

Projeto ABCIA

Módulo 2 - Aula 2

Machine Learning

Prof. Msc. Acauan C. Ribeiro



Apresentação do Professor

- Bacharel em Ciência da Computação (UFRR) - 2010
- Especialização em Engenharia de Sistemas - 2012
- Mestrado em Mineração Visual de Dados (IC-UNICAMP) - 2017
- Certificação em Inteligência Artificial (Huawei - HCIA) - 2021 🏆

Professor do curso de Ciência da Computação/UFRR desde 2012, atuando nas disciplinas e pesquisa de **Banco de Dados, Estrutura de Dados, Desenvolvimento Web e Visualização de Informações.**



Prof. Msc. Acauan C. Ribeiro

- Identificar finalidade e limitações dos algoritmos de Machine Learning a partir das noções de **modelo**, **predição** e **classificação**.
- Dados de Treino e Dados de Teste
- Identificar a finalidade da noção de viés na avaliação da qualidade de um modelo de ML.
- Identificar a finalidade da noção de variabilidade nos erros do modelo
- *Underfitting* e *Overfitting*
- Identificar características de um modelo perfeito de ML (baixo viés e baixa variabilidade de erros)

O que é aprendizagem?

"**Aprendizado** é uma **mudança** relativamente permanente no conhecimento ou no comportamento produzido pela **experiência**." (1)

Memorizar

*Observar e
explorar*

Errar



Repetir

*Identificar
padrões*

(1) Livro: How Learning Works: Seven Research-Based Principles for Smart Teaching (Ambrose et al, 2010)

Exemplo



Imagine que temos a missão de ensinar uma criança de 2 anos a identificar um conjunto de animais.

Exemplo



Cachorro: Au! Au!

Gato: Miau!

Leão: Rugido. Roaarr!
Grrrr!

**Mas e o jacaré? Ou a
Girafa?**

Uma possível abordagem, que podemos observar que é aplicada em desenhos animados, brinquedos e livros infantis para essa tarefa é a **identificação pelo som** que o animal emite. A ideia é fazer uma **correlação** entre uma **característica marcante**, as vezes única e aquele animal.

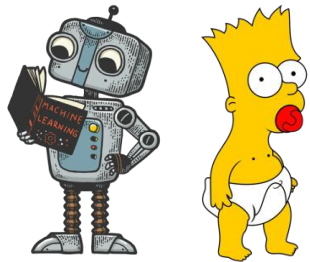
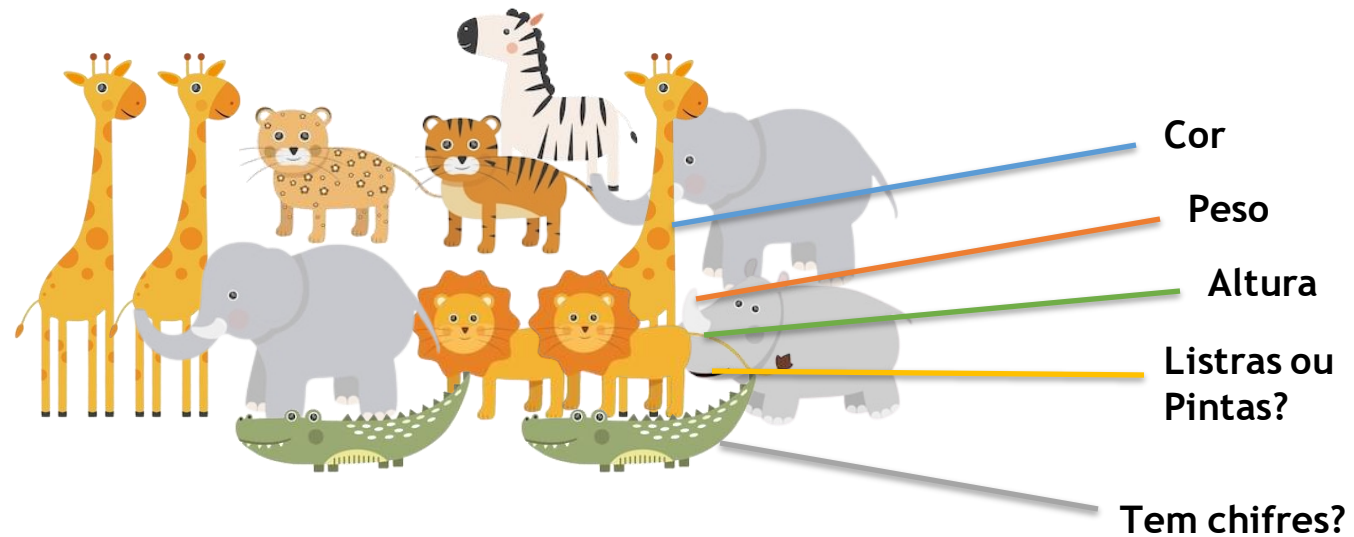


Exemplo



Agora temos que ensinar uma máquina a fazer essa classificação. Poderíamos utilizar o recurso dos sons, porém estaríamos adicionando um grau de complexidade desnecessário. Existem diversas abordagens.

Exemplo

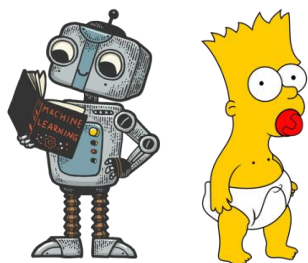


Uma possível abordagem é identificar algumas características desse nosso conjunto de dados..

Exemplo



Cor == Amarelo?



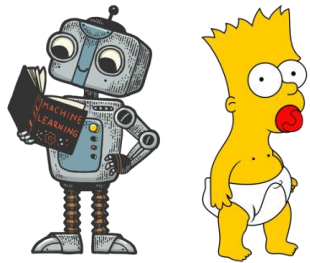
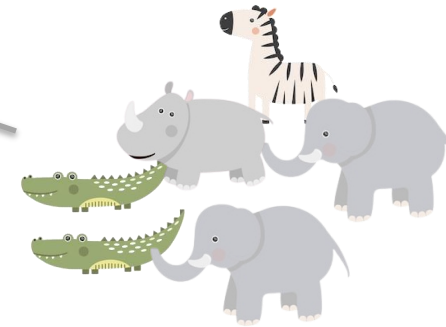
Exemplo



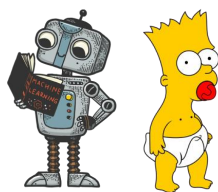
Cor == Amarelo?

Sim

Não



Exemplo



Cor == Amarelo?

Sim

Não



Altura > 3m

Sim

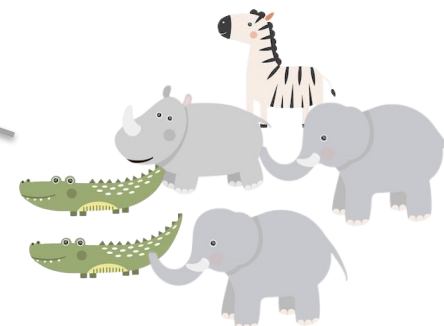
Não



Girafa



Felino



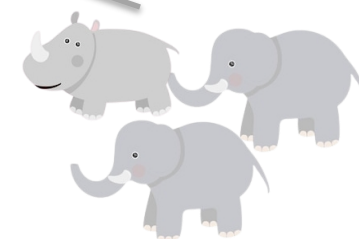
Peso > 1 tonelada

Sim

Não



??



??

Exemplo



Cor == Amarelo?

Sim

Não



Altura > 3m

Peso > 1 tonelada

Sim

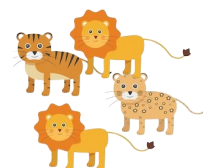
Não

Sim

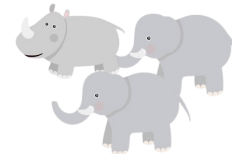
Não



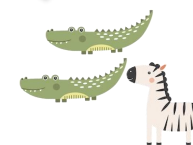
Girafa



Felinos



Tem tromba?



Tem Listras?

Sim

Não

Sim

Não

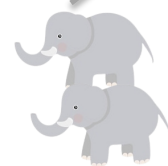
Sim

Não

Tigre



Leão



Elefante

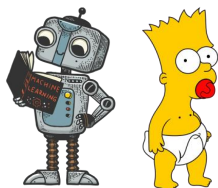
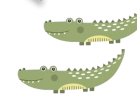
Rinoceronte



Zebra

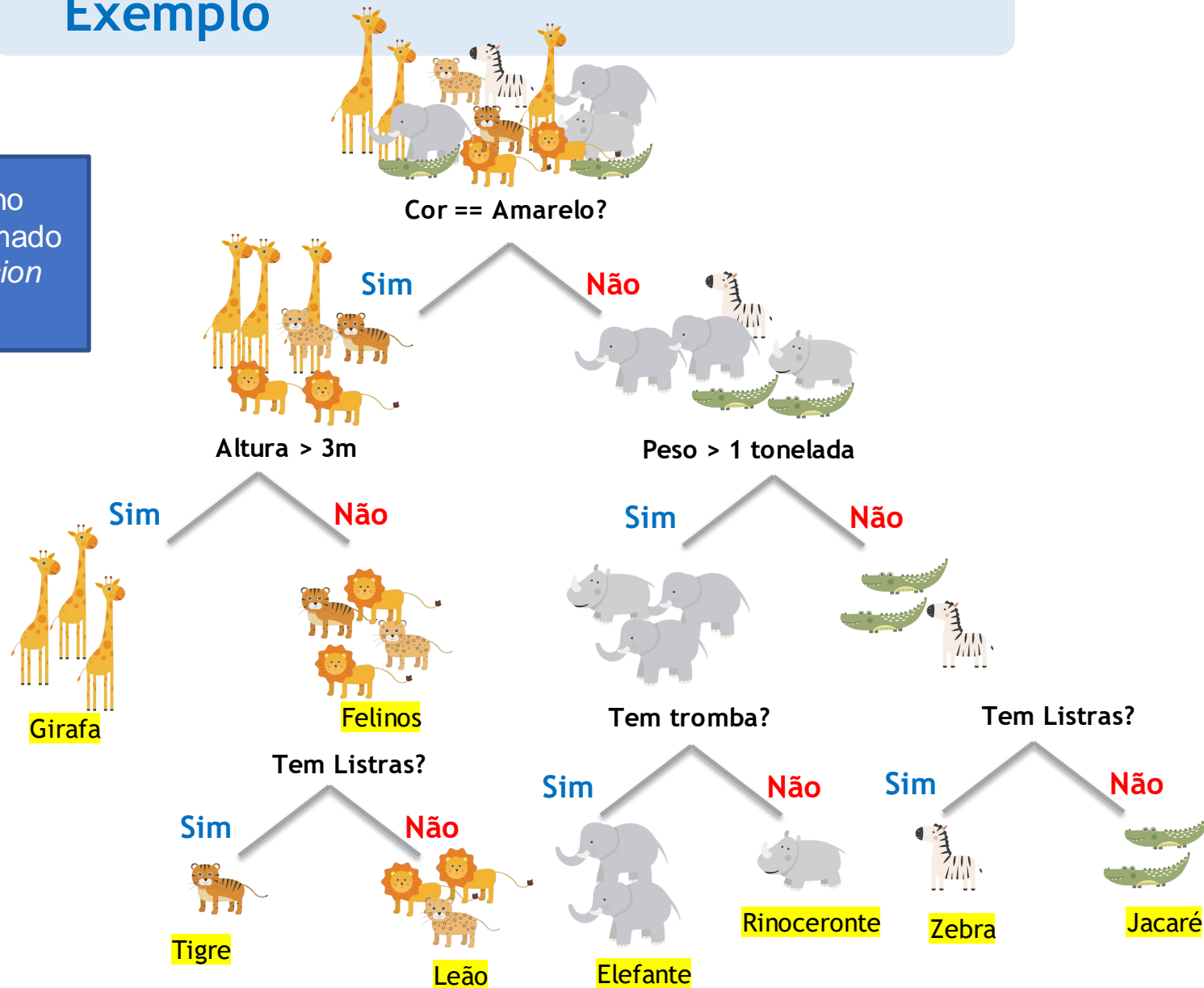


Jacaré



Exemplo

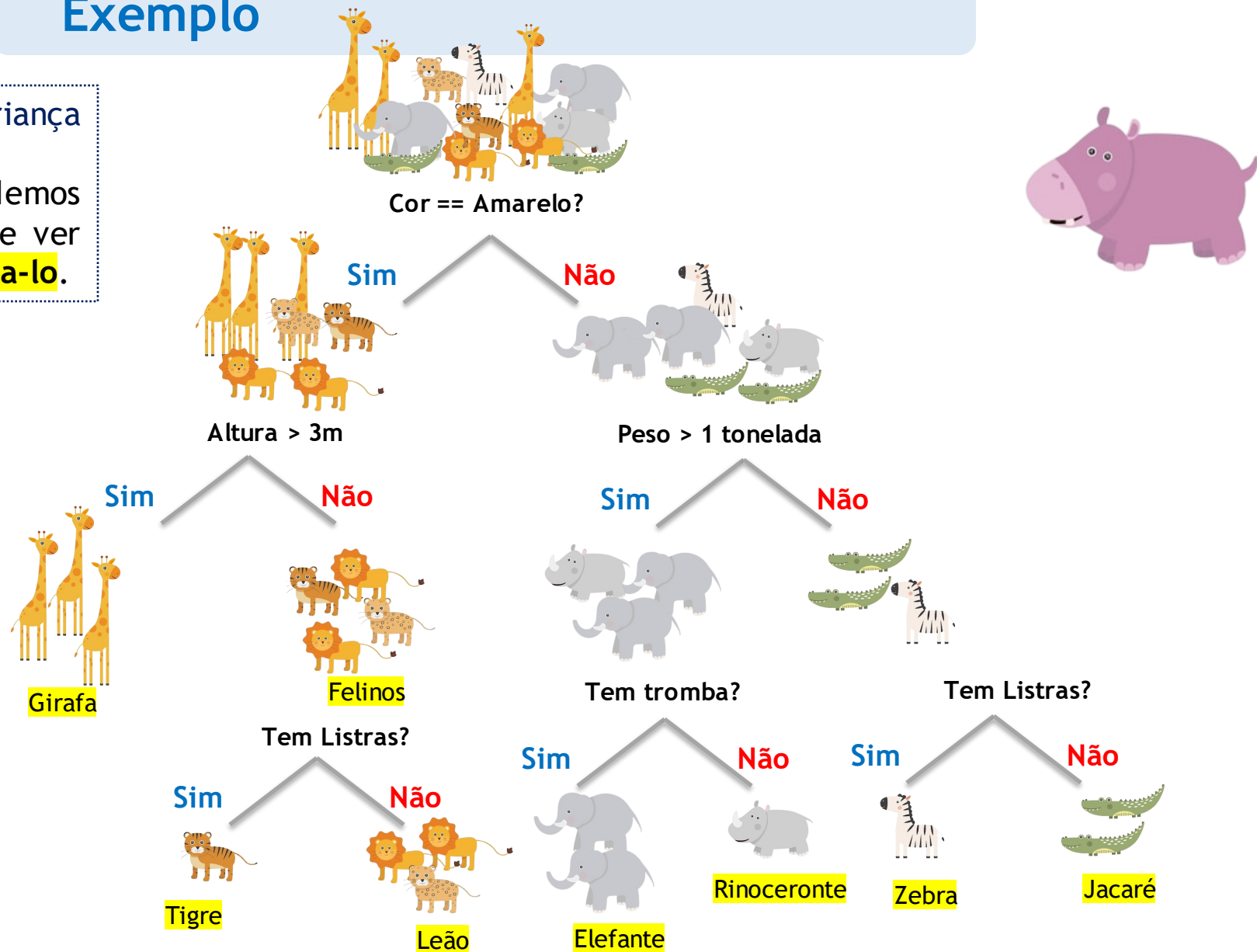
Este é um modelo utilizado no aprendizado de máquina. Chamado "Árvore de Decisão" (Decision Tree).



Exemplo

Como saber se a criança aprendeu?

Realizamos **TESTES**. Podemos colocar um novo animal e ver onde a criança irá **classifica-lo**.

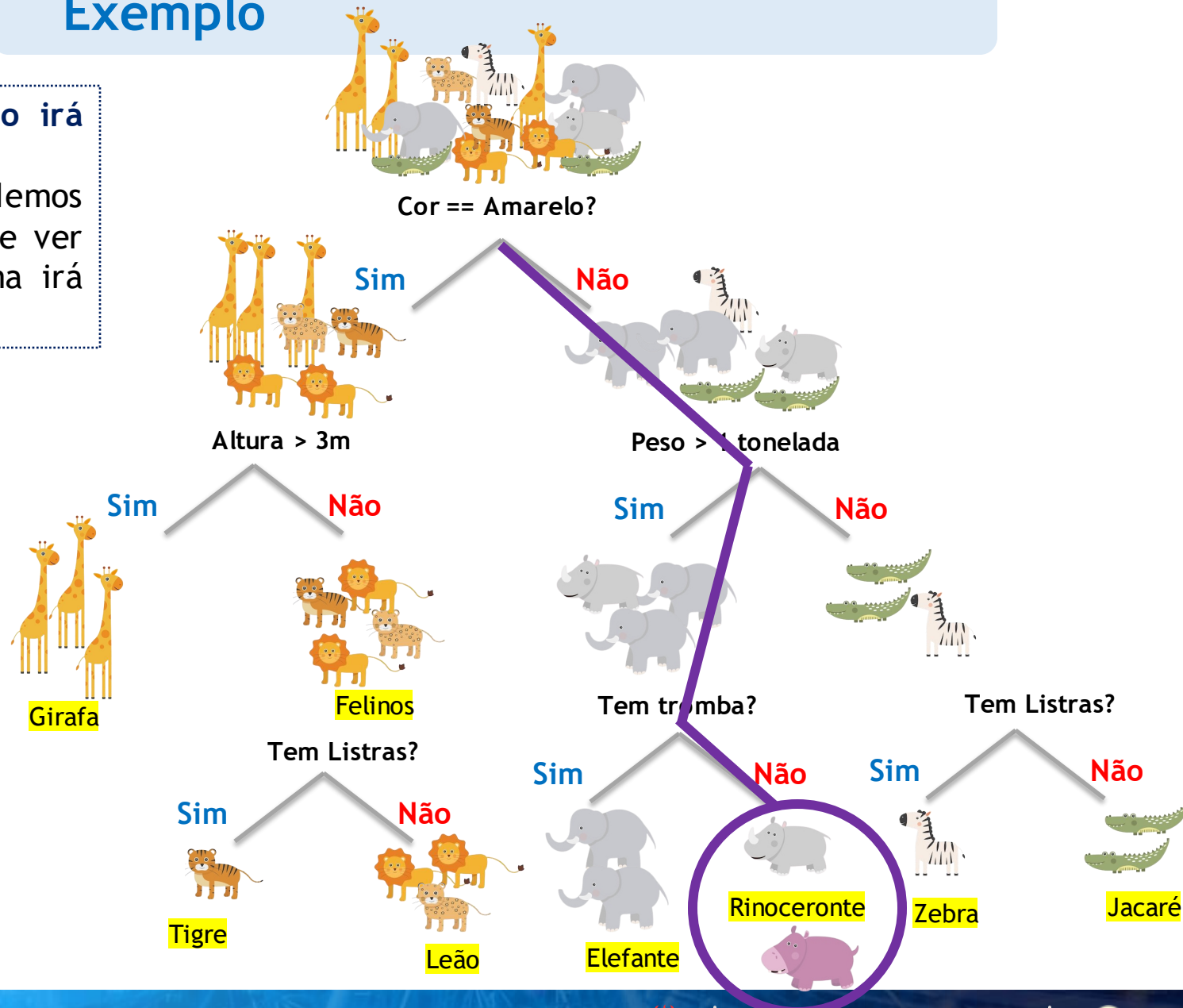
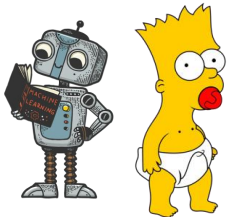


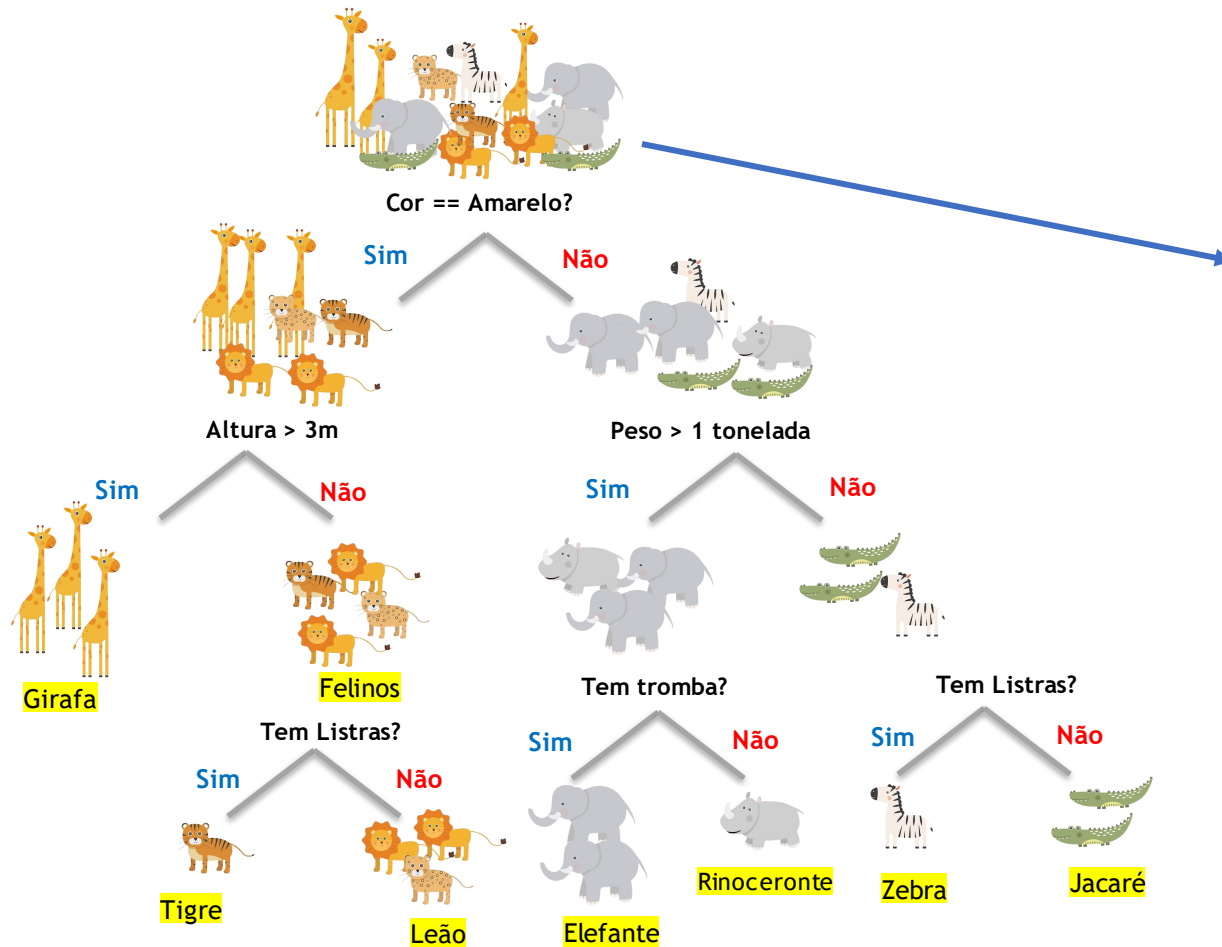
?

