

---

Aprendizagem Semi-Supervisionada  
aplicada à Engenharia Financeira

*Levi Henrique Santana de Lelis*

---

# Aprendizagem Semi-Supervisionada aplicada à Engenharia Financeira

*Levi Henrique Santana de Lelis*

**Orientador:** *Prof. Antônio de Pádua Braga*

Dissertação submetida ao Programa de Pós-graduação em Engenharia Elétrica da UFMG como requisito parcial para obtenção do grau de mestre em Engenharia Elétrica.

**UFMG - Belo Horizonte**  
**06/2007**

# Agradecimentos

---

**A**gradeço aos meus pais, Vicente e Maria, pelo constante apoio. Não tenho dúvidas de que sem o incentivo e carinho deles, dificilmente teria chegado até aqui. Agradeço também aos meus irmãos Davi e Marina, por sempre estarem presentes e me ajudando de todas as formas possíveis. Agradeço também à Thuane, pela dedicação total e carinho.

Impossível deixar de lembrar dos amigos nesse momento tão importante, portanto, agradeço aos meus amigos de Viçosa e Belo Horizonte, em especial ao Rafael, Léo e Guilherme, que estiveram presentes em todas as etapas do desenvolvimento deste trabalho.

Um muito obrigado a todos amigos do LTC pelas ótimas discussões promovidas sobre esse empolgante campo de pesquisa. Um agradecimento especial ao Cristiano que participou de forma direta no trabalho, me dando conselhos valiosos.

Aos amigos que fiz no CPDEE durante as disciplinas, um muito obrigado especial. Momentos difíceis que passamos juntos nunca serão esquecidos, principalmente pela amizade construída nessa fase. Dudu, Vinsaid, Roger, Luciana, Louback, JK e Daniel, um forte abraço.

Obrigado ao Zeluis pela formação que me proporcionou, por apontar o caminho e me apoiar.

Obrigado ao Braga por esse quase um ano e meio de intenso aprendizado, não só técnico, mas humano também. Graças a ele subi mais um degrau e agora consigo ter uma vista mais interessante do mundo.

Por fim, obrigado ao CNPq pelo apoio financeiro.

---

---

Ao saudoso Tio Geraldo.

# Resumo

---

A aprendizagem semi-supervisionada se tornou, recentemente, em uma boa alternativa para aumentar a capacidade de generalização de modelos de aprendizagem de máquina. A abordagem é utilizada, geralmente, em problemas que as amostras rotuladas são difíceis de serem obtidas, e as sem rótulo, por sua vez, são abundantes e facilmente coletadas; o algoritmo de treinamento semi-supervisionado busca, então, gerar a superfície de separação entre as classes com base também nas amostras sem rótulo. Um exemplo clássico é o de classificação de sites na web: gasta-se muito tempo para se rotular uma amostra de forma confiável. Entretanto, um crawler é capaz de coletar inúmeras amostras sem rótulo rapidamente.

Este trabalho mostra uma aplicação diferenciada do aprendizado semi-supervisionado, pois nesse caso, possui-se um extenso conjunto de treinamento rotulado (séries temporais financeiras são facilmente encontradas para download na internet), no entanto, a qualidade desse conjunto é colocada em questão: a rotulação é confiável? Através da abordagem semi-supervisionada foi possível minimizar os dados ruidosos do conjunto de treinamento, melhorando os resultados obtidos.

# Abstract

---

---

Semi-supervised learning had become, recently, a good alternative to improve generalization capacity in machine learning models. The approach is generally used in problems that labeled samples are hard to be obtained and unlabeled, in turn, are plenty and easily collected; the semi-supervised training algorithm tries to generate the separation surface between the two classes based also on unlabeled data. A classic example is the web sites classification: give a trustworthy label to a sample is time consuming. Nonetheless, a crawler can quickly collect a great number of unlabeled samples.

This work shows a different application of semi-supervised learning, because in this case, it is possessed a long labeled training set (financial time series can be easily downloaded from internet), however, the quality of this set is put under prove: the labels are trustworthy? Through the semi-supervised approach it was possible minimize the noise data from training set, improving the results obtained.

# Conteúdo

---

Resumo . . . . .	iv
Abstract . . . . .	v
Sumário . . . . .	vii
Lista de Figuras . . . . .	xi
Lista de Tabelas . . . . .	xii
<b>1 Introdução</b>	<b>1</b>
1.1 Abordagem utilizada . . . . .	2
1.2 Organização do texto . . . . .	4
<b>2 Pré-processamento dos dados</b>	<b>6</b>
2.1 Teoria do Mercado Eficiente . . . . .	6
2.2 Métodos clássicos . . . . .	7
2.3 Análise Fundamentalista . . . . .	9
2.4 Análise Técnica . . . . .	10
2.4.1 Filosofia da Análise Técnica . . . . .	11
2.4.2 Diferenças entre Análise Fundamentalista e Técnica . . . . .	12
2.4.3 Janela de tempo para Análise Técnica . . . . .	13
2.4.4 Teoria de Dow . . . . .	14
2.4.5 Trabalhos que utilizaram Análise Técnica . . . . .	18
2.4.6 Padrão Japoneses de <i>Candlesticks</i> . . . . .	23
2.5 Abordagem utilizada neste trabalho . . . . .	30
2.5.1 Automatização do rotulamento . . . . .	30
2.6 Conclusões . . . . .	32
<b>3 Aprendizagem Semi-Supervisionada aplicada à Engenharia Financeira</b>	<b>34</b>
3.1 Dados Conflitantes . . . . .	34
3.2 SVM . . . . .	37

3.2.1	Características das SVMs . . . . .	37
3.2.2	Aprendizado Estatístico . . . . .	38
3.2.3	SVMs de Margem Rígida . . . . .	46
3.2.4	SVMs com margens suavizadas . . . . .	47
3.2.5	SVMs não lineares . . . . .	48
3.3	Aprendizagem Semi-supervisionada . . . . .	50
3.3.1	<i>Transductive Support Vector Machines</i> . . . . .	53
3.3.2	Aprendizagem semi-supervisionada aplicada à Engenharia Financeira . . . . .	55
3.4	Treinamento dos modelos . . . . .	57
3.5	Conclusão . . . . .	61
<b>4</b>	<b>Resultados</b>	<b>63</b>
4.1	Estratégia de Investimento . . . . .	63
4.2	Simulação da bolsa de valores . . . . .	66
4.3	A simulação . . . . .	68
4.4	Investindo em um simulador mais real . . . . .	69
4.4.1	Negócios realizados no Em Ação . . . . .	70
4.5	Otimização dos investimentos . . . . .	72
4.6	Seleção de Portfólio . . . . .	74
4.7	Exemplo de utilização da Seleção de Carteira . . . . .	77
4.8	Teoria de Markowitz aplicada ao sistema de previsão . . . . .	79
4.9	Conclusão . . . . .	81
<b>5</b>	<b>Discussões e Conclusões</b>	<b>83</b>
5.1	Abordagens falhas . . . . .	83
5.2	Resultados secundários . . . . .	85
5.3	A abordagem utilizada . . . . .	86
5.3.1	Utilização como um sistema tutor . . . . .	86
5.3.2	Treinamento independente . . . . .	87
5.3.3	Desbalanceamento das classes . . . . .	87
5.3.4	A não necessidade de retreino . . . . .	88
5.4	Sugestões para trabalhos futuros . . . . .	89
5.4.1	Ainda sobre Análise Técnica . . . . .	89
5.4.2	Combinando Análise Técnica e Fundamentalista . . . . .	89
	<b>Referências</b>	<b>95</b>
	<b>A Gráficos de investimentos em 2006</b>	<b>96</b>
	<b>B Gráficos de investimentos em 2007</b>	<b>112</b>

# Lista de Figuras

---

---

2.1	Lojas Americanas - LAME4 - em tendência de alta . . . . .	12
2.2	Gráfico em barras da Cemig - CMIG4 - em 2007 . . . . .	14
2.3	Tendência primária, secundária e terciária . . . . .	15
2.4	As três fases em um tendência de alta - PMAM4 . . . . .	16
2.5	Divergência entre o volume e preço, indicando uma reversão . . .	17
2.6	Sinais de final de tendência . . . . .	17
2.7	Divergência entre o gráfico de preço e o indicador de índice de força relativa - VALE5 . . . . .	20
2.8	Vários sinais de compra e venda são falsos quando a série possui uma tendência definida . . . . .	21
2.9	Média Móvel de 15 dias de ITSA4 . . . . .	21
2.10	Formato de um <i>candlestick</i> . . . . .	24
2.11	CMIG4 no gráfico de <i>candlestick</i> . . . . .	24
2.12	Dias de indecisão . . . . .	25
2.13	<i>Doji Candlesticks</i> . . . . .	25
2.14	Nuvem Negra, um padrão de reversão . . . . .	27
2.15	Linha Perfurante, um padrão de reversão . . . . .	27
2.16	Estrela da Manhã, um padrão de reversão . . . . .	28
2.17	Estrela da Tarde, um padrão de reversão . . . . .	28
2.18	Triângulo, uma figura de reversão . . . . .	29
2.19	Ombro-Cabeça-Ombro, importante figura de reversão . . . . .	30
3.1	Sequência de Martelos Invertidos, indicando futura queda no papel	35
3.2	Sequência de Martelos, indicando futura valorização no papel . .	36
3.3	Um exemplo de <i>overfitting</i> . . . . .	40
3.4	Provavelmente a melhor solução para o problema . . . . .	41
3.5	Um exemplo de <i>underfitting</i> . . . . .	41
3.6	Todas as dicotomias possíveis para um separador linear . . . . .	43

3.7	Com quatro amostras, um separador linear não é capaz de separar todos os casos . . . . .	43
3.8	Relação entre Risco Funcional, Empírico e Complexidade do modelo . . . . .	44
3.9	Duas soluções, a de margem máxima e uma separação qualquer.	45
3.10	Problema linearmente separável, porém, com amostras ruidosas .	47
3.11	Mapamento do espaço de entrada para um outro espaço em que a separação linear seja aplicada . . . . .	49
3.12	Algoritmo de <i>clustering</i> semi-supervisionado. Ponto em triângulo indica uma classe, em cruz outra. Os demais pontos são as amostras sem rótulo. . . . .	51
3.13	Amostras rotuladas de duas classes . . . . .	52
3.14	Superfície de separação utilizando apenas dados rotulados no treinamento - abordagem indutiva . . . . .	52
3.15	Amostras rotuladas e conjunto de trabalho sem rótulo de duas classes . . . . .	53
3.16	Superfície de separação utilizando dados rotulados e não rotulados no treinamento - abordagem transdutiva . . . . .	53
3.17	Algoritmo em português do TSVM Light . . . . .	55
3.18	Interface em JSE - $JSVM^{lisc}$ . . . . .	57
3.19	Curva ROC da abordagem supervisionada dos modelos de compra, primeira; e venda, segunda . . . . .	59
3.20	Curva ROC da abordagem semi-supervisionada dos modelos de compra, primeira; e venda, segunda . . . . .	60
4.1	Simulador da Bolsa de Valores . . . . .	67
4.2	Pontos de compra e venda da USIM5 - aplicação realizada no simulador Em Ação . . . . .	71
4.3	Lucro em um <i>pull back</i> da TAMM4 . . . . .	72
4.4	Rendimento da carteira na simulação do Em Ação, em março de 2007 . . . . .	72
4.5	Análise de correlação no gráfico de Retorno por Risco . . . . .	76
4.6	<i>Candlestick</i> - AMBV4 . . . . .	81
4.7	<i>Candlestick</i> - SUBA3 . . . . .	82
A.1	Comprou 1459 ações SDIA4 em: 20060116 por: 6.85. Vendeu 1459 ações SDIA4 em: 20060202 por: 7.5 . . . . .	97
A.2	Comprou 232 ações SUBA3 em: 20060208 por: 47.0. Vendeu 232 ações SUBA3 em: 20060220 por: 58.49 . . . . .	97
A.3	Tentou aplicar em ELET6 em: 20060209 mas não tinha dinheiro	98

A.4	Tentou aplicar em GOLL4 em: 20060209 mas não tinha dinheiro	98
A.5	Comprou 252 ações SUBA3 em: 20060301 por: 53.75. Vendeu 252 ações SUBA3 em: 20060307 por: 49.25 . . . . .	99
A.6	Comprou 2162 ações SDIA4 em: 20060309 por: 5.75. Vendeu 2162 ações SDIA4 em: 20060321 por: 5.75 . . . . .	99
A.7	Comprou 2182 ações SDIA4 em: 20060324 por: 5.69. Vendeu 2182 ações SDIA4 em: 20060406 por: 6.38 . . . . .	100
A.8	Comprou 2377 ações SDIA4 em: 20060502 por: 5.85. Vendeu 2377 ações SDIA4 em: 20060509 por: 6.15 . . . . .	101
A.9	Comprou 179 ações LAME4 em: 20060516 por: 81.45. Vendeu 179 ações LAME4 em: 20060529 por: 83.5 . . . . .	102
A.10	Comprou 2 ações SDIA4 em: 20060516 por: 5.63. Vendeu 2 ações SDIA4 em: 20060602 por: 5.97 . . . . .	102
A.11	Comprou 373 ações ELET6 em: 20060616 por: 40.01. Vendeu 373 ações ELET6 em: 20060704 por: 45.2 . . . . .	103
A.12	Tentou aplicar em LAME4 em: 20060616 mas não tinha dinheiro	103
A.13	Tentou aplicar em ELET6 em: 20060619 mas não tinha dinheiro	104
A.14	Tentou aplicar em ELET6 em: 20060621 mas não tinha dinheiro	104
A.15	Comprou 1 ação SDIA4 em: 20060622 por: 5.22. Vendeu 1 ação SDIA4 em: 20060703 por: 5.71 . . . . .	105
A.16	Comprou 3132 ações SDIA4 em: 20060710 por: 5.38. Vendeu 3132 ações SDIA4 em: 20060719 por: 6.1 . . . . .	105
A.17	Comprou 434 ações ELET6 em: 20060824 por: 43.95. Vendeu 434 ações ELET6 em: 20060911 por: 41.4 . . . . .	106
A.18	Tentou aplicar em ELET6 em: 20060825 mas não tinha dinheiro	106
A.19	Tentou aplicar em SDIA4 em: 20060825 mas não tinha dinheiro	107
A.20	Tentou aplicar em ELET6 em: 20060828 mas não tinha dinheiro	107
A.21	Comprou 450 ações ELET6 em: 20060922 por: 39.88. Vendeu 450 ações ELET6 em: 20061005 por: 48.0 . . . . .	108
A.22	Tentou aplicar em LAME4 em: 20060922 mas não tinha dinheiro	108
A.23	Comprou 1 ação SDIA4 em: 20060925 por: 5.71. Vendeu 1 ação SDIA4 em: 20061003 por: 6.06 . . . . .	109
A.24	Comprou 537 ações SUBA3 em: 20061009 por: 40.2. Vendeu 537 ações SUBA3 em: 20061108 por: 47.75 . . . . .	109
A.25	Comprou 4175 ações SDIA4 em: 20061113 por: 6.14. Vendeu 4175 ações SDIA4 em: 20061206 por: 7.02 . . . . .	110
A.26	Comprou 450 SUBA3 em: 20061212 por: 65.0 . . . . .	110
A.27	Tentou aplicar em GOLL4 em: 20061226 mas não tinha dinheiro	111

B.1	Comprou 154 ações SUBA3 em: 20070122 por: 64.6. Vendeu 154 ações SUBA3 em: 20070206 por: 69.9 . . . . .	113
B.2	Comprou 1031 ações ITSA4 em: 20070316 por: 10.47. Vendeu 1031 ações ITSA4 em: 20070323 por: 11.8 . . . . .	113
B.3	Comprou 1063 ações ITSA4 em: 20070503 por: 11.43. Vendeu 1063 ações ITSA4 em: 20070521 por: 12.42 . . . . .	114
B.4	Tentou aplicar em SUBA3 em: 20070518 mas não tinha dinheiro	114

# Lista de Tabelas

---

---

3.1 Modelos da abordagem supervisionada selecionados para testes .	59
3.2 Modelos da abordagem semi-supervisionada selecionados para testes . . . . .	61
3.3 Matriz de confusão do modelo de venda selecionado, abordagem supervisionada . . . . .	61
3.4 Matriz de confusão do modelo de venda selecionado, abordagem semi-supervisionada . . . . .	61
4.1 Valor do <i>moving stop loss</i> utilizado em cada um dos modelos . . .	66
4.2 Resultados dos investimentos por modelo . . . . .	68
4.3 Corretagem para investimentos no simulador da Folha de São Paulo . . . . .	70
4.4 Retornos, de 5 dias, em 2005 . . . . .	78
4.5 Retorno esperado e risco . . . . .	78
4.6 Resultado da carteira sem otimização . . . . .	78
4.7 Resultado da otimização . . . . .	78
4.8 Resultado da carteira otimizada . . . . .	79
4.9 Proporção de participação de cada empresa nas carteiras . . . . .	79
4.10 Resultado da aplicação da Teoria de Markowitz ao sistema de previsão . . . . .	80

---

# Introdução

---

**D**esde 1890, quando um grupo de agentes liderados por Emílio Rangel Pestana inaugurou a Bolsa livre, o que viria a se tornar mais tarde a bolsa de valores do estado de São Paulo, muito se modificou na forma de operar e realizar negócios [7]. Uma das mudanças mais notáveis foi, sem dúvida, a utilização da tecnologia para agilizar e tornar as transações mais seguras. Em 1970, os boletos que eram utilizados para realizar as transações foram substituídos por cartões perfurados, os negócios passaram a ser registrados de forma eletrônica, o que marcou o início da digitalização da bolsa e hoje, todos os negócios realizados são feitos por meio de um sistema eletrônico. Mais tarde, no final da década de 90, os *home brokers* passaram a ser utilizados [7], com isso o mercado de ações se popularizou entre os pequenos investidores, que puderam investir com menor custo de corretagem e de qualquer lugar que possuísse conexão com a internet.

As inovações tecnológicas que envolvem esse milionário mercado, vão além de facilitar os investimentos. O número de sistemas investidores, também conhecidos como *trading systems* tem crescido com o passar do tempo e, junto com o crescimento numérico de robôs que operam na bolsa, vem aumentando também a tecnologia utilizada para melhorar o desempenho desses investidores de silício [37].

Inicialmente, os *trading systems* constituíam-se de regras simples, do tipo se-então, que eram passadas ao computador. Pouco eficientes, essas regras, se aplicadas isoladamente, dificilmente traziam lucros. Começou-se então, a criar regras baseadas em análise técnica, como o cruzamento de Médias Móveis como sinal de reversão de tendência [2]. Tornou-se possível também a

---

utilização de grande quantidade de dados históricos para a previsão das tendências futuras: regressões lineares e não-lineares passaram a ser aplicadas no mercado financeiro [42]. O auxílio de modelos computacionais passou a ser primordial para a Análise Técnica [37], tornou-se muito mais fácil, prático e confiável a análise dos dados através dos inúmeros programas disponíveis no mercado. Além dos sistemas on-line na internet que possibilitam traçar, em instantes, Médias Móveis, Bandas de Bollinger, curvas do Estocástico [37], dentre outros, os *home brokers* geralmente já vêm acompanhados de todas essas ferramentas. No entanto, diferentemente dos sistemas de suporte à decisão, como os citados anteriormente, a programação dos *trading systems* é um pouco mais complicada, pois regras devem ser definidas como sinais de compra e venda. Essas regras tendem a ser difíceis de serem definidas, uma vez que necessitam de um especialista para registrar esses gatilhos. Ainda assim, mesmo com a participação de um especialista, a Análise Técnica possui componentes de difícil modelagem.

Um exemplo desta dificuldade de transformar a análise técnica em regras que fossem entendidas pelo computador está nos padrões japoneses de *candlestick*, em que palavras subjetivas fazem parte da regra, como: um pouco acima, muito abaixo. Os advérbios de intensidade são difíceis de serem quantificados e dependem, na maioria das vezes, apenas da experiência do grafista que analisa o gráfico. A Inteligência Computacional, no entanto, possibilitou a criação de sistemas que fizessem uso desses padrões. Alguns trabalhos que modelaram os padrões de *candlestick* com lógica nebulosa, como o de Lee [29].

Melhor ainda seria se o sistema pudesse observar os indicadores e padrões, de forma a aprender, por si só, como são definidos os gatilhos de compra e de venda. Nesse caso, além de não necessitar do especialista para definir os padrões, o que possibilita ao sistema ter desempenho superior ao que teria, caso fosse fechado nos conhecimentos de seu professor; não seria necessário tempo, por parte do usuário, para traçar retas e procurar por padrões nos gráficos, a fim de determinar um sinal de reversão de tendência.

## 1.1 Abordagem utilizada

Neste trabalho, desenvolveu-se um sistema de investimento baseado em Análise Técnica e em modelos de Aprendizado de Máquina. O sistema desenvolvido pode ser utilizado não só como suporte à tomada de decisão, em relação à compra e venda, mas também, como um sistema de investimento inteligente capaz de tomar as decisões e agir de forma autônoma. Diariamente,

após o fechamento do pregão, a BOVESPA disponibiliza em seu *site* na internet, um arquivo com a cotação de todas as empresas negociadas naquele dia. O sistema faz o *download* desse arquivo e, posteriormente, o seu *parser*, armazenando os dados em um banco de dados local. Com base nas informações do dia, o sistema diz quais ações são boas opções de compra e quais deveriam ser vendidas.

Para chegar-se ao produto final, um *trading system* completo<sup>1</sup>, algumas etapas tiveram de ser vencidas, sendo a mais interessante delas, quando detectou-se dados conflituosos no conjunto de treinamento. Estes conflitos poderiam prejudicar seriamente o treinamento do modelo e até inviabilizar a abordagem, portanto, precisavam ser minimizados.

Para que os sinais de compra e venda pudessem ser gerados, transformou-se a previsão de séries temporais em um problema de classificação. Essa transformação tornou-se possível ao se considerar a teoria dos padrões de *candlestick* [40], que se baseiam na detecção, na série de preço, de desenhos que, por sua vez, caracterizam a psicologia dos investidores em um dado momento. Os *candlesticks* têm como principal objetivo detectar pontos de reversão nas séries temporais; como esses momentos em que a tendência troca de direção caracterizam pontos de entrada ou saída do mercado, o foco do trabalho passou a ser detectar esses padrões.

Os padrões de *candlestick* ofereciam a generalidade desejada para o modelo, uma vez que são aplicáveis em qualquer série temporal de ações e em qualquer situação (por não terem o seu desempenho influenciado pela não estacionaridade das séries [40]), não se sabia, no entanto, se seriam gerados sinais de reversão em número suficiente para sustentar bons investimentos. No trabalho de Lee [29], citado anteriormente, utilizou-se um especialista para gerar a base de dados para treinar o modelo, o que o deixou limitado a um pequeno número de padrões - apenas aqueles conhecidos pelo especialista. Para que o sistema fosse capaz de operar apenas através dos padrões de *candlestick*, seria necessário que ele tivesse a liberdade de captar todos os padrões possíveis de reversão nas séries fornecidas para treinamento. Para que isso se tornasse possível, foi necessário desenvolver um sistema de rotulação automática das séries, para gerar um bom conjunto de treinamento.

Devido ao fato dos padrões de *candlestick* ocorrer, em alguns casos, dias antes ou depois da reversão da tendência, a transformação das séries de preço em amostras rotuladas inseriu muito ruído no conjunto de treinamento.

---

<sup>1</sup>Entende-se como completo o sistema que cuida de todas as fases de um investimento, desde a seleção de quais empresas irão compor a carteira de ações, até a geração de sinais de compra e venda

Dessa forma, a rotulação automática das amostras de treinamento podia não estar realizando o seu trabalho corretamente.

Atingiram-se bons resultados quanto à minimização dos dados ruidosos através da abordagem semi-supervisionada, que utiliza, além do conjunto rotulado de amostras, um conjunto sem rótulo, conhecido como grupo de trabalho. Este tipo de abordagem é geralmente utilizada em problemas em que as amostras de treinamento são escassas e caras, fazendo com que a utilização de amostras sem rótulos, que geralmente são abundantes e baratas, ajude na capacidade de generalização do modelo. No caso do modelo descrito neste trabalho, as amostras são abundantes, o que já o torna um caso especial da aprendizagem semi-supervisionada. No entanto, apesar de abundante, a rotulação das amostras não se mostrou confiável, apresentando um grande número de conflitos.

Nos testes realizados, o sistema agiu de forma autônoma, uma vez que o autor não possui nenhuma experiência de investimento em bolsa de valores. No entanto, acredita-se, também, na eficiência do sistema como apenas um suporte para o investidor. Os resultados trazidos pelo sistema foram significativos, no período de um ano, quando considerou-se uma carteira em que as ações foram criteriosamente selecionadas. O modelo transdutivo mostrou um rendimento de 192% no ambiente de simulação criado com exclusividade para este propósito. No simulador da Folha de São Paulo em parceria com a BOVESPA, o Em Ação <sup>2</sup>, o sistema também apresentou bons resultados: aproximadamente 30% em três meses de investimentos.

## 1.2 Organização do texto

Antes de qualquer tentativa de construção de um sistema de previsão, faz-se necessária uma revisão de literatura sobre a previsibilidade do mercado, tentando responder à seguinte questão: é possível prever os movimentos do mercado de modo a obter lucros? No capítulo 2, além de responder a essa pergunta, será feita uma revisão bibliográfica das principais abordagens encontradas na literatura para construção de sistemas de suporte à decisão, sistemas tutores e agentes inteligentes. No mesmo capítulo será feita uma introdução sobre a Análise Técnica, que é a principal ferramenta utilizada para pré-processar os dados nesse tipo de sistema.

Após o pré-processamento dos dados de entrada do sistema, será feita, no capítulo 3, uma revisão bibliográfica do modelo de aprendizado de máquina utilizado, o SVM. Nesse capítulo aborda-se-á a estratégia para que ele soluci-

---

<sup>2</sup><http://www.emacao.com.br>

onasse o problema de dados conflitantes, decorrente da separação de dados, descrita no capítulo 2, utilizando a abordagem semi-supervisionada.

No capítulo 4 serão mostrados os resultados alcançados pelo sistema, que investiu de forma independente, assim como resultados, tanto do simulador construído pelo autor, apenas para este propósito, como resultados do simulador do Folha em Ação. Até então, terá sido descrita a construção de um sistema gerador de sinais de compra/venda de ações, e, na seção referente à Otimização de Portfólio, será descrito um método de seleção de ações para compor a carteira operada pelo sistema, de forma a otimizar os lucros.

Finalmente, no capítulo 5, serão tecidas algumas discussões a respeito do desenvolvimento e resultados obtidos por este trabalho, propondo, obviamente, sugestões para trabalhos futuros.

---

# Pré-processamento dos dados

---

**U**m dos grandes desafios em desenvolver um sistema que seja capaz de prever acontecimentos futuros, em séries temporais, está em como pré-processar os dados que servirão de entrada para o treinamento e, posteriormente, para as tomadas de decisões. Além da entrada, deve-se definir a saída, o que espera-se do sistema, como resposta para uma determinada situação: poderia ser o valor da série daqui a uma semana, um mês ou quem sabe um ano? A infinidade de possibilidades que cerca tais decisões torna este projeto grandioso e desafiador.

No entanto, antes de tentar qualquer abordagem de previsão de séries financeiras e geração de sinais de reversão, faz-se necessário responder a seguinte questão: é possível prever os movimentos do mercado de modo a obter lucros?

## 2.1 Teoria do Mercado Eficiente

A teoria do Mercado Eficiente foi desenvolvida por Eugene Fama no início dos anos 60 na *University of Chicago Graduate School of Business* em sua tese de doutorado [14]. Ele defendia que informações que poderiam afetar o preço das ações se espalhavam rapidamente e eram incorporadas ao preço sem atraso. O mercado era eficiente e, por isso, não era possível superá-lo, ou em outras palavras, o lucro obtido através de aplicações viriam apenas por sorte. Por isso não era possível prever reversão de tendências ou valor de fechamento de um determinado dia. A melhor previsão para o dia seguinte era considerada como o valor de hoje, ou seja, depois do fechamento de um dia,

o dia seguinte era marcado por um ruído branco, cuja definição é conhecida como *random walk*.

Na primeira edição de seu livro, em 1973, Malkiel, afirma que um chimpanzé vendado seria capaz de selecionar uma carteira de ações tão boa quanto um especialista [34]. A verdade é que, se os investidores daquela época lucravam apenas por sorte, a sorte deles não dava sinais de que estava para ir embora. E até onde se sabe, nenhum chimpanzé, vendado ou não, enriqueceu seu dono aplicando na bolsa de valores.

Na década de 70, apesar de várias tentativas, os economistas não conseguiam refutar a hipótese do mercado eficiente, o que criou uma certa distância entre a prática e a teoria, pois investidores seguiam “com sorte” e lucrando, enquanto não se conseguia provar que era possível prever os movimentos do mercado. Mais tarde, na década de 80, percebeu-se que os testes não estavam sendo realizados da forma correta. Assumia-se a hipótese como verdadeira e os testes raramente tinham poder de refutar esta hipótese nula [42]. Foi então que propôs-se um modelo mais geral para representar as séries temporais financeiras, onde a *random walk* seria apenas um caso especial. A partir deste modelo, conseguiu-se negar a aleatoriedade para a maioria das séries financeiras [42]. Apesar desses trabalhos e outros do final da década de 80, como de Lo e Mackinlay [32], existem pesquisadores que ainda defendem a teoria. Um exemplo está em uma publicação mais recente de Malkiel [35] que já não afirma ser um macaco capaz de construir uma carteira de ações tão rentável quanto um especialista, mas, que o mercado é ainda muito mais eficiente do que previsível.

As provas matemáticas que rejeitam a Teoria do Mercado Eficiente estão além do escopo deste trabalho; no entanto, através das referências, é possível verificar as provas e constatar que é possível construir um sistema que seja capaz de prever movimentos do mercado e aplicar na bolsa de valores de forma a obter lucros. Serão analisados, nas próximas seções, os trabalhos encontrados na literatura que apresentaram bons resultados em tentativas de prever movimentos futuros do mercado.

