

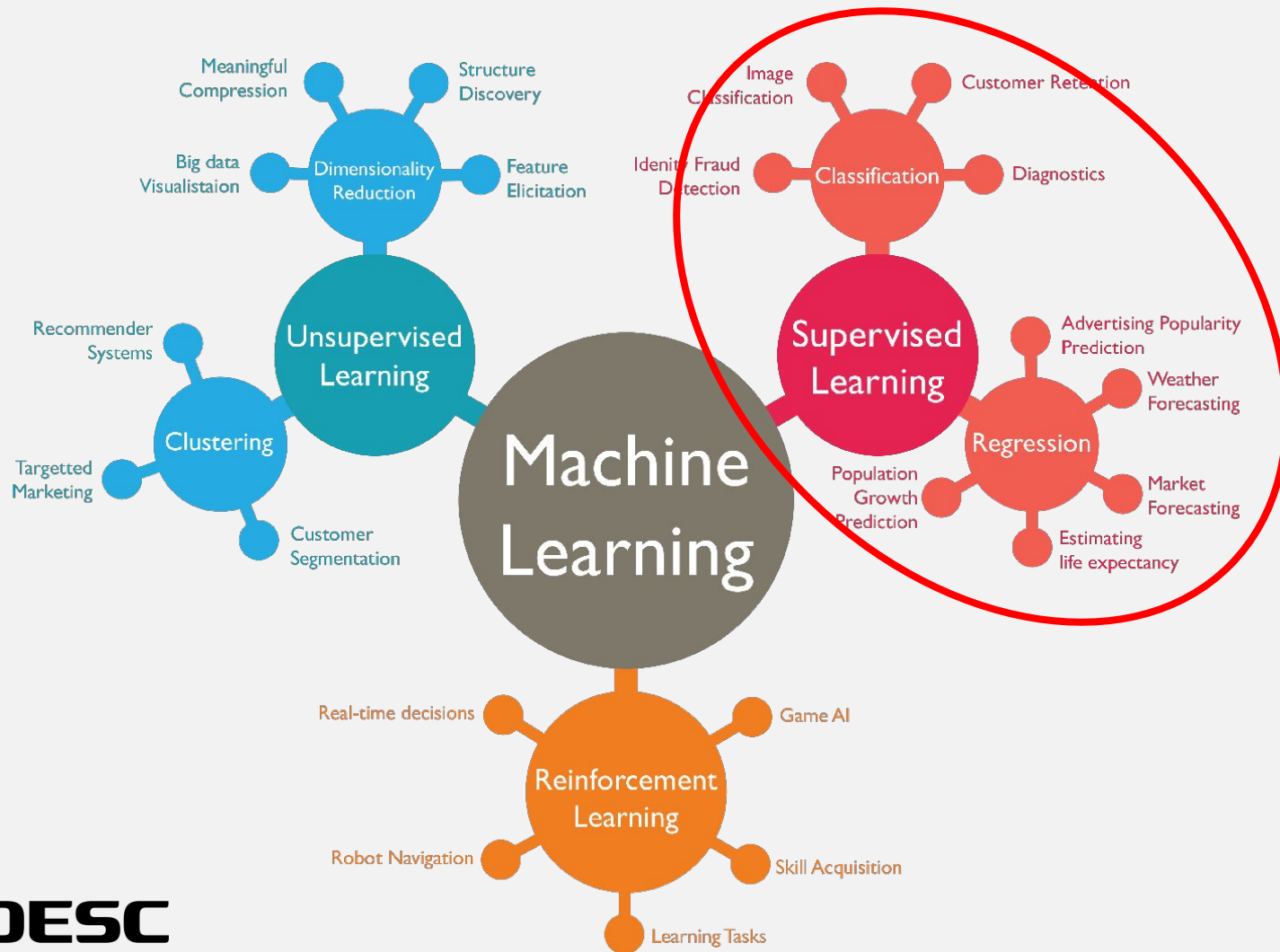
# Modelos de Aprendizagem Supervisionada

Definições e exemplos de modelagens

Douglas Macedo Sgrott  
Orientador: Rafael Parpinelli  
31/05/2021

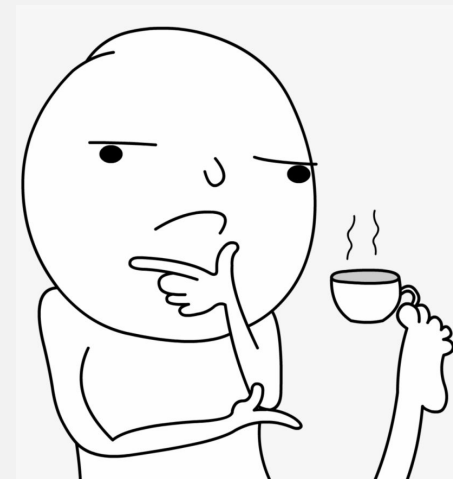
# Aprendizagem Supervisionada

- Uma sub-área da Aprendizagem de Máquina (Machine Learning)



# ■ Aprendizagem Supervisionada

- O que é aprendizagem?



# ■ Aprendizagem Supervisionada

- O que é aprendizagem?
- E por quê essa pergunta é importante?



# **Aprendizagem Supervisionada**

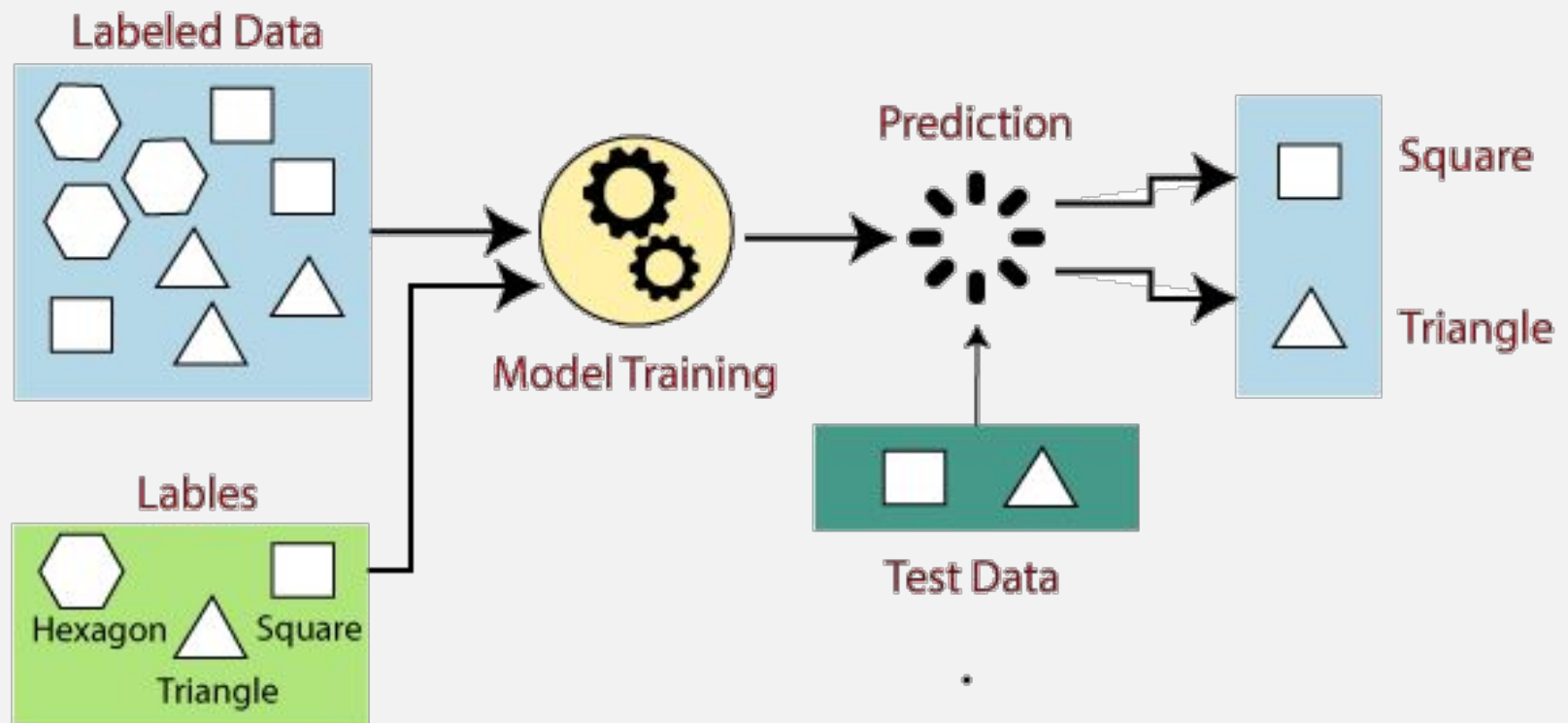
- O que é aprendizagem?
  - Existem várias definições, inclusive:
    - O ato de obter conhecimento (dic. Cambridge)
    - Modificação da tendência de um comportamento através da experiência (dic. Merriam-Webster)

# Aprendizagem Supervisionada

- O que é aprendizagem?
  - Existem várias definições, inclusive:
    - O ato de obter conhecimento (dic. Cambridge)
    - Modificação da tendência de um comportamento através da experiência (dic. Merriam-Webster)
    - “A persisting change in **human performance** or performance potential...[which] must come about as a result of ...” From Psychology of Learning for Instruction by M. Driscoll
    - “A change in **human disposition** or capability that persists over a period of time and ...” — From The Conditions of Learning by Robert Gagne
    - “Learning is the relatively permanent change in a **person’s** knowledge or behavior due to ... —From Learning in Encyclopedia of Educational Research, Richard E. Mayer

