

Introdução à Redes Neurais

Panorâma, conceitos e prática

Enzo L. Fernandes

Universidade Estadual Paulista – *Campus* Botucatu

November 8, 2021

1 Aprendizado de Máquina

Motivação

2 Conceitos

Pipeline de Projetos utilizando Aprendizado de Máquina

3 Redes Neurais

Parte 1: Aprendizado de Máquina

Demanda e Interesse

Aplicações

Parte 2: Conceitos

Exemplo

Deseja-se *classificar* uma planta a partir de uma imagem em duas classes: margarida ou girassol.

Abordagens possíveis:

- Criar um conjunto de regras
- Contratar especialistas
- ... ?
- Machine Learning!



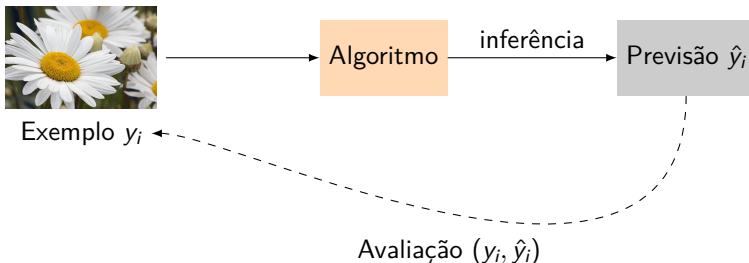
(a) Margarida



(b) Girassol

Mas então, como isso funciona?

Exemplo com uma instância de treinamento:



$$\text{Avaliação } (y_i, \hat{y}_i) = \begin{cases} y_i = y, & \text{✓} \\ y_i \neq y, & \text{✗} \rightarrow \text{aplicar correções} \dots \end{cases}$$

O que é Aprendizado de Máquina?

“Um programa é dito aprender a partir da **experiência E**, em relação a uma classe de **tarefas T**, com medida de **desempenho P**, se seu desempenho em **T**, medido por **P**, melhora com **E**.”

– Tom Mitchell

No exemplo do Girassol/Margarida teremos que:

E: imagens de flores, cada qual com o respectivo rótulo;

T: associar corretamente cada imagem à respectiva espécie;

P: quantidade de imagens classificadas corretamente.

Exemplo

Deseja-se *estimar o preço* de uma casa a partir de informações obtidas na tabela abaixo:

Metros Quadr.	Andares	n° Banheiros	Esquina	DDD	Preço (K R\$)
210	2	4	"Nao"	15	350
315	2	5	"Sim"	11	515
120	1	2	"Nao"	14	250
80	1	1	"Nao"	11	180
900	3	7	"Sim"	11	850
245	2	3	"Sim"	14	450
215	1	3	"Sim"	11	310
190	1	2	"Não"	15	220

Exemplo

Deseja-se *estimar o preço* de uma casa a partir de informações obtidas na tabela abaixo:

Metros Quadr.	Andares	n° Banheiros	Esquina	DDD	Preço (K R\$)
210	2	4	"Nao"	15	350
315	2	5	"Sim"	11	515
120	1	2	"Nao"	14	250
80	1	1	"Nao"	11	180
900	3	7	"Sim"	11	850
245	2	3	"Sim"	14	450
215	1	3	"Sim"	11	310
190	1	2	"Não"	15	220

E: informações sobre características de residências;

T: estimar o valor de uma residência dadas as suas características;

P: erro entre o valor estimado e o valor real da residência.

Exemplo

Notação

Cada **coluna** da tabela corresponde a uma **atributo** ou *feature*.

Metros Quadr.	Andares	n° Banheiros	Esquina	DDD	Preço (K R\$)
210	2	4	"Nao"	15	350
315	2	5	"Sim"	11	515
120	1	2	"Nao"	14	250
80	1	1	"Nao"	11	180
900	3	7	"Sim"	11	850
245	2	3	"Sim"	14	450
215	1	3	"Sim"	11	310
190	1	2	"Não"	15	220

Exemplo

Notação

Cada **linha** da tabela corresponde a uma **amostra** (ou instância de treinamento).