## 2.2 Métodos clássicos

Inicialmente, talvez por influência dos métodos tradicionais (regressão) de previsão de séries temporais, usava-se como entrada apenas atrasos temporais da própria série analisada. Janelas de atraso de diferentes tamanhos eram utilizadas, e dava-se destaque quando encontravam-se relações não-lineares com a própria série. Uma alternativa muito mais eficiente passou

a ser utilizada, mercados com alta correlação passaram a ser considerados como possíveis entradas para o modelo [42], [11], no caso do trabalho de Refenes, em uma das várias abordagens realizadas pelo autor, algumas séries temporais são selecionadas como entrada para o modelo, como por exemplo, U.S. S&P, German Dax, French Cac, Dutch Eoe e Swiss SMI, e com estes índices tentou-se prever valores da FTSE. Através de uma análise de correlação, verificam-se as séries que possuem influência sobre a FTSE, ou são influenciadas por ela. A partir disso, através de Redes Neurais Artificiais [8], os autores procuraram, por relações lineares e não lineares, prever, com boa precisão, índices futuros. Abordagem parecida foi realizada no trabalho de Castro, quando, através de um estudo estatístico das séries temporais de mercados financeiros de todo o mundo, verificou-se quais tinham influência significativa no índice IBOVESPA. Dessa forma foi possível, utilizando também Redes Neurais Artificiais, prever, com ótima precisão, o valor de fechamento da BOVESPA. Além de ser uma informação preciosa para que os investidores tomassem suas decisões, a previsão do índice BOVESPA poderia ser utilizada em operações *intraday* (comprar e vender determinada ação em um mesmo dia), quando em uma previsão de fechamento em alta, ações que possuem alta correlação com o índice poderiam ser compradas, logo após a abertura do pregão, e vendidas no final do dia, caso a alta indicada pelo previsor se confirmasse.

Embora seja possível, é muito difícil lucrar com esse tipo de abordagem, pois, caso fosse utilizado, por exemplo, o sistema de previsão do índice BOVESPA para operações *intraday*, além de alto índice de acerto, o lucro só seria possível em investimentos com grandes variações (grande retorno). Pois, todo investimento necessita de um retorno mínimo para compensar os gastos decorrentes da aplicação do capital. Investimentos intradiários, nesse aspecto, carecem de especial cuidado, já que a carga tributária para este tipo de operação é excessivamente alta, e, se somados aos custos de corretagem, torna-se difícil a obtenção de lucros.

Caso as estratégias clássicas sejam aplicadas para investimentos de médio ou longo prazo, a série analisada deve ser específica, como o FTSE ou IBOVESPA nos exemplos anteriores. Com isso, cada empresa analisada, possivelmente, necessitaria de um modelo específico de aprendizado de máquina, o que seria um complicador para o desenvolvimento do sistema. Além disso, as séries são não estacionárias, ou seja, em um ano, a empresa analisada poderá ter obtido um bom desempenho, mas no ano seguinte poderá estar em uma tendência de baixa e, mesmo prevendo corretamente, não seria possível obter lucros com aquela empresa.

A não estacionaridade das séries temporais financeiras traz uma outra importante consequência a esse tipo de sistema: a necessidade de re-treinamento. Ao treinar um modelo de aprendizado de máquina para prever valores futuros de uma série temporal, esta-se captando as relações entre as entradas do modelo e a saída desejada. No entanto, as séries são não estacionárias e as relações existentes no momento do treinamento serão, provavelmente, diferentes das relações entre as mesmas variáveis depois de um certo tempo. Nesse caso seria necessário retreinar o modelo, de forma a captar os novos padrões. Todavia, até detectar a necessidade de retreino, muitos prejuízos podem ter sido contabilizados.

## 2.3 Análise Fundamentalista

Parte dos modelos de aprendizado de máquina procura se espelhar na natureza do desenvolvimento de seus processos de aprendizagem, como, por exemplo, Redes Neurais Artificiais, que se baseiam no neurônio natural [8], nos Algoritmos Genéticos, que tomaram como exemplo a capacidade adaptativa dos seres vivos segundo a teoria de Charles Darwin [19]; e na lógica fuzzy que busca valores matemáticos para os advérbios de intensidade [55]. Além da construção do modelo capaz de aprender, procura-se buscar inspiração na forma como o aprendizado e a tomada de decisão ocorrem na realidade. Investidores, geralmente, não utilizam apenas uma técnica, mas certamente, um arsenal de técnicas para melhorar a precisão de seus investimentos. Além de procurar por padrões que ocorreram no passado e podem voltar a ocorrer no futuro, os investidores ficam atentos às notícias que podem influenciar a variação de seus papéis na bolsa.

É muito comum, empresas de capital aberto disponibilizar na internet uma seção exclusiva para informação aos investidores, com notícias e relatórios financeiros sobre a empresa. Outra fonte de informação muito comum entre os investidores são os relatórios diários de *sites* especializados, contendo um resumo da economia mundial. Com base nessas informações consultadas pelos investidores, surgiram trabalhos [43], [47] e [12] que procuraram extrair informações importantes dessas fontes. No trabalho de Seo, as informações são recolhidas de sites de notícias bem conceituados, como CNN, Reuters, Financial Network, dentre outros; no caso de Thomas, a fonte é um dos fóruns mais movimentados do Estados Unidos, o *ragingbull.com*; Duarte, por sua vez, propôs um classificador de notícias sobre economia baseado em Redes Neurais Artificiais. Os três trabalhos tratam a previsão financeira como um problema de classificação de texto comum, em que níveis de classificação do

tipo bom, ruim e razoável são dados às empresas.

Provavelmente os melhores *trading systems* do futuro farão uso da análise fundamentalista, pois informações extremamente importantes podem ser encontradas na internet, antes mesmo de fazerem grandes efeitos no mercado. Uma das questões que permanece aberta é a fonte escolhida para se classificar. Quando, por exemplo, em 2006, anunciou-se a compra da Perdigão por parte da Sadia na Folha de São Paulo <sup>1</sup>, o valor da ação da Perdigão já tinha sofrido um ajuste positivo de, aproximadamente, 10%. Normalmente as informações chegam com um certo atraso nos grandes meios de comunicação, e as pessoas que detêm a informação privilegiada conseguem uma maior margem de lucro.

O ideal seria o agente ter mais de uma fonte de consulta. Se as informações podem chegar com atraso nos grandes meios, elas podem, simplesmente não chegar em meios menos convencionais como *blogs* e fóruns de discussão. Por estes motivos, a utilização de análise fundamentalista é muito mais do que um problema de classificação de texto. O agente coletor de informações deve ser eficiente, pois o volume de informações é muito grande e parte das informações não é de importância alguma. Além disso, o agente deve “entender” mais de uma língua, uma vez que os mercados mundiais estão fortemente interligados, e notícias externas podem ter maior impacto que notícias do país de origem da bolsa. Para atender a todos esses requisitos, é necessário recurso computacional, tanto para varrer a internet quanto para armazenar as informações que possam ser relevantes. Muita pesquisa tem sido desenvolvida nesta área de recuperação de informação, tendo maior iniciativa empresas como Google e Yahoo.

## 2.4 Análise Técnica

Uma alternativa que está entre os métodos tradicionais e a Análise Fundamentalista é a Análise Técnica. Através dela é possível pré-processar os dados de forma a facilitar a aprendizagem do modelo. Essa é a grande vantagem que o uso da Análise Técnica traz em relação aos métodos tradicionais. E quanto à Fundamentalista, sem dúvida, a vantagem é a facilidade de utilização, já que os indicadores e osciladores técnicos tratam-se, na sua maioria, de números, diferentemente da grande quantidade de texto utilizada na Fundamentalista.

Antes de discutir-se sobre quais as técnicas foram selecionadas para se utilizar neste trabalho e como se procedeu a seleção, é necessário discutir os princípios da Análise Técnica.

---

<sup>1</sup><http://www.folha.com.br>

### 2.4.1 Filosofia da Análise Técnica

Análise Técnica é o estudo da dinâmica do mercado, principalmente através de gráficos, com o propósito de prever tendências futuras de preço [40]. No caso do lote comum de ações, que foi o utilizado neste trabalho, a Análise Técnica trabalha com os valores históricos de preço (abertura, mínimo, máximo e fechamento) e o volume de negociações para fazer previsões a respeito do futuro.

Essa técnica parte de algumas premissas, são elas:

1. A dinâmica do mercado desconta tudo.
2. Os preços se movem em tendências.
3. A história se repete.

#### *A dinâmica do mercado desconta tudo*

Talvez esta premissa - a dinâmica do mercado desconta tudo - seja a base de toda a Análise Técnica. Com ela acredita-se que todas as informações necessárias para operar já estão contidas no preço, e, em um gráfico, já tem-se contido toda a análise fundamentalista. Ou seja, caso o gráfico esteja em uma tendência de alta, não importa se a causa foi política, econômica ou psicológica, o que realmente importa para os analistas técnicos é que a tendência é de alta. Pode-se afirmar, diante disso, que os grafistas (também são conhecidos por este nome) fazem, mesmo que indiretamente, análise baseada nos fundamentos. Um gráfico não faz com que haja, por exemplo, uma mudança de tendência, ele simplesmente reflete os fundamentos naquele momento.

Como uma regra, os grafistas não se importam com a razão de o mercado estar se movendo em determinada direção, afinal, não é importante saber porque o mercado está indo, mas simplesmente para onde ele está indo. E o que as ferramentas de Análise Técnica fazem é auxiliar o investidor a entender esses movimentos - quando estão para acontecer, qual a variação esperada - sem se preocupar com as razões que levaram a eles.

#### *Os preços se movem em tendências*

O corolário desta premissa, baseado na primeira lei de Newton, poderia ser: “É mais provável que uma tendência em movimento continue na direção original do que haver uma reversão” [40]. Por isso, caso uma tendência de alta seja detectada, talvez seja interessante aplicar o capital nesta ação, até que ela mostre sinais evidentes de reversão. Grande parte das ferramentas



Figura 2.1: Lojas Americanas - LAME4 - em tendência de alta

técnicas são seguidoras de tendência, o objetivo é determinar se uma série está em alta. A figura 2.1 mostra uma ação em tendência de alta.

### *A história se repete*

Ao estudar Análise Técnica, estuda-se, na verdade, a psicologia do ser humano. Sabe-se que o ser humano se comporta de maneira semelhante em determinadas situações, então, se no passado ocorreu determinado padrão no gráfico que precedeu uma tendência de alta, possivelmente, ao repetir novamente esse padrão, o futuro será semelhante ao passado. Este trabalho se orienta basicamente por essa premissa, ou seja, pela psicologia dos investidores.

### 2.4.2 *Diferenças entre Análise Fundamentalista e Técnica*

O princípio da Análise Fundamentalista diz que, se através de um estudo que determina o valor real de uma ação tem-se que aquela empresa está subvalorizada, então, compra-se papéis daquela empresa, até que ele se valorize e chegue ao valor considerado justo. No caso de a empresa estar super valorizada, espera-se até que ela atinja valores abaixo daqueles, para comprar. Na Análise Técnica não existe a possibilidade de se calcular um valor considerado justo, a análise é feita apenas com base em valores passados. Isso se mostra bastante útil quando existem grandes bancos manipulando o preço das ações, apesar de o preço já estar baixo, segundo a análise fundamentalista, ele poderá cair ainda mais, devido à manipulação. Já a Análise Técnica, enquanto não houver mostras significativas de que há uma reversão por vir, ela não indica compra do papel.

Murphy [40] afirma que é duvidoso que existam investidores que utilizem apenas Análise Fundamentalista como fonte de informações para tomada de

decisão sobre os investimentos. Ele utiliza alguns argumentos para sustentar essa afirmação, sendo um deles é que, historicamente, em grandes quedas, não há motivo fundamentalista aparente no início do movimento. E, talvez o argumento mais forte, é pelo fato de a Análise Fundamentalista ser considerada um subconjunto da Técnica [40]. Com isso, ao utilizar Análise Técnica, mesmo que indiretamente, está-se utilizando Análise Fundamentalista.

Por esses motivos os trabalhos encontrados na literatura, baseados em classificação de texto - Análise Fundamentalista - são, na sua maioria, apenas suporte à tomada de decisão, sendo de difícil aplicação direta no mercado de ações. O objetivo deste trabalho foi criar um sistema que, além de dar suporte à tomada de decisão, fosse possível aplicar de forma autônoma, identificando os momentos de compra e venda. Por isso, a Análise Técnica mostrou-se como uma boa opção de pré-processamento dos dados de entrada.

### 2.4.3 Janela de tempo para Análise Técnica

Um dos primeiros parâmetros a ser definido antes de se construir qualquer *trading system* é qual será o seu horizonte de investimento, ou, em outras palavras, depois de uma compra, quanto tempo deve-se permanecer com as ações, visando sempre a maior margem de lucro. Respondendo a esta questão, provavelmente, irá se restringir as possibilidades de técnicas que poderão ser utilizadas como entrada do sistema. Análise Técnica oferece grande flexibilidade quanto ao horizonte de investimento, podendo ser utilizada em estratégias de *day trade*, em que, os papéis são comprados e vendidos no mesmo dia, incluindo investimentos de longo prazo, do tipo “comprar e esperar”.

Em um gráfico de preço, onde se tem a função  $f(.)$  de preço por tempo, ou,  $p = f(t)$ , para cada instante  $t$  têm-se representado os valores de abertura, fechamento, mínimo e máximo do preço  $p$ . A figura 2.2<sup>2</sup> mostra o gráfico de preço das ações preferenciais da Cemig - CMIG4 - no ano de 2007. O traço para a esquerda, em cada barra, indica o preço de abertura, para a direita o de fechamento, o máximo e mínimo são representados pelo valor máximo e mínimo da barra.

No caso da figura 2.2, cada barra representa um dia de negociação, desde o momento em que abre o pregão até o seu fechamento, mas poderia, perfeitamente, representar outras janelas de tempo. Não existem restrições para o tempo representado pela barra de preço. Um negociador intradiário, ou *intraday trader*, geralmente usa uma janela de 15 minutos e, com base nesses valores procura descobrir os padrões que revelam o momento de compra e

---

<sup>2</sup>Extraído de <http://br.advfn.com>

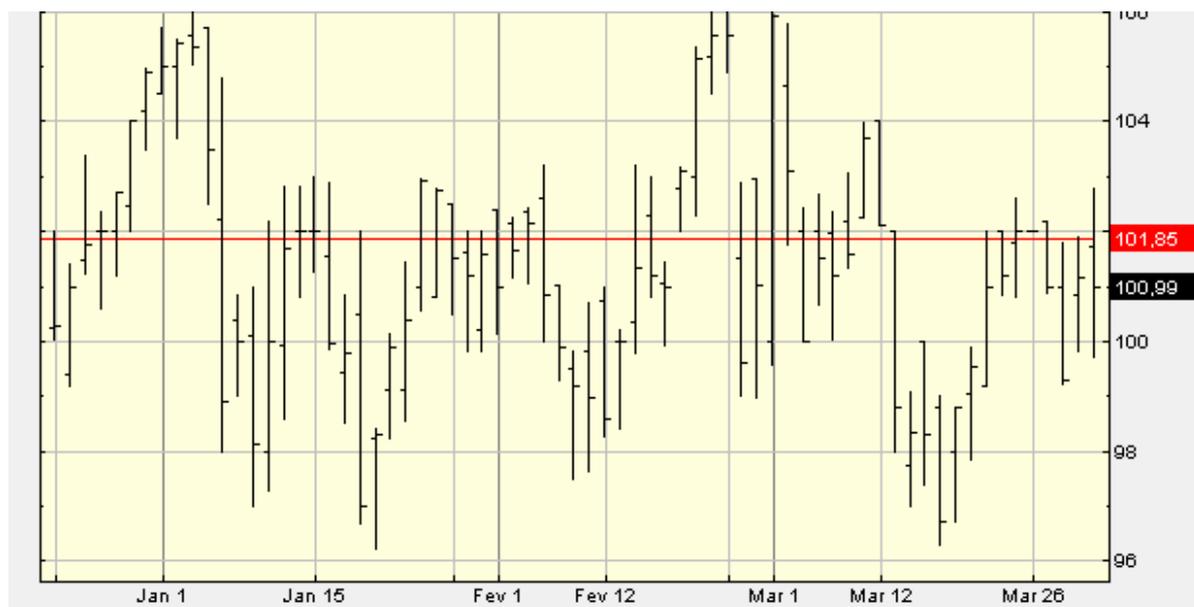


Figura 2.2: Gráfico em barras da Cemig - CMIG4 - em 2007

venda da ação. A utilização, neste trabalho, da menor janela de tempo possível, geraria mais transações, e, possivelmente, um lucro maior. Entretanto, existem várias restrições, como disponibilidade de dados históricos e alta taxa tributária para a utilização de dados intradiários, por esse motivo utilizou-se valores diários.

### 2.4.4 Teoria de Dow

No final do século XIX e início do XX Charles Dow publicou em uma sequência de editoriais, para o *Wall Street Journal*, teorias que se tornaram a base de vários indicadores técnicos[40]. Por isso, antes de qualquer técnica grafista, é interessante mencionar as principais idéias que compõem a Teoria de Dow.

*O movimento pode ser dividido em três tendências*

Segundo Charles Dow, o movimento do preço de uma ação pode ser dividido em três tendências diferentes, a primária, também conhecida como de longo prazo, tendo a duração aproximada de um a dois anos [40]. A secundária dura alguns meses e é considerada como uma tendência de médio prazo. Por último, a terciária, que geralmente possui duração de alguns dias ou semanas [37]. Na figura 2.3 é possível conferir a diferença entre as três tendências.

Neste trabalho, optou-se por investir com base na tendência terciária, que é a de curto prazo. O principal motivo desta escolha é que seria possível lucrar mesmo em uma tendências primária ou secundária de baixa, pois, é muito comum em uma tendência bem definida ocorrer o que os analistas chamam de



Figura 2.3: Tendência primária, secundária e terciária

*pullback*, ou um movimento breve na direção contrária a da tendência principal. Dessa forma é possível entrar e sair rapidamente em uma ação em queda e ainda obter lucro no breve movimento contrário. No capítulo referente aos resultados será mostrado um caso, em que conseguiu-se quase 4% de lucro em uma tendência secundária de forte baixa.

### *A tendência principal tem três fases*

A tendência principal pode ser dividida, segundo Dow, em três fases diferentes [37]:

1. Acumulação - Nesta fase apenas uma pequena minoria começa a se posicionar favorável à tendência, provavelmente trata-se de um grupo de pessoas que detém informação privilegiada.
2. Alta Sensível - Uma leve alta já pode ser notada, e indicadores técnicos já conseguem perceber a reversão na tendência, sinalizando o momento ideal para os operadores, seguidores de tendência, entrarem no mercado.
3. Euforia - Indica o final da tendência de alta, é quando os investidores leigos entram no mercado.

A figura 2.4 mostra claramente as três fases da tendência de alta.

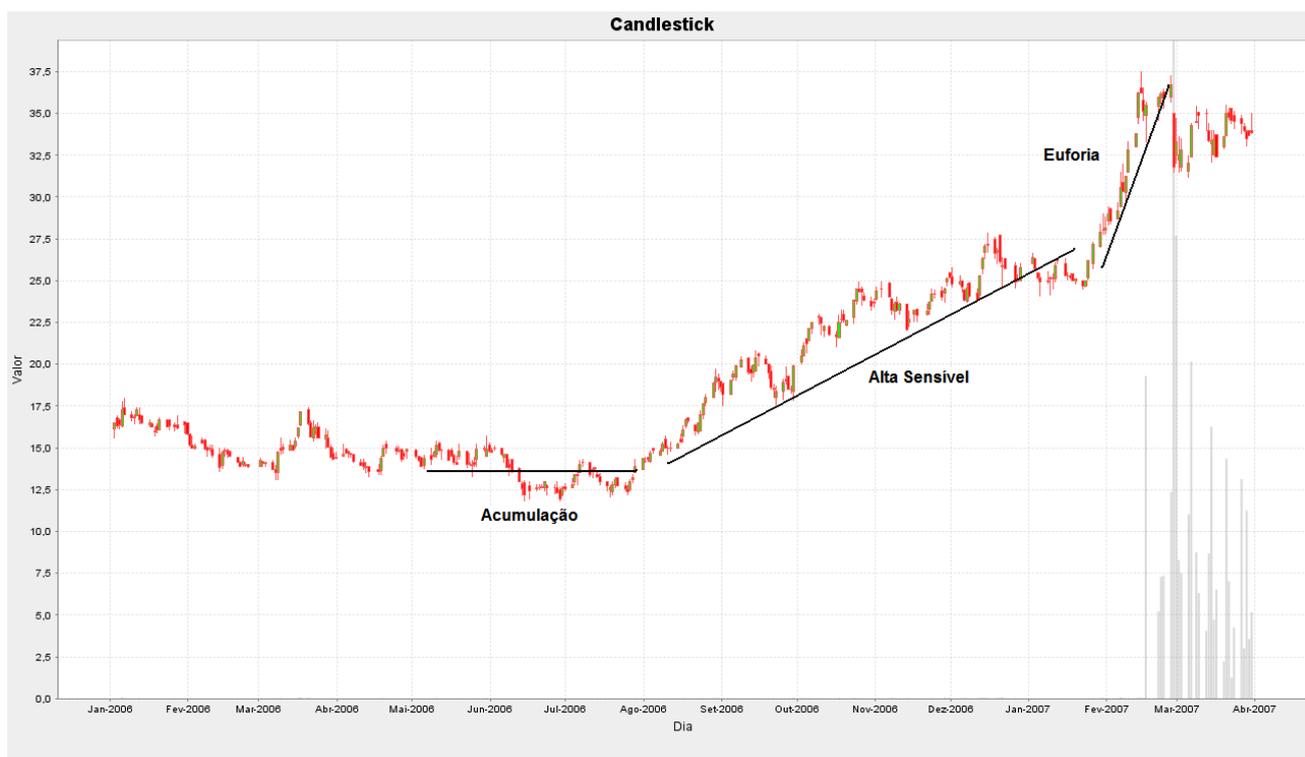


Figura 2.4: As três fases em um tendência de alta - PMAM4

Analogamente, as três fases se aplicam a uma tendência de baixa, com os seguintes nomes: Distribuição, Baixa Sensível e Pânico.

*O volume confirma a tendência*

Apesar de considerar como um indicador secundário, o volume, para Dow, representa um confirmador de final ou início de tendência [40]. Para manter uma tendência, seja ela de alta ou baixa, é necessário que haja um número suficiente de investidores negociando e mantendo a tendência. Quando há divergência entre o volume e o preço, a indicação é de que uma reversão está próxima [37]. A figura<sup>3</sup> 2.5 mostra a divergência entre o volume e o valor do Índice Bovespa, indicando uma reversão de tendência terciária.

Padrões de reversão de tendência sem a confirmação através do volume podem indicar *pullbacks* rápidos, provavelmente causados pela manipulação do papel por grandes investidores. Como já dito anteriormente, este trabalho tentará aproveitar as tendências terciárias, que também podem ser causadas por manipuladores, por isso o volume não foi utilizado como entrada do sistema de predição de reversão de tendência.

<sup>3</sup>Extraído de <http://www.infomoney.com.br>



Figura 2.5: Divergência entre o volume e preço, indicando uma reversão

*A tendência acaba com sinais definitivos de reversão*

Segundo Dow, é preferível esperar um pouco para se ter certeza que houve uma reversão na tendência a sair antes do final da onda de alta. Esperar um pouco mais significa diminuir a margem de lucros, por outro lado, sair antes de a tendência acabar pode significar um abatimento maior ainda [37]. O investidor que segue esse tipo de estratégia, de entrar somente quando tiver certeza de que há uma tendência de alta e sair somente quando também tiver confirmações de que a tendência acabou é conhecido como seguidor de tendência. Existem vários indicadores técnicos que auxiliam os investidores a seguir as tendências da melhor forma possível.

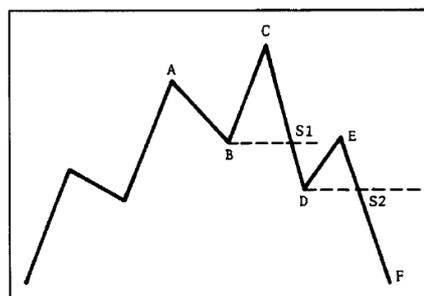


Figura 2.6: Sinais de final de tendência

A figura<sup>4</sup> 2.6 mostra momentos que podem ser considerados sinais claros de final de uma tendência de alta. Quando entre C e D perde-se o ponto S1, alguns seguidores de tendência já dariam como certa a reversão. No entanto, alguns investidores preferem esperar até que o ponto S2 seja perdido, caracterizando, dessa forma, a reversão para a tendência de baixa.

### *Críticas quanto à teoria de Dow*

A maior crítica sofrida pela Teoria de Dow é que, na média, perde-se de 20% a 25% em cada investimento, por entrar e sair no mercado tardiamente. No entanto, nunca foi pretensão de Dow antecipar uma reversão de tendência para se conseguir o lucro máximo com aquele movimento do mercado.

O objetivo deste trabalho não é construir mais um *trading system* seguidor de tendência, pretendeu-se construir um modelo de previsão, que antecipasse as reversões de tendência terciárias. Por esse motivo, grande parte das técnicas grafistas que surgiram, a partir da Teoria de Dow, foram descartadas como possibilidade de entrada para o sistema.

### 2.4.5 *Trabalhos que utilizaram Análise Técnica*

Alguns trabalhos que apresentaram bons resultados e utilizaram Análise Técnica como fonte de pré-processamento para as entradas dos dados nos modelos de predição serão citados nesta seção.

Tsaih, Hsu e Lai [49] construíram um sistema de suporte à decisão, para investimentos na S&P 500, baseados em regras pré-definidas por especialistas. Se algum gatilho fosse disparado, com base nestas regras, os dados que o dispararam seriam utilizados como entradas de uma Rede Neural Artificial, as saídas, em forma de um problema de classificação, traziam informações sobre as movimentações futuras do mercado. Indicadores como Índice de Força Relativa, Médias Móveis e Estocástico foram utilizados [49]. Faz-se necessário, já que alguns serão utilizados mais adiante nesse texto, o detalhamento de cada um desses indicadores.

### *Índice de Força Relativa*

Este índice compara as forças de compradores e vendedores com base no preço de fechamento da ação, trazendo assim, informações relevantes para os analistas técnicos. As equações 2.1 e 2.2 detalham esse indicador [37]. *FR* define a razão entre as médias das variações de alta e baixa em um dado

---

<sup>4</sup>Extraído de [40]

intervalo de tempo (utiliza-se, geralmente, 14 dias). A equação 2.2 transforma os valores de  $FR$  (para facilitar a interpretação do indicador) para valores entre 0 e 100.

$$FR = \mu_{va} / \mu_{vb} \quad (2.1)$$

onde,

$\mu_{va}$  = Média das variações de alta

$\mu_{vb}$  = Média das variações de baixa

$$IFR = 100 - (100 / (1 + FR)) \quad (2.2)$$

Com base no índice de força relativa, pode-se perceber quem está prevalecendo no mercado, se a força dos comprados, ou a força dos vendidos. Quanto mais o  $IFR$  encontra-se próximo de 100, maior a força dos comprados, por sua vez, quanto mais próximo de 0, maior a força dos vendidos. Por esse motivo, esse indicador pode ser utilizado como um alerta de que o papel está sobrecomprado ou sobrevendido. Valores próximos a 100, indicam que o papel já foi muito comprado e mostra que o momento de reversão, nem que seja de um *pullback*, está próximo. Quando próximo a 0, a informação é que ele já foi muito vendido e um momento de reversão para alta está próximo.

Uma outra forma de utilização do índice de força relativa é quando há divergência entre ele e o gráfico de preço. Quando tem-se uma tendência de alta, e o índice de força relativa, que também era de alta começa a cair, a demonstração é que o final da tendência está próxima, pois os comprados estão perdendo força para os vendidos. A figura 2.7 mostra um exemplo que esta divergência antecipou o momento de reversão.

#### Estocástico

O indicador Estocástico relaciona o preço de fechamento atual com as maiores máximas e menores mínimas de um dado intervalo de tempo [37]. Acredita-se que, quando em uma tendência de alta, o valor da máxima se aproxima do fechamento, e o mesmo ocorre em uma tendência de baixa, quando o fechamento se aproxima da mínima. A fórmula 2.3 mostra os detalhes do indicador.

$$\%K = \left( \frac{F_t - Min_{t-N}}{Max_{t-N} - Min_{t-N}} \right) * 100 \quad (2.3)$$

onde,

$F_t$  = Fechamento Atual



Figura 2.7: Divergência entre o gráfico de preço e o indicador de índice de força relativa - VALE5

$Min_{t-N}$  = Menor Mínimo em uma janela de tempo de N dias

$Max_{t-N}$  = Maior Máximo em uma janela de tempo de N dias

O valor de  $N$  utilizado, geralmente, é de 14 dias.

Utiliza-se uma média do indicador estocástico, conhecida como  $\%D$  para indicar sinais de compra e venda. Quando o estocástico cruza sua média de baixo para cima, indica um momento de compra; na situação reversa, quando o cruzamento ocorre de cima para baixo, um sinal de venda é lançado. No entanto, essa utilização do estocástico só funciona bem quando utilizado em séries estacionárias, caso haja uma tendência definida, muitos sinais falsos serão lançados.

Assim como o índice de força relativa, o Estocástico pode ser utilizado como um indicador de sobrecomprado e sobrevendido. Os investidores geralmente utilizam as seguintes faixas:

$$r = \begin{cases} \text{sobrecomprado se } \%K \geq 80 \\ \text{sobrevendido se } \%K \leq 20 \end{cases} \quad (2.4)$$

### Médias Móveis

Segundo Charles Dow, o que importa é a média dos movimentos e não as grandes variações ruidosas que ocorrem, mesmo em tendências bem definidas



Figura 2.8: Vários sinais de compra e venda são falsos quando a série possui uma tendência definida

[40]. Ao retirar a média do preço de uma ação, o que se faz é retirar uma curva mais comportada que define aquele movimento. A figura 2.9 mostra a média móvel de 15 dias de ITSA4.



Figura 2.9: Média Móvel de 15 dias de ITSA4

As médias móveis possuem diversas aplicações na Análise Técnica. Com ela pode-se determinar, mesmo com um atraso considerável, reversão de tendências. Essas reversões podem ser determinadas quando cruzam-se médias móveis de diferentes janelas de tempo, ou até mesmo quando há interceptação da média com o valor de preço. Além de determinar os momentos de reversão, pode-se utilizar as médias como suporte ou resistência para os papéis. Por

exemplo, na figura 2.9, existe uma tendência de que o preço não caia abaixo da média móvel de 15 dias. Nesse caso, considera-se a média como um suporte. No entanto, o preço indo abaixo dessa média, indica que a ITSA4 está realmente em uma tendência de baixa e, no futuro, para mostrar a volta a tendência de alta, o preço deve "vencer" a média móvel de baixo para cima. Neste último caso, ela estará representando uma resistência para a subida do papel.

Tsaih, Hsu e Lai [49] transformaram os valores destes indicadores em valores binários, segundo certas condições. Por exemplo, ele utiliza 4 variáveis de entrada somente para o estocástico, RSI1, RSI2, RSI3 e RSI4. A primeira delas terá valor igual a 1, caso o estocástico tenha caído de 100 e valerá -1, caso contrário. A mesma estratégia foi utilizada para as médias móveis, a fim de detectar cruzamentos com a séries de preço. É um tipo de abordagem que simplifica as entradas e já pré-processam os dados de forma a facilitar a aprendizagem do modelo, no caso, Redes Neurais Artificiais. No entanto, este pré-processamento pode limitar o modelo, neste mesmo trabalho, no caso do Índice de Força Relativa, a divergência entre a variação do preço e a variação do índice não podem ser captados pela rede, devido à característica discreta dos valores de entrada.