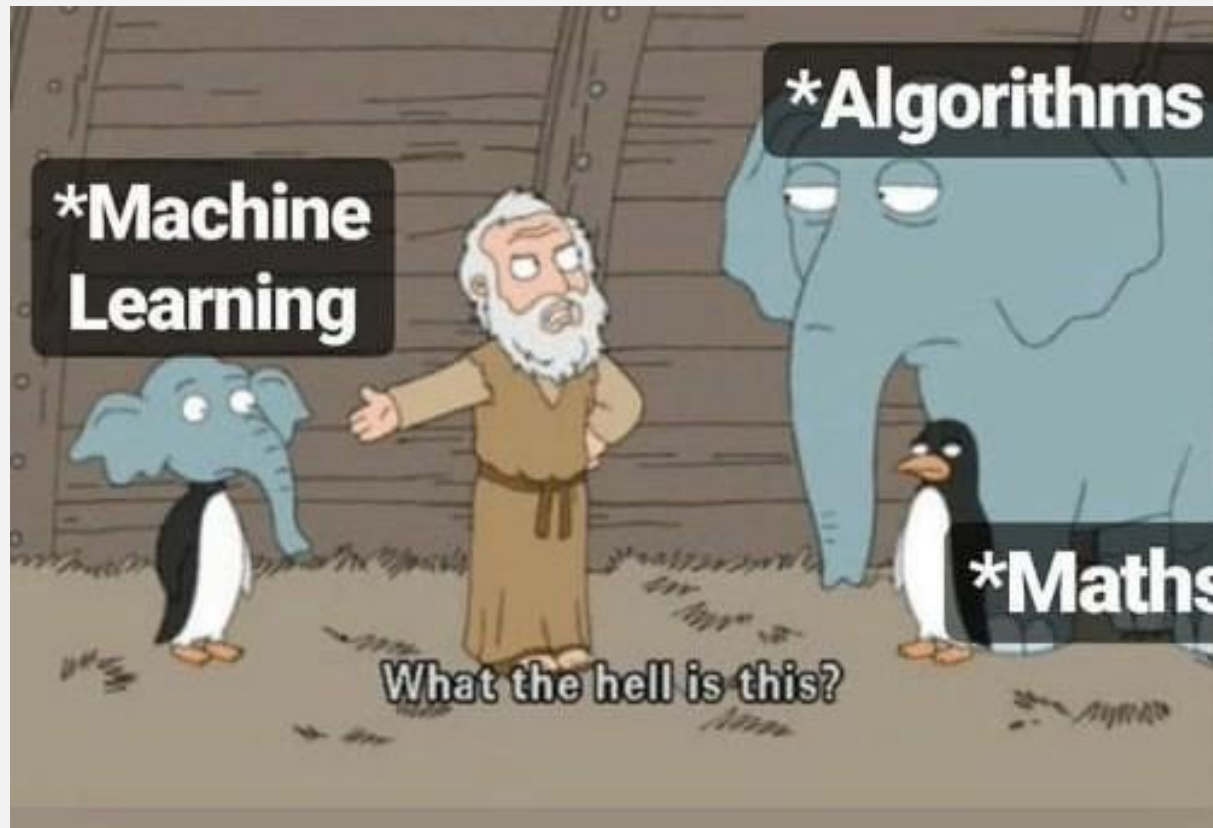
# Aprendizagem Supervisionada

- Consiste em “aprender” uma função que mapeia dados de entrada em relação a saída através de **exemplos (labels ou rótulos)**



# Alguns esclarecimentos...

- Machine Learning NÃO é novidade.
- Machine Learning é fruto do avanço de várias outras áreas.
- O amadurecimento de ML vem de uma demanda por algoritmos de modelagem mais sofisticados.





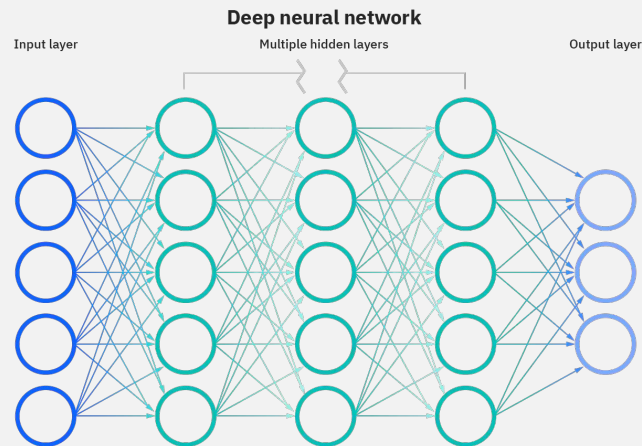
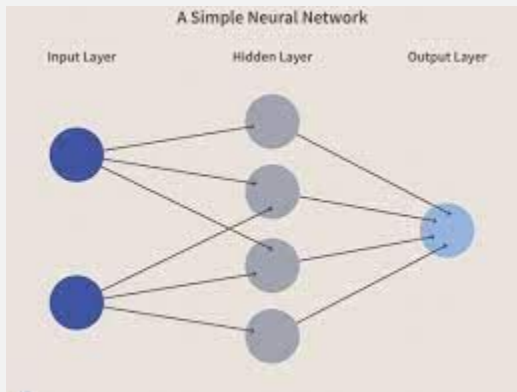
# Alguns esclarecimentos...

- Então por quê se fala tanto de Machine Learning hoje em dia?

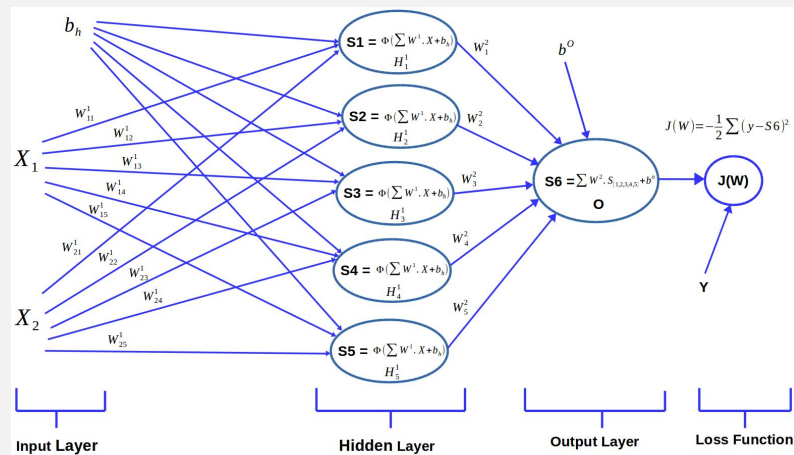
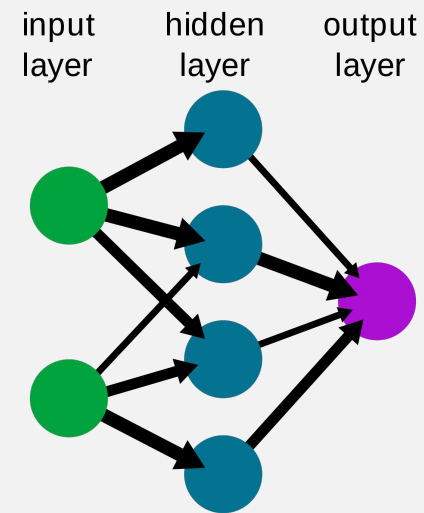


# Alguns esclarecimentos...

- Então por quê se fala tanto de Machine Learning hoje em dia? **Marketing**
- Resultados de uma pesquisa por “Neural Network” no Google

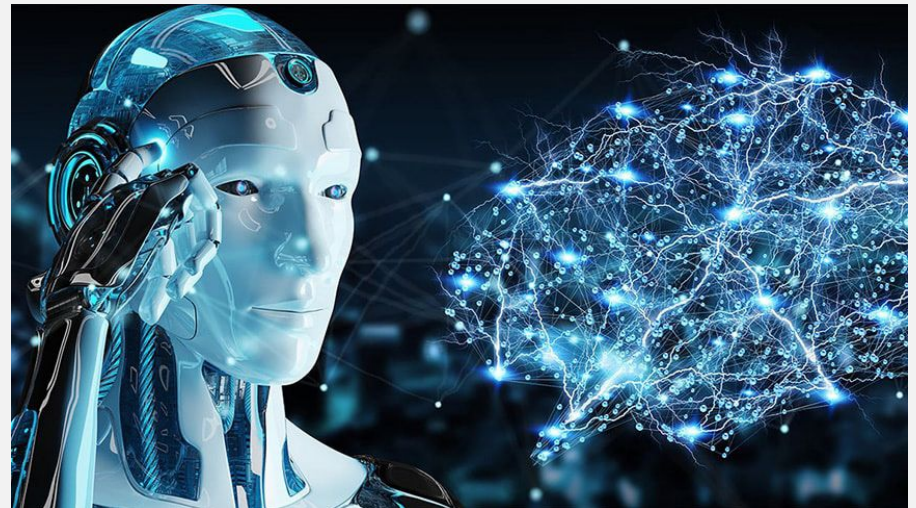
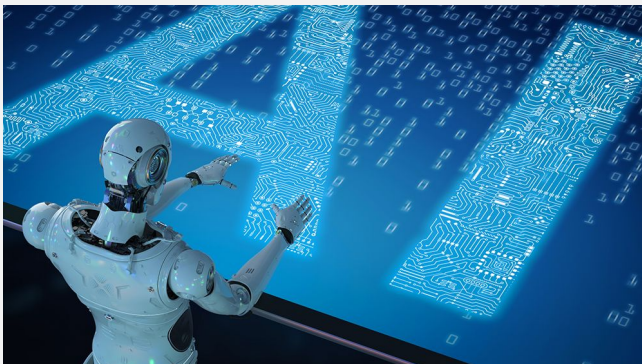
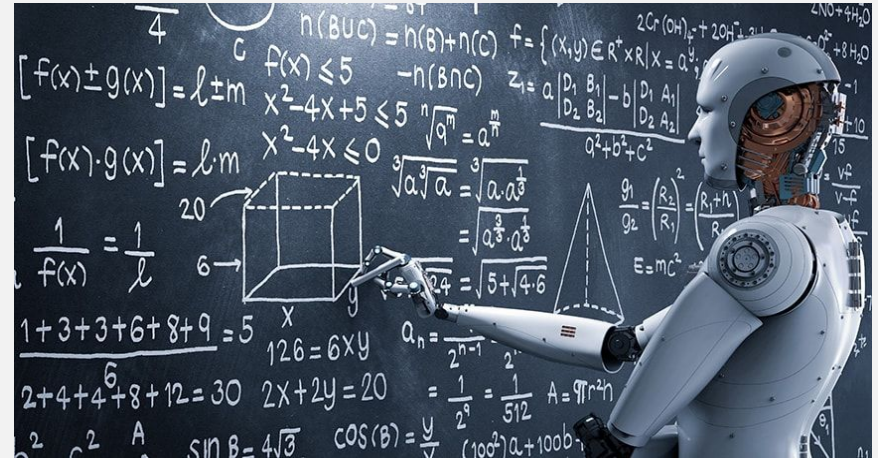


A simple neural network



# Alguns esclarecimentos...

- Então por quê se fala tanto de Machine Learning hoje em dia? **Marketing**
- Resultados de uma pesquisa por “Machine Learning” no Google





# Alguns esclarecimentos...

- Então por quê se fala tanto de Machine Learning hoje em dia? **Mídia**

<https://www.forbes.com> › [tonybradley](#) › 2017/07/31 › f... ▼

## Facebook AI Creates Its Own Language In Creepy Preview Of ...

Jul 31, 2017 — In a glimpse at what the beginning of the technological singularity might look like, researchers at **Facebook** shut down an artificial intelligence ...

<https://www.snopes.com> › [Fact Checks](#) › [Technology](#) ▼

## FACT CHECK: Did Facebook Shut Down an AI Experiment ...

Aug 1, 2017 — **Facebook** AI Robots Shut Down After They Secretly 'Invent Their Own Language' ... **Facebook** shuts AI system after bots speak **their own language**, defy ... Implanting **neural nets**, within our brains that are connected to it, etc.

# Alguns esclarecimentos...

- Então por quê se fala tanto de Machine Learning hoje em dia? **Mídia**

<https://www.forbes.com> › [tonybradley](#) › 2017/07/31 › f... ▼

## Facebook AI Creates Its Own Language In Creepy Preview Of ...

Jul 31, 2017 — In a glimpse at what the beginning of the technological singularity might look like, researchers at **Facebook** shut down an artificial intelligence ...