Exemplo

Como saber se o modelo irá funcionar?

Realizamos **TESTES**. Podemos colocar um novo animal e ver onde a criança ou sistema irá **classifica-lo**.





1

Definição e análise do conjunto de entrada

Cor	Altura	Peso	Tromba	Listras	Rótulo
Amarelo	5m	400kg	0	0	Girafa
Cinza	1,8m	1,5t	1	0	Elefante
Branco	2m	600kg	0	1	Zebra
Amarelo	1,6	480kg	0	0	Girafa
Verde	1m	300kg	0	0	Jacaré
...
Cinza	2m	2t	0	0	Rinoceronte



Cor == Amarelo?

Sim

Não

Aplicações baseadas em **Aprendizado de Máquina (AM)** utilizam heurísticas que buscam por modelos capazes de representar o **conhecimento presente em um conjunto de dados**.

Em geral, os conjuntos de dados são estruturados em **formato tabular**, uma matriz atributo-valor, em que cada linha representa um objeto (instância ou exemplo) e cada coluna representa um atributo (característica ou variável).

1

Definição e análise do conjunto de entrada

Cor	Altura	Peso	Tromba	Listras	Rótulo
Amarelo	5m	400kg	0	0	Girafa
Cinza	1,8m	1,5t	1	0	Elefante
Branco	2m	600kg	0	1	Zebra
Amarelo	1,6	480kg	0	0	Girafa
Verde	1m	300kg	0	0	Jacaré
...
Cinza	2m	2t	0	0	Rinoceronte

Gira

Tigre

Leão

Elefante

Rinoceronte

Zebra

Jacaré



Cor == Amarelo?

Sim

Não

Atributos preditivos: cujos valores descrevem características dos objeto;

Atributo alvo: cujo valor rotula o objeto, com uma classe ou valor numérico.

Essas denominações têm por origem o frequente uso dos valores dos atributos preditivos de um objeto para predizer o valor de seu atributo alvo. Nem todos os conjuntos de dados possuem atributo alvo. Quando possuem, são chamados de **conjuntos de dados rotulados**.

1

Definição e análise do conjunto de entrada

Cor	Altura	Peso	Tromba	Listras	Rótulo
Amarelo	5m	400kg	0	0	Girafa
Cinza	1,8m	1,5t	1	0	Elefante
Branco	2m	600kg	0	1	Zebra
Amarelo	1,6	480kg	0	0	Girafa
Verde	1m	300kg	0	0	Jacaré
...
Cinza	2m	2t	0	0	Rinoceronte

Qualidade e Tipos de Dados

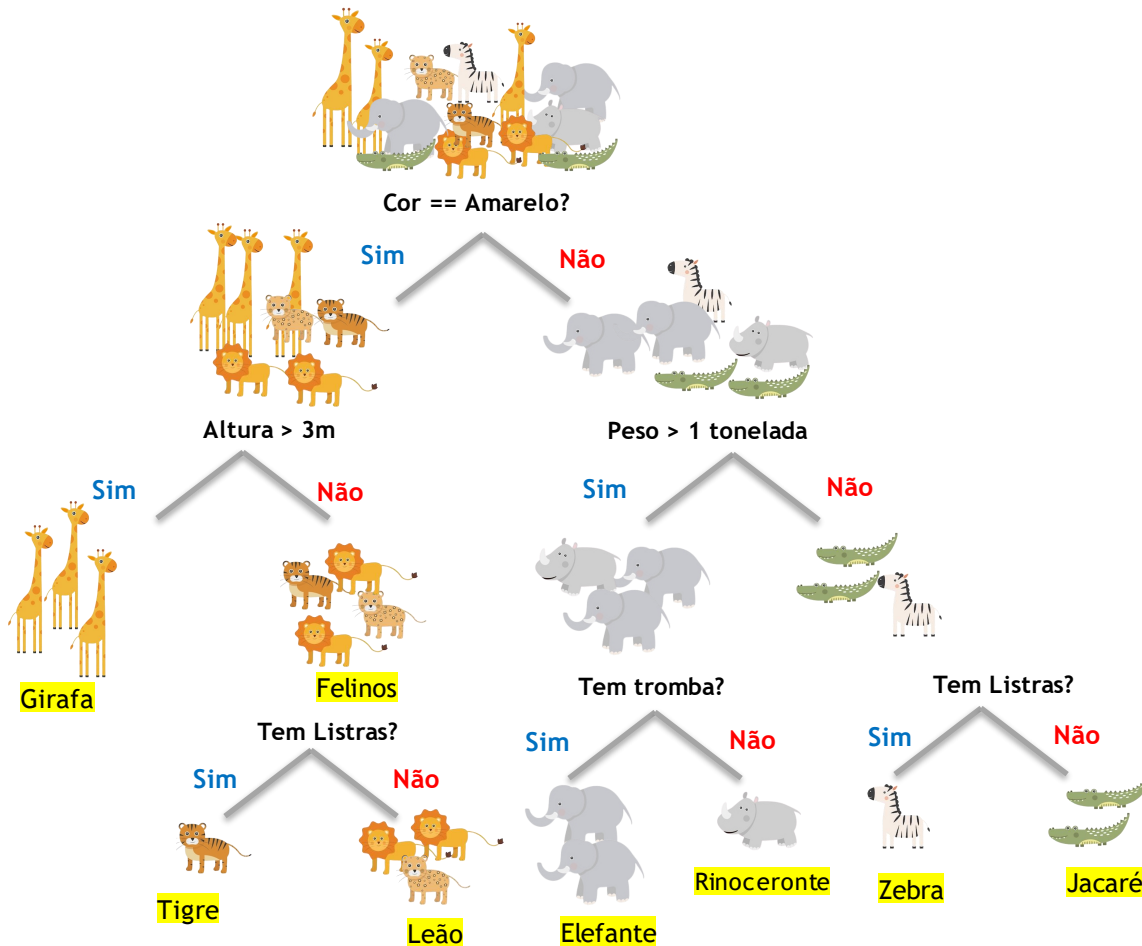


2

Definição do modelo

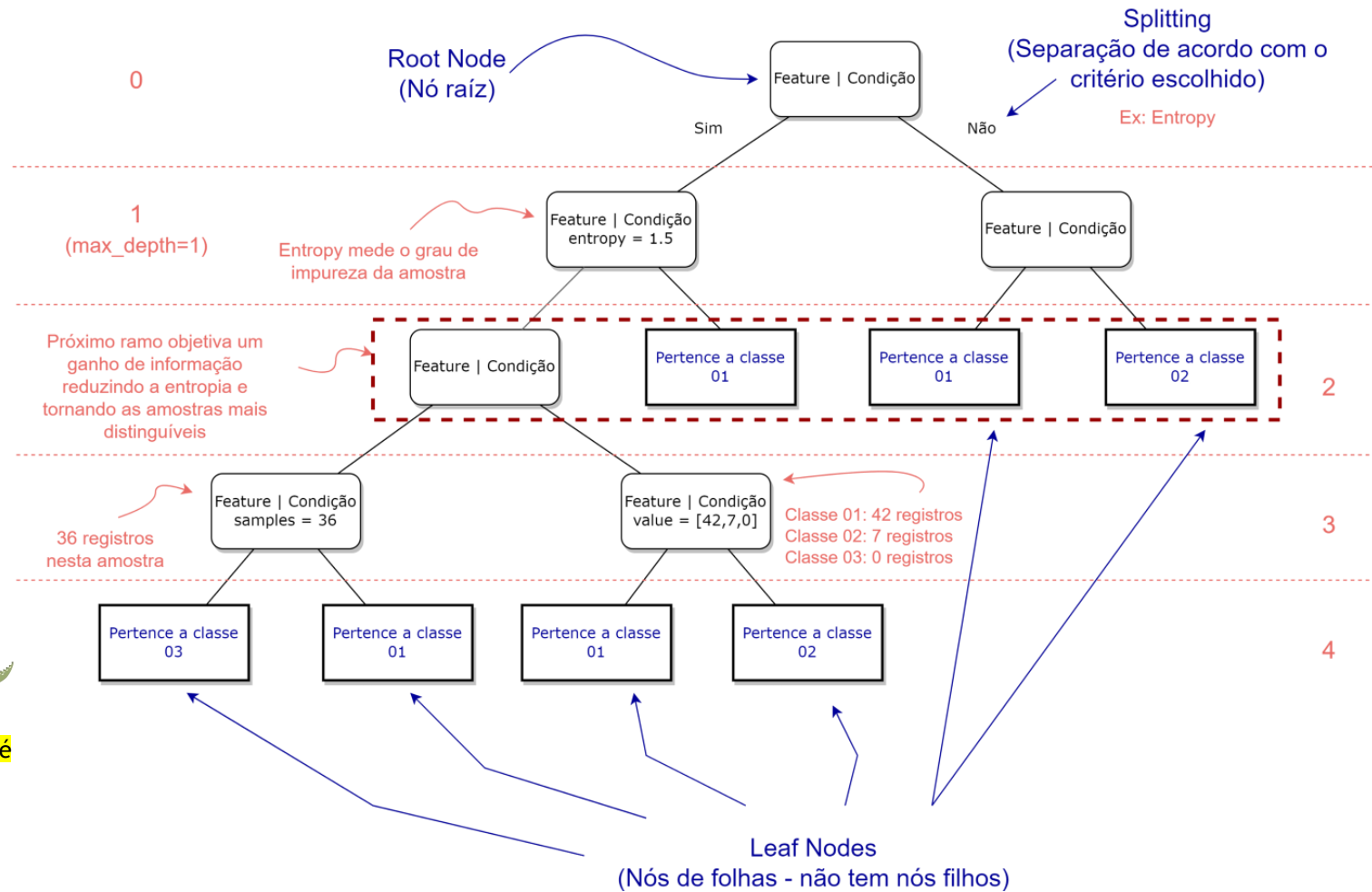
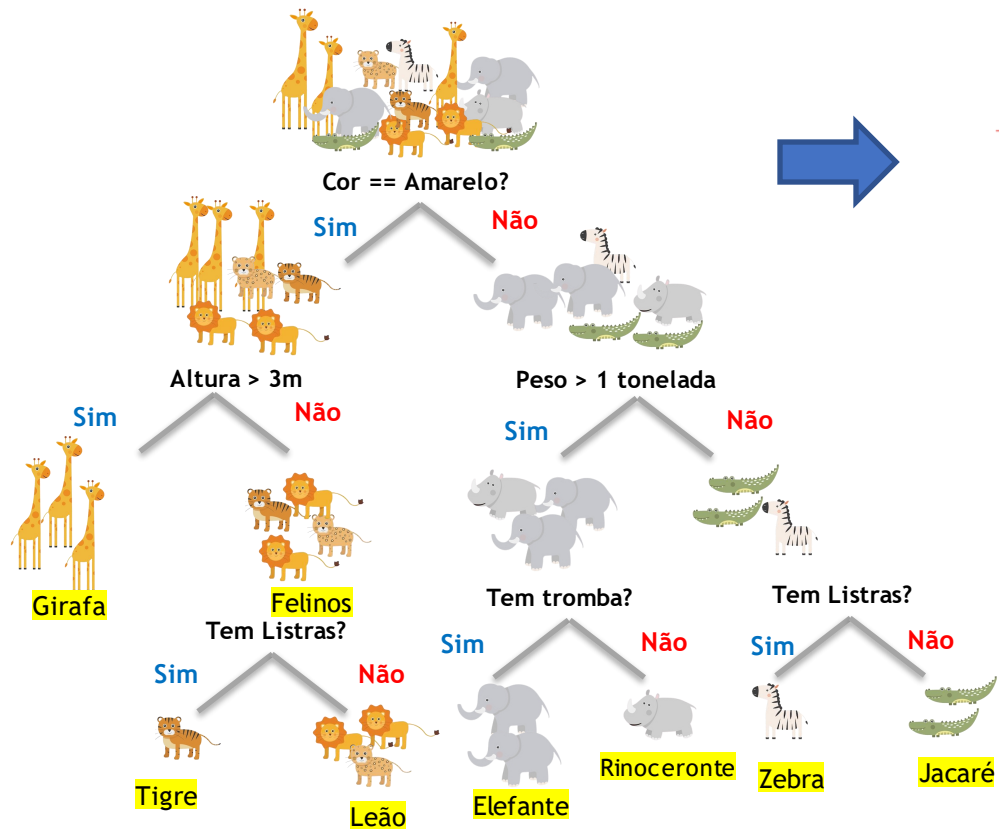
Árvore de Decisão: Baseia-se na separação binária (True | False) de condições utilizando os atributos – *features* – dos dados. Utiliza algoritmos como: **ID3, C4.5 e o CART**

Esses modelos de **aprendizado supervisionado** são empregados em tarefas tanto de **regressão** quanto **classificação**. O objetivo do método é criar **uma árvore que aprende** com os dados por meio de regras básicas, chamadas de *if-then rules*. O modelo identifica qual é o **melhor atributo para separar os dados utilizando o critério escolhido**, que, neste caso, pode ser *entropy* ou *gini*.



2

Definição do modelo



Definição de Aprendizado de Máquina

"A Capacidade de melhorar o desempenho na realização de alguma tarefa por meio da experiência." - *Mitchell (1997)*

"Classifica-se a **aprendizagem de máquina** como uma ramificação da **Inteligência Artificial** que, por sua vez, é uma das áreas da **Ciência da Computação**. De modo geral, **Machine Learning (ML)** é uma automatização de modelos analíticos, a qual permite a criação de modelos que analisam conjuntos de dados e aprendem a reconhecer padrões/realizar previsões; ML é baseada na aplicação de algoritmos que dão aos modelos a capacidade de identificação e/ou previsão, a depender do tipo de aprendizado envolvido na aplicação."

A ideia básica é: Em vez de falar para a máquina o que e como ela vai fazer. Você passa um conjunto de dados e ela "sozinha ou com ajuda" vai aprender como fazer determinada **ação**.

Programação Clássica vs Aprendizado de Máquina



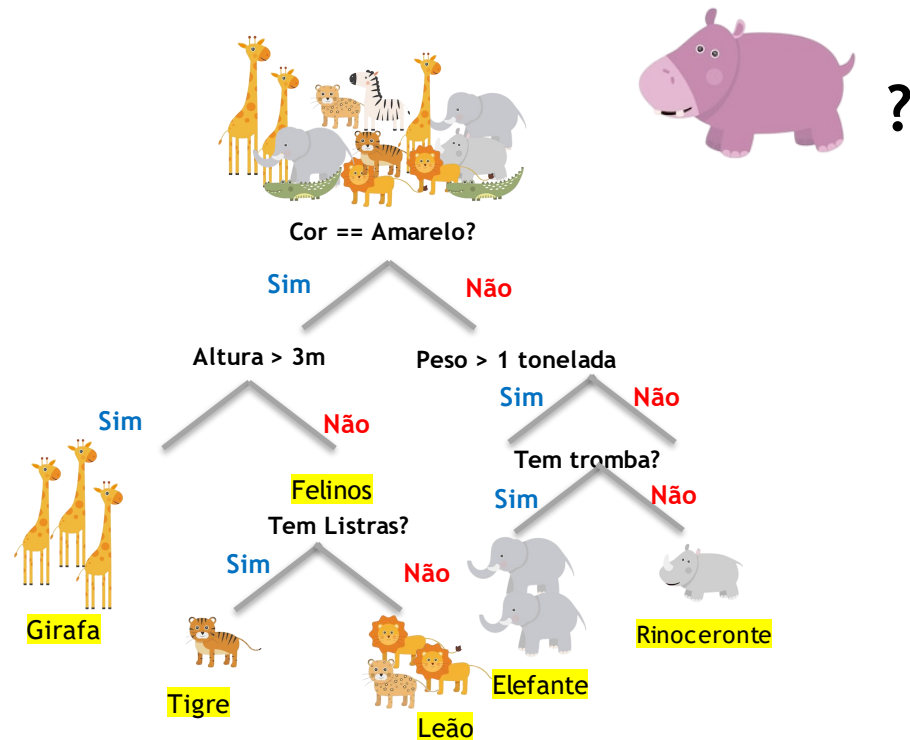
Fonte: traduzido e adaptado por Nei Grando, de "Paradigm shift in Machine Learning, by Chollet, 2017" in "A European Perspective (2018), by European Comission"

Tipos de Aprendizado de Máquina



Machine Learning: Fundamentos

Inferência/Indução: Aprender a partir de experiências passadas. Para tal, utilizando um princípio de inferência denominado **indução**, que permite extrair conclusões genéricas a partir de um conjunto particular de exemplos.



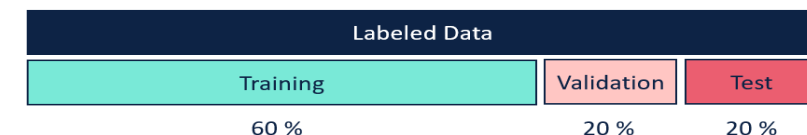
Cor	Altura	Peso	Tromba	Listras	Rótulo
Amarelo	5m	400kg	0	0	Girafa
Cinza	1,8m	1,5t	1	0	Elefante
Branco	2m	600kg	0	1	Zebra
Amarelo	6m	340kg	0	1	Girafa
Verde	1m	200kg	null	0	Pato
...
Cinza	2m	2t	0	0	Rinoceronte

Dados de Treino
80%
(Training Dataset)

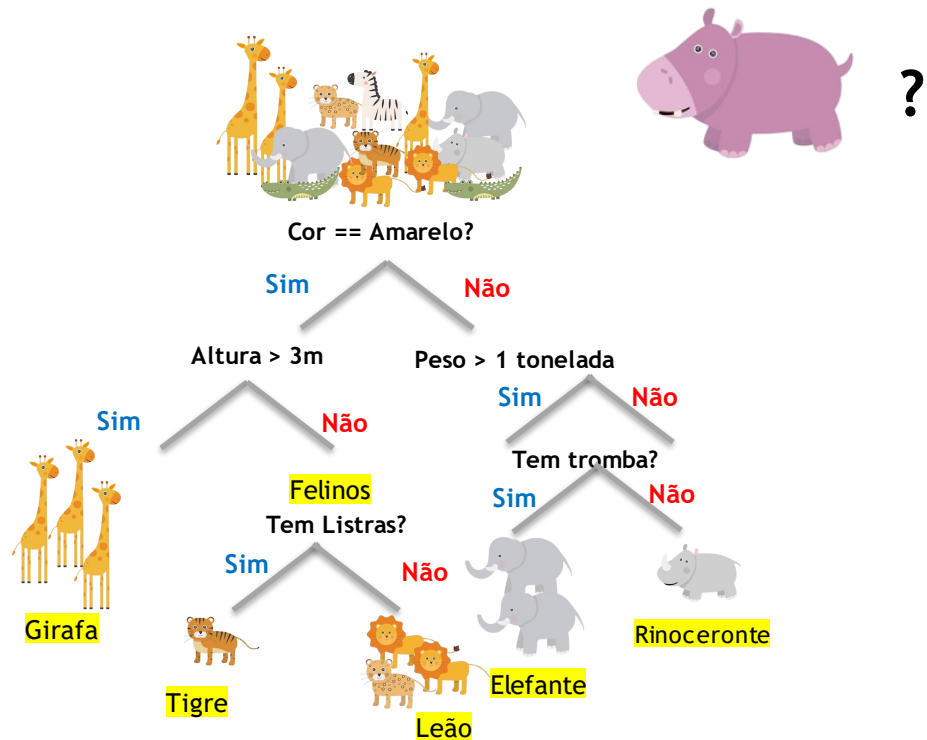
Cor	Altura	Peso	Tromba	Listras
Amarelo	5m	400kg	0	0
Cinza	1,8m	1,5t	1	0
...
Cinza	2m	2t	0	0

Dados de Teste (20%)
(Test Dataset)

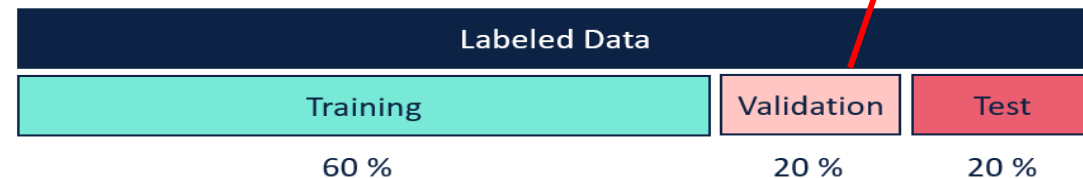
Outras referências:



Machine Learning: Fundamentos

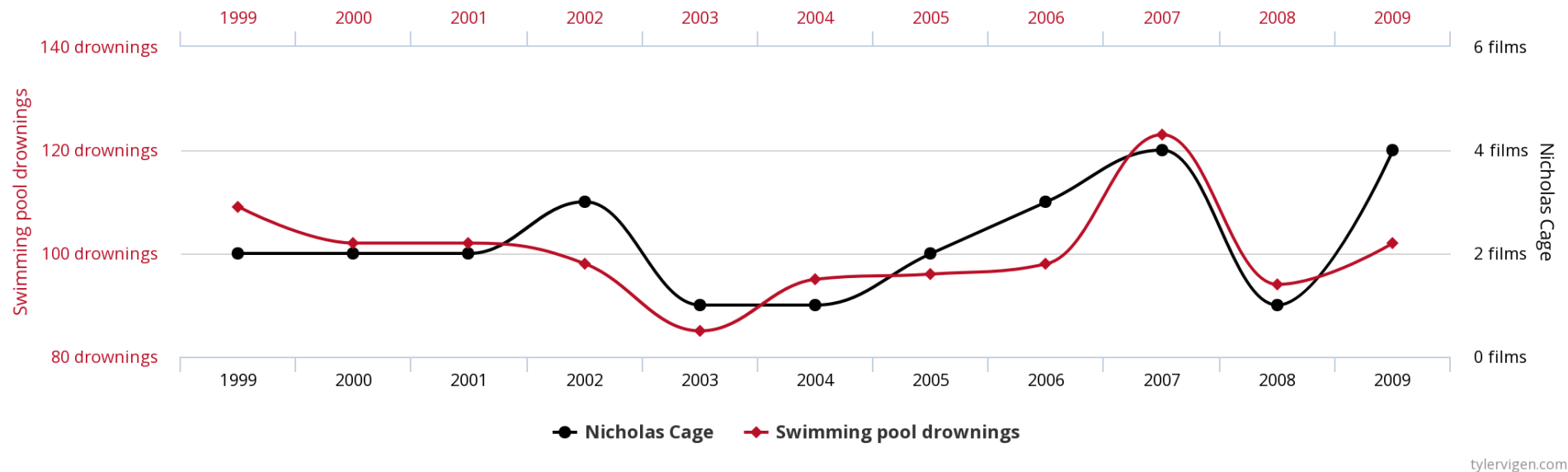


Generalização: Propriedade de um modelo continuar a ser válido para **novos objetos** dá-se o nome capacidade de generalização. Essa capacidade é necessária para que o modelo seja **válido** para novos dados



Machine Learning: Fundamentos

Number of people who drowned by falling into a pool
correlates with
Films Nicolas Cage appeared in



<https://www.tylervigen.com/spurious-correlations>

Correlação

- Relação estatística entre variáveis
- Mede a interdependência entre elas
- Mesmo eventos sem relação de causa podem ter alto coeficiente de correlação