Com exceção das Médias Móveis, os dois osciladores - Índice de Força Relativa e Estocástico - poderiam ter sido utilizados como entrada sem transformação alguma no modelo. Ambos variam entre 0 e 100, ou seja, são genéricos para qualquer tipo de índice ou ação, portanto, a sua utilização não faria o sistema perder em generalidade. No caso das médias, algumas ações custam poucos centavos, enquanto outras, como da AMBV4, custam em torno de R\$1000,00, neste caso, algum tipo de transformação faz-se necessária.

Alguns trabalhos encontrados na literatura [23] [54] [57] utilizaram como entrada do sistema de previsão, vários indicadores técnicos, como, Estocástico, Momento, Índice de Força Relativa, Disparidade e outros. O grande problema em utilizar uma série de indicadores está na perda de generalidade do sistema.

No caso de Jae Kim, [23] em que o objetivo fora prever movimentos futuros da KOSPI - Índice da Bolsa Koreana - o sistema ficou limitado somente àquele índice. Nesse caso, o sistema foi desenvolvido com apenas este propósito, tanto que, no treinamento, utilizaram-se apenas dados daquele índice.

Em uma das abordagens descritas no livro de Zirilli [57] dissertou-se sobre a construção de um modelo em que o objetivo era prever o valor de fechamento de determinadas ações, dado os indicadores técnicos e o valor de abertura do dia em questão. Ao utilizar o valor de abertura da ação, elimina-se o *gap*, ou

buraco, que é causado por acontecimentos noturnos, entre os pregões. Dessa forma, era possível obter lucros da seguinte forma: caso o valor fosse acima do valor de abertura, comprava-se a ação; caso contrário, mantinha-se fora do mercado.

Os trabalhos, descritos acima, apresentaram bons resultados, no entanto, não aproveitaram uma das maiores vantagens da Análise Técnica, que é a generalidade. Este foi um dos principais objetivos do sistema desenvolvido neste trabalho: desenvolver um sistema tão geral quanto possível. Caso, por exemplo, utilizasse a abordagem de Zirilli para prever valores de fechamento das ações preferencias da Petrobrás, o modelo resultante lidaria apenas com aquelas ações e, provavelmente, não apresentaria bons resultados com outros papéis.

### 2.4.6 Padrão Japoneses de Candlesticks

Uma das chaves para a generalidade do sistema desenvolvido neste trabalho foi encontrada nos padrões Japoneses de *candlestick*. Devido a sua importância no trabalho, selecionou-se uma seção para revisão bibliográfica dos trabalhos que também utilizaram essa técnica.

*Candlesticks* podem ter dois significados diferentes dependendo do contexto, um é forma de mostrar o gráfico de preços e outro, são os padrões que podem ser encontrados nesses gráficos.

#### O gráfico

A figura 2.2 mostra o gráfico de preço das ações preferenciais da Cemig - CMIG4. Nele é possível observar os valores de abertura, fechamento, máximo e mínimo da ação. No entanto, a visualização do gráfico é um pouco dificultada pelo formato das barras, às vezes é difícil observar a que preço se deram a abertura e fechamento. Os *candlesticks* facilitam a visualização da mesma informação, é como se o gráfico "saltasse aos olhos", facilitando a leitura dos valores.

Assim como o gráfico de barras, o *candle* é formado pelos valores de abertura, fechamento, máximo e mínimo, como mostrado na figura 2.10. O corpo claro, ou vazado, indica um dia de alta. O corpo escuro, ou preenchido, indica um dia de baixa. O gráfico da CMIG4 em *candlestick* pode ser observado na figura<sup>5</sup> 2.11.

É interessante notar como no gráfico em *candlestick* fica mais fácil a visualização da variação do preço do papel. Anteriormente foi dito que usa-se

---

<sup>5</sup>Extraído de <http://br.advfn.com>

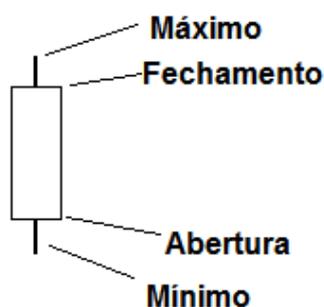


Figura 2.10: Formato de um *candlestick*



Figura 2.11: CMIG4 no gráfico de *candlestick*

o termo *candlestick* com dois significados diferentes. O primeiro, e talvez o mais comum deles já foi apresentado. Resta agora, apresentar os padrões mostrados por estes gráficos.

### Os padrões

O corpo de um *candle* é definido pelo espaço entre o valor de abertura e fechamento, o cabelo, ou sombra, são as linhas que definem o valor máximo e mínimo do dia. Cada combinação diferente destas figuras possui um significado diferente.

É conhecido como um dia longo, aquele que possui um corpo grande, ou em outras palavras: a diferença entre o preço de abertura e fechamento é grande. Analogamente conhece-se como dia curto aquele em que a abertura está próxima do fechamento, fazendo com que o *candle* tenha um corpo pequeno.

Ressalta-se que até agora não se fez menção ao tamanho das sombras; elas desempenham, no entanto, importante papel na formação dos padrões. Dias curtos em que as sombras são aproximadamente do mesmo tamanho que do corpo, por exemplo, representam dias de indecisão por parte dos investidores, figura<sup>6</sup> 2.12.



Figura 2.12: Dias de indecisão

Dias em que o preço de abertura é igual ao preço de fechamento são conhecidos como *Doji Candlesticks* e cada um deles tem um significado diferente. Na figura<sup>7</sup> 2.13 pode-se observar quatro padrões diferentes de *doji*, quando o cabelo, ou a linha que indica o valor máximo, é tão grande quanto a sombra, indica um dia de indecisão. Quando o *doji* apresenta uma longa linha supe-

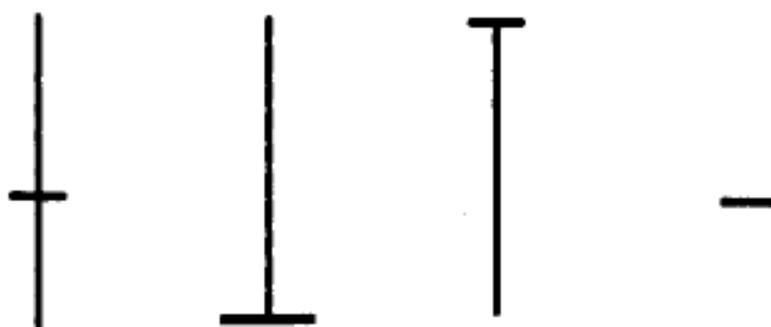


Figura 2.13: *Doji Candlesticks*

rior e nenhuma inferior, indica-se tendência de baixa e, quanto maior a linha, mais forte será essa tendência. Mostra-se como uma tendência de baixa porque, em algum ponto do dia, o valor do preço esteve alto, mas começou a cair e fechou em forte baixa em relação àquele momento de alta. De forma análoga ocorre quando existe uma longa linha inferior, mas nenhuma superior, o que mostra forte recuperação no intradiário e possível tendência de alta.

---

<sup>6</sup>Extraído de [40]

<sup>7</sup>Extraído de [40]

As leituras dos padrões são feitas de forma semelhante às da figura 2.13: isoladamente, ou combinando uma sequência de *candles*.

Até mesmo os nomes utilizados para identificar os padrões possuem um efeito psicológico, por exemplo, Homem Enforcado ou Nuvem Negra demonstram sinais de apreensão, quando o investidor deveria se proteger, provavelmente retirando o seu dinheiro do mercado. Como era de se esperar, esses dois exemplos indicam reversão para tendência de baixa. Além do reflexo emocional contido nos nomes, eles mostram grande influência dos tempos de guerra que viveu o Japão nos séculos XVI e XVII, como Três Soldados ou Lápide. Os padrões surgiram logo após esse período conturbado, no século XVIII, quando o Japão foi unificado e o comércio de arroz se transformou em um ótimo negócio; deu-se, então, o início da Análise Técnica e as tentativas de prever preços futuros, no caso, de arroz [41].

Existem padrões que determinam reversão de tendência e outros que, por sua vez, determinam continuação de tendência. Alguns dos padrões de reversão são os mesmos, ou seja, possuem o mesmo desenho, tanto para uma mudança de alta para baixa quanto de baixa para alta. O tipo de reversão que ocorrerá será determinado pela tendência em que a série se encontra. Por exemplo, não faz o menor sentido aparecer um padrão de reversão para alta, em uma série que já está em alta. Por isso a tendência torna-se muito importante para a identificação computadorizada dos padrões *candlesticks*.

Os padrões de continuidade, apesar de minoria, são muito importantes para que se faça um bom investimento, pois, sair de uma tendência de alta, antes do momento correto, pode ser tão prejudicial quanto deixar de pegá-la. Pior que isso, talvez seja comprar em baixa esperando por uma reversão de alta e a reversão não aparecer, nesse ponto é que os prejuízos aparecem.

Na literatura encontram-se registrados, aproximadamente, quarenta padrões de reversão que, geralmente, possuem entre um a cinco *candles*. A realização deste trabalho sinaliza que esse número seja algumas vezes maior do que o registrado, possivelmente por desconhecimento, ou falta de capacidade do ser humano em captar esses padrões, e também, por haver, padrões que são utilizados, mas não estão catalogados.

### *Nuvem Negra*

Como dito anteriormente, este padrão demonstra uma tendência de reversão para baixa. Ele ocorre quando em alta: um papel abre o pregão com o seu preço acima do fechamento anterior, indicando ainda a tendência de alta. No entanto, no decorrer do dia o preço cai, fechando, pelo menos, até na altura da metade do *candle* anterior. A figura 2.14 mostra os detalhes do padrão.



Figura 2.14: Nuvem Negra, um padrão de reversão

#### *Linha Perfurante*

Geralmente os padrões de *candlestick* apresentam padrões similares, tanto para reversão de alta para baixa, quanto de baixa para alta. O correspondente da Nuvem Negra de reversão de baixa para alta é a Linha Perfurante que, de maneira análoga, mostra o momento de reversão, quando os comprados passam a superar a força dos vendidos. As similaridades com o padrão Nuvem Negra podem ser conferidos na figura 2.15.

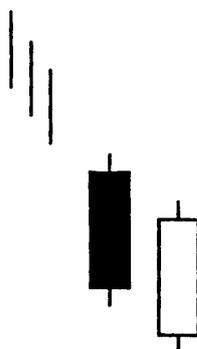


Figura 2.15: Linha Perfurante, um padrão de reversão

#### *Outros Padrões*

Mais exemplos de padrões de reversão podem ser conferidos nas figuras 2.16 e 2.17.

É interessante observar, mais uma vez, os nomes, como "Estrela da Manhã" e "Estrela da tarde". O primeiro é uma conotação do surgimento dos primeiros raios do sol, trazendo esperança aos investidores por uma mudança para uma tendência de alta. O segundo, por sua vez, traz a escuridão com o por-do-sol, juntamente com a preocupação dos investidores em presenciar uma reversão para baixa.

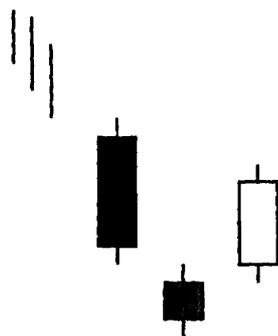


Figura 2.16: Estrela da Manhã, um padrão de reversão



Figura 2.17: Estrela da Tarde, um padrão de reversão

Existem publicações de trabalhos cujos sistemas, baseados em *candlesticks*, foram desenvolvidos utilizando-se lógica nebulosa [29]. Com ajuda de um especialista, construiu-se uma base de dados confiável que poderia ser utilizada como um suporte à decisão do investidor. Este é o tipo de sistema conhecido como caixa branca que, além de sua confiabilidade no suporte à decisão, permite que o sistema seja utilizado como ferramenta educacional, pois um investidor inexperiente terá oportunidade de aprender os padrões ao utilizar a aplicação. Os autores demonstraram a grande utilidade do modelo, ao construir um sistema tutor com interface gráfica baseado no trabalho descrito anteriormente [30]. Essa abordagem possui, entretanto, algumas desvantagens significativas: além de necessitar de um especialista em *candlesticks*, o sistema estaria limitado aos conhecimentos desse especialista. Não podendo, desta forma, extrapolar para os padrões ainda não aprendidos ou catalogados pelos investidores. Por isso, neste trabalho, não se utilizou abordagem parecida com a utilizada por Lee e colegas [29] [28] [30]. Sabe-se que existem aproximadamente 50 padrões de *candlestick* catalogados, entre padrões de reversão e continuação. Caso o sistema fique limitado a esses padrões, provavelmente não seriam detectados pontos de reversão suficientes para obtenção

de lucro satisfatório.

### Figuras

Além dos padrões discutidos anteriormente, é interessante citar outros trabalhos que detectam tipos diferentes de padrões nos *candles*. Por exemplo, Kamijo e Tanigawa, construíram um sistema que detectava os padrões de triângulo nas séries de preço [27]. Figuras como os triângulos (ver figura 2.18<sup>8</sup> para mais detalhes) precedem grandes variações no mercado, podendo tanto ser uma mudança para tendência de alta, como mudanças para tendência de baixa.

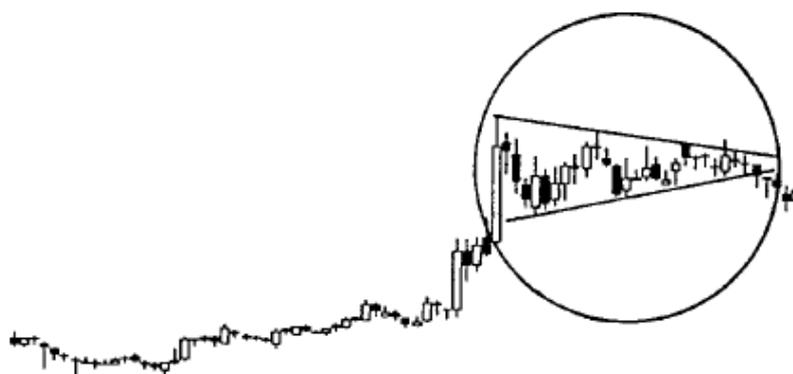


Figura 2.18: Triângulo, uma figura de reversão

O trabalho de Kamijo e Tanigawa, apesar de mostrar bom desempenho, das 16 figuras testadas no sistema, acertou 15, apresenta as mesmas desvantagens do sistema proposto por Lee [28], que é a necessidade de um especialista para formar a base de conhecimento, fazendo com que o desempenho do sistema fique limitado ao especialista. Além disso, os triângulos geralmente não fornecem informação suficiente para gerar um sinal de compra/venda, sendo assim, mais adequados para sistemas de suporte à decisão.

Além dos triângulos, existem outras figuras que podem ser aprendidas por sistemas inteligentes, como o mostrado pela figura 2.19, o Ombro-Cabeça-Ombro é uma importante figura de reversão. Existem também as Bandeiras, Flâmulas, Topos Duplos, Ombro-Cabeça-Ombro Invertido e outras.

O grande desafio em construir um modelo que seja capaz de aprender tais figuras está na diferença de dimensão das entradas. A figura 2.19 mostra uma figura que foi “desenhada” com seis dias de negócios, mas poderia ter sido construída com um número indefinido de dias, dificultando a modelagem do sistema.

---

<sup>8</sup>Extraído de [27]

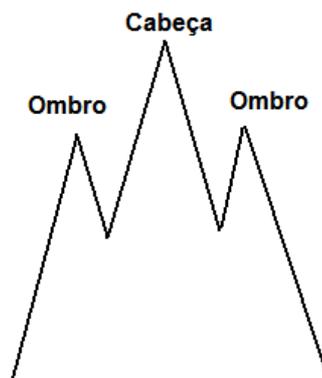


Figura 2.19: Ombro-Cabeça-Ombro, importante figura de reversão

## 2.5 Abordagem utilizada neste trabalho

No geral, sistemas construídos com auxílio de um especialista para formar a base de dados são de suporte à decisão e, portanto, não são capazes de operar de forma autônoma, objetivando lucros. O objetivo deste trabalho é que o sistema aprenda com dados históricos e seja capaz de lucrar com o que aprendeu, não sendo requisito essencial que ele possa ser utilizado como um sistema tutor, ou que seja possível extrair as regras que foram aprendidas por ele. Durante a fase de testes, como explicado anteriormente, o sistema foi utilizado como um *trading system* autônomo, que identificava os momentos de comprar e vender. Mas sabe-se também, apesar de não ter sido testado, que ele poderá ser utilizado como suporte à decisão.

Os sistemas podem ser classificados em duas categorias: aqueles que selecionam quais ações utilizar para investir e os que dizem o momento de compra e venda dado um conjunto pré-definido de ações [31]. Preocupou-se com as duas categorias de sistema, neste trabalho. Em um primeiro momento será detalhado como foi treinado o modelo para indicar o momento correto de compra e venda, posteriormente, no capítulo referente a resultados, na seção de Otimização, será detalhado sobre a escolha das ações que irão compor a carteira de investimentos do sistema.

### 2.5.1 Automatização do rotulamento

Questões importantes a respeito da estratégia de investimento já foram respondidas, como por exemplo, o horizonte temporal de investimento. Como o objetivo é investir em mudanças na tendência terciária, o horizonte temporal é de curto prazo, estimando-se investimentos de na média entre 5 a 10 dias. Espera-se com isso que seja possível, apesar de não aconselhável, obter lucro em ações que estão em tendência primária ou secundária de baixa, no

momento de seu *pullback*.

Como citado anteriormente, esperava-se que o sistema descobrisse novos padrões nos *candles*, de forma a aumentar o número de sinais de reversão gerados e com isso fosse possível construir um *trading system* baseado apenas nos padrões de *candlestick*. Para que isso ocorresse, foi necessária uma grande quantidade de dados para efetuar o treinamento. Para tanto, utilizaram-se séries de empresas com capital aberto na BOVESPA desde 1986 até o ano de 2005, sendo que o ano de 2006 fora reservado para validação e testes<sup>9</sup>.

Um *parser* foi desenvolvido na linguagem Java<sup>10</sup> para separar todos esses dados e armazená-los de forma organizada e de fácil consulta em um Banco de Dados relacional. Para tanto, utilizou-se o HSQLDB<sup>11</sup>, que é um gerenciador de banco de dados desenvolvido também em Java e de código aberto. Bastante leve, ele pode ser embutido na aplicação e distribuído juntamente com a mesma. Para agilizar o desenvolvimento do sistema, facilitando as consultas, inserções, modificações e a própria criação do esquema do banco de dados, utilizou-se o Hibernate 2<sup>12</sup>, juntamente com XDoclet<sup>13</sup> e Ant<sup>14</sup>. Através destas tecnologias, inseriu-se uma camada entre a aplicação e o banco de dados, fazendo com que ele se parecesse com um banco de dados Objeto Relacional [1]. Com isso, o esquema do banco de dados saiu automaticamente a partir das classes Java. Todo código desenvolvido nessa fase foi reaproveitado na construção do *trading system*.

Ao observar algumas dessas séries, notou-se que, possivelmente, nem todas iriam contribuir para a aprendizagem do modelo: a maioria delas tratava-se de empresas de terceira linha e os *candlesticks* poderiam não se aplicar a elas. As empresas de terceira linha possuem baixo volume de negócios e às vezes possuíam somente um negócio por dia, não sendo possível, desta forma, detectar a psicologia do mercado a partir de um volume tão baixo de compra/venda. Selecionaram-se, então, uma a uma, as séries que poderiam contribuir para a aprendizagem do modelo. Da seleção, 730 séries foram armazenadas no banco de dados local.

Com base nessas séries, dever-se-iam separar os padrões de reversão daqueles que representavam continuidade. A utilização de um especialista foi descartada, tanto pelos motivos apresentados anteriormente, como pela li-

---

<sup>9</sup>Os dados históricos podem ser baixados diretamente do site da BOVESPA, <http://www.bovespa.com.br>

<sup>10</sup><http://www.sun.com/>

<sup>11</sup><http://hsqldb.org/>

<sup>12</sup><http://www.hibernate.org/>

<sup>13</sup><http://xdoclet.sourceforge.net/>

<sup>14</sup><http://ant.apache.org/>

mitação do sistema, e pelo tempo gasto para alguém analisar todas as 730 séries. A primeira decisão tomada em relação ao rotulamento das duas classes - reversão e continuação -, foi a separação do problema em dois. Apesar de existirem padrões que são os mesmos para reversão de baixa para alta e de alta para baixa, como o Martelo, também existem aqueles que são únicos para cada um dos dois tipos de reversão, como a Linha Perfurante e Nuvem Negra. Assim, para facilitar a aprendizagem, selecionou-se um modelo para aprender as reversões e continuidades, quando estava-se em tendência de alta e um outro para aprender os padrões quando em tendência de baixa.

O grande desafio de investidores que utilizam análise técnica para investir é determinar qual a tendência da série em um dado momento. Caso o papel esteja entrando em uma tendência de alta, provavelmente será interessante comprar ações daquela empresa. Se uma tendência de baixa for detectada, a decisão a ser tomada é vender os papéis, antes que prejuízos tenham que ser contabilizados. Vários indicadores, como o de Médias Móveis, auxiliam na detecção de tendências. O trabalho de rotulamento dos exemplos do conjunto de treinamento foi realizado através da tendência da série temporal. Diferente da tendência buscada pelos investidores, que pode variar de um dia para o outro, com o filtro de Hodrick-Presscott [22], pode-se retirar a tendência das séries históricas de forma exata, algorítmica, uma vez que já se têm todos os valores disponíveis. O procedimento utilizado para rotular o conjunto de dados foi o seguinte: o ponto de inflexão da tendência seria classificado como pertencente à classe de reversão e os demais pontos como de continuação.

## 2.6 Conclusões

Neste capítulo revisaram-se as principais abordagens encontradas na literatura para se construir um sistema inteligente que auxilie em investimentos no mercado de ações, tanto de sistemas de suporte à tomada de decisão quanto de sistemas que sejam capazes de investir de forma autônoma. Propôs-se uma nova abordagem para a construção de um sistema independente do tipo caixa-preta, baseado nos padrões de *candlestick*, que até o momento, na literatura, só foram utilizados por sistemas tutores e de apoio à decisão.

Para a primeira parte do desenvolvimento do sistema, que consiste em construir um modelo capaz de indicar os momentos de compra e venda das ações, já concluiu-se a separação das amostras que compõem o conjunto de treinamento. No entanto, esta separação apresenta algumas falhas, devido, principalmente, à automatização do processo. Essas falhas não ocorrem quando o trabalho de formação da base de treinamento é feito por um especi-

alista, como visto nos trabalhos encontrados na literatura e discutidos neste capítulo. A solução parcial para essas falhas está intrinsecamente ligada à escolha do modelo de aprendizado de máquina utilizado que, por este motivo, será abordada no próximo capítulo, realizando-se uma revisão bibliográfica do modelo de aprendizado de máquina adotado, aplicado à previsão de séries financeiras.

---

# Aprendizagem Semi-Supervisionada aplicada à Engenharia Financeira

---

No capítulo anterior foi descrito como o conjunto de treinamento  $\{\vec{x}_i, y_i\}_{i=1}^N$ , onde  $\vec{x}_i$  representa o vetor de características de entrada, ou, os valores de abertura, fechamento, máximo e mínimo de preço normalizados em uma janela temporal de cinco dias. A saída desejada, ou o rótulo de cada vetor é dado por  $y_i$  e  $N$ , o número de amostras. Descreveu-se o rotulamento das séries do conjunto de treinamento, utilizando o filtro de Hodrick-Presscott[22].

Diferentemente da abordagem que utiliza um especialista para separar os pares de treinamento, a apresentada no capítulo anterior não demanda tempo e não fica limitada aos conhecimentos do especialista. No entanto, a grande questão, muitas vezes difícil de ser respondida, é se estas classes estão corretamente separadas, e se o conjunto de dados é representativo. De nada adiantaria possuir um método eficiente de separação das classes para treinamento se a separação não for correta.

## 3.1 Dados Conflitantes

Para que a classificação ficasse correta era necessário que em todo ponto de inflexão da tendência retirada da série pelo filtro de Hodrick-Presscott houvesse um padrão de reversão. No entanto, não há garantias de que os padrões ocorram no ponto de inflexão. Algumas vezes, o padrão demora uns 2 ou até 3 dias para confirmar a mudança na tendência. A figura 3.1 mostra um exemplo

desse conflito.

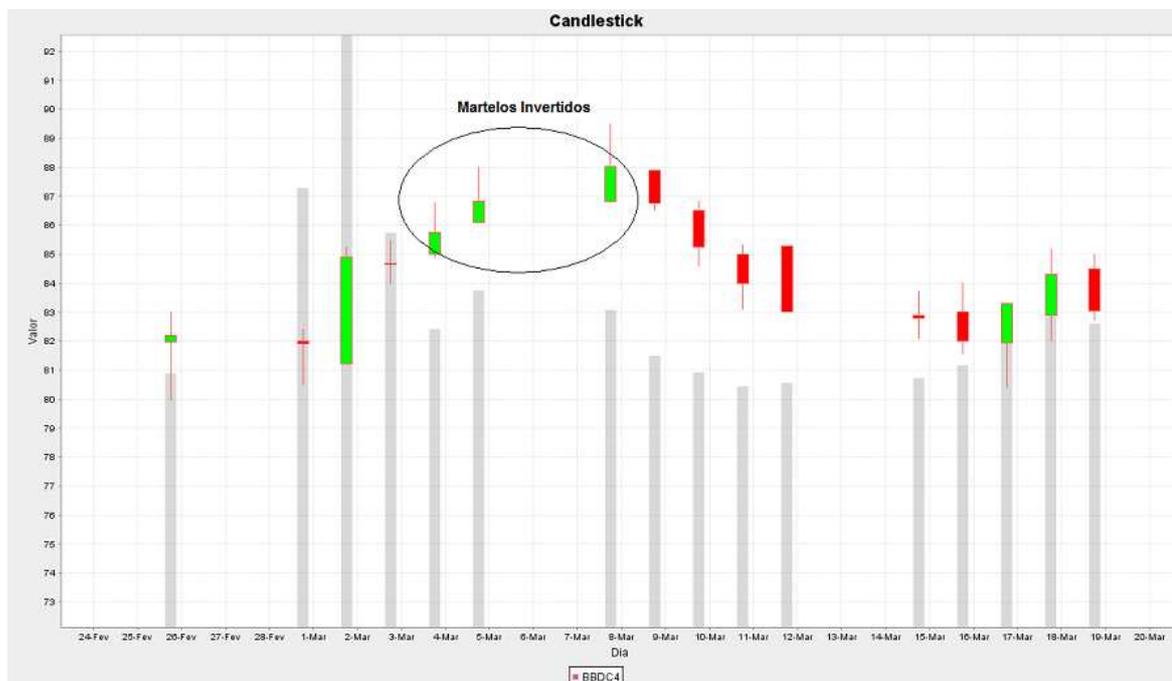


Figura 3.1: Sequência de Martelos Invertidos, indicando futura queda no papel

O ponto de inflexão indica que o dia 8 de março representa um padrão de reversão, o que é plausível, uma vez que representa um Martelo Invertido, que é um padrão catalogado e muito popular entre os investidores. Entretanto, não seria condizente classificar também os outros dias como padrões de reversão, já que também são Martelos Invertidos?

Esse foi, indubitavelmente, o maior desafio enfrentado ao aplicar-se essa abordagem, em que os dados são extraídos automaticamente das séries temporais. Chegou-se a um impasse, caso as amostras, imediatamente antes ao ponto de inflexão, fossem classificadas como pertencentes à classe de Continuação, estar-se-ia classificando exemplos como os das figuras 3.1 e 3.2, em que uma sequência de martelos antecipa a alta, de forma incorreta, o que poderia prejudicar a aprendizagem do modelo.

Por outro lado, caso fossem rotulados como pertencentes à classe de reversão, muitos outros pontos, talvez em maior número ainda, seriam classificados incorretamente. Poder-se-ia imaginar que os exemplos mostrados são casos especiais e muito raros, no entanto, os dois exemplos foram coletados da mesma empresa, Bradesco, e em um pequeno intervalo de tempo. O que mostra que em determinadas séries, esse tipo de situação é muito mais comum do que se imagina. Mais uma vez, constata-se que analisar série por série não seria viável, pois padrões desconhecidos não seriam levados em consideração pelo especialista.

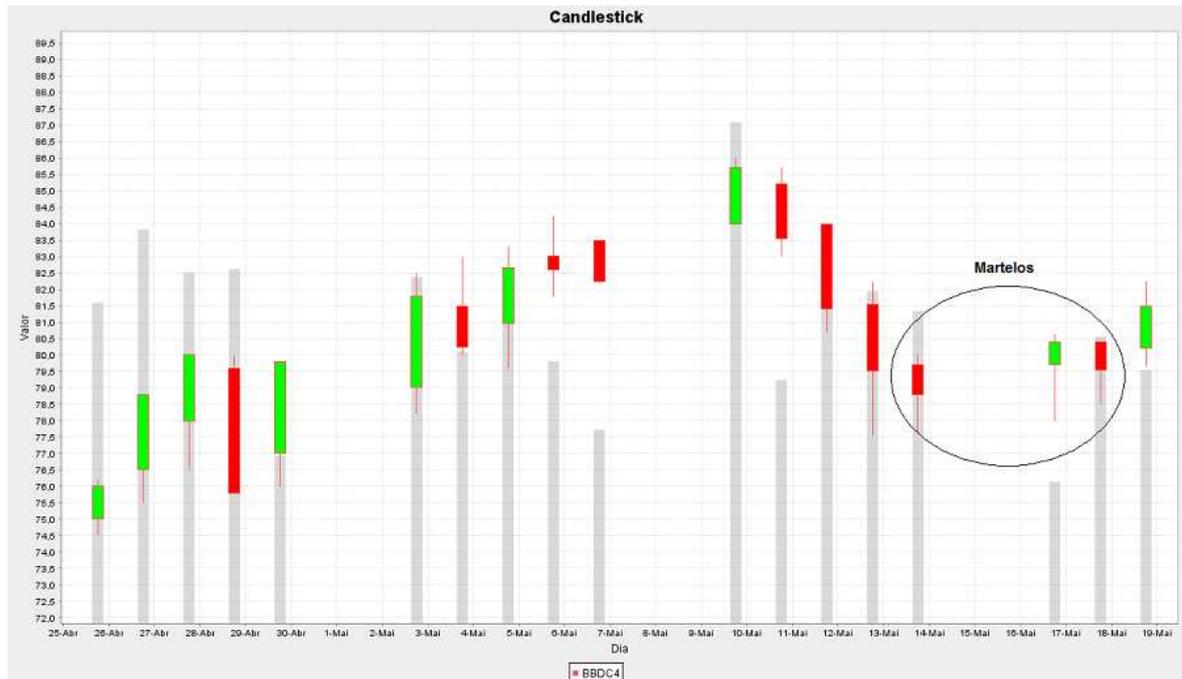


Figura 3.2: Sequência de Martelos, indicando futura valorização no papel

É não é somente nessa situação que poder-se-ia estar causando rotulação incorreta. Muitas vezes um padrão de reversão aparece e a reversão propriamente dita só mostra sinais reais depois de um tempo, geralmente um dia, e sem nenhum outro padrão aparecer no ponto real de inflexão.

