

CENTRO UNIVERSITÁRIO FEEVALE

FERNANDO RAFAEL STAHNKE

USO DE *DATA MINING* PARA IDENTIFICAÇÃO DE PADRÕES NO MERCADO DE
CAPITAIS

Anteprojeto de Trabalho de Conclusão

Novo Hamburgo, maio de 2008.

FERNANDO RAFAEL STAHNKE

fstahnke@gmail.com

**USO DE DATA MINING PARA IDENTIFICAÇÃO DE PADRÕES NO MERCADO
DE CAPITAIS**

Centro Universitário Feevale
Instituto de Ciências Exatas e Tecnológicas
Curso de Ciência da Computação
Anteprojeto de Trabalho de Conclusão

Professor orientador: Juliano Varella de Carvalho

Novo Hamburgo, maio de 2008.

RESUMO

Atualmente na economia mundial existem diversas formas de captação de recursos utilizadas pelas empresas, e neste cenário, destaca-se a abertura de capital na qual as empresas lançam ações ao público, para serem negociadas nas bolsas de valores. O uso de *Data Mining* neste contexto, mais especificamente no mercado à vista, visa definir uma nova forma de efetuar a chamada “Análise Técnica”, que atualmente estuda os gráficos formados pelos preços das ações disponíveis no mercado. A aplicação de técnicas de mineração de dados baseia-se no fato de que estes gráficos são construídos exclusivamente a partir dos preços das ações num período de tempo. Logo, pretende-se com a aplicação destas técnicas a identificação de padrões de comportamentos, e assim, determinar a tendência futura e a frequência de ocorrência.

Palavras-chave: *Ação*. Mercado à vista. Mineração de Dados. Descoberta de Conhecimento. Padrões.

SUMÁRIO

MOTIVAÇÃO	5
OBJETIVOS	9
METODOLOGIA.....	10
CRONOGRAMA.....	11
BIBLIOGRAFIA.....	12

MOTIVAÇÃO

Enriquecer com investimento em bolsas de valores é o sonho de muitos investidores, mas comprar e vender ações no momento certo requer cautela e informações confiáveis. Hoje, devido a facilidade do acesso a estas instituições através de sistemas como o *home-broker*, que permite a negociação de ações pela internet, a participação de investidores considerados pessoa física corresponderam por 25% do giro financeiro total da Bovespa (Bolsa de Valores de São Paulo) no mês de fevereiro, do ano corrente, com uma movimentação total de R\$ 116,58 bilhões de reais (FOLHA ONLINE, 2008).

As operações de compra e venda de ações nas bolsas de valores de todo o mundo oscilam de acordo com o objetivo de seus investidores, perspectivas e conhecimento para a tomada de decisão, onde as metas são diversas, ou seja, a curto, médio ou longo prazo (SANTOS, 2007). Inclusive estas decisões interferem nas relações e nos dados da base a ser analisada.

No Brasil, a Bovespa concentra todas as negociações de ações do país por um preço, que é estabelecido através da lei de oferta e demanda (ELDER, 1993). Ações são títulos de renda variável emitidos por Sociedades Anônimas que representam a menor fração do capital social da empresa, tornando o possuidor deste título dono de uma fração do capital social da empresa que o emitiu. Pode ser de dois tipos: **ordinárias**, que proporcionam ao acionista direito de voto em assembleias e participação nos resultados das empresas ou **preferenciais**, onde o acionista tem a prioridade no recebimento de dividendos e no reembolso de capital, sem direito a voto (BOVESPA, 2007).

Atualmente, na Bovespa, existem os mercados a termo, à vista e de opções para as negociações de ações, sendo o mercado à vista foco deste trabalho. Este mercado caracteriza-se pela entrega dos títulos vendidos no 2º dia útil após a realização do negócio (liquidação física) e pela liquidação financeira (pagamento e recebimento) no 3º dia útil após a negociação dos títulos (BOVESPA, 2007). Neste mercado existem duas linhas de análises, sendo elas as análises técnica e fundamentalista, variando em fundamentação e em termos práticos, no período de tempo de investimento, a curto ou longo prazo.