Metros Quadr.	Andares	n° Banheiros	Esquina	DDD	Preço (K R\$)
210	2	4	"Nao"	15	350
315	2	5	"Sim"	11	515
120	1	2	"Nao"	14	250
80	1	1	"Nao"	11	180
900	3	7	"Sim"	11	850
245	2	3	"Sim"	14	450
215	1	3	"Sim"	11	310
190	1	2	"Não"	15	220

Exemplo

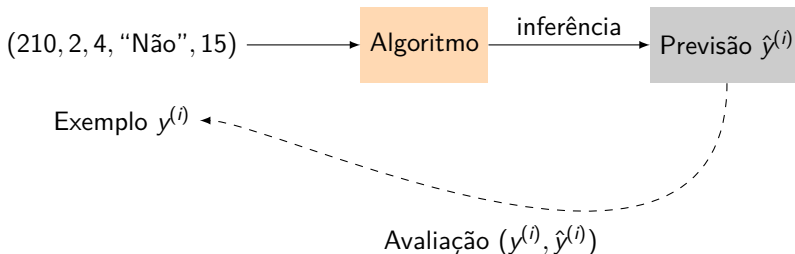
Notação

Atributo Meta (ou *target*) é o atributo que se deseja prever durante o treinamento dos modelos.

Metros Quadr.	Andares	n° Banheiros	Esquina	DDD	Preço (K R\$)
210	2	4	"Nao"	15	350
315	2	5	"Sim"	11	515
120	1	2	"Nao"	14	250
80	1	1	"Nao"	11	180
900	3	7	"Sim"	11	850
245	2	3	"Sim"	14	450
215	1	3	"Sim"	11	310
190	1	2	"Não"	15	220

Exemplo

Exemplo com a primeira instância de treinamento da tabela de dados:



$$\text{Avaliação } (\mathbf{y}, \hat{\mathbf{y}}) = \text{MSE } (\mathbf{y}, \hat{\mathbf{y}}) = \sum_{i=1}^m \left(\hat{y}^{(i)} - y^{(i)} \right)^2$$

Formalismos: Entrada e Saída

Os dados obtidos obtidos *a priori* – também chamados de *dataset* – correspondem à:

- \mathbf{x} : matriz $m \times n$ em que cada linha corresponde a uma amostra de treinamento e cada coluna um atributo.
- \mathbf{y} : matriz $m \times k$ em que cada linha corresponde a um valor esperado (ou valores esperados) para cada linha de \mathbf{x} .

Ou seja, $h_{\theta}(\cdot)$ será tal que:

$$h_{\theta}: \mathbb{R}^n \rightarrow \mathbb{R}^k$$

Formalismos: Entrada e Saída

Em cada matriz \mathbf{x} e \mathbf{y} temos elementos $e_j^{(i)}$ nos quais:

- i corresponde à i -ésima amostra do conjunto e;
- j corresponde à j -ésimo atributo da amostra.

Assim temos:

$$\mathbf{x} = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & \cdots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} & \cdots & x_n^{(2)} \\ & \vdots & & \ddots & \vdots \\ x_1^{(m)} & x_2^{(m)} & x_3^{(m)} & \cdots & x_n^{(m)} \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} y_1^{(1)} & \cdots & y_k^{(1)} \\ y_1^{(2)} & \cdots & y_k^{(2)} \\ \vdots & \ddots & \vdots \\ y_1^{(m)} & \cdots & y_k^{(m)} \end{bmatrix}$$

Formalismos: Treinamento

O processo de treinamento é feito ajustando os valores dos *parametros internos* θ em função de uma **função de custo** C que compara os valores previstos \hat{y} e os valores esperados y .

C depende do algoritmo utilizado e do tipo de tarefa realizada.

- classificação – e.g. Erro Quadrático Médio
- regressão – e.g. Entropia Cruzada

Desta forma o treinamento consiste em otimizar os valores dos parâmetros internos θ de modo a minimizar C :

$$\min_{\theta} C(\theta)$$

Tipos de Aprendizado

Os tipos de aprendizado podem ser divididos em:

- **Aprendizagem Supervisionada**
 - Regressão
 - Classificação
- **Aprendizagem Não-Supervisionada**
 - Redução de Dimensionalidade
 - Análise de Agrupamento
- **Aprendizagem por Reforço**
 - Agentes Autônomos

Parte 3: Redes Neurais