

# Microsoft Innovation Center

## Mineração de Dados (Data Mining)

André Montevecchi

[andre@montevecchi.com.br](mailto:andre@montevecchi.com.br)

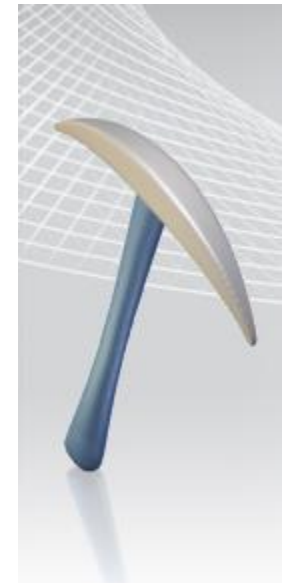
# AGENDA

- Introdução
- Objetivo
- BI e Mineração de Dados
- Aplicações
- Exemplos e Cases
- Algoritmos para Mineração de Dados



# INTRODUÇÃO

- ◉ O que é **Knowledge Discovery in Databases (KDD)**?

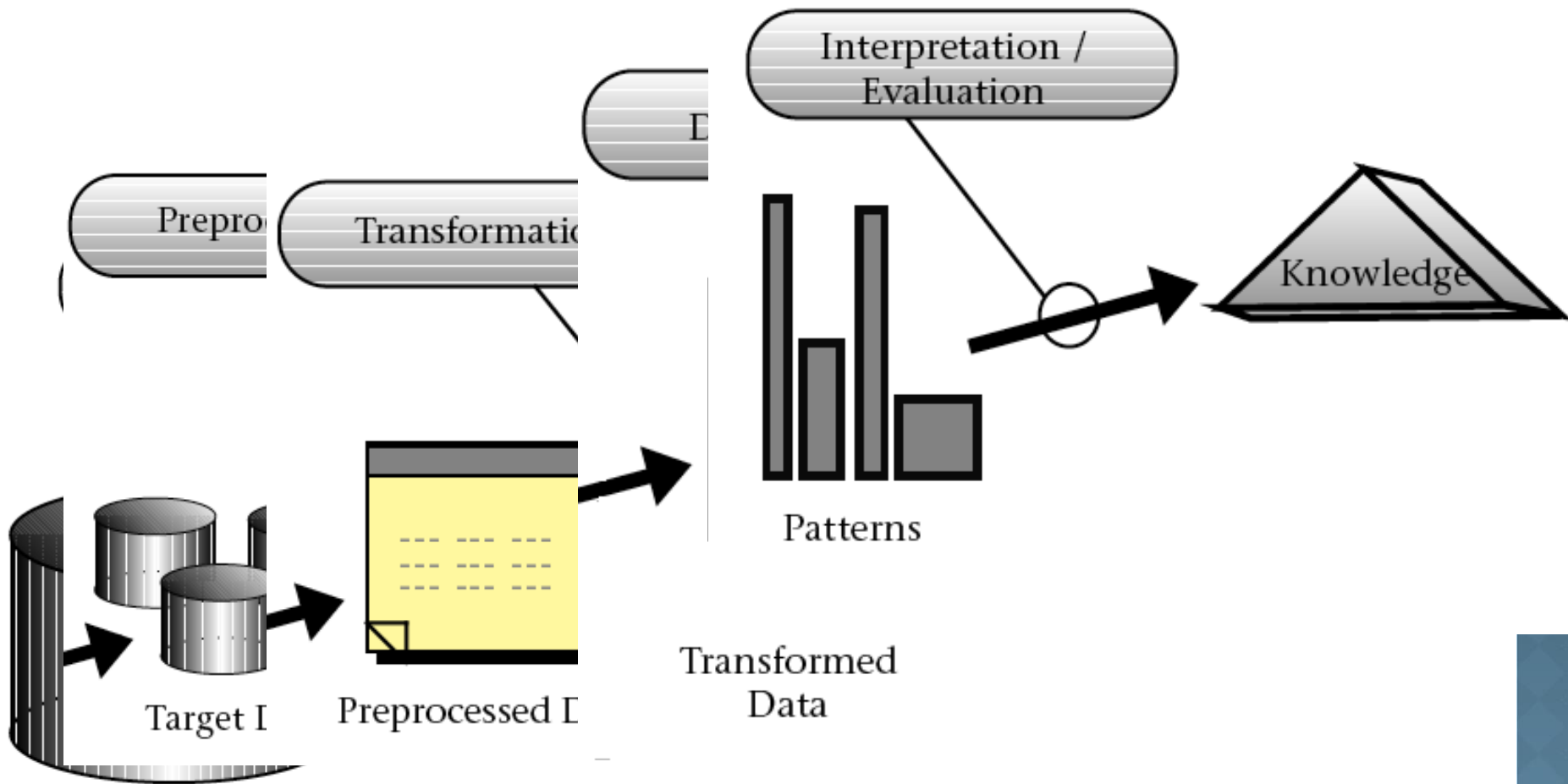


“Processo não-trivial de identificação de padrões válidos, novos, potencialmente úteis e finalmente compreensíveis a partir dos dados”

(Usama Fayyad)



# INTRODUÇÃO



# INTRODUÇÃO

- ⦿ O que é Mineração de dados?

“Mineração de dados é a aplicação de algoritmos específicos para extrair padrões e informações relevantes dos dados.”

(Usama Fayyad)

- ⦿ Data Mining é um passo do processo de KDD.
- ⦿ A mineração de dados (MD) surge para essa finalidade e pode ser aplicada tanto para a pesquisa científica como para impulsionar a lucratividade da empresa.

# OBJETIVO

- ◉ **Porque utilizar Data Mining?**
- ◉ Dados crescem nas empresas
- ◉ Identificar tendências e oportunidades escondidas
- ◉ Competitividade



# BI E DATA MINING

- Business Intelligence (BI), ou Inteligência de negócios, busca acessar, analisar, prever, relatar e compartilhar as informações necessárias para alcançar os objetivos de sua empresa.
- Mineração de Dados busca conhecimento novo e útil acerca do seu meio ambiente.



natizada (Batata-f



# APLICAÇÕES

## ⦿ Marketing:

- Traçar perfil de clientes
- Identificar produto de maior potencial de compra
- Gerar padrões de comportamento de consumidores por regiões

## ⦿ Finanças:

- Detectar padrões de fraudes no uso de cartões de crédito
