

Aprendizagem de máquina

Inteligência Artificial

André Câmara



Slides adaptados do professor Valmir Macario (Deinfo-UFRPE)

Aplicações - AM



Aprendizagem

- A palavra "aprendizagem" tem muitos significados diferentes:
 - **Memorizar** algo;
 - **Aprender** fatos através da observação e exploração;
 - **Melhorar** movimento e/ou habilidades cognitivas através da prática;
 - **Organização** de novos conhecimentos em **representações** eficazes.



Aprendizagem indutiva

- Dada uma coleção de exemplos de f , retornar uma função h que se aproxime de f .
- A função h é chamada **hipótese**
 - Não é fácil saber se uma h específica é uma boa aproximação de f
 - Uma boa h irá **generalizar** bem – prever corretamente exemplos ainda não vistos
 - Esse é o **problema fundamental da indução**



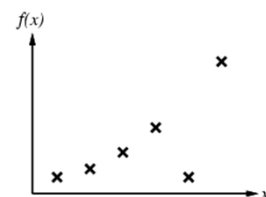
Aprendizagem

- "Aprendizagem é o processo pelo qual uma entidade **adquire** conhecimento". Rich, 1983.
- "Aprendizagem é uma forma de **inferência** que objetiva começar com informações sobre o domínio e então, **estendê-las** de alguma forma". Ginsberg, 1993.
- Para Russel e Norvig: "Aprender significa se comportar melhor ao adquirir **experiência**. Todo aprendizado aprende a representação de uma função."



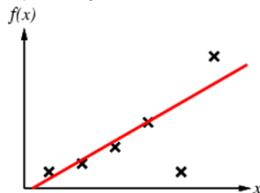
Aprendizagem Indutiva

- **Exemplo:** Ajustar uma função de uma única variável a alguns pontos de dados
 - Exemplos são pares $(x, f(x))$ de números reais

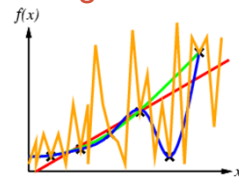


Aprendizagem Indutiva

- Construir/ajustar h para coincidir com f no conjunto de treinamento
- (h é consistente se coincide com f em todos os exemplos)
- Por exemplo, o ajuste da curva:



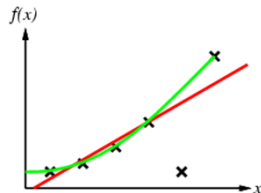
Inductive learning method



- **Ockham's razor:** preferir a hipótese mais simples, consistente com os dados.

Aprendizagem Indutiva

- Construir/ajustar h para coincidir com f no conjunto de treinamento
- (h é consistente se coincide com f em todos os exemplos)
- Por exemplo, o ajuste da curva:

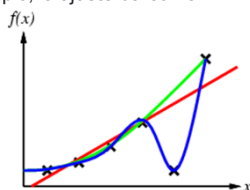


Questões: on-line x off-line

- **Aprender de uma vez ou aos poucos?**
 - **Incremental (on-line):** atualiza hipótese a cada novo exemplo
 - mais flexível, ... porém
 - ordem de apresentação é importante (*backtracking*)
 - é difícil revisar as crenças
 - **não incremental (off-line):** gera h a partir de todo conjunto de exemplos
 - mais eficiente e prática
 - mais usado!

Aprendizagem Indutiva

- Construir/ajustar h para coincidir com f no conjunto de treinamento
- (h é consistente se coincide com f em todos os exemplos)
- Por exemplo, o ajuste da curva:



Aprendizagem de Máquina

- **Aprendizagem Supervisionada:**
 - Ocorre nas situações em que é possível perceber tanto as entradas como as saídas;
 - Frequentemente as saídas são fornecidas por um supervisor(especialista) humano;
 - Envolve aprendizagem de uma função a partir de exemplos de suas entradas e saídas.
- **Aprendizagem Não-Supervisionada:**
 - Envolve a aprendizagem de padrões na entrada, quando não são fornecidos valores de saída específicos.
- **Aprendizagem por Reforço:**
 - O agente deve aprender a partir do reforço (recompensa) ou punição.

