

Aprendizagem de Máquina



George Darmiton da Cunha Cavalcanti

Tsang Ing Ren

CIn/UFPE

Roteiro

- Por que Aprendizagem de Máquina?
- O que é um problema de aprendizagem?
- Um exemplo:
 - Aprendendo a jogar Damas
- Quais questões devem ser feitas (e respondidas) acerca de Aprendizagem de Máquina?

Por que Aprendizagem de Máquina?

- Recentes avanços em algoritmos e em teoria
- Aumento do poder computacional
- Aplicações comerciais e industriais
- A sociedade produz muita informação
 - Fonte: *business*, ciências, medicina, economia, geografia, meio-ambiente, esportes, ...

Termos usados na área

- ❑ Instância, padrão, exemplo, amostra
- ❑ Atributo, característica
- ❑ Os atributos podem ser:
 - Qualitativos (Nominal e Ordinal)
 - Quantitativos (Numéricos, Intervalares)
- ❑ *Missing Values*

Aprendizagem de Máquina

Algoritmos para adquirir e estruturar conhecimento a partir de exemplos

□ Exemplo: regras if-then

```
If tear production rate = reduced  
    then recommendation = none  
Otherwise, if age = young and astigmatic = no  
    then recommendation = soft
```

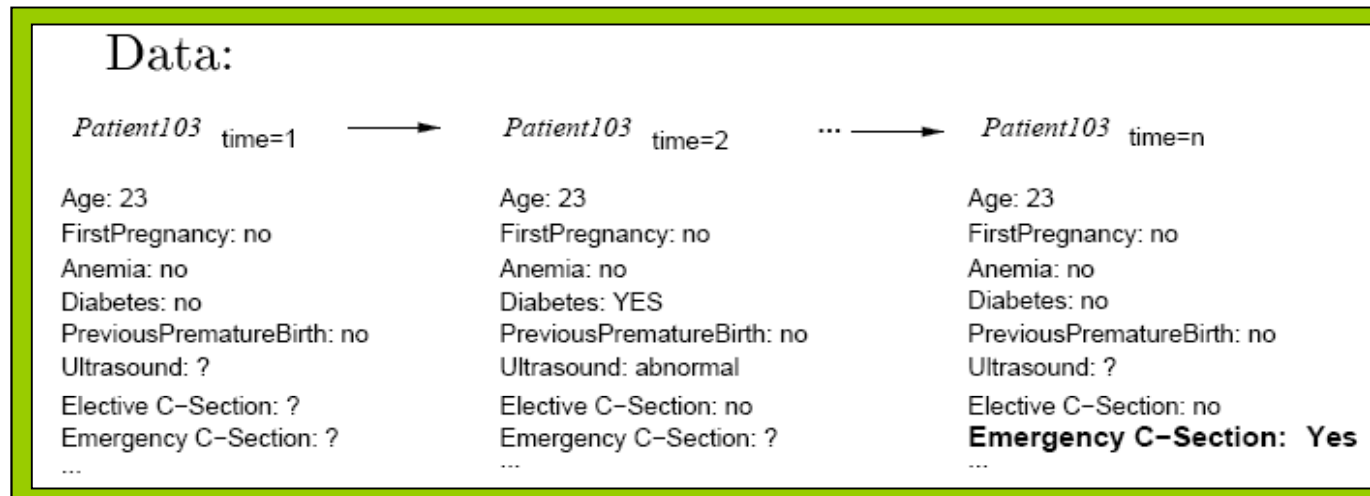


Age	Spectacle prescription	Astigmatism	Tear production rate	Recommended lenses
Young	Myope	No	Reduced	None
Young	Hypermetrope	No	Normal	Soft
Pre-presbyopic	Hypermetrope	No	Reduced	None
Presbyopic	Myope	Yes	Normal	Hard
...

Por que Aprendizagem de Máquina?