<https://www.snopes.com> › [Fact Checks](#) › [Technology](#) ▼

## FACT CHECK: Did Facebook Shut Down an AI Experiment ...

Aug 1, 2017 — **Facebook** AI Robots Shut Down After They Secretly 'Invent Their Own Language' ... **Facebook** shuts AI system after bots speak **their own language**, defy ... Implanting **neural nets**, within our brains that are connected to it, etc.

Bob: "I can can I I everything else."

Alice: "Balls have zero to me to me to me to me to me to me to me to me to."

# Alguns esclarecimentos...

- Então por quê se fala tanto de Machine Learning hoje em dia? **Mídia**

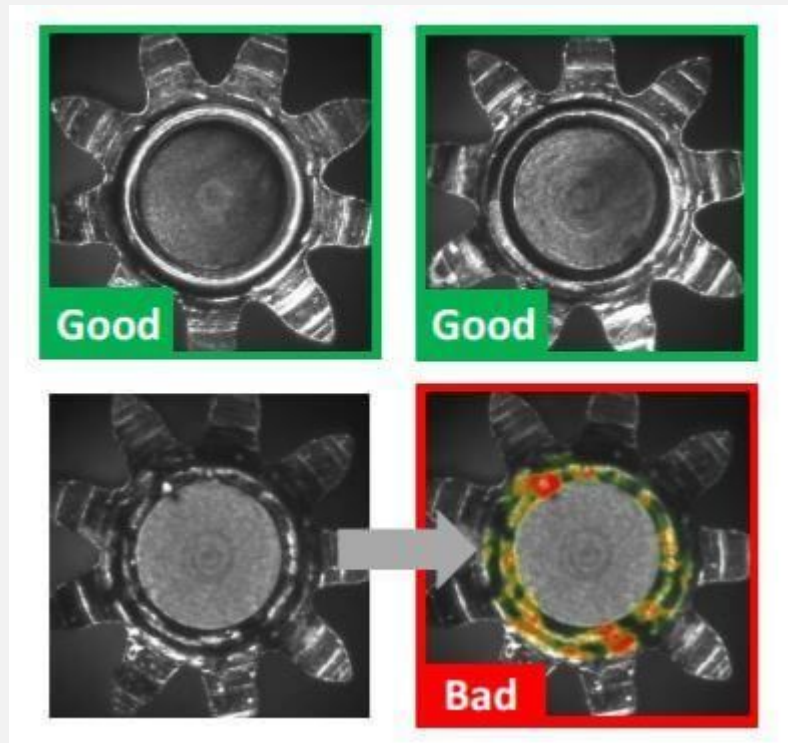
Bob: "I can can I I everything else."

Alice: "Balls have zero to me to me to me to me to me to me to me to me to me to."



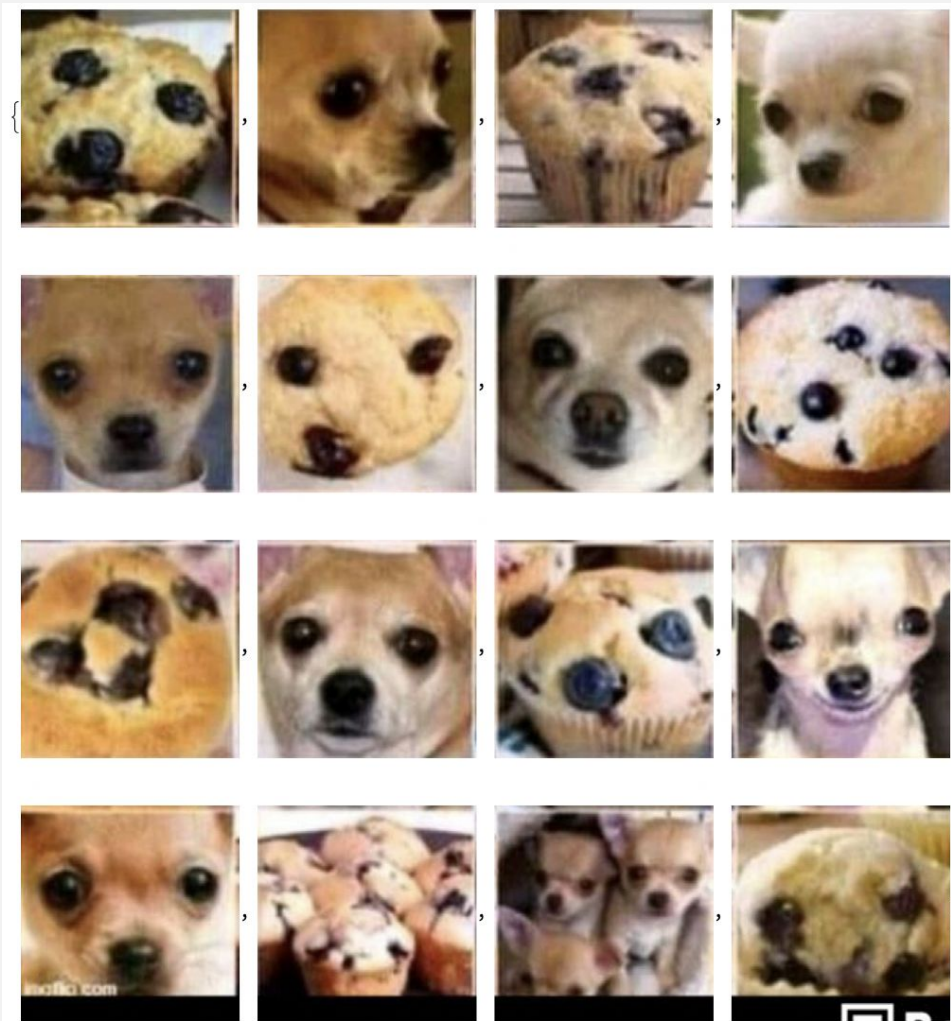
# Alguns esclarecimentos...

- Então por quê se fala tanto de Machine Learning hoje em dia? **Demanda**



# Alguns esclarecimentos...

- Machine Learning não é a solução de todos os problemas.





# Aprendizagem Supervisionada: Modelos

- Um dos principais objetivos é gerar um modelo a partir de dados.
- Existem VÁRIOS tipos de modelos e algoritmos.
- Qual escolher?



# Aprendizagem Supervisionada: Modelos

- Um dos principais objetivos é gerar um modelo a partir de dados.
- Existem VÁRIOS tipos de modelos e algoritmos.
- Qual escolher? Resposta: **No Free Lunch Theorem**



# Aprendizagem Supervisionada: Modelos

- Qual tipo de modelo treinar?
  - DEPENDE!



# Aprendizagem Supervisionada: Modelos

- Qual tipo de modelo treinar?
  - DEPENDE!
- Qual é o tipo de dado?
  - Linear
  - Não linear
- Qual a prioridade?
  - Acurácia
  - Precisão
  - Velocidade de treino
  - Velocidade de inferência
  - Simplicidade
  - **Interpretabilidade**
- Qual a finalidade?
  - Regressão
  - Classificação



# Aprendizagem Supervisionada: Modelos

- Qual tipo de modelo treinar?
  - DEPENDE!
  - Aplicar Navalha de Ockham

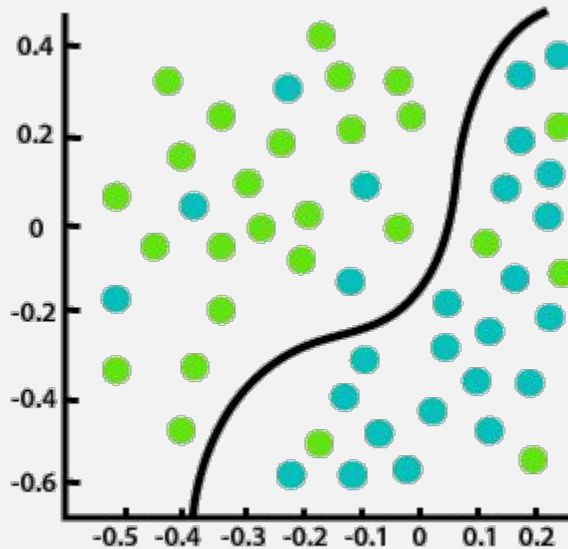
**The virtue of simplicity: On machine learning models in algorithmic trading**

learning-powered trading algorithms. The analysis shows that machine learning quants use Ockham's razor – things should not be multiplied without necessity – as a heuristic tool to prevent excess model complexity and secure a certain level of human control and interpretability in the modelling process. I argue that

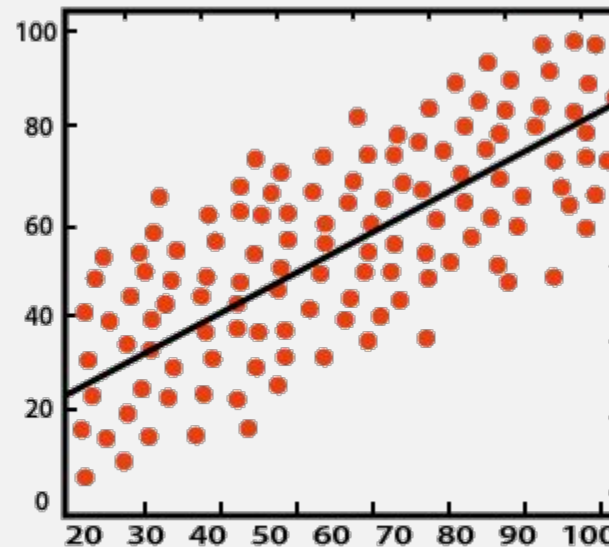


# Aprendizagem Supervisionada: Tipos

- Se divide principalmente em 2 tipos de aplicações:
  - Regressão
  - Classificação



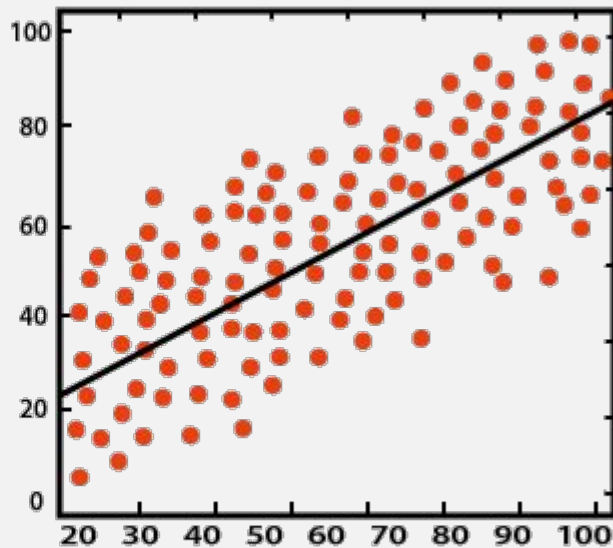
Classification



Regression

# Regressão

- Objetivo: Gerar um modelo (analítico ou não) que melhor represente a distribuição de dados contínuos.
- Realizar previsões numéricas sobre algum fenômeno.