Foi então que decidiu-se realizar testes com as duas possibilidades de rotulação incorreta. Uma delas seria classificar os pontos imediatamente antes e depois de um ponto de inflexão, além do próprio ponto de inflexão, como pertencentes à classe de reversão. A outra seria classificar esses pontos duvidosos como pertencentes à classe de continuação. Ao comparar os resultados preliminares das duas abordagens, percebeu-se que o número de falsos positivos quando se treinava considerando os pontos como pertencentes a classe de reversão era muito maior. Este resultado, apesar de óbvio, foi bastante importante para se considerar apenas a rotulação do conjunto de treinamento dos pontos dúbios como pertencentes à classe de continuação.

É um consenso entre os investidores que negócios que tendem a ser mal-sucedidos devem ser sempre evitados. É consenso também que, apesar de evitá-los a todo custo, eles vão acontecer. E para minimizar as perdas de um negócio mal realizado existe o que é conhecido como *stop-loss*: um gatilho de venda que é disparado quando o preço de uma ação chega abaixo de um piso predeterminado. Na construção de um *trading system* gerador de sinais de compra e venda, que trata, como neste caso, a questão como um problema de classificação, o número de falsos positivos deve ser o menor possível, para

evitar o uso do *stop-loss*.

Assim como acontece com aplicações de detecção de *spams* [48], passar um email não desejado para a caixa de entrada tem um custo muito menor do que, por exemplo, jogar na caixa de *spams*, onde ele provavelmente nunca seria lido, uma grande proposta de emprego. Neste tipo de aplicação o falso positivo deve ser evitado a todo custo, mesmo que se prejudique o desempenho global do sistema.

Ao utilizar mais pontos de reversão no treinamento do modelo de classificação é natural que classifiquem-se mais amostras como reversão e, conseqüentemente, gere um número maior de falsos positivos. Por esse motivo, a possibilidade de classificar os pontos duvidosos como reversão foi inicialmente descartada.

A abordagem semi-supervisionada foi a solução encontrada para minimizar este pré-processamento incorreto da base de dados. Nessa abordagem as amostras sem rótulo podem melhorar o desempenho de modelos de aprendizagem de máquina. O TSVM [3] foi utilizado para treinar os modelos semi-supervisionados deste trabalho e, antes de dissertar-se sobre o uso de amostras sem rótulo no treinamento, fazem-se necessárias uma breve introdução sobre aprendizado estatístico, minimização de risco estrutural e máquinas de vetores de suporte (SVM). Posteriormente será detalhada a abordagem semi-supervisionada.

## 3.2 SVM

As Máquinas de Vetores de Suporte, ou *Support Vector Machines - SVM* [6], foram escolhidas como o modelo de aprendizado de máquina a ser utilizado neste trabalho, principalmente por dar suporte ao aprendizado transdutivo, sendo que a implementação feita em [24] é livre para aplicações científicas e já oferece a implementação do TSVM, ou *Transductive Support Vector Machines*.

Nessa seção, detalhar-se-á sobre a base estatística que deu origem às SVMs, assim como suas principais características.

### 3.2.1 Características das SVMs

Uma das características que torna a SVM [6] e o MOBJ [46] tão atrativos é a grande capacidade de generalização apresentada por estes dois modelos. Sabe-se que os conjuntos de treinamento  $\{\vec{x}_i, y_i\}_{i=1}^N$  nem sempre representam, para problemas reais, fidedignamente o problema que está sendo modelado. Dessa forma, exige-se muito da máquina de aprendizado para que haja uma

boa generalização. Um modelo apresenta boa generalização quando é capaz de classificar corretamente amostras  $\vec{x}_i$  não apresentadas a ele durante o processo de treinamento. Através da minimização do risco estrutural, que será discutido em mais detalhes a seguir, esses dois modelos atingem ótima generalização em aplicações práticas.

Além de boa generalização, a SVM apresenta grande robustez quando os vetores de entrada são de grandes dimensões [33]. Este problema é conhecido no meio acadêmico como a Maldição da Dimensionalidade, quando o desempenho <sup>1</sup> dos modelos caem exponencialmente com o aumento da dimensão dos vetores de entrada [16]. Eliminando a Maldição da Dimensionalidade, a SVM torna-se uma excelente opção para aplicações como reconhecimento de padrões em imagens, quando, geralmente, a dimensão das amostras são grandes [33].

Um dos maiores inconvenientes em se utilizar Redes Neurais Artificiais é o fato de o treinamento envolver a minimização de uma superfície que, muitas vezes, possui vários mínimos locais [8]. Nesse caso, não se sabe *a priori*, ao finalizar um treinamento, se aquele é o melhor resultado possível. Deve-se, então, retreinar até convencer-se de que o resultado apresentado é conveniente e satisfatório. Em problemas pequenos e quando se pode utilizar algoritmos eficientes de otimização, como o Levenberg-Marquardt [38], este problema não chega a ser um empecilho. Entretanto, em problemas com número elevado de amostras, como o descrito neste trabalho, a necessidade de retreino para confirmação de resultados pode inviabilizar o projeto. No caso da SVM, o treinamento do modelo se reduz a um problema de Programação Quadrática e, neste caso, existem vários algoritmos bem consolidados na literatura que poderiam ser utilizados para encontrar o único mínimo, o global, da função [9].

Como a SVM surgiu a partir da Teoria de Aprendizado Estatístico [51], a ser descrita adiante, ela possui uma fundamentação matemática e estatística muito clara, o que facilita a sua aceitação em alguns ramos da Ciência. E são estes fundamentos que serão discutidos a seguir.

### 3.2.2 *Aprendizado Estatístico*

A revisão sobre Aprendizado Estatístico que será feita nesta seção terá como pano de fundo apenas problemas de classificação, pois facilita a compreensão dos conceitos e preenche os requisitos deste trabalho, em que foi construído um sistema de previsão financeira através de um modelo de aprend-

---

<sup>1</sup>Desempenho nesse caso, refere-se a tempo computacional requerido no treinamento

dizado de máquina de classificação.

Um modelo de aprendizagem através de exemplos pode ser descrito utilizando três componentes[51]:

- Uma distribuição fixa, mas não conhecida,  $P(x)$ , gera, aleatoriamente, os vetores  $\vec{x}$ .
- Um supervisor, ou professor, retorna um rótulo para cada valor de  $\vec{x}_i$ . A distribuição  $P(y|x)$  também é fixa e desconhecida.
- Uma máquina de aprendizado capaz de implementar um conjunto de funções,  $f(x, \omega), \omega \in \Lambda$ .

Deve-se, então, durante o treinamento determinar qual valor de  $\omega$  para que a máquina de aprendizado  $f(x, \omega)$  tenha o melhor desempenho. Esta escolha é feita com base no conjunto de treinamento, ou, nos  $N$  pares de exemplo,  $(\vec{x}_1, y_1), \dots, (\vec{x}_N, y_N)$ , que são amostras independentes e identicamente distribuídas (i.i.d) segundo a distribuição  $P(x, y)$  mostrada a seguir.

$$P(x, y) = P(x)P(y|x) \quad (3.1)$$

### *Risco Funcional*

Ao treinar um modelo, procura-se selecionar o  $\omega$ , sendo que,  $\omega \in \Lambda$ , de tal forma que se obtenha o menor erro possível, mas qual seria a melhor forma de medir este erro? O risco funcional mede o risco real que o modelo apresenta. Ou,

$$R(\omega) = \int L(y, f(x, \omega))dP(x, y) \quad (3.2)$$

Onde a função  $L(\cdot)$  representa uma função de perda. Quanto mais a escolha do  $\omega$  minimizar o erro funcional, melhor terá sido a escolha. No entanto, apesar de fixa, a função de distribuição  $P(x, y)$  não é conhecida. Este risco não pode, portanto, ser utilizado no treinamento de modelos de aprendizado de máquina. Usa-se uma aproximação do risco funcional, conhecido como risco empírico.

### *Risco Empírico*

O risco empírico é calculado com base nos  $N$  pares de exemplo do conjunto de treinamento  $(x_1, y_1), \dots, (x_N, y_N)$ . E pode ser calculado da seguinte forma:

$$R(\omega)_{emp} = \frac{1}{N} \sum_{i=1}^N (y_i - f(x, \omega))^2 \quad (3.3)$$

Onde  $y_i$  representa o valor desejado ou a classe a que esta amostra pertence e  $f(x, \omega)$ , o valor retornado pelo modelo de aprendizagem. Alguns modelos, como Redes Neurais Artificiais trabalham, com a minimização do risco empírico, no entanto, se apenas ele for levado em consideração, pode acontecer de o modelo memorizar o conjunto de treinamento e, se, como geralmente acontece, o conjunto for limitado, o modelo apresentará problemas de generalização. Algumas técnicas, como validação cruzada, melhoram a generalização em RNAs [8].

No entanto, se o conjunto não for limitado, tem-se, pela teoria dos grandes números, que o risco empírico converge para o funcional, como mostrado pela equação 3.4.

$$P(|R - R_{emp}| > \epsilon) \rightarrow 0, \text{ quando } N \rightarrow \infty \quad (3.4)$$

A grande preocupação é que em problemas reais, como dito anteriormente, o número de exemplos para treinar o modelo é, geralmente, limitado.

Quanto maior a complexidade de um modelo, ou seja, quanto maior for o conjunto  $\Lambda$ , onde os valores de  $\omega$  são selecionados, maior será a capacidade do modelo em separar os dados do conjunto de treinamento. No entanto, nem sempre é interessante separá-los completamente, pois, os pares  $(x_i, y_i)$  contêm ruído que, se forem incorporados pelo modelo podem prejudicar a sua generalização. A figura 3.3 mostra um exemplo em que a complexidade do modelo foi excessiva para o problema em questão, fazendo com que o ruído fosse incorporado pelo modelo. Este tipo de situação é conhecido como *overfitting*.

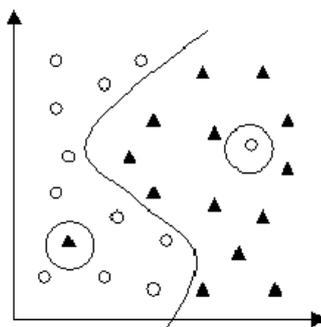


Figura 3.3: Um exemplo de *overfitting*

Pode-se observar amostras com ruído nas duas classes da figura 3.3: na classe da esquerda, representada por pequenos círculos, existe uma amostra

com o rótulo da classe da direita, representada, por sua vez, por pequenos triângulos; assim como existe uma amostra com o rótulo de bola no meio da classe de triângulos. Provavelmente as rotulações destas duas amostras estão incorretas. O modelo de aprendizagem de máquina deveria saber lidar com este tipo de situação, o que não aconteceu nesse caso, figura 3.3. A melhor solução para este problema, provavelmente, é o mostrado na figura 3.4, onde o ruído do conjunto de treino não é incorporado ao modelo.

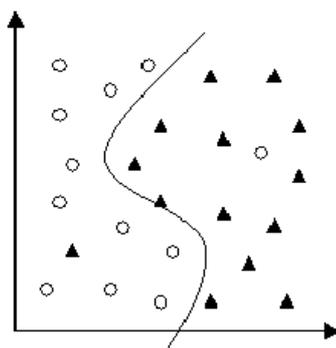


Figura 3.4: Provavelmente a melhor solução para o problema

No entanto, se o modelo de aprendizado de máquina não oferecer parâmetros livres suficientes para aprender o problema, pode ocorrer o problema oposto ao mostrado na figura 3.3, conhecido como *underfitting*. Um exemplo de *underfitting* é mostrado na figura 3.5, onde o modelo não foi capaz de separar corretamente as duas classes.

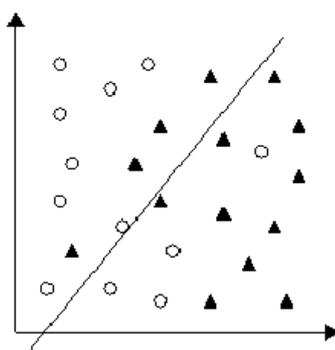


Figura 3.5: Um exemplo de *underfitting*

Essa questão, manter o equilíbrio entre o nível de complexidade do problema com a capacidade de solução do modelo, de forma a obter o *fitting*(3.4), é conhecido na literatura como o dilema da variância e polarização [18].

Definir a complexidade, ou o número de possibilidades que podem ser encontradas no conjunto  $\Lambda$ , é essencial para o bom desempenho de um modelo

de previsão. Deve-se encontrar, pois, o menor conjunto possível que seja capaz de solucionar o problema. O ajuste dessa complexidade é conhecida na literatura como Minimização do Risco Estrutural, e é através do controle do Risco Estrutural que as SVMs atingem boa generalização.

### *Dimensão VC*

A dimensão VC [52] é uma forma de medir a capacidade de um modelo separar determinado conjunto de dados. Dado um problema linearmente separável e sendo  $\Lambda$  o conjunto de superfícies de separação que um modelo consegue gerar, a dimensão VC deste modelo é o máximo de pontos que podem ser particionados pelas funções ou superfícies de separação contidas em  $\Lambda$ .

Pode-se concluir que, quanto maior a dimensão VC de um modelo de aprendizado de máquina, maior será a sua capacidade, pois ele poderá separar um maior número de pontos. Essa definição ficará mais clara com um exemplo, transcrito a seguir [21].

Seja  $\Lambda$  o conjunto de funções sinal,

$$\Lambda = \text{Sinal}(\vec{x}_i \vec{w} + b) \quad (3.5)$$

onde a função  $\text{Sinal}(\cdot)$  assume valor igual a 1, caso o parâmetro  $\vec{x}_i \vec{w} + b$  seja maior que zero e, caso contrário possuirá valor igual a -1.

A dimensão VC desta máquina de aprendizado é dada por,

$$V(\Lambda) = m + 1 \quad (3.6)$$

onde  $m =$  dimensão de  $\vec{x}_i$

O resultado apresentado pela equação 3.6 pode ser conferido na figura 3.6, em que um problema com vetor de entrada de duas dimensões, faz com o que, um modelo de separação linear (mostrado na equação 3.5) tenha uma dimensão VC igual a 3. Neste caso, todas as dicotomias possíveis,  $2^N = 8$ , onde  $N = m + 1$ , são mostradas na figura.

Caso seja acrescentada mais uma amostra ao problema anterior, a máquina descrita por 3.5 já não será capaz de separar todos os casos, como mostra a figura 3.7

É interessante notar que mostrou-se, assim como o problema clássico do ou-exclusivo [8], que um separador linear não é capaz de separar todas as possibilidades, quando o número de amostras é maior que 3, necessitando de uma máquina de maior capacidade, ou, com uma maior dimensão VC.

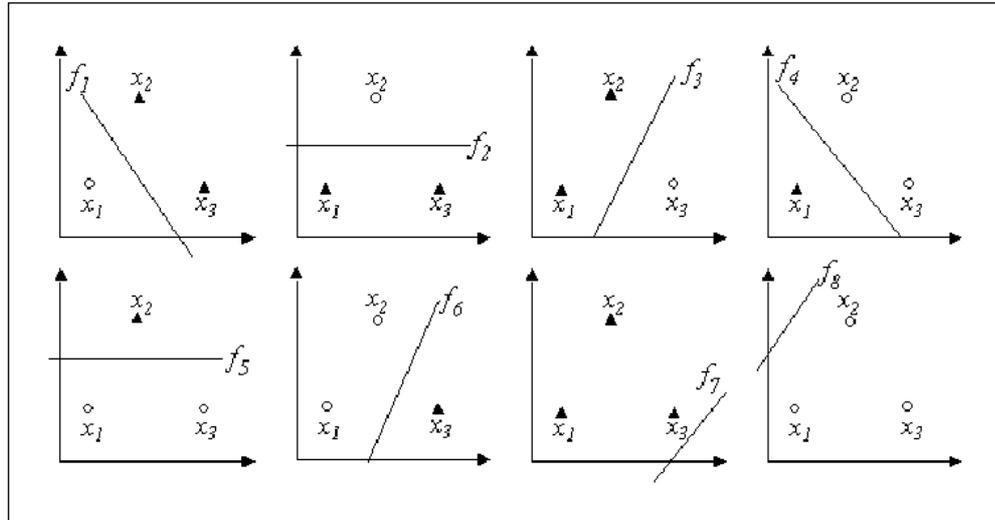


Figura 3.6: Todas as dicotomias possíveis para um separador linear

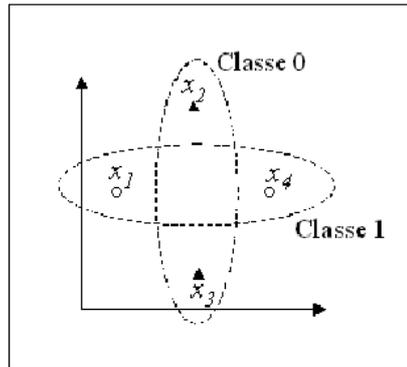


Figura 3.7: Com quatro amostras, um separador linear não é capaz de separar todos os casos

#### Limite Superior do Risco Funcional

Apesar de a dimensão VC ser apenas uma medida teórica e, em muitos modelos, ser difícil de ser calculada, ela serve de base para algumas conclusões importantes que serão discutidas nesta seção.

Vapnik apresentou um teorema, do qual deduziu-se um limite superior para o risco funcional [53].

$$R_{func} \leq R_{emp} + \sqrt{\frac{c}{n} \left( h + \ln\left(\frac{1}{\delta}\right) \right)} \quad (3.7)$$

onde  $n$  é o número de amostras,  $h$  a dimensão VC,  $c$  uma constante universal e  $1 - \delta$  a probabilidade de se limitar superiormente o Risco Funcional.

Desta fórmula, tiram-se conclusões importantes, a partir do segundo termo da parte direita da equação 3.7. Uma conclusão que já foi discutida anteriormente e mostrada pela equação 3.4 é que, quando  $n \rightarrow \infty$  o risco empírico

converge para o risco funcional. E a conclusão mais importante, neste ponto do texto, diz respeito ao valor da dimensão VC: quanto menor for o valor que ela assumir, menor será também o segundo termo da equação e, o risco empírico se aproximará do funcional. Mostra-se desta forma, a importância em minimizar a complexidade do sistema, ou, em outras palavras, o risco estrutural.

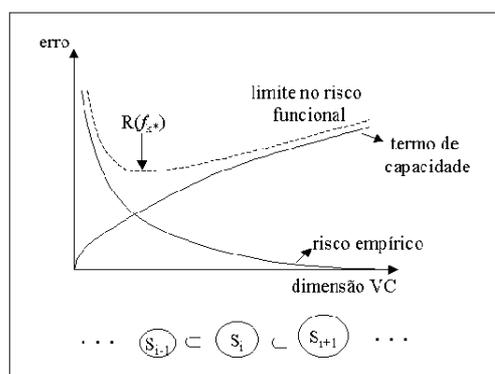


Figura 3.8: Relação entre Risco Funcional, Empírico e Complexidade do modelo

A figura 3.8 mostra a relação entre o risco funcional, empírico e a complexidade do modelo. Tem-se que a complexidade do sistema cresce da esquerda para a direita e, junto com o seu crescimento, inicialmente minimiza-se o risco funcional e empírico, até um ponto ótimo  $f_x^*$ . A partir deste ponto, o limite do risco funcional começa a aumentar e o empírico continua diminuindo, caracterizando uma situação de *overfitting*. O ponto  $f_x^*$  é o que deve ser buscado durante o treinamento, como o ponto em perfeito equilíbrio entre a variância e a polarização [33].

Apesar de o limite delimitador do risco funcional ser, muitas vezes, difícil de ser calculado, ele fornece uma boa base teórica para a construção de modelos de aprendizado de máquina [39].

### Margem Máxima

O algoritmo de treinamento da SVM procura entre as superfícies de separação das classes, aquela que possui a maior margem de separação. A figura 3.9 mostra dois exemplos, o da direita mostra uma superfície de separação qualquer para um problema linearmente separável.

O exemplo da esquerda possui margem máxima e é a superfície de separação procurada pela SVM. A margem pode ser definida como a menor distância entre um ponto de uma classe e a superfície de separação do problema.

Para a dimensão VC do separador linear descrita anteriormente através da

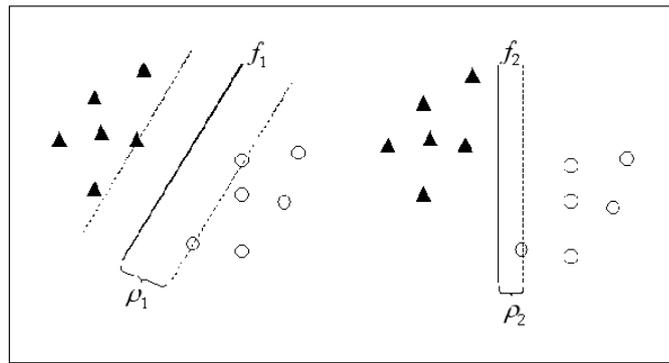


Figura 3.9: Duas soluções, a de margem máxima e uma separação qualquer.

equação 3.6 não levou-se em consideração a margem da superfície de separação. A margem considerada poderia ser qualquer uma, como a exibida no exemplo da direita na figura 3.9. Ao estabelecer uma margem de tamanho mínimo a ser respeitada, restringe-se a capacidade de separação do modelo e, conseqüentemente, diminui a dimensão VC do modelo.

Smola e outros [45] mostraram como se dá esta relação entre a margem da superfície de separação e a dimensão VC do modelo. Chegou-se assim, a seguinte relação,

$$h \leq \min\left\{\frac{R^2}{\rho^2}, m\right\} + 1 \quad (3.8)$$

Onde, seja  $R > 0$  tal que  $\|\vec{x}\| \leq R$  e  $\rho$  a margem do classificador.  $R$  também pode ser visto com uma forma geométrica n-dimensional que circunscreve os vetores de entrada  $\vec{x}$ .

Pode-se perceber que, a dimensão VC de um classificador pode ser menor, com o aumento de margem da superfície de separação. No entanto, ao aumentar a margem, aumenta-se o que é conhecido como erros marginais, ou seja, alguns exemplos do conjunto de treinamento não classificados corretamente de acordo com a margem pré-determinada. Esta relação mostra a importância do tamanho da margem no dilema da variação e polarização, pois, aumentando-se a margem, diminui-se a dimensão VC do modelo e, conseqüentemente, tende a uma situação de *underfitting*. Por outro lado, se a restrição da margem não foi verificada, a dimensão VC não será minimizada e poderá ocorrer *overfitting*.

O hiperplano de separação é considerado ótimo quando ele maximiza o tamanho da margem e, ao mesmo tempo, minimiza os erros marginais [33].

### 3.2.3 SVMs de Margem Rígida

Para se encontrar a margem de separação ótima para problemas linearmente separáveis, o problema pode ser visto como um problema de programação quadrática, descrita no algoritmo abaixo.

1. Dado um conjunto de treinamento linearmente separável,  $\Lambda = \{(x_1, y_1), \dots, (x_N, y_N)\}$
2. Seja  $\alpha^* = (\alpha_1^*, \dots, \alpha_n^*)$  a solução para o seguinte problema de otimização:

3. Maximizar:

$$\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j \vec{x}_i \cdot \vec{x}_j$$

Sujeito às seguintes restrições:

$$\sum_{i=0}^n y_i \alpha_i = 0 \text{ e sendo que, } \alpha_i \geq 0, \alpha_1, \dots, \alpha_n$$

4. O par  $(\vec{w}^*, b^*)$  que define o plano ótimo, pode ser encontrado de acordo com as equações abaixo,

- 5.

$$\vec{w}^* = \sum_{i=1}^n \alpha_i^* y_i \vec{x}_i$$

- 6.

$$b^* = -\frac{1}{2} \{ \max_{\{i|y_i=-1\}} (\vec{w}^* \cdot \vec{x}_i) + \min_{\{i|y_i=1\}} (\vec{w}^* \cdot \vec{x}_i) \}$$

Os valores de  $\alpha_i$  serão diferentes de zero para as amostras do conjunto de treinamento que estão sob as margens. Estas amostras são conhecidas como vetores suporte, pois é através deles que obtém-se os valores de  $\vec{w}^*$  e  $b^*$ , formando a margem de separação ótima. É interessante notar que o treinamento depende então, apenas das entradas que compõem o conjunto de vetores suporte. Desta forma, caso o conjunto de treinamento seja constituído apenas pelos vetores suporte, a superfície de separação resultante seria a mesma se tivesse sido utilizado todo o conjunto de treinamento.

Embora não sejam muito úteis, por resolver apenas problemas linearmente separáveis, as SVMs de margem rígida são importantes para que as definições dos tópicos a seguir sejam desenvolvidas com mais clareza.

### 3.2.4 SVMs com margens suavizadas

Os dados do conjunto de treinamento geralmente apresentam ruído, ocasionado por pequenas falhas na extração do conjunto. Caso uma amostra ruidosa seja incorporada como um vetor suporte para formar a margem de separação, o desempenho do modelo poderá estar prejudicado por aquela amostra. Algumas vezes, o ruído chega a levar a amostra para a área que pertence a outra classe, como se vê na figura 3.10, onde as amostras destacadas pelas setas exibem amostras com o rótulo incorreto do conjunto de treinamento.

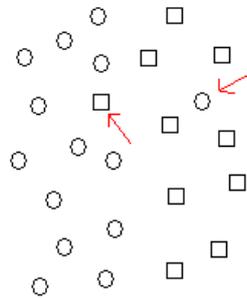


Figura 3.10: Problema linearmente separável, porém, com amostras ruidosas

As SVMs de margens rígidas não seriam capazes de lidar com este tipo de problema. Para tanto, bastou incorporar uma variável de folga ao problema de programação quadrática, de forma que o modelo aceitasse a classificação incorreta de algumas amostras do conjunto de treinamento.

Desta forma, o algoritmo pode ser descrito como,

1. Dado um conjunto de treinamento linearmente separável,  $\Lambda = \{(x_1, y_1), \dots, (x_N, y_N)\}$
2. Seja  $\alpha^* = (\alpha_1^*, \dots, \alpha_n^*)$  a solução para o seguinte problema de otimização:
3. Maximizar:

$$\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j \vec{x}_i \cdot \vec{x}_j$$

Sujeito as seguintes restrições:

$$\sum_{i=0}^n y_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C$$

4. O par  $(\vec{w}^*, b^*)$  que define o plano ótimo, pode ser encontrado de acordo com as equações abaixo,

5.

$$\vec{w}^* = \sum_{i=1}^n \alpha_i^* y_i \vec{x}_i$$

6.

$$b^* = -\frac{1}{2} \{ \max_{\{i|y_i=-1\}} (\vec{w}^* \cdot \vec{x}_i) + \min_{\{i|y_i=1\}} (\vec{w}^* \cdot \vec{x}_i) \}$$

Onde o parâmetro  $C$  determina a margem aceitável para aquele problema. Caso um multiplicador de Lagrange assuma o valor de  $C$ , a amostra correspondente àquele ponto estará entre as margens. Para valores entre 0 e  $C$ , definem-se as amostras que estão em cima da margem, permitindo, desta forma, uma folga para que o modelo aceite classificações erradas, como as mostradas na figura 3.10.

Apesar de aceitar uma gama maior de problemas, esse algoritmo não é poderoso o suficiente para resolver problemas não lineares. A SVM capaz de resolver este tipo de problema será discutido na próxima seção.

### 3.2.5 SVMs não lineares

Nesta seção será descrito, um algoritmo semelhante ao mostrado nas seções anteriores, com uma pequena, mas poderosa modificação, que habilita este modelo de aprendizado de máquina tratar problemas que não são separáveis linearmente.

Uma rede neural de duas camadas pode ser utilizada para lidar com problemas não lineares, desde que a função de ativação da camada intermediária insira alguma não linearidade ao modelo. O que acontece é que os vetores de entrada são mapeados para um outro espaço de característica pela camada intermediária. Nesse outro espaço, o problema torna-se linearmente separável, ficando a cargo do neurônio da última camada solucionar o problema. Na SVM ocorre processo semelhante, como mostra a figura 3.11.

Utilizando uma função qualquer  $\Phi$ , reecre-se o algoritmo anterior da seguinte forma,

1. Dado um conjunto de treinamento qualquer,  $\Phi(\Lambda) = \{(\Phi(x_1), y_1), \dots, (\Phi(x_N), y_N)\}$
2. Seja  $\alpha^* = (\alpha_1^*, \dots, \alpha_n^*)$  a solução para o seguinte problema de otimização.
3. Maximizar:

$$\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j \Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j)$$

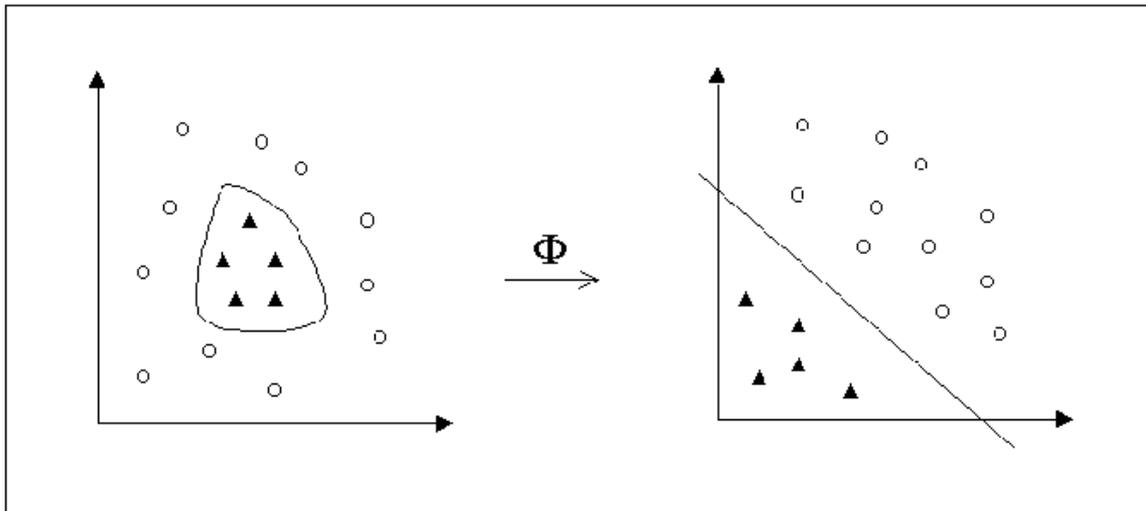


Figura 3.11: Mapeamento do espaço de entrada para um outro espaço em que a separação linear seja aplicada

Sujeito as seguintes restrições:

$$\sum_{i=0}^n y_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C$$

4. O par  $(\vec{w}^*, b^*)$  que define o plano ótimo, pode ser encontrado de acordo com as equações abaixo,

5.

$$\vec{w}^* = \sum_{i=1}^n \alpha_i^* y_i \Phi(\vec{x}_i)$$

6.

$$b^* = -\frac{1}{2} \{ \max_{\{i|y_i=-1\}} (\vec{w}^* \cdot \Phi(\vec{x}_i)) + \min_{\{i|y_i=1\}} (\vec{w}^* \cdot \Phi(\vec{x}_i)) \}$$

Percebe-se que o único ajuste, ainda, a ser feito segundo o algoritmo demonstrado acima é o produto interno de  $\Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j)$ . Para tanto, define-se uma função *kernel*  $K(\vec{x}_i, \vec{x}_j) = \Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j)$ .

