

Aprendizado de Máquina

Introdução

Luiz Eduardo S. Oliveira

Universidade Federal do Paraná
Departamento de Informática
<http://lesoliveira.net>

Introdução

- Desde a invenção dos computadores temos nos perguntado se os computadores são capazes de aprender.
- Infelizmente ainda estamos muito longe do aprendizado humano
- Algoritmos eficientes em certos tipos de tarefas de aprendizagem e um entendimento teórico tem surgido nos últimos anos.

Quando é necessário utilizar aprendizagem de máquina?

Sempre que não for possível escrever um algoritmo determinístico para resolver o problema. Nesses casos o conhecimento deve ser extraído a partir de exemplos.

O que é

Aprendizagem é uma propriedade essencialmente humana

Aprender

Significa **mudar para fazer melhor** quando uma situação similar acontecer

Aprender não é memorizar e sim ter a capacidade de generalizar um dado comportamento para uma nova situação

- Computadores memorizam facilmente.
- Dificuldade em generalizar.

Tipos de Aprendizagem

Supervisionada

- Fornecemos a “resposta correta” durante o treinamento.
- Classes são conhecidas a priori
- Ajustamos os pesos em função das respostas corretas que conhecemos

Não Supervisionada

- Classes não são conhecidas
- Algoritmo deve definir quais são as classes em função de um determinado critério
 - ▶ Descoberta de Conhecimento

Duas Fazes da Aprendizagem

Treinamento (supervisionado)

- Apresentados exemplos ao sistema
- Sistema aprende a partir de exemplos
- O sistema modifica gradualmente seus parâmetros ajustáveis para que a saída se aproxime da saída desejada.
 - ▶ Métodos paramétricos de aprendizagem supervisionada, como por exemplo Redes Neurais, Máquinas de Vetores de Suporte, etc.

Utilização

- Também conhecida como fase de testes.
- Novos exemplos jamais vistos aparecem
- Desejável que o sistema tenha um bom poder de generalização

Problemas de Aprendizagem

Classificação

- Problema clássico de aprendizagem supervisionada
- Dada uma entrada, atribui a mesma a uma das n classes pré-definidas problema em questão.
- Ex: OCR

Regressão

- Fazer uma predição a partir de um exemplo
- Ex: Predizer o valor de uma ação dados os valores de dias anteriores.

Problemas de Aprendizagem

Por Reforço (*Reinforcement Learning*)

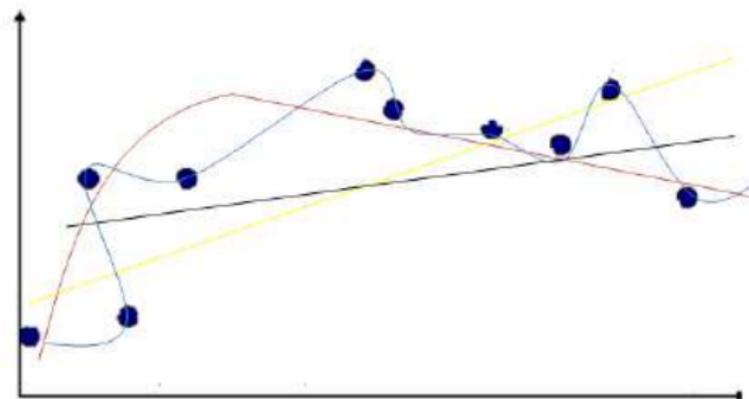
- Não damos a boa resposta, mas dizemos se a decisão tomada pelo algoritmo foi boa ou ruim.
- Ex: Navegação de Robôs

Incremental

- Dados podem não estar disponíveis para aprendizagem ao mesmo tempo
- Chegada em lotes
- Aparição de novas classes
- Alguns modelos não são adequados para esse tipo de aprendizagem
 - ▶ Catastrophic forgetting

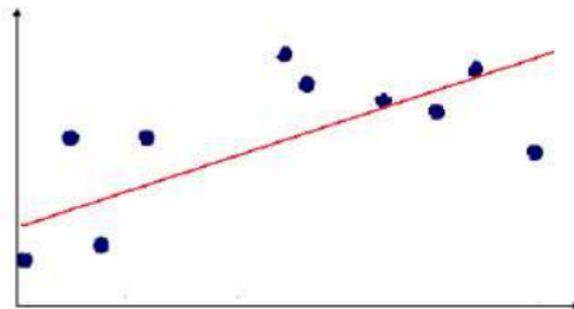
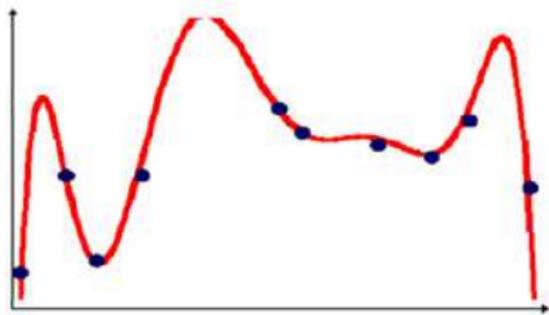
Porque não é Trivial

- Dado uma quantidade finita de dados para o treinamento, temos que derivar uma relação para um domínio infinito.
- Existe um número infinito de tais relações.