A Máquina Aprendendo

- Especifique seu problema como uma tarefa de aprendizagem;
 - Conjunto de exemplos de seu problema.
- Escolha o método de aprendizagem;
- Aplicar o método de aprendizagem;
- Avaliar os resultados e o método.



Conceitos e Definições

- **Exemplo** (caso, registro ou dado) é uma tupla de valores de atributos
 - Um paciente, dados médicos sobre uma determinada doença
- **Atributo**: descreve uma característica ou um aspecto de um exemplo.
 - Nominal: cor
 - Contínuo: peso
- **Classe**: atributo especial (aprendizado supervisionado), denominado rótulo ou classe.
 - Classificação: discreto ou nominal C1, C2, ... Ck
 - Regressão: valores reais



Representação de Exemplos

- Conjunto de exemplos;
- Atributos de domínio;
- Erros nos exemplos.



Atributos de Domínio

- Conceitos que **descrevem** os exemplos
 - (Alguns) serão encontrados na solução para o problema.
- Alguns conceitos ...
 - são necessários para especificar exemplos:
 - Exemplo: dados do pixel para reconhecimento de escrita.
 - são atributos dos exemplos (funções)
 - número_de_pernas(humano) = 2;
 - revestimento(truta) = escama
 - especificam categorizações binárias:
 - e_homeotérmico(humano); põe_ovos(galinha);
- Questões sobre os conceitos de domínio:
 - O que será mais útil na solução?
 - O que pode ser descartado sem preocupação?
 - Quais são binários, quais são funções?



Conceitos e Definições

- **Conjunto de exemplos**:
 - Um conjunto de exemplos⁴ composto por exemplos contendo valores de atributos bem como a classe associada.

Tempo	Temperatura	Humidade	vento	Joga
Sol	85	85	Não	Não
Sol	80	90	Sim	Não
Nublado	83	86	Não	Sim
Chuva	70	96	Não	Sim
Chuva	68	80	Não	Sim
Chuva	65	70	Sim	Não
Nublado	64	65	Sim	Sim
Sol	72	95	Não	Não
Sol	69	70	Não	Sim
Chuva	75	80	Não	Sim
Sol	75	70	Sim	Sim
Nublado	72	90	Sim	Sim
Nublado	81	75	Não	Sim
Chuva	71	91	Sim	Não



Conceitos e Definições

- **Ruído**:
 - É comum no mundo real, trabalhar com dados imperfeitos.
 - Podem ser derivados:
 - do próprio processo que gerou os dados
 - do processo de aquisição de dados
 - do processo de transformação
 - ou mesmo devido a classes rotuladas incorretamente
 - por exemplo, exemplos com os mesmos valores de atributos mas com classes diferentes



Conceitos e Definições

- **Overfitting** (“superajustamento”):

- É dito que ocorreu um *overfitting* do classificador quando este obtém um ótimo resultado sobre os dados utilizados no treinamento, mas não é capaz de generalizar para classificar corretamente novos exemplos



Aplicando Algoritmos de Aprendizagem

- Divisão do conjunto de exemplos (dados):

- Conjunto de Treinamento
- Conjunto de Validação
- Conjunto de Teste

- Tarefas:

- Classificação;
- Agrupamento;
- Regressão;



Conceitos e Definições

- **Missing Values** (Valores perdidos/faltos): em geral, indicados por valores fora do escopo

- Tipos: desconhecidos, não registrados, irrelevantes

- Razões

- Mau-funcionamento do equipamento
- Mudanças na definição do experimento
- Incapacidade de medição

- Valores perdidos podem, de fato, significar alguma coisa

- A maioria dos métodos de aprendizado não assumem isto
- No entanto, este tipo de informação pode ser codificado como um valor adicional



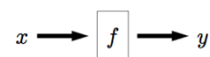
Exemplo: classificação de spam

- **Entrada:** x = email

From: pliang@cs.stanford.edu Date: September 23, 2015 Subject: CS221 announcement Hello students, I've attached the answers to homework 1...	From: a9k62n@hotmail.com Date: November 1, 2012 Subject: URGENT Dear Sir or maDam: my friend left sum of 10m dollars...
--	---