Ao utilizar funções *kernel*, que devem ser definidas, a princípio, de modo empírico, insere-se uma não linearidade à SVM, permitindo um mapeamento do espaço de característica original, onde o problema pode ser não linearmente separável, para um outro espaço de característica, desta vez, linearmente separável. Neste ponto, SVMs de margens suaves poderiam ser utilizadas para solucionar o problema.

Deve-se, portanto, ao utilizar a SVM como modelo de aprendizado de máquina, definir qual o parâmetro C será utilizado e, também, qual função *kernel*

será empregada.

Mostrou-se nesta seção uma pequena introdução sobre aprendizado estatístico, cuja teoria subsidiou a construção das SVMs. Além disso, discutiu-se sobre a minimização do risco estrutural nas SVMs, através da maximização das margens e discutiu-se sobre a utilização de *kernels* para inserir não linearidade ao modelo.

Na seção seguinte serão discutidos a utilização da aprendizagem semi-supervisionada em outros trabalhos da literatura e ganhos, ao se utilizar este tipo de abordagem. Finalmente, será mostrado como o uso de amostras sem rótulo contribuiu para melhorar a performance do *trading system* descrito neste trabalho.

### 3.3 *Aprendizagem Semi-supervisionada*

Até o momento, no texto, utilizou-se a palavra transdutiva com o mesmo significado de semi-supervisionada, no entanto, elas apresentam uma leve diferença em suas definições. Aprendizado semi-supervisionado é quando utilizam-se dados com e sem rótulo em qualquer instante do projeto de um modelo. Essa definição vem, provavelmente, de um paralelo com o aprendizado supervisionado, em que todos os dados possuem rótulo; com o não supervisionado, em que nenhum dado possui rótulo. Um modelo transdutivo lida apenas com dados com e sem rótulo, não trabalhando com amostras nunca vistas. Já o indutivo é capaz de lidar com amostras nunca vistas, independentemente se o tipo de aprendizado é supervisionado, semi-supervisionado ou não supervisionado[56]. Dessa forma, o modelo utilizado neste trabalho, o *Transductive Support Vector Machines* na verdade, não é transdutivo, mas sim indutivo e semi-supervisionado pois, ele foi utilizado para classificar dados nunca vistos pelo modelo, após o treinamento com dados com e sem rótulo[56]. O nome TSVM vem da idéia original de utilizá-lo apenas para dados observados, embora não seja essa a sua maior utilização. Para entender melhor a diferença entre indutivo e transdutivo, faz-se uma analogia: o transdutivo seria uma prova que se leva para casa para resolver(estuda-se o suficiente para resolver apenas as questões da prova); já o indutivo seria a prova que você faz em sala de aula(estuda-se para resolver qualquer questão que possa, eventualmente, aparecer na prova) [56].

Na descrição das Máquinas de Vetores Suporte utilizaram-se apenas exemplos de aprendizagem supervisionada, ou seja, para cada valor de entrada com distribuição  $P(x)$ , existia um supervisor que rotulava cada uma daquelas entradas de acordo com uma probabilidade condicional  $P(y|x)$ . Em outras

palavras: o conjunto de treinamento era dado por,  $(x_1, y_1), \dots, (x_N, y_N)$ , sob a função de distribuição  $P(x, y)$ . A diferença é que na aprendizagem semi-supervisionada, além do conjunto de treinamento usual, descrito pelas equações anteriores, tem-se um conjunto de dados sem rótulo, ou,  $(x_1), \dots, (x_n)$ , também descritos pela função de distribuição  $P(x)$ , sendo que  $n$  é o número de pontos sem rótulo, ou grupo de trabalho, como também é conhecido.

No capítulo 2 descreveu-se a dificuldade em separar grande quantidade de dados com rótulos confiáveis para formar uma base de treinamento em tamanho suficiente para a construção de um *trading system* baseado apenas em *candlesticks*. Essa dificuldade não é exclusividade de modelos de previsão financeira, normalmente é difícil, caro, ou demorada a formação de uma boa base de dados rotulados para qualquer problema de aprendizado de máquina. No entanto, geralmente a obtenção de dados de entrada  $\vec{x}$  sem rótulos é bem mais simples, o que torna o conjunto de trabalho muito maior que o conjunto de aprendizado supervisionado, ou  $n \gg N$ .

Por causa dessa limitação dos dados rotulados, a capacidade de generalização de modelos pode ficar prejudicada, mesmo em algoritmos bem estabelecidos e que geralmente oferecem boa generalização, como o MOBJ e SVM. A utilização do conjunto de trabalho surgiu como uma alternativa para, em alguns casos, melhorar a generalização do sistema [3] [50].

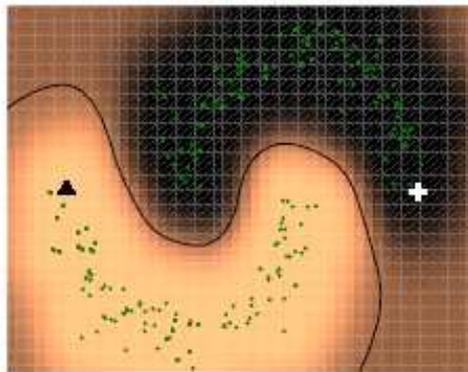


Figura 3.12: Algoritmo de *clustering* semi-supervisionado. Ponto em triângulo indica uma classe, em cruz outra. Os demais pontos são as amostras sem rótulo.

A figura 3.12 mostra um exemplo em que têm-se apenas duas amostras rotuladas, uma na forma de triângulo, à esquerda, e a outra em cruz, à direita. Todos os outros pontos são amostras sem rótulo. Através dos pontos sem rótulo, pode-se inferir que a superfície de separação entre as duas classes estará na região com menor densidade de amostras entre as duas classes. As figuras 3.13, 3.14, 3.15 e 3.16 mostram um outro exemplo em que o conjunto de trabalho melhorou, significativamente, a capacidade de generalização do

sistema.

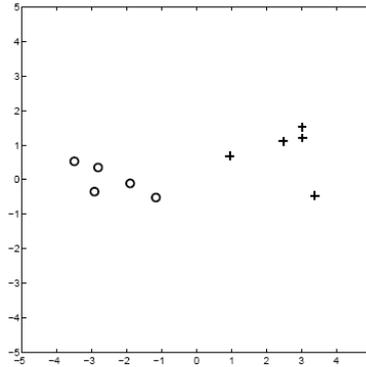


Figura 3.13: Amostras rotuladas de duas classes

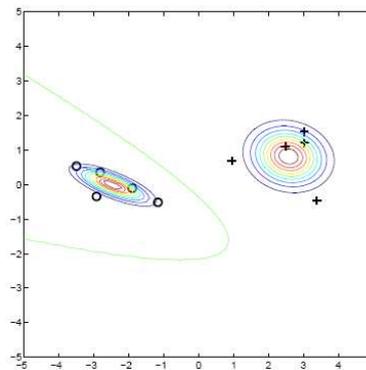


Figura 3.14: Superfície de separação utilizando apenas dados rotulados no treinamento - abordagem indutiva

Apesar de ser uma técnica relativamente nova, já encontram-se na literatura algumas aplicações que obtiveram sucesso em combinar amostras com rótulo, com um conjunto de trabalho sem rótulo. Blum e Mitchell [5] utilizaram como exemplo a classificação de sítios na internet. Neste problema em particular, muito tempo deveria ser despendido para rotular os exemplos a serem utilizados no treinamento, o que poderia inviabilizar o projeto. Exemplos sem rótulos poderiam, no entanto, ser facilmente coletados por um *crawler*. Seo, Giampapa e Sycara [43] utilizaram dados sem rótulo para treinar o modelo de aprendizado proposto naquele trabalho. Como explicado no capítulo 2, Seo e co-autores, utilizaram texto de fontes confiáveis, como CNN e Reuters para verificarem se empresas comentadas por estas fontes eram boas ou não para se investir. O modelo de previsão financeira, baseado em Análise Fundamentalista tornou-se, então, um problema de classificação de texto. Depois do trabalho de Joachims [25], a utilização de amostras sem rótulo em problemas de classificação de texto se popularizou e grande parte dos modelos produzidos para este fim passaram a utilizar a aprendizagem semi-supervisionada.

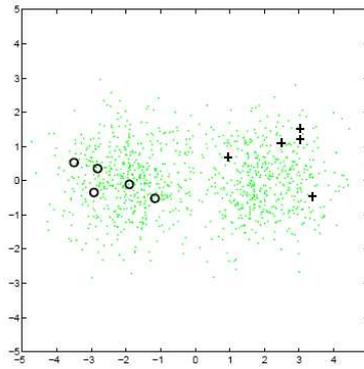


Figura 3.15: Amostras rotuladas e conjunto de trabalho sem rótulo de duas classes

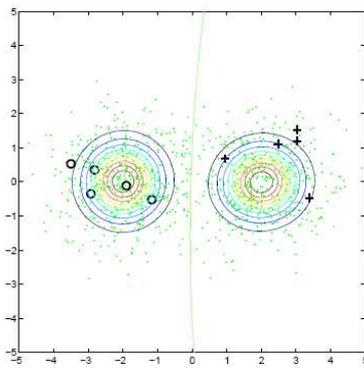


Figura 3.16: Superfície de separação utilizando dados rotulados e não rotulados no treinamento - abordagem transdutiva

#### 3.3.1 *Transductive Support Vector Machines*

A seção anterior mostrou como amostras sem rótulo podem, intuitivamente, melhorar a generalização de um modelo de aprendizado de máquina. Nesta seção será detalhada como a abordagem transdutiva foi implementada no SVM Light [25], utilizada neste trabalho.

A abordagem algorítmica, força bruta, para um problema de treinamento de aprendizado transdutivo seria: testar cada uma das possibilidades de rotulação do conjunto de trabalho. No entanto, para realizar cada um desses testes, seriam necessários  $2^n$  treinamentos do modelo, onde  $n$  é o número de amostras sem rótulo. A idéia do algoritmo força bruta é testar todas as combinações possíveis de rotulamento do conjunto de trabalho e para cada combinação, realizar um treinamento indutivo, como o mostrado na seção sobre SVMs. Nesse caso, esse procedimento teria a complexidade temporal de  $\Omega(2^n)^2$  treinamentos indutivos, caso o problema considerado tenha apenas 2 classes.

---

<sup>2</sup>Para mais detalhes sobre complexidade assintótica e limite inferior, consultar [10]

Algoritmos com complexidade temporal exponencial, como o de treinamento força bruta de um problema transdutivo, não são viáveis para problemas com muito mais que 10 entradas [10] [25]. A utilização de amostras sem rótulo necessita então, de uma heurística, ou seja, um algoritmo aproximado, que não necessariamente irá retornar o resultado ótimo para o problema. A idéia é limitar o número de combinações a serem testadas, de forma a diminuir a complexidade temporal do problema.

No SVM light, inicialmente, dá-se rótulo a todo o conjunto de trabalho com base na solução indutiva gerada através do conjunto rotulado de treinamento. Entretanto, essa rotulação não ocorre simplesmente com a divisão das amostras pelo hiperplano gerado pelo treinamento,  $\vec{w} \vec{x} + b = 0$ , pois o usuário do SVM Light define qual a porcentagem do conjunto de treinamento pertence à classe positiva e qual fatia compõe a classe negativa. Por esse motivo, supondo que  $num_+$  defina o número que amostras que devem ser classificadas como pertencentes à classe positiva, os  $num_+$  maiores valores de  $\vec{x}$  aplicados à  $\vec{w} \vec{x} + b$  serão rotulados como classe positiva. O restante do conjunto de trabalho, por sua vez, será rotulado como pertencentes à classe negativa.

A partir dessa solução inicial, iterativamente, troca-se os rótulos de amostras do conjunto de trabalho e obtem-se uma nova solução. A troca de rótulos deve ser feita de uma forma criteriosa, do contrário, o número de treinamentos indutivos a serem realizados será igual ao algoritmo de força bruta. Joachims [25] definiu que trocar-se-ia de rótulo as amostras que possuissem a soma da variável de folga (rever seção sobre SVMs de margens suavizadas) maior que dois e que as duas folgas (correspondentes às duas amostras) fossem maior que zero. Com isso trocava-se amostras que estavam do lado errado do hiperplano de separação, ou que simplesmente não respeitavam o tamanho da margem. Após a troca de rótulos realiza-se novamente o treinamento da SVM. O processo é repetido até que não exista mais amostras que possam ter os rótulos trocados. Nesse instante, incrementa-se o valor do peso das amostras do conjunto de trabalho no treinamento e todo o processo descrito anteriormente é repetido. Quando o valor da influência do conjunto de trabalho chega ao máximo (definido pelo usuário) o treinamento é finalizado. O algoritmo pode ser observado na figura 3.17.

Na conclusão de seu artigo que apresenta o TSVM para a classificação de texto [25], Joachims levanta uma série de questões que até então não haviam sido respondidas. Uma delas é se TSVM Light, apresentado nesta seção, apresentaria bons resultados caso fosse aplicado a outros pontos de testes, diferentes daqueles utilizados no treinamento. Na próxima seção aborda-se-á como o TSVM foi aplicado de forma indutiva e semi-supervisionada a um

```
“Calcular solução inicial desconsiderando o conjunto de trabalho”.
“Classificar o conjunto de trabalho de acordo com a porcentagem de amostras positivas e negativas selecionadas pelo usuário”

Enquanto(“Influência do conjunto de trabalho não chegou ao máximo”)
    “Treinar considerando o conjunto de trabalho rotulado anteriormente”

    Enquanto(“Houver amostras cuja a soma das folgas dê mais que dois e as duas folgas sejam maiores que zero”)
        “Troque as amostras”
        “Retreine”

    Fim

    “Incremente a influência do conjunto de trabalho”

Fim

Retorne(“Os rótulos do conjunto de trabalho”)
```

Figura 3.17: Algoritmo em português do TSVM Light

problema de previsão financeira.

#### 3.3.2 Aprendizagem semi-supervisionada aplicada à Engenharia Financeira

No princípio deste capítulo descreveu-se como o conjunto de treinamento, cuja rotulação foi descrita no capítulo 2, possui algumas falhas que podem atrapalhar o treinamento de um modelo de aprendizado de máquina. Essas falhas ocorrem devido à rotulação automática, que se baseia na tendência da série temporal. Como se vê, pode ocorrer mais de um padrão de *candlestick* antes da reversão; assim como o padrão, por sua própria característica, pode aparecer somente após o ponto de inflexão da tendência. Dessa forma, o esquema de automatização da rotulação adotada poderia estar inserindo muito ruído ao conjunto de treinamento. Apesar de a SVM lidar bem com algumas amostras ruidosas, caso elas apareçam em excesso, poderiam prejudicar seriamente o desempenho do modelo.

Os trabalhos que justificam a utilização da aprendizagem semi-supervisionada sempre se apoiam na dificuldade em conseguir amostras com rótulo em número considerável, por muitas vezes ser uma tarefa demorada, outras, por ter um alto custo financeiro. No caso deste trabalho, as amostras rotuladas existem, e em grande número, por isso foi desenvolvido um método de rotulação automática que, em minutos, separa uma grande quantidade de dados. No entanto está se questionando a qualidade dessas amostras, se são realmente confiáveis ou se poderão prejudicar o treinamento.

Foram treinados, então, quatro modelos: dois para os pontos de compra e outros dois para os pontos de venda. Em um par deles, compra/venda, utilizou-se apenas amostras com rótulo, o outro, semi-supervisionado, foi treinado com amostras sem rótulo.

Como o treinamento de uma TSVM exige a regulação de alguns parâmetros,

de forma empírica, desenvolveu-se uma aplicação que poderia documentar os resultados e automatizar a variação dos parâmetros. A construção e utilização desta aplicação será descrita na próxima seção.

#### *JSVM<sup>litc</sup>*

O *JSVM<sup>litc</sup>* foi construído com o intuito de facilitar o treinamento de SVMs e TSVMs. Como descrito no tópico correspondente à SVM, o treinamento deste modelo muitas vezes é trabalhoso e pode demandar muito tempo. Deve-se, além da função de *kernel*, definir o valor de C, que neste caso foi determinado empiricamente. No caso da detecção dos padrões de *candlestick*, têm-se muito mais exemplos da classe de continuação do que da classe de reversão, ou seja, é um problema de classificação desbalanceado, o que se faz ajustar mais um parâmetro da SVM, que seria o parâmetro de peso da classe positiva (ponto de reversão) no erro do treinamento. Com o ajuste deste parâmetro, pode-se resolver o problema de balanceamento das classes mas, como será visto adiante, o desbalanceamento das classes, neste caso, é interessante e vantajoso. Por fim, tem-se o parâmetro do peso dos exemplos sem rótulo para cada classe, no caso da abordagem semi-supervisionada.

Variar cada um desses parâmetros empiricamente, documentar os resultados, para depois ser feita uma análise, são tarefas extremamente trabalhosas e demoradas. A automatização das chamadas da SVM Light [24], com as variações dos parâmetros já sendo feitas de uma forma pré-programada, agilizaria muito o desenvolvimento dos modelos. Um outro requisito importante para esta ferramenta, seria a geração de relatórios, de todos os modelos gerados, com curvas ROC e matrizes de confusão [15].

A partir desses requisitos é que surgiu o *JSVM<sup>litc</sup>*. O que criou-se foi basicamente uma interface gráfica utilizando Java - JSE, onde o valor inicial e final de cada parâmetro, assim como o seu incremento poderiam ser definidos. Nesta mesma interface selecionavam-se o arquivo que continha o conjunto de treinamento e todos os arquivos de validação, lembrando que estes arquivos deveriam estar no formato aceitado pela SVM [24]. Cada modelo, após ser treinado, era testado com o conjunto de validação; a partir deste conjunto as curvas ROC e matrizes de confusão eram construídas. Após terminar a execução para todos os valores estipulados pelo usuário, o programa gera um relatório completo em HTML, com os gráficos ROC, gerados pelo JFreechart<sup>3</sup>. Selecionava-se também o melhor de todos os modelos, baseados na distância que cada ponto correspondente a cada modelo representava na curva ROC. O

---

<sup>3</sup><http://www.jfree.org/jfreechart/>

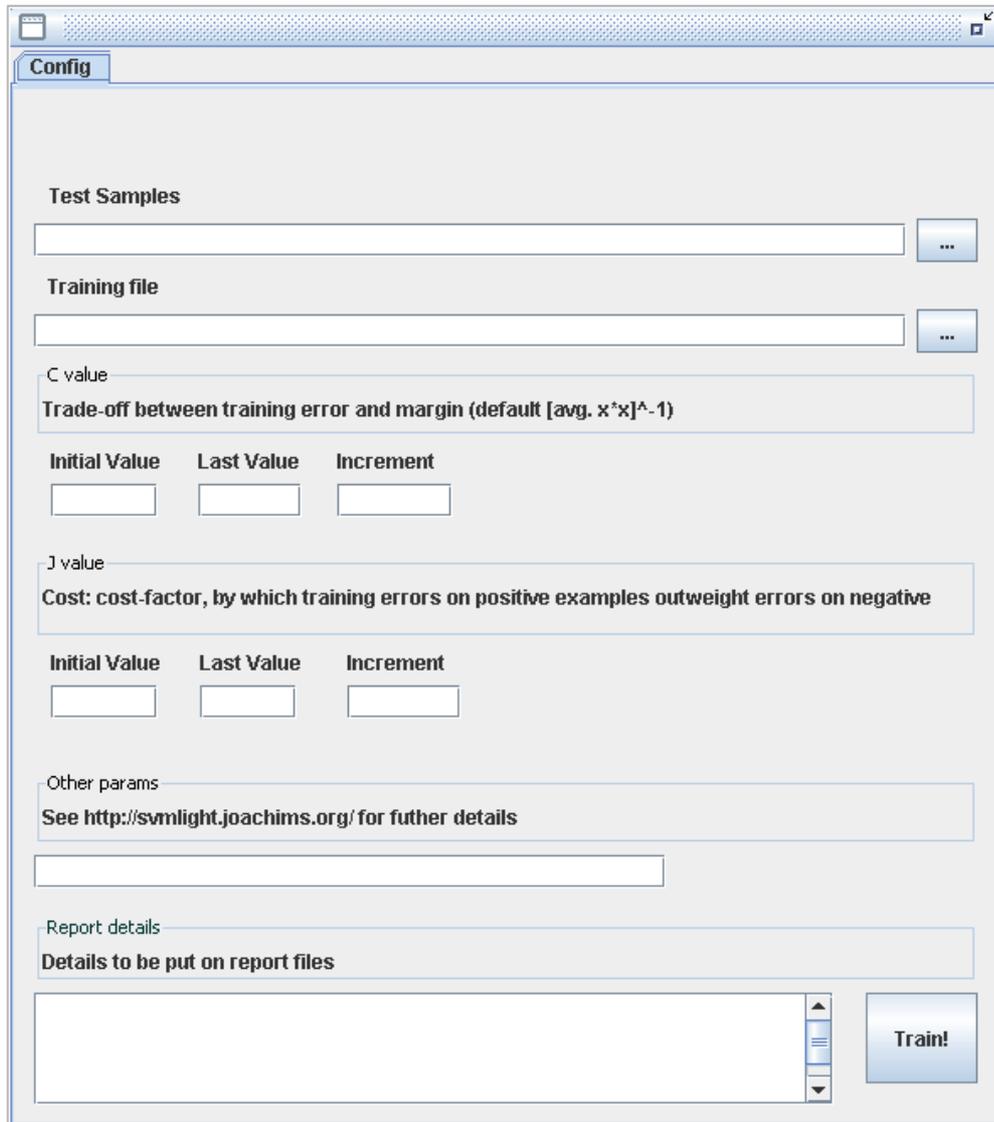


Figura 3.18: Interface em JSE -  $JSVM^{lite}$

modelo que obtivesse maior distância da reta  $Sensibilidade = 1 - Especificidade$  [15] seria selecionado como o melhor. A figura 3.18 mostra a interface do  $JSVM^{lite}$ .

### 3.4 Treinamento dos modelos

Como dito no capítulo 2, utilizaram-se 730 séries da BOVESPA entre os anos de 1986 e 2005 para treinar os quatro modelos de classificação de pontos de reversão. Como conjunto de validação a ser fornecido para o  $JSVM^{lite}$  selecionaram-se algumas séries de 2006, são elas: CGAS5, CMIG3, CMIG4, CPFE3, CPLE6, CPSL3, CRUZ3, CSAN3, CSMG3, CSNA3, CTNM4, CYRE3, DASA3, DURA4, ENBR3, GETI3, GETI4, GFSA3, GGBR3, GGBR4, GOAU3,

GOAU4, GUAR3, LREN3, PCAR4, PMAM4, SUBA3, TAMM4, VALE3, VALE5. Com base na performance de cada modelo treinado, com as séries descritas acima, selecionaram-se os modelos que seriam utilizados na fase de testes.

Trata-se, claramente, de um problema de classificação altamente desbalanceado - como citado anteriormente -, em que o número de exemplos da classe positiva (pontos de reversão) é significativamente menor. Ao variar o peso do erro de um exemplo positivo no treinamento, varia-se também o número de classes positivas geradas, como resposta e, conseqüentemente, compromete-se a confiabilidade do sistema. Seis pares de modelos - compra/venda - foram gerados. O modelo 1 é aquele a cujo peso de exemplos positivos foi dado o menor valor, ou seja, menos sinais de reversão foram gerados; em contrapartida, esses sinais eram mais confiáveis. O modelo 6, por sua vez, tem o maior peso para exemplos da classe positiva, gerando mais sinais de compra e venda, com a desvantagem de serem sinais menos confiáveis - com maior número de falsos positivos. Os outros 4 são modelos intermediários a esses dois.

Para cada valor de peso do erro causado por uma amostra positiva no treinamento, variaram-se os valores de  $C$ , sempre utilizando como função de *kernel* a RBF. A partir disso, selecionou-se o melhor valor de  $C$  para aquele valor de peso, baseado nos resultados apresentados pelos modelos com o conjunto de validação. Os pontos na curva ROC dos melhores modelos, gerados para compra/venda da abordagem supervisionada e compra/venda da abordagem semi-supervisionada, podem ser conferidos nas figuras 3.19 e 3.20, respectivamente.

Uma observação sobre a curva ROC de todos os modelos - tanto supervisionado quanto dos semi-supervisionados, é que os pontos dúbios, que foram enquadrados como pertencentes ao conjunto de trabalho, não foram considerados no cálculo da especificidade e sensibilidade. Isso torna as curvas ROC apenas uma aproximação do que seria, caso os pontos fossem todos rotulados e houvesse garantias quanto a confiabilidade do conjunto. Dessa forma, não se tem nenhuma informação quanto ao desempenho das abordagens com relação aos pontos dúbios; só se saberá sobre o seu desempenho, quando os modelos se sujeitarem a testes de investimentos, tema de discussão do próximo capítulo.

Os seis modelos de compra foram escolhidos em cada uma das abordagens de aprendizado, nas faixas onde a taxa de falso positivo fosse aceitável. Nos modelos de compra da abordagem supervisionada (figura 3.19) os modelos 1 e 2 foram desconsiderados, pois o número de sinais de compra gerados não era suficiente para gerar lucro considerável, assim, como os três últimos modelos da abordagem semi-supervisionada (figura 3.20). Nesse caso, o número de

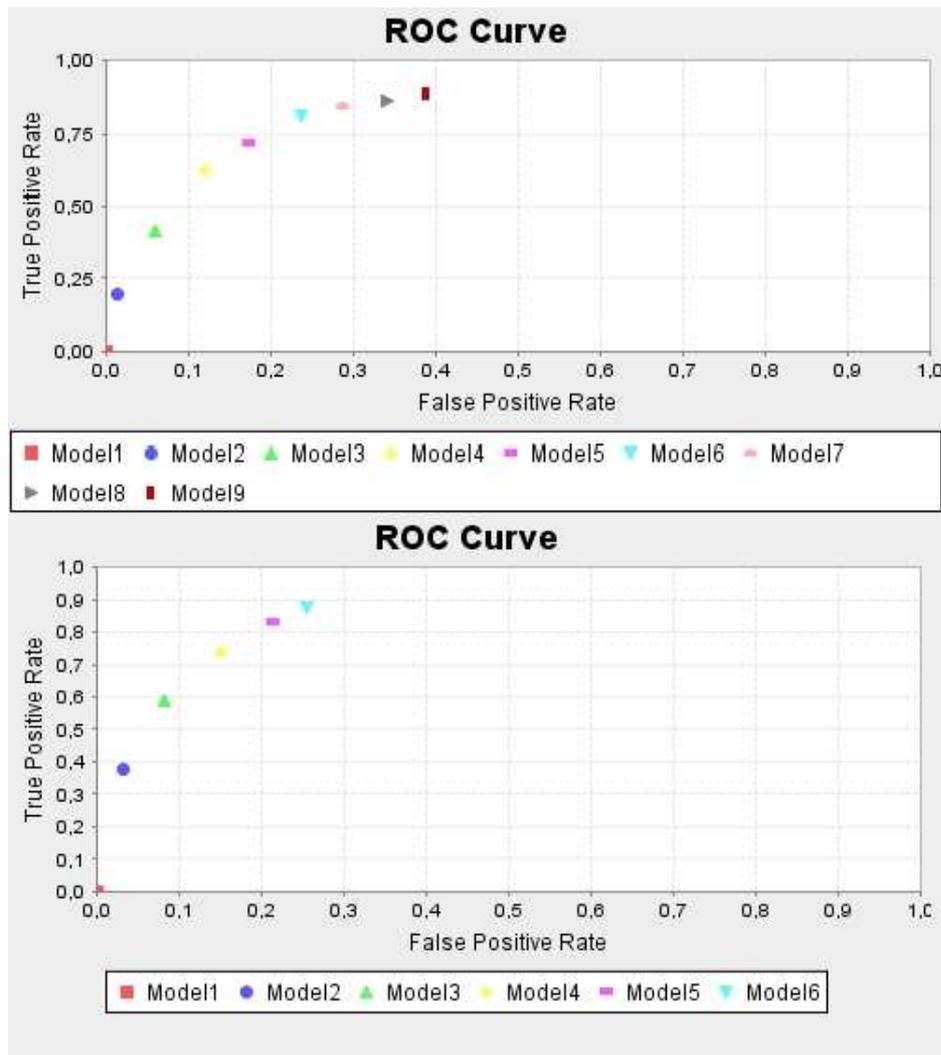


Figura 3.19: Curva ROC da abordagem supervisionada dos modelos de compra, primeira; e venda, segunda

falsos positivos foi considerado alto e, por isso, o número de negócios com prejuízos seria, também, alto, diminuindo a margem de lucro do modelo.

Para os modelos de compra, selecionaram-se então, para comparação, pontos que tinham posição mais próxima possível na curva ROC. Os modelos selecionados estão listados nas tabelas 3.1 e 3.2.

