

KDD E MINERAÇÃO DE DADOS:

Introdução e Motivação

Prof. Ronaldo R. Goldschmidt

ronaldo@de9.ime.eb.br

rribeiro@univercidade.br

geocities.yahoo.com.br/ronaldo_goldschmidt

INTRODUÇÃO E MOTIVAÇÃO

Fatos: Avanços em TI e o crescimento exponencial de BDs

- Internet
- SGBDs
- Dispositivos de memória (maior capacidade e menor custo)
- Leitoras de códigos de barras
- Sistemas de Informação em geral
- Etc...

INTRODUÇÃO E MOTIVAÇÃO

Exemplos:

- FedEx
- UPS
- Wal-Mart
- NASA
- Projeto Genoma
- Caixa Econômica
- Banco do Brasil
- Dentre muitos outros ...

INTRODUÇÃO E MOTIVAÇÃO

Posicionamento:

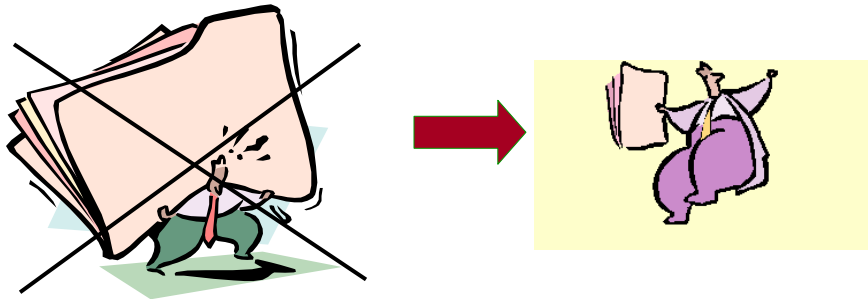
Hoje, a maioria das organizações produz **mais informações** em uma semana **do que muitas pessoas** poderiam ler em toda vida.

“Mais dados implicando em menos informações.”

INTRODUÇÃO E MOTIVAÇÃO

Necessidade:

Ferramentas **inteligentes** que auxiliem na **análise de dados** e na **busca por conhecimentos** em **GRANDES** bases de dados.



INTRODUÇÃO E MOTIVAÇÃO

Avanços em TI



Crescimento Exponencial de BDs



Necessidade de Ferramentas para Análise Grandes BDs

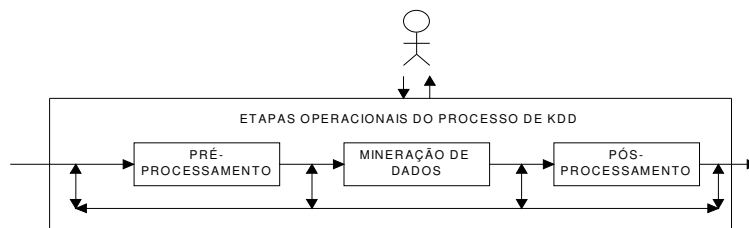


Área da Descoberta do Conhecimento em Bases de Dados (KDD)

INTRODUÇÃO E MOTIVAÇÃO

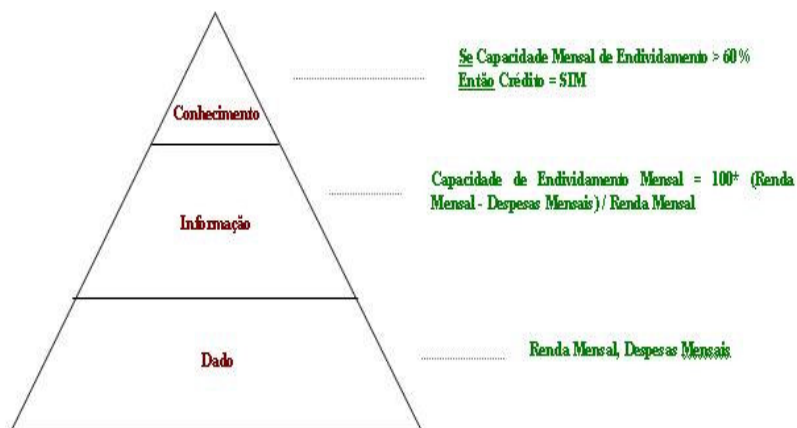
- **KDD – Knowledge Discovery in Databases**

