de risco baseada nos erros de predição de preditores neurais auto-regressivos, dentro do arcabouço Normal multivariado. Para tal, nós propusemos o *modelo de seleção de carteiras baseado em erros de predição*, que utiliza o retorno predito por preditores RNAR-RM como estimador do retorno esperado das ações e as variâncias e covariâncias das séries dos erros de predição como medida de risco das ações.

Nós conduzimos um grande número de experimentos com dados reais do mercado de ações brasileiro, por meio dos quais simulamos investimentos com o modelo de seleção de carteiras baseado em erros de predição e com o modelo média-variância. Nossos resultados experimentais validaram a hipótese formulada.

8.3 Trabalhos Futuros

Nossos trabalhos futuros incluem:

- A investigação de melhores preditores neurais auto-regressivos e métodos de treinamento, incluindo:
 - A otimização da topologia e demais parâmetros estruturais e de treinamento das redes neurais, abrangendo a utilização de entradas adicionais relevantes, a determinação das defasagens temporais (*time lags*) e a obtenção de parâmetros de treinamento, de forma individualizada para cada série de retornos das ações [60, 61, 63, 136] — uma estratégia promissora é atacar estes problemas através de métodos evolutivos [137];
 - A implementação de métodos de regularização para evitar a saturação dos pesos dos neurônios e melhorar a generalização do preditor [88];

- O desenvolvimento de métodos de treinamento mais eficientes, explorando as possibilidades do treinamento incremental;
 - A utilização do modelo de seleção de carteiras baseado em erros de predição em uma aplicação de combinação de preditores (*pool of forecasts*) [69], explorando diversos métodos de predição conjuntamente para produzir predições mais precisas para os retornos futuros das ações;
 - A implementação de algoritmos de treinamento paralelos utilizando processadores gráficos (*Graphics Processing Units* — GPUs) [138];
 - A ampliação da comparação do desempenho dos preditores RNAR-RM com outros preditores e métodos de predição descritos na literatura, além daqueles reportados em Freitas, De Souza e Almeida [33, 34].
- O refinamento do modelo de seleção de carteiras baseado em erros de predição para incorporar as fricções do mercado e produzir estratégias de investimento mais aderentes aos problemas reais de investimento. Neste sentido, nós investigaremos o uso de meta-heurísticas para a resolução do problema inteiro-quadrático misto com restrições não-lineares formulado pelo modelo de seleção de carteiras baseado em erros de predição estendido (ver Seção 7.2.3) [135].
 - A comparação do desempenho do modelo de seleção de carteiras baseado em erros de predição com outros modelos de seleção de carteiras além do modelo média-variância, incluindo outras estratégias de investimento, como por exemplo, os *trading systems*; e, também, com outros universos de ações que não aquele restrito às ações integrantes do índice de mercado.
 - Investigação de outros métodos de predição de retornos adequados ao modelo de seleção de carteiras baseado em erros de predição.

Referências Bibliográficas

- 1 ABREU, P. F. S. P.; STEPHAN, C. *Análise de Investimentos*. Rio de Janeiro: Editora Campus Ltda, 1982.
- 2 FISCHER, D. E.; JORDAN, R. J. *Security Analysis and Portfolio Management*. 6. ed. [S.l.]: Prentice-Hall International, 1995.
- 3 ELTON, E. J. et al. *Modern Portfolio Theory and Investment Analysis*. 7. ed. [S.l.]: John Wiley & Sons, Inc., 2007.
- 4 SHARPE, W. F.; ALEXANDER, G. J.; BAILEY, J. V. *Investments*. 6. ed. Upper Saddle River, New Jersey: Prentice-Hall, Inc., 1999.
- 5 MARKOWITZ, H. M. *Portfolio Selection: Efficient Diversification of Investments*. 2. ed. New York: John Willey & Sons, 1991.
- 6 MARKOWITZ, H. M. The utility of wealth. *The Journal of Political Economy*, The University of Chicago Press, v. 60, n. 2, p. 151–158, abr. 1952. ISSN 0022-3808.
- 7 MARKOWITZ, H. M. Portfolio selection. *Journal of Finance*, New York, VII, n. 1, p. 77–91, mar. 1952.
- 8 MARKOWITZ, H. M. *Portfolio Selection: Efficient Diversification of Investments*. 1. ed. New York: John Willey & Sons, 1959.
- 9 MASPERO, D. *Portfolio Selection for Financial Planners*. 2004. Newfin Working Paper 3/04. Disponível em: <http://www.uni-bocconi.it/doc_mime_view.php?doc_id=31612>. Acesso em: Abril de 2008.
- 10 FAMA, E. F. Portfolio analysis in a stable paretian market. *Management Science*, v. 11, n. 3 Series A, p. 404–419, 1965.
- 11 KON, S. J. Models of stock returns — a comparison. *The Journal of Finance*, American Finance Association, v. 39, n. 1, p. 147–165, mar. 1984. ISSN 0022-1082.