- **Saída:** $y \in \{\text{spam}, \text{not spam}\}$

- **Objetivo:** obter um **preditor** f

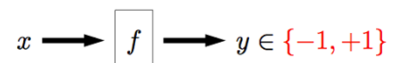


APLICANDO ALGORITMOS DE APRENDIZAGEM

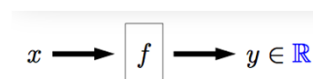


Tipos de tarefas de predição

- **Classificação binária** (e.g. spam / not spam)



- **Regressão** (e.g., localização, ano, idade, etc)



Tipos de tarefas de predição

- **Multiclass classification:** y é uma categoria (e.g., imagem → gato, porta, cachorro, ...)
- **Ranking:** y é uma permutação (e.g., lista de páginas web → ranking)
- **Predição estruturada:** y é um objeto composto por várias partes (uma sentença/frase)



Precisão do classificador

- Complemento da taxa de erro denotado por $acc(h)$

$$acc(h) = 1 - err(h)$$



Avaliação de Sistemas de Aprendizado

- Experimental
 - Conduzir experimentos controlados com validação cruzada para comparar métodos em uma variedade de bases de dados.
 - Coletar dados sobre o seu desempenho, ex.: acurácia de teste, tempo de treinamento, tempo de teste.
 - Analisar diferenças quanto a significância estatística.
- Teórica
 - Analisar algoritmos matematicamente e provar teoremas sobre sua:
 - Complexidade computacional
 - Abilidade de se adaptar aos dados de treinamento
 - Número de exemplos de treinamento necessários para se aprender uma função correta



Erro de regressão

- Calcula-se a distância entre o valor real com o atributo y_i pela hipótese induzida. Pode-se usar: erro médio quadrado (mse) ou a distância absoluta média (mad)

$$mse-err(h) = \frac{1}{n} \sum_{i=1}^n (y_i - h(x_i))^2$$

$$mad-err(h) = \frac{1}{n} \sum_{i=1}^n |y_i - h(x_i)|$$



Generalizar é Difícil

- **Não queremos aprender por memorização**
 - Boa resposta sobre os exemplos de treinamento somente.
 - Fácil para um computador.
 - Difícil para os humanos.
- **Aprender visando generalizar**
 - Mais interessante.
 - Fundamentalmente mais difícil: diversas maneiras de generalizar.
 - Devemos extrair a essência, a estrutura dos dados e não somente aprender a boa resposta para alguns casos.



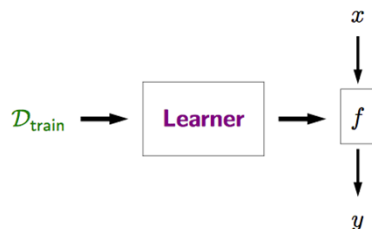
O que fazer em caso de muitos erros?

- Precisa de mais características?
- Características existentes descrevem bem o problema?
- Parâmetros dos algoritmos de aprendizagem foram inicializados corretamente?
- Que tal avaliar outro algoritmo?



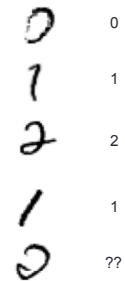
Resumo

- Framework de A.M.

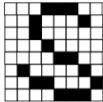



Reconhecedor de números

- Configuração:
 - Obter uma grande coleção de imagens, cada uma rotulada com um dígito
 - Nota: alguém tem que rotular a mão todos os dados!
 - Quer aprender a prever rótulos de novas imagens
- Características: Atributos que ajudem a tomar uma decisão de qual dígito a imagem representa
 - Pixels: (6,8)=ON
 - Padrões de Forma: NumComponentes, NumLoops, entropia
 - ...

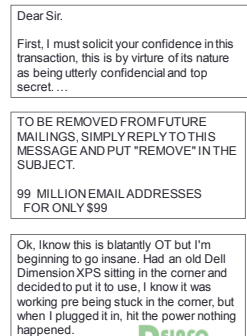


Exemplo Reconhecimento de Caracteres

- Positivo:
 - Esta é a letra S:
 
- Negativo:
 - Esta não é a letra S:
 
- Conceito de domínio:
 - Informação do pixel
- Categorias:
 - Pares (Matriz, Letra)
 - Exemplos positivos e negativos
- Tarefa
 - Classificar (Categorizar) corretamente um exemplo não visto entre uma das 26 categorias (A-Z)