“É um **processo**, de várias etapas, não trivial, **interativo e iterativo**, para **identificação** de **padrões compreensíveis, válidos, novos** e potencialmente **úteis** a partir de bases de dados.” [Fayyad et al., 1996]



INTRODUÇÃO E MOTIVAÇÃO

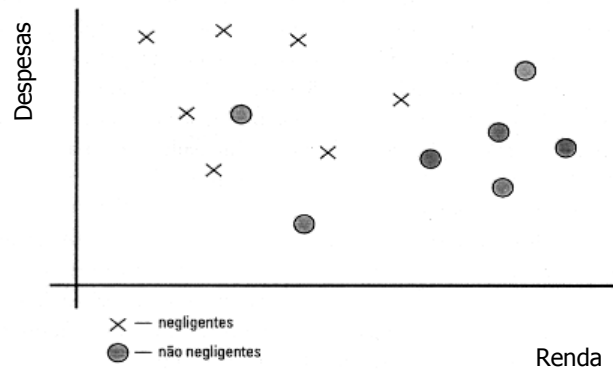
Hierarquia Dado - Informação - Conhecimento:



INTRODUÇÃO E MOTIVAÇÃO

Exemplo de aplicação na área de concessão de crédito:

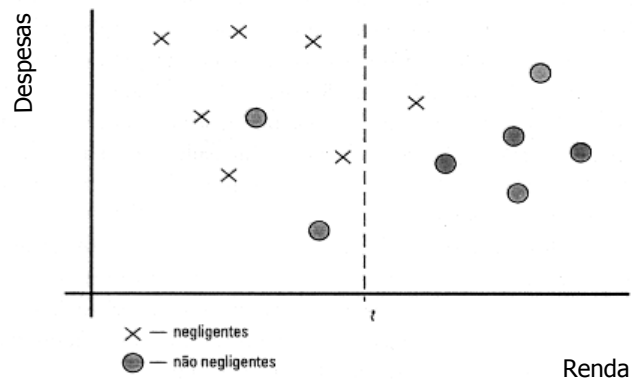
Conjunto de dados (Fatos)



INTRODUÇÃO E MOTIVAÇÃO

Exemplo de aplicação na área de concessão de crédito:

Padrão: Se renda > R\$ t Então Crédito = SIM (Cto)



INTRODUÇÃO E MOTIVAÇÃO

Outros Exemplos de Áreas de Aplicação:



Comércio



Energia



Finanças



Telecomunicações



Medicina



Meio-Ambiente



Educação



Indústria

INTRODUÇÃO E MOTIVAÇÃO

- **Comércio / Marketing**

Perfil do Consumidor (Marketing Direto), Promoção de Produtos, Segmentação de Mercado, etc;...

- **Finanças**

Análise de Investimentos, Análise de Crédito, Detecção de Fraudes em compras de Cartão de Crédito, etc;...

- **Medicina**

Diagnóstico e Prevenção de Doenças, Detecção de Fraudes em Planos de Saúde, etc;...

INTRODUÇÃO E MOTIVAÇÃO

- **Educação**

Análise de Matrículas e Demandas por Escolas, Evasão Escolar;...

- **Energia**

Previsão de Demanda, Distribuição de Recursos;...

- **Telecomunicações**

Detecção de falhas, Dimensionamento de Sistemas de Comunicação, Detecção de Fraudes;...

INTRODUÇÃO E MOTIVAÇÃO

- **Meio Ambiente**

Monitoramento ambiental, Prevenção de desequilíbrios ecológicos;...