- 12 FAMA, E. F. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, American Finance Association, v. 25, n. 2, p. 383–417, maio 1970. ISSN 0022-1082.
- 13 FAMA, E. F. Efficient capital markets II. *Journal of Finance*, v. 46, n. 5, p. 1575–1617, dez. 1991.
- 14 FAMA, E. F. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, v. 49, p. 283–306, 1998.
- 15 MALKIEL, B. G. The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, American Economic Association, v. 17, n. 1, p. 59–82, 2003. ISSN 0895-3309.
- 16 COSTA JR, N. C. A.; CERETTA, P. S. Efeito dia da semana: Evidência na América Latina. *Teoria e Evidência Econômica*, v. 8, n. 14, p. 27–35, maio 2000.
- 17 TETLOCK, P. C. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, v. 62, n. 3, p. 1139–1168, out. 2007.
- 18 EDMANS, A.; GARCÍA, D.; NORLI, Ø. Sports sentiment and stock returns. *Journal of Finance*, v. 62, n. 4, p. 1967–1998, out. 2007.
- 19 BECHER, D. A.; JENSEN, G. R.; MERCER, J. M. Monetary policy indicators as predictors of stock returns. *Journal of Financial Research*, n. 4, p. 357–379, 2008.
- 20 BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. *Time Series Analysis: Forecasting and Control*. 3. ed. [S.l.]: Prentice-Hall, Inc., 1994.
- 21 MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Séries Temporais*. 1. ed. [S.l.]: Editora Edgar Blücher Ltda, 2004.
- 22 MORETTIN, P. A. *Econometria Financeira — Um Curso em Séries Temporais Financeiras*. 1. ed. [S.l.]: ABE — Associação Brasileira de Estatística, 2004.
- 23 WHITE, H. Economic prediction using neural networks: The case of IBM daily stock returns. In: *Proceedings of the IEEE International Conference on Neural Networks*. [S.l.: s.n.], 1988. p. 451–458.
- 24 KASSICIEH, S.; PAEZ, T.; VORA, G. Data transformation methods for genetic-algorithm-based investment decisions. In: *Proceedings of the Thirty-First Hawaii International Conference on System Sciences*. [S.l.: s.n.], 1998. v. 5, p. 122–127 vol.5.
- 25 MAYER, H. A.; SCHWAIGER, R. Evolutionary and coevolutionary approaches to time series prediction using generalized multi-layer perceptrons. In: ANGELINE, P. J. et