A análise fundamentalista baseia-se nos fatores econômicos e fundamentais das empresas, como ramo de atuação, perspectivas de mercado, imagem, além das análises dos demonstrativos financeiros e de relatórios de administração emitidos por elas, sendo utilizada em aplicações a longo prazo.

Dentre os tipos de análises existentes, a análise técnica é a mais utilizada pelos chamados investidores pessoa física, atraídos por sua aparente simplicidade, pois é basicamente a análise do gráfico histórico de preços da ação de uma empresa. Neste tipo de gráfico são desconsideradas todas as outras informações relevantes que influenciam no preço da ação, dispensando a necessidade de avaliação de outras informações (MATSURA, 2007). Este tipo de análise permite ao investidor buscar a ação que tem maior potencial de ganhos a curto e médio prazo.

Conforme Deschatre (1997), o mercado de capitais possui fases com ciclos de movimentos, dependentes uns dos outros. Cada fase do mercado possui seu próprio ciclo de acumulação, determinação, distribuição e nova acumulação. Acumulação é a parte do ciclo onde os investidores adquirem títulos. Determinação é o intervalo entre a aquisição e a venda dos títulos que acontece na chamada distribuição do ciclo. Após o período de distribuição, ocorre novamente a nova fase de acumulação, fechando o ciclo.

Como a análise técnica é baseada em indicadores estatísticos, a maior parte dos softwares encontrados no mercado já apresenta a maioria dos indicadores conhecidos (DESCHATRE, 2005). Segundo Matsura (2007), a crescente complexidade dos instrumentos de negociação do mercado financeiro, assim como o acesso a novas tecnologias de processamento da informação, estimulam o desenvolvimento de novos sistemas de análise e operação, inclusive com o uso de técnicas de Inteligência Artificial.

O uso de técnicas de mineração de dados e inteligência artificial no mercado financeiro para determinação de estratégias e previsão das tendências de alta e baixa já é realidade desde 1986 nos Estados Unidos. Primeiramente, através da aplicação de sistemas baseados em regras heurísticas, com o uso de sistemas especialistas. Após com o uso de

análise estatística e, finalmente com o uso de redes neurais artificiais. Todas as opções com suas respectivas vantagens e desvantagens (CARVALHO, 2005).

Segundo Bressan (2004) o uso de redes neurais é uma técnica de reconhecimento de padrões constantemente abordada na exploração de capitais. Em pesquisa, NTUNGO e BOYD (Apud BRESSAN, 2004) verificaram o desempenho de redes neurais e outro modelo conhecido como ARIMA em estratégias de negociação nos mercados futuros de milho e prata, com uso de dados semanais. A conclusão dos autores é que ambos os resultados são equivalentes em termos de retornos financeiros para os dois modelos de maneira positiva para estes mercados.

Diversos trabalhos nesta área foram desenvolvidos, alguns inclusive com resultados expressivos, como por exemplo, o estudo desenvolvido por BERGENSON e WUNSCH (Apud ZANETI; ALMEIDA, 1998), onde com o uso de um sistema composto por redes neurais híbridas, obtiveram um retorno de 660% em 25 meses com uma aplicação de US\$ 10.000 entre janeiro de 1989 e janeiro de 1991. Um dos aspectos negativos deste método, segundo os autores, foi o grande número de horas gasto para a construção do modelo. Estes fatos ressaltam a existência de um vasto campo de estudo acerca da comparação do desempenho de modelos preditivos para previsão de ativos em bolsas de valores.

Conforme Carvalho (2005), Mineração de Dados ou *Data Mining* (DM) é um conjunto de técnicas reunidas da Estatística e da Inteligência Artificial com objetivo de descobrir conhecimento novo que esteja “escondido” em grandes massas de dados. Já Harrison (1998) define DM como um processo de exploração e análise de uma grande massa de dados que através de meios automáticos ou semi-automáticos, objetivam a descoberta de padrões e regras significativas.