- ❑ Três nichos para aprendizagem de máquina
 - Mineração de dados: uso de dados históricos para melhorar a tomada de decisão
 - ❑ Registros médicos → Conhecimento médico
 - Softwares que não são desenvolvidos pela computação convencional
 - ❑ Direção autônoma
 - ❑ Reconhecimento de fala
 - Programas que se adaptam
 - ❑ Leitor de notícias que seleciona os interesses dos usuários

Mineração de dados: um exemplo



- Dado (padrões de treinamento):
 - 9.714 registros de pacientes
 - Cada registro de paciente contém 215 características
- Aprender a prever:
 - Classes de futuros pacientes que possuem alto risco de se submeter a uma operação Cesariana

Mineração de dados: um exemplo

□ Uma das 18 regras aprendidas

If No previous vaginal delivery, and
Abnormal 2nd Trimester Ultrasound, and
Malpresentation at admission
Then Probability of Emergency C-Section is 0.6

Over training data: $26/41 = .63$,
Over test data: $12/20 = .60$

Análise de crédito: um exemplo

Data:

Customer103: (time=t0)

Years of credit: 9
Loan balance: \$2,400
Income: \$52k
Own House: Yes
Other delinquent accts: 2
Max billing cycles late: 3
Profitable customer?: ?

...

Customer103: (time=t1)

Years of credit: 9
Loan balance: \$3,250
Income: ?
Own House: Yes
Other delinquent accts: 2
Max billing cycles late: 4
Profitable customer?: ?

...

...

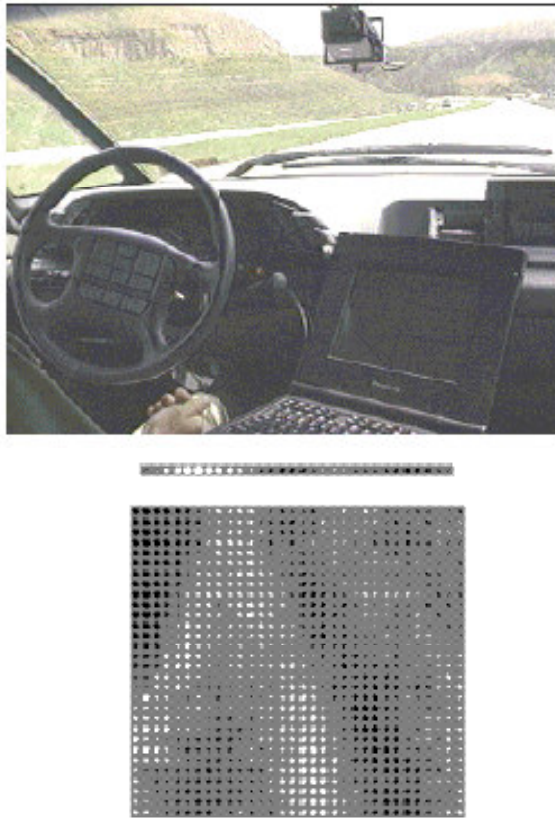
Customer103: (time=tn)

Years of credit: 9
Loan balance: \$4,500
Income: ?
Own House: Yes
Other delinquent accts: 3
Max billing cycles late: 6
Profitable customer?: No

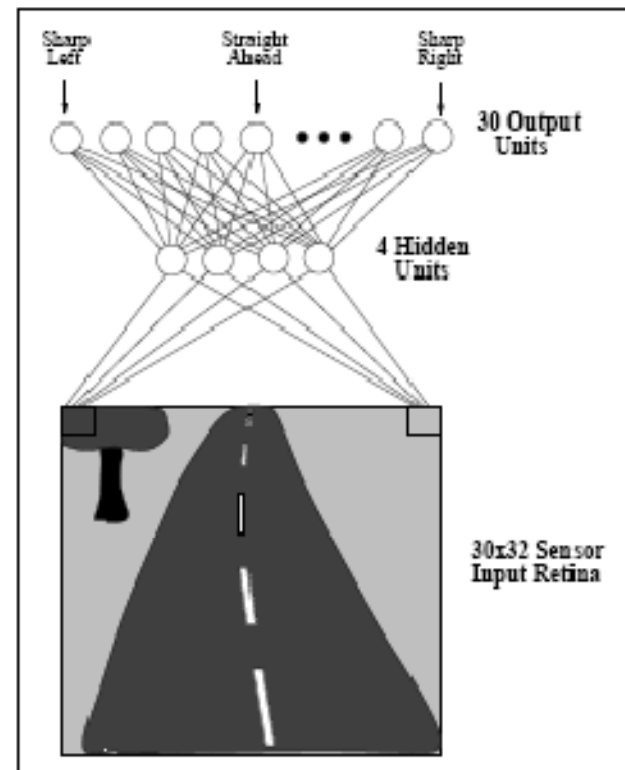
...

