

USP LABIC

Data Mining (Suporte)

Solange Oliveira Rezende

USP São Carlos / ICMC
Departamento de Ciências de Computação e Estatística
Laboratório de Inteligência Computacional
<http://www.icmc.usp.br>

© Solange Oliveira Rezende

USP LABIC

Tecnologias que dão suporte para Data Mining

sop

USP LABIC

Suporte para Data Mining

- ♦ **Aprendizado de Máquina**
Muito importante em todas as etapas do processo.
- ♦ **Estatística**
Apoia, especialmente, os métodos na etapa de preparação de dados.
- ♦ **Ferramentas de Visualização**
Importante na etapa de pré-avaliação, possibilitando verificar o que foi extraído, e de que forma está organizado.
- ♦ **Banco de Dados e Data Warehouse**
Algumas ferramentas são utilizadas para auxiliar a manipulação dos dados.

sop

USP LABIC

Técnicas Estatísticas

- ♦ As técnicas estatísticas são fundamentais para o processo de DM
- ♦ Amostragem dos dados (seleção e preparação)

sop

USP LABIC

Técnicas Estatísticas

- ♦ Essas técnicas, juntamente com técnicas de IA, especialmente as que manipulam incerteza, provêem mecanismos para:
 - ✦ Evitar *overfitting* nos dados
 - ✦ Tratar ruídos dos dados
 - ✦ Manipular conjunto de dados incompletos
 - ✦ Incorporar conhecimento de fundo

sop

USP LABIC

Ferramentas de Visualização

- ♦ É uma ferramenta exploratória na análise dos dados
- ♦ Permite o aumento da capacidade de análise e de interpretação dos resultados obtidos
- ♦ Os principais tipos dessas ferramentas são:
 - ✦ linguagens de programação especializadas
 - ✦ ferramentas Graphic User Interface (GUI)

sop



Data Warehouse e OLAP

- ♦ Já apresentadas na aula DM-DW-olap.ppt

SOF