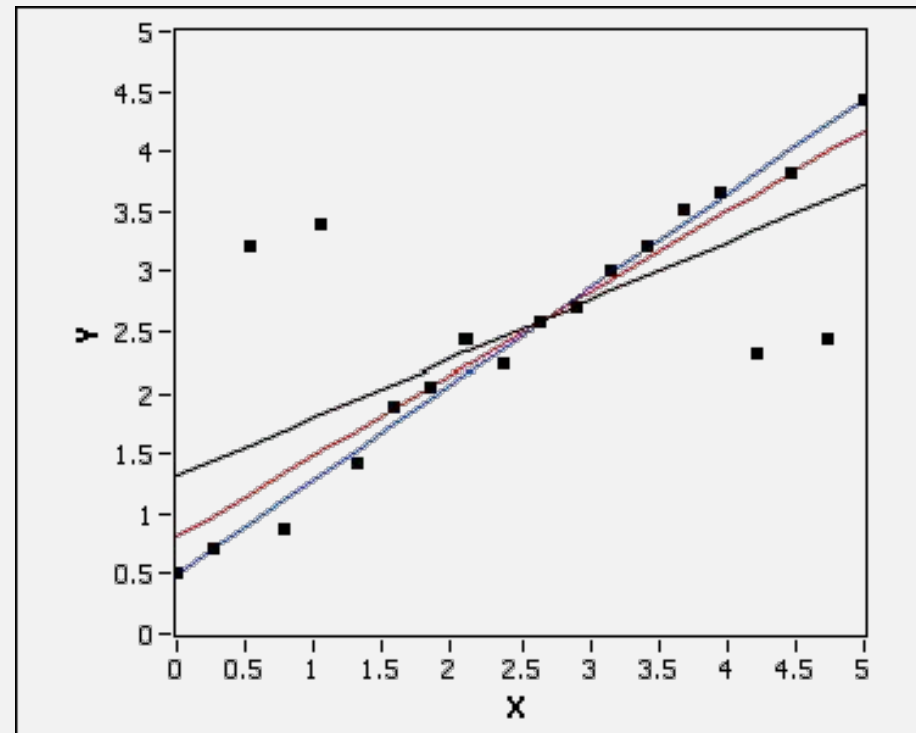
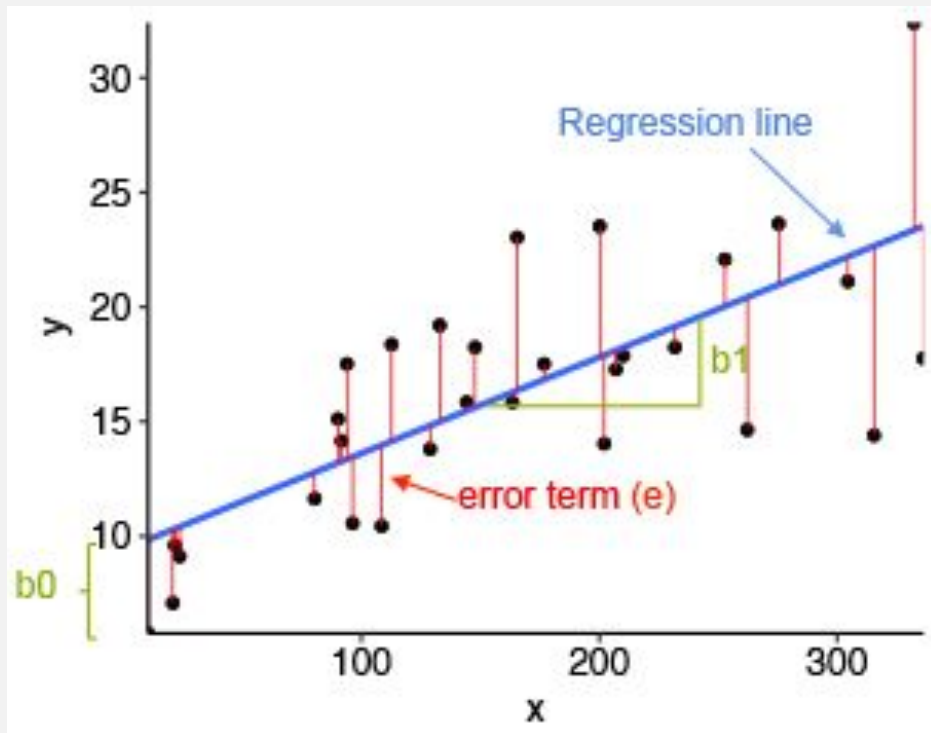


Regression

G Gasolina	5,199
G PETROBRAS GRID	5,299
E Etanol	3,999
G PETROBRAS Radiump	6,290
Diesel S10	4,199

# Regressão Linear

- **Ordinary Least Squares (OLS)**
  - Qual a maneira mais simples de ajustar uma reta sobre um conjunto de pontos?





# Regressão Linear

- **Ordinary Least Squares (OLS)**

- Consiste em ajustar uma reta com  $p$  variáveis de entrada ( $x$ ) e  $p+1$  parâmetros

$$\hat{y}(w, x) = w_0 + w_1x_1 + \dots + w_px_p$$

Square Loss Function  
(função de custo)

de modo a minimizar o erro

$$\min_w ||Xw - y||_2^2$$

# Regressão Linear

- **Ordinary Least Squares (OLS)**

- Consiste em ajustar uma reta com  $p$  variáveis de entrada ( $x$ ) e  $p+1$  parâmetros

$$\hat{y}(w, x) = w_0 + w_1x_1 + \dots + w_px_p$$

Square Loss Function  
(função de custo)

de modo a minimizar o erro

$$\min_w ||Xw - y||_2^2$$

- **Ridge Regression**

- Mesma ideia do OLS, mas com um parâmetro  $\alpha \geq 0$  adicional para evitar que os parâmetros assumam valores MUITO altos:

$$\min_w ||Xw - y||_2^2 + \alpha ||w||_2^2$$

- Quanto maior o valor de  $\alpha$ , menor o valor dos parâmetros.

# Regressão Linear

- Ordinary Least Squares

Vantagens	Desvantagens

# Regressão Linear

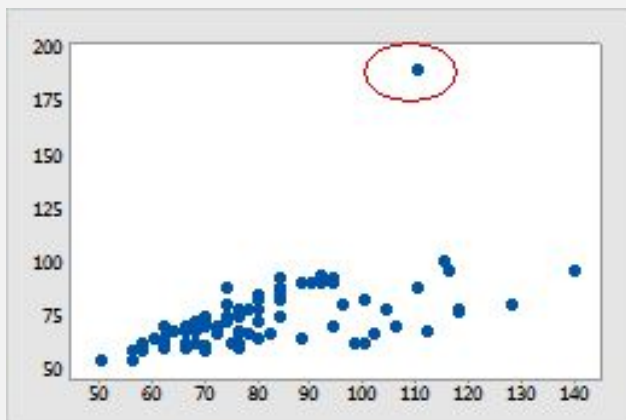
- **Ordinary Least Squares**

Vantagens	Desvantagens
Extremamente simples	
Fácil de interpretar	
Dependendo das condições (Gauss-Markov), produz a melhor regressão linear	

# Regressão Linear

- **Ordinary Least Squares**

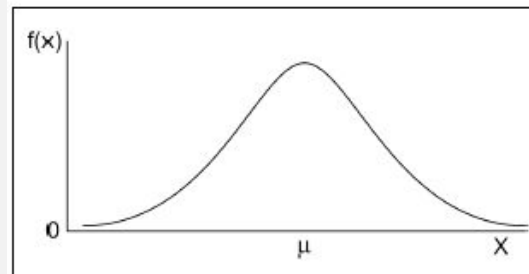
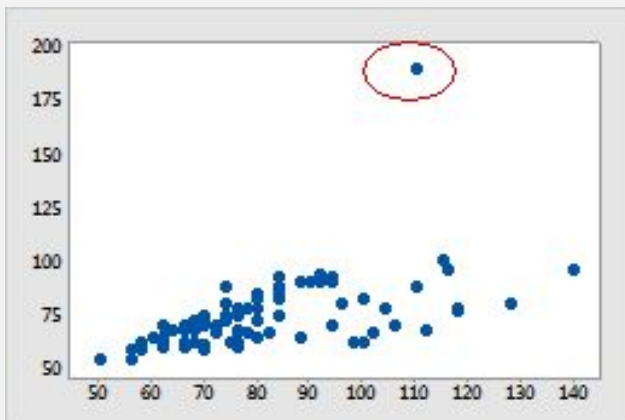
Vantagens	Desvantagens
Extremamente simples	Sensível em relação a outliers
Fácil de interpretar	
Dependendo das condições (Gauss-Markov), produz a melhor regressão linear	



# Regressão Linear

- **Ordinary Least Squares**

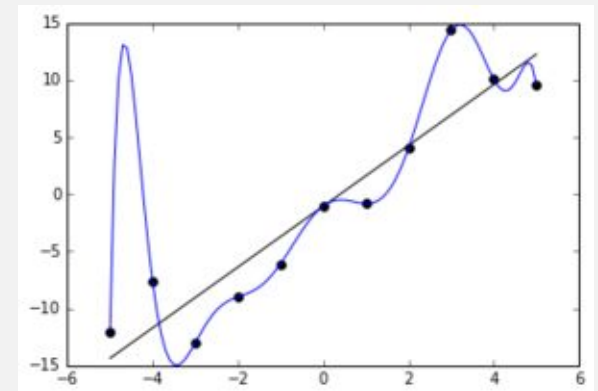
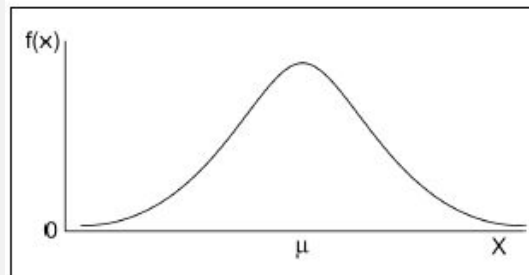
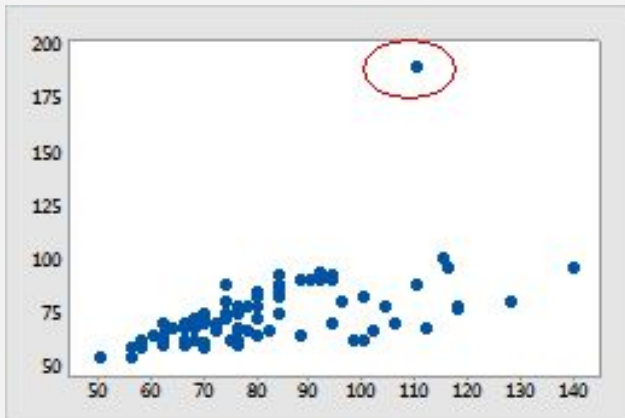
Vantagens	Desvantagens
Extremamente simples	Sensível em relação a outliers
Fácil de interpretar	Pode ter resultados ruins para dados que não seguem uma distribuição normal
Dependendo das condições (Gauss-Markov), produz a melhor regressão linear	