- Regras aprendidas dos dados sintéticos:
- [*Deny Credit Card application*]
If Other-Delinquent-Accounts > 2, and Number-Delinquent-Billing-Cycles > 1
Then Profitable-Customer? = No
- [*Accept Credit Card application*]
If Other-Delinquent-Accounts = 0, and (Income > \$30k) OR (Years-of-Credit > 3)
Then Profitable-Customer? = Yes

Problemas difíceis para serem resolvidos pela computação convencional



Stanford's New Driverless Car
GM researching driverless cars



ALVINN [Pomerleau, 1989]
drives 70 mph per 90 miles on public highways

Disciplinas Relevantes

- ❑ Inteligência Artificial
- ❑ Métodos Bayesianos
- ❑ Teoria da Complexidade
- ❑ Controle
- ❑ Teoria da Informação
- ❑ Filosofia
- ❑ Psicologia e neurobiologia
- ❑ Estatística

O que é um problema de aprendizagem?

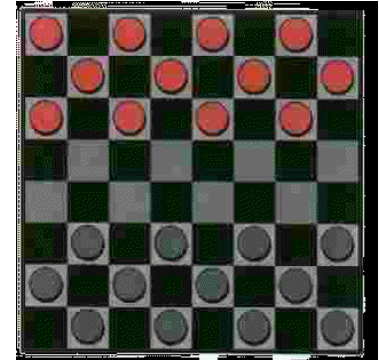
- Aprender = Melhorar a realização de alguma tarefa através da experiência
 - Melhorar a tarefa T
 - Em relação a medida de desempenho P
 - Baseado na experiência E

- Exemplo: Aprender a jogar damas
 - T : jogar damas
 - P : % de vitórias em um torneio
 - E : oportunidade de jogar contra si mesmo

Aprendendo a Jogar Damas

- T : jogar damas
- P : % de vitórias em um torneio

- Perguntas pertinentes
 - O que significa experiência?
 - O que exatamente deve ser aprendido?
 - Como o problema deve ser representado?
 - Qual algoritmo de aprendizagem deve ser empregado?



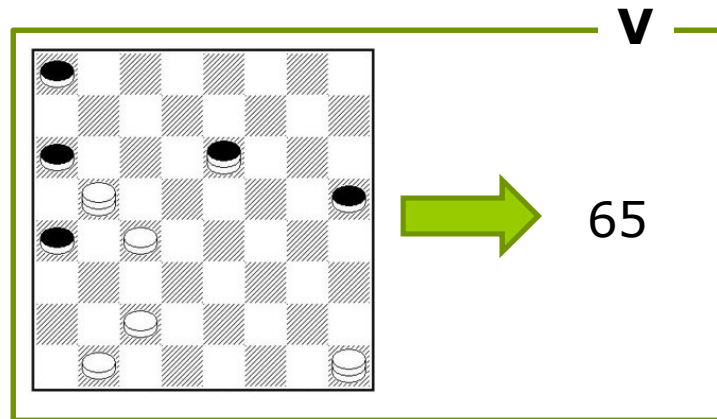
Tipos de Treinamento

- ❑ Direto ou indireto?
 - Exemplo: jogo de damas
 - Direto
 - ❑ Para cada estado é informado o movimento correto
 - Indireto
 - ❑ Seqüência de movimentos e resultado final
 - ❑ Problema: determinar a importância de cada movimento

- ❑ Questão: as estratégias escolhidas irão levar ao objetivo?

