

**iniciar nova gravação?**

# Aprendizado de Máquina Supervisionado

## Modelo para Regressão



# Por onde começamos?

Começaremos os nossos estudos em **Modelos de Regressão** analisando o caso mais simples possível: a **regressão linear com uma variável**

Começaremos os nossos estudos em **Modelos de Regressão** analisando o caso mais simples possível: a **regressão linear com uma variável**

- A regressão linear com uma variável consiste em encontrar uma reta que modela o comportamento dos dados.

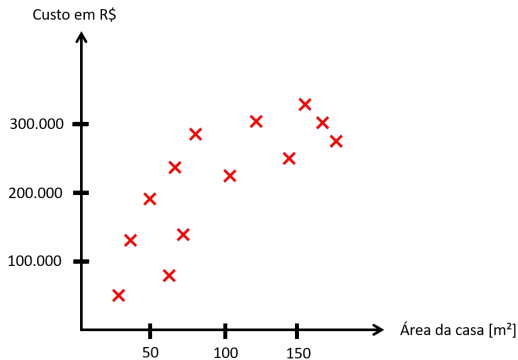
Começaremos os nossos estudos em **Modelos de Regressão** analisando o caso mais simples possível: a **regressão linear com uma variável**

- A regressão linear com uma variável consiste em encontrar uma reta que modela o comportamento dos dados.
- Apesar de simples, tal modelagem é vastamente utilizada e pode sim ser considerado um tipo algoritmo de ML, uma vez que o modelo é diretamente treinado a partir de dados.

Começaremos os nossos estudos em **Modelos de Regressão** analisando o caso mais simples possível: a **regressão linear com uma variável**

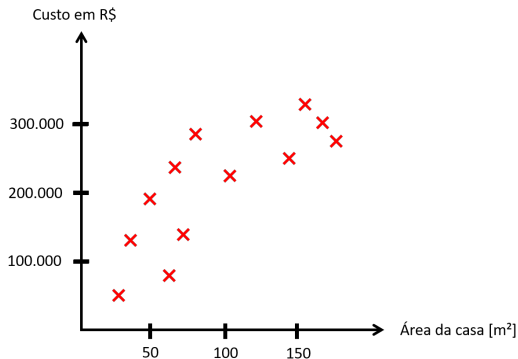
- A regressão linear com uma variável consiste em encontrar uma reta que modela o comportamento dos dados.
- Apesar de simples, tal modelagem é vastamente utilizada e pode sim ser considerar um tipo algoritmo de ML, uma vez que o modelo é diretamente treinado a partir de dados.
- Muitos conceitos que serão vistos aqui serão também aplicados nos algoritmos mais avançados vistos futuramente na disciplina

## Exemplo



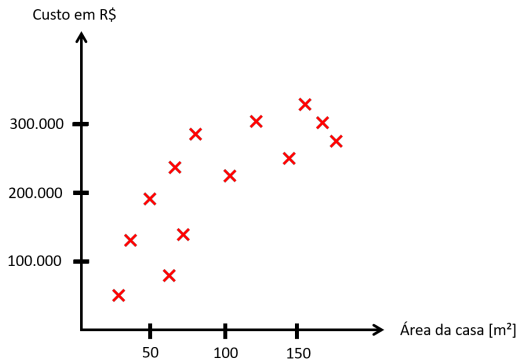
- Cada X na figura acima representa uma casa com uma certa área  $x_i$  e que foi vendida por um valor  $y_i$

## Exemplo



- Cada X na figura acima representa uma casa com uma certa área  $x_i$  e que foi vendida por um valor  $y_i$
- Imagine agora que um amigo seu quer vender uma casa na mesma região. A casa possui  $100 m^2$ . Quanto ele pode cobrar pelo imóvel?





- Cada X na figura acima representa uma casa com uma certa área  $x_i$  e que foi vendida por um valor  $y_i$
- Imagine agora que um amigo seu quer vender uma casa na mesma região. A casa possui  $100 m^2$ . Quanto ele pode cobrar pelo imóvel?
- Usando um algoritmo de regressão, é possível ajustar uma reta para esses dados. Essa reta poderá auxiliar o seu amigo.