- al. (Ed.). *Proceedings of the Congress on Evolutionary Computation*. Mayflower Hotel, Washington D.C., USA: IEEE Press, 1999. v. 1, p. 275–280. ISBN 0-7803-5537-7.
- 26 CORTEZ, P.; ROCHA, M.; NEVES., J. Evolving time series forecasting neural network models. In: *Proceedings of International Symposium on Adaptive Systems: Evolutionary Computation and Probabilistic Graphical Models (ISAS 2001)*. Havana, Cuba: [s.n.], 2001.
- 27 CAO, L. J.; TAY, F. E. H. Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks*, v. 14, n. 6, p. 1506–1518, nov. 2003.
- 28 MIRMIRANI, S.; LI, H. C. Gold price, neural networks and genetic algorithm. *Computational Economics*, Kluwer Academic Publishers, Hingham, MA, USA, v. 23, n. 2, p. 193–200, 2004. ISSN 0927-7099.
- 29 HUARNG, K.; YU, T. H. Ratio-based lengths of intervals to improve fuzzy time series forecasting. *IEEE Transactions on Systems, Man and Cybernetics — Part B: Cybernetics*, v. 36, n. 2, p. 328–340, abr. 2006. ISSN 1083-4419.
- 30 MOODY, J.; SAFFELL, M. Learning to trade via direct reinforcement. *IEEE Transactions on Neural Networks*, v. 12, n. 4, p. 875–889, jul. 2001.
- 31 HUARNG, K.; YU, T. H.; HSU, Y. W. A multivariate heuristic model for fuzzy time-series forecasting. *IEEE Transactions on Systems, Man and Cybernetics — Part B: Cybernetics*, v. 37, n. 4, p. 836–846, ago. 2007. ISSN 1083-4419.
- 32 FREITAS, F. D. *Ambiente Computacional para o Problema de Seleção de Carteiras de Investimento Baseado na Predição de Retornos*. Dissertação (Mestrado) — Programa de Pós-Graduação em Informática, Universidade Federal do Espírito Santo — PPGI/UFES, 1997.
- 33 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. Autoregressive neural network predictors in the brazilian stock markets. In: *VII Simpósio Brasileiro de Automação Inteligente (SBAI)/II IEEE Latin American Robotics Symposium (IEEE-LARS)*. São Luis, Brasil: [s.n.], 2005. p. 1–8. ISBN 85-85048-55-7.
- 34 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. Avaliação de preditores neurais auto-regressivos no mercado de ações. In: *VII Congresso Brasileiro de Redes Neurais (CBRN)*. Natal, Brasil: [s.n.], 2005. p. 1–6. ISSN 1808-8589.
- 35 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. A prediction-based portfolio optimization model. In: *5th International Symposium On Robotics and Automation — ISRA 2006*. Hidalgo, Mexico: [s.n.], 2006. p. 520–525. ISBN 970-769-070-4.

- 36 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. Um modelo de otimização de carteiras baseado em predição. In: *XXXVIII Simpósio Brasileiro de Pesquisa Operacional (SBPO)*. Goiânia-GO, Brasil: [s.n.], 2006. p. 1–8.
- 37 SHARDA, R.; PATIL, R. B. A connectionist approach to time series prediction: An empirical test. *Journal of Intelligent Manufacturing*, v. 3, n. 5, p. 317–323, 1992.
- 38 LOWE, D. Novel exploitation of neural network methods in financial markets. In: *International Conference on Neural Networks, IEEE World Congress on Computational Intelligence*. [S.l.: s.n.], 1994. p. 3623–3628 vol.6.
- 39 HANSEN J.V.; NELSON, R. Neural networks and traditional time series methods: a synergistic combination in state economic forecasts. *IEEE Transactions on Neural Networks*, v. 8, n. 4, p. 863–873, jul. 1997. ISSN 1045-9227.
- 40 FREITAS, F. D.; De SOUZA, A. F.; ALMEIDA, A. R. Prediction-based portfolio optimization model using neural networks. *Neurocomputing*, 2009. (in-press).
- 41 MOODY, J. Economic forecasting: Challenges and neural solutions. In: *International Symposium on Artificial Neural Networks*. Hsinchu, Taiwan: [s.n.], 1995. Keynote Talk.
- 42 HELLSTRÖM, T. Data snooping in the stock market. *Theory of Stochastic Processes*, v. 5, n. 21, p. 33–50, out. 1999.
- 43 HENDRY, D. F.; CLEMENTS, M. P. Pooling of forecasts. *Econometrics Journal*, v. 7, n. 1, p. 1–31, jun. 2004.
- 44 FAMA, E. F. Mandelbrot and the stable paretian hypothesis. *The Journal of Business*, The University of Chicago Press, v. 36, n. 4, p. 420–429, out. 1963. ISSN 0021-9398.
- 45 FAMA, E. F. The behavior of stock-market prices. *The Journal of Business*, The University of Chicago Press, v. 38, n. 1, p. 34–105, jan. 1965. ISSN 0021-9398.
- 46 FAMA, E. F. Random walks in stock market prices. *Financial Analysts Journal*, v. 21, p. 55–59, 1965.
- 47 MANDELBROT, B. *Citation Classic Commentaries — The Variation of Certain Speculative Prices*. [S.l.]: Eugene Garfield, jun. 1982. 20 p.
- 48 MANDELBROT, B. The variation of certain speculative prices. *The Journal of Business*, The University of Chicago Press, v. 36, n. 4, p. 394–419, out. 1963. ISSN 0021-9398.
- 49 MANDELBROT, B.; FISHER, A.; CALVET, L. *A Multifractal Model of Asset Returns*. [S.l.], set. 1997.