- Identificar clientes que estão tendendo a mudar de administradora de cartão de crédito
- Identificar regras de estoque a partir dos dados do mercado



# APLICAÇÕES

- ◉ Telecomunicações:
  - Identificar fraudes em ligações telefônicas, principalmente em celulares
- ◉ Medicina:
  - Prever quais pacientes têm maior probabilidade de contrair uma determinada doença
- ◉ Agricultura:
  - Identificar as áreas mais propícias a um bom cultivo de determinado vegetal

# CASES MICROSOFT



## ○ ABS-CBN Interactive (ABSi)

- Empresa de entretenimento nas Filipinas
- Objetivo:
  - Vender ring tones e outros conteúdos para celulares
  - Descobrir quais produtos oferecer para determinados consumidores
- Solução:
  - Utilização de data mining do SQL SERVER 2005 para determinar recomendação de produtos
- Resultado:
  - Melhores recomendações de produtos para seus clientes
  - Dobrou a taxa de resposta de campanhas de marketing

# CASES MICROSOFT



## ○ Clalit Health Services

- Plano de saúde com 3.7 milhões de associados (60% da população de Israel)
- Objetivo:
  - Identificar quais associados mais se beneficiarão de programas de prevenção de doenças
- Solução:
  - Utilização base sócio-demográfica e registros médicos para identificar os associados com maior probabilidade de desenvolverem problemas de saúde
  - Uma vez identificados, são encaminhados a programas de prevenção
- Resultado:
  - Melhora da qualidade de vida
  - Redução de custos

# ALGORITMOS

- ◉ DM da Microsoft possui 9 algoritmos implementados
- ◉ Facilidade de uso
- ◉ Ambiente familiar
- ◉ Possibilidade de usar o EXCEL
- ◉ DM – Parte do SQL SERVER sem custos adicionais
- ◉ Demo EXCEL!