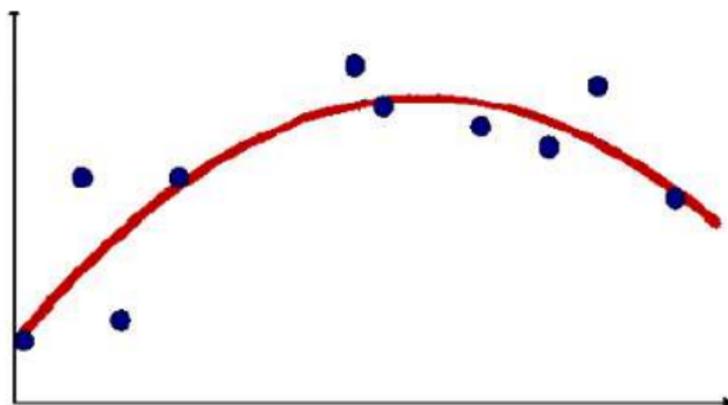
Ajuste dos Modelos

- Qual relação seria mais apropriada?
- A resposta mais simples seria aquela que tem melhor poder de generalização



Ajuste de Modelos

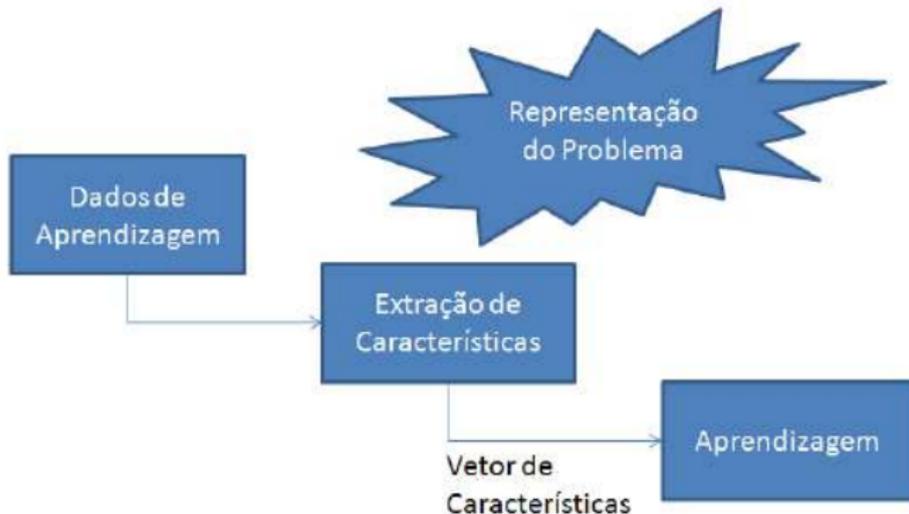
Um bom modelo



A validação cruzada é uma boa estratégia para evitar problemas de *over-fitting*.

- Utiliza uma base de dados auxiliar chamada de “Base de Validação”

Estrutura Básica



Extração de Características

Objetivo

Caracterizar um objeto através de medidas, as quais são bastante similares para objetos da mesma classe, e bastante diferentes para objetos de outras classes.

- Características devem ser discriminantes e invariantes.
- A maioria dos modelos de aprendizagem exigem que as características sejam normalizadas.

Extração de Características

Qualquer medida que se possa extrair de um determinado objeto.

- Simbólicas
 - ▶ Ex: Cor
- Numéricas contínuas
 - ▶ Peso
- Binárias
 - ▶ Determinam a presença ou ausência de uma determinada característica.

Alguns algoritmos de aprendizagem de máquina são mais indicados para determinados tipos de características.

