

GT-JeDi - Curso de Desenv. de Jogos IA para Jogos



Fernando Osório

2006/2 - A10

APRENDIZADO: *Machine Learning e Data Mining*



➤ Machine Learning

Conceitos sobre **Machine Learning** (ML = Aprendizado de Máquina)

- Aprendizado Indutivo
- Aprendizado Supervisionado
- Classificação de Padrões
- Aproximação de Funções (Regressão)
- Aprendizado Semi-Supervisionado
- Aprendizado Não Supervisionado
- Técnicas de Aprendizado:

Redes Neurais Artificiais (RNA ou ANN - Artificial Neural Nets),

Árvores de Decisão (IDT - Induction of Decision Trees),

Algoritmos Genéticos (AG ou GA - Genetic Algorithms)

Raciocínio Baseado em Casos (CBR - Case-Based Reasoning)

➤ Data Mining

Conceitos sobre **Mineração de Dados e Descoberta de Conhecimentos (KDD)**

- Agrupamento (Clustering)
- Associações de padrões (Regras de Associação - Algoritmo Apriori)
- Análise de Dados

APRENDIZADO:
Machine Learning e Data Mining



Referência de Base: [Rez 2003]
Sistemas Inteligentes: Fundamentos e Aplicações.
 Solange Rezende (Ed.). Editora Manole, 2003. 525 p.
Capítulo 4: Conceitos sobre Aprendizado de Máquina
 Maria Carolina Monard e José Augusto Baranauskas



Machine Learning

Figura 4.1: A hierarquia do aprendizado [Rez 2003, pg. 91]

APRENDIZADO:
Machine Learning e Data Mining



Aprendizado Indutivo

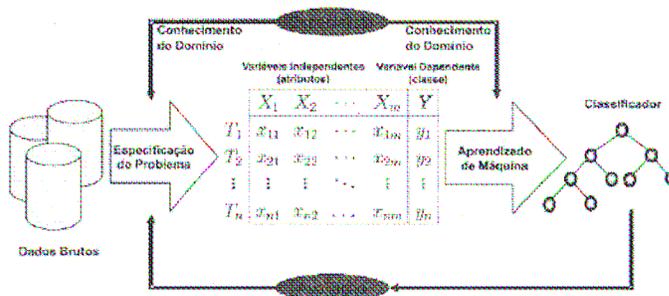


Figura 4.2: Processo de classificação [Rez 2003, pg. 92]

APRENDIZADO: *Machine Learning e Data Mining*



Aprendizado Indutivo:

Conceitos Importantes...

- Base de Aprendizado (Base de Exemplos):

Exemplos, Atributos, Valores Qualitativos e Quantitativos, Classes

Distribuição das Classes, Quantidade de Exemplos e de Atributos

Tipos de Classificadores (Binário, 1-de-N, Conj. de Classes Aberto ou Fechado, ...)

- Algoritmo de Aprendizado:

Simbólico, Estatístico, Conexionista, Evolutivo, "Lazy Learning"

- Aprendizado e Otimização: função alvo e estimativa de erro

- Avaliação de Performance:

Resposta: Acerto, Erro, Sem Resposta (faixas de tolerância - limiares)

Acerto e Erro: Precisão (acurácia), Estimativa de Erro

Avaliação do Erro de Classificação, Erro de Aproximação (MSE) e Erro Majoritário

Curva de Pareto, Curva ROC (FP - Falso Positivo x FN - Falso Negativo)

Matriz de Confusão, Matriz de Contingência

Generalização: Overfitting e Underfitting

Validação Cruzada (Cross-Validation) e N-Fold Cross-Validation

Qual é o melhor algoritmo? Existe "o melhor" de todos?

APRENDIZADO: *Machine Learning e Data Mining*



➤ Referências

* Disciplinas da Unisinos:

> Redes Neurais (Graduação em Computação)

<http://www.inf.unisinos.br/~osorio/neural.html>

> Sistemas Adaptativos (Mestrado em Computação)

<http://www.inf.unisinos.br/~osorio/sistadap.html>

* Livros:

Tom Mitchell - Machine Learning

Solange Rezende - Sistemas Inteligentes

Obras específicas de Redes Neurais, Algoritmos Genéticos e Árvores de Decisão

* Softwares:

SNNS (Neural), C4.5 (Arv. Decisão), GALib (Genético)

Weka (Diversos Algoritmos)