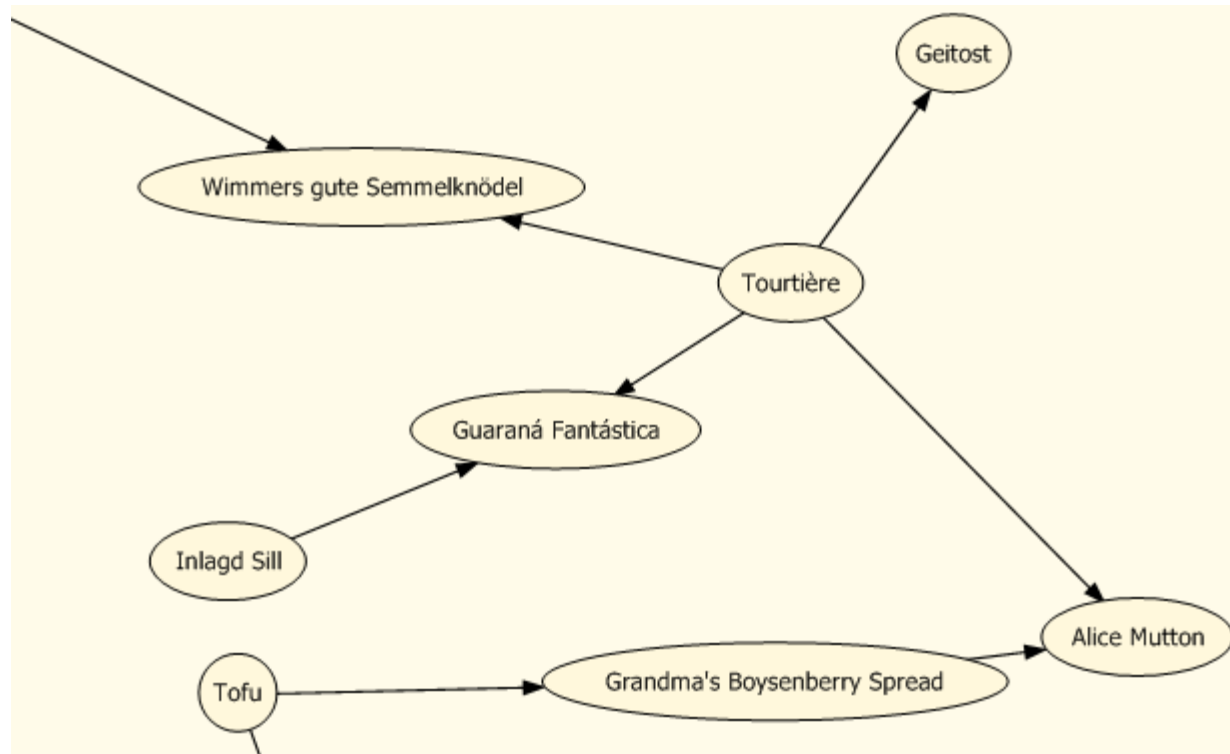


# ALGORITMOS

- ⦿ Microsoft Regras de associação
- ⦿ Microsoft *Clustering*
- ⦿ Microsoft Árvore de Decisão