Extração de Características

Globais

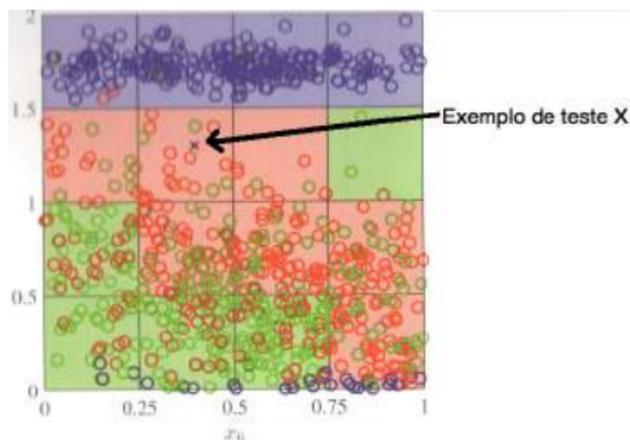
- Extrair características de uma maneira holística.
- Maneira como os humanos reconhecem objetos
- Gestalt (Percepção)
- Em alguns casos a segmentação para extração de características locais não é fácil, como por exemplo, texturas.

Locais

- Segmentar em partes menores para então extrair características.
- Identificar a estrutura do objeto a ser reconhecido e então extrair partes de interesse.

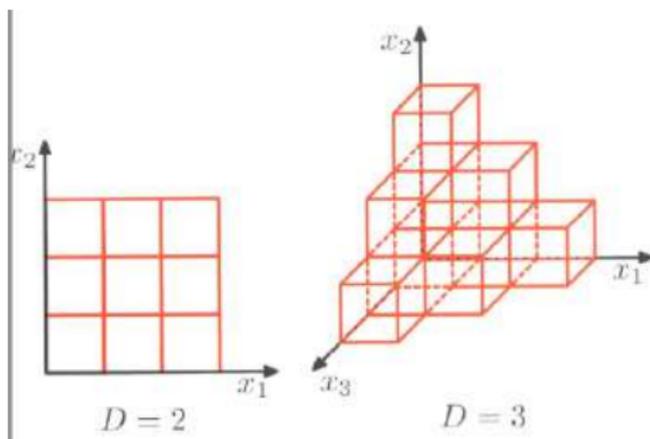
Maldição da Dimensionalidade

- Em geral, na busca por melhores resultados na aprendizagem, criam-se vetores de características com mais informações.
- Para ilustrar o problema, vamos considerar o exemplo da Figura abaixo, com duas características x_6 e x_7 e três classes.
- Para classificar o padrão x , dividimos o espaço em células de tamanho igual e atribuímos a x a classe mais frequente dentro da célula.



Maldição da Dimensionalidade

- Conforme adicionamos características ao vetor, o número de células cresce de forma exponencial.
- Em função disso, quanto mais características temos, mais base de dados temos que ter para preencher todas as células.
- Na maioria dos problemas, entretanto, a quantidade de dados disponível para a aprendizagem é limitada.



Tamanho da Base de Teste

Qual é o tamanho ideal de uma base de teste para avaliar o algoritmo proposto?

- O número de exemplos na base de teste deve ser inversamente proporcional ao erro do melhor classificador.
- Mas qual é o melhor classificador se ainda estamos construindo o sistema?
 - ▶ O ser humano.
- A regra usada para estimar o tamanho mínimo de uma base de teste (n) é dada por $n = 100/p$, em que p é o erro do ser humano.
- Considere por exemplo que o ser humano tenha um erro de 1% em um dado problema (e.g., reconhecimento de caracteres)
- Uma base de teste adequada seria 10000 caracteres = $100/0.01$

Métricos e Não Métricos

Para características contínuas, em geral os métodos chamados **métricos** são mais adequados

- kNN
- Redes Neurais
- Classificadores Lineares

Entretanto, algumas vezes as características disponíveis são nominais, simbólicas, etc..

Nesses casos métodos **não métricos** são mais adequados

- Árvores de decisão
- Aprendizagem conceitual.

Reconhecimento de strings

- Encontrar um determinado segmento de DNA dentro de uma cadeia muito maior

Múltiplos Classificadores

Em diversas situações, a combinação de diversos *experts* pode trazer benefícios na resolução de um problema

Sistemas com múltiplos classificadores podem ser divididos em

- Combinação
 - ▶ Classificadores treinados manualmente e então combinados utilizando uma regra de fusão qualquer, como por exemplo o Voto Majoritário.
- *Ensembles* (Conjuntos)
 - ▶ Classificadores criados automaticamente através de técnicas de aprendizagem de máquina, como por exemplo *Bagging*, *Boosting*, Seleção de Características.
- Seleção Dinâmica de Classificadores
 - ▶ Consiste em selecionar em função dos outros classificadores aquele que melhor classifica um dado padrão.