Causalidade

- Uma relação causal entre variáveis
- É também uma correlação (o inverso não se aplica)
- Diferentes tipos
- Diversos métodos para testar
- Perguntas essenciais
 - Inferência causal
 - Descoberta causal

Machine Learning: Fundamentos

Dado o tamanho em metros quadrados de uma casa queremos saber o preço médio dela para Boa Vista-RR em 2023:

casa	m ²	preço
A	40	R\$ 273 mil
B	60	R\$ 410 mil
C	100	R\$ 684 mil
D	250	R\$ 1,7 milhões

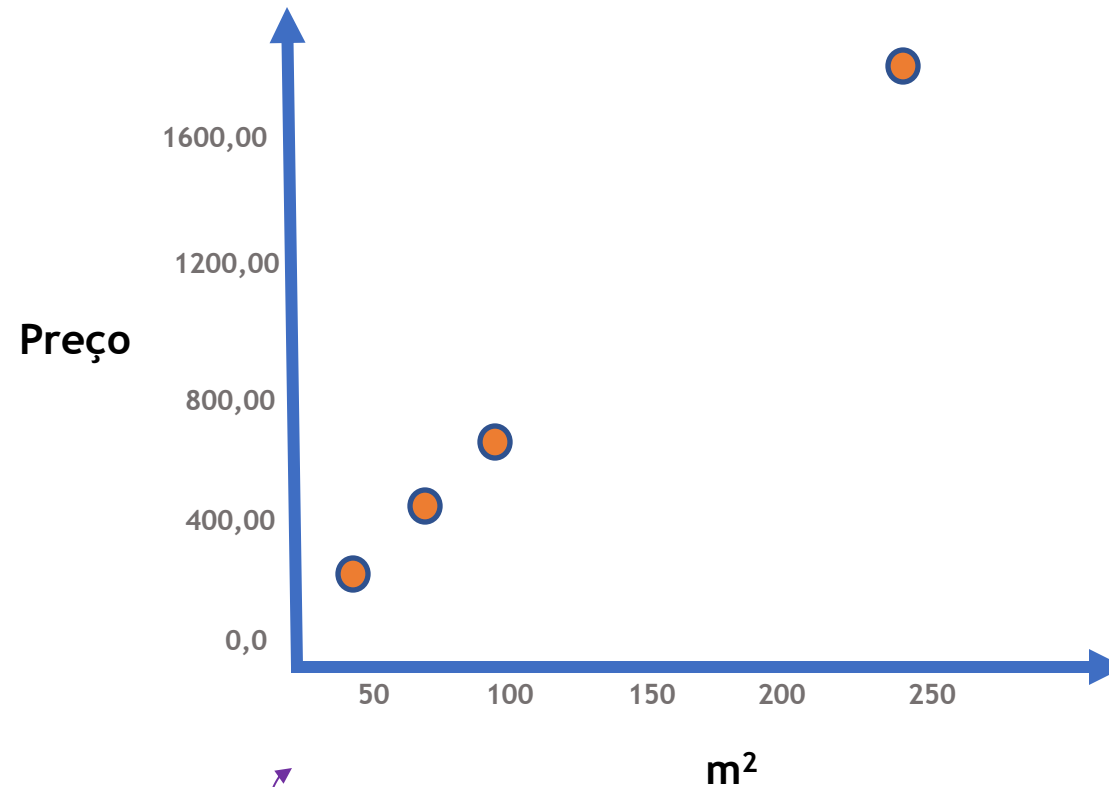
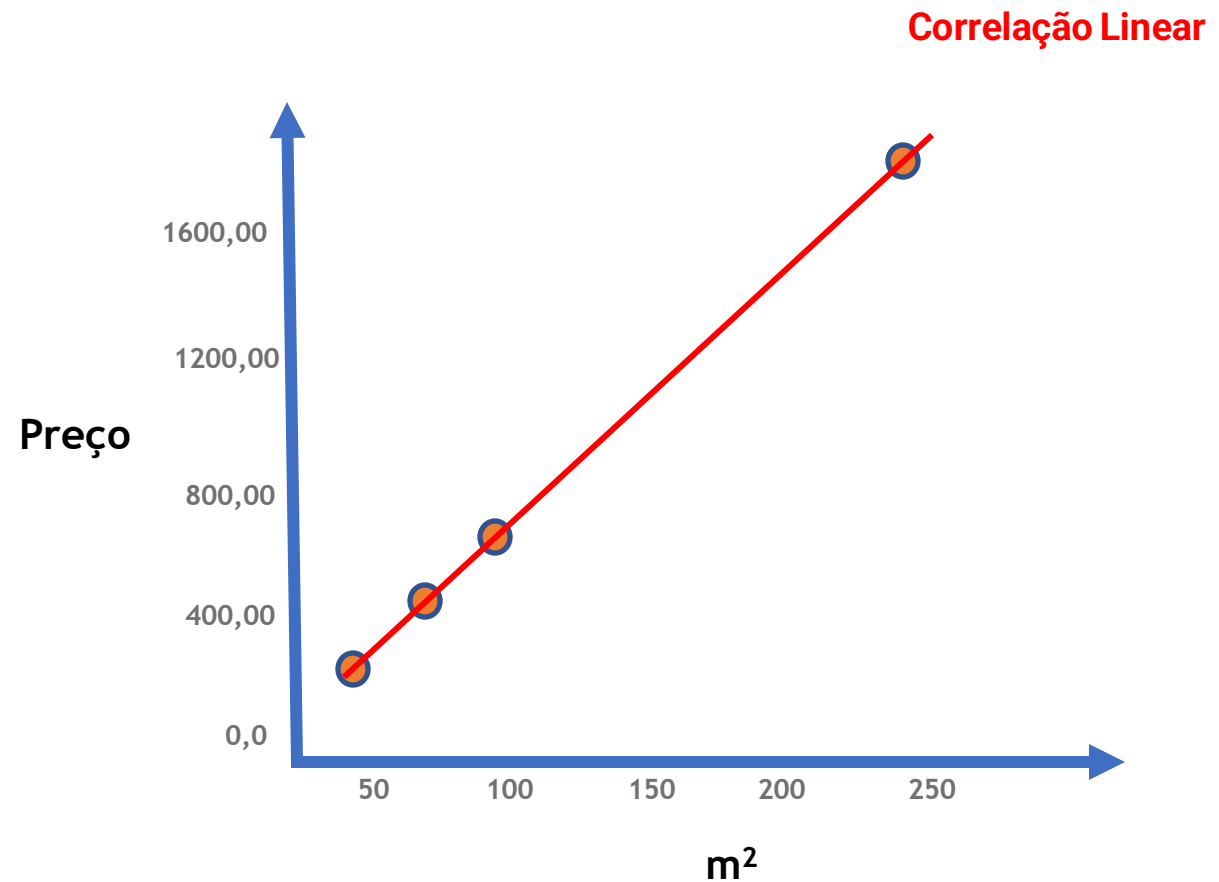


Gráfico de dispersão (scatterplot)

Machine Learning: Fundamentos

Dado o tamanho em metros quadrados de uma casa queremos saber o preço médio dela para Boa Vista-RR em 2023:

casa	m ²	preço
A	40	R\$ 273 mil
B	60	R\$ 410 mil
C	100	R\$ 684 mil
D	250	R\$ 1,7 milhões



O '**r de Pearson**' será sempre um número entre -1 e 1. Se for zero, então, não há correlação *linear* entre as variáveis. Mas atenção: é possível que exista uma correlação não linear.

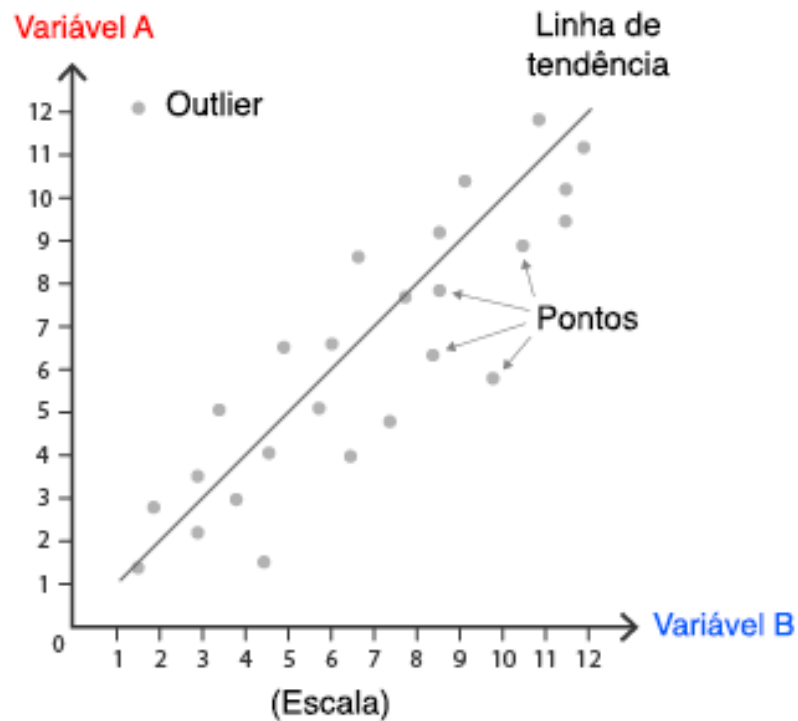
Machine Learning: Fundamentos

Variável dependente e independente: As variáveis podem ter 3 tipos de interações entre si. A correlação entre elas pode ser: **positiva**, **negativa** ou **neutra**.

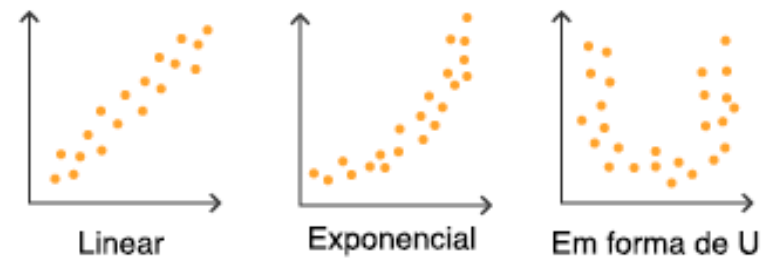
Tipo	Descrição	Exemplos comuns
Positiva	Quando o valor de uma variável aumenta, o outro também sobe.	<ul style="list-style-type: none">Altura das pessoas VS o tamanho do sapatoTempo de viagem de um avião VS distância percorrida
Neutra	Não há relação entre as duas variáveis.	<ul style="list-style-type: none">Consumo de manga por pessoa VS notas em matemática.Altura das pessoas VS quantidade de habitantes de uma cidade.
Negativa	Quando o valor de uma variável sobe, o outro desce.	<ul style="list-style-type: none">Distância percorrida VS tempo até a chegada.Refeições feitas em casa VS dinheiro gasto em restaurantes.

Machine Learning: Fundamentos

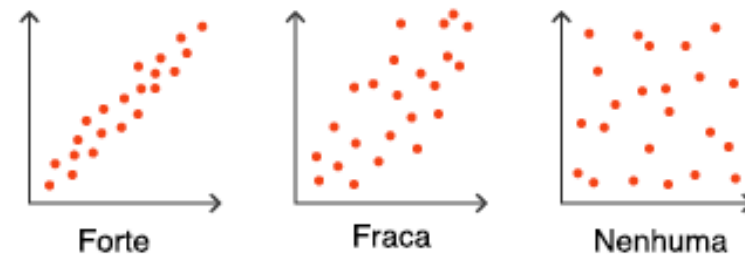
Em geral, nos gráficos, as **variáveis dependentes** são representadas no eixo vertical (Y) e as **independentes** como o eixo horizontal (X).



Tipos de correlação:



Força da correlação:

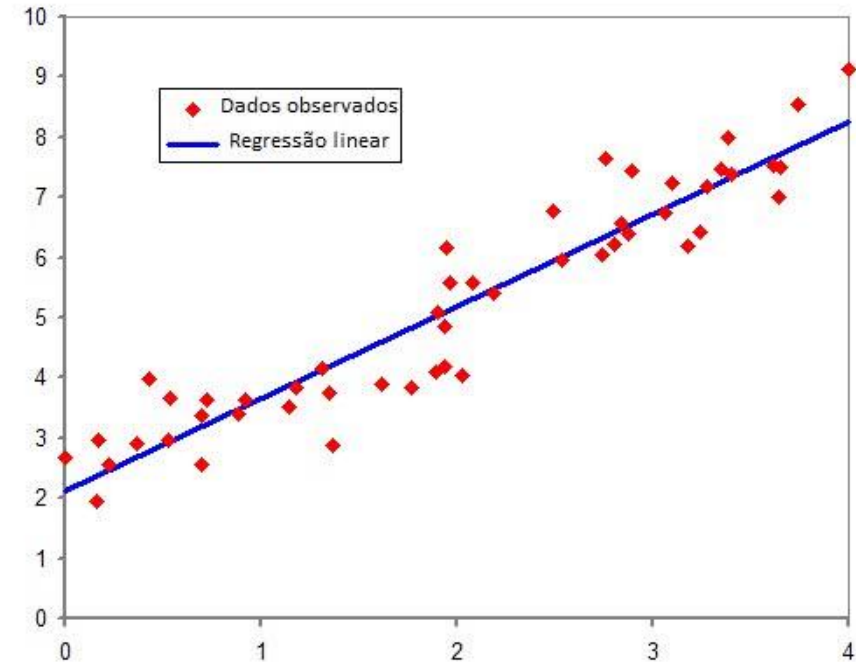


Regressão Linear

Regressão linear é denominada dessa forma por ser uma reta traçada a partir de uma relação em um **diagrama de dispersão**.

Essa reta resume uma relação entre os dados de duas variáveis e também **pode ser utilizada para realizar previsões**.

Ou seja: dado X e Y, quanto que X explica Y. Para isso, a regressão linear utiliza os pontos de dados para encontrar a melhor linha de ajuste para modelar essa relação.



A linha traçada pode ser representada pela equação, $Y_i = \alpha + \beta * X_i + \epsilon_i$, onde y é a variável explicada (dependente) e representa o que o modelo tentará prever; α é a constante, representa a interceptação da reta com o eixo vertical; β representa a inclinação em relação à variável explicativa; X é a variável explicativa (independente) e ϵ representa os valores residuais e possíveis erros.

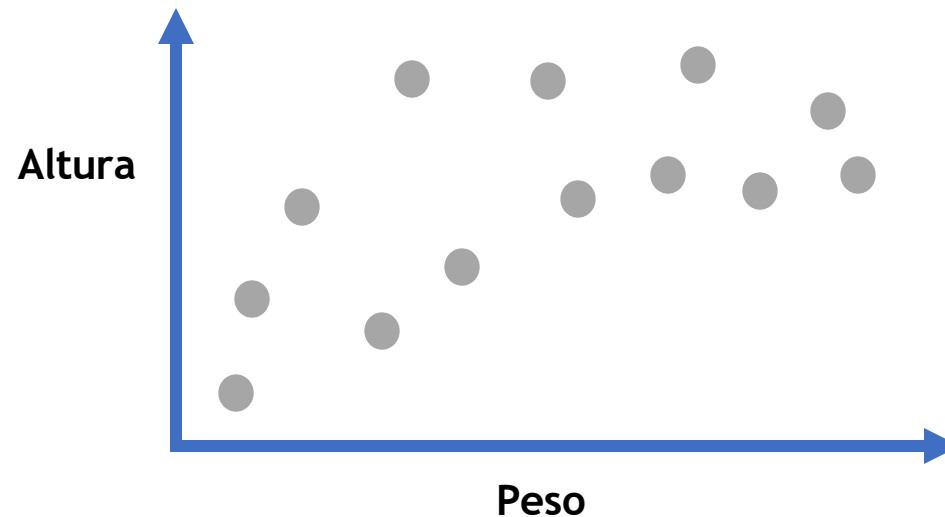
Regressão Linear - Exemplo de Código

```
1. from sklearn import linear_model, datasets
2.
3. #digit dataset from sklearn
4. digits = datasets.load_digits()
5.
6. #create the LinearRegression model
7. clf = linear_model.LinearRegression()
8.
9. #set training set
10. x, y = digits.data[:-1], digits.target[:-1]
11.
12. #train model
13. clf.fit(x, y)
14.
15. #predict
16. y_pred = clf.predict([digits.data[-1]])
17. y_true = digits.target[-1]
18.
19. print(y_pred)
20. print(y_true)
```

Na implementação em **Python**, utilizando essa ou outra biblioteca, você não precisa criar o algoritmo ou dominar a fórmula matemática. É necessário apenas determinar corretamente as variáveis e avaliar, anteriormente, se a **regressão linear** é a melhor forma de realizar essa previsão.

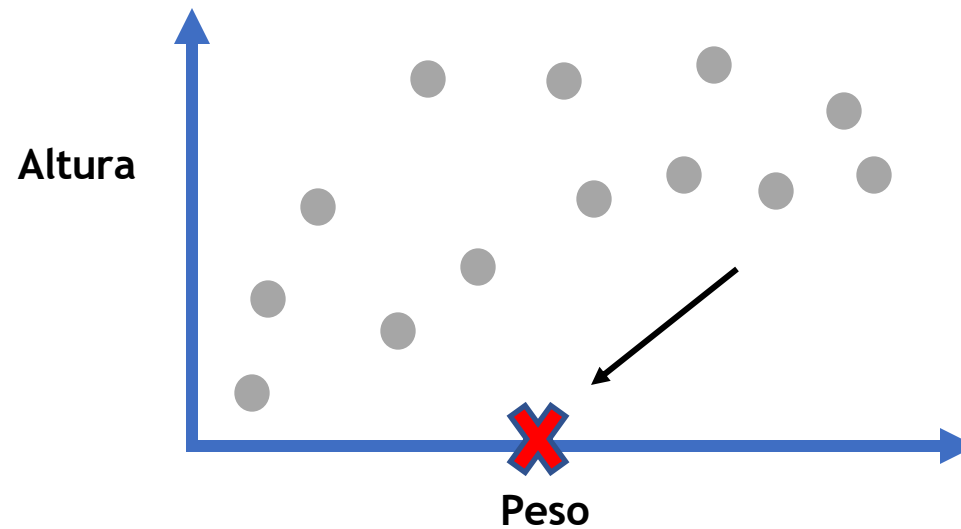
Machine Learning: Fundamentos

Imagina que temos um conjunto de dados de uns animais para estudo. No caso ratos. Dado um conjunto de dados, queremos prever a **altura** do **rato** dado seu **peso**.



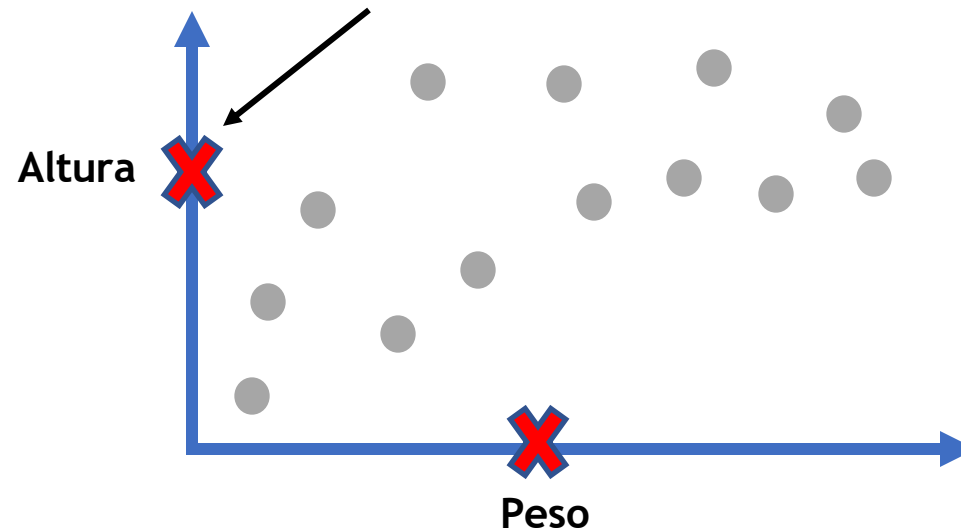
Machine Learning: Fundamentos

Dado um conjunto de dados, queremos prever a **altura** do rato dado seu **peso**.



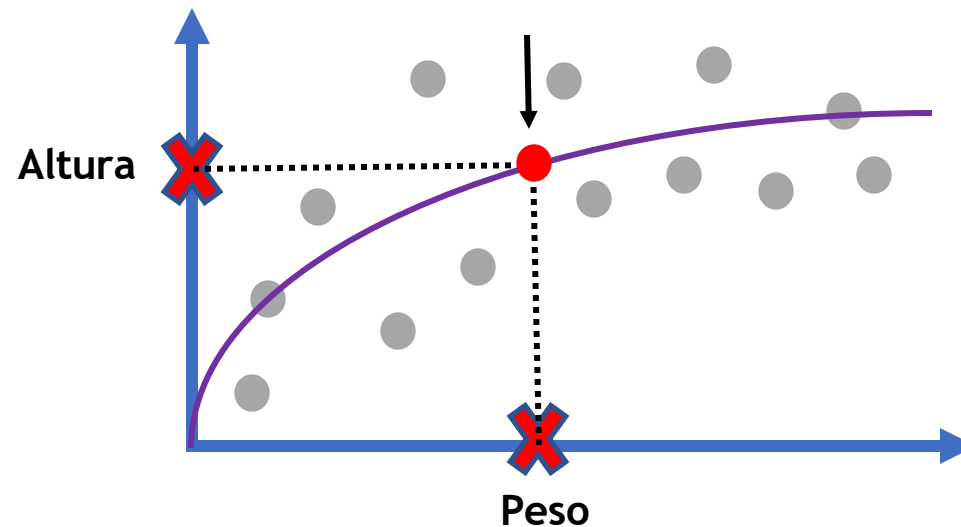
Machine Learning: Fundamentos

Dado um conjunto de dados, queremos prever a **altura** do rato dado seu **peso**.



Machine Learning: Fundamentos

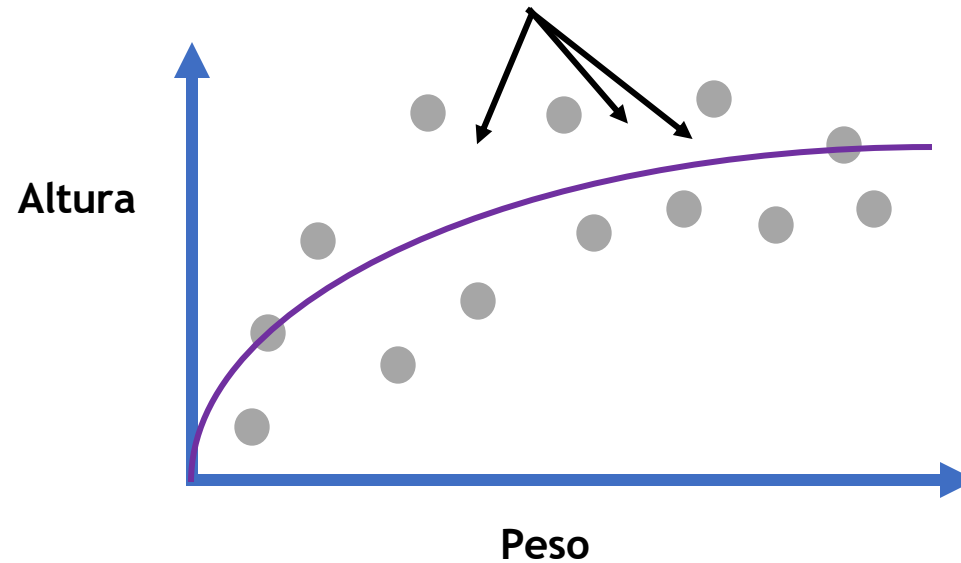
Dado um conjunto de dados, queremos prever a **altura** do rato dado seu **peso**.