Segundo Carvalho (2005), a tarefa de mineração de dados é uma etapa do processo de descoberta de conhecimento, composto por cinco fases distintas, sendo elas: identificação do problema; descoberta do conhecimento; análise das relações descobertas; uso das relações descobertas e avaliação dos resultados.

A identificação do problema é a primeira fase do processo de DM, onde o problema é identificado e os objetivos a serem alcançados são definidos. Após, os dados passam pela etapa de descoberta de conhecimento, onde são definidos os dados úteis à extração de conhecimento, validados e tratados, tornando os dados pré-processados compatíveis com as entradas dos algoritmos de mineração de dados. (REZENDE, 2005; CARVALHO, 2000). Nesta etapa são definidos o algoritmo (técnica) e a tarefa a serem utilizadas para a exploração

eficiente da massa de dados. Identificadas as relações, é efetuada a análise das relações descobertas quanto a sua relevância para o escopo da aplicação, sendo esta uma tarefa não automatizada. (NAVEGA 2007).). O uso das relações descobertas constitui a tarefa de definição da forma de utilização destas. A avaliação dos resultados verifica se os objetivos propostos na identificação do problema foram alcançados, e desta forma, solucionando o problema.

As tarefas de mineração de dados a serem utilizadas e o conjunto de ações a serem monitoradas serão definidos durante a pesquisa bibliográfica, identificando as necessidades para que o processo possa ser executado de maneira coerente com a realidade. Da mesma forma, a técnica de mineração de dados também será definida posteriormente, dentre as técnicas existentes conforme citado por Goldschmidt (2005) e Carvalho (2005).

Sendo assim, a motivação deste trabalho consiste em explorar a base de dados do mercado à vista da Bovespa, baseando-se nos fundamentos da análise técnica, minimizando riscos e elevando a lucratividade de investidores, através do uso de técnicas de mineração de dados, do uso de ferramentas gratuitas, e da identificando padrões, levando-se em consideração aspectos variáveis como preço de abertura e fechamento, médias de variação por período, carteira de ações, valor financeiro, etc.

Seguindo o fundamento desta análise, todas as interferências de mercado e ambiente estão já agregadas ao valor do preço da ação, fatores como interferências políticas, macro-econômicas e análises de perspectivas de mercado, balanços financeiros, imagem e ramo de atuação das empresas utilizadas para investimento não são analisados. Entretanto, com base em trabalhos precedentes, algumas variáveis podem ser utilizadas para estudo, visando abranger as interferências no mercado, como por exemplo, a utilização do índice Dow Jones (No Brasil, o Ibovespa), CDI, dólar comercial, entre outros. (ZANETI; ALMEIDA, 1998)

Por fim, os resultados obtidos serão testados num ambiente *on-line* que simula o comportamento do mercado acionário brasileiro, validando assim as regras, os padrões e a metodologia empregada.

OBJETIVOS

Objetivo geral

Aplicar técnicas de *Data Mining* em uma base de dados históricos obtidos na Bolsa de Valores de São Paulo (BOVESPA) a fim de identificar padrões de comportamento e tendências nas oscilações do mercado a vista, e auxiliar os investidores que utilizam análise técnica na tomada de decisões com este conhecimento.

Objetivos específicos

- Descrever conceitos de Mineração de Dados buscando a identificação de padrões na base de dados definida.
- Definir e modelar a estrutura da base de dados;
- Executar o pré-processamento dos dados armazenados;
- Estudar uma ferramenta que será utilizada para a mineração de dados;
- Interpretar o conhecimento descoberto através da mineração de dados;
- Demonstrar que a análise técnica pode ser auxiliada por outra forma que não seja a visualização de gráficos, utilizada atualmente;
- Determinar a tendência dos preços de ações, indicando ao investidor a probabilidade desta tendência.
- Apresentar um relatório final descrevendo os resultados obtidos.