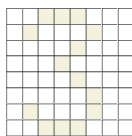
Outro Exemplo: Filtro de Spam

- Entrada: x = e-mail
- Saída: y = "spam" ou "email"
- Setup:
 - Obter uma grande coleção de e-mails, cada um identificado como "spam" ou "email"
 - OBS: Alguém tem que rotular esses dados na mão!
 - Quer aprender a prever rótulos de novos e-mails futuros.
- Características: Os atributos utilizados para tomar a decisão: email / spam
 - Palavras: FREE!
 - Padrões de Texto: \$dd, CAPS
 - Non-text: SenderInContacts

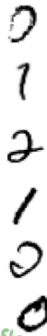


Reconhecedor de números

- Entrada: x = matriz de pixels

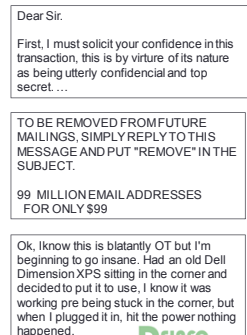


- Saída: y = um número 0-9



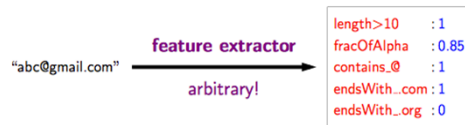
Filtro de Spam

- Classificador
- Dados:
 - Coleção de e-mails rotulados, spam or email
- Classificador
 - Aprende com o conjunto de treinamento
 - (Ajusta com o conjunto de validação)
 - Testa em novos e-mails



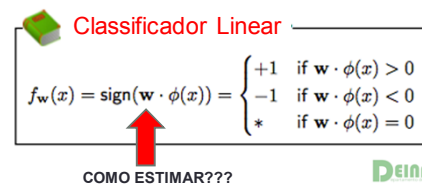
Extração de características

- **Exemplo:** dada uma string x , prever se é um endereço de email
- **Pergunta:** quais propriedades são relevantes para prever y ?
- Extrator de características:



Preditor linear

- Dados:
- Weight vector $\mathbf{w} \in \mathbb{R}^d$
- Feature vector $\phi(x) \in \mathbb{R}^d$



Vetor de características Notação

- Matematicamente, não requer os nomes dos atributos



$$\phi(x) = [\phi_1(x), \dots, \phi_d(x)].$$

$$\phi(x) \in \mathbb{R}^d \quad \square \text{ Ponto em um espaço hiperdimensional}$$



Loss functions

Definição:

Uma função de custo $\text{Loss}(x, y, \mathbf{w})$ mede o quão insatisfeito você ficaria se você utilizasse \mathbf{w} para fazer uma predição sobre x , quando a saída correta é y .

- *Loss minimization framework*
- Atacar o problema de aprendizagem como um problema de otimização



Preditor linear

Weight vector $\mathbf{w} \in \mathbb{R}^d$

Feature vector $\phi(x) \in \mathbb{R}^d$

length>10	: -1.2
fracOfAlpha	: 0.6
contains_@	: 3
endsWith_.com	: 2.2
endsWith_.org	: 1.4

length>10	: 1
fracOfAlpha	: 0.85
contains_@	: 1
endsWith_.com	: 1
endsWith_.org	: 0

- **Score:** combinação ponderada das features (atributos)

$$\mathbf{w} \cdot \phi(x) = \sum_{j=1}^d w_j \phi(x)_j$$

Ex.: $-1.2(1) + 0.6(0.85) + 3(1) + 2.2(1) + 1.4(0) = 4.51$



Recursos

- Bases de Dados
 - UCI Repository: <http://www.ics.uci.edu/~mllearn/MLRepository.html>
 - UCI KDD Archive: <http://kdd.ics.uci.edu/summary.data.application.html>
 - Statlib: <http://lib.stat.cmu.edu/>
 - Delve: <http://www.cs.utoronto.ca/~delve/>
 - LETOR: <http://research.microsoft.com/en-us/um/beijing/projects/letor/index.html>
- Ferramenta
 - WEKA: <http://www.cs.waikato.ac.nz/ml/weka/>
 - Matlab: <http://www.mathworks.com/products/matlab/>
 - RapidMiner: <http://rapid-i.com/content/view/181/190/>