No cenário ideal poderíamos traçar a equação da reta e obter a previsão exata e correta.

Machine Learning: Fundamentos

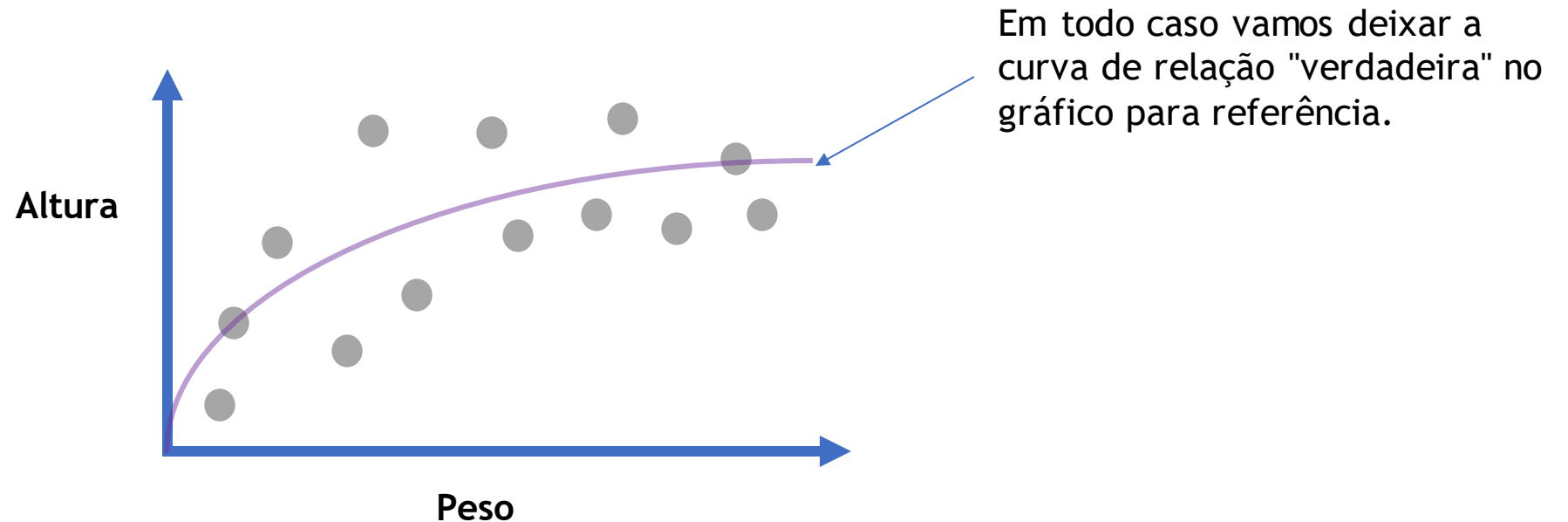
Dado um conjunto de dados, queremos prever a **altura** do rato dado seu **peso**.



Mas nesse caso não sabemos a fórmula, então vamos usar dois métodos de **machine learning** para **aproximar a relação** desses pontos

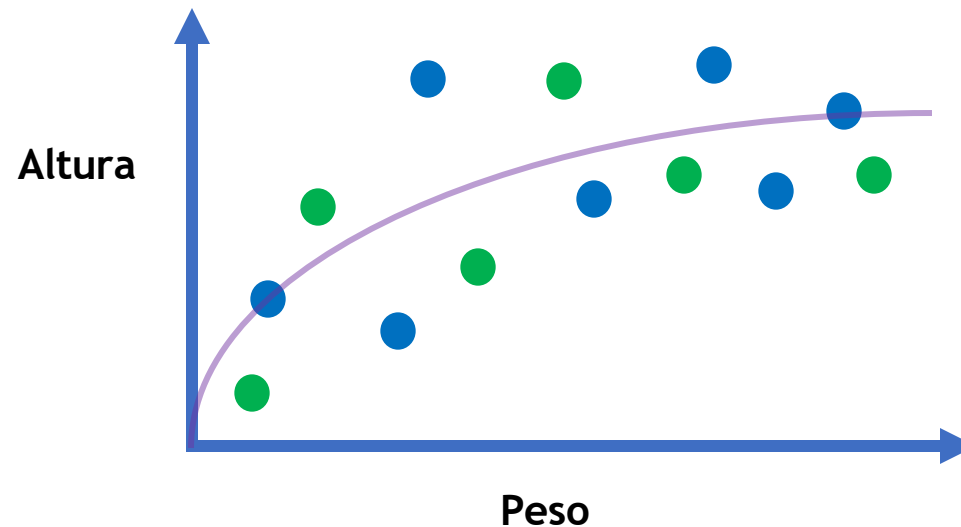
Machine Learning: Fundamentos

Dado um conjunto de dados, queremos prever a **altura** do rato dado seu **peso**.



Machine Learning: Fundamentos

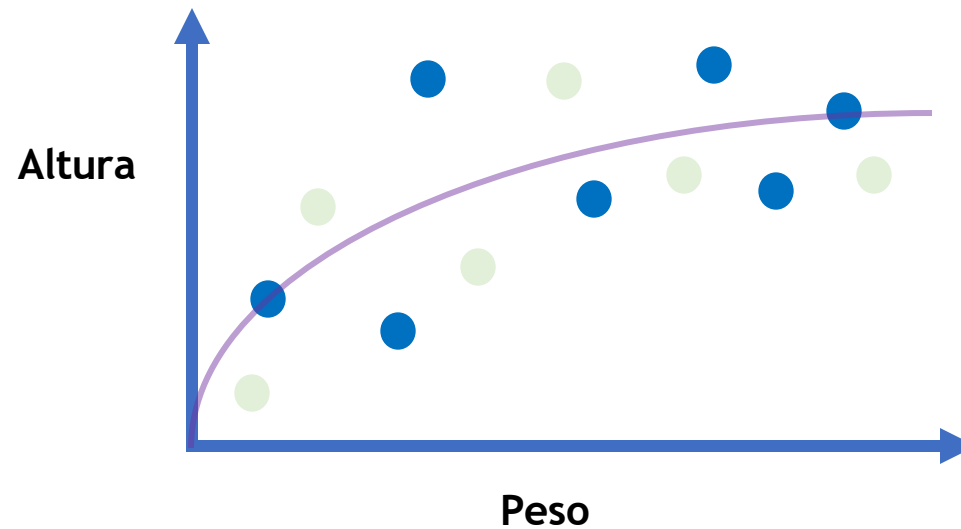
Dado um conjunto de dados, queremos prever a **altura** do rato dado seu **peso**.



Primeiramente dividimos os dados em 2 datasets, um para **treinamento** e o outro para **teste**.

Machine Learning: Fundamentos

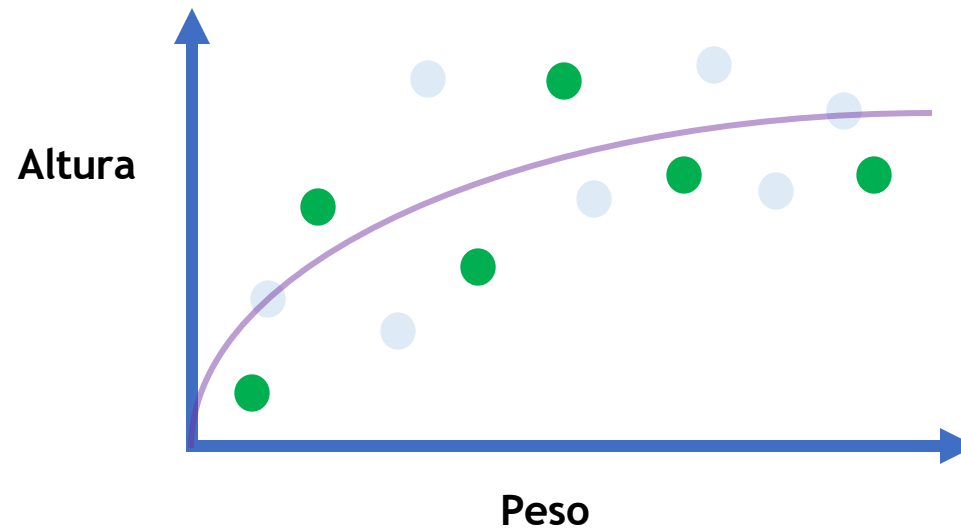
Dado um conjunto de dados, queremos prever a **altura** do rato dado seu **peso**.



Os **pontos azuis** são dos **dados de treinamento**.

Machine Learning: Fundamentos

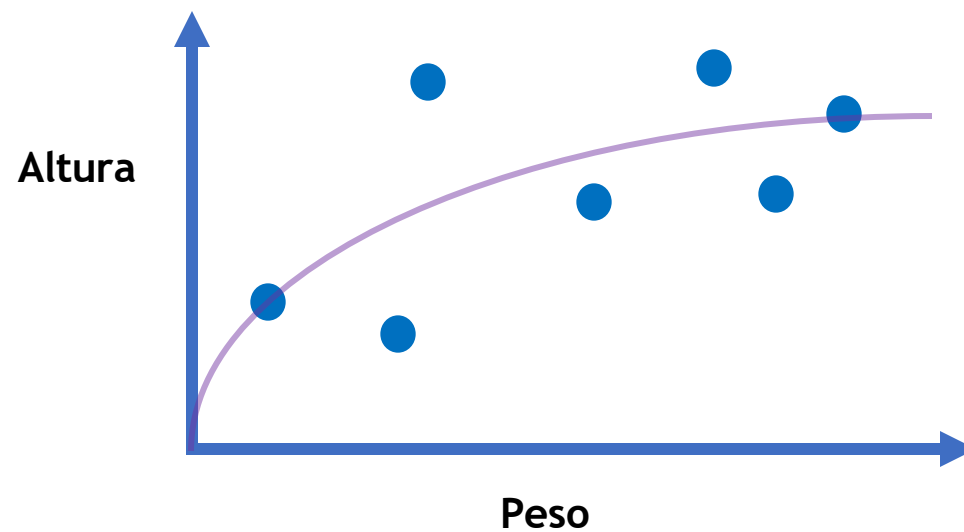
Dado um conjunto de dados, queremos prever a **altura** do rato dado seu **peso**.



Os **pontos verdes** são dos **dados de teste**.

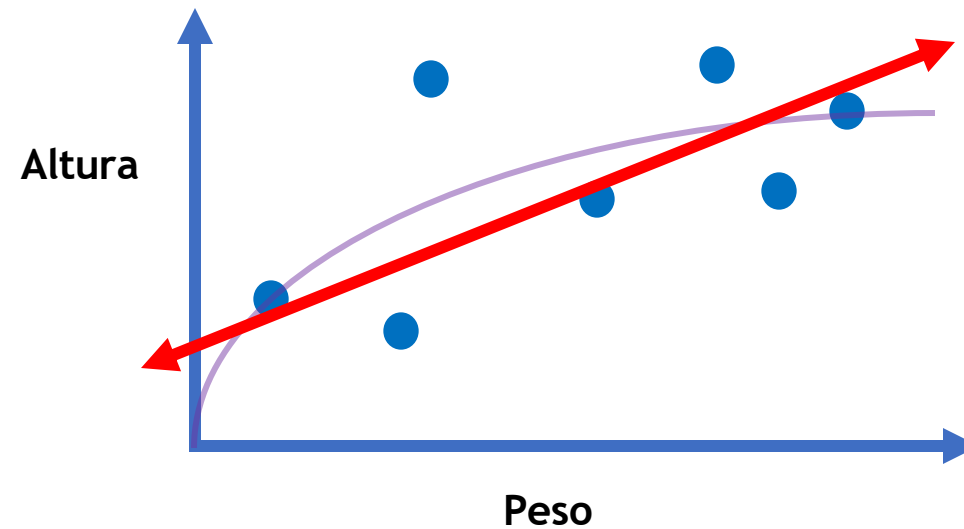
Machine Learning: Fundamentos

Aqui estão comente os dados de treinamento.



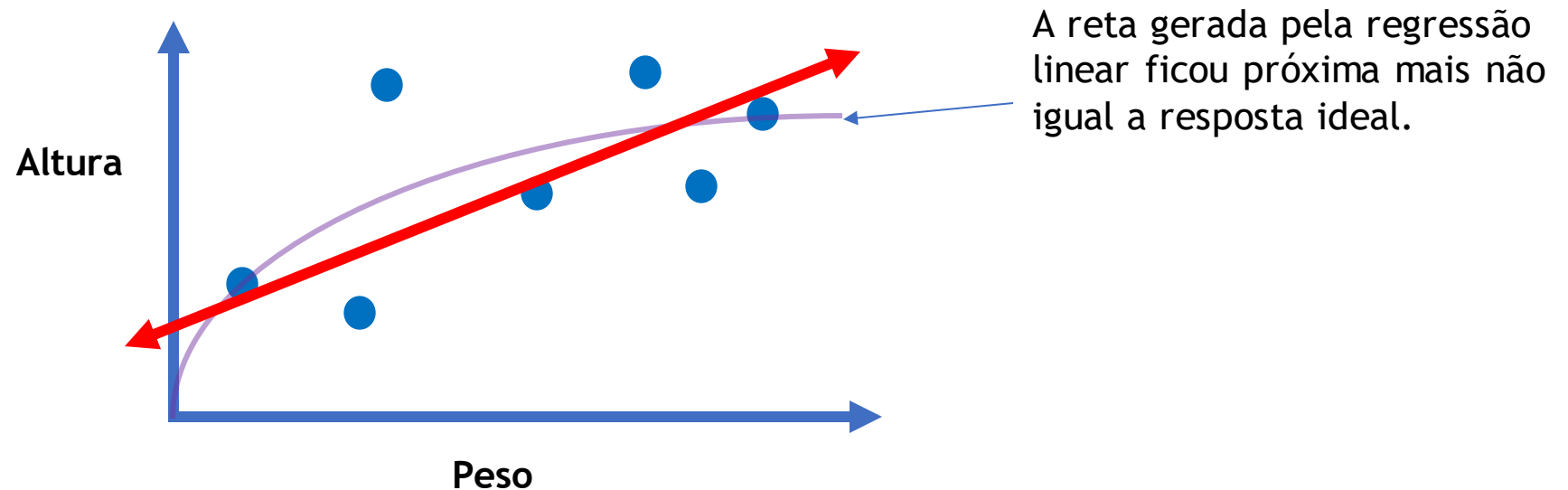
Machine Learning: Fundamentos

O primeiro algoritmo de **Machine Learning** utilizado será a **Regressão Linear**, conhecida também como "*Mínimos quadrados*"



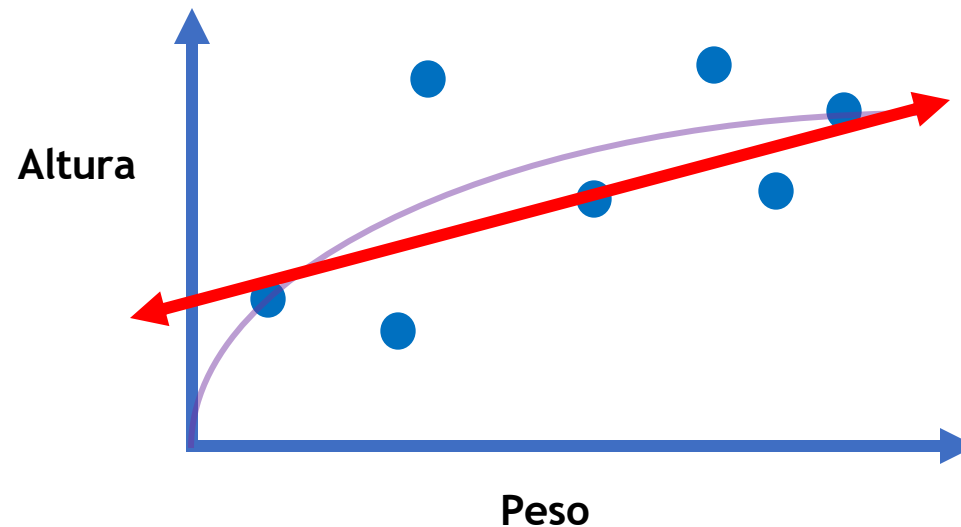
Machine Learning: Fundamentos

O primeiro algoritmo de **Machine Learning** utilizado será a **Regressão Linear**, conhecida também como "*Mínimos quadrados*"



Machine Learning: Fundamentos

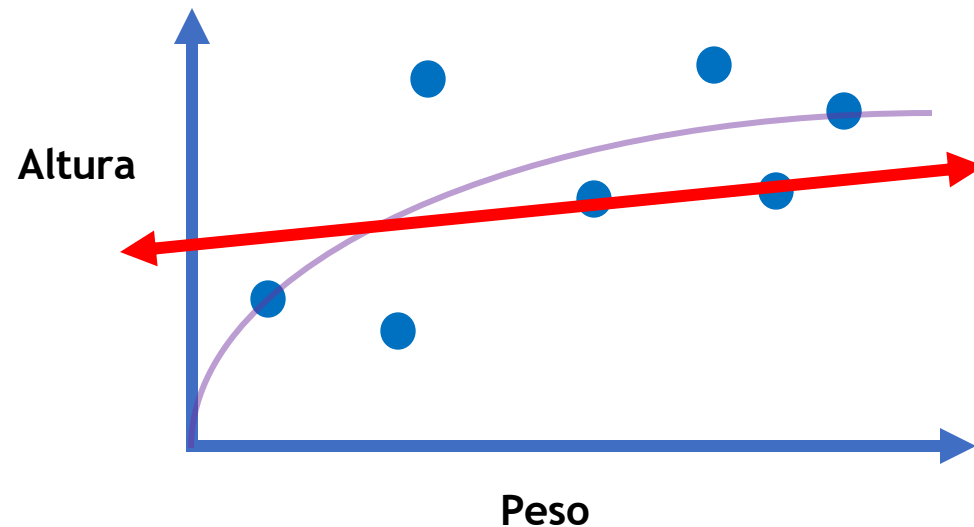
O primeiro algoritmo de **Machine Learning** utilizado será a **Regressão Linear**, conhecida também como "*Mínimos quadrados*"



Não importar o quanto se tende ajustar "fit".. Essa solução nunca será uma curva.

Machine Learning: Fundamentos

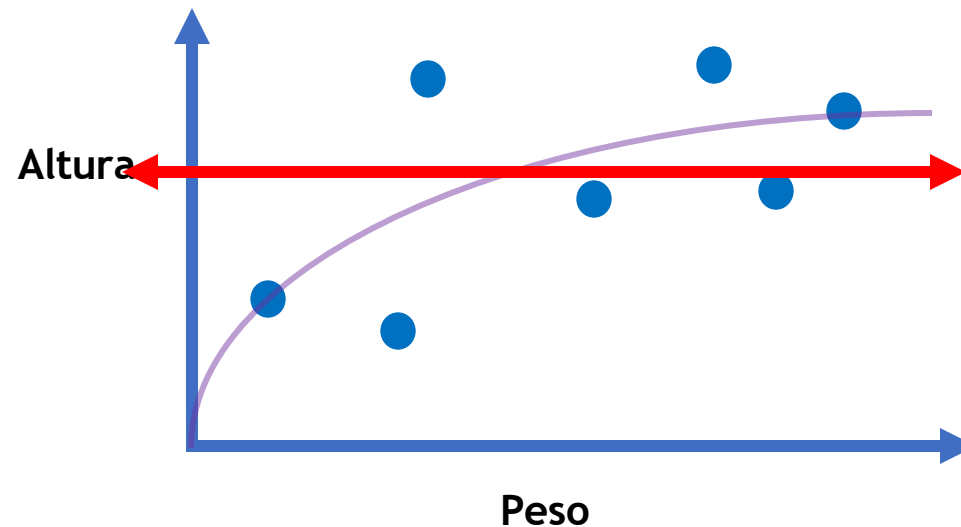
O primeiro algoritmo de **Machine Learning** utilizado será a **Regressão Linear**, conhecida também como "*Mínimos quadrados*"



Não importar o quanto se tende ajustar "fit".. Essa solução nunca será uma curva.

Machine Learning: Fundamentos

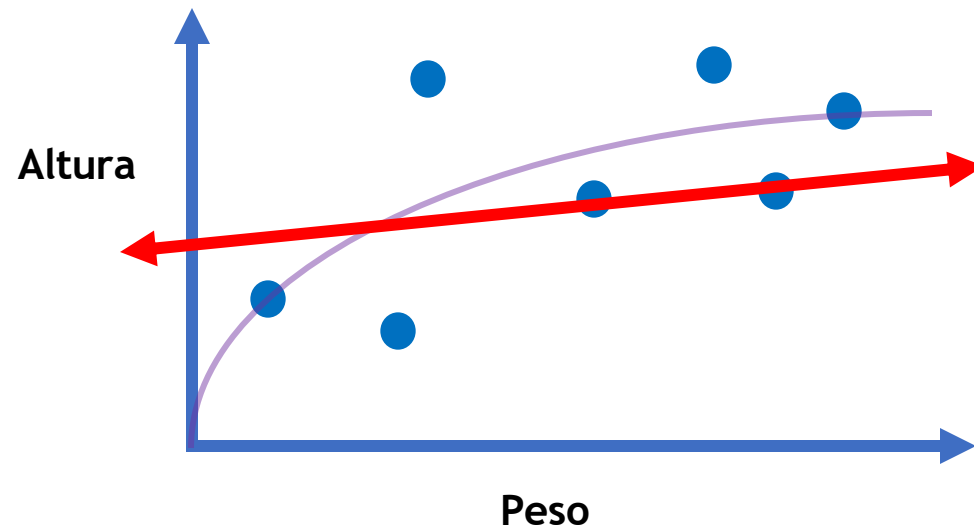
O primeiro algoritmo de **Machine Learning** utilizado será a **Regressão Linear**, conhecida também como "*Mínimos quadrados*"



Não importar o quanto se tende ajustar "fit".. Essa solução nunca será uma curva.

Machine Learning: Fundamentos

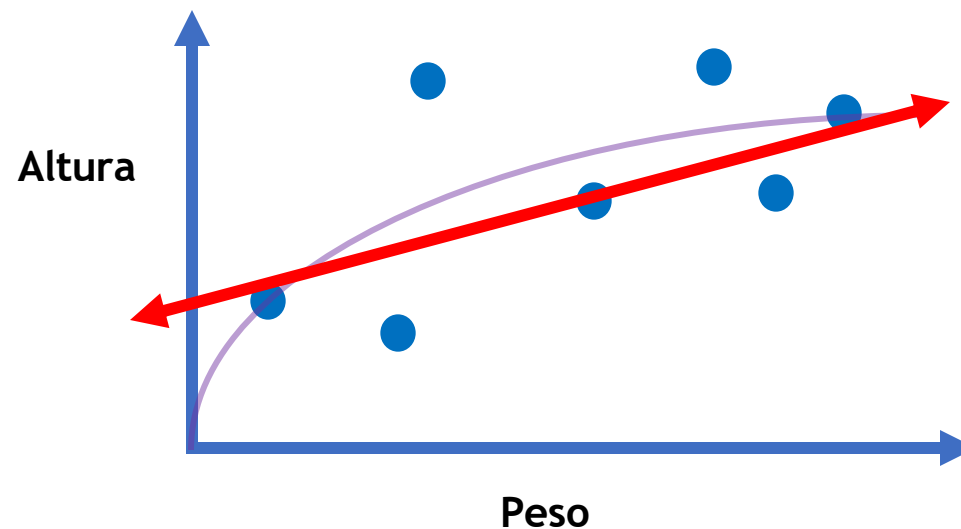
O primeiro algoritmo de **Machine Learning** utilizado será a **Regressão Linear**, conhecida também como "*Mínimos quadrados*"



Não importar o quanto se tende ajustar "fit".. Essa solução nunca será uma curva.