Tabela 3.1: Modelos da abordagem supervisionada selecionados para testes

	<b>Modelo1</b>	<b>Modelo2</b>	<b>Modelo3</b>	<b>Modelo4</b>	<b>Modelo5</b>	<b>Modelo6</b>
<b>Distância</b>	0.25	0.35	0.38	0.40	0.39	0.37
<b>Especificidade(%)</b>	93.9	87.9	82.6	76.1	71.2	65.7
<b>Sensibilidade(%)</b>	41.7	62.0	71.6	81.2	84.4	86.6

A linha **distância** nas tabelas 3.1 e 3.2 indica a distância daquele modelo com relação a reta  $Sensibilidade = 1 - Especificidade$ . Quanto maior esta distância, mais próximo o modelo estará do canto superior esquerdo do gráfico

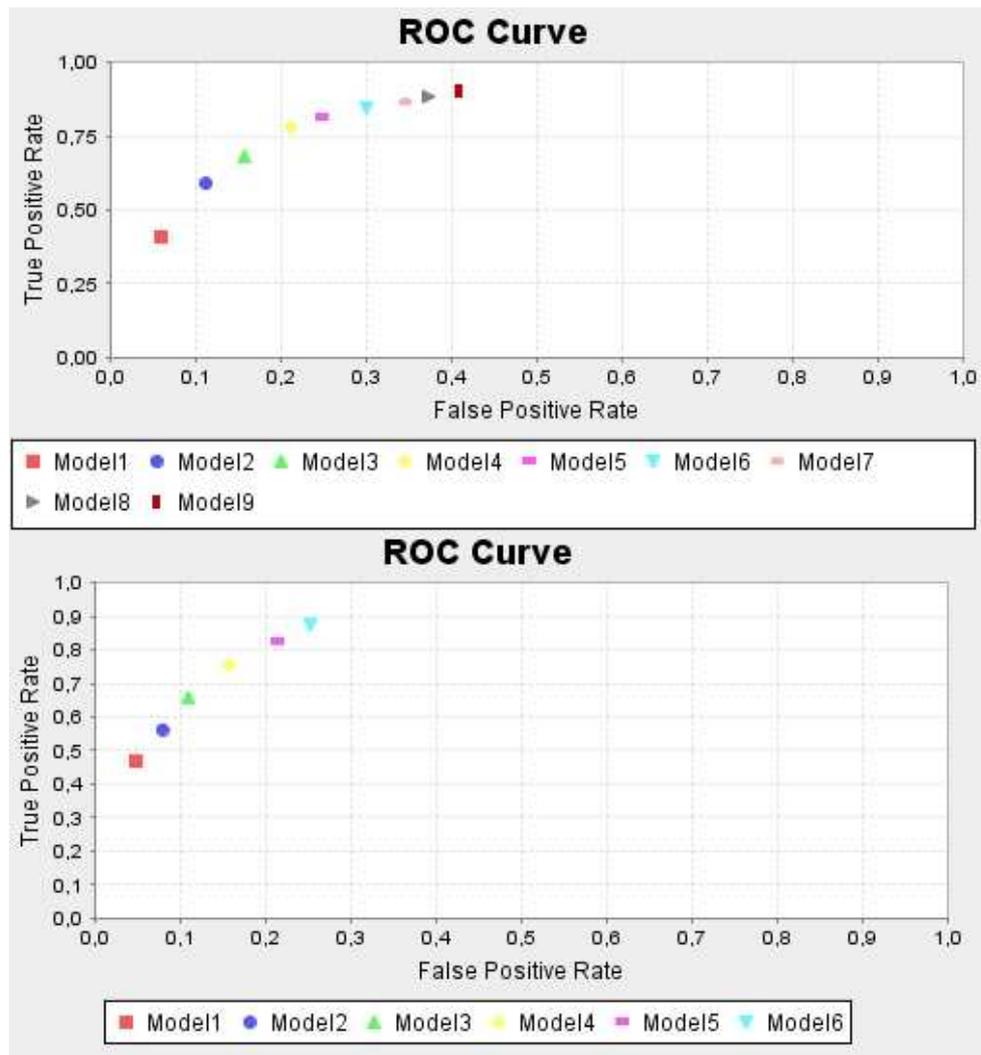


Figura 3.20: Curva ROC da abordagem semi-supervisionada dos modelos de compra, primeira; e venda, segunda

ROC e, conseqüentemente, melhor será o seu desempenho.

Os modelos de venda foram selecionados de forma análoga, com a diferença que selecionou-se apenas um modelo por abordagem. Adotou-se esta estratégia pela própria característica de investimento do sistema: compra-se quando o sistema detecta um padrão de *candlestick*. Um fato comum, principalmente para os modelos que apresentam baixa taxa de verdadeiros positivos (verificar figuras 3.19 e 3.20 para mais detalhes), é que várias tendências de alta terciária são perdidas; isso não afeta, no entanto, o desempenho do sistema, desde que ele seja capaz de detectar tendências de alta em número suficiente para obter lucros. Por outro lado, ao entrar em uma tendência de alta, a detecção do momento de reversão para a tendência de baixa torna-se primordial, pois todo o lucro obtido pode ser perdido, caso ela não seja detectada. Por esse motivo, o modelo de venda foi fixo para as duas abordagens, e os modelos

Tabela 3.2: Modelos da abordagem semi-supervisionada selecionados para testes

	<b>Modelo1</b>	<b>Modelo2</b>	<b>Modelo3</b>	<b>Modelo4</b>	<b>Modelo5</b>	<b>Modelo6</b>
<b>Distância</b>	0.24	0.33	0.36	0.40	0.38	0.36
<b>Especificidade(%)</b>	94.0	88.6	84.2	75.2	69.9	65.2
<b>Sensibilidade(%)</b>	40.6	58.8	67.9	81.8	84.4	86.6

selecionados possuíam alta taxa de verdadeiro positivo - mesmo que isso implicasse em uma alta taxa de falsos positivos - pois é preferível sair antes do final da tendência de alta com algum lucro, a tentar ficar até o final, mas com grandes chances de ficar sem lucro. As matrizes de confusão dos dois modelos são apresentadas nas tabelas 3.3 e 3.4.

Tabela 3.3: Matriz de confusão do modelo de venda selecionado, abordagem supervisionada

	<b>Positivo Predito</b>	<b>Negativo Predito</b>
<b>Verdeiro Positivo</b>	177	48
<b>Verdadeiro Negativo</b>	323	1139

Tabela 3.4: Matriz de confusão do modelo de venda selecionado, abordagem semi-supervisionada

	<b>Positivo Predito</b>	<b>Negativo Predito</b>
<b>Verdeiro Positivo</b>	177	48
<b>Verdadeiro Negativo</b>	323	1139

### 3.5 Conclusão

Discutiu-se, neste capítulo, uma possível solução para diminuir a quantidade de ruído inserida no conjunto de treinamento. Através da teoria do aprendizado semi-supervisionado, quando amostras sem rótulo podem melhorar a capacidade de generalização do modelo de aprendizagem de máquina, por trazer informações relevantes quanto à distribuição das classes. No entanto, a utilização do aprendizado semi-supervisionado neste trabalho não objetivou aumentar o conjunto de treinamento de forma a definir melhor a função de distribuição das classes, mas sim, solucionar o problema de padrões conflituosos, que se mostraram muito comuns nas séries temporais analisadas.

Para selecionar quais modelos seriam utilizados, o  $JSVM^{litc}$  desconsiderou as amostras duvidosas do conjunto de validação, isso faz com que, até o momento, não se tenha informação alguma quanto ao desempenho das duas abordagens. Esse "empate, *a priori*" das duas abordagens fica mais claro ainda

quando são selecionados os modelos com base na distância que cada um deles representa na curva ROC até a reta  $Sensibilidade = 1 - Especificidade$ . Os valores comparados são muito próximos para as duas abordagens, o que pode ser verificado nas tabelas 3.1 e 3.2.

Somente após a utilização das duas abordagens em uma simulação de um *trading system* que poder-se-á concluir, com base nos retornos, se a abordagem semi-supervisionada conseguiu minimizar o problema dos padrões conflitantes de *candlestick*. A simulação e posterior avaliação são temas de discussão do próximo capítulo.

---

# Resultados

---

**O**s capítulos anteriores deste trabalho ocuparam-se do problema de previsão financeira. Dissertou-se sobre as dificuldades gerais de se construir um *trading system* baseado em técnicas de inteligência computacional e, especialmente, sobre as dificuldades verificadas no decorrer da pesquisa que subsidia esta dissertação. A maior delas, sem dúvida, revelou-se a detecção de dados conflituosos no conjunto de treinamento. Toda a abordagem que vinha sendo desenvolvida, até aquele ponto, poderia ser perdida ou, então, os resultados a serem apresentados não seriam tão satisfatórios quanto se esperava.

No capítulo anterior, foi proposta uma nova abordagem para minimizar os dados conflituosos, a fim de diminuir o ruído do conjunto de treinamento e, possivelmente, aumentar a capacidade do sistema que viesse a utilizar aquele modelo. Treinou-se um par de modelos, compra/venda, utilizando a abordagem semi-supervisionada. Para que os testes pudessem ser realizados, foi necessário desenvolver um *trading system* completo, juntamente com um simulador da bolsa de valores. O desenvolvimento destas ferramentas serão descritos nos tópicos a seguir.

## 4.1 *Estratégia de Investimento*

O modelo que gera os sinais de compra e venda representa, apesar de mais importante, apenas uma parte do sistema de investimento, sendo preciso elaborar, ainda, uma estratégia de investimento, dado os sinais de compra/venda. Esse tipo de sistema pode ser comparado a um projeto de auto-

móvel, o motor tem que funcionar da melhor maneira possível, mas de nada adiantaria se as suspensões travassem constantemente, ou se a estrutura não oferecesse uma boa aerodinâmica.

A escolha da estratégia de investimento a ser utilizada pelo *trading system*, deve estar, intimamente ligada ao tempo de duração de cada negócio e ao tipo de sinal de compra/venda que é gerado. O modelo de detecção de padrões de reversão reconhece, principalmente, reversão na tendência terciária, que geralmente resulta em um negócio de curto prazo, de 5 a 10 dias. No entanto, pode acontecer de o modelo detectar uma reversão de tendência secundária, fazendo com que o tempo em que o capital fica aplicado na ação cresça para aproximadamente um mês.

Como o tempo de duração de cada negócio é relativamente curto, o tempo em que se fica sem possuir nenhuma ação encarteirada é relativamente alto. Ou seja, existe uma tendência a ficar mais tempo com o capital resgatado do que aplicado. Devido à alta taxa de aproveitamento do modelo (vide capítulo 3 para mais detalhes) é intuitivo que deve-se aproveitar da melhor maneira possível, quando o modelo detecta um ponto de reversão.

Por esses motivos, a estratégia de investimento utilizada neste trabalho pode ser definida como uma estratégia do "tudo ou nada". Pelo número de sinais de entrada no mercado ser relativamente baixo, quando um sinal é lançado, deve-se investir todo o dinheiro disponível naquela oportunidade, pois não se sabe, *a priori*, quando um novo sinal será lançado, e, se o modelo possui um bom aproveitamento; espera-se que, na média, os lucros superem os prejuízos.

A utilização deste tipo de estratégia necessita, entretanto, de alguns comentários. Investir todo o dinheiro em uma única ação, dado um sinal de compra, e resgatar todo o investimento, dado um sinal de venda, exige que as ações analisadas preencham um requisito. Na bolsa de valores, sempre que ocorre uma transação, obrigatoriamente, houve o desejo de alguém de vender, e de outra pessoa o de comprar. Então, de nada adiantaria tentar aplicar todo o capital em uma empresa que ninguém quer vender, ou pior, tentar vender todas as ações encarteiradas de uma empresa que ninguém quer comprar. É extremamente importante que as ações analisadas possuam boa liquidez. Além de boa liquidez, o capital investido nesta estratégia não deve ser nem muito baixo, devido às taxas de corretagem, e também não pode ser muito alto, para não prejudicar a liquidez, pois, mesmo ações com grande volume de negócios podem ter que esperar por muito tempo no *book* de ofertas, caso a transação possua cifras muito grandes. Nas simulações construídas para este trabalho, utilizou-se o valor inicial de R\$10000,00, mas acredita-se que é

possível investir valores superiores com a mesma eficiência.

Como especificado no capítulo anterior, o modelo de venda escolhido nas duas abordagens deveria ter alta taxa de verdadeiro-positivo, mesmo que isso viesse a implicar em um aumento no número de falsos-positivos. Isso se faz necessário, pois sair antes do final de uma tendência mas com lucro é muito mais interessante que tentar ficar até o final e correr o risco de ficar sem lucro. No entanto, apenas uma alta taxa de verdadeiros-positivos no modelo de detecção de pontos de venda, provavelmente, não seria suficiente, pois, pode acontecer de em uma tendência de alta, após uma valorização considerável da ação, ocorrer uma grande queda. Para esse tipo de situação e para os casos onde o modelo de compra erra e a tendência de alta não aparece, existe o *stop loss* [37], instrumento já mencionado anteriormente.

O *stop loss* funciona como um gerenciador de riscos. Em alguns negócios realizados pelo sistema, a tendência de alta demora um tempo para entrar e, até que isso venha a acontecer, prejuízos passageiros irão aparecer na carteira do *trading system*. Até quanto deste prejuízo é, porém, aceitável? Até que ponto deve-se esperar para que haja uma reação do papel e este passe a dar lucros. Esse valor de folga, que seria um prejuízo aceitável pela carteira, é definido no *stop loss*. Por exemplo, caso o valor seja definido como  $X\%$ , e, após uma compra, o valor do papel atinja  $X\%$  abaixo do valor de compra, um sinal de venda será lançado.

Existe uma outra versão do *stop loss* que se move juntamente com o preço do papel. Caso o preço da ação se valorize, o patamar de referência para a perda passa a ser o novo preço valorizado do papel. Essa versão é conhecida como *moving stop loss* [37].

Existem algumas técnicas para se determinar o valor do *stop loss*, através de resistências, médias móveis e outras. Mais uma vez, incorporar esses instrumentos técnicos poderia fazer com que o sistema perdesse em generalidade. Por isso o valor do *moving stop loss* foi determinado empiricamente. Apesar de determinado empiricamente, o *stop loss* de um *trading system* pode funcionar muito melhor que de um investidor profissional que tem à sua disposição todas as técnicas possíveis. Isso acontece, simplesmente pelo fato de o investidor *in silico* não ser afetado por emoções.

Como explicado no capítulo anterior, foram selecionados seis modelos diferentes de cada abordagem para a geração de sinais de compra. O modelo 1 é aquele em que são gerados menos sinais de reversão, no entanto, são sinais mais confiáveis. O modelo 6 é o oposto, gera mais sinais, apesar de menos confiáveis. Os outros quatro modelos são intermediários a estes dois. Pelo fato de o modelo 6 apresentar mais situações de prejuízo, pelo número maior

de falsos positivos, o seu *moving stop loss* deverá ser mais "justo", para evitar grandes perdas. Analogamente, o modelo 1 terá o valor mais folgado, com os outros modelos possuindo valores intermediários aos dois extremos. Os valores correspondentes a cada um dos modelos podem ser observados na tabela 4.1.

Tabela 4.1: Valor do *moving stop loss* utilizado em cada um dos modelos

<b>Modelos</b>	<b>Moving Stop Loss (%)</b>
1	8
2	7
3	6
4	5
5	4
6	3

## 4.2 Simulação da bolsa de valores

Para se avaliar um *trading system*, é necessário que se faça o teste por um longo período de tempo, para confirmar, estatisticamente a performance do mesmo. Caso fosse utilizado por apenas uma semana, por exemplo, em uma situação de sorte poder-se-ia ter ótimos resultados que não viriam a se repetir nunca mais. Por isso, para que o teste fosse possível por um longo período de tempo e fosse, concomitantemente, rápido de ser realizado, construiu-se um simulador da bolsa de valores.

O objetivo, ao construir o simulador, era que fosse possível testar o modelo durante o ano de 2006. Para tanto, após a construção de um simulador, com interface gráfica para facilitar o seu uso, algumas ações com boa liquidez foram selecionadas para compor o conjunto de testes do modelo. Foram elas: AMBV4, ELET6, EMBR3, GOLL4, ITAU4, LAME4, UOLL4 e USIM5. É oportuno ressaltar que nenhuma delas foi utilizada no conjunto de treinamento ou validação. No simulador da bolsa, as ações que iriam ser analisadas como possíveis geradoras de sinais de compra e venda, poderiam ser selecionadas facilmente através da interface gráfica, como pode ser observado na figura 4.1.

Assim como a aplicação "Consultora", que baixa diariamente os dados do BOVESPA e onde o modelo faz previsões para apenas um dia, o Simulador da Bolsa utiliza do mesmo banco de dados e tecnologias descritas nos capítulos anteriores. A aplicação que faz previsão para o dia será descrita, em maiores detalhes, mais adiante.

Pode-se observar na figura 4.1 todas as funcionalidades criadas nesse ambiente artificial para testar modelos de previsão de pontos de reversão em

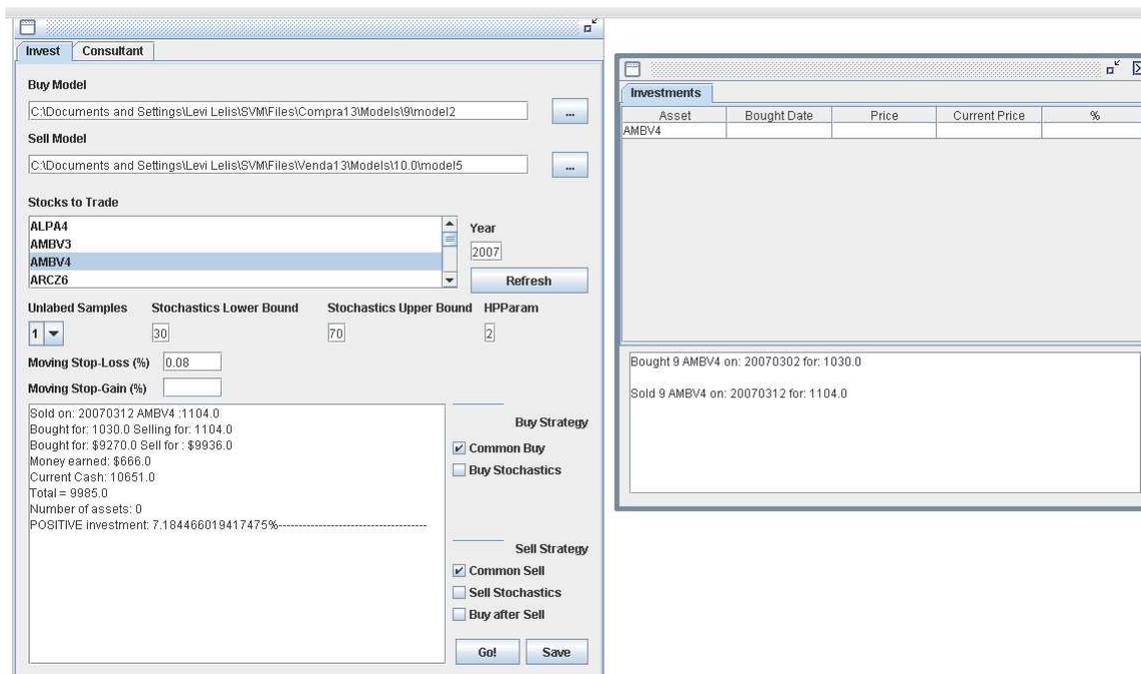


Figura 4.1: Simulador da Bolsa de Valores

tendências de séries financeiras. Na parte superior da aplicação selecionam-se os modelos de compra e venda que são utilizados para aquela simulação. Os modelos são os arquivos gerados pelo *SVM light* [24] contendo os vetores suporte, que foram selecionados durante o treinamento. Mais abaixo, é possível selecionar as ações que serão utilizadas em investimentos, esta seleção é fundamental para o desempenho do sistema, detalhes sobre ela serão discutidos ainda neste capítulo, na seção sobre otimização. Seleciona-se também o ano em que espera-se realizar a simulação. Além disso, é possível selecionar, o valor do *moving stop loss* utilizado. Por último, seleciona-se a estratégia de investimento a ser utilizada.

Algumas simplificações foram levadas em consideração na construção do ambiente de simulação da bolsa de valores. Ao escolher apenas ações com boa liquidez, considerou-se que todas as ações poderiam ser compradas automaticamente ao valor do preço de abertura do pregão após aquele que indicou o momento de compra. Para a venda, considerou-se possível vender sempre no valor de fechamento do pregão que gerou o sinal de reversão. Utilizou-se essa estratégia para venda, pois, além de ser possível realizar este tipo de transação, utilizando-se o *after market*, ela é interessante para evitar que uma abertura em forte baixa no dia seguinte atrapalhe o desempenho do sistema.

Com relação às taxas de investimentos, foi utilizado um valor de R\$15,00 por par de negócios, compra/venda como valor de corretagem. Sendo que os emolumentos e impostos foram desconsiderados, assim como os dividendos.

Até o momento no texto descreveu-se toda a estrutura do simulador que fora construído para testar os modelos de previsão de reversão de tendência. Restando agora, detalhar e comentar os resultados obtidos.

### 4.3 A simulação

A simulação, objetivando testar e avaliar os modelos, foi realizada no período de 16 de janeiro a 19 outubro de 2006, com as oito ações descritas no tópico anterior. Esperava-se, ao selecionar as oito ações, que as chances de gerar sinais de compra aumentassem e com isso, maximizar o tempo em que o sistema ficaria com o capital aplicado.

Além das duas abordagens, construiu-se um outro modelo de investimento baseado apenas em Análise Técnica, ou seja, sem a utilização de técnicas de Inteligência Computacional. Esse modelo de investimento, baseado no oscilador Estocástico, funcionava da seguinte forma: quando o indicador estava em uma das faixas de sobrecomprado, ou sobrevendido, e ultrapassava essa faixa para a intermediária, um sinal de compra, ou venda, era lançado. Em outras palavras, dada a faixa do Estocástico predileta dos investidores, que é de 20, para indicar papel sobrevendido e 80 para sobrecomprado, ao papel sair dessa faixa, um sinal era gerado. Por exemplo, caso a ação estivesse com o valor de 10 no estocástico e passasse no dia seguinte para 25, era gerado um sinal de compra.

Para o modelo sem técnicas de Inteligência Computacional, variou-se também o valor do *moving stop loss*, para verificar se alguma melhora poderia ser trazida ao seu desempenho. O resultado do desempenho das duas abordagens e do sistema baseado apenas em Análise Técnica pode ser conferido na tabela 4.2

Tabela 4.2: Resultados dos investimentos por modelo

Modelos	Rendimento(%)		
	Semi-Supervisionado	Supervisionado	Estocástico
1	78.86	60.65	-3.10
2	62.77	30.36	-1.80
3	60.28	36.13	2.10
4	52.52	50.35	-2.60
5	60.67	49.72	-0.01
6	40.83	31.89	-4.20

Percebe-se claramente, pelos resultados, que a abordagem semi-supervisionada trouxe melhoras significativas ao modelo de previsão. Estas melhoras, provavelmente, estão ligadas aos dados conflitantes de *candlestick* uma vez que, em todos os testes de modelos equivalentes, a abordagem que utilizou

amostras sem rótulo no treinamento obteve um desempenho superior. Se antes, apenas através das matrizes de confusão, não era possível verificar se havia diferenças entre os modelos, depois dos testes já é possível fazer uma inferência de que a rotulação automática utilizada na abordagem supervisionada inseriu muito mais ruído no conjunto de treinamento do que a semi-supervisionada.

É oportuno ressaltar também que, utilizando-se apenas análise técnica no investimento, no caso o oscilador Estocástico, não foi possível realizar lucros. Mesmo a abordagem supervisionada, com bastante ruído no conjunto de treinamento, obteve resultados expressivamente melhores, mostrando como a Inteligência Computacional, combinada aos indicadores da Análise Técnica, pode aumentar a margem de lucro dos grafistas.

Apesar de esse simulador ser um ótimo parâmetro para verificação do desempenho do sistema, ele possui muitas simplificações. Estas simplificações não colocam em dúvida o desempenho do modelo, mas é interessante que o sistema seja testado em um ambiente que levasse em consideração fatores que podem influenciar no resultado final dos investimentos, como a liquidez dos papéis operados. Para esses testes, utilizou-se o Folha em Ação - <http://www.emacao.com.br> - que é uma parceria da BOVESPA com a Folha de São Paulo, bastante popular, possuindo milhares de participantes é o melhor laboratório de treinamento para investimentos na bolsa. A seguir, discutem-se, detalhadamente, os resultados obtidos.

## 4.4 *Investindo em um simulador mais real*

O Folha em Ação utiliza dados reais da BOVESPA, com um atraso de apenas 15 minutos, tornando-se assim, um simulador bastante real, mesmo para negociações *intraday*. O grande inconveniente da utilização desse simulador está nas limitações aplicadas a ele para que haja um maior equilíbrio na competição promovida entre os participantes. Começa-se com aproximadamente R\$200000,00, sendo que 50% deste valor já vem aplicado em ações da BOVESPA que têm boa liquidez. Limita-se, então, que cada participante movimente apenas R\$100000,00 diários, para compra ou para venda. Esta limitação atrapalha, consideravelmente, a estratégia “tudo ou nada” descrita no capítulo anterior, pois, ao investir todo o capital - R\$200000,00 - não é possível que a aplicação seja feita em apenas um dia. O mesmo ocorre quando deve-se vender, o resgate do capital deve ser realizado por partes.

Além da limitação de movimentação de caixa, os valores de corretagem são excessivamente altos. Sugere-se que, ao utilizar o *trading system* descrito

neste trabalho, que se procure trabalhar com corretagem fixa e investir um valor que valha por aquela corretagem. A tabela com os valores de corretagem utilizados pela Folha em Ação estão na tabela<sup>1</sup>

Tabela 4.3: Corretagem para investimentos no simulador da Folha de São Paulo

<b>Valor da Operação(R\$)</b>	<b>Taxa de Corretagem</b>
Até R\$ 135,07	2,70 (parte fixa)
Acima de R\$ 135,08 até R\$ 498,62	2%
Acima de R\$ 498,63 até R\$ 1.514,69	1,5% + R\$ 2,49
Acima de R\$ 1.514,70 até R\$ 3.029,38	1% + R\$ 10,06
Acima de R\$ 3.029,39	0,5% + R\$ 25,21

Além do valor de corretagem, apresentado pela tabela 4.3, deve-se considerar o valor dos emolumentos que, apesar de pequenos, pode influenciar no resultado final. O simulador da Folha pode ser considerado como um limite inferior para o desempenho do sistema, já que o ambiente é altamente desfavorável para a estratégia adotada pelo *trading system* descrito neste trabalho. Já o simulador construído apenas para testar o sistema pode ser considerado como um limite superior, onde tem-se um ambiente favorável ao modelo de previsão.

Apesar de ser considerado como um limite inferior para o desempenho do sistema, aprendeu-se uma estratégia muito interessante quando testava-se o sistema no simulador. Por não poder realizar toda a compra de uma só vez, comprava-se metade em um dia (após o sinal de compra ter sido gerado) e a outra metade no dia seguinte. O modelo apresenta bom desempenho, mas nem todos os negócios realizados por ele começam com um lucro imediato; muitas vezes o preço cai um pouco mais, até entrar realmente na tendência de alta. Ao dividir a compra total em duas partes, pode-se melhorar o preço médio da compra, caso haja uma queda no segundo dia de compra. Por outro lado, caso ocorra um aumento no preço, irá aumentar também o preço médio, diminuindo os lucros. Acredita-se que, para investidores mais cautelosos, essa estratégia possa ser bem interessante.

#### 4.4.1 Negócios realizados no Em Ação

Os testes no simulador da Folha de São Paulo foram realizados durante o mês de março de 2007. Quando, além do negócio principal, que se deu através da Usiminas, realizaram-se outros com a finalidade de testar o sistema apenas. Um deles foi com ações da TAM que, em tendência primária de baixa, ocorreu um *pull back* e foi possível obter lucros.

---

<sup>1</sup>Extraído diretamente do *site* do simulador, <http://www.emacao.com.br>

O tempo em que se ficou encarteirado com ações da Usiminas - USIM5 - foi relativamente alto, aproximadamente um mês; no entanto, esse tipo de negócio pode acontecer, sendo mais comum, quando o mercado está com alta volatilidade, que era o caso no mês de março de 2007. A aplicação ocorreu alguns dias após a queda histórica da bolsa de Xangai que, derrubou as bolsas do mundo inteiro.

Os *candlesticks*, na figura 4.2, mostram os pontos de compra e venda da Usiminas e a alta volatilidade presente entre os pontos de compra e venda.

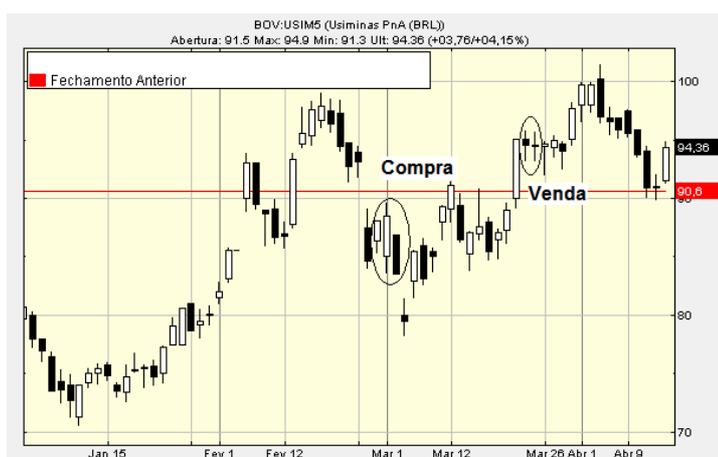


Figura 4.2: Pontos de compra e venda da USIM5 - aplicação realizada no simulador Em Ação

Como citado no início deste texto, escolheu-se aplicar segundo a tendência terciária para poder aproveitar, inclusive, pequenas movimentações contrárias à tendência secundária e primária, conhecidas como *pull backs*. Ocorreu um caso em que se ganhou ao entrar em um *pull back* das ações preferenciais da TAM - TAMM4 -, que se encontrava em uma forte tendência primária de baixa. Ao detectar um momento de reversão, aplicou-se no papel, alguns dias depois, apareceu um novo padrão de reversão, desta vez indicando venda, ou retorno à tendência principal de baixa. Os detalhes podem ser conferidos na figura 4.3. Conseguiu-se dessa forma, através de um investimento de curto prazo, um lucro de quase 4%, em uma ação em que a maioria dos investimentos perdiam, pois até o final do mês de março, a TAM não mostrou sinais evidentes do final da tendência de baixa.

Ao final do negócio realizado com a Usiminas, retirou-se, diretamente do *site* do simulador, uma imagem do rendimento daquele negócio na carteira. Como foi o único negócio da carteira no mês, ele representa o desempenho total no mês. A carteira pode ser visualizada na figura 4.4.

O questionamento mais natural que se aflora neste ponto do desenvolvimento do *trading system*, é se é possível melhorar ainda mais os resultados

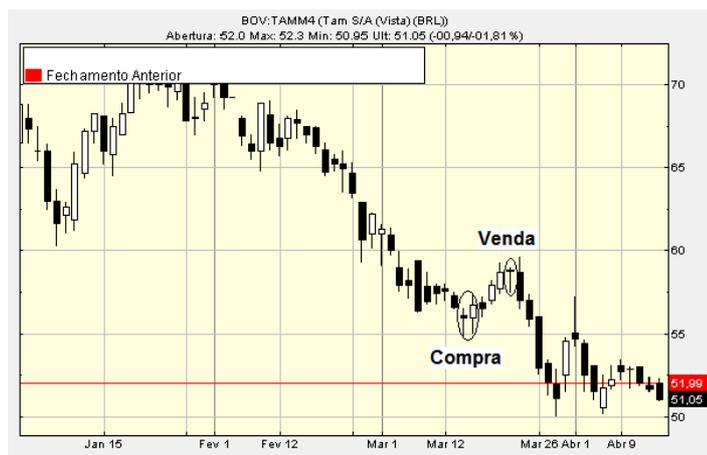


Figura 4.3: Lucro em um *pull back* da TAMM4

AÇÃO	Nome	Sector	Data/Hora Compra	Qtd.	Valor Compra Médio	Valor Atual	Valor Total	Ganho Perda	%
USIM5	USIMINAS PNA	c v META	01/03 18:59	200	87,94	95,20	19.040,00	1.452,00	8,26
<b>Total Ações</b>									
Total						<b>17.588,00</b>	<b>19.040,00</b>	<b>1.452,00</b>	<b>8,26</b>

Figura 4.4: Rendimento da carteira na simulação do Em Ação, em março de 2007

trazidos pela abordagem semi-supervisionada. Seria possível utilizar uma estratégia de investimento que otimizaria os lucros? Ou será que uma seleção mais criteriosa das ações que ficam disponíveis para análise do modelo não traria resultados melhores? Estas perguntas são respondidas na próxima seção, em que será detalhada uma das estratégias de otimização de carteiras mais antigas e como ela pode ser aplicada a este sistema.