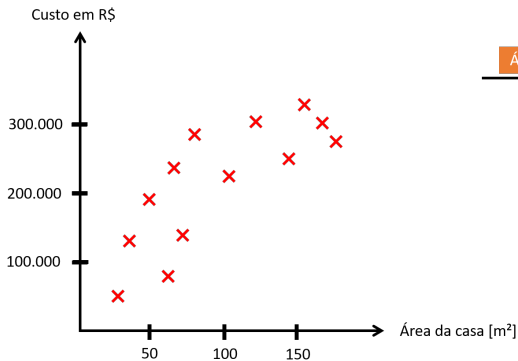


Tabela com os dados

Área da casa [m²]	Custo em R\$
32	51.000
149	265.000
78	110.000
...	...
220	315.000

Pergunta:

Qual será o número de linhas da tabela?

Aproveitaremos esse exemplo para introduzirmos alguns termos comumente usados na área de Aprendizado de Máquina.

Conjunto de dados de treinamento:

Corresponde aos dados que serão usados para treinar o modelo. Por exemplo, a tabela toda abaixo.

**Tabela com os dados**

Área da casa [m <sup>2</sup> ]	Custo em R\$
32	51.000
149	265.000
78	110.000
...	...
220	315.000

Observação

Note que a casa do seu amigo não está na tabela acima. Isso porque ela ainda não foi vendida e ninguém sabe por quanto ela será de fato vendida.

## Variável de entrada $x$

É a primeira coluna da tabela. Também chamada de **característica**, ou, **característica de entrada**. Por exemplo, para uma das amostras de treinamento da tabela, temos

$$x = 149$$

## Tabela com os dados

Área da casa [m <sup>2</sup> ]	Custo em R\$
32	51.000
149	265.000
78	110.000
...	...
220	315.000

## Variável de saída $y$

É a segunda coluna da tabela. Também chamada de **variável alvo** ou **variável de saída**. Por exemplo, para a segunda amostra de treinamento da tabela (onde  $x = 149$ ), temos

$$y = 265.000$$

## Tabela com os dados

Área da casa [m <sup>2</sup> ]	Custo em R\$
32	51.000
149	265.000
78	110.000
...	...
220	315.000

## Pergunta:

O número total de amostras de treinamento " $m$ " é igual ao número de linhas da tabela?

## Amostras de treinamento

Usaremos a notação  $(x^{(i)}, y^{(i)})$  para nos referirmos à  $i$ -ésima amostra de treinamento ( $i$ -ésima linha da tabela). Por exemplo, a 2ª amostra de treinamento da tabela é

$$(x^{(2)}, y^{(2)}) = (149; 265.000)$$

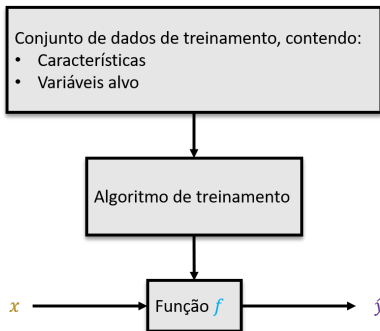
## Tabela com os dados

Área da casa [m²]	Custo em R\$
32	51.000
149	265.000
78	110.000
...	...
220	315.000

## Observação

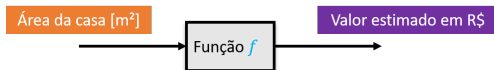
A nossa representação não significa exponenciação. Ou seja,  $x^{(2)} \neq x^2$

## Aprendizado Supervisionado



- A função  $f$  é também chamada de **modelo**
- Dada uma variável de entrada  $x$ , a variável  $\hat{y}$  representa uma previsão (ou, estimativa) para a variável alvo  $y$ .





Pergunta:

Como representar a função  $f$ ? Qual é a fórmula matemática que  $f$  irá assumir?

Resposta:



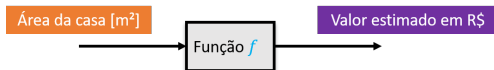
Pergunta:

Como representar a função  $f$ ? Qual é a fórmula matemática que  $f$  irá assumir?

Resposta:

Supondo que queremos uma reta, podemos definir

$$f_{w,b}(x) = wx + b$$



Pergunta:

Como representar a função  $f$ ? Qual é a fórmula matemática que  $f$  irá assumir?

Resposta:

Supondo que queremos uma reta, podemos definir

$$f_{w,b}(x) = wx + b$$

- Primeiro, precisaremos definir quanto valerá  $w$  e quanto valerá  $b$ .



## Pergunta:

Como representar a função  $f$ ? Qual é a fórmula matemática que  $f$  irá assumir?