# Regressão Linear

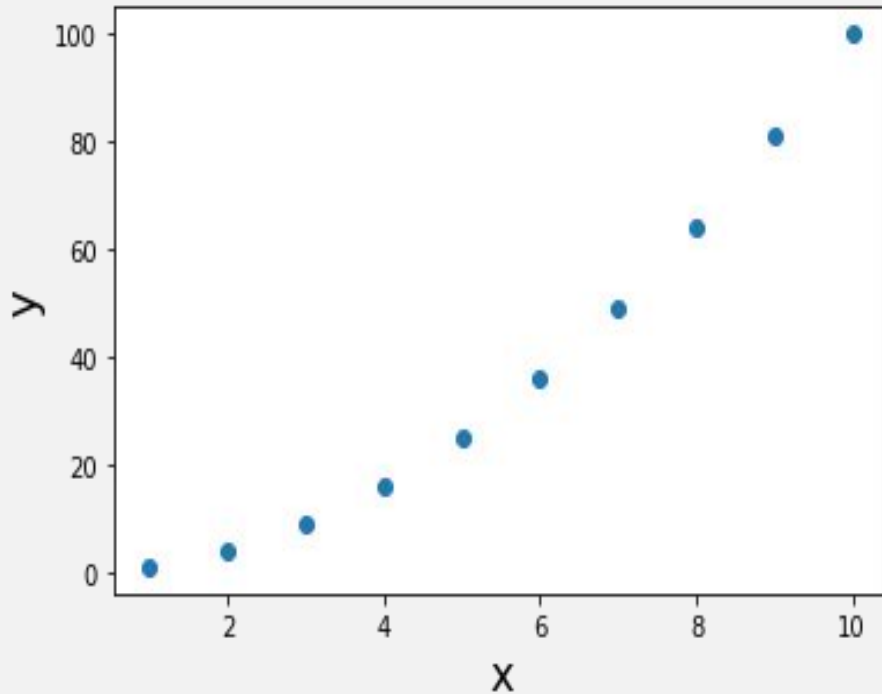
- **Ordinary Least Squares**

Vantagens	Desvantagens
Extremamente simples	Sensível em relação a outliers
Fácil de interpretar	Pode ter resultados ruins para dados que não seguem uma distribuição normal
Dependendo das condições (Gauss-Markov), produz a melhor regressão linear	Tendência de gerar overfit/sobreajuste (Ridge regression pode aliviar este problema)



# Momento filosófico

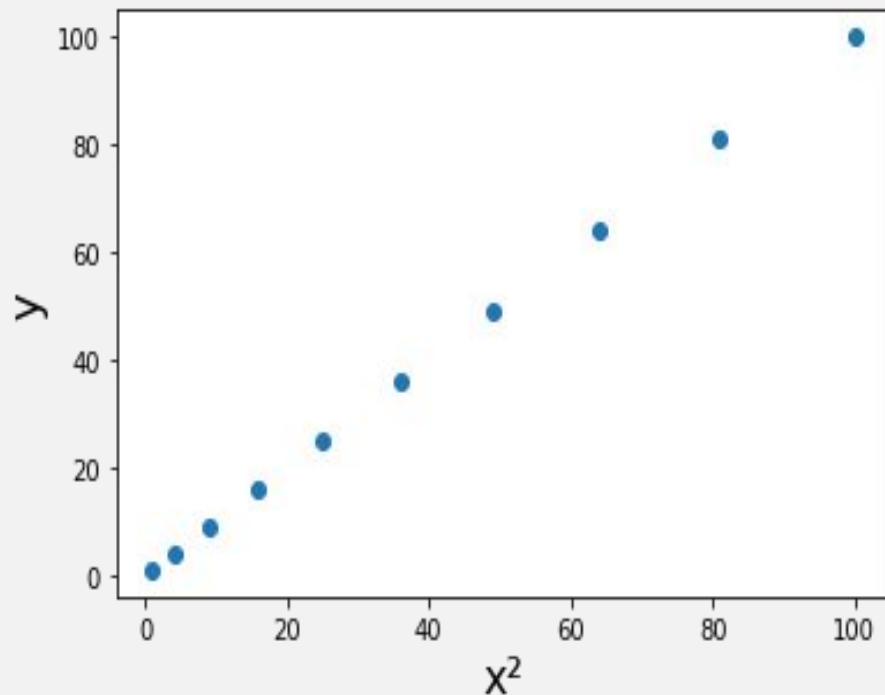
- Estes dados são lineares?





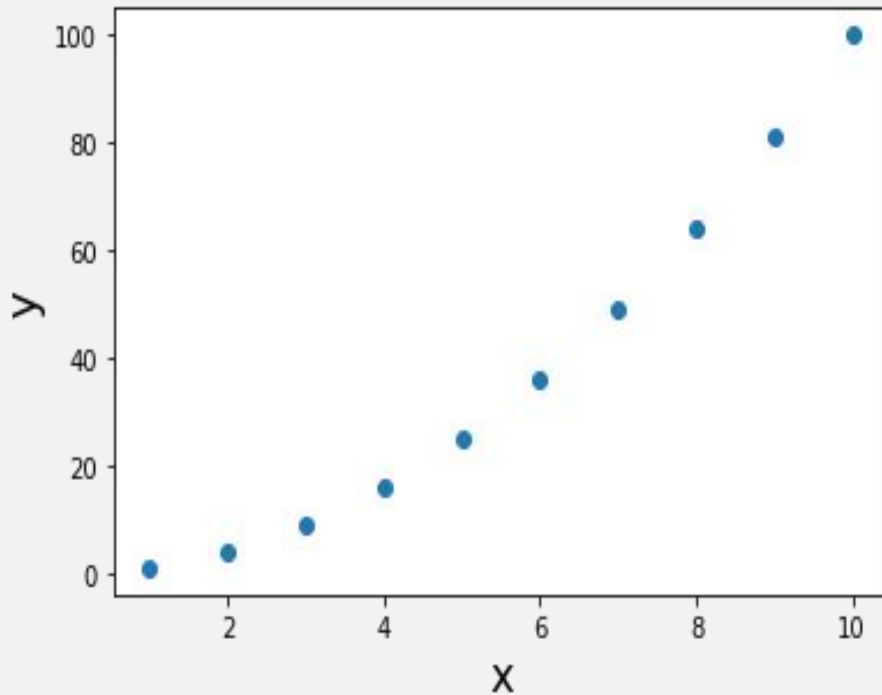
# Momento filosófico

- Estes dados são lineares?

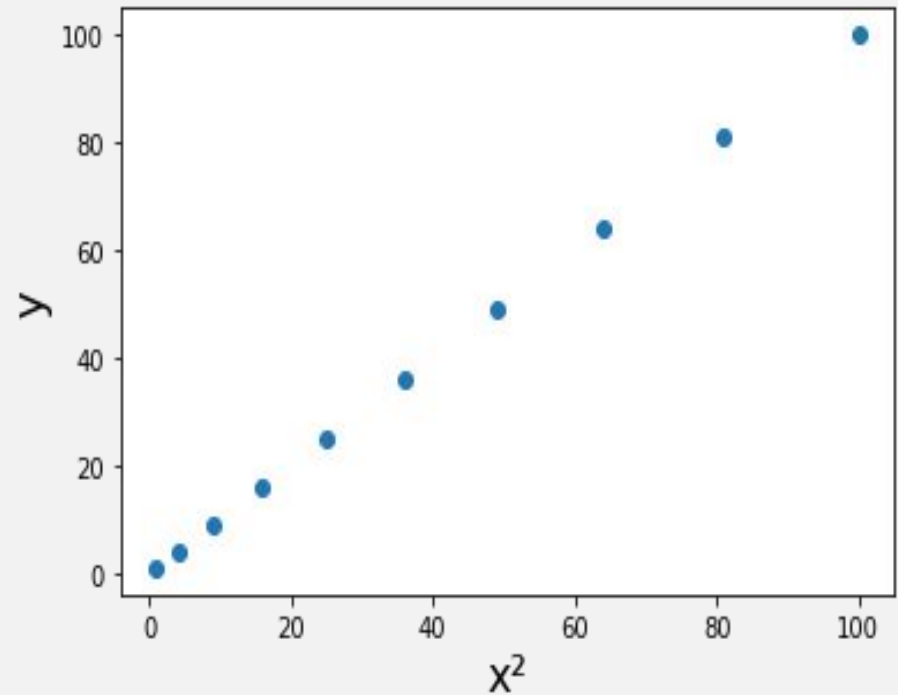


# Momento filosófico

- Estes dados são lineares?



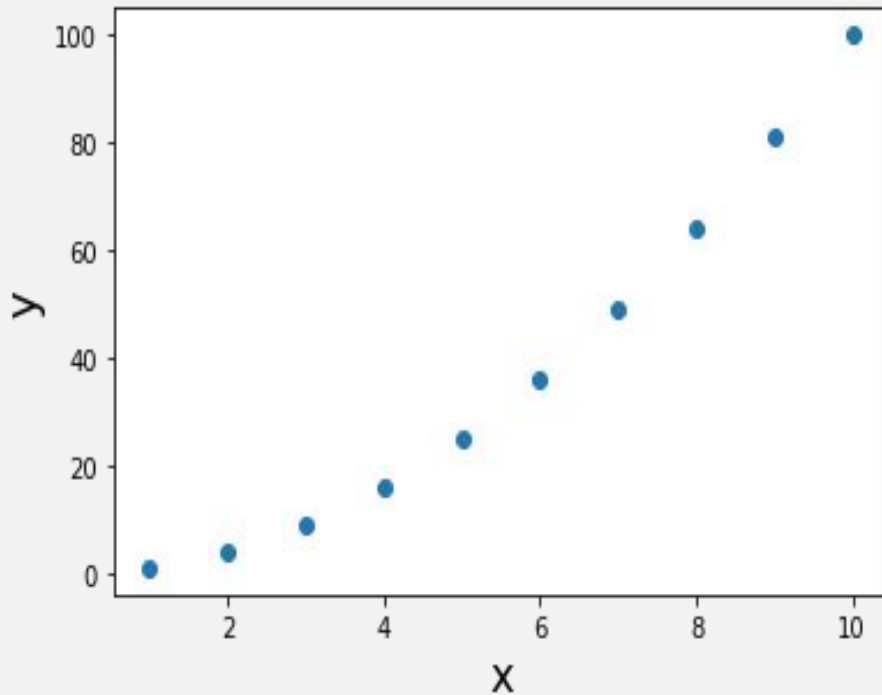
Y é não linear em relação a X



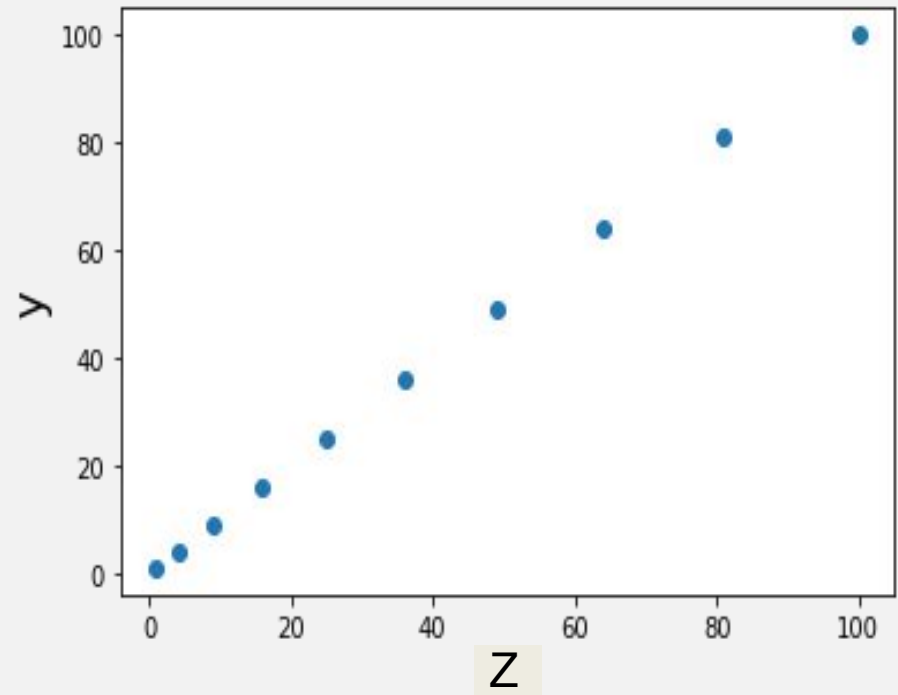
Y é linear em relação a  $X^2$

# Momento filosófico

- Estes dados são lineares?