# MICROSOFT REGRAS DE ASSOCIAÇÃO

- Encontrar dependência entre atributos



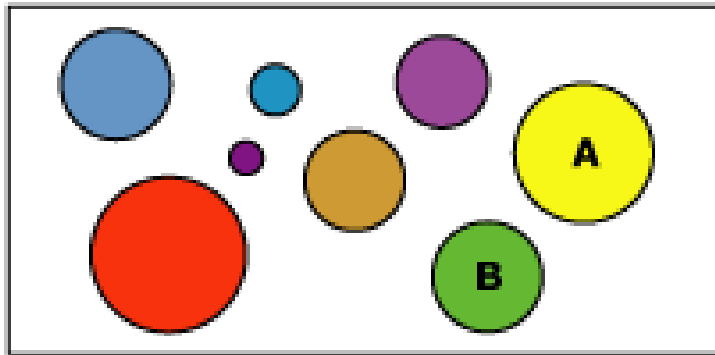
# MICROSOFT REGRAS DE ASSOCIAÇÃO

## Aplicações:

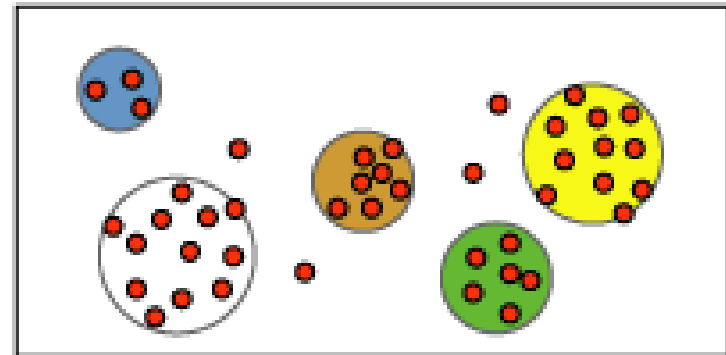
- ◉ Recomendação
- ◉ Projeto de Catálogos
- ◉ Mala direta de Marketing direcionado
- ◉ Planejamento de promoções e vendas

# MICROSOFT CLUSTERING

- Identificar automaticamente grupos de afinidades.



**A** = Commuters who drive to work  
**B** = Commuters who bicycle to work





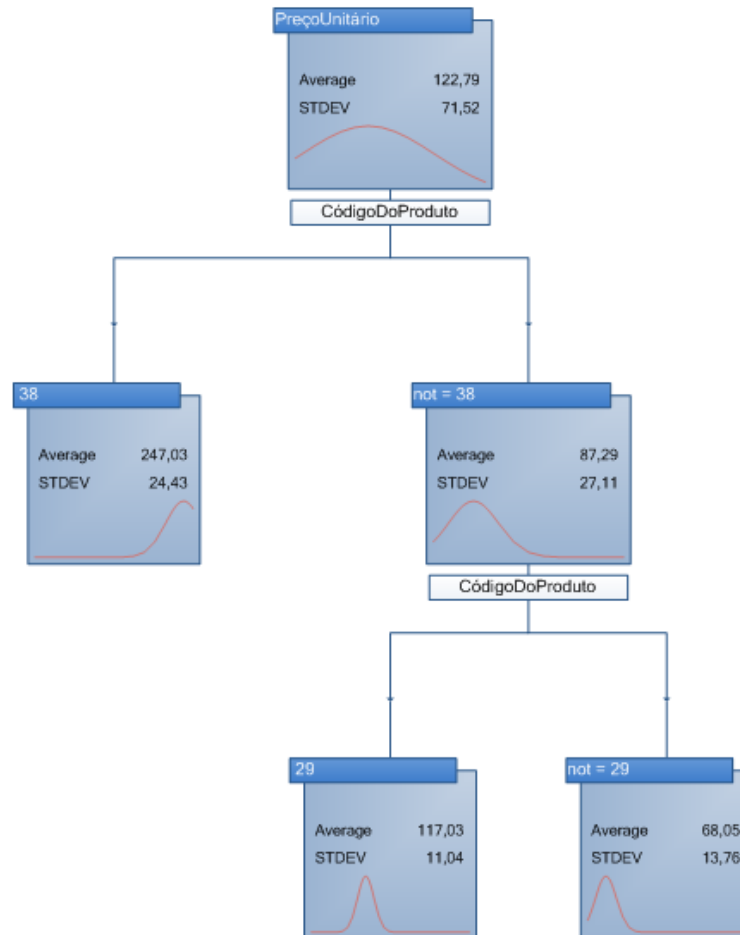
# MICROSOFT CLUSTERING

## Aplicações:

- ⦿ Marketing: identificar grupos de clientes distintos
- ⦿ Agricultura: identificação de áreas de uso de terra
- ⦿ Finanças: Identificar transações suspeitas (Outliers)

# MICROSOFT ÁRVORE DE DECISÃO

- Analisa características e determina a classe mais provável.



# MICROSOFT ÁRVORE DE DECISÃO

## Aplicações:

- ◉ Identificar as características dos clientes antigos que possam indicar se há chance deles realizarem compras futuramente.
- ◉ Classificar clientes em potencial, clientes leais e clientes problemáticos.
- ◉ Análise de crédito.



**Obrigado!**

**André Montevercchi**  
**andre@montevercchi.com**

***Microsoft***<sup>®</sup>

*Your potential. Our passion.*<sup>™</sup>