Machine Learning: Fundamentos

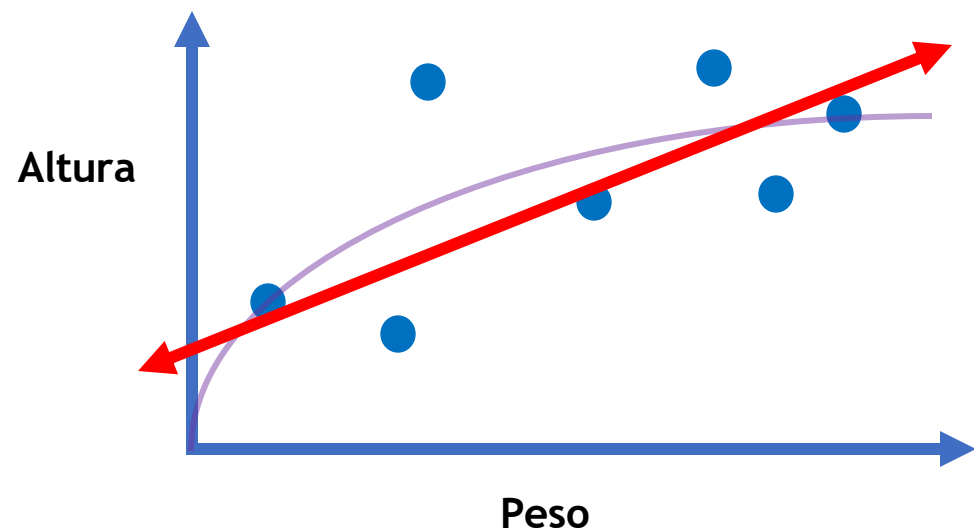
O primeiro algoritmo de **Machine Learning** utilizado será a **Regressão Linear**, conhecida também como "*Mínimos quadrados*"



Não importar o quanto se tende a ajustar "**fit**"... Essa solução nunca será uma curva.

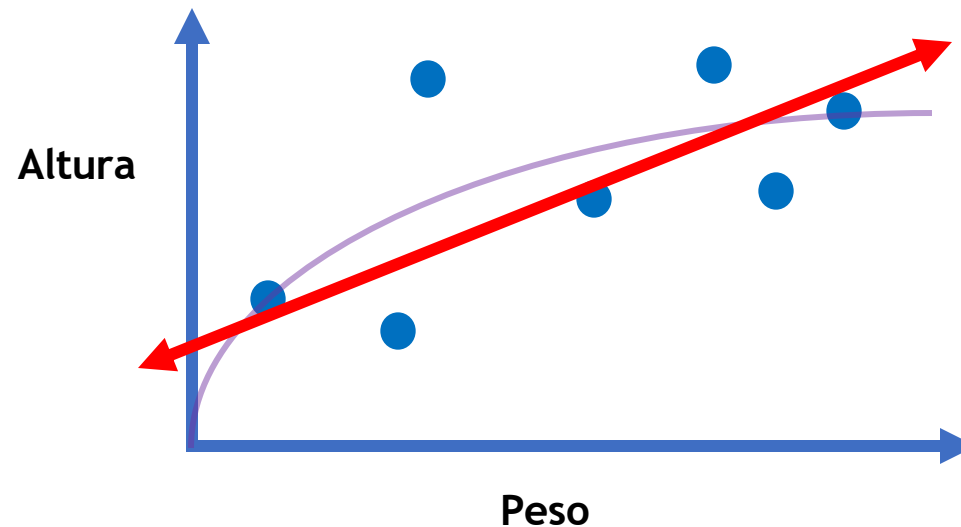
Machine Learning: Fundamentos

A incapacidade de um algoritmo de aprendizagem de máquina, como a "Regressão Linear", de conseguir representar a verdadeira relação esperada é chamada de **Viés (Bias)**.



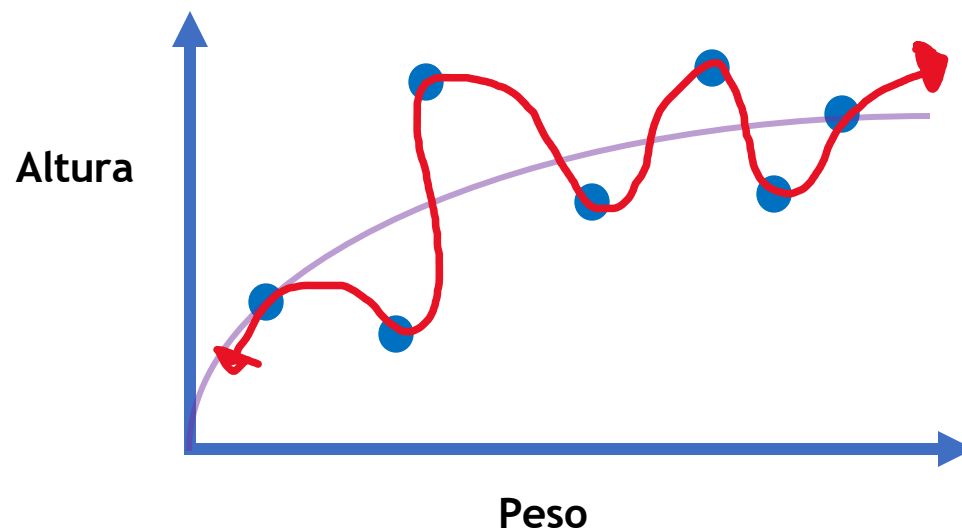
Machine Learning: Fundamentos

Como a linha reta não pode ser curvada, como a relação "verdadeira". Dizemos que ela possui uma quantidade relativamente alta de **viés**.



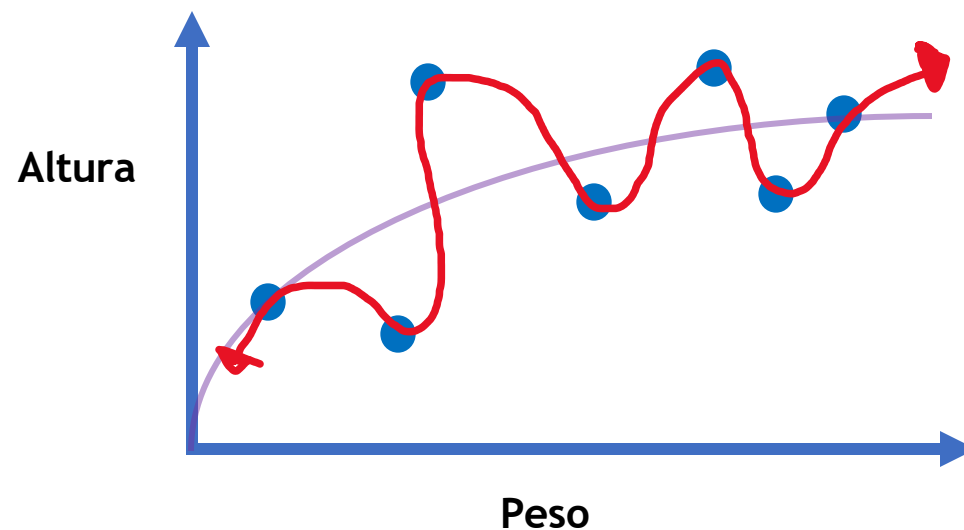
Machine Learning: Fundamentos

Um outro método de aprendizado de máquina poderia ajustar uma "linha sinuosa" ao conjunto de treinamento. Essa linha cheia de curvas é super flexível e segue o conjunto de treinamento ao longo do arco da relação verdadeira.



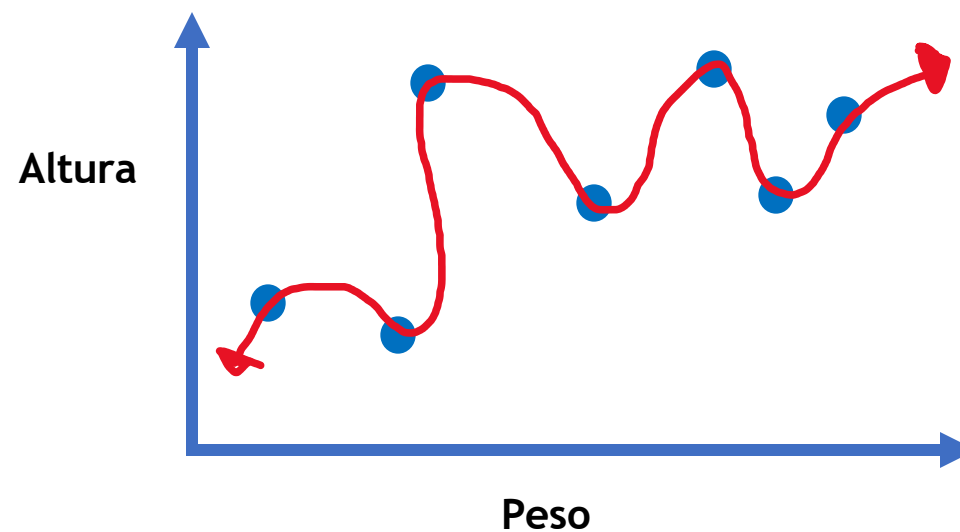
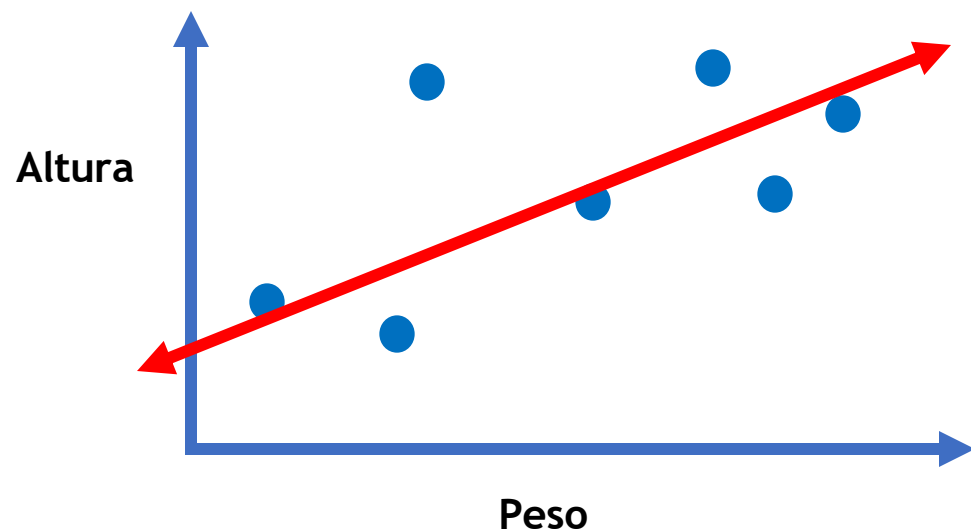
Machine Learning: Fundamentos

Como a linha sinuosa consegue seguir o arco da relação verdadeira entre peso e altura, dizemos que ela possui **pouco viés**.



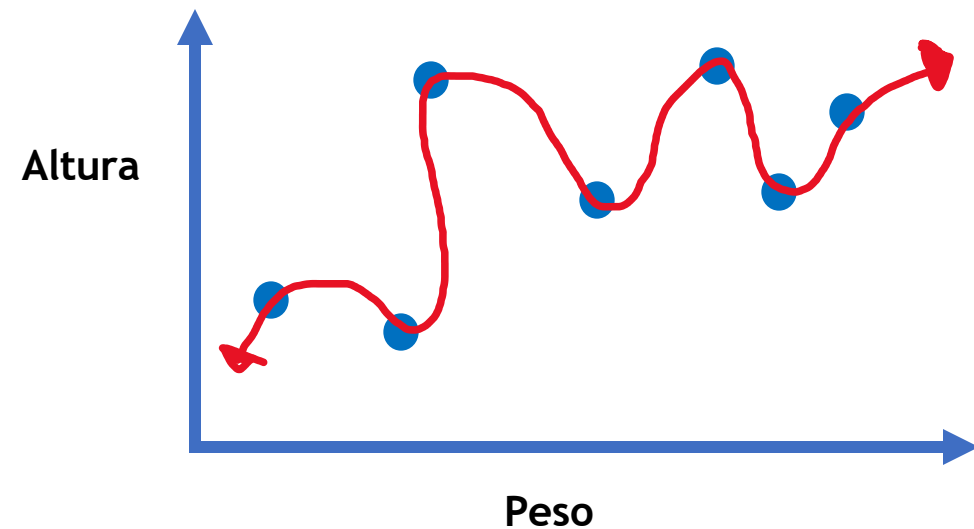
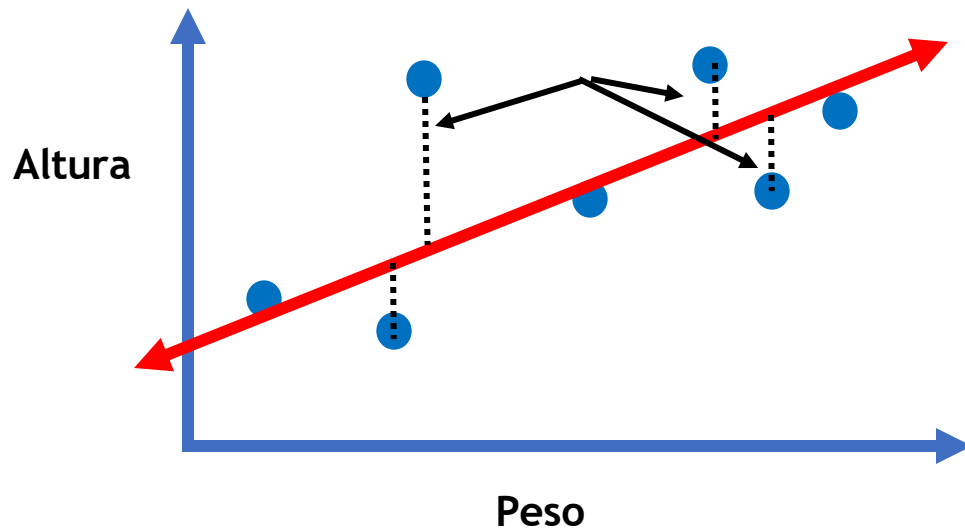
Machine Learning: Fundamentos

Podemos comparar o quanto a **linha reta** e a **linha com curvas** se **ajusta** ao **dataset de treinamento** calculando a soma dos quadrados.



Machine Learning: Fundamentos

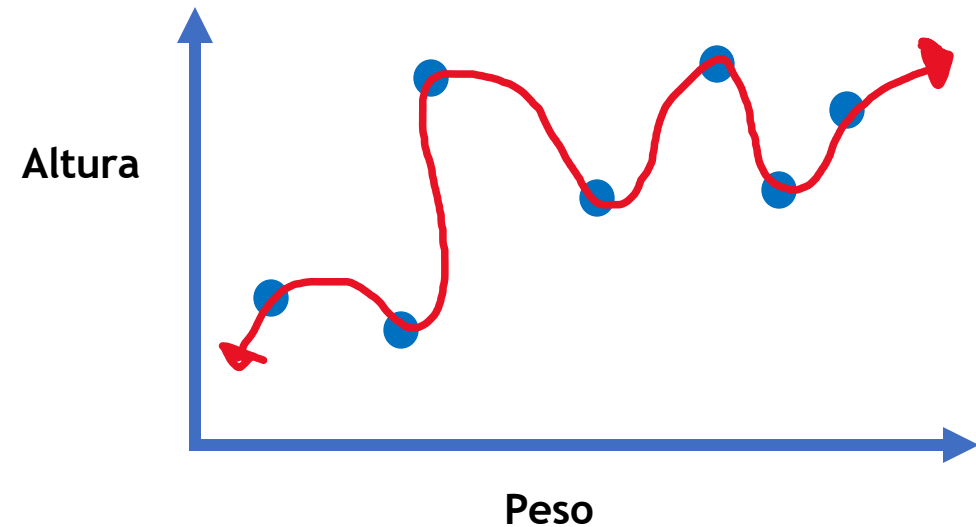
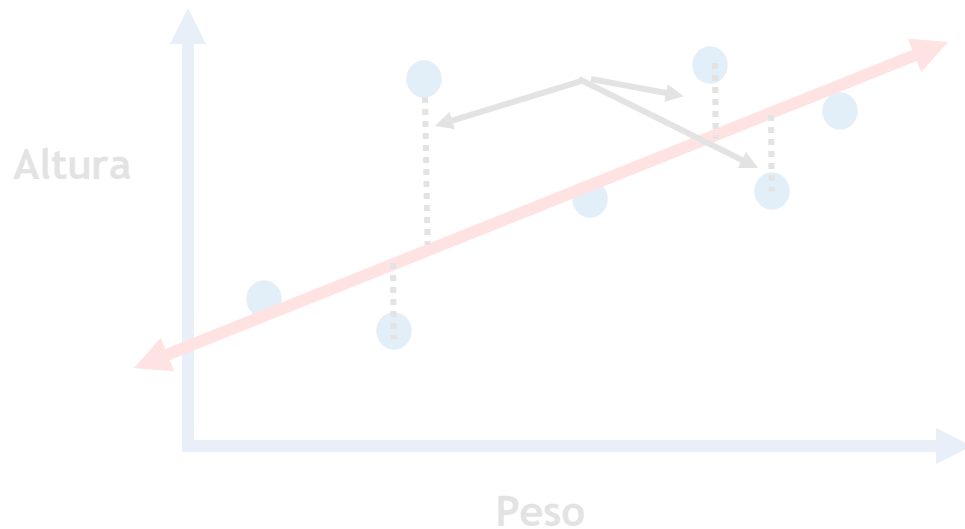
Em outras palavras, medimos todas as distâncias das linhas de ajuste aos dados, elevamos ao quadrado e somamos todas elas.



OBS: Essas distâncias são elevadas ao quadrado, assim as distâncias "negativas" não anulam os efeitos das positivas.

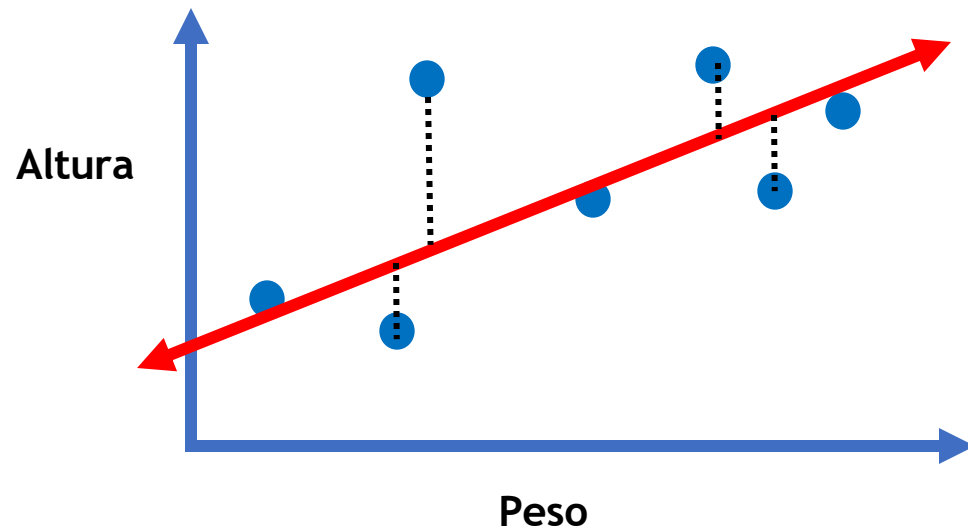
Machine Learning: Fundamentos

Note que a linha com curvas se ajusta tão bem aos dados, que as distâncias entre ela e os dados são todas ZERO.

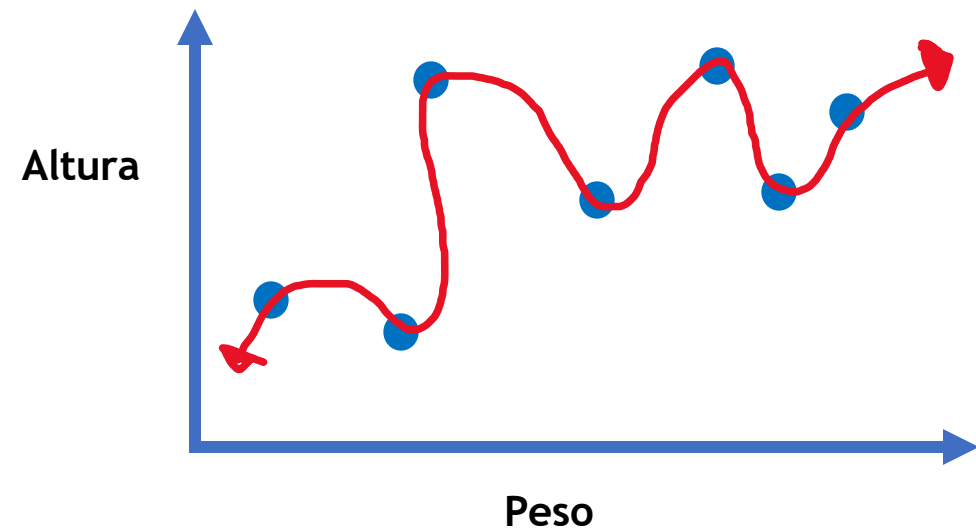


Machine Learning: Fundamentos

Na disputa para ver se a linha reta se ajusta aos dados do conjunto de treinamento melhor que a linha curva, a **Linha com curvas VENCE!**

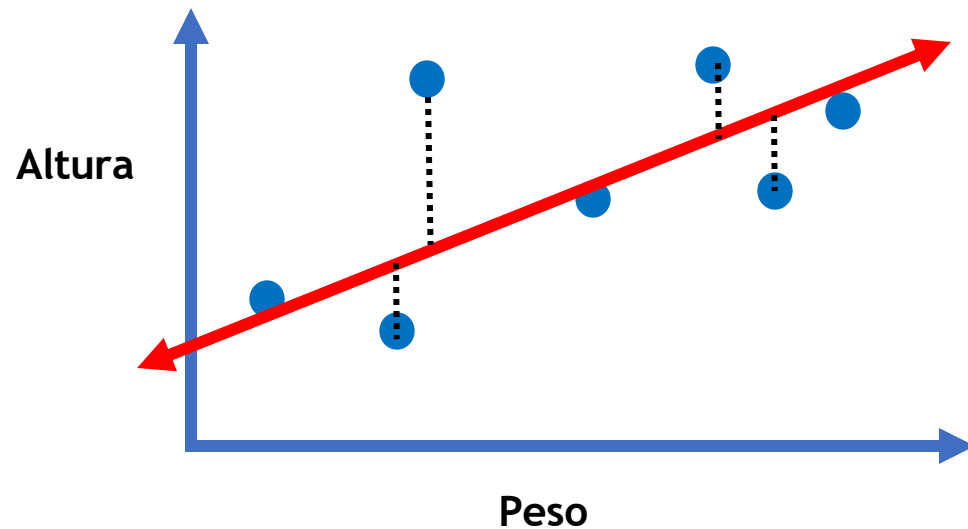


Vs.