- 50 SHARPE, W. F. A simplified model for portfolio analysis. *Management Science*, v. 2, n. 9, p. 277–293, 1963.
- 51 SHARPE, W. F. A linear programming approximation for the general portfolio selection problem. *Journal of Financial and Quantitative Analysis*, n. 6, p. 1263–1275, 1971.
- 52 KONNO, H. Piecewise linear risk functions and portfolio optimization. *Journal of Operations Research Society of Japan*, n. 33, p. 139–156, 1990.
- 53 KONNO, H.; YAMAZAKI, H. Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science*, New York, v. 37, n. 5, p. 519–531, maio 1991.
- 54 KONNO, H.; SUZUKI, S. A fast algorithm for solving large scale mean-variance models by compact factorization of covariance matrices. *Journal of Operations Research Society of Japan*, n. 35, p. 93–104, 1992.
- 55 KONNO, H.; PLISKA, S.; SUZUKI, K. Optimal portfolio with asymptotic criteria. *Operations Research*, n. 45, p. 187–204, 1993.
- 56 KONNO, H.; SHIRAKAWA, H.; YAMAZAKI, H. A mean-absolute deviation-skewness portfolio optimization model. *Annals of Operations Research*, v. 45, n. 1, p. 205–220, dez. 1993.
- 57 FRIED, J. Forecasting and probability distributions for models of portfolio selection. *Journal of Finance*, XXV, n. 3, p. 539–554, jun. 1970.
- 58 De SOUZA, A. F.; FREITAS, F. D. Supercomputer performance with a workstation cluster on dense matrix multiplication. In: *IASTED International Conference on Parallel and Distributed Computing and Systems*. Washington D.C., USA: [s.n.], 1995. p. 257–261.
- 59 CHEN, Y. et al. Time-series forecasting using flexible neural tree model. *Information Sciences*, v. 174, n. 3–4, p. 219–235, ago. 2005.
- 60 CHEN, Y.; YANG, B.; ABRAHAM, A. Flexible neural trees ensemble for stock index modeling. *Neurocomputing*, v. 70, n. 4–6, p. 697–703, jan. 2007.
- 61 LIU, F.; NG, G. S.; QUEK, C. RLDDE: A novel reinforcement learning-based dimension and delay estimator for neural networks in time series prediction. *Neurocomputing*, v. 70, n. 7–9, p. 1331–1341, mar. 2007.
- 62 LEE, J. W. et al. A multiagent approach to Q-Learning for daily stock trading. *IEEE Transactions on Systems, Man and Cybernetics — Part A: Systems and Humans*, v. 37, n. 6, p. C1–853, nov. 2007. ISSN 1083-4427.

- 63 THAWORNWONG, S.; ENKE, D. The adaptive selection of financial and economic variables for use with artificial neural networks. *Neurocomputing*, v. 56, p. 205–232, 1 2004.
- 64 PANTAZOPOULOS, K. et al. Financial prediction and trading strategies using neurofuzzy approaches. *IEEE Transactions on Systems, Man and Cybernetics — Part B: Cybernetics*, v. 28, n. 4, p. 520–531, ago. 1998. ISSN 1083-4419.
- 65 HELLSTRÖM, T. Predicting a rank measure for stock returns. *Theory of Stochastic Processes*, v. 22, n. 6, p. 64–83, 2000.
- 66 HELLSTRÖM, T. Optimization of trading rules with a penalty term for increased risk-adjusted performance. *Advanced Modeling and Optimization*, v. 2, n. 4, p. 135–149, nov. 2000.
- 67 HELLSTRÖM, T. Outlier removal for prediction of covariance matrices with an application to portfolio optimization. *Theory of Stochastic Processes*, v. 22, n. 6, p. 47–63, 2000.
- 68 LAZO, J. G.; VELLASCO, M. M. B. R.; PACHECO, M. A. C. A hybrid genetic-neural system for portfolio selection and management. In: *Sixth International Conference on Engineering Applications of Neural Networks (EANN 2000)*. Kingston University, UK: [s.n.], 2000.
- 69 LEUNG, M. T.; DAOUK, H.; CHEN, A. Using investment portfolio return to combine forecasts: A multiobjective approach. *European Journal of Operational Research*, v. 134, n. 1, p. 84–102, out. 2001.
- 70 HUNG, K.; CHEUNG, Y.; XU, L. An extended ASLD trading system to enhance portfolio management. *IEEE Transactions on Neural Networks*, v. 14, n. 2, p. 413–425, mar. 2003.
- 71 XU, L.; CHEUNG, Y. Adaptive supervised learning decision networks for trading and portfolio management. *Journal of Computational Intelligence in Finance*, v. 5, n. 6, p. 11–16, nov. 1997.
- 72 SHARPE, W. F. Mutual fund performance. *Journal of Business*, v. 39, n. 1, p. 119–138, jan. 1966.
- 73 SORTINO, F. A.; MEER, R. van der. Downside risk — capturing what's at stake in investment situations. *Journal of Portfolio Management*, p. 27–31, 1991.
- 74 SÁ, G. T. *Mercado de Ações e Bolsa de Valores*. Rio de Janeiro: Aplicação Editora Técnica, 1987.