Y é não linear em relação a X



Y é linear em relação a Z

# Regressão polinomial

- Funções não lineares podem ser LINEARIZADAS!

Suponha a função não linear:

$$\hat{y}(w, x) = w_0 + w_1x_1 + w_2x_2 + w_3x_1x_2 + w_4x_1^2 + w_5x_2^2$$

Se aplicarmos a seguinte transformação nas variáveis independentes

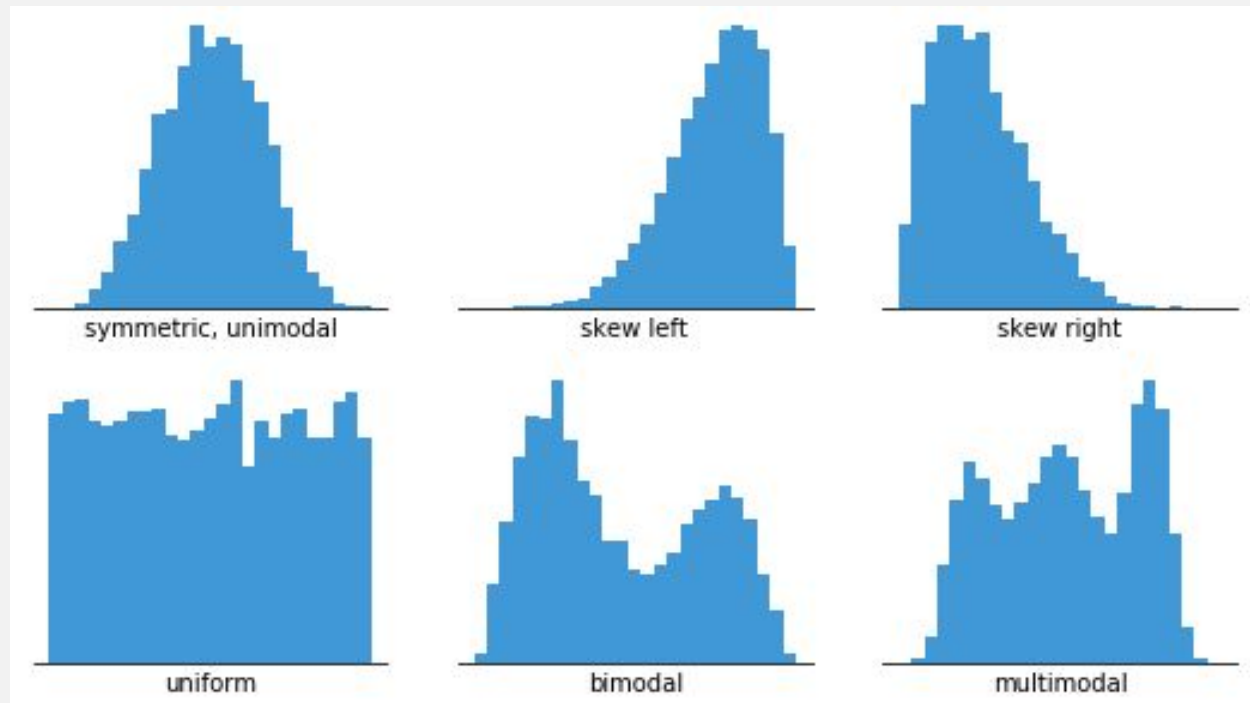
$$z = [x_1, x_2, x_1x_2, x_1^2, x_2^2]$$

Obtém-se uma função de classe linear:

$$\hat{y}(w, z) = w_0 + w_1z_1 + w_2z_2 + w_3z_3 + w_4z_4 + w_5z_5$$

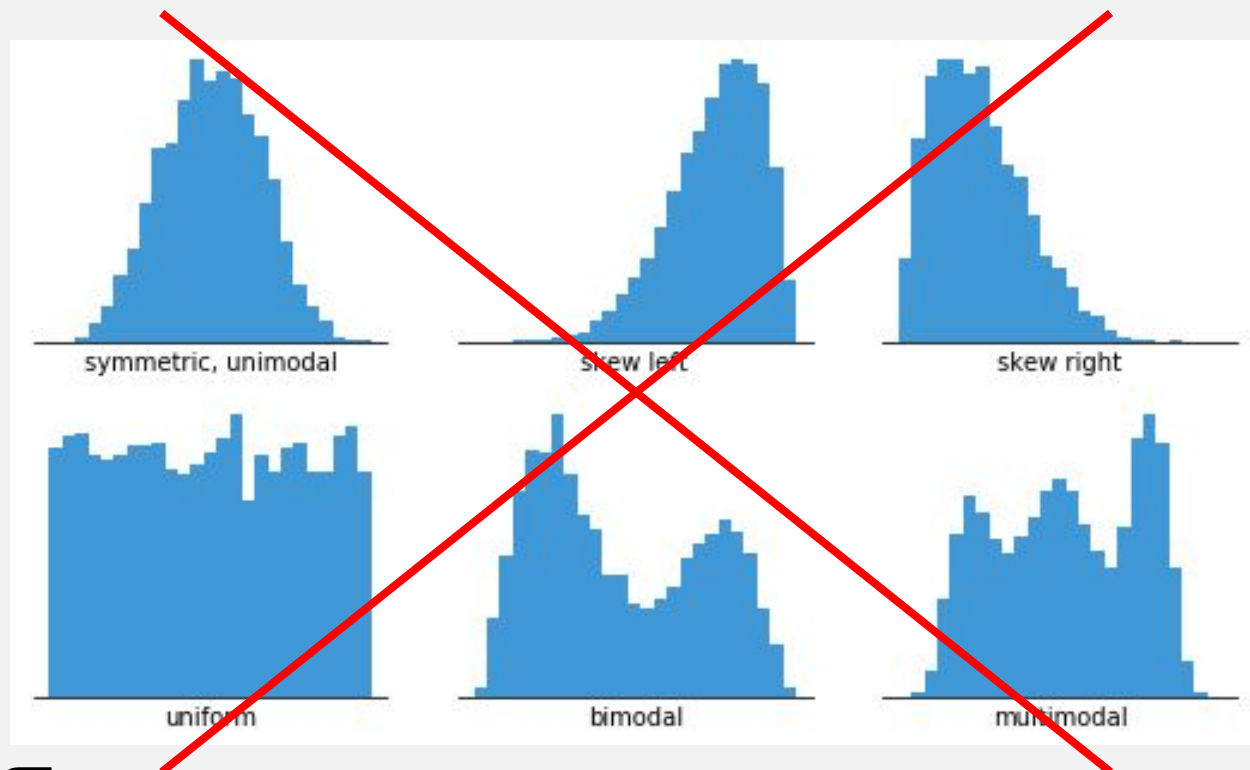
# Regressão - Decision Trees

- É um tipo de modelo **não-paramétrico** que aprende simples **decisões** a partir de um conjunto de dados
- O que significa “modelo não-paramétrico”?



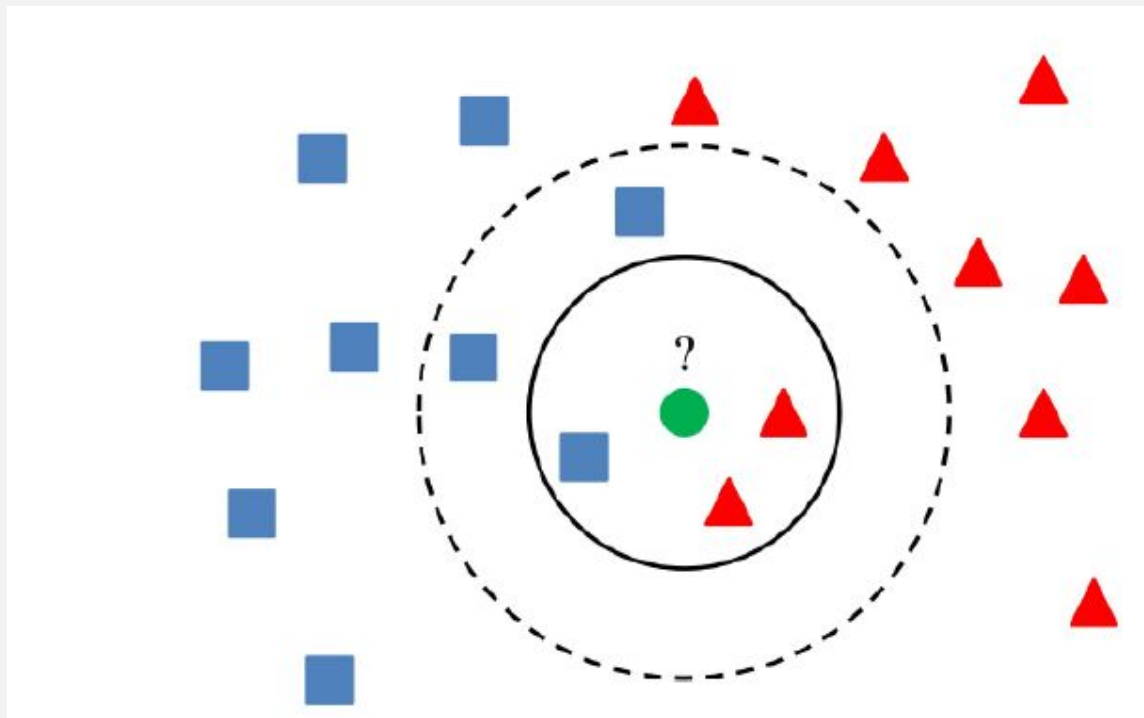
# Regressão - Decision Trees

- É um tipo de modelo **não-paramétrico** que aprende simples **decisões** a partir de um conjunto de dados
- O que significa “modelo não-paramétrico”?