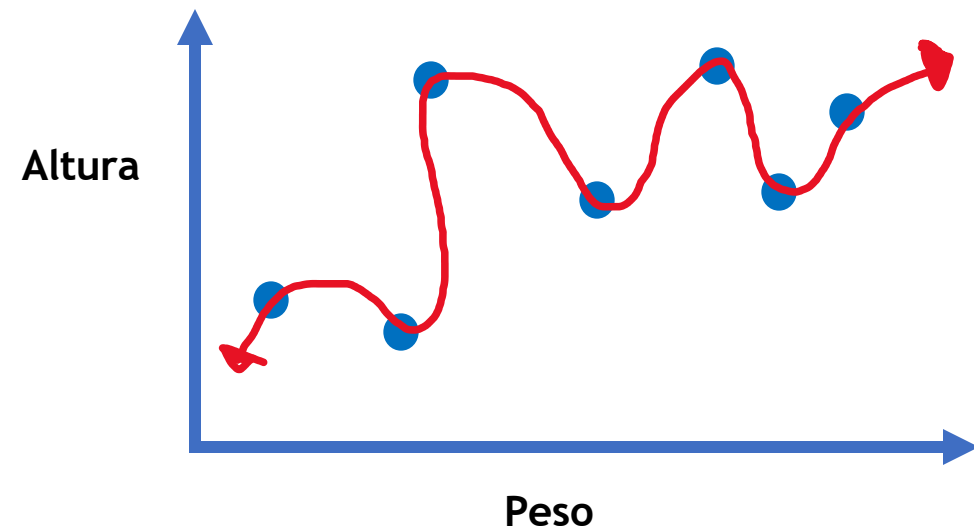


Machine Learning: Fundamentos

Mas veja que até agora somente calculamos a soma dos quadrados para o **Dataset de Treinamento**.

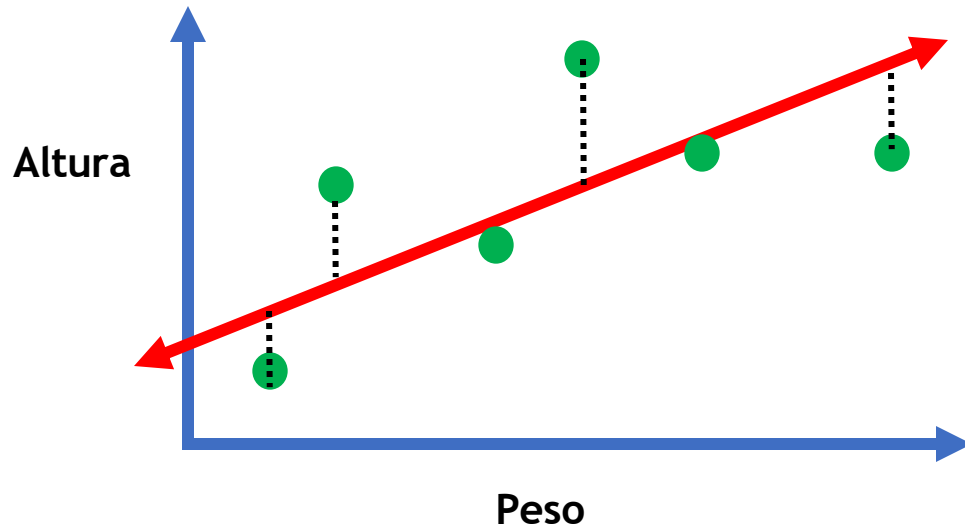


Vs.

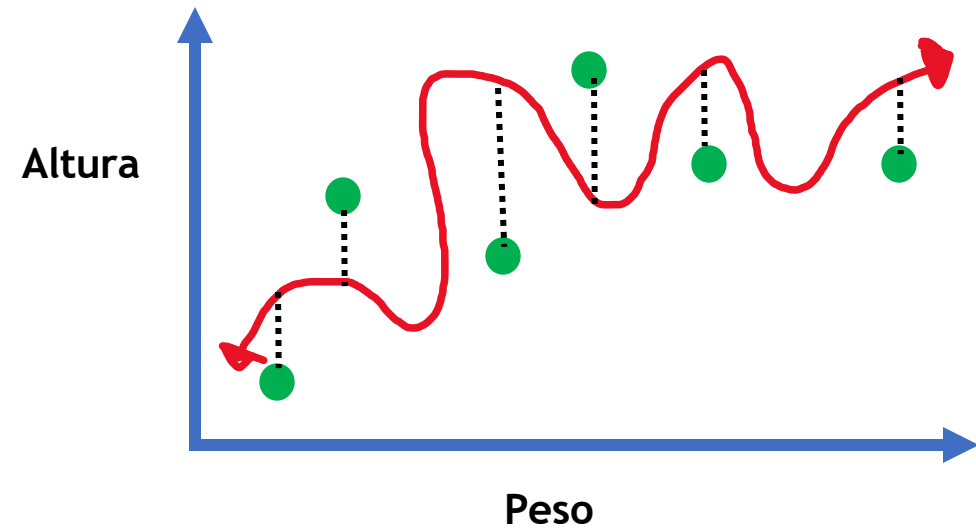


Machine Learning: Fundamentos

Nós também temos que realizar esse procedimento para o **Dataset de Teste**.

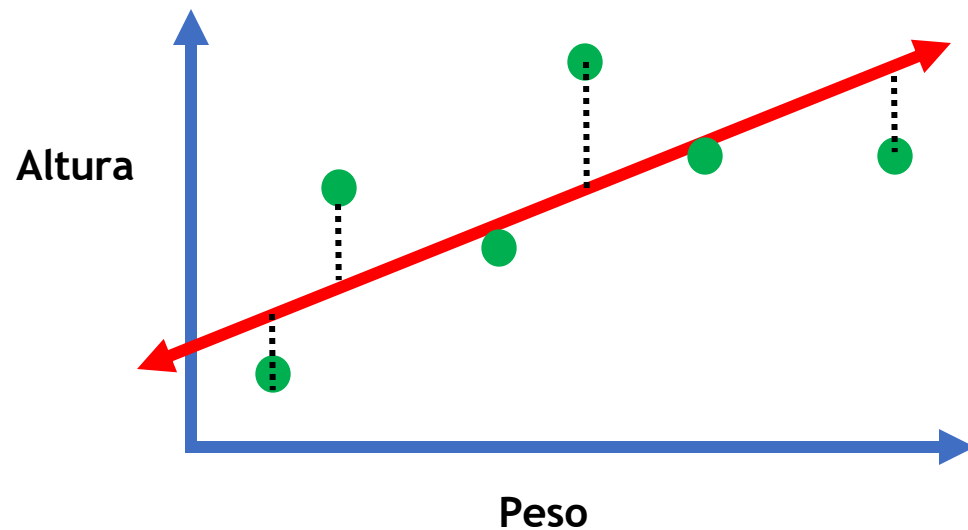


Vs.

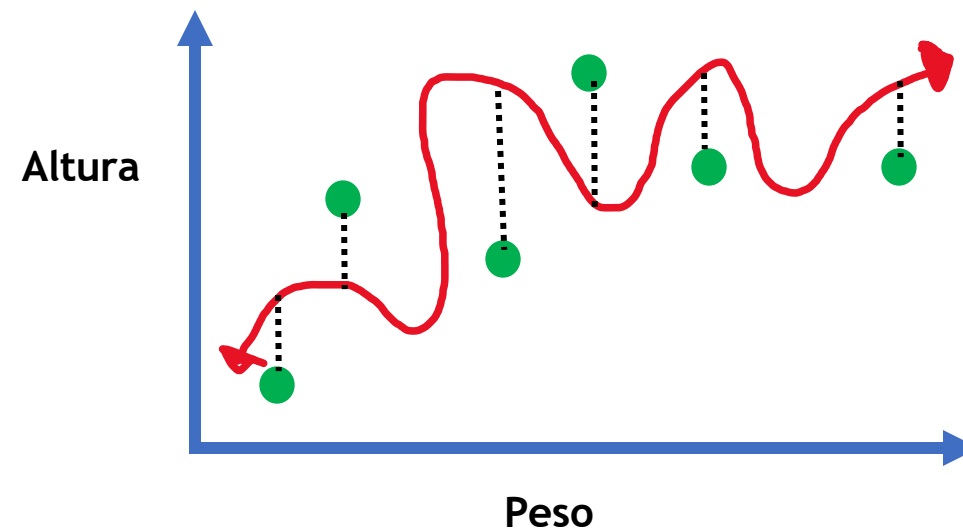


Machine Learning: Fundamentos

E agora nesta disputa podemos ver que a Linha Reta se **ajusta melhor** aos **Dataset de Teste** que a Linha com Curvas.

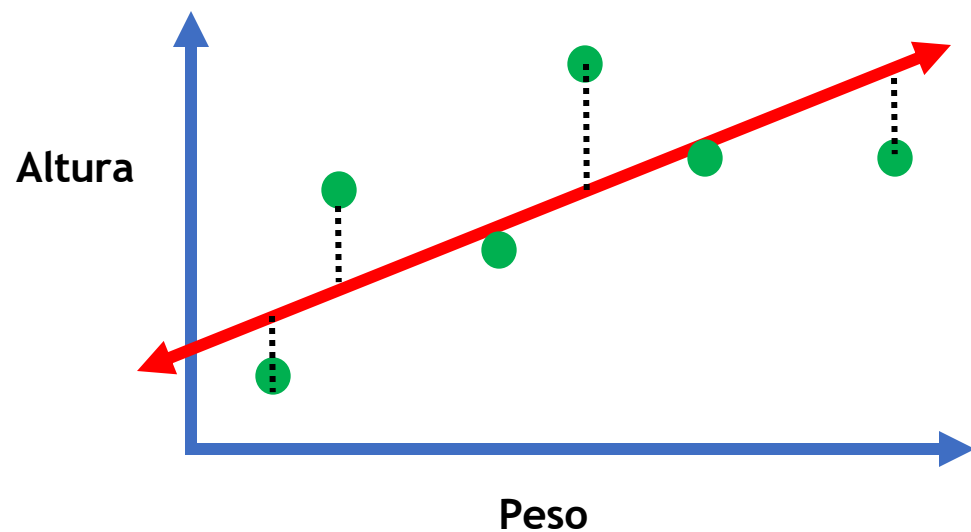


Vs.

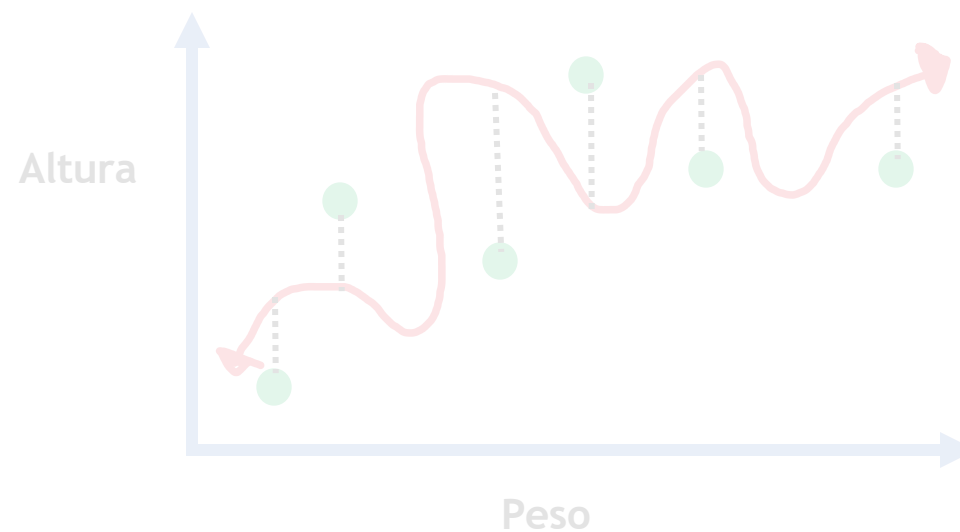


Machine Learning: Fundamentos

A linha reta **VENCE** nos dados de teste.

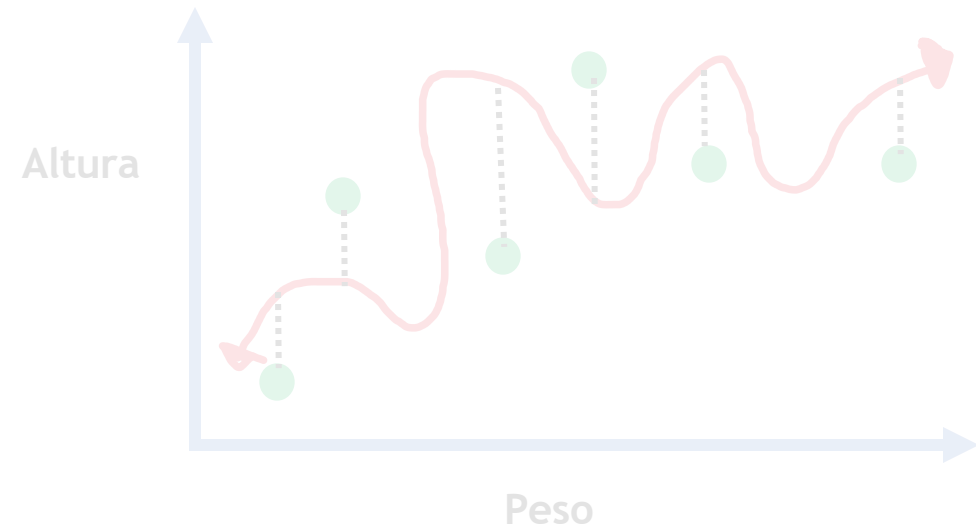
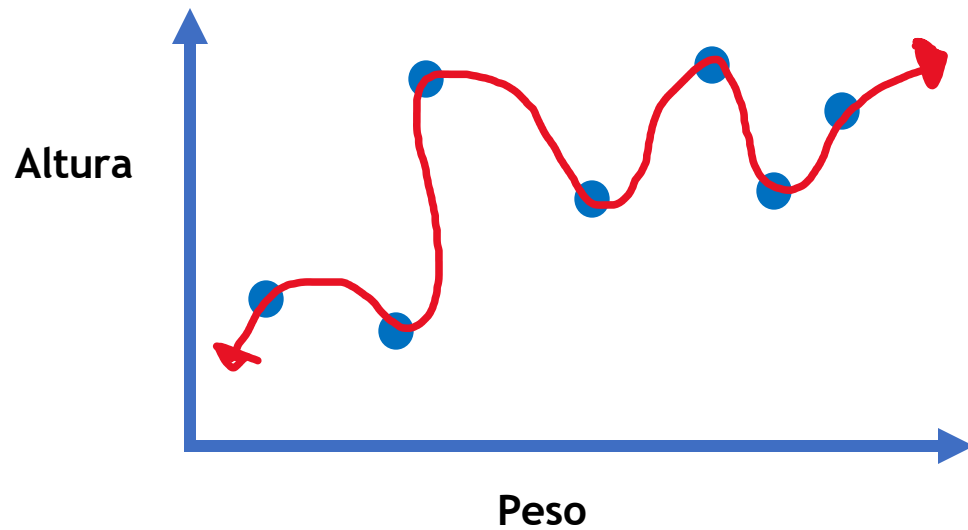


Vs.



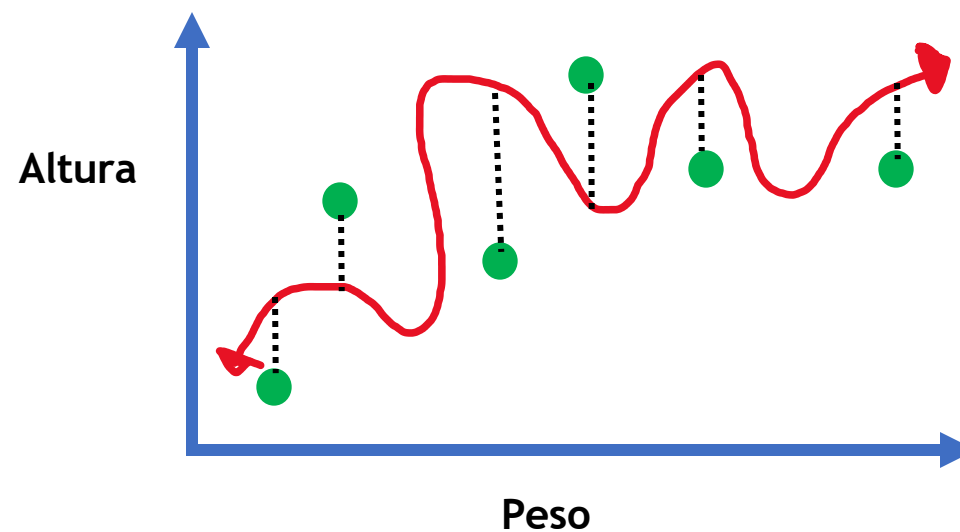
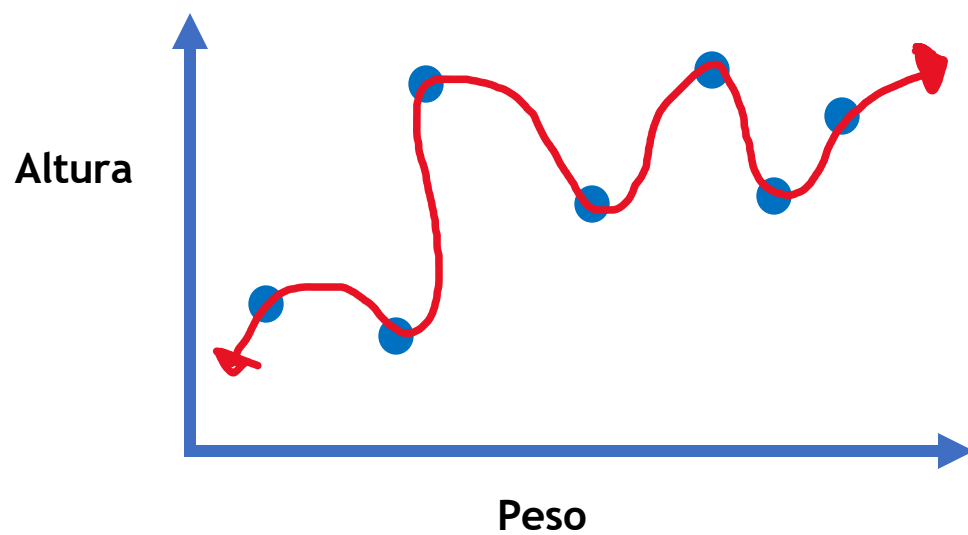
Machine Learning: Fundamentos

Apesar da Linha com Curvas realizar um ótimo trabalho na **fase de treinamento**...



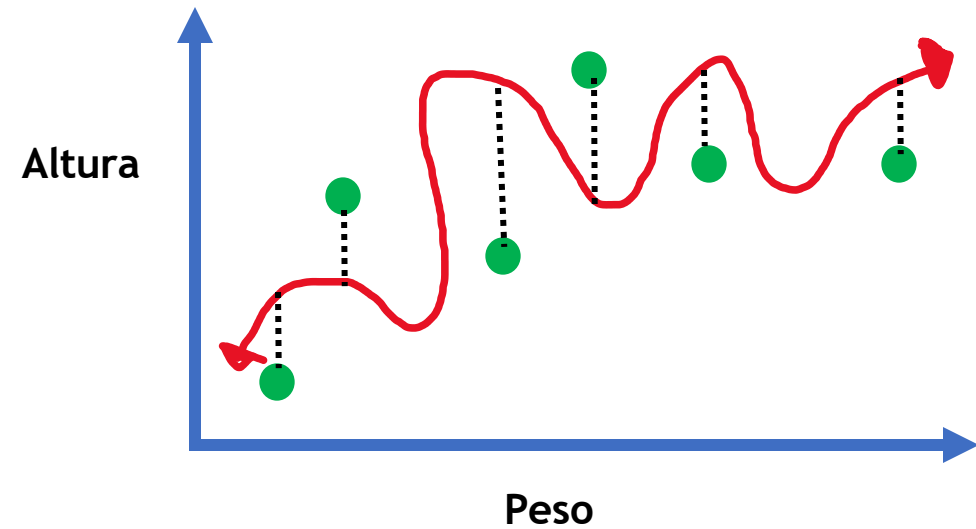
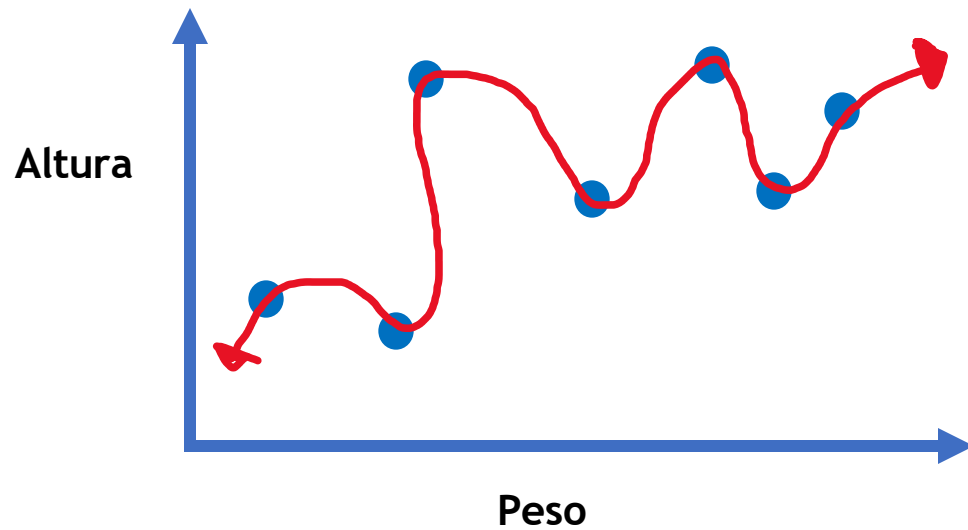
Machine Learning: Fundamentos

...ela foi muito ruim na **fase de testes**...