METODOLOGIA

A metodologia a ser utilizada para conclusão deste trabalho está organizada da seguinte forma:

1. Pesquisa bibliográfica em livros, artigos, revistas, trabalhos de conclusão e páginas da internet para o desenvolvimento da fundamentação teórica;
2. Pesquisa de softwares existentes para Análise Técnica;
3. Elaboração do anteprojeto e definir o escopo do trabalho com base na pesquisa realizada e auxílio do professor orientador;
4. Análise e estudo de uma Ferramenta para mineração de dados, como trabalham os algoritmos, testes iniciais e verificação de resultados;
5. Estudo da base de dados escolhida para a mineração de dados;
6. Definição e modelagem da estrutura escolhida para o tratamento de dados;
7. Redação do Trabalho de Conclusão I;
8. Revisão e entrega do Trabalho de Conclusão I;
9. Utilização de uma ferramenta para mineração de dados;
10. Relatar descobertas encontradas no processo de mineração de dados;
11. Conclusão. Desenvolvimento do relatório apresentando os resultados;
12. Redação do Trabalho de Conclusão II;
13. Revisão e entrega do Trabalho de Conclusão II;
14. Banca. Defesa da monografia perante a banca avaliadora.

CRONOGRAMA

Trabalho de Conclusão I

Etapa	Meses				
	Mar/08	Abr/08	Mai/08	Jun/08	Jul/08
1					
2					
3					
4					
5					
6					
7					
8					

Trabalho de Conclusão II

Etapa	Meses				
	Ago/08	Set/08	Out/08	Nov/08	Dez/08
9					
10					
11					
12					
13					
14					

BIBLIOGRAFIA

- AURÉLIO, M; VELLASCO, M; LOPES, C.H. **Descoberta de Conhecimento e Mineração de Dados**. [on line],1999. Disponível em < www.ica.ele.puc-rio.br/cursos/download/DM-apostila1.pdf>. Acesso em 20 Mar2008.
- CARVALHO, Juliano V. de. **Reconhecimento de Caracteres Manuscritos Utilizando Regras de Associação**. Campina Grande: 2000, Tese (Dissertação de Mestrado) - Centro de Ciências e Tecnologia,UNIVERSIDADE FEDERAL DA PARAIBA,2000.
- CARVALHO, Luis Alfredo Vidal de. **Data Mining : A Mineração de Dados no Marketing , Medicina, Economia, Engenharia e Administração**. Rio de Janeiro: Ciência Moderna, 2005.
- DESCHATRE, Gil Ari. **Ganhe nas bolsas com o seu micro**. Rio de Janeiro: Ciência Moderna, 1997.
- ELDER, Alexander.**Como se transformar em um operador e investidor de sucesso**. Rio de Janeiro:Campus,2005.
- GOLDSCHMIDT,Ronaldo;PASSOS,Emmanuel. **Data Mining: Um Guia Prático**. Rio de Janeiro:Campus,2004.
- MATSURA,Eduardo. **Comprar ou Vender?**.Rio de Janeiro: Saraiva, 2007.
- NAVEGA, Sergio. **Princípios Essenciais do Data Mining**. [online].2002. Disponível em <<http://www.intelliwise.com/reports/i2002.htm>>. Acesso 26.Mar.2008.
- PAULOS, John Allen. **A lógica do mercado de ações**. Rio de Janeiro:Campus,2004.
- PRODANOV, Cleber C. **Manual de Metodologia Científica**. 3. ed. Novo Hamburgo: Editora Feevale. 2006.

SANTOS, Tiago. **Simulação Multiagentes Aplicado ao Mercado de Capitais**. Novo Hamburgo:2007, Instituto de Ciências Exatas e Tecnológicas, CENTRO UNIVERSITÁRIO FEEVALE, 2007.

ZANETI, L.A.; ALMEIDA, F.C. de. **Exploração do uso de redes neurais na previsão do comportamento de ativos financeiros**. Terceiro Seminário em Administração. [online]. Out.1998, Disponível em : <
<http://www.fundacaofia.com.br/proinfo/artigos/artigotelebras.pdf>>. Acesso em: 26.Mar.2008.