# Regressão - Decision Trees

- É um tipo de modelo **não-paramétrico** que aprende simples **decisões** a partir de um conjunto de dados
- O que significa “modelo não-paramétrico”?



# Regressão - Decision Trees

Vantagens	Desvantagens
Fácil de entender e PRINCIPALMENTE de interpretar.	Muito suscetível a overfit.
Precisa de pouco pré processamento.	Bastante sensível aos dados ou parâmetros do modelo.
Boa complexidade computacional de inferência	Predição não é contínua, e sim linear em trechos. Péssimo para extrapolação.
Capaz de lidar com múltiplas saídas (múltiplos regressores)	Existem conceitos que não funcionam bem na forma lógica de árvore de decisão.
Consegue lidar tanto com variáveis numéricas como categóricas	Dados desbalanceados geram árvores tendenciosas.



# Regressão - Random Forest

- Se uma **ÁRVORE** de decisão (Decision Tree) é um modelo de regressão, o que seria uma **FLORESTA** aleatória (Random Forest)?

# Regressão - Random Forest

- É um tipo de modelo não-paramétrico que aprende simples decisões a partir de um conjunto de dados utilizando **VÁRIAS árvores de decisão diferentes**
- **Random Forest = um Ensemble (agrupamento) de Decision Trees**

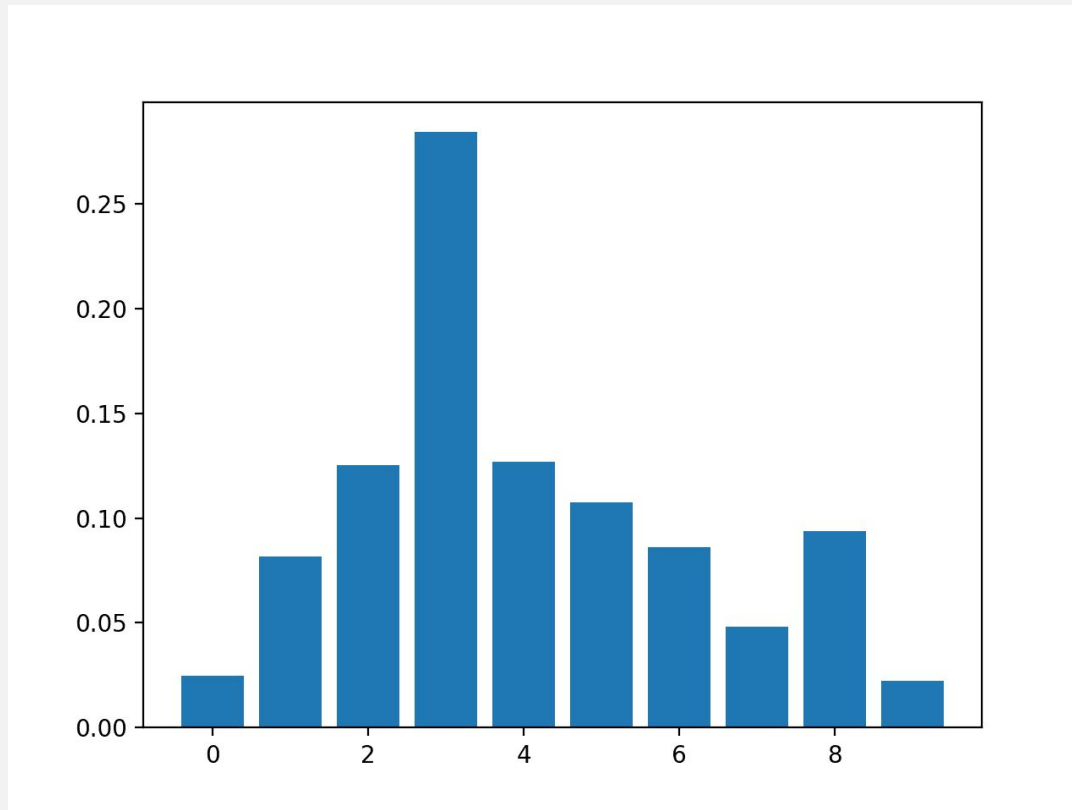


Random  
Forest

Decision  
Trees

# Regressão - Random Forest

- Interpretar uma Random Forest usando CART (Classification And Regression Tree) pode ser inviável.
- Mas ainda é possível explorar outras técnicas de interpretação dos resultados.

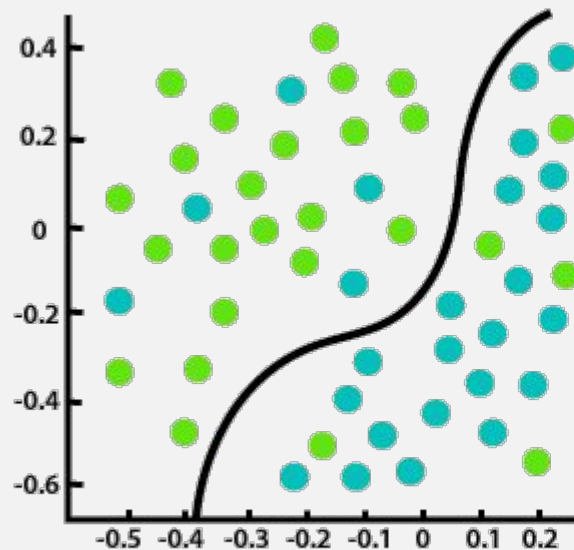


# Regressão - Random Forest

Vantagens	Desvantagens
Fácil de entender e <b>PRINCIPALMENTE de interpretar.</b>	<b>Muito</b> suscetível a overfit
Precisa de pouco pré processamento.	<b>Bastante sensível aos dados ou parâmetros do modelo.</b>
Boa complexidade computacional de inferência	<b>Predição não é contínua, e sim linear em trechos. Péssimo para extrapolação.</b>
Capaz de lidar com múltiplas saídas (múltiplos regressores)	Existem conceitos que não funcionam bem na forma lógica de árvore de decisão.
Consegue lidar tanto com variáveis numéricas como categóricas	Dados desbalanceados geram árvores tendenciosas.

# Classificação

- Objetivo: Prever a categoria de uma observação dada a partir dos dados de entrada.



Classification



CARO

CARO

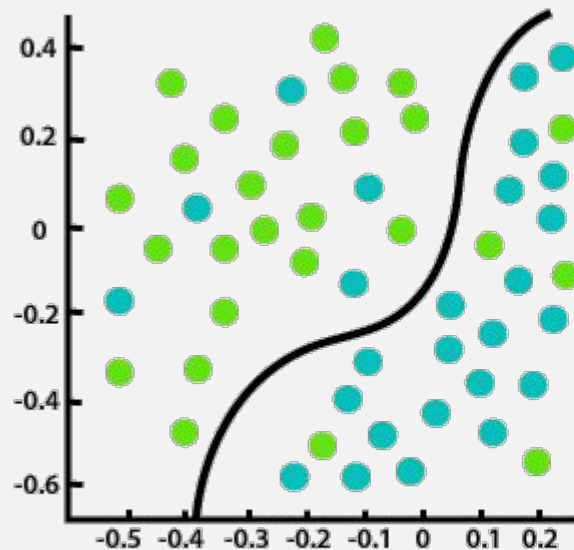
BARATO

ABSUUUURDO

MAIS OU MENOS

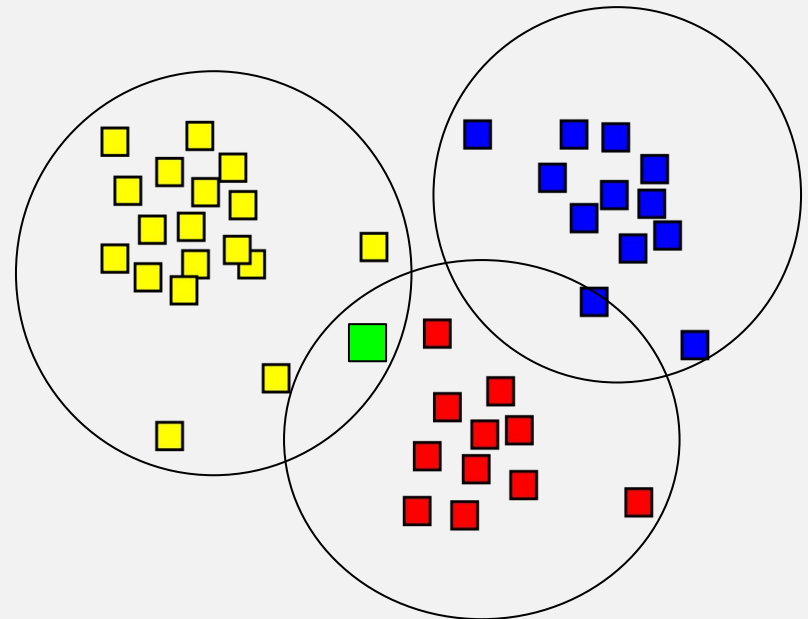
# Classificação

- Objetivo: Prever a categoria de uma observação dada a partir dos dados de entrada.
- Modelos de classificação podem ser analíticos ou não.



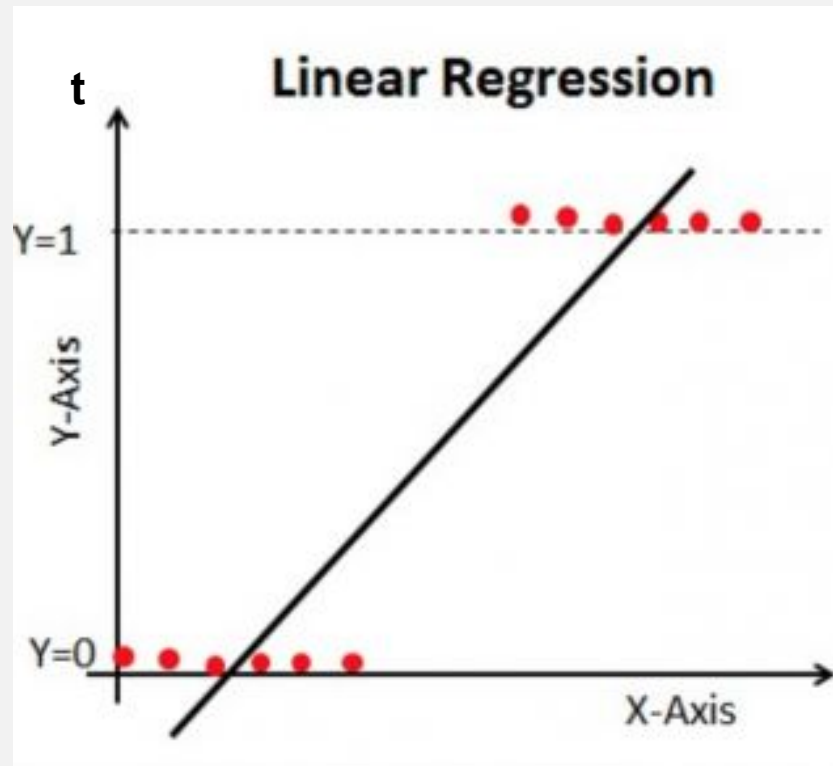
Classification

- Classificações podem ser:
  - Binárias ou
  - Multi-classe
  - Multi-label / Multi-rótulo



# Classificação - Logistic Regression

- Como determinar uma função que determina a probabilidade de um dado?