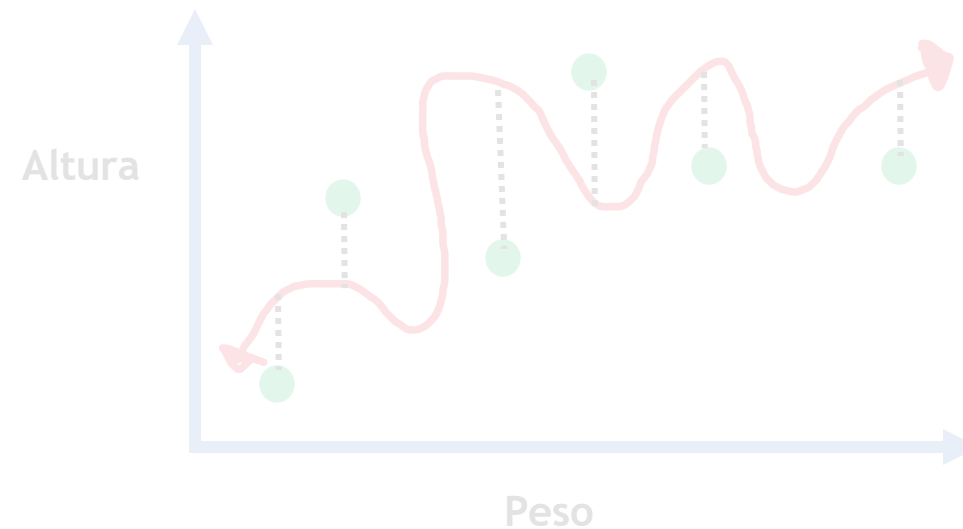
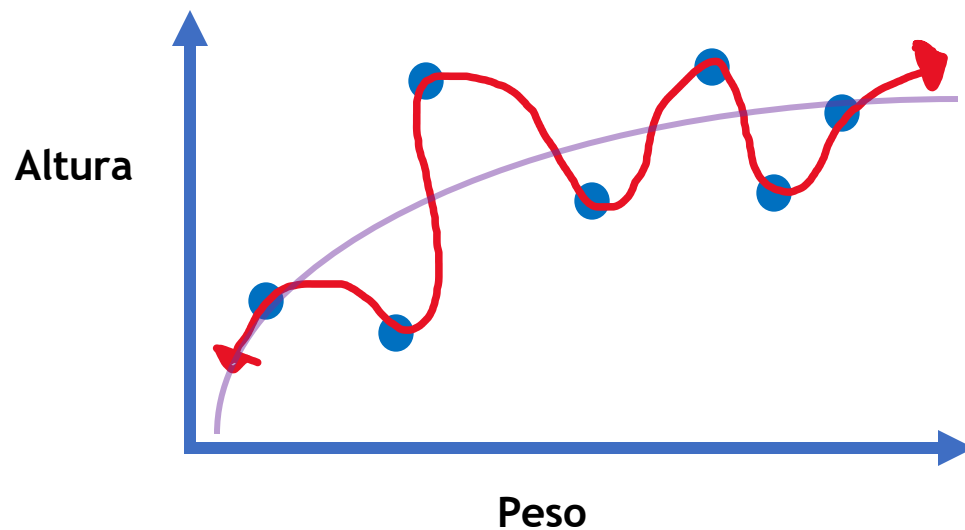
Machine Learning: Fundamentos

Nos termos de Aprendizagem de Máquina, a diferença de ajustes entre os conjuntos de dados é chamado de **Variância**.



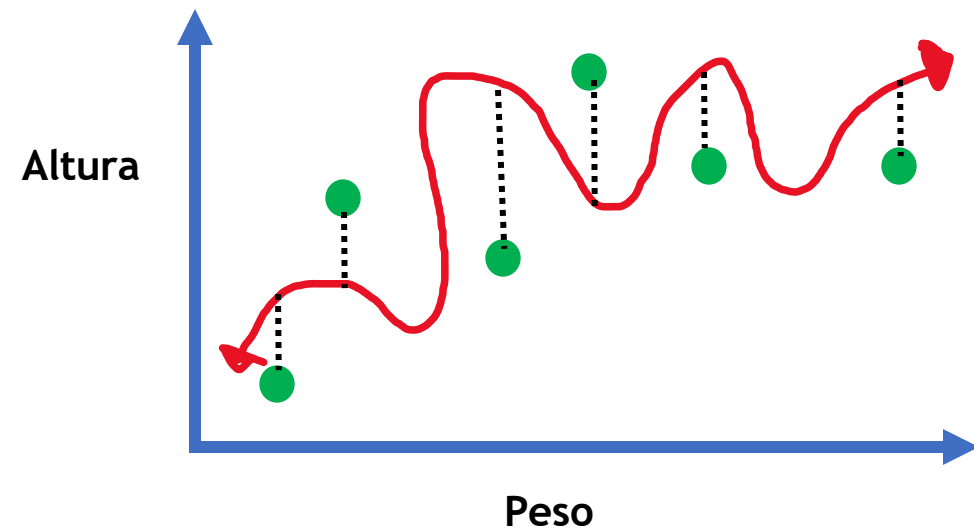
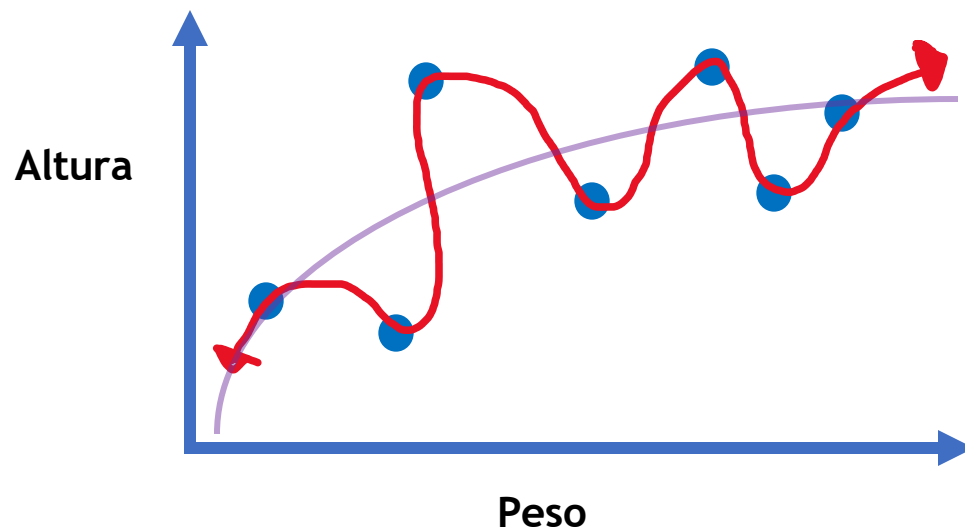
Machine Learning: Fundamentos

A Linha com Curvas possui um **baixo viés**, pois é bastante flexível e consegue se ajustar bem à curva da relação "resposta".



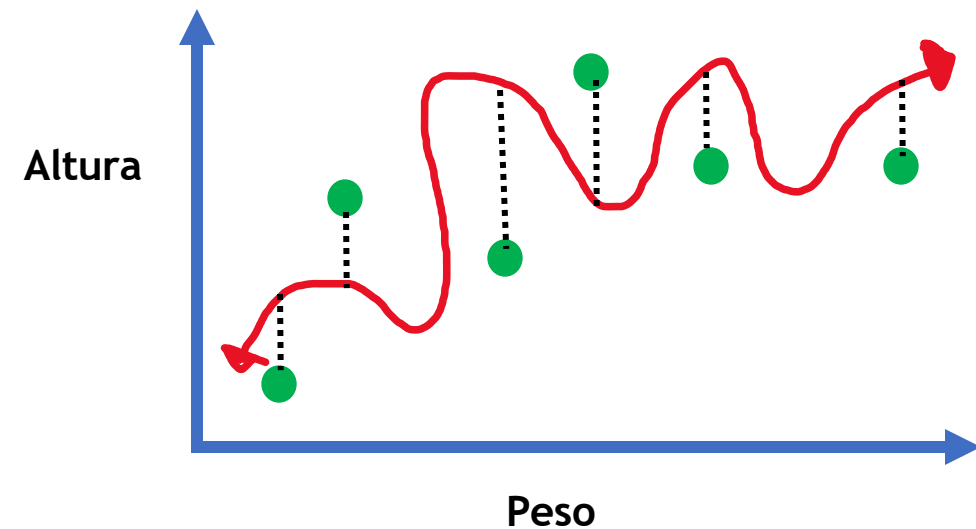
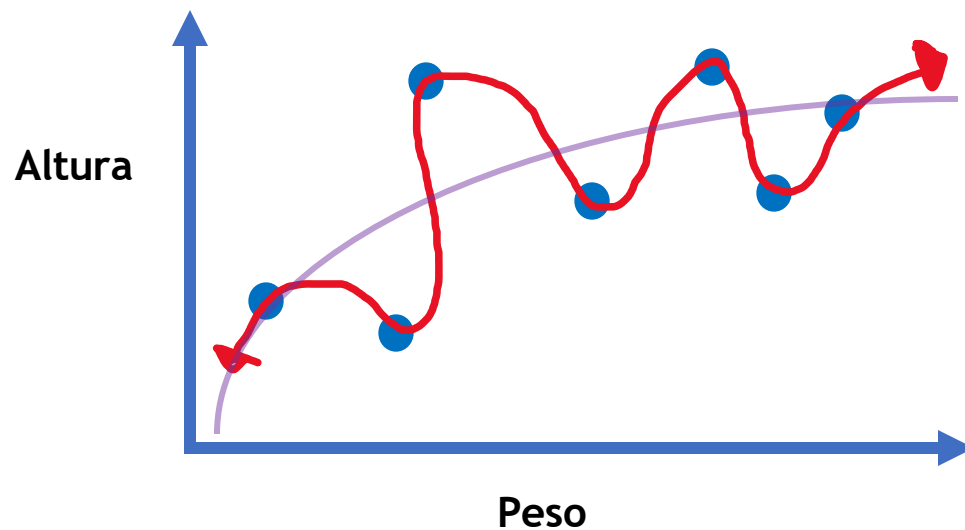
Machine Learning: Fundamentos

Mas Linha com Curvas possui uma alta **variância** pois a diferença da soma dos quadrados entre os dois conjuntos de teste é muito alta.



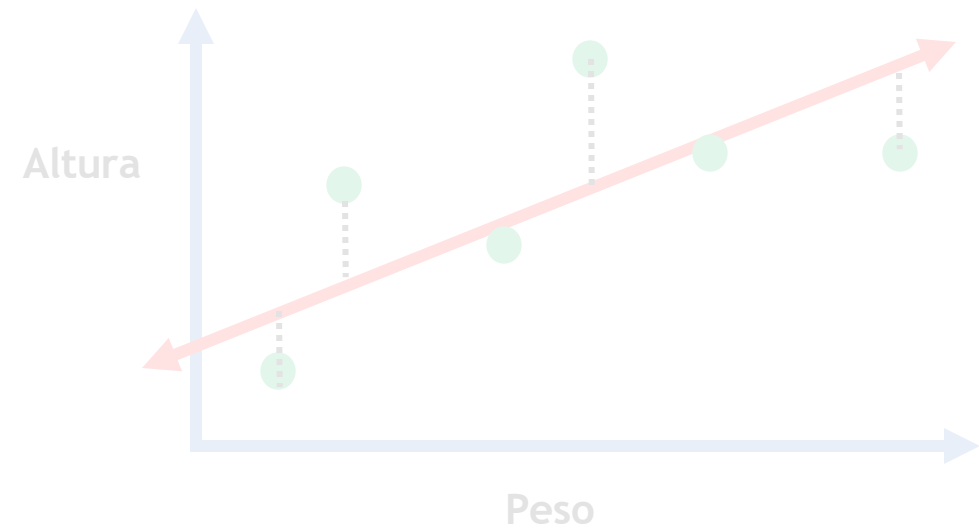
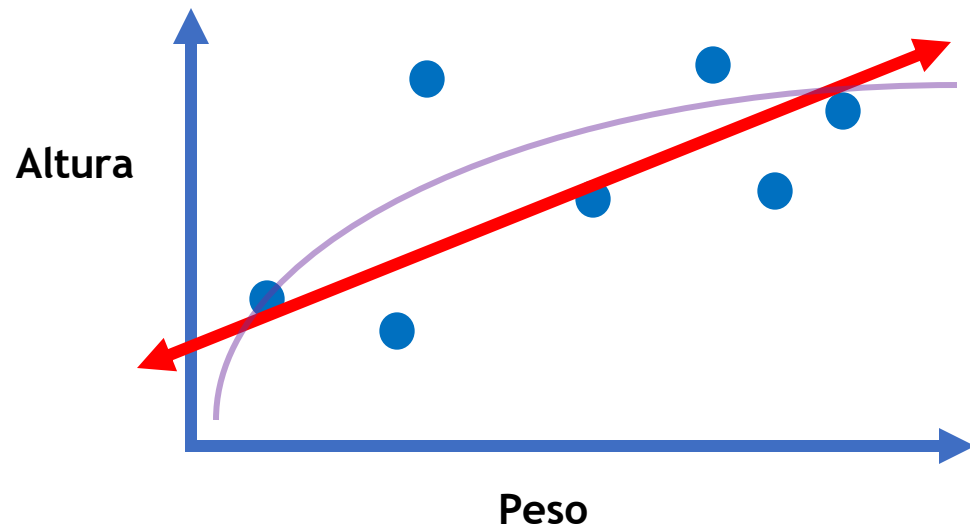
Machine Learning: Fundamentos

É difícil prever o comportamento da **Linha com Curvas** na prática, pois ela pode se ajudar bem no momento de **treino** ou de **teste** a depender do conjunto específico que está sendo analisado.



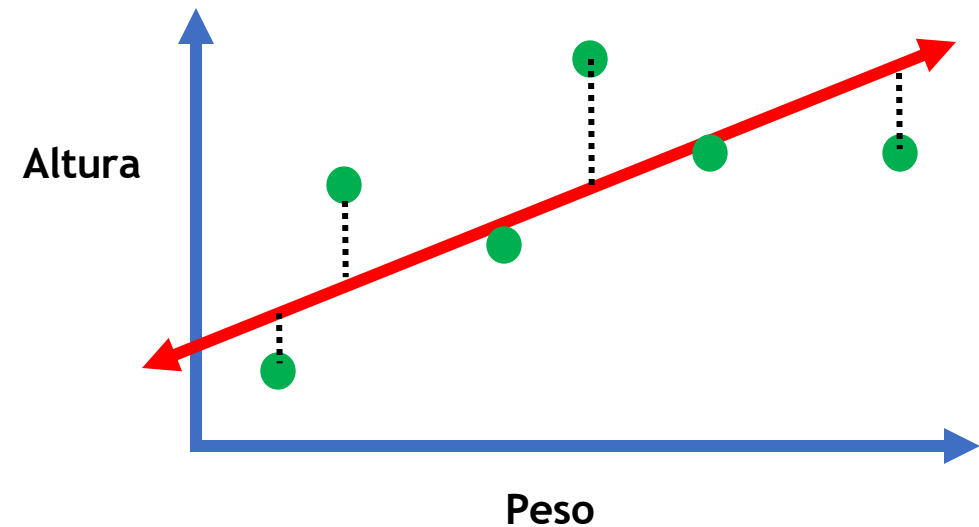
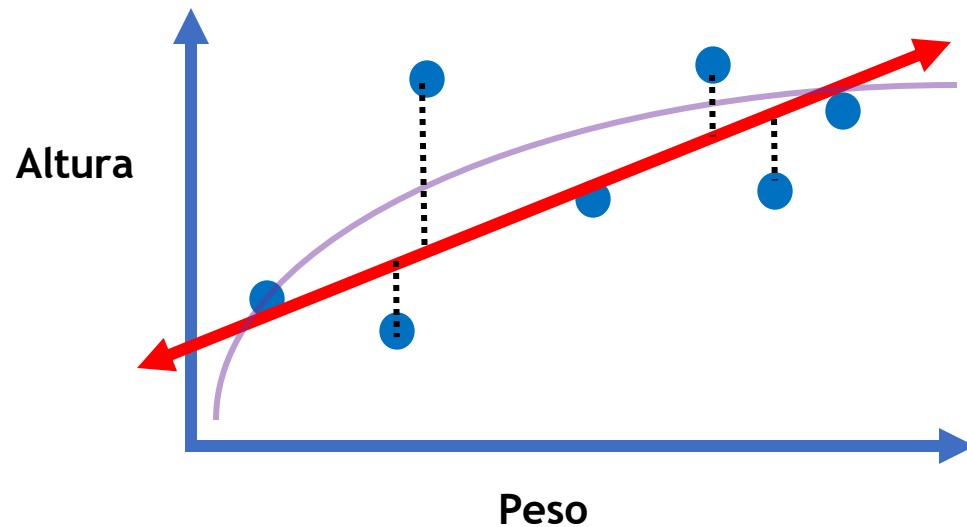
Machine Learning: Fundamentos

Por outro lado, a **Linha Reta** possui um **viés** relativamente alto, pois ela não é capaz de representar bem curva da relação entre peso e altura.



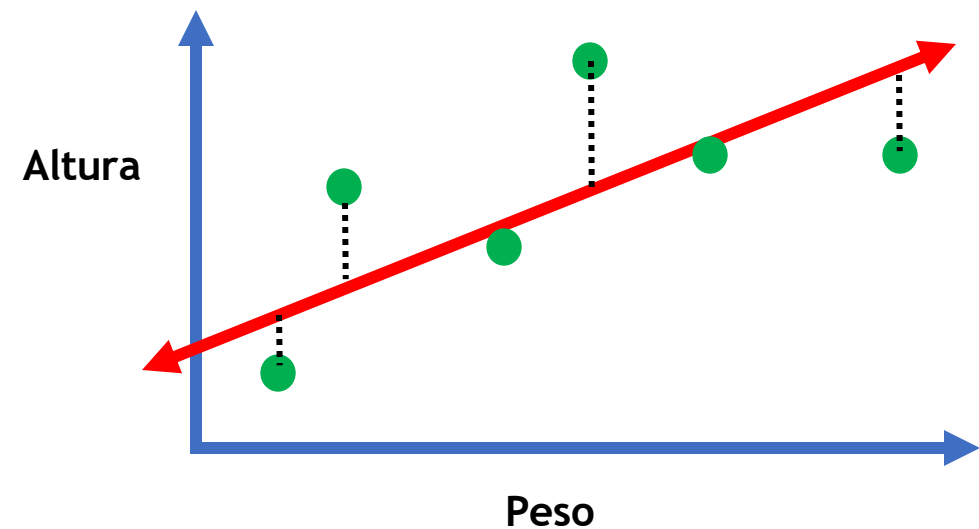
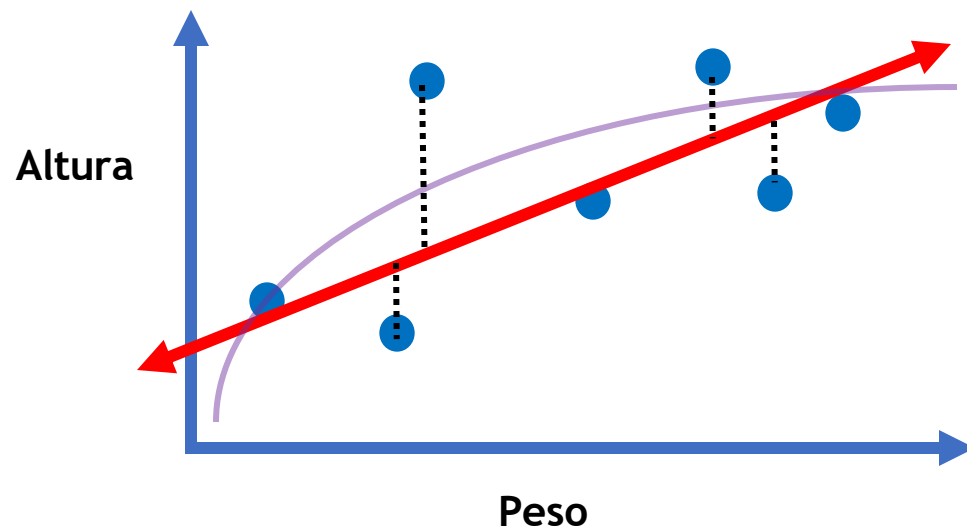
Machine Learning: Fundamentos

Mas a **Linha Reta** possui uma **variância relativamente baixa**, porque a **Soma dos Quadrados** resulta em valores **similares** para conjunto de dados diferentes.



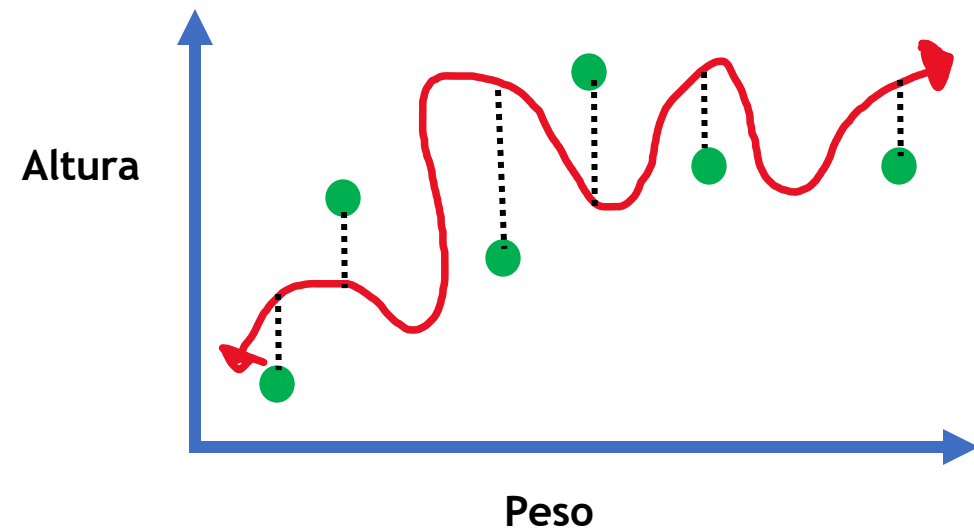
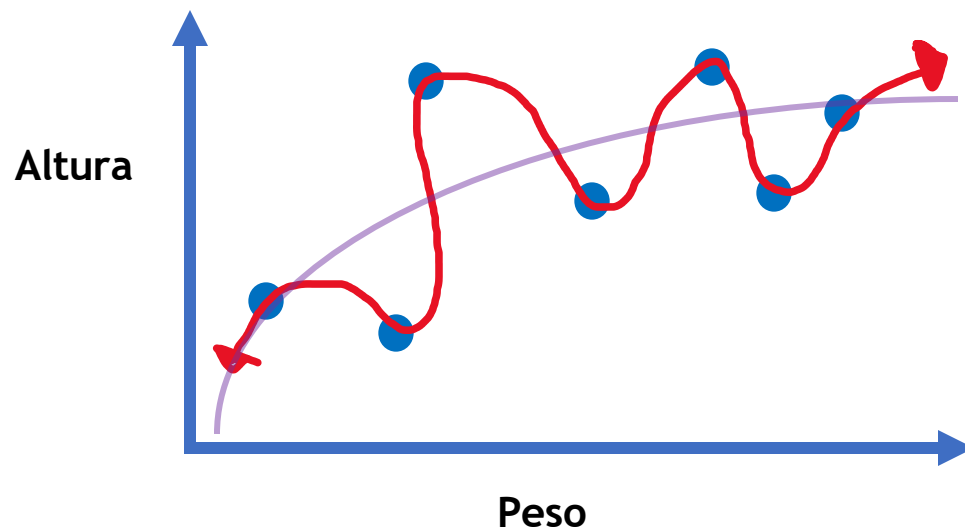
Machine Learning: Fundamentos

Em outras palavras, a **Linha Reta** poderá nos dar **BOAS** previsões, mas não previsões **ÓTIMAS**. Mas ainda assim elas seriam consistentemente boas previsões.



Machine Learning: Fundamentos

Como a Linha com Curvas se ajusta muito bem ao conjunto de treinamento, mas não ao conjunto de testes. Dizemos que a Linha com Curvas está **SUPERAJUSTADA** (**overfit**)



Machine Learning: Fundamentos

Overfitting: Quando um modelo apresenta uma baixa **capacidade de generalização**, a regra está superajustada aos dados. Isso significa que o modelo memorizou ou se especializou nos dados de treinamento.

Underfitting: Na situação inversa, quando o modelo possui **baixa capacidade preditiva para os dados de treinamento**, pode ter ocorrido um subajuste (underfitting). Subajuste geralmente ocorre quando os dados de treinamento disponíveis são pouco representativos ou quando o modelo usado é muito simples e não captura os padrões existentes.

