

Machine Learning 01

Conteúdo

1. Introdução
2. Aprendizagem computacional
3. Preparação dos dados
4. Regressão
5. Classificação
6. Redes Neurais e Deep Learning

Introdução

Aprender a aprender

O problema de aprendizagem

- O que é um sofá?
- Problema da Netflix
- “Engenharia reversa”
- Aprender para prever

Componentes da Aprendizagem

- Entrada X
- Função Alvo $f: X \rightarrow y$
- Conjunto de Dados D
- Dados de exemplo (X_n, y_n) em D , gerados por f
- Conjunto de hipóteses H
- Algoritmo de aprendizado A
- Função de aproximação $g: X \rightarrow y$ em H , tal que $g \approx f$

Exemplos?

O que não é aprendizagem

- Design de modelo
- Problema das moedas

O problema de aprendizagem:

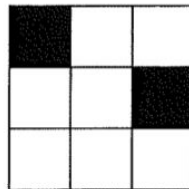
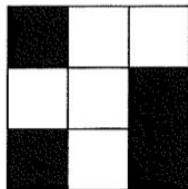
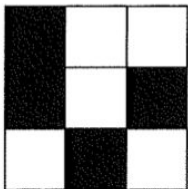
- Precisa de dados
- Precisa que f exista
- Parte de que f é desconhecida

Tipos de aprendizagem

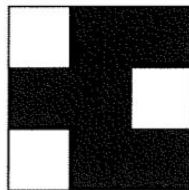
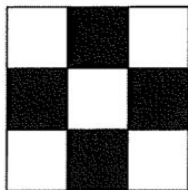
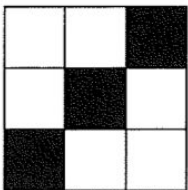
- Aprendizagem Supervisionada
- Aprendizagem por Reforço
- Aprendizagem Não Supervisionada

Tipos de aprendizagem

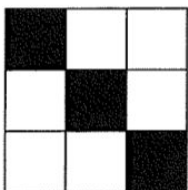
- Aprendizagem Supervisionada
- Aprendizagem por Reforço
- Aprendizagem Não Supervisionada



$f = -1$



$f = +1$



$f = ?$

É possível aprender?

- Se f existe
- Se há dados suficientes para $\text{supor } g \approx f$
- Deve haver informação extraída em D que seja útil FORA de D

Componentes da aprendizagem

Não temos provas, mas temos convicções

O problema de generalização

- Probabilidade de um evento
- Probabilidade em população e amostra
- Erro \times nº de exemplos
- Saber uma probabilidade é melhor do que não saber nada

O problema de generalização

- Dados de treinamento e dados de teste
- Os dados de treinamento e os dados de teste obedecem a uma mesma distribuição (também desconhecida para o problema de aprendizagem)

Componentes da Aprendizagem

- Entrada X
- Função Alvo $f: X \rightarrow y$
- Conjunto de Dados D
- Dados de exemplo (X_n, y_n) em D , gerados por f
- Conjunto de hipóteses H
- Algoritmo de aprendizado A
- Função de aproximação $g: X \rightarrow y$ em H , tal que $g \approx f$
- Distribuição dos pontos P

O problema de generalização

- Dados de treinamento e dados de teste
- Os dados de treinamento e os dados de teste obedecem a uma mesma distribuição (também desconhecida para o problema de aprendizagem)
- A entrada do algoritmo de aprendizagem é um espaço de hipóteses, e não uma hipótese só
- Existe um erro para as hipóteses em D (E_{in}) e um fora de D (E_{out}).

Componentes da Aprendizagem

- Entrada X
- Função Alvo $f: X \rightarrow y$
- Conjunto de Dados D
- Dados de exemplo (X_n, y_n) em D , gerados por f
- Conjunto de hipóteses H
- Algoritmo de aprendizado A
- Função de aproximação $g: X \rightarrow y$ em H , tal que $g \approx f$
- Distribuição dos pontos P
- Erro da hipótese em D E_{in}
- Erro da hipótese fora de D E_{out}

Trabalhando fora dos dados

- Dados em D e dados fora de D
- Os dados em D e os dados fora de D obedecem a uma mesma distribuição (também desconhecida para o problema de aprendizagem)
- A entrada do algoritmo de aprendizagem é um espaço de hipóteses, e não uma hipótese só
- Existe um erro para as hipóteses em D (E_{in}) e um fora de D (E_{out}).
- Uma hipótese é boa se $E_{out} \approx 0$.
- MAS

Um problema pode ser considerada boa sem que E_{out} seja próximo de zero?
Qual? Como?

Complexidade das funções

- Quão complexo deve ser o conjunto de hipóteses H ?
- E a função g ?
- Qual a relação da complexidade da função f com o problema de aprendizado?

ERROR:

object 'slide18' has no attribute 'title'

- Erros comuns sem qualquer aplicação

Erro na aproximação de f

- Quanto g aproxima f ?
- É necessário que haja uma medida
- A mesma medida não serve para todas as instâncias de todos os problemas.

Exemplo: o banco e o supermercado

- Como saber o erro sem saber qual é a função f ?

Erros nos dados: ruído

- Os dados podem ser ambíguos
- Os dados podem ser inexatos
- Nossa função f se torna uma distribuição de probabilidade. Ao invés de $f(X) = y$, temos $P(y | X)$.
- Isso não é um problema para o modelo, mas com certeza é um problema para a aproximação do erro.

Componentes da Aprendizagem

- Entrada X
- Função Alvo $f: X \rightarrow y$
- Conjunto de Dados D
- Dados de exemplo (X_n, y_n) em D , gerados por f
- Conjunto de hipóteses H
- Algoritmo de aprendizado A
- Função de aproximação $g: X \rightarrow y$ em H , tal que $g \approx f$
- Distribuição dos pontos P
- Erro da hipótese em D E_{in}
- Erro da hipótese fora de D E_{out}
- Erro de aproximação entre f e g
- Ruído nos dados de exemplo