## 4.5 Otimização dos investimentos

A estratégia de investimento utilizada para os modelos de previsão de reversão de tendência se mostrou, apesar de simples, bastante eficiente. É intuitivamente perceptível que, para maximizar a margem de lucros do sistema, é necessário aplicar todo o dinheiro disponível, desta forma, poder-se-á aproveitar as tendências de alta, obtendo-se rendimento máximo da carteira. Entretanto, da mesma forma que um investimento correto gera lucro máximo, um investimento incorreto gera prejuízo na mesma proporção, já que todo o dinheiro está aplicado e, assim, todo ele sofrerá a desvalorização. Uma vez que a margem de acerto do sistema é alta, vide capítulo anterior, faz-se com que, na média, os lucros superem os prejuízos.

Neste momento torna-se pertinente levantar o seguinte questionamento:

seria possível propor outra estratégia que otimizaria os investimentos? Pela própria característica do sistema de gerar sinais, muitas vezes esparsos, de compra e venda, a estratégia de aplicar todo o capital disponível na empresa em que o sinal foi gerado, mostra-se como uma ótima estratégia, por maximizar o tempo do dinheiro aplicado e conseqüentemente maximizar os lucros. Variações na estratégia de investimento poderiam ser testadas em dias em que são gerados mais de um sinal de compra. Neste caso, o investimento poderia ser dividido entre as empresas ou, até mesmo, selecionar uma das ações e investir todo o capital nela.

O sistema de detecção de padrões de *candlesticks* trabalha com um atraso de apenas 5 dias, desta forma, ele não tem informações históricas para saber selecionar em quais ações investir, dados vários sinais de compra. Para otimizar as aplicações, sendo  $p$  o preço de fechamento da ação em um dado instante  $t$  de tempo, a saída  $S$  do sistema teria que ser dada por:

$$S = \{r, d\} \quad (4.1)$$

onde,

$$r = \begin{cases} 1 & \text{se ponto de reversão} \\ -1 & \text{se ponto de continuação} \end{cases} \quad (4.2)$$

e,

$$d = \frac{\partial p}{\partial(t+1)} \quad (4.3)$$

Com base na inclinação da tendência no próximo intervalo de tempo, seria possível otimizar a porcentagem de investimento que seria feita em cada papel. No entanto, os padrões de *candlestick* não determinam qual o tamanho da tendência, ou, em outras palavras, qual o valor do papel em um dado dia do futuro; os padrões apenas indicam se haverá ou não uma reversão de tendência. Para determinar o valor esperado no final da tendência, poderia utilizar-se outras técnicas de Análise Técnica, como Suporte e Resistências Históricas [37]. O grande problema em incorporar essas análises seria a possível perda de generalidade do sistema. Os padrões de *candlestick* são genéricos, independem da empresa ou da época em questão, o que coloca o sistema proposto neste trabalho em vantagem considerável com relação a outros *trading systems* que também fazem uso da Inteligência Computacional. Estes, na sua maioria, necessitam de retreino de tempos em tempos. O momento de retreinar é muito complicado de se determinar, pois é difícil discernir se o sistema está apenas passando por uma fase ruim ou se o seu desempenho realmente caiu e um

outro treinamento é necessário.

A solução para melhorar os investimentos do sistema foi encontrada na teoria de seleção de portfólio, desenvolvida por Harry Markowitz em 1952 [36].

## 4.6 Seleção de Portfólio

É consenso entre os investidores que suas carteiras devem maximizar os lucros e, ao mesmo tempo, minimizar o risco [36]. Para tornar mais claras essas definições, toma-se, como exemplo, uma variável aleatória  $Y$ , sendo que ela pode assumir valores finitos,  $y_1, y_2, \dots, y_N$ . Supondo que cada valor  $y_i$  ocorra com uma probabilidade  $p_i$ . O valor esperado, ou média de  $Y$ , é dado por:

$$[y] = p_1y_1 + p_2y_2 + \dots + p_Ny_N \quad (4.4)$$

Um investimento livre de riscos é aquele que sempre resulta no valor esperado, ou  $p_1 = 1$  e  $p_2, \dots, p_N = 0$ , fazendo com que  $[y] = y_1$ . O risco pode, pois, ser calculado como a dispersão do retorno com relação ao seu valor esperado. Uma medida de dispersão comumente utilizada é a variância, dada por:

$$\sigma^2 = p_1(y_1 - E)^2 + p_2(y_2 - E)^2 + \dots + p_N(y_N - E)^2 \quad (4.5)$$

O risco de um investimento pode ser dividido em duas categorias: o diversificável e o não diversificável. O não diversificável trata do risco que atinge o mercado como um todo, não dependendo, portanto, da saúde financeira da empresa ou do seu segmento de atuação [26]. Um exemplo interessante foi o ocorrido no dia 26 de fevereiro de 2007, quando o governo Chinês tomou algumas medidas de forma a desacelerar o crescimento da economia do seu país, causando uma queda de aproximadamente 8% na bolsa de Xangai. Este se juntou a outros fatores, como o fato de a maioria das bolsas em todo o mundo já estar nos seus valores próximos ao máximo. O que se viu, nesse processo, foi um efeito dominó, que derrubou as bolsas do mundo inteiro.

Como o próprio nome já diz, o risco diversificável pode ser minimizado ao diversificar a carteira de investimentos. Por exemplo, se o preço do petróleo aumenta, existe uma tendência de que as empresas petrolíferas tenham os seus papéis valorizados. No entanto, as transportadoras, pelo aumento do custo de seu principal insumo, tendem a ter uma desvalorização. Ao utilizar uma carteira composta por empresas desses dois segmentos, o investimento tende a ter o seu risco - variância - minimizado, causado pela alta do preço do barril de petróleo.

A idéia da minimização do risco (a partir de agora, risco estará se refe-

rindo ao diversificável) se baseia na diversificação dos investimentos. Como citado no exemplo sobre o aumento do preço do barril de petróleo, supôs-se que as ações que compunham a carteira tinham correlação negativa, quando uma empresa desvalorizava, a outra valorizava, fazendo com que o investimento como um todo não sofresse variações. Os investidores frequentemente utilizam uma analogia com cestas e ovos para descrever este processo de diversificação: divide-se os ovos em várias cestas para, caso uma delas caia, nem todos os ovos estarão perdidos. Na estratégia de investimento utilizada nesse trabalho, descrita anteriormente, todos os ovos eram colocados em apenas uma cesta, a diferença é que eles eram trocados de cesta constantemente, em investimentos de curto para médio prazo. A idéia é que o modelo de previsão de reversão escolha qual a cesta do momento, de modo a obter maior lucro.

Para calcular o retorno esperado de uma carteira de ações, basta somar os retornos esperados multiplicados pela participação de cada empresa na carteira, ou,

$$[X] = \sum_{i=1}^n X_i \mu_i \quad (4.6)$$

onde,

$X_i$  = participação do ativo  $i$  na carteira

$\mu_i$  = retorno esperado do ativo  $i$

No entanto, o cálculo da variância da carteira não se dá de uma forma direta, para tanto, faz-se necessária a definição de covariância. A covariância entre as empresas  $i$  e  $j$ , por exemplo, é dada, em termos da correlação,  $\rho_{ij}$ , por,

$$\sigma_{ij} = \rho_{ij} \sigma_i \sigma_j \quad (4.7)$$

somando par a par cada uma das covariâncias, tem-se o risco da carteira:

$$R = \sum_{i=1}^n \sum_{j=1}^n X_i X_j \sigma_{ij} \quad (4.8)$$

Esta foi a definição feita por Markowitz [36], que lhe rendeu mais tarde o prêmio Nobel de Economia.

Pode-se notar por 4.8 e 4.7 que quanto maior a correlação entre as séries de preço das empresas, maior será o risco da carteira. Por outro lado, caso haja empresas com correlação negativa, o risco será diminuído, esta é

a minimização de risco por diversificação. A figura <sup>2</sup> 4.5 mostra a relação do risco versus retorno de uma carteira com duas ações, de acordo com a correlação entre elas. Observa-se que, caso elas possuam correlação igual a -1, o risco pode ser anulado, e, no outro extremo, com correlação igual a 1, o risco cresce proporcionalmente com o retorno, não havendo assim, possibilidade de minimizar o risco diversificável combinando as duas empresas.

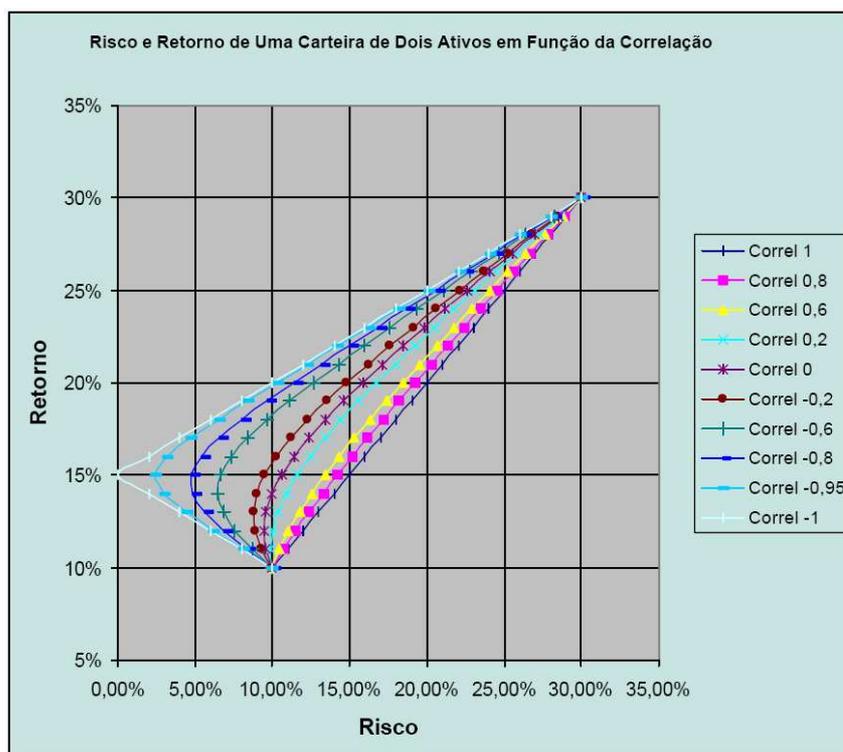


Figura 4.5: Análise de correlação no gráfico de Retorno por Risco

À primeira vista, pode parecer paradoxal a utilização da teoria de seleção de Markowitz, que se baseia na diversificação, em um sistema que não mantém investimentos em mais de uma ação por vez. No entanto, o que a seleção de ações faz, é selecionar as melhores empresas em que suas séries de preço são pouco correlacionadas. Quando duas empresas fortemente correlacionadas são utilizadas pelo sistema de previsão, o que pode vir a acontecer é a empresa com menores retornos históricos, menor saúde financeira, gerar o sinal de compra alguns dias antes da empresa que possui retorno médio mais elevado. Desta forma, todo dinheiro será investido no primeiro sinal de compra e, com isso, o melhor investimento, que viria depois, não poderia ser realizado. Ao utilizar a seleção de Markowitz, a tendência é de que seja gerado apenas um sinal por vez e, esses sinais são mais confiáveis, uma vez que apenas empresas com bons históricos são disponibilizadas para análise do sistema.

<sup>2</sup>extraído de [26]

É claro que, para aumentar o lucro, é necessário maximizar o tempo do capital aplicado; com a minimização do risco, isso acontece naturalmente, pois, espera-se que haja sempre alguma empresa na carteira saindo de sua tendência de baixa e revertendo para uma tendência de alta.

## 4.7 Exemplo de utilização da Seleção de Carteira

Nesta seção mostrar-se-á como a seleção de carteira pode ser utilizada para melhorar o retorno de uma carteira com estratégia de longo prazo, em que aplica-se o capital e espera-se por um longo período [26]. Neste tipo de estratégia, o investidor compra as ações de uma empresa com boas perspectivas e espera até que os papéis se valorizem.

Para a seleção da carteira, dados históricos, de junho de 2005 a dezembro do mesmo ano, das seguintes empresas foram utilizados: AMBV4, ELET6, EMBR3, GOLL4, ITAU4, LAME4, USIM5, SDIA4 e SUBA3. A tabela 4.4 mostra o retorno de cada papel, que foi calculado dentro de um horizonte de 5 dias.

O valor esperado de retorno para o horizonte de tempo, neste caso de 5 dias, é dado pela média dos retornos, o risco, por sua vez, pelo desvio padrão. Os valores são mostrados na tabela 4.5.

Ao utilizar-se de uma carteira igualmente distribuída entre as ações da tabela 4.5, chegou-se aos resultados exibidos na tabela 4.6.

Surge, assim, o objetivo de melhorar o retorno esperado da carteira com o menor risco possível. Para isso, formulou-se um problema de otimização, em que o retorno foi fixado por um limite inferior e dever-se-ia minimizar o risco da carteira. O problema foi modelado da seguinte forma:

$$\text{Minimizar: } \sum_{i=1}^n X_i \sum_{j=1}^n X_j \sigma_{ij} \quad (4.9)$$

Sujeito às seguintes restrições:

$$\sum_{i=1}^n X_i = 1 \quad (4.10)$$

$$\sum_{i=1}^n X_i \mu_i \geq 2 \quad (4.11)$$

Fixou-se o valor desejado de retorno como 2%. Ao resolver este problema de otimização, encontraram-se as proporções de participação na carteira mostradas na tabela 4.7.

Ressalta-se que algumas ações tiveram suas participações zeradas na carteira. Isso se deu pelo fato de existirem empresas que possuíam alta correlação

Tabela 4.4: Retornos, de 5 dias, em 2005

Retorno(%)								
AMBV	ELET6	EMBR3	GOLL4	ITAU4	LAME4	USIM5	SDIA4	SUBA3
0,72	1,75	2,19	6,43	0	0,93	6,2	-1,28	-0,59
0,14	4,59	-1,2	0,26	-5,66	-1,18	-14,97	-6,67	-10,45
1,47	-1,87	-0,64	-5,4	0,47	2,59	-1,63	1,97	11,49
0,16	7,9	3,96	7,03	1,55	-1,15	10,74	7,1	3,51
-9,04	1,64	6,94	-6,49	7,08	6,74	-8,87	10,94	-5,03
5,48	-3,86	4,13	-1,53	-1,01	2,17	-1,75	3,04	7,06
2,95	-4,72	-3,31	1,33	-1,35	-2,47	0,73	0,68	-2,96
0,9	2,08	-5,1	-3,09	-0,23	-2,11	-5,15	-4,43	0
0,9	-6,56	-2,81	-2,49	-0,46	-0,43	-1,17	2,24	0
0,69	2,34	-1,73	-1,75	1,72	-4,74	2,32	5,09	8,03
-1,82	-2,72	4,93	1,32	0	4,65	-3,19	-4,54	8,86
-0,96	-7,2	-1,68	1,57	-4,87	-1,54	-6,07	-0,21	-5,33
2,18	12,5	1,67	4,85	8,98	3,26	10,8	4,55	3,53
-1,6	-2,22	0,76	4,62	-3,23	-1,47	2,29	1,43	3,99
-0,26	-4,48	5,23	2,32	3,01	0,67	-0,83	-1,57	0,95
0	0,57	1,72	1,05	-1,65	-2,66	1,94	2,61	8,42
-0,26	4,88	-0,32	-1,63	3,1	1,02	2,97	2,53	2,58
5,23	2,92	0,06	-0,23	4,66	-1,13	-1,64	6,14	-0,78
2,11	6,61	-1,88	-4,42	-4,09	-1,97	1,23	9,47	-2,29
-0,61	4,76	-0,5	3,85	6,61	0,45	7,13	2,83	4,85
-0,73	9,3	-1,2	-2,83	-3,59	3,03	-8,65	0	1,27
3,89	-2,88	-1,76	-3,01	0	0,96	-3,24	3,26	13,01
-0,67	7,53	-2,22	4,11	2,55	-2,37	3,3	-7,87	0,99
-5,59	-9,11	0,56	3,29	-2,17	-2,6	-10,73	-6,18	-1,65
6,53	4,99	-1,71	3,45	5,7	7,41	9,59	-0,19	-1,96
2,14	5,06	4,87	14,92	4,63	5,92	8,22	9,41	11,33
2	-0,24	-0,94	6,26	3,94	1,42	3,01	0,17	-1,63
2,33	-2,53	3,16	3,65	3,17	11,13	8,41	-4,73	3,59
2,81	0,13	5,8	1,01	1,71	7,86	1,67	9,74	9,58
4,35	2,93	0,97	14,12	2,13	1,61	6,65	10,88	9,2
-2,73	-2,31	7,35	4,79	-3,5	-3,63	1,83	-4,84	8,11
1,16	-2,07	6,79	12,67	-0,7	11,06	4,16	10,5	-4,34
3,37	2,83	-1,62	4,52	1,06	1,36	3,53	-0,78	-4,37

Tabela 4.5: Retorno esperado e risco

	AMBV	ELET6	EMBR3	GOLL4	ITAU4	LAME4	USIM5	SDIA4	SUBA3
<b>Retorno (%)</b>	0,83	0,99	0,98	2,26	0,9	1,36	0,87	1,86	2,39
<b>Risco</b>	3,05	5,02	3,32	5,14	3,57	4,05	6,22	5,39	5,75

Tabela 4.6: Resultado da carteira sem otimização

Retorno	Risco
1,37%	6,65

Tabela 4.7: Resultado da otimização

	AMBV	ELET6	EMBR3	GOLL4	ITAU4	LAME4	USIM5	SDIA4	SUBA3
<b>Participação (%)</b>	0	7	0	33	0	17	0	13	3

positiva entre si, e, através da otimização, seleccionaram-se as que ofereciam melhor relação retorno e risco para o desempenho desejado de 2%. Os resultados para esta configuração de carteira é apresentado na tabela 4.8.

Tabela 4.8: Resultado da carteira otimizada

Retorno	Risco
2%	9,43

Pode-se observar que, apesar da ligeira alta no risco, obteve-se uma melhora considerável no retorno da carteira, apenas modificando as participações das empresas. Resultados diferentes poderiam ser obtidos bastando variar o lucro desejado. Poder-se-ia também obter uma carteira com o mesmo lucro da anterior, tabela 4.6, mas, com um risco menor. Nesse caso esperava-se que o lucro aumentasse e para isso, foi necessário um leve crescimento no risco.

Com base na otimização realizada com valores históricos de 2005, testou-se o rendimento realizado pelas duas carteiras com a estratégica de longo prazo de aplicar e esperar, durante todo o ano de 2006. Na tabela 4.9 encontram-se os resultados dos investimentos para as duas carteiras.

Tabela 4.9: Proporção de participação de cada empresa nas carteiras

	AMBV	ELET6	EMBR3	GOLL4	ITAU4	LAME4	USIM5	SDIA4	SUBA3	Rendimento
Participação Otimizada(%)	0	7	0	33	0	17	0	13	3	17,15
Participação Não Otimizada(%)	0,11	0,11	0,11	0,11	0,11	0,11	0,11	0,11	0,11	14,76

Nota-se, portanto, uma ligeira melhora no resultado da carteira otimizada. Melhoras significativas foram, na realidade, notadas quando aplicou-se a seleção de ações no sistema de detecção de pontos de reversão.

## 4.8 Teoria de Markowitz aplicada ao sistema de previsão

Uma das grandes desvantagens, até o momento, do sistema de previsão de pontos de reversão em séries financeiras é o fato de o sistema não fornecer, como saída, a proporção que deve ser feita do investimento no caso de mais de um sinal de compra ser gerado ao mesmo tempo. Dever-se-ia investir tudo em apenas uma das empresa? Dividir igualmente? Qual seria o critério utilizado para a tomada desta decisão? Em uma situação mais pessimista, caso seja gerado apenas um sinal, a atitude mais sensata seria investir todo o dinheiro na empresa em que o sinal fora detectado, no entanto, o que garante que no dia seguinte uma outra empresa não geraria um sinal de compra, e aquele

sinal não seria a melhor opção? Como explicado anteriormente, a Teoria de Seleção de Portfólio de Markowitz foi utilizada para minimizar este problema.

Ao utilizar todas as ações citadas no item anterior, o sistema de investimento obteve 129,66% durante o ano de 2006 (período entre 19 de janeiro e 31 de dezembro de 2006). Ao restringir os investimentos do sistema para somente aquelas ações que obtiveram participação maior que zero na otimização - apresentada anteriormente -, o sistema obteve, no mesmo período, um rendimento de 192,78%. Detalhes podem ser conferidos na tabela 4.10

Tabela 4.10: Resultado da aplicação da Teoria de Markowitz ao sistema de previsão

	<b>Investimentos Não Realizados</b>	<b>Erros</b>	<b>Acertos</b>	<b>Rendimento(%)</b>
<b>Participação Otimizada(%)</b>	10	2	14	192,78
<b>Participação Não Otimizada(%)</b>	17	3	16	129,66

A coluna “Investimentos não realizados” na tabela 4.10 indica os sinais de compra que foram gerados, mas não puderam ser realizados pelo fato de o sistema já estar com todo o capital aplicado. Quanto maior o número de investimentos não realizados, indica um maior número de empresas correlacionadas existentes na carteira e, conseqüentemente, mais escolhas por ações inadequadas podem ser realizadas. Um exemplo de escolha errada de investimento, e que justifica, em partes, a diferença nos rendimentos das duas carteiras, se deu no dia 8 fevereiro de 2006, quando foram gerados sinais de compra da AMBV4 e SUBA3, e o sistema não otimizado escolheu, de forma aleatória, AMBV4 como seu investimento. Nessa ocasião, o investimento gerou um rendimento de 3.86% entre os dias 8 e 16 daquele mês. Os momentos de compra e de venda podem ser verificados na figura 4.6. A seta para cima indica o momento de compra e para baixo o momento de venda.

O sistema que passou pela seleção de ações não possui AMBV4 na carteira, por isso, ao invés de investir todo o capital nesta empresa, o sistema apostou em SUBA3. Esse investimento se mostrou perfeito, tanto no momento de entrada como no de saída, o que gerou um rendimento de 24.44% entre os dias 8 e 20 de fevereiro. O momento de entrada e saída estão indicados pelas setas na figura 4.7. Deve-se notar o padrão de reversão de *candlestick*, martelo, no dia indicado pelo sistema como de venda.

O número de erros em uma carteira otimizada tende a diminuir, pois, além de as empresas apresentarem bom desempenho de investimento em conjunto, elas apresentam bons históricos de investimento, o que justifica o melhor

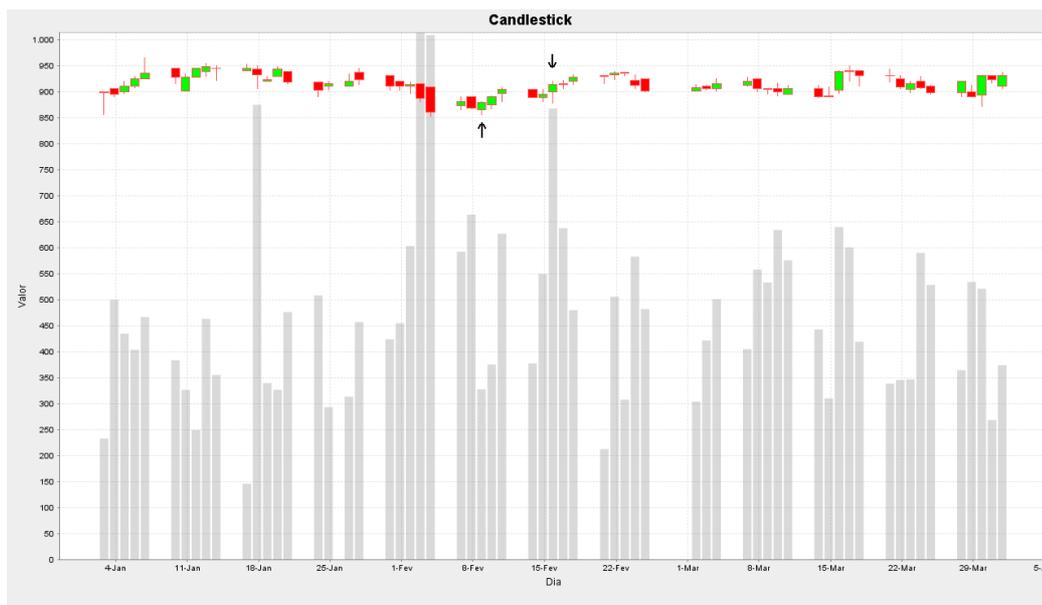


Figura 4.6: *Candlestick* - AMBV4

aproveitamento da carteira otimizada. A carteira que não passou pelo processo de otimização já apresenta ótimo desempenho (84,21%) nos seus investimentos, no entanto, essa taxa tornou-se melhor ainda com a teoria de Markowitz [36]: 87,5%

Os gráficos em *candlesticks* de todos os investimentos (momento de entrada e saída) das carteiras otimizadas de 2006 e 2007 podem ser conferidos nos Apêndices.

## 4.9 Conclusão

Neste capítulo pôde-se perceber a sensível melhora que a abordagem semi-supervisionada trouxe ao modelo de previsão de reversão de tendência. Mostrou-se também, como é possível lucrar satisfatoriamente sem a necessidade de entender sistematicamente sobre Bolsa de Valores ou Análise Técnica: tudo que se tem a fazer é seguir criteriosamente os sinais de compra e venda lançados pelo sistema, mesmo que seja em um ambiente completamente desfavorável à estratégia de aplicar todo o capital em apenas uma empresa por vez, como é do ambiente da Folha em Ação.

Um outro tema, bastante interessante, também abordado neste capítulo, foi a teoria de seleção de portfólio, desenvolvida por Harry Markowitz, em 1952 [36]. Essa teoria, desenvolvida em uma época que o mercado era considerado extremamente eficiente e que só era possível lucrar com investimentos a longo prazo, ou em qualquer outro caso, apenas por sorte; se mostrou atual e apli-

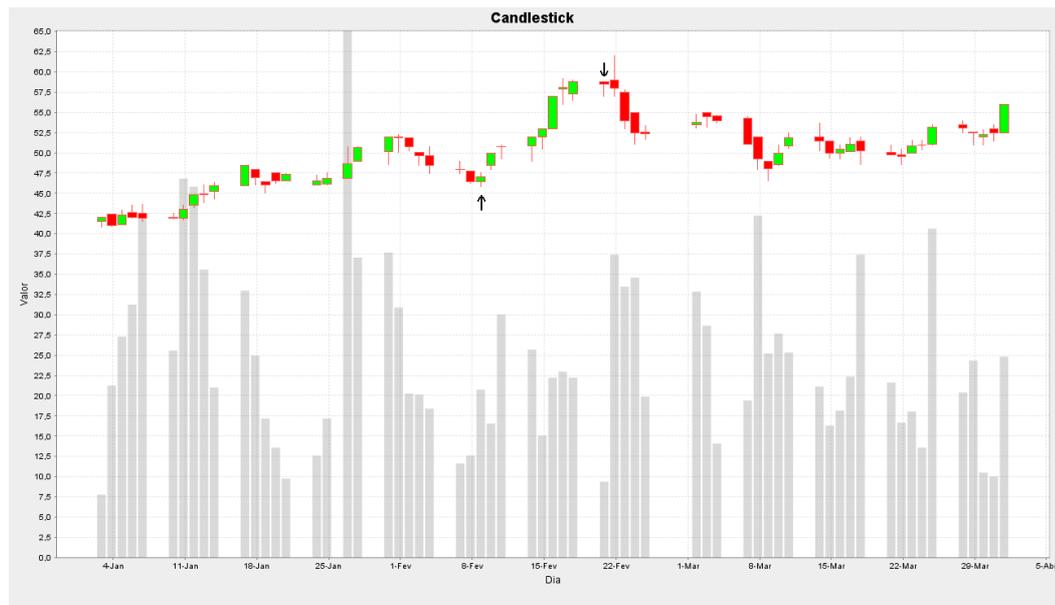


Figura 4.7: *Candlestick* - SUBA3

cável ao sistema descrito nesse trabalho, melhorando significativamente o seu retorno.

No capítulo seguinte serão levantadas algumas questões sobre este trabalho, assim como conclusões e sugestões para trabalhos futuros, envolvendo previsão financeira com técnicas de Inteligência Computacional.

---

## Discussões e Conclusões

---

A proposta deste trabalho de construir um *trading system* baseado em técnicas de Inteligência Computacional e em Análise Técnica realizou-se com êxito. Verificaram-se várias possibilidades de utilização do estudo dos gráficos para construção de tal sistema. No entanto, surgiu, no decorrer da pesquisa, um problema com relação à utilização de algumas dessas abordagens: a diminuição da generalidade do modelo. O principal requisito do sistema de previsão, dissertado nesse trabalho, preza pela generalidade do modelo, eliminando, assim, uma série de possibilidades de entrada para o modelo. Nesse capítulo serão discutidas as tentativas realizadas em busca de o modelo autônomo e genérico. Serão feitas também algumas sugestões de trabalhos futuros na área de previsão financeira.

### 5.1 Abordagens falhas

Até chegar ao resultado apresentado nos capítulos anteriores, várias tentativas de se contruir um modelo de previsão foram realizadas. Apesar de, aparentemente não apresentar bons resultados, é produtormente comentar algumas delas, para, em trabalhos futuros, os testes realizados sem sucesso neste trabalho não venham a se repetir.

Uma das abordagens cuja utilização mostrou-se desnecessária, sendo, no entanto, de grande importância para decisões futuras, se deu com o uso de um comitê de Redes Neurais Artificiais [44]. Cinco redes neurais foram separadas, cada uma delas representando uma janela temporal dos padrões de *candlestick*. O desempenho final da combinação dos modelos não apresentou

melhoras e concluiu-se que a janela de tamanho cinco representava todos os padrões. Com isso eliminou-se a possibilidade de usar mais de modelo para cada janela de padrões.

Outro expediente, extremamente trabalhoso, e que não apresentou bons resultados, foi a tentativa de gerar os padrões de *candlestick* artificialmente, sem a ajuda de um especialista. A idéia era formar as figuras básicas, segundo aquelas descritas na literatura, e incluir ruído às amostras [4], criando assim, uma base de dados que pudesse treinar o sistema. No entanto, não se conseguiu definir testes para verificar de forma eficiente, se o padrão havia sido aprendido pelo modelo.