Underfitting x Overfitting

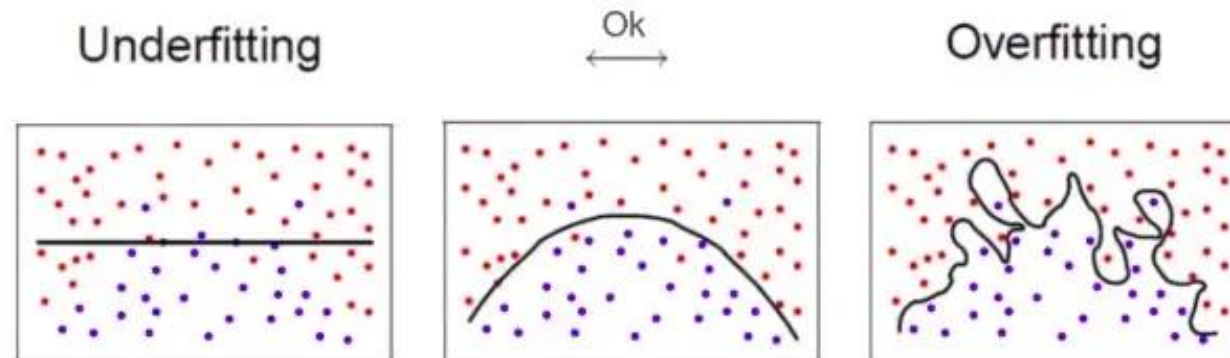
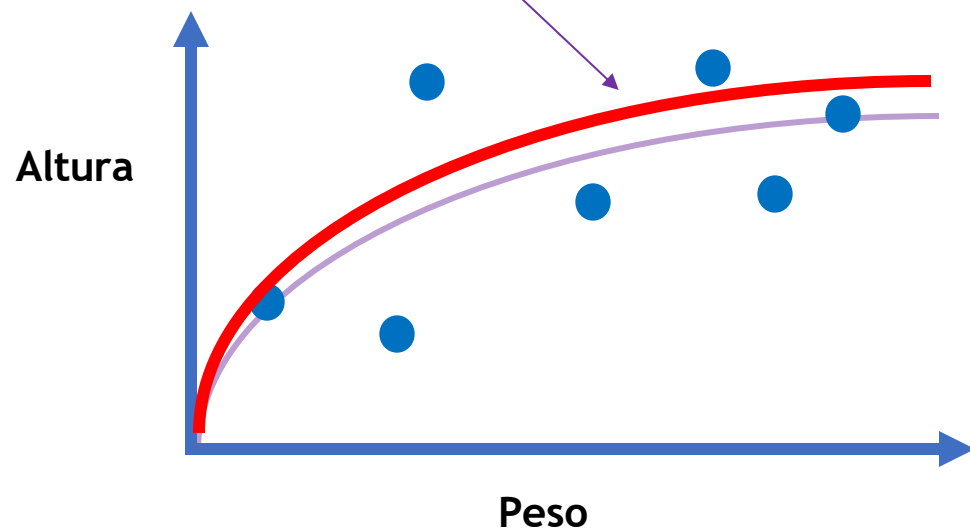


Imagem fonte: <https://www.datascienceacademy.com.br>

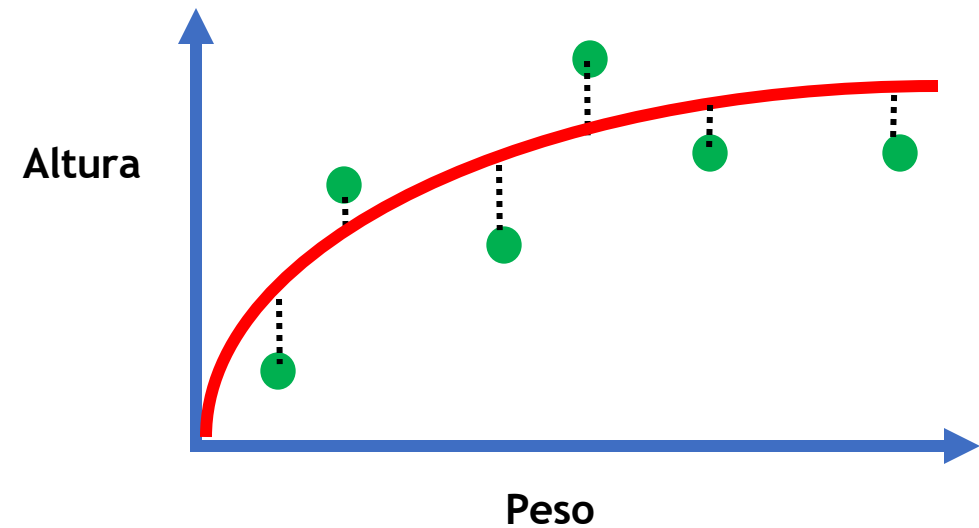
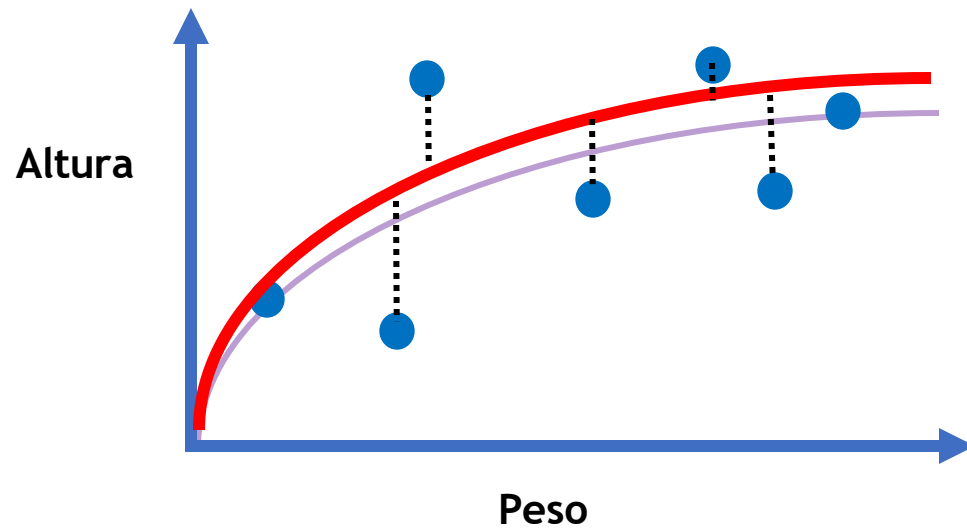
Machine Learning: Fundamentos

Em Machine Learning, o algoritmos ideal possui **Baixo Viés** e consegue adequadamente representar a relação verdadeira...



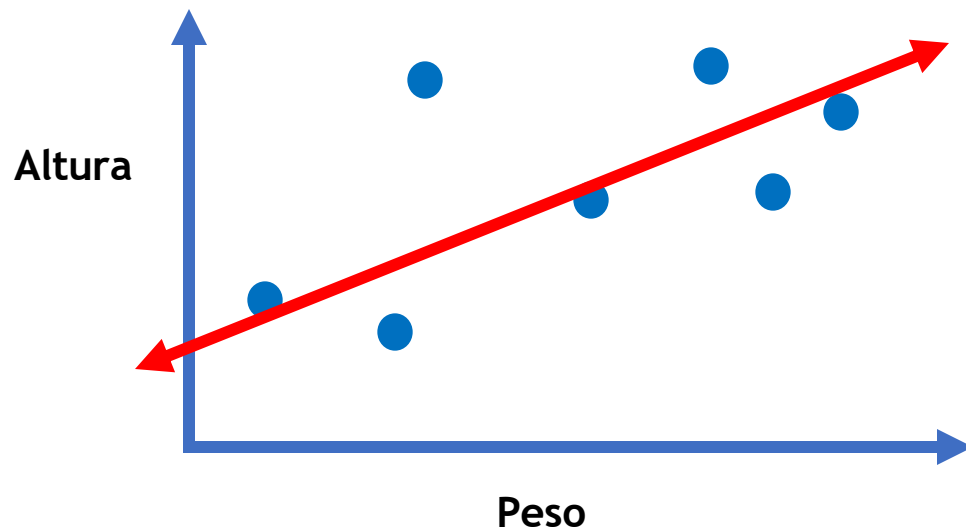
Machine Learning: Fundamentos

... e possui **baixa Variabilidade**, por produzir previsões consistentes para conjuntos de dados diferentes.



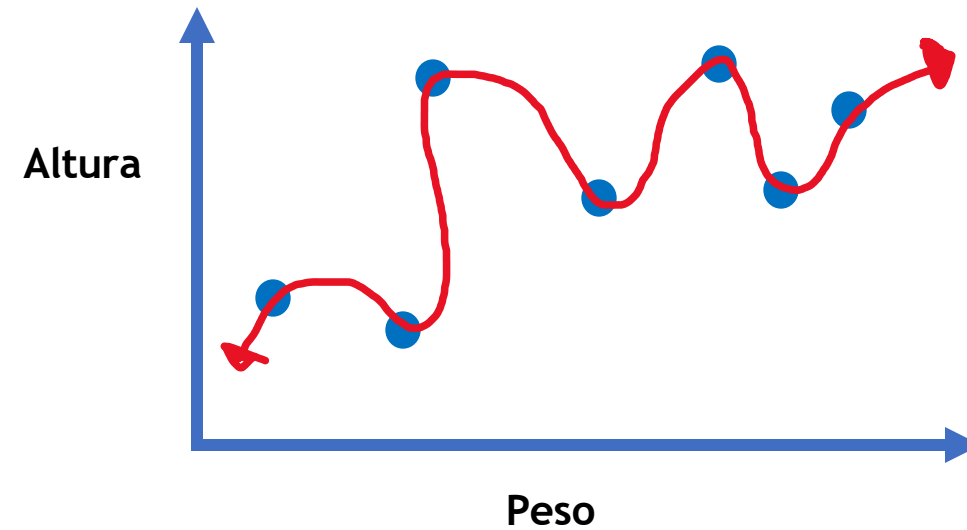
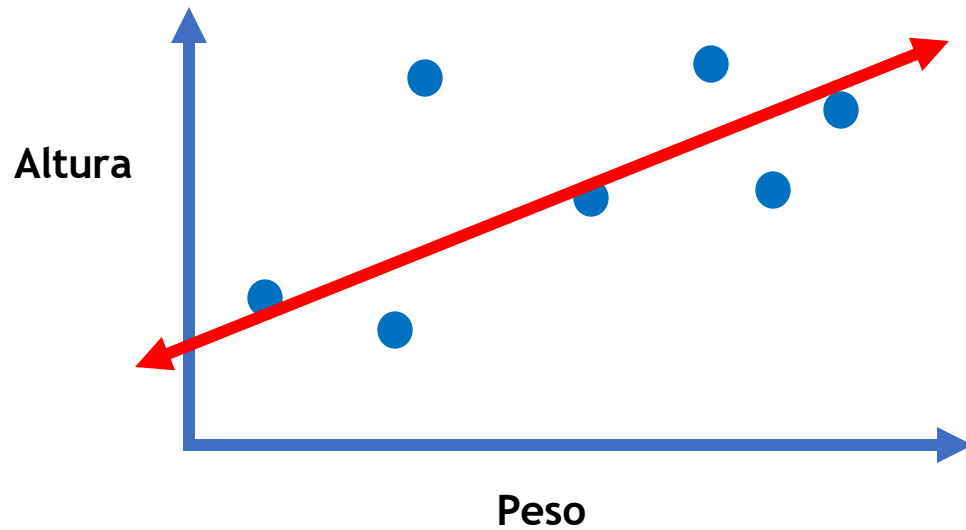
Machine Learning: Fundamentos

Isso pode ser feito ao encontrar o equilíbrio ideal entre um **modelo simples**...



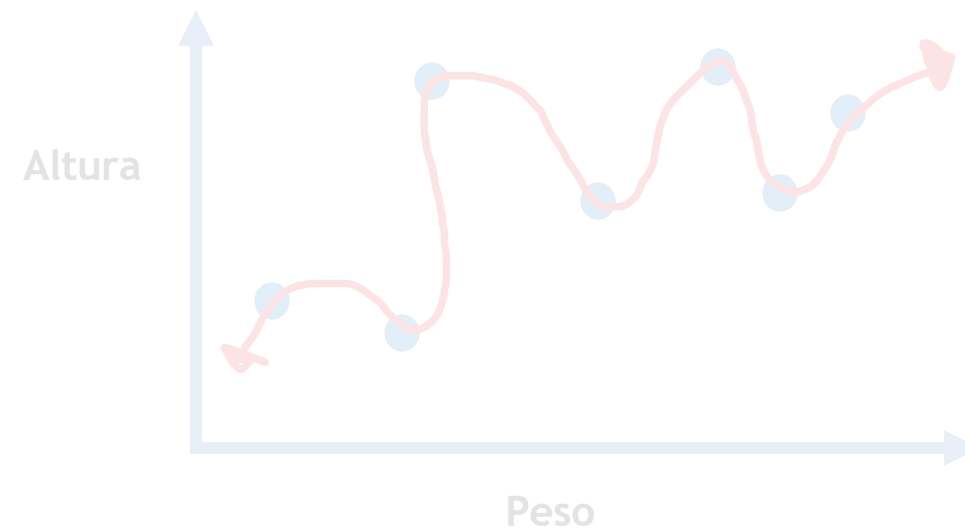
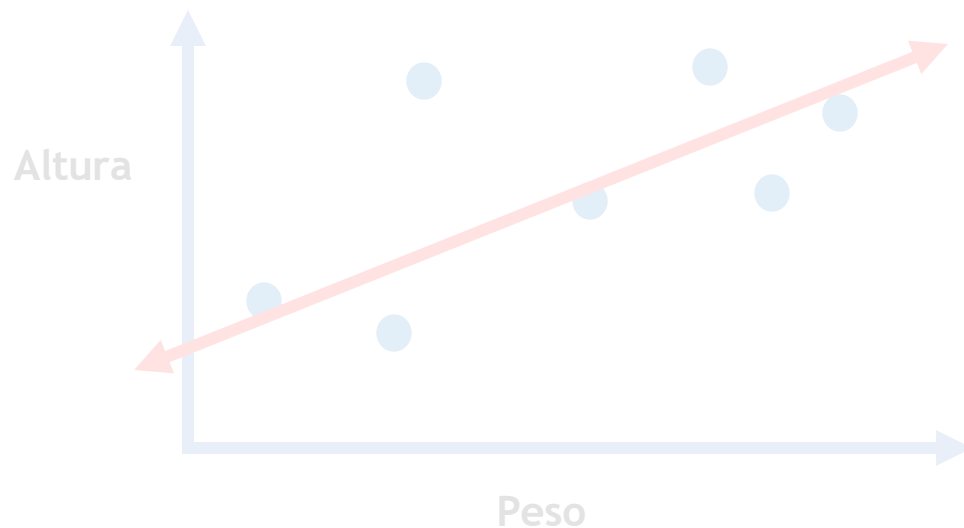
Machine Learning: Fundamentos

... e um modelo complexo.



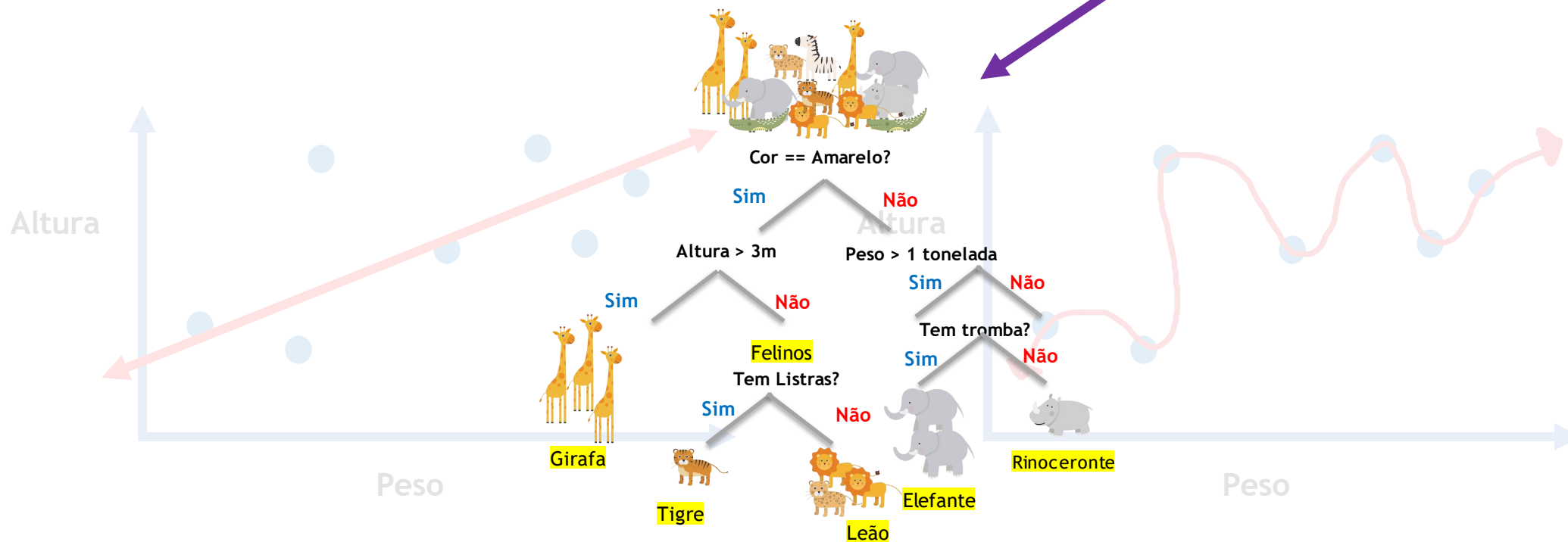
Machine Learning: Fundamentos

Para tentar encontrar um equilíbrio ideal entre um modelo simples e um modelo complexo, são utilizadas técnicas como: **Regularization**, **Boosting** e **Bagging**



Machine Learning: Fundamentos

Para tentar encontrar um equilíbrio ideal entre um modelo simples e um modelo complexo, são utilizadas técnicas como: **Regularization**, **Boosting** e **Bagging**.



Examples of Confirmation Bias



Not seeking out
objective facts

Não buscar
fatos objetivos



Interpreting information to
support your existing belief

Interpretando informações
para apoiar sua crença
existente



Only remembering details
that uphold your belief

Lembre-se apenas de
detalhes que
sustentam sua crença



Ignoring information that
challenges your belief

Ignorar informações
que desafiem sua
crença

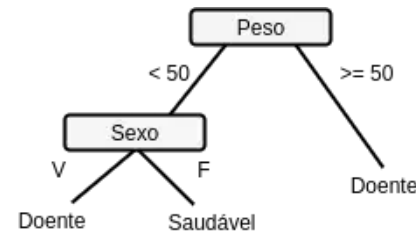
verywell

Viés Indutivo

Quando um algoritmo de AM está aprendendo a partir de um conjunto de dados de treinamento, ele está **procurando uma hipótese**, no espaço de possíveis hipóteses, capaz de descrever as relações entre os objetos e que melhor se ajuste aos dados de treinamento.

A representação utilizada define a preferência ou **viés** (*bias*) de representação do algoritmo e pode restringir o conjunto de hipóteses que podem ser induzidas pelo algoritmo.

Cada algoritmo utiliza uma forma ou representação para descrever a hipótese induzida. Por exemplo, **redes neurais artificiais** representam uma hipótese por um conjunto de valores reais, associados aos pesos de suas conexões da rede. **Árvores de decisão** utilizam uma estrutura de árvore em que cada nó interno é representado por uma pergunta referente ao valor de um atributo e cada nó externo está associado a uma classe.



(a) Árvore de decisão

Se Peso ≥ 50 então Doente
Se Peso < 50 e Sexo = M então Doente
Se Peso < 50 e Sexo = F então Saudável

(b) Conjunto de regras

0,45	-0,40	0,54	0,12	0,89	0,37
0,45	-0,48	0,54	0,12	0,89	0,37
-0,29	-0,40	0,54	0,12	0,89	0,37

(c) Redes Neurais

Viés Indutivo

Além do **viés de representação**, os algoritmos de AM possuem também um **viés de busca**. O viés de busca de um algoritmo é a forma como o algoritmo busca a hipótese que melhor se ajusta aos dados de treinamento. Por exemplo, o algoritmo *ID3*, que é utilizado para indução de árvores de decisão, tem como **viés de busca a preferência por árvores de decisão com poucos nós**.

Assim, **cada algoritmo de AM possui dois vieses**, um viés de representação e outro de busca. O viés é necessário para restringir as hipóteses a serem visitadas no espaço de busca. **Sem viés não haveria aprendizado/generalização**. Os modelos seriam especializados para exemplos individuais.



Exercícios

Vamos Treinar?

1) Se um modelo tem um grande desvio no conjunto de testes e uma pequena variação, isso significa que o modelo?

- ☐ Sobreajuste (Overfitting)
- ☐ Pode ser o excesso de ajuste pode estar subjacente
- ☐ Apenas ajuste
- ☐ Underfitting
- ☐ NA

Vamos Treinar?

1) Se um modelo tem um grande desvio no conjunto de testes e uma pequena variação, isso significa que o modelo?

(X) Sobreajuste (Overfitting)

() Pode ser o excesso de ajuste pode estar subjacente

() Apenas ajuste

() Underfitting

() NA

2) O rótulo previsto pelo algoritmo de regressão é?

- ☐ Auto-deformação
- ☐ Discreto
- ☐ Tipo de deformação
- ☐ Contínuo
- ☐ NA

2) O rótulo previsto pelo algoritmo de regressão é?

- ☐ Auto-deformação
- ☐ Discreto
- ☐ Tipo de deformação
- ☒ Contínuo
- ☐ NA

3) Qual das alternativas a seguir não é uma política de aprendizado integrada em algoritmos de aprendizado de máquina?

- ☐ Bagging (Ensacamento)
- ☐ Boosting (Impulsionando)
- ☐ Stacking (Empilhamento)
- ☐ Marking (Marcação)
- ☐ NA

3) Qual das alternativas a seguir não é uma política de aprendizado integrada em algoritmos de aprendizado de máquina?

- ☐ Bagging (Ensacamento)
- ☐ Boosting (Impulsionando)
- ☐ Stacking (Empilhamento)
- ☒ Marking (Marcação)
- ☐ NA

4) Os seguintes indicadores de avaliação pertencentes ao algoritmo de regressão são?

- ☐ Taxa de recordação
- ☐ Matriz de confusão
- ☐ Erro quadrado médio
- ☐ Precisão
- ☐ NA

Vamos Treinar?

4) Os seguintes indicadores de avaliação pertencentes ao algoritmo de regressão são?

- ☐ Taxa de recordação
- ☐ Matriz de confusão
- ☒ Erro quadrado médio
- ☐ Precisão
- ☐ NA

5) *(Múltipla escolha)* Os seguintes indicadores de avaliação pertencentes ao algoritmo de regressão são?

- ☐ Aumentar a quantidade de dados
- ☐ Engenharia de recursos
- ☐ Reduza os parâmetros de regularização
- ☐ Adicione recursos

Vamos Treinar?

5) (*Múltipla escolha*) Os seguintes indicadores de avaliação pertencentes ao algoritmo de regressão são?

☒ Aumentar a quantidade de dados

☒ Engenharia de recursos

☐ Reduza os parâmetros de regularização

☒ Adicione recursos



UFRR



Softex



Até a próxima aula!

Continue praticando :)