- 75 BRASIL. *Código Civil*. 1. ed. Brasília: Câmara dos Deputados, Coordenação de Publicações, 2002. (Série Fontes de Referência. Legislação; nº 43).
- 76 BOVESPA. *Dicionário de Finanças*. 2008. Disponível em: <<http://www.enfin.com.br/bolsa/main.php>>. Acesso em: Abril de 2008.
- 77 BOVESPA. *Mercados da BOVESPA*. 2008. Disponível em: <<http://www.bovespa.com.br/Mercado/RendaVariavel/Mercados/MercadosBovespa.asp>>. Acesso em: Agosto de 2008.
- 78 BOVESPA. *BOVESPA a Bolsa do Brasil — BOVESPA*. 2008. Disponível em: <<http://www.bovespa.com.br>>.
- 79 ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of the united kingdom inflation. *Econometrica*, v. 4, n. 50, p. 987–1007, 1982.
- 80 BOLLERSLEV, T. Generalized autoregressive conditional heteroscedasticity model. *Journal of Econometrics*, n. 31, p. 307–327, 1996.
- 81 FILDES, R.; LUSK, E. J. The choice of a forecasting model. *OMEGA International Journal of Management Science*, v. 12, n. 5, p. 427–435, 1984.
- 82 MEADE, N.; SMITH, I. M. D. Ararma vs arima — a study of the benefits of a new approach to forecasting. *OMEGA International Journal of Management Science*, v. 13, n. 6, p. 519–534, 1985.
- 83 SUTTON, R. S. Learning to predict by the method of temporal differences. *Machine Learning*, v. 3, p. 9–44, 1988.
- 84 LUDERMIR, T. B. et al. Weightless neural models: A review of current and past works. *Neural Computing Surveys*, v. 2, p. 41–61, 1999.
- 85 SPECHT, D. F. Probabilistic neural networks. *Neural Networks*, Elsevier Science Ltd., Oxford, UK, UK, v. 3, n. 1, p. 109–118, 1990. ISSN 0893-6080.
- 86 MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Journal Bulletin of Mathematical Biology*, Springer, New York, v. 5, n. 4, p. 115–133, dez. 1943.
- 87 WASSERMAN, P. D. *Neural Computing: Theory and Practice*. [S.l.]: Van Nostrand Reinhold, 1989.
- 88 HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2. ed. [S.l.]: Prentice-Hall, Inc., 1999.

- 89 RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, n. 323, p. 533–536, 1986.
- 90 HECHT-NIELSEN, R. Theory of the backpropagation neural network. In: *International Joint Conference on Neural Networks — IJCNN*. [S.l.: s.n.], 1989. p. 593–605 vol.1.
- 91 MINSKY, M. L.; PAPERT, S. A. *Perceptrons*. Cambridge, MA, U.S.A.: MIT Press, 1969.
- 92 LEIGHTON, R. R. *The Aspirin/MIGRAINES Neural Network Software, Release V7.0*. [S.l.], 1994.
- 93 NELSON, C. R. *Applied Time Series Analysis for Managerial Forecasting*. 1. ed. San Francisco, USA: Holden-Day, Inc., 1973.
- 94 NARENDRAN, R.; GANAPATHY, V.; SOMASUNDARAM, M. A study on the effects of recency factors on prediction in real-world domains. In: *International Conference on Neural Networks, IEEE World Congress on Computational Intelligence*. [S.l.: s.n.], 1994. p. 3646–3650 vol.6.
- 95 REFENES, A. et al. Financial time series modelling with discounted least squares backpropagation. *Neurocomputing*, n. 14, p. 123–138, 1997.
- 96 MARKOWITZ, H. M. *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1990*. 1990. Disponível em: <http://nobelprize.org/nobel_prizes/economics/laureates/1990/>. Acesso em: Abril de 2008.
- 97 MARKOWITZ, H. M. Foundations of portfolio theory. *Journal of Finance*, New York, v. 46, n. 2, p. 469–477, jun. 1991.
- 98 ROY, A. D. Safety-first and the holding of assets. *Econometrica*, v. 5, n. 20, p. 431–449, jul. 1952.
- 99 ONG, S.; RANASINGHE, M. Portfolio variance and correlation matrices. *Journal of Real Estate Portfolio Management*, v. 6, n. 1, p. 1–6, 2000.
- 100 MARKOWITZ, H. M. The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly*, n. 3, p. 111–133, mar. 1956.
- 101 HAMZA, F.; JANSSEN, J. Linear approach for solving large-scale portfolio optimization problems in a lognormal market. In: *IAA/AFIR Colloquium*. Nürnberg, Germany: [s.n.], 1996. p. 1019–1039.
- 102 DUARTE JR, A. M. A comparative study of downside risk and volatility in asset allocation. *Investigación Operativa*, v. 4, n. 3, p. 213–228, 1994.