- **Indústria**

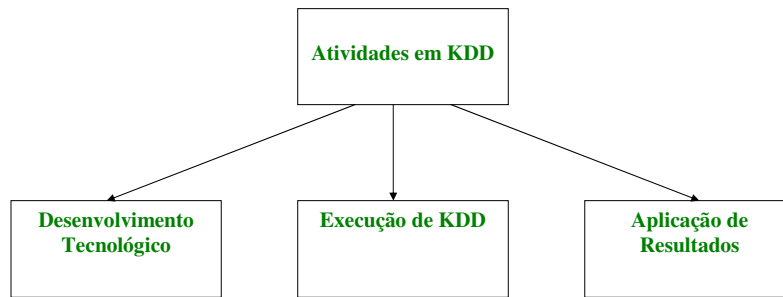
Previsão de demanda, Planejamento da produção e distribuição;...

- **Área Social**

Caracterização de Perfil para Reintegração Social;...

INTRODUÇÃO E MOTIVAÇÃO

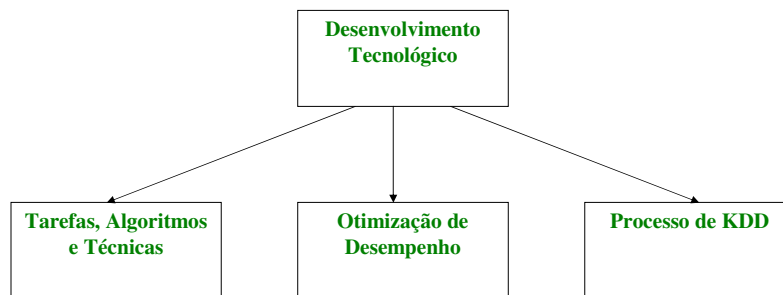
Atividades em KDD - uma Taxonomia



[Goldschmidt et al., 2002a]

INTRODUÇÃO E MOTIVAÇÃO

Atividades em KDD - uma Taxonomia



[Goldschmidt et al., 2002a]

INTRODUÇÃO E MOTIVAÇÃO

Tendências e Perspectivas:

- **Mineração de Dados em Larga Escala**
- **Mineração de Dados Paralela e Distribuída**
- **Mineração de Dados Ativa**
- **Mineração de Dados na WEB**
- **Assistência ao Controle do Processo de KDD**

INTRODUÇÃO E MOTIVAÇÃO

Bibliografia:

- **Data Mining: Um Guia Prático – Ronaldo Goldschmidt e Emmanuel Passos, Editora Campus, 2005.**
- **Data Mining: Concepts and Techniques. Jiawei Han e Micheline Kamber, Morgan Kaufman, 2001.**
- **Data Mining: Técnicas e Aplicações para o Marketing Direto - Fernanda Cristina Naliato do Amaral - Ed. Berkeley, 1ª Edição - 2001.**

INTRODUÇÃO E MOTIVAÇÃO

Bibliografia:

- Predictive Data Mining: a practical guide - Shaolom M. Weiss / Morgan Kaufmann, 1ª Edição - 1998.
- Data Mining: A Mineração de Dados no Marketing, Medicina, Economia, Engenharia e Administração, Luiz Alfredo Carvalho, Érica, 2001.
- Data Mining and Knowledge Discovery with Evolutionary Algorithms, Alex Freitas, Springer-Verlag, 2002.

INTRODUÇÃO E MOTIVAÇÃO

Bibliografia:

- Sistemas Inteligentes – Fundamentos e Aplicações, Solange Rezende, São Paulo: Manole, 2003.
- Neural Networks: A Comprehensive Foundation S. Haykin - Marmillan College Publisng Company, 2ª Edição - 1999.
- C45: Programs for Machine Learning - J. R. Quinlan - Morgan Kanfmann, 1ª Edição - 1993.

INTRODUÇÃO E MOTIVAÇÃO

Bibliografia:

- Handbook of Genetic Algorithms L. Davis - Thonson Comp. Press 1ª Edição - 1990.
- Genetic Algorithms + Data Structures = Evolution Programs - Michalewicz, Springer-Verlag 1ª Edição - 1994.
- An Introduction to Fuzzy Sets Analysis and Design, Pedrycz e Gomide, MIT, 1998.
- Fuzzy Logic for Business, Finance and Management, Gomide e Gomide, World Scientific, 1997.