Escolhendo a função objetivo

- ❑ Função que dado um tabuleiro escolhe o melhor movimento
 - $ChooseMove: Board \rightarrow Move$
- ❑ Uma alternativa é encontrar uma função que atribui um valor (*score*) dado um tabuleiro
 - $V: Board \rightarrow \mathcal{R}$



Possíveis definições para a função objetivo

- ▣ Se b é um estado final de vitória, então $V(b) = 100$
- ▣ Se b é um estado final de derrota, então $V(b) = -100$
- ▣ Se b é um estado final de empate, então $V(b) = 0$
- ▣ Se b não é um estado final, então $V(b) = V(b')$, sabendo que b' é o melhor estado final que pode ser alcançado a partir do estado b
- ▣ Essa estratégia gera respostas corretas, entretanto não é factível devido ao custo computacional

Escolhendo a função objetivo

- ❑ Conjunto de regras?
- ❑ Árvore de Decisão?
- ❑ Redes Neurais Artificiais?
- ❑ Funções Polinomiais?
- ❑ Máquinas de Vetores de Suporte
- ❑ Modelos Escondidos de Markov
- ❑ ...

Uma representação para a função objetivo

$$V'(b) = \omega_0 + \omega_1 \cdot bp(b) + \omega_2 \cdot rp(b) + \omega_3 \cdot bk(b) + \omega_4 \cdot rk(b) + \omega_5 \cdot bt(b) + \omega_6 \cdot rt(b)$$

$bp(b)$: número de peças pretas no tabuleiro b

$rp(b)$: número de peças vermelhas no tabuleiro b

$bk(b)$: número de rainhas pretas em b

$rk(b)$: número de rainhas vermelhas em b

$bt(b)$: número de peças vermelhas ameaçadas pelas pretas (i.e., as quais podem ser capturadas na próxima jogada)

$rt(b)$: número de peças pretas ameaçadas pelas vermelhas

Obtendo exemplos de treinamento

$V(b)$: a função objetivo verdadeira

$V'(b)$: a função aprendida

$V_{train}(b)$: o valor de treinamento

Uma regra para estimar os valores de treinamento:

$$V_{train}(b) \leftarrow V'(Sucessor(b))$$

Regra para ajustar os pesos

□ Regra de ajuste dos pesos LMS

- LMS: *Least Mean Squares*

□ Repita

- Selecione um exemplo de treinamento b aleatoriamente

- i) Calcule $error(b)$

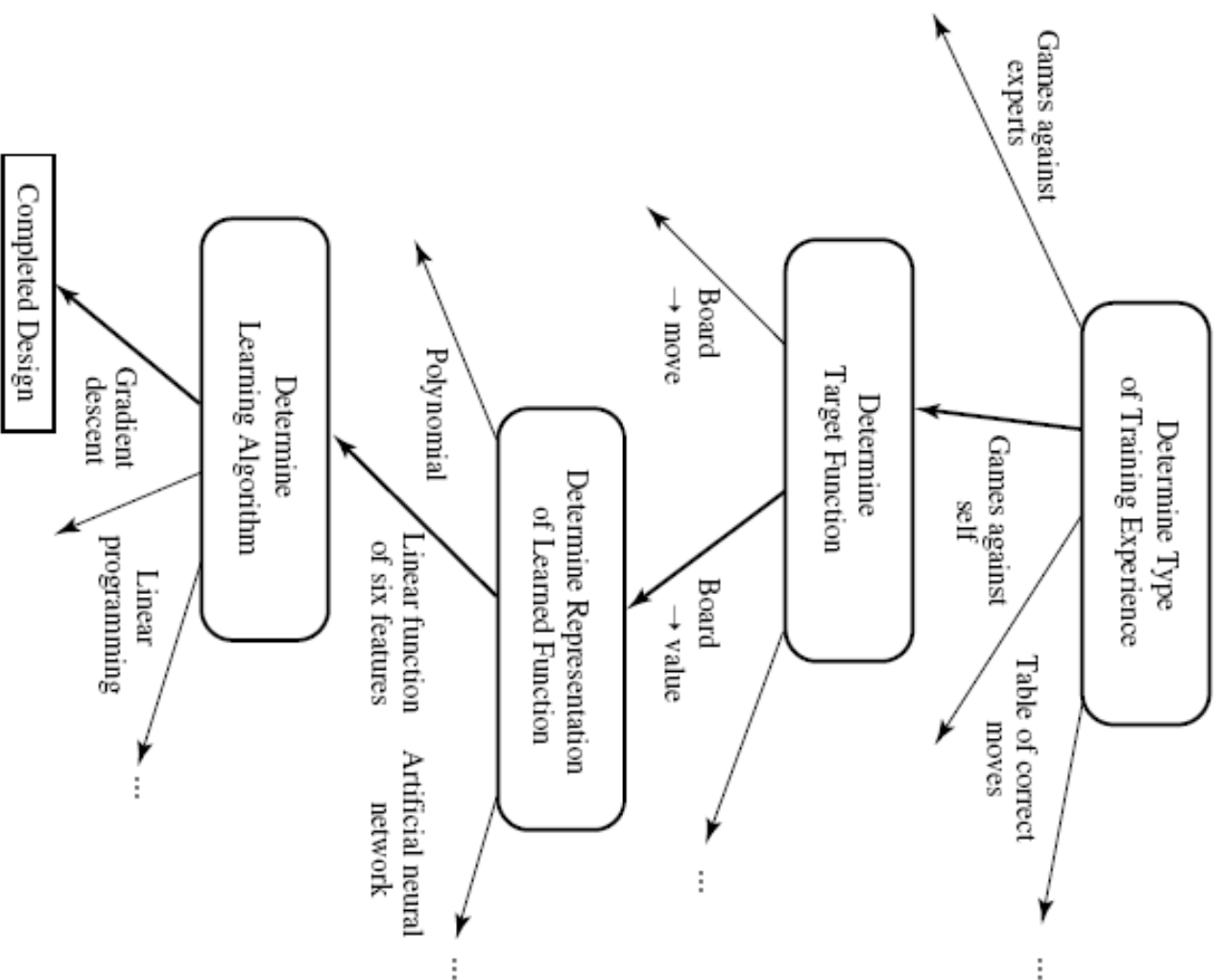
$$error(b) = V_{train}(b) - V'(b)$$

- ii) Para cada característica f_i ajuste o peso w_i

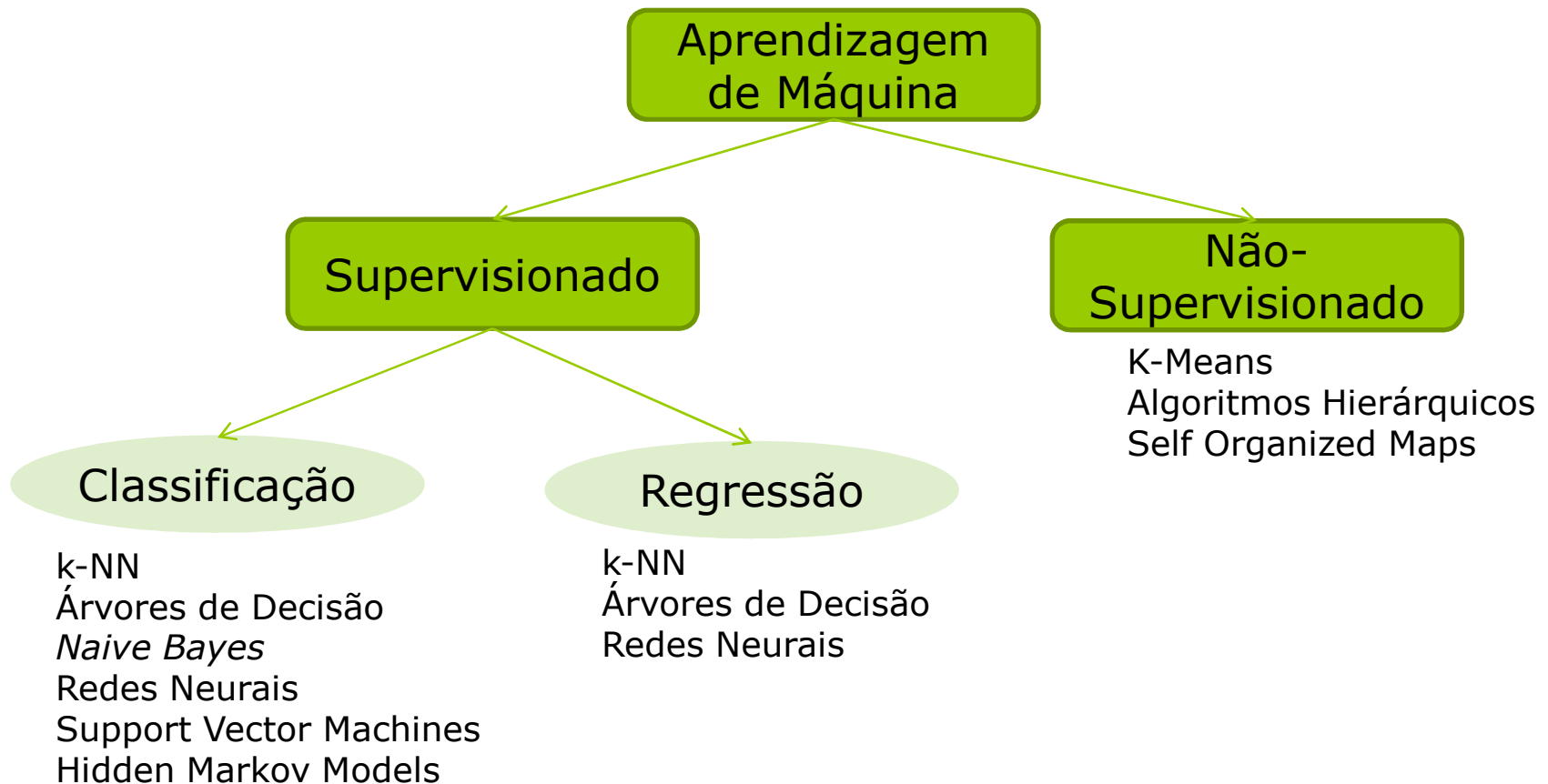
$$\omega_i \leftarrow \omega_i + c \cdot f_i \cdot error(b)$$

c é uma constante pequena (ex: 0.1) para moderar a taxa de aprendizagem

Opções de Projeto



Tipos de Aprendizagem



Alguns pontos em Aprendizagem de Máquina

- ❑ Quais algoritmos aproximam funções bem (e quando)?
- ❑ Como o número de exemplos de treinamento influencia o desempenho?
- ❑ Como a complexidade das hipóteses impacta no desempenho?
- ❑ Como dados ruidosos influenciam na precisão?
- ❑ Quais são os limites teóricos da aprendizagem?
- ❑ Como inserir conhecimento *a priori*?
- ❑ Quais pistas podemos capturar de sistemas biológicos?
- ❑ Como sistemas podem se adaptar?

Aprendizagem e Ética



- ❑ Algoritmos, como os de Mineração de Dados, são freqüentemente usados de forma discriminatória
 - Exemplo:
 - ❑ Empréstimo bancário: uso de algumas informações (ex. sexo, religião, cor) não é eticamente correto
- ❑ Questões importantes:
 - Quem deve ter acesso aos dados?
 - Para que propósito os dados são coletados?
 - Que tipo de conclusão pode ser legitimado pelos dados?
- ❑ Estatística apenas não é suficiente!

Informações Básicas

- Site da disciplina
 - <http://www.cin.ufpe.br/~if699>
- Horário
 - Terças (10:00-8:00) e quintas (8:00-10:00)

Bibliografia

- Tom Mitchell. **Machine Learning**. McGraw-Hill. 1997.
- S. Theodoridis and K. Koutroumbas. ***Pattern Recognition***. Academic Press. 2006.
- Christopher M. Bishop. **Pattern Recognition and Machine Learning**. Springer. 2006
- Richard O. Duda, Peter E. Hart and David G. Stork. **Pattern Classification**. Wiley-Interscience. 2000