- 103 BOVESPA. *Índice BOVESPA — Definição e Metodologia*. 2008. Disponível em: <<http://www.bovespa.com.br/Pdf/Indices/IBovespa.pdf>>. Acesso em: Abril de 2008.
- 104 LANGBEIN, J. H.; POSNER, R. A. Market funds and trust-investment law. *American Bar Foundation Research Journal*, v. 1, n. 1, p. 1–34, 1976.
- 105 LANGBEIN, J. H.; POSNER, R. A. Market funds and trust-investment law: II. *American Bar Foundation Research Journal*, v. 1, n. 1, p. 1–43, 1977.
- 106 ZHAO, T. C.; OVERMARS, M. *Forms Library — A Graphical User Interface Toolkit for X*. 2000. Disponível em: <<http://savannah.nongnu.org/projects/xforms/>>. Acesso em: Abril de 2008.
- 107 LINDO Systems Inc. *LINDO API*. 2008. Disponível em: <<http://www.lindo.com/downloads/>>. Acesso em: Abril de 2008.
- 108 GALASSI, M. et al. *GNU Scientific Library Reference Manual — Revised Second Edition (v1.8)*. 2. ed. [S.l.], ago. 2006.
- 109 WILLIAMS, T.; KELLEY, C. *gnuplot — An Interactive Plotting Program*. 2007. Disponível em: <<http://www.gnuplot.info/>>.
- 110 ELMASRI, R.; LEE, J. Y. Implementation options for time-series data. In: ETZION, O.; JAJODIA, S.; SRIPADA, S. (Ed.). *Temporal Databases — Research and Practice (Lecture Notes in Computer Science)*. Berlin / Heidelberg: Springer, 1998. v. 1399/1998, p. 115–128.
- 111 FREITAS, F. D. et al. Portfolio selection with predicted returns using neural networks. In: *IASTED International Conference on Artificial Intelligence and Applications*. Marbella, Spain: [s.n.], 2001. p. 99–103.
- 112 LAWRENCE, S.; GILES, C. L.; TSOI, A. C. Lessons in neural network training: Overfitting may be harder than expected. In: *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI-97)*. [S.l.: s.n.], 1997. p. 540–545.
- 113 DROSSU, R.; OBRADOVI, Z. Regime signaling techniques for non-stationary time series. In: *HICSS '97: Proceedings of the 30th Hawaii International Conference on System Sciences*. Washington, DC, USA: IEEE Computer Society, 1997. p. 530. ISBN 0-8186-7743-0.
- 114 MOODY, J. Prediction risk and architecture selection for neural networks. In: CHERKASSKY, V.; FRIEDMAN, J. H.; WECHSLER, H. (Ed.). *From Statistics to Neural Networks: Theory and Pattern Recognition Applications*. [S.l.]: Springer, NATO ASI Series F, 1994.