## Resposta:

Supondo que queremos uma reta, podemos definir

$$f_{w,b}(x) = wx + b$$

- Primeiro, precisaremos definir quanto valerá  $w$  e quanto valerá  $b$ .
- Assim, para um dado valor para  $x$ , a função  $f_{w,b}(x) = wx + b$  poderá ser computada.

# Exemplo



## Pergunta:

Como representar a função  $f$ ? Qual é a fórmula matemática que  $f$  irá assumir?

## Resposta:

Supondo que queremos uma reta, podemos definir

$$f_{w,b}(x) = wx + b$$

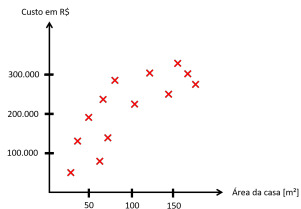
- Primeiro, precisaremos definir quanto valerá  $w$  e quanto valerá  $b$ .
- Assim, para um dado valor para  $x$ , a função  $f_{w,b}(x) = wx + b$  poderá ser computada.
- Para  $w$  e  $b$  escolhidos de forma apropriada,  $f_{w,b}(x)$  nos fornecerá um valor apropriado para  $\hat{y}$

## Simplificando a notação:

Em alguns casos, usaremos a notação simplificada

$$f(x) = wx + b$$

Podemos aplicar **regressão linear com uma única variável** quando temos um conjunto de dados conforme abaixo



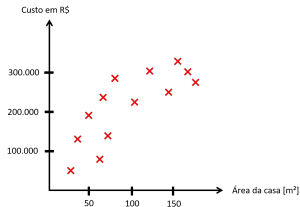
Isso significa que encontraremos um modelo do tipo

$$f(x) = wx + b$$

para esses dados.

## Observações

Podemos aplicar **regressão linear com uma única variável** quando temos um conjunto de dados conforme abaixo



Isso significa que encontraremos um modelo do tipo

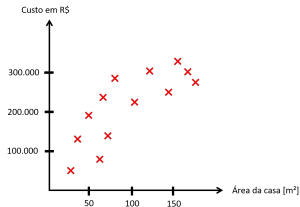
$$f(x) = wx + b$$

para esses dados.

## Observações

- Se chama Regressão pois  $y$  pode assumir infinitos valores possíveis

Podemos aplicar **regressão linear com uma única variável** quando temos um conjunto de dados conforme abaixo



Isso significa que encontraremos um modelo do tipo

$$f(x) = wx + b$$

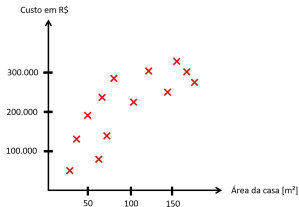
para esses dados.

## Observações

- Se chama Regressão pois  $y$  pode assumir infinitos valores possíveis
- É linear pois não temos termos não lineares em relação aos parâmetros do modelo, ou seja, termos do tipo  $b^2$ ,  $w^3$ ,  $bw$ ,  $\sqrt{b}$ , etc.



Podemos aplicar **regressão linear com uma única variável** quando temos um conjunto de dados conforme abaixo



Isso significa que encontraremos um modelo do tipo

$$f(x) = wx + b$$

para esses dados.

## Observações

- Se chama Regressão pois  $y$  pode assumir infinitos valores possíveis
- É linear pois não temos termos não lineares em relação aos parâmetros do modelo, ou seja, termos do tipo  $b^2$ ,  $w^3$ ,  $bw$ ,  $\sqrt{b}$ , etc.
- É uma regressão baseada em uma única variável pois temos apenas uma variável de entrada  $x$  (área da casa)

**De olho no código!**

Vamos agora ver no código como podemos tentar ajustar uma reta para um determinado conjunto de dados.

Clique no link abaixo para acessar o código:

```
https://colab.research.google.com/github/xaximpvp2/master/blob/main/codigo\_aula3\_representacao\_  
do\_modelo.ipynb
```

## Parte 1

Rode todo o código. Responda às questões nele contidas e complete-o, se necessário.

## Parte 2

- 1) Pensando em como seria um conjunto de medições para um resistor de  $50\ \Omega$ , crie um conjunto de dados contendo 4 amostras de tensão *versus* corrente.
- 2) Insira esses dados nos códigos da Parte 1, faça as adaptações necessárias e verifique os novos resultados.