Dissertou-se, até aqui, sobre a utilização de aprendizado semi-supervisionado e supervisionado, não apresentando nenhuma solução que utilizasse aprendizado não supervisionado. No entanto, tentou-se a construção de um modelo de previsão com base no agrupamento de classes, que é uma abordagem típica de um modelo não supervisionado. Essa abordagem também sem sucesso aparente, se deu através do algoritmo de k-médias [20]. Separaram-se os padrões de reversão com base na tendência da série, assim como foi realizado na abordagem apresentada nos capítulos anteriores. Além da janela temporal de cinco dias, acrescentou-se uma janela de também cinco dias, da tendência da série, com o objetivo de diferenciar os padrões que possuem o mesmo desenho em reversões de alta para baixa e baixa para alta. Conseguiu-se alguns resultados, no entanto, foram prematuramente abandonados. Mais à frente, ainda nesse capítulo, discutir-se-á sobre abordagens que não apresentam bons resultados aparentes e que podem, entretanto, gerar bons sistemas de investimento.

Embora a maioria das abordagens realizadas neste trabalho trate a questão de previsão como um problema de classificação, tentou-se, também, a utilização de Redes Neurais Artificiais para aplicar regressão nos valores da tendência. Caso um modelo fosse capaz de prever a mudança na tendência da série, já seria suficiente para obter lucros satisfatórios. No entanto, essa abordagem não apresentou bons resultados, mesmo quando tomava-se como parâmetro apenas a mudança de sinal da derivada da tendência. Ainda que tivesse obtido bons resultados, este modelo falharia, gravemente, quanto ao requisito de generalidade, necessitando de constante retreino.

Por fim, tentaram-se várias abordagens para operações *intraday*, entretanto, sem uma boa base de dados com todas as variações que ocorrem em um dia, tornou-se bastante difícil a construção de um sistema com esta característica. Nas tentativas realizadas, utilizaram-se vários indicadores técnicos, sendo que eles foram simplesmente incorporados ao modelo, na esperança de

acrescentar informações relevantes a ele [57].

## 5.2 Resultados secundários

O modelo de classificação mostrado neste trabalho trouxe alguns resultados secundários, como por exemplo, a aplicação construída para simular os investimentos. Ela possui base para ser transformada em um *framework* para testes de modelos de previsão de séries financeiras. O programa já está conectado a uma base de dados com séries de ações da BOVESPA de 1986 a 2007 e, na sua construção utilizaram-se padrões de projetos [17], que auxiliaram no desacoplamento entre as partes do sistema. Dessa forma, torna-se fácil a inclusão de uma nova estratégia de investimento ou mesmo um novo modelo de previsão, sem precisar modificar o restante do sistema: basta conectar as partes, e o *framework* se encarregará do seu funcionamento.

Algumas modificações ainda se fazem necessárias para que se possa classificá-lo como um *framework* para simulação de *trade systems*; no entanto, a base já está concluída e, no caso de um novo projeto na área de previsão financeira, o seu desenvolvimento seria simplificado.

A abordagem não-supervisionada utilizando o k-médias, citada anteriormente, não trouxe, aparentemente, bons resultados, pelo fato de o modelo ter sido, provavelmente, abandonado de forma prematura. Por se tratar de um problema de classificação de dificuldade elevada, todas as abordagens deveriam ser testadas em simulações de investimentos, antes de serem abandonadas. Os resultados apresentados inicialmente, em matrizes confusão, podem não ser animadores, mas os apresentados por investimentos relativos àquela matriz podem ser interessantes. Todas as abordagens descritas na seção anterior, poderiam ter apresentado bons resultados, mas, provavelmente, como não se tinha, ainda, um sistema capaz de testá-las de forma eficiente, algumas abordagens foram perdidas.

A construção de um sistema de previsão para séries financeiras exige grande disciplina do pesquisador, todos os resultados devem ser armazenados para posterior análise, assim como os códigos devem ser bem documentados e organizados [57]. O *JSVM<sup>litc</sup>* contribuiu muito para documentar e facilitar análises a respeito dos modelos treinados, através de sua interface. Apesar de ainda não poder ser considerado *user-friendly*, com algumas pequenas modificações, poderá se tornar em uma grande ferramenta. Com isso, todos os projetos que fazem uso da implementação do SVM desenvolvida por Joachims [24], deveriam, seriamente, considerar a utilização do *JSVM<sup>litc</sup>* como uma ferramenta para agilizar o desenvolvimento de modelos.

## 5.3 A abordagem utilizada

Zhu, [56], afirma que, possivelmente, seres humanos aprendem em um ambiente semi-supervisionado. Em várias situações acumulamos dados “sem rótulo” e depois fazemos a ligação, muitas vezes inconscientemente, destes dados com os que possuem rótulo, gerando conhecimento. Um exemplo que deve ser ressaltado, são de crianças com 17 meses de idade que conseguem relacionar melhor o nome com o objeto correspondente, se elas escutaram o nome do objeto várias vezes antes. Crianças que nunca ouviram a palavra, apresentaram mais dificuldade em entender o seu significado. Neste caso, as palavras escutadas pela criança seriam amostras sem rótulo que vão sendo armazenadas de alguma forma e, na hora de aprender o seu significado, os exemplos sem rótulo agilizam a aprendizagem da criança [13]. Deste ponto de vista, o tipo de aprendizagem utilizada neste trabalho aproxima-se da que ocorre com os seres humanos.

Todos os modelos de classificação que utilizam treinamento semi-supervisionado o fazem, geralmente, pela escassez de amostras rotuladas [56]. No caso do modelo construído nesse trabalho, as amostras eram abundantes, porém a qualidade era bastante duvidosa, devido aos padrões conflitantes de *candlestick*. Através da utilização de dados sem rótulo, melhoraram-se consideravelmente os ruídos do conjunto de treinamento o que, conseqüentemente, trouxe melhores rendimentos para o *trading system* que utilizou o modelo semi-supervisionado. Este é, portanto, um caso especial da utilização de dados sem rótulo no treinamento pois, buscou-se através deles, não uma melhor representatividade da distribuição das duas classes, mas sim uma maior qualidade do conjunto.

E, apesar da abordagem semi-supervisionada ter apresentado resultados superiores à supervisionada, é importante destacar que os desempenhos das duas são muito superiores a um sistema baseado em regras se-então que utilizou apenas os indicadores técnicos: o sistema baseado apenas no indicador estocástico não conseguiu obter lucro algum.

### 5.3.1 Utilização como um sistema tutor

Como já citado no início do texto, não é objetivo que o sistema de investimento, descrito neste trabalho, fosse utilizado como um sistema tutor, onde investidores inexperientes conseguiriam aprender os padrões de reversão e continuação ao utilizá-lo. No entanto, é possível aplicá-lo com esse fim, e padrões que ainda não foram catalogados podem agora ser observados pelos

usuários do sistema. Ao utilizar o simulador do Folha em Ação, já foi possível se familiarizar com alguns padrões e até mesmo antecipar algumas decisões do *trading system*. Acredita-se que, com um tempo de uso, caso se tenha atenção às decisões do sistema, seja possível aprender a operar com o sistema.

#### 5.3.2 *Treinamento independente*

Na construção do modelo, em momento algum, houve a participação de um especialista em Análise Técnica. É interessante notar que o sistema aprendeu de forma independente, aplicando, também, de forma autônoma. Houve sempre, como pano de fundo, os padrões de *candlestick*, no entanto, o que se passaram ao modelo foram apenas valores que representam a janela temporal do *candle* e, a partir disto, o modelo foi capaz de reconhecer os padrões que precediam uma reversão de tendência. É importante ressaltar que, na fase de testes, alguns padrões, catalogados na literatura, foram capturados pelo sistema, e devido ao seu ótimo desempenho, essa detecção, de padrões já registrados, pelo modelo, pôde ser vista como uma validação da Análise Técnica e dos padrões de reversão de *candlesticks*. A identificação, por exemplo, do Martelo, ou Homem Enforcado pelo modelo prova que os padrões são figuras que realmente representam reversão e que a teoria por trás delas, como uma medida da psicologia dos investidores, é realmente válida.

#### 5.3.3 *Desbalanceamento das classes*

Um outro ponto que não pode deixar de ser discutido refere-se ao desbalanceamento entre as classes tratadas neste problema. Como o número de amostras de reversão é muito menor do que amostras de continuação, é natural que se pense em utilizar alguma técnica disponível na literatura para balancear as classes, de forma a obter um melhor resultado do modelo de previsão. No entanto, apenas através do uso do parâmetro, que dá pesos ao valor de erro de cada amostra em treinamento do SVM, foi possível conseguir bons resultados de balanceamento. Desde o início, quando separaram-se seis modelos de compra com pesos diferentes para o erro de amostras positivas no treinamento, já esperava-se que o modelo com o menor peso apresentasse melhores resultados. O modelo 1, que apresenta o peso mais próximo da distribuição normal, ou seja, representa ainda um grande desbalanceamento no número de amostras das classes, gerou melhores investimentos. Apesar de sua matriz de confusão não apresentar o melhor desempenho global, em termos de retornos, apresentou, tanto na abordagem semi-supervisionada, quanto na supervisionada, melhores resultados. Isso mostra que, dependendo da aplica-

ção, a diferença no número de amostras no treinamento pode ser vantajoso. Esse tipo de situação ocorre quando um falso positivo possuiu alto custo. Outros exemplos seriam: sistema de detecção de *spam* e operações irregulares com cartão de crédito.

#### 5.3.4 A não necessidade de retreino

Talvez a maior desvantagem de se utilizar um sistema inteligente de investimento esteja na necessidade de, periodicamente, retreinar o modelo de previsão. A grande parte dos sistemas de previsão para séries temporais são treinados para captar relações entre séries correlacionadas. Geralmente faz-se um estudo estatístico para verificar quais seriam as melhores entradas para prever movimentos futuros de determinada série.

O grande problema desse tipo de abordagem está no fato de as relações entre as séries possuírem uma certa sazonalidade. Em outras palavras, de tempos em tempos, a relação entre as entradas e saídas desejadas do modelo é alterada, devido à própria dinâmica das séries. Isso faz com que esses sistemas tenham um prazo de validade que, infelizmente, é difícil de ser determinado. É possível fazer algumas previsões, com base em experiências anteriores da validade desses sistemas, a taxa de acerto desse período não é, no entanto, assegurada.

Justamente por não se saber, ao certo, por quanto tempo um sistema de previsão financeira irá durar, prejudica-se consideravelmente o seu desempenho. É normal que estes modelos apresentem erros de previsão, tanto é que, ao se desenvolver um sistema, deve-se levar em consideração estratégias que minimizem as perdas, como o *stop loss*, descrito no capítulo 2. A grande dificuldade está em determinar quando o modelo simplesmente errou, ou quando o modelo perdeu a sua performance original, pelo fato da dinâmica do mercado ter se alterado. Caso uma série de erros seja confundida com necessidade de retreinamento, o sistema pode estar sendo desperdiçado, fazendo-se com que todo o seu esforço de produção seja jogado fora. Por outro lado, caso se insista em utilizar um sistema que realmente precisa de retreinamento, todo o benefício que ele eventualmente trouxe pode vir a ser perdido.

A solução poderia ser, pensando no caso de empresas, a contratação de um profissional competente para a detecção da necessidade de retreinamento do sistema. No entanto, apesar de tornar possível a sua utilização, a disponibilização de um profissional para este fim pode ter um alto custo que inviabilizaria a utilização do modelo. É por esses fatos que, muitas vezes, a utilização de um modelo de previsão, principalmente da área financeira, torna-se inviável

economicamente.

Pensando na dificuldade do retreinamento, o principal foco do *trading system* desenvolvido neste trabalho foi que o seu modelo de previsão fosse genérico o suficiente para qualquer situação. Ainda não é possível assegurar a generalidade do modelo. Os testes realizados até o momento, entretanto, indicam que, no mínimo, a validade do modelo é superior a outros modelos apresentados na literatura. Pode-se afirmar isto, pois, após a utilização do modelo por 1 ano e 4 meses, ele ainda apresenta bons resultados. E acredita-se, devido ao próprio tempo de utilização dos padrões de *candlestick* e pelo intervalo temporal dos dados usados em treinamento, de 1986 a 2005, que o sistema provavelmente não necessitará de retreino.

## 5.4 Sugestões para trabalhos futuros

A grande diversidade de possibilidades que cerca o problema de previsão financeira e o torna tão excitante e desafiador, faz com que se possa imaginar inúmeras outras possibilidades de construção de modelos. Algumas delas serão sugeridas neste tópico.

### 5.4.1 Ainda sobre Análise Técnica

A adição de outros indicadores técnicos ao modelo deve ser criteriosa, pois em alguns casos, pode acontecer de o desempenho ser significativamente inferior. Seria muito simples, caso acrescentar novas variáveis de entrada, treinar, validar e testar fossem tarefas simples e rápidas de serem realizadas. Elas, no entanto, necessitam de muito tempo, principalmente de processamento de máquina. Por isso, acrescentar uma nova variável, torna-se mais seguro quando há alguma teoria sustentando a sua adição.

Devido à flexibilidade temporal dos *candles*, é natural cogitar a possibilidade de construção de trabalhos com investimentos em outro tipo de tendência, como a primária ou secundária. Possivelmente, o esquema desenvolvido neste trabalho se repetirá, necessitando apenas de troca dos dados utilizados no treinamento, de validação e de testes para *candles* com a configuração temporal desejada.

### 5.4.2 Combinando Análise Técnica e Fundamentalista

Uma grande possibilidade para desenvolvimento de sistemas futuros está ligada à combinação de várias estratégias. E, das várias combinações possíveis, talvez que venha a oferecer um melhor desempenho seja aquela que

combine análise técnica à fundamentalista, assim como vários investidores fazem no dia-a-dia.

Como discutido anteriormente, alguns autores acreditam que a Análise Fundamentalista seja um subconjunto da técnica [40], as notícias fundamentalistas podem levar um certo tempo para se agregarem à série, tornando-se, assim, possível de se detectarem no preço, as mudanças causadas por fatores políticos, econômicos ou outra circunstância qualquer. Caso o sistema seja capaz de antecipar mudanças fundamentais, antes mesmo que ela se agregue ao preço, pode-se tirar muita vantagem desta situação.

Mesmo que já tenha sido incorporado ao preço, alguns padrões podem ser difíceis de serem detectados. Talvez com uma classificação textual em um portal de notícias, torne-se mais fácil a detecção destes padrões. Da mesma forma que se sugeriu anteriormente, muita pesquisa ainda pode ser desenvolvida nesta área, que é um campo que está começando a ser explorado. Acredita-se que muitos trabalhos inovadores estão por vir, principalmente através de técnicas como as descritas no capítulo 2, envolvendo classificação de texto.

# Bibliografia

---

- [1] <http://www.hibernate.org/>, 2007.
- [2] Jonathan Batten and Craig Ellis. Technical trading system performance in the australian share market: Some empirical evidence. *Asia Pacific Journal of Management*, 13(1):87–99, 1996.
- [3] Kristin Bennett and Ayhan Demiriz. Semi-supervised support vector machines. In Michael J. Kearns, Sara A. Solla, and David A. Cohn, editors, *NIPS*, pages 368–374. The MIT Press, 1998.
- [4] Chris M. Bishop. Training with noise is equivalent to Tikhonov regularization. *Neural Computation*, 7(1):108–116, 1995.
- [5] Blum and Mitchell. Combining labeled and unlabeled data with co-training. In *COLT: Proceedings of the Workshop on Computational Learning Theory*, Morgan Kaufmann Publishers, 1998.
- [6] Bernhard E. Boser, Isabelle Guyon, and Vladimir Vapnik. A training algorithm for optimal margin classifiers. In *COLT*, pages 144–152, 1992.
- [7] BOVESPA, março 2007. <http://www.bovespa.com.br>.
- [8] A. P. Braga, T. B. Lurdemir, and A. C. P. L. F. Carvalho. *Redes Neurais Artificiais: teorias e aplicações*. LTC, 2000.
- [9] Bernardo Carvalho. Novas estratégias para detecção automática de vetores de suporte em least squares support vector machines. Master's thesis, CPDEE - UFMG, 2005.
- [10] T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*. MIT Press, 1990.

- 
- [11] Cristiano Leite de Castro. Sistema inteligente para previsão on-line de Índice ibovespa baseado em redes neurais artificiais. Master's thesis, CPDEE - UFMG, 2003.
- [12] Eber Machado Duarte. Construção de um agente inteligente baseado em redes neurais artificiais para coleta e classificação de informações disponíveis na internet. Master's thesis, CPDEE - UFMG, 2001.
- [13] K. Graf Estes, J. L. Evans, M. W. Alibali, and J. R. Saffran. Can infants map meaning to newly segmented words? statistical segmentation and word learning. *Psychological Science*, 2006.
- [14] Eugene Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417, May 1970. available at <http://ideas.repec.org/a/bla/jfinan/v25y1970i2p383-417.html>.
- [15] T. Fawcett. Roc graphs: Notes and practical considerations for researchers, 2004.
- [16] J. H. Friedman. On bias, variance, 0/1 - loss, and the curse-of-dimensionality. *Data mining and Knowledge Discovery*, 1(1):54–77, 1996.
- [17] Erich Gamma. *Design patterns: elements of reusable object-oriented software*. Addison-Wesley professional computing series. Addison-Wesley, pub-AW:adr, 1995. See book review [?].
- [18] Stuart Geman, Elie Bienenstock, and Ren Doursat. Neural networks and the bias/variance dilemma. *Neural Comput.*, 4(1):1–58, January 1992.
- [19] David E. Goldberg. *Genetic algorithms in search, optimization, and machine learning*. Reading, Mass. : Addison-Wesley Pub. Co., 1989, 412p. CALL NUMBER: QA402.5 .G635 1989, 1989.
- [20] J. A. Hartigan and M. A. Wong. Algorithm AS136. A  $K$ -means clustering algorithm. *Applied Statistics*, 28:100–108, 1979.
- [21] Simon Haykin. *Neural networks: a comprehensive foundation*. Prentice Hall, 1994.
- [22] R. J. Hodrick and E. C. Prescott. Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29, 1997.
- [23] Kyoungjae Kim. Artificial neural networks with evolutionary instance selection for financial forecasting. *Expert Syst. Appl.*, 30(3):519–526, 2006.

- 
- [24] T. Joachims. Making large-scale support vector machine learning practical. In A. Smola B. Schölkopf, C. Burges, editor, *Advances in Kernel Methods: Support Vector Machines*. MIT Press, Cambridge, MA, 1998.
- [25] T. Joachims. Transductive inference for text classification using support vector machines. In *Proc. International Conference on Machine Learning (ICML)*, 1999.
- [26] Cleber G. Junior, Edson O. Pamplona, and José Arnaldo Barra Montevechi. Seleção de carteiras através do modelo de markowitz para pequenos investidores (com o uso de planilhas eletrônicas). In *IX Simpep*, 2002.
- [27] K. Kamijo and T. Tanigawa. Stock price pattern recognition: A recurrent neural network approach. In *IEEE International Joint Conference on Neural Networks (4th IJCNN'90)*, volume I, pages I-215–I-222, San Diego, 1990. IEEE. NEC.
- [28] C. L. Lee, A. Liu, and W. Chen. Pattern discovery of fuzzy time series for financial prediction knowledge and data engineering. *IEEE Transactions*, 18:613–625, 2006.
- [29] Chiung-Hon Leon Lee, WenSung Chen, and Alan Liu. An implementation of knowledge based pattern recognition for financial prediction. *Cybernetics and Intelligent Systems - IEEE Conference*, 1:218–223, 2004.
- [30] Chiung-Hon Leon Lee, WenSung Chen, and Alan Liu. Candlestick tutor: An intelligent tool for investment knowledge learning and sharing. *icalt*, 00:238–240, 2005.
- [31] K.H. Lee and G.S. Jo. Expert system for predicting stock market timing using a candlestick chart. *Expert Systems with Applications*, (16):357–364, 1999.
- [32] A. Lo and A. C. Mackinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *Rev. Financ. Stud.*, 1:41–66, 1988.
- [33] Ana Carolina Lorena. Introdução às máquinas de vetores suporte. Technical report, ICMC, 2003.
- [34] Burton G. Malkiel. *Random Walk Down Wall Street: Including a Life-Cycle Guide to Personal Investing*. W W Norton & Co Inc, May 1990.

- 
- [35] Burton G. Malkiel. The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1):59–82, Winter 2003. available at <http://ideas.repec.org/a/aea/jecper/v17y2003i1p59-82.html>.
- [36] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.
- [37] Eduardo Matsura. *Comprar ou Vender? Como investir na bolsa utilizando Análise Técnica*. Editora Saraiva, 2 edition, 2006.
- [38] J. J. Moré. The Levenberg-Marquardt algorithm: Implementation and theory. In G.A. Watson, *Lecture Notes in Mathematics* 630, pages 105–116. Springer-Verlag, Berlin, 1978.
- [39] Muller, Mika, Ratsch, Tsuda, and Scholkopf. An introduction to kernel-based learning algorithms. *IEEE TNN: IEEE Transactions on Neural Networks*, 12, 2001.
- [40] J. Murphy. *Technical Analysis of the Financial Markets*. New York Institute of Finance, 1999.
- [41] Steve Nison. *Japanese Candlestick Charting Techniques*. New York Institute of Finance, 1991.
- [42] A. Refenes, A. Burgess, and Y. Bentz. Neural networks in financial engineering: a study in methodology. *Neural Networks, IEEE Transactions on*, 8(6):1222–1267, 1997.
- [43] Young-Woo Seo, Joseph Andrew Giampapa, and Katia Sycara. Text classification for intelligent portfolio management. Technical Report CMU-RI-TR-02-14, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, May 2002.
- [44] Amanda Sharkey. *Multi-Net Systems*, chapter Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems, pages 1–30. Springer-Verlag, 1999.
- [45] A. Smola, P. Bartlett, B. Scholkopf, and D. Schuurmans. Advances in large margin classifiers, 2000.
- [46] Teixeira, Braga, Takahashi, and Saldanha. Improving generalization of mlps with multi-objective optimization. *Neurocomputing*, 35(1-4):189–194, 2000.

- 
- [47] James D. Thomas and Katia Sycara. Integrating genetic algorithms and text learning for financial prediction. In Alex A. Freitas, William Hart, Natalio Krasnogor, and Jim Smith, editors, *Data Mining with Evolutionary Algorithms*, pages 72–75, Las Vegas, Nevada, USA, 8 2000.
- [48] Konstantin Tretyakov. Machine learning techniques in spam filtering, 1995.
- [49] Ray Tsaih, Yenshan Hsu, and Charles C. Lai. Forecasting s&p 500 stock index futures with a hybrid ai system. *Decis. Support Syst.*, 23(2):161–174, 1998.
- [50] V. Vapnik. *Statistical Learning Theory*. Wiley, 1998.
- [51] V. N. Vapnik. An overview of statistical learning theory. *Neural Networks, IEEE Transactions on*, 10(5):988–999, 1999.
- [52] Vladimir Vapnik, Esther Levin, and Yann Le Cun. Measuring the VC-dimension of a learning machine. *Neural Computation*, 6(5):851–876, 1994.
- [53] Vladimir N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.
- [54] J. T. Yao, C. L. Tan, and H.-L. Poh. Neural networks for technical analysis: A study on klc1. *International Journal of Theoretical and Applied Finance*, 2(2):221–241, 1999.
- [55] L. A. Zadeh. Knowledge representation in fuzzy logic. In R. R. Yager, editor, *An Introduction to Fuzzy Logic Applications in Intelligent Systems*, pages 1–25. Kluwer, Boston, 1992.
- [56] Xiaojin Zhu. Semi-supervised learning literature survey. Technical Report 1530, Computer Sciences, University of Wisconsin-Madison, 2005. [http://www.cs.wisc.edu/~jerryzhu/pub/ssl\\_survey.pdf](http://www.cs.wisc.edu/~jerryzhu/pub/ssl_survey.pdf).
- [57] Joseph S. Zirilli. *Financial Prediction Using Neural Networks*. International Thomson Computer Press, Boston, MA, USA, 1996.

## Gráficos de investimentos em 2006

---

---

### **Formato das legendas:**

- <Operação><Número de ações><Nome da ação> em: <Data: AAAAMMDD> por: <Preço>
- Tentou comprar: utilizado quando um sinal de compra foi lançado, mas o capital já estava aplicado.

A seta para cima indica momento de entrada no mercado e a para baixo momento de saída.



Figura A.1: Comprou 1459 ações SDIA4 em: 20060116 por: 6.85. Vendeu 1459 ações SDIA4 em: 20060202 por: 7.5



Figura A.2: Comprou 232 ações SUBA3 em: 20060208 por: 47.0. Vendeu 232 ações SUBA3 em: 20060220 por: 58.49



Figura A.3: Tentou aplicar em ELET6 em: 20060209 mas não tinha dinheiro



Figura A.4: Tentou aplicar em GOLL4 em: 20060209 mas não tinha dinheiro



Figura A.5: Comprou 252 ações SUBA3 em: 20060301 por: 53.75. Vendeu 252 ações SUBA3 em: 20060307 por: 49.25



Figura A.6: Comprou 2162 ações SDIA4 em: 20060309 por: 5.75. Vendeu 2162 ações SDIA4 em: 20060321 por: 5.75



Figura A.7: Comprou 2182 ações SDIA4 em: 20060324 por: 5.69. Vendeu 2182 ações SDIA4 em: 20060406 por: 6.38



Figura A.8: Comprou 2377 ações SDIA4 em: 20060502 por: 5.85. Vendeu 2377 ações SDIA4 em: 20060509 por: 6.15



Figura A.9: Comprou 179 ações LAME4 em: 20060516 por: 81.45. Vendeu 179 ações LAME4 em: 20060529 por: 83.5



Figura A.10: Comprou 2 ações SDIA4 em: 20060516 por: 5.63. Vendeu 2 ações SDIA4 em: 20060602 por: 5.97



Figura A.11: Comprou 373 ações ELET6 em: 20060616 por: 40.01. Vendeu 373 ações ELET6 em: 20060704 por: 45.2



Figura A.12: Tentou aplicar em LAME4 em: 20060616 mas não tinha dinheiro



Figura A.13: Tentou aplicar em ELET6 em: 20060619 mas não tinha dinheiro



Figura A.14: Tentou aplicar em ELET6 em: 20060621 mas não tinha dinheiro



Figura A.15: Comprou 1 ação SDIA4 em: 20060622 por: 5.22. Vendeu 1 ação SDIA4 em: 20060703 por: 5.71



Figura A.16: Comprou 3132 ações SDIA4 em: 20060710 por: 5.38. Vendeu 3132 ações SDIA4 em: 20060719 por: 6.1



Figura A.17: Comprou 434 ações ELET6 em: 20060824 por: 43.95. Vendeu 434 ações ELET6 em: 20060911 por: 41.4

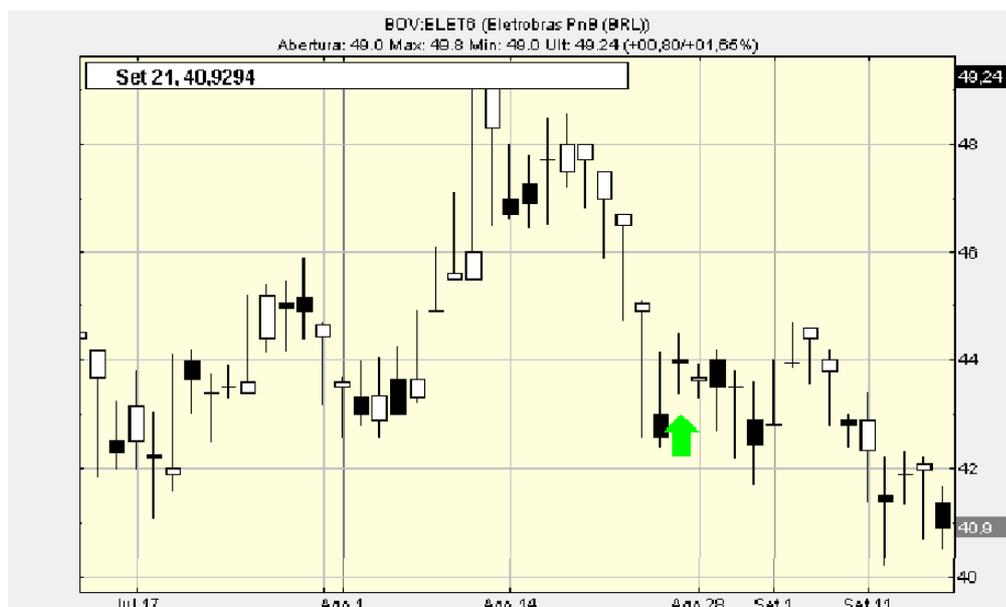


Figura A.18: Tentou aplicar em ELET6 em: 20060825 mas não tinha dinheiro



Figura A.19: Tentou aplicar em SDIA4 em: 20060825 mas não tinha dinheiro



Figura A.20: Tentou aplicar em ELET6 em: 20060828 mas não tinha dinheiro



Figura A.21: Comprou 450 ações ELET6 em: 20060922 por: 39.88. Vendeu 450 ações ELET6 em: 20061005 por: 48.0



Figura A.22: Tentou aplicar em LAME4 em: 20060922 mas não tinha dinheiro



Figura A.23: Comprou 1 ação SDIA4 em: 20060925 por: 5.71. Vendeu 1 ação SDIA4 em: 20061003 por: 6.06



Figura A.24: Comprou 537 ações SUBA3 em: 20061009 por: 40.2. Vendeu 537 ações SUBA3 em: 20061108 por: 47.75



Figura A.25: Comprou 4175 ações SDIA4 em: 20061113 por: 6.14. Vendeu 4175 ações SDIA4 em: 20061206 por: 7.02



Figura A.26: Comprou 450 SUBA3 em: 20061212 por: 65.0



Figura A.27: Tentou aplicar em GOLL4 em: 20061226 mas não tinha dinheiro

---

## Gráficos de investimentos em 2007

---

**Formato das legendas:**

- <Operação><Número de ações><Nome da ação> em: <Data: AAAAMMDD>  
por: <Preço>
- Tentou comprar: utilizado quando um sinal de compra foi lançado, mas o capital já estava aplicado.

A seta para cima indica momento de entrada no mercado e a para baixo momento de saída.



Figura B.1: Comprou 154 ações SUBA3 em: 20070122 por: 64.6. Vendeu 154 ações SUBA3 em: 20070206 por: 69.9



Figura B.2: Comprou 1031 ações ITSA4 em: 20070316 por: 10.47. Vendeu 1031 ações ITSA4 em: 20070323 por: 11.8



Figura B.3: Comprou 1063 ações ITSA4 em: 20070503 por: 11.43. Vendeu 1063 ações ITSA4 em: 20070521 por: 12.42



Figura B.4: Tentou aplicar em SUBA3 em: 20070518 mas não tinha dinheiro