- 115 ARMSTRONG, J. S.; COLLOPY, F. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, v. 8, n. 1, p. 69–80, jun. 1992.
- 116 GRINBLATT, M.; TITMAN, S. Performance measurement without benchmarks: An examination of mutual fund returns. *The Journal of Business*, The University of Chicago Press, v. 66, n. 1, p. 47–68, jan. 1993. ISSN 0021-9398.
- 117 GRAUER, R. R.; HAKANSSON, N. H. Applying portfolio change and conditional performance measures: The case of industry rotation via the dynamic investment model. *Review of Quantitative Finance and Accounting*, v. 17, n. 3, p. 237–265, 2001.
- 118 SPIEGEL, M. R.; STEPHENS, L. J. *Schaum's Outline of Theory and Problems of Statistics*. 3. ed. [S.l.]: McGraw-Hill, 1998.
- 119 STATMAN, M. How many stocks make a diversified portfolio? *The Journal of Financial and Quantitative Analysis*, University of Washington School of Business Administration, v. 22, n. 3, p. 353–363, set. 1987. ISSN 0022-1090.
- 120 SHAFER, G.; VOVK, V. *Probability and Finance: It's Only a Game!* 1. ed. [S.l.]: Wiley-Interscience, 2001. ISBN 978-0471402268.
- 121 HUANG, L.; LIU, H. Rational inattention and portfolio selection. *Journal of Finance*, LXII, n. 4, p. 1999–2217, ago. 2001.
- 122 BRANDT, M. W. Dynamic portfolio selection by augmenting the asset space. *Journal of Finance*, LXI, n. 5, p. 2187–2217, out. 2006.
- 123 BOYLE, P.; TIAN, W. Portfolio management with constraints. *Mathematical Finance*, n. 3, p. 319–343, jul. 2007.
- 124 XU, J. Price convexity and skewness. *Journal of Finance*, v. 62, n. 5, p. 2521–2552, out. 2007.
- 125 HU, O. Applicability of the fama-french three-factor model in forecasting portfolio returns. *Journal of Financial Research*, v. 30, n. 1, p. 111–127, 2007.
- 126 KRITZMAN, M.; MYRGREN, S.; PAGE, S. Optimal execution for portfolio transitions. *Journal of Portfolio Management*, p. 33–39, Spring 2007.
- 127 AIT-SAHALIA, Y.; BRANDT, M. W. Variable selection for portfolio choice. *Journal of Finance*, LVI, n. 4, p. 1297–1351, ago. 2001.
- 128 GUTIERREZ, R. C.; KELLEY, E. K. The long-lasting momentum in weekly returns. *Journal of Finance*, v. 63, n. 1, p. 415–447, fev. 2008.

- 129 OLIVEIRA, A. L. I.; MEIRA, S. R. L. Detecting novelties in time series through neural networks forecasting with robust confidence intervals. *Neurocomputing*, v. 70, n. 1-3, p. 79–92, dez. 2006.
- 130 FABOZZI, F. J.; KOLM, P. N. Robust portfolio optimization: Recent trends and future directions. *Journal of Portfolio Management*, p. 40–48, Spring 2007.
- 131 DeGENNARO, R. P. Market imperfections. *Journal of Financial Transformation*, v. 14, p. 107–117, ago. 2005.
- 132 DeGENNARO, R. P.; ROBOTTI, C. Financial market imperfections. *Economic Review - Federal Reserve Bank of Atlanta*, Third Quarter 2007. Disponível em: <http://www.frbatlanta.org/filelegacydocs/erq307_DeGennaro.pdf>.
- 133 BOVESPA. *Tabela de Taxas e Emolumentos para Operações de Investidores*. 2008. Disponível em: <<http://www.bovespa.com.br/Investidor/CustosOperac/CustosOperacionais.htm>>. Acesso em: Abril de 2008.
- 134 RFB. *Aplicações Financeiras - Renda Fixa e Renda Variável*. 2008. Disponível em: <<http://www.receita.fazenda.gov.br/pessoafisica/irpf/2008/perguntas/aplicfinanrenfixarenvariavel.htm>>. Acesso em: Abril 2008.
- 135 SCHYNS, M. *Modelling Financial Data and Portfolio Optimization*. Tese (Doutorado) — Faculté d'Economie, de Gestion et de Sciences Sociales, Ecole d'Administration des Affaires — Université de Liège, 2001.
- 136 RIVALS, I.; PERSONNAZ, L. Neural-network construction and selection in nonlinear modeling. *IEEE Transactions on Neural Networks*, v. 14, n. 4, p. 804–819, jul. 2003. ISSN 1045-9227.
- 137 YAO, X. Evolving artificial neural networks. *Proceedings of the IEEE*, v. 87, n. 9, p. 1423–1447, set. 1999.
- 138 NVIDIA. *NVIDIA CUDA technology*. 2008. Disponível em: <http://www.nvidia.com/object/cuda_home#>.