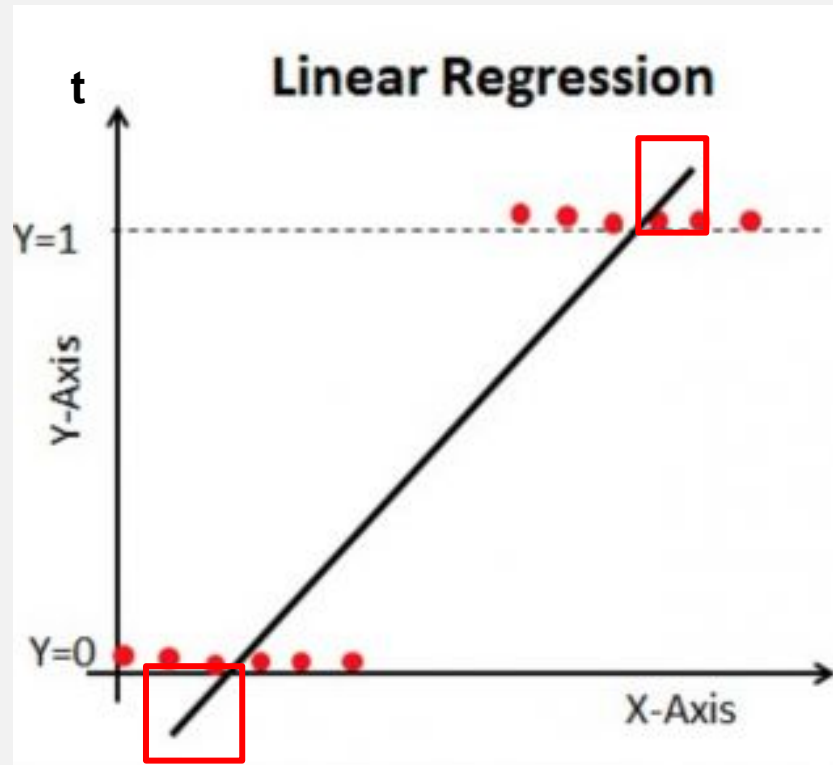


$$t = \beta_0 + \beta_1 x$$

Onde está o problema?

# Classificação - Logistic Regression

- Como determinar uma função que determina a probabilidade de um dado?

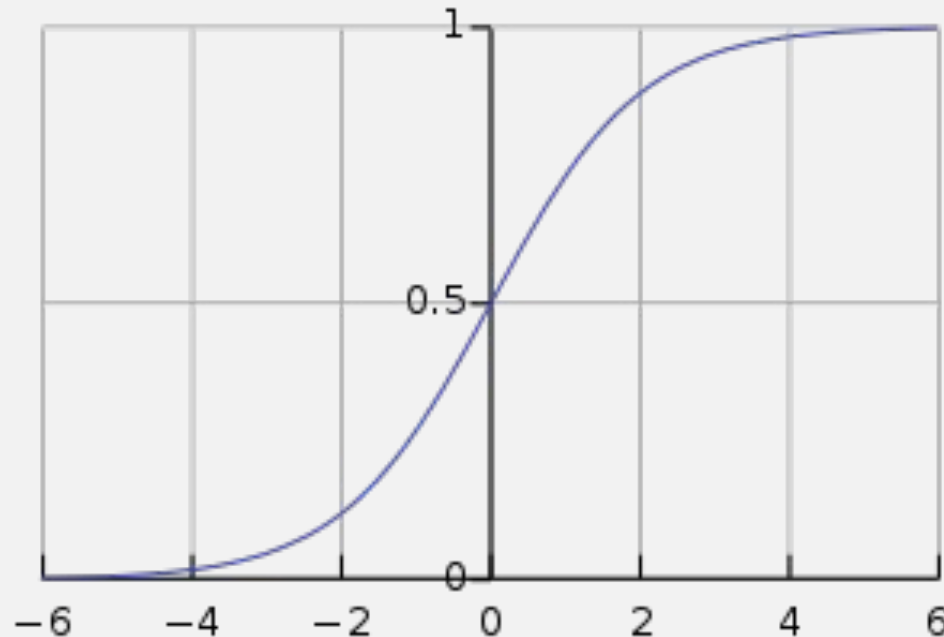


$$t = \beta_0 + \beta_1 x$$



# Classificação - Logistic Regression

- Definimos uma nova função: Função Sigmóide



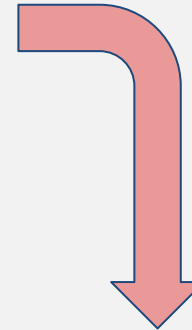
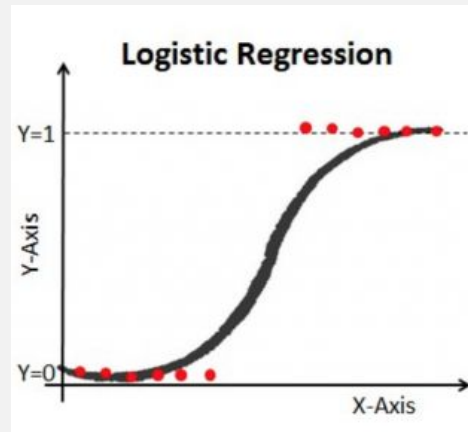
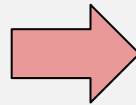
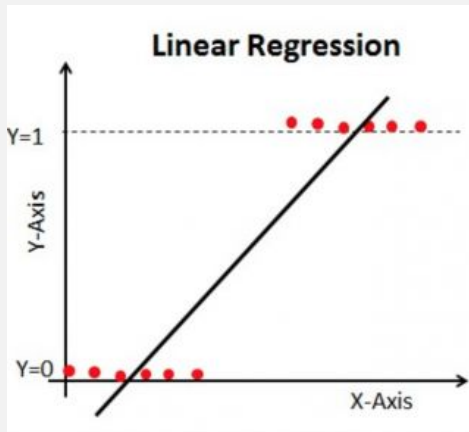
$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}}$$

# Classificação - Logistic Regression

- Como definir a fronteira de decisão?

$$t = \beta_0 + \beta_1 x$$

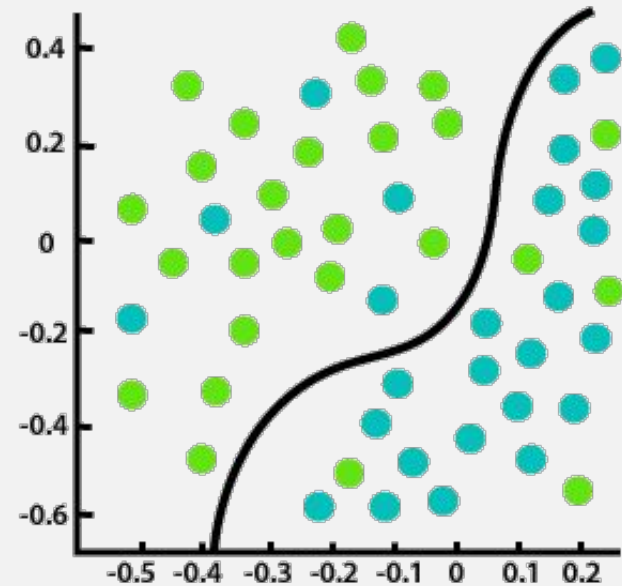
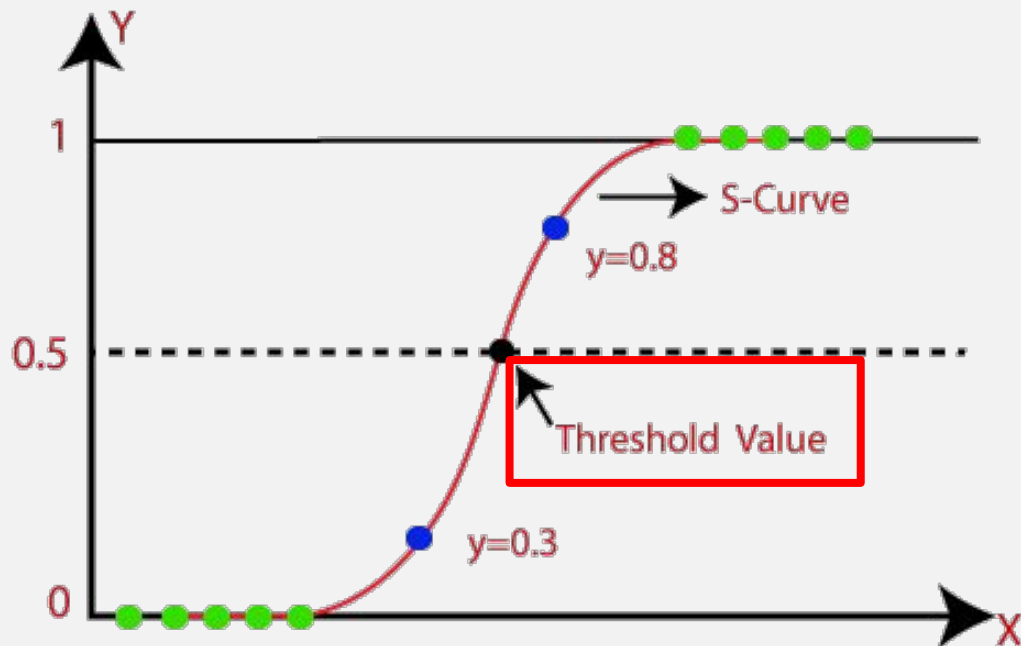
$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}}$$



$$p(x) = \sigma(t) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

# Classificação - Logistic Regression

- Como definir a fronteira de decisão?



Classification

# Classificação - Logistic Regression

Vantagens	Desvantagens
Implementação é simples	A fronteira de decisão é linear.
Inferência é rápida	Suscetível a overfit caso haja poucos dados ou o problema tenha uma alta dimensão.
Fácil de interpretar (coeficiente dos modelos podem ser interpretados como uma métrica de importância da variável de entrada)	Pode ter problemas com variáveis de entrada correlacionadas
Pode ser expandido para múltiplos rótulos.	Parte do suposição de linearidade entre variável de entrada e saída.

# Resumo

- **Machine Learning não faz milagre (Mas faz bastante mágica)**
- **Tudo (ou quase) depende do contexto e dos dados**
- **Existem 2 principais tipos de modelagem supervisionada: Regressão e Classificação**
- **É necessário matemática, mas não é suficiente**

# Referências

- <https://deepai.org/machine-learning-glossary-and-terms/non-parametric-model>
- <https://www.geeksforgeeks.org/advantages-and-disadvantages-of-